ABSTRACT

ADEDIPE, FOLUKEMI EBUNOLUWA. Use of Transcriptome Sequencing and RNA Interference for Target-site Screening in Two Agricultural Pests: Western Corn Rootworm and Kudzu Bug. (Under the direction of Dr. Marcé Lorenzen).

The western corn rootworm, *Diabrotica virgifera virgifera*, and the kudzu bug, *Megacopta cribraria* are both agricultural pests in the United States. Currently, control of these species is achieved through the use of genetically engineered (GE) crops and foliar and in-furrow insecticides. However, an increase in public concern about the safety of insecticides, along with the evolution of resistance to insecticides and to GE crops, has resulted in a need to find alternative control methods. In this dissertation I report on the use of transcriptome sequencing and RNA interference (RNAi) for target-site screening in these two agricultural pests. In my study, I assembled Illumina sequencing reads to create combined transcriptome assemblies for each species, and then generated searchable databases for each. I then surveyed the respective databases for potential target genes and tested these via RNAi to assess loss-of-function phenotypes. These data represent a valuable resource that will aid future research on functional genomics in these economically-important species and could provide insight for new pest management methods.

Use of Transcriptome Sequencing and RNA Interference for Target-site Screening in Two Agricultural Pests; Western Corn Rootworm, and Kudzu Bug

by
Folukemi Ebunoluwa Adedipe

A dissertation submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Entomology

Raleigh, North Carolina
2018

APPROVED BY:

Marcé Lorenzen Chair of Advisory Committee	Brian Wiegmann
Yasmin Cardoza	Dominic Reisig
Alberto Bressan External member	

BIOGRAPHY

I have many memories of farmers while I was in high school, which have served as the motivation that drives my career goals today. Growing up in Nigeria, I saw how farmers labored and suffered, only to get little income from their produce, due to poor farming methods, a lack of proper storage facilities, and the lack of appropriate pest control. Because farmers do not have suitable storage and pest control methods, they need to sell produce quickly at low prices during the harvest season, only to suffer from a lack of food and income during the dry season.

I have long desired to help farmers increase productivity within a system that is environmentally friendly and sustainable. I knew that achieving this goal would require skills and knowledge I lacked; so, I undertook undergraduate studies in Post-Harvest (Storage) Technology from Federal University of Technology in Akure, Nigeria, and graduate studies (Master's degree) in Integrated Pest Management from West Virginia University (WVU). However, I realized that to accomplish my goals, I needed further education, which is why I came to NC State in Fall 2013 to pursue a Ph.D. in Molecular Entomology under the supervision of Dr. Marcé Lorenzen.

My research focuses on finding alternatives to traditional pesticides. I have been using a combination of transcriptomics and RNA interference, a highly-effective species-specific gene-silencing mechanism, to gain insight into gene function in the western corn rootworm and the kudzu bug. I hope that my research will help establish new target sites for insect control. In my free time, I enjoy cooking for my family and watching movies.

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my outstanding supervisor Marcé Lorenzen. Since I met her in 2013, she has been an inspiring teacher and mentor. Her support and constant keen interest in my research made everything possible. I have greatly valued her set of skills, guidance and support. I would also like to thank all my committee members, Brian Wiegmann, Yasmin Cardoza, Dominic Reisig, and Alberto Bressan for their insightful advice, willingness to help, and as well for their support and guidance. I have enjoyed their valuable expertise and experiences, and my debt to them is enormous.

Colleagues in schools and departments have had a huge influence, which has reflected in this study. I thank all of the past and present members of the Lorenzen lab who have not only provided their expertise but also their friendship. In particular, I acknowledge Dr. Nathaniel Grubbs for taking time out of his busy schedule to answer my research questions and help improve my writing skills; also, I want to thank Dr. Clay Chu for his valuable help and contributions (particularly with western corn rootworm) to my research. Thanks to my fellow graduate students Steven Reyna and Nicole Gutzmann for being very supportive throughout my studies. I would also like to thank Will Klobasa, Stephanie Neal, Terri O’Leary and other lab members who provided help whenever I needed assistance. Everyone’s knowledge, support, guidance and friendship over the years have been invaluable.

I have always valued the contribution of my parents. They made enormous sacrifices to give me the best education and have always given me their unconditional love and support. My deepest thanks go to my husband Dr. Oluwatosin Adedipe and my children Iseoluwa, Ewaoluwa, and Oluwadarasimi, who have always had faith in me, loved me, and who have supported me throughout my years of study. They have made huge sacrifices by allowing me to pursue my

dream and study full-time, for that, and for everything else, I will always be grateful. I dedicate this work to them.

TABLE OF CONTENTS

LISTS OF TABLES

LISTS OF FIGURES

Chapter 1

Chapter 3

Chapter 4

CHAPTER 1

Introduction and Literature Review of Western Corn Rootworm, *Diabrotica virgifera virgifera*, and Kudzu Bug, *Megacopta cribraria*

Pest status of western corn rootworm

The western corn rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae), is one of the most economically important agricultural pests in the United States (US). These beetles reproduce in corn (*Zea mays* L.), the most widely-planted field crop in the US, generating about $46 billion dollars annually (USDA-ERS 2018), with overall losses due to rootworm damage and control that are estimated around $1 billion per year in the US alone (Gray et al., 2009). *D. virgifera virgifera* are reported to have originated in Guatemala, Central America (Melhus et al., 1954) where they have been a problem for over 5,000 years (Melhus et al., 1954; Gray et al., 2009). The first report of *D. virgifera virgifera* as a pest of corn in the US was in 1940 (Wangila et al., 2015). However, their appearance was first reported in Kansas, USA in 1868 (Chiang 1973), and they are currently found in all corn-growing states throughout the US (Krysan and Miller 1986). They are highly successful at invading new corn-growing regions and are currently spreading through Europe (Baufeld and Enzian 2001), having already invaded at least 22 European countries (EPPO 2011; Lemic et al., 2015).

D. virgifera virgifera are problematic because the larvae feed on corn roots, subsequently reducing the plant’s ability to take up water and nutrients (Hou et al., 1997; Riedell 1999; Riedell 1993; Dunn and Frommelt 1998). This can directly lower yields and, depending on the level of root damage, can cause plants to fall over (lodge) in strong winds or heavy rain, which results in mechanical harvest losses. Although adult feeding mainly occurs on corn silks and is generally considered to cause only minor damage to the grain, the bigger problem is that adults can potentially transmit and spread maize chlorotic mottle virus (Jensen 1985) and corn stalk rot fungi (Gilbertson et al., 1986). Despite the fact that more insecticides are used in the US to manage *Diabrotica spp* than any other single pest (EPA 2005), their biology, ecology and behavior have been largely understudied. This is in large part due to the complexities associated with rearing and studying these soil dwelling insects for which no long-term artificial diet exists.

Rootworm Biology

D. virgifera virgifera are univoltine and overwinter as eggs in the soil, with the eggs typically hatching in late spring (Pruess et al. 1968; Gray et al. 2009). Depending on temperature they generally complete the rest of their life cycle within 4-6 weeks in the field (Jackson and Elliot 1988, Woodson and Jackson 1996). The larvae are the most economically damaging stage (Branson and Krysan 1981; Oyediran et al. 2004) with first-instar larvae feeding on root hairs (Chiang 1973), and second- and third-instar larvae tunneling into the vascular cylinder of the root. These feeding activities can result in pruned root systems (Figure 1), which in turn reduces the host plant’s ability to obtain nutrients from the soil which reduces yields and increases the odds of plant lodging (Branson and Krysan 1981; Riedell 1993). Pupation occurs in the soil with males emerging several days earlier than females (Jackson and Elliott 1988). Thirteen days after females enclose, they become sexually mature and start to lay eggs, and can continue to lay eggs for a couple of months (Branson and Johnson 1973; Hill 1975). Studies have shown that female *D. virgifera virgifera* live longer than males (Ball and Hill 1952; Branson and Johnson 1973, Hill 1975), and that in the laboratory their mean fecundity is about 1000 eggs (Hill 1975), depending on the available nutrition. Adults mainly feed on the vegetative and reproductive tissues and appear to mate in the field they emerge from (Nowatzki et al. 2002). However, when food becomes scarce, adults will move to nearby corn fields (Ball and Hill 1952; Levine and Oloumi-Sadeghi 1996).

Rootworm Management

Sampling and monitoring: There are several methods by which the number of *Diabrotica spp* eggs, larvae and adults can be estimated. For eggs, the most common method is soil sampling. Eggs are extracted from the soil sample by washing, sieving, and floating the eggs (Chandler et al., 1966). Larval sampling is generally carried out through plant-area soil sampling, using soil cubes and soil cores (Fisher and Bergman 1986). Techniques for separating larvae from the soil or plant roots is similar to that of eggs because the process also uses water to float the larvae by first mixing the soil or roots with water (Fisher and Bergman 1986) and gathering them using a Berlese-type funnel (Fromm et al., 1998). However, sampling soil for eggs and larvae is extremely labor intensive and to date no meaningful egg and larval thresholds have been set. Sampling for pupae is also difficult because they are inconspicuous in earthen cells, they are very fragile, and the pupal stage is relatively short (Chiang 1973). Therefore, the most frequently used monitoring method is based on the number of adults present. This can be done by counting how many adult beetles are on representative corn plants, caught on yellow sticky traps (Hein and Tollefson 1985a, Kuhar and Youngman 1998), or captured in emergence traps (Hein and Tollefson 1985b).

Synthetic insecticidal control: Insecticides such as cyclodienes (Ball and Weekman 1962), organophosphates (Meinke et al., 1998), and carbamates (Meinke et al., 1998), are recommended for *D. virgifera virgifera* control (Gassmann and Weber 2012). These chemicals can be applied to the soil at planting to prevent larvae from feeding on corn roots or can be applied aerially to reduce the adult population before oviposition takes place (Meinke et al., 1998). However, *D. virgifera virgifera* have evolved resistance to many of these chemicals. For example, Ball and Weekman (1962) documented resistance to cyclodiene in *D. virgifera virgifera* larvae as far back as 1959. Interestingly, this resistance persisted, even years later after the insecticides were no longer in use (Parimi et al., 2006; Meinke et al., 1998). The genetic basis of resistance to cyclodienes such as aldrin, is most likely a mutation in a subunit of the gamma-aminobutyric acid receptor (Wang et al., 2013).

Crop rotation and resistance: To lessen selective pressure on resistance genes in *D. virgifera virgifera* populations, alternative management strategies such as crop rotation were implemented. Since 1912, farmers have found that rotating corn with soybeans is an effective method of *D. virgifera virgifera* control. When overwintering eggs hatch in soybeans fields, neonates die of starvation unless they are able to locate a suitable host shortly after hatching (Gillette, 1912). Surprisingly, *D. virgifera virgifera* have developed resistance to crop rotation, an unusual trait that was documented in East-central Illinois as early as 1987 (Levine et al., 2002; Knolhoff et al., 2006). This is one of the reasons that *D. virgifera virgifera* are thought to be highly adaptive, allowing them to overcome different pest-control methods through a variety of resistance mechanisms more rapidly than other pest species.

Bt corn and resistance: *Bacillus thuringiensis (Bt)* Berliner is a gram positive, soil dwelling bacterium which produces spores that contain crystal (cry) proteins. These cry proteins are toxic to some insect species and have been used as bioinsecticides since the late 1950s, but, their high cost and variable efficacy once limited their use to a small niche market (Cannon 1993). However, everything changed when genetically engineered (GE) corn (aka *Bt* corn) expressing one or more cry proteins was approved (EPA, 2003). Importantly, due to their specificity, each cry protein (aka *Bt* toxin) only affects a few target species. For example, Cry3Bb1, mCry3A and Cry34/35Ab1 are only effective for control of coleopterans such as *D. virgifera virgifera*, but not lepidopterans. Studies have shown that adoption of *Bt* corn has significantly reduced crop damage, and increased yields (Tabashnik et al., 2008).

Originally many thought that as long as *Bt* corn produced a high enough dose of Bt toxin to kill most of the rootworms that feed on it, resistance would never develop (James, 2009). However, this is not the case, some populations have developed resistance to one, or more *Bt* toxins. For example, field populations of *D. virgifera virgifera* resistant to Cry3Bb1 and mCry3A have been discovered in Iowa (Gassmann et al 2011; Monsanto 2011). This might be due to the continual usage of Bt corn, causing a high selective pressure (Tabashnik et al., 2008), and moreso, low levels of *Bt* expression can also cause a similar effect (Diverly et al., 2016; Tabashnik et al. 2013). To slow resistance development to *Bt*, a refuge strategy has been implemented (Gould, 1998). This method involves growing some non-*Bt* corn in, or near the field having Bt corn to enable survival of non-resistant (aka susceptible) beetles. Since only a few resistant individuals are expected to emerge from a *Bt* field, the refuge helps maintain a high percentage of susceptible alleles in the population, thus delaying widespread resistance (Gould, 1998). Yet, the basic biological question of what makes them so adaptive remains unanswered. To date, *Bt* corn and crop rotation are the major control methods for this pest; however, new methods are being developed.

One promising new pest control method is based on RNA interference (RNAi), a naturally occurring phenomena that can silence genes. RNAi was first used in the laboratory for understanding gene function, but was quickly coopted as a potential means of crop protection (Baum et al., 2007). Since RNAi has been shown to work exceptionally well in *D. virgifera virgifera* (Baum et al., 2007; Alves et al., 2010), I wanted to look at the impact of RNAi on other target genes. However, a major requirement for RNAi is prior knowledge of each target gene's sequence. Since the sequence of the *D. virgifera virgifera* genome is not yet publically available I utilized Next Generation Sequencing (NGS) technology to obtain transcriptomic data for several *D. virgifera virgifera* life stages. The data were then surveyed for members of the ABC superfamily of genes and a subset of these were tested via RNAi.

Pest Status of Kudzu Bug

The kudzu bug, *Megacopta cribraria*, (Hemiptera: Plataspidae), a native of Asia and Australia, was first discovered in northern Georgia during the fall of 2009, and until that time *M. cribraria* had not been seen in the United States (Suiter et al., 2010). However, since then they have been reported in thirteen additional states: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and the District of Columbia (Figure 2), (www.kudzubug.org 2015; Gardner et al., 2013). Their rapid spread has been attributed to being strong fliers (Suiter et al., 2010).

While not a major pest in its home range, *M. cribraria* is a pest of soybeans (*Glycine max* L. Merrill); the second most planted field crop in the US, generating about $42 billion dollars annually (USDA-ERS 2018). The primary host plants that support *M. cribraria* reproduction and nymphal development are the invasive kudzu plant, *Puereria montana* (Loureiro) Merrill variety *lobata* (Willdenow), and soybean (Suiter et al. 2010, Zhang et al. 2012). There has been speculation that *M. cribraria* was deliberately brought to the US to control the kudzu plant, and *M. cribraria* has been shown to reduce kudzu biomass by as much as 33% (Zhang et al., 2012); however, there is no evidence that they can be used as a biocontrol to reduce the distribution of this invasive weed in the United States. In fact, the presence of vast acres of kudzu in US has helped the *M. cribraria* population to grow and expand (Blaustein 2001). Moreover, (Blount et al., 2015) described *M. cribraria* as a potential pest of several agronomically important legume crops, including soybean. *M. cribraria* have indeed been shown to be a serious pest in soybeans, with densities up to 50 or more adults per plant (Suiter et al., 2010), and if untreated, infestations could potentially devastate soybean production, causing yield losses as high as 60% (Seiter et al., 2013). The sustainability of this crop is not only of economic importance but is also vital to overall food security (USDA-ERS 2016).

Apart from being an agricultural pest, *M. cribraria* is also a nuisance pest. The problem stems from *M. cribraria* forming overwintering aggregations both in and around homes (Suiter et al., 2010). In the fall, they congregate on outside walls, where they defecate, leaving brown spots and are an annoyance to residents (Dhammi et al., 2016; Eger et al., 2010). Similar to stink bugs, when disturbed *M. cribraria* produces a defensive chemical which can be harmful to skin and eyes (Ruberson et al., 2013).

Kudzu Bug Biology

M. cribraria undergo one to three generations per year in China and Japan (Dhammi et al., 2016; Eger et al., 2010). Similarly, in the US, Zhang et al. (2012) reported that in Georgia, *M. cribraria* undergo two and maybe a third generation on kudzu. The authors further reported a significant variation between years. The adults have been found overwintering in leaf litter and under tree bark near kudzu patches and soybean fields, as well as in and around urban structures (Lahiri et al., 2015; Halbert & Eger 2010). Adults are not active until temperatures warm in the spring (Zhang et al., 2012; Reisig et al., 2012), at which time they disperse to kudzu, where they come together, feed, mate and lay eggs (Zhang et al., 2012; Reisig et al., 2012). Adults will also converge on volunteer or early-planted soybeans when available (Del Pozo-Valdivia and Reisig

2013). While they usually complete their first generation on kudzu (Del Pozo-Valdivia and Reisig 2013, Seiter et al., 2013), the second generation can develop on either kudzu or soybeans (Zhang et al., 2012, Seiter et al., 2013a).

M. cribraria nymphs and adults feed by piercing the plant's epidermis and sucking sap from the phloem ((Tayutivutikul and Yano 1990, Thippeswamy and Rajagopal 2005, Suiter et al., 2010), these feeding activities deprive the host of water and nutrients required for growth (Stubbins et al., 2017), reducing the number of pods per plant and number of beans per pod (Seiter et al., 2013; Waldvogel and Alder, 2011). Since *M. cribraria* do not overwinter in soybean fields it is not surprising that they infest the field by first occupying the edges, and later move into the field (Seiter et al., 2013).

Although studies have reported *M. cribraria* on a variety of host plants from several legume crops such as pigeon pea (*Cajanus cajan* L.), black eye pea (*Vigna sinensis* L.), lima bean (*Phaseolus lunatas* L.), pinto bean (*Phaseolus vulgaris* L.) (Medal et al., 2013), wisteria (*Wisteria sinensis* Sims) (Ruberson et al., 2013), white sweet clover (*Melilotus alba* Medikus), white clover (*Trifolium repens* L.), red clover (*Trifolium pretense* L.), alfalfa (*Medicago sativa* L.), perennial peanut (*Arachis glabrata* Benth), and American joint vetch (*Aeschynomene Americana* L.) (Medal et al., 2016), Eger et al., (2010) indicated that many of these are likely incidental occurrences. Therefore, to date the primary host plants for *M. cribraria* reproduction and nymphal development in the US are thought to be the invasive kudzu plant, and soybeans (Suiter et al., 2010, Zhang et al., 2012).

Interestingly, *M. cribraria* deposits a bacterical capsule containing *Candidatus Ishikawaella capsulata* along with their eggs (Jenkins et al., 2010; Brown et al., 2014). Moreover, it has been shown that *M. cribraria* relies on these endosymbiotic bacteria for proper development (Fukatsu and Hosokawa 2002; Hosokawa et al., 2006). Specifically, nymphs reared in the absence of this bacterium were shown to develop slower, have smaller body sizes and have higher death rates (Jenkins et al., 2010; Brown et al., 2014). Research over the past several decades has demonstrated that many of the phenotypic properties displayed by insects are due to inputs they receive from their associated symbionts (Fukatsu and Hosokawa 2002; Hosokawa et al., 2006; Seiter et al., 2015).

Kudzu Bug Management

Chemical control: In the US, insecticides are used to control *M. cribraria* (Seiter et al., 2014; Brown et al., 2015), In China, 85% of the *M. cribraria* populations found in soybean fields are controlled with the use of pyrethroids such as cypermethrin, deltamethrin, sumicidin and organophosphate insecticides (Zhang et al., 2012). However, insecticidal sprays alone may not be sufficiently effective to stop yield loss in soybeans when *M. cribraria* population sizes are high. It is not surprising that research has shown that yield loss from *M. cribraria* feeding is directly linked to the number of nymphs and adults per plant (Seiter, et al., 2013). Therefore, the economic threshold for insecticide application in soybean fields has been set at five adults per plant, or a single nymph per sweep (Stubbins et al., 2015; Dhammi et al., 2016). Moreso, a recent study reported that applying insecticide immediately after detecting the presence of one nymph per sweep was enough to keep the population below the economic threshold (Seiter et al., 2015). However, there are drawbacks to insecticidal sprays, including potential hazards to humans and the environment. For example, increased use of insecticides could have a negative impact on pollinators and other beneficial insects (Vanbergen et al., 2013). Therefore, finding alternative methods for managing this agricultural pest are needed.

Biological control: Given the fact that *M. cribraria* is exotic to the US, there is very little information about what natural enemies may impact it in the US. Several biological control agents such as adult parasitoids have been identified that appear to reduce the populations of *M. cribraria*. Specifically, two egg parasitoids, *Paratelenomus saccharalis* (Dodd) (Hymenoptera: Platygastridae), and *Ooencyrtus nezarae* (Hymenoptera: Encyrtidae), have been shown to be effective in parasitizing *M. cribraria* eggs in the US (Ruberson et al., 2013; Ademokoya et al., 2017). For example, study showed up to 40% parasitism in *M. cribraria* adults, eggs and in Soybean (Knight et al., 2017). Some even speculate that these parasitoids may have contributed to the sudden decrease in *M. cribraria* populations over the past few years (Ademokoya et al., 2017). Interestingly, many other insect species have been found to feed on *M. cribraria* including, the big-eyed bugs (*Geocoris punctipes*), the insidious flower bug (*Orius insidiosus*), the spined soldier bug (*Podisus maculientris*), the lady beetle (*Hippodamia convergens*), the assassin bug (*Zelus renardii*), the striped lynx spider (*Oxyopes salticus*) and the green lynx spider (*Peucetia viridians*) (Greenstone et al., 2014). Other natural enemies include the entomopathogenic fungus, *Beauveria bassiana* (Hypocreales: Clavicipitaceae), and two native dipteran species *Strongygaster triangulifera* and *Phasia robertsonii* (Golec et al., 2013, Ruberson et al., 2012). The fungus has been reported to infect both *M. cribraria* nymphs and adults (Golec et al., 2013, Ruberson et al., 2012); while the dipteran species have only been reported to parasitized *M. cribraria* adults. However, despite the current reduction in *M. cribraria* population size and the availability of both chemical and biological control methods, it is wise to investigate alternate control strategies since resistance is inevitable.

The ability of insects to evolve resistance to insecticides, including *Bt*, demonstrates a continuing need to understand pest biology, not only on the organismal level, but also at the molecular level.

My research focuses on the use of RNAi, a sequence-specific technique, (Fire et al., 1998) used for assessing gene function in arthropods due to its power, speed, and relative ease of use to knock-down transcript levels of essential genes in *M. cribraria*. Results from these studies are expected to provide insights into the role of potential target genes in *M. cribraria* and may help establish new target sites for insect control.

The objectives of my Ph.D. research were to obtain genetic (transcriptomic) data for *D. virgifera virgifera* and *M. cribraria*, establish custom databases, and survey these for target genes which I then characterized via RNAi to determine their potential for use in new control methods. This dissertation is composed of two manuscripts, chapters 3 and 4, each is written in the format of the journal to which it will be submitted.

Figure 1: Corn root damaged by the western corn rootworm, *Diabrotica virgifera virgifera.* The primary damage done by *Dvv* feeding on corn roots reduces corn plants structural support and makes it more susceptible to lodging. Undamaged corn roots (left) and damaged corn roots (right). Figure adapted from Jim Kalisch, University of Nebraska-Lincoln.

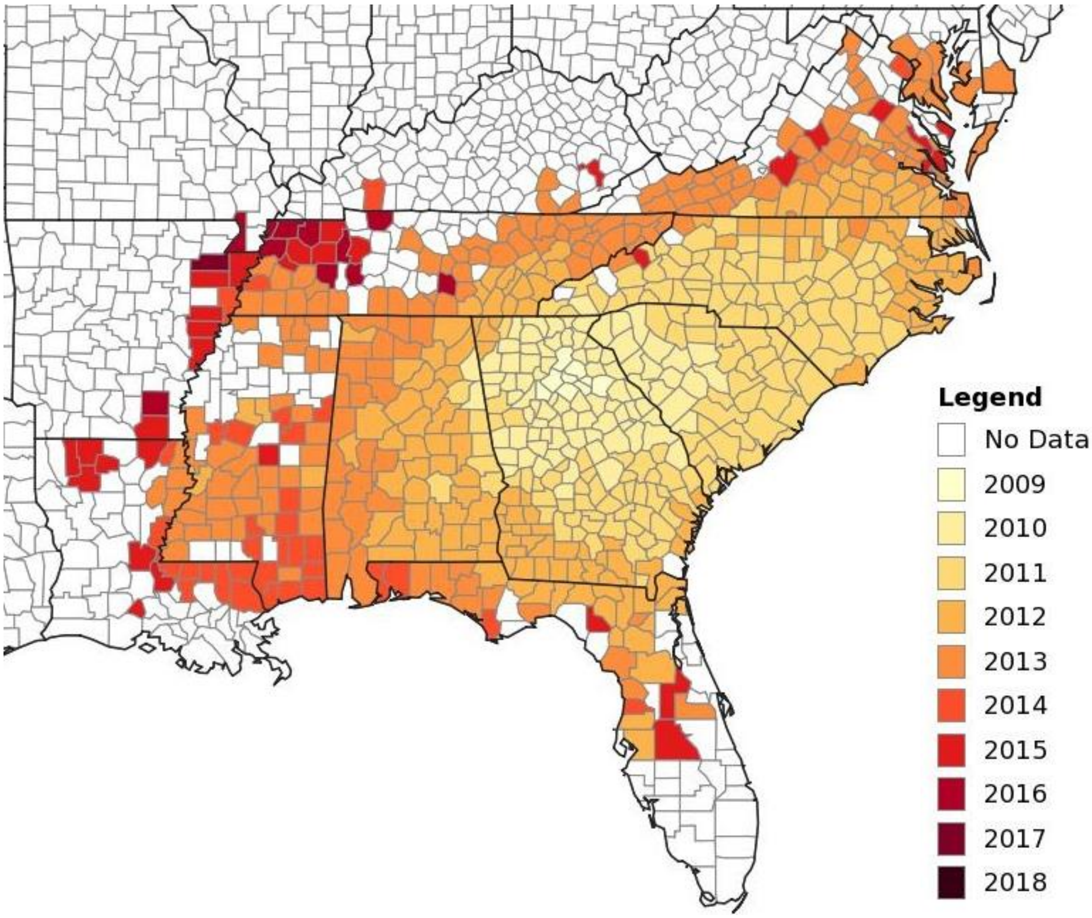


Figure 2: Distribution of the kudzu bug, *Megacopta cribraria* in the southeast U.S from 2009-2018. Map is adapted from University of Georgia (available at https://www.kudzubug.org/distribution-map/)

Figure 3: **Soybean stems damaged by *M. cribraria***. The primary damage is when nymph and adults feed on stems and petioles. Damaged stems exhibit brown lesions and often develop sooty mold. Figure adapted from https://soybeans.ces.ncsu.edu

REFERENCES

Ademokoya, B., Balusu, R., Ray, C., Mottern, J. & Fadamiro, H. (2017) The First Record of *Ooencyrtus nezarae* (Hymenoptera: Encyrtidae) on Kudzu Bug (Hemiptera: Plataspidae) in North America. *Insect Science*, **18**, 1–7.

Alves, A.P., Lorenzen, M.D., Beeman, R.W., Foster, J.E., & Siegfried, B.D. (2010) RNA interference as a method for target-site screening in the Western corn rootworm, *Diabrotica virgifera virgifera. Insect Science*, **10**, 162.

Ball, H. & Weekman, G. (1962) Insecticide resistance in adult western corn rootworm in Nebraska. *Journal of Economic Entomology* **55**, 439-441.

Ball, H. & Hill, R. (1952) You can control corn rootworm. Nebraska Export Station Quart. Winter, 53.

Baufeld, P. & Enzian, S. (2001) Simulations model for spreading scenarios of western corn rootworm (*Diabrotica virgifera virgifera*) in case of Germany. Proceedings of the XXI IWGO Conference, Legnaro, Italy.

Blaustein, R. J. (2001) Kudzu's invasion into Southern United States life and culture, In J. A. McNeeley (ed.), The great reshuffling: human dimensions of invasive species. IUCN/The World Conservation Union, Switzerland/Cambridge. pp. 55–62.

Blount, J.L., Buntin, G.D. & Roberts, P.M. (2016) Effect of planting date and maturity group on soybean yield response to injury by the kudzu bug, *Megacopta cribraria* F. (Hemiptera: Plataspidae). *Journal of Economic Entomology*, **109**, 207–212.

Branson, T., & Krysan, J. (1981) Feeding and oviposition behavior and life-cycle strategies of Diabrotica- an evolutionary view with implications for pest-management (Coleoptera, Chrysomelidae). *Journal of Economic Entomology*, **10**, 826-831.

Brown, S.A., Kerns, D.L., Williams, T.S., Emfinger, K. & Jones, N. (2015) Evaluation of foliar insecticides for kudzu bug control in soybeans. *Arthropod Management*, **40**, 1.

Brown, A.M., Huynh, L.Y., Bolender, C.M., Nelson, K.G. & McCutcheon, .JP. (2014) Population genomics of a symbiont in the early stages of a pest invasion. *Mol. Ecol*. **23,** 1516-1530. doi:10.1111/mec.12366.

Cannon, R. (1993) Prospects and progress for *Bacillus thuringiensis*-based pesticides. *Pesticide Science* **37**, 331-335.

Chandler, J.H., Musick, G.J. & Fairchild, M.L. (1966) Apparatus and procedure for separation of com rootworm eggs from soil. *Journal of Economic Entomology,* **59**, 1409-1410.

Chiang, H.C. (1973) Bionomics of the northern and western corn rootworms. *Annual Review Entomology*, **18**, 47-72.

Chu, C., Spencer, J., Curzi, M., Zavala, J. & Seufferheld, M. (2013) Gut bacteria facilitate adaptation to crop rotation in the western corn rootworm. Proceedings of the National Academy of Sciences **110**, 11917-11922.

Del Pozo-Valdivia, A.I., Seiter, N.J., Reisig, D.D., Greene, J.K., Reay-Jones, F.P.P. & Bacheler, J.S. (2016) *Megacopta cribraria* (Hemiptera: Plataspidae) Population dynamics in soybeans as influenced by planting date, maturity group, and insecticide use. *Journal of Economic Entomology*, **109**, 1141–1155.

Del Pozo-Valdivia, A.I. & Reisig, D.D. (2013) First-generation *Megacopta cribraria* (Hemiptera: Plataspidae) can develop on soybeans. *Journal of Economic Entomology* **106**, 533–535.

Dhammi, A., Krestchmar, J.B., Ponnusamy, L., Bacheler, J.S., Reisig, D.D., Herbert, A., Del Pozo-Valdivia A.I., & Roe, A.M. (2016) Biology, Pest Status, Microbiome and Control of Kudzu Bug (Hemiptera: Heteroptera: Plataspidae): A New Invasive Pest in the U.S. *International Journal of Molecular Science*, **17**, 1570.

Dively, G. P., Venugopal, P. D. & Finkenbinder, C. (2016) Field-Evolved Resistance in Corn Earworm to Cry Proteins Expressed by Transgenic Sweet Corn. *PLoS One*, **11**, e0169115

Dunn, J.P. & Frommelt, K. (1998) Effects of below-ground herbivory by *Diabrotica virgifera virgifera* (Coleoptera) on biomass allocation and carbohydrate storage of maize. *Applied Soil Ecology*, **7**, 213-218.

Eger, J.E. Jr., Ames, L.M., Suiter, D.R., Jenkins, T.M., Rider, D.A. & Halbert, S.E. (2010) Occurrence of the Old-World bug *Megacopta cribraria* (Fabricius) (Heteroptera: Plataspidae) in Georgia: A serious home invader and potential legume pest. *Insecta Mundi*, **121**, 1–11.

EPA (2002) Environmental Protection Agency. Transmittal of meeting minutes of the FIFRA Scientific Advisory Panel meeting held August 27-29.

EPA (2005) Environmental Protection Agency. EPA annual pesticide report. Available from http://www.epa.gov/oppfead1/annual/2005/05annualrpt.pdf

EPPO (2011) European and Mediterranean Plant Protection Organization.*Diabrotica virgifera*. Availablefromhttp://www.eppo.org/QUARANTINE/Diabrotica_virgifera/diabrotica_virgifera.htm

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., & Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391,** 806-11.

Fisher, J. & Bergman, M. (1986) Field sampling larvae and pupae, pp. 83-99, In Krysan, J. L., & Miller, T. A. (eds.), Methods for the study of pest *Diabrotica*. Springer-Verlag, New York.

Fromm, E.A., Bernklau, E.J. & Bjostad, L.B. (1998) Simple technique for scouting corn rootworm larvae. (http://www.colostate.edu/Depts/Entomology/posters/fromm981108.html).

Fukatsu, T. & Hosokawa, T. (2002) Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, *Megacopta punctatissima*. *Applied Environmental Microbiology*, **68**, 389–396.

Gardner, W.A., Blount, J.L., Golec, J.R., Jones, W.A., Hu, X.P., Talamas, E.J., Evans, R.M., Dong, X., Ray, C.H. Jr. & Buntin, G.D. (2013) Discovery of *Paratelenomous saccharalis* (Dodd) (Hymenoptera: Platygastridae), an egg parasitoid of *Megacopta cribraria* F. (Hemiptera: Plataspidae) in its expanded North American range. *Journal of Entomological Science*, **48**, 355–359.

Gassmann, A., Petzold-Maxwell, J., Keweshan, R. & Dunbar, M. (2011) Field-evolved resistance to *Bt* maize by Western corn rootworm. *PLoS one*,**6**, e22629.

Gilbertson, R.L., Brown, W.M. Jr., Ruppel, E.G. & Capinera J.L. (1986) Association of com stalk rot Fusarium spp. and western com rootworm beetles in Colorado. *Phytophathology*, **76**, 1309-1314.

Gillette, C. (1912) *Diabrotica virgifera* Leconte as a corn root-worm. *Journal of Economic Entomology* **5,** 364-366.

Golec, J.R., Hu, X.P., Ray, C. & Woodley, N.E. (2013) *Strongygaster triangulifera* (Diptera: Tachinidae) as a parasitoid of adults of the invasive *Megacopta cribraria* (Heteroptera: Plataspidae) in Alabama. *Journal of Entomological Science* **48**, 352–354.

Gould, F. (1998) Sustainability of transgenic insecticidal cultivars: Integrating pest genetics and ecology. *Annual Review of Entomology* **43**, 701-726.

Gray, M., Sappington, T., Miller, N., Moeser, J. & Bohn, M. (2009) Adaptation and invasiveness of western corn rootworm: intensifying research on a worsening pest. *Annual Review of Entomology*, **54**, 303-321.

Greenstone, M.H., Tillman, P.G. & Hu, J.S. (2014) Predation of the newly invasive pest *Megacopta cribraria* (Hemiptera: Plataspidae) in soybean habitats adjacent to cotton by a complex of predators. *Journal of Economic Entomology*, **107**: 947–954.

Gustin, R.D. & Schumacher, T.E. (1989) Relationship of some soil core parameters to movement of first-instar western com rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology*, **18**, 343-346.

Halbert, S. & Eger, J.E. (2010) Pest alert - Bean plataspid, Megacopta cribraria (Fabricius) (Hemiptera: Plataspidae) an exotic legume pest established in Georgia. Florida Department of Agriculture and Consumer Services. DACS-P-01725

Hein, G. L. & Tollefson, J. J. (1985a) Use of the Pherocon AM trap as a scouting tool for predicting damage by com rootworm (Coleoptera: Chrysomelidae) Larvae. *Journal of Economic Entomology*, **78**, 200-203.

Hein, G.L. & Tollefson, J.J. (1985b) Absolute sampling technique for com rootworm (Coleoptera: Chrysomelidae) adult emergence that adjusts to fit common-row spacing. *Journal of Economic Entomology*, **78**, 1503-1506.

Hill, R.E. (1975) Mating, oviposition patterns, fecundity, and longevity of the western com rootworm. *Journal of Economic Entomology*, **68**, 311-315.

Hosokawa, T., Kikuchi, Y., Nikoh, N., Shimada, M. & Fukatsu, T. (2006) Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. *Plos Biology* **4**, 1841-1851.

Hou, X.L., Meinke, J. & Arkebauer, T.J. (1997) Soil moisture and larval western com rootworm injury: Influence on gas exchange parameters in com. *Agronomy journal*, **89,** 709-717.

Ismael, Y. Bennett, R. & Morse, S. (2002) Farm-level economic impact of biotechnology: smallholder *Bt* cotton farmers in South Africa. *Outlook on Agriculture* **31**, 107-111.

Jackson, J.J. & Elliot, N.C. (1988) Temperature dependent development of immature stages of the western com rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae). *Environmental Entomology,* **17**, 166-171.

Jenkins, T.M., Eaton, T.D., Suiter, D.R., Eger, J.E., Ames, L.M. & Buntin, G.D. (2010) Preliminary genetic analysis of a recently-discovered invasive true bug (Hemiptera: Heteroptera: Plataspidae) and its bacterial endosymbiont in Georgia, USA. *J. Entomol. Sci.* **45** 1-2.

Jensen, S.G. (1985) Laboratory transmission of maize chlorotic mottle vims by three species of com rootworms. *Plant Disease,* **69**, 864-868.

Knight, I.A., Roberts, P.M., Gardner, W.A., Oliver K.M., Reay-Jones, F.P.F., Reisig, D.D. & Toews, M.D (2017) Spatial Distribution of *Megacopta cribraria* (Hemiptera: Plataspidae) Adults, Eggs and Parasitism by *Paratelenomus saccharalis*(Hymenoptera: Platygastridae) in Soybean. *Environmental Entomology*, **46**, 1292–1298

Knolhoff, L., Onstad, D. Spencer, J. & Levine, E. (2006) Behavioral Differences Between Rotation-resistant and Wild-type *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae) *Environmental Entomology,* **35**, 1049-1057.

Krysan, J.L. & Miller, T.A. (1986) Methods for the study of pest *Diabrotica*. Springer Verlag, West Hanover, MA.

Kuhar, T.P. & Youngman, R.R. (1998) Olson yellow sticky rap: Decision-making tool for sampling western corn rootworm (Coleoptera: Chrysomelidae) adults in field corn. *Journal of Economic Entomology*, **92***,* 957-963.

Lemic, D., Mikac, K., Kozina, A. & Renata, B. (2015). Monitoring techniques of the western corn rootworm are the precursor to effective IPM strategies. *Pest Management Science*. **72**, 10.1002/ps.4072.

Levine, E., Spencer, J.L., Isard, S.A., Onstad, D.W. & Gray, M. E. (2002) Adaptation of the western corn rootworm to crop rotation: evolution of a new strain in response to a management practice. *American Entomologist* **48**, 94-107.

Levine, E. & Oloumi-Sadeghi, H. (1996) Western corn rootworm (Coleoptera: Chrysomelidae) larval injury to corn grown for seed production following soybeans grown for seed production. *Journal of Economic Entomology*, **89**, 1010-1016.

Medal, J., Halbert, S., Cruz, A.S., Smith, T. & Davis, B.J. (2016) Greenhouse study to determine the host range of the kudzu bug, *Megacopta cribraria* (Heteroptera: Plataspidae). *Florida Entomologist* **99,** 303–305.

Medal, J., Cruz, A.S., Williams, K., Fraser, S., Wolaver, D., Smith, T. & Davis, B.J. (2015) First record of *Paratelenomus saccharalis* (Hymenoptera: Platygastridae) on kudzu bug *Megacopta cribraria* (Heteroptera: Plataspidae) in Florida. *Florida Entomologist,* **98**, 1250–1251.

Medal, J., Halbert, S., Smith, T. & Santa Cruz, A. (2013) Suitability of selected plants to the bean plataspid, *Megacopta cribraria* (Hemiptera: Plataspidae) in no-choice tests. *Florida Entomologist,* **96**, 631–633.

Meinke, L., Siegfried, B., Wright, R. & Chandler, L. (1998) Adult susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to selected insecticides. *Journal of Economic Entomology*, **91**, 594-600.

Melhus, I., Painter, R. & Smith, F. (1954) A search for resistance to the injury caused by species of Diabrotica in the corns of Guatemala. Iowa State College *Journal of Science*, **29**, 75-94.

Metcalf, R. J. (1986) *in* Krysan, J.L. & Miller T.A. [eds.], Methods for the study of pest *Diabrotica*. Springer, New York. pp. 7-15.

Monsanto (2011) Gassmann research is strong reminder of need for comprehensive IPM. Availableathttp://www.monsanto.com/newsviews/Pages/gassmann-research-reminder-of-need-for-comprehensive-IPM.aspx, Accessed December 19 2011.

Nowatzki, T.M., Tollefson, J.J. & Calvin, D.D. (2002) Development and validation of models for predicting the seasonal emergence of com rootworm (Coleoptera: Chrysomelidae) beetles in Iowa. *Environmental Entomology*, **31**, 864-873.

O'Neal, M.E., Gray, M.E. & Smyth, C.A. (1999) Population characteristics of a western corn rootworm (Coleoptera: Chrysomelidae) strain in east-central Illinois corn and soybean fields. *Journal of Economic Entomology*, **92**, 1301-1310.

Oyediran, I., Hibbard, B. & Clark, T. (2004) Prairie grasses as hosts of the western corn rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology*, **33**, 740-747.

Parimi, S., Meinke, L.J., French, B.W., Chandler, L.D., & Siegfried, B.D. (2006) Stability and persistence of aldrin and methyl-parathion resistance in western corn rootworm populations (Coleoptera: Chrysomelidae). *Crop Protection*, **25**, 269-274.

Pruess, K., Witkowski, J.F., & Raun, E.S. (1974) Population suppression of western corn rootworm by adult control with ULV malathion. *Journal of Economic Entomology*, **67**, 651-655.

Pruess, K., Weekman, G. & Somerhalder, B. (1968) Western corn rootworm egg distribution and adult emergence under two corn tillage systems. *Journal of Economic Entomology* **61**, 1424-1427.

Reisig, D. & Bacheler, J. (2002) Kudzu Bug (*Megacopta cribraria*), a New Potentially Devastating Pest of Soybeans. Available online: https://organicgrains.ces.ncsu.edu/2012/12/kudzu-bug-megacopta-cribraria-a-newpotentially-devastating-pest-of-soybeans/

Riedell, W.E. & Reese, R.N. (1999) Maize morphology and shoot CO2 assimilation after root damage by western corn rootworm larvae. *Crop Science*, **39,** 1332-1340.

Riedell, W.E. (1993) Advances in understanding corn rootworm damage effects on maize physiology. Proc. Washington, DC: *Annual Corn Sorghum Research Conference*, **22,** 76-90.

Ruberson, J.R., Takatsu, K., Buntin, G.D., Eger, J.E., Jr., Gardner, W.A., Greene, J.K., Jenkins, T.M., Jones,W.A., Olson, D.M. & Roberts, P.M. (2013) From Asian curiosity to eruptive American pest: *Megacopta cribraria* (Hemiptera: Plataspidae) and prospects for its biological control. *Applied Entomology Zoology*, **48**, 3–13.

Sappington, T., Siegfried, B. & Guillemaud, T. (2006) Coordinated *Diabrotica* genetics research: Accelerating progress on an urgent insect pest problem. *American Entomologist* **52**, 90-97.

Seiter, N.J., del-Pozo Valdivia, A.I., Greene, J.K., Reay-Jones, F.P.F., Roberts, P.M., Reisig, D.D. (2016) Management of *Megacopta cribraria* (Hemiptera: Plataspidae) at different stages of soybean (Fabales: Fabaceae) development. *Journal of Economic Entomology*, **109**, 1167–1176.

Seiter, N.J., del-Pozo Valdivia, A.I., Greene, J.K., Reay-Jones, F.P.F., Roberts, P.M. & Reisig, D.R. (2015) Action thresholds for managing *Megacopta cribraria* (Hemiptera: Plataspidae) in soybean based on sweep-net sampling. *Journal of Economic Entomology*, **108**, 1818–1829.

Seiter, N.J., Greene, J.K. & Reay-Jones F.P.F. (2014) Aggregation and oviposition preferences of *Megacopta cribraria* (Hemiptera: Plataspidae) in laboratory bioassays. *Journal of Entomological Science* **49,** 331–335.

Seiter, N.J., Greene, J.K., & Reay-Jones, F.P. (2013) Reduction of soybean yield components by *Megacopta cribraria* (Hemiptera: Plataspidae). *Journal of Economic Entomology*, **106**, 1676–1683.

Suiter, D. R., Eger, J.E. Jr., Gardner, W.A., Kemerait, R.C., All, J.N., Roberts, P.M., Greene, J.K., Ames, L.M., Buntin, G.D. & Jenkins, T.M. (2010) Discovery and distribution of *Megacopta cribraria* (Hemiptera: Heteroptera: Plataspidae) in northeast Georgia. *Journal of Integrated Pest Management*, **1**, 1–4.

Spike, B.P. & Tollefson, J.J. (1991) Yield response of com subjected to western com rootworm (Coleoptera: Chrysomelidae) infestation and lodging. *Journal of Economic Entomology*, **84**, 1585-1590

Stubbins, F. L., Agudelo, P., Reay-Jones, F.P.F. & Greene, J. K. (2015) First report of a mermithid nematode infecting the invasive *Megacopta cribraria* (Hemiptera: Plataspidae) in the United States. *Journal of Invertebrate Pathology,* **127,** 35–37.

Tayutivutikul, J. & Yano, K. (1990) Biology of insects associated with the kudzu plant, Pueraria lobata (Leguminosae) *Megacopta punctissimum* (Hemiptera, Plataspidae). *Japanese Journal of Entomology* **58**, 533–539.

Thippeswamy, C., & Rajagopal, B.K. (2005) Comparative biology of *Coptosoma cribraria* Fabricius on field bean, soybean, and redgram. *Karnataka Journal of Agricultural Science*, **18**, 138–140.

United State Department of Agriculture, Economic Research Service. Farm Income and Wealth Statistics. http://www.ers.usda.gov/data-products/ farm-income-and-wealth-statistics.aspx

Vanbergen, A.J. & the Insect Pollinators Initiative (2013) Threats to an ecosystem service: pressures on pollinators-REVIEWS. *Front Ecol Environ*, 11, 251–259.

Waldvogel, M., & Alder, P. (2012) Kudzu bug: A nuisance and agricultural pest. (http://www.ces.ncsu.edu/depts/ent/notes/Urban/kudzubug.htm)

Wang, H., Coates, B., Chen, H., Sappington, T., Guillemaud, T. & Siegfried, B. (2013) Role of a gama-aminobutyric acid (GABA) receptor mutation in the evolution and spread of *Diabrotica virgifera virgifera* resistance to cyclodienes insecticides. *Insect Molecular Biology* **22**, 473-484.

Wangila, D.S., Gassmann, A.J., Petzold-Maxwell, J.L., French, B.W. & Meinke, L.J. (2015) Susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to *Bt* corn events. *Journal of economic entomology*, **108**, 742–751.

Woodson, W.D. & Jackson, J.J. (1996) Developmental rate as a function of temperature in northern corn rootworm (Coleoptera: Chrysomelidae). *Annual Entomology Society*, **80**, 226-230

Zhang, Y., Hanula, J.L. & Horn, S. (2012) The biology and preliminary host range of *Megacopta cribraria* (Heteroptera: Plataspidae) and its impact on kudzu growth. *Environmental Entomology*, **41**, 40–50.

CHAPTER 2

Use of RNA Interference in Agriculture

INTRODUCTION

As the global population increases, the demand for food also increases. Insect damage is one of the major problems faced globally and can substantial reduce both pre- and post-harvest production if left untreated (Oerke, 2006; Fox and Fimeche, 2013). While insect pests are controlled through various methods, including insecticide applications, heavy or improper use can lead to major problems. One obvious issue is the evolution of insecticide resistance. For example, over 100 insect species are now known to be resistant to one or more insecticides (Elzen and Hardee, 2003). Another issue is the impact insecticides have on human health, and beneficial insects, including predators, parasitoids, and pollinators (Vanbergen et al., 2013). Due to indirect effects on human health and the environment it is impossible to know the total economic impact these control measures have. However, unlike conventional insecticides, new technologies based on RNA interference (RNAi) appear to be species-specific. In fact, RNAi can be designed to target a single pest species, while sparing beneficial ones. This is due to RNAi's sequence-dependent mode of action (Zhang et al., 2015; Chen et al., 2015). As long as the DNA sequence of the target gene is unique (significantly different from that found in beneficial species) the double-stranded RNA (dsRNA) should only impact the pest it was designed to target. However, RNAi-based insecticides can also be designed to target a broad range of insects, either by choosing more conserved sequences (Powell et al., 2016; Baum et al., 2007), or by deploying multiple dsRNAs at once. The active ingredient is RNA, therefore, is organic, biodegradable, and could be more economical than traditional insecticides if produced by microorganisms (e.g., yeasts). Clearly, there is a need for alternative management strategies that are sustainable and environmentally friendly for controlling agricultural pests, and RNAi may be such an approach.

RNAi is a natural process cells use to control gene expression. It is thought to have evolved from an early defense mechanism against viruses, where double-stranded RNA (dsRNA) were recognized as foreign to the cell and destroyed along with mRNA, and other single-stranded RNAs. Such destruction leads to the post-transcriptional inhibition of gene expression and protein synthesis. The power of this technology lies in its molecular mechanism, specifically, its ability to destroy only transcripts from those genes it is designed to target. The introduction of dsRNA that match part of a target gene’s sequence into the cell will trigger an RNAi response causing the destruction of the homologous mRNA and a loss-of-function phenotype (Castel and Martienssen 2013; Ketting, 2011). This review provides background on naturally occurring RNAi, how it was discovered, and the potential RNAi holds for use in pest control with a focus on the use of RNAi in agriculture.

The reason researchers are fascinated by the potential use of RNAi for insect control, is the fact that it is species-specific (Baum et al., 2007; Whyard et al., 2009; Bellés, 2010; Terenius et. al., 2011), can be delivered by a variety of methods (Baum et al., 2007; Scott et al., 2013; Terenius et al., 2011), and can possibly serve as an eco-friendly pest management tool that can be modified for use in a wide range of insect pests (Borovsky, 2005; Gordon and Waterhouse, 2007; Price and Gatehouse, 2008). Unlike conventional pesticides, RNAi appears to have minimal off-target effects (Birmingham et. al., 2006).

RNAi DISCOVERY

Prior to its discovery (Fire et al., 1998), researchers observed and reported that RNA inhibited protein expression in plants and fungi (Napoli et al. 1990; Romano and Macino 1992). Napoli et al., (1990) wanted to deepen the pigment in petunia petals by introducing a transgene, *chalcone synthase* (*CHS*) known to have an important role in floral pigmentation. However, rather than finding petunias with darker color, they found plants with white flowers or ones having patterns of white on a wild-type pigmented background- phenotypes not seen in the control plants. This result was unexpected, so they termed this phenomenon co-suppression as the transgene introduced created a block in anthocyanin biosynthesis. Romano and Macino (1992) discovered a similar thing in a fungus, *Neurospora crassa*, a process they called quelling. In this case, the transformation was carried out with constructs of the *albino* gene (*Al-3*) known to be involved in the carotenoid biosynthetic pathway. In this study, exogenous sequences were introduced into ectopic locations, about 36% of *N. crassa* transformants showed an albino phenotype. A few years later, Guo and Kemphues (1995) reported mRNA silencing by sense and antisense RNA in the roundworm, *Caenorhabditis elegans*. They observed in their studies that the introduction of a *proteinase-activated receptor* gene (*par-1*) known to play an important role in germline development resulted in degradation of the *par-1* message. While all of these processes resulted in gene silencing, their mechanism was not fully understood. It wasn't until Fire and Mello's landmark publication (1998) describing the effect of injecting dsRNA into *C. elegans* that the cause of this phenomenon was fully recognized. Fire and Mello were later awarded the Nobel Prize in Medicine for this discovery.

Since then three classes of noncoding (nc) RNAs have been identified and reported to be involved in post-transcriptional regulation of gene transcripts (Koch and Kogel 2014; Mercer, 2009). MicroRNAs (miRNAs) are known to interfere with mRNA translation (Cartel and Sontheimer, 2009; Yeo and Chong, 2011), and PIWI-interacting RNAs (piRNAs) known as endogenous small noncoding RNAs act as guardians in the genome, by protecting it from invasive transposable elements in the germline (Siomi et al., 2011; Thomson and Lin, 2009).

Moreso, long noncoding RNAs (lncRNA) (Mercer et al., 2009) have been identified, but their specific function is still unknown (Kaikkonen et al., 2011).

Since its discovery in 1998, RNAi has played an essential role in functional genetics studies. It has been used in many model systems including *Drosophila melanogaster* (Roignant et al., 2003; Bischoff et al., 2006; Miller et al., 2008), *Tribolium castaneum* (Tomoyasu and Denell, 2004; Fujita et al., 2006; Arakane et al., 2008; Konopova and Jindra, 2008; Minakuchi et al., 2009; Parthasarathy and Palli, 2009) and *Bombyx mori* (Quan et al., 2002; Ohnishi et al., 2006; Hossain et al., 2008) to study loss-of-function phenotypes for a wide-array of genes. Such functional genomic studies provide researchers with a greater understanding of the genes involved in biological phenomena like physiology, embryology, reproduction and behavior in both model and non-model insects (Bellés, 2010).

In addition to the benefit of using RNAi as a tool for assessing a gene's role in biological processes, there are also many other applications for RNAi including medical therapy for viral infection (Shankar et al., 2005), and genetic diseases (Shankar et al., 2005); however, discussion of RNAi technology in medical field is beyond the scope of this review. Here I will focus on the use of RNAi in agriculture, and specifically on its use in pest management (Baum et al., 2007; Mao et al., 2007).

RNAi MECHANISMS

Since the discovery of RNAi, significant efforts have been made to describe the molecular mechanisms of RNAi (Aravin, et al., 2007). Studies performed on several model insects, including *D. melanogaster, T. castaneum* and the silkworm, *Bombyx mori* whose genomes have been sequenced revealed that the RNAi pathway is composed of two phases, the initiator phase, and the effector phase, and together they result in post-transcriptional gene silencing (PTGS) Vaucheret et al., 2001; Vaucheret and Fagard, 2001; Castel and Martienssen, 2013). The initial step starts with the introduction of long dsRNA into a cell. These long dsRNA molecules are cut into small interfering RNAs (siRNA) by an enzyme called Dicer (Shreve et al., 2013; Swevers et al., 2013; Hammond et al., 2000). Dicer encodes a multidomain protein containing an ATP-dependent RNA helicase, PAZ domain, two tandem RNase III domains, and a dsRNA-binding domain (Bernstein et al., 2001; Knight and Bass 2001). The 21- to 25-bp products of Dicer activity are referred to as short interfering RNAs (siRNAs) (Hamilton and Balcombe, 1999). They are thought to serve as "guides" to bring nuclease machinery to the target mRNA; each siRNA associates with a protein complex called RNA-induced silencing complex (RISC), which then binds to one of the strands in the RNA fragments, allowing it to search for similar or identical mRNA (Hamilton and Balcombe, 1999). For this to occur, the unincorporated sequence must be cleaved from the siRNA duplex by the RNase H activity of an Argonaute (Ago) protein (Matranga et al., 2005; Rand et al., 2005). If the incorporated sequence is the antisense strand, then it guides RISC to its homologous target mRNA where the activity of the Ago protein causes the destruction of the target mRNA (Hammond et al. 2001; Song et al., 2004).

In addition to the mechanism described above, RNAi components have also been shown to play a role in two other processes; (1) post-transcriptional silencing (PTGS) through the action of microRNAs (miRNAs) (Ouellet et al., 2006), and transcriptional gene silencing (TGS) by chromatin modification (Niwa and Slack 2007). Regardless of whether PTGS or TGS, gene silencing by RNAi pathway involves the production of small RNAs utilizes Ago family members (Zamore 2009). The main components of RNAi are well conserved across phyla, however, the mechanism and systemic spread of silencing are highly diverged and gene regulation is mediated by a variety of RNA-based products (siRNAs, miRNA) (Shreve et al., 2013; Swevers et al., 2013). RNAi is also likely to have significant implications for the biological role of heterochromatin and genome maintenance (Lippman and Martienssen 2004). Finally, miRNAs play important roles in development and fundamental cellular processes, influencing the expression of an estimated 30 percent of all protein coding genes (Ouellet et al., 2006).

SYSTEMIC RNAi

In some genetic model organisms, the application of RNAi as a tool is less problematic, because the cells have the ability to up-take dsRNA from the extracellular environment and in some cases spread the dsRNA (or its signals) to neighboring cells (Yu et al., 2012; Voinnet and Baulcombe 1997, Fire et al., 1998). This process is referred to as systemic RNAi (Huvenne and Smagghe, 2010; Whangbo and Hunter, 2008). The phenomenon that the RNAi effect could be seen throughout the organism regardless of the site of injection was first discovered in *C. elegans* (Fire et al., 1998). Later, it was also discovered that a systemic RNAi response could be triggered when *C. elegans* were soaked in a solution of dsRNA (Tabara et al., 1998) or fed transgenic *Escherichia coli* expressing dsRNA (Timmons and Fire 1998).

Studies on the systemic RNAi mechanism in insects are very important as they broaden our knowledge on RNAi for pest control. In *C. elegans*, two types of dsRNA uptake mechanisms have been identified and studied (Whangbo and Hunter, 2008). The first is mediated by a transmembrane protein called SID-1 (systemic-interference-defective). It is believed to act as a passive dsRNA channel that shows a preference for long dsRNA molecules (Hunter et al., 2006). In *C. elegans*, Sid-1 is essential for systemic RNAi and plays a role in dsRNA uptake and subsequent spreading of the RNAi signal in both somatic and germ-line cells (Winston et al., 2002). The second method uses a protein known as SID-2, which is mainly found in the intestinal tissue of the roundworm and facilitates environmental RNAi; it may induce receptor-mediated endocytosis of dsRNAs, delivering them to the cytoplasm where they trigger RNAi pathway (Whangbo and Hunter, 2008).

Several other genes such as *RNA spreading defective* (*Rsd-2, Rsd-3, and Rsd-6*) are also known to play a role in the systemic spread of dsRNA in *C. elegans*, which was previously attributed to endocytosis, suggesting that transport of dsRNA in *C. elegans* could also use a Sid-1-independent mechanism. However, in insects, Sid-1-like (sil) proteins vary in copy number and sequence conservation across orders (Gu, et.al., 2012). For example, Diptera do not possess *sil* genes, while *T. castaneum* and *B. mori* each have three sil homologues (Tomoyasu et al., 2008). The *T. castaneum sil* genes share more identity with *C. elegans TAG-130* gene, which is not required for systemic RNAi compared to SID-1 (Tomoyasu et al., 2008). However, the relationship between *sid-1* and *sil* genes is unknown, as is the relationship between *sil* genes and systemic RNAi (Tomoyasu et al., 2008). This is partially due to the lack of one-to-one orthology between *sid-1* and *sil* genes. However, the conclusion so far is that the *sil* genes are functional *sid-1* orthologs.

Hence, flour beetles and honeybees have systemic RNAi, while dipterans, which lack *sil* genes have poor systemic RNAi. The existence of a Sid-2-mediated mechanism is also questionable, as *sid-2* orthologs have not been found in any animal genomes other than *C. elegans*, but this may be due to the rapid evolution of *sid-2* homologs. Furthermore, the presence of *Rsd* homologs does not seem to determine whether systemic RNAi occurs in insects; it is possible that the expression level and tissue specificity of these genes are responsible for the RNAi efficiency and the dsRNA uptake from the cell (Tomoyasu et al., 2008). While systemic RNAi has been documented in several insects (Bucher et al., 2002; Tomoyasu and Denell 2004), it is assumed that all animals with systemic RNAi response use the mechanism described in *C. elegans*. However, it is clear that the systemic RNAi response is not universal, as the leading insect model, *Drosophila melanogaster*, lacks a systemic RNAi response (Roignant et al., 2003).

METHODS OF dsRNA DELIVERY TO INSECTS

Several methods have been employed to deliver dsRNAs to insects, these include microinjection, feeding, and soaking (Scott et al., 2013; Yu et al., 2012). Determining which method to use one must consider the end goal (e.g. pest control), the insect species and/or life stage, as well as when and where the target gene is expressed. Almost all entomological RNAi experiments rely on extracellular dsRNA, whereby the dsRNAs are delivered to the hemolymph or midgut so that the cells can take up the dsRNAs (Yu et al., 2012). However, the delivery of dsRNA into individual cells of a multicellular organism has been problematic, because some species lack systemic RNAi, therefore posing challenges for the use of RNAi as a pest control tool (Zotti and Smagghe 2015).

Microinjection was the first method used for dsRNA delivery in both *C. elegans* (Fire et al., 1998) and *D. melanogaster* (Kennerdell and Carthew, 1998). In insects, this method delivers dsRNAs corresponding to the target gene directly into the hemolymph or embryo, and has advantages over other methods. In some organisms, it is possible to deliver dsRNA directly by injecting dsRNA into any stage (Kennerdell and Carthew 1998; Brown et al., 1999), unlike oral RNAi where dsRNA can only be administered to larvae, nymphs and adults.

Secondly, oral RNAi enables delivery of dsRNAs in a high-throughput manner, while avoiding the mechanical damage caused by microinjection (Scott et al., 2013). In addition, RNAi delivered by feeding is of particular interest for insect control in agriculture as they can act as a species-specific insecticide (Baum et al., 2007). The first study on oral RNAi in an insect was demonstrated in a lepidopteran, the light brown apple moth, *Epiphyas postvittana*, with temporary knockdown of a gut gene, *Carboxylesterase 1* (*EposCXE1*) and a pheromone binding protein gene (*EposPBP1*) by feeding larvae *EposCXE1 and EposPBP1* dsRNA (Turner et al., 2006). Since then, several studies have reported a knockdown through oral RNAi in Lepidoptera (Bautista et al., 2009; Mao et al., 2007; Whyard et al., 2009), Coleoptera (Baum et al., 2007; Whyard et al., 2009), Diptera (Walshe et al., 2009; Whyard et al., 2009), Hemiptera (Price & Gatehouse, 2008), and Isoptera (Zhou et al., 2008), suggesting that knockdown of specific targets may be feasible in some insects.

Feeding insects genetically modified bacteria (microbe-based RNAi) that express dsRNA targeting host genes has been documented in various species to be effective (Taracena et al., 2015). While it requires several steps, it has been shown to be very effective in insects that exhibit a robust RNAi response (Zhu et al., 2011). In insects, Murphy et al., (2016) were the first group that proved yeast can be genetically modified to deliver dsRNA to an insect pest through feeding. Because yeast is naturally growing on the surface of fruit crops and are known to be attractive to Drosophila species, this group took advantage of the microbe's relationship with Drosophila to create a yeast-based biopesticide. The yeast expressed dsRNA targeting *y-Tubulin*, required for nuclear division and microtubule assembly, and was fed to *Drosophila suzukii*. The authors found a reduction in population size, as well as in reproductive fitness. dsRNA has also been delivered to larval mosquitos through ingestion using nanoparticles that is composed of chitosan (Zhang et al., 2010). Zhu et al., (2011) reported that bacteria expressing dsRNA fed to insects induce RNAi effect (Tian et al., 2009). Delivering dsRNAs in bacteria and yeast will likely increase the environmental stability of the dsRNAs, while lowering the overall production cost. While dsRNA synthesis in vitro could be expensive, dsRNA synthesis and delivery via bacteria and yeast are more practical for field applications (Zhu et al., 2011). Some crops have been genetically modified to express dsRNA targeting a vital gene within a given pest species (Baum et al., 2007; San Miguel and Scott 2015). Baum (2007) was the first study to demonstrate the potential for RNAi in insect control via feeding dsRNAs expressed in planta that targeted essential insect genes.

RNAi by soaking is mainly conducted in cell lines. However, it has been suggested to be valuable as a high-throughput tool for large-scale gene expression studies in *C. elegans* (Maeda et al., 2001). While all of these methods have been successful, they all have certain disadvantages. Microbe-based RNAi methods of dsRNA delivery and chemically modified siRNAs do not provide continuous expression, so silencing is only for a short period. However, transgenic approaches could allow for continuous expression, and therefore silencing can be sustained.

IDENTIFICATION OF TARGET GENES

Since RNAi is species specific, appropriate target genes must be identified for each new species to be controlled by this strategy. One way to identify potential target genes is through homology to ones used in previous studies (She et al., 2009; Lorenzen et al., 2002). For the purpose of pest control, scientists have been using RNAi to target vital genes such as *Vacuolar-type H+ -ATPase* (*V*-ATPase) in a variety of pest species, including the western corn rootworm, *Diabrotica virgifera virgifera* and the Colorado potato beetle, *Leptinotarsa decemlineata* (Baum et al., 2007; Mao et al., 2015). Not only is *V*-ATPase necessary for survival, performing a wide range of functions including the transport of protons across membranes, but it is also highly conserved, making it relatively easy to identify in genomic or transcriptomic sequence data. As mentioned earlier, the successful use of RNAi for pest control rests on the ability of ingested dsRNAs to trigger gene knock down. Since V-ATPase plays an essential role in the insect midgut, keeping an alkaline environment in the lumen (Wieczorek et. al., 2009), dsRNAs don't have to travel far to cause serious disruption of essential functions.

Another common target is a gene that encodes a vacuolar-sorting protein (*Snf7*) (Rusten 2008) Snf7 is a component of the ESCRT-III complex (endosomal sorting complex required for transport), and like V-ATPase, is highly conserved due to its conserved function (i.e. required for sorting of cell membrane receptors (Rusten 2008). As a result of their involvement in biological processes, Bolognesi et al., (2012) characterized the ortholog of WCR larvae *DvvSnf7* gene, which lead to a significant high mortality in targeted species even at low concentration. And as mentioned earlier, Murphy and colleagues (2016), successfully knockdown *y-Tubulin* when fed genetically modified yeast that express dsRNA targeting *y-Tubulin* to *Drosophila suzukii.* Their RNAi analysis showed a reduction in population size as well as reproductive fitness in adults. A result that is not surprising since *y-Tubulin* is responsible for nuclear division and microtubule assembly. All of the above-mentioned genes will function well as targets for RNAi-based control (Baum et al., 2007).

An added advantage of using highly conserved genes is that they can be easily recognized in database searches. BLAST analysis of NCBI and other genomic and transcriptomic databases can reveal potential target genes for a wide number of species; however, if the target species lacks genomic and transcriptomic data, these sequences can be easily obtained by extracting RNA and performing RNA-Seq. This type of next generation sequencing is a rapid method of obtaining a wealth of genetic data and allows identification of novel transcripts; especially in non-model insects that lack a sequenced genome (Wang et. al., 2011). Homologous genes can be quickly identified based on identity to known targets (e.g. V-ATPase), and offers an easy way to select candidate genes for RNAi analysis. In addition, publically available sequence data, such as that found at NCBI can be used to determine if a particular target sequence is unique, not expected to harm other species.

FACTORS AFFECTING THE EFFICIENCY OF RNAi IN INSECT CONTROL

According to the review literature, there are several factors known to influence the silencing effect and efficiency of RNAi as a control tactic. As detailed below, some factors known to have a sizable impact on the efficiency of RNAi are: dsRNA concentration, sequence composition, length of sequence, life stage and in the case of protein-coding genes – rate of target proteins' turnover rate.

For each RNAi target, the optimal dsRNA concentration has to be determined (Meyering-Vos and Muller, 2007; Shakesby et al., 2009). For example, exceeding the optimal dsRNA concentration does not result in more silencing; however, it can decrease the exposure time required to get a phenotype (Bolognesi et. al., 2012, Saleh et. al., 2006), suggesting a relationship between dsRNA concentration and exposure time.

The dsRNA sequence of a particular gene could possibly silence a gene in a non-target species. Such off target effects can be minimized through database search as mentioned above; however, in the absence of genomic resources, lab experiments would be required to determine off-target effects. Studies have reported an off-target silencing in the triatomid bug *R. prolixus,* when two homologous *nitroporin* genes were silenced (Araujo et al., 2006). In addition, dsRNA targeting *V-ATPase* in the Colorado potato beetle *(Leptinotarsa decemlineata*) also silenced the orthologous gene in *D. virgifera virgifera* (Baum et al., 2007). Although this result implies that

RNAi could be used to control two or even more coleopteran pests, it could also affect non-target insects leading to a biosafety issue.

The length of the dsRNA fragments plays an essential role and determines the efficiency of uptake and gene silencing in organisms (Mao et al., 2007). The length of the dsRNA is directly involved in the success of gene silencing, and frequently ranges between 300 and 520bp. In most experiments, insects are fed long dsRNAs (Huvenne and Smagghe 2010). While RNAi can be triggered by a dsRNA as short as 50 nt in length, some experiments have shown that long dsRNAs are more efficiently taken up compared to siRNAs (Saleh et al. 2006, Bolognesi et. al., 2012). This could be due to the fact that a long dsRNA with the highest match of mRNA can be processed into siRNA. Therefore, giving a higher identity of siRNAs to cause specific silencing of the target gene (Auer and Frederick 2009). Studies have reported silencing of genes in varieties of insects by incorporating siRNA into diet instead of dsRNA (Kumar et al., 2009, Upadhyay et al., 2011).

The silencing effect of RNAi targeting *aquaporin* (*AQP2*) in aphid (*Acyrthosiphon pisum* pea) persists for five days and is then reduced (Shakesby et al., 2009). A similar effect was reported by Turner et al., (2006), where a transient effect of dsRNA against the pheromone binding protein in the light brown apple moth (*Epiphyas postvittana*) was thought to be correlated with the turnover rate of the target protein.

Early stage juvenile insects show more silencing effects than later stages (Araujo et al., 2006); however, older stages are easier to handle. For instance, no silencing effect was observed after treating 4th instar nymphs of *Rhodnius prolixus* with *nitropin* dsRNA compared to 42% silencing when using 2nd instar larvae (Araujo et al., 2006). Also, in the case of the fall armyworm (*Spodoptera frugiperda*), a stronger silencing effect was observed in 5th instar larvae when treated with allatoregulating genes (*AT 2*) compared to adult moths (Griebler et al., 2008).

PROSPECTS OF RNAi IN INSECT CONTROL

After a decade of using RNAi as a method to study the function of various genes, there is now significant growth in the use of RNAi for the control of insect pests (Baum et al., 2007; Mao et al., 2007). To date, RNAi has been used in more than 30 species, from eight insect orders (see Table 1). Most of the published studies on RNAi have focused on injection of dsRNA, with few published reports looking at the effectiveness of feeding dsRNA. However, successful feeding of dsRNA has been documented in several insect species including, *Diabrotica virgifera virgifera* (Chitvan et al., 2015; Baum et al., 2007), *Leptinotarsa decemlineata* (Zhu et al., 2010), *Epiphyas postvittana* (Turner et al. 2006*), Aedes aegypti* (Singh et al. 2013), *Helicoverpa armigera* (Mao et al., 2007) and *Drosophila suzukii* (Murphy et al., 2016). Although, Murphy et al., (2016) reported that ingestion of yeast expressing dsRNA targeting *D. suzukii yTub23C* during the larval stage decreased survival to adulthood, a closely related species, *Drosophila melanogaster*, lacks a systemic RNAi response (Roignant et al., 2003). This might explain why Murphy et al (2016) found so much variation in their results. While both injection and feeding have been documented in a variety of insects, there is evidence that in some species, RNAi is not very efficient. Allen and Walker (2011) demonstrated that repeated delivery of dsRNA corresponding to tarnished plant bug, *Lygus lineolaris*, *Inhibitor of Apoptosis* (*IAP*) was not successful. Their further analysis revealed that the dsRNA had been degraded prior to contact with *Lygus lineolaris* cells. So, while this technology has proved promising in some insect species, others have less successful stories.

In 2007, two studies reported the concept of plants expressing dsRNA from hairpin vectors that targeted economically important agricultural pests (see Table 2), the cotton bollworm, *Helicoverpa armigera* (Mao et al., 2007), and western corn rootworm, *D. virgifera virgifera* (Baum et al., 2007). Baum et al., (2007) demonstrated that silencing of a *V-ATPase* gene in the midgut cells of the western corn rootworm protected the plant and led to larval stunted growth and mortality. This was the first study to show that RNAi can be an effective insect control method. Similar results were obtained when the *cytochrome P450 monooxygenase (CYP6AE14)* gene, known to help insects detoxify plant secondary metabolites, was silenced in midgut cells of the cotton bollworm. In this case, the insect's ability to detoxify these compounds was drastically reduced (Mao et al., 2007). Cotton plants engineered to produce dsRNAs targeting *CYP6AE14* were fed to larvae, and successfully decreased endogenous *CYP6AE14* mRNA levels in the insects, and resulted in stunted larval growth due to the presence of gossypol in cotton plants (Mao et al., 2007). Mao et al. (2013), show that co-expressing *CYP6AE14* dsRNA with a plant cysteine protease, and enzyme that damages the insect's peritrophic matrix- a barrier which prevents large molecules from getting into the midgut cells, could protect the plant from bollworms. A similar approach has been exploited in plants, with the aim of developing plants that are more resistant to nematodes. Engineered plants producing dsRNA targeting genes in the root knot nematode have been reported (Yadav et al., 2006; Huang et al., 2006). Arabidopsis expressing a dsRNA targeting *CYP9A14*, when silenced, reduced larval resistance to deltamethrin, a pyrethroid insecticide that can protect cotton from insect attack (Tao et al., 2012). In a related study, injection of dsRNAs targeting *CYP6B7* into the midgut of resistant cotton bollworm larvae restored susceptibility to the pyrethroid insecticide deltamethrin (Tang et al., 2012). These studies suggest that targeting detoxification genes such as those in the cytochrome

P450 family could serve as an effective method to increase the lifespan of insecticide chemistries that are at, or near the end of their effectiveness due to resistance development.

CONCLUSIONS AND FUTURE PERSPECTIVES

RNAi has potential for becoming a successful approach for insect pest management. However, several issues need to be addressed before this technology can be used in the field. Although RNAi is known for its specificity (Bellés, 2010), studies have shown that off-target gene silencing can occur (Xu et al., 2006). The specific case of a single dsRNA causing knock down of a specific target in two pest species, such as *Diabrotica virgifera* and *Leptinotarsa decemlineata* (Baum et al., 2007) is not a problem; however, if a pest-specific target triggers RNAi in a non-target species, this will likely have a negative effect on RNAi as a pest control method. Moreover, RNAi targeting a particular insect pest could have an adverse impact on beneficial organisms. Therefore, species-specific biology such as genes involved in mate recognition might be good targets for pest management, since this would reduce the possibility of silencing genes in a non-target species. Extensive studies are needed to completely understand the factors behind off-target silencing and to prevent potential hazards to the environment.

Similar to insecticides and *Bt* plants, there is also some possibility that some insects may evolve resistance to a particular dsRNA sequence due to mutations in the target gene. In addition, resistance due to a mutation in Dicer or another part of the RNAi pathway could be a huge hindrance for RNAi as a pest management tool (Carthew and Sontheimer 2009). Even creating another dsRNA targeting a different portion of the same gene or a new target gene might not overcome this resistance barrier. Therefore research into RNAi resistance should be investigated before this technology is accepted on a large scale.

To further our understanding of the dsRNA uptake and processing mechanisms in insects, RNAi pathways in different insect species and orders need to be identified and their roles studied. Differences, such as what makes *Tribolium castaneum* have strong systemic RNAi, when *Drosophila* appears deficient in systemic RNAi needs to be understood (Roignant et al., 2003). Likewise we need to understand why *Drosophila*, which is not responsive to environmental dsRNA when unpackaged, has significant gene knockdown and high mortality when fed dsRNAs in transfection solution (Whyard et al. 2009), or in the case of *D. suzukii*, is responsive to dsRNA when expressed in yeast (Murphy et al., 2016). Taken together this suggests that the presence of a robust systemic RNAi response may not be a precondition for RNAi-mediated insect control. Therefore, a better understanding of how dsRNA gets into midgut cells might increase our ability to improve oral RNAi in other insect species. Despite various setbacks associated with this technology, perhaps in the future, RNAi will be a widely-accepted tool for used in integrated pest management.

Table 1: Overview of oral and injection-based RNAi against a number of target genes in several species

ORDER	SPECIES	TARGET GENE	STAGE	REFERENCES
Coleoptera	*Aethina tumida*	*Vacuolar ATPase A*[1]	Larvae	Powell et al., 2016
Coleoptera	*Tribolium castaneum*	*Distal-less*[1]	Adult	Bucher et al., 2002
Coleoptera	*Leptinotarsa decimlineata*	*Vacuolar ATPase E*[2]	Larvae	Zhu et al., 2010
Coleoptera	*Diabrotica virgifera virgifera*	*Hunchback*[1, 2], *brahma* [1, 2]	Larvae	Chitvan et al., 2015
Coleoptera	*Diabrotica undecimpunctata*	*Subunits of vacuolar ATPase*[2]	Larvae	Baum et al., 2007
Coleoptera	*Harmonia axyridis*	*Distal-less* [1]	Larvae	Niimi et al., 2005
Coleoptera	*Protaetia brevitarsis*	*Pbtf* [1]	Larvae	Kim et al., 2008
Diptera	*Drosophila suzukii*	*Y-Tubulin* [2]	Adult	Murphy et al. 2016
Diptera	*Anopheles gambiae*	*Defensin* [1]	Adult	Blandin et al. 2002
Diptera	*Aedes aegypti*	*Chitin*[1]	Adult	Kato et al. 2006
Diptera	*Culex pipiens*	*Foxo*[1]	Adult	Sim & Denlinger 2008
Diptera	*Drosophila melanogaster*	*Beltless* [1]	Adult	Dzitoyeva et al., 2003
Diptera	*Sarcophaga peregrina*	*P120-catenin* [1]	Larvae	Nashikawa & Natori 2001
Diptera	*Lutzomyia longipalpis*	*Xanthine dehydrogenase*[1]	Adult	Sant' Anna et al., 2008
Diptera	*Glossina morsitans*	*Transferrin* [1, 2]	Adult	Lehane et al., 2008
Lepidoptera	*Spodoptera litura*	*Aminopeptidase* [1]	Larvae	Rajagopal et al. 2002
Lepidoptera	*Epiphyas postvittana*	*Carboxylesterase 1*[2]	Larvae	Turner et al. 2006
Lepidoptera	*Manduca sexta*	*Vacuolar ATPase E*[1]	Larvae	Whyard et al., 2009
Lepidoptera	*Hyalophora cecropia*	Peptidoglycan recognition protein [1]	Pupae	Bettencourt et al. 2002
Lepidoptera	*Bombyx mori*	Creb binding protein [1]	Larvae	Tabunoki et al., 2004
Lepidoptera	*Helicoverpa armigera*	*Cytochrome P450* [2]	Larvae	Mao et al., 2007
Hemiptera	*Oncopeltus fasciatus*	*Hunchback* [1]	Adult	Liu and Kaufman 2004
Hemiptera	*Rhodnius prolixus*	*Nitrophorin 2* [1, 2]	Nymphs	Araujo et al., 2006
Hemiptera	*Bemisia tabaci*	*Chickadee* [1]	Adult	Ghanim et al., 2007
Hemiptera	*Acyrthosiphon pisum*	*Ap-crt* [1] *& Ap-cath-L* [1]	Nymphs	Jaubert et al., 2007
Orthoptera	*Locusta migratoria*	*Hunchback* [1]	Nymphs	He et al., 2006
Orthoptera	*Gryllus bimaculatus*	*Sulfakinin* [1,2]	Nymphs	Meyering & Muller 2007
Hymenoptera	*Apis mellifera*	*Vitellogenin* [1,2]	Workers	Amdam et al., 2003
Hymenoptera	*Nascoonia vitripennis*	*Epoxide hydrolase* [1]	Pupae	Abdel-latief et al., 2008
Isoptera	*Reticulitermes flavipes*	*Cellulase hexamerin* [1,2]	Workers	Zhou et al., 2006
Blattodea	*Blatella germanica*	*BgEcR-A* [1]	Larvae	Cruz et al., 2006
Blattodea	*Periplaneta americana*	*Engrailed* [1]	Larvae	Marie et al., 2000

[1]dsRNA delivered by injection. [2]dsRNA delivered by feeding

Table 2.O verview of the published studies on the use of plant-based RNAi against different insect pests

ORDER	SPECIES	TARGET GENE	STAGE	REMARKS	REFERENCES
Coleoptera	*Diabrotica virgifera virgifera*	*Vacuolar ATPase*	Larvae	Mortality	Baum et al. 2007
Coleoptera	*Leptinotarsa decemlineata*	*Vacuolar ATPase A & E*	Larvae	Mortality	Zhu et al. 2011
Coleoptera	*Diabrotica undecimpunctata howardii*	*Vacuolar ATPase*	Larvae	Mortality	Baum et al., 2007
Lepidoptera	*Helicoverpa armigera*	*HaEcR*	Larvae	Molting defect	Zhang et al. 2015
Lepidoptera	*Helicoverpa armigera*	*HR3*	Larvae	Deformities	Xiong et. al. 2013
Lepidoptera	*Helicoverpa armigera*	*HaAK*	Larvae	Deformities	Liu et al. 2015
Lepidoptera	*Helicoverpa armigera*	*CYP6AE14*	Larvae	Reduced growth	Mao et al. 2011
Hemiptera	*Myzus persicae*	*MpC001, Rack1*	Nymphs	Progeny reduced	Pitino et al. 2011
Hemiptera	*Bemisia tabaci*	*Hacuolar ATPase*	Adults	Mortality	Thakur et. al. 2014
Hemiptera	*Myzus persicae*	*Serine protease*	Adults	Progeny reduced	Bhatia et al. 2012
Hemiptera	*Myzus persicae*	*Hunchback*	Nymphs	Inhibited reproduction	Mao et. al. 2014

REFERENCES

Allen, M.L. & Walker, W.B. (2012) Saliva of *Lygus lineolaris* digests double stranded ribonucleic acids. *Journal of Insect Physiology*, **58**, 391–396.

Alvarado, A.S. & Newmark, P.A. (1999) Developmental Biology Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc. Natl. Acad. Sci. USA* **96**, 5049–5054.

Amdam, G.V., Simoes, Z.L., Guidugli, K.R., Norberg, K. & Omholt, S.W. (2003) Disruption of vitellogenin gene function in adult honeybees by intra-abdominal injection of double-stranded RNA. *BMC biotechnology*, 3, 1472-6750.

Araujo, R.N., Santos, A., Pinto, F.S., Gontijo, N.F., Lehane, M.J. & Pereira, M.H. (2006) RNA interference of the salivary gland *nitrophorin 2* in the triatomine bug *Rhodnius prolixus* (Hemiptera: Reduviidae) by dsRNA ingestion or injection. *Insect Biochemistry and Molecular Biology*, **36**, 683-693.

Aravin, A. A, Hannon, G. J. & Brennecke, J. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science*, **318**, 761–4.

Auer, C. & Frederick, R. (2009) Crop improvement using small RNAs: applications and predictive ecological risk assessments. *Trends Biotechnol*. **27**, 4-5.

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P., Ilagan, O., Johnson, S., Plaetinck, G., Munyikwa, T., Pleau, M., Vaughn, T. & Roberts, J. (2007) Control of coleopteran insect pests through RNA interference. *Nature Biotechnol*ogy, **25,** 1322–1326.

Bautista, M. & Tanaka, T. (2009) RNA interference mediated knockdown of a cytochrome P450, *CYP6BG1*, from the diamondback moth, *Plutella xylostella*, reduces larval resistance to permethrin. *Insect Biochem Mol Biol*, **39**, 38–46.

Bellés, X. (2010) Beyond Drosophila: RNAi in vivo and functional genomics in insects. *Annual Review of Entomology*, **55**, 11–28.

Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 3–6.

Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., Marshall, W. S. & Khvorova, A (2006) 3' UTR seed matches, but not overall identities, are associated with RNAi off-targets. *Nature Methods,* **3**, 199 – 204.

Blair, C.D., Sanchez-Vargas, I., Franz, A.W.E. & Olson, K.E. (2006) Rendering Mosquitoes Resistant to Arboviruses through RNA Interference. *Microbe*, **1,** 466-470

Blandin, S., Moita, L.F. & Kocher, T. (2002) Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the Defensin gene. *EMBO Reports*, **3**, 852-856.

Bolognesi, R., Ramaseshadri, P., Anderson, J., Bachman, P., Clinton, W. & Flannagan, R. (2012) Characterizing the mechanism of action of double-stranded RNA activity against Western corn rootworm (*Diabrotica virgifera virgifera* LeConte). *Plos One*, **7**, 3-4.

Boutros, M., Kiger, A.A., Armknecht, S., Kerr, K., Hild, M., Koch, B., Haas, S.A., Paro, R., Perrimon, N. & Heidelberg Fly Array Consortium (2004) Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science*, **303**, 832–835.

Brown, S.J., Mahaffey, J.P., Lorenzen, M.D., Denell, R.E. & Mahaffey, J.W. (1999) Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. *Evolution & Development*, **1**, 11-15

Bucher, G., Scholten, J. & Klingler, M. (2002) Parental RNAi in *Tribolium* (Coleoptera). *Current biology*, **12**, 85-86.

Carthew, R.W. & Sontheimer, E.J. (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642–655.

Castel, S.E. & Martienssen, R.A. (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet*. **14,** 100–112.

Chen, A., Zheng, W., Zheng, W. & Zhang, H. (2015) The effects of RNA interference targeting Bactrocera dorsalis ds-Bdrpl19 on the gene expression of rpl19 in non-target insects. *Ecotoxicology,* **24,** 595–603.

Choi, J-H., Kijimoto, T., Snell-Rood, E., Tae, H., Yang, Y. (2010) Gene discovery in the horned beetle *Onthophagus taurus*. *BMC Genomics*, **11**, 703.

Cruz, J., Martín, D., & Bellés, X. (2007) Redundant ecdysis regulatory functions of three nuclear receptor HR3 isoforms in the direct-developing insect *Blattella germanica*. *Mech Dev*., **124,** 180-9.

Dietz, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., Couto, A., Marra, V., Keleman, K. & Dickson, B.J. (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature*, **448,** 151-156.

Eleftherianos, I., Gökçen, F., Felföldi. G., Millichap, P.J., Trenczek, T.E., ffrench-Constant R.H & Reynolds, S.E. (2007). The immunoglobulin family protein Hemolin mediates cellular immune responses to bacteria in the insect *Manduca sexta*. *Cell Microbiology*, **9,** 37-47

Elzen, G.W. & Hardee, D.D. (2003) United States Department of Agriculture-Agricultural Research Service research on managing insect resistance to insecticides. *Pest Manag.,* Sci. **59,** 770–776.

Fairbairn, D.J., Cavallaro, A.S., Bernard, M., Mahalinga-Iyer, J., Graham, M.W. & Botella, J.R. (2007) Host-delivered RNAi: an effective strategy to silence genes in plant parasitic nematodes. *Planta*, **226**, 1525–1533.

Félix, M. (2008) RNA interference in nematodes and the chance that favored Sydney Brenner. *Journal of Biology*, **7**, 34

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., & Mellow, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391,** 806–811.

Fox, T. & Fimeche, C. (2103) "Global food Waste Not, Want Not." ImechE, England and Wales.

Franz, A.W.E., Sanchez-Vargas, I., Adelman, Z.N, Blair, C.D., Beaty, B.J., James, A.A. & Olson, K.E. (2006) Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. Proceedings of the National Academy of Sciences of the United States of America, **10**, 4198-4203

Futahashi, R., Tanaka, K., Matsuura, Y., Tanahashi, M., Kikuchi, Y., & Fukatsu, T. (2011) *Laccase2* is required for cuticular pigmentation in stinkbugs. *Insect Biochemistry and Molecular Biology*, **41**, 191–6.

Gavilano, L.B., Coleman, N.P., Burnley, L.E., Bowman, M.L., Kalengamaliro, N.E., Hayes, A., Bush, L. & Siminszky, B. (2006) Genetic engineering of *Nicotiana tabacum* for reduced nornicotine content. *J. Agric. Food Chem*., **29,** 9071–9078.

Ghanima, M., Kontsedalova, S., & Czosnekb, H. (2007) Short Communication: Tissue-specific gene silencing by RNA interference in the whitefly *Bemisia tabaci* (Gennadius). *Insect Biochemistry and Molecular Biology* **37**,732–738.

Griebler, M., Westerlund, S.A., Hoffmann, K.H., Meyering-Vos, M. (2008) RNA interference with the allatoregulating neuropeptide genes from the fall armyworm *Spodoptera frugiperda* and its effects on the JH titer in the hemolymph. *Insect Physiology*, **54**, 997–1007.

Guo, S., & Kemphues, K.J. (1995) *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative *Ser/Thr* kinase that is asymmetrically distributed. *Cell*, **81,** 611-20.

Gupta, P.K. (2006) RNAi for Crop Improvement. *Current Science* **91**, 1443-1446.

Gu, L., Knipple D.C. (2012) Recent advances in RNA interference research in insects: Implications for future insect pest management strategies. *Crop Protect*. **45,** 36-40.

Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R., & Hannon, G. J. (2001) *Argonaute2*, a link between genetic and biochemical analyses of RNAi. *Science*, **293**, 1146–50.

Hammond, S.M., Bernstein, E., Beach, D., & Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, **404**, 293–6.

Hamilton, A.J. & Baulcombe, D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, **286**, 950–952.

Huang, G., Allen, R., Davis, E.L., Baum, T.J. & Hussey, R.S. (2006) Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conserved and essential root-knot nematode parasitism gene. *Proc. Natl Acad. Sci. USA*, **26,** 14302–14306.

Hunter, CP., Winston, W.M, Molodowitch, C., Feinberg, E.H, Shih, J., Sutherlin, M., Wright, A.J., & Fitzgerald, M.C. (2006) Systemic RNAi in *Caenorhabditis elegans*. Cold Spring Harb Symp *Quant Biol*. **71,** 95-100.

Huvenne, H., & Smagghe, G. (2010) Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: a review. *Journal of Insect Physiology,* **56,** 227–35.

Kaikkonen, M.U., Lam, M.T. & Glass, C.K. (2011) Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc. Res.*, **90,** 430– 440.

Kamthan, A., Chaudhuri, A., Kamthan, M. & Datta, A. (2015) Small RNAs in plants: recent development and application for crop improvement. *Front Plant Sci*., **6,** 208.

Kennerdell, J.R. & Carthew, R.W. (1998) Use of dsRNA-mediated genetic interference to demonstrate that *frizzled and frizzled 2* act in the wingless pathway. *Cell*, **95**, 1017-1026.

Ketting, R.F. (2011) The many faces of RNAi. *Dev. Cell*, **20,** 148–161.

Khajuria, C., Vélez, A.M., Rangasamy, M., Wang, H., Fishilevich, E., Frey, M., Carneiro, N., Gandra, P., Narva, K., & Siegfried, D. (2015) Parental RNA interference of genes involved in embryonic development of the western corn rootworm, *Diabrotica virgifera virgifera* LeConte. *Insect Biochemistry and Molecular Biology* **63,** 54–62.

Knight, S.W. & Bass, B.L. (2001) A Role for the RNase III Enzyme DCR-1 in RNA Interference and Germ Line Development in *Caenorhabditis elegans*. *Science,* **293,** 2269–2271.

Koch, A., & Kogel, K.H. (2014) Review article New wind in the sails: improving the agronomic value of crop plants through RNAi-mediated gene silencing. *Plant Biotechnology*, **12,** 821–831.

Krishnan, P., Gireesh-Babu, p., Rajendran, K. & Chaudhari, A. (2009) RNA interference-based therapeutics for shrimp viral diseases. *Dis Aquat. Org*., **86,** 263–272.

Kumar, M, Gupta, G.P, Rajam, M.V 92009). Silencing of *acetylcholinesterase* gene of *Helicoverpa armigera* by siRNA affects larval growth and its life cycle. *J Insect Physiol*. **55**, 3-8.

Lakatos, L., Szittya, J., Silhavy, D., & Burgyán, J. (2004) Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombus viruses *EMBO,* **23,** 876–884.

Lee, K.S., Kim, B.Y., Choo, Y.M., Yoon, H.J. Kang, P.D. Woo, S.D. Sohn, H.D. Roh, J.Y. Gui, Z.Z. Je, Y.H. & Jin, B.R. (2009) Expression profile of cathepsin B in the fat body of *Bombyx mori* during metamorphosis. *Biochemistry & molecular biology,* **154,** 188-194.

Lee, R.C., Hammell, C.M., & Ambros, V. (2006) Interacting endogenous and exogenous RNAi pathways in *Caenorhabditis elegans*. *RNA,* **12,** 589–597.

Lehane, M.J., Gibson, W. & Lehane, S.M. (2008) Differential expression of fat body genes in Glossina morsitans morsitans following infection with *Trypanosoma brucei brucei*. *International journal for parasitology,* **38,** 93-101.

Li, J., Todd, T.C., Oakley, T.R., Lee, J. & Trick, H.N. (2010) Host-derived suppression of nematode reproductive and fitness genes decreases fecundity of *Heterodera glycines Ichinohe*. Planta, 232: 775–785.

Lilley, C.J., Davies, L.J. & Urwin, P.E. (2012) RNA interference in plant parasitic nematodes: a summary of the current status. *Parasitology,* **139,** 630–640.

Lippman, Z., & Martienssen, P. (2004) Review article: The role of RNA interference in heterochromatic silencing. *Nature,* **431**, 364-370.

Liu, P.Z. & Kaufman, T.C. (2004) Hunchback is required for suppression of abdominal identity, and for proper germband growth and segmentation in the intermediate germband insect *Oncopeltus fasciatus. Development,* **131**, 1515-1527.

Lorenzen, M., D., Brown, S.J., Denell, R.E. and Beeman, R.W. (2002) Transgene expression from the *Tribolium castaneum* Polyubiquitin promoter. *Insect Molecular Biology,* **11**, 399.

Lynch, J.A. & Desplan, C. (2006). A method for parental RNA interference in the wasp Nasonia vitripennis. *Nature Protocols,* **1,** 486-494

Mao, J., Zhang, P, Liu, C., & Zeng, F. (2015) Co-silence of the coatomer β and v-ATPase A genes by siRNA feeding reduces larval survival rate and weight gain of cotton bollworm, *Helicoverpa armigera. Pestic Biochem Physiol.*, **118**, 71-6.

Mao, Y.B., Xue, X.Y., Tao, X.Y., Yang, C.Q., Wang, L.J. & Chen, X.Y. (2013) Cysteine protease enhances plant-mediated bollworm RNA interference. *Plant Mol. Biol.*, **83,** 119–129.

Mao, Y.B., Tao, X.Y., Xue, X.Y., Wang, L.J. & Chen, X.Y. (2011) Cotton plants expressing CYP6AE14 double-stranded RNA show enhanced resistance to bollworms. *Transgenic Res.*, **20,** 665–673.

Mao, Y.B., Cai, W.J., Wang, J.W., Hong, G.J., Tao, X.Y., Wang, L.J., Huang, Y.P. & Chen, X.Y. (2007) Silencing a cotton bollworm *P450 monooxygenase* gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nature Biotechnology*, **25,** 1307–1313.

Mathur, G., Sanchez-Vargas, I., Alvarez, D., Olson, K.E., Marinotti, O. & James, A.A. (2010) Transgene-mediated suppression of dengue viruses in the salivary glands of the yellow fever mosquito, *Aedes aegypti*. *Insect molecular biology,* **19,** 753- 763

Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., & Zamore, P.D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into *Ago2*-containing RNAi enzyme complexes. *Cell.* **123,** 607-620.

Mercer, T.R., Dinger, M.E. & Mattick, J.S. (2009) Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.,* **10,** 155–159.

Meyering-Vos, M. & Müller, A. (2007) RNA interference suggests sulfakinins as satiety effectors in the cricket *Gryllus bimaculatus*. *Insect* physiology, **53**, 840-8.

Miguel, K.S., & Scott, J.G. (2016) The next generation of insecticides: dsRNA is stable as a foliar-applied insecticide. *Pest Manag Sci*, **72**, 801–809.

Murphy, K.A., Tabuloc, C.A., Cervantes, K.R., & Chiu, J.C. (2016) Ingestion of genetically modified yeast symbiont reduces fitness of an insect pest via RNA interference. *Scientific Reports,* **6**, 22587.

Mutti, N.S., Park, Y., Reese, J.C., & Reeck, G.R. (2006) RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrthosiphon pisum*. *Insect Science*, **6,** 1-7.

Napoli, C., Lemieux, C., & Jorgensen, R. (1990) Introduction of a *Chimeric Chalcone Synthase* Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant Cell,* **2**, 279–289.

Niimi, T., Kuwayama, H., Yaginuma, T. (2005) Larval RNAi applied to the analysis of postembryonic development in the ladybird Beetle, *Harmonia axyridis*. *Insect Biotechnology and Sericology* **7,** 95-102.

Nishikawa, T., & Natori, S. (2001) Targeted disruption of a pupal hemocyte protein of Sarcophaga by RNA interference. *European Journal of Biochemistry*, **268**, 5295–5299.

Niwa, R. & Slack, F.J. (2007) The evolution of animal microRNA function. *Curr Opin Genet Dev.*, **2,** 45-50.

Nunes, C.C. & Dean, R.A. (2012) Host-induced gene silencing: a tool for understanding fungal host interaction and for developing novel disease control strategies. *Mol. Plant Pathol.*, **13,** 519–529.

Oerke, E.C. (2006) Crop losses to pests. *Agric. Sci.*, **144,** 31–43.

Ouellet, D.L., Perron, M.P., Gobeil, l., Plante, P. & Provost, P. (2006) MicroRNAs in Gene Regulation: When the Smallest Governs It All. *Biomed Biotechnol.* **2000,** 69616.

Pauchet, Y., Wilkinson, P., Munster, M., Augustin, S., & Pauron, D., (2009) Pyrosequencing of the midgut transcriptome of the poplar leaf beetle *Chrysomela tremulae* reveals new gene families in Coleoptera. *Insect Biochem Mol Biol* **39,** 403–413.

Pitino, M., Coleman, A.D., Maffei, M.E., Ridout, C.J. & Hogenhout, S.A. (2011) Silencing of aphid genes by dsRNA feeding from plants. *PLoS One*, **6**, e25709.

Powell, M.E., Bradish, H.M, Gatehouse, J.A, Fitches, E.C (2016) Systemic RNAi in the small hive beetle *Aethina tumida* Murray (Coleoptera: Nitidulidae), a serious pest of the European honey bee Apis mellifera. *Pest Manag Sci.*, **73**, 53-63

Price, D.R. & Gatehouse, J.A. (2008) RNAi-mediated crop protection against insects. *Trends Biotechnology*, **26,** 393–400.

Rajagopal, R., Sivakumar, S., Agrawal, N., Malhotra/ P., & Bhatnagar, R. (2002) Silencing of Midgut *Aminopeptidase N* of *Spodoptera litura* by Double-stranded RNA Establishes Its Role as *Bacillus thuringiensis* Toxin Receptor. *Biological Chemistry*, **277**, 46849–46851.

Rand, T.A., Petersen, S., Du, F. & Wang, X. (2005) *Argonaute2* Cleaves the Anti-Guide Strand of siRNA during RISC Activation. Cell, **123,** 621–629.

Rusten, T.E. & Simonsen, A. (2008) ESCRT functions in autophagy and associated disease. *Cell Cycle*, **7**, 1166–1172.

Saleh, M-C, Hekele, A., Gillis, A., Foley, E. & O'Farrell, P.H. (2006) The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. *Nature Cell Biol*ogy **8,** 7-9.

Sant'Anna, M.R., Alexander, B., Bates, P.A. & Dillon, R.J. (2008). Gene silencing in phlebotomine sand flies: *Xanthine dehydrogenase* knock down by dsRNA microinjections. *Insect Biochem Mol Biology* **38**, 652-60.

San Miguel, K. & Scott, J. G. (2015) The next generation of insecticides: dsRNA is stable as a foliar-applied insecticide. *Pest Management Science,* **72,** 801-9.

Seyhan, A.A. (2011) RNAi: a potential new class of therapeutic for human genetic disease. *Human Genetics*, **130**, 583–605.

Scott, J. G., Michel, K., Bartholomay, L. C., Siegfried, B. D., Hunter, W. B., Smagghe, G., Douglas, A. E. (2013) Towards the elements of successful insect RNAi. *Insect Physiology*, **59**, 12–21.

Shankar, P., Manjunath, N. & Lieberman, J. (2005) The prospect of silencing disease using RNA interference. *JAMA.* **293**, 1367-1373.

She, R., Chu, J.S.-C., Wang, K., Pei, J., & Chen, N. (2009) GenBlastA: Enabling BLAST to identify homologous gene sequences. *Genome Research*, **19**, 143–149

Shreve, J.T., Shukle, R.H., Subramanyam, S., Johnson, A.J., Schemerhorn, B.J., Williams, C.E. & Stuart, J.J. (2013) A genome-wide survey of small interfering RNA and microRNA pathway genes in a galling insect. *Insect Physiology*, **59,** 367–76.

Sim, C. & Denlinger, D.L. (2008*) Insulin signaling and FOXO* regulate the overwintering diapause of the mosquito *Culex pipiens*. *Proc Natl Acad Sci U S A.* **105**, 77-81.

Sindhu, A.S., Maier, T.R., Mitchum, M.G., Hussey, R.S. & Davis, E.L. (2009) Effective and specific in planta RNAi in cyst nematodes: expression interference of four parasitism genes reduces parasitic success. *Exp. Bot.*, **60**, 315–324.

Singh, A. D., Wong, S., Ryan, C. P. & Whyard, S. (2013) Oral Delivery of Double-Stranded RNA in Larvae of the Yellow Fever Mosquito, *Aedes aegypti*: Implications for Pest Mosquito Control. *Insect Science*, **13**, 69.

Siomi, M.C., Sato, K., Pezic, D. & Aravin, A.A. (2011) PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.*, **12**, 246– 258.

Siritunga, D. & Sayre, R.T. (2003) Generation of cyanogen-free transgenic cassava. *Planta.* **217,** 67-73.

Soares, C.A.G., Lima, C.M.R., Dolan, M.C., Piesman, J., Beard, C.B. & Zeidner, N.S. (2005) Capillary feeding of specific dsRNA induces silencing of the *isac* gene in nymphal *Ixodes scapularis* ticks. *Insect molecular biology*, **14**, 443-452.

Song, J., Smith, S.K., Hannon, J. G., Leemor, J. (2004) Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science*, **305,** 14-34.

Swevers, L., Huvenne, H., Menschaert, G., Kontogiannatos, D., Kourti, a, Pauchet, Y., Smagghe, G. (2013) Colorado potato beetle (Coleoptera) gut transcriptome analysis: expression of RNA interference-related genes. *Insect Molecular Biology*, **22,** 68–84.

Roignant, J.Y., Carré, C., Mugat, B., Szymczak, D., Lepesant, J.A. & Antoniewski, C. (2003) Absence of transitive and systemic pathways allows cell-specific and isoform specific RNAi in *Drosophila*. *RNA*, **9**, 299-308.

Romano, N., & Macino, G. (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Molecular Microbiology*, **6,** 3343–53.

Tabara, H., Grishok, A., & Mello, C. C. (1998). RNAi in *C. elegans*: Soaking in the Genome Sequence. *Science***, 282,** 430 – 431.

Tabunoki, H., Higurashi, S., Ninagi, O., Fujii, H., Banno, Y. & Nozaki, M. (2004) A carotenoid-binding protein (CBP) plays a crucial role in cocoon pigmentation of silkworm (*Bombyx mori*) larvae. *FEBS letters*, **567,** 175-178.

Tang, T., Zhao, C., Feng, X., Liu, X. & Qiu, L. (2012) Knockdown of several components of cytochrome P450 enzyme systems by RNA interference enhances the susceptibility of *Helicoverpa armigera* to fenvalerate. *Pest Manag. Sci***., 68,** 1501–1511.

Tao, X.Y., Xue, X.Y., Huang, Y.P., Chen, X.Y. & Mao, Y.B. (2012) Gossypol-enhanced *P450* gene pool contributes to cotton bollworm tolerance to a pyrethroid insecticide. *Mol. Ecol.*, **21,** 4371–4385.

Taracena, M.L., Oliveira P.L., Almendares O., Umaña C., Lowenberger C., Dotson EM., Paiva-Silva GO., Pennington P.M. (2015) Genetically modifying the insect gut microbiota to control Chagas disease vectors through systemic RNAi. *PLoS Negl. Trop. Dis.*, **9,** e0003358.

Terenius, O., Papanicolaou, A., Garbutt, J.S., Eleftherianos, I., Huvenne, H., Kanginakudru, S., Smagghe, G. (2011) RNA interference in Lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design. *Insect Physiology*, **57**, 31–45.

Thomson, T. & Lin, H. (2009) The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu. Rev. Cell Dev. Biol.*, **25,** 355–376.

Tian, H., Peng H., Yao, Q., Chen, H., Xie, Q., Tang, B. & Zhang, W. (2009) Developmental control of a lepidopteran pest *Spodoptera exigua* by ingestion of bacteria expressing dsRNA of a non-midgut gene. *PLoS One* **4,** e6225.

Timmons, L. & Fire, A. (1998). Specific interference by ingested dsRNA. *Nature*, **395,** 854

Tomoyasu, Y., Miller, S.C., Tomita, S., Schoppmeier, M., Grossmann, D., Bucher, G. (2008) Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in *Tribolium*. *Genome Biol*., 9, R10

Tomoyasu, Y. & Denell, R.E. (2004) Larval RNAi in *Tribolium casteneum* (Coleoptera) for analyzing adult development. *Development genes and evolution*, **214,** 575-578.

Turner, C.T., Davy, M.W., MacDiarmid, R.M., Plummer, K.M., Birch, N.P. & Newcomb, R.D. (2006) RNA interference in the light brown apple moth, *Epiphyas postvittana* (Walker) induced by double-stranded RNA feeding. *Insect molecular biology***, 15,** 383-391.

Upadhyay, S.K, Chandrashekar, K., Thakur, N, Verma, P.C, Borgio, J.F, Singh, P.K, (2011) RNA interference for the control of whiteflies (*Bemisia tabaci*) by oral route. *Bioscience,* **36,** 5-6.

Vanbergen, A.J. & the Insect Pollinators Initiative (2013) Threats to an ecosystem service: pressures on pollinators-REVIEWS. *Front Ecol Environ*, 11, 251–259.

Vaucheret, H., B!eclin, C. & Fagard, M. (2001) Post-transcriptional gene silencing in plants. *Cell Science*, **114,** 3083–3091.

Vera, J.C., Wheat, C.W., Fescemyer, H.W., Frilander, M.J., Crawford, D.L. (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Molecular *Ecology*, **17,** 1636–1647.

Wang, Y., Zhang, H., Li, H., Miao, X (2011) Second-generation sequencing supply an effective way to screen RNAi targets in large scale for potential application in pest insect control. *PLoS One* **6,** e18644.

Walshe, D.P., Lehane, S.M., Lehane, M.J. & Haines, L.R. (2009) Prolonged gene knockdown in the tsetse fly *Glossina* by feeding double stranded RNA. *Insect molecular biology*, **18,** 11-19.

Wei, Z., Yin, Y., Zhang, B., Wang, Z., Peng, G., Cao, Y., & Xia, Y. (2007) Cloning of a novel protease required for the molting of *Locusta migratoria manilensis*. *Dev Growth Differ*., **49**, 611–621

Whangbo, J. S., & Hunter, C. P. (2008) Environmental RNA interference. *Trends in Genetics: TIG*, **24,** 297–305.

Whyard, S., Singh, A.D. & Wong, S. (2009) Ingested double-stranded RNAs can act as species-specific insecticides. *Insect Biochemistry and Molecular Biology*, **39,** 824-832.

Wieczorek, H., Beyenbach, K.W., Huss, M., Vitavska, O. (2009) Vacuolar-type proton pumps in insect epithelia. *Experimental Biology,* **212**, 1611-1619.

Winston, W.M., Sutherlin, M., Wright, A.J., Feinberg, E.H. & Hunter, C.P. (2007) *Caenorhabditis elegans SID-2* is required for environmental RNA interference. *Proceedings of the National Academy of Sciences of the United States of America*, **104,** 10565-10570.

Winston, W.M., Molodowitch, C. & Hunter, C.P. (2002) Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science***, 295,** 2456–2459.

Xu, L., Duan, X., Lv, Y., Zhang, X., Nie, Z., Xie, C., Ni, Z. & Liang, R. (2014) Silencing of an aphid *carboxylesterase* gene by use of plant-mediated RNAi impairs *Sitobion avenae* tolerance of Phoxim insecticides. *Transgenic Res.,* **23,** 389–396.

Yadav, B.C., Veluthambi, K. & Subramaniam, K. (2006) Host-generated double stranded RNA induces RNAi in plant-parasitic nematodes and protects the host from infection. *Mol. Biochem. Parasitol.* **148,** 219–222.

Yeo, J.H. & Chong, M.M. (2011) Many routes to a micro RNA. *IUBMB Life*, **63,** 972–978.

Yu, N., Christiaens, O., Liu, J., Niu, J., Cappelle, K., Caccia, S., Smagghe, G. (2012) Delivery of dsRNA for RNAi in insects: an overview and future directions. *Insect Science,* **20,** 4–14.

Zamore, P.D., Tuschl, T., Sharp, P.A. & Bartel, D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101,** 25-33.

Zhang, H., Li, H., Guan, R., & Miao, X. (2015) Lepidopteran insect species-specific, broad-spectrum, and systemic RNA interference by spraying dsRNA on larvae. *Entomol. Exp. Appl.*, **155,** 218–228.

Zhang, H., Li, H.C. & Miao, X.X. (2013) Feasibility, limitation and possible solutions of RNAi-based technology for insect pest control. *Insect Sci.,* **20,** 15– 30.

Zhang, X., Zhang, J. & Zhu, K. Y. (2010). Chitosan/double-stranded RNA nanoparticle-mediated RNA interference to silence chitin synthase genes through larval feeding in the African malaria mosquito (*Anopheles gambiae*). *Insect Mol. Biol.* **19,** 683–693.

Zhou, X., Wheeler, M.M., Oi, F.M. & Scharf, M.E. (2008). RNA interference in the termite *Reticulitermes flavipes* through ingestion of double-stranded RNA. *Insect Biochemistry and Molecular Biology,* **38,** 805-815.

Zhu, F., Xu, J., Palli, R., Ferguson, J., & Palli, S. R. (2011) Ingested RNA interference for managing the populations of the Colorado potato beetle, *Leptinotarsa decemlineata*. *Pest Management Science,* **67,** 175–82.

Zotti, M. J. & Smagghe, G. (2015) RNAi Technology for Insect Management and Protection of Beneficial Insects from Diseases: Lessons, Challenges and Risk Assessments. *Neotrop Entomol.* **44,** 197-213.

CHAPTER 3

Insights into the *Diabrotica virgifera virgifera* ATP-binding Cassette Transporter Gene Family Through Transcriptomics

Folukemi Adedipe[1], Nathaniel Grubbs[1], Brad Coates[2], Brian Wiegmman[1] and Marcé D. Lorenzen[1]

[1]Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA

[2]Corn Insects and Crop Genetics Research Unit, ARS, USDA, Ames, IA 50011, USA

ABSTRACT

The western corn rootworm, *Diabrotica virgifera virgifera*, is a pervasive pest of cultivated maize in North America and Europe, which has evolved to survive exposure to multiple insecticidal agents. Members of the ATP-binding cassette (ABC) transporter family play diverse cellular roles in arthropods, including xenobiotic detoxification. In this study, 65 putative *D. virgifera virgifera* ABC transporters, belonging to subfamilies A to H (*Dvv*ABC-A to *Dvv*ABC-H), were identified within transcriptome assemblies from the combined sequencing reads from Illumina libraries prepared from four RNA samples (embryos, larvae, adult male, and adult female). Eight *D. virgifera virgifera* ABC genes, with direct orthology to *Tribolium castaneum* gene models and known to have obvious knockdown phenotypes, were tested for functional conservation using RNA interference, which resulted in phenotypes similar to those seen in *T. castaneum*. Interestingly, depletion of *DvvABCB_19147* and *DvvABCG_3712* transcripts in adult females produced detrimental reproductive and developmental phenotypes (egg-laying and -hatching defects), demonstrating their potential as insect birth control targets. This study provides insights into the developmental and physiological roles of *D. virgifera virgifera* ABC transporters and suggests candidate target sites for the development of novel insect control tactics.

Keywords: ATP-binding cassette (ABC) transporter, phylogenetic, transcriptome, RNA interference (RNAi), corn rootworm

INTRODUCTION

The western corn rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae), is a major pest of maize in North America (Levine and Oloumi-Sadeghi, 1991; Levine et al., 2002). It is considered a notorious pest due to its ability to overcome various management methods. Costs from managing this pest, as well as from crop losses due to the damage it causes are estimated to be over one billion dollars annually in North America (Marra et al., 2012). Common management practices include the use of insecticides, crop rotation, and transgenic maize expressing crystal proteins from the soil bacterium *Bacillus thuringiensis* (*Bt*). Unfortunately, *D. virgifera virgifera* has evolved resistance to crop rotation (Gray et al., 2009; Krysan, 1986; Sammons et al. 1997), to many chemical insecticides (Meinke et al., 1998; Miller et al. 2009), and to a number of *Bt* proteins (Gassmann *et al*., 2011; Narva *et al*., 2013). However, recent studies have shown that *D. virgifera virgifera* is highly sensitive to a new management tool, RNA interference (RNAi), which could be used as an alternative technology to protect corn from rootworm damage (Baum *et al*., 2007; Rangasamy and Siegfried, 2012; Li *et al*., 2015). Nonetheless, the ability of this insect to adapt to control methods demonstrates a fundamental need for a greater understanding of *D. virgifera virgifera* biology and molecular genetics.

Unfortunately, the genetic resources available for *D. virgifera virgifera* are limited, so progress toward understanding multigenic traits, like those that contribute to resistance evolution, will require the development of fundamental genetic and genomic tools for this species. Complicating the development of these tools is the fact that *D. virgifera virgifera* has a large genome, about 2.58 GB (Coates et al., 2012; Sappington et al., 2009), which is ten times bigger than in species like *Drosophila melanogaster* and *Tribolium castaneum*, which are rich with genetic resources. However, focusing on the transcriptome could provide valuable information, even in the absence of a completed genome assembly. Indeed, RNA-Seq has become a quick, cost effective method for obtaining a wealth of genetic data in non-model insects that lack a sequenced genome (Wang et al., 2011).

ATP-binding cassette (ABC) transporters comprise one of the largest protein families; and are present in all organisms from prokaryotes to eukaryotes (Higgins, 1992). Functional studies on ABC-family genes of several arthropod species have relied on genomic data, including *T. castaneum* (Broehan et al., 2013*), Bombyx mori (*Liu et al., 2011), *D. melanogaster (*Dean et al., 2001), *Bemisia tabaci* (Tian et al., 2017), *Daphnia pulex* (Sturm et al., 2009), and *Tetranychus urticae* (Dermauw et al., 2014); however, a significant proportion of the reported ABC transporters in each of these species could have been discovered in the transcriptomic data alone.

ABC-family proteins are structurally characterized by two domains: the highly conserved nucleotide-binding domain (NBD) and the variable transmembrane domain (TMD) (George and Jones, 2012). The NBD binds and hydrolyzes ATP to provide the energy required for driving a substrate across a cellular membrane, while the TMD creates the channel through which the substrate is transported (Rees et al., 2009). Each NBD possesses several highly-conserved, characteristic motifs including; Walker A, Walker B, Q-loop, D-loop, H-loop, and the ABC signature motif, while each TMD is made up of five to six transmembrane α-helices, and dictate substrate specificity (Rees et al., 2009). To function as transporters, ABC proteins require two NBDs and two TMDs to work together. Some ABC transporters, called full-transporters (FT) contain all the necessary domains (2 TMDs and 2 NBDs), in a single protein. Half-transporters (HT), on the other hand, only have a single TMD and a single NBD, so must work as pairs, either as homodimers or heterodimers, with another half transporter molecule to form a functional unit (Higgins, 1992; Dean et al., 2001; Rees et al., 2009). Because of its conserved nature, the sequence of the NBD has been used to classify this superfamily into eight subfamilies designated A to H (ABCA to ABCH) (Dean and Annilo, 2005).

Among insects, ABC transporters have been implicated in diverse functions such as, development (Broehan et al., 2013), transportation of eye pigments (Ewart et al., 1994; Mackenzie et al., 1999; Liu et al., 2014; Mackenzie et al., 2000), and resistance to insecticides (Aurade et al., 2010; Buss and Callaghan, 2008). Interestingly, based on transcriptomic data, Bretschneider et al., (2016) suggested that ABC transporters are involved in the degradation of plant secondary metabolites in the cotton bollworm, *Helicoverpa armigera*. And Broehan et al., (2013) reported the effects of RNAi knock-down of several ABC transporters in the red flour beetle, *T. castaneum*, which included growth arrest, eye pigmentation defects, abnormal cuticle formation, egg-laying and egg-hatching defects, and mortality.

As mentioned above, there are many transcriptomic and functional genomic studies on ABC genes of several arthropod species, but, none in *D. virgifera virgifera*. Characterizing the ABC genes in *D. virgifera virgifera* is particularly important since members of the ABC family have been implicated in the mode of action and resistance to *Bt* toxins in Lepidoptera (Guo et al., 2015). As a first step toward understanding the function of these transporters in this species, we mined the transcriptome data (RNA-Seq) for ABC transporter genes and characterized several via RNAi. Specifically, eight ABC transporters were identified as putative orthologs to those previously reported to have a defining RNAi phenotype in the model coleopteran species, *T. castaneum* (Broehan et al., 2013; *DvvABCA_18330*, *DvvABCB_19147*, *DvvABCE_2830*, *DvvABCF_2701*, *DvvABCG_3712*, *DvvABCG_49457*, *DvvABCG_13829* and *DvvABCH_18290*). Subsequent RNAi-mediated knockdown demonstrated conservation of function with *T. castaneum*, as well as established potential new insecticidal targets for the control of this devastating agricultural pest.

MATERIALS AND METHODS

Insect rearing

All *D. virgifera virgifera* used in this study were obtained from a laboratory-established colony at North Carolina State University (see Chu et al., 2017). Eggs deposited in an oviposition chamber (agar plate with cheese cloth) were collected weekly, pipetted into soil-filled containers, and kept at 26°C. Experimental larvae were reared on roots of germinated corn seeds in 16-oz containers, while adults were maintained in a 30cm^3 BugDorm (MegaView Science, Taiwan) at 26°C, 70% relative humidity with an L14:D10 photoperiod and fed an artificial diet as needed. All injected beetles were kept and reared in a container containing corn seedling under similar conditions to allow further observation.

RNA isolation and Illumina sequencing

In an effort to identify as many *D. virgifera virgifera* ABC transporters as possible, we isolated RNA from four samples. Specifically, four distinct sets of insects were collected by separating three life stages (embryos, larvae, and adults) and then separating adults by sex. Total RNA was isolated using the RNeasy Mini Kit (Qiagen, Hilden, Germany) and treated with DNase I (Qiagen) according to the manufacturer’s instructions. The RNA was submitted to the Genomic Sciences Laboratory (North Carolina State University, NC USA) for quality assessment, poly(A) selection, fragmentation, size selection (600- to 700-bp) and Illumina TruSeq library preparation and sequencing (Illumina MiSeq sequencer, 2 x 300-bp).

Transcriptome assembly and annotation

Raw reads (all life stages) underwent an initial quality assessment using FastQC (Andrews, 2010), and were trimmed based on quality sores, and any remaining Illumina adapter sequences removed. Trimmed reads were *de novo* assembled into contigs using SeqMan NGen (DNASTAR, Inc; Madison, Wisconsin, USA) to generate a combined assembly and the relative completeness evaluated by comparison with the universal single copy orthologs from Arthropoda obtained from OrthoDB v 9 using BUSCO v 3 (Waterhouse et al., 2012). SeqMan NGen (DNASTAR) assemblies were evaluated for total number and median length (N50) of contigs. The derived transcript sequences from SeqMan NGen (DNASTAR) were imported into Blast2GO v 4.0 (Conesa et al., 2005; Götz et al., 2008) and annotations acquired with BLASTx comparisons to the non-redundant (nr) arthropod-specific database at the National Center of Biotechnology Information (NCBI). The combined graphs were created at level 2 for Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories from Blast2GO.

Bioinformatics analysis of the *D. virgifera virgifera* ABC transporter superfamily

A searchable database was created from the combined *D. virgifera virgifera* assembly and the database searched using BlastStation software (TM Software Inc.; Arcadia, CA, USA). The deduced amino-acid sequence of each *T. castaneum* ABC transporter was used as query to identify the corresponding *D. virgifera virgifera* genes (Broehan et al., 2013). Homologous sequences were selected based on their sequence identity and E-values ($<10^{-6}$). Each of the selected *D. virgifera virgifera* sequences were then compared (BLASTx) to the complete nr database at NCBI (http://blast.ncbi.nlm.nih.gov/), retaining only those matching ABC transporters. This process was repeated until no new ABC transcripts were found in the *D. virgifera virgifera* database. The number and positions of transmembrane domains were assessed with NCBI's Conserved Domain Search Service. Finally, each putative *D. virgifera virgifera* ABC gene was assigned to a subfamily (A-H) based on BLASTx analyses, as well as phylogenetic inferences.

To ensure all potential ABC transporters were identified we generated a second set of assemblies, one was produced using Trinity (Grabherr et al., 2011) and another using SOAPdenovo (Xie et al., 2014). Independent databases were then generated from each (Trinity and SOAPdenovo) and all ABC-related sequences identified. The complexity of each ABC gene set was reduced by clustering allelic variants (sequence), and a non-redundant, comprehensive file of *D. virgifera virgifera* ABC contigs generated. The assembly of origin is denoted by D (DNAStar), T (Trinity), and S (SOAPdenovo).

Phylogenetic trees were constructed using the NBD, the most conserved region among all ABC transporters, corresponding to the *D. virgifera virgifera* ABC. Sequence alignments were performed with ClustalW using MEGA 7 (Tamura et al. 2013),) and subjected to phylogenetic analysis with 1000 bootstrap replications. The phylogenetic trees were constructed using the Neighbor-joining method, but the Best Model was determined for each tree by the program. Subfamily-specific clustering was then compared with that of previous phylogenetic analyses of ABC transporters from *T. castaneum, B. mori* and *D. melanogaster*. To help determine orthology, an unrooted phylogenetic tree showing the phylogenetic relationships of the selected *D. virgifera virgifera* genes and the full set of *T. castaneum* ABC transporters was constructed. The deduced amino-acid sequence of each *T. castaneum* ABC gene was obtained from NCBI, and aligned to the full-length of the *D. virgifera virgifera* deduced amino-acid sequence when possible (note that some *D. virgifera virgifera* contigs failed to represent full-length transcripts). We used *ABCD2* from *T. castaneum* and *Anoplophora glabripennis* as the outgroup. Sequence alignments were performed as described above.

Gene expression analyses by RT-PCR

Relative levels of gene expression were determined for each of the selected ABC genes (*DvvABCA_18330, DvvABCB_19147, DvvABCE_2830, DvvABCF_2701, DvvABCG_3712, DvvABCG_49457, DvvABCG_13829* and *DvvABCH_18290*), by semi-quantitative RT-PCR to ensure injection of dsRNAs would be performed before peak expression occurred. RNA was extracted from each developmental stage, with adults separated by sex and total RNA reverse transcribed into cDNA using the Superscript™ III First-Strand Synthesis System (Invitrogen) along with an anchored poly(T) primer. Each cDNA pool represents a specific developmental stage, or sex. These pools were then individually used as template for PCRs using each of eight pairs of gene-specific primers. PCR expression profiling was performed by using each of the 4 prepared cDNA pools, Master Mix and the gene-specific primer set designed to amplify the *D. virgifera virgifera* ABC transcript of interest. The highly-conserved housekeeping gene, *DvvRPS6*, was used to assess template quality and quantity. The PCRs were performed in a T100 Thermal Cycler (Bio-Rad) with the following cycling conditions: (95°C for 3 min), 25× (95°C for 30s, 58°C for 30s, 72°C for 10s), (4min incubation at 72°C). Amplification products from each of the different life stages and sexes were then visualized and compared using gel electrophoresis.

Primer design and *in vitro* dsRNA synthesis

Primer sets used to generate dsRNA were designed (Table 1) to regions encoding the gene-specific TMD domains, and not in regions encoding the conserved NBD domains, to avoid potential off-target effects. Primers were designed using Vector NTI Advance (VNTI) software (Invitrogen), for all ABC genes whose orthologs are known to produce obvious RNAi phenotypes in *T. castaneum* (Broehan et al., 2013). Partial cDNAs were amplified for the selected genes (*DvvABCA_18330, DvvABCB_19147, DvvABCE_2830, DvvABCF_2701*, *DvvABCG_3712*, *DvvABCG_49457*, *DvvABCG_13829* and *DvvABCH_18290*), as previously described. Fully-nested PCR was performed with an initial denaturation of (95°C for 3 min), 35× (95°C for 30s, 58°C for 30s, 72°C for 10s), (4min incubation at 72°C). The PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions, then ligated into the pGEM-T vector (Promega), and the resulting plasmids were used to transform competent *E. coli* cells. All positive clones were cultured in a selective LB medium. The recombinant plasmid DNAs were isolated using the Plasmid Mini Kit (Qiagen), and the inserts sequenced and confirmed by BLAST analysis (http://blast. ncbi.nlm.nih.gov). Each ABC-containing plasmid was individually used as template in PCR reactions containing the following primers; T7 as a forward primer (due to location in pGEM), and a pGEM-specific reverse primer that was tailed with T7. This enabled all PCRs to be performed with the same set of primers. PCR was performed with an initial denaturation of (95°C for 3 min), 35× (95°C for 30s, 58°C for 30s, 72°C for 10s), (4min incubation at 72°C). The PCR products were confirmed by agarose gel electrophoresis, purified, and ~1µg of each was used as template for dsRNA synthesis using the MEGAscript T7 *in vitro* Transcription Kit

(Ambion). Each of the synthesized dsRNAs were purified using the MEGAclear Kit (Ambion) and concentration determined using a Nanodrop 1000 (Thermo Scientific, Waltham, MA, USA).

RNAi bioassay

RNAi assays were conducted by individually injecting dsRNA corresponding to each of the 10 specific *D. virgifera virgifera* ABC genes into the hemocoel of 3rd-instar larvae, pre-pupae and/or newly-eclosed female adults. Before microinjection, experimental insects were anesthetized on ice for 30 min, then injected with a gene-specific dsRNA at a concentration of 1-2µg/µl. For each treatment, the entire process was replicated three times, as three independent biological replicates, with at least 20 individuals in each replicate. Following injection, larvae and pre-pupae were allowed to recover at room temperature for one hour, and then moved to germinated corn for further monitoring and phenotypic analysis. Phenotypes were observed daily using a stereomicroscope, and transcript levels assessed after 5 days by RT-PCR using RNA isolated from pools of injected individuals (one individual per replicate, for a total of three individuals per RT-PCR).

Treated females were kept in an oviposition chamber (agar plate with cheese cloth) and maintained on an artificial diet. Two days post-injection, females were mated to untreated males, and generally started to lay eggs ~10 days later. To determine egg viability, eggs were harvested from the oviposition chamber and placed on moistened filter paper in Petri dishes and held at 26°C, 70% relative humidity with an L14:D10 photoperiod. Females were allowed to lay eggs over a two-week period, and eggs were counted every other day to assess the rate of egg laying. Hatch rate counts were made every other day, beginning ten days after the first egg lay (22 days post injection) and continuing for four weeks until no further hatching was observed.

RESULTS

Transcriptome sequencing, assembly, and annotation

Illumina (MiSeq) sequencing generated 22,296,672 raw sequence reads across four libraries. Of these, 13,070,671 assembled in a single-combined transcriptome producing 25,296 contigs with an average length of 2,463 bp. BLAST-based annotation was performed with Blast2GO (Conesa et al., 2005) using BLASTx against the arthropod-specific section of NCBI's non-redundant (nr) protein database. Among the 25,296 contigs derived from the *D. virgifera virgifera* transcriptome, BLASTx analysis revealed over 18,343 contigs that shared significant identity (*E*-values < 10^{-6}) with known arthropod proteins, and a subset of these received GO mapping and annotation terms (Figure 1). Sequences lacking significant identity to known arthropod proteins likely represent poorly conserved and/or novel sequences, as well as non-coding RNAs. The distribution of top BLAST hits by species showed that *T. castaneum* was the most frequent, representing 65% of the matches (Figure 2). The primary reason for this is the extensive genomic/transcriptomic data available for *T. castaneum.* BLAST-based annotation was performed using BLASTx in Blast2GO against the NCBI nr database. Among ontologies acquired by GO mapping at level 2, a majority of the associated GO terms were assigned to cell structural component, metabolic process, and catalytic activity respectively for GO Cellular Component (C), Biological Process (P) and Molecular Function (Figure 3). The SeqMan NGen assembly showed a high degree of completeness based on a BUSCO score of 955, or 98.6% of the 1,066 arthropod reference set (v. 9.0) being represented.

Annotation and phylogeny of ABC transporter genes

Because of the prominent role ABC transporters play in resistance to insecticides, we used our database of *D. virgifera virgifera* transcriptome contigs to identify the ABC genes for this species. To do this, we used the BLAST function of BlastStation to search the *D. virgifera virgifera* transcriptome database with each of the predicted protein sequences of the 74 known *T. castaneum* ABC genes (Broehan et al., 2013; Grubbs et al., 2015). Our analysis identified 65 putative *D. virgifera virgifera* ABC genes. To determine the evolutionary relationships of the *D. virgifera virgifera* ABC genes, phylogenetic analysis was carried out based on ClustalW multiple sequence alignment of amino-acid sequences of the identified NBDs. We found that the sequences formed clades according to the eight known ABC transporter subfamilies (A to H) (Figure 4; Table 2). The most similar sequences from BLASTx analyses were putative transporters from *T. castaneum* and *A. glabripennis.* Although these results suggest that *D. virgifera virgifera* shared a lineage with both *T. castaneum* and *A. glabripennis*, it is important to note that these species have extensive publicly available genomic data (Broehan et al., 2013; McKenna et al., 2016). In addition, an unrooted tree constructed using only the eight genes selected for RNAi analysis reveals that these genes are more evolutionarily conserved and clearly group into clusters with *T. castaneum* (Figure 5). Moreso, the phylogenetic tree reveals the differences between each of the selected genes and their outgroup, indicating that the *D. virgifera virgifera* genes selected for RNAi analysis are indeed orthologous to the intended *T. castaneum* genes.

Each *D. virgifera virgifera* ABC gene was named based on subfamily as well as *D. virgifera virgifera* contig (Table 2 – list of ABCs). Broehan et al., (2013) identified 73 ABC transporters in *T. castaneum*, and a 74th ABC transporter was reported by Grubbs et al. in 2015, but our study only identified 65 ABC transporters in *D. virgifera virgifera*. However, since our survey is only based on low-coverage transcriptomic data, it is entirely possible that the *D. virgifera virgifera* genome harbors additional ABC genes, which are either transcribed at relatively low levels, or are expressed at time points not covered by our RNA-Seq libraries. The ABCA subfamily is comprised of full-transporters that are among the well-characterized ABCs; the *D. virgifera virgifera* ABCA subfamily contains four members. The ABCB subfamily contains seven members, which includes two types of proteins: full transporters and half transporters with a TMD-NBD arrangement. As expected, the ABCC subfamily proteins are always full transporters, and 32 were identified in *D. virgifera virgifera*. The ABCD, ABCE and ABCF subfamilies each contained only a small number of ABC members, with clear orthologous relationships to *T. castaneum*. The ABCD subfamily contained two ABC proteins, which were TMD-NBD half transporters. Only one ABCE and three ABCF genes were identified, and each had two NBD motifs with no TMDs, suggesting they are unlikely to function as transmembrane transporters. The *D. virgifera virgifera* ABCG group is the second largest subfamily, with 12 members, most of which were half transporters with only a single NBD. The *D. virgifera virgifera* ABCH subfamily contains four members with a reverse domain organization similar to the ABCG subfamily.

Gene expression across developmental stages

Broehan et al., (2013) published a comprehensive analysis of the ABC genes of *T. castaneum*, in which they tested the developmental function of each of the 73 genes they identified using RNAi. However, only ten of these genes produced obvious phenotypes. Because *D. virgifera virgifera* is a difficult species to rear (Chu et al., 2014), we chose to focus our functional analysis of *D. virgifera virgifera* ABC genes to those orthologous to the ten TcABC genes that gave observable phenotypes: *DvvABCA_18330, DvvABCB_19147, DvvABCE_2830, DvvABCF_2701, DvvABCG_3712, DvvABCG_49457, DvvABCG_13829* and *DvvABCH_18290* (Table 2). Interestingly our initial survey (DNASTAR assembly) failed to find an ortholog of *T. castaneum scarlet*, so while *TcABCG-9A* and *TcABCG-9B* represent two separate, but related eye-color genes in *T. castaneum* (*white* and *scarlet*, respectively) we only identified the *white* ortholog in the initial *D. virgifera virgifera* transcriptome assembly. Therefore, our analysis is of eight *D. virgifera virgifera* genes. We used RT-PCR to examine the developmental expression profiles of these eight *D. virgifera virgifera* ABC genes, not surprisingly, all of the transcripts were expressed in all of the examined developmental stages, except for ABCG group which exhibit stage-specific transcription in eggs, and adult (Figure 6). The low expression of *white* transcripts in *D. virgifera virgifera* stages is similar to what (Bohen et al., 2013) reported for *white*, in *T. castaneum*. While *white* expression peaks during the larval and pupal stages due to its role in the transport of pigments (Sullivan et al., 1980), its expression level remains relatively low.

RNAi bioassays

Insects were microinjected with dsRNAs at the growth stages specified in Table 3, and five days post injection, RNA was isolated from both the experimental and control groups. Transcript levels were then assessed via semi-quantitative RT-PCR and treated samples compared to the control group. As seen in Figure 7, the level of each target transcript was significantly reduced at 5 days post-injection when compared to buffer-injected controls. Moreover, each dsRNA treatment resulted in defined phenotypes among each specific dsRNA treated cohort (Table 4; Figure 8). Specifically, knockdown of *DvvABCA_18330* led to mortality among approximately 60% of the treated pre-pupae compared to 5% for the buffer treated control. In addition, the adults that survived pre-pupal injection and successfully eclosed showed defects in wing and elytra development (Figure 8A), while buffer-injected controls did not. Importantly, these results are similar to those reported for the *T. castaneum* orthologs (Broehan et al., 2013). Subfamily A transporters have been implicated in lipid transport in mammals, which can impact cell physiology (Quazi and Molday, 2011), so it is conceivable that *DvvABCA_18330* also plays a role in lipid transport like its orthologs, *TcABCA-9A/9B*, in *T. castaneum* (Broehan et al., 2013).

In insects, ABCB genes have been implicated in insecticide resistance (Labbe et al., 2011), while in humans, subfamily B genes are known to be mitochondrial transporters with roles in iron metabolism and transportation of Fe/S protein precursors (Pondarré et al., 2006; Krishnamurthy et al., 2006). Injection of *DvvABCB_19147* dsRNA into larvae resulted in 100% mortality, and injections into pre-pupae led to defects in their development, which occurred during the pupal-adult molt and ultimately resulted in mortality. Injection of *DvvABCB_19147* dsRNA into pre-pupae caused severe abnormalities, leaving individuals partially trapped in their pupal cuticle (Figure 8B). To evaluate the effects of *DvvABCB_19147* RNAi on female fecundity, dsRNA was injected into newly-eclosed adult females. Once they recovered, these females were mated with un-injected males. Interestingly, drastic reduction in egg laying was observed compared to untreated females (Figure 9). Injection of this dsRNA also affected ovary development, causing the ovaries to be underdeveloped and resulting in the failure to produce eggs (Figure 8J).

RNAi knockdown of *DvvABCE_2830* and *DvvABCF_2701* (the ortholog of *TcABCE1* and *TcABCF1*; Broehan et al., 2013) resulted in the deaths of all injected subjects. Prior to death, it was noted that body mass was reduced in the treated individuals and was significantly less than the body mass of larvae treated with buffer alone (Figure 8G, 8H). Injection of *DvvABCE_2830* and *DvvABCF_2701* dsRNA separately into pre-pupae both caused 100% mortality with no adult eclosion (results not shown).

Injection of dsRNA targeting *DvvABCG_3712, DvvABCG_49457,* and *DvvABCG*_13829 resulted in phenotypes similar to those seen in *Tribolium* (Broehan et al 2013) when performing RNAi with dsRNAs targeting the corresponding *T. castaneum* orthologs. Specifically, injection of dsRNA targeting *DvvABCG_13829*, which are most closely related to eye-color genes, gave a white-eye phenotype (Figure 8E). While injection of *DvvABCG_3712* dsRNA into pre-pupae caused developmental defects and resulted in 80% mortality (Figure 8C). Adult females treated with *DvvABCG_3712* dsRNA produced fewer eggs than females injected with buffer (Figure 10), and the eggs were not viable and failed to hatch (Figure 11), even after an additional four weeks of observation (Figure 8I). Injection of *DvvABCG_49457* dsRNA into larvae and pre-pupae resulted in molting defects in both stages, where about 80% of injected larvae died during molt, and development of the surviving pupae was arrested at the pupal stage (Figure 8D).

Injection of dsRNA specific for *DvvABCH_18290* into young larvae and pre-pupae arrested their development as they were preparing to molt and resulted in 100% mortality (Figure 8F), and injected larvae appeared to desiccate prior to death (personal observation). In other species, orthologs of *DvvABCH_18290* are known to transport cuticular lipids that are deposited in the outer epicuticle layer to form a waterproof barrier (Guo et al., 2015; Broehan et al., 2013). Therefore, desiccation is not an unexpected loss-of-function phenotype for these genes.

DISCUSSION

D. virgifera virgifera is a serious agricultural pest, and since its genome sequence is not yet publically available, the transcriptomic data presented here is a particularly valuable genetic resource. Moreover, due to the importance of ABC transporters in resistance development, the 65 ABC transporter genes we identified are expected to be useful in downstream studies. For example, in recent years ABC transporters have become a major focus for research in arthropods because of their roles in xenobiotic transport and insecticide resistance (Lage et al., 2003; Tian et al., 2017; Labbe et al., 2011; Merzendorfer et al., 2014; Dermauw et al., 2014). In particular, ABC transporters have been implicated as targets for *Bt* toxins (Heckel, 2015). Gahan et al., (2010), reported *Bt* resistance to have been genetically linked to *HvABCC2* in *Heliothis virescens* and *HaABCC2* in the cotton bollworm, *Helicoverpa armigera* (Xiao et al., 2014). Recently, Tay et al., (2015), demonstrated that a member of subfamily A, *HaABCA2*, is involved in *Bt* mode of action in *H. armigera*, Thus, highlighting the importance of studying this protein superfamily in *D. virgifera virgifera*.

Here we used RNAi to determine if eight *D. virgifera virgifera* genes, each a putative one-to-one ortholog of a *Tribolium* ABC transporter with a documented RNAi phenotype, served a similar function in *D. virgifera virgifera*. Interestingly, our RNAi results do indeed demonstrate that these eight genes play similar roles in both species. While some loss-of-function phenotypes included a visible developmental defect such as loss of eye pigmentation, others simply caused growth arrest, and/or death. However, one of our most interesting findings was of two ABC transporters whose loss-of-function phenotype results in females that either fail to lay eggs, or whose eggs never hatch. These genes are exciting new targets for RNAi-based management of this pest.

Given the species specificity of RNAi technology, ABC transporters could be ideal targets due to the fact that they may have less sequence conservation than some of the vital genes currently being tested. RNAi against several of our targeted *DvvABC* genes resulted in lethality. For example, targeting genes necessary for survival may also be useful in pest control. *DvvABCA_18330* knockdown gave the same phenotype as seen in *T. castaneum* (Broehan et al., 2013), and resulted in significant rates of death during the pupal-to-adult molt. Severe phenotypes were also seen with RNAi of *DvvABCE_2830* and *DvvABCF_2701*. ABCE and ABCF subfamilies are highly conserved across all phyla; they lack TMDs and therefore do not function as transporters. Instead, ABCE proteins play a role in ribosome recycling and regulation of protein translation (Barthelme et al., 2011), whereas ABCFs play a role in translation (Paytubi et al., 2009). Because of their role in translation and ribosome biosynthesis, ABCE and ABCF proteins are essential for all life stages (Barthelme et al. 2011). These genes are highly conserved across taxa in both sequence and RNAi phenotype (Broehan et al., 2013). Given their likely role in basic cell function, the lethal phenotype from RNAi is not surprising. Targeting *DvvABCG_49457* and *DvvABCH_18290* both caused molting defects that resulted in mortality. This result was consistent with RNAi knock-down of *TcABCG8* in *T. castaneum*, but Broehan et al., (2013) also reported that injected individuals prematurely developed compound eyes during the larval stage; we did not observe an eye phenotype in *D. virgifera virgifera*. This could be because most of the injected *D. virgifera virgifera* larvae failed to develop into the next stage and eventually died, so they may not have had the opportunity to develop compound eyes.

The ABCG proteins are half-transporters that form the second largest ABC subfamily (19 members) in *D. virgifera virgifera* (Table 2). In insects, some of the first ABCGs to be characterized were pigment transporters (*brown*, *scarlet* and *white* genes) in *D. melanogaster* (Sullivan et al., 1974; Mackenzie et al., 1999). Mutants of *white* are characterized by white eyes (i.e. complete loss of eye pigmentation), *brown* mutants by dark brown eyes (i.e. loss of red pigments), and *scarlet* mutants by bright red eyes (i.e. loss of brown pigments) (Mackenzie et al., 1999). Studies have revealed that some ABCG proteins perform a physiological role due to their ability to transport lipids, sterols, and drugs (Tarr et al. 2009). RNAi targetting *DvvABCG_13829* gave white eye-color phenotypes, this is consistent with Broehan et al., (2013). Our findings support that *DvvABCG_13829* is part of the ommochrome pathway, and is likely acting as a heterodimer to import ommochrome pigments into the eyes. The *white* gene is necessary to transport both pigment types in *D. melanogaster*, thus loss of this gene's function results in loss of all visible eye pigments. Not surprisingly, loss of *white* function in *D. virgifera virgifera* results in white eyes, just as it does in *D. melanogaster* and *T. castaneum* (Grubbs et al., 2015). A more interesting result would be RNAi targeting the *ABCG-9A* transcript (*scarlet*) since loss of *scarlet* function in *T. castaneum* results in a white-eyed phenotype, rather than the red-eyed phenotype seen in flies (Lorenzen et al., 2002; Broehan et al., 2013; Grubbs et al., 2015). While the same is likely true in *D. virgifera virgifera*, this hypothesis hasn't yet been confirmed.

In *D. melanogaster* a third ABCG transporter, Brown, is required for proper pigmentation of the eye. In flies, Brown dimerizes with White and transports pteridine-based pigments into the eye. Although a *brown* gene has been found in the *T. castaneum* genome, its function is not known (Grubbs et al., 2015). While our initial survey of the *D. virgifera virgifera* transcriptome failed to idenitify *scarlet* and *brown*, a later assembly revealed a *scarlet* ortholog. In summary, this study provides the first insights into the biology of ABC transporters in *D. virgifera virgifera* and provides a much-needed resource for comparative genomics/transcriptomics with other species and will also greatly aid annotation of the *D. virgifera virgifera* genome when the sequence becomes available.

Table 1: Primer sequences used for dsRNA synthesis and RT-PCR analysis

Gene	Sequences	Purpose
DvvABCA_18330 -F	TAATACGACTCACTATAGGGGAGATGGTGTGACCTGGAGC	dsRNA synthesis
DvvABCA_18330 -R	TAATACGACTCACTATAGGGACGGTGAGCTCGTCGAAAAT	dsRNA synthesis
DvvABCA_18330 -F3	GCATGAGGTTGGATGTCATC	RT-PCR
DvvABCA_18330 -R3	GTCCAGCGGTACTTAAGAGAG	RT-PCR
DvvABCB_19147-F	TAATACGACTCACTATAGGGGGATGCGTTGGTGTGTATGC	dsRNA synthesis
DvvABCB_19147-R	TAATACGACTCACTATAGGGAGGATTCGCCCTTTCGATGG	dsRNA synthesis
DvvABCB_19147-F3	TGGTCGGCAAATGGTAGGAG	RT-PCR
DvvABCB_19147-R3	CACCGTAATGGAGGTTATGCC	RT-PCR
DvvABCE_2830-F	TAATACGACTCACTATAGGGCCTGGGCAACCTTCAACTCA	dsRNA synthesis
DvvABCE_2830-R	TAATACGACTCACTATAGGGTTCCTGGACGGTTTTGTCCC	dsRNA synthesis
DvvABCE_2830-F3	ACACGATCGGCCAAGTAGGTC	RT-PCR
DvvABCE_2830-R3	GGATCCTCCAGATTGGACAGA	RT-PCR
DvvABCF_2701-F	TAATACGACTCACTATAGGGGCCGCTGCTAATAGTGGAGA	dsRNA synthesis
DvvABCF_2701-R	TAATACGACTCACTATAGGGCTGGTCGTGGGAAACGATCA	dsRNA synthesis
DvvABCF_2701-F3	CAGGTACGTGATGAATCTCCAG	RT-PCR
DvvABCF_2701-R3	TGCCAAGGCCACTCTTGCTT	RT-PCR
DvvABCG_3712-F	TAATACGACTCACTATAGGGACGGTATGTCCGCCAATGTT	dsRNA synthesis
DvvABCG_3712-R	TAATACGACTCACTATAGGGGCTCTCCACATGCTACCTCC	dsRNA synthesis
DvvABCG_3712-F3	GCACGTCAGCTTCGGTCTAT	RT-PCR
DvvABCG_3712-R3	GGGGAAAGACAGTCGTCCTG	RT-PCR
DvvABCG_49457-F	TAATACGACTCACTATAGGGTGCTCTCGAAGGAACACTGG	dsRNA synthesis
DvvABCG_49457-R	TAATACGACTCACTATAGGGAAGCCATACAGCGTTGTCGA	dsRNA synthesis
DvvABCG_49457-F3	CGTAGGCAGTATCGAGATGACG	RT-PCR
DvvABCG_49457-R3	CAACCAACAGCACGATCTGTC	RT-PCR
DvvABCG_13829-F	TAATACGACTCACTATAGGGCGACCTCTGGCTTGGATTCT	dsRNA synthesis
DvvABCG_13829-R	TAATACGACTCACTATAGGGCCATTCACCAACCTAGCTCC	dsRNA synthesis
DvvABCG_13829-F3	CGTTCGCCTCAGAAATGCTG	RT-PCR
DvvABCG_13829-R3	CCATCAGCATCCTCTCTGGTT	RT-PCR
DvvABCH_18290-F	TAATACGACTCACTATAGGGGTGTTACACCGGCTACCCAA	dsRNA synthesis
DvvABCH_18290-R	TAATACGACTCACTATAGGGCCCGTTCTAACCGCCTCTTT	dsRNA synthesis
DvvABCH_18290-F3	CCCCATTTTACGGTACTTGCTAGG	RT-PCR
DvvABCH_18290-R3	AGCGGTAAGCTCCTGGCAGAA	RT-PCR
DvvRPS6-F	TTGAAGAAGAAGAGGTGC	RT-PCR
DvvRPS6-R	ACCTACGCCTCTTCAACATC	RT-PCR

Underlined sequence corresponds to T7 promoter

Table 2: Classification of *D. virgifera virgifera* ATP binding cassette (ABC) transporters

	***D. v. virgifera* transcript**			**Nearest *T. castaneum* ortholog**		
Subfamily	**Gene ID**	**Size (AA)**	**Contig Name**	**Accession**	**E-value**	**Identity(%)**
DvvABC-A	TBD	1644	*18330*	XP_008199153.1	0.0	60
DvvABC-A	TBD	1708	*50718*	KXZ75645.1	0.0	58
DvvABC-A	TBD	1757	*49125*	XP_008195104.1	0.0	42
DvvABC-A	TBD	1641	*266167*	XP_008195056.1	0.0	53
DvvABC-B	TBD	1247	*21313*	XP_967244.3	0.0	60
DvvABC-B	TBD	1257	*17742*	XP_008191266.1	0.0	63
DvvABC-B	TBD	667	*19147*	XP_015834348.1	0.0	71
DvvABC-B	TBD	716	*39715*	XP_001813375.1	0.0	75
DvvABC-B	TBD	834	*9796*	XP_008194672.1	0.0	76
DvvABC-B	TBD	657	*13664*	XP_008190820.1	0.0	60
DvvABC-B	TBD	682	*17837*	XP_972133.2	0.0	69
DvvABC-C	TBD	1268	*41801*	XP_969849.1	0.0	37
DvvABC-C	TBD	1257	*44708*	XP_015836131.1	0.0	44
DvvABC-C	TBD	1252	*48952*	XP_971802.2	0.0	48
DvvABC-C	TBD	1285	*17573*	XP_015836083.1	0.0	44
DvvABC-C	TBD	1556	*51687*	XP_008197283.1	0.0	73
DvvABC-C	TBD	1297	*21020*	EFA03948.2	0.0	52
DvvABC-C	TBD	1234	*222633*	XP_971802.2	0.0	46
DvvABC-C	TBD	1343	*18126*	EFA04472.2	0.0	57
DvvABC-C	TBD	1374	*49513*	XP_008193834.1	0.0	44
DvvABC-C	TBD	1343	*14070*	XP_008194093.1	0.0	51
DvvABC-C	TBD	1350	*22628*	XP_008193834.1	0.0	56
DvvABC-C	TBD	1345	*20002*	EFA04472.2	0.0	58
DvvABC-C	TBD	1364	*7536*	XP_008193834.1	0.0	61
DvvABC-C	TBD	1377	*47333*	XP_008193834.1	0.0	59
DvvABC-C	TBD	753	*49618*	XP_015835265.1	0.0	76
DvvABC-C	TBD	1535	*45163*	XP_008192060.1	0.0	61
DvvABC-C	TBD	1082	*43960*	XP_973693.2	0.0	48
DvvABC-C	TBD	1224	*48940*	XP_015836125.1	0.0	42
DvvABC-C	TBD	1165	*217405*	XP_015835266.1	0.0	50
DvvABC-C	TBD	871	*10132*	XP_008200429.1	0.0	55
DvvABC-C	TBD	1258	*48300*	XP_971632.1	0.0	44
DvvABC-C	TBD	1324	*47673*	XP_969781.1	0.0	72
DvvABC-C	TBD	1258	*5345*	XP_015836093.1	0.0	40
DvvABC-C	TBD	1331	*22413*	XP_969781.1	0.0	55
DvvABC-C	TBD	1260	*18709*	XP_008193834.1	0.0	53
DvvABC-C	TBD	1329	*21941*	XP_969849.1	0.0	48
DvvABC-C	TBD	1324	*15305*	XP_969781.1	0.0	64
DvvABC-C	TBD	1320	*12562*	XP_972534.1	0.0	49
DvvABC-C	TBD	1307	*10642*	XP_008195764.1	0.0	60
DvvABC-C	TBD	1308	*41602*	XP_008193119.1	0.0	50
DvvABC-C	TBD	1318	*12703*	KYB27686.1	0.0	54
DvvABC-C	TBD	1310	*14968*	EFA03655.1	0.0	49
DvvABC-D	TBD	755	*11014*	XP_971218.1	0.0	77
DvvABC-D	TBD	658	*11628*	XP_015838765.1	0.0	78
DvvABC-E	TBD	609	*2830*	XP_968009.1	0.0	92
DvvABC-F	TBD	922	*2701*	XP_971562.1	0.0	90
DvvABC-F	TBD	624	*802*	XP_966990.1	0.0	89
DvvABC-F	TBD	711	*9935*	XP_972814.1	0.0	84

Table 2. Continued

D. v. virgifera transcript				Nearest *T. castaneum* ortholog		
Subfamily	**Gene ID**	**Size (AA)**	**Contig**	**Accession**	**E-value**	**Identity (%)**
DvvABC-G	TBD	660	9811	XP_008192053.1	0.0	69
DvvABC-G	TBD	668	3712	XP_001813184.1	0.0	77
DvvABC-G	TBD	720	14042	XP_973458.1	0.0	76
DvvABC-G	TBD	652	10897	XP_008192849.1	0.0	63
DvvABC-G	TBD	641	22358	XP_015834971.1	0.0	63
DvvABC-G	TBD	604	23081	XP_971735.1	0.0	54
DvvABC-G	TBD	638	13051	XP_015834339.1	0.0	60
DvvABC-G	TBD	622	38769	XP_973526.1	0.0	54
DvvABC-G	TBD	658	13829	NP_001034521.1	0.0	59
DvvABC-G	TBD	941	49457	XP_968472.1	0.0	73
DvvABC-G	TBD	643	36869	XP_968555.2	0.0	71
DvvABC-G	TBD	651	79525	NP_001306193.1	0.0	61
DvvABC-H	TBD	714	20789	XP_015839275.1	0.0	43
DvvABC-H	TBD	704	18290	XP_008198313.1	0.0	84
DvvABC-H	TBD	796	5118	XP_973444.1	0.0	43
DvvABC-H	TBD	763	11818	XP_967359.1	0.0	72

Table 3: Results of RNAi of select ABC transporters

DvvABCA_18330	Pre-pupal	60%	Deformed wings & elytra	8A
DvvABCB_19147	Larval	100%	Lethal	NS
	Pre-pupal	100%	Defect in pupal-adult molt	8B
	Eclosed females	0%	Malformed ovaries; low egg lay	8J
DvvABCE_2830	Larval	100%	Lethal	8G
DvvABCF_2701	Larval	100%	Lethal	8H
DvvABCG_3712	Pre-pupal	80%	Lethal; pupal developmental arrest	8C
	Eclosed females	0%	Prevented embryonic development	8I
DvvABCG_13829	Pre-pupal	0%	Pigmentation defect; white eyes	8E
DvvABCG_49457	Larval	100%	Lethal at molting	NS
	Pre-pupal	80%	Lethal pupal developmental arrest	8D
DvvABCH_18290	Larval	100%	Lethal at molting	NS
	Pre-pupal		Lethal at adult-pupal transition	8F

NS = image not shown

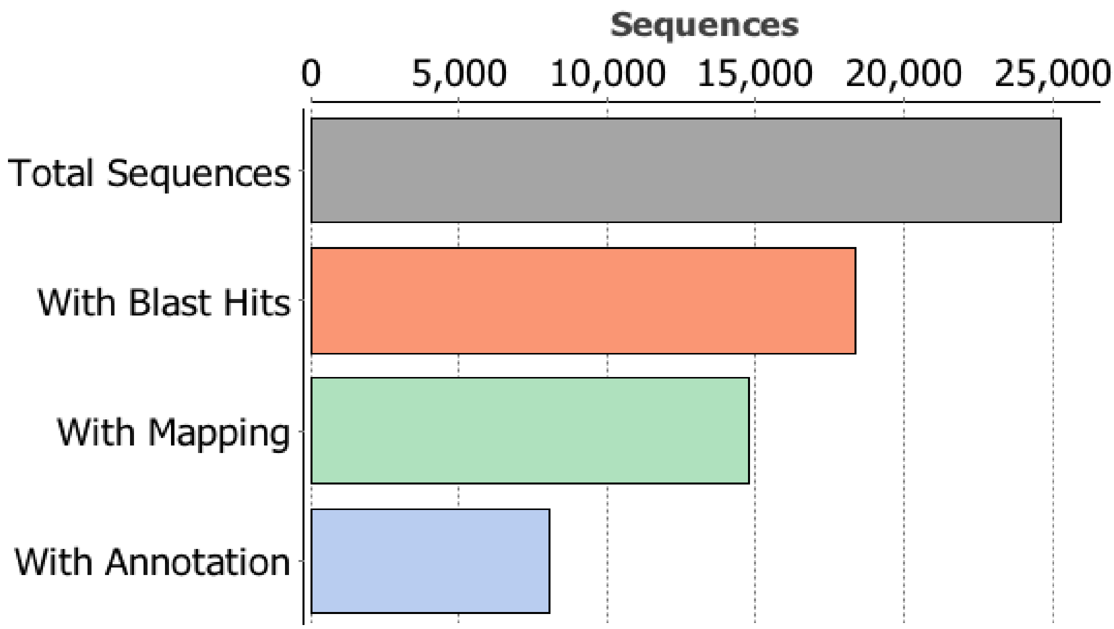


Figure 1: Blast2GO annotation results for the combined *D. virgifera virgifera* transcriptome. Among the 25,296 contigs used for the BLASTx search against the NCBI nr protein database, BLASTx analysis revealed over 18,343 contigs shared significant identity with known arthropod proteins.

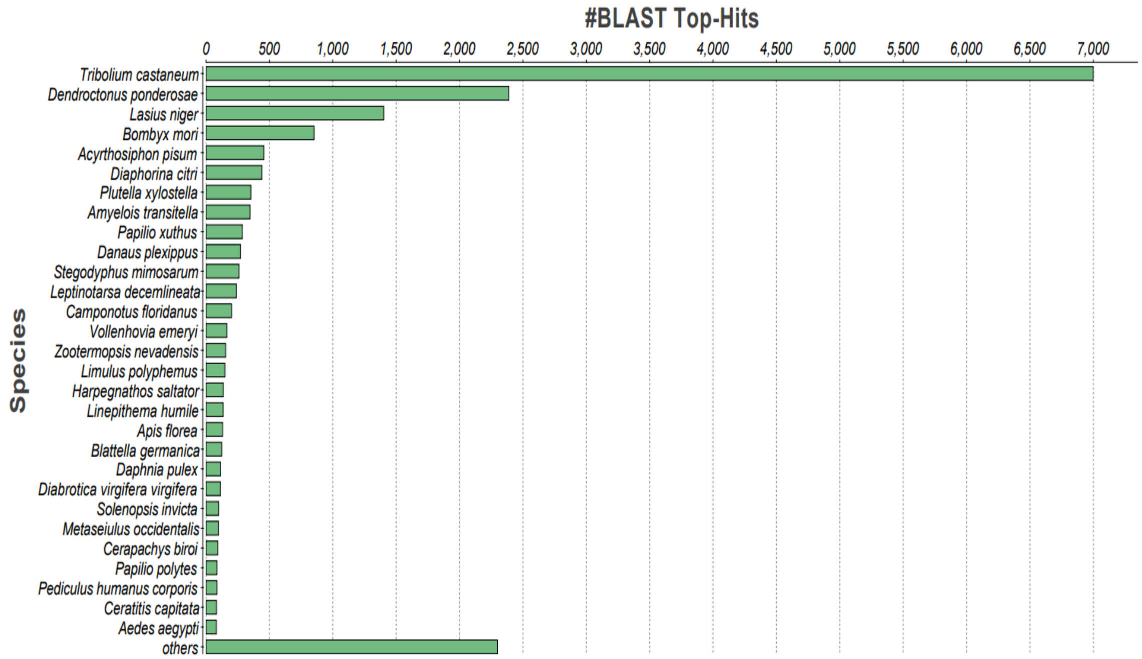


Figure 2: Top-Hits species distribution from Blast2GO. The most represented species in the top scoring BLASTx-matches the *D. virgifera virgifera* contigs when compared to an arthropod-specific database.

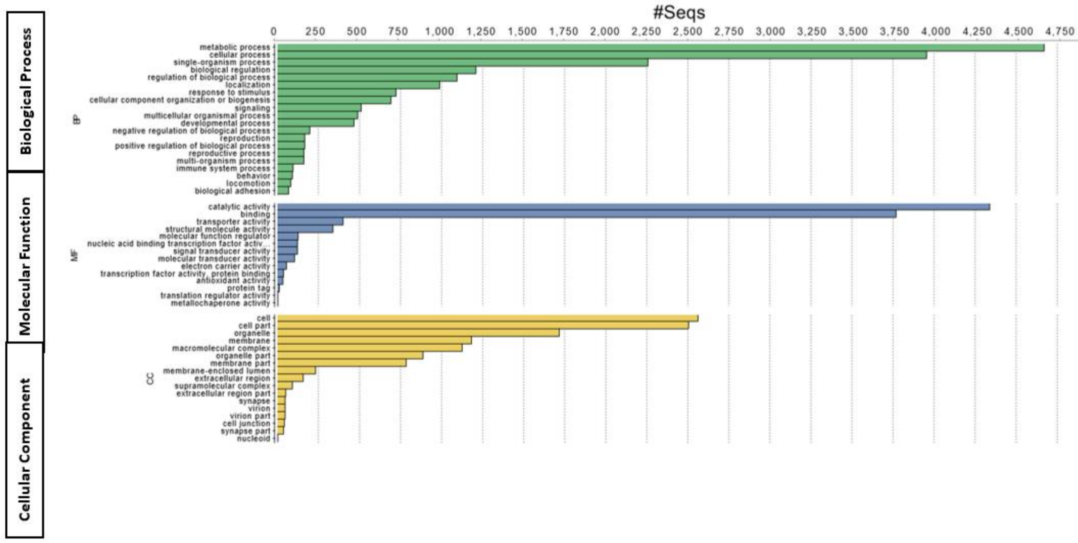


Figure 3: Gene ontology classification of the *D. virgifera virgifera* transcriptome.

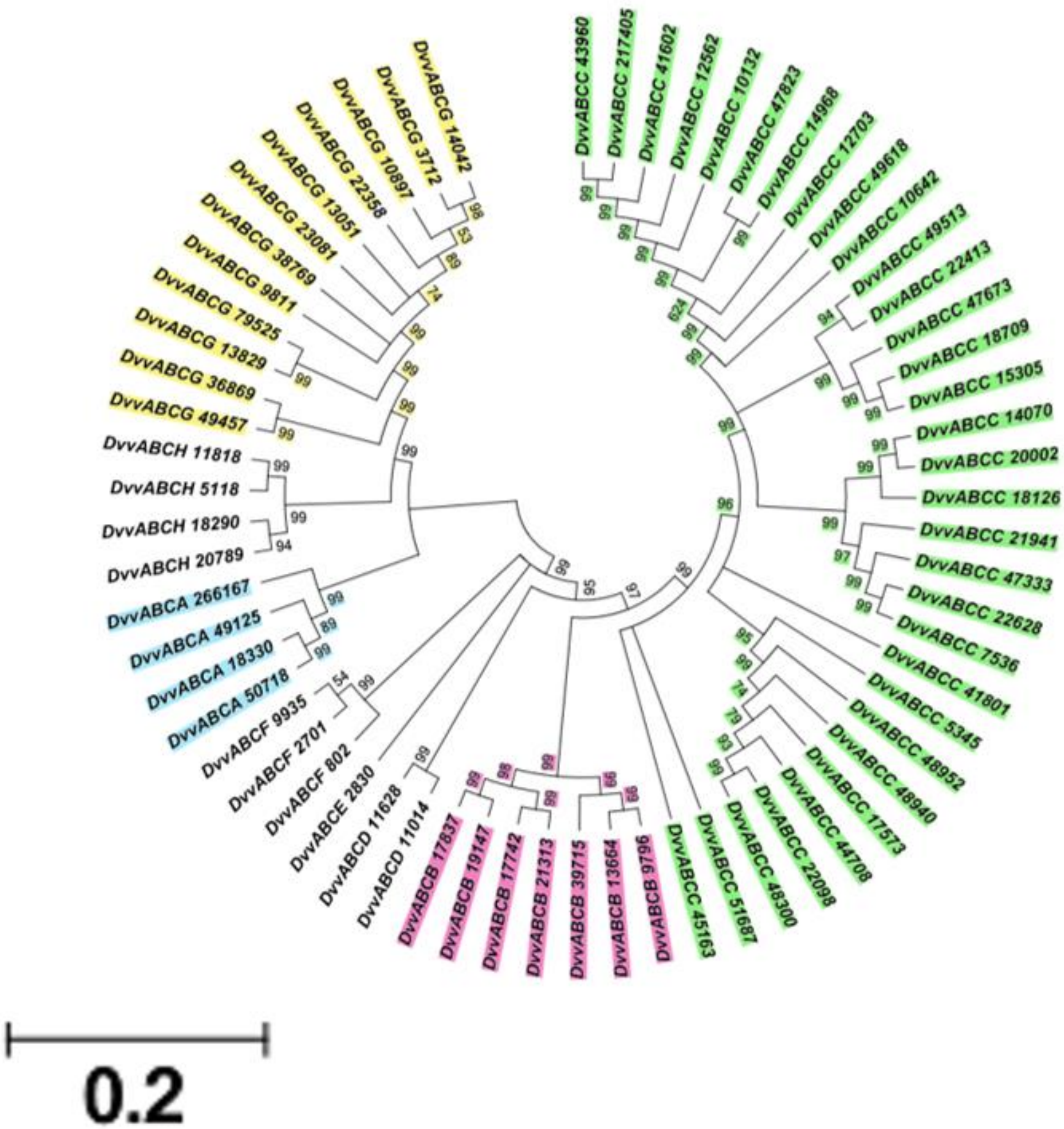


Figure 4: Phylogeny of ABC transporters in *D. virgifera virgifera*. The scale bar indicates an evolutionary distance of 0.5 amino acid substitutions per site. Bootstrap values are given in percentages at the internodes.

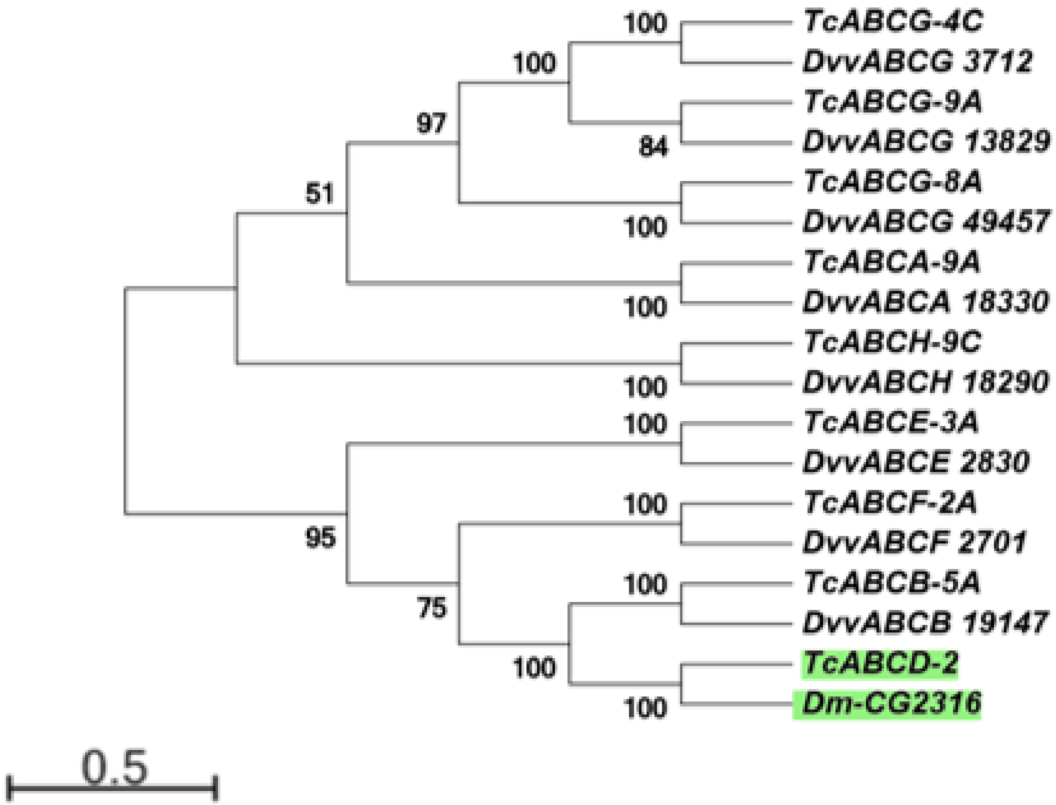


Figure 5: Phylogenetic relationships of select *D. virgifera virgifera* ABC transporters targeted with RNAi to their *T. castaneum* orthologs. The scale bar indicates an evolutionary distance of 0.5 amino acid substitutions per site. Bootstrap values are given in percentages at the internodes. *T. castaneum* and *Drosophila melanogaster* were used as an outgroup.

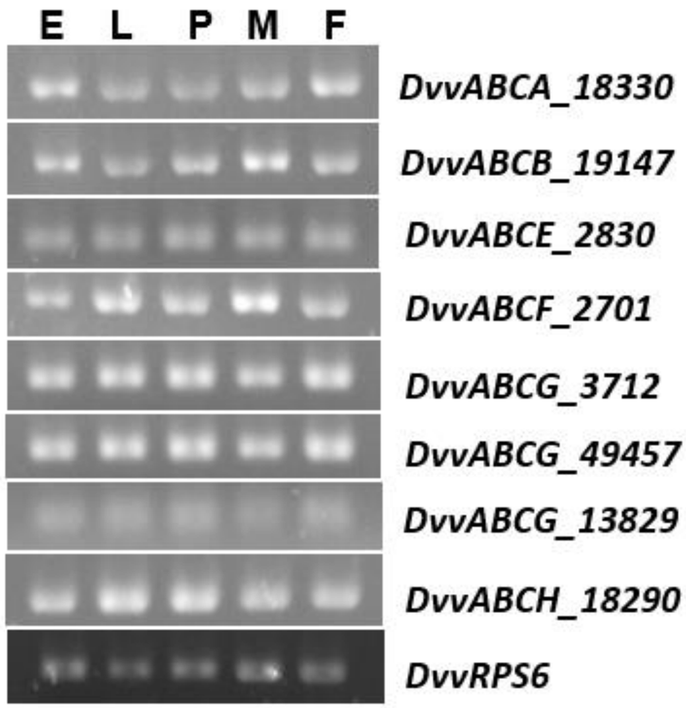


Figure 6: Developmental expression profile of select *D. virgifera virgifera* ABC transcripts. RT-PCR results from RNA isolated from E (eggs); L (larvae); P (pupae); M (Male); F (female). *RPS6* was used as a positive control.

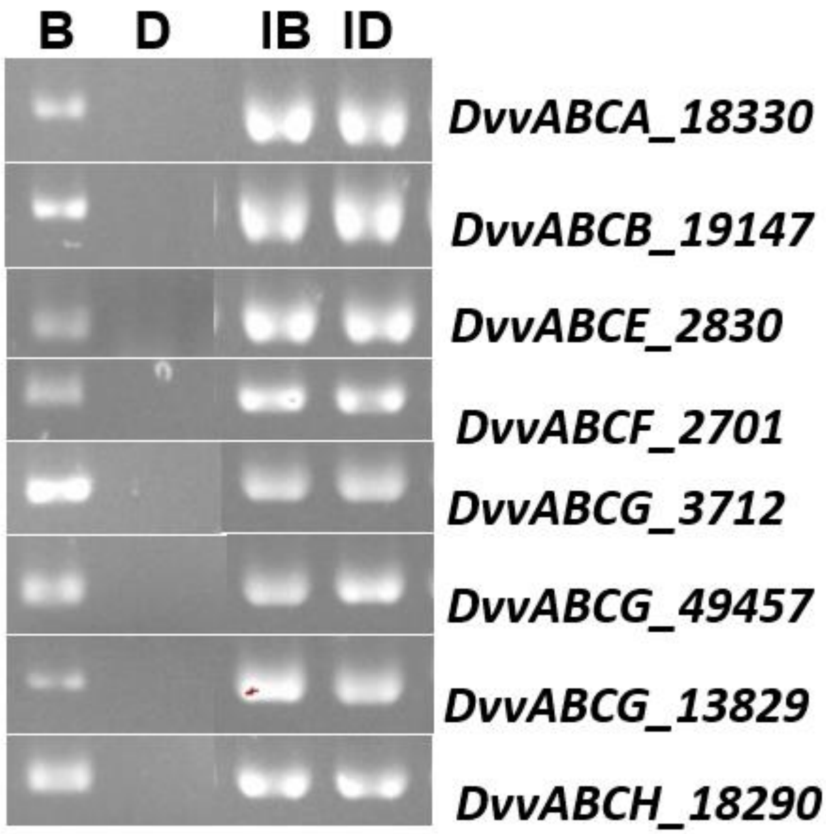


Figure 7: RNAi gene knockdown confirmed by RT-PCR. Target RNA levels were evaluated five days after injection. Gene names given next to gel samples. The lanes are as follows: RNA from pools of buffer injected (B) individuals, dsRNA injected (D) individuals, and the highly-conserved housekeeping gene, *DvvRPS6*, was used to assess template quality in buffer injected (IB) as well as dsRNA injected (ID) individuals.

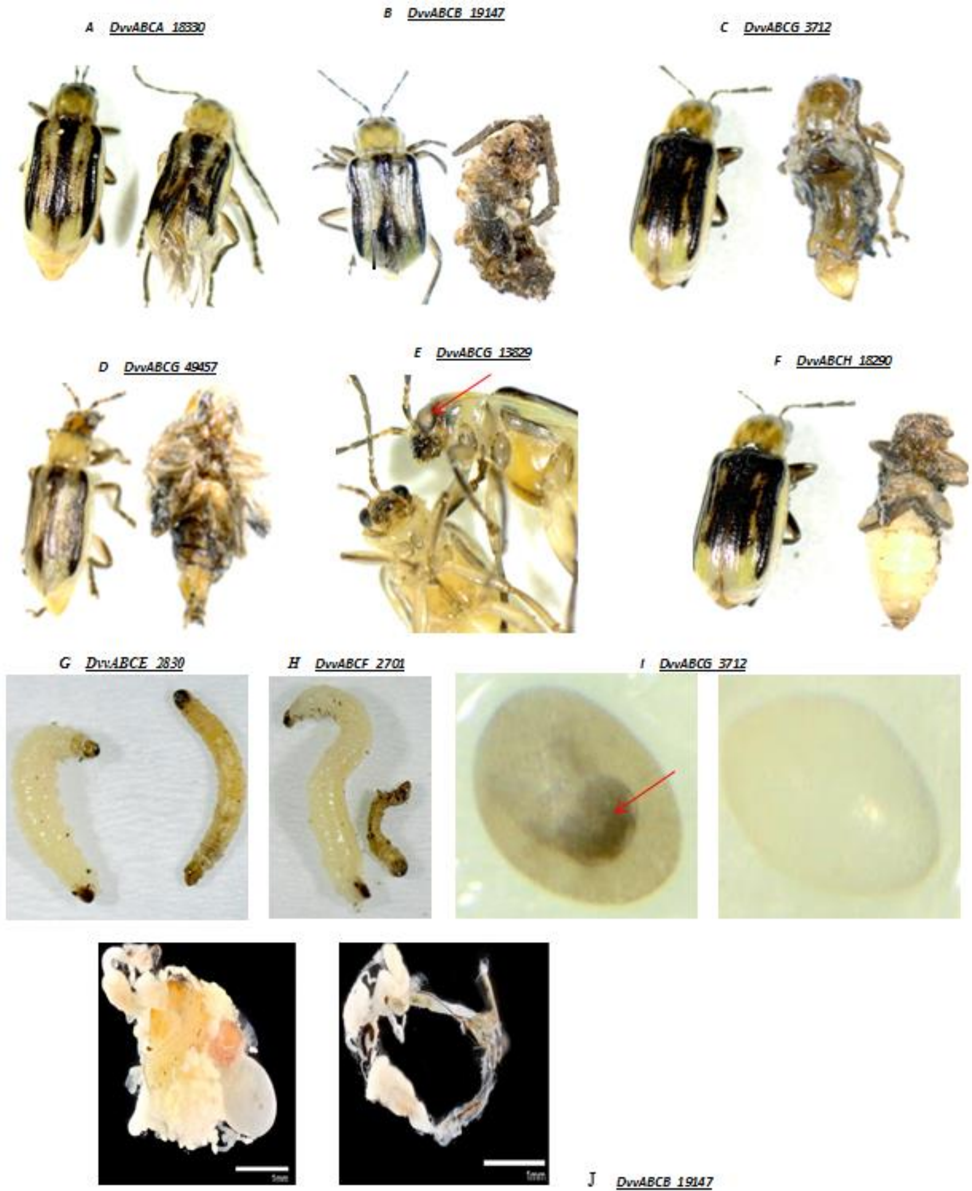


Figure 8: Developmental effects of ABC transporter knockdown by RNAi. Buffer-injected control (right) of comparable stage individuals are shown next to experimentals in each image. (A) *DvvABCA-18330* – pre-pupal injection caused defects in adult wing development. (B) *DvvABCB-19147* –pre-pupal injection caused molting defects during eclosion. (C, D&F) *DvvABCG-3712, DvvABCG-49457* and *DvvABCH-18290* – larval and pre-pupal injections caused severe molting defects during eclosion (individuals injected as pre-pupae shown). (E) *DvvABCG-13829* – pre-pupal injections caused loss of eye pigment (arrow). (G-H) *DvvABCE-2830* and *DvvABCF-2701* – larval injection caused reduction in body mass and death before molting. (I) *DvvABCG-3712* – injections into adult females prevented embryonic development in oviposited eggs (arrow indicates developed embryo). (J) *DvvABCB-19147* – injections into adult females caused disruption of ovary development.

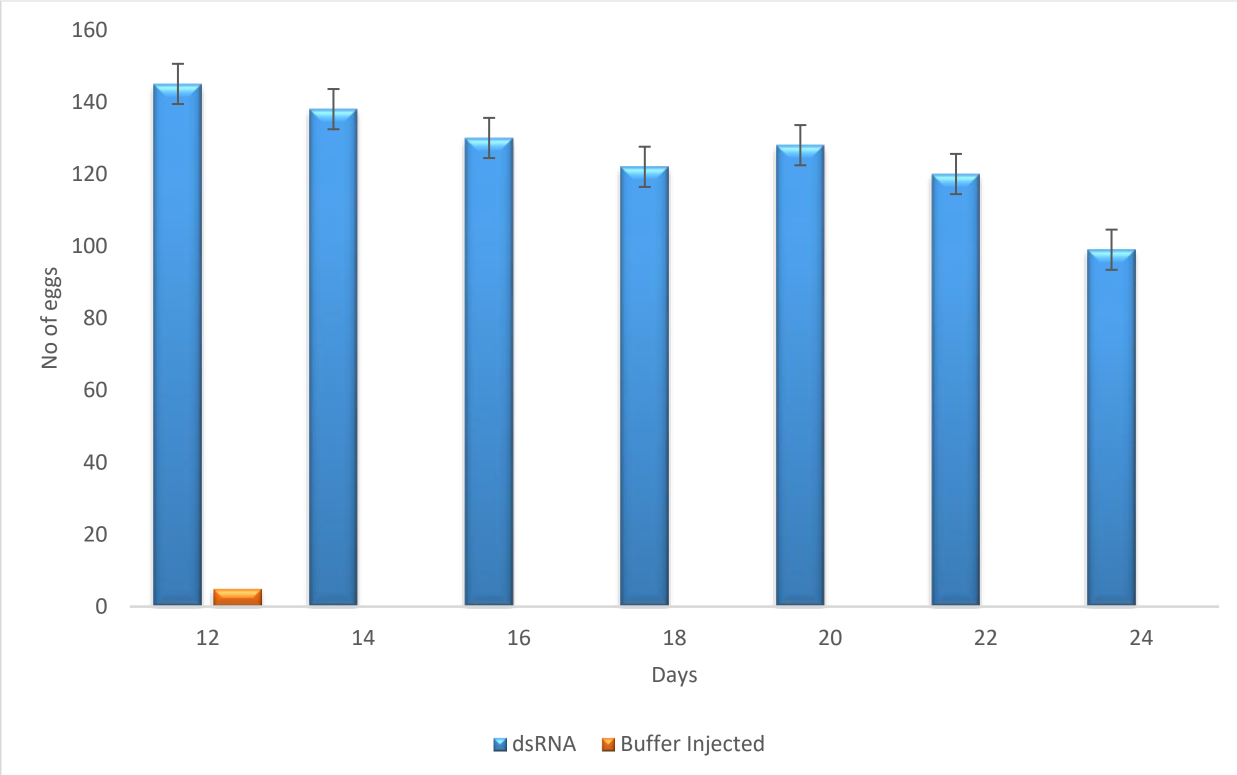


Figure 9: Effects of *DvvABCB_19147* RNAi on oviposition. Females injected with *DvvABCB_19147* dsRNA failed to lay eggs, while buffer-injected females lay near-normal numbers. The eggs laid within a period of two weeks were counted every other day.

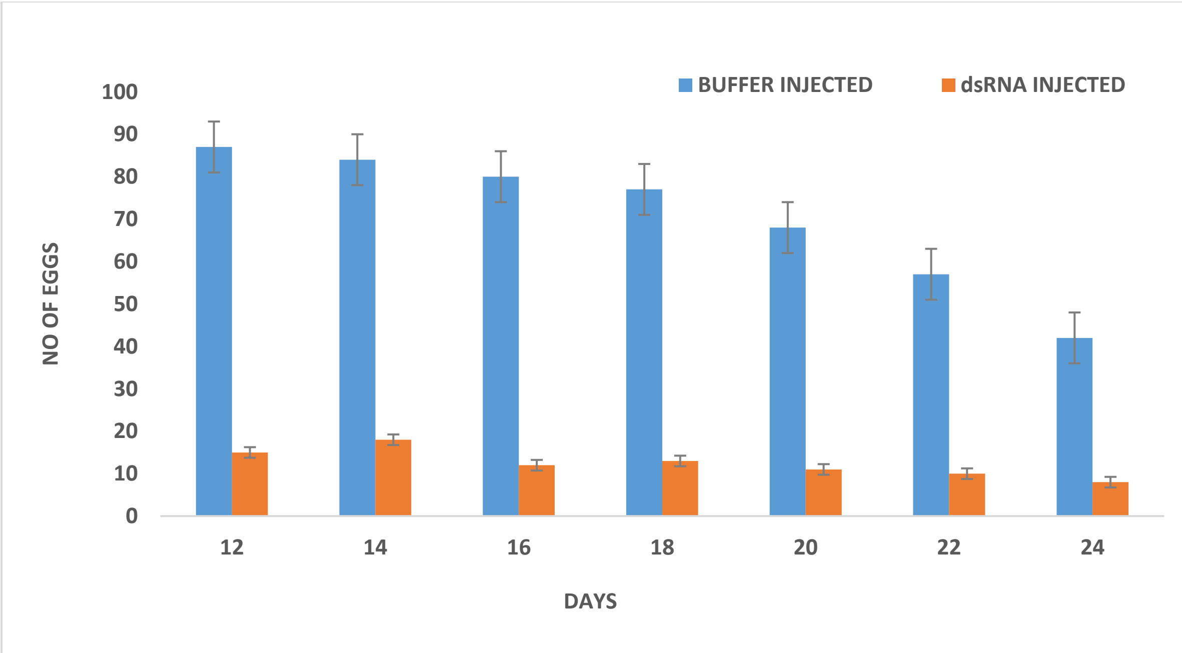


Figure 10: Effects of *DvvABCG_3712*RNAi on egg laying. Females injected with *DvvABCG_3712* dsRNA failed to lay eggs, while buffer-injected females lay near-normal numbers. The eggs laid within a period of two weeks were counted every other day.

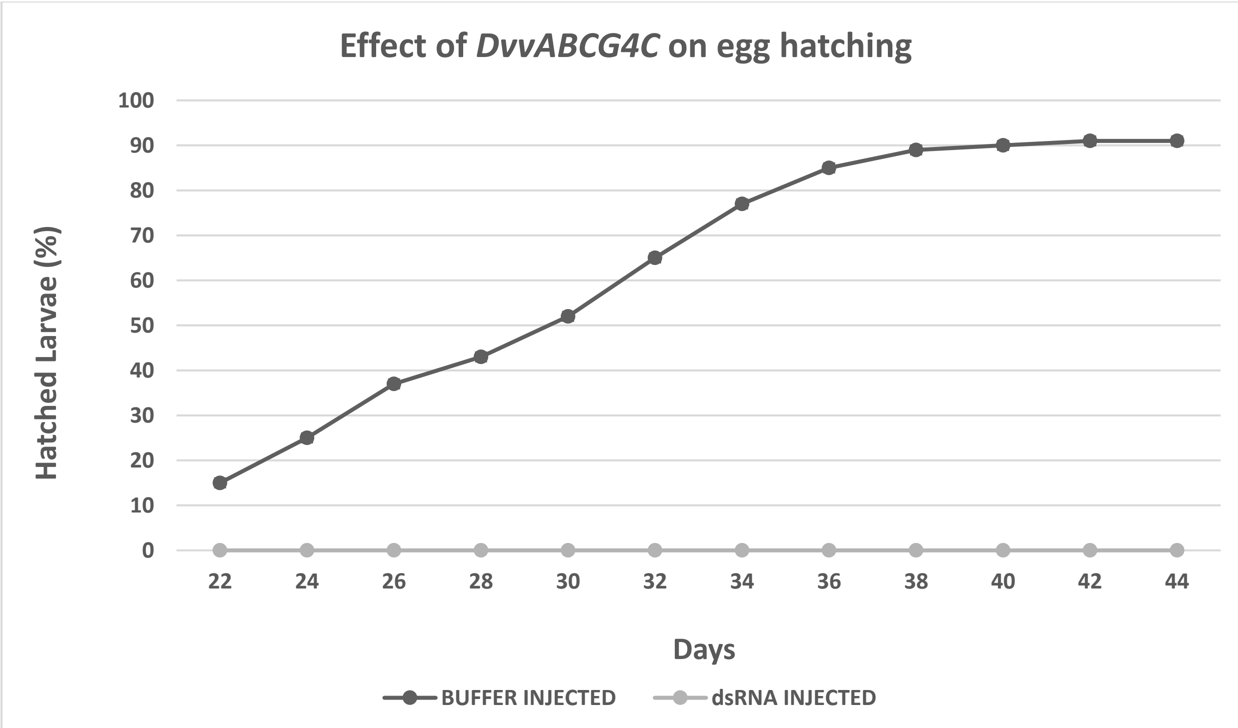


Figure 11: Effect of *DvvABCG_3712* RNAi on egg hatching. Eggs laid by females treated with *DvvABCG-4C* dsRNA failed to hatch, while those laid by buffer-injected females hatched successfully.

REFERENCES

Aurade, R.M., Aurade, Jayalakshmi, S. & Sreeramulu, K. (2010) P-glycoprotein ATPase from the resistant pest, *Helicoverpa armigera*: purification, characterization and effect of various insecticides on its transport function *Biochem. Biophys. Acta Biomembr*. **1798**, 1135-1143

Barthelme, D., Dinkelaker, S., Albers, S.V., Londei, P., Ermler, U. & Tampé, R (2011) Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. *Proc Natl Acad Sci U S A*. **108**, 3228–33.

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P. & Ilagan, O. (2007) Control of coleopteran insect pests through RNA interference. *Nat Biotechnol* **25**, 1322–1326.

Bretschneider, A., Heckel, D.G. & Vogel, H. (2016) Know your ABCs: Characterization and gene expression dynamic s of ABC transporters in the polyphagous herbivore *Helicoverpa armigera*. *Insect Biochem Mol Biol*. **72**, 1–9

Broehan, G., Kroeger, T., Lorenzen, M. & Merzendorfer, H. (2013) Functional analysis of the ATP-binding cassette (ABC) transporter gene family of Tribolium castaneum. *BMC Genomics,* **14**, 6

Buss, D. & Callaghan, A. (2008) Interaction of pesticides with p-glycoprotein and other ABC proteins: a survey of the possible importance to insecticide, herbicide and fungicide resistance *Pestic. Biochem. Physiol*., **90**, 141-153

Chu, C.C., Sun, W., Spencer, J.L., Pittendrigh, B.R., & Seufferheld, M.J. (2014) Differential of RNAi treatments on field population of the western corn rootworm. *Pesticide biochemistry and physiology,* **110**, 1-6.

Coates, B.S., Alves, A.P., Wang, H.C., Walden, K.K.O., French, B.W. & Miller, N.J. (2012) Distribution of genes and repetitive elements in the *Diabrotica virgifera virgifera* genome estimated using BAC sequencing. *J Biomed Biotechnol*. 604076, 9pp

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, ***21,*** 3674-3676.

Dean, M., Annilo, T. & (2005) Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annu. Rev. Genomics Hum. Genet***. 6,** 123-142.

Dean, M., Rzhetsky, A. & Allikmets, R. (2001) The human ATPbinding cassette (ABC) transporter superfamily. *Genome Res* **11,** 1156–1166.

Dermauw, W. & Van Leeuwen, T. (2014) The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. *Insect Biochem. Mol. Biol*. **45,** 89–110.

Ewart, G.D., Cannell, D., Cox, G.B. & Howells, A.J. (1994) Mutational analysis of the traffic ATPase (ABC) transporters involved in uptake of eye pigment precursors in *Drosophila melanogaster*. Implications for structure-function relationships. *J. Biol. Chem.* **269,** 10370-10377.

Gahan, L.J., Pauchet, Y., Vogel, H. & Heckel, D.G. (2010) An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLoS Genet*. **6**, e1001248.

Gassmann, A., Petzold-Maxwell, J., Keweshan, R. & Dunbar, M. (2011) Field-evolved resistance to *Bt* maize by Western corn rootworm. *PLoS one* **6**, 622-629.

George, A.M. & Jones, P.M. (2012) Perspectives on the structure-function of ABC transporters: the Switch and Constant Contact models. *Prog Biophys Mol Biol* **109**, 95–107.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J. & Conesa, A. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research*, ***36,*** 3420-3435.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. & Chen, Z. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, ***29***, 644-652.

Gray, M.E., Sappington, T.W., Miller, N.J., Moeser, J. & Bohn, M.O. (2009) Adaptation and invasiveness of western corn rootworm: intensifying research on a worsening pest. *Annu Rev Entomol* **54**, 303–321.

Grubbs, N., Haas, S., Beeman, R.W. & Lorenzen, M.D. (2015) The ABCs of eye color in *Tribolium castaneum*: orthologs of the *Drosophila* white, scarlet, and brown Genes. *Genetics*. **99**, 749-59.

Guo, Z., Kang, S., Zhu, X., Xia, J., Wu, O., Wang, S., Xie, W. & Zhang, Y. (2015) Down-regulation of a novel ABC transporter gene (Pxwhite) is associated with Cry1Ac resistance in the diamondback moth, *Plutella xylostella* (L.). *Insect Biochem. Mol. Biol*. **15**, 30-40

Heckel, D.G. (2015) Roles of ABC proteins in the mechanism and management of *Bt* resistance. In: Soberon, M., Gao, A., Bravo, A. (Eds.), Bt Resistance: Characterization and Strategies for GM Crops Producing *Bacillus Thuringiensis* Toxins. CABI, Wallingford, UK, 98-106.

Higgins, C.F. (1992) ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8,** 67–113.

Labbe, R., Caveney, S., Donly, C. & (2011) Genetic analysis of the xenobiotic resistance associated ABC gene subfamilies of the Lepidoptera. *Insect Mol Biol*. **20,** 243–56.

Lage, H. (2003) ABC-transporters: implications on drug resistance from microorganisms to human cancers. *Int. J. Antimicrob. Agents* **22**, 188-199.

Levine, E., Spencer, J.L., Isard, S.A., Onstad, D.W. & Gray, M. E. (2002) Adaptation of the western corn rootworm to crop rotation: evolution of a new strain in response to a management practice. *American Entomologist* **48**, 94-107.

Levine, E., Oloumi-Sadeghi, H., 1991) Management of diabroticite rootworms in corn. Annu. Rev Entomol. 36, 229 – 255.

Li, H., Khajuria, C., Rangasamy, M., Gandra, P., Fitter, M., Geng, C., Woosely, A., Hasler, J., Schulenberg, G., Worden, S., McEwan, R., Evans, C., Siegfried, B. & Narva, K.E., (2015) Long dsrna but not sirna initiates rnai in western corn rootworm larvae and adults. *J Appl Entomol,* **139**, 432–445.

Liu, S., Zhou, S., Tian, L., Guo, E., Luan, Y., Zhang, J. & Li, S. (2011) Genome-wide identification and characterization of ATP-binding cassette transporters in the silkworm, *Bombyx mori*. *BMC Genomics*. **12**, 491.

Lorenzen, M.D., Brown, S.J., Denell, R.E. & Beeman, R.W. (2002a) Cloning and characterization of the *Tribolium castaneum* eye-color genes encoding tryptophan oxygenase and kynurenine 3-monooxygenase. *Genetics* **160**, 225–234.

Krishnamurthy, P.C., Du, G., Fukuda, Y., Sun, D., Sampath, J., Mercer, K.E., Wang, J., Sosa-Pineda, B., Murti, K.G. & Schuetz, J.D. (2006) Identification of a mammalian mitochondrial porphyrin transporter. *Nature.* **443**, 586–9.

Krysan, J.L. (1986) Introduction: biology, distribution, and identification of pest *Diabrotica*. In Methods for the Study of Pest *Diabrotica*, Springer, New York, 1–23.

Mackenzie, S.M., Brooker, M.R., Gill, T.R., Cox, G.B., Howells, A.J. & Ewart, G.D. (1999) Mutations in the white gene of *Drosophila melanogaster* affecting ABC transporters that determine eye colouration. *Biochim Biophys Acta*. **1419**, 173–85.

Mao, Y.B., Cai, W.J., Wang, J.W., Hong, G.J., Tao, X.Y., Wang, L.J., Huang, Y.P. & Chen, X.Y. (2007) Silencing a cotton bollworm *P450 monooxygenase* gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nature Biotechnol.* **25**, 1307–1313.

Marra, M.C., Piggott, N.E. & Goodwin, B.K. (2012) The impact of corn rootworm protected biotechnology traits in the United States. *AgBioForum*. **15**, 217-230.

McKenna, D.D., Scully, E.D., Pauchet, Y., Hoover, K., Kirsch, R., Geib, S.M., Mitchell, R.F., Waterhouse, R.M., Ahn, S.J., Arsala, D., Benoit, J.B., Blackmon, H., Bledsoe, T., Bowsher, J.H., Busch, A., Calla, B., Chao, H., Childers, A.K., Childers, C., Clarke, D.J., Cohen, L., Demuth, J.P., Dinh, H., Doddapaneni, H., Dolan, A., Duan, J.J., Dugan, S., Friedrich, M., Glastad, K.M., Goodisman, M.A., Haddad, S., Han, Y., Hughes, D.S., Ioannidis, P., Johnston, J.S., Jones, J.W., Kuhn, L.A., Lance, D.R., Lee, C.Y,, Lee, S.L., Lin, H., Lynch, J.A., Moczek, A.P., Murali, S.C., Muzny, D.M., Nelson, D.R., Palli, S.R., Panfilio, K.A., Pers, D., Poelchau, M.F., Quan, H., Qu, J., Ray, A.M., Rinehart, J.P., Robertson, H.M., Roehrdanz, R., Rosendale, A.J., Shin, S., Silva, C., Torson, A.S., Jentzsch, I.M., Werren, J.H., Worley, K.C., Yocum, G., Zdobnov, E.M., Gibbs, R.A. & Richards, S. (2016) Genome of the Asian longhorned beetle (Anoplophora glabripennis), a globally significant invasive species, reveals key functional and evolutionary innovations at the beetle-plant interface. *Genome Biol.* **11**, 227.

Meinke, L., Siegfried, B., Wright, R. & Chandler, L. (1998) Adult susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to selected insecticides. *J. of Economic Entomol.* **91**, 594-600.

Merzendorfer, H. (2014) ABC transporters and their role in protecting insects from pesticides and their metabolites. *Adv. Insect Physiol*. **46,** 1–72.

Miller, N.J., Guillemaud, T., Giordano, R., Siegfried, B.D., Gray, M.E., Meinke, L.J. & Sappington, T.W. (2009) Genes flow and adaptation of *Diabrotica virgifera virgifera*. *Agric For Entomol*. **11**, 47–60.

Narva, K.E., Siegfried, B.D. & Storer, N.P. (2013) Transgenic approaches to western corn rootworm control. *Adv Biochem Eng Biotechnol* **136**, 135–162.

Paytubi, S., Wang, X., Lam, Y.W., Izquierdo, L. & Hunter, M.J. (2009) ABC50 promotes translation initiation in mammalian cells. *J Biol Chem* **284**, 24061–24073.

Pondarré, C., Antiochos, B.B., Campagna, D.R., Clarke, S.L., Greer, E.L., Deck, K.M., McDonald, A., Han. A.P., Medlock, A., Kutok, J.L., Anderson, S.A., Eisenstein, R.S. & Fleming, M.D. (2006) The mitochondrial ATP-binding cassette transporter Abcb7 is essential in mice and participates in cytosolic iron-sulfur cluster biogenesis. *Hum Mol Genet*. **15**, 53–64

Quazi, F. & Molday, R.S. (2011) Lipid transport by mammalian ABC proteins. *Essays Biochem* **50**, 265–290.

Rangasamy, M. & Siegfried, B.D. (2012) Validation of RNA interference in western corn rootworm *Diabrotica virgifera virgifera* LeConte (Coleoptera: Chrysomelidae) adults. *Pest Management Science,* **68**, 587–591.

Rees, D.C., Johnson, E. & Lewinson, O. (2009) ABC transporters: the power to change. *Nat. Rev. Mol. Cell Biol.* **10**, 218-227.

Richards, S., Gibbs, R.A., Weinstock, G.M., Brown, S.J. & Denell, R. (2008) The genome of the model beetle and pest *Tribolium castaneum. Nature* **452**, 949-955.

Sammons, A., Edwards, R., Bledsoe, L., Boeve, P. & Stuart, J. (1997) Behavioral and feeding assays reveal a western corn rootworm (Coleoptera: Chrysomelidae) variant that is attracted to soybean. *Environ Entomol*. **26**, 1336–1342

Sappington, T.W., Siegfried, B.D. & Guillemaud, T. (2006) Coordinated *Diabrotica* genetics research: accelerating progress on an urgent insect pest problem. *Am Entomol* **52**, 90–97.

Saurin, W., Hofnung, M. & Dassa, E. (1999) Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J Mol Evol*. **48**, 22–41.

Sturm, A., Cunningham, P.& Dean, M. (2009) The ABC transporter gene family of *Daphnia pulex*. *BMC Genomics*. **10**, 170.

Sullivan, D.T., Bell, L.A., Paton, D.R. & Sullivan, M.C. (1980) Genetic and functional analysis of tryptophan transport in Malpighian tubules of Drosophila. *Biochem Genet* **18**, 1109–1130

Sullivan, D.T., Grillo, S.L. & Kitos, R.J. (1974) Subcellular localization of the first three enzymes of the ommochrome synthetic pathway in *Drosophila melanogaster*. *J Exp Zool*. **188**, 225–33.

Tarr, P.T., Tarling, E.J., Bojanic, D.D., Edwards, P.A. & Baldan, A. (2009). Emerging new paradigms for ABCG transporters. *Biochim Biophys Acta*. *1791*, **584**–593.

Tay, W.T., Mahon, R.J., Heckel, D.G., Walsh, T.K., Downes, S., James, W., Lee, S.-F., Reineke, A., Williams, A.K. & Gordon, K.H.J. (2015) Insect resistance to *Bacillus thuringiensis* toxin Cry2Ab is conferred by mutations in an ABC transporter subfamily a protein. *PLoS Genet*. **11**, 23.

Tian, L., Song, T., He, R., Zeng, Y., Xie, W., Wu, Q., Wang, S., Zhou X. & Zhang Y. (2017) Genome-wide analysis of ATP-binding cassette (ABC) transporters in the sweetpotato whitefly, *Bemisia tabaci BMC Genomics*. **18**, 330.

Tyzack, J.K., Wang, X., Belsham, G.J. & Proud, C.G. (2000) ABC50 interacts with eukaryotic initiation factor 2 and associates with the ribosome in an ATP-dependent manner. *J Biol Chem.* **275**, 34131–9

Vélez, A., M., Fishilevich, E., Matz, N., Storer, N., P., Narva, K. E. & Siegfried, B.D. (2017) Parameters for Successful Parental RNAi as An Insect Pest Management Tool in Western Corn Rootworm, *Diabrotica virgifera virgifera*. *Genes*. *8*, 7. http://doi.org/10.3390/genes8010007

Wang, Y., Zhang, H., Li, H. & Miao, X. (2011) Second-generation sequencing supply an effective way to screen RNAi targets in large scale for potential application in pest insect control. *PLoS One* **6,** 618-644.

Waterhouse, R.M., Fredrik, Tegenfeldt, Jia Li, Evgeny, M.Z, & Evgenia, V.K. (2012) "OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs." *Nucleic acids research* **41**, D358-D365.

Zhao, Z., Fang, L.L., Johnsen, R. & Baillie, D.L. (2004) ATP-binding cassette protein E is involved in gene transcription and translation in *Caenorhabditis elegans. Biochem Biophys Res Commun*. **323**, 104–111.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., Huang, W., He, G., Gu, S., Li, S. & Zhou, X., (2014) SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics*, *30*(12), pp.1660-1666.

Xie, X., Cheng, T., Wang, G., Duan, J., Niu, W. & Xia, Q. (2012) Genome-wide analysis of the ATP-binding cassette (ABC) transporter gene family in the silkworm, *Bombyx mori. Mol Biol Rep*. **39**, 7281–7291.

Xiao, Y., Zhang, T., Liu, C., Heckel, D.G., Li, X., Tabashnik, B.E. & Wu, K. (2014) Missplicing of the ABCC2 gene linked with Bt toxin resistance in *Helicoverpa armigera. Sci. Rep. 4 art*. 6184.

CHAPTER 4

Transcriptome Analysis and Systemic RNAi Response in the Kudzu Bug, *Megacopta cribraria*

ABSTRACT

The kudzu bug, *Megacopta cribraria* (Hemiptera: Plataspidae), a native of Asia, arrived in the southeastern United States in 2009 and has rapidly become an economically important pest of soybeans, *Glycine max* (L.) Merrill. While pyrethroid-based insecticides are effective against kudzu bugs, increased use could lead to resistance, and cause additional harm to non-target species such as pollinators. Here we investigate the potential of RNA interference (RNAi) as a possible control strategy for *M. cribraria*. Since limited genetic data was available for this species, we first isolated and sequenced RNA from four groups of individuals; mixed-staged embryos, mixed-staged nymphs, adult females and adult males. The combined *M. cribraria* transcriptome assembly consists of 12,857 contigs, 11,224 of which had at least one significant match in NCBI's non-redundant protein database. Importantly, we were able to identify orthologs of many of the RNAi pathway genes, including *Dicer2* and *Argonaute-2*, however, no transcript sharing significant homology with the *systemic RNA interference deficient-1* (*sid-1*) gene was found. Although sid-1 plays a key role in double-strand RNA (dsRNA) uptake in *Caenorhabditis elegans*, it is unclear what role *sid-1* like genes play in insects. To assess RNAi functionality in *M. cribraria* we next surveyed the transcriptome for potential RNAi target genes and tested one of these, *catalase*, via microinjection of target-specific dsRNA into forth-instar nymphs. The resultant phenotypes (mortality and inhibition of nymphal growth) and RT-PCR results indicate that we achieved gene-specific knockdown of *catalase*, demonstrating that the RNAi pathway is functional in *M. cribraria*, and can induce a systemic response. This is the first published transcriptomic data from this species; these findings represent a valuable resource to aid future research on functional genomics in this economically important species and could provide insight into new pest control methods.

INTRODUCTION

The kudzu bug, *Megacopta cribraria* Fabricius (Hemiptera: Plataspidae), is endemic to Asia, but was discovered in Georgia, USA during the fall of 2009 (Suiter et al., 2010). *M. cribraria* has since been reported in twelve additional states (Alabama, Arkansas, Delaware, Florida, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia), as well as the District of Columbia (www.kudzubug.org 2018; Gardner et al., 2013). This rapid spread was attributed to it being a strong flyer (Suiter et al., 2010). While not a major pest in its home range, *M. cribraria* is a pest in southern regions of the U.S. Its pest status is in large part due to its impact on soybeans (*Glycine max* L. Merrill); the second most planted field crop in the U.S. generating about $42 billion dollars annually (USDA-ERS 2018). If untreated, *M. cribraria* infestations can devastate soybean production, causing yield losses of up to 60% (Seiter et al., 2013). The sustainability of this crop is not only of economic importance but is also vital to overall food security (USDA-ERS 2016). In addition, *M. cribraria* is also a threat to other legume crops and various fruit trees (Eger et al., 2010, Ruberson et al. 2013, Seiter et al., 2013).

The primary reproductive hosts of *M. cribraria* in the United States are the invasive kudzu plant, *Puereria montana* (Loureiro) Merrill variety *lobata* (Willdenow), and soybeans; the insects can complete two generations per year in the southeastern U.S. (Suiter et al., 2010, Zhang et al., 2012). The first generation is typically completed on kudzu, or early-planted soybeans (Del Pozo-Valdivia and Reisig 2013, Seiter et al., 2013), while the second generation can develop on kudzu or soybeans (Zhang et al., 2012, Seiter et al., 2013a). Interestingly, *M. cribraria* development may also occur on secondary hosts such as lima bean (*Phaseolus lunatus* L. [Fabales: Fabaceae]) and pigeon pea (*Cajanus cajan* L. [Fabales: Fabaceae]) (Zhang et al., 2012, Medal et al., 2013, Ruberson et al., 2013, Seiter et al., 2013, Blount et al., 2015). Apart from being an agricultural pest, *M. cribraria* is also a nuisance pest. The problem stems from the insects forming overwintering aggregations both in and around homes (Suiter et al., 2010).

M. cribraria nymphs and adults feed by piercing the plant's epidermis and sucking sap from the phloem (Tayutivutikul and Yano 1990, Thippeswamy and Rajagopal 2005, Suiter et al., 2010). These feeding activities deprive the host of water and nutrients required for growth (Stubbins et al., 2017), and therefore weakens and stunts the plant, reducing the number of pods per plant and number of seeds per pod (Seiter et al. 2013; Waldvogel and Alder, 2011). Although insecticide applications have been an effective method for reducing pest pressure from *M. cribraria* in soybean fields (Miao et al., 2016; Seiter et al., 2015), continuous usage will likely hasten resistance development, and result in increased negative pressure on beneficial species, such as pollinators and parasitoids (Gay 2012; Ademokoya et al., 2017). Also, control of this species may be more challenging than control of leaf-feeding insects because unlike leaf feeders that ingest topical insecticides along with leaf tissue, this mode of feeding could greatly reduce the amount of topical insecticide *M. cribraria* ingests (Stubbins et al., 2017).

Biological control methods are also available. Specifically, two egg parasitoids, *Paratelenomus saccharalis* (Dodd) (Hymenoptera: Platygastridae), and *Ooencyrtus nezarae* (Hymenoptera: Encyrtidae), have been shown to be effective in parasitizing *M. cribraria* eggs in the U.S. (Ruberson et al., 2013; Ademokoya et al., 2017). Some speculate that these parasitoids have contributed to the sudden decrease in *M. cribraria* populations over the past few years (Ademokoya et al., 2017). However, despite the current reduction in *M. cribraria* population size and the availability of both chemical and biological control methods, it is wise to investigate alternate control strategies since resistance is inevitable.

One promising alternative to chemical-based control is RNA interference (RNAi). RNAi is a naturally occurring phenomenon that results in gene-specific silencing (Price and Gatehouse 2008; Smagghe et al., 2010). When a double-stranded RNA (dsRNA) is introduced into the cell, its recognized by an enzyme called Dicer and processed into small interfering RNAs (siRNAs). After processing, one strand of the siRNA is assembled into an RNA-induced silencing complex (RISC) in conjunction with the Argonaute multidomain protein, which is responsible for recognition and degradation of mRNA transcripts that are homologous to the dsRNA (Hannon 2002; Kim et al., 2009). This technique has been shown to be an effective control method for some agricultural pests (Baum et al., 2007; Yu et al., 2013), including some hemipterans (Price and Gatehouse 2008; Smagghe et al., 2010). Moreso, RNAi has proven to be an effective tool for studying gene function in some hemipterans. For example, Futahashi et al., (2010) investigated the biological role of *laccase2* in three hemipteran, *Riptortus pedestris* (Alydidae), *Nysius plebeius* (Lygaeidae) and *M. punctatissima* (Plataspidae). In this study, when *laccase2*-specific dsRNA was injected into last-instar nymphs, their cuticular pigmentation after adult molting was impacted. The results highlight the conserved biological role of *laccase2*, as well as the efficacy of RNAi in these species.

Interestingly, *M. cribraria* and *M. punctatissima* are considered by some to be distinct species (Hosakawa et al., 2014; McPherson 1941), while others suggest that they are conspecific (i.e. belong to same species) and consider them to be synonymous with one another (Eger et al., 2010). Since their identity in the U.S. is still unclear (Dhammi et al., 2016) I sought to perform RNA-Seq on individuals from the U.S. population and compare available sequence data to ascertain if there are significant differences and perhaps help address this controversy. In the process I also surveyed the transcriptomic data for key genes involved in the RNAi pathway and tested RNAi functionality by microinjecting dsRNA targeting *catalase* (*cat*), a highly-conserved gene known to be vital in other organisms (eg. *Drosophila melanogaster*). Our results both confirm the functionality and utility of RNAi in this species as well as demonstrate that *cat* expression is required for *M. cribraria* survival. Also, this study shows that the *M. punctatissima laccase-2* gene sequence is 99% identical to our *M. cribraria laccase-2* sequence.

MATERIALS AND METHODS

Insect collection and rearing

M. cribraria adults were collected from a patch of wild kudzu, *Puereria montana* (Loureiro) Merrill variety *lobata* (Willdenow), growing on the North Carolina State University campus in Raleigh, NC. Adults were maintained in the lab on soybean, *Glycine max L*. Merrill, under a controlled environment of 28^{0}C, 60% relative humidity (RH) and 14:10 light: dark (L:D) cycle in 60-cm^3 cages (BugDorm, MegaView Science, Taichung, Taiwan). The soybean plants used as host were grown and maintained under standard greenhouse conditions in a greenhouse at NCSU.

RNA isolation and Illumina sequencing

Total RNA was isolated from four separate flash-frozen (liquid nitrogen) *M. cribraria* samples (mixed-stage embryos, mixed-stage nymphs, adult males and adult females), using the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) and treated with DNase I (Qiagen) according to the manufacturer’s instructions. The RNA was submitted to the Genomic Sciences Laboratory (North Carolina State University, NC USA) for quality assessment, 2x poly(A) selection, fragmentation, size selection (600- to 700-bp) and separate Illumina TruSeq libraries were then generated and sequenced on an Illumina MiSeq sequencer (2 x 300-bp).

Transcriptome assembly and annotation

Raw reads underwent an initial quality assessment using FastQC (Andrews, 2010), and were trimmed based on quality scores, and all remaining Illumina adapter sequences removed. Trimmed reads (all life stages) were *de novo* assembled into contigs using SeqMan NGen (DNASTAR, Inc; Madison, Wisconsin, USA) to generate a combined assembly. This dataset was annotated using Blast2GO software (4.0) (Conesa et al., 2005; Götz et al., 2008) with BLASTx comparisons to the non-redundant (nr) arthropod-specific database at the National Center of Biotechnology Information (http://blast.ncbi.nlm.nih.gov/). The combined graphs were created at level 2 for Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories from Blast2GO.

RNAi pathway and *Catalase* sequence identification

A searchable database was generated from the *M. cribraria* combined assembly and the database searched (tBLASTn) using BlastStation software (TM Software Inc. Arcadia, CA, USA). The deduced amino-acid sequence of each *Cylas puncticollis*, the African sweetpotato weevil *core* RNAi gene (Prentice et al., 2015) was used as bait to identify the corresponding *M. cribraria* RNAi gene. Putative orthologs were identified via initial BLAST searches based on sequence identity (E-values less than 10^{-6}) and homology was determined via reciprocal BLAST comparison (BLASTx) to all known arthropod proteins in the nr database at NCBI. Only sequences having high-scoring hits to core RNAi genes were retained.

Phylogenetic trees were constructed using the putative *M. cribraria* (*Mc*) orthologs of *Dcr-2*, and *Ago-2*, along with aligned orthologous proteins from other species, the following species were selected: *Halyomorpha halys* (*Hh*), *Acyrthosiphon pisum* (*Ap*), *Aphis glycines* (*Ay*), *Tribolium* *castaneum* (*Tc*), and *Bemisia tabaci* (*Bt*). Sequence alignments were performed with ClustalW using MEGA 7 (Tamura et al., 2013), and subjected to phylogenetic analysis using maximum likelihood tree search and 1000 bootstrap replications in Mega 7. An optimal model of sequence evolution was determined for each data set by likelihood ratio tests and AIC within the Mega 7 program.

The *M. cribraria* transcriptomic database was queried with the deduced amino-acid sequence of the *T. castaneum catalase* (*Tccat*) gene. The top hit was assessed as described above for the core RNAi genes. A phylogenetic tree was constructed using cat proteins from the following species: *T. castaneum* (*Tc*), *H. halys* (*Hh*), *D. melanogaster* (*Dm*), *Bombyx mori* (*Bm*), *A. pisum* (*Ap*), *B. tabaci (Bt) Helicoverpa armigera (Ha),* and *Spodoptera litura (Sl).* Sequence alignments were performed as outlined above and subjected to phylogenetic analysis with 1000 bootstrap replications.

Primer design and *in vitro* dsRNA synthesis

Primers used to generate *Mccat*-specific dsRNA **(**Table 1) were designed using Vector NTI Advance (VNTI) software (Invitrogen), and PCR performed using 100ng of *M. cribraria* cDNA as template. The *Mccat* gene was amplified from developmental stage-specific cDNA template using cycling conditions of 95°C for 3 min, 40 cycles of 95°C for 30s, 58°C for 30s, and 72°C for 10s, and finally 4 min at 72°C. The PCR product was purified using a QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions, then ligated into the pGEM-T vector (Promega, Madison, WI, USA), with the resulting plasmids used to transform chemically competent *E. coli* (Top10; Invitrogen). Recombinant clones were selected on LB agar containing ampicillin as described by the manufacturer. Plasmid DNA was isolated using a

Plasmid Mini Kit (Qiagen), and the insert sequenced at the Genomic Sciences Laboratory (NCSU), and confirmed by BLAST analysis (http://blast. ncbi.nlm.nih.gov). The cloned *Mccat* gene fragment was used as template in PCR reactions containing the following primers; T7 as a forward primer (due to location in pGEM), and a pGEM-specific reverse primer that was tailed with T7. PCR was performed with an initial denaturation of 95°C for 3 min, the 40 cycles of 95°C for 30s, 58°C for 30s, and 72°C for 10s, then 4 min at 72°C). The PCR products were confirmed via agarose gel electrophoresis and purified as described above and quantified on a Nanodrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) and ~1.0 µg was used as template for dsRNA synthesis using the MEGAscript T7 *in vitro* Transcription Kit (Ambion, Foster City, CA, USA). Synthesized dsRNAs were purified using the MEGAclear Kit (Ambion) and concentration determined using a Nanodrop 1000 spectrophotometer (Thermo Scientific).

RNAi bioassay

Insects were anesthetized on ice for 30 min before microinjecting *Mccat*–specific dsRNA (concentration = 1 µg/µl) into the hemocoel of fourth-instar nymphs. This process was replicated three times, with at least 30 individuals per replicate. The same procedure was also conducted for control injections (buffer alone). Following injection, nymphs were allowed to recover at room temperature, before returning to soybean plants where they were held for monitoring phenotypes and downstream molecular analysis. Phenotypes were observed daily using a stereomicroscope, and transcript levels were assessed by RT-PCR using total RNA pooled from three injected individuals (one per replicate) at day 1 through day 5 time points. The earliest time point was 24-hours post injection, and the latest was five days post injection. The *M. cribraria ribosomal* *protein 6* (*McRSP6*) gene was used as an internal standard to estimate the relative expression of mRNA. The sequence of the *McRSP6* primers is given in Table 1.

RESULTS AND DISCUSSION

Transcriptome sequencing, assembly, and annotation

To generate a fairly representative combined transcriptome for *M. cribraria*, four separate samples of RNA (mixed-staged embryos, mixed-staged nymphs, adult females and adult males) were each sequenced using the Illumina (MiSeq™ 2000) platform. A total of 23,983,499 raw sequence reads of ~300 bp in length were obtained across the four libraries. Of these, 9,875,298 assembled into a single-combined transcriptome producing 12,857 contigs with an average length of 2008 bp and an average GC content of 48%. BLAST-based annotation was performed with Blast2GO (Conesa et al., 2005) using BLASTx against the arthropod-specific section of NCBI's non-redundant (nr) protein database. Among the 12,857 contigs derived from the *M. cribraria* transcriptome, BLASTx analysis revealed that 11,224 of these shared significant identity (*E*-values < 10^{-6}) with known arthropod proteins, and a subset of these received GO mapping and annotation terms (Figure 1). Sequences lacking significant identity to known arthropod proteins likely represent poorly conserved or novel sequences, as well as non-coding RNAs. The distribution of top BLAST hits by species showed that the brown marmorated stink bug, *Halyomorpha halys* was the most frequent, representing (>9000) of the matches (Figure 2).

Blast2GO also grouped annotated contigs into three broad Gene Ontology (GO) categories based on predicted function: Biological Processes, Molecular Function and Cellular Component (Fig. 3). The most dominant terms for Biological Process were cellular, metabolic and single organism processes, while the most frequent for Molecular Function were binding, catalytic and transporter activity. Lastly, the most frequent terms for Cellular Component were cell, organelle, and membrane (Fig. 3). As expected, GO analysis of the *M. cribraria* combined transcriptome indicates that the majority of biological processes, molecular functions, and cellular components are similar to those found in other published transcriptomes (Kaur et al., 2017; Sparks et al., 2014; Zhang et al., 2013). The SeqMan NGen assembly was assessed for its degree of completeness using BUSCO, and received a score of 2006, or 75% of the 2,675 genes in the arthropod reference set v. 9.0 were represented as either single copy or duplicated, thus indicating that a reasonable proportion of genes in the arthropod reference were represented within *M. cribraria* transcriptomic data.

Identification of RNAi-related genes

A searchable database was created, and the combined transcriptome was queried for the presence of core RNAi pathway genes (Prentice et al., 2015). BLAST analysis revealed 31 putative homologs in the *M. cribraria* transcriptome database **(**see Table 2, Appendix C). To confirm their identity the putative *M. cribraria* homologs were compared (BLASTp) to known arthropod proteins in the NCBI database. Twenty-nine of these shared the greatest identity to sequences from *H. halys*, while 2 shared more identity to *Operophtera brumata* and *Locusta migratoria* (Table 2). Although we could not identify orthologs for all 47 known RNAi pathway genes, there are many reasons this could occur.

Although our RNA-Seq analysis only identified 31 of 47 RNAi pathway genes, the remaining 16 genes could have been missed due to the fact that our survey is based on low-coverage transcriptomic data. It is entirely possible that these genes are in the *M. cribraria* genome, but are either transcribed at relatively low levels, or are expressed at time points not covered by our

RNA-Seq libraries. While these sequences were not present in our transcriptome dataset, many of these will likely be detected when the *M. cribraria* genome is sequenced.

Chief among the sequences we failed to identify is what is thought to be the insect version of the *systemic RNA interference deficient-1* (*sid-1*) gene. While *C. elegans* possesses *sid-1*, a gene that encodes a pivotal protein involved in dsRNA uptake, insects do not appear to have a clearly assignable one-to-one ortholog of *sid-1*, but rather have a gene known as *sid-1-like*. Previous analysis of several insect genomes and transcriptomes shows that the number of *sid-1-like* genes varies among insect species (Price and Gatehouse, 2008). For example, dipterans like *D. melanogaster* lack *sid-1-like* genes, while some hemipterans, hymenopterans, orthopterans and phthirapterans have been reported to have a single *sid-1-like* gene (Tomoyasu et al., 2008). Among lepidopterans the number of *sid-1-like* genes are even more diverse, with *Spodoptera exigua* having only one and *B. mori* having three copies (Price and Gatehouse, 2008). Interestingly, Sparks et al (2014), reported that the *H. halys* transcriptome lacks *sid-1-like* genes. It was first thought that the presence or absence of these genes explained the robust and systemic RNAi response in *T. castaneum* and the lack of systemic RNAi in *D. melanogaster*, respectively. However, there are reports that suggest the *sid-1-like* genes found in *T. castaneum* may not be critical for the systemic RNAi response observed in these beetles, but this was based on RNAi-mediated knockout of the *sid-1-like* genes and then assessing the impact via RNAi (Gourley et al., 2003). Since this did not affect the systemic RNAi response the authors concluded that the *sid-1-like* genes found in *T. castaneum* are not responsible for the movement of dsRNA across cell membranes (Gourley et al., 2003). This conclusion would be stronger if knockout was through CRISPR/Cas9 rather than RNAi, since the effect of CRISPR/Cas9 is permanent, and RNAi is not. Another gene, *FBX011*, is thought to be involved in dsRNA uptake in some species

(Sabin et al., 2009) and a *FBX11* ortholog was found in *M. cribraria* transcriptome. The core components of the RNAi machinery are proteins that, together with the small RNA fragments, are involved in gene silencing. Besides these core genes, many different factors have been identified that play a role in the efficiency of RNAi. Since the discovery of RNAi, significant efforts have been made to describe the molecular mechanisms of RNAi (Aravin, et al., 2007). Studies performed on several model insects, including *D. melanogaster*, *T. castaneum* and the silkworm, *Bombyx mori*, have revealed that the RNAi pathway is composed of two phases, the initiator phase, and the effector phase, and together they result in post-transcriptional gene silencing (PTGS) Vaucheret et al., 2001; Vaucheret and Fagard, 2001; Castel and Martienssen, 2013). The main components of RNAi are well conserved across phyla; however, the mechanism and systemic spread of silencing are highly diverged and gene regulation is mediated by a variety of RNA-based products (siRNAs, miRNA) (Shreve et al., 2013; Swevers et al., 2013). While systemic RNAi has been documented in several insects (Bucher et al. 2002; Tomoyasu and Denell 2004); it is assumed that all animals with systemic RNAi response use the mechanism described in *C. elegans*. Nevertheless, it seems clear that the systemic RNAi response is not universal, as the leading insect model, *D. melanogaster*, lacks a systemic RNAi response (Roignant et al., 2003). Therefore, more comprehensive searches for the proteins involved in the systemic RNAi mechanism in insects continues to be critically important for broadening our knowledge of RNAi and its possible applications in pest control.

In this study, we found putative orthologs for 31 RNAi pathway genes in the *M. cribraria* transcriptome. Our results, together with previous studies, highlight the conserved biological role of the RNAi pathway among hemipterans species (Upadhyay et al., 2013; Bansal and Andy, 3013; Clauvis et al., 2016). Phylogenetic analyses (Figure 4A, B) revealed the relationship of the major RNAi genes in *M. cribraria* to those found in other insect species. For example, phylogenetic analysis shows that *McDcr-2* and *McAgo-2* are grouped into clusters along with orthologs from other hemipterans as expected (Figure 4A & B), indicating that RNAi core genes are evolutionarily conserved within sampled Hemiptera (Figure 4). As expected, we did not find a *sid-2* ortholog, which to date appears to only be present in nematodes. Phylogenetic analyses of the *catalase* (*cat*) gene shows a similar relationship among sampled species. *Mccat* joins a cluster with the two sampled hemipterans indicating that these genes are evolutionarily conserved in the order, well separated from sequences present in Diptera and Lepidoptera (Figure 5).

Transcripts encoding sid-1 or sid-1-like, which are thought to be involved in dsRNA uptake were not found in the *M. cribraria* transcriptome. While these have been reported in other insect systems (excluding *Drosophila*), hemipterans, *A. pisum* and *H. halys*, have been shown to possess systemic RNAi, but no *sid-1-like* gene has been found in their genomes/transcriptomes (Mogilicherla et al., 2018; Lu et al., 2017; Sparks et al., 2014; Mutti et al., 2006). This suggests that sid-1-like may not be necessary for systemic RNAi and that these species may have other pathways for dsRNA uptake. Additionally, some recent studies have shown that the delivery of dsRNA either by microinjection or by feeding leads to gene knockdown. Injection of dsRNA targeting the homeotic gene *sex combs reduced* (*Scr*), which is responsible for determining segmental identity in early embryos, into *H. halys* females resulted in a range of defects in offspring, and death was observed prior to hatching (Lu et al., 2017). In addition, Mogilicherla et al., (2018) demonstrated that they could induce mortality in *H. halys* by injectiong or feeding nymphs dsRNAs targeting genes known to cause mortality in other insects. Thus the lack of an apparent *sid-1* ortholog in the *H. halys* transcriptome had no negative impact on RNAi silencing in their bioassays. Similarly, while transcripts encoding *sid-1* were not found in this study, gene knockdown via RNAi was confirmed in *M. cribraria.* While this implies that *sid-1* may not be necessary for this process in *M. cribraria,* with additional mechanisms and gene interactions remaining to be described in Hemiptera, it is also possible that a *sid-1* ortholog simply has not yet been found or recognized in Hemiptera.

RNAi Bioassay

To determine if the RNAi pathway is functional and capable of inducing a systemic RNAi response in *M. cribraria* as it does in *M. punctatissima* (Futahashi et al., 2010), we injected dsRNAs targeting the *M. cribraria catalase* (*Mccat*) gene into fourth-instar nymphs and assessed knockdown both phenotypically and molecularly at several time points. The dsRNA-treated individuals showed a significant decrease in transcript accumulation, while no change is seen in the buffer-injected controls (Figure 6). The decrease can be detected 3 days post injection and lasted through day 5 (Figure 6). The housekeeping gene *RPS6* maintained its expression level independent of the treatment (Figure 6). Moreover, knockdown of *Mccat* led to mortality among treated nymphs relative to buffer injected controls, with nymphal mortality being significantly higher in treated individuals compared to those injected with buffer alone. By day 5, mortality had reached an average of 75% in treated nymphs compared to 5% for buffer treated controls (Figure 7). In addition, nearly all dsRNA-treated individuals displayed a reduction in body size, and none survived to adulthood (data not shown); this is likely due to a reduction in *cat* expression since it is required for survival. The control group, however, showed no signs of growth inhibition. Taken together, injection of dsRNAs targeting *Mccat* confirms the functionality and utility of RNAi in this species and shows that *cat* expression is required for survival in *M. cribraria*.

Since RNAi successfully reduced *cat* expression and resulted in significant mortality, we believe *cat* could be an excellent target for controlling *M. cribraria*. Moreover, *cat* is known to be an important antioxidant enzyme, like superoxide dismutase which helps defend against oxidative stress (Li et al., 2005). Moreover, the efficacy of *cat* gene silencing has been observed and reported in several species, including the palm weevil (*Rhynchophorus ferrugineus*) (Al-ayedh et al. 2016), the sand fly (*Lutzomyia longipalpis*) (Diaz-Albiter et al., 2011), the tobacco cutworm (*S. litura*), (Haiming et al., 2013), the fruit fly (*D. melanogaster*) (Mackay et al., 1989), and the kissing bug (*R. prolixus*) (Paes et al., 2001), and in each case the authors found a significant reduction in growth and an increase in mortality. Therefore, targeting the *cat* gene has great potential for insect control due to its essential role in homeostasis.

It should be noted that *M. cribraria* and the closely related species *M. punctatissima* are morphologically indistinguishable and this has led to controversy over which species is found in the U.S., or whether both species are present (Eger et al., 2010). Interestingly, Jenkins & Eaton (2011) conducted a molecular analysis confirming that genes sampled from primary and secondary endosymbionts of these species are highly similar, suggesting a shared origin and identity. Also, Hosakawa et al., (2014) reported that *M. cribraria* populations found in the U.S. are related to *M. punctatissima* populations found in the Kyushu region of Japan, and that the two species share many features, so they cannot be morphologically distinguished from one another. Comparison of the *M. punctatissima laccase-2* gene and the *laccase-2* sequence found in our *M. cribraria* transcriptome database shows that these sequences share 99% identity, with only four nucleotide differences (Figure 8). While this high similarity in a single gene is not sufficient evidence to confirm species identity, it does support the idea, as proposed by Dhammi et al., (2016), that the two forms are better considered a single species. More extensive population genetic sampling will be required to more fully resolve the species boundaries within this group.

This study presents transcriptomic data generated from various *M. cribraria* developmental stages. These datasets represent a valuable genetic resource for screening potential genes for the development of biotechnological tools for insect control. Through transcriptomic analysis I identified 31 RNAi pathway genes and performed RNAi bioassays to demonstrate RNAi efficiency in *M. cribraria*. My RNA-Seq data will enable others to perform a wide range of functional genomic studies on this pest and offers insights into stage-specific expression. My work also characterized the *cat* gene and demonstrated its possible use for biological control of *M. cribraria*, and *RP6S* for use as a control gene when performing RT-PCR. In addition, since no transcriptomic or genomic data was available for this species, my RNA-Seq data provides a valuable resource for comparative genomics/transcriptomics with other hemipteran species and will also greatly aid annotation of the *M. cribraria* genome when the sequence becomes available. Given the fact that insecticide resistance is likely to evolve in major crop pests, RNAi offers a completely new mode of action, making it a promising approach for controlling *M. cribraria*. Although *in planta* RNAi (dsRNA delivered by the plant) would be the best option for RNAi-based control (Baum, et al., 2007; Mao, et al., 2007), my tests or oral RNAi in *M. cribaria* failed to demonstrate target gene knock down. This could be due to not having a high enough titer of dsRNA, or it could be due to nucleases in the insect gut, or some other potential factor, thus leaving additional research to be performed by others.

Table 1: Primer sequences used for dsRNA synthesis and RT-PCR analysis

Gene	Primer sequences	Purpose
MccatF	TAATACGACTCACTATAGGGCCTCGCCACATCTCCGATAG	dsRNA synthesis
MccatR	TAATACGACTCACTATAGGGGGGTCTCTAGCTGTGTCAGC	dsRNA synthesis
McRSP6F	ATGGCAATGCGATGCCTCTT	RT-PCR
McRSP6R	ACCTACGCCTCTTCAACATC	RT-PCR

Underlined sequence corresponds to T7 promoter

Table 2: Overview of RNAi genes identified in the transcriptome of *M. cribraria.*

RNAi core genes	Contig	Size (AA)	ComparIson to *Cylas brunneus*	First hit Blastp
miRNA				
Dcr-1	11725	1090	E= 0.0; bits = 748	*AVK59457.1 Nezara viridula*
Ago-1	7709	901	E= 0.0; bits = 1664	*XP_024214308.1 H. halys*
Loquacious	1167	384	E = 1E-98; bits = 351	*AVK59462.1 Nezara viridula*
Expotin-5	8269	1889	E= 0.0; bits = 722	*XP_014241155.1 Cimex lectularius*
SiRNA				
Dcr-2	7670	1603	E = 9E-175; bits =610	*XP_014275310.1 H. halys*
Ago-2	7506	1234	E= 0.0; bits = 741	*AVK59467.1 Nezara viridula*
piRNA				
Ago-3	1655	483	E = 6E-118; bits = 422	*XP_014276831.1 H. halys*
Aubergine/piwi	241	791	E = 3E-172; bits = 601	*AGJ83331.1 Oncopeltus fasciatus*
Antiviral				
Ars2	4509	757	E = 8E-106; bits = 381	*Hh(XP_014277996.1)*
Egghead	3764	457	E= 0.0; bits = 632	*Hh(XP_014283435.1)*
Nucleases				
dsRNAse	3484	409	E = 2E-38; bits = 156	*Hh(XP_014282547.1)*
Exsome	2692	809	E = 2E-144; bits = 509	*Hh(XP_014279792.1)*
Poly(A) polymerase	7462	601	E= 0.0; bits = 713	*Hh(XP_014282208.1)*
dsRNA uptake				
SID1-related	No Hit			
FBX011	3711	871	E= 0.0; bits = 1395	*Hh(XP_014287303.1)*
Hermansky(CG4966)	5140	405	E = 1E-61; bits = 235	*Hh(XP_014288755.1*
RISC				
Tudor-SN	2299	907	E=8.00E-05; bit= 43.9	*(BAW35372.1) Locusta migratoria]*
Vasa intronic gene	2606	395	E =8E-28; bits = 123	*XP_014276922.1 H. halys*
FXMRI	1649	563	E = 1E-166; bits = 583	*XP_014287676.1 H. halys*
P68 RNA helicase	211	533	E =2E-179; bits = 625	*XP_014287495.1 H. halys*
Translin factor X	5300	282	E =1.00E-13; bits = 178	*XP_024218689.1 H. halys*
Armitage	5309	797	E =1E-170; bits = 597	*Hh(XP_014289817.1)*
Homeless(spindle-E)	4827	1771	E= 0.0; bits = 639	*Hh(XP_014286769.1)*
Maelstrom	9069	477	E = 5E-40; bits = 162	*Hh(XP_014290039.1)*
RNA helicase Belle	3595	678	E= 0.0; bits = 692	*Hh(XP_014279436.1)*
PRP16, mut6 homolog	10285	1203	E =0.0; bits = 1561	*Hh(XP_014279344.1)*
Gawky	602	1452	E = 3E-51; bits 201	*Hh(XP_014288686.1)*
Staufen	522	734	E = 4E-141; 385	*Hh(XP_014282520.1)*
Clip1	10636	425	E = 9E-163;569	*Hh(XP_014270973.1)*
Elp-1	10864	630	E= 4.00E-15; bits =80.9	*Hh(XP_014290480.1)*
GLD-1 homolog	3261	342	E =8E-121; bits = 469	*Hh(XP_014271454.1*
ACO-1 homolog	3707	893	E= 0.0; bits = 1368	*Hh(XP_014275296.1)*

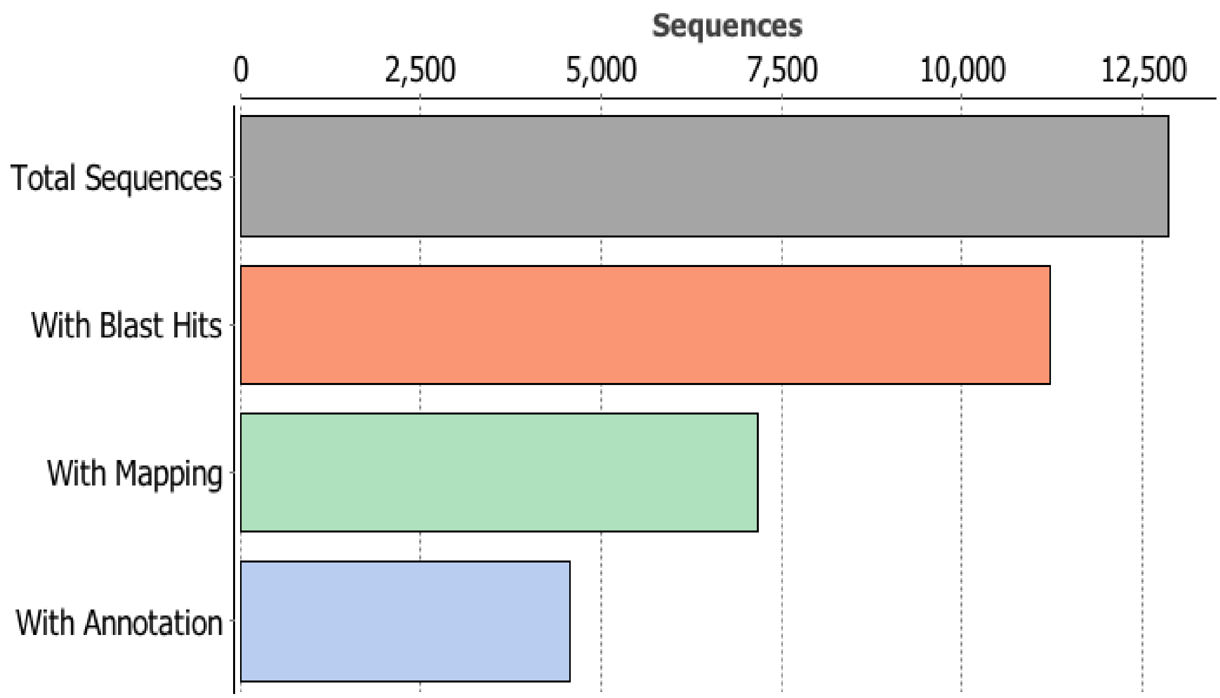


Figure 1: Blast2GO annotation of the *M. cribraria* transcriptome. Among the 12,857 contigs used for the BLASTx search against the NCBI nr protein database, BLASTx analysis revealed over 11,224 contigs shared significant identity with known arthropod proteins.

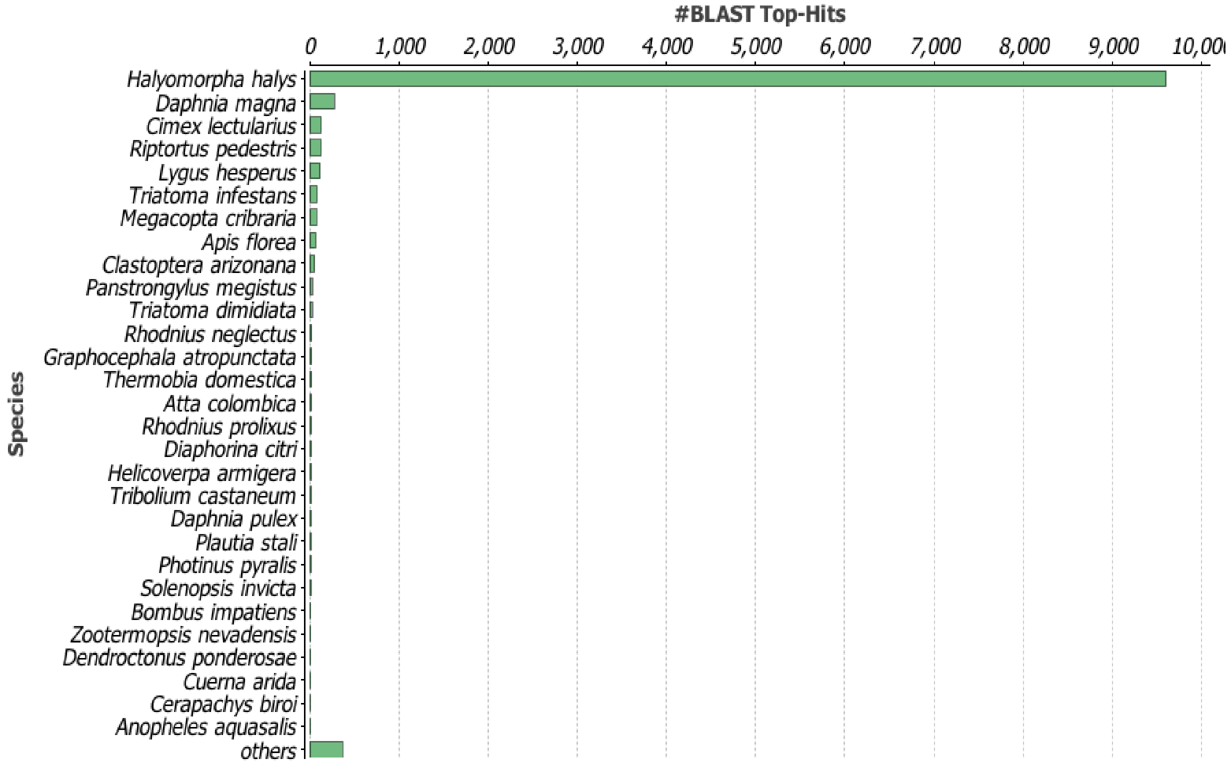


Figure 2: Top-Hits species distribution of *M. cribraria* transcriptome. The most represented species in the top scoring BLASTx-matches when compared to an arthropod-specific database.

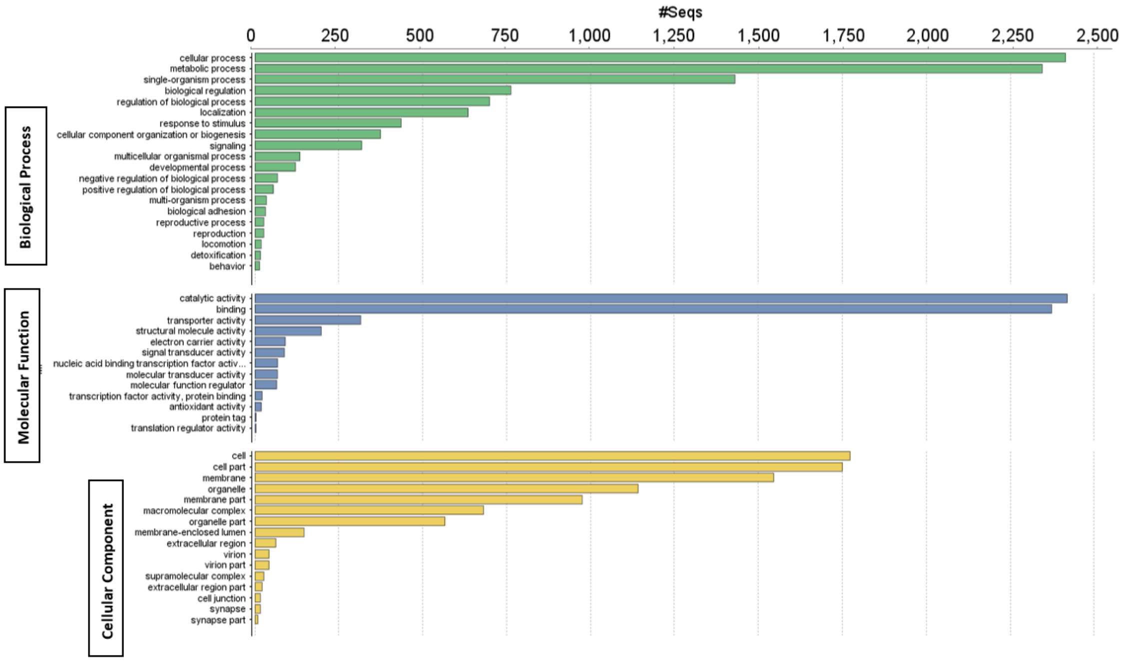


Figure 3: Gene ontology classification of the *Megacopta cribraria* transcriptome.

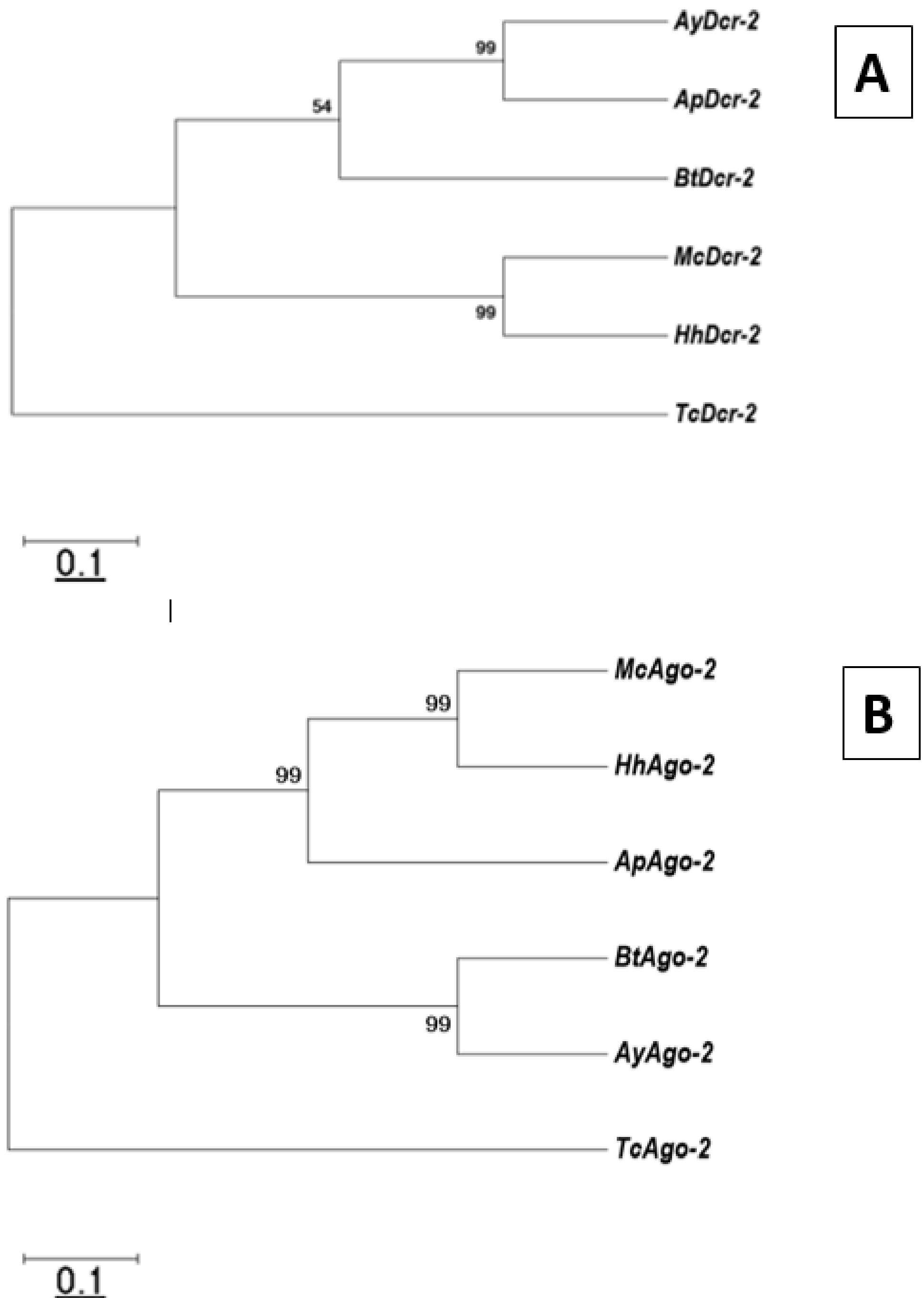


Figure 4: Phylogenetic trees of RNAi core genes constructed using maximum likelihood and 1000 bootstrap replicates. Protein sequences of Dcr2 (A) and Ago2 (B) from several species were compared. *Halyomorpha halys* (*Hh*), *Acyrthosiphon pisum* (*Ap*), *Aphis glycines* (*Ay*), *Tribolium castaneum* (*Tc*), *Megacopta cribraria* (*Mc*) *and Bemisia tabaci* (*Bt*). The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per site. Bootstrap values are given in percentages at the internodes.

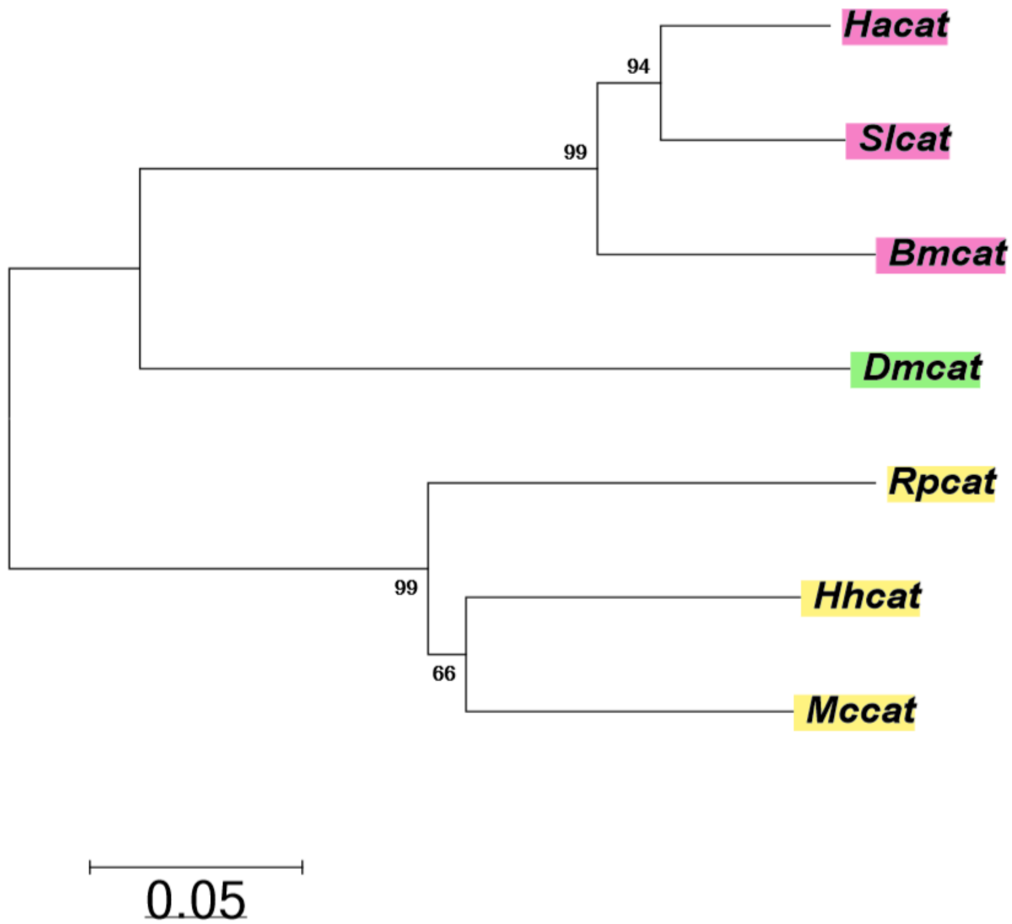


Figure 5: Phylogenetic relationships among insect species catalases found by maximum likelihood search with 1000 bootstrap replicates. *Halyomorpha halys* (Hh), *Drosophila melanogaster* (Dm), *Bombyx mori* (Bm), *Helicoverpa armigera* (Ha), *and Spodoptera litura* (Sl), *Riptortus pedestris* (Rp), and *Megacopta cribraria* (Mc). The scale bar indicates an evolutionary distance of 0.05 amino acid substitutions per site. Bootstrap values are given in percentages at the internodes.

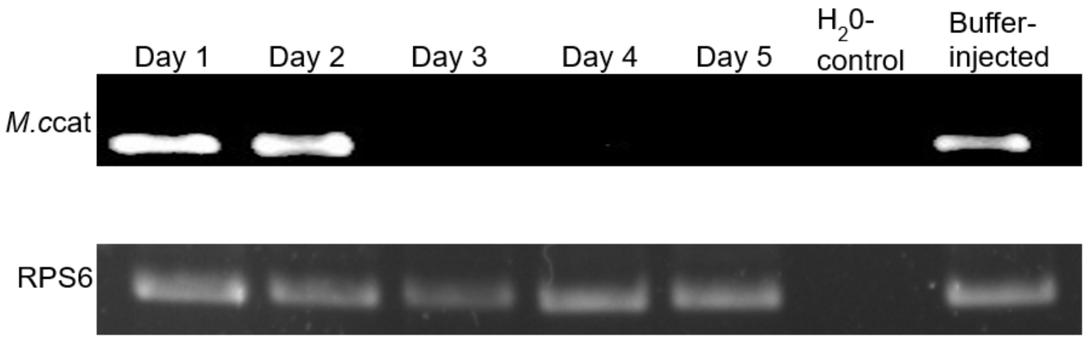


Figure 6: Effect of RNAi on the expression levels of the *cat* gene in *M. cribraria*. (Top) Knockdown for *Mccat* transcript determined by RT-PCR, showing day 1 through day 5 post injection, water control, and buffer-injected control on day 5 post injection. (Bottom) Internal loading control, *RPS6*.

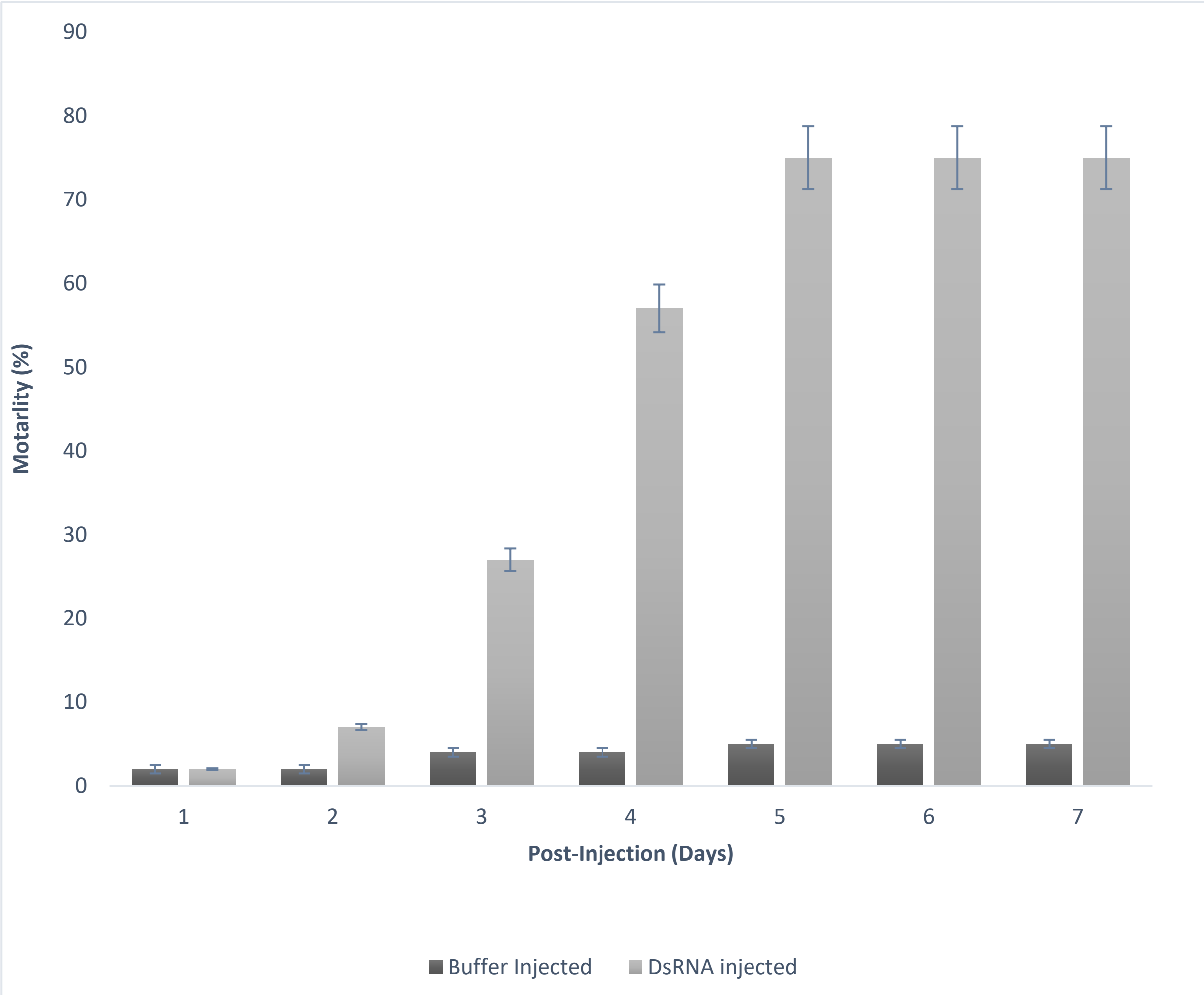


Figure 7: Effects of *Mccat* RNAi on the development of *M. cribraria*. dsRNAs were injected into fourth-instar nymph, and the mortality rates were recorded daily starting from 24 hours post injection until 7 days post injecton. Buffer was injected as a control.

```
MpLac-2         ACAGCAGCATTACGTGAGTGGTCAAGTCGTAATCGTAGAGATGGGAGTTGGGGTCCAGTG
McLac-2         ACAGCAGCATTACGTGAGTGGTCAAGTCGTAATCGTAGAGATGGGAGTTGGGGTCCAGTG
                ************************************************************

MpLac-2         AGGGTGCCTGACGGATTGCGATACTGCCGTAGATACCATCCATCTTCTGGAGCCCAGTGT
McLac-2         AGGGTGCCTGACGGATTGCGATACTGCCGTAGATACCATCCATCTTCTGGAGCCCAGTGT
                ************************************************************

MpLac-2         GAGCATGCCAGAAATGTGTGCCTGCGTTTCCTGCAATCCACTGATACCTGAACGTGTTAC
McLac-2         GAGCATGCCAGAAATGTGTGCCTGCGTTTCCTGCAATCCACTGATACCTGAATGTGTTAC
                ****************************************************.*******

MpLac-2         CTTCGGAGATGGGGCACTGAGTGACGAAAGGTACACCGTCAGATATCTGAGTACCCCTCT
McLac-2         CTTCGGAGATGGGGCACTGAGTGACGAAAGGTACACCGTCAGATATCTGAGTACCCCTCT
                ************************************************************

MpLac-2         GGTGGATTCCGTGCCAATGTATGGTCAGTTCCAATCCTTGCATTGCATTGTGTACGTCGA
McLac-2         GGTGGATTCCGTGCCAATGTATGGTCAGTTCCAATCCTTGCATTGCATTGTGTACGTCGA
                ************************************************************

MpLac-2         TGACGACTTTGTCTCCTTCACAGACCTGTATGCTTGGACCTGGAAGCATTCTGTTGATAG
McLac-2         TGACGACTTTGTCTCCTTCACA-ACCTGTATGCTTGGACCTGGAAGCATTCTGTTGATAG
                ********************** *************************************

MpLac-2         TAAGCAATCCCCTTTCAACACCATCTGCGAGAACACACTGACAGTGACTCCATACCGTGT
McLac-2         TAAGCAATCCCCTTTCAACACCATCTGCGAGAACACACTGACAGTGACTCCATACCGTGT
                ************************************************************

MpLac-2         TGGTGGCGTTAGGTGTGCACACCTGGCAAGCCGCTCCTAGGACGGTGTAAAGCTCGGCTG
McLac-2         TGGTGGCGTTAGGTGTGCACACCTGGCAAGCCGCTCCTAGGACGGTGTAAAGCTCGGCTG
                ************************************************************

MpLac-2         TGAAGTGGTAGTAGCACACTTTTGGTGGCTCCCCTTCTCTGCAAGCCCTGGCGCATTCCT
McLac-2         TGAAGTGGTAGTAGCACACTTTTGGTGGCTCCCCTTCTCTGCAAGCCCTGGCGCATTCCT
                ************************************************************

MpLac-2         CGGGTGACGAAAGGATAGGGTTCCTCTGAAGTTCTTCCAGTCCTCTCCTATCGAAGTCGA
McLac-2         CGGGTGACGAAAGGATAGGGTTCCTCTGAAGTTCTTCCAGTCCTCTCCTATCGAAGTCGA
                ************************************************************

MpLac-2         GCGAGGATGCCGGAGCGTACTTGCTGGGATTGGTGTCAACGCTTTCCGGTCCGAACCCAG
McLac-2         GCGAGGATGCCGGAGCGTACTTGCTGGGATTGGTGTCAACGCTTTCCGGTCCGAACCCAG
                ************************************************************

MpLac-2         GACCGAGGGGGTGAGTTTGGAGGAGGCCATGGCTTGAAGCGAAGAAATCTGCCGATGTGT
McLac-2         GACCGAGGGGGTGAGTTTGGAGGAGGCCATGGCTTGAAGCGAAGAAGTCTGCCGATGTGT
                **********************************************.*************

MpLac-2         CGTCAGAGAAAGAAGACTGTCCGGGGGCCAGTTCCTGTCGCCACCATGCGGCCGGCGTCG
McLac-2         CGTCAGAGAAAGAAGACTGTCCGGGGGCCAGTTCCTGTCGCCACCATGCGGCCGGCGTCG
                ************************************************************

MpLac-2         AAGGGTCACCTGGTCGTAATGGTCCTGTTCGTTTCTTGACGCGGGTTGAAGTTGGCACTT
McLac-2         AAGGGTCACCTGGTCGTAATGGTCCTGTTCGTTTCTTGACGCGGGTTGAAGTTGGCACTT
                ************************************************************

MpLac-2         TGGTTGCGCTGGCCTCGAGGGCGATGGCAAAGAGGAGAAGAGCCGGTAGGAGCCGCATGT
McLac-2         TGGTTGCGCTGGCCTCGAGGGCGATGGCAAAGAGGAGAAGAGCCGGTAGGAGCCGCATGT
                ************************************************************

MpLac-2         CTTGCACTGACCGGGTGTAAGGGGAGACTGC
McLac-2         CTTGCACTGACCGGGTGTAAGG---------
                **********************
```

Figure 8: Alignment of the nucleotide sequence of the *Laccase-2* gene from *M. punctatissima* and *M. cribraria*. These alignments show that *Mplac-2* (top row) shares 99% sequence identity with *Mclac-2* (bottom row) orthologs and (*) indicate the conserved region. Four nucleotides in *Mplac-2* differed from *Mclac-2*.

REFERENCES

Ademokoya, B., Balusu, R., Ray, C., Mottern, J. & Fadamiro, H. (2017) The First Record of *Ooencyrtus nezarae* (Hymenoptera: Encyrtidae) on Kudzu Bug (Hemiptera: Plataspidae) in North America. *Insect Science*, **18**, 1–7.

Al-Ayedh, H., Rizwan-ul-Haq, M., Hussainb, A & Aljabr, A.M. (2016) Insecticidal potency of RNAi-based catalase knockdown in *Rhynchophorus ferrugineus* (Oliver) (Coleoptera: Curculionidae). *Pest Manag Sci* **72**, 2118–2127.

Arakane, Y., Muthukrishnan, S., Beeman, R.W., Kanost, M.R., Kramer, K.J. (2005) *Laccase 2* is the *phenoloxidase gene* required for beetle cuticle tanning. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11337-11342.

Bartel, D.P. (2009) MicroRNAs: Target recognition and regulatory functions. *Cell*, **136**, 215–239.

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P., Ilagan, O., Johnson, S., Plaetinck, G., Munyikwa, T. & Pleau, M. (2007) Control of coleopteran insect pests through RNA interference. *Nat Biotechnol*. **25,** 1322–1326.

Blount, J.L., Buntin, G.D. & Roberts, P.M. (2016) Effect of planting date and maturity group on soybean yield response to injury by the kudzu bug, *Megacopta cribraria* F. (Hemiptera: Plataspidae). *Journal of Economic Entomology***,** **109**, 207–212.

Brown, S.A., Kerns, D.L., Williams, T.S., Emfinger, K. & Jones, N. (2015) Evaluation of foliar insecticides for kudzu bug control in soybeans. *Arthropod Management Tests*, **40**, 1.

Cahais, V., Gayral, P., Tsagkogeorga, G., Melo-Ferreira, J. & Ballenghien, M. (2012) Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. *Molecular Ecology Resources*, **12,** 834-845.

Caihong, T., Wee, T. T, Hongqiang, F., Ying, W., Yongmin, H. & Guoping, L. (2015) Transcriptome from different Developmental stages. *Scientific Report* **5**, 11042.

Clauvis N.T., Taning, E.C., Andrade, W.B., Hunter, O.C. & Guy, S. (2016) *Asian Citrus Psyllid* RNAi Pathway – RNAi evidence. *Scientific Report* **6**, 38082.

Cristofoletti, P.T., Ribeiro, A.F., Deraison, C., Rahbe, Y., Terra, W.R. (2003) Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrthosiphon pisum. Insect Physiol*. **49**, 11–24.

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J. & Talon, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676.

de Carvalho, D.B., Congrains, C., Chahad-Ehlers, S., Pinotti, H. & Brito, R.A. (2017) Differential transcriptome analysis supports *Rhodnius montenegrensis* and *Rhodnius robustus* (Hemiptera, Reduviidae, Triatominae) as distinct species. *PLOS ONE* **12**, e0174997.

de Loof, A. (2008) Ecdysteroids, juvenile hormone and insect neuropeptides: Recent successes and remaining major challenges. *Gen Comp Endocrine*, **155**, 3–13.

Del Pozo-Valdivia, A.I., Seiter, N.J., Reisig, D.D., Greene, J.K., Reay-Jones, F.P.P. & Bacheler, J.S. (2016) *Megacopta cribraria* (Hemiptera: Plataspidae) Population dynamics in soybeans as influenced by planting date, maturity group, and insecticide use. *Journal of Economic Entomology*, **109**, 1141–1155.

Dhammi, A., Krestchmar, J.B., Ponnusamy, L., Bacheler, J.S., Reisig, D.D., Herbert, A., Del Pozo-Valdivia A.I., & Roe, A.M. (2016) Biology, Pest Status, Microbiome and Control of Kudzu Bug (Hemiptera: Heteroptera: Plataspidae): A New Invasive Pest in the U.S. *International Journal of Molecular Science*, **17**, 1570.

Diaz-Albiter, H., Mitford, R., Genta, F.A., Sant' Anna, M.R. & Dillon, R.J. (2011) Reactive oxygen species scavenging by catalase is important for female *Lutzomyia longipalpis* fecundity and mortality. *PLoS One* **6**, e17486.

Dittmer, N.T., Suderman, R.J., Jiang, H., Zhu, Y.C., Gorman, M.J., Kramer, K.J., Kanost, M.R., (2004) Characterization of cDNAs encoding putative laccase-like multicopper oxidases and developmental expression in the tobacco hornworm, *Manduca sexta*, and the malaria mosquito, *Anopheles gambiae*. *Insect Biochem. Mol. Biol*. **34**, 29-41.

Dutton, A., Klein, H., Romeis, J. & Bigler, F. (2002) Uptake of *Bt*-toxin by herbivores feeding on transgenic maize and consequences for the predator *Chrysoperla carnea*. *Ecol. Entomol*. **27,** 441–447.

Eger, J.E. Jr., Ames, L.M., Suiter, D.R., Jenkins, T.M., Rider, D.A. & Halbert, S.E. (2010) Occurrence of the Old-World bug *Megacopta cribraria* (Fabricius) (Heteroptera: Plataspidae) in Georgia: A serious home invader and potential legume pest. *Insecta Mundi*, **121**, 1–11.

Elias-Neto, M., Soares, M.P.M., Simoes, Z.L.P., Hartfelder, K., Bitondi, M.M.G., (2010) Developmental characterization, function and regulation of a *Laccase2* encoding gene in the honey bee, *Apis mellifera* (Hymenoptera, Apinae). *Insect Biochem. Mol. Biol*. **40,** 241-251.

Elzen, G.W. & Hardee, D.D. (2003) United States Department of Agriculture-Agricultural Research Service research on managing insect resistance to insecticides. *Pest Manag. Sci*. **59,**770–776.

Fire A., Xu, S., Montgomery, M.K., Kostas, S.A. & Driver, S.E. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811.

Firmino, A.A.P., Fonseca, FC.A., de Macedo, L.L.P., Coelho, R.R., de Souza, Jr., J.D.A. & Togawa, R.C. (2013) Transcriptome analysis in cotton boll weevil (*Anthonomus grandis*) and RNA interference in insect pests. *PLoS ONE*, **8**, e85079.

Futahashi, R., Tanaka, K., Matsuur, Y., Tanahashi, M., Kikuchi, Y. & Fukatsu, T. (2011) *Laccase2* is required for cuticular pigmentation in stinkbugs. *Insect Biochemistry and Molecular Biology* **41**, 191-196.

Gay, H. (2012) Before and after Silent Spring: from chemical pesticides to biological control and integrated pest management–Britain, 1945–1980. *Ambix* **59,** 88–108.

Gazara, R., Cardozo, K., Bellieny-Rabelo, C., Ferreira, D., Terra, C., & Venancio, T. M. (2017) De novo transcriptome sequencing and comparative analysis of midgut tissues of fournon-model insects pertaining to Hemiptera, Coleoptera, Diptera and Lepidoptera. *Gene* 627*,* 85-93.

Gourley, B.L., Parker, S.B., Jones, B.J., Zumbrennen, B. & Leibold, E.A. (2003) Cytosolic Aconitase and Ferritin are regulated by iron in *Caenorhabditis elegans*. *Biol Chem* **278**, 3227–3234.

Haiming, Z., Xin, Y., Zhen, H., Meiying, H., Shaohua, C. & Rizwan-ul-Haq, M. (2013) RNAi-mediated knockdown of catalase causes cell cycle arrest in SL-1 cells and results in low survival rate of *Spodoptera litura* (Fabricius). *PLoS ONE* **8,** e59527 (2013).

Hannon, G.J. (2002) RNA interference. *Nature* **418**, 244–251.

Head, G., Brown, C.R., Groth, M.E. & Duan, J.J. (2001) Cry1Ab protein levels in phytophagous insects feeding on transgenic corn. Implications for secondary exposure risk assessment. *Ent. Exp. Appl.*, **99**, 37–45.

Hartig, J.V., Tomari, Y. & Forsteman, K. (2007) piRNAs—The ancient hunter of genome invaders. *Genes Dev* **21,** 1707–1713.

Jaubert-Possamai, S.L., Trionnaire, G., Bonhomme, J., Christophides, G.K., Rispe, C. & Tagu, D. (2007) Gene knockdown by RNAi in the pea aphid *Acyrthosiphon pisum*. *BMC Biotechnology*, **7,** 63.

Ji, R., Yujun, W., Yanbin, C., Meiping, Z., Hong-Bin, Z., Li, Z., Jichao, F. & Keyan, Z. (2016) Transcriptome Analysis of Green Peach Aphid (*Myzus persicae*). *Plant Science*, **7,** 1562.

Kaur, N., Chen, W., Zheng, Y., Hasegawa, D.K., Ling, K., Fei, Z. & Wintermantel, W.M. (2017) Transcriptome analysis of the whitefly, *Bemisia tabaci* MEAM1 during feeding on tomato infected with the crinivirus, *Tomato chlorosis virus,* identifies a temporal shift in gene expression and differential regulation of novel orphan genes. *BMC Genomics*, **18**, 370.

Keeling, C.I., Henderson, H., Li, M., Yuen, M. & Clark, E.L. (2012) Transcriptome and full-length cDNA resources for the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major insect pest of pine forests. *Insect Biochem Mol Biol*, **42**, 525-536.

Kim, V.N., Han, J. & Siomi, M.C. (2009) Biogenesis of small RNAs in animals. ***Nature*** **10**, 126–139.

Klowden, M.J. (2007) Endocrine systems. In: Klowden, M.J, editor. Physiological systems in insects. 2nd ed. Burlington, MA: Academic; p. 1–49.

Koch, A. & Kogel, K.H. (2014) New wind in the sails: improving the agronomic value of crop plants through RNAi-mediate gene silencing. *Plant Biotechnology*, **12**, 821–31.

Liu, Q. & Paroo, Z. (2010) Biochemical principles of small RNA pathways. *Annu Rev Biochem* **79**, 295–319.

Li, H., Jiang, W., Zhang, Z., Xing, Y. & Li, F. (2013) Transcriptome analysis and screening for potential target genes for RNAi-mediated pest control of the beet armyworm, *Spodoptera exigua. PLoS ONE*, **8**, e65931.

Li, H., Chougule, N.P. & Bonning, B.C. (2011) Interaction of the *Bacillus thuringiensis* delta endotoxins Cry1Ac and Cry3Aa with the gut of the pea aphid, *Acyrthosiphon pisum* (Harris). *J. Invert. Path.* **107**, 69–78.

Yong, L., Chen, M., Reding, K. & Pick, L. (2017) Establishment of molecular genetic approaches to study gene expression and function in an invasive hemipteran, *Halyomorpha halys*. *EvoDevo* **8,** 15.

Mackay, W.J. & Bewley, G.C. (1989) The genetics of catalase in *Drosophila melanogaster*: isolation and characterization of acatalasemic mutants. *Genetics* **122**, 643–652.

Mao, Y.B., Cai, W.J., Wang, J.W., Hong, G.J., Tao, X.Y., Wang L.J., Huang, Y.P. & Chen, X.Y. (2007) Silencing a cotton bollworm *P450 monooxygenase* gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nat. Biotechnol*. **25**, 1307–1313.

Mutti, N.S., Park, Y., Reese, J.C. & Reeck, G.R. (2006) RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrthosiphon pisum*. *Journal of Insect Science,* **6**, 1–7.

Matsumoto, Y., Suetsugu, Y., Nakamura, M. & Hattori, M. (2014) Transcriptome analysis of the salivary glands of *Nephotettix cincticeps* (Uhler). *Journal of Insect Physiology*, **71**, 170-176.

Martin, J.A & Wang, Z. (2011) Next-generation transcriptome assembly. *Nature Reviews Genetics* **12**, 671–682.

Meister, G. & Tuschi T. (2004) Mechanisms of gene silencing by double stranded RNA. *Nature*, **431**, 343–9.

Niu, B.L., Shen, W.F., Liu, Y., Weng, H.B., He, L.H., Mu, J.J., Wu, Z.L., Jiang, P., Tao, Y.Z. & Meng, Z.Q. (2008) Cloning and RNAi-mediated functional characterization of *MaLac* 2 of the pine sawyer, *Monochamus alternatus*. *Insect Mol. Biol*. **17**, 303-312.

Paes, M., Oliveira, M. & Oliveira, P. (2001) Hydrogen peroxide detoxification in the midgut of the bloodsucking insect, *Rhodnius prolixus*. *Arch Insect Biochem Physiol* **48,** 63–71.

Pitino, M., Coleman, A.D., Maffei, M.E., Ridout, C.J. & Hogenhout, S.A. (2011) Silencing of aphid genes by dsRNA feeding from plants. *PLoS One*, **6**, e25709.

Prentice, K., Pertry, I., Christiaens, O., Bauters, L., Bailey, A. & Niblett, C. (2015) Transcriptome Analysis and Systemic RNAi Response in the African Sweetpotato Weevil (Cylaspuncticollis, Coleoptera, Brentidae). *PLoSONE*, **10**, e0115336.

Price, D.R.G., & Gatehouse, J.A. (2008) RNAi-mediated crop protection against insects. *Trends in Biotechnology*, **26**, 393–400.

Bansal, R. & Michel, A.P. (2013) Core RNAi Machinery and Sid1, a Component for Systemic RNAi, in the Hemipteran Insect, *Aphis glycines*. *Int. J. Mol. Sci***. 14,** 3786-3801.

Raps, A., Kehr, J., Gugerli, P., Moar, W.J., Bigler, F. & Hilbeck, A. (2001) Immunological analysis of phloem sap of *Bacillus thurigiensis* corn and of the nontarget herbivore *Rhopalosiphum padi* (Homoptera. Aphididae) for the presence of Cry1Ab. *Mol. Ecol.* **10**, 525–533.

Rosenfeld, J.A., Reeves, D., Brugler, M.R., Narechania, A., Simon, S., Durrett, R., Foox, J., Shianna, K., Schatz, M.C., Gandara, J., Afshinnekoo, E., Lam, E.T., Hastie, A.R., Chan, S., Cao, H., Saghbini, M., Kentsis, A., Planet, P.J., Kholodovych, V., Tessler, M., Baker, R., DeSalle, R., Sorkin, L.N, Kolokotronis, S., Siddall, M.E., Amato, G. & Mason, C.E. (2016) Genome assembly and geospatial phylogenomics of the bed bug *Cimex lectularius*. *Nature communications*, **7**, 10164.

Ruberson, J.R., Takatsu, K., Buntin, G.D., Eger, J.E., Jr., Gardner, W.A., Greene, J.K., Jenkins, T.M., Jones, W.A., Olson, D.M. & Roberts, P.M. (2012) From Asian curiosity to eruptive American pest: *Megacopta cribraria* (Hemiptera: Plataspidae) and prospects for its biological control. *Applied Entomology Zoology*, **48**, 3–13.

Sabin, L.RM, Zhou, R., Gruber, J., Lukinova, N. & Bambina S. (2009) *Ars2* regulates both miRNA- and siRNA-dependent silencing and suppresses RNA virus infection in *Drosophila*. *Cell* **138**, 340–351.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D.R. & Dean, D.H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. *Mol. Biol. Rev.* **62**, 775–806.

Seiter, N.J., Greene, J.K., Reay-Jones, F.P.F., Roberts, P.M. & All, J.N. (2015) Insecticidal control of *Megacopta cribraria* (Hemiptera: Plataspidae) in soybean. *Journal of Entomology Science*, **50**, 263–283.

Seiter, N.J., Greene, J.K. & Reay-Jones, F.P. (2013) Reduction of soybean yield components by *Megacopta cribraria* (Hemiptera: Plataspidae). *Journal of Economic Entomology*, **106**, 1676–1683.

Shakesby, A.J., Wallace, I.S., Isaacs, H.V., Pritchard, J., Roberts, D.M. & Douglas, A. E. (2009) A water-specific aquaporin involved in aphid osmoregulation. *Insect Biochem. Molecular Biology*, **39**, 1–10.

Sparks, M.E., Shelby, K.S., Kuhar, D. & Gundersen-Rindal, D.E. (2014) Transcriptome of the Invasive Brown Marmorated Stink Bug, *Halyomorpha halys* (Stål) (Heteroptera: Pentatomidae). *PLOS ONE* **9,** e111646.

Stevenson, P.C., Muyinza, H., Hall, D.R., Porter, E.A. & Farman, D.I. (2009) Chemical basis for resistance in sweetpotato *Ipomoea batatas* to the sweetpotato weevil *Cylas puncticollis*. *Pure Appl Chem*, **81,** 141–151.

Stubbins, F.L., Mitchell, P.L., Turnbull, M.W., Reay-Jones, F.P.F. & Greene, J.K. (2017) Mouthpart morphology and feeding behavior of the invasive kudzu bug, *Megacopta cribraria* (Hemiptera: Plataspidae). *Invertebr. Biol*. DOI:10.1111/ivb.12184.

Swevers, L., Huvenne, H., Menschaert, G., Kontogiannatos, D. & Kourti, A. (2013) Colorado potato beetle (Coleoptera) gut transcriptome analysis: expression of RNA interference-related genes. *Insect Mol Biol* **22,** 668–684.

Terenius, O., Papanicolaou, A., Garbutt, J.S., Eleftherianos, I., Huvenne, H. & Kanginakudru, S. (2011) RNA interference in Lepidoptera: An overview of successful and unsuccessful studies and implications for experimental design. *Journal of Insect Physiology*, **57**, 231–45.

Tomoyasu, Y., Miller, S.C., Tomita, S., Schoppmeier, M. & Grossmann, D. (2008) Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in *Tribolium*. *Genome Biol* **9**, R10.

Upadhyay, S.K., Dixit, S., Sharma, S., Singh, H. & Kumar, J. (2013) siRNA Machinery in Whitefly (*Bemisia tabaci*). *PLoS ONE* **8,** e83692.

Vera, J.C, Wheat, C.W, Fescemyer, H.W, Frilander, M.J & Crawford, D.L. (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology,* **17**, 1636-1647.

Waldvogel, M. & Alder, P. (2011) Kudzu Bug—A Nuisance and Agricultural Pest. North Carolina State University Department of Entomology Residential, Structural and Community Pests Insect https://www.ces.ncsu.edu/depts/ent/notes/Urban/kudzubug.html.

Wang, Y., Zhang, H., Li, H. & Miao, X. (2011) Second-generation sequencing supply an effective way to screen RNAi targets in large scale for potential application in pest insect control. *PLoS One,* **6**, e18644.

Wang, Z., Gerstein, M. & Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Rev Genetics*, **10**, 57–63.

Wheat, C.W. (2010) Rapidly developing functional genomics in ecological model systems via 454 transcriptomes sequencing RID E-7137-2010. *Genetica* **138,** 433–451.

Wonhoon, L., Joongnam, K., Chansik, J., Kim, H., Si, H. L., & Seunghwan L. (2009) Complete Mitochondrial Genome of Brown Marmorated Stink Bug *Halyomorpha halys* (Hemiptera: Pentatomidae), and Phylogenetic Relationships of Hemipteran Suborders. *Molecular Cells*, **28**, 155-165.

Xu, H.J., Chen, T., Ma, X.F., Xue, J., Pan, P.L & Zhang, X.C. (2013) Genome-wide screening for components of small interfering RNA (siRNA) and micro-RNA (miRNA) pathways in the brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae). *Insect Molecular Biology*, **22**, 35–47.

Yu, N., Christiaens, O., Liu, J., Niu, J. & Cappelle K. (2013) Delivery of dsRNA for RNAi in insects: an overview and future directions. *Insect Sci* **20**, 4–14.

Zhang, M., Zhou, Y., Jones, H.D., Gao, Q., Wang, D. & Ma, Y. (2013) Identifying potential RNAi targets in grain aphid (*Sitobion aveane* F.) based on transcriptome profiling of its alimentary canal after feeding on wheat plants. *BMC Genomics,* **14**, 560.

Zhang, H., Li, H.C. & Miao, X.X. (2013) Feasibility, limitation and possible solutions of RNAi-based technology for insect pest control. *Insect Science*, **20**,15–30.

Zhang, F., Guo, H., Zheng, H., Zhou, T. & Zhou, Y. (2010) Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV). *BMC Genomics*, **11**, 303.

Zhu, L., Zhang, Y., Guo, W., Xu, X.J. & Wang, Q. (2014) De novo assembly and characterization of *Sophora japonica* transcriptome using RNA-seq. *Biomed Research International*, 750961.

CHAPTER 5

Conclusion and Applications

CONCLUSIONS

Both *D. virgifera virgifera* and *M. cribraria* are agricultural pests in the United States, and while *D. virgifera virgifera* currently posses a greater threat to food security than *M. cribraria*, the latter has the potential to become a significant pest. Currently both species are fairly well controlled through the use of genetically engineered (GE) crops or foliar insecticides. While these control measures are effective today, increased use could lead to resistance, and as mentioned earlier, there are already populations of *D. virgifera virgifera* that are resistant to crop rotation, some insecticides and one or more *Bt* toxins. Since issues surrounding food security will only increase as the earth's population increases, finding new ways to control agricultural pests is of paramount importance.

My goal has been to learn, master, develop and test new control methods, ones expected to be more eco-friendly than traditional insecticides, and hopefully less harmful to non-target species, such as pollinators. Many studies, including those I performed during my dissertation work, suggest that RNA interference (RNAi) may be a viable control method for insect species such as *D. virgifera virgifera*. For example, I was able to interrupt the *D. virgifera virgifera* life cycle by introducing double-stranded RNA (dsRNA) designed to target critical ABC transporter genes. However, while ingestion of dsRNAs by *D. virgifera virgifera* can trigger "oral RNAi," this has not been shown to work in *M. cribraria*. Rather, in *M. cribraria* I was only able to trigger RNAi through injection of dsRNA.

In both cases I had to perform RNA-Seq to acquire target sequences, because neither *D. virgifera virgifera* nor *M. cribraria* had publically available genome assemblies, or high-throughput transcriptome data. To ensure the widest selection of transcripts, I sequenced a variety of developmental stages, and after selecting my targets, double-checked each target gene's expression profile. This allowed me to microinject each of the dsRNAs into *D. virgifera virgifera* or *M. cribraria* shortly before each gene product was required, thus generating a stronger phenotype.

Interestingly, my RNAi results from both *D. virgifera virgifera* and *M. cribraria* produced obvious knockdown phenotypes that were not only gene specific but also produced similar knockdown phenotypes as seen for the corresponding genes in other species. Of these, the most interesting phenotypes were those that impacted reproduction. Specifically, I tested dsRNAs targeting ABC transporter genes that resulted in egg-laying and -hatching defects in *D. virgifera virgifera*. This suggests their potential as insect birth control targets. In *M. cribraria*, injection of dsRNAs targeting an essential gene, *Mccat*, resulted in significant mortality and inhibition of nymphal growth.

However, while my results demonstrate the feasibility of RNAi as an alternative method for controlling these agricultural pests, to date oral RNAi has not been demonstrated in *M. cribraria*. Since RNAi needs to be triggered through ingestion of dsRNA to be practical in a field setting, I did attempt to feed dsRNAs to *M. cribraria*. My results were inconclusive, so were not included in Chapter 4, but it is unclear why my tests failed. One possibility is that the dsRNAs were broken down in the gut, but another possibility is that the concentration of dsRNA was too low. There have been reports of oral RNAi in insect species that had been thought to be immune to this technique. The difference was the use of very high concentrations of dsRNA. Again, it is possible that the dsRNAs are being broken down in the gut, but that having very high concentrations of dsRNA allows enough dsRNA to survive and trigger an RNAi effect.

As mentioned above, the fact that oral RNAi works exceptionally well in *D. virgifera virgifera* has been known for at least a decade, so it is thought that RNAi could offer an additional means of crop protection in the very near future. For example, the EPA recently registered Monsanto's new and innovative plant-protectant, SmartStax PRO, (EPA, 2017). In addition to expressing Cry3Bb1 and Cry34Ab1/Cry35Ab1, SmartStax PRO also expresses dsRNA that targets *Snf7* (Head et al 2017). Since SmartStax PRO pyramids *Bt* traits with an RNAi-based trait, it is expected to increase control of corn rootworms, even those that have developed resistance to *Bt* and traditional insecticides.

While RNAi has the potential of becoming a successful approach for insect control, practical use of RNAi in the field requires that dsRNAs function when ingested. In the case of Monsanto's SmartStax PRO, transgenic plants have been engineered to express dsRNAs targeting *Snf7* in the corn roots (EPA, 2017; Baum et al., 2007), and corn rootworms are highly susceptible to oral RNAi, thus making this product highly practical for field use. Another potential method for use in the field is application of engineered microorganisms (bacteria or yeast) designed to express pest-specific dsRNAs. Studies have shown that ingestion of bacteria or yeast designed to express dsRNAs can trigger RNAi in some pest species, even ones such as *D. suzukii*, which is known to lack systemic RNAi (Murphy et al., 2016; Zhu et al., 2011).

To further our understanding of dsRNA uptake and processing mechanisms in insects, RNAi pathways in different insect species and orders should be studied. Differences, such as what makes *T. castaneum* have strong systemic RNAi when *D. melanogaster* appears deficient needs to be understood. Likewise, we need to understand why *D. melanogaster,* which is not responsive to environmental dsRNA when unpackaged, has significant gene knockdown and high mortality when fed dsRNAs in transfection solution, or in the case of *D. suzukii*, is responsive to dsRNA when expressed in yeast. Taken together, this suggests that the presence of a robust systemic RNAi response may not be a precondition for RNAi-mediated insect control. Therefore, a better understanding of how dsRNA gets into midgut cells might increase our ability to improve oral RNAi in other insect species. Despite various setbacks associated with this technology, perhaps in the future, RNAi will be a widely-accepted tool for use in pest management.

However, several issues should be addressed before this technology gets used in the field. Although RNAi is known for its specificity, some studies have shown that off-target gene silencing can occur (Carthew and Sontheimer 2009). However, some instances of a single “species-specific” dsRNA triggering RNAi in more than one species may be less problematic than others. For example, Baum and co-workers (2007) administered a single dsRNA targeting the *DvvVATPase* gene to both *D. virgifera virgifera* and *Leptinotarsa decemlineata.* While this single dsRNA was effective in both pests, this has not been seen as a problem because both species are pests. However, if a pest-specific target triggers RNAi in a non-target species, and that non-target species is beneficial, such as the honey bee, this would have a very negative impact on the use of RNAi as a control method in the field.

Therefore, targeting highly conserved genes such as *polyubiquitin* (100% conservation between insects and humans at the amino-acid level) would be catastrophic to beneficial insects. Perhaps genes that are truly species-specific are better targets, or genes involved in species-specific biology, such as those involved in mate recognition, since this would reduce the possibility of silencing genes in non-target species. Clearly extensive studies would be needed to completely understand the factors behind off-target silencing and to prevent potential hazards to the environment.

Similar to insecticides and GE plants, there is also the possibility that insects will evolve resistance to a particular dsRNA sequence, due to mutations in the target gene, or due to mutations in one or more genes in the RNAi pathway. While creating transgenic plants expressing a different dsRNA, or targeting a different portion of the same gene or a new target gene, would overcome one of these resistance barriers, loss of the RNAi pathway would be devastating for RNAi as a pest-management tool. Therefore, research into mechanisms of RNAi resistance should be investigated before this technology is accepted on a large scale.

REFERENCES

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P., Ilagan, O., Johnson, S., Plaetinck, G., Munyikwa, T., Pleau, M., Vaughn, T. & Roberts, J. (2007) Control of coleopteran insect pests through RNA interference. *Nature Biotechnol*ogy, **25,** 1322–1326.

Carthew, R.W. & Sontheimer, E.J. (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642–655.

EPA, (2017) www.epa.gov/pesticide-registration/registration-rnai-control-corn-rootworm

Head, G.P., CarrolI, M.W., Evan, S.P., Willse, A.R., Clark, T.L., Storer, N.P., Flannagan, R.D., Samuel, L.W. & Meinke, L.J. (2017) Evaluation of SmartStax and SmartStax PRO maize against western corn rootworm and northern corn rootworm: efficacy and resistance management. *Pest management science*, **73,** 1883-1899.

Murphy, K.A.,Tabuloc, C.A., Cervantes, K.R., & Chiu, J.C. (2016) Ingestion of genetically modified yeast symbiont reduces fitness of an insect pest via RNA interference. *Scientific Reports,* **6**, 22587.

Zhu, F., Xu, J., Palli, R., Ferguson, J., & Palli, S. R. (2011) Ingested RNA interference for managing the populations of the Colorado potato beetle, *Leptinotarsa decemlineata*. *Pest Management Science*, **67,** 175–82.

APPENDICES

Appendix A: DvvABC naming chart

Subfamily	*Dvv* name	Contig numbers
A	DvvABCA_49125	DvvABC-A T49125_c1_seq2
A	DvvABCA_50718	DvvABC-A T50718_c0_seq1
A	DvvABCA_18330	DvvABC-A D18330
A	DvvABCA_266167	DvvABC-A C266167_7.0
B	DvvABCB_21313	DvvABCB_D21313+T40801_c0_seq1
B	DvvABCB_17742	DvvABC-B D17742
B	DvvABCB_19147	DvvABC-B D19147
B	DvvABCB_39715	D vvABC-B T39715_c0_seq2
B	DvvABCB_9796	DvvABC-B D9796
B	DvvABCB_13664*	DvvABC-B D13664+T8676_c0_seq1*
B	DvvABCB_17837	DvvABC-B T17837_c0_seq2
C	DvvABCC_41801	DvvABC-C T41801_c0_seq1
C	DvvABCC_44708	DvvABC-C T44708_c0_seq1
C	DvvABCC_48952*	DvvABC-C T48952_c0_seq2+T48952_c0_seq1*
C	DvvABCC_17573	DvvABC-C T17573_c0_seq1
C	DvvABCC_51687	DvvABC-C T51687_c0_seq1
C	DvvABCC_21020*	DvvABC-C S21020Locus_43217_0+T47823_c0_seq1*
C	DvvABCC_222633*	DvvABC-C C222633_3.0+T22098_c0_seq1*
C	DvvABCC_18126	DvvABC-C D18126+D19164
C	DvvABCC_49513	DvvABC-C T49513_c0_seq6
C	DvvABCC_14070	DvvABC-C D14070
C	DvvABCC_22628	DvvABC-C D22628
C	DvvABCC_20002	DvvABC-C D20002
C	DvvABCC_7536	DvvABC-C D7536
C	DvvABCC_47333	DvvABC-C T47333_c0_seq11
C	DvvABCC_49618*	DvvABC-C T49618_c0_seq2*
C	DvvABCC_45163	DvvABC-C T45163_c0_seq1
C	DvvABCC_43960*	DvvABC-C T43960_c1_seq1+D21270+T49513_c0_seq3*
C	DvvABCC_48940	DvvABC-C T48940_c0_seq1+D21306

(*) Represent short sequences

Appendix A: DvvABC naming chart continued

Subfamily	*Dvv* name	Contig numbers
C	DvvABCC_217405*	DvvABC-C C217405_3.0+D12550+S19317Locus_19671_1*
C	DvvABCC_10132*	DvvABC-C T10132_c0_seq1+D20603+T57212_c0_seq1*
C	DvvABCC_48300*	DvvABC-C T48300_c0_seq1+D21892*
C	DvvABCC_47673	DvvABC-C T47673_c0_seq1
C	DvvABCC_5345	Dvv ABC-C D5345
C	DvvABCC_22413	DvvABC-C D22413+T49258_c0_seq1
C	DvvABCC_18709*	DvvABC-C D18709*
C	DvvABCC_21941	DvvABC-C D21941
C	DvvABCC_15305	DvvABC-C D15305
C	DvvABCC_12562	Dvv ABC-C D12562+D20321
C	DvvABCC_10642	DvvABC-C D10642
C	DvvABCC_41602	DvvABC-C T41602_c0_seq1
C	DvvABCC_12703	DvvABC-C D12703
C	DvvABCC_14968	DvvABC-C D14968
D	DvvBCD_11014	DvvABC-D D11014
D	DvvABCD_11628	DvvABC-D D11628
E	DvvABCE_2830	DvvABC-E D2830
F	DvvABCF_2701	DvvABC-F D2701
F	Dvv BCF_802	DvvABC-F D802
F	DvvABCF_9935	DvvABC-F D9935
G	DvvABCG_9811	DvvABC-G D9811
G	DvvABCG_3712	DvvABC-G D3712
G	DvvABCG_14042	DvvABC-G D14042
G	DvvABCG_10897	DvvABC-G D10897
G	DvvABCG_22358	DvvABC-G D22358
G	DvvABCG_23081	DvvABC-G D23081
G	DvvABCG_13051	DvvABC-G D13051
G	DvvABCG_38769	DvvABC-G T38769_c0_seq1

(*) Represent short sequences

Appendix A: DvvABC naming chart continued

Subfamily	*Dvv* name	Contig numbers
G	DvvABCG_13829	DvvABC-G D13829
G	DvvABCG_49457	DvvABC-G T49457_c0_seq1
G	DvvABCG_36869	DvvABC-G T36869_c0_seq1
G	DvvABCG_79525	DvvABC-G T79525_c0_seq1
H	DvvABCH_20789	DvvABC-H D20789
H	DvvABCH_5118	DvvABC-H D5118
H	DvvABCH_18290	DvvABC-H D18290
H	DvvABCH_11818	DvvABC-H D11818

(*) Represent short sequences.

Appendix B.
Sequences of ATP binding cassette (ABC) transporters genes in *D. v. virgifera*

```
>DvvABC-A T49125_c1_seq2
MSLLQLEVVVWKNYIIRKRHWFLTIIESFLPVCLFLLIAYGRSQITGLNKIEVPNPTYGRIDGINYNIDVAGTYILYTPYNEYYTNITTKMLTKLGM
PSDHIQRFGSENELLGYYYKNPNYTTVAIIFDDKNPKHFNYKIRYHKNYYQPEFLVTDRMYQRVYGYSPGSEYVYRDEDFLKIQLALDMSFLEENLR
SSNIEDTYSIQIQEFPYPPHKVDSAITTLFLEFLPLITLFSFIFLCPAVLKRVVEEKHSGTKELMKMVGMKSWMLWFGWFIYAMIPMFFAVSVISIF
MKVPMFGSDSPLVEFANGGILFVFLILYCMAAVAFCFAISSFFSRPTIALVAGILVWILSFFIPKYACGLDEANKLSWLSNVLLNLLPNMSLHYGYS
IISVFEEREVGINWSNFFKPGSGSSDDITMLNVYVMLIVDIVIYTIFTFYMDGVNPGKYGVRKSILFPLQNFMKLCRKPSITTVPVDSETVPLEKVE
AGHNLAKGIQINHLHKRYKQKQAVNNLNLDIYKNQITVLLGHNGAGKSTTMSIITGLIPATSGSVTINGLDINTDMDEIRKSLGLCPQHNLLFTDLT
VKEHLLFFAKLKGKSTKEANIEAKSLLAKLNMPDKEHSMAQTLSGGMQRKLCLAMALIGDSKVLILDEPSSGMDPQSRRELWDLLLQWRGEKTILIT
THFMEEADALGDWIAIMSDGSLNCYGTPMFLKKKYDTGYHLNLMIEEDADIDAISRRVKHFMPDAHLTSNNGNNLVFVLPYQNTNMTGLLGDLEKNR
SELHLSNISITITTLEDVFLKTRQEIEHSSDDTSSVATQDDIQPHASLTLITLLKKKVNFSLKKWSTYIVPVGVAVVLFSLTVFLSSGNDYYSERGP
ELPLKLSSYKQTSVYYSGDTSNEKINTLMKYYMSAVESQRSLPFKVDNVEKSILQRSLENIAYYKEHMIAGANFVYDRDFTLVALYNGIATHSTPIS
LNLITDSLAKTLLGPEYGISVSNWPLPQVQERLSSQEYSEAKVAVLWLILLPVGCLFIHGIFIIFPHTEISTRFLQIQYMAGVKPFFYWLVNWIADM
TFYIFLMFILSFLLWICSPVFQHNGTELGYLFSIFLCYGIAGIPFAYIFSRKKTASGAFALFVIMGMFLGIILTLTIAVLLESQDEYYVNIGNKIKY
VCFFLLPQVSLSDALVSFIRRTVNIYNFKISPQRLRVMCNQMESHACCVESSIECQNYKSYNNLFSEHYMFMIGCGVFYLTINIILDTYFMKKLKAM
VMRQCNLILKSFKDKDTALIPKNDKYVNEDVGDGYNTLRAKKIMKLYAGKQIVKNINFTLKHGHCLGILGVNGAGKTTTFKMLTREEVVDDGEIKIE
LDNNKQPLDITGSEYLETIGYCPQSDSLNFVLTGRQLLSTVAKLRGVTDETMIDRFLEAFDLKQYADIPCGHYSGGNKRKLSLAISLIGNRKFVLLD
EPTNGVDPATRRKCWDLIKLMKGNSGNKIGFILTSHSMTECEALCDELKIMKKGSFVEEGRLVDLKNRYGGFTLKLKLVSNQGQPINLVDDDVDEVD
GVISNKFKSVDDLKSYFTNHDRGEIKDEHSGLLHVYIKDKTKKWSDIFQEVEALKTHNSHLIEDYAISEASLEDVFLKVAREDEEDVTKKK
>DvvABC-A T50718_c0_seq1
MGEKLDKFLLLMWKNWLLQYRKPIQTIIEIVAPVLFSFLLVYIRSLSDPTYEHERIYKPFCAFEGNFNLSLPLFCNDPDAVQAGPSLFSSMQTGDSN
FSLTILYTAPEDKYYDDVIGFFKLIPKVNTQRVISSGELERISLAGNSSYYGIEFGGDRDNVEVNIRFPGETVYQLDSIGRQNWRTNLVYPVFQSSG
PRSPDNATGSAPNYFAEGFLAVQHFVTTALLLADKGITHEDPQEIFNFLMDKSNPFPLIKMQRFPFASWFSDQLLTALTSMLGMIVMLSFVYTCINT
VKSITTEKEKQLKESMKIMGLPNWLHWLAWFVKCFIFVIISVALMVVLLKVRWYSNTEFTVFTYADPFVLFVFLLLYCCSMITFCFALSVFFSKANT
AATMAGLAWFISYAPYLFMQNNYTDLALSTKLLTSLLPNTAMAFGFQVVLMYEGTGDGVTWSTLFKPNTPDDTLSLGLIWIMIIINTIIYLLIALYI
EAVFPGEYGVAEKWYFPFTKKYWCGKGAVSDYNNYNDHSANEFFETEPNLRPGIQIFNLKKAFGKKTAVRNLSLNMFENQITVLLGHNGAGKTTTMS
MLTGMISPTNGTAKINGHDIRLDIAGVRNSLGLCPQHNIIFDELTVEEHLYFFSKLKGLSKSKITAEIDKYVKLLDLEPKRKEKSSTLSGGMKRKLC
VGMALCGNSKVVMLDEPTAGMDPSARRVLWELLQKQKEGRTILLSTHFMDEADLLGDRIAIMAGGELQCCGSSFFLKKKYGAGYSLIMDKAKTCNPH
KVTELLKKFIPEIEIHSNVGSELTYLLSENNAPVFEAMLKQIEQESNELGIRSYGISLTTMEEVFMKVGADHGQEEIYNEKQKKDKKEKNPILTSAD
LGPESLSPSYTGGLHLIVNQFVAMLMKKIIGTFRSWVLLAIQIFLPMINVIIVMSVPSGSVPSVLPAMPLNLARFTDPITLIERFDDPQNYTNNFYG
VLNGYGMSPTNIDNITRIMLDLTATNPNVVRRSYIIGSTFTEQEFEFGIIKTIHPVITAWFNNDPYHSPGIALGMALSSVYNHYTNGHIEFVNKPLP
FKAATQIDNIVGGQSQGFQFAFNIGFSMAFISSFYVIFLIRENVSKSKHLQFVSGVKVYIFWFVNIICDMFVYLLVCSVLLITIYCFQQDGFKTSGD
MGRLFFLLLLFGWSFMPIYYVASLIFTVPSTGYTRMTLVGIFIGNAAFLLVEVLKTQSTNLKRIGETLDNVFLVFPHYSLATGINKCYAIYSYNTLC
ETVFDSCAQNNFTKDECISKFPSTVADICTNLNDNYFSWNNNGIAKNVTYSLISGVLWSILLFIIEYKFIARLMYYINQKFFPKQPILIQDEDDDVS
KEKERIHMATDHDIRQTNILVVKDLTKYYNNFLAVNGLSIGIEKSECFGLLGINGAGKTTTFKMMSGDETVSYGDAWIEGKSVKTQLKQVQRNIGYC
PQFDALLDDMTARETIIMYCLLRGIEFKRTQRIANFLSRDFDFHRHLDKKVKEMSGGNKRKLSTVLSLIGDPPVLFLDEPTTGMDPATKRYLWDSLC
KIRDNGKCIVLTSHSMEECEALCTRIAIMVNGNFKCLGSTQHLKNKFAEGYTLTIKLKKLPESGGLVHADTESLEKYIKDKFPYAHLREKHQELLYY
YITDTSMAWSTMFGILERAKRSDLNIEDYSLGQSSLEQVFLTFTKHQNPEGDDVKKKN
>DvvABC-A D18330
MSAESNSNRKRNHSYTVLQDGSDIGFQEPRGAMAENLDKFILLMWKNWLLQYRKPVQTAVEILAPVIFSILLVVIRSLSDPVRHETVVYPPFCTIPI
ALRENKTGITICPSYDGVPYMDRNQSDGGNSSNNPFEKFALVYSPSNPPIDQVMNYFRLAFKNVVALESSQALEKYFITNSSNITFAGIQFDDSYKT
LKSLDDVKHLQVSIRFPGETRLKLDPFNYNNWRTNLIFPIYQQPGPRLYNLITGAAPSYYREGFLGLQYYLTLSVLLAKSNVTYATDDYLDLISWVI
ENQFPLVNMRRFPNAPWYEDILLTALKSLIGIIIMLSFVYTCINTVKSITTEKEKQLKESMKIMGLPNWLHWTAWFVKCFMFLLISSVLMVIFLKVR
WYTNTNFTVFTKADPFVLLLFLMFYNCATITFCFALSVLFNKANTAATIAGMVWFLSYSPYLFMANVYDTLTLTSKLVASIGSNTAMAFGFQVILMY
EGTGEGIQWNNIFTPNTPDDSLTLGLILIMLTVDSIMYLLIALYIEALFPGEFGVPQPWYFPFTAQYWCGHPIYRGVEDFDNGAIKGEFFETEPENL
KPGIQIRNLKKVFNQKVAVRNLSLNMYEDQITVLLGHNGAGKTTTMSMLTGMITPNGGTAKISGYDIRTDMEGVRKSLGLCPQHNIIFDELTVAEHI
YFFSKLKGMRKGEIKSEIAKYVDLLELQEKRNSKASTLSGGMKRKLCVGVALCGNSKVVMLDEPTAGMDPAARRALWNLLETQKDGRTILLTTHFMD
EADILGDRIAIMAGGQLQCCGSSFFLKKKYGAGYSLIMDKSQECDPRRVTQLLKKYIPDIEINSNVGSELTYLLVDDHVHVFEPMLRELETESEMLG
IRSYGISLTTLEEVFMKVGADHGQEEMYNHEHGNIVQNGTAQNGFTNGINGTHKTNNGTHTMVPTYSNGFSLLLNQIIAMLLKKFVSTVRSWILLGI
QVMMPTLFLIIAFVVARKNKMTGNLPAMPLSLSKFENPVTLVENGTSDYLPYYMKVLEDYGYPATIVDNITSVLLDKTEHHPILVTRRYQAAATFGE
SGIPDLPFNLPNLTAWFNNNPFHSPAVSLSLMLNSIYRKLGGCDDCTIEFTNSPLPYSAATQASQLLTVQNIGFQLSFNIGFSMSFVASFYVLFVIR
ENRCKSKHLQFVSGVKVYVFWLTAAFCDMLTYLFTVFVLMITMVMFQEDGFKSGSDISRMFFILFYFGWAFLPMFYLSSYFFQVPSTGYTRMTLVSI
FGGNAAFLVVQVLQSPGLDLQYIGNALHWLFLIFPHYSLATGINESFKVYAYNNICANLLKTCEEQHIPKKTCISIMRNERIREICEDADLNYFKWK
APGIARNMVYSFLTGIILFALLLAIEYKIFSRVHYYLTQKHFTKKPIPVEDEDSDVSKERERIHVATEIDIKQNYTLAVKDLTKYYKNFLAVNGLCV
GIRKFECFGLLGINGAGKTSTFMMMTGDTSISYGDAWVNGKSIKQHLEEVQKIIGYCPQFDALLDDMTAEESIIMFAMLRGLPFKDTFKLADYLSKE
FDFTRHLKKKVKELSGGNKRKLSTAIALIGDPPLLYLDEPTTGMDPATKRHLWNALCKIRDRGKCIVLTSHSMEECEALCTRIAIMVNGNFKCLGST
QHLKHKFAEGYTLTIKLKKIAESSSEGLSETEPIEKFICQRFPGAQLRERHQELLNYYITNKSVPWSKMFGILEKGKRSDLNIEDYSLGQCSLEQVF
LLFTKHENST
>DvvABC-A C266167_7.0
MAYGGKSLFFSQLKAMLKRNLLLKKRQKLKTIGEILFPIYFLALLVVIELVLPDPNLPEISTPRGEEYLFKSFDNGESHKIAYAPNTTTNINFVNNV
TETWKKMSNSKSSLEWVPYSAEDKVQEAYDQDHKFAPMAVIFKSDVPYVHTSLSYQIRTNPYYISTPSSNDIGWDKQSCRGGTNPNSNMEDGSTCPA
NTYYFSGFLALQTLLDYTRIQLDNPTQDNIPYISLEIFPKAATTVGDFANMIRIIIPIYMVLALSQFVTYLIILIVGEKEKKIKEGMKLMGLMDKVF
WLSWLIIYACFVILLATICVLLLFAFKVFVHVNLLLIFILMVLYGISLIMFAFMLTPFFDKARTAGVLSSFILVLIIVLYFVEQIVSGKHPQYLWAL
SLLSPAGFASALDKVAISDIQGKGLDFSNMWDNHGSGVPFGGSLVMIMVDIVLYAIIGWYLDNVIPSSYGVKRSPIFFLMPSYWRSSHNQIPPNQLN
TESSVDVEPVPREMQDKEAIRIINLNKSFTQCRKPTVTALDGINLSIYKGQITAILGHNGAGKTTLFNILTGLSSPTSGSALVFGYDVSNPNEMDKI
RRMTGVCPQHDILFDDLTPREHLEFFAAIKGISNRQSAIEKIIREIDLLDKIDTASRSLSGGQKRKLSIGIALIGDPKIIILDEPTAGVDPYSRRHL
WNVLQNVRRDKVILLTTHFMDEADILADRKAVVSKGKIRCCGSSLFLKNKFGIGYHLTFVLEDKSNENAINHLVLQFVKMARKDRRHGKELSFILPH
NAVENFASLFAAIEHEISIKSDLGISSYGVSMTTLEEVFLGLQKEEEYGDATVYQDLPQSSEGISNENRALTEGIRSLEAFRCTPSTSQNLRTLIGL
RLLRLRREKRKLWMVIILPIIFTGLGLYLNKVIDTGKKYEPPKTYPMPLDLSHYLGYNISIYNGSPADLTTFKNQLEKNGAELDTYDGEFTSLLDIA
```

PHYASFNIHDFSSSSQNISVLYNGSYSNNLPMFINLISNTFYSMAATTGKIAVTTYPFELPPTNVSPTPASPGNFILGMIFLFAPIVLAVDMVYERE
IKARNQLRVNGLPFTVYFTSFFLVQILLMATITILLIILIVVMKPPTFSNSSITVLAIWVILYCPASVLFCSCVSYMFDKSESAQSVMPNVSTLLGL
IPYISVMYAKESVATILHYVFAFTDMMYVPYGMLYYIQKINLECISDPHCEGTTFSSFMIPEIIVLFVTLIIQIPLLFLIMLILDVKKNGGQVLDIF
RSKKISENIVEESRDVGVHEDRDVKNERQRVNNLIRDPQNNRSVITVENLHKVYQKGVKSGICSRSAESLKVAIKSISLAVDSAEVFGLLGHNGAGK
TTAMKIITAEEAPTRGRVQIVGRNITSSANTAFQYLGYCPQHDAQWKNITVKEHLELYSHIRGIPKNEIKRTVDLYLTGLQIHEHKDKQAEKCSGGT
RRKLSYAMAMIGNPKIVLLDEPSTGMDPQSKRFLWDTVLASFRGSRGAILTTHSMEEADALCSRIGIMVNGELRCLGSSQHLKNLYGAGYNLEVKLA
TQQGTDLRQKLKELENYVFHLFPNAVLQETFGDRLIFSVPQQSVPSLANCFRSLEDGKRRLNIEEYSFSQTTLEQVFLKFAQEGEHED
>DvvABC-B D21313+T40801_c0_seq1
MGNKKNENDLKCEKNKTNEISYLRLFQYCTTLDIIFMSIGAICAALSGIVQPYSMTLFGDVTGAIVTYASNYNESLSEPEKTLLADELINAVWLFGM
KSVGVGIGVILTTYISTVLFIYSASRQIFKIRKAFLEKTLNQDIAWFDQNRTGDFASTFTQNISKLEEGIGEKIGTFLFFESTFVAGCVLGLVKGWK
LALVCMVSLPLSTTIMTIISWISTKFSAQEMESYGEAGTIAEEVFTAIKTVVAFDGQDKEIERYNKHLVDAKKNNLKRSFFTAVSNGCLWFFVYACY
ALSFWYGVTLILNDRHLPEHEQVYTPANMVSVFFSTLIATWNFGLGAPLLETFGTAKGAAQKIFFVLDSKPVIRKFANSGIVINDAALEVTFEDVHF
SYPSRSDVKILRGINLTISKGETVALVGKSGCGKSTCIQLLQRFYDPDMGQIKINGIDIKQLNLDWLKQKISVVSQEPDLFSTTIAENIRYGKLNAT
QQEIENAAKKADIHTFIQTLPRGYQTVLGERGTQLSGGQKQKIAIARALVRTPEFLLLDEATSALDTTSEAEIQEALDSIRGTCTIIIVAHRLSTIR
QASKIVVINEGKVLEMGTHTELMDMKGAYHNLVISQGLTETLEEKGNRRRSRKFSEANKSINEKEDEELEDTQINQPVAKNILWKVLKLNASEWFYI
LIGCLSSLITGASLPIYGLVFGGIMGIFANDNDGEVRSESNMYCLYFLILGVVTGVAMFWQTTSFSVAGEHLTLKIRSKTFEAMLRQEIGWYDQKSN
GVGALCARLAGDAVAVQGAAGPQIGTTINFISTFILTCTFSFYFEWRTSFVLFSLCPVIFFSVYFEQKVLQEDATKNQKMLEASAKLAVEAIGNIRT
VVSLGCEKVFMEQYIKELLPYQKMARKKSHYRGIIVGLARSLMLFAYVAGIRYGINLIISGDCPYGTIFIVCEVMIVGTWSVGNALSLSPNFQKGLV
AACRIITLLERQPVVQNMPDALNFLWEDENVEYSEVYFSYPTRPSIPILKALNLLIPKGKTVALVGSSGCGKSTIIQLLERFYDPSYGKVEISDKNI
RYVDLKSLRSQLGIVSQEPNLFDKTIAENIAYGANYKQVEMDAIIDAAKSANIHEFILNLPKGYETKVGSKGTQLSGGQKQRIAIARALLRNPKILL
LDEATSALDNESEKIVQEALDNARKSRTCITIAHRLTTIQDADVICVLKEGNVVEMGTHKELLEKQGLYYKFYKLQSVESIS
>DvvABC-B D17742
MTEEKKHSXKDKEKIGIDAQFVNSEEPKEKIKNVSFPQMFRYASTYDKFLMVVGLISATGTGVLQPLNTILFGSLTGDIIAYATSIQINLPADQKKI
AEDNFFDGIRYFALMNSLIGIGMFVFSYLATVTFNYSAMRQIFKIRSAYLKSILNQDVGWYDINQTGDFASRMSDDLFKFEDGIGEKVPVFWSFQVV
FLTSLIIALVKGWELALICLTSLPATLITIGIVGLLTTKLAKNELEAYGAAGXIAEEALSLIRTITAFGGQKKEVDRYNKNLVEAKNNNIKRSMFSA
LGFGLLWFMIYASYALAFWYGVKLVLRDRTATNQIYTPSNMVTVFFSVMTGSMNFGIASPYIEAFGISKAAASKIFSVIDHKPTINLSKGNGKTLNV
LIGNIQFKDVAFRYPSRKDVPILKGLSLNIKSGDTVALVGSSGCGKSTVIQLLQRLYDADSGEVTIDGKNIKEYDLTWLRSQIGVVGQEPILFGTSI
LENIRYGKDGVTEEDVIQAAKKANAHNFIKALPNGYNTLVGEKGAQLSGGQKQRIAIARALVRNPTLLLLDEATSALDNTSEAKVQAALDAASVECT
TIIVAHRLSTIRGANKIIVLSQGVVVEEGTHEELMELKQEYYRLVTAQVKSSEQFEVAEKKKVVRAISLAESSTGSDHNIEATKEDNEDDFNENKDV
SVFEILKMNAPEWPYILFAGLGSIVVGCGMPVFAVLFGSILGTLANGDPDFVRSETNKYCLYFVLGGLITMVSVFTQMYLLGIAGEKMTERVRSRLF
KAMIYQEIGFFDKKTNGVGALCAKLSSDASNIQGATGIRVGTILQSIATFCLAIGLSMYYEWKLGLVTAAFTPVILIAMFFERRNTRGGNDSRDSAL
QKSTRTAVEAVGNIRTVASLGLEEKFQQLYESELMPHYKSSLKTVHWRAIVFGLSRSLLFFAYATAMYYGGFLIRDGLPYDRVFKVSQAQIMGTVSI
ANSLAFSPNFAKGVAAAKKVKSFLSRIPLIRDLPSSRQMVKASGNFSFSEIEFTYPTRPNVLILKGLNLDILNGKTVALVGESGCGKSTIIQLIERF
YDPRSGEVKMDGVDLKDISLDSLRSHMGIVSQEPNLFNKSIAENIAYGDNSREVSMDEIIKAAKNANIHNFITGLPKGYETKLGEKAVQLSGGQKQR
IAIARALVRNPKVLLLDEATSALDTESEKVVQEALDQAKKGRTCVTIAHRLTTIQDADLICVVANGVIAESGSHQELLQKEGLYYKLYTQKT
>DvvABC-B D19147
MNVYSVIVSSHNLSKSLKLLKLVNSRRYFSNMSQKISQNKTYISVEIPETTKGASRGFFSRIFKRNKEVKEKQEPRLKKSELKRLFSLAKPEKWKLT
TAIGFLVVSSTVTMSIPYSLGKILDIIYVGSGDSDAARARLNQVCGILLGVFVLGAICNYARVYLMSTAGYRMTNALRRQVFGAILRQEQGWFDKRP
TGELVNRLSADTQIVGSALSQNISDGLRSLVMVFAGTGMMLYMSPQLALVGLAIVPPVAVVAVLYGRFVRKISRKVQDSLADSTKVAEERIANIRTV
KSFAQEPREITSYNISIENVLKYCYKEAKARALFYGMTGFSGHIIIISVLYYGGVMVSSNTITVGNLSSFLLYAAYIGISVGGLSSFYSELNKSLGA
ATRIWEIIDREPTIPTQGGLIPMTDVEGHIEFKNVKFCYPSRSDIEIFKSLVLDIQPGKTLAVVGPSGSGKSTLAALLLRLYDPIDGAVYIDNQNIK
ELDPAWIKKHIGTVSQEPILFSCSIKENILYGADDPQKINDDEFIRICKEANVYEFVQNLPEGFETVVGERGVMLSGGQKQRVAIARALIKNPKILL
LDEATSALDAQSEHYVQEALDRVMKGRTVLTIAHRLSTIQNADIIAVLQNGQIVEQGKYEQLLQSNGPFRELVQHQTFSQLMDD
>DvvABC-B T39715_c0_seq2
MATVLNLSNIPFKAINRHLKTVVYQKQCIFKVYGQNTKNVYLSQFYYTKEYSTKTEKKSNSSKGSGGILAGILSKIPSSAPKAQPIRGCFHPGASTL
TRETISLQKQTPITGRQMVGAMLQYIWPENDKSIRDRVKLAVSLLIGAKVMNVCVPFIFKYSVDYLNVGNALNMETAPQTVATVATSLLLGYGIARA
GAAGFNELRNAVFAKVAQHSIRKIAKNVFIHLHNLDLTFHLQRQTGALSKTIDRGSRGINFVLSAMVFNVVPTMFELALVSSILGIKCGAMFAGISL
GCVGVYAMYTLSVTQWRTKFRVYMNKAENEAGNKAIDSLINYETVKYFNNEKYEAERYDSALQKYESASLKTASSLALLNFGQNAIFSAALSGIMIL
AANEIVKGNMTVGDLVMVNGLLFQLSIPLGFLGSVYREVRQALIDMQTMFTLMTMDSAVKSKPEAPYLHVDSKSSPIKFENVSFEYGPGKTIFKDLN
LTIEPGKKVAIVGGSGSGKSTLVRLLYRFYEPSKGRILIGNQDIRDVDLDSLRRAIAIVPQDSVLFHDTIRHNLHYGDLQATEEEIMNAAKLAEIHS
SIVTWPQGYNTQVGERGLKLSGGEKQRVAIARAILKNSPILVFDEATSSLDSITEHNILMALRNATKGRTSICIAHRLSTIMDADEIIVLENGQVAD
RGTHKSLLESHSGLYWKLWETQNQSHTEPEKKKSAV
>DvvABC-B D9796
MLNISYCPPNISVWDVWVDHGIPQCFMNTVTSSVLAIYILIAGSIQLHIYRRWGVESGTRFLPRGRLYYLQTFLILFIPVLEIVRFILTATVYDDKH
IYGYMIVSLVLTTFAFPFSLWVLKVELYNILPSVQTRGHGIVLLLFWTLAFISENLAFINLTQENWWFKLKDLTDQLEMALFILRYIACLLIFFLGL
KAPGIVNEVDYYMLGGHQRNVINPNNENASTWKNFWKKVKILAPFLWPKKNFTLQFKVLICFILLLGGRAVNFVPIYQKLIVDSMEETVAKMLFRWD
WVLIYVGLKFLQGGGTGGMGLLNNVRSFLWIRVQQYTTREVEVELFRHLHSLSLKWHLGRKTGEVLRVMDRGTDSINNLLNYIIFSIFPTIVDIIVA
IVFFVSAFNIWFGLIVFTTMILYIVLTIVITEWRTKFQRRMNLADNETRSRSVDSLLNFETVKYYGAENYEVDAFREAVLKFQDEEFKSSITLNILN
TVQNVIICGGLLAGSLLCVYMVVEPKTLKAGDYVLFATYIVQLYVPLNWFGTYYRAIQKNFVDMENMFDLLREEQEIIDAPGATNLSVPRGHVEFKN
VSFGYLPEKLILKNITFSVPSGKTVALVGPSGSGKSTIIRLLFRFYDVDTGVIVIDGQNIKTVTQESLRRAIGVVPQDTVLFNNTVEYNIQYGRLTA
TVTDVIEAARGADIHEKILTFPEAYETKVGERGLRLSGGEKQRVAIARTLLKAPNIILLDEATSALDTQTERNIQESLNRMCANKTTIIVAHRLSTI
IHADEILVLQEGEIVERGKHDHLIGQEGIYANMWRQQLENKDKESLENSSESAKSK
>DvvABC-B D13664+T8676_c0_seq1*
AFTKLNMTATQVEDDQEKKQKNNVKFDNIAVNLKFKNKENTSTWSNFWKKLKILSPFLWPKKDFVLQLKVLVCFLLLGGGRVVNLFVPIYQKLIVDS
MQGTVEKMVFRWDWVVVYVGLKFLQGGGTGGMGLLNNLRSFLWLKVQQYTTREVEVELFRHLHSLSLKWHLGRKTGEVLRVMDRGTDSINNLLNFLI
FSILPTIVDITIAVIYFVSAFNVWFGLIVFTTMVLYIVLTIIITEWRTKYRRRMITAENETRTRSVDSLLNFETVKYYGTEKFEVQAFREAVLKFQV
EEFKSNVTSNILSTVQNIIVCSGLLAGSLLCVYMIVEPHTLQPGDYVLFATYIVQLYIPLNMFGNYYSMIQRNFVNMENMFDLLRQEQEVVDPPDAN
DIVVSRGGVEFKDVTFRYLPEKVVLRNISFSVPSGKTVALVGPSGSGKSTIIRLLFRFYDVESGLIVIDGQHINTVTQESLRRSIGVVPQDTVLFNN
TIKYNIQYGRLTASDADVIEAARGADIHDKILTFPESYETNVGERGLRLSGGEKQRVAIARTLLKAPNIILLDEATSALDTQTERNIQESLNRMCTN
KTTIIVAHRLSTIIHADEILVLQEGSIVERGKHEDLITQGGIYASMWEQQLQNKDKEPAENKPESSEPKKRPPKR
>DvvABC-B T17837_c0_seq2

MWRLIQQNLHNKLLFHNSAIFKKYFCQGHIKNYISRTVSQKPVPKSFIPSKVATFGVTLTGGLIIKLYISHNGVLCEAKKTRMAGYENKSDKNVKFD
WARLWKYLKPHIWYFLAAIVGALAVALLNIQIPQVIGGVVNVLAKFSESRDSELFLNEMKRPVIKLISMYLAQSVCTFFYIFMLSNLGERMAYKMRT
DLFESILKQDIAFFDQQRTGEIINRLTADVQDFKSSFKQIVSGGLRAATQIVGCSVSLIMLSPQMTFISLLCIPSVIAVGTVFGSLLRSVSRRAQAQ
VEKTTAVADEAVSNIRTVRAFAMEDQEKELFNTEADRAMVLNEDLGLGIGLFQAGTNMFINGMVLSTLYMGGYLLSTNQLSAGEVMAYLVASQTIQR
SLAQISLLFGSVVRGVAAGSRVFEYINKTPKMALTGGKVLPYESVKGDIEFKNVCFAYPTRSQQIILQGFNLSVPSGKTVAIVGASGNGKSTVVALL
ERFYDVKDGSITLDGHDIRSLDPSWLRGRVLGLISQEPVLFGTTIMENIRYGKPDASDDEVKEAALLANADEFITSFPKGYNTPVGERGVTLSGGQK
QRIAIARALLKNPVVLLLDEATSALDTESEKIVQQALERARTGRTVIVIAHRLSTIQNADLIVVLNKGKIVEMGTHESLKKLGGYYWSLAYQQQNSP
AG
>DvvABC-C T41801_c0_seq1
MDVATETKEISPEYEACLLSRIFFLWTIPFFRKILNGTLKLSYFAVNAEDDAKILLKQLEEKWAAEVKRSKENKEKPNLKKVVMGIFRKQLAIIYTF
TGVNFVILRVITCVLLTILNRTFTEEYRFRERIMLGGMVVVLVVINYYSDYIFHYQLRKIAIRMKAGLSSLIYNKLLVLHLQSMEEATVGKIINILA
DDLDRIVQYTTYLGYLILTPLQILTLYASMWYLLGTAAFSGALFFTILLPIMVLGSKISSVYRLKASKIADSRVNMMKEIISGIKAIKMYVWELSFT
KIINDYRKEELRNITKYLLFQIGYYHFAVLKLASAISVLTMLYVNGILFPEIIFSCIQFLVILKKGVLAFMPQAMLLQFEITNVFVRISELLLMNEV
DIDSKINADVSGIQLDQIKASWSTKSSFVISSNNINIPSRSLTVIIGPVGSGKSTLLQLLLGELKLSSGNIQMNGSVSYSGQESWLFTSSLKNNILF
GNDYDYNKYQQVVSACSLLPDFDQLRNGDLSCVGEQGVLLSGGQKARINLARAVYRDADIYLLDDPLSAVDSEVGKRIFDECVKSYLRDKTRLLVTH
HLQYLQEADQIIIVDKGNLESFDSLSEIKFKKPEIFEWIQKLNKADETKETTVTSKLNLKDESDQKLNYVCNENRTKITPRRALFEYIKSGSCIIPL
FLTVFSAQLLLGLTDYWLLYWTKQSYSTHAESKNVTENSNVHHMRDIQWTTEFNTNFTIYCSIVSSAFLMLILSNILLTKFFNTASKNMYKNMYKAV
ISAPLSFFYTHPSGRILNRFSSDTTAVDVRFADNLKDISQVYFMGIGSIILVIVSDYYMSFFVLIVFIYAKLLKWYQYIGSELKHLEITAKSPIYTY
INNTLSGITTIRATKNQEKLQKEFIILLNNHTSVLRLNVLCGAAFSVIVDTVCAFFLFGVIFALILLNEYDHKITPGVAGLAILQVFQLMGALQFAT
QKSLDFVQYLMSIERMIEFSKIEPEEFSETKKNGVYNFQPGYSGAEVVFKNVSMFYNKGSLVLNGINVSIASGEKVGVVGRTGAGKSSLINVLLRLS
EFSGSVQINGVDTKTMPLEILRKRVSIIPQDPVLFTNTVRYNLDPFGEFTDDNLWTVIKQVELKDCITSLDSVVTSGGNNFSTGEKQLICLARAILR
KNELAILDEATANVDQRTDELIQRTLKMRFANFTLITIAHRLQTIMQTDKVLLMDNGYLVECDHPYRLILKRGKFYDLVQQTGDATAAHLEDMAYK
HFTQHHS
>DvvABC-C T44708_c0_seq1
MDSSKIKKRRLNPRDSANIWSLVTFSYITDLLKRGVQKDLEEEDLYEVSKNCASKHCAEQIEKKWKKGKKDGKYPSLFLTLCKVFGWKYFLWGITQL
FCELFRSIFEPNAITNLVSYFQPGQTKLSKSDAYFNAAILLFSNFLQKIYFFNYDLFLFIMGIQVRTALCSLLFRKVLKLSPAAMTKASLGNIVTLI
TKDVQQIRRSMFAFNDIWVFTVVISVTCYLLYARLGVLTFFAVGMHFCIIPIEFFLGKMVTKMRNETSEMTDERIQVSQEVLSATKIIKMYTWEDYF
YQKINLARTREVAKMVIVFYLRMILMLLGAISTKIGFYLMIMGYIWMEQPPDASIIFFITAHYDNIVIFFGYLLPEAVGRFAELSVTIKRVNRVLIA
EELDQDHQSKQNNVKPYVELKNITVCVGKEELLSKISFKVDSGLTLVTGKVGSSKSTLLKAILKEYPLSSGEVVCSGTVSYASQHPWCFPSTIKQNI
LFGEKYNEKRYQEVLKVCALEYDLNLFEKGDETILTDNGQNLSKGQQARVNLSRAIYRQADIYLLDDCLSALDAHVHQFIFKKCIKEFLKEKICILV
TQSKSTLTQADAVYVLDKGHIIGSYNPQENTEKEVNNFADTIYKTKEFLIKENGLMESMEVDRFLETEQSNYKNIYEEEMKKGAVDKSVYIKYISYG
GGAALFTFALLLIGSKQAAESYSEKLISVWSDDKQKVLNIKANISNIIGQNITQLSTNLTQAEVQAASTFQMYTVMVILSLVLELLKTFAFLDFCKR
ASINIHKAMIKNVLHSVMAFFDTHFIGNILNRFSQDVINVDENLPFQLLNCLEVAINVGAAATLLMTVNPYFFFYISVTFGVMLLFLKLYLPITRNL
RRLEASTRSPMIGHLNAALEGVTTVRAYKAENMIIDEYERHQDVFTSAHYSLLCFKFGMGFYMSMLAGVMVTLVICSFVLFETDATAGSIGLALTQV
IYLGALIQGAVRCWADLETLMTATERALEYTGLKSETTQGSVPKHWPKDGTISFQNVSLSYNSTQRILNNLSFQVQAKEKIGIVGRTGAGKSSIIST
IFRLYEVDGKITIDGVDIKLLPLKHLRKHLAIIPQDPILFSGTIRSNLDPFGEFQEKDLWEALEKANLKTSITNLDIKVSSYASNFSLGQKQMVCLA
RAILIKSKIVILDEATSTMDHETENLIQETIKHNFSDSTVLTIAHRLQSILECDKVLVLDRGQIKEFGTPKELLENKIGHFSKMVAQGDMSS
>DvvABC-C T48952_c0_seq2+T48952_c0_seq1*
MDSSKRNAKETNPRTKASIFRLLTFFYNYPLFKKARKKGLDDSDIYEIPKWLASEQLGNNLEDAWMRQRKKKGEDASLVRCLISCFGMQYLILGLIQ
LVVKTVLVFIQPRALSKVVAYYAPNQTDVTTKDLYIYASLVVGLNVFSVIYNHNYQQFTTEVGIRVRTSVAALVYRKAVKLGPNAWNHVTVGKIVTL
ITKDVFAFEMALIFVNDMWIGVIQTIIITGVIFNRIQWSVFGGIGFYMLTIPLQLFVGKVVSSKRVQSAKRTDERLQLTTETIRNIKTIKMYTWENF
FTTKLNELRKLELQNLSPVFYLKSLVLIVGSTATSLSFFFMIMTYIWSGHFTDAETVYFIQTCYQSLKSFITVSIPIGIAQCSDLRASLKRLSHFLK
LEEVVDRRSQTISPRVYMRHVSAKVGEKTILNDISFSAEKGVNLITGNIGSGKSSIIKVLLGEYPVSSGQMAIDGTISYASEEPWLFPSTLRQNILF
GEPYNEKRYNEVLKLCALNIDLKKLPKGDKTIVGDRGVNFSKGQQSRINLARAIYRISDIYLIDECLAGLDSKVNYYIFRNCIMDFLKDKVVIMVSN
NINHIKLLYGNNTLVVEDGRTLSLEKQKETLDKRITYFIDDVEMNYFDDDIDVTDEDILEDEANERTQLIGGDKNENKNLYDEEKKSGTVSLKVYLR
YYKYAGGIIMLVILAIIFIAAQAALSYSEKLVSKWVNLEPNITNLTLSNQTDTEEYINIINKRDNYLVMYTFLTVVMVVLTFTRIYMNFFFAIRASR
NLHRKMLKSVINAFMSFFDKHFVGNVVNRFSKDLMTMDEVIPLNTYEIFRQTLGLLGILYLIISVNKLFIIPSIFLFVKLYFIQKFYLPTGRGIKRL
DAATRSPMIGYLNATLEGITAVRAFEKQPLLISEFDQHLDHYTSTSFMMTCAIRFYGFIMDMVSTTFFAAIVIKFVCFKTDAQAGDVGLAITQAMLL
TGFLQYVIRQYTEIENNMTGIERVLEYTDIESEDKLQGRTLKNWPSFGEIKYENVSLVYNSSEQRVLKNISFTIKSKEKIGIVGRTGAGKSSIISVL
FRLYDIEGRILIDNEDIKNLCLKYLRSSIGIIPQDPILFSGSIRANMDPSGTHTDAEIWSAIEKVHLKNLFTCLDDVIVENGSRYSSGQRQLICLAR
ALVSKNKIIVLDEATSNMDPDTCSLLQRTIKTHFSQCTVlTIAHKLNTVLDCDRIMVVDHGEIIEFDTPQALRTKQGGIFNKMIEHS
>DvvABC-C T17573_c0_seq1
MDHCEKVQRPPHPRENANIFSLLTFGYTGKLFTRGFKQDLEDDDLYDVIKKCRSKKCTDKLEHQYNSRSKTKDPNKKVSIFKIIWHMYGLRYILLGL
FHMSGRLLTSTLEPDALSHLVGYFKPGQTKMTIHDAMYFAGIMIGVKAFHTIYFQNYHIYLTELALQIRISFSSLIYRKALKLSPKALEDTSLGNIV
TVINKDVQQFEHSIWMFNDLWISVLQTFVMCYLIYQRTGVASVVAILMLVLVIPVQGYVAKIIKNLRLKMSRRTDERLQRTQESLSTIKTIKMYTWE
QVFADKIGEARFKELKIKLLKSAYAKISLMIMSSLVGKFAFYAMLMIYLYIHEDMSAEDIFYIMRIFGTLRFTMAMAFSMGFTRIGELSASLKRINRI
LELEELPDVIDKPDDDPQIDLRNVSVNLRNKDILKNVDLKLEKGLNVLTGQLGCGKSSLIKVVLRDYPILDGGEVRTRGRKSYASQDPWLFPSSIKQ
NILFGEKYDFKKYQQVVAACALEYDFKILEKGDDTIVADRGMNLSKGQQARINLARAIYRDSDIYLIDDALTALDTRVQEQIFTQCIQGLLKDKCVV
LVTHNAKHIHAADKLVILHDGAIKYIGDQANATEDILLEALEDEEIEEVVMETEESKVIDEKTELLEKPQLRKRQVYHENKKQGSVDFDLYIQYIKM
GGGFIFAILLLFTFAGATLTESTSQKMLTNWINEKSTIQGLKEKHLKNTSINFEDIDLTNQTAIYYNSTYNISAEVIRNIGRLEVKATKSLNLYTIL
VIGYSFVELLKRYIILRVGFRASVNLHKKMVTSIVHSTMAFFDSFFIGNILNRFSQDLSIVDEHLSMMMSHMVDALFHLVGVVGLIATINWKFIIPA
VVLAVFSLLLRSIYIRTSRSLKRLEAATRSPLVGHLNSTMEGLTTIRAYKAQDILKNEFDRHQDLFSSAFYTSICAKAAFSFLMEISSIAFTTTVIV
RFLFFDTGTNSGDVGLTLNQAGMLSGIVHMGLAAWSELENSMTSVERAMEYTTIEGESNTGTDNIKWPTNGEIVYQDVSMTYTNSHEKVLKDISFTV
KAGTKIGIVGRTGAGKSSIISTLYRLYNYQGKIFVDGVELKQLSLKFLRQHISIIPQDPIMFSGTVRSNIDPLKEFSDEEIWKTLHKVQLDSVVPKL
EVDVDDVNFSTGGQRQLICLARAIIRKNKIMVLDEATANMDPETEQVAQKIIEENFSSCTVLIIAHRLDAILDCDKILVLDKGNVIEFDSPKVLLDNK
TSLFSEMMRNSHLGDGEEKQKSN
>DvvABC-C T51687_c0_seq1
MANLTELTTLPTTVTELPKLTSNSSQQNATIEEDAYAALSSFCGSPFWDTNLTWNTNDPEFTPCFEKTVLVWIPCIFLWVFAGLEVFYIFNSKRRNI
PWNWLNLSKLAITATLIVLTISDIVNSFKTAGNRGSEVNDVDIYSPLIKILTFALTGVFIFYNRKHGLQTSGLLFLFWFFVVLCAIPQFRTEIRRSQ
RNEILPDYYYFYTSYLIYFPLVVLMFLLNCFADKAPLERKYPQSQNPSPEESASFLSRLLFAWFDPLAWKGFRKPLETKDLWDINVEDSSRELVPVF
EKYWSQTLKKAESTSPTQVAHAKFKSDSASVDFVNNKKKKQASILPALFKCFGPIYFFGALLKLIQDLLTFVSPQILGFLISYTRFNQEMWKGYLYA

VLLFITATIQTLFLAQYFNRMFVVGMRIRTVLVSTIYRKALKISNSARKESTVGEIVNLMAVDAQKFMDLIGYLNMIWSAPLQICLSLYFLWKELGP
SVLAGLAVMIILIPVNGFIANKIKKLQVKQMKNKDERVKLMNEILNGIKVLKLYAWEHSFEDQVLKIRNKEIKVLKQAAYLNAGTSFIWSCAPFLVS
LVTFATYVLVDEKHVLDANKAYVSISLFNIIRFPLSMLPMMLSNLVQTWVSVKRINKFMNAEELDPDNVLHEPSEEPLTIENGTFSWGEGPILKDIN
INIKKKTLTAIVGTVGSGKSSLISAFLGEMDKLSGRVNSYGKVAYVSQQAWIQNATLRDNILFGKPYDKTLYDAVVEACALNPDLEMLPAGDQTEIG
EKGINLSGGQKQRVSVARAVYSNADIYFLDDPLSAVDSHVGKHIFEKVIGPHGLLKSKTKILVTHGITYLPQTEKIIVLKDGRISETGTYQELLDRK
GAFAEFLLQHITEEAETEAELDELKDQLADTPLSQEVARQLVRHRSRVSESQSETGSDHIGNGSIQRLNSVDKSNHKLSIDDGKKGPKKGEKLIETE
KAETGNVSWAVYKHYLKSIGLVFMLATLIFNLVYQGFSVGSNVWLGLWSDDQKIVVNDTVDTARRDMYLGVYGALGLGQVVTILFASLALYVGSLNS
ARALHNLILGNVLKAPCTTFFDVTPLGRILNRFSKDIDTLDNVLPMTMRGWITCFYSVVGTLAVISYSTPVFILVIIPLCFIYYFMQRFYVATSRQV
KRLESVSRSPIYSHFGESVSGAHAIRAYNQQERFIVESENKVDMNQICYYPGIISNRWLAVRLEMIGNLIIFFAALFAVLKKDQAPGLVGLSVTYSL
QITQTLNWLVRMTSDVETNIVAVERIKEYGEAPQEAAWEIPSKQTSPAWPEEGTVQFKNYSVRYRPGLDLVLHDLDFEIKGGEKVGIVGRTGAGKSS
LTLALFRIIEAAQGNILIDGIRIDELGLHTLRSRLTIIPQDAVLFSGTLRINLDPFNKHSDEDVWSLEHAHLKEFVKGLPAGLNHEVTEGGENLSVG
QRQLICLARALLRKTKILILDEATAAVDLETDDLIQKTIRTEFKECTVLTIAHRLNTIMDSDRVIVLDKGRIAEFDSPARLLANENSIFASMSRDAG
LA
>DvvABC-C S21020Locus_43217_0+T47823_c0_seq1*
MEEVSFLEQKKRKRHPIETTNCVSKLFFCWLPVYLFRGLKNEVTEDDMYVTVEKQKSLYLGTKLEQAWKKQLKKKRPSLLLAIGSVFKWELAIYAVF
NAWFELLRIAQPFLISKLVSYFQDGAKSDNIVVCAITLIIVTFAQVVSIHHYQLKVMVLGMKIRVAACALIYRKALKLSKTALSQTTVGQMVNLLSN
DVGRFDYSGQHIHYLWLSPCVGIACAILLYKETGITGLSGSMFLLCFVPAQVYMAKLTSQFRLKTALKTDERIRLMNEIISGIQVIKMYTWEKPFAK
LIEFIRKKEIDQIRHTSIIKALTITFNIALSRAAILVCILTYILSGNVLTASYAFTVTSYYTYMRAIITLQFPQAMTQFAETLVSISRIQKFLLFEE
LDVKYTPMHKQINNNEILQNNKPDVIEVRPVGIKIKNAAVKWIKKHPENTIEQINFEARSNQLVALVGPVGAGKSTLLQVILKELKPLEGTVKVTGT
VSFASQEPWVFASSVRQNILFGETYNENKYYEVLRVCALEKDLKLFPHGDRTLIGERGTSLSGGQRARINLARAIYKEADIYLLDDPLSAVDTQVGK
QIFNRCICNYLRKKCVVLVTHQLQYLYRAKFIYLIDDGRIRASGTFENLKNGDNAFTKLLATATEMDQLEHDRKLSKSESISSINYEEEDYISIDQQ
KETIGSGGVSWRVYGNYLKAGGNLFKTLILALSFVGTQVLVSLSDIFLTLWVNIEQWRRKQAVFINETSNSSNLDTILKYKDQSTESYWPHVMDQIS
PLHIYSFLVIATVILAVARSLASFSYFLTASTNVHNAMFRKIISSPMLFFNSNPSGRILNRFSKDIGILDELLPLATADSTWVGLTVVATTIVISLL
NPWILIPTAVIVCIFYKIKQIFLVSSRNIKRIEAVTRSPIFTHLAASLQGLPTIRAFGAEQILTQEFDNFQDAYTSSYFMFLTASRGFGFWLDLHCA
FYISLVVISIVFSQSDSLGGNVGLSLTQAITLSGMFQWVMREWSELENQMTSVERTQEYTDLPIESDNERKQPPADWPSSGSLTFKNMSLRYSLEAP
CVLKNLNFEIKPREKIGIVGRTGAGKSSLIQALFRLAINEGSLIIDNVDINTVELKVLRSKISIIPQEPVLFSGTLRKNLDPFDDYSDKVLWAALDE
VELKHAVEEMRHNLYGKMAEGGSNFGVGQRQLLCLARAIIRNNKILVLDEATANVDPMTDAIIQKTIREKFAECTVLTIAHRLNTVMDSNKVLVMNA
GEAVEFDHPYTLLQNKKTIFYGLVRNTGKNMAAHL
>DvvABC-C C222633_3.0+T22098_c0_seq1*
MDHCRKEQKKTNPRETANIFSLLTFAYVGSLFKKAFKLDLHEDDIYEVLNQHRAKKCGDLLEREWLSEKNRSKKPSVSRVMWRCYGKRYLFRGFIDF
VFNTINSIIRPYVIAHFVAYFAPGQTVLTRNDAYFYGCGILVIHFVALFYNHNYMIWIQMFAIEMRTAFSSLLYRKALKLSSAGTTDTNLGNIVTLI
TRDVQAFVTVIFSVNDAWVSITQVLVICYLLFNKVGVVSFVGIGILISVFPFQVYISTWVTKFRLQTCQKSDERLQLSQEILSTIRIIKMYTWELFF
TQKVTKARKEEMRKLVLGFYCRRVIIITGVMFLNFGFYMLIMACIWAGVSTDTTIIFYILSNFRDLRHFLAIVMPMGMGNVSEFVSSYKRLVKALKS
EELVEEITQEDPKSPPFIKLKEATVKVRGHEILKGVSFQISPGLTVITGPIGCGKSTLFKTILKDIPLESGSVLTQGRISYASQDPWLFPSTIRQNI
LFGEPYQAKRYNEVITACALTFDLSLFEHGDETILTDRGLNLSKGQQARINLARAIYRSSDIYLLDDCLTALDASVQVYIFNECIKKFLKNKSCVLV
SQNPSHIQQADYVIILEDGTVKDVGKPSEQIIREAKEVVDHGVDLQKQEKQQIKEDNSADFDNEVTDDIKLLEKDEKSTKKIYSEVKKQGSVDWRVY
KQYFIYGGGFLFVLFNISMRGSTQFAVSYAERLLTKWVDKKQIVLNIEKNISTNPLSNDTLFPSLSLAKSIEQDTFRLYSVFVMLGTGMELLTTYAL
LEFCRRASVNLHKAMIKKIINSVMMFFDTHLIGNILNRLSQDMLSVDEYLSAVLDQCCRTIFNVGGIVFLVTLVNPSGAFIIALIFALAFLMRKMYL
PAGRSLKRLEAASRSPMIGHLNATLEGVTTVRASKVQGILIEEYNRHLDLYTSASFTSMCSARAFGFFIDLICSLFIVFVVIKFLFFNKDNTAGDIG
LTITQVSALSISITWCVRQFTELENYMTSLERILEYTNIKTENKEGAEINNWPSLGSISYQNVYLSYEGNTHYVLKDLNFEIKPTEKIGIVGRTGAG
KSSIISTLFRLYEVQGKILIDGVDIKTLSLDFLRKHIVVIPQDPTIFSGTIRSNLDPLNEFEDKDLWNALQTVGINDSITNLEQPTNSNLLTFSSGQ
KQLLCLARALLRKNKIVVMDEATANMDHETDKLLHKIIKENFFDCTILTIAHRLHSVLGCDKVMVLDRG
>DvvABC-C D18126+D19164
MDIGFKLKHENPKKQAGLVKKLFFGWLVKLVKQGTKKQLEISDLYEPLDKDNSKTLGDCLERHWKNEILKSQIKKTSPSLLKAIVKAFYFEFLLYGI
AWFVLNVLLRCSQPIILFHFIALFSGENREENQGDMYIYGGLLILVSVLSIFFMHHLQIGLASIGMRVRVACSSLVYRKITKLSHQTLGQTAVGQVV
NLMSNDVHRFDLVLLPLHAFWAIPFQFVILSYFIWQQVQIASLAGLVSMVIISLPLQGYLGKLMGTLRANIAKKTDNRVKLMSEIISGIQVIKMYAW
EKPFEKVIEIARKNEIRCVTLTSYLRGIFASCMVFLERMSLCFTLICYVLLGNNITAEKVFSLAQAFNILQLSMAIWYPLAVSHGAEALISIKRLKA
FLTLEEKEVSRIKGLSTPGVVMSNVSSSWCDAGETLQDISLNIPPGFLCVVIGPLGAGKSSLLQLLLGELAIKTGTVLMGGEISYCSQEPWLFQSTI
RNNILFGRPFDKQLYEKVVKVCALERDFQQFPERDETVVGERGVSLSGGQRARINLARAIYRQADVYLLDDPLSAVDTHVGKHLFNQCIVKHLRGKT
RILVTHQLQYLKKANLIVVLNEGKIETVGTFEQLSRSKLDFAKIIVDSVEPGDKHEETTESSDLTNTVSNPRKASVTSTKSDLSESLEYFEENNFSD
EVIEHDPQHTAWKEYFQATRRTGLLIVVFTMLVLAQTICSGTDLWVAFWTNQEVIRHSTSRPLESDVSSDEPILYEATPFDSSLNLTHNYEYKVEEP
VVTSDGIFDYVYINNHVYHLVKTGYAIGFYGFLIVGVIILTLFRSMMFVKVCMIASVNIHSKMFSTLLRAPMRFFDTNASGRILNRFSKDMGSIDEI
LPRVLLESVQIFLVLIGILVNVSISSPYAIIAMLLLGICFLKLRSCYLSIAMSLKHIEGKVKSPMFSHVNSSLHGMATIRASNTEDILIKEFDEHQN
VHTSAWYLTIVCISAFGLWMDIVCIIFLTCVILTFIFLQNFFKVNSSLVGLAISQSMTLTGMLQYGMKQTAEVINQLTAVERVLQYTHINTEGPFES
PIETRPVEPWPKFGRVDYNHVYLKYSEDDPPVLRNVQFTILPGQKVGIVGRTGAGKSSLVAALFRLVNFQGTISIDGIDTKRMGLTYLRKKISIIPQ
EPVLFSATLRNNLDPFEEFTDEQIWKALEQVELKDISHSLDLMICEGGSNFSLGQRQLICLARAVLRNNRILVLDEATANVDHRTDSLIQTTIRNKF
KDYTVITIAHRLNTIMDYDKVIVMSYGRVVEFGHPHQLLQLPDGHFHRMVLETGPVMSLQLKDVAMVAYSNLLKQEAASIR
>DvvABC-C T49513_c0_seq6
MDSTEVQNNPNPRETAGILSIMFFGFTIPTFRTGMKKQIGVEDVYNTLRSDRSEYLGDRLEKNWNRELELVKQNKKEKPSLFKAVVRTFWLEYLLIG
IIAAVSDIGIRSFQPSMLGGVLDYYTPGSKKSKLEAYFYGSGMVIMSLTSSILMNQYMTNAAHAGMKVRSSVCALIYRKATRLSRTALGDTPAGKVV
NLLSNDVARFDFASMMIHQLWIGPVSTIPVIYFLYQRAQLPGIAGILVILIMSPLQATSAKLATKFRKMTALRTDERIRLMDEVILGIQVIKMYAWE
IPFQKVIGLARKAEIKIIKKNSYIRALFMTLGMVTTKLALFATLVTITLTGGEITAAKVFVLISYFNMLSMTLTGILTRGIAEIAEMFTSIARLQKF
LENDEYKYDQNKISSDPNENSKSMLSLKNIYAKWDASLSDCVLNDITLELNKGKLLGVIGPVGAGKSSLLQTLLGELDIDKGSINANGTYSYASQEP
WIFSSTIRQNILFGADYDRKRYKEVIRVCALERDLEQFKDRDMTVVGDKGASLSGGQKARINLARAVYREKDIYLLDDPLSAVDIHVSKVLYDECIN
GFLKNKTRVLVTHQVHYLKNASHIIILNNGRIEGEGSFNDLANSDNLYAQLLTKEPEITEEEKVKQIEFARQVSRQSRKSRKSSMTSIISEKTIVEG
LMEDSDEDSEDEENTKKLEMRDMQEESSKGKVLGSLLLRYFRTGANCFNVFLILSLYVLVQVSSSMVDYFLSIWTNIEDERTSASILLESNSSSTLL
EDISTTIASNNINSTVVLADFPSTIPSTTLSDAALNTLNTTVKLEPQAWIQFPAWTTEECLYIYGILISISILMTLGRSFSFYSLVMSCSRNLHQIL
FKGVIYAKMRFFDTNPSGRVLNRFAKDIGSVDESIPRVLLDGGQMILQIFGSLILVGVANPYAIAMALILAIAAFFLRNIYLKTSKNLKRVEGILKS
PVFTHLSATLQGLPTIRALRTEKILQAEFDKIQNYHISAGFMNFTCGAAFGISLSMCSNIFIAVLIFSILTFDKAMGLTGGTVGLAVSQASTLTGMI
QMAMRMTADISNQMMSVERVLEYEQLPPEPQPKNPKPLPKNWPSSGKIDFQNVKLQYFEGGPTIIKDLSLNVKATEKVGIVGRTGAGKSSLIAALFR
LANVEGKIAIDDIDTRDIYLRDLRSKISIIPQDPVLFSGTLRYNLDPFEEYSDEQLYKVLDEVELKDPNNIINRLENRVMDRGSNYSVGQRQLICLA

RAIIRNNRILMLDEATANVDPQTDALIQKTIRTKFANCTVLTVAHRLNTIMDSDKVLVMDQGTAAEFDHPHLLLQNSNGIFSKMVEDTGPGTCEQLK
TVAKNTFDKNLLIEE
>DvvABC-C D14070
MDVVFKLKTENPKRRANFLSKLLFAWLVPIIKKGNKQTLEVVDLYEALDADKSKRLGDALEKHWDRQVLLAKEKSKKPSLMWALWKTFRLELIMYGI
LWCIQNVVLTSLKPLLIAQLVALFINYSDDVVTDIYIYSAAFIGVNVLIVFIYHHTAFGLQALGMKIRVAISSLVYRKITKLNQKALGQTAAGQIAN
LLSNDVMRFDLVVVPLHALWVMPIQVAVLLFIMWQQVGISSVAGVVAMAVVSLPVQGYLARLQGKFRERISKKTDKRVKVMSEVVSGIQVIKMYAWE
KPFEAVIKLARGTEVKDITAASYLRGVYSSCMVFLDRMALFFTIVCYILLGNLITADIVFSLAQTFNILQMAMAIYWPWAVSQGAEALVSVKRVQEF
LVMEEKEESKIDDLGKSGVILSKVNASWTPTTSTLQDISLQIPSGTLCAVVGPVGAGKSSLLQLLLGELPTKSGKVSIGGEISYSSQEPWLFQSSVR
KNILFGKPYEKTWYDKVVKVCALERDFEQFPHKDRTIVGEKGVSLSGGQRARINLARAIYRNADVYLLDDPLSAVDTHVGKHLFENCIVNHLRGKTR
ILVTHQIQYLKKANLIVVVNDGKIEAQGTFNELMEYNLDFTKLLVAADETDDKEEDADSEGKLGKSGRASARSRRKSSAVSALSDFSESMYDDPEVI
EEESGNTNIGTTFKQYILSTKSLCFVTFTTILLVIAQGINVLVDLWVTYWTSQEQIRHSNGTVISLQSADINEVELYPINQTTKIDATHRLVYNIEN
NGSQILKQISFDDIFDKIDLHGQIQKIIKTDYAIYIYSALMALAIILTIIRSFLFYKCCMLSSANLHKSIFHTLLKAPMRFFDTNPSGRILNRFSKD
MGAIDEVLPKVLLDSCQIFVVMAGILVSITVSNYYILILLVPMGYIFLKFKDWYVATAKVLKHIEGITKSPVFSYVNSTLNGITTIRAANAQNILKD
EFDENQDAHTSAWYLTIGCMSSFGLWLDIVSIIFMACVIGCFIILTKFTNVDGSLVGLAVSQGSVLIGLSQYGVRQLADFINQLTSVERTLQYTKIE
EGPFETPKEKIPKGVWPKEGLLEFRNLSLIYVVNDPPVLNNLNFVIKPGQKVGIVGRTGAGKSSLIAALFRL
APTKGSIFVDGVDTKNLGLTDLRKKISIIPQEPVLFSATLRYNLDPFNEFDDEKIWKALEQVDLRESIDSLDFQVAEGGGNFSLGQRQLVCLARAVL
KNNKILVLDEATANVDPRTDALIQATIRKRFKDCTVLTIAHRLNTIMDSDKVLVMSFGSMIEFDHPHNLLEIPDGHFHRMVLETGAGMTQQLKSIAK
AAYNNKIE
>DvvABC-C D22628
MDSTKKHVKVSPEESASPFSVLFYWWILPFLKYGYQNDIGMKDVYNTTQADQSGPLGDELQENWEREILSYIDKQKNKPSLKNAIFRTFWKSFILSG
AAIFVQFIIIKTLQPVVLAKYINFFDTNNKPYLGWIWGCGVVLLALANVVLYHSTMLATQRIGMRIRTAVSSLTYRKLLKLNHKSLGETAAGQLVNL
MSNDVQRFDVCAASIHFIWIMPIYAVLTFYILYIYVGIIAAVTGMAFISLESIPLQGTISRWQGVLRYKIALRTDKRIKLMSELTSGIQVIKMYAWE
KPFEKIVEMSRKYEIDVIAKTSYLYGILSATSVFTERLILYVTLIPFVLLGHRLTGGIAFSLANLFNNIQLVMAINFPRALSSYNEANVSIARLEKF
LLLEEVEEETVVNEKHGDHVGCINLNNVTASWSPKSIVPTLIDIDLHLRCGTLCCVVGNVGSGKSSLLQLFLRELPITSGYMNIAGKISYASQEPWL
FVSNVKDNILFGKSFNKKRYQDVIKVCSLERDLKQLPYGDKTLVGERGKSLSGGQKARINLARAVYTEADIYLFDDPLSAVDTKVGKHLFDECITKY
LNGKTRILVTHQLQYMKKSDLIVIINNGKIDKVAKFHDLTEHELNLLQQTPEIDDKDKEKMPTILETKIPKVSSTATLQSASSLASSIPTEEPGETG
ELIEKGDLSTSLYWEYFRSGTGIGFLLFTGFMFIFSQIITNASDLWLSHWTNVEARRYVSTLNLSTEFISTTISSINNHTVTDMLKSVAFTNTNPQE
IITTADPSVSASKLSSNMSSLELENFKSEGYYIWVYSALIIAVVILQIWRCFLYYQVCMSSSKALHNKMFHNILQAPMRFFDTNPSGRILNRFSKDM
GAVDELLPSCQIDAIQILMVFVGILVMVVIVNPWMILTTIIIGPVLFILRKMYLKTAQSVKRLEGISKAPVFSHISASLFGITTIRASNAEKMVTTE
FDILQDQHSSTWFLFLVSGRVFGFYLDVICCIFLAIVTIQFLLFRDENTLSGNVGLAISHSYILTGMVQMGIRQSVEVASHMISVERILQYTKLEKD
GVFESLPAKKPPRDWPNKGKIIFKNTFLRYALNMTPSLRDLSIDIKSGEKVGIVGRTGAGKSTLIASLFRLAPVDGEIIIDDIETGGIGLHDLRTNI
SIIPQDPVLFSASVRYNLDPFEKHSDEILWKALENVELKGVVTDLNQPVSEGGSNFSAGQRQLICLARAIVRNNKILVMDEATANVDPQTDALIQKT
IRERFRDCTVLTVAHRLNTIMDSDRVLVMDAGQAVEFDHAYRLLQNSDGYFAKFLKDAGPPMANKLRDIAKEDYYKKHTSPEVIVDDD
>DvvABC-C D20002
MDIGYKLKAENPKQNANIFSKIFFGWMIPLIRRGTKQNLEINDLYKTLKKDQSKRLTDALEKNWQKQVDKAKKKGGKPSVLMAISRTFAFEFMMYGI
LWAIQNVVLMSLKPILIAQLIELFTDDTSTRFREMYFFSTSLILVSLLIVFFFHHTNFGLQAIGMRIRVATSSLIYRKITRLNQKSLGETATGQIVN
LLSNDVQRFDMVVVPLHALWVMPLQVAILMFIIWNQVGISSLAGVISMAIIALPVQGYMAKLMGMLRQKVSGKTDTRVKLMNEVIGGIQVIKMYAWE
KPFEKVIKQARSSEIGDITKASYLRGVFSSFIVFLDRVALFFTVMTYVLLGNVISADIVFSLAQTFNILQTAMAIWYPVAISVGAEALVSAKRIQNF
LIMEEREEASIEKIDKPGIILSNVFASWTTKGRTLQDISFQVPPGTLCAVVGPVGAGKSSLLQLLLGELPLKHGRVQLGGEVSYSSQEPWLFQSTVR
NNILFGSPYEKRWYEKVVKVCALERDMEQFPQGDKTIVGEKGVSLSGGQRARINLARAIYRQADVYLMDDPLSAVDTHVGRHLFDQCILHHLRGKTR
ILVTHQLQYLKKAGLIVVLNDGKIEAQGTFEELMDSKMDFTKLLVAADETGEKHDKQEDADPEPVDFTRKFSSTRRFSVLSDASGEISLSMRSVDMT
SDKNGEEEESAGDGKPFKDYLFATKNICFVIFVCVLMIMAQAFVVGVDLWLTFWTSQEAIRHENGTIVESTSPTVEIIPLHGDNFSYSYNYNYSYIP
ENNTNSFNNTKSFNINDIFDTVNVDGQLKKIIKTNWALYFYSGLIGLAIVFTLTRSLLFFKGCMMASVNLHSSMFHMLLKAPMRFFDTNPSGRILNR
FSKDMGAIDELLPMGFLDTMQIMLALCGILVNITVSNAYIVIAIAILAAVFLKFRSWYISSARVLKHLEGITKSPVFSHINATLNGIITIRASNAQD
VLIEEFDENQDANTSAWYLTIACMNSFGLWLDFLAIIFLAIVTFCFVILRKFTDVNGSLVGLAVSQCSALTGMLQFGMRQTAEIINQLTSVERVMQY
TKLDTEGPFDTPEENRPRGVWPKRGQIEFRNLSLKYVENDPPVLRNLNFMITPGQKIGIVGRTGAGKSSLISALFRLAPLEGAIYIDGVNTKNLGLT
DLRRKVSIIPQEPVLFSASLRYNLDPFNEFDDDKIWDALEQVELRDSVDSLDFHVAEGGGNFSLGQRQLVCLARAVLKNNKVLVLDEATANVDPRTD
ALIQATIRKRFKDCTVLTIAHRLNTIMDSDKVLVMSFGNMIEFDHPHKLLQIPDGHFHKMLLETGPVMSAQLKDVAMRAYQQE
>DvvABC-C D7536
MDSTKKYVKVSPEQKANLISKLFYWWFLPFFKFGYKHDITIKDVYNTTQGDLSAALGDALQKNWEDEIRNHEDKKQKKPSLKNAILKTFWKSYSRSG
IALGFQFVVIRMLQPIVLAEFINYFDSNQEAYIGWWLATGVIGMAFLNVIITHSCTLDTQRVGMRVRIAVCSLIYRKLLKLSHNSLGQTASGQLVNL
LSNDVQRFDLAAQFIHYAWIMPLTAGISFYILYRYVGIIAAVTGMVFITLESLPLQGSFSKWQGKLRYKIALKTDKRVKLMSEITSGIQVIKMYAWE
KPFEKVVELSRKYEIDIITTTSYCYGVLSAMGIFTERMILYITVITFVLVGQRLTGDVVFSLAQLFNTVQLIMAIFFPRALSFYSEAKVSITRLEDF
LLLDENEKIPEPEKISNPDELGEIQLTNISASWASKPIAPTLMDLNLHIQPGTLCCVVGNVGCGKSSLLQLLLRELPASRGQMKINGKISYASQEPW
LFVSNVKENILFGKPFLKNRYHDVVKVCSLERDFKQFPFGDRSLVGERGTSLSGGQRARINLARAVYTEADIYLFDDPLSAVDTKVARHLFDECISK
YLYGKTRILVTHQLQFMKKADLIIIINNGQIEKIAKFNELSENDLNALQQEPEADDKEKEKVPETGDKIRKNSTVPHFQSMSSLASSVFSDDPNEED
ELIEKGAISNATYVEYWKSGGGVLFLFFTIFIFLVAQMITNASDLWLTHWTNNEAKRYHLSDSLHNATSNKTLERPLLMSTGLSDEMSTSATLLTDS
TPSADSSLLSLSQKRYLTDhIRNATLEELNDIPSTEYYIYVYTGIILASVVFLTWRSFLYYQICMTASKVLHNKMFNNVLQAPMRFFDTNPSGRILN
RFSKDMGAVDELLPRCQIDAIQIFMVMIGILAMVFIVTPWMIIPAVILAPLYYFFRVVYLTSAQSLKRLEGVSRAPVFSHISASLYGITTIRASNAE
KMITTEFDILQDQHTSTWYLFIVSSTAFGFYLDVMSCFFLAIVTYQFLLFRTENTLSANVGLVISQSLILTGMVQYGVRQSAEVASNMISVERVLQY
TKLDKEGPFETLPGKKPPRNWPEKGRIIFKNTYLKYAPELPPVLKDLNIEVNSGEKIGIVGRTGAGKSTLIASLFRLAPIEGTISIDDVDTAEIGLN
DLRLNISIIPQEPILFSASLRYNLDPFEKHGDEVLWKALEDVELKGAISDLNQVVSEGGSNFSAGQRQLICLARAIIRNNKVLVMDEATANVDPQTD
GLIQKTIRERFQDCTVLTIAHRLNTIMDSDRVLVMDAGQAMEFDHPYQLLQNPEGYFSKMVQETGPAMAELLSNVAKNDYIKKKGtLTSEVPQNLAI
EDNKD
>DvvABC-C T47333_c0_seq11
MESSNSVPGAQTETEEPILRRSTRVRKPVTRLNLWFLPFFKYGYKNDVELKDIYNATKPDMSESLGNQLQKNWEEQIKKCDQSQNKKKPSLKSAIVK
TYLKMYTASGVMIFLQFIVIRMLQPIVLAEYINFFDKKHGIKPNEFGWLLASGVVIMAFLNVILLHNTFLNTQRVGMRVRIAVGSLLYRKLLKLNHT
SLGKTAAGQLVNLLSNDLQRFDYAAVFLHYIWVMPINSAISFYIMYRTCNLVAALTGMGAIIVEALPLQSYLSRLQGQLRGKIAVRTDTRIKLMSEI
TTGIQVIKMYAWEKPFEKVVELARRLEIDSITKSSYIHGMLSAMSVFTERFALFIVVITFVLMEGRISGDIVFSMAQLVNTVQMIFAMFFPRALSTY
AEAKVSIKRIEEFLLLEETTPKKKLGENDTLAGKPGEIELFNVTASWSENPITPTLVDLNLHLTPGTLCCVVGNVGSGKSSTLHLLLKELSPTNGKV
NVYGNVAYASQEPWLFVSNVRENILFGRPYDKKRYQEVIRVCSLERDLKELPHGDKTLVGERGAALSGGQKARVNLARAVYSDADIYLFDDPLSAVD

TKVGRHLFDECMVKYLKNKTRILVTHQLQYMKKADIIVILNNGRIEKAANFSALSQDELNVLQQEPEPEEDENEVKKERSKVERSQSTISEKTLRSE
HYQSNTSLSVSEEPEDKSNERGEEEVEQGTVPFSIFVEYFRGAGVFFLTFTAVAFIAGQAFTSITDLWLTHWTNVEGKYYADMLAANVENTSETNRN
NLTNTLTTLLPLTTSADPVLLTLPTLPTNISVMESMLTISSVIDTTTVRIEAEQVYDQNRYIIIYTILMFSSMILLTLRSFLYFIVCMSSSRTLHNK
MFSNILQAPMRFFDTNPSGRILNRFSKDMGAIDEMLPKTGIDTIQIFTVLVGVLVLVFIVTPWMIIPTILIGFLFYYCRNIYLASAQSIKRWEGVAK
APVFSHVSATLQGISTIRASNSEEMVIKEFDVLQDQHTGTYVCFLVSSTAFGLYLDIISTLFLGLLTYQFLIFSVDSTAGGNVGLVLSQSLILTGML
QYGVRQTAEAISNMISVERVLQYTKLEKEGPFESLTKPPRDWPEKGKIIFKNTFLRYSLDVPPALKNLNIEVNSGEKVGIVGRTGAGKSTLIASLFR
LAPIEGTISIDNLNTAEIGLKDLRLNVSIIPQEPVLFSESVRYNLDPFEKYSDEILWKAIESVELKHAITDLNQTVSEGGSNFSAGQRQLLCLARAI
VRNNKVLVMDEATANVDPGTDALIQKTIRERFRDCTVLTIAHRLNTVMDSDRVLVMDAGEAVEFDHPFKLLQNPDGYFSRMLRETGPAMEEMLKAVA
EEDYNKKYGGIKGKKGE
>DvvABC-C T49618_c0_seq2*
INLARAVYKQADVYLLDDPLSAVDTHVGKEIFEQCITGYLKNKTVILVTHQIQYLQELDNILYLEDGIPKTGSFRELQASGLDFTKSLGESTEEEET
TKESEVMRQLSIKSVASVEVEEPKEVQEQKGTGTIGGYVYRAYFRAGGNCCVIFIFFFLFLLSQMFASASDYYLSYWVNLEQQDSAANSTQSTGLSS
NETITNNFWHFSRETSIYIYSGLIVLLVIITLVRSFTFFAVCMRASTRLHDNMFVSITRATMRFFNTNASGRILNRFSKDMGSIDELLPTALIDCLQ
IGLALLGIIVVVAVVTPFLMIPTVAIGIIFFFFRIFYLRTSRNIKRIEGTTRSPVFSHLNASLLGLTTIRAFQAEEILKKEFDNHQDLHSSAWYAFI
STSRAFGYWLDLLCMVYVTLVTFSFLVIGNEKFGGNVGLAITQAIGLTGMFQWGMRQSTELENQMTSVERVLQYNDIEKEGALESQPDKKPKPSWPE
KGKLEFRDLYLKYFPQDPFVLKDLSFVVKPKEKIGIVGRTGAGKSSLINALFQLTETIGDILIDEVNISQIGLHDLRTKISIIPQEPVLFSGTMRKN
LDPFDEFSDAELWSVLEEVELKEAVENLVAGLNSKMSEGGTNFSVGQRQLVCLARAILRKNKLLVLDEATANVDPQTDALIQQTIRTKFAECTVLTI
AHRLNTVMDSDKVLVMDAGTLKEFDHPHVLLQDKDGIFYGMVQQTGKTMAETLHQLAKYSYEAMSHLHPDSFR
>DvvABC-C T45163_c0_seq1
MTSDVIGRNWNWTSLCGPEGFQIWSQQRHDFGQCFQQLCFQIPVLSLFAVLSSYYFGRQTGFVSRGKLQRYCIITRNVVSLLLMFYPIMHAYNTIMT
KTPNRENASFLVCAVQAISWCSHFLYTVGLRSRLGKSQRGPTAMGIVWCMVFAMTIVSLRSAYLENSQRPTENTKMELGITLYYTVLQIIYALSLIP
GGSTTTLNFPERYTEITERQPLINANAYGRFSEEGDPNDLGVAMEDTNWLSRLSFSWVKSLVNKGVEDKIVTSDDLYDLPDSISSDNNSNRLENYLR
IEHDTDSSREVIVIQPRKSSSLLAALHRCYAWQFYSVGILRLVADCSNFAGPILLNKLVDFIEKRSIDIKWGYLYAVLLVVSTSMSALCDSHFNFKM
SVIGLKMRGAIINTIYRKTLSVRSTVLMSKLSVGEIMNYMSTDTDRIVNSCPSFHAVWSIPFQLFVSLYLLYNQVGLAFLAGVMFSIILIPINKCIA
NKIGDLSTKMMEQKDGRVKLTSEILRGIKAIKLYVWEQHFIRLITRQRDAELKYLKGRKYLDALCVYFWATTPVVISILTFVTYVLMGNKLTAATVF
TSIALLNMLVSPLNAFPWVLNGLTEAWVSLKRIQKLLDLPYMNLNEIYDHKELEDNTDQNTEISISNASFSWEKATPVVQVPKGKGKGKRTLSKRGQ
GESSVQRRETIIFKLRDITLKIKKGEFIGIMGTVGCGKSTLLSSILGELAVESGSLAVRQVETGFGLVTQQPWLQRGTVRDNILFGKPYEEKKYKDI
IFACGLIEDMQILPSGDKTEVGDSGGTLSGGQKARIALARAIYQEKQVYLLDDILSAVDVKVGKHIFNHCIIGLLQGKTRLLCTHHVKYLLHADRIL
LLENGAIKSIGKPVDVLRNIDDTLPVDLELEDSIQSEYTSSECTLENSLCVDNAQEKDKDIELFKEVSERGNLDFTVIISYWKGISHFVSISILISM
TLMQVSRNFTDWWLANGVTSPVTNSTNLTVYLATATDYIDYVEDDDNMSEFLKVYVELACVNTLFTLIRSFIFAYGGILAATKFHKVLLKTVLRARC
TFFDVTPIGRIINRFSSDTYTVDDSLPFILNIFLAQLFGLLGSLFITMYGLPWICVFLIPLVPIYTYLLNQYRITSRELKRISSVTLSPIYNHFNET
LQGLTTIRAMRVIHKFKHDNNQHLEANLKAQFASQAAARWLGLRLQFIGVIIITGVSFIAVIQHQYDIADPGFIGLAISYALGITGSLSGVVNSFTE
TEREMVAVERINQYVESIPMESKYFVMDPPFAWPSQGVISFEHVTLHYRKHLPPSLMNLSFDTRPYEKLGVVGRTGAGKSSIISVLFNLVDITAGVI
TIDAVDIKKISLATLRSRMFCIPQDPFLFSGTLKENLDPLEEFKDDEIWNALGRVNLTDVINSLGGLDYKIDGSGSNFSVGQKQLVCLARAVLHNAK
ILCIDEATANVDEETDRLIQNTIRTAFRSSTVITIAHRIQTIFDSDRVLVMHQGEVVEFDKPDVLMADSSTYFYRLVQEA
>DvvABC-C T43960_c1_seq1+D21270+T49513_c0_seq3*
MEEVKVGGVPEKKLNPKDRSNFISRLFFCYLLPYFVKGYKRELTEDDMYRHRSEHDSSKLGQRLEERWIKHVAKNKNPAFWKVLMGTFFWEIFGQDI
LVILCEAVRMAQPFIISKLLIVYEKDPKENINEVYLYSGLIVATSFISVVLVHRFCLGLFQIGMKMRIASCSLIYRKALRLSKSALAETTIGQMVNL
LSNDVGRFDTAPIHLHHLYIAPIQALIVMVFLYIVVGWTGLLGTFFLLLSIPLQSWLGKKTSQFRLKTATRTDERVRLMNEIISGIQVIKMYTWEYP
FAKLVEYVRGNEIKFIRKTSVIRGILISIAMTLNDIAVAICLVTYVLTGNPLTASYAFTVTSYYRLMGSLTYFLPMAVSQGSEMWISMKRIEKFLQY
DELNIENLKNSLNGKQNGKQDLQELTGSDKGSIHIKKASAKWLKSQYENNLENIDMDVAAGSVAAVVGSVGSGKTTLLHIIMKELELESGSVNVTGK
ISYASQEPWLFGGSIRQNILFGETYDQLKYNEVVRVCALERDFTLFPYGDRTLAGERGVTLSGGQRARINLARAVYKDADIYLLDDPLSAVDTHVGK
HIFEECICEYLGSKCVVLVTHQLQYLKSIQNIYLLKNGQIAVSGSYNNIKDSNTHYSELLHDIEEEEEEIRRKSKVVQTKEEEAEGEAGPVLQREGK
STGIISGRVYRKYAQAGGHIFKAFSILFLFILSQALLSVCQYFVTFWVNIQQWKSSQNILNSTDTGVVVENVTAPLNLTTPPQSTDNTGFADWWLTP
VFVSENTSFYFIGLNMVMVIIMNIRSSSFYAWAITASTKMHNMMFENIVYSPMRFFNINPSGRILNRFSKDIGALDEVLPMTIMDTLGIGLEVVAIC
FVLGSLTYWIMIPTILIGILFYVMKVVFLETSRDVKRVESVNRSPIYTHLTATLQGLTTIRAFKAEGILKNEFDSYQNNHSAAFFMYLASTRTFGFW
LDFICVIYTALVLLSLIFIESETFGGNMGLALTQTLALTGMFQWGMRQWSELENQMTSVERVQEYADLKHEDDNYTVEIDETWPEKGKIEFKNLSLQ
YSPNDPPVLKNLSF
>DvvABC-C T48940_c0_seq1+D21306
MVNKIMEKPKEKFNPLSKVNIFSYLTFGYVLELIKKGWKHDLNESDLYGLPKKCKSKFNGERAVKQWNKNPSLFKLLLSKFGLAYLLFCLAHVTWVE
IRSIVRPYGMSKLIAYFDKKQQTISETQAFYCASLVIFIKFFSVFYTANINVFELKYLLRLKTSLQSLLYEKLLKLSSNSILETSSGNLITVMTKDI
YSIETNLWVFKEFVIFFVQTTTVFYLLYSRIGVTAFVAVGMFAVALPVQGFLCKLVTKLRLTVGKHSDERLQITQEALTAIRVIKMYTWERFFIGKI
SKERSKELSALLKMFFVHFLIVIIGIFLSGIVFLALILTYTALGYDTNTELIFYIDSLFGELTLAMGIMIPINMSRTAELIASLKRINKVFQCQEIS
KIEKIYKDALIDINNVSVEIKNQKVLDDVSLKISKPGLHVITGVVGSGKSSIFKVIIGLHNLSEGYCKLGGTISFAAQEPWLFPSSIKNNILFGEKY
DKARYEEVLKICGLDYDLAFMQNGDETIVAEAGLNLSKGQQARVNLARCIYKESDIYLLDDPLTALDANVQDVIFEQCILKFLKEKICILISQNPKH
IENAENIFTLSHGKLVGDTNIIPAPTPDIGEEQKEPVQVSEAKKQSEAAKSSHEDDQYTNDDLFAEKNNTGKVGADVYKSYMRFGGGICFVVFIILL
YTGSQVAHFTSEKLKAKWVDVQANDGSTMLPILGGPYKKQDLLLFTTFTTVFSSLYNLVTLYSLLRFCRNASYNVHKLMVSRITAAVMTFFDTNYLG
NVVNRFSYDLNILDERLPMMFLHLFRAAFYCLGILVLIATVNWVFLFPSIIFISLVVTLRILYINTGRNLKRLEAATRSPLVGHINATLDGLSTIRA
FDAEKRLTNEFYRHQDLYSSAVFTMRLSHTALIFYMGALSATFSTIIMAKFLFFPDDTSAGNVGLALTNIIRLSDMILWGINEWIEIETNMTSVERC
LEYTKVKQEDDSGKTPKDWPTEGGVIYKQVELKYGDNKIPVLKNINFTIKPKHNIGIVGRTGAGKSSIISTLFRLYEYGGNIEIDGVDIKTISLSCL
RKNIAIIPQDPILFSGTLRENIDPYKEHADDKIWSLLNQLSIENIQTLDSTVANFSAGQKQLICMARAALNNCRIIVLDEITANMDTENDNLIHSLV
EKIFKNCTILTIAHRLNFILSCDMVIVMENGQIVEFDEPKTLLENEKSLFAKICDKASS
>DvvABC-C C217405_3.0+D12550+S19317Locus_19671_1*
MEESKGQTAEQKLNPIDRSSYFSNLLFCWLLPYFVKGYKRELTEDDMYRHRSEHDSSKLGDRLEERWIKHIAKNKNPAYYKVLIGTFIWEIFVLNVL
VVLIEAIRMTQPFIISKLLTIYEKDPKENINDVYLYSGLIIATSLVSVILLHKFNFAMMQVGMKMRIASCSLIYRKALRLSKSALAETTIGQMVNLL
SNDVGRFDQAAHHLHYFYIAPIQALIVMVFLYLFAGWTALLGTIFLLLSIPLQSWLGKKTSQFRLKTATRTDERVRLMNEIISGIQVIKMYTWEYPF
AKLVELVRGNEVKFIRKTSVIRAILMSLAMTLNDVAVAICLVTFVLTGNPLTASYAFTVTSYYRLMGTLTYFFPQAVSQGSEMYISMKRIQKFLQYD
ELKTDAYFIPENSLNGNKNGKLAMKELPSSDKGSIHIRNASAKWLKSQPENNLEDIDMDVTPGNLAAVVGSVGSGKTTLLHIIMKELELQRGSVTVS
GKISYASQEPWLFGGSVRQNILFGEKYDQVKYNEVVRVCALQRDFTLFPHGDRTLAGERGVTLSGGQRARINLARAVYKEADIYLLDDPLSAVDTHV
GKQIFEECICGYLGSKCVVLVTHQLQYLKKIPNIYLLKNGQVAVSGSYDDIRNSNTDYSKVLANIEEEEEEIRRKSRIVQAKQEKDEEEEEVQVLQR
EAKSTGNISGRVYLKYAQAGGHLFKALCLIILFVMGQALLSMCQYFVTFWVNIQQWKALQNIASPTAPASLENMTAPLDITSPSPQPVDHTGFADWW

LTPLFISENTVIYFSALNILMVVIMNIRSSSFYAWSITASTKMHNTMFENIVYSPMRFFTINPSGRILNRFSKDIGALDEVLPMTFMDTVQIGLEVA
AICFVIGSLTYWIMIPTICIGILFYVMRVIFLETSRDVKRVESVTRSPIYTHLTASLQGLTTIRAFKAEDILKTEFDKYQNTHSAAFFMYLGANRTF
GFWLDFICVIYIALVILSLIFVESETFGGNMGLALTQIMGLTRMFQWGMRQWSELENQMTSVERVQEYADLKHEEDNHTVEPDKTWPDQGKVEFKNM
SLQYSPDDPPVLKNLSFVIQPSEKVGIVGRTGAGKSSLIQAVFRLTHIDGSILIDDIDTKSVALKKLRSKISIIPQEPVLFSGTLRNNLDPFDEYKD
>DvvABC-C T10132_c0_seq1+D20603+T57212_c0_seq1*
ISMDVSSKSLVAVVGAVGSGKSSLFQVILQELEPSEGYIEVNGKISFASQEPWLFGGTIRQNILFGQPYDKQRYEEVLMVCALQRDLTLLPYGDKTR
IGERGITLSGGQRSRINLARAIYKEADIYLLDDPLSAVDTHVGKHLFEECICRFLADKCVVLVTHQLQYLKTVANIILLENGMVRASGSYEDIIKSN
SIFTKLLENLEKEEEDEKNQKHVKNKLYIENQGEKEAPALIKEEKGVGKVTTYVYMNYIRAGGNLCKVVILLLAFILSQVLDTSSEYFVTIWVNIEQ
WVIQTYNSTTVLDKSNQSSTVIANPYDLTQDHWLTSFFTGNITVYVYATFVLLLLIVTYTRSIFFYRWGLSSSIVMHNKMFDNIIYSPMHFFNNNPS
GRILNRFSKDIGIVDEVLPTTILDTLQIGLSVLSICLIIGTLSPWMLIPTIIVLVIFYFMRFIFLETSRDVKRVESINRSPVYSHLTASLQGLTTIR
AFKAQDILAVEFDNLQNIHSAAFFMYLGATRTFAFWLDFFCVVYIGLVLVALLVVKEEKFGGNVGLALTQSMSLTGMFQYGIRQWSELENQMTSVER
VQEYADLTKEIDEGVDPPKTWPEQGQIVFNKLSLGYSKTLPLVLKNLSFIVNPKEKIGIVGRTGAGKSSIIQAMFRLTFTEGSIIIDGIDSKLIPLR
KMRSKISIIPQEPVLFSGTVRYNLDPFNEYPDEILWKALEDTELKYVVSDLPEGLDNKMAEGGINFSVGQRQLFCLARAIVRQNKILILDEATANVD
LHTDGLIQKTIRKKFADCTVITVAHRLHTIMDSDKVLVMDAGEILEYDSPYILLQNTNGVFYSLLMQTGTNTAQNLIEIANKAYMLSQAANASS
>DvvABC-C T48300_c0_seq1+D21892*
MDHCQRERRKPNPRETANIFSFLTFAYVGGLFRKAGKADLTEEDIFEVITNCKSKTCGDYLEQEWMNEKKLEKHPSLVRVLWNCYGLKYLLLGFIDF
VVSTARSIMTPYFISHLVGYFSPGQTTYTREDGYFYGYSYLIYICLDTVYNHQYLLWVQTLGMEMRTAFASLLYRKALRLSPVAVAETSLGNIVTLI
TRDVQSFVAVIFTINDLWIAITQTVIICYMLYSKIGVASFVGIGVLLSSIPLQILITTFIAKCRLAGCKKTDERLQLTQEILTTIKIIKMYTWENFF
SRKVSQARKVEMKKLTLAFYLKRVLILNGIiFTNLGFFMVILATIWTGVSTDTTVIFFVLSNFRYLRGWLGGAVPFGLGLGSEFVASFKRISKALNA
EELNKDQHSEQLIEKPRVELDEISVQLKGEPVLDNVSFTITSGLTMITGAVGSGKSSILKVILKDLPLTKGFVETQGRISYASQDPWLFPSSIRQNI
LFGEPYNAQRYQEVVKACALLFDFNLFENGDETILTDNGQNLSKGQQARINLARAVYRNSDIYLLDDSLTALDASVQDHIFHECIQKFLKGKICVLV
CQTASQIKEADNVIIMENGRIKDMGKPTARIIQQSYMIADQDDSIEKEVIHDSNQPTIEESSEKTGLIEAEQSTARTKIYKEQKKEGSVDWYVYKKF
MIYGGGVFLILLNIFMRGFTQFTTTSADRLLTWVDKKQAVMTIQNSIANISTEQFSYMNITLETAQAEEHTTFRLYYLCLLASSSLELLTTWLLLSF
CKTASINIHKALVYQINNSVMRFFDTHLIGNILTRFSQDMMNIDESIAYNLDQFCRMLFSMGGILTLTIFINPAFAVCVIVVTIMLFWVRSCYLPAG
RSMKRLEASSRSPMLGHLNATLEGLTTVRASNVQSILIEEYDRHLDLYTSANFTWYMSSRAFGFFVDMLSNFFLIFIIMDYLFFKRDITAGDIGLSL
TQISQLSIVLTWGIRVFTELENHMTSLERVIEYTDIKTEKKDGTTVQNWPSTGSISYQNVYLAYGDSDSFVLKNLNFEIKPGEKIGVVGRTGAGKSS
ILATLFRLYEVKGKIIIDGVDIKTLALDYLRKNIVVIPQDPILFSGTIRSNLDPLNEFEDKDLWHTLERVGINTSIASLEQPINSSVMNFSSGQKQL
LCLARAILRKNKLVVMDEATANMDHETDRMLHKIISDNFADCTVLTIAHRLHSVLVCDKVMVLDRGEMKEFDQPQELLKNRNGMFYKMVKQA
>DvvABC-C T47673_c0_seq1
MDIAKENYNPNPRDGANVFSVLFFSYTIEMFRKGYRKVLDVDDLYNPIRSDRSTVLGDRLEKKWNSHLLKCNKTRKKPSLLKILVSTFWPEYLYLGC
ILALMDLLVRLSQPIMLGNMLSYFKPGTETTKEEALWYAGAVVVLNGISALLINQYIMNAFHYGMKVRAAVCSLIYRKALRLSHTALGDTASGKIVN
LLSNDVSRFDIVSIFIHHMWVAPTSAIFILFFMYREAGYAAILGILTVFLVVPLQSYTGRLSAIYRKQTALRTDERVRLMDEVISGVQVIKMYAWEK
PFKHLIKLARKAELKIVTKSSYVRALFMTFNLFTTRAALFATLLTMVLTDQQITATKVFVFMSYFNILSQTMSAMFVRGIAEIAELIVAIKRVQEFM
LNEEFRPVVKYQNNNDKIFFNKTLINLKDLTVKWNLATKDDALKKINLAVNEGKLIGIIGTVGSGKSSLLQTILGELEVIYGSIEVNGTVSYASQEA
WVFAATIRQNIVFGGKYDKHRYNQVVKACALEKDFKQFENGDLTIVGERGASLSGGQKARINLARALYREADIYLLDDPLSAVDIHVSKHLYDECIN
GYLAHKTRVLVTHQVHHLKDADHIVILNNGTIENEGSFEELSVSDNLYAKLLTSEVELTDEEKQKQTDTAKISRKISIRRSRTSLVSAASELSLTDA
IIKEATGSDDEEEPEIRLKDLQEESSKGKVKESLFWSYLIHGGNIFFVAFVLLLYILSQVAASGVDYFVSMWVNIEEFRNQTLSRTSANKTEEVSIR
STVEWSTETCIYIYAGGLIALFVIAFARSMLFYKLAMWSSQKLHDVMFHSVIAATMRFFDTNPSGRILNRFSKDIGAIDEWLPKAILDASQIILMMV
GSLILVAIINPYFLILVFFMGIFFLSMRHVYLKTSKNIKRLEGIMRSPVFTHLNATLQGLTTIRSFGAQPILMNEFDKHQDSHTSAWFMYIAASSAF
GFYLDIICFIFIAIITFYLLNFGEALSLKGGEVGLAITQATALTGMVQWGIRQSAEITNQLMSVERVLEYKSLPAEIQPVVPKKPLPKWPEKGNVSF
RNMGLKYIENGPLTLNNLNLSIQSNEKVGIVGRTGAGKSSLIAALFRLAKIDGEINIDGIDTKDLLLEDLRSKISIIPQDPVLFSGNLRYNLDPFEE
YPDELLYKAIEDVELKDPANIINRLENKVMDRGSNYSVGQRQLICLARAIIRNNKILVLDEATANVDPQTDALIQKTIRKKFADCTVLTVAHRLNTI
MDSDKVLVMASGTMVEFDHPHLLLQNSNGVFYKMVAEAGKSLGDQLRKIAKDSYQRTLSLPE
>DvvABC-C D5345
MNKNKLSDNRQEHPIQNSNFLSRIIFLYTLPIFWKTHKRGTFEEDELYEVLKKCKAEEAGNEVEKCIKDDVKNHGYLSLFRILCKCYGLKYFLLGMI
QVIVRSIVIITIPLTLSKVIKYFQPGQDELSRSSAIYCAILLVLLNFLSVTYLHNYLLVLSELGIKVRTAFCSAIFRKSLRLTQKSLLEVSSGKIVT
LMTKDVAAIENFIFYFNDVWVGIVQAIIVFYLIFNKMGVGSLIGIGFFLLVIPAQVLLGKLISKLRLESSKKTDERLQTTKEMLSGISTIKMNVWEF
EFEHKINQTRKKEMKAILKIFLAKTTVLLIGSLTSKIAYYLLLTTFILLGNTISAELVYYITTLFLRIRHALNVAIPIGVTVTADMSAAMKRIQQMM
TAEEVIVNDEIEVDQKKSYVSLNNLSIEVDDKVLLKDITLRFDSGLYIISGSTGSGKTVLLKSILGEFDTKSDHIIKYGTISYAAQEPWIFPATFKQ
NIVFDEPFDEERYNRILDICALGFDIEHFTLANNLKIGDKGANLSKGQKSRINLARALYRNADIYLIDDCLSSLDVNVQRHVFQKCIREFLSDKICI
FVTQNLQYTNGNDKIIVLNEGKTFVSQNLSEEDDQDIKTILKKEQQLKNLIEVNNNADEVNNDETTEESSLLKKLNIGRNNMYQENKKSGKVDFYTY
RKYGNYGGGILVIAIVLVFFIGSQFSKSYTEKILSNWVNHIHHNTTNINEENTVWDFSLKIYSAFTLSTTILSIFSAFLLFNFTRKASIKLHRCLSN
TVINATMKFFDSHLIGNIINRFSKDLSIVDEQIPFAIYEFLEIVLSLCGVIILITLVNVGFIIPTAVVFILLYFLRKLYLPTGRNLQRLDTAVRSPM
LGHLNSSLEGLSTIRAFKAEETISKEFDRHQDLYTSANYTSKCCQRAFGYYLDVCCTTFIGAIIAAFIIVGKETGVGDVGLAITQAFMLNGIAQYAI
RLWADIENKMTSVERILEYVGIKTEDNDLRDIFKDWTSNNIIQYKNVNFRYDNSQQYILKNINITINEGEKVGIIGRTGSGKSSLITLLLRLFDYEG
KITIGNEDIQSLPLTLLRSKISIISQDPVLFSGTIKSNMAPKVFGVYPDEDIWKVIHKLELTTIIDSLDFQICEGGKNLSAGQRQLICLARCLLHQN
KIILLDEATSEMDQGSESLFNNILEKQILDCYSFRSTVIVVAHSLKTILNCDRVLVVQNGEIVENGAVSTLAQDEHSLFYNMLKSSGLLTIKC
>DvvABC-C D22413+T49258_c0_seq1
MDSTKETYNPNPRDSSNIFSILFFGFTYPIFKKGMKKDFDVEDLYNPLKNDRSTLLGDKLERNWDKQKLKPKGKPSLLRAVAITYWYEYLRLGFLTL
IGDVFIRVSQPYVLGLLLTYYNPHSKTTKEQALGYAGMIVTFNTLTSFIKNQYMMNSMHAGMRVRTSVCALIYRKAIKLSTTALGKSSVGKIVNLLS
NDVSRFDSASMLLHQMWVGPVSAIIIMYIIFQDIGWSGATGVVTIVTIMPMQAYIGKLSAKYRRIIAGKTDERIRLMNEVIGGIQVIKMYAWEIPFT
KLISLARKAEINVIRRSAYIRGVFMAFNLFNNRFALFATLATMVFTNKAITAARVYVFMSYYQILANTLAGVFVNGLTQIAELLVSINRIEEFLENE
EHKELPALKRSVISNVNEEMVACKNMVASWNENSLDPVLKNLNFKLGKNTLLGVIGPVgSGKSSLLQSILGELDIVDGSLAVHGTLSYASQEPWIFS
GSIRQNILFGSDFDKARYDEVLKVCELRKDFDQFPDRDFTLIGEKGANLSGGQKARINLARAIYKDADIYLLDDPLSAVDSIVSKVLYEDCINGFLA
KKARILVTHQVYYLKTADHILVLNNGGIEIEGTYSELLKSDNPLTVHLTEEIEEILKRQKSQDSAIEEQEAKTMSTMDLSKVNGKMKAIVKEMQEET
SKGKVKGSLFLEFFKSGMSICEIIFLVTLLTIAQASATVIDWFISFWTNIEEYQMTLNSTAANVLNSTVVDALNSTAVDAINSTAVDVLNSTTVNAL
NSTVVDESVPFYMNWSSQTCLIIYGAILVFCLVTTLSRSLSFYRFILNCSENLHGMLFTGVTNTYMRFFDKNPSGRILNRFSKDIGSVDEILPRMLF
EASRVTLKMLGHLVLVLYVNPASVIVVIILGILFSFIRIVYLRSSNNIKRLEGRMKSPVFSHLTATLEGLTTIRAFKAQNILRAEFDKHQDSHTSAW
FMFISTSSAFGFSLDVICLMFIGSLTFSLIGLGEYFNLTGGDVGLAINQASSLTQNIQFLVRFSADISNQLMSVERILEYKELIPEQQPEKLLLPPK
SWPDKGIVNIEHLNMKYIDDGPTILKDVSVKINPKEKIGIVGRTGAGKSSLISAIFRLTPLEGKIYIDDINTKDITLKQLRSKVSIIPQDPILFSGT

LRYNLDPFDEYSDETLYRALNEVELKDPSNIINRLENRVMDRGSNYSVGQRQLICLARAIIRNNKILVLDEATANVDPQTDALIQRTIREKFADCTV
LTVAHRLNTIIDNDRILVLEAGEIVEFDHPYLLLQNKFGVFRKMVEETGTAMLRQFLETSSQNYQKLID
>DvvABC-C D18709*
MEMTKEKYNPNPREKANILSTLFFGYTYDVFKKGLSKTLEVDDLYNPLKIDRSKILGDHLQSNWDKQLEKAKKTNSKPSLLKAILMTYWLEYLKLGI
FDITTDLVLRLIQPLMLGKLLDHFKPEAQVSKNEALMYAGIIVAINVLNFLIGNQYMVEGFHSGMRIRAACCAVIYRKSLKLSKTALGETASGKLVN
LLSNDVSRFDLVSLLIHQIWIAPVLSLIVMVLLYQRTGYAGIVGVVAVFVIVPIQTYTGKLSAKYRKQTAMKTDERVRLMDEIISGIQVIKMYAWEI
PFRKVIRIARRNEIKIITKSAYVRASFMALNLFTTRLALFCTLLTVILSNEPITASMVFVVMSYFNVISMAMSTMFTRGVSEIAECLVAVRRIRDFL
LNEEYDPNRSTSGMNGNVKSVDDFKEIISLKNLTVKWNLSFSDNALENINLNVQDGQLIGIIGPVGSGKSSLLQTLLGELDITQGEMRVQGQISYAS
QEPWVFAATVRQNILFGEEYDKKRYQEVIQACSLEKDFEQFPNGDLTLVGDRGSSLSGGQKARINLARAVYREADVYLLDDPLSAVDIHVSKHLYEK
CINGYLASRTRILVTHQVHYLKDADNIIILNNGRIEDEGTFNFLANSDNVYAKLLTAEPENKEEKKPERQKYSRQLSQRSRKDSMSSIISELSIADT
LLSNDVDFDEEEAEKEPEFDVKDLQEQSSKGKVGGSLLFKYMLAGSNVFAVFICVLLYLGTQLAASGTDYWVSYWVNVEEFRNSSESLNSTAPRFII
PSIELTTDNCLYIYSVILGALFVLAMTRSFFFYKMAMLSSKKLHGTIFDNVIDATMRFFDTNPGGRILNRFSKDMGAVDELLPKAILDSSQILLSMC
GSLVLIIVVNPYFLILIGVLSAVFGVMRHIYLKSSKNIKRLEGIMRSPVFTHLRATIEGLTTIRAFGAQSTLMDEFDHHQDYHSSAWYMFIVSSTAF
GFYLDCFCTIFLAVLTFCLLLFGETFNLRGGEVGLAITQATALANFLQWGMRQSAEVTNQLMSVERILEYKSLDKEPQPVGPKKPAKEWPQKGEITF
KDTCLRYFEGGPLVLKHLNLSIQPKEKVGVVGRTGAGKSSLIQALFRLAPIEGTIKIDEIDTKDITLNDLRLKISIIPQDPVLFSGTLRYNLDPFEE
YTDEVLYKAIEDVELRDPANVINRLENRVMDRGSNYSVGQRQLICLARAILKNNKVLMLDEATANVDPQTDALIQKTIRKKFSDCTVITVAHRLN
>DvvABC-C D21941
MDSTKKHNNPSPQISANFLSKIFFCWIFPLFKTGFKKKLQPSDIYNTLNSDTCHATINTLQRHWENELMLHKAGKKDKPSLKLALFKTYAFPYSLQG
VMVFLQVVLIKTLQPLVLAELLKYFDKTQKYDMFGEYSGWILGTATVFLAFIFALSYHHSTLGSQRIGMRARAACSALVYRKLLKLSQASLNKTPGG
KLVNLLSNDLQRFDIASMYLNFIWVMPFQAAICFYIMYRSVGIAALAGTAFMLFEgIVMQGYLSRLQGTLRSKIAEKTDFRVKLMNELVAGIKVIKM
YAWENYFAKVVETARTEEVRLISKTSAIKGLSLAFIVVTERIALYLTVITFVLLGHQISADKVFLSAQLYNSLQLYSCIMFPYALAGYAEVKVSLRR
LEQFLLLEENNVQVQNTVVQQIGTIQAREACATWNEDLKDDTLNNLNLKLrAGKLCCVVGTVGSGKSSLLQMLLGELPIKSGQLEVSGDLSYASQEP
WLFVSSVRENILFGKPYIKNRYDEVVQVCALETDFQQFPFGDKTIVEERGVSLSGGQRARLNLARAVYTSADIYLLDDPLSAVDTRVGKHLFEKCIK
GFLGSKTRILVTHQLQFLKDADVIIVFEKGKIKKMGTFDELSEDYLKSLQENRSEEQIEDEKKDVSLNTRKQSNMEEIKQVLSQNEEGPKENNEETC
EGSIPFSTYGKYFRFGTSCFGFCFMVILFVLAQAASNAGDLWVTYWVNTVAANSTTTPPLEIVVANNTSSGNYSKYNYSVVGENIPLEERFTHDSNY
YIIIYSVTICLAVFLTPIRSLNFYRIIMNASRNLHNSMFSKILEAPMRFFDTNPSGRILNRFsNDMGIIDELLPKAMLDGTQVLLVLVGILVLVFIK
ILWMIIPAVVIGILFYYLRLFFLKTTQDVKRLEGVSRAPVYSHVIATLDGMTTIRASKCEDMVIQEFDGLMDNSSGAWYLYIASCEVFGFYLDCIST
VFIAIVTYQFLIFDNPDPSSAGNAGLVLSQCLILTGMLQMGVRQTAEVANNMTSVERVLEYTTLEKETEGKKQLSVNNLIPRYKSLDKDWPQAGEVE
FKDVFLRYTLESAPVLKNLNLVFRAGEKIGVVGRTGAGKSSLVSALFRLSPLEGSVKVDGVDTVNIELSKLRSTISIIPQEPVLFSATVRYNLDPFN
TVSDEQIWDALEKVELKHSITDLEMEIREGGSNFSTGQRQLMCLARAIVRNNKILVLDEATANVDPSTDALIQKTIRKNFKYCTVITIAHRLNTIMD
SDKVLVMDAGQAVEFEHPHILLKDYEGYFSRMLKETGTATEKALKKIAEEHYRQLFLDEQeASQEGL
>DvvABC-C D15305
MEMTKEKYNPNPRETANIISALFFGYTYDVFKKGLSKTLEVDDLYNPLKSDRSKVLGDNLQRNWDKQLEKSKKSNSKPSLLKAIVMSFWLEYAKLGI
YTITTDLVLRLIQPLILGNLLDYFKPESQVSENEALTYAGILVAINILNILMSNQYMLEVFHSGMRIRASCCAVIYRKSLKLSKTALGETASGKLVN
LLSNDVSRFDIVSLFIHQLWMAPVLSLIVMGILYQRTGYAGVVGVVAVFVIVPFQTYTGKLSAIYRKQTAFKTDERVRLMDEIISGIQVIKMYAWEI
PFRKIIRIARRNEIKIITKSAYVRATFMALNLFTTRLALFCTLLTVILSNEEITASKVFVVMGYFNVLSMAMSTMFSRGVSEMAEVMVAIRRIRDFL
INEEYDPNRTNFAMNGNVNSVDDFKEIISLQNLTVKWNASFSDNALENININISDGQMVGIIGPVGSGKSSLLQTLLGELDITEGSMRVRGQISYAS
QEPWVFASTVRQNILFGEEYDKKRYHEVIEACSLEKDFEQFPNGDLTLVGDRGSSLSGGQKARINLARAVYREADVYLLDDPLSAVDIHVSKHLYEK
CINGYLANRTRILVTHQVHYLKDADNIIILNNGRIEDEGTFNFLANSDNVYAKLLTAEPENVEEKKQDRSKYNRQLSIRSRKDSLASIVSELSIADT
LISNDIDFEEEEAEKQLEFDVKDLQEQSSKGKVGGSLLFKYMLAGSNSFAVLICVILYLGTQLAASGVDYWVSYWVNVEEFRSSSESSNSTTPTFII
PSIELTTNNCLYIYGFLLGALFVLAMTRSFFFYKLAMWSSKKLHATIFDNIVTATMRFFDTNPGGRILNRFSKDMGCVDELLPKAILDSSQMLLSMC
GSLLLIVVVNPYFLILIGALSVVFGFMRHVYLKSSKNIKRLEGIMRSPVFTHLRATIEGLTTIRAFGAQSTLMDEFDHHQDYHSGAWYMFIVSSNAF
GFYLDCFCSLFLAALTFSLLLFGEAFNLKGGQVGLAITQATALTSLLQWGMRQSAEVTNQLMSVERVLEYQNLEKEPQPVVPQKPAKEWPQKGEITF
KDTCLKYFEGGPLVLKHLDLKIQPKEKVGVVGRTGAGKSSLIQALFRLAPIEGSIKIDDIDTKDISLNDLRLKISIIPQDPVLFSGTLRYNLDPFEE
YSDEVLYKAIQDVELKDPANVINRLENRVMDRGSNYSVGQRQLICLARAILKNNKVLMLDEATANVDPQTDALIQKTIRRKFSDCTVITVAHRLNTI
MDSDKVLVMDAGQISEFDHPHLLLQNKNGVFYSMVAETGRTTAEQLRKIASDSYQKLNALPE
>DvvABC-C D12562+D20321
MEEIHSKAKAKKKNPLDTANIFSSIFFFWLLPFFVKGYKKDLTEDDMYEHRNCHESGKLGDQLQVKWNKQLRKKKPSLIKACASSFVLDLLGINCLV
VIAEVVRISQPFLIAELLKVYESKQLDEERSEIYLYASLIVLTSFVAANMGHNFNLRLMQLGMQMRVASCSLIYRKALKLSKSALAETTIGQMVNLL
SNDVSRFDLAVYFFHNIYIGPIEILIVMYLLYVNVGGAALAGAIVLLLFIPFQSWLGKKTSQFRLQTANRTDERVRLMNEIIAGIQVIKMYTWEQPF
AKLVEFSRKKEMKYIRYTSVIRSVLMSCIIMLHRSAIAASILVYVFTGNALTASYAYTVTSYYRLLYTVTNFLPTAISQAAELYVSSRRIQTFLLYD
EVEDENYIALDEKSHKENGIKSLELTIRNKEPGIHLKNASAKWLKSSPENNLEKINMDVTPGNVVAVVGPVGSGKTTLLHIILKELELQSGSVDVHG
VVSYASQEPWLFGGSIRQNIIFGQKFDQSKYDEVVKVCALQRDFTLFPHGDRTLAGERGVSLSGGQRARINLARAVYKDADIYLLDDPLSAVDAHVG
KQLFEDCVTGYLSSKCVVIVTHQLQYLKKLKHIYLMRDGKVKLSGTYQDLKNSNTEYSKLLTDIKDEEETRKMSRTRTLSREEEMFEDNEIQILAKE
AQSTGKISGRVYMNYIKASGHLFHIFLISMIFITSQFMESFSEYYVTYWVNIQQVNSQNLTETVTSNSNITTNLNITSNSSLPKALRTTVQPENPYE
GWWLTPLFTSEYTAIYYLLVVISLIIMILGRSLYFYHLSLTSSTTLHNQMFQNIIYSPMRFFNINPSGRILNRFSKDIGVMDENLPMTLMDTVQIGL
YVFSICLVIGSLTIWIVIPTVLLAVLFYLMRIVFVQTSRDVKRIEAVTRSPIFTHLSASLQGLTTIRAFKAQEILKKEFDNYQNLNTSAYFIFAGAN
RTFGFWLDFVCVVYIGLVLAALLLIKSEQFGGNIGLALSQAMALMGMFQWMRQWSEFENQMTSVERVQEYADLKQEIDQPSNKPRETWPEKGEVEFS
NMSLKYSPDDPFVLKNLTFVVKPKEKVGIVGRTGAGKSSLIQALFRLTHIEGSILIDGVDTKTISLKKLRSNISIIPQEPVLFSGTLRKNLDPFDEY
KDEVLWSALEEVELKHAVQELPAGLDNKMAEGGSNFSVGQRQLLCLARAIIRNNKILVLDEATANVDPHTDGLIQMTIRKKFENCTVLTIAHRLHTI
MDSDKVLVMDAGKMVEFDHPYSLLQNKQGTFYSLVMKTGKGTAQTLMSIAEESQKTK
>DvvABC-C D10642
MDTGNKTTKPRNPAENTNPISFLFFFYMFPIFKKTYKYKLTEEELFSPLKEHTSSRLGSKLEKAWKEEYRIHKKTALHRALFRIFGLRYTVLGLIRL
FDELLLIVVMPYCIRTLVAYLEAGQTKITKDEALIYAAALVITLLLDAVMQQPNYMGLQHIAMKIRVACSSLIYRKTLRFSREALGNTTVGQLVNLL
SNDVSKFDQLFGLTHYAWIGPIQVALGTWLLYREIGVSAFFGMAFLVAFVPLQIWLAKKMSVMRLKTALRTDERVKLMNEIISGIQVIKMYCWEKPF
AHVIDLARRSEMRAIRSHSCLLGILYSFEVFVSRTAIFVSIVGYVLLGNYVSADKVFAITAIYNQMRTIITIIFSLSITALAEMHVTIDRIHKLMIF
DEREQDSEDGYEKMNGNYQSKLNGLDGMKNGINGNGTAIELIKKVKEPKLMLSGVSAKWLAQSPENNLSDITFNVPPNKMLAIIGPVGSGKSSIINL
ILKELPVKSGKLEIDGKVSYASQEPWLFAASVRQNILFGEEYDEERYKLVVEVCALKSDFALFPHGDKTLVGEKGKALSGGQKARVNLARCVYKKAD
IYLLDDPLSAVDANVGKHLYDRCIKQFLSNKICVLVTHQLQYLRNADKIIIMKDGKMEMTGSYTELKKSGLDFAKVMEEFNEEAEEDKRMKSIKSKA
SIYDEPMEDEEDQVLEKEMQEKGTIKAATYYLYLKAGGGICSMLALGFLFIICQVVANAGEYYVTYWVNLEQDFSEKQRMNLTAPNETINRDLIMWS
YTALIVGNIIISVVKAVYFMIFFVIASKNLHKYIFDKLIKATMRFYNTNPSGRILNRFSKDLGTVDEYLPSVIIDVIEIALLLLGAITLSAIVEPLL

TVPAVVLMIIFYLLKIVYSETSRSVKRVEAITKSPMLSHLTASVNGLSTVRAFHAEKMLTEEFDNYQDSHSAAWFLYLASSKCFGLWLDIICIVFIA
VAVFSLLLFRDTIHGGDLGLVITQYLGLMGSLQWGMRQWTELENNMTSVERILEYTRLETEPERKEPKNIPQPWPEKGLVEFRDVSLRYSPQDPPVL
KGLNFTVQPKEKIGIVGRTGAGKSSTITALFQLYPLEGTVVVDGVDTTQIPLDLVRTNISIIPQEPVLFSGKMRENLDPFENYSDDVLWNALDQVEL
KDVISELPAGLNTEVTEGGNNFSVGQRQLVCLARALIRNNKILVMDEATANVDPHTDSLIQKTIRDKFADCTVLTIAHRLHTVMDSDKILVMNSGRV
EEFNHPYLLLQNVHGVLHNLVDATGPSTAKNLENIAKESYEKKRV
>DvvABC-C T41602_c0_seq1
MEEIRNDQPPKKSNPIERTNAISRLFFCWLIPYFAKGYKRDLTEDDMYKHRSEHDSGSLGARLERRWSKHIKNNKNPAYWKVLIGTFFCEIFLLNIL
AFTTEGIRMAQPFLISKLLKVYEEIPTEESMNDIYLYAGLIIVTSFLNVIMVHRFNLAVMQIGMKMRIASCSLIYRKSLKLSKSALAETTIGQMVNL
LSNDVGRFDQATHHLHHFYLAPIQTTVVMVLLYLVVGWTALLGTVFLLLSIPLQSWLGKKTSQFRLKTATRTDERVRLMNEIISGIQVIKMYTWEYS
FAKLVEYVRKMEMKYIRKTSLIRGLLMSMTTVLNKGAVAISIITYVISGNTLTASYAYTVASYYRLLGSITMWFPQAVSQGSEMFISMKRIQSFLLF
DELEADKYFEPGQELKEVKKQHKYITDKVLQTSNGSIHIKKASAKWLKSQPDNNLENINMDVTAGDVVAVVGSVGSGKSTLLHIIMKELELQSGSVD
ITGTVSYASQEPWLFGGSIRQNILFGEEFDQEKYDEVVKVCALERDFTLFPHGDRTLAGERGVTLSGGQRARINLARAVYKDADIYLLDDPLSAVDT
HVGKHLFEECICGYLGNKCVVLVTHQLQHLKNIDKIYLMTNGKIAVSGSYSDIRESNTEYSKLLADIEEEEEETRRKSQRVATKSEKPEKEGAVQVL
QRESTGSGNISDHVYMSYARAGGNLIKVFLLIFMFFFGQALDSLSDYFVTFWVNIQQWNSTKVIINATDPDAVVYHSFENWWLTPVFTSKNTVYFYT
ALVASVVIIVNIRSLSFYLWCIGASTKLHNKMFMNIVYSPMRFFNTNPSGRILNRFSKDIGSVDEVLPTTLLDTVQIALAVAAITTVIGSLTYWIMI
PTVAIAVFFYFLRVIFLETSRDVKRVESVTRSPIYTHLTASLQGLTTIRAFRAQDILKEEFDRYQNLHSAAFFMYLAANRSFGFWLDFICVIYVGLV
IVALVFVKSEQYGGNMGLALTQTLGLTGMFQWGMRQWSELENQMTSVERVQEYADLKQEDDEHSIEPPKTWPHAGEIQFDNMSLYYTTDDPPVLKNL
SFTIKPSEKIGIVGRTGAGKSSLIQALFRLTHIEGRILVDDFDTKKISLKVLRSKISIIPQEPVLFSGTLRNNLDPFDEYTDEALWNALEEVELKQA
VEELPAGLANKMAEGGTNFSVGQRQLVCLARAIVGNNKILVLDEATANVDPYTDALIQKTIRKKFANCTVLTIAHRLNTIMDSDKVLVMDAGQAVEF
DHPFTLLQNINGVFYSLVMQTGKATAKNLFAIAEQSKMLRDNNTVF
>DvvABC-C D12703
MDDGNCDELKNKGKNPQQSSNFLSNLFFCWGLKVFYKGWNKVLNEDDLYKPLEEHESHYLGDQLEAFWKLEKLEHSHPCLIRPIWKLFKKDILIHAC
CTFVLEFCIKLSQPLLLKKLIEYYEPNQTSVTLTEAYIYSLFIVLCSLFYVIFCHSFHLSLQHLGMKIRIACCSLIYRKSLKLSKKALVNTTIGQMI
NLLSNDVNRFDNLFRYWHFLWVGPLESIVVMYLIFEVAGYAGLAGFlIMLLFIPFQMVMGKLTTQFRQKTAGKTDERVLLMSEVICGIQVIKMYTWE
KSFSKLIEHIRNSEIKQIRLTSYIRALHMSFSKFITRAAVFLCIITYTFTGKRLNAGYVYVISPFYNILKSAVSNDFPQAIIQTAEALVSIARIEEF
LQFDEVYYEPIREPYQAKKITSGTSLIHSSKKKKSVGVYLENVSAKWIDTQEENTLMNINFNVGPQQLVAIVGAVGSGKTSLLQTIMKELPLSQGNK
DTVGKISYASQEPWLFASTIKQNILFGEPWDPKKYERVIKVCALERDLSLLAHGDKSMVGDRGVALSGGQRARVNLARAVYKDADIYILDDPLSAVD
THVGKQIFEDCICSYLRGKCTVLVTHQLQYLKNVHKIYLLERGKIVVSGTFKEISESNTDFARLLVSNVEEEESDTIDEESSEIEYFENMSDEKPTE
ITEETVVVKLTSRAYSGYLKYAGGWFYGFWVLMLFVLTQFLASGSDYFVNFWVNREQETFQEEEGNKTLLLGYSDDVYMEENYNYVWKYSLGHLTPI
QPDFFNSDRCIIIYSCIILSVITVTVMRSLCFYKMCMNASVKLHNTMLHKICNGTMFFFNTNPSGRILNRFSKDMGCIDETMPNVLVDTIQIALNVL
AVNIVIGTVNPWILISTLAIGCLFYLFKVVYVSSSRNLKRMEAATRSPVFSHIHASLQGLTTVRAFGAQEILRNEFDKHQDLHSSAYYLFVCCNRAF
GFWLDLICVGYIALVTFSFFIFGSAYGGNVGLAVTQAISLTGMFQWGMRQWSELDNSMTSVERVLDYINIKQESLEKQKEPPKKWPEYGMLEFQSVY
LRYATNLPYVLNNLTFKIKPKEKVGIVGRTGAGKSTLISALFRLTDIEGKILIDGICTSDIPLISMRSKIAIIPQSPVLFSGTVRSNLDPMDEYDDK
DLWRALEEVALKDVVSEMEMGLESKIGEAGSNLSLGQRQLVCLARAIVRKNKILVLDEATANVDPKTDALIQSTIRNKFANCTVLTIAHRLHTVMDS
DKVLVMDAGSAVEFNHPHILLQNTEGIFCSLVKQTGKSMAENLCKIAEESYMNLQK
>DvvABC-C D14968
MEEVKKNRIQKKPHPLKRTNLFSHIFFCWLPGFLSKGLRKDLDENDMYQTRHSQQSRYLADQLVVAWKNELKKSKPSFLWALFVVFRLELFYYAFFN
MFSDFIKIAQPMLISRLVSYFQPGATTNQTEIYINAFLIIVASLIQVTSVHNYQMLVMSLGMKVRVASCALIYRKALKLSKTSLAETTIGQMVNLLS
NDVGRFDFSGQHIHNVWLAPCETVVVMILLYFYVGPTGLIGCVFLLSFIPFQMYMAKLTSQYRLKTAIRTDERVRLMNEIINGIQVIKMYTWEKPFA
KLVEMVRKREITEIGHTSTIRAIMMSFNLTMSRSAVFLCVMTYVLTGNTLTASYAFTVTSFYAFLRVSVTQQFPQAITQFAETKVSISRIQKFLMYD
ELDRRTDSEKNGLNGVTEKISNKTDAKEVGVKIKNASVKWIKSLPENTLEEITFEAKSNQLIAVVGTVGGGKSTLLHVILKELDPIEGSVEVNGSIS
YASQEPWIFGGSIRQNIIFGQKYDEIKYEEVLRVCALEKDLALFPHADRTLVGERGVMLSGGQRARINLARAIYKESDIYLLDDPLSAVDTHVGKQL
YTNCITGYLKHKCVILVTHQLQYLRTANCIYLFEDGKIRASGTYQALKNSDSAFTKLLASSKDEEKKDGARRISKAESVESEASEAEHEAIEQKREE
RATGTVSKRVYANYINAGGYWLKSVVLLCTFVAAQVFGSLTDIFLTTWVNVEQWRVDNKNNGAFTNKTLNESLANATYSTNDNKPPFWSTVLTEDNT
LLIYSCLVILTIILAVTRSITFFRFCLRASTNLHNSMFGKIIFSPMLFFNTNPSGRILNRFSKDIGALDESLPNCLVDTIGIGLIVTGTTLVIASVN
PWVLLPTGVILVIFYFIRQAFLASSRDIKRVEAVTRSPIFTHLSASLQGLTTIRAFRAEEILTKEFDHFQDAYTAAYYMFLTANRGFGFWLDLHCVI
FIGMVVVSILFIQKETFGGNVGLSLTQAITLSGMFQWGMRQWSELENQMTSVERVQEYADLPKEKDEIKKEPPPNWPSMGHMKFENMSLRYSDDSPY
VLKNLNFEVKPKQKIGIVGRTGAGKSSLIQAVFRLAQNEGHIFIDGIDSKSVELRLLRSKISIIPQEPVLFSGTLRKNLDPFDEHKDEVLWDALEEV
ELKHAVDELPAGLDSKMAEGGSNFSVGQRQLVCLARAIIRRNKILVLDEATANVDPMTDSIIQTTIRLKFSECTVLTIAHRLHTIMDSDKVLVMDAG
EAVEFDHPHKLLKKKGVFYGLVRQTGTTMAENLQGIAEESFYKQIKDR
>DvvABC-D D11014
MPTVISKFLEQGESTFKQNKSVFSGAFIAAVLCTYAYKVGYPFVDSLIHKPKDNLNNNHIVQKDLQRKNGLVKNKKLKGRLKNSIPNFNLAFILQFI
KLVRIMIPSFICTETVLLSGHTTFLFLRTFLSIYVANLEGAIVKYIVMKDPQNFVKQLGKWFAVAIPATFINSMIRYLESRIALSFRTRLVDHSYKL
YFKNQSYYRVTVLDGRLDNCAQRLTDDIETVANTVSHLYGQITKPCFDILLMVIALANLVKSRHSNLVIGPVIICGVVMFSALLLRFVSPRFGHLVA
QEAEKKGYLRHVHGRIVSNAEEIAFYGGHQVEESQLRQAFRVLSKHLEHMFGVKLWFIMLEQFLMKYVWSGAGIIVVSLPILLAAGNRRKISNKSLL
SIPDFSSKSESLIAEDTADDQIEDSVSERTHYFTTSKNLLITGSDAVERLMSSYKNIVELAGHTARVANMFEVLEEASNGIYHKTLVAKKEKSADFE
IEFRGDQPLAKGKIIYSTNNEIILKNVPIVTPNCDIVCPSLSLELTPGQHLLITGPNGCGKSSLFRILSGLWPIYGGELHTPKNSMFYIPQRPYMVI
GNLRDQVIYPDTYTDMVKKGITEENLLKIMTMVHLDHIVERDGFYEMKDWTDILSGGEKQRMAIARLFYHKPKYALLDECTSAVSIDVESFIYQSAI
DMGITLLTITHRPTLWKFHTHILQFDGTGSWEFSQLNHTSRLTLKKEKEDLLKAENNEERSKRLDELNKLLGEDS
>DvvABC-D D11628
MAPNYSKVLNKHKALAGAGALGTILLIILKYRNKQGKLNKKKVQKAVEEEVKYLISEKGETKIKAQVDKKFFSQLSQLFGIACPGWTSQESGLFFLI
ALSLVSRSMCDLWLINHGTKIESSIIAMDPTLFKQRLLYYVLAIPIISVVNNVLKYSIGALKIQLRTNMTRHLYEEYLKNYTYYRISNLDNRISNAD
QLLTTDIDKFCEGVTDLYCNTAKPLLDICIYVYKLSTTLGGGTPGIMLVYLLVSGVLLTNLRKPTARLTAGEQKLEGEFRHINSRLITHSEEVAFYN
GNSREKATLMASYNKLLNHLRKFLRFRVAMGVVDNIVAKYFAGVVGFWVVSLPFMTSGHSFRLLGTNERSRLYYTYGRMLVKLAEAIGRLVLAGRDL
TRLAGFTARVTQLKTVLSELNSGKYQRTMVSGSESLQINGGkLIFRNNIIKFHKVPLITPNGDVLINEITFEINSGMNVLVCGPNGAGKSSLFRILG
ELWPLFGGELTKPPRGKLFYIPQRPYMTLGCLRDQLTYPHSGAEAARRGTTDAKLEEYLQRVQLGYILEREGGLDAVADWLDVLSGGEKQRIAMARL
FYHQPQFAILDECTSAVSVDVEGSMYKYCRDVGITLLTVSHRKSLWQHHEYVLHLDGRGGYSFKPIDNCDEQFGS
>DvvABC-E D2830
MSKRKGMEESDKLTRIAIVNADKCKPKRCRQECKKSCPVVRLGKLCIEVVPNSKIATISEELCIGCGICVKKCPFEAIAIINLPSNLQKETTHRYGK
NSFKLHRLPIPRPGEVLGLVGTNGIGKSTALKILAGKQKPNLGRYMDPPDWTEILSHFRGSELQNYFTKILEDDLKALIKPQYVDQIPKAVKGTVGQ
LLDRKNELDNMNQICGMLDLLHIKEREIAALSGGELQRFACAMVCIQNGDIFMFDEPSSYLDVKQRLNAARTIRSLIDPSKFIIVVEHDLSVLDYLS

DFICCLYGVPGAYGVVTMPFSVREGINIFLDGFVPTENLRFRDESLVFKVAESATEEEIKRMNHYEYPTMTKTMGSFELKVAQGQFSDSEILVLLGE
NGTGKTTFIRMLAGNLEPDSGSGELPQLHISYKPQKISPKSTGLVRQLLHEKIRDAYIHPQFIADVMKPLKIEDIIDQEVQNLSGGELQRVAMTLCL
GKPADVYLVDEPSAYLDSEQRLVAAKVIKRFILHAKKTGFVVEHDFIMATYLADRVIVFEGSPSVKTTAHAPQTLLAGMNRFLELLGITFRRDPNNF
RPRINKLESVKDVEQKRAGQYFFLED
>DvvABC-F D2701
MSKKRGAKKGKNLDDDFEETSSIISEKEKINSKTPKNKTTKKGKKGKDDWSDDEEIEQGKEVNQSEPAESKPVAKKKGKKGKGNKNDDWSDKEEADI
KLSDSETESMPVAVKKSAKKNKKKKDDWSDKEVDIQLSESDIEEAPKIVKKSAKKNKKKDDWSDKEDIEIETKLSDSEEELAPAIVKKSAKKNKKNK
NQVRDESPDIEEMEAHDEGFKEDDVSEEEIVKPVKQDKKKSKGKESKKDESKVDVEVKEEIKESKEDSPIVEDKVIENGKPEKQNKELENKKDNDVE
ELTEKISTTQISDDKLDESKEKKLTHKEKKKMKKLQEYEKQMETMLKKGGQGHSELDSNFTVSQTQKTAGQLAAFENAVDIKVENFSISAKGNDLFV
NANLLIAQGRHYGLVGPNGHGKTTLLRHVAQRAFDIPPNIDILYCEQEVVADDNTAVETVLAADVKRNDLLAECKKLEAAANSGDLEIQERLNEVYS
ELKAIGADSAEPRARRILAGLGFDKEMQDRATKNFSGGWRMRVSLARALYIEPTLLLLDEPTNHLDLNAVIWLDNYLQAWKKTLLIVSHDQSFLDNV
CNEIIHLDNKKLYYYKGNYSMFKKMHVQKKREMIKEYEKQEKRIKELKSSGSSKKQAEKKQKEALTRKQEKNRTKIQKQEEDTTPTELLQRPKDYLV
KFRFPEPPPLQPPVLGLHNTRFAYPGQKPLFVDTDFGIDMSSRVAIVGPNGVGKSTFLKLLTGDLSPDKGENRKNHRLRIGRFDQHSGEHLTAEETP
SEYLMRLFDLPYEKARKQLGTFGLASHAHTIKMKDLSGGQKARVALAELCLNAPDVLILDEPTNNLDIESIDALAEAINEYTGGVIIVSHDERLIRE
TNCSLYVIEDQTINELEGDFDDYRKELLESLGEVINSPSIAANAAVAQ
>DvvABC-F D802
MPSDAKKREQQRKKDAAKARQAGKKNEKQKGEEQNDTLKTNGLSNGATNGSSELSAEEALCAKLEADARLNAEARACTGSLAVHPKSRDVKIDTFSI
TFHGCEMLQDALLELNCGRRYGLLGLNGSGKSTILAVLGNREVPIPDHIDIFHLTREMPASDKTALECVMEVDEERVRLEKLAEELVACEDDESQEQ
LMDIYERLDDMAADTAEARAANILHGLGFTREMQNKKTKDFSGGWRMRIALARALYVKPHLLLLDEPTNHLDLDACVWLEEELRNYKRILVLISHSQ
DFLNGVCTNIIHINKKRLKYYTGNYDAFVKTRMELLENQMKQYNWEQDQINHMKNYIARFGHGSAKLARQAQSKEKTLAKMVAQGLTEKVTSDKIVT
FYFPSCGTIPPPVIMVQNVSFRYNDSTPLIYKNLEFGIDLDTRLALVGPNGAGKSTLLKLLYGDLTPTEGMIRKNSHLRIARYHQHLHELLDLDLSP
LEYMMKEFPDIKEKEEMRKIIGRYGLTGRQQVCPIRQLSDGQRCRVVFAWLAWQVPHMLLLDEPTNHLDMETIDALADAINEFEGGMVLVSHDFRLI
SQVAEEIWVCEKGTVTKWQGDILSYKDHLKTKILKDAAKRK
>DvvABC-F D9935
MGTCSEYIKNVFPAIDEESKQYVEGVLLNGADDFEDSEEVYDAVGEVLKEISNDKSEDDIRNICNDLLCMLKPDKGEKATNGAMKVLNAPVHLGSMV
DNTDTNIDDVKSIWLIQRDDSLKVDARKLEKAEAKLQEKLDKRTKEIKVIAPPKLQTATASQVTSKKDSKLEAKGTNRTQDIRIENFDVAYGDRVLL
QGADLTLASGRRYGLVGRNGLGKSTLLRMISGSQLRIPSHISILHVEQEVVGDDTVALDSVLECDTVREELLKKEKEISAAINSGSVDPQLNSQLTE
VYNQLQNIEADKAPARASIILNGLGFTSEMQQNATKTFSGGWRMRLALARALFSRPDLLLLDEPTNMLDIKAIIWLENYLQNWPTTLLVVSHDRNFL
DTVPTDILHLHSQRIEAYRGNYEQFEKTKTEKLKNQQREYEAQMQQRQHVQEFIDRFRYNANRAALVQSKIKMLEKLPELKPIVKETEVVLRLPETE
PLSPPILQLDEILFRYNSERVIFSNVNLGATMDSRICIVGDNGAGKTTLLKIIMGILSPTSGMRNVHRNLKFGYFSQHHVDQLDMNVNSVELLQQTY
PGKPIEEYRRQLGSFGVSGDLALQTVSSLSGGQKSRVAFATMCMGRPNFLVLDEPTNHLDIETIEALGKALKKYTGGVILVSHDERLIRMVCSELWV
CGNGSVKSVEGGFDEYRKIVEQELEAAAQSK
>DvvABC-G D9811
MIGTEEITRYIDDDKIKINPSHNYTAPSSIPLTENVNMTDTKSLAINYNNNYRFKSTQENMSSVFPKKTAVDLNFEDITFYSTSWSITKFKKETKKI
LHGVSGQFKSGELSVIMGPSGAGKSTLLNVLAGYITKGSTGTVKLNDVVRDQSPRYRKLSAYIPQDEELRMALTAKEAMTFAAHLKLGYRVSNDYKF
KQIAGILKMLGLEECQHTLTAQLSGGQRKRLAVALELLSNPPILFLDEPTTGLDSLSCTQCVSLFKNLASEGRTVIATVHQPSALIFEMFDKLYALS
EGKCIYDGRVSDVVPYFEKLNLKCPPYHNPADYLIEVSIGDHNANINKLADSVKSYQYKSMESGENDDLKGRDLFEEKKVGLYSDETKAVPAAIIMQ
FLLLYKRNLLIIKRCYGPSLNRVLAHIVIGLIFGYLYRNVGSAADTVLANYVYLYGTLLLTVYTGKMPVTLSFPLEMKILSREHFNRWYKLTPYLLS
VILVEIPFQVICTWMYIAVSYWLTNQPLDFRLFLFVIFVTACSLCGQSMGYFIGATTPVKVAVFIAPVLACFLSVFGFCIRAIDTSTMFKPIFFISY
FRAAFQSVVYSIYGFNREILVCPIEEEYCHYKDPHKFLSEMDILDVDLVSNFILIVIVWCVMHAVTYLTLWLRLNKR
>DvvABC-G D3712
MATEGLLKNKSDVKLQIVSNQPKKLVHLPERPKVDLAFSDLEYVVKQGSKEKKILKNVSGMLRSGELCAIMGPSGAGKSTLLNILTGYKTIGVKGHV
LMNGSDRDLSQFRKLSAYIMQDNQLHANLRVDEAMAVAAALKIGAKSPKDREDIINEILDTLGLLDHKKTMTSGLSGGQKKRLSIALELVSNPPVMF
FDEPTSGLDSSSCFQCISLLKTLAKGGRTIICTIHQPSARLFEMFDQLYTLADGQCVYQGSTTFLVPFLASLQLQCPSYHNPASYIIEVACGEYGDH
TRTLVNAIENGKNDIRDASQIVGFKMSDGLNNAYQYAKDNLKTLISDELSKEGSTVNGNRKSNNVNEDTGIDSKDVEKANVDSALLNTSVVVKQPRY
GNSEFQQFFIILKRALLFSRRDWTLMYLRLFAHILVGFLIGALYFKIGNDGSKVLSNLGFLFFNMLFLMYTSMTITILSFPLEMPVLLKEHFNRWYS
LRSYYLAITISDMPFQTIFCILYVTIVYFMTSQPLDPARFGMFLLSSLLVSFVAQSVGLVVGAAMNVQNGVFLAPVMSVPFLLFSGFFVSFDAIPIY
LRWITYLSYIRYGFEGTALATYGYNRPNLDCFAQYCHFKKALTTLEELDMDKSSYWFDILALVVIFFFLRVSAYLFLKWKLKSNH
>DvvABC-G D14042
MADDSKCIKINVPLDPEDTNQAVAYTKTVHAPPLLKETESLNNASVLSGSQTNIYNGSSPITISGAGALRKVPNSSPSNHKRPMIALTHLPKRPPVD
ITFTDLSYSVSEGRKRGYKTILKCISGKCKSGELTAIMGPSGAGKSTIMNILAGYKTSNLSGQVMINGKERNLRRFRKMSCYIMQDDCLSPHLTVKE
AMMVSANLKLGKTVTLSEKKVVINEIIENLGLQTCIDINSSNLSGGQRKRLSIGLELVNNPPVMFFDEPTSGLDSSSCFQCLCLLKSLARGGRTIIC
TIHQPSARLFEMFDHLYMMAEGQCIYRGPVLGLVPFLSSMGLNCPSYHNPADYVMEVACGEHGDYVQKLVVAVNAGRCTKFATPDHRSSKIVSNDIA
KEANGKNSSGDVISVPNGSVKPTTPTTPVTCTTSLLDSSENLSPTEKNGFSTTGLQQFTILLKRSMYMILMDKTLTRMRLVSHFVIGCLIGLIYYDI
GQDAAKVTSNAGCLFFCVMFMMYTAMMPTILTFPLEMSVTVREHLNYWYSLKAYYMAKTLADIPFQVVMTLCYIIGVYFITSQPLDVTRFGMILLVT
VLTALVSQSFGLLIGAAFNIEGGVFLGPISTIPMVLFSGFFTNLNDIPFYLRWLPYLSYLKYGFEACMIAIYGLDRPKLTCNIEYCHFKYPKKFLEQ
MSMKDDMVSYFIDVGVLGGLFIFLRVIAYFMLRIKLMQNR
>DvvABC-G D10897
MEAVLHESTFVGQSKDAAAVVKSNGVNGDSVNFSVKVNLNQEKTLICDDGTNGKMCQNTVDLEFQNISFRATQGNIFTKRTTKEILHQVGGRFQPGQ
LIAIMGPSGAGKSTLLDVLSGYRIRGVTGSVYVNGQPRSLKEFRKTSCYITQDDRLQPLLTVDENMWVAADLKLPSSVSNSNKRSIINNILKTLNLE
NTKKVRAAGLSGGQKKRLSIALELVNNPMVMFLDEPTTGLDSSSCSTCITLLKKLTMQGKTIICTIHQPSATLFAMFDQVYVVGNGYCLYQGSTNKL
VPFLQDCGFPCPQYHNPADYVIELACIEHGPEKPEYMRERIQNGQSYQYFNEPEKMKSTICNYPGPKPIKRVvSSDaSSFPTSQFHQLGVLMRRGFI
KAYRDKMLTYLRIGTNIIVGLMLGSLYWKAGSDGSKVLDNFNLLFAILLHHMMSTMMLTILTFPQEMSILIKEHFNRWYSLKMYYTSVTIVDIPVSV
IGCFLFSLIIYYMTDQPNDKARFLIFFVTCMLVVLVAQGVGLTIGAYFDVVNGTFVGPTLMVPMMMFSGFGVRLRDLPTIMYYGSFTSYLRYGLEGV
VQAVYGMDRGILQCPEDKFCFYKYPKTFLEIVDVRSDQFDNDIIALLLFLFVLRIAAYVVLRYKLASVR
>DvvABC-G D22358
MELHSRNTSLRHLPKRPPVDIEFHDVSYTVPQGRKGSKLILRSVNGSFQAGQLTAIMGPSGAGKSTLLNILAGYKTQGATGQILINGATRNLKQFRK
MSRYIMQEDMIQPLLTVEEAMMIAANLKLGNTLSLSDKSSAIDEILSLLRLDKAKRTGTSRLSGGERKRLSIALELLNNPPVLFLDEPTTGLDDLSC
SQCISLLKKIAEGGRTVICSIHTPSAKIFSQFDNVYILSDGQCVYQGYGPEVVSYLSKVGIECPKTYNPADFIIEVCCNEYGNFQERMVSVIDNGRN
IYRTNGNLPQLQEVELANDSLENRSSISSISIREEVYADLSSSDFNHESSWFSQFWILTTRLWMQMWRDKTYLIMRTVLYIILALLIGSLYYKMGQD
GSKTIFNFGFYYCCIIFFMYIPMMPILLQFPQEIQLVKREHFNKWYRLSAYFSALSFSTVPVQLCLGVVYVSCVYLLTDQPLEFRRMSMFFFICILT

SVISESLGLLISAQLKVVNAVFMGPVSAVPFMLLAVYGFGSGYDTIPSIIKFFMHFSYLRYSLEGLIHAMLKDREKLSCPDTEEYCIYTDLNLFVRD
MGMENTIYWVDVlVLIFILILFRGGSYYLLRQRLTPNKTFRALQYIGRLVKSQFGLAR
>DvvABC-G D23081
MELQQFASNNLRISDKILDVEFQDISFVTKEKNGYKKIIDGVSGKFHCGHLTAIMGPSGAGKTSLLNILTGYQVTGTTGTIKCNSSSRKQKGVLQYK
KESCYILQDDSLPNLFTVEECMMIASKLKIANMAKKAREFLINEILTNLSLLKAKNTRCQSLSGGQKKRLSIALELVDNPPILFLDEPTTGLDSAST
TQCVDLLKKLANGGRIVICTIHQPNTQTYEMFDQVYMLAKGRCVYQGPSTNTVPFLASVGLHCPQYHNPADYIMEVVSGEYGDHIDQLAVAAQDKKW
QNIPTIKLSDTPAIDSKDNNIIYSDENVTLSKSPSEWKRFFILLQRSSVQLYRDWTISQLKLVLHLLVGLFLGITFQNCGRDATKVISNLGFLQVGI
VYLAYTSMMPAVLKFPTELVILKKESFNNWYKLTTYYAAFLVFDIPQQMLFSTVYCIGCYFVSDQPLEVdRFFSVLFVLVLASLSSSGFGLILGTIT
NPINGVFFGAVGLCFFISVGGFFIMFTHMSNVMYLFSYISYISFSVEGVMQAIYGYGRGQLHCPEEAEFCQYVSSEVLLEDIGMSKPNYWIDIIYLT
CTFLTFRTIAFVTLKRKLANP
>DvvABC-G D13051
MSEASGSAMEVLLPKFKPLTQLNTIAKKPPIDVEFCDLTYSISDSSYKGGWRQLLKSINGKFRSGELTAILGPSGAGKSTLLNILAGYMTAGVKGSV
KINGKPRDMRIFTKLSSYIMQEDLVQPRLSVRESMMVAANLKLSASIGHTQKVAVVHEVIQLLGLEKCYDTKTEYLSGGQRKRLSVALELVNNPPVI
FLDEPTTGLDNVSIKQCIDLLKKITRLERTVICTIHQPPASLFQIFDQVYIMANGYCVYNGSPNQLVPFMSSVNCVCPETSTPADFIIEVIQTNQDN
IPILQNQIQNGKINMKDKKLKPLQSHKTLGIYEIYQETTQTGMHIHDIEYPTSFWTQFTVLLCRMALQMKRNKSMWIIQFFHHVISATLVGGIFYQI
GNEASQVLPIFKYCVTINVFFVYTHVMVPVLLFPIEVKLLKREYFNRWFSLKPYFLASTIVNIPMLVGYGMIFISIVFFMTGQPIEWQRFLMFTMIA
INVGFCSQGLGYAIGSNCGILSGSVVAPCWLAILLALSVYGMGYKDGIEPVMKAFMSLSYVRYGLVGISSTLLNDRAEMDCNDIYCHYKNPEKLLED
MGMSQNVPLHQFAYILGYTVLFRIIAYISLKYRMTSELRNKLVYYAAKIVKQKET
>DvvABC-G T38769_c0_seq1
MGENTVCDILGTVDPPQESNKIHDEDFRLHSLELPENQVNDPASLINVGFRHISYSVYDGILFSRTKRKLLDDITGNFPGGQLTAIMGPSGAGKTCL
MNILSGYATTGILGEVLVNDEPRNNVAFRKQSCYIMQNDDLQPLLTVLESMRVAADLMLTASKTDKDKKIASILKSVTLWEIKHTRTDALSGGQKKR
LSVALELLRDPQVMFFDEPTSGLDSLNSIRLVKLLKEMSESGKTIICTIHQPSATIFKLFDHLYVVSSGKCIYQGSPNNLLPYFEDFDLICPTQHNP
ADFLLEVASGDYGDFTHALSTKSNNGINESYAKLSASTMKLNSIKEPNIGEDFLAATKRTIVNEIEYTWDECGLNSYPTSTINQFVVLTKRSFLMLS
RDRTLTYFRLGTHSAIALFLGILYFGVGLDAANINDNFSFMFFTVMFLMMTAFNCVVTTFPSELPIIIKEHFNKWYAIRSYYAAVSISDITVQILAT
VLYATITYLCTQQPMEWSRVSSFLFICILVSVISQSWALVVGSCLSVTNGVVIGPFFLLPFIMFSGYFVQLRDCPHQLKWMFDISFPRYALEGLVLT
IFGYDRGKLPCESKDFCLYVYPEQFIKDKDMENASYTIG
>DvvABC-G D13829
MAVYNRGGREEESSASESLSLAARLGLRVVRDYGTDFRSVGRVIPIEERMKFTWNDINVFAMTKTIKNKFCCFGFMSHGYDEHILKSVDGVACSGEL
LAILGSSGSGKTTLLNALTYLPMDGMVVSGMRCINGVPVDGQRLRNVSAFVQQDDCFITTLTVREHLVFQALVRMERGVSYNQRLIRVDEVLSELLL
KKCENTIIGSTDKGGISGGERKRLTFASEMLTNPHIMFCDEPTSGLDSFMALQVVQALKAMAQNGRTVICTIHQPSSELYVLFDKIMLMTEGITCFL
GTREDADGFFITMGAGCPRNYNPADYFVKLLSVIPDREESCRQAIALIADKFHNSNLGRRLAADSAYIRAQEEVDNAVWLSNNLRPYRNSCWAQFRA
VLWRSWISMLKDPLIIKVRFLQTVITSLLIGMIYYGQELNEEGVMNINGVLFIFLTNMTFQNVYAVIHVFTAELPVFLREHRSGMYRTDVYFISKTL
AELPFFIIIPVAFTTVCYYLIGLNGTLTKYFITCGIVILVANAALSYGYLVSCISRNTSMALTLGAPLVIPFLLFGGYFMNLGSLPHYLKWLSYFSW
FRFGNEALMINQWENITIDCSTNSTICPKNGHVILETYNFAEENFTVDIIALCMLILGFRFFAYLALLNKTCSC
>DvvABC-G T49457_c0_seq1
MMRSRTPSQGGGSQGFEMERKYSVPSNPESRAFSGGTTSEDLHAWSIYRQNLNSDFTDSALGSTDKSPLPYGNFQLRDTTVQSILSHPRYGPKSALG
SNMYTYLKFGLPRVFPPNHNGSNRSGTPQHFRRNSSTRPHNIRRSKAGSHGPRDGSSGYDSSDNETSHNYKQNRKYRSDPDFRMQNVYHAEQTSPGI
PLAAMHQGDIRHSNSQWNRNKSISEANLLALGYNRSCHNSERHLIDPRRNSVADYGHHNHDVVDHSDMGHMGRPMSKTESHFSVPHSRRGPPSILRT
DYLNQDDESGTTFMFPHLQAHGLGIFPSSQSCTKSRQHLLLNEISFEIRGGEIMAIMTTSEEEGTALLDIIAGFSSPALGTIFLNGHSVRAHTLKSR
VAYVQNDLNLCKDMTVVQTLRLHYDLKKPTEKLGYLKIESMDRINVLIDDLGLEQVRNTKVSMMTISERRRLNVACHLILDTDIVLLDQPTKSMDIF
DTFFLVEYLKQWASGGAGSTLGRIVILTMHPPTYEIFTMLSRILLVSAGRTMYSGRRRDMLPYFALVEYPCPAFKNPSDYYLDLVTLDDLAAEAMLE
SSQRIEQLAEIFRQKQEPLSDPGPPSSLPLTVRNCNCFVAAFALFTKSMIYTQPATFLSWLTVIVLSASLSLILGAIFWDIPSTDHQLILNDRYGYH
YSVMCIVHWPLLLAMTVNEVRRNRKVIERDIKDGLYGRVIYIITKSIINIFPSLFVWLIYVVPSYSMTGLYMQHLNNYDGFYIYIGVMLLYLSCIQI
FLMAFIYTVPLSNTATIFCGTVLSAFFLSAGYSLHLKDIPMYLQWIEKISPSEWLIPYLLNRELSTEAIQSLQGTITTLCRNKQIQHQDIIVQLPCP
PPNGTNSLKSFGYLKSDNLTFDYGNPVIAMGVFYCIFFVISCFMFALNLCRSRRRRRQDTKNDANKP
>Dvv ABC-G T36869_c0_seq1
MLSGGYILELCNVFYSGQVEKRSVLQRMVGNIKTAVILKDVSMLVHSGEVLAVLGSKGSGKKALLDVISRRAQGPIRGQIYLNNHPVSMCLFQQKCA
YVTHKCDFIPGLNVEQTLYYTPTKFTGYLKMSKVKQVIADLALSQVAKKCVEDLTKSEYRRLMIGVQLIKNPVVLLLDEPTWDLDPLNTYLIISILS
NAAKKYGTAIILTMEKPRSDVFPFLDRVLYLCLGDVVYTGGTKQMLEYFNVIGFPCPQLENPLMYYLCLSTVDRRSRERFVESNYQIAALVEKFKNE
GVIFQKSPSMISPNPNHEHGQHEKVPFMHGRPGRFSTGWTIYVRLLAATVSFKRAGLRQTFLRTFALPLYFFLMWLFYREMKDWQHTFISRNGLILN
CLCCVYFVGIINTILIYPIYRTRYYQDTQEGLYGGTLFLLTYNLVSLPFSFLSTVLSAVIIFPLIISFDNPVDFVYFTLILWACYIFAEQQTMAILM
IVKDYMKAAIFSIYLTVVCITLGSGILRSMKGLPEWLYYATYGTQARYAAAYLNRKVFTHSALSKALPFDLLHNCTQISFETSILNGANNAYCRYAS
GQAFLTERYSRDPTEVIFGVLEEDFNVGLTCGFALGMIVFNLFLYLIPLPSFIKAKFRE
> Dvv ABC-G T79525_c0_seq1
MLDVEETPLNLFQSHGSSGSYNEFDIYASASLAKKRQRTYSRWSPIEEGVTLVWENLSVYSSTTKNGQMQHKQIINGVTGAVKAGSLVAIMGSSGAG
KSTLMTALGYRTEGSILTEGNILINGRQIGDYMKYLSGFMHQEDMLLSYLTVREHMNIMANLKLDRRLSGNDKKQLIYDILRQLGLMKCIDLKIGGI
DQAKSLSGGEKKRLAFATELLTDPPLLFCDEPTTGLDSYSAQKLVVIMNQMAITGKTILCTIHQPSSDIFAMFSQLILVADGRIAYMGSTNNALDFF
ERMGYVCPTSYNPADFYIKTLSTTPGYEDNCRQTVKRICDQFAVSDEAKEVEIVVQYELHMGRVATQRKFELRQNFKEIRWLSKLFWLTYRWILEIY
RNPSLEAMKIAQRMLIGFIVGFCYLGTDALTQNGVQSVTGIIFMFVSENTFNPMYSVLHQFPSYLPLFLREYKSGLYHPATYYLSRILSLLPGFILE
PLIFVITAYWLAGLRTSGYAFLMTLSVIVLTMNVSSACSVMFSNAFDSVPTALAYLVPFDYAHMITSGVFVKLSTLPLVFSWIKYLSWLMYSTEALY
ILQWKDVTNITCVNPPEGFSCMNNGQDVLARYDFKETDFQFDIFGMLVLLLCFYILGYFFLWRRVRYH
>DvvABC-H D20789
MHDDYAVFVDRVVKKYGQKEVLTGICMKVERGSIYGLLGASGCGKTTLLSSIVGRKKIDGGEIWVLGGKPGEAGSGVPGPRVGYMPQDIALVGEFTV
KDAIYYFGRIFSMEDSLIAKRYRNLHTLLELPPDDRYLKNCSGGQQRRVSLAASLVHKPELLIMDEPTVGVDPVLRDRIWKHLVDITKKDNTSVIIT
THYIEECRQANKIGLMREGKLLAEESPTRLLTLFNSETLEDVFLLLSKRQEEGRLQELTSHRVVDDQNNSMLANDTTGSTTSVATSVSTFEIGHGST
DILAKKKILKARNALNKSRMKALFDKNLKQFYRNITGIIFLMTFPILQVGVFMGAVGGDIRSIPLGIVNDEAMSVTCPGFSFNGTATATDDRACQLR
NISCRFLSYLDHPMIEKVHFETLEDAKDAVLHGKIVGAMYMSSNFTSFLEERIDKGKDIEKDILSLSEIKVWMDMSNRQIGATLKYKLIDLYTKFQN
SLFDDCDFVPGFGDLPVNINFIYGDGDEPYTVFMIPGSLITIMFFMGAIMTSQIIITDRHDGVWDRSIVAGVTSLEITITHLVLQASICIIQTAELL
VVVYLIYQQEYSGSLWLMYVMVYLQGICGMAYGFWVSVISTDHSMANTVLTGIFLPMMMLSGLMWPTEGMPPALRIFSRCLPFTMAIESLRNVSKRG
WSIDNFQVYSGMGVGFMWTVFFGVLSVYLIKKKR
>DvvABC-H D5118

MFSWNNLVNVLGRLLGIYRRALSIEGTTMGDRVEPQVPGPDLVGVNVERPPLVHQQSTVWNRRQNAVSVRHAYKHYGSKKKPNHVLSDLNMTVGKGT
IYGLLGASGCGKTTLLSCIVGRRRLNTGEIWVLGGKPGTKGSGVPGKRVGYMPQEIALYGEFTIKETMMYFGWIFGMESKEIYERLQFLLNFLDLPS
QNRMVKNLSGGQQRRVSFAVALMHDPELLILDEPTVGVDPLLRQSIWNHLVQITKDGNKTVIITTHYIEEARQAHTIGLMRSGKLLAEESPQVLLQM
YQCTSLEDVFLKLSRKQVQVGAAANDQNMANNISLATLNWSKKESVSVTEESGVVGLNFHQSKEVLVPDSNGHLDMGKPPSSATHGIKEACDDCSSC
SDFTTTGKLRALLQKNFLRMWRNVGVMLFIFALPVMQVILFCLAIGGDPKGLKLAIVNHEKNFTNLTYQECEYEKGCKFGNLSCRYVDTINTTTIIK
QYYATPEEAKEAVRTGNAWGALYFTENFTDALVARMALGKDADEETLDQSEVRVWLDMSNQQIGIILQRDLQLAFQNFTKDIFRDCEYNEELAEIPI
SFKEPIYGSNQPSFTDFVAPGVILTIVFFLAVALTSSALIIERMEGLLDRSWVAGVTPSEILFSHVITQFVVMCGQTTLVLIFMIIVFQVECKGDIF
LVVVITILQGLCGMCFGFVISAICELERNAIQLALGSFYPTLLLSGVIWPIEGMPTVLRYISTFLPLTLATTSLRSMMTRGWSIGEPDVYYGFIATI
VWIVLFLTISLLVLRSKRG
>DvvABC-H D18290
MNTMNNHLALYVKDVEKSFGSNKVLDKLCLNVQKGTIYGLLGSSGCGKTTLLRCIIGRSNVNKGDIYVLGEKQDGSVVLADRVGYMPQDISLVGELS
ARDSIYFFGRIYNLKDDIIKSRCDELIELLDLPTDNRCVRDCSGGEQRRVSFAVTLVHKPELLILDEPTVGTDSILRNRIWKYLTKITEEDKTTVII
TTHYIEETRQADKIALMRKGKLLTEEHPETLLERYGCDLLEEVFVILSSKQEETEQTSGVDNLAFQSSQTGINESTERTNAALDVISEAPQQTRTER
KKQSIFEMIRFKSLLFKNYKKLKSNLISTIFLVVFPALQVLFFLTAVGRETKGNPFGIVNNERSQEFCKTFPRNESAVPYDKFSCHINSASCLFEDY
LDVPLFYMIRFDSKYEALSALESGKIRGFLHMHQNLSELIEARVNYGVDKYDLPDEEMQVYMDMTSIQVGGIIKERIADKYQLFHEKIYDDCRYLSK
AVQIPLRVEESFYGNKNEEYVIFITPGIITSMIFFMGVLMTCVIILDEKCEGIWDRSIVAGVTALELSLSHLLYQNVFMIIISLDFLLFVFVIFGQP
YIGSLWLIFALLYLQGLAGIMLGFAISVLSDNHVTANTIVTGIFNPLVLVSGILWPAEALPKYIYFVIVWLPLAIPTSALRNVIRKGWTLTNLHTLH
SIGILSFWILFLLALNIYLTKTKR
>DvvABC-H D11818
MERPSDIELRRRQFKTQASTLNSRRQQAVCVRRAVKKYGTKSNPYVILDNLNMTVPKGCIYGLLGASGCGKTTLLNCIVGRKRLNSGELWVLGGTPG
SRGSGVPGPRVGYMPQETALNGEFTIRETLKYFGWISRMTTDEVEAKVDFFINFLMLPDADRQVKTLSGGQQRRVSLAATLLHEPELLILDEPTVGV
DPLLRTNIWNHLVEITKYGRTTVIITTHYIDETRQAHLIGLMRGGYFLAEESPERLITQFGVDSLEDVFLKLSVMQNMGKRRRSSIAKSVVETITVP
ELGGAVNPAAVLDDDLGEISGEFGDSVSVTSRTGRRVSIAPEPTAEHIIPELPPDEEPEVSFKDYFKFVQSHHMRALIWKNFLWMWRNIPMMMFIIG
LPISQTILFCWTIGHDPQGIKVSVVNNEINFPDETCHQQPLNCNGTKISCNYLNEVAKSYSITWEFMKTEEEARHRVERGKSWGVVVVPHNFTDALW
SRIENSRDTPPEDIMASTISVYEDKSNENIATFLTRDMLYGFLTFISEFVKSCGYNEKSVGVPIRYNEPIYGYQDPDFTDFAAPGVVLTIIFFMAVA
LTSGSMIMERNEGILERSLVNGISGTELLFSQIITQFVVMLGQSIAVLLVAFILFGLTQRGDWTTVTTLTVLAGICGMCFGFVVACSCENERSATYL
AMGSFLPIVMLCGIIWPIEAMHKYISWMSVVLPLTQATESLRCILARGWPISNPTVYNGFISLTIWIFIFLTISILLLKFKKG

DvvABC Cdna sequences

>Dvv ABC-A T49125_c1_seq2
CAATAAACGCTCATACAAACTTTCAAAACAACCAGTACAGTACCCCTCACTCCAGAAACCCCCAAAAAAACTATATGAATTTCGCACCTTTCCACCA
GGATAGCTCCTTAATTAAAAATTATTGAATATAGAAAACAACTTGCAAACTGAAAGATGGAAATCTTAAATAAAAAATTATTTTGAAGTGCGTGGTT
TGTCCAAATATAATTTAAAAAAATGCGATGCCACAAAGAAGTTTAAAACAATATTTGGTTTCTCAACTACACAGGTGTCCAGTCTGTGTGCTTACTT
GTTTTGTTAAATGTAACTGTTTGCAAGAATAAATATTAAATGTTACCAAAAAAATAAAAAATCACTGTATTCTACTATTTCTTTTTGGTAACATCTT
CTTCATCTTCTCTGGCAACTTTCAAAAAGACGTCTTCTAACGAAGCTTCACTGATTGCATAATCTTCAATTAAATGCGAATTGTGTGTTTTGAGTGC
CTCCACTTCTTGGAAAATATCGCTCCATTTTTTAGTTTTATCCTTAATGTAAACATGGAGGAGACCTGAATGTTCATCTTTGATTTCACCGCGATCA
TGATTCGTAAAATAGGATTTAAGATCATCGACACTTTTAAATTTGTTACTTATTACTCCATCTACCTCGTCTACGTCATCATCTACCAGATTTATAG
GCTGACCTTGATTAGACACAAGTTTTAATTTTAATGTAAATCCACCATATCTATTCTTTAAGTCCACCAATCTTCCTTCTTCAACAAACGAACCTTT
CTTCATTATTTTCAGCTCGTCGCATAAAGCTTCGCATTCAGTCATACTATGAGATGTTAAAATGAATCCAATCTTGTTTCCACTATTGCCTTTCATG
AGTTTAATCAGATCCCAACACTTCCTTCTAGTAGCCGGATCTACGCCGTTGGTGGGCTCGTCCAATAATACAAACTTACGGTTACCTATAAGGGATA
TGGCCAAGCTTAGTTTTCTTTTGTTTCCTCCACTATAATGTCCGCACGGGATATCCGCATATTGTTTTAAGTCAAAAGCTTCCAAAAAGCGATCAAT
CATTGTTTCATCAGTAACACCTCGTAACTTTGCCACGGTACTTAGTAGTTGTCTGCCAGTAAGGACAAAGTTTAGCGAATCTGACTGAGGACAATAA
CCTATTGTCTCCAAATACTCGCTTCCTGTGATGTCCAGAGGTTGTTTATTATTGTCTAATTCGATCTTAATTTCTCCATCATCAACCACTTCCTCTC
TTGTAAGCATTTTAAAGGTTGTAGTTTTACCAGCCCCATTTACACCTAATATACCCAAACAGTGACCGTGCTTAAGAGTGAAATTTATATTTTTAAC
TATCTGTTTTCCGGCATATAATTTCATTATTTTCTTGGCTCTTAAAGTATTGTAGCCGTCACCAACATCTTCGTTAACATATTTGTCATTTTTGGGG
ATTAGTGCTGTATCTTTGTCTTTGAATGATTTCAAAATAAGATTACATTGACGCATAACCATAGCTTTTAATTTCTTCATAAAATATGTGTCTAGAA
TAATGTTTATCGTTAGATAAAAAACTCCGCATCCTATCATAAACATGTAATGTTCGCTGAATAAGTTGTTGTAGCTCTTATAATTTTGACATTCGAT
TGAACTTTCAACGCAGCAAGCGTGTGACTCCATCTGATTACACATAACTCTCAAACGTTGGGGAGATATTTTGAAATTGTAAATATTAACTGTGCGT
CTAATAAACGAGACCAATGCATCTGATAATGAAACTTGCGGTAATAAGAAGAAACACACATATTTAATTTTATTTCCAATATTTACGTAGTATTCAT
CTTGCGATTCTAGTAATACGGCTATGGTTAGTGTAAGAATGATGCCCAAGAACATCCCCATGATGACGAATAACGCAAAGGCACCAGATGCTGTCTT
CTTTCGGCTGAATATATACGCAAAAGGTATCCCAGCAATACCATAACACAGAAATATGGAAAATAAATAACCGAGTTCGGTACCGTTATGTTGGAAT
ACGGGAGAACATATCCACAGTAAGAAACTTAATATGAACATTAAAAATATGTAAAATGTCATATCGGCAATCCAATTAACAAGCCAATAGAAGAATG
GTTTCACGCCAGCCATGTACTGTATTTGTAAAAATCTCGTCGATATTTCAGTATGTGGGAATATGATAAAAATGCCGTGAATAAAAAGACATCCCAC
CGGTAACAGGATGAGCCACAGAACAGCAACTTTTGCTTCCGAATACTCCTGTGATGAGAGTCGCTCTTGGACCTGCGGTAAAGGCCAATTACTGACA
GATATTCCATATTCGGGACCTAATAACGTTTTAGCCAGCGAGTCGGTAATTAGATTTAAAGAAATCGGAGTACTGTGAGTAGCTATACCATTGTATA
ATGCTACTAATGTAAAATCACGATCATAAACGAAATTAGCTCCAGCAATCATATGTTCTTTATAGTAAGCTATGTTTTCCAGACTTCTTTGCAAAAT
AGATTTTTCAACGTTATCAACTTTGAATGGTAAACTTCTTTGCGATTCTACAGCTGACATGTAATATTTCATTAACGTATTAATTTTTTCGTTTGAT
GTGTCCCCGCTGTAGTAAACCGATGTTTGTTTGTAGCTGCTGAGTTTGAGGGGTAGTTCTGGTCCCCTTTCACTATAGTAGTCATTTCCAGAACTTA
AAAATACCGTCAGAGAAAAAAGGACTACAGCTACACCCACAGGTACAATGTATGTTGACCACTTTTTTAGAGAGAAATTCACTTTCTTTTTGAGTAA
AGTAATAAGAGTCAAAGAGGCATGAGGTTGGATGTCATCTTGTGTGGCCACTGATGAAGTGTCGTCTGATGAATGTTCAATCTCTTGGCGTGTTTTT
AGGAATACATCTTCCAAAGTCGTTATCGTTATCGATATATTACTAAGGTGTAACTCGGATCGATTTTTCTCCAAATCTCCAAGAAGCCCCGTCATGT
TGGTATTCTGATAGGGTAAAACAAATACCAAGTTGTTTCCATTATTTGACGTCAAATGCGCATCGGGCATAAAATGTTTAACTCTTCTACTAATGGC
ATCTATATCTGCGTCTTCTTCAATCATTAAATTTAAGTGGTAACCTGTATCGTATTTCTTTTTCAAAAACATTGGAGTACCATAACAGTTTAGACTA
CCGTCAGACATTATTGCTATCCAGTCTCCTAGTGCATCTGCTTCTTCCATAAAATGTGTGGTAATAAGGATTGTTTTCTCCCCTCGCCATTGCAACA
ATAGATCCCAAAGCTCTCTTCGAGATTGGGGATCCATTCCCGAAGAAGGTTCATCCAAAATTAATACCTTTGAATCACCGATAAGGGCCATTGCCAA

GCATAGCTTTCTTTGCATACCACCTGAGAGTGTCTGAGCCATCGAATGCTCTTTATCGGGCATATTTAATTTTGCCAATAAACTTTTTGCTTCTATG
TTGGCTTCTTTCGTAGATTTTCCTTTCAGCTTGGCGAAAAATAAGAGATGTTCTTTAACTGTCAAATCGGTAAACAGAAGATTATGTTGGGGACATA
GTCCCAAACTCTTCCTGATTTCGTCCATGTCAGTATTAATATCCAATCCATTTATTGTCACCGAACCAGACGTAGCTGGAATCAGTCCTGTTATTAT
TGACATGGTGGTTGACTTCCCTGCTCCATTGTGACCAAGTAGTACAGTGATTTGATTTTTGTAAATGTCTAAATTCAGGTTATTGACAGCTTGTTTT
TGTTTATACCTTTTATGAAGATGGTTAATTTGAATACCCCTTAGCTAAATTGTGACCTGCTTCAACCTTCTCCAAAGGAACTGTTTCAGAATCTACTG
GCACTGTGGTGATAGATGGTTTTCTACAGAGCTTCATAAAATTTTGTAAAGGGAATAATATACTCTTTCGGACCCCGTATTTACCAGGATTAACTCC
ATCCATATAAAATGTGAAAATGGTGTAGATAACAATATCGACTATTAACATCACATAAACGTTTAACATGGTGATATCATCTGATGATCCACTACCA
GGCTTGAAGAAATTTGACCAATTGATGCCGACTTCTCTTTCTTCAAATACAGAAATGATGGAATATCCGTAATGAAGGGACATATTTGGTAGTAGGT
TTAACAGAACGTTGGATAACCAAGACAATTTGTTTGCTTCGTCTAGTCCACATGCGTATTTTGGTATGAAAAATGATAAAATCCATACGAGAATTCC
TGCGACTAATGCAATTGTGGGTCGGCTGAAAAACGAACTGATGGCAAAACAAAATGCTACTGCAGCCATACAATACAATATTAGAAACACGAATAGT
ATTCCGCCGTTGGCAAATTCTACAAGTGGAGAGTCAGACCCAAACATTGGTACTTTCATAAAAATAGATATCACACTAACGGCGAAAAACATCGGTA
TCATGGCATAAATAAACCAGCCAAACCATAACATCCAGGATTTCATTCCTACCATTTTCATTAGTTCCTTTGTTCCTGAATGTTTCTCTTCAACAAC
TCTCTTAAGTACCGCTGGACATAAAAATATAAAACTAAACAATGTGATCAGCGGCAAAAATTCCAAAAACAATGTTGTTATAGCAGAGTCTACTTTA
TGGGGAGGGTATGGAAACTCTTGTATTTGTATGGAGTAAGTGTCCTCTATATTGCTACTTCTCAAATTTTCCTCCAAGAATGACATGTCGAGGGCTA
GTTGAATTTTCAAAAAAATCTTCATCCCGATAAACATACTCACTGCCGGGGCTATAACCGTAGACTCTTTGGTACATGCGATCGGTGACCAAAAACT
CCGGCTGGTAGTAGTTTTTGTGATACCTTATCTTGTAGTTAAAATGTTTGGGGTTTTTGTCATCAAATATAATTGCAACTGTTGTATAGTTTGGATT
TTTATAGTAGTATCCCAATAATTCATTTTCTGATCCAAACCTTTGTATATGATCACTAGGCATCCCCAACTTGGTTAACATTTTGGTAGTAATGTTG
GTATAATATTCATTATAGGGTGTGTATAATATATACGTTCCAGCCACATCTATGTTATAATTTATACCATCTATCCTGCCATACGTGGGATTGGGAA
CTTCTATTTTGTTCAAGCCTGTGATTTGACTTCTGCCGTAGGCAATCAATAAAAATAGGCATACTGGCAGGAATGATTCTATAATTGTTAGAAACCA
ATGCCTTTTTCTTATTATGTAGTTTTTCCAAACAACGACTTCCAGTTGTAAAAGACTCATTGTTAAAACCAAAGAAAACACCCTCCTAGAAAACCAT
AATACAAAAATGGATTTTGTAGGAAATTTTTTTTTGATTTTTATTTTAATGTTTATAACACATACACTTTCGAACTTTACATGACGCGATTTACAA
AATAACTAATGATAATTATTGATAACGATAATGAACATGACAGAAATATAATTTATTTTTTTATG
>Dvv ABC-A T50718_c0_seq1
TGTTTTATTTATATTTCAATTATTATTCTAATTCTTCTTCTTAACATCGTCACCTTCTGGATTTTGGTGTTTTGTAAACGTTAGGAATACCTGTTCT
AGACTACTTTGTCCCAAAGAATAGTCTTCAATATTCAGATCACTTCGCTTAGCTCTTTCTAATATACCAAACATAGTGGACCATGCCATGGAAGTGT
CGGTTATATAATAATACAATAGTTCTTGATGTTTCTCTCTAAGATGAGCATAAGGAAATTTATCTTTGATATACTTCTCCAGTGACTCTGTATCTGC
ATGTACCAGACCACCGCTCTCTGGCAATTTCTTTAGTTTTATGGTTAAGGTGTATCCCTCCGCAAATTTGTTTTTAAGATGCTGTGTGGATCCAAGA
CACTTGAAATTGCCGTTCACCATAATGGCTATTCGAGTGCACAAGGCTTCACATTCTTCCATGCTGTGAGAAGTTAGAACGATGCACTTGCCGTTGT
CTCGAATTTTGCATAATGAATCCCACAGGTAACGTTTCGTTGCTGGATCCATACCTGTTGTAGGTTCGTCCAAAAATAGTACTGGCGGATCCCCTAT
CAACGACAAAACAGTACTTAACTTCCTCTTATTTCCACCGCTCATCTCTTTAACTTTCTTGTCCAAATGACGATGAAAATCGAAGTCTCGAGACAAG
AAATTCGCAATTCTTTGGGTTCTTTTAAACTCTATTCCTCTGAGTAGGCAGTACATGATAATCGTTTCCCTTGCTGTCATATCATCCAGCAAAGCGT
CGAATTGAGGACAGTAGCCGATGTTCCGTTGTACTTGCTTCAGTTGAGTTTTTACACTCTTTCCTTCGATCCATGCATCACCGTAAGATACAGTTTC
ATCTCCGCTCATCATTTTAAATGTTGTAGTTTTTCCTGCACCATTAATACCCAACAAACCAAAGCACTCAGACTTTTCAATTCCAATAGACAACCCA
TTCACAGCTAAAAAGTTGTTGTAGTATTTGGTTAAATCTTTCACGACTAAAATGTTAGTTTGACGTATATCATGATCTGTGGCCATGTGAATCCTTT
CTTTTTCTTTGGATACGTCGTCGTCTTCGTCTTGAATGAGAATTGGTTGTTTCGGGAAGAATTTTTGGTTGATGTAGTACATCAGTCTGGCTATGAA
TTTGTATTCGATGATAAATAGTAATATTGACCACAAGACACCAGAAATCAGTGAGTACGTTACATTTTTTGCGATACCATTATTGTTCCATGAAAAG
TAGTTGTCGTTAAGATTGGTACAAATATCAGCGACAGTGGATGGAAATTTTGAAATGCACTCATCTTTAGTAAAATTATTTTGAGCACACGAGTCAA
ATACTGTTTCACAAAGAGTGTTGTATGAATAGATAGCGTAGCATTTGTTTATTCCTGTAGCTAAAGAATAGTGCGGAAACACAAGAAAAACGTTATC
CAAGGTTTCTCCTATACGTTTTAGGTTCGTGCTTTGCGTTTTAAGCACTTCAACTAAAAGAAAAGCTGCATTCCCTATAAATATACCCACGAGAGTC
ATCCTTGTGTAACCCGTCGATGGAACTGTGAATATTAATGATGCTACATAATAAATTGGCATAAACGACCAACCAAACAAAAGTAATAGGAAAAATA
GCCGACCCATATCTCCACTTGTCTTAAACCCATTCTTGTTGGAAACAATATATAGTGATCAATAATACTGAACATACCAGTAAGTAGACAAACATATC
ACAAATTATATTAACAAACCAAAAAATGTAAACCTTCACTCCTGATACAAACTGCAGGTGTTTTGACTTGGATACGTTTTCTCTTATCAAGAATATA
ACATAAAATGATGAGATAAACGCCATACTGAATCCAATATTAAATGCAAACTGGAATCCCTGAGATTGTCCACCAACAATGTTGTCTATTTGGGTTG
CAGCTTTAAAAGGTAATGGTTTGTTAACAAATTCAATATGTCCGTTAGTGTAGTGATTGTAAACGGAAGACAAAGCCATGCCTAAGGCAATACCAGG
TGAGTGGTAAGGATCATTGTTAAACCATGCCGTAATAACAGGATGAATTGTCTTGATGATACCAAATTCAAATTCTTGCTCTGTGAAAGTTGATCCT
ATAATGTAACTTCTACGCACTACGTTTGGATTAGTTGCAGTCAGGTCTAACATTATACGGGTAATATTGTCAATATTTGTTGGAGACATTCCATATC
CATTTAAAACACCGTAGAAGTTATTCGTATAATTTTGTGGATCGTCAAATCGCTCAATTAAAGTAATAGGGTCAGTGAATCTTGCTAAATTTAAAGG
CATTGCCGGCAAAACAGAGGGAACACTTCCTGAAGGTACACTCATAACTATAATAACGTTTATCATAGGCAAAAATATCTGAATCGCCAACAAAACC
CATGATCTAAAGGTTCCTATAATCTTCTTCATAAGCATTGCGACGAACTGATTGACAATCAAATGTAAACCACCAGTGTACGAGGGAGATAAAGATT
CTGGACCTAAATCAGCAGAAGTGAGTATAGGATTTTTTTCTTTTTTGTCCTTTTTCTGTTTTTCATTGTAAATTTCTTCTTGTCCATGGTCTGCACC
AACTTTCATGAAAACTTCTTCCATAGTGGTTAAAGAAATACCATAGCTCCTGATTCCAAGTTCATTTGATTCTTGTTCTATTTGTTTTAACATTGCT
TCGAATACAGGAGCATTATTTTCAGATAACAAGTAAGTTAGTTCACTTCCAACATTGCTATGGATCTCTATCTCTGGTATAAACTTCTTCAGCAGTT
CTGTAACTTTATGTGGATTGCAAGTTTTCGCTTTATCCATAATAAGACTGTATCCCGCACCGTACTTCTTCTTCAAGAAGAAGCTGGATCCACAGCA
TTGTAATTCTCCACCAGCCATGATTGCTATTCTATCTCCTAAAAGATCAGCTTCGTCCATAAAATGTGTACTGAGCAATATTGTTCGACCCTCTTTT
TGTTTCTGTAATAATTCCCACAACACTCGTCTCGCTGATGGATCCATTCCAGCAGTAGGTTCGTCCAACATAACAACCTTCGAATTTCCACACAAGG
CCATTCCAACACAAAGTTTTCTTTTCATTCCCCCTGATAGAGTACTGGATTTCTCCTTCCTCTTTGGCTCCAGATCAAGCAGCTTGACGTATTTATC
TATTTCTGCTGTTATTTTGCTTTTAGACAAACCTTTGAGTTTGCTGAAGAAGTACAAGTGCTCTTCGACGGTGAGCTCGTCGAAAATGATGTTGTGT
TGTGGACAAAGGCCCAAGCTATTCCTGACTCCTGCGATGTCTAACCGGATATCATGACCATTGATTTTTGCTGTGCCGTTTGTTGGGCTTATCATTC
CTGTTAACATGGACATCGTCGTAGTTTTTCCAGCTCCGTTGTGACCCAACAAAACTGTGATTTGATTTTCAAACATATTTAAGCTTAGATTCCGTAC
AGCCGTTTTCTTGCCAAAAGCTTTTTTCAGATTGAATATCTGAATCCCAGGTCTAAGGTTTGGTTCGGTTTCAAAGAATTCATTTGCTGAATGATCA
TTATAGTTATTATAATCTGATACAGCTCCTTTTCCACACCAGTATTTCTTTGTGAATGGGAAGTACCATTTTTCAGCTACTCCATATTCTCCTGGAA
ATACTGCCTCTATGTAGAGAGCAATTAATAGGTAGATTATTGTGTTTATTATGATCATTATCCATATTAATCCCAGGCTCAAGGTGTCATCTGGGGT
GTTTGGTTTAAACAGAGTGCTCCAGGTCACACCATCTCCAGTTCCTTCATACATAAGTACAACTTGAAATCCAAAAGCCATAGCAGTATTTGGCAAC
AAACTTGTTAATAGCTTAGTAGATAATGCAAGATCAGTGTAGTTATTCTGCATGAAGAGATAAGGAGCATAAGATATGAACCATGCCAAACCTGCCA
TAGTTGCTGCAGTATTTGCTTTGGAAAAGAAGACGCTCAGAGCAAAGCAAAAAGTGATCATAGAGCAACAATATAGCAGGAGAAATACAAACAAAAC
GAATGGGTCTGCATAGGTGAATACTGTGAATTCTGTATTCGAGTACCACCTTACTTTGAGAAGTACTACCATAAGTGCTACCGAAATTATAACAAAA
ATGAAGCACTTGACAAACCAGGCCAACCAATGTAGCCAATTAGGTAGTCCCATAATTTTCATTGATTCCTTTAACTGTTTTTCTTTTTCTGTGGTGA
TAGATTTGACTGTATTTATACAAGTGTAGACAAAACTAAGCATGACAATCATTCCAAGCATGCTCGTAAGTGCTGTTAGAAGTTGATCACTGAACCA
ACTTGCGAATGGAAAGCGTTGCATTTTTATGAGAGGAAATGGATTAGACTTGTCCATCAAAAAATTAAATATTTCCTGTGGATCTTCGTGTGTAATA
CCCTTATCTGCTAGAAGAAGAGCAGTAGTCACAAAATGTTGTACGGCCAGAAATCCTTCAGCAAAATAATTTGGTGCTGATCCAGTAGCATTGTCGG
GGGATCTTGGACCAGAGCTTTGAAATACGGGGTACACTAAGTTGGTCCTCCAATTTTGACGTCCAATCGAATCTAATTGATACACTGTTTCTCCTGG

AAATCTTATATTAACCTCTACATTGTCTCTATCCCCTCCAAACTCGATACCATAATAGGATGAGTTTCCAGCTAATGATATTCTTTCTAATTCGCCT
GAGCTTATTACACGCTGAGTGTTGACTTTTGGTATTAACTTAAAGAAGCCTATTACATCGTCGTAGTATTTATCTTCTGGAGCAGTATATAATATTG
TAAGGCTGAAATTGCTATCACCAGTTTGCATACTACTAAATAATGATGGTCCTGCTTGCACAGCATCTGGGTCATTACAAAAAAGCGGTAGGCTAAG
GTTGAAGTTTCCCTCGAAAGCACAAAATGGCTTATATATTCTTTCATGTTCATATGTCGGATCCGACAAACTTCTGATGTAGACTAAAAGGAATGAA
AAGAGAACAGGAGCGACAATTTCTATTATGGTTTGAATTGGTTTTCTATATTGCAGCAGCCAATTCTTCCACATTAACAACAAAAATTTATCCAATT
TTTCTCCCATTATTCCTTTATTACCGGTATTAGTGCAATTTGATCAGATCTATTCACTCACTACACTTCAC
>Dvv ABC-A D18330
CTAAAACCATCAGTTCAAGTTCATCTTTGTGGTGGAACATGTATACAAGACAGTATTAAAGTTTATGTGATTATAAATGACATAATAGTTCCTATTA
TAATTTATGTGAAATCGGTAATCGGTAAATATAACAATGAGTGCGGAGTCTAATTCAAATAGAAAAAGAAATCATAGCTACACTGTACTGCAAGATG
GATCTGACATTGGATTTCAAGAACCAAGAGGTGCCATGGCGGAGAACCTGGACAAGTTTATTCTTCTTATGTGGAAGAATTGGTTGCTTCAGTATAG
AAAACCCGTACAAACTGCGGTAGAAATATTGGCGCCGGTTATCTTTTCCATCTTACTTGTAGTAATAAGAAGTCTATCGGATCCAGTTCGTCACGAA
ACTGTAGTGTATCCGCCCTTTTGTACCATTCCTATAGCTCTAAGAGAGAACAAAACTGGGATCACGATTTGCCCTTCTTATGATGGAGTACCTTATA
TGGACAGAAATCAATCAGATGGCGGAAACAGTTCTAACAATCCATTTGAGAAATTTGCACTTGTCTACTCACCGTCCAACCCTCCAATAGACCAAGT
GATGAATTATTTTCGACTCGCTTTTAAAAATGTCGTGGCCTTGGAGAGCTCTCAGGCTCTGGAGAAATATTTCATAACAAATTCGTCCAACATTACA
TTCGCTGGTATACAGTTCGATGATTCTTATAAGACCTTAAAAAGTTTAGACGACGTTAAGCACCTACAAGTATCAATAAGATTTCCAGGAGAAACCA
GGTTAAAGTTAGATCCTTTTAATTACAATAACTGGAGAACGAATCTCATTTTTCCAATATATCAGCAGCCAGGACCTCGACTTTATAACTTGATTAC
AGGAGCAGCTCCAAGTTATTACAGGGAAGGTTTTTGGGCCTACAATATTACCTAACACTTAGTGTTTTACTGGCGAAGAGTAATGTAACTTATGCT
ACAGATGATTATCTGGATCTTATATCCTGGGTAATAGAAAACCAGTTTCCTTTAGTCAACATGCGAAGGTTCCCGAATGCACCATGGTACGAAGATA
TTCTGCTGACAGCTTTGAAGTCTCTGATTGGTATTATAATAATGTTAAGTTTTGTCTACACTTGTATTAATACTGTCAAGTCAATAACCACGGAAAA
GGAAAAGCAACTTAAAGAATCCATGAAGATAATGGGTCTCCCAAACTGGTTGCACTGGACAGCATGGTTTGTGAAATGCTTCATGTTTCTTTTAATA
TCATCAGTATTGATGGTAATATTTTTAAAAGTAAGATGGTACACCAATACGAATTTCACTGTATTTACCAAAGCTGATCCGTTTGTTCTACTTTTAT
TTTTAATGTTTTACAACTGCGCCACTATTACTTTTTGCTTCGCACTTAGCGTCCTGTTTAACAAAGCTAATACTGCAGCTACTATAGCAGGTATGGT
CTGGTTTTTATCATACTCACCATATTTATTTATGGCGAATGTATATGACACCTTAACGTTAACTTCGAAATTAGTAGCAAGTATTGGATCAAATACT
GCCATGGCATTTGGTTTTCAAGTAATTCTGATGTACGAGGGAACTGGCGAAGGAATTCAGTGGAACAACATTTTTACACCAAACACCCCTGATGATT
CCTTAACTTTAGGGTTAATTCTAATTATGTTAACGGTTGACTCAATTATGTACTTATTAATTGCATTATACATTGAGGCTCTTTTCCCCGGAGAATT
CGGAGTTCCACAACCTTGGTATTTTCCATTCACTGCACAATATTGGTGTGGACACCCTATATACAGAGGTGTCGAAGATTTTGATAATGGTGCTATA
AAAGGAGAATTTTTCGAAACTGAACCAGAAAACCTAAAACCAGGAATTCAAATCCGAAATCTCAAAAAAGTTTTTAACCAAAAAGTAGCTGTTAGAA
ATCTCTCTCTAAACATGTACGAAGACCAAATAACAGTACTTCTAGGACACAACGGTGCTGGTAAAACTACAACGATGTCGATGCTAACAGGAATGAT
CACACCAAACGGAGGTACTGCAAAGATTAGCGGGTATGATATAAGGACCGATATGGAAGGCGTAAGGAAAAGCTTAGGACTTTGTCCGCAGCACAAT
ATTATCTTCGATGAGCTCACTGTGGCTGAGCATATTTACTTCTTTAGTAAGCTCAAAGGAATGAGAAAGGGTGAAATTAAGTCTGAGATAGCAAAAT
ATGTAGATCTTCTCGAATTACAAGAAAAGAGAAACTCCAAAGCATCGACTCTCTCAGGCGGTATGAAAAGGAAACTCTGCGTGGGTGTAGCCCTTTG
TGGTAATTCTAAAGTTGTTATGCTTGATGAACCAACTGCTGGTATGGATCCTGCTGCAAGAAGAGCTCTCTGGAATTTGCTAGAAACACAGAAAGAT
GGACGTACTATACTTCTTACAACTCATTTTATGGACGAAGCTGACATCCTTGGTGATCGTATTGCAATTATGGCTGGTGGACAGCTGCAATGTTGTG
GATCAAGTTTTTTCCTCAAGAAAAAATATGGTGCTGGTTACAGTCTTATTATGGACAAGTCACAAGAATGTGACCCTAGACGAGTTACACAGCTGTT
GAAAAAGTATATACCAGACATTGAGATAAATAGTAACGTAGGTTCAGAATTGACGTATCTGCTGGTAGATGATCATGTCCATGTATTTGAACCTATG
CTTAGAGAACTTGAAACTGAATCCGAGATGTTGGGAATTAGAAGCTATGGTATTTCGCTGACGACACTAGAAGAAGTATTTATGAAAGTTGGAGCAG
ACCATGGTCAAGAGGAAATGTACAATCATGAACATGGCAACATTGTTCAGAATGGGACTGCACAAAATGGATTTACCAATGGGATAAACGGGACACA
CAAAACTAACAATGGTACTCACACCATGGTTCCGACATATTCAAACGGATTCTCTCTATTACTGAACCAGATTATAGCCATGTTATTGAAAAAGTTC
GTTTCAACTGTCCGCTCGTGGATCTTACTGGGCATACAAGTAATGATGCCCACTCTGTTTCTTATCATTGCCTTTGTGGTAGCCCGCAAAAACAAAA
TGACAGGAAATCTTCCTGCCATGCCTCTAAGCTTAAGTAAATTTGAGAACCCTGTGACTTTAGTAGAAAATGGGACATCGGATTATTTGCCTTACTA
CATGAAAGTTTTAGAGGATTATGGATACCCTGCAACCATCGTAGATAATATTACTTCAGTATTATTGGATAAGACAGAACATCACCCAATTCTCGTT
ACTAGAAGATACCAAGCAGCAGCAACGTTCGGCGAATCTGGAATACCTGATCTACCATTTAACCTCCCGAATTTGACAGCTTGGTTCAATAACAACC
CCTTTCACTCACCAGCGGTATCTTTGTCTTTGATGCTAAATAGCATTTACCGAAAACTAGGCGGTTGTGATGACTGTACTATAGAGTTTACCAACAG
TCCTCTACCTTACAGCGCCGCGACGCAAGCGTCGCAGTTGTTGACAGTACAAAACATTGGATTCCAGTTGTCGTTTAATATAGGATTTAGTATGTCA
TTTGTTGCTTCCTTCTATGTTTTATTTGTTATTAGGGAGAATCGTTGTAAGTCTAAGCATCTGCAGTTTGTGTCAGGAGTGAAAGTATACGTCTTTT
GGTTAACTGCAGCCTTCTGTGATATGTTGACGTATCTGTTTACAGTGTTTGTACTGATGATCACTATGGTTATGTTCCAAGAAGATGGGTTTAAGTC
AGGGAGTGATATAAGTCGAATGTTTTTCATCCTGTTTTACTTCGGATGGGCATTTTTACCAATGTTTTACCTAAGTTCGTATTTCTTTCAAGTTCCT
TCTACGGGTTACACTAGAATGACGCTTGTTAGTATCTTTGGAGGAAACGCTGCTTTCCTTGTGGTACAAGTCCTTCAAAGTCCAGGATTAGACCTGC
AATATATAGGAAATGCTTTGCATTGGCTGTTCCTTATATTTCCGCATTACTCGTTGGCAACTGGAATAAATGAAAGTTTCAAAGTGTATGCTTACAA
TAACATCTGCGCAAACCTTCTTAAAACGTGCGAGGAACGCACATTCCGAAGAAAACTTGTATATCGATAATGCGTAATGAAAGAATTCGTGAAATT
TGTGAAGACGCTGATCTGAATTACTTTAAATGGAAGGCACCCGGAATTGCCAGGAACATGGTTTATTCATTTTTAACAGGCATTATTTTATTCGCGC
TTCTTTTGGCCATTGAGTATAAGATTTTCTCTAGAGTACACTATTACTTAACTCAAAAGCATTTTACCAAAAAACCTATACCTGTCGAAGACGAAGA
TTCTGACGTTTCGAAGGAACGAGAAAGAATACATGTTGCTACAGAGATTGACATTAAACAAAATTATACGTTGGCGGTCAAAGATTTGACTAAGTAC
TATAAGAACTTTTTGGCTGTGAACGGGCTGTGCGTTGGAATAAGAAAGTTCGAATGCTTTGGATTGCTTGGAATCAACGGTGCTGGAAAAACTAGTA
CCTTCATGATGATGACTGGTGATACTAGTATTTCATATGGTGACGCCTGGGTAAATGGTAAAAGTATTAAACAACATTTGGAAGAAGTTCAAAAAAT
AATCGGATATTGTCCCCAATTTGATGCTTTACTGGACGACATGACAGCAGAAGAGTCGATTATAATGTTTGCTATGCTAAGAGGTTTACCTTTCAAA
GATACTTTTAAATTAGCTGATTATTTGTCTAAAGAATTCGATTTTACCAGACATTTGAAAAAGAAAGTGAAAGAACTAAGCGGTGGCAACAAGAGGA
AATTAAGTACAGCGATTGCGTTGATTGGGGATCCTCCTTTGTTGTATCTAGATGAACCAACCACAGGAATGGACCCTGCCACCAAACGTCACCTCTG
GAATGCCTTATGCAAAATCCGAGACAGAGGTAAATGTATAGTTCTCACATCGCACAGCATGGAAGAATGCGAAGCTCTCTGCACCCGAATAGCCATA
ATGGTAAACGGCAACTTCAAATGTCTCGGGTCCACGCAGCATCTCAAACACAAATTCGCAGAAGGGTACACTTTGACGATCAAATTAAAGAAAATCG
CGGAGAGTTCCAGCGAAGGCCTCTCAGAAACTGAACCTATAGAGAAGTTCATTTGCCAAAGGTTCCCAGGAGCCCAGCTTCGGGAGAGACACCAGGA
GTTATTAAATTATTATATAACTAATAAGTCGGTGCCTTGGTCAAAGATGTTTGGCATCTTGGAGAAGGGAAAGAGAAGTGATCTCAATATAGAAGAT
TATTCATTAGGACAATGTAGTTTAGAACAGGTGTTCCTGTTATTTACTAAACATGAAAATAGTACATGATTTAGTTCCCTTAAGAGATGCGTCAGTG
CCTTTGCTCTGTGTTATTTTCTCATTGTGATAATATATTATTTTTAGTTATTAATAAGATTCCTCATTGGGGCACTACCGTTTAAAGAAACCAATAA
AGGATGATGTGTCGGTTTTTAATCAAAATAATGTGTTTCAATAGAATATGATAGTAAGGTAATATAATTAGCCCAATAAATGACCGTTTTCGAGTGT
AATTTCCAGGGGCRACTCCGAATTGCATGAARGTTTGGATTTGGATTCTACTTRCCCTCCACTTCAGRGTTGAATTTGTGCCGTTGGTTGCTTTTAC
TTGGGGGTGACATTTACCCCTTCTCGGGGGTAAAAAACGCGTG
>Dvv ABC-A C266167_7.0
CCAATCAGTAGATTATACAACTCATATAATCAGAAGATCTAAGTCAGCATGCAAAATGGCCTACGGCGGCAAATCACTTTTTTTCTCGCAACTGAAA
GCAATGCTCAAAAGGAATTTATTACTAAAAAAGAGGCAAAAATTAAAAACGATAGGTGAAATCTTGTTCCCCATCTACTTTCTAGCCCTTCTGGTTG
TTATCGAGTTGGTTCTACCTGATCCAAATTTACCAGAAATCAGTACTCCAAGAGGAGAAGAATATTTATTTAAATCTTTTGACAACGGTGAATCTCA

TAAAATAGCATATGCACCAAATACAACAACAAATATCAACTTTGTCAACAATGTAACCGAAACATGGAAAAAGATGTCCAATTCCAAATCGAGTCTA
GAATGGGTCCCTTATAGTGCAGAGGATAAGGTCCAGGAAGCTTATGATCAAGATCATAAATTTGCACCGATGGCAGTGATCTTTAAGAGTGATGTTC
CTTATGTGCATACCTCATTAAGTTACCAGATTCGTACTAATCCCTACTATATCAGTACCCCTAGTTCTAATGATATTGGCTGGGATAAACAGTCTTG
TAGGGGTGGTACAAACCCAAATTCAAACATGGAAGACGGCAGTACATGTCCAGCTAATACTTACTACTTTTCTGGATTTTTGGCTCTTCAAACCCTC
TTAGACTATACGCGAATACAACTTGATAATCCAACTCAAGATAATATACCGTATATATCTTTGGAAATATTCCCAAAAGCTGCAACTACTGTTGGTG
ATTTTGCAAACATGATTCGCATTATAATTCCAATTTATATGGTACTAGCTCTTTCTCAATTCGTTACGTATCTAATAATTTTAATAGTTGGAGAAAA
GGAGAAAAAGATTAAGGAAGGAATGAAACTTATGGGATTGATGGACAAAGTGTTCTGGTTATCATGGTTGATTATATACGCATGTTTTGTAATTCTT
TTGGCTACAATATGCGTTTTGCTGTTATTTGCCTTCAAGGTATTTGTTCATGTAAACCTTCTACTAATCTTCATCTTAATGGTCCTCTATGGTATTT
CCCTCATTATGTTTGCATTTATGTTAACACCATTTTTCGACAAAGCGAGGACGGCTGGAGTACTAAGTTCCTTTATATTGGTACTAATAATTGTTTT
ATATTTCGTCGAGCAAATTGTCTCCGGTAAACATCCTCAATATTTATGGGCACTGTCCCTTTTAAGTCCGGCTGGATTTGCATCAGCTCTTGATAAA
GTAGCAATATCAGATATTCAAGGAAAAGGATTAGACTTCAGTAATATGTGGGATAACCATGGTTCGGGGGTTCCTTTTGGTGGCAGTTTGGTTATGA
TAATGGTCGATATAGTATTGTATGCCATAATAGGCTGGTATCTTGATAATGTCATTCCAAGCTCATATGGTGTAAAAAGATCGCCAATATTTTTCCT
GATGCCATCTTATTGGAGATCCAGTCATAATCAGATCCCTCCCAACCAACTTAATACAGAAAGTTCTGTAGATGTGGAGCCGGTGCCAAGGGAAATG
CAAGACAAAGAAGCAATCAGAATTATAAACTTGAATAAAAGTTTTACGCAATGCCGTAAACCAACCGTTACTGCGCTCGATGGTATCAATTTGTCGA
TTTACAAAGGCCAAATCACTGCAATTCTGGGTCATAATGGTGCCGGAAAGACTACGCTCTTTAATATTTTGACTGGATTGTCCTCGCCTACTAGTGG
GAGTGCTTTGGTTTTTGGATATGATGTCAGTAACCCCAATGAAATGGACAAAATAAGAAGAATGACAGGAGTTTGCCCGCAGCATGATATCCTGTTC
GACGATCTGACACCAAGAGAACACTTGGAATTCTTTGCAGCCATAAAAGGAATATCGAACAGACAGTCCGCAATCGAGAAGATAATAAGAGAAATAG
ATTTACTTGACAAAATTGACACAGCCAGTAGAAGCTTAAGCGGGGGACAAAAACGGAAACTCTCTATTGGTATTGCTCTCATTGGTGATCCAAAAAT
CATTATACTTGATGAACCAACGGCTGGCGTTGATCCCTATTCAAGAAGACATTTATGGAATGTTTTGCAAAATGTAAGACGCGACAAAGTGATTTTA
TTGACTACACATTTTATGGACGAAGCTGATATTTTAGCTGATAGAAAAGCTGTTGTTTCCAAAGGAAAAATAAGATGTTGCGGAAGCTCATTATTTT
TGAAAAACAAATTTGGAATCGGCTACCATTTGACATTTGTCCTGGAGGATAAGTCAAATGAAAATGCTATAAACCATTTGGTTCTGCAATTCGTCAA
AATGGCAAGGAAAGATAGACGACACGGAAAAGAATTGAGTTTTATATTGCCACATAATGCTGTTGAAAATTTTGCTTCACTTTTTGCGGCCATAGAA
CACGAGATCAGCATTAAATCCGATTTAGGTATTTCCAGTTATGGCGTTTCAATGACTACTTTGGAAGAAGTTTTTCTTGGTTTGCAAAAAGAAGAGG
AATATGGCGACGCAACGGTATACCAGGATTTGCCCCAGAGCTCAGAAGGGATTAGCAACGAAAACAGAGCTCTTACCGAAGGTATTCGTTCATTGGA
AGCATTTCGCTGTACTCCTAGTACTTCGCAAAATTTAAGAACGTTGATAGGCCTACGACTTTTGAGACTACGACGTGAGAAAAGGAAGCTGTGGATG
GTAATCATCTTGCCAATTATTTTTACAGGTCTAGGACTGTACTTAAATAAGGTTATTGACACTGGAAAAAAATACGAACCACCTAAAACATATCCAA
TGCCATTAGACTTGAGCCATTACCTAGGCTATAATATTAGTATCTACAATGGGAGTCCGGCAGATTTAACTACATTTAAGAATCAGTTAGAGAAAAA
TGGAGCAGAGCTAGATACTTATGATGGAGAATTTACCTCATTGTTAGACATTGCTCCACATTATGCATCTTTTAACATTCACGATTTTAGTTCTTCT
TCTCAAAATATATCCGTACTTTACAATGGCAGCTATTCCAATAACTTACCAATGTTTATTAATCTTATCAGCAATACTTTTTACAGTATGGCCGCAA
CAACTGGAAAAATCGCAGTTACCACATACCCTTTTGAACTACCTCCAACTAATGTGAGTCCTACTCCTGCTAGTCCAGGGAACTTCATTTTGGGTAT
GATTTTCCTGTTTGCGCCAATCGTACTAGCAGTTGATATGGTGTATGAAAGAGAGATAAAAGCCAGAAATCAGTTAAGAGTTAATGGACTTCCTTTC
ACCGTTTACTTTACAAGCTTTTTCCTGGTTCAAATCTTATTAATGGCAACCATAACTATATTGTTAATAATTCTTATTGTTGTTATGAAACCACCGA
CTTTTTCAAATTCCTCAATAACAGTTTTAGCAATATGGGTTATTCTCTATTGCCCCGCTTCAGTTTTGTTTTGTTCCTGCGTTAGCTACATGTTTGA
TAAGTCGGAGTCTGCGCAATCGGTTATGCCAAACGTATCTACTCTTCTAGGATTAATACCATATATATCTGTGATGTATGCAAAAGAAAGTGTTGCA
ACTATTTTGCATTACGTATTTGCATTTACTGATATGATGTATGTACCTTATGGTATGCTTTATTATATACAAAAGATTAATTTGGAATGTATATCTG
ATCCCCATTGCGAGGGAACTACATTCTCCAGCTTCATGATACCGGAAATTATAGTACTATTTGTTACGCTGATAATTCAAATACCTCTCTTGTTCCT
GATTATGCTTATTCTGGATGTTAAGAAAAATGGAGGACAAGTTTTGGATATATTTAGATCAAAGAAAATTTCCGAAAATATTGTTGAAGAGTCCAGA
GATGTCGGAGTACATGAGGACAGAGATGTTAAAAATGAGAGACAGCGAGTAAACAATTTAATTAGGGATCCTCAAAATAACCGCTCTGTTATAACAG
TTGAGAATCTACATAAGGTTTATCAAAAAGGTGTTAAATCTGGTATTTGCTCACGTTCTGCGGAAAGTCTGAAGGTCGCAATAAAGTCTATATCATT
GGCAGTAGATTCGGCAGAAGTATTCGGACTTCTTGGACACAACGGTGCTGGAAAAACTACGGCAATGAAAATAATTACAGCTGAAGAGGCACCAACA
CGAGGACGGGTACAAATTGTTGGAAGAAACATAACCTCCAGCGCTAATACAGCATTCCAGTATTTGGGCTACTGTCCACAACATGATGCCCAATGGA
AAAATATAACTGTGAAAGAACATCTAGAATTGTATTCTCACATAAGGGGTATTCCGAAGAATGAAATTAAAAGAACTGTTGACTTGTACTTAACGGG
TCTTCAAATTCACGAACATAAAGATAAACAGGCTGAAAAATGTTCGGGAGGTACCAGAAGGAAGCTCAGTTATGCTATGGCTATGATTGGCAATCCG
AAAATAGTTCTTCTCGATGAGCCTAGTACTGGAATGGACCCGCAAAGTAAAAGATTTTTATGGGATACTGTTTTAGCTAGTTTTAGGGGTTCAAGGG
GAGCTATTCTTACAACACACTCCATGGAAGAAGCTGATGCGTTATGTTCGAGAATTGGCATTATGGTAAATGGAGAACTTAGATGTCTAGGATCAAG
TCAACATCTTAAAAATCTATATGGAGCTGGATACAATTTAGAAGTTAAATTGGCAACACAGCAGGGTACAGATTTACGCCAAAAATTAAAAGAATTA
GAAAACTATGTGTTCCATCTTTTTCCCAATGCAGTATTACAGGAAACTTTTGGCGACAGGCTCATATTCAGCGTGCCACAACAAAGTGTTCCTTCCC
TGGCCAATTGTTTTAGGAGTTTAGAGGACGGTAAGCGTCGATTAAATATTGAAGAATACAGCTTTAGCCAGACAACATTGGAACAAGTTTTCCTAAA
ATTCGCCCAAGAAGGTGAACATGAAGATTAATTATGTGATAATGGTATATTACATTACTATTACTATTTATTTGAAGACAT
>Dvv ABC-B D21313+T40801_c0_seq1
TGTGTTTTTAATATTAGTTCAACTAAACTAAATACTAATTTTTAATAGAAATATACAAATTTTTTTTTAAGAAATACTCTCTACCGATTGTAATTTA
TAAAATTTATAATATAAACCCTGTTTTTCCAGAAGCTCTTTATGAGTTCCCATTTCTACTACATTTCCTTCTTTTAGTACACAAATAACATCGGCAT
CTTGAATTGTCGTCAAGCGATGAGCTATGGTTATACACGTTCGACTTTTTCTAGCATTGTCCAAAGCTTCTTGTACTATCTTTTCACTTTCATTATC
TAATGCTGATGTTGCCTCATCCAATAAAAGTATTTTTGGATTTCGCAGCAAAGCTCGTGCTATGGCAATACGTTGTTTTTGTCCTCCAGAGAGCTGA
GTTCCTTTACTACCTACTTTTGTTTCGTAACCCTTTGGAAGATTTAAAATAAATTCATGAATATTTGCAGATTTAGCAGCATCAATAATAGCATCCA
TTTCCACTTGTTTGTAATTTGCACCATATGCTATATTCTCAGCTATAGTTTTGTCGAATAGATTTGGCTCTTGTGAGACTATTCCCAACTGAGATCT
TAACGACTTCAAATCAACATATCGAATATTTTTATCACTAATTTCCACTTTTCCATAACTAGGATCGTAGAATCTTTCCAGCAGTTGAATAATAGTT
GATTTTCCGCAGCCACTTGATCCTACCAAAGCAACTGTCTTTCCTTTTGGGATTAAGAGGTTTAATGCTTTTAGAATTGGTATTGAAGGTCTCGTAG
GGTATGAAAAGTAAACTTCAGAATATTCTACATTCTCATCTTCCCATAAAAAGTTTAAGGCGTCTGGCATGTTTTGTACGACTGGTTGCCTCTCTAA
AAGCGTTATTATTCTGCATGCCGCAACAAGACCTTTTTGGAAATTTGGAGAAAGAGACAATGCGTTTCCAACTGACCATGTACCGACTATCATTACT
TCACAAACTATAAAAATTGTTCCATAGGGACAATCACCGGAAATGATTAGGTTAATGCCATATCTTATTCCCGCTACATAGGCAAACAGCATTAGAC
TTCTAGCTAAACCTACTATTATGCCACGATAATGTGACTTTTTTCTTGCCATTTTCTGATACGGTAATAACTCTTTGATATACTGTTCCATAAATAC
TTTTTCACAACCTAGAGAAACCACTGTTCTTATATTACCGATAGCTTCTACAGCCAACTTTGCAGATGCTTCTAACATTTTTTGGTTTTTCGTGGCA
TCCTCTTGTAATACCTTCTGCTCAAAATATACAGAAAAGAATATTACGGGACACAAAGAAAAAAGCACAAAGCTTGTCCTCCACTCAAAATAAAACG
AGAACGTACACGTAAGTATGAATGTTGAAATAAAGTTTATTGTTGTTCCTATCTGTGGCCCAGCTGCCCCTTGGACAGCTACAGCATCTCCAGCCAG
ACGTGCACACAGTGCCCCAACTCCATTGCTTTTTTGGTCATACCAGCCAATTTCCTGTCTTAACATTGCTTCAAATGTTTTCGACCTAATTTTTAGA
GTTAAATGTTCTCCAGCAACAGAAAAACTGGTTGTCTGCCAAAACATTGCAACTCCAGTGACTACACCTAGTATTAGGAAATATAAACAATACATAT
TACTCTCACTCCGTACTTCTCCATCATTATCATTTGCAAAAATTCCCATTATTCCTCCAAAAACTAATCCATAAATTGGTAAAGAAGCTCCCGTTAT
GAGAGAAGATAAACAGCCAATAAGAATATAAAACCATTCTGAGGCGTTTAATTTTAAAACCTTCCATAAAATGTTCTTCGCAACTGGTTGATTAATC
TGTGTATCTTCAAGTTCTTCGTCTTCTTTTTCATTTATAGACTTGTTAGCTTCACTGAATTTTCTGCTCCGTCGTCTATTACCTTTTTCTTCTAAAG
TTTCAGTTAGTCCTTGAGAAATTACTAGATTGTGATAAGCTCCTTTCATATCCATTAGTTCAGTATGTGTGCCCATTTCTAAAACCTTTCCTTCATT

GATTACCACGATTTTACTTGCTTGTCGTATAGTTGATAGCCGATGTGCAACTATTATTATTGTACAGGTTCCTCGTATTGAATCTAAAGCCTCTTGT
ATTTCTGCCTCACTTGTCGTATCGAGAGCTGAAGTCGCTTCATCTAGAAGTAGAAACTCGGGAGTTCGTACTAGAGCTCTAGCTATGGCTATCTTTT
GTTTCTGGCCACCAGATAACTGAGTTCCTCGTTCTCCTAGAACCGTTTGGTAACCCCGCGGCAAAGTCTGTATAAATGTGTGGATGTCAGCTTTCTT
CGCTGCATTCTCTATTTCTTGTTGAGTAGCATTTAGTTTTCCATATCTTATGTTTTCGGCGATGGTAGTGGAAAATAAATCTGGCTCTTGACTTACA
ACGGAAATCTTTTGTTTAAGCCAGTCTAAATTTAGCTGCTTAATATCTATACCATTTATCTTTATTTGACCCATATCTGGATCGTAAAACCTCTGAA
GTAATTGAATACAGGTAGATTTACCACATCCAGATTTTCCAACTAATGCAACTGTCTCACCTTTACTTATTGTAAGATTTATACCTCTAAGAATCTT
TACATCTGATCTTGAAGGATAACTAAAATGCACATCTTCAAATGTTACTTCCAAAGCGGCATCATTTATTACAATTCCACTGTTTGCAAATTTTCTG
ATAACAGGTTTGCTATCTAATACGAAAAATATTTTTTGTGCTGCCCCTTTGGCTGTTCCAAAAGTCTCCAGTAGTGGAGCTCCTAATCCAAAATTCC
AAGTAGCTATTAACGTTGAAAAGAACACAGATACCATATTTGCCGGTGTATAAACTTGTTCGTGTTCTGGAAGATGTCTGTCATTCAGTATTAATGT
TACTCCATACCAAAACGACAAAGCGTAGCACGCGTATACGAAAAACCATAAGCATCCATTACTTACAGCAGTAAAAAAGCTACGTTTTAAGTTATTC
TTTTTGGCATCAACGAGGTGTTTATTATATCTTTCAATTTCTTTGTCTTGTCCATCAAAAGCAACAACAGTTTTGATTGCGGTGAAGACCTCTTCAG
CGATGGTACCAGCTTCTCCGTATGATTCCATTTCTTGCGCTGAAAATTTTGTAGATATCCATGATATTATTGTCATAATTGTAGTCGATAAAGGAAG
GGATACCATGCACACCAATGCTAATTTCCATCCTTTAACTAGACCCAAAACACATCCTGCTACAAAGGTTGACTCAAAAAATAGAAAAGTTCCAATT
TTCTCACCAATTCCCTCTTCTAATTTTGAAATGTTTTGTGTGAATGTACTTGCAAAGTCTCCAGTACGATTCTGGTCAAACCACGCAATATCTTGAT
TAAGCGTTTTCTCAAGAAAAGCTTTTCTGATTTTGAAAATCTGTCTAGATGCGCTGTAAATAAAAAGGACCGTTGATATATAGGTAGTTAAAATAAC
GCCTATCCCAACGCCTACCGACTTCATGCCAAACAGCCAAACGGCATTTATTAATTCATCTGCTAACAAAGTCTTTTCTGGCTCACTAAGTGATTCG
TTGTAGTTTGAAGCGTATGTTACTATTGCACCGGTAACATCTCCAAATAATGTCATAGAATATGGTTGAACTATTCCACTTAAGGCAGCACAAATAG
CTCCAATTGACATAAAAATAATATCCAAAGTAGTACAGTACTGAAACAAACGAAGATAACTTATTTCATTAGTTTTGTTCTTTTCACATTTTAAATC
ATTCTCATTTTTTTTGTTTCCCATTTTGCATTCGGTTTGCACTAACTACCAACTTTCCTTCTATTTTGTGTATAACCATAAAATTGTTATTGCAGTT
TCTTAACTATGTAGAATAAAGTCTGTCATGAGGCATTAGTTAGTATTAAAATGCATCCATAAAAACTAATAGCCAGACAGAATGATTTATCTAATTG
TCTATATCAAAAACGAATACAACTGATTCATTTTTTAAATACGAATACGAACGCAAGTCGGTCCAATTAAAATTTGTTCAGCAAATAAAAATCTAAT
TCTAGCTCAATGT
>Dvv ABC-B D17742
GGATGACGTCACTAGTAYGATATTCATACCAAAAAGTCCCAATTTAAAAATAAAAATCGACCTGTCTGATGACTTCTCCTCAAATTTGCCCAYTCTC
GAGATAATGAAATAGGTAATGCAAACTCAAACGTCCTCACTGTATAAAAAAACAACCATGCACAAATGCTCCGGTCAAATTTGACGCTACCTTCCGG
ACTATATGTTCCAATGCCCAAATTAATCAAACTAACTTCTTACGAGATGTAGCACTGACTTGAAAATTTGATTTTTACGTTTTTTGAGTATATAATT
TGTAGTACAGTCCCTCCTTTTGCAGAAGTTCTTGATGGCTTCCAGATTCAGCAATTACACCATTTGCTACTACGCATATCAAGTCTGCATCTTGTAT
GGTTGTTAATCTGTGGGCAATCGTAACACATGTCCTTCCTTTTTTAGCTTGATCTAGAGCTTCTTGAACAACCTTCTCACTTTCAGTGTCTAGTGCA
GATGTGGCTTCATCAAGGAGTAAAACTTTTGGATTTCTTACTAACGCTCTGGCAATAGCAATTCTTTGTTTTTGTCCACCAGACAACTGAACAGCCT
TTTCTCCTAGTTTTGTTTCATATCCCTTCGGTAATCCAGTTATAAAATTATGAATATTGGCATTTTTAGCAGCTTTAATTATCTCATCCATAGAAAC
CTCCCTTGAATTATCTCCATAAGCAATATTTTCAGCAATAGACTTATTGAACAGATTTGGTTCTTGTGATACGATACCCATATGCGACCTCAACGAA
TCTAGCGATATATCTTTTAAATCTACACCATCCATTTTTACTTCACCAGATCTAGGATCGTAAAATCTCTCGATGAGTTGAATTATCGTAGATTTTC
CGCAACCACTTTCTCCCACTAGAGCTACAGTTTTTCCATTTAGAATGTCCAAATTAAGTCCTTTTAAAATTAGTACGTTGGGTCGTGTTGGATACGT
AAATTCGATTTCTGAAAAACTGAAGTTGCCACTTGCCTTAACCATTTGTCGTGAAGAAGGTAAATCACGAATAAGTGGAATTCTTGAAAGGAAGCTT
TTAACTTTTTTTGCTGCTGCAACACCCTTAGCAAAGTTTGGAGAAAATGCCAGTGAGTTGGCAATTGAAACTGTCCCCATAATTTGTGCTTGAGATA
CCTTAAACACTCTATCGTATGGCAATCCATCCCTTATCAAAAATCCGCCGTAGTACATTGCGGTAGCATAAGCAAAGAACAACAAACTTCTGGAAAG
ACCAAACACGATTGCTCTCCAGTGAACTGTTTTTAACGAAGATTTATAATGTGGCATGAGCTCAGATTCATATAGCTGCTGGAACTTTTCTTCTAAG
CCTAAGGAGGCAACTGTTCGTATATTCCCAACGGCTTCTACTGCAGTCCTTGTTGATTTTTGTAGTGCGGAATCTCTTGAATCGTTTCCACCTCTGG
TGTTTCTTCTTTCAAAGAACATAGCGATCAAAATAACAGGTGTGAAAGCTGCTGTAACTAAACCCAATTTCCATTCATAATACATGGAAAGACCAAT
TGCTAAGCAAAATGTGGCTATTGATTGAAGAATGGTTCCAACCCGAATTCCCGTTGCCCCTTGAATATTAGAAGCATCAGAAGACAATTTGGCACAA
AGCGCTCCAACCCCATTTGTTTTCTTGTCAAAGAAGCCTATTTCCTGGTATATCATTGCCTTGAATAATCGACTCCTAACTCTTTCTGTCATTTTTT
CACCAGCAATTCCTAACAGGTACATCTGAGTAAATACTGATACCATAGTTATGAGTCCTCCCAGCACAAAATATAAGCAATATTTGTTGGTTTCAGA
TCTTACAAAGTCAGGATCACCATTTGCTAATGTCCCCAAAATGCTACCAAACAAGACTGCAAAAACGGGCATTCCACAACCAACAACAATTGATCCA
AGACCAGCAAATAAAATGTAAGGCCATTCTGGAGCATTCATCTTCAATATTTCAAAAACGGAAACGTCTTTATTTTCATTAAAATCGTCTTCATTAT
CTTCTTTTGTAGCCTCTATGTTATGGTCAGAGCCTGTGCTACTTTCTGCCAAGCTAATAGCTCTTACCACTTTCTTCTTTTCAGCGACTTCAAACTG
TTCCGATGATTTGACTTGGGCTGTTACTAATCTATAATACTCTTGTTTAAGTTCCATTAGTTCTTCGTGGGTTCCTTCTTCTACCACGACTCCTTGA
GAAAGTACGATAATTTTGTTTGCACCTCGAATTGTTGATAGCCTGTGTGCCACAATTATAGTTGTACATTCAACACTTGCCGCATCTAGAGCTGCTT
GTACTTTAGCTTCACTGGTATTATCCAGTGCTGAAGTAGCCTCGTCCAAAAGTAATAAAGTAGGATTTCTTACCAAAGCTCTAGCTATAGCTATCCT
TTGTTTTTGTCCACCTGAGAGTTGAGCCCCTTTCTCTCCAACCAATGTATTATATCCGTTAGGAAGAGCTTTAATAAAGTTATGCGCATTAGCTTTT
TTAGCTGCTTGAATGACATCTTCCTCTGTTACTCCATCTTTACCGTACCTAATATTTTCTAAAATGGATGTTCCAAAGAGAATTGGTTCTTGACCTA
CAACTCCTATTTGGCTTCTTAACCAAGTTAAATCGTATTCTTTAATATTTTTTCCGTCTATGGTGACCTCTCCAGAGTCCGCATCGTATAACCGTTG
AAGTAACTGAATAACCGTAGATTTTCCACAGCCCGAACTACCTACAAGTGCCACAGTGTCTCCAGATTTAATGTTCAAACTTAATCCTTTTAATATC
GGTACATCTTTCCTAGATGGATAACGGAAAGCCACATCTTTAAATTGTATATTTCCTATGAGTACGTTTAGCGTTTTTCCATTTCCTTTTGATAAAT
TAATTGTAGGTTTGTGATCGATTACACTAAAAATTTTTGACGCCGCCGCTTTTGAGATTCCAAATGCCTCTATATATGGAGAGGCAATACCAAAGTT
CATAGATCCTGTCATGACACTAAAGAAAACAGTCACCATATTACTAGGAGTATATATTTGATTCGTAGCTGTCCTGTCCCTTAAAACTAATTTTACA
CCATACCAGAATGCCAGAGCATAACTGGCATAGATCATGAACCACAATAAACCGAAACCTAATGCTGAAAACATCGATCTTTTTATATTATTGTTTT
TGGCTTCAACTAAATTTTTATTATATCTGTCTACTTCTTTCTTCTGACCTCCAAAAGCCGTTATGGTTCTAATTAAAGACAATGCTTCTTCGGCTAT
YGATCCAGCAGCGCCATATGCTTCCAGTTCATTTTTTGCTAATTTTGTAGTGAGAAGTCCTACAATTCCGATTGTTATTAATGTAGCGGGTAGAGAG
GTTAAACATATTAAAGCAAGTTCCCATCCTTTCACAAGAGCTATGATAAGTGAAGTTAAAAAAACAACTTGAAAACTCCAGAATACCGGTACTTTTT
CACCTATGCCATCTTCAAATTTGAACAAATCATCTGACATCCTACTGGCAAAATCTCCAGTCTGATTAATATCGTACCAACCTACATCTTGGTTTAA
AATGGATTTCAAATATGCTGACCTTATTTTGAATATCTGTCTCATAGCAGAATAGTTAAAAGTGACTGTTGCTAGGTAACTGAACACAAACATTCCT
ATTCCGATCAGAGAATTCATTAGGGCAAAATATCGGATACCATCAAAAAAATTGTCTTCTGCTATTTTCTTCTGATCGGCAGGTAAATTTATCTGTA
TCGACGTAGCATATGCAATGATATCTCCTGTGAGGCTACCAAAGAGAATCGTATTTAGGGGTTGTAAGACTCCTGTTCCAGTTGCCGATATAAGTCC
TACTACCATTAAAAATTTATCATAAGTACTTGCATACCTAAACATCTGAGGAAAAGATACATTCTTAATTTTTTCCTTTGGTTCTTCACTATTAACA
AATTGGGCATCAATACCAATTTTCTCTTTATCCTTTRTACTATGTTTTTTTTCTTCTGTCATTGTAATAATTTTAGTRTAATGTACGCAATATCAAA
GGTATTTTAGTTGTGT
>Dvv ABC-B D19147
CCTCATTTCTGGCGGTCAATGTATTACATTCACATCCTATTGTTTAGTTAACAATGATAAAATATGTGATAGTCATTGCAAACCAATGTGTAAACGT
TGATGTAAATAATATATACATAATGAGATAATTATTATTTATCGATAGGATTACTGCCAAAACCAGACAAGAATGAATGTATATTCAGTGATAGTCA
GTAGTCATAATCTATCGAAGTCGCTCAAGCTTTTGAAATTGGTTAATAGTAGGAGATACTTTAGCAACATGTCTCAGAAAATAAGTCAAAATAAAAC
ATATATTTCAGTAGAATTCCGGAAACAACCAAAGGTGCCTCAAGAGGATTTTTCAGTAGAATTTTTAAAAGAACAAAGAAGTAAAAGAAAAACAA

GAACCTAGATTGAAAAAATCCGAACTCAAGAGATTGTTTTCTCTTGCCAAACCTGAAAAATGGAAGTTAACTACAGCAATAGGTTTCCTTGTAGTGT
CGAGTACAGTAACAATGTCAATCCCTTATAGTTTAGGAAAAATATTAGATATTATATATGTGGGTAGTGGAGACTCAGATGCAGCGAGAGCACGGTT
AAACCAAGTGTGTGGTATATTGCTAGGTGTGTTTGTGTTAGGGGCCATTTGTAATTATGCCAGAGTTTACTTAATGTCTACAGCAGGATATAGAATG
ACAAATGCTTTACGAAGACAAGTTTTTGGTGCAATATTAAGACAAGAACAAGGTTGGTTCGACAAAAGACCTACAGGTGAACTTGTGAACAGACTGT
CAGCAGACACACAAATAGTAGGATCTGCTTTATCTCAGAATATATCTGATGGTTTAAGGTCTCTAGTCATGGTATTTGCTGGTACAGGTATGATGTT
GTACATGTCTCCACAGTTGGCATTGGTTGGTTTAGCAATAGTACCTCCAGTAGCAGTTGTTGCTGTGCTTTATGGTAGATTTGTGAGAAAGATTTCA
AGGAAAGTCCAAGATTCGTTAGCAGATTCTACAAAGGTGGCAGAGGAAAGAATTGCGAATATAAGAACAGTGAAATCATTTGCACAAGAACCTAGAG
AAATAACTAGTTATAATATTTCAATTGAAAATGTTCTTAAGTATTGTTACAAAGAAGCAAAAGCAAGAGCTTTATTTTATGGCATGACAGGGTTTAG
TGGTCATATTATAATAATATCAGTTTTGTATTATGGTGGGGTGATGGTGTCTTCAAATACCATAACAGTAGGTAATCTCTCATCATTCCTACTATAT
GCCGCATACATTGGTATTTCTGTTGGTGGATTATCCAGTTTTTATTCAGAACTAAACAAATCTTTAGGGGCAGCAACCAGAATTTGGGAAATAATTG
ACAGAGAACCCACAATTCCTACCCAAGGAGGTTTAATTCCAATGACAGACGTTGAGGGTCATATTGAGTTCAAAAACGTTAAGTTCTGTTATCCTTC
AAGAAGTGACATTGAAATTTTCAAAAGTTTAGTTTTGGACATCCAACCAGGCAAAACGCTCGCCGTTGTTGGTCCTAGTGGTTCGGGCAAGAGTACT
TTAGCGGCGCTTTTATTAAGACTATATGATCCTATAGATGGGGCAGTATACATTGATAACCAGAATATTAAGGAGTTAGATCCTGCATGGATTAAGA
AACACATCGGTACCGTTAGTCAAGAACCCATACTTTTTTCTTGTTCGATAAAGGAAAATATACTGTATGGAGCAGATGATCCTCAGAAAATCAATGA
TGACGAATTTATAAGAATTTGTAAAGAAGCCAATGTTTACGAATTTGTTCAAAACCTACCTGAAGGTTTTGAAACTGTTGTAGGAGAGAGAGGGGTT
ATGCTTAGTGGAGGCCAAAAACAAAGAGTAGCAATAGCTAGAGCTCTTATCAAAAATCCAAAGATACTACTGTTGGATGAAGCCACTAGTGCGCTGG
ATGCCCAGTCTGAGCATTATGTACAGGAAGCTTTAGATAGAGTTATGAAAGGCAGAACCGTTTTAACCATAGCTCATAGATTATCTACTATTCAAAA
TGCCGATATTATTGCCGTTCTACAAAATGGACAAATTGTAGAACAAGGAAAGTACGAGCAGTTACTACAAAGTAATGGACCTTTTAGAGAGTTGGTG
CAACATCAAACGTTTTCGCAATTAATGGATGACTGATTTTTAATCACGGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGCCG
CCGTATCATTAAAAAAAAAAAAAAGA
>Dvv ABC-B T39715_c0_seq2
GTTAACATTCAAAATATGTTTGGTGCGAAATTTTATACACCCCATAATTTATTATTATTTTCTCTAAAACAACTGCAACACATTTAACACAATTTAA
AATCAAATATTAATTGCTAATTGCATAGCCAAGATCAGATTTAATAACGAGTTTCATTTTTCATTTAATAAATTTGGATCACGTTATTCTCAGAGAA
TTCGTTTTATGAATTTATATGACCTAAATTTGAACTTATATATATTTATCTATTGAAACGGGTGTAATTTCAAAATTTTTATACTATGAAACAAAAT
GACGTTTGATGTCTTATATATTATGTGCAGTCCGACAGGCACCAACATCAGATACGTGAACCTGCGTAAAAAACGAGACATACGTGACGTTTCCTCC
GATTTCACATCACGTGTTTTTATGTCAAACATGTTTTACTTAAATTGTAAGAGTTTGGATTGACAACTCTTAAATTATATATATACCCTACTAAAAG
AAAATAAATTTCTTTCGTCATACTGCTGATTTTTTCTTTTCTGGTTCAGTGTGCGATTGATTTTGCGTTTCCCACAATTTCCAATATAGTCCGGAAT
GACTTTCCAGTAGTGACTTGTGTGTTCCTCTGTCAGCGACTTGACCATTTTCCAGTACTATGATTTCGTCTGCGTCCATGATGGTGCTAAGGCGGTG
AGCTATGCAGATACTGGTACGTCCCTTAGTTGCATTTCTCAGTGCCATTAGTATATTATGCTCAGTTATGGAGTCTAAAGAACTAGTGGCTTCGTCA
AACACCAGAATAGGGGAGTTTTTCAAAATTGCTCTTGCTATAGCTACTCTCTGCTTCTCGCCGCCACTTAACTTCAATCCTCTTTCACCTACTTGCG
TGTTGTATCCTTGGGGCCAGGTCACAATGGAACTGTGAATTTCTGCGAGTTTTGCAGCGTTCATTATTTCTTCTTCTGTAGCTTGAAGATCACCGTA
ATGGAGGTTATGCCTAATTGTATCATGGAATAAAACGCTATCTTGAGGCACAATGGCGATTGCACGGCGCAAACTGTCCAGGTCTACATCTCTTATA
TCTTGATTTCCTATGAGGATTCGCCCTTTCGATGGTTCATAAAATCTATACAGCAGTCTAACAAGAGTCGATTTTCCTGATCCGGATCCACCTACAA
TAGCTACTTTCTTGCCAGGTTCTATTGTTAAATTTAAATCTTTGAAAATTGTTTTTCCGGGTCCATACTCAAAAGACACGTTTTCGAATTTTATCGG
TGAAGATTTTGAATCGACATGAAGGTATGGGGCTTCAGGTTTACTCTTTACGGCACTATCCATTGTCATCAGAGTGAACATAGTTTGCATATCTATT
AGTGCCTGTCTCACTTCTCTATATACACTTCCCAGAAATCCTAAAGGAATTGATAATTGAAACAGCAGTCCATTTACCATCACTAAATCTCCAACTG
TCATGTTCCCTTTAACAATTTCATTAGCTGCTAAAATCATAATCCCACTAAGAGCGGCGCTGAAGATGGCATTTTGTCCAAAATTTAACAGTGCTAA
ACTAGAAGCTGTCTTCAAACTGGCAGATTCATATTTTTGTAATGCAGAATCGTATCGTTCAGCTTCATACTTTTCATTGTTGAAATACTTAACAGTT
TCATAATTGATTAAAGAATCTATGGCTTTATTTCCAGCTTCATTTTCAGCTTTATTCATGTAAACCCTGAACTTTGTCCTCCACTGCGTCACTGACA
ATGTGTACATAGCATACACACCAACGCATCCCAAAGAAATTCCTGCAAACATAGCACCACATTTAATTCCAAGAATAGATGAAACCAAGGCCAATTC
AAACATTGTTGGAACTACGTTGAAAACCATAGCTGACAATACAAAATTTATTCCTCTTGAACCTCTATCGATTGTTTTGGAAAGGGCACCTGTTTGT
CTTTGTAAGTGAAATGTTAGGTCTAGATTATGTAAGTGAATGAACACATTTTTCGCAATTTTCCTAATAGAATGCTGAGCAACTTTAGCAAAAACAG
CATTCCTTAATTCGTTAAAACCTGCAGCTCCAGCTCTTGCAATACCATATCCTAATAATAAACTAGTTGCTACAGTGGCTACCGTTTGTGGAGCAGT
TTCCATATTTAAGGCATTTCCAACATTTAAATAATCCACAGAATATTTAAAAATAAATGGGACACACACATTCATGACCTTGGCTCCAATAAGTAAA
GACACTGCCAATTTTACCCTGTCTCTTATGGATTTATCATTTTCTGGCCATATATATTGTAACATAGCTCCTACCATTTGCCGACCAGTGATAGGAG
TCTGTTTTTGCAAGGATATAGTTTCTCTGGTCAGAGTGGAGGCCCCTGGATGAAAACATCCTCGAATAGGTTGTGCTTTGGGAGCACTACTAGGTAT
TTTACTTAAAATGCCTGCTAAAATGCCACCACTACCTTTGCTACTGTTTGATTTTTTCTCAGTCTTTGTAGAATACTCTTTGGTATAATAAAATTGG
CTAAGGTATACATTTTTTGTATTTTGTCCATAAACTTTAAATATGCATTGTTTTTGGTAAACTACTGTTTTTAAATGGCGATTTATAGCTTTAAAAG
GAATATTACTTAAATTTAACACGGTAGCCATTGTACTGATTTATTGAACTACACATAAATTTAATGAACTACTGGAAAGAAATATTAAATCATATCT
TTTACCTTTAAATAAAT
>Dvv ABC-B D9796
AACAACACCAAACAACAGTACAATAACCTACAATTTTTATCCTATTGTGTACTATGACATTTTGGTATGACTAAACAATAAATGTAAATATACAACA
TAATTTTGAAYAGACTCCAACATTCACATTCAACACATTTCCATACATAAATGACGAATGAATGAAGATTATTTTTTAGTTTGACAATAATAATATG
CTGAACATATCATACTGCCCACCAAATATTTCAGTATGGGATGTATGGGTGGATCATGGAATCCCACAGTGTTTCATGAACACTGTGACTTCTTCTG
TATTAGCAATATATATACTAATAGCTGGTTCAATACAACTACATATTTACAGACGATGGGGTGTTGAATCTGGAACCAGGTTTTTGCCTAGAGGAAG
ATTATATTATTTACAAACCTTTTTAATATTGTTTATACCAGTGTTAGAAATAGTTAGATTTATTTTAACCGCAACTGTTTATGATGATAAACATATT
TATGGATATATGATTGTTTCTCTAGTATTAACGACATTTGCTTTTCCGTTTTCTTTATGGGTTTTAAAAGTGGAGTTATATAACATTCTTCCATCAG
TTCAAACAAGAGGCCATGGGATAGTTTTATTGTTATTTTGGACATTGGCCTTTATATCAGAAAATCTAGCGTTTATCAATTTGACCCAGGAGAACTG
GTGGTTTAAACTGAAGGATTTGACCGATCAGCTGGAAATGGCATTATTTATTTTACGATATATTGCTTGCCTACTTATCTTTTTCCTGGGCCTCAAA
GCTCCCGGAATAGTAAACGAAGTTGATTATTATATGTTAGGTGGTCATCAAAGGAATGTTATTAATCCGAACAATGAAAATGCCTCTACATGGAAGA
ATTTTTGGAAAAAGGTGAAGATATTAGCGCCTTTCTTATGGCCAAAGAAAAATTTCACATTACAGTTTAAAGTACTGATTTGCTTCATATTATTATT
AGGAGGAAGGGCGGTTAATTTGTTTGTGCCTATCTATCAAAAATTAATTGTTGATAGCATGGAAGAAACTGTGGCAAAGATGCTTTTCAGATGGGAC
TGGGTTCTCATATATGTTGGTTTAAAATTTCTTCAAGGAGGTGGCACTGGTGGCATGGGTCTATTAAATAATGTGAGATCGTTCTTATGGATTAGAG
TCCAACAGTACACAACAAGGGAAGTTGAAGTAGAATTATTTAGGCACTTACATAGTCTTTCGTTGAAGTGGCATCTAGGAAGAAAAACGGGGGAAGT
GTTGAGAGTGATGGACAGGGGGACAGATAGCATCAATAACTTATTGAATTACATTATATTTTCCATTTTTCCAACTATAGTGGATATAATAGTAGCA
ATTGTCTTCTTCGTTTCTGCCTTTAATATATGGTTTGGACTGATCGTTTTTACTACAATGATACTTTACATAGTTCTAACGATCGTTATTACCGAAT
GGAGAACAAAATTTCAGCGTCGTATGAATTTAGCAGATAATGAAACTCGAAGCAGAAGTGTAGATTCTCTTCTCAATTTTGAAACAGTTAAATATTA
CGGAGCTGAGAATTACGAAGTGGATGCTTTTAGGGAAGCTGTTCTTAAATTCCAAGATGAAGAATTTAAGTCAAGTATCACTCTTAATATCTTAAAC
ACAGTACAAAATGTTATAATATGCGGAGGGCTATTAGCGGGAAGTTTGCTCTGTGTTTATATGGTAGTAGAACCTAAAACTCTAAAGGCAGGCGATT
ATGTATTATTTGCTACATATATAGTACAACTTTACGTGCCGCTTAATTGGTTTGGTACCTACTACAGGGCCATTCAAAAGAATTTTGTAGATATGGA
AAACATGTTTGACCTTCTCAGAGAAGAACAAGAAATTATTGACGCTCCAGGAGCTACTAATCTATCTGTTCCAAGGGGCCATGTTGAGTTTAAAAAC

GTCAGTTTTGGATATCTTCCTGAGAAGTTGATATTAAAAAATATAACGTTCAGTGTTCCTTCTGGAAAAACTGTTGCACTAGTAGGACCATCAGGAA
GTGGCAAGAGTACTATTATAAGACTCCTCTTCAGGTTTTATGATGTCGATACAGGAGTTATTGTTATAGATGGGCAAAATATTAAGACTGTTACTCA
AGAATCACTAAGAAGAGCTATCGGCGTAGTTCCACAGGATACAGTATTATTTAATAATACTGTAGAATACAATATACAATATGGTAGACTAACAGCA
ACAGTGACTGATGTGATAGAAGCCGCTAGAGGTGCTGATATACACGAAAAAATTCTTACCTTTCCAGAAGCTTATGAAACCAAGGTTGGAGAGAGAG
GACTAAGACTAAGTGGTGGTGAGAAGCAGAGAGTGGCTATAGCTAGAACGCTATTGAAGGCTCCAAATATAATCTTACTTGATGAAGCTACCAGTGC
CTTAGATACTCAAACGGAAAGAAATATACAAGAATCTCTGAATAGAATGTGTGCAAACAAAACCACGATAATAGTAGCACATAGACTCTCCACCATT
ATTCATGCAGATGAGATATTAGTTTTACAAGAAGGAGAAATAGTGGAAAGAGGAAAACACGACCACCTGATAGGCCAAGAAGGAATATACGCTAACA
TGTGGCGACAGCAATTGGAGAATAAAGATAAAGAGTCTTTGGAAAATAGCTCGGAGAGTGCTAAATCTAAATAGCCATAGCGCAGCTCGAACATAGC
TAAATTGTGGTACTGAATTGAAATATTCTATTTAATTATAAGTAAAATAGTTAAAATATGGTATTAGAAAACAATATTGTTTTTTATTTAATAAAAA
AATAAA
>Dvv ABC-B D13664+T8676_c0_seq1*
CTTCTTTTAGGCGGCCGTTTTTTAGGTTCAGAACTTTCTGGTTTATTTTCTGCAGGCTCTTTGTCTTTATTTTGTAGCTGTTGTTCCCACATACTTG
CATATATTCCGCCTTGTGTTATTAAATCTTCGTGCTTTCCTCGCTCCACGATGCTTCCTTCTTGTAAAACTAATATCTCATCTGCGTGGATAATGGT
AGATAACCTGTGTGCTACTATGATTGTAGTTTTATTTGTACACATTCTATTCAGAGATTCCTGTATATTTCTTTCAGTTTGTGTATCTAAGGCACTT
GTAGCTTCGTCTAATAAGATTATATTTGGTGCTTTTAATAATGTTCTTGCTATGGCTACTCTCTGTTTTTCACCGCCACTAAGTCTAAGTCCTCTCT
CTCCAACATTTGTTTCATATGATTCTGGGAACGTGAGAATTTTATCGTGTATGTCAGCGCCTCTAGCGGCTTCGATTACATCGGCATCGGACGCTGT
TAGTCGTCCGTACTGTATATTGTATTTTATAGTATTATTAAATAGTACTGTGTCCTGTGGTACCACACCGATAGACCTTCTTAATGACTCTTGAGTA
ACAGTATTAATATGCTGTCCGTCTATTACTATTAGTCCTGACTCTACATCATAGAATCTGAACAGAAGTCTTATAATTGTACTCTTACCACTTCCTG
ATGGACCCACCAATGCTACCGTTTTGCCAGATGGGACAGAAAACGAAATATTTCTTAAAACCACTTTTTCAGGGAGATATCTGAAGGTAACGTCTTT
AAATTCAACACCACCTCTTGAAACAACTATATCATTAGCGTCTGGAGGATCAACAACTTCCTGTTCTTGTCTAAGAAGGTCGAACATATTTTCCATG
TTCACAAAATTCCTTTGAATCATACTGTAGTAATTACCAAACATGTTAAGAGGTATGTAAAGTTGAACTATATAGGTTGCAAATAATACGTAATCTC
CGGGTTGTAGAGTGTGAGGTTCCACTATCATGTATACACACAATAAACTTCCTGCCAACAGTCCACTGCATACTATAATGTTCTGTACAGTACTTAA
AATATTTGATGTAACGTTTGACTTAAACTCTTCTACTTGAAATTTGAGAACGGCTTCTCTAAACGCTTGTACCTCAAATTTTTCTGTACCGTAATAT
TTAACAGTTTCAAAATTAAGCAAAGAGTCAACACTTCTCGTTCTAGTTTCATTTTCTGCCGTGATCATGCGACGTCGATACTTTGTTCTCCATTCAG
TAATTATAATTGTAAGTACAATGTATAATACCATCGTAGTAAAAACGATTAGCCCGAACCATACGTTGAAGGCTGACACAAAGTAAATGACAGCTAT
CGTTATATCTACTATAGTCGGCAAAATGGAGAATATGAGAAAGTTTAATAAGTTGTTAATACTGTCTGTTCCTCTATCCATTACTCTCAAAACTTCA
CCTGTTTTTCGACCAAGATGCCATTTCAATGAAAGACTATGCAAGTGTCTAAATAGTTCCACTTCAACTTCTCTTGTTGTATATTGCTGTACTTTGA
GCCACAAGAATGATCTCAAATTATTTAACAATCCCATTCCTCCTGTTCCTCCTCCTTGTAGAAACTTTAAGCCAACATATACAACAACCCAGTCCCA
TCTAAAAACCATTTTTTCTACGGTACCCTGCATACTATCAAcGATTAACTTTTGGTAGATGGGAACAAACAAATTTACCACTCTCCCGCCACCCAAT
AACAAAAAACAAACTAGCACTTTTAACTGCAGCACGAAGTCCTTCTTTGGCCAGAGGAATGGCGATAAGATCTTTAACTTTTTCCAAAAGTTGCTCC
ATGTCGATGTGTTTTCTTTGTTCTTAAATTTCAAATTAACCGCAATGTTGTCAAACTTGACGTTGTTCTTCTGCTTCTTTTCTTGATCATCTTCGAC
TTGTGTAGCTGTCATGTTAAGTTTTGTAAATGC
>Dvv ABC-B T17837_c0_seq2
GTTTTATTTATATTATACATACTTATTTACAATATTAACAAATAGATTATCCAGCTGGACTGTTTTGTTGTTGGTATGCTAATGACCAATAGTAGCC
TCCCAGTTTCTTCAGCGATTCATGAGTACCCATCTCTACAATTTTTCCTTTATTTAATACAACAATCAAATCTGCATTCTGGATTGTAGACAATCTA
TGAGCTATCACTATCACCGTTCTACCTGTCCGTGCTCGTTCCAATGCTTGCTGAACTATTTTTTCTGATTCTGTATCCAAAGCACTTGTAGCTTCAT
CCAAAAGTAAAACAACTGGATTTTTAAGAAGAGCACGGGCAATGGCAATTCTCTGTTTCTGGCCCCCGGATAAAGTCACTCCTCTTTCACCAACAGG
TGTGTTATAGCCCTTCGGAAAACTGGTTATAAATTCATCCGCATTCGCCAGTAGTGCAGCTTCTTTTACCTCATCATCTGAAGCATCAGGTTTGCCG
TAGCGAATATTTTCCATTATTGTAGTACCAAACAATACTGGTTCCTGGCTTATTAAACCCAGTACTCTACCTCTTAACCACGAAGGATCCAATGATC
TTATATCGTGTCCATCCAATGTTATCGAACCATCTTTTACATCATAAAATCTTTCTAAAAGTGCTACCACTGTCGATTTTCCATTACCAGACGCCCC
TACAATAGCAACTGTTTTTCCTGATGGTACGCTAAGATTAAAACCTTGCAGAATTATCTGCTGACTTCTAGTTGGATAAGCAAAACAAACGTTCTTA
AATTCAATATCCCCCTTCACTGATTCATATGGTAATACTTTTCCTCCTGTAAGTGCCATTTTAGGAGTTTTATTAATGTATTCGAAAACTCTCGAAC
CAGCTGCTACGCCCCTTACTACAGAGCCAAAAAGTAATGATATTTGAGCCAAAGAGCGTTGTATGGTTTGAGATGCTACTAAATATGCCATAACTTC
CCCAGCAGATAACTGATTGGTTGATAGTAAGTAACCACCCATGTACAAAGTAGATAACACCATACCATTTATAAACATATTGGTTCCAGCTTGAAAT
AAACCAATGCCTAATCCCAAATCCTCATTTAAAACCATTGCTCTATCAGCTTCTGTGTTAAATAATTCCTTTTCTTGGTCCTCCATGGCAAACGCTC
GAACTGTTCTGATATTGCTCACCGCCTCGTCCGCTACGGCAGTTGTTTTTTCTACCTGAGCTTGAGCTCTTCTTGAAACTGATCTTAGTAGAGAGCC
GAAAACTGTTCCTACGGCTATAACAGATGGTATGCAAAGTAGAGAAATGAATGTCATTTGAGGTGATAGCATAATTAAAGATACTGAACAACCGACA
ATCTGCGTGGCCGCTCGCAAACCTCCAGAAACGATTTGTTTGAAGCTACTTTTAAAGTCTTGAACATCTGCTGTTAATCTGTTAATAATTTCACCTG
TTCTTTGCTGATCAAAAAATGCTATATCCTGTTTTAATATAGATTCAAATAAATCTGTCCTCATTTTATATGCCATACGTTCTCCTAAATTCGAAAG
CATAAAGATGTAGAAAAATGTACATACACTCTGAGCAAGATACATGGATATTAACTTTATAACAGGTCTCTTCATTTCATTAAGAAATAATTCACTA
TCCCTGCTTTCACTAAATTTTGCTAGTACATTCACCACACCACCTATAACTTGAGGTATTTGTATATTTAATAAAGCCACAGCCAATGCCCCCACAA
TTGCTGCTAAGAAGTACCAAATATGTGGTTTCAAGTACTTCCATAGCCTTGCCCAATCAAATTTAACATTTTTGTCTGATTTATTTTCATAACCAGC
CATTCGAGTCTTCTTTGCTTCGCATAATACCCCATTGTGTGATATGTACAATTTTATAATTAGACCACCAGTAAGTGTAACCCCAAATGTTGCAACT
TTTGATGGTATAAATGATTTTGGAACTGGTTTTTGGGACACTGTTCTGGAGATATAATTTTTAATATGACCCTGGCAGAAATATTTCTTAAATATTG
CACTATTGTGGAATAATAATTTGTTATGTAAATTTTGTTGAATTAATCTCCACATTATTGTTTATATATTTCTTTTAAAATTTGTTGGCTAGAAGAA
AACCAAAATTTATTTAAATAAAAATTAAAAACTATAATAATTAGCTATGAATGTATTTTACATTTTCATTAGAATTTTTTGAATATTAAAATTTGT
GGTA
>Dvv ABC-C D14968
CTGTCGTTATGATCTCATAAAAGTTCTTAATTTAAAATTGGCTAATAATCACAGGAACATTTTTTTGAAGACCTATCTATCTTTAATTTGTTTATAA
AAACTTTCTTCTGCTATTCCTTGTAAGTTCTCTGCCATTGTGGTACCAGTTTGTCGAACTAATCCATAGAATACCCCTTTCTTCTTAAGCAGCTTGT
GTGGATGATCGAATTCAACAGCTTCTCCAGCGTCCATAACTAAAACCTTATCGGAATCCATGATTGTGTGCAATCTGTGTGCTATCGTTAACACAGT
ACATTCAGAAAATTTCAGTCTTATAGTGGTTTGAATGATGGAATCAGTCATTGGGTCTACGTTGGCAGTTGCTTCGTCGAGTACCAGAATCTTGTTT
CTTCTAATTATCGCTCTAGCTAAGCAGACCAGCTGCCTTTGGCCAACACTGAAGTTTGAACCGCCTTCGGCCATTTTGCTGTCCAGACCTGCTGGTA
ATTCATCAACGGCGTGTTTTAGTTCTACTTCCTCCAAAGCATCCCACAGAACCTCATCTTTGTGTTCATCGAACGGATCTAAATTTTTCCTTAGCGT
TCCTGAAAACAATACTGGTTCTTGGGGTATGATTGATATCTTAGATCTAAGCAATCTCAATTCTACGCTCTTACTATCTATACCATCTATAAATATG
TGTCCCTCATTTTGAGCCAACCTAAATACAGCTTGTATCAAAGAAGATTTACCAGCACCGGTCCGACCAACTATACCAATCTTCTGTTTAGGTTTGA
CTTCAAAGTTGAGATTTTTCAACACGTATGGACTATCATCAGAATATCTCAAGGACATGTTCTCAAACTTCATATGGCCCATGCTTGGCCAATTAGG
GGGTGGCTCCTTTTTAATCTCGTCTTTTTCCTTTGGAAGGTCTGCGTATTCTTGTACTCTTTCCACTGATGTCATTTGATTTTCGAGTTCACTCCAT
TGACGCATACCCCATTGGAACATTCCAGAAAGTGTTATTGCTTGAGTCAATGATAAGCCGACATTCCCACCAAAAGTTTCTTTCTGTATAAACAATA
TGCTTACGACCACCATTCCTATAAAAATTACACAGTGTAGATCTAGCCAAAACCCAAATCCACGATTAGCAGTCAAAAACATATAGTACGCTGCAGT
ATAAGCATCTTGAAAGTGATCGAATTCTTTAGTAAGAATCTCTTCAGCTCTAAAAGCTCTTATAGTTGTTAGACCTTGCAAAGATGCTGATAAATGC

GTAAATATTGGACTACGAGTTACAGCTTCTACTCTTTTAATGTCTCTGCTTGATGCTAAAAATGCTTGCCTTATAAAGTAGAAAATAACTAATATTA
CGCCGGTCGGTAGGAGAACCCATGGATTTACTGATGCAATAACTAAAGTTGTACCAGTAACAATCAGACCAATCCCAATCGTATCTACCAAGCAGTT
AGGCAAAGACTCATCCAAAGCTCCAATATCTTTTGAAAATCTATTCAAAATTCTTCCAGATGGATTAGTATTAAAGAACAACATCGGAGAAAAAATA
ATTTTCCCAAACATAGAATTGTGTAGATTCGTAGAAGCTCTCAAGCAAAATCTGAAGAACGTTATCGATCTTGTCACTGCTAGTATGATGGTAAGTA
TTACTAAGCAACTGTATATTAATAAGGTGTTGTCTTCCGTTAACACTGTACTCCAAAATGGAGGTTTATTGTCATTTGTAGAATAGGTTGCGTTTGC
TAAAGATTCATTAAGAGTTTTATTAGTAAAGGCTCCATTGTTTTTGTTGTCTACTCGCCATTGTTCTACGTTGACCCAAGTTGTTAAAAATATATCA
GTCAAACTTCCGAATACTTGAGCAGCTACGAAAGTACATAACAGCACCACAGATTTTAACCAGTAACCACCAGCATTAATATAATTTGCATAAACCC
TTTTCGATACTGTACCAGTAGCCCTTTCTTCCCTTTTTTGCTCTATAGCTTCATGTTCGGCTTCACTAGCTTCGGACTCTACCGACTCCGCCTTAGA
AATTCTACGTGCGCCATCTTTTTTCTCTTCATCTTTTGAAGAAGCTAGAAGTTTAGTAAAGGCACTGTCTGAATTCTTTAAAGCCTGGTATGTTCCT
GAGGCTCTGATTTTTCCGTCTTCGAATAAATAGATGCAGTTAGCAGTACGAAGGTACTGCAGCTGATGCGTCACTAGAATTACACATTTATGTTTCA
AATAGCCAGTTATGCAATTGGTGTATAATTGTTTGCCTACGTGAGTATCTACAGCAGAAAGGGGATCGTCTAGCAAATAAATGTCAGATTCTTTATA
GATAGCACGTGCTAAGTTAATCCTTGCTCTTTGTCCTCCACTCAACATAACTCCTCTTTCCCCTACTAGGGTTCTGTCTGCGTGAGGAAATAACGCA
AGATCTTTCTCCAGAGCGCAGACTTCTAAAACTTCCTCATACTTAATTTCATCGTACTTCTGACCAAGATAATGTTCTGTCTTATGCTACCACCAA
ATATCCAAGGCTCTTGAGACGCATAAGATATAGAACCGTTAACTTCAACACTACCCTCTATAGGATCTAATTCTTTTAAAATGACATGTAACAATGT
AGATTTTCCTCCACCCACTGTTCCTACCACAGCTATTAACTGATTTGATTTGGCTTCGAATGTTATCTCTTCCAAAGTATTTTCTGGTAAAGATTTG
ATCCATTTAACAGAGGCGTTTTTAATTTTAACGCCTACTTCTTTCGCGTCCGTTTTATTTGATATCTTTTCAGTGACACCATTTAAACCGTTCTTTT
CTGAGTCAGTTCTTCTGTCTAATTCGTCGTACATCAAAAACTTTTGTATCCTAGATATCGACACCTTGGTTTCTGCAAATTGCGTAATGGCTTGAGG
AAATTGTTGGGTGACAGATACTCTTAGGAATGCGTAAAACGATGTTACTGTAAATGCGTAAGAAGCTGTCAAGGTGTTGCCTGTGAGGACGTATGTC
ATCACACAGAGGAAGACAGCTGATCGACTCATTGTAAGGTTGAACGACATCATAATGGCTCGGATAGTAGATGTGTGTCCTATTTCGGTAATTTCTC
TCTTTCTAACCATTTCCACTAATTTTGCGAACGGCTTTTCCCACGTGTACATTTTTATCACTTGTATTCCGTTAATAATTTCATTCATCAGTCTTAC
TCTCTCATCTGTTCGTATGGCAGTTTTGAGCCTATATTGTGACGTTAATTTGGCCATATACATTTGAAACGGTATAAAGCTTAGTAGAAAAACGCAT
CCAATTAATCCGGTTGGTCCAACATAAAAATATAGCAAAATCATAACAACAACTGTTTCACAAGGTGCTAACCATACGTTATGAATGTGTTGACCTG
AAAAGTCAAACCTTCCTACATCATTAGATAACAGATTTACCATTTGTCCAATTGTCGTCTCTGCTAATGATGTTTTACTTAATTTTAATGCTTTCCT
ATAGATCAACGCACATGATGCAACTCTAACCTTCATTCCTAAGGACATGACCAACATCTGGTAGTTATGGACGGACGTGACTTGGATGAGACTTGCT
ACGATTATTAAAAATGCGTTGATGTAGATTTCTGTTTGGTTGGTTGTGGCACCTGGTTGGAAATATGACACTAGTCTGGAGATCAACATTGGTTGTG
CAATTTTTATAAAATCTGAAAACATATTAAAAAGGCATAATAAAATAATTCTAGCCTAAATACTACAAAAAGGGCCCATAAAAAAGACGGCTTCGA
TTTTTTAAGTTCGTTTTTCCATGCTACTACTAATTGATCTGCTAAATATCTGGATTGTTGGGAATGTCTAGTTTGGTACATATCATTTTCGTCTAAA
TCTTTCCTTAATCCTTTAGAGAGAAATCCTGGTAACCAGCAAAAAAATATGTGTGAAAATAAATTAGTTCTTTTCAATGGATGCGGTTTCTTTTGTA
TTCGATTTTTTTAACTTCTTCCATTATCTCTTTTGTACTTTAACATAAAATAAGTAGGTAAACAGCTGACCACTCTTTCCTGACACAACTCTTTCC
ACAGATCACTTACAAGTTCAATGTCATTTGTAAAATTATTCCGAAACACAGTTTTTGATAACTTTTAGGGCAAATTATATCACGACAGTGTTTTGCG
TTCGTTTGTATCTTTG
>Dvv ABC-C D12703
CTGTACTAATGTGTGCAGATCTTTACTTTTGGAGGTTCATATAACTCTCTTCGGCAATTTTACATAAATTCTCAGCCATCGATTTTCCAGTTTGTTT
AACCAAACTACAGAATATTCCCTCTGTATTTTGCAACAAAATGTGCGGATGATTGAATTCTACCGCTGAACCTGCATCCATAACTAATACCTTATCA
GAGTCCATGACTGTATGCAATCTGTGTGCTATTGTTAGAACGGTACAATTGGCAAATTTGTTTCGTATAGTAGACTGTATTAAAGCGTCTGTTTTTG
GATCAACATTGGCTGTAGCTTCATCCAAAACAAGGATCTTATTTTTTCTGACTATAGCTCTAGCCAAGCACACGAGTTGACGTTGACCCAAACTAAG
GTTTGATCCTGCTTCACCGATTTTGCTTTCTAAGCCCATTTCCATTTCGGAGACAACATCCTTTAGAGCTACTTCTTCTAGAGCTCGCCATAGGTCT
TTATCATCGTATTCATCCATAGGATCTAAGTTTGATCGAACTGTACCAGAAAATAATACAGGTGACTGGGGTATAATAGCTATCTTTGAACGCATAC
TGATTAAGGGGATATCACTAGTACAGATTCCATCAATCAATATCTTTCCTTCAATATCTGTTAATCTGAAAAGAGCAGAAATCAACGTTGATTTTCC
CGCTCCTGTTCTTCCTACAATTCCAACCTTCTCCTTAGGTTTAATCTTAAAAGTAAGATTGTTCAAAACATACGGTAGGTTTGTGGCATACCGTAAA
TAAACAGATTGGAACTCAAGCATGCCATATTCTGGCCATTTTTTGGGCGGTTCTTTTGTTTCTGTTTTATATTAATGTAATCAA
GGACTCTTTCCACTGACGTCATCGAGTTGTCCAATTCGCTCCATTGGCGCATGCCCCATTGGAACATGCCTGTTAAACTTATGGCTTGGGTGACTGC
TAGACCTACGTTGCCGCCATAAGCACTGCCGAAAATGAAGAAACTGAAAGTAACCAATGCAATATAGCCAACGCAAATAAGATCTAACCAGAAGCCG
AAGGCTCGATTGCAGCACACAAATAGATAATACGCTGAGCTGTGCAAATCCTGATGTTTGTCGAATTCATTTCTTAAGATTTCTTGAGCTCCAAAAG
CTCGTACTGTAGTAAGTCCTTGCAGAGATGCGTGTATGTGAGAAAAGACTGGACTTCGTGTTGCTGCTTCCATACGTTTCAAATTTCTACTTGAAGA
AACGTAAACAACTTTGAATAAATAGAACAGACAACCTATTGCTAGAGTTGATATCAGAATCCAAGGATTGACAGTTCCAATTACAATGTTAACAGCT
AGAACATTTAGAGCAATCTGAATTGTGTCGACTAAAACGTTGGGCATTGTTTCATCTATACATCCCATATCTTTCGAAAATCTGTTGAGTATTCTGC
CAGAAGGGTTTGTGTTGAAGAAGAACATAGTACCATTACAAATCTTATGCAACATAGTATTGTGCAGTTTTACAGATGCGTTCATACACATTTTATA
GAAACATAAAGACCTCATAACTGTTACAGTGATCACAGATAAAATAATACAGCTGTATATTATAATGCATCTATCACTATTGAAGAAATCAGGCTGT
ATAGGTGTCAGATGTCCTAGAGAATATTTCCAAACATAATTGTAGTTTTCTTCCATATAAACGTCATCTGAGTATCCTAGAAGTAATGTTTTATTTC
CTTCTTCTTCTTGAAAGGTTTCTTGTTCTCTATTTACCCAGAAATTTACAAAATAGTCCGATCCGCTAGCCAAGAACTGAGTCAGCACAAACAGCAT
GAGCACCCAAAATCCATAAAACCAACCACCTGCATATTTCAAGTACCCAGAATAGGCTCTAGATGTCAATTTAACGACTACAGTTTCCTCAGTAATT
TCTGTGGGCTTCTCATCGGACATATTTTCAAAATATTCTATTTCTGACGATTCTTCATCTATCGTATCACTTTCCTCTTCTTCAACGTTGCTGACGA
GTAGTCGCGCGAAGTCGGTGTTGGATTCCGATATTTCTTTGAATGTTCCACTTACGACGATCTTTCCACGTTCTAGTAGGTATATTTTGTGTACGTT
TTTTAAATATTGCAGTTGATGGGTAACAAGAACAGTGCATTTTCCTCGCAAATAACTGCAAATACAGTCCTCAAATATTTGTTTACCTACGTGAGTA
TCGACGGCTGAAAGCGGATCATCTAATATATAAATATCTGCATCCTTATAGACAGCCCCTTGCCAAATTGACTCTAGCCCCTTTGGCCTCCACTAAGAG
CAACTCCACGATCACCAACCATACTCTTATCTCCATGAGCTAATAATGACAGATCTCGTTCCAAAGCACAGACTTTGATAACTCGCTCATACTTTTT
GGGATCCCATGGTTCACCAAATAAAATATTTTGTTTGATGGTACTGGCAAACAACCATGGCTCTTGTGAAGCATAGGAGATTTTACCAACAGTATCT
TTGTTGCCTTGGCTGAGTGGTAGCTCTTTCATGATGGTTTGGAGAAGAGATGTTTTACCAGATCCCACTGCTCCTACTATAGCTACCAGTTGTTGAG
GACCAACGTTAAAATTTATATTCATCAAAGTGTTTTCTTCTTGTGTATCAATCCATTTAGCTGAGACGTTTTCTAAATACACACCTACAGATTTCTT
TTTTTTAGAACTATGTATAAGCGATGTACCAGATGTTATCTTCTTAGCTTGGTATGGTTCTCGTATAGGCTCATAATAAACCTCATCAAACTGAAGA
AACTCTTCTATTCTAGCTATAGATACCAAGGCTTCCGCAGTTTGAATTATAGCTTGAGGAAAGTCATTACTGACTGCTGATTTTAAAATGTTATAGA
ACGGAGAGATAACGTAAACGTACCCAGCATTAAGTCTTTTTCCAGTGAAAGTATAAGTTATAATACATAAAAATACTGCTGCTCTTGTTATAAATTT
GCTGAAAGACATATGAAGAGCTCGAATGTAGGAAGTAAGGCGAATTTGTTTTATTTCTGAATTCCTTATATGTTCGATCAGTTTGGAAAAAGATTTT
TCCCAAGTATACATCTTGATCACTTGTATCCCACAAATGACTTCACTCATGAGTAAAACCCCTTTCATCTGTTTTACCAGCTGTTTTCTGCCTGAATT
GCGTTGTTAGCTTTCCCATCACCATTTGAAAGGGAATAAATAGAAGCATAATYAAGAATCCAGCTAGCCCGGCATATCCAGCTACCTCAAAAATCAA
ATACATCACAACTATGCTTTCCAACGGACCAACCCAAAGAAAATGCCAATATCTAAAAAGATTGTCAAACCTGTTAACATCATTAGACAGCAGGTTT
ATCATCTGACCGATCGTTGTGTTCACTAAAGCTTTTTTGCTTAATTTCAAACTTTTTCTGTAAATCAACGAACAACAAGCTATTCTGATTTTCATTC
CAAGGTGTTGCAGTGAAAGATGGAATGAATGACAGAAAATTACGTAAAAAAGTGAACACAAAACTATAAACAGGGAGTATATGTATGCTTCTGTGAG
AGTTACTGACGTCTGGTTTGGTTCGTAATATTCTATAAGCTTCTTTAAGAGTAAAGGCTGTGATAACTTTATACAAAATTCTAAAACAAAGTACAA
CAGGCGTGAATGAGGATATCCTTTTTGAATAGTTTCCAAATCGGACGTATGAGGCATGGGTGTGAGTGTTCTAGTTTTTCTAACTTCCAAAACGCTT

CGAGCTGATCTCCGAGGTAATGAGACTCATGTTCTTCAAGGGGCTTGTATAAGTCGTCTTCATTTAGTACTTTGTTCCATCCCTTGTAAAATACTTT
TAAACCCCAACAAAAAAATAAATTTGACAAAAAGTTTGAGCTTTGTTGTGGATTTTTACCCTTATTCTTGAGTTCATCACAATTGCCATCATCCATA
ATTCATATACCAATTGTTTTAATTTTATATTTGCCACGGGTAATTTAAATTTAAATAAAAAAACACGTTAAAACTGTAAAACCAAACACAAAATGC
TTCACTCTGTCCTAAAGTTGTTCGTTGAGCCGTGAACCATAAAATGGCAAACTTTTTAAAACCGATTAAATTGGGTTTTGACAAACTCTTAATTACT
ATCAGTTTTTATAGTATTATTTTTCTTTTTCGATAATTATTTCAGAATTGTTCGTTACAGTTTGTCAGTTAAGTATCTAAAGTTATCGCTAGTATTC
GRTATAACGCTACGCTTGGTCGKCTGTAACAGACGATCAGTTGTGTTGTGTGGATTCAGACAACTGTTT
>Dvv ABC-C T41602_c0_seq1
TTATAAAAAAATATTAGATTACAACGTAACTAAACACTAAAAAACTATTAAATTAATTAAGTGTACATATATGACGTTTATTTAAATATACAACTCT
CTATATGTTTCACAGTTCTAATAATTCATTATAATTTTTTCAAGTCGATTCTTTTTCTTTAGAAAACAGTATTGTTATCTCTAAGCATCTTGCTTTG
TTCTGCTATTGCGAATAAATTCTTTGCTGTTGCTTTGCCTGTTTGCATAACTAGAGAATAGAAAACACCATTGATATTTTGTAAAAGAGTAAAAGGA
TGGTCAAACTCTACTGCTTGTCCAGCATCCATCACTAGAACTTTGTCCGAATCCATAATAGTATTCAATCTGTGAGCGATTGTTAATACCGTACAAT
TAGCAAACTTCTTTCTGATTGTTTTTTGAATTAAAGCGTCTGTGTAGGGATCAACGTTTGCAGTTGCTTCATCCAGGACTAAAATTTTGTTGTTGCC
AACTATTGCTCTTGCTAGACAGACCAATTGTCTCTGTCCCACACTGAAATTAGTACCTCCTTCAGCCATTTTATTAGCTAAACCTGCTGGCAGTTCT
TCTACTGCTTGCTTTAATTCCACCTCTTCCAATGCGTTCCATAGGGCCTCATCAGTATACTCGTCGAAGGGATCTAAGTTGTTTCTTAAAGTTCCAG
AAAATAGTACTGGTTCTTGAGGTATAATTGAAATTTTTGATCTAAGAACCTTTAAGGATATTTTCTTGGTATCAAAATCGTCCACCAAAATCCTGCC
CTCAATATGGGTTAATCGGAAGAGAGCTTGAATTAAAGATGACTTTCCTGCTCCAGTTCGTCCTACAATACCAATTTTTTCTGAAGGTTTGATAGTA
AAACTGAGATTTTTAAGTACAGGAGGATCATCAGTTGTGTAATAAAGCGACATATTATCAAATTGTATTTCTCCAGCATGTGGCCAAGTTTTCGGTG
GTTCGATGCTATGCTCATCGTCTTCCTGTTTTAAGTCTGCGTATTCCTGAACTCTTTCGACGGATGTCATTTGGTTTTCAAGTTCACTCCATTGCCG
CATGCCCCATTGGAACATGCCAGTGAGTCCTAATGTTTGAGTGAGTGCTAGACCCATATTGCCACCATATTGTTCACTCTTTACAAAAACTAACGCC
ACTATGACCAAACCAACGTATATAACGCATATGAAATCTAACCAAAAACCAAAAGATCTATTCGCCGCAAGATACATAAAGAATGCAGCACTGTGTA
AGTTTTGATATCTATCGAATTCTTCTTTAAGAATATCTTGTGCCCGAAATGCTCTTATGGTCGTTAATCCCTGCAAGGAAGCTGTAAGGTGAGTATA
AATTGGACTTCGAGTAACTGATTCGACTCGTTTTACATCACGACTTGTTTCTAAGAAAATTACTCTCAAAAAGTAGAAGAAAACTGCTATGGCCACC
GTGGGTATCATAATCCAATATGTAAGACTTCCTATTACAGTAGTGATAGCGGCCACAGCCAGAGCTATTTGTACTGTATCCAGTAAGGTCGTTGGTA
AGACCTCATCAACTGAACCGATATCTTTAGAAAATCTATTAAGTATCCTTCCAGACGGGTTGGTATTAAAAAATCTCATTGGGCTATATACTATATT
CATGAACATCTTATTGTGCAGTTTGGTTGATGCACCTATACACCACAAGTAAAAGCTTAGTGATCTAATGTTAACGATAATGACGACTGATGCTACT
AGAGCAGTATAAAAATAAACAGTATTCTTCGAAGTGAACACAGGGGTCAGCCACCAATTTTCGAATGAGTGATAAACTACTGCGTCTGGATCTGTTG
CATTTATTATAACTTTTGTTGAGTTCCATTGTTGGATATTTACCCAGAAAGTTACAAAATAATCCGATAAACTATCCAAAGCCTGGCCAAAGAAAAA
CATGAAGATCAACAGAAACACTTTAATTAGATTTCCTCCCGCTCTGGCATAACTCATGTAAACGTGGTCAGATATATTACCACTTCCTGTACTTTCG
CGTTGTAAGACCTGGACAGCACCTTCTTTCTCGGGCTTTTCAGATTTAGTCGCAACTCGCTGAGATTTTCTTCTGGTTTCTTCTTCTTCCTCTTCAA
TGTCAGCTAACAATTTACTGTATTCTGTATTCGACTCTCTTATATCAGAGTAAGATCCTGATACTGCTATCTTACCGTTTGTCATTAAGTAGATTTT
ATCAATGTTCTTTAAATGTTGTAATTGATGCGTTACTAATACGACACACTTATTACCGAGATATCCACAAATACATTCTTCAAATAGATGTTTGCCT
ACATGAGTGTCAACAGCTGAGAGGGGGTCATCTAATAAATATATATCAGCATCTTTATAGACTGCTCGAGCTAAGTTTATTCTTGCTCGTTGTCCGC
CACTAAGAGTAACTCCCCTTTCTCCTGCTAAAGTTCTATCCCCGTGCGGGAATAATGTGAAATCTCTTTCTAAAGCACATACTTTAACTACTTCATC
GTATTTTTCCTGGTCGAACTCTTCTCCAAAAAGAATGTTTTGCCTTATGCTGCCACCAAACAACCATGGTTCTTGGGAAGCATAAGACACTGTACCA
GTTATATCTACGGAGCCACTTTGAAGTTCTAATTCTTTCATGATTATGTGCAATAACGTTGATTTCCCACTGCCTACAGATCCAACAACAGCTACAA
CGTCACCCGCTGTAACGTCCATGTTAATATTTTCCAAATTATTGTCTGGTTGTGATTTTAGCCACTTTGCCGATGCCTTTTTGATATGAATACTCCC
ATTTGACGTTTGTAAAACTTTATCAGTAATATATTTATGTTGTTTTTTGACTTCCTTCAGTTCTTGTCCAGGTTCAAAATATTTGTCAGCCTCTAGT
TCGTCAAACAGTAAAAAGCTTTGGATTCTCTTCATTGATATAAACATTTCGGAACCTTGAGATACTGCCTGGGGGAACCACATGGTTATAGAACCTA
GAAGTCGATAGTAGGAAGCGACTGTGTAGGCGTATGAAGCAGTTAGAGTATTTCCTGAGATTACATAAGTAATTATGCTGATCGCAACAGCACCTTT
ATTTAACACTGTAGTCATAGACATGAGAAGTCCACGAATTAAAGAGGTTTTTCGGATATACTTCATTTCCATTTTTCTTACATATTCTACTAGTTTT
GCAAAGGAATACTCCCAAGTGTACATTTTAATGACCTGGATACCAGAAATAATTTCGTTCATTAGTCTTACTCTTTCATCTGTACGTGTAGCAGTTT
TTAACCTAAATTGTGATGTTTTTTTACCCAACCATGATTGGAGTGGTATGGACAGTAACAAGAACACGGTTCCTAGTAGTGCTGTCCATCCAACAAC
AAGATACAATAGCACCATCACGACGGTTGTCTGTATGGGTGCCAAATAAAAATGATGAAGGTGATGTGTAGCTTGATCGAATCTTCCTACATCATTC
GATAATAAATTTACCATTTGACCTATTGTAGTCTCCGCTAAAGCAGATTTGCTGAGTTTTAAAGACTTTCTATAGATTAGTGAACAAGATGCTATTC
TCATCTTCATTCCTATTTGCATTACGGCCAAATTGAATCTGTGTACCATTATTACATTAAGAAATGAAGTAACTATGATTAAACCGGCGTACAAATA
AATGTCGTTCATGCTCTCTTCTGTTGGTATTTCTTCATACACCTTTAGCAGTTTTGAAATCAGAAACGGTTGAGCCATTCTTATACCTTCAGTTGTA
AAAGCCAATATGTTTAATAAGAATATTTCGCAGAAAAATGTTCCTATGAGTACTTTCCAGTAAGCTGGGTTTTTGTTATTTTTAATGTGTTTGCTCC
ATCTTCGCTCCAATCTTGCACCCAGCGATCCTGAATCATGTTCAGATCTATGTTTATACATATCATCTTCAGTTAGATCTCTTTTATACCCTTTGGC
AAAGTAAGGAATAAGCCAGCAGAAAAATAGTCTGGATATTGCATTTGTTCTTTCTATTGGATTCGATTTTTTCGGAGGTTGATCATTTCTTATTTCC
TCCATTTTTATTGTCTACCGTATATTGTTGCAATCATCCAAATATTTAACGTGTGTGCACTATTCAAATTTATCATGAATGCACACCGGTTTTGGCA
ATTACAATGCAAT
>Dvv ABC-C D10642
TTGGCAAGTAACAAAAATTATAACTGGTCTAAATACTGTTCTTGAAAAGTATGAAAAGATCACAGCAATCTTAAAATTAATTCCATGAGATAAAATT
CTTAAACTCTTTTCTTCTCGTAACTTTCTTTGGCTATGTTTTCTAAATTCTTTGCAGTAGAAGGCCCAGTAGCGTCAACCAAATTGTGAAGCACACC
ATGTACGTTTTGTAGAAGAAGATATGGATGATTAAATTCCTCCACTCTGCCGGAATTCATAACCAATATTTTGTCAGAATCCATTACCGTATGTAGC
CTGTGGGCAATTGTCAGCACCGTACAATCCGCAAACTTATCCCTTATGGTTTTCTGAATCAAAGAATCGGTGTGTGGATCCACATTCGCGGTAGCTT
CGTCCATGACCAAAATCTTATTATTTCTAATAAGAGCTCTCGCCAGACAAACCAACTGCCTCTGGCCAACACTGAAGTTGTTTCCTCCTTCTGTAAC
TTCTGTATTGAGGCCTGCTGGCAGTTCGGATATTACATCTTTCAGTTCAACTTGATCTAAGGCGTTCCATAATACGTCATCACTGTAGTTTTCGAAA
GGATCCAAGTTTTCACGCATTTTACCTGAGAAAAGCACTGGTTCTTGTGGTATAATCGAAATATTAGTCCTAACCAAATCCAACGGTATTTGCGTCG
TATCTACACCATCGACCACAACAGTTCCTTCAAGTGGATAGAGTTGAAAGAGAGCCGTTATGGTTGACGATTTACCAGCCCCAGTCCTTCCAACTAT
GCCAATTTTCTCCTTGGGCTGTACTGTAAAGTTCAGTCCCTTAAGTACAGGCGGATCCTGTGGACTGTATCGCAAACTGACGTCTCTGAATTCAACA
AGACCCTTCTCCGGCCACGGCTGAGGAATATTTTTAGGTTCTTTCCGTTCAGGTTCTGTTTCCAAACGGGTGTATTCCAAAATTCTTTCTACTGACG
TCATGTTGTTTTCTAGTTCGGTCCATTGTCGCATTCCCCACTGGAGTGAGCCCATTAGACCCAGATACTGAGTGATGACTAAACCTAAATCACCTCC
ATGGATTGTGTCTCGGAACAACAAGAGCGAAAATACTGCCACAGCTATGAAGACAATACAGATTATATCTAACCAAAGTCCAAAACATTTACTGGAA
GCCAAATATAAGAACCATGCTGCACTGTGGCTATCCTGGTAGTTGTCGAATTCTTCCGTTAACATTTTTTCAGCGTGGAAAGCCCTAACTGTGCTCA
ATCCATTAACTGAAGCTGTGAGGTGACTCAACATTGGACTTTTTGTAATTGCTTCAACTCTTTTAACACTTCTACTGGTTTCAGAATACACTATTTT
CAATAAATAGAAAATTATCATTAGCACCACTGCAGGTACAGTCAATAAGGGCTCTACTATTGCAGATAGTGTAATAGCTCCAAGAAGAAGTAATGCA
ATCTCAATAACATCAATGATAACAGAAGGCAAATATTCATCCACAGTTCCTAGATCCTTCGAAAATCTGTTTAGGATCCTTCCAGATGGATTTGTGT
TGTAAAATCTCATTGTCGCTTTGATTAGTTTGTCAAATATATATTTATGTAGATTCTTAGAAGCTATTACGAAAAAAATCATAAAGTACACTGCCTT
GACTACCGATATGATTATGTTACCAACGATAAGGGCAGTGTAAGACCACATGATCAGATCCCTGTTTATAGTTTCATTGGGGGCGGTTAAATTCATT
CTTTGCTTCTCGCTAAAATCTTGTTCTAAGTTCACCCAATAAGTAACATAGTATTCTCCGGCATTAGCAACTACTTGACAAATAATAAACAGGAAGC

CCAGTGCTAACATCGAACAAATCCCACCGCCTGCTTTTAAGTACAAATAATAAGTAGCTGCCTTAATGGTACCCTTCTCCTGCATTTCCTTCTCCAG
AACTTGGTCCTCTTCATCTTCCATTGGCTCATCGTAAATTGACGCTTTTGATTTAATTGACTTCATCCGTTTGTCTTCTTCGGCTTCCTCGTTAAAT
TCTTCCATAACTTTGGCGAAATCTAGACCGCTCTTCTTTAGTTCTGTGTAGCTTCCAGTCATCTCCATCTTACCATCTTTCATGATAATGATCTTAT
CGGCGTTTCTTAAGTACTGTAGCTGATGTGTAACCAGAACACATATTTTGTTTGACAAGAACTGTTTGATGCATCTATCGTAGAGATGTTTTCCTAC
GTTAGCATCAACAGCCGATAGAGGATCGTCTAGAAGATATATGTCGGCTTTTTTGTACACGCATCTAGCCAAATTCACTCTAGCTTTTTGCCCTCCA
CTTAGAGCTTTTCCTTTCTCTCCAACCAAAGTTTTATCTCCGTGGGGGAATAAGGCAAAATCTGATTTTAAGGCACATACTTCTACAACCAGTTTGT
ATCTTTCTTCGTCATATTCCTCACCGAACAAGATGTTTTGACGTACGCTCGCAGCAAAAAGCCATGGCTCCTGGGAGGCATAAGACACTTTCCCATC
TATTTCTAGCTTTCCTGATTTTACTGGTAGTTCTTTTAGTATCAAGTTTATGATGCTGCTCTTGCCGCTACCAACTGGTCCAATGATCGCGAGCATT
TTATTAGGAGGGACATTGAATGTTATGTCACTTAGGTTATTTTCAGGAGACTGTGCTAACCACTTGGCAGAAACTCCAGATAGCATCAATTTAGGCT
CTTTGACTTTTTTAATTAATTCAATAGCTGTACCGTTTCCGTTAATACCGTTCTTCATGCCATCTAGTCCGTTCAGTTTAGATTGATAGTTACCATT
CATCTTTTCATAACCATCTTCAGAGTCTTGTTCTCGTTCGTCAAATATCATTAGCTTATGTATCCTATCTATAGTAACGTGCATCTCGGCCAATGCT
GTTATACTTAGAGAAAATATGATGGTAATGATTGTTCTCATCTGGTTGTAGATGGCGGTAATAGCAAAAACTTTATCGGCTGATACATAATTACCGA
GAAGGACGTAGCCTACGATGCTCACAAATATAGCAGTTCTGGACACGAATACTTCGAATGAATAGAGTATTCCTAATAGGCAGGAGTGTGATCTGAT
TGCACGCATTTCGGATCTCCGTGCCAAATCAATAACATGAGCGAACGGTTTTTCCCAGCAATACATTTTAATCACCTGAATACCGGAGATAATCTCG
TTCATTAGTTTGACCCTTTCGTCCGTCCTCAAAGCTGTTTTCAGACGCATCACAGACATTTTTTTGGCTAACCAAATTTGAAGCGGGACGAACGCCA
CTAGGAACGCCATACCAAAAAAGGCACTCACTCCAATCTCTCTATACAATAGCCACGTCCCAAGAGCGACTTGGATTGGCCCTATCCAGGCGTAATG
AGTCAAACCAAATAACTGGTCAAATTTGCTTACGTCATTGGATAATAAATTTACTAATTGACCTACAGTAGTGTTTCCTAAAGCTTCTCGGCTAAAC
CTCAAAGTTTTCCTATATATCAAAGACGAGCAAGCCACCCGTATCTTCATGGCGATGTGCTGTAAGCCCATATAGTTTGGTTGTTGCATCACAGCAT
CTAATAAAAGTGTAATCACAAGTGCAGCTGCATAAATTAACGCCTCATCTTTCGTTATTTTGGTCTGTCCTGCTTCCAAATATGCAACTAGCGTTCG
TATACAGTAAGGCATTACAACTATTAACAAGAGTTCGTCGAACAATCTGATTAACCCTAAGACCGTATATCTAAGTCCAAATATCCTAAATAAAGCC
CTATGTAGTGCAGTCTTTTTGTGTATCCTATATTCTTCTTTCCATGCTTTTTCTAATTTTGATCCTAGTCTACTAGATGTGTGCTCCTTCAACGGAG
AAAATAGTTCTTCTTCTGTTAATTTGTACTTGTACGTTTTTTTGAAGATAGGGAACATGTAGAAAAAGAAGAGAAAAGAAATCGGATTGGTATTTTC
CGCAGGGTTGCGGGGCTTAGTGGTCTTGTTCCCCGTGTCCATTGTTGCTTTAAACACGGTATATCACAGTAAAATTCGTAAGTACACTTCGTATGTG
TTTCGAGAGTCCGCGAACTGCAGAGCGGACAATCGCACACATCTGTCCCAAT
>Dvv ABC-C D12562+D20321
AAAACACCTGGACATGACCCTTGACGTGAGGCTCCACTATAAATCTCACATAAAACAGAAACGACCAGAGCTTGGAATAAAGTACAAAAAGATTTAC
CGATTGCTACGCAAAAGTTAAAAGTTAACCACGTACAATAAACTGATTCGATATAAGCAAATACTAAATCCTATCTGGACCTATGGCATACAACTTT
GGGAATGTACCAGTAAAAGAAACACTACAGTTATACAGCGATTCCAAAACAAGACACTGAGGAACATCGTCGATGCACCTTGGTATTGTCAAAATGG
TGATATCTATTGAGACCTTAAAATTAATACTGTGAATTAAACCACGTATAAAATTACCGGCGGTCACGAACAGCGGCTAATTCAACATGTGAACGTC
GAAACCATCTATCTCCTGGACACCAAGGATATAGTGAAAAGGCTCAAAAGAACAAAGCCTTTTGAGTYACTGAATTAATTGATACTACAAAGCAGAG
TAATGTGCGGGAAATAACCTACATAATATGTACAGCATATATTATATGCTTTTTCTCATGAGCATCTTTCAGTGCGTCACAGTTTTTCGATTTCATT
CTAACGCATTAAATTGTATGTGACAGAAAAAAACGAAAAAATCTCGAAATGGAGAACCTTTAGTTTCAAATAACTCAAATGTGGTGCAATTTTTTGG
GAAAACTTAACAGACATGTTTAGAAGTTCACTAAAAAACCTTTCAAATGAGCTCGGATAAAAGCTTTTTTTGCATAAGAACTGACTGACTTATGACG
AAAATAAAGTCGATCCCTGCTTTTTTTCGGAATGTAACAATTAAACACTCGTCATTACAATCCTAATAGAAATGAATAGCTTCCCTCTTCAAATTAA
CTTTATTTAGGTATAGTTGATACGATCCACGTGATTTGACCGGTTTAGAATGCTTAGTTTAGAAAAAATTGTTGATTCGATCGAAAATGACATTATT
ATAATTTAAAAAAAAGTGCATTTTTCTTAAATAAATCTAAAAGTATTCATAATATGCAACAATAATTTAAATATTCATGGTAGATTTTGTTTACTGA
AAATTTTGATTTTTCTTTTTGAATCATGAATACTTCCGACCTCCATATCTATCGACTTCCATAGTTATTTTGATTTCATTTCCTCGATGGGATTGT
TTTTACCCCCAGGACAAAATCTACTAGAGGGTAATTTTTGAAAAAGAGGGTCAACCGAGCTTACATCCAAATTTTCATTAAAATCGGTGTTGATGCT
TAAAATTCCACGGTTTTACAGTTTTTTAGCGCTGATTACTGCTCTAATAAATAAAAAAAGTTTCCTATAGTTTGTCCTAAAATATAAGTAGCAGCCC
CTCATAAGTTTTTAAATTTACATATCTATAGGTTTGTTACACTAAACTAAGCACCATATTGTGATGTCATAAGCGTTTAAACACTGGTCAATTATTC
AAGAATTCCTACCTATCTTATATACATACAACTAATAGGAACAATGTATATGTTACATTAATACAAAAATACTTTCACTGGAAATTTTAAATTCTCT
ATTATGTATATTTAAAGAGAATTAATTGGAGATAATTAATTAATTATTAATTAATTTCAATACTAAAATGTATGCTTAGTCTCTATTTAGTTTTCTG
GCTTTCTTCTGCTATTGACATTAGAGTTTGAGCCGTCCCCTTGCCCGTCTTCATCACCAATGAATAGAAAGTTCCTTGTTTATTCTGCAATAGCGAA
TACGGATGATCAAACTCCACCATTTTCCCAGCATCCATCACCAAAACTTTATCTGAATCCATTATAGTATGTAATCTATGCGCAATCGTCAACACTG
TACAGTTTTCGAATTTCTTCCGGATGGTCATTTGGATCAATCCATCTGTGTGGGGATCTACGTTCGCTGTGGCTTCATCTAATACAAGAATTTTGTT
GTTTCGGATTATAGCTCTTGCTAGACATAGCAACTGTCTTTGGCCTACACTGAAATTGGAACCTCCTTCCGCCATTTTATTGTCCAGACCAGCTGGC
AACTCTTGGACGGCGTGTTTCAGTTCAACTTCTTCCAGGGCGCTCCACAGAACCTCGTCCTTGTATTCGTCAAAAGGATCCAAATTTTTCCGCAACG
TTCCAGAAAACAGCACGGGCTCCTGCGGTATAATAGAAATATTTGACCTAAGTTTTTTCAAAGAGATAGTCTTCGTATCAACACCATCGATAAGAAT
ACTGCCTTCAATGTGGGTTAGTCGGAACAAAGCTTGAATTAAAGAAGACTTTCCAGCTCCAGTTCGTCCTACTATACCAACTTTTTCTTTAGGTTTG
ACCACAAAAGTCAGGTTTTTCAACACAAATGGATCATCAGGAGAATATTTCAACGACATGTTGCTGAACTCTACTTCTCCTTTTTCGGGCCAGGTTT
CTCGTGGTTTGTTCGAGGGTTGGTCTTTTTCTTGCTTTAGATCTGCGTATTCTTGTACTCTTTCCACAGACGTCATCTGATTTTCAAACTCGCTCCA
CTGTCGCATTCCCCATTGGAACATGCCCATTAAAGCCATAGCCTGTGAGAGTGCTAGACCTATATTACCTCCAAATTGCTCACTTTTAATCAATAAT
AAAGCTGCCAGTACCAATCCAATATAAACCACACAAACAAAATCAAGCCAAAATCCAAACGTTCTGTTAGCACCTGCGAAGATGAAATAAGCCGAAG
TATTCAAATTTTGATAATTGTCAAATTCCTTTTTGAGTATTTCTTGGGCTTTAAAAGCTCTAATTGTTGTCAGACCTTGTAATGATGCAGAAAGATG
TGTGAAAATTGGACTTCTGGTAACAGCTTCTATTCGCTTAACATCTCTACTGGTTTGGACGAAAACTATCCTCATTAGGTAGAACAATACAGCAAGC
AATACAGTTGGTATAACAATCCATATGGTTAGACTGCCAATTACCAAACAGATAGAGAAAACATATAAGCCAATCTGAACAGTATCCATAAGAGTCA
TTGGTAAATTTTCATCCATAACACCAATATCTTTCGAGAACCTGTTTAATATTCTTCCAGAAGGATTGATGTTGAAGAACCTCATAGGACTGTAAAT
AATATTCTGAAACATCTGGTTATGTAGTGTTGTCGACGATGTTAAGGATAAATGATAAAAATAAAGTGACCGTCCCAAGATCATGATGATCAAAGAT
ATTACTACGAGTAAGTAGTAGATGGCAGTATATTCGCTGGTGAACAAGGGCGTTAACCACCAACCTTCGTACGGGTTTTCAGGTTGAACGGTGGTTC
TCAGCGCTTTCGGTAAAGATGAGTTGGATGTTATATTTAAATTTGTTGTTATATTTGAATTGGATGTAACTGTTTCAGTCAGGTTTTGGGAATTCAC
TTGTTGAATGTTGACCCAGTAAGTAACGTAGTACTCAGAGAAACTCTCCATAAACTGACTGGTAATAAAAATCATACTGATCAGAAATATATGGAAA
AGGTGCCCCGAAGCTTTAATATAGTTCATATACACACGACCAGAAATTTTGCCTGTACTTTGAGCCTCCTTTGCCAATATTTGAATTTCGTTATCTT
CGAACATTTCTTCCTCCCTAGAAAGCGTTCTTGTTCTCGACATTTTCCTAGTTTCCTCCTCGTCCTTTATATCAGTAAGAAGTTTACTGTATTCTGT
ATTGGAATTTTTTAAATCCTGATAGGTACCAGACAATTTAACCTTTCCATCTCTCATTAAGTATATATGTTTAAGTTTTTTTAAGTACTGTAACTGA
TGAGTTACTATAACTACACATTTGCTACTAAGATACCCGGTAACACAGTCTTCGAATAGTTGTTTTCCAACATGGGCGTCAACTGCTGATAAAGGAT
CATCTAGTAGATATATATCAGCGTCTTTATACACAGCTCTAGCTAAATTAATCCTAGCTCTTTGTCCCCCACTAAGTGACACCCCTCTTTCCCCAGC
TAAAGTCCTGTCACCATGAGGAAATAAAGTAAAATCTCTTTGTAACGCGCATACTTTAACTACTTCGTCGTATTTACTTTGATCGAACTTTTGTCCA
AAAATAATATTTTGTCTGATACTACCTCCAAATAACCAAGGTTCTTGTGATGCGTAAGAAACAACCCCGTGGACGTCCACAGACCCTGACTGTAATT
CCAGTTCCTTTAAAATAATGTGTAACAAAGTAGTTTTACCGCTACCTACAGGACCCACAACAGCCACAACATTGCCCGGAGTTACGTCCATGTTAAT
TTTCTCTAAATTATTTTCAGGAGAAGATTTTAACCATTTCGCTGAAGCGTTTTTTAGGTGTATACCTGGTTCTTTGTTTCTAATGGTGAGTTCCAAC
GATTTAATTCCGTTCTCTTTGTGTGATTTTTCGTCTAATGCTATATAGTTCTCATCCTCGACTTCGTCGTATAATAAAAATGTTTGGATTCTCCGTG

AGGAGACATATAACTCCGCAGCTTGGGAAATGGCCGTGGGTAAGAAATTGGTAACTGTATACAAAAGCCTGTAGTATGATGTTACGGTATAGGCATA
GCTAGCAGTTAAGGCGTTTCCAGTAAATACGTATACTAGTATGCTAGCTGCTATAGCACTCCTGTGAAGCATAATGATACAAGACATTAGCACCGAT
CTGATCACTGAGGTATATCGGATGTACTTCATCTCCTTCTTTCTAGAAAATTCTACTAATTTAGCAAATGGTTGTTCCCAAGTGTACATTTTGATTA
CTTGTATTCCCGCTATAATTTCGTTCATCAGTCTTACTCTCTCGTCAGTTCTGTTGGCTGTTTGAAGTCTAAATTGTGATGTCTTCTTACCCAACCA
AGATTGAAATGGAATAAACAATAATAATACTATAGCTCCTGCTAACGCAGCTCCTCCCACGTTTACGTACAATAAGTACATAACAATTAATATTTCT
ATCGGGCCTATGTATATGTTATGGAAAAAGTAAACAGCCAAGTCAAATCGGGACACATCGTTTGAAAGCAAGTTGACCATTTGCCCTATTGTAGTTT
CAGCCAAAGCAGACTTGCTTAATTTTAGAGCTTTTCTGTATATGAGGGAACATGACGCCACACGCATTTGCATACCTAACTGCATAAGTCGTAAATT
AAAGTTGTGCCCCATGTTGGCTGCTACAAAGGAAGTTAGGACGATTAAAGAAGCATACAGGTAGATTTCACTTCTTTCTTCATCTAGCTGCTTGGAC
TCATACACTTTTAAAAGTTCTGCAATGAGGAATGGCTGAGATATCCTTACGACCTCTGCAATGACTACGAGACAATTGATACCTAATAGATCTAAAA
CAAAACTAGAGGCACATGCTTTTATCAATGACGGCTTCTTCTTTCGTAGTTGCTTATTCCATTTTACTTGCAGCTGGTCACCTAACTTGCCCGATTC
GTGGCAATTTCGATGTTCATACATATCATCTTCGGTGAGATCTTTCTTGTAACCTTTCACAAAAAATGGAAGAAGCCAAAAAAAAAATATTGATGAA
AATATATTAGCAGTGTCTAAAGGATTTTTCTTCTTTGCTTTGGCTTTGCTGTGTATTTCTTCCATATTTAAGTTGTGTTGTTATGTACATGTAGTCC
TCTTTATAGATATTACCGAACTGATTCACTTCACTTTAAGGGTTAAATTACTTGTCACACTATTTTTTATAATTAGAGGGATTCCTATATTCTTAAT
ACATTTTTTTAACAATTTTGTACATTCTAAC
>Dvv ABC-C D15305
CACATAACCWTATGCCCGATGTGTGTAATTTCACATTTCTTTGTTATATTTTCTTTTGGAATGATGTACTAAAGGGCCGGTTGTTCGAACGCCAATC
AACAATGATCACTATCAAATAATTATTTACTGTCAATGTCAACTTTGTTTGGGTTGTTAAAAAGTCTGTGGAGTCTATAATCAATTAATATAACAAT
ARTTATTAACATAATTAATAATAAATCTCATAATTRTAATTAATTATGTTTTYAGCAACCCAAACAAAGTTGACATTCGACAGTTTTGGTGACAGCA
ATTAAATATTTGACAATAATCATTGCTGACTAGCGTTCGAACAACCGGCCCTAAGGTATTAACGAATTCATTAAACTCAATATATCGGCACATAGAA
CCTAATAAAAAAATTTGTAAAATGGTAATCTTGAACATAAGATATGCGCCTTAGTTAGCAGTTATCGCATTGAGTAAAATATAAGACTAGAGTTTAA
AATTAATTAACCTTAATATAATTAACGTTAAAATAAATCAATGTAACATACTTAATGGCAACAAAATTTTTAACAATCTAATTGTTTTTGTACTCTT
TGATTTAGTAATTTTTACTCTGGTAATGCATTTAGTTTTTGATAACTATCTGAAGCAATTTTTCTTAGCTGTTCAGCTGTTGTTCTACCAGTTTCAG
CAACCATACTGTAAAAAACTCCGTTTTTATTTTGAAGGAGTAAGTGAGGGTGATCAAATTCAGATATTTGTCCAGCGTCCATAACTAATACTTTATC
AGAGTCCATGATTGTGTTCAGTCTATGTGCGACGGTTATCACAGTGCAATCTGAGAACTTTCTTCTGATAGTTTTCTGGATTAAAGCGTCAGTCTGT
GGATCTACATTAGCCGTAGCTTCATCCAACATAAGAACTTTGTTATTTTTAAGAATAGCTCTTGCCAAGCAAATCAACTGTCTCTGTCCAACACTGT
AATTGGAACCTCTATCCATGACTCTATTTTCCAATCGGTTAATAACATTAGCAGGATCTTTTAGTTCTACGTCTTGTATAGCTTTATACAAAACTTC
GTCTGAATATTCTTCGAATGGGTCCAAGTTGTATCTTAAAGTTCCTGAAAATAGCACTGGGTCTTGGGGAATAATCGAGATTTTTAAACGGAGATCA
TTGAGGCTAATGTCTTTGGTGTCGATGTCATCGATTTTAATTGAACCTTCAATTGGGGCTAGTCTAAAGAGTGCCTGAATTAAGGAAGATTTTCCTG
CACCGGTTCTGCCAACAACTCCAACCTTTTCCTTTGGTTGAATTTTAAGATCAAGATGTTTCAAAACCAATGGTCCACCCTCAAAGTATTTTAAACA
AGTATCTTTAAAAGTGATTTCTCCTTTCTGAGGCCATTCTTTTGCTGGTTTTTGAGGTACGACCGGTTGTGGTTCTTTTTCCAAATTTTGGTACTCC
AATACCCTTTCCACGCTCATAAGTTGATTGGTTACCTCGGCAGATTGACGCATACCCCATTGTAGTAAACTTGTTAATGCTGTTGCTTGAGTGATGG
CCAGACCTACTTGACCTCCCTTCAGATTAAAGGCCTCCCCAAATAATAGTAAACTAAAAGTAAGTGCTGCCAAAAAAAGAGAACAGAAACAGTCCAA
ATAAAATCCAAATGCATTACTGGAAACTATAAACATGTACCATGCACCTGAATGATAATCTTGGTGATGATCAAACTCGTCCATAAGAGTACTTTGG
GCTCCAAAAGCTCTTATGGTTGTCAAACCTTCAATGGTAGCTCTCAAATGGGTAAACACGGGACTTCTCATTATTCCTTCTAACCTCTTAATATTCT
TCGAAGATTTAGGTAGACATGCCTCATAAAACCAAAAACTACACTCAAGGCGCCAATCAGAATTAAGAAGTAAGGATTGACTACGACTATCAGTAA
GAGAGATCCACACATAGACAAAAGCATCTGTGAAGAATCCAATATTGCTTTCGGTAGTAATTCGTCAACACATCCCATATCTTTAGAAAACCTGTTG
AGAATTCTTCCTCCTGGATTGGTATCGAAGAATCTCATTGTAGCAGTGACAATATTATCAAAAATTGTCGCATGCAGTTTTTTGGAGCTCCACATAG
CAAGTTTATAAAAGAAGAAAGATCTTGTCATAGCAAGAACAAAAAGAGCTCCTAAAAGGAACCCATATATATAAAGGCAATTATTTGTTGTCAGTTC
AATGGATGGAATAATGAATGTAGGTGTAGTAGAATTAGAAGATTCGCTGGAGCTCCTGAATTCCTCTACGTTGACCCAATAGCTTACCCAGTAGTCT
ACCCCACTGGCAGCTAATTGAGTTCCTAGATACAATATAACGCATATTAAAACAGCAAATGAATTAGAACCTGCCAACATATATTTAAACAACAGAG
AACCGCCCACTTTTCCTTTTGATGATTGTTCCTGAAGGTCTTTGACATCAAATTCGAGTTGCTTTTCAGCTTCTTCTTCTTCAAAGTCGATATCATT
CGATATGAGAGTATCCGCAATGCTTAATTCACTGACTATACTTGCCAGTGAATCTTTTCGACTTCTTATTGAGAGTTGCCTATTATATTTAGATCTA
TCTTGTTTCTTTTCTTCTACATTTTCTGGCTCAGCAGTTAAAAGTTTTGCATAAACATTGTCACTGTTTGCCAAAAAATTGAAGGTTCCTTCATCTT
CTATGCGTCCATTATTTAAAATAATAATGTTATCCGCATCTTTCAAATAGTGTACTTGATGGGTAACTAAAATTCTCGTTCTGTTGGCCAAATATCC
ATTAATACATTTTTCATATAGATGTTTGGAAACGTGTATGTCCACAGCTGACAACGGATCATCCAAGAGATATACATCGGCTTCTCTATACACAGCT
CTTGCTAAATTTATTCTAGCTTTTTGGCCGCCACTCAAAGATGATCCCCTATCACCTACTAGTGTTAAATCTCCATTTGGGAACTGCTCAAAATCTT
TCTCTAAAGAGCAAGCTTCTATAACTTCATGGTATCTCTTTTTGTCGTATTCTTCACCAAAAAGAATGTTTTGACGAACAGTGGACGCAAAAACCCA
CGGCTCCTGAGAAGCATAGGATATTTGACCTCGAACTCTCATACTGCCTTCGGTTATGTCCAGTTCCCCCAGTAATGTTTGCAATAATGAACTCTTC
CCGCTACCAACAGGTCCAATAATTCCAACCATCTGTCCATCAGAAATATTTATATTTATATTCTCAAGAGCATTGTCACTGAAGCTTGCATTCCATT
TTACAGTCAAATTTTGCAAACTTATAATTTCTTTAAAATCGTCTACACTGTTTACATTGCCATTCATAGCGAAATTTGTTCTATTAGGATCGTATTC
TTCGTTTATGAGAAAATCTCTAATTCTTCGAATAGCTACCATAACCTCTGCCATTTCAGATACTCCTCTACTAAACATGGTGGACATAGCCATTGAA
AGAACGTTGAAATAACCCATTACAACAAAAACCTTCGACGCTGTGATTTCTTCGTTCGACAAAATAACTGTCAGCAAAGTACAGAACAATGCCAATC
TGGTGGTGAAAAGATTCAATGCCATAAACGTAGCTCTCACATAAGCAGATTTTGTAATAATTTTTATTTCATTTCGTCTGGCAATTCGGATTATTTT
CCTGAACGGTATTTCCCATGCGTACATTTTGATAACTTGAATTCCTGAGATGATTTCGTCCATTAGTCTGACTCTCTCGTCTGTTTTGAAAGCAGTC
TGTTTTCTGTATATTGCAGATAATTTTCCGGTGTAGGTTTGAAAGGGAACTATCACAAATACTGCCACAACTCCCACTACTCCTGCATACCCAGTTC
TCTGGTACAGTATACCCATTACTATAAGTGACAATACAGGTGCCATCCATAATTGGTGAATGAATAAACTAACAATATCAAATCTACTGACATCGTT
GGATAATAGGTTGACCAGTTTTCCGGATGCTGTTTCTCCCAAGGCTGTTTTGCTCAATTTTAATGACTTTCTGTATATCACAGCACAACAAGATGCT
CTTATCCTCATGCCACTGTGAAAAACCTCTAGCATATATTGATTTGACATCAAAATGTTCAAGATGTTTATGGCCACTAGGATACCTGCATACGTAA
GAGCCTCATTTTCGGAAACTTGTGATTCTGGTTTAAAATAATCCAATAGGTTCCCAAGTATTAGAGGTTGAATGAGTCTTAGCACTAAATCTGTCGT
GATAGTATATATTCCTAGCTTTGCGTACTCCAACCAAAATGACATAACAATGGCTTTCAATAAACTTGGTTTGGAATTTGATTTCTTAGATTTTTCT
AGTTGCTTGTCCCAATTTCTTTGTAAATTATCTCCCAAGACTTTGCTTCTGTCACTCTTTAACGGATTATATAAATCGTCAACTTCCAGCGTCTTCG
ATAAGCCCTTCTTAAATACATCATACGTATATCCAAAAAATAGAGCCGATATAATATTCGCAGTCTCCCGTGGATTGGGATTATATTTCTCTTTGGT
CATTTCCATCTTTGCGATTTCTTAACTATTCTAATTTTCTAAGTAAGTATTCGTAAGTAGTCACAGGACAATTTTTTGTTTTTAATTTAACATGAAT
TGGTTATCTCGGGAATTGATACAACCGCTATTTAAGATATGTTTATCTATCATTAGTGTACATGTTC
>Dvv ABC-C D21941
CGGACTTTTCCTTTTGACCACTTCTTCTTTTTGGCTTTGCCACCTCCAGACCCCTCTTTCTTCTTCTGGGTCTTTTGGGGTTGCTTCGCCGATGACT
TAGAGTCCTTCTTGGGAGGCATSTTGCTCATCCAGAAATAACTGTCTATAATGTTCCTCCGCTATCTTTTTAAGCGCCTTCTCAGTCGCTGTCCCAG
TTTCTTTTAACATTCTACTGAAATATCCTTCATAATCTTTTAACAGTATGTGCGGATGCTCGAATTCTACAGCTTGACCGGCGTCCATTACCAAAAC
TTTATCGGAATCCATTATTGTATTTAGTCTATGTGCTATGGTTATCACAGTACAATATTTAAAGTTCTTTCGAATTGTCTTCTGAATTAAGGCATCG
GTACTGGGATCAACATTGGCCGTGGCTTCATCTAAAACAAGTATTTTGTTGTTTCTTACAATCGCTCTAGCTAGACACATTAATTGCCTTTGACCGG
TACTGAAATTTGAACCACCCTCCCGTATTTCCATTTCAAGGTCAGTAATCGAATGTTTCAACTCTACCTTTTCTAATGCATCCCATATCTGCTCATC

ACTAACAGTGTTGAATGGATCCAGATTATATCTAACAGTCGCTGAGAACAGAACTGGTTCTTGTGGTATTATAGAAATTGTTGATCTCAATTTACTA
AGCTCTATATTGACAGTATCCACGCCATCAACTTTTACTGATCCTTCTAATGGTGAAAGTCTGAACAATGCGGATACAAGTGAAGATTTTCCAGCAC
CAGTCCTGCCAACTACACCAATTTTTCTCCAGCTCTAAACACAAGATTTAAATTCTTTAGAACGGGAGCACTTTCTAATGTGTACCTCAGGAAAAC
GTCTTTAAACTCTACTTCTCCCGCTTGTGGCCAATCTTTATCTAGAGATTTATATCTAGGTATTAAATTATTCACTGACAGTTGTTTCTTTCCTTCT
GTTTCTTTTTCCAAAGTTGTATATTCAAGTACTCTCTCTACACTGGTCATATTGTTAGCTACTTCTGCAGTTTGGCGGACGCCCATCTGTAACATTC
CTGTTAATATCAAGCACTGGGATAAGACAAGGCCAGCGTTTCCTGCAGAACTGGGATCTGGGTTATCAAAAATAAGGAACTGGTACGTAACTATAGC
TATAAAAACTGTACTTATGCAATCAAGATAAAAGCCAAATACTTCACAGGAAGCAATATAGAGGTACCAAGCTCCTGAACTATTGTCCATAAGTCCG
TCAAATTCTTGGATAACCATATCTTCGCACTTAGATGCTCTTATGGTTGTCATTCCGTCTAGGGTGGCGATAACATGGGAATATACTGGAGCTCTAC
TGACTCCTTCCAGGCGCTTCACGTCTTGAGTCGTTTTTAAAAAGAACAATCTGAGATAATAGAACAAAATTCCTATGACTACTGCTGGTATTATCAT
CCACAATATTTTAATAAACACCAGTACCAGGATACCTACAAGGACCAACAAAACTTGTGTACCATCAAGCATCGCTTTTGGTAAAAGCTCGTCAATG
ATTCCCATGTCATTYGAAAATCTGTTCAGGATTCTTCCTGAAGGATTCGTGTCAAAAAATCTCATAGGTGCTTCCAAAATCTTACTGAACATACTGT
TGTGCAAGTTTCTTGAGGCATTCATAATAATTCTGTAGAAATTTAAAGATCTGATTGGAGTCAGAAAAACAGCTAGACATATTGTGACACTGTATAT
AATTATATAATAATTTGAGTCGTGGGTAAACCGTTCTTCAAGTGGAATATTTTCACCCACTACCGAATAGTTGTACTTGCTGTAATTTCCAGACGAG
GTATTATTTGCAACAACTATTTCCAAGGGTGGTGTTGTTGTACTATTAGCTGCCACTGTATTTACCCAATAAGTCACCCACAGGTCTCCTGCATTAG
ATGCCGCTTGGGCCAAAACGAATAAAATTACCATAAAACAAAATCCAAAACAACTGGTACCGAAACGGAAGTATTTGCCATAAGTTGAAAAGGGAAT
TGATCCTTCGCAGGTTTCTTCATTATTTTCCTTGGGCCCTTCTTCATTTTGAGATAACACCTGTTTTATTTCTTCCATATTAGACTGTTTTCTAGTA
TTTAAAGATACATCCTTTTTCTCATCTTCAATCTGTTCTTCACTTCTATTTTCTTGCAATGACTTTAGGTAATCCTCGGACAGTTCGTCAAAGGTTC
CCATTTTCTTGATTTTTCCCTTTTCGAAAACTATAATAACATCAGCATCTTTTAAAAACTGCAACTGATGTGTCACCAAAATCCTAGTTTTACTGCC
CAAAAATCCTTTAATGCATTTTTCAAAAAGATGTTTACCAACTCGTGTATCCACCGCAGATAAAGGATCGTCTAGTAGGTAAATATCAGCGCTAGTG
TACACAGCCCTCGCTAAGTTTAATCTTGCTCTTTGACCACCACTCAGGGAAACCCCCCGTTCTTCTACAATGGTTTTATCGCCAAAAGGAAATTGCT
GAAAGTCGGTTTCTAGAGCACATACTTGAACGACTTCGTCATATCTGTTTTTAATGTAAGGTTTTCCAAACAATATATTTTCTCTTACGCTAGATAC
AAATAACCAAGGTTCTTGGGAAGCGTAAGACAAGTCTCCTGACACCTCTAATTGGCCACTTTTGATTGGAAGCTCACCGAGTAACATTTGCAGTAAA
GAAGATTTGCCACTACCCACCGTCCCCACCACGCAACATAGTTTTCCTGCARTTAACTTTAGATTTAAGTTATTCAGAGTGTCATCCTTCAAATCCT
CATTCCAAGTAGCACAAGCTTCTCTGGCTTGTATGGTTCCTATTTGCTGGACCACAGTATTTTGCACCTGAACATTATTTTCTTCCAAAAGTAAAAA
TTGTTCGAGTCTTCTTAAAGAAACCTTAACTTCGGCATATCCAGCCAATGCATAAGGAAACATGATGCAACTGTACAATTGTAAGGAATTGTATAAT
TGGGCGCTAAGGAAAACTTTGTCAGCCGATATTTGATGTCCTAGCAGAACAAACGTTATGACTGTCAGGTATAATGCAATTCTCTCAGTAACTACTA
TAAATGCAAGTGACAATCCCTTAATAGCTGACGTTTTAGATATTAAACGTACTTCTTCAGTCCTTGCGGTTTCCACAACCTTTGCGAAATAATTTTC
CCATGCATACATTTTAATTACTTTAATACCTGCCACTAACTCATTCATCAATTTTACTCTAAAATCAGTCTTTTCTGCTATTTTTGAACGAAGTGTC
CCTTGAAGTCTTGATAAATAACCTTGCATAACAATYCCTTCGAAAAGCATAAATGCTGTTCCCGCCAGAGCAGCTATTCCAACACTTCGGTACATAA
TGTAGAAGCAAATAGCAGCTTGGAATGGCATTACCCAAATAAAATTTAAATACATTGAGGCAATATCGAATCTTTGTAAATCGTTGGAAAGAAGATT
TACCAATTTACCTCCTGGGGTTTTGTTTAAAGATGCTTGGCTCAATTTGAGTAACTTTCTATAAACTAACGCTGAACATGCTGCTCTTGCTCTCATT
CCAATTCTTTGTGATCCCAAAGTTGAATGGTGATAAGAGAGAGCAAATATAAAGGCTAAAAACACAGTGGCTGTTCCTAATATCCATCCTGAATACT
CACCGAACATGTCGTACTTTTGGGTTTTGTCGAAATATTTTAATAGTTCTGCAAGAACCAGAGGTTGAAGTGTCTTTATTAAAACAACTTGCAGAAA
TACCATTACTCCTTGTAATGAATATGGAAAGGCGTATGTTTTAAAAAGTGCCAATTTCAAACTAGGTTTATCCTTTTTTCCGGCTTTGTGTAACATT
AATTCATTTTCCCAATGTCTTTGAAGCGTATTGATTGTAGCGTGACATGTGTCTGAGTTGAGCGTGTTGTAGATATCCGACGGTTGTAATTTTTTCT
TAAAACCAGTTTTAAAAAGTGGAAATATCCAGCAGAAAAATATTTTTGATAAAAAGTTTGCCGATATTTGAGGTGATGGATTGTTGTGTTTTTTGGT
AGAATCCATTATTAAAGGAAACAATGTCTACCTAAGAATTCTTCTGAATTTAAATTTGAAGCACTATCTAGAACAGCGTGTAAGCACTTTTCGCAAG
TTTTATGGCTACG
>Dvv ABC-C D18709*
ATTTAACCTGTGTGCGACGGTTATCACAGTGCAATCTGAGAACTTCTTTCTGATAGTTTTCTGGATTAAAGCATCAGTTTGTGGATCCACATTGGCT
GTGGCTTCATCTAACATCAGAACTTTATTATTTTTCAGAATTGCTCTTGCTAAGCAAATTAACTGTCTTTGCCCAACACTGTAATTGGAACCTCTAT
CCATGACTCTATTTTCTAATCGGTTAATAACATTAGCAGGATCTCTTAACTCTACATCTTCTATAGCTTTATATAAAACTTCGTCTGTATATTCTTC
GAATGGGTCCAAGTTGTACCTTAAAGTTCCTGAAAATAGGACTGGGTCTTGTGGAATAATAGAGATTTTCAAACGGAGATCATTGAGGGTAATATCT
TTGGTATCGATCTCATCGATTTTGATAGTTCCTTCAATTGGAGCCAGTCTAAATAGTGCTTGAATTAAAGAAGATTTTCCTGCACCTGTTCTGCCAA
CAACTCCAACCTTCTCCTTGGGTTGAATGCTTAGGTTAAGATGTTTCAAAACCAATGGTCCCCCTTCAAAATACCTTAAACATGTATCTTTAAAAGT
AATTTCTCCTTTCTGTGGCCATTCTTTCGCAGGTTTTTTAGGACCGACTGGTTGTGGTTCTTTGTCTAAACTTTTGTATTCTAATATTCTCTCTACA
CTCATAAGTTGATTGGTTACTTCGGCAGATTGGCGCATACCCCATTGTAAAAAGTTTGCCAGTGCTGTTGCTTGAGTAATCGCCAAACCTACTTCAC
CTCCTCTTAAATTAAAAGTTTCTCCGAATAACAATAAACAAAAAGTAAGTACTGCTAAAAATATTGTACAGAAACAATCCAAGTAAAATCCAAATGC
AGTACTGGAAACTATAAACATATACCATGCACTTGAATGATAGTCTTGATGATGATCAAATTCGTCCATAAGAGTGCTTTGCGCTCCAAAGGCTCTT
ATAGTTGTCAGACCTTCAATTGTAGCTCTTAAATGTGTAAAAACTGGGCTTCTCATTATTCCTTCTAACCTTTTAATATTCTTAGAGGATTTTAAAT
ATATATGTCTCATTACACCAAACACTGCACTTAAAACTCCTATCAAAATTAAGAAGTAAGGGTTAACTACAATAATCAGTACAAGAGATCCACACAT
AGATAGAAGAATCTGTGAAGAATCCAATATTGCTTTAGGCAATAATTCATCGACAGCTCCCATATCCTTAGAAAACCTGTTAAGAATTCTTCCTCCA
GGATTTGTGTCGAAGAATCTCATTGTAGCATCAATAACATTATCGAAAATTGTGCCATGAAGTTTTTTCGAGCTCAACATAGCCATTTTATAAAAGA
AGAAAGACCTTGTCATAGCAAGGACAAAAAGAGCTCCTAAGATAACTGAATATATATAAAGGCAATTATCTGTTGTTAATTCGATGGATGGGATGAT
GAACCTAGGTGCAGTAGAATTCAAAGATTCAGACGAGTTTCTGAATTCCTCTACATTCACCCAGTAACTTACCCAGTAATCCGTCCCACTGGCAGCT
AACTGAGTACCGAGATATAGTAGAACGCATATGAAAACAGCAAATACATTAGAACCTGCTAACATGTATTTAAATAACAAAGAACCGCCCACTTTTC
CCTTCGATGATTGTTCCTGAAGGTCTTTGACATCAAACTCTGGTTCTTTTTCAGCTTCCTCTTCGTCAAAGTCGACATCGTTTGATAAAAGAGTATC
CGCAATACTCAATTCACTTATTATACTTGACATTGAATCTTTTCGACTTCTTTGTGATAGTTGCCTAGAATATTTCTGTCTCTCTGGTTTCTTCTCT
TCCTTATTTTCTGGTTCAGCTGTTAAGAGTTTTGCGTACACATTGTCACTATTTGCCAAAAAGTTGAAAGTTCCTTCATCTTCGATGCGACCATTAT
TTAAAATAATGATGTTGTCAGCATCCTTCAGATAGTGTACTTGATGGGTAACTAAAATTCTCGTCCTGCTTGCCAAATAGCCATTAATACATTTTTC
ATATAAATGTTTGGAAACATGTATATCTACAGCGGATAACGGATCATCCAAGAGATAGACATCTGCTTCTCTATACACTGCTCTTGCTAAATTAATC
CTAGCCTTCTGACCACCACTCAAAGATGATCCTCTATCACCTACGAGAGTTAAATCTCCATTTGGAAACTGTTCAAAATCTTTTTCTAAAGAACAAG
CTTGTATAACTTCCTGGTATCTCTTTTTATCATATTCTTCACCAAAAAGAATGTTTTGACGAACAGTGGCTGCAAATACCCACGGTTCCTGGGAAGC
ATAGGAGATTTGACCTTGTACTCTCATTTCGCCTTGGGTGATGTCCAGTTCACCTAGTAAGGTTTGCAATAAAGAACTCTTTCCACTACCTACAGGT
CCAATAATTCCTATCAACTGTCCATCTTGAACATTTAAATTTATATTCTCGAGAGCATTGTCGCTAAAACTTAAATTCCATTTTACAGTTAAATTTT
TCAAACTGATAATTTCCTTAAAATCGTCTACACTTTTGACATTGCCGTTCATACCAGAAGTTGATCTATTGGGATCGTACTCTTCGTTTAGAAGGAA
ATCTCTGATTCTCCGAACAGCTACCAAACATTCTGCTATTTCTGATACTCCTCTAGTAAACATGGTGGACATAGCCATTGAAATAACGTTGAAGTAA
GACATTACAACGAAAACCATTGAGGCTGTGATTGGTTCGTTCGACAAAATAACAGTAAGCAAAGTACAGAATAACGCCAATCTAGTTGTGAACAGGT
TCAATGCCATAAATGAAGCTCTCACATACGCAGATTTTGTAATGATTTTGATTTCATTTCGTCTAGCAATCCGTATGACCTTCCTGAACGGTATTTC
CCATGCGTACATTTTGATTACTTGAATTCCTGAGATGATTTCGTCCATTAGCCTGACTCTCTCGTCTGTTTTCATAGCTGTCTGTTTTCTGTATTTT
GCAGATAGCTTTCCAGTGTATGTTTGAATAGGAACTATCACAAAAACTGCCACAACCCCAACTATTCCAGCATATCCAGTTCTCTGGTATAATAAAA

CCATTACTATAAGTGACAATACAGGTGCAATCCATATTTGATGAATGAGTAAACTGACGAGATCAAATCGACTGACATCGTTGGATAGGAGGTTGAC
CAGTTTCCCTGAAGCTGTTTCTCCCAAGGCTGTTTTGCTCAATTTTAAGGACTTTCTATATATAACAGCACAACAAGCTGCTCTTATCCTCATGCCA
CTGTGGAAACCCTCTACCATATACTGATTTCCGATCAAAAAGTTCAATACGTTTATAGCCACTATGATTCCTGCATACATGAGTGCCTCATTTTTTG
AAACTTGTGCTTCAGGTTTAAAATGGTCCAACAGCTTTCCAAGCATTAATGGTTGGATGAGTCTAAGGACTAAATCTGTTGTTATGTCGAATATTCC
TAACTTCAAATACTCTAACCAATATGTCATAAGAATGGCTTTCAGTAAACTTGGTTTTGAGTTTGTTTTCTTGGCTTTTTCTAGTTGCTTGTCCCAG
TTGCTTTGTAAATGATCTCCCAATATTTTGCTTCGATCGATCTTTAATGGATTATATAAATCGTCAACTTCTAGAGTTTTCGATAAGCCCTTTTTAA
ACACATCATATGTATATCCAAAAAATAGAGTCGATAAAATATTCGCTTTCTCCCGTGGATTGGGATTATATTTCTCTTTGGTCATTTCCATCTTTGC
GGCCTGGCAATTCCTCTAGTTGCTACTCAAGTATATCCTGTAACTACCCTCGTTTTCAATTTCGCATGATATTTTCCACTGGACTGCATAAAGATCG
CCAAGATTTATGTGTATGCTTCTATGTATTATATCACTACAACCAAATCAAGTGAGGTAGGTACCTTGCGTTTTTACAACCTCATCATCAGCGTCAC
AATGTGTATCTGTATAATTATTATTACGGCAAGGTTTTTCTTGATTGGTTTGTTCTGTATTTACATCGACCTGTTGGTTATCTCCGGCCTTGATGTA
ACCGCTATGTA
>Dvv ABC-C D22413+T49258_c0_seq1
CTAATGACTAATGAATTTTTAATAGAAGGTCAATACCGATAAATCCGGAGGTTCCAAAGTTTCAATGAACAGAGGCAATGGATTCCACAAAGGAAAC
CTACAATCCCAATCCAAGGGACAGTTCCAATATATTCTCAATATTATTTTTTGGGTTCACGTATCCCATATTCAAGAAGGGCATGAAGAAAGACTTT
GACGTGGAAGATTTGTATAACCCATTGAAAAATGACAGAAGTACTCTTTTGGGAGATAAGTTAGAAAGGAATTGGGACAAACAAAAACTTAAACCAA
AAGGCAAACCCAGTCTCCTCCGAGCAGTAGCAATTACATATTGGTACGAATATCTACGATTAGGGTTCCTCACATTGATAGGAGATGTTTTTATTCG
AGTAAGCCAACCTTACGTTTTGGGCTTGTTGCTCACCTATTACAACCCACATTCCAAAACCACCAAGGAACAAGCTCTAGGGTACGCTGGGATGATA
GTTACTTTTAATACTCTAACGTCGTTTATAAAGAATCAGTATATGATGAATTCTATGCACGCTGGTATGCGAGTAAGGACTTCGGTGTGTGCTCTTA
TTTATAGAAAGGCTATCAAATTAAGCACTACTGCCCTCGGCAAAAGTTCAGTAGGGAAGATTGTTAATCTTCTTTCGAACGATGTAAGCAGATTCGA
TTCAGCCTCGATGTTACTGCACCAAATGTGGGTTGGACCAGTTTCTGCAATAATAATTATGTATATCATTTTTCAAGATATTGGATGGTCCGGAGCA
ACTGGAGTTGTTACCATTGTTACAATTATGCCCATGCAAGCGTATATTGGCAAATTATCAGCAAAATATAGAAGAATAATCGCTGGTAAAACAGATG
AAAGAATAAGACTGATGAACGAGGTCATAGGGGGCATTCAGGTGATTAAGATGTATGCGTGGGAAATACCGTTTACAAAATTAATCAGCCTTGCACG
AAAAGCTGAAATTAATGTTATTAGAAGAAGCGCGTACATAAGAGGTGTATTTATGGCTTTCAATTTATTCAACAATAGATTTGCTTTGTTTGCAACT
TTGGCTACGATGGTATTTACTAACAAAGCTATCACGGCAGCAAGAGTGTACGTTTTTATGTCATATTACCAAATCCTAGCTAACACACTAGCTGGAG
TATTCGTCAATGGTCTTACACAGATTGCTGAACTTCTAGTTTCCATTAACAGAATCGAAGAGTTTCTCGAAAATGAAGAACACAAAGAATTACCTGC
CCTAAAAAGATCTGTAATCTCGAACGTCAATGAAGAGATGGTAGCTTGTAAGAATATGGTTGCGTCTTGGAATGAAAACTCACTTGATCCTGTTTTG
AAAAACCTAAATTTTAAACTTGGAAAAAATACTCTACTTGGTGTTATTGGTCCTGTTGGcAGTGGGAAAAGTTCATTGCTACAGTCTATTTTGGGTG
AACTTGATATAGTTGACGGAAGCCTAGCCGTCCACGGCACACTCTCCTACGCCTCCCAAGAACCTTGGATCTTTTCCGGCAGTATTAGGCAAAACAT
TCTTTTTGGTTCCGACTTTGACAAAGCCCGCTATGATGAAGTATTAAAAGTATGCGAACTCAGAAAAGATTTTGACCAATTCCCCGACCGAGATTTC
ACTCTGATAGGAGAGAAAGGCGCCAATCTCAGTGGTGGTCAAAAAGCTAGGATTAATCTAGCTCGTGCTATTTACAAAGATGCTGATATATATTTAC
TGGATGATCCACTTTCTGCCGTTGACAGTATTGTTTCCAAAGTATTATATGAAGACTGCATTAATGGATTTTTAGCTAAGAAAGCAAGAATATTAGT
AACCCACCAAGTTTATTATTTAAAAACTGCTGATCACATTCTCGTTCTCAATAATGGAGGCATCGAAATTGAAGGCACCTATTCAGAACTACTAAAA
AGCGACAACCCTCTCACTGTCCACTTAACCGAAGAAATCGAAGAAATTTTAAAACGTCAAAAGTCCCAAGATTCTGCTATAGAAGAACAGGAAGCTA
AGACTATGAGCACGATGGATCTTTCCAAAGTGAATGGAAAAATGAAAGCCATTGTCAAAGAAATGCAAGAAGAAACATCTAAAGGAAAAGTTAAAGG
TTCCTTATTCCTCGAATTCTTTAAGTCAGGCATGAGCATCTGCGAAATAATATTTTTAGTTACTCTTCTTACTATAGCTCAAGCCAGCGCCACTGTT
ATAGATTGGTTCATTAGTTTCTGGACGAATATCGAAGAGTACCAGATGACATTGAACTCTACCGCAGCTAATGTATTAAACTCTACCGTAGTTGATG
CATTAAACTCTACCGCAGTTGATGCGATAAACTCTACCGCAGTTGATGTGTTAAATTCTACCACTGTGAATGCATTAAACTCTACTGTAGTTGATGA
AAGTGTGCCCTTTTACATGAATTGGAGTAGCCAAACTTGTTTGATAATTTATGGTGCAATTTTAGTCTTTTGCTTGGTTACAACGTTATCCAGGTCA
CTTAGTTTTTACAGATTTATTCTAAATTGTTCTGAGAATTTACACGGCATGTTGTTCACTGGAGTAACTAATACATATATGAGATTTTTTGATAAAA
ATCCAAGTGGAAGAATCTTGAATCGGTTTTCAAAGGATATTGGTTCTGTAGATGAAATCTTGCCAAGGATGTTATTCGAAGCTTCTAGGGTCACCCT
CAAAATGCTTGGACACTTGGTATTAGTCCTTTACGTAAACCCGGCATCAGTTATCGTTGTAATAATATTGGGAATTTTGTTTTCCTTTATTAGGATA
GTCTATCTTCGATCATCCAACAATATTAAGCGGCTCGAGGGAAGAATGAAAAGTCCAGTCTTTAGTCATCTGACAGCTACATTGGAAGGACTCACCA
CCATAAGAGCCTTTAAAGCTCAAAACATCCTTAGAGCAGAATTCGACAAGCACCAAGATAGCCACACCAGTGCTTGGTTCATGTTTATTAGTACCAG
TTCTGCGTTTGGATTTTCATTGGATGTCATCTGTTTGATGTTCATTGGATCGTTAACCTTCAGTTTGATTGGGCTGGGTGAATATTTTAATTTAACC
GGTGGTGACGTGGGTCTAGCTATTAATCAAGCTTCGTCGCTTACTCAAAACATACAATTTCTGGTTCGTTTCTCTGCTGATATCTCCAATCAACTAA
TGTCAGTGGAAAGAATTTTAGAATATAAAGAATTGATACCAGAACAGCAACCGGAAAAACTACTGTTACCTCCAAAAAGTTGGCCGGATAAAGGGAT
TGTTAATATAGAACATTTAAATATGAAATATATAGATGATGGTCCCACTATTCTAAAAGATGTCAGTGTTAAGATTAATCCAAAAGAAAAGATTGGT
ATTGTGGGCAGAACTGGAGCAGGAAAATCTTCATTAATATCAGCCATATTCCGATTGACTCCTCTAGAAGGTAAAATCTACATTGACGATATCAATA
CCAAAGACATAACACTAAAACAGCTAAGATCTAAAGTTTCAATTATACCCCAAGATCCCATCTTGTTCTCCGGCACCTTGAGATATAATCTAGATCC
CTTTGATGAATACAGCGATGAAACTTTGTACAGAGCCTTGAACGAAGTCGAATTAAAGGATCCGAGCAACATTATTAATCGTTTGGAAAACAGGGTT
ATGGACAGAGGATCAAACTATAGTGTTGGGCAGAGGCAGCTAATATGTTTGGCGAGAGCTATTATAAGAAATAATAAGATTTTGGTCTTGGATGAAG
CTACTGCAAACGTGGATCCACAGACTGATGCTCTAATCCAAAGGACTATTAGAGAAAAGTTCGCAGATTGTACAGTTTTAACAGTGGCACATCGGTT
GAACACAATCATCGACAATGACCGAATATTGGTGCTTGAAGCAGGAGAAATTGTCGAATTTGACCATCCGTACTTATTATTACAGAACAAATTTGGA
GTATTCAGAAAGATGGTGGAAGAAACTGGAACAGCTATGCTCAGACAGTTTTTGGAAACTTCCAGTCAGAATTATCAAAAATTAATTGATTAAGTGA
TCTGTTAAGTTCCTGTATTGTACTTGTATTATTTTTT
>Dvv ABC-C D5345
CAGAGGATCAGTATAATAACTTAAAATATAAAATGATGAAAATAACGTGAAAATAATATTTTTTTTAAATTTCGATAGTTTATATAATGATAAAGAA
TAATACATAAGTTTAAAAAATACTGCCTTTGTAATATTACCCAATTATACATTATATTTTATGTATATCAACATTTAATAGTCAATAACCCAGAACT
TTTTAGCATATTATAAAATAGGCTATGTTCATCTTGTGCCAAAGTACTAACTGCGCCATTTTCAACAATTTCTCCATTTTGTACCACTAGGACTCTA
TCACAATTTAGAATAGTTTTTAGACTATGGGCCACAACTATAACTGTACTTCGAAAACTATAACAGTCCAAAATTTGTTTCTCCAGAATGTTGTTGA
ACAAAGACTCAGATCCCTGATCCATCTCTGAAGTTGCTTCATCCAAAAGAATGATTTTGTTTTGATGCAACAAGCATCTGGCAAGACAGATTAATTG
CCTTTGTCCAGCACTTAAATTTTTTCCTCCTTCGCATATTTGAAAATCTAGAGAATCTATTATTGTTGTTAACTCCAGCTTATGGATGACCTTCCAA
ATGTCTTCGTCTGGGTATACACCGAAAACTTTAGGTGCCATATTGCTTTTTATTGTTCCCGAAAAAAGCACAGGATCTTGAGAAATTATGGATATTT
TTGAACGAAGTAATGTAAGCGGAAGACTTTGGATATCTTCATTACCAATGGTGATTTTCCCTTCGTAATCAAAGAGCCGAAGTAATAAAGTTATCAA
TGAAGACTTTCCTGAACCTGTTCTTCCAATTATTCCAACTTTTTCTCCTTCATTAATTGTAATATTAATGTTTTTTAAAATATACTGTTGACTATTA
TCGTAACGAAAATTTACATTTTTGTATTGGATTATATTGTTTGAAGTCCAATCTTTAAATATGTCCCTAAGGTCATTGTCTTCTGTTTTTATACCAA
CGTATTCCAAAATCCGTTCTACCGATGTCATTTTATTCTCAATATCTGCCCAAAGCCTAATGGCGTATTGGGCAATGCCATTCAACATAAACGCTTG
TGTTATTGCTAAACCAACATCCCCAACACCAGTTTCTTTTCCAACAATGATGAACGCTGCAATAATAGCACCTATAAATGTAGTACAACAAACATCC
AAGTAGTACCCAAAAGCCCTCTGGCAGCATTTTGAAGTATAGTTTGCAGATGTATATAAATCTTGATGTCTATCAAATTCTTTGCTAATGGTTTCTT
CGGCTTTAAATGCTCTAATAGTTGAAAGTCCTTCGAGAGATGAGTTTAAATGGCCAAGCATTGGACTTCGAACTGCTGTATCCAGTCTCTGCAGATT
TCGACCAGTTGGCAAGTAAAGTTTCCTTAAAAAATACAAAAGTATGAAAACTACAGCAGTTGGTATAATGAACCCAACGTTTACTAATGTTATTAGA

ATTATTACTCCGCATAAACTCAAAACAATTTCTAAAAATTCATATATTGCAAAAGGAATTTGTTCATCTACAATGCTAAGATCCTTCGAAAATCGGT
TAATAATGTTTCCAATTAGGTGACTATCAAAAAATTTCATAGTTGCATTGATTACAGTGTTACTTAAGCATCGATGTAGTTTTATAGAAGCCTTTCT
TGTAAAATTAAATAACAAAAAGGCTGAGAATATGGAAAGAATCGTGGTTGACAGTGTAAACGCTGAATATATCTTCAAAGAAAAATCCCACACTGTA
TTTTCTTCATTGATATTCGTTGTGTTGTGATGAATGTGATTTACCCAGTTGCTTAAAATCTTTTCAGTATAACTTTTTGAAAATTGTGAACCAATAA
AAAATACCAATACAATTGCAATGACCAAAATTCCTCCACCATAATTTCCATATTTTCTATAGGTATAAAAATCAACTTTTCCAGACTTTTTATTTTC
TTGGTACATATTATTTCTTCCAATATTTAATTTCTTCAACAGCGAAGATTCTTCAGTGGTTTCATCATTGTTAACCTCATCAGCGTTGTTATTAACT
TCGATCAGATTCTTTAACTGTTGTTCTTTTTTTAAAATGGTTTTGATATCCTGGTCATCCTCTTCGGATAGGTTTTGTGATACAAATGTTTTACCTT
CGTTCAATACAATTATTTTGTCATTTCCGTTAGTATATTGAAGGTTTTGAGTAACAAAGATACATATTTTATCTGAGAGGAACTCTCTGATACACTT
TTGAAATACGTGGCGTTGAACGTTGACATCCAAAGAGGATAAACAATCATCTATTAAGTAAATATCAGCATTTCGATACAAAGCTCTTGCTAAATTT
ATTCTAGACTTTTGTCCTTTACTGAGGTTGGCACCTTTGTCACCTATTTTTAGGTTGTTTGCTAGGGTAAAATGCTCAATATCAAAACCCAAAGCAC
ATATATCCAAAATTCGGTTATAACGTTCTTCGTCGAAAGGCTCATCAAAAACAATATTTTGTTTAAAAGTTGCTGGGAAAATCCATGGTTCCTGCGC
AGCATAAGAAATAGTTCCATATTTAATTATATGATCACTTTTTGTATCAAATTCACCCAAGATGCTTTTTAAAAGAACCGTTTTGCCACTACCAGTC
GAACCACTAATAATGTATAGACCAGAATCAAATCTTAAAGTAATATCTTTTAGTAGCACTTTATCATCTACTTCAATTGATAAATTGTTCAGTGAAA
CATATGATTTCTTTTGATCAACCTCAATTTCGTCATTAACAATTACTTCTTCAGCAGTCATCATCTGCTGAATTCTTTTCATAGCTGCAGACATGTC
AGCTGTAACAGTAACACCGATGGGTATAGCGACATTGAGGGCATGTCTTATTCTTAAAAACAAAGTTGTGATGTAGTATACCAGCTCTGCTGAAATA
GTGTTGCCCAAAAGGATGAAAGTGGTTAAAAGTAGATAGTAGGCTATCTTTGATGTTAGACTTCCAATTAGCAGAACGGTGGTTTTGGCAAGAAATA
TTTTCAATATCGCTTTCATTTCCTTCTTCCTCGTTTGGTTAATTTTGTGTTCAAATTCAAACTCCCATACATTCATTTTTATTGTACTGATTCCAGA
CAACATTTCTTTGGTAGTTTGCAATCGTTCATCAGTCTTCTTACTAGATTCCAGTCTTAATTTTGATATCAACTTTCCCAATAAAACTTGAGCAGGA
ATAACTAACAGGAAGAATCCAATTCCAATTAAAGAACCAACGCCCATTTTGTTAAATATTAAATAAAACACAATAATCGCTTGCACTATTCCTACCC
ACACGTCATTGAAATAGAAAATGAAGTTTTCTATCGCTGCTACATCTTTCGTCATCAAAGTTACAATTTTTCCTGAGCTCACTTCTAACAGCGATTT
CTGGGTCAGTCGTAAAGATTTTCTAAATATGGCCGAACAGAATGCAGTTCGTACCTTGATTCCCAATTCAGATAGAACCAGAAGGTAGTTGTGTAGG
TAAGTCACACTGAGAAAATTCAGGAGTACAAGTAGGATGGCGCAGTATATGGCTGAGCTTCTGGATAATTCGTCTTGACCAGGTTGAAAGTACTTTA
TAACTTTGGATAACGTCAAAGGGATGGTTATTATAACTATAGATCTAACAATCACTTGTATCATTCCTAGTAAAAAATATTTTAAACCATAGCATTT
ACATAATATTCGAAAAAGAGACAAGTATCCATGATTTTTTACGTCATCTTTAATACATTTTTCGACCTCATTTCCGGCTTCTTCTGCTTTGCATTTC
TTGAGTACTTCATAAAGTTCATCCTCTTCGAAGGTTCCTCTTTTGTGTGTTTTCCAAAATATTGGCAAGGTGTATAAAAATATGATTCTAGATAAGA
AGTTGGAATTCTGTATGGGATGTTCCTGTCTGTTATCACTTAGTTTATTCTTGTTCATGATTATCCCTGACCATCCTCACATCGCCACAAGCTTGCA
AATATTTGTCAACTGTACGTTATTACGCAATAACAACTTACAATACGTCTATG
>Dvv ABC-C T47673_c0_seq1
CTCAGAGCAACAGTGGCCTAACCCACAAATTGAACATTTTTAGATTAATATATAAATAAAAGCAGCAACATTATGATTCTGAAATTGTTTAGCCTCC
CCGGTCTCCAGATTTGACGCCACTTGATTTTTCTTGTGACCTACGGTAACCTTTCTGGTTACCTACAGGAAAAGGACTATTTTTAAAGGCTGTTTAC
AGATATTGCTTGAAACCGATCATACCTAACATTTGCTGTAGGTACCTCAAATTAAGTGATATATAATTTCCAACGTCAAATATATCGTCATAACATA
AAAATACATTTGTAATCATTATTAATGGTATACCTATACCTAGGTAGATATACCAAGAATTATTTTCGAAATCTCATAACGTGTCAAAATATTAAAA
CTAATCCCTACTCAGGGAGTGATAGAGTCCTTTGATAACTATCTTTGGCAATTTTCCTAAGTTGATCACCAAGACTCTTGCCTGCTTCAGCTACCAT
TTTATAAAAGACACCATTAGAATTCTGAAGTAGTAGATGTGGATGATCGAATTCTACCATAGTGCCCGATGCCATCACAAGTACTTTATCCGAATCC
ATAATAGTGTTCAGTCTATGAGCAACTGTAAGTACCGTACAATCTGCAAACTTTTTTCTTATAGTTTTCTGGATTAAAGCATCAGTCTGTGGATCCA
CATTGGCAGTCGCTTCATCCAACACTAAAATTTTGTTATTTCTAATAATAGCTCTGGCCAAACAGATCAACTGCCTTTGACCAACGCTGTAATTGGA
ACCTCTGTCCATCACTTTATTTTCTAATCTATTTATTATATTGGCAGGATCTTTTAGTTCGACATCCTCGATGGCTTTGTATAAAAGTTCATCGGGG
TATTCTTCGAAAGGATCCAGATTATATCTTAAGTTTCCCGAAAATAGGACAGGATCTTGGGGAATGATTGATATTTTTGAGCGGAGATCTTCCAAAA
GTAAGTCTTTTGTGTCTATTCCGTCTATGTTAATTTCTCCATCGATTTTGGCTAGTCTAAATAGAGCTGCTATCAGAGACGATTTTCCAGCGCCTGT
TCTACCGACTATTCCAACCTTTTCATTGGACTGAATAGATAAGTTTAGATTATTCAATGTCAAGGGTCCATTTTCTATATATTTTAATCCCATATTT
CTAAATGAAACGTTTCCCTTTTCTGGCCATTTCGGCAATGGCTTTTTAGGTACCACTGGCTGTATTTCGGCAGGTAAGGACTTGTATTCCAATACTC
TTTCTACACTCATTAGCTGGTTAGTAATCTCTGCAGACTGACGAATACCCCATTGAACCATACCTGTCAGTGCTGTTGCTTGTGTTATAGCAAGACC
GACTTCACCACCTTTTAGTGATAACGCTTCTCCAAAATTAAGTAAATAAAATGTAATAATGGCTATAAATATGAAACATATGATGTCCAAGTAGAAT
CCAAACGCAGAACTTGCAGCAATATACATGAACCAGGCACTAGTATGGGAATCTTGATGCTTGTCGAATTCGTTCATTAGAATGGGTTGTGCACCAA
ATGACCTAATCGTTGTTAATCCCTGGAGAGTAGCATTTAAGTGGGTAAATACAGGACTTCTCATTATTCCTTCCAGTCTTTTGATATTTTTCGATGT
TTTCAGGTACACATGTCTCATTGATAAGAAGAAAATTCCCATGAAGAAAACTAATATAAGGAAATATGGGTTTATTATTGCAACTAATATCAACGAT
CCAACCATCATTAAAATAATCTGCGAAGCATCTAAGATAGCTTTGGGCAACCATTCATCGATGGCTCCTATATCTTTCGAGAACCTATTTAAAATTC
TTCCACTGGGATTCGTATCGAAAAATCTCATAGTAGCTGCGATGACACTATGGAACATGACGTCATGTAATTTTTGAGAACTCCACATGGCTAGCTT
GTAAAATAGCATCGATCTAGCGAATGCTATCACGAATAGAGCTATTAATCCTCCAGCGTATATGTAAATACATGTTTCGGTTGACCATTCTACGGTA
GACCTAATACTGACTTCTTCTGTTTTGTTTGCACTTGTTCTAGATAACGTTTGGTTTCTAAATTCTTCTATATTTACCCACATGCTGACAAAATAGT
CCACACCACTGGCTGCCACTTGCGAAAGGATATACAACAGAAGAACGAAAGCCACGAAAAATATATTACCTCCGTGGATTAAATAACTCCAGAACAG
TGATTCTTTCACTTTACCCTTCGAGGATTCTTCCTGTAAGTCTTTCAGTCTTATTTCTGGTTCTTCTTCATCATCACTTCCAGTAGCTTCTTTTATA
ATAGCATCGGTGAGACTAAGCTCACTAGCGGCGCTGACTAACGACGTTCGCGACCTCCTGATTGATATTTTTCTGCTAATTTTAGCTGTGTCCGTCT
GTTTCTGTTTTTCTTCATCTGTAAGTTCAACTTCGGAAGTCAGCAATTTGGCATATAAATTATCACTAACTGAAAGCTCTTCAAAACTGCCTTCATT
TTCTATAGTTCCGTTATTCAAAATAACAATATGATCAGCATCTTTTAAGTGATGAACTTGGTGGGTAACCAATACTCTAGTTTTATGAGCTAAATAC
CCATTTATACATTCATCATAAAGATGCTTTGAAACATGAATATCCACAGCAGATAAAGGATCATCTAACAAATATATATCAGCCTCCCGGTACAACG
CCCTTGCCAAATTGATCCTAGCCTTTTGCCCTCCGCTCAAAGACGCCCCCCGTTCTCCTACGATGGTGAGGTCTCCGTTTTCAAATTGTTTAAAATC
CTTTTCTAGGGCACACGCTTTGACCACTTGATTGTATCGGTGTTTGTCGTATTTGCCACCAAAAACTATGTTCTGTCGGATGGTAGCTGCAAATACC
CATGCTTCTTGGGAGGCGTAAGAGACGGTGCCGTTAACTTCAATACTGCCATAGATCACCTCGAGCTCTCCTAAAATCGTCTGTAACAGCGAACTTT
TTCCGCTTCCGACTGTTCCAATGATACCAATCAACTTCCCCTCATTAACAGCCAAATTTATTTTTTTTAGAGCGTCATCTTTGGTGGCTAAATTCCA
TTTGACGGTCAAGTCTTTTAAGTTAATAAGAGTTTTGTTGAAGAATATTTTGTCGTTATTGTTTTGGTATTTTACCACTGGACGGAATTCCTCATTT
AACATGAATTCCTGGACTCTTTTGATTGCTACTATTAGTTCTGCAATTTCCGCTATACCTCTAACAAACATGGCAGACATCGTTTGGGAAAGTATGT
TGAAGTAGGACATGAAGACGAATACCTTGGTAGCTGTGATCTGCTGGTCAGTCAATACCATTGTCAATAATGTAGCGAATAAAGCAGCTCTGGTAGT
AAATAAGTTGAAGGTCATAAACAAAGCTCTTACATACGACGACTTTGTTACAATTTTCAGTTCAGCTTTCCTCGCTAATTTGATCAAATGTTTGAAA
GGCTTTTCCCATGCGTACATCTTGATTACTTGGACACCAGATATGACTTCATCCATCAGTCGGACTCTCTCGTCCGTTCTTAGAGCAGTTTGTTTTC
TATAAATTGCAGAAAGTCGACCAGTGTATGATTGCAAAGGTACGACTAGGAACACCGTCAGAATACCCAGGATTGCTGCATAGCCAGCTTCTCTGTA
CATGAAGAACAAAATAAATATGGCCGAAGTGGGTGCTACCCACATGTGGTGGATAAAAATTGATACTATGTCAAACCTGCTGACGTCATTTGATAAT
AAATTAACTATTTTACCAGAAGCAGTGTCTCCTAATGCTGTGTGACTTAATCGAAGTGCCTTTCTATATATCAGGGAACAAACTGCTGCTCTAACTT
TCATACCATAATGAAAGGCATTCATGATATACTGATTGATTAATAAAGCGCTTATTCCATTTAGAACTACTACAGCACCTGCGTACCAAAGTGCTTC
TTCCTTAGTAGTCTCTGTTCCTGGTTTGAAGTACGACAACATGTTACCCAACATGATGGGCTGTGATAACCTAACTAAGAGATCCATTAGAGCTAAT
ATACAGCCTAGATATAAGTATTCTGGCCAAAATGTGGACACTAGTATTTTCAGCAAGCTCGGTTTTTTTCGGGTTTTGTTGCATTTTAATAAATGTG

AATTCCATTTTTTTTCTAATCTATCTCCTAATACTGTACTTCTATCACTTCTTATTGGATTATACAGATCATCCACGTCCAAAACTTTGCGATATCC
TTTCCGGAACATCTCGATTGTGTAAGAAAAAAAGAGGACGGAAAATACGTTGGCTCCATCCCTGGGATTGGGATTGTAATTTTCTTTTGCTATATCC
ATGATTATTCACAGAGTGTTTACTTAAAACCATTGAACTAAACACAAATTTTATTAAGGTTAACTTAGGCACTAAACTAAAACTCGCGTGAACGGAA
CACAAAACACTTATCAAAAATTATATAATTAACATTAATTAAGAGGGCTCAAAATGTATCACGTTATCAAGACGATTTTAAAAACTTTCAATTACCG
ACACACTTCTTTCTTCCTAAAGATTAACCATAACTGACACACAGTGACTTAAACGATAATAAAAAGATTGTTTGGGCTGTTTCACTTTATGGCGGCA
TTAAGAGTCGGATGGTCTCTTTTATGCAGATGATTGTAATTGGAGGCGGCTGTTAAAAGTTATTATCTGGGTTGAATTCCTGGTAGTTAATTTGAAC
GGCCTGAGGATATTCCCCGTGAGTATTATCTTCCAACTGGACCAATTTGGGTTTTTTGTGCTGCCGTTGTTGCGTATGAGTAATTTGTTATAGCAGC
AGAAGTACAATTGTGATTGAAGTCTGCTAATATTCAATTGTACAGCAAACAAAATTTAATCTCCCACGGAAAAACAATTAAAACTATTTCTCACTTG
TTTTATTTTCGCGGTAACTGTTTGAAATTTACGTCAACTCACGCCGTGGTTTGTTTGTTCGCTTAAAATAAACGTCCTCTCCGTTA
>Dvv ABC-C T48300_c0_seq1+D21892*
AAATTTTATTTAAAATCTAAATAATTTAGTAATCCTGCTTGCTTAACCATTTTGTAAAACATTCCATTCCTATTCTTAAGTAATTCTTGGGGTTGAT
CAAATTCTTTCATTTCTCCTCTATCTAATAGCATTACTTTATCACATACTAACACAGAATGTAATCTGTGTGCTATCGTTAACACCGTACAATCAGC
AAAATTATCACTAATAATCTTTGTGTAACATCCTATCAGTCTCATGGTCCATGTTCGCTGTGGCCTCATCCATTACTACTAACTTATTCTTACGTAAG
ATAGCTCTAGCTAAACAGAGCAGTTGCTTTTGCCCAGAACTGAAATTCATGACGCTGCTATTAATTGGTTGTTCTAAACTGGCAATAGATGTATTAA
TTCCAACTCTTTCCAGTGTATGCCAGAGATCTTTATCTTCAAATTCATTCAGAGGGTCTAGATTTGATCTGATAGTTCCTGAAAATAAAATCGGATC
TTGTGGTATCACAACAATATTTTTTCTCAAATAGTCCAAAGCCAGTGTCTTAATATCAACTCCATCGATAATTATCTTGCCTTTGACTTCGTACAAC
CTAAAAAGAGTAGCAAGAATTGACGATTTCCCGGCTCCAGTACGACCAACCACACCAATTTTTTTCTCCAGGTTTAATTTCAAAATTTAAGTTTTTC
AACACAAAGCTGTCACTATCACCGTATGCCAAATATACGTTTTGGTAAGAAATACTTCCTGTACTGGGCCAATTTTGAACTGTGGTGCCGTCCTTCT
TTTCTGTTTTGATATCTGTATATTCTATAACTCTTTCAAGTGACGTCATGTGGTTTTCCAATTCGGTAAACACTCTAATACCCCAGGTTAGTACTAT
AGATAATTGCGAAATCTGCGTTAAAGATAAACCAATATCACCAGCAGTAATATCTCTTTTAAAGAAAAGATAATCCATTATGATGAAAATCAAGAAG
AAATTACTGAGCATATCGACGAAAAATCCAAAAGCTCTCGAGCTCATGTACCAAGTAAAATTGGCAGACGTGTATAAATCTAGATGTCTGTCATATT
CCTCGATGAGGATAGATTGAACATTGGATGCTCTCACGGTAGTTAAGCCTTCCAGTGTGGCGTTCAGATGGCCTAACATGGGACTGCGAGATGATGC
CTCCAATCTCTTCATACTTCGTCCTGCTGGCAAATAACAACTCCTTACCCAAAACAACATGATAGTCACAACAATGACACATACTGCAAATGCAGGG
TTGATGAAAATCGTCAGGGTTAATATTCCACCCATACTGAACAGCATCCTGCAAAACTGATCCAGGTTATATGCAATACTTTCATCAATATTCATCA
TATCTTGGGAAAACCTAGTCAAGATATTTCCGATCAAGTGGGTATCGAAAAATCGCATAACCGAATTGTTAATCTGATACACCAAGGCCTTGTGGAT
ATTGATCGAAGCCGTTTTGCAGAAACTTAAAAGTAGCCATGTCGTAAGTAATTCCAATGATGAAGAGGCTAGTAAACATAAATAATAGAACCTAAAA
GTGGTATGTTCCTCGGCTTGAGCAGTCTCCAGTGTTATATTCATGTAGCTGAATTGTTCTGTTGAGATGTTTGCAATACTATTCTGTATAGTCATGA
CTGCTTGTTTTTTGTCGACCCATTTTGTTAAAAGTCTATCCGCTGAAGTAGTAGTAAATTGAGTGAACCCTCTCATAACTATGTTCAAAAGAATCAA
GAATACTCCTCCTCCATAAATCATGAACTTTTTATACACGTACCAATCAACAGAACCTTCCTTCTTCTGTTCTTTATAAATTTTAGTTCTTGCTGTG
CTTTGTTCTGCTTCTATAAGTCCCGTTTTTTCACTCGATTCTTCTATAGTTGGCTGGTTGGAATCATGAATGACTTCTTTTTCTATACTATCATCCT
GATCTGCTATCATATAACTTTGTTGAATAATTCTGGCAGTAGGTTTACCCATGTCTTTAATTCTGCCGTTTTCCATTATGATCACGTTATCAGCTTC
TTTTATCTGAGAAGCTGTTTGGCACACTAAAACACATATTTTTCCTTTTAAAAATTTCTGGATGCATTCGTGAAAAATGTGATCTTGAACTGAAGCA
TCCAGTGCGGTCAAACTGTCATCTAAAAGGTATATATCGCTGTTTCTGTATACTGCTCTTGCCAGGTTTATTCTGGCTTGTTGACCCTTGCTCAGGT
TTTGACCATTGTCAGTTAATATAGTCTCGTCTCCGTTTTCAAAGAGGTTAAAATCAAATAGAAGAGCACAGGCCTTGACTACTTCTTGATACCGTTG
TGCATTGTATGGTTCTCCAAATAAAATATTCTGCCTAATCGATGAAGGAAACAGCCACGGGTCTTGGGAAGCATAGGAAATCCTTCCCTGTGTTTCA
ACGAAACCTTTCGTCAACGGCAAATCCTTCAGAATAACTTTTAAAATGGAGCTCTTACCAGATCCCACAGCACCAGTAATCATAGTCAAACCAGAAG
TTATCGTAAAGGAAACATTGTCCAAAACTGGTTCTCCTTTGAGTTGCACTGAGATTTCATCGAGCTCTACACGCGGTTTTTCTATGAGTTGCTCACT
GTGTTGATCTTTATTTAATTCTTCAGCATTTAGTGCTTTAGATATTCTCTTAAATGATGCTACGAACTCGCTGCCTAAGCCTAAACCAAATGGAACT
GCTCCTCCCAACCATCCTCGAAGGTATCTGAAGTTAGACAGGACGAAGAAAATCACTGTTGTGTCGGTTGATACACCTGTCCAAATTGTGGCCAAGA
TAACCATGAAGAAACCGAGGTTTGTAAATATTATTCCGTTAAGGATTAGAACTCTTTTGAGGTAGAATGCTAATGTAAGCTTTTTCATTTCGACTTT
CCTGGCCTGTGAAACTTTTCTGGAAAAGAAGTTTTCCCAAGTATACATTTTGATTATTTTTATCGTCGTAAGAATTTCTTGAGTAAGCTGCAACCTT
TCATCAGTTTTTTTACATCCTGCCAACCTACATTTTGCAATGAAAGTAGTTATAAGTATTTGTAAAGGGATAGATGACAACAAAACTCCAATCCCCA
CGAACGAGGCAACTCCAATTTTAGAGTAAAGCATGTAACATATAATAACTGTCTGTGTTATAGCAATCCACAAATCATTGATGGTAAAAATTACTGC
GACAAACGACTGTACGTCTCTTGTTATCAGTGTTACGATGTTTCCCAAAGAGGTTTCTGCTACAGCGACAGGTGATAACCTCAGAGCTTTCCTGTAT
AGCAAAGATGCAAATGCAGTCCTCATTTCCATTCCAAGCGTTTGGACCCACAACAAATACTGATGATTATAAACAGTGTCCAAACAGATATATATCA
GGTAGGAATATCCATAAAAGTATCCATCTTCTCGAGTATATGTCGTTTGTCCTGGCGAGAAGTATCCAACCAAATGACTTATGAAGTAAGGTGTCAT
GATACTTCTTGCTGTAGAAACTACAAAATCAATAAATCCAAGCAGAAGATATTTCAAACCATAACAATTCCATAGTACTCTAACCAACGAAGGATGT
TTCTCCAATTTTTTCTCATTCATCCATTCTTGTTCCAAATAATCACCACAAGTTTTTGACTTGCAGTTTGTGATGACTTCAAAAATGTCTTCCTCAG
TGAGGTCAGCTTTACCTGCTTTTCGAAACAATCCTCCAACATATGCAAAAGTTAAAAAAGAGAAGATGTTAGCTGTCTCTCTAGGATTAGGTTTTCT
CCTCTCTCTTTGACAATGATCCATTGTTCACAGATACTTAACATTGACTGACTAGACT
>Dvv ABC-C T10132_c0_seq1+D20603+T57212_c0_seq1*
GTGGGTTTGTTAAGATAAAAATCATTTTAATATTTCTTCCTGATGTTATTAGGTAAGATAAACTTATGTACATGCTTTAATAATGTGGCTTAAACAG
AATACATTGCTGTCTTCTTGTACTATGTTATGTTCTTTCAATTAGGACTGGAATTATTACTGGAAATACACAGGTTACATAATTTAAAAGAAATTAT
AATTTCACATAAATATAAATTACATTTTTTTACAAACACCAATTAAAAACATGAAGGAAATATTCAGCTACTTGCATTCGCGGCTTGACTTAGCATA
TAAGCTTTATTGGCAATTTCTATTAGGTTCTGAGCTGTATTTGTGCCGGTCTGCATCAGTAAGGAATAGAAAACACCATTTGTGTTTTGTAAAAGAA
TATAAGGACTGTCATATTCTAGGATTTCTCCAGCGTCCATAACCAATACTTTATCTGAATCCATAATCGTATGCAACCTATGTGCCACAGTTATCAC
TGTGCAGTCTGCAAATTTCTTCCTTATAGTTTTTTGAATGAGTCCATCTGTATGAAGATCAACATTAGCTGTAGCTTCATCTAAAATTAATATTTTG
TTTTGGCGAACTATGGCTCTAGCTAAACAAAATAACTGACGTTGTCCTACACTAAAATTTATTCCACCTTCTGCCATTTTATTATCTAGGCCTTCCG
GAAGGTCACTTACAACATATTTAAGCTCTGTATCTTCCAGAGCCTTCCATAATATCTCATCAGGATATTCGTTAAATGGATCCAAGTTATACCTAAC
TGTACCAGAAAACAGTACTGGTTCTTGAGGTATAATGGATATCTTTGACCTCATTTTACGTAATGGGATCAATTTTGAATCAATACCATCAATAATA
ATGCTACCTTCCGTAAAAGTCAACCGAAACATTGCTTGAATAATAGAGGATTTTCCAGCTCCAGTTCTTCCGACGATGCCAATCTTTTCCTTCGGGT
TAACAATAAAACTAAGATTCTTTAGAACCAAAGGTAACGTCTTGGAGTACCCGAGAGATAATTTGTTAAAAACAATTTGCCCTTGTTCTGGCCAAGT
TTTAGGCGGATCTACACCTTCGTCAATTTCTTTCGTTAGATCTGCATATTCTTGGACCCTCTCTACTGAAGTCATTTGGTTTTCTAATTCACTCCAC
TGCCTTATACCATACTGGAACATTCCCGTGAGACTCATGGATTGAGTTAAAGCCAACCCAACATTTCCACCAAATTTTTCTTCTTTTACAACCAAAA
GAGCAACTAAAACTAGCCCAATATAAACCACACAGAAAAAGTCTAACCAGAAGGCAAAAGTTCTAGTGGCCCCTAGATACATAAAAAAATGCTGCGCT
ATGAATATTTTGCAGATTGTCAAATTCCACAGCCAATATATCTTGTGCTTTAAATGCTCTTATCGTAGTTAGACCTTGAAGAGATGCCGTCAGATGG
GAATATACAGGACTACGGTTGATCGATTCCACTCTCTTGACATCTCGGCTGGTCTCCAAGAATATAAACCTCATAAAATAAAATATTACTAATACTA
TTATAGTCGGTATTAGCATCCAAGGACTTAAAGTTCCAATTATGAGACATATGGATAATACTGATAAACCGATCTGAAGTGTGTCCAAAATAGTAGT
TGGTAACACCTCATCTACAATTCCAATATCTTTGGAAAATCTATTCAAAATTCTCCCAGACGGATTATTATTAAAGAAATGCATAGGACTGTATATG
ATGTTGTCAAACATTTTATTATGCATAACTATAGAAGAAGACAGACCCCATCTGTAGAAAAATATTGAACGTGTGTAAGTGACTATAAGCAATAATA
AAACAAAAGTCGCATACACATATACAGTAATATTTCCGGTAAAGAAAGAGGTCAACCAGTGATCCTGTGTCAAGTCGTATGGATTAGCAATTACCGT

GGAAGATTGGTTGGATTTGTCCAAAACTGTTGTTGAATTATATGTTTGTATCACCCACTGCTCAATATTAACCCATATTGTTACAAAATATTCTGAA
GAAGTGTCGAGAACCTGACTCAGAATAAAAGCTAAGAGTAGTATTACAACTTTACATAAGTTTCCACCAGCTCTAATATAATTCATATAAACGTAAG
TAGTAACTTTCCCTACACCCTTCTCTTCTTTTATTAAAGCAGGAGCTTCTTTTTCTCCCTGGTTTTCGATATATAGTTTATTTTTAACGTGTTTCTG
ATTCTTTTCATCTTCTTCCTCTTTTTCTAGATTTTCTAGGAGTTTGGTAAATATACTATTCGATTTTATAATATCTTCGTATGAACCTGATGCCCTT
ACCATGCCATTCTCCAGCAAAATAATATTTGCTACAGTTTTTAAATATTGCAATTGATGTGTGACTAGAACCACACATTTGTCAGCTAAAAATCTAC
AAATGCATTCTTCAAACAAATGTTTTCCAACATGAGTATCTACCGCGGACAGCGGATCATCAAGAAGATATATATCTGCTTCTTTATATATTGCTCT
AGCTAAATTTATTCTAGATCTTTGCCCTCCACTAAGAGTAATTCCTCTCTCACCTATTCGCGTTTTATCACCATATGGTAAAAGAGTTAAATCTCTT
TGTAAAGCACATACCATGAGTACTTCTTCGTACCTCTGTTTGTCGTAGGGTTGTCCAAAGAGAATATTTTGCCTAATGGTGCCACCAAACAACCACG
GTTCTTGTGATGCAAATGATATTTTTCCATTCACTTCAATGTAACCCTCTGAAGGTTCTAATTCTTGAAGAATTACTTGAAATAGTGATGATTTTCC
ACTACCTACTGCTCCTACAACGGCCACCAAACTTTTCGAAGATACGTCCATTGAAAT
>Dvv ABC-C C217405_3.0+D12550+S19317Locus_19671_1*
TCATCTTTATACTCATCGAAAGGATCCAAATTATTTCTCAATGTTCCAGAGAATAGGACTGGCTCTTGAGGAATAATGGAAATTTTAGATCTGAGCT
TCTTCAAAGCCACAGATTTTGTATCAATATCATCAATAAGTATGCTACCATCGATATGAGTCAACCGAAAAACAGCTTGAATTAACGAAGACTTTCC
TGCACCTGTTCTTCCCACTATACCCACTTTCTCAGAAGGTTGGATAACAAAACTGAGATTCTTAAGAACTGGGGGATCGTCTGGTGAATACTGAAGG
GACATATTTTTAAATTCTACTTTTCCTTGATCTGGCCACGTCTTATCTGGTTCTACAGTGTGGTTGTCTTCTTCGTGTTTTAAGTCAGCATATTCTT
GTACCCTTTCCACTGATGTCATTTGGTTTTCGAGTTCAGACCATTGTCGCATGCCCCATTGGAACATTCTCGTCAGACCCATGATTTGAGTAAGAGC
AAGACCCATGTTTCCTCCAAATGTTTCGCTTTCAACAAAAATCAGCGACAATATAACTAAGGCAATGTATATTACACATATAAAGTCCAACCAGAAT
CCAAAGGTTCTATTTGCTCCTAGATACATGAAAAATGCAGCGCTGTGAGTATTTTGGTATTTATCAAATTCAGTTTTGAGTATATCTTCAGCTTTGA
ATGCTCTGATAGTAGTCAGTCCTTGCAGCGATGCTGTTAAATGAGTGTAGATGGGACTTCGCGTGACGGATTCCACCCTCTTAACGTCTCGACTGGT
TTCTAAAAATATGACCCTCATTACATAAAACAATATACCTATGCAGATAGTGGGTATCATGATCCAGTATGTCAAACTCCCTATCACGAAACAGATT
GCAGCTACCTCGAGACCAATTTGAACAGTATCCATGAAAGTCATTGGTAAAACTTCATCTAAGGCACCGATATCTTTGGAAAACCTATTAAGTATTC
TCCCAGAAGGATTTATAGTGAAGAACCTCATGGGACTATAAACTATGTTCTCGAACATGGTATTGTGCATTTTTGTTGAGGCAGTTATAGACCATGC
GTAGAATGAAGACGATCTAATGTTCATAATAACGACCATAAGAATGTTTAGAGCGCTAAAGTAGATTACAGTGTTTTCGCTTATAAACAATGGGGTT
AGCCACCAGTCTGCGAATCCTGTATGATCAACTGGTTGAGGAGAAGGACTGGTAATGTCTAACGGTGCAGTCATATTTTCTAAAGACGCAGGTGCTG
TTGGGGATGCAATATTTTGCAAAGCTTTCCATTGTTGAATATTGACCCAGAAAGTAACGAAATATTGGCACATACTCAACAAAGCTTGGCCCATTAC
AAACAGAATTATTAAACAGAGAGCCTTAAACAGATGCCCACCTGCTTGAGCGTATTTTAGATACACACGTCCAGAAATATTACCAGTACTCTTAGCT
TCTCTTTGTAACACCTGCACTTCTTCTTCTTCTTCATCTTTCTCTTGTTTTGCTTGAACAATTCTGGATTTCCTTCTTATTTCTTCCTCTTCCTCTT
CTATATTTGCCAAAACTTTACTGTAGTCTGTATTGGAGTTTCTTATGTCATCATATGATCCAGAAACAGCTACTTGACCATTCTTTAATAAATATAT
ATTCGGTATTTTCTTTAAATATTGTAGCTGATGCGTCACCAAGACGACGCATTTGCTTCCAAGGTATCCACAGATACATTCTTCAAATATCTGTTTA
CCAACATGAGTATCAACGGCGGATAGGGGATCATCTAGTAAGTAAATATCAGCTTCTTTGTAAACAGCTCTAGCTAAGTTGATTCTAGCTCTTTGTC
CTCCACTTAAAGTAACTCCTCGTTCACCCGCCAAGGTTCTATCTCCATGAGGAAATAGAGTAAAATCTCTTTGTAAGGCACACACTCGTACTACTTC
GTTGTATTTAACTTGGTCGTATTTTTCTCCAAATAGAATATTTTGCCTTACGCTACCTCCAAACAGCCAAGGTTCTTGAGATGCATAAGAAATTTTA
CCACTTACAGTAACAGATCCACGTTGGAGCTCGAGCTCCTTCATTATTATGTGCAGTAAGGTTGTCTTTCCACTTCCTACAGATCCAACCACAGCTG
CTAGATTACCAGGTGTTACATCCATGTCAATGTCTTCTAAATTGTTCTCGGGTTGCGACTTCAACCATTTAGCAGACGCGTTTCTAATATGGATGCT
TCCCTTATCGGAACTAGGTAATTCTTTCATAGCAAGTTTGCCGTTTTTGTTACCGTTTAGAGAGTTTTCAGGAATGAAATKAGCGTCAGTCTTGAGC
TCATCATACTGCAAGAATTTCTGGATTCTCTTCATCGATATGTACATCTCGGAGCCTTGAGAAACGGCCTGAGGGAAGAAGTAAGTCAATGTGCCCA
TCAGACGGTAATATGAAGTAACTGTGAAGGCATAGGAAGCTGTTAGAGGATTTCCTGTCAAAACAAATGTTACCAAGCATATGGCAACAGCTACGTC
ATTTAAAGTCATGGCCAAGGACATCAGAATTGCACGAATGACGGATGTTTTCCTGATGAACTTTACTTCATTCCCTCTAACTAACTCAACTAACTTG
GCAAATGGGTATTCCCAAGTGTACATTTTGATTACTTGAATTCCAGATATGATTTCATTCATTAGCCTTACCCTTTCATCGGTACGAGTTGCAGTTT
TCAATCTGAACTGCGAAGTTTTCTTGCCCAACCATGATTGTAAGGGTATAGAGAGTAACAAAAATATAGTTCCAAGAAGTGCTGTCCATCCAGCAAA
GAGATATAAAAAAACCATCACAATAAGTGCCTGTATTGGAGCGATATAGAAGTAATGAAGATGATGTGCAGCTTGATCAAATCTTCCTACATCATTA
GAAAGTAAATTTACCATCTGCCCAATGGTAGTTTCTGCTAGAGCAGATTTGCTAAGTCTTAAAGCTTTCCTATATATCAAAGAACACGAAGCTATTC
TCATCTTCATACCAACTTGCATCATTGCAAAATTGAATTTATGGAGCAGGATTACGCTGACCAATGAAGTGGCTATTATCAATCCTGAGTACAAATA
AACATCATTGATATTTTCTTTGGGATCCTTTTCATAAATCGTTAGTAGTTTTGATATTATAAATGGTTGGGTCATTCTTATCGCCTCTATTAATACA
ACAAGAACATTGAGAACAAATATTTCCCAAATAAATGTTCCTATTAACACCTTATAATATGCGGGATTCTTGTTCTTAGCTATGTGTTTGATCCACC
TTTCCTCTAGCCTATCTCCAAGCTTACTGGAATCATGTTCACTTCGGTGTCTGTACATATCATCTTCGGTTAGTTCTCGTTTGTAGCCTTTAACGAA
GTAAGGAAGTAACCAACAAAATAATAAATTGGAGAAATAACTTGATCGATCTATTGGATTAAGTTTTTGTTCTGCAGTCTGGCCTTTTGATTCCTCC
ATTTTGTAATTTTAGCTCTAAATTCAGTGCCTACAAGCCAGGAAATATCAGGAACATCAGGACATAAAGTAGTTTTAAGTAAAGCACTTGAACAAAA
CAAAACGATTTTTTATCACATAAAATAATTGATACCTAATTGAAGTGAGAAGTGTATGTTTGTGATATGA
>Dvv ABC-C T48940_c0_seq1+D21306
ATTTATTAGGAACTGGCTTTATCACAAATTTTTGCAAATAAGGATTTTTCATTTTCTAATAGCGTCTTCGGCTCATCGAACTCAACAATTTGTCCAT
TTTCCATCACAATTACCATATCACAGCTCAGAATAAAGTTTAACCTATGAGCTATCGTTAAGATAGTACAATTCTTGAATATTTTCTCGACCAACGA
ATGTATCAAGTTATCATTTTCCGTGTCCATGTTCGCGGTAATCTCGTCCAAGACTATAATCCTGCAGTTATTCAAAGCCGCTCTGGCCATGCAGATT
AGCTGTTTTTGTCCTGCACTAAAGTTGGCAACAGTAGAGTCCAAAGTTTGTATGTTTTCAATCGAGAGTTGGTTAAGCAGAGACCAAATCTTATCGT
CTGCGTGTTCTTTATAGGGATCTATGTTTTCTCGAAGTGTGCCACTGAATAATATAGGGTCTTGCGGTATTATGGCGATGTTTTTCCTCAAGCAACT
CAAGGATATAGTTTTGATATCCACGCCATCAATTTCAATATTTCCTCCGTATTCATATAGCCTAAACAACGTTGAAATAATTGAAGATTTTCCTGCC
CCAGTTCTTCCTACAATACCAATGTTGTGTTTGGGTTTTATTGTAAAATTAATATTTTTCAAGACTGGAATCTTGTTATCACCGTATTTTAATTCAA
CTTGCTTGTAAATGACTCCACCTTCTGTTGGCCAGTCCTTTGGTGTTTTGCCTGAGTCATCTTCTTGTTTCACTTTTGTGTACTCCAAACAACGTTC
CACTGACGTCATATTAGTTTCTATTTCTATCCACTCGTTGATACCCCATAAAATCATATCTGAGAGTCTTATTATATTTGTAAGCGCAAGCCCTACG
TTTCCAGCGGAAGTATCGTCTGGGAAAAACAGGAATTTCGCCATAATGATGGTGGAAAACGTAGCACTGAGTGCTCCCATATAAAATATCAAAGCAG
TATGGCTCAATCTCATAGTGAAAACTGCCGACGAATACAAATCCTGATGTCGATAGAATTCATTTGTCAGCCTTTTTTCAGCATCGAATGCTCGGAT
AGTTGACAAACCATCGAGCGTTGCATTAATGTGGCCCACCAAAGGGCTTCTCGTTGCAGCTTCTAGTCTTTTTAAATTACGACCCGTGTTGATATAT
AGAATTCTTAACGTCACAACGAGGCTGATAAAAATTATTGAAGGGAAGAGGAATACCCAATTAACTGTTGCTATGAGTACCAGGATACCTAAGCAGT
AAAATGCGGCCCTAAATAGATGAAGAAACATCATAGGTAACCTCTCGTCTAAAATATTCAAGTCATACGAAAACCTATTGACCACATTGCCTAAATA
ATTTGTATCAAAAAATGTCATTACTGCTGCTGTGATTCGAGAAACCATCAGTTTATGTACATTGTAGGAGGCATTCCTGCAAAATCTTAACAACGAG
TATAGAGTCACTAAATTGTACAGAGAAGAAAATACAGTGGTAAACGTGGTGAATAACAGAAGATCTTGCTTTTTGTAGGGTCCTCCAAGAATGGGGA
GCATTGTGCTGCCATCGTTAGCTTGTACGTCAACCCACTTTGCCTTTAATTTCTCTGAAGTAAAATGAGCTACTTGGGAGCCAGTATACAGTAAAAT
TATGAACACGACAAAGCAAATTCCTCCTCCGAATCTCATGTAAGACTTATATACATCCGCCCCAACCTTTCCCGTGTTATTTTTCTCGGCGAATAAA
TCGTCATTTGTATATTGATCGTCTTCGTGACTAGATTTAGCGGCCTCTGATTGTTTCTTAGCTTCAGACACCTGAACGGGTTCTTTTTGTTCCTCAC
CGATGTCTGGGGTGGGAGCCGGAATTATGTTAGTGTCACCTACCAGCTTTCCATGACTTAGAGTAAAAATGTTCTCAGCGTTTTCGATATGTTTTGG
ATTCTGAGAAATCAAAATACAAATCTTCTCTTTTAAGAATTTAAGAATACATTGTTCAAATATCACATCTTGGACGTTGGCATCTAGAGCTGTAAGC

GGATCATCCAGCAAATAAATGTCACTCTCTTTGTATATACACCTGGCCAAATTAACTCTAGCTTGTTGTCCTTTACTAAGATTAAGTCCAGCTTCAG
CTACAATAGTCTCGTCTCCGTTTTGCATAAATGCCAGATCGTAATCTAAACCACAAATTTTGAGAACTTCTTCGTAGCGTGCTTTGTCATATTTTTC
ACCGAAAAGAATGTTGTTTTTAATCGATGATGGAAATAGCCATGGTTCTTGAGCTGCAAACGATATGGTACCGCCTAATTTGCAGTATCCTTCTGAC
AAATTGTGGAGTCCAATGATAACTTTAAATATCGAACTCTTTCCTGACCCAACAACTCCTGTGATCACGTGCAGTCCAGGTTTCGAGATTTTTAGAG
AAACATCATCAAGAACCTTCTGATTTTTTATTTCTACCGAAACATTGTTAATATCAATCAAAGCATCTTTGTAAATCTTTTCTATTTTCGAAATTTC
TTGGCATTGGAAGACTTTGTTGATTCGTTTCAAAGATGCAATCAACTCTGCTGTTCTACTCATGTTTATGGGGATCATTATCCCCATGGCCAATGTG
AGTTCTCCGAAGAGGGAATCTATGTAGAAGATCAACTCGGTGTTGGTGTCGTAACCTAGCGCAGTGTAAGTAAGGATCAAAGCCAGGAAAACTATTC
CTGACAAAAAAATTCCGATGATGACGATAAGGAAGTGAACAAAAAACATCTTTAGAAGAGCTGACAATTCTTTGCTTCTTTCTTTAGATATTTTTCC
AATAAAAAACCTTTCCCAAGTATACATTTTGATGACACGAATTGCTGTAAGTGCTTCCTGCGTTATCTGTAGCCTTTCATCAGAGTGCTTTCCTACT
GTTAGCCGCAGCTTTGTTACTAATTTACAGAGAAATCCTTGTACGGGAAGAGCTACGGCAAACATTCCAACGGCCACGAATGCAGTTACCCCAATCC
TGCTATATAATAAATAGAACACGGTCGTCGTCTGTACGAAAAATATTACAAATTCTTTAAACACCCAAAGATTCGTCTCAATACTGTAGATATCTTT
TGTCATGACTGTTATTAGATTTCCAGAGCTGGTCTCTAAGATCGAGTTGGAGGAAAGCTTCAGCAACTTTTCATACAGGAGGGACTGCAGCGATGTC
TTGAGTCTTAATAAATATTTCAGTTCGAATACGTTGATGTTAGCGGTGTAAAATACCGAGAAGAATTTGATAAATATTACTAAACTCGCGCAGTAGA
ATGCTTGTGTTTCTGATATTGTCTGCTGTTTTTTGTCGAAATATGCAATTAATTTGCTCATTCCATACGGTCGTACTATACTGCGGATTTCGACCCA
TGTTACATGAGCTAAACAAAATAACAGGTAGGCTAATCCGAATTTGCTCAACAATAGTTTAAATAACGACGGATTTTTATTCCACTGCTTCACTGCT
CGCTCTCCGTTAAATTTTGATTTACATTTTTTTGGTAATCCATATAAATCACTCTCGTTCAAATCATGTTTCCAACCTTTTTTTATTAGCTCCAAAA
CATATCCAAATGTCAAATATGAAAATATATTAACTTTTGATAAAGGATTGAATTTCTCCTTTGGTTTTTCCATAATCTTATTTACCATGACTGGTAG
CCATTAAATCTTTCACTTAAAGGATAACGCTTACAACACAAACCCAAAAAAATTTATACCATTCAGTTTGGAATCTATGTTATTCACAAACATTAGA
TACAAGTGAATCATTATTAGTTGTTTT
>Dvv ABC-C T43960_c1_seq1+D21270+T49513_c0_seq3*
AAAACTGAGATTCTTAAGAACTGGAGGATCATTTGGTGAATACTGAAGAGACAAATTTTTAAATTCTATTTTGCCTTTTTCTGGCCATGTTTCATCT
ATTTCTACAGTATAATTATCATCTTCGTGTTTTAAGTCAGCGTATTCTTGGACCCTTTCCACTGAGGTCATTTGGTTTTCTAGTTCAGACCATTGTC
GCATCCCCCACTGAAACATTCCCGTCAGAGCTAAGGTTTGCGTGAGGGCAAGACCCATATTTCCTCCAAATGTTTCACTTTCAATGAAAATCAATGA
CAATAGTACTAAGGCAGTGTATATTACACATATAAAGTCCAGCCAGAAACCAAAGGTCCTAGTGGATGCAAGATACATGAAAAATGCTGCGCTATGA
TTATTTTGATAACTATCAAACTCATTTTTGAGTATTCCTTCAGCTTTGAATGCTCTTATAGTAGTCAATCCTTGAAGTGTAGCTGTTAGATGAGTAT
AGATAGGACTTCGGTTTACTGATTCCACTCTCTTAACATCTCGACTAGTTTCTAAAAATACGACCTTCATTACATAGAACAGTATCCCTATAAGGAT
AGTAGGTATCATTATCCAGTAGGTTAAACTACCTAGTACGAAACATATCGCAACAACCTCAAGGCCAATTCCAAGAGTATCCATGATAGTCATTGGT
AAAACTTCATCTAAAGCACCGATGTCTTTGGAAAACCTATTCAGAATTCTTCCAGAAGGATTTATATTGAAAAACCTCATTGGGCTGTAAACTATGT
TCTCAAACATCATATTGTGCATTTTTGTTGATGCTGTTATAGCCCACGCATAAAACGAAGACGACCTGATATTCATAATAATTACCATCACCATGTT
TAAACCAATGAAGTAGAATGAAGTGTTTTCACTCACAAAAACTGGGGTTAGCCACCAATCAGCAAAACCTGTATTGTCGGTCGATTGAGGAGGAGTG
GTCAAATTTAACGGTGCGGTAACATTTTCTACTACAACCCCAGTGTCTGTTGAGTTTAGTATGTTTTGTGAAGATTTCCATTGTTGAATATTGACCC
AGAAAGTAACAAAATATTGACATACGCTTAACAAAGCTTGGCTAAGTATAAAGAGGAATAATATAGAAAAAGCCTTAAAGATATGTCCACCAGCTTG
AGCATACTTTCGATAAACACGTCCAGAAATAATTCCTGTACTTTTACCTTCTCTTTGTAATACAGGACCTGCTTCTCCTTCAGCTTCTTCTTCTTTT
GTTTGAACAACCTTAGACTTCCTTCTTATCTCTTCTTCCTCCTCTTCTATATCATGCAACAGTTCACTATAATGTGTATTAGAATCTTTTATATTAT
TATAGGACCCGGAGACAGCTATTTGTCCATTTTTTAATAAATATATATTCTGAATACTTTTTAAATACTGTAGTTGATGAGTGACCAGTACGACGCA
CTTGCTTCCAAGGTATTCACAAATACATTCTTCAAATATATGTTTACCAACGTGTGTATCCACAGCGGATAAAGGATCATCTAGTAAGTAGATATCA
GCGTCTTTGTAAACTGCCCTAGCTAAGTTTATTCTAGCTCTTTGTCCTCCACTTAAGGTAACACCTCGTTCACCCGCTAAGGTTCGATCACCATAAG
GGAATAGAGTAAAGTCCCTTTCTAGTGCACATACACGAACTACTTCGTTGTATTTGAGTTGGTCATATGTTTCTCCAAATAGGATATTTTGCCTAAT
ACTACCTCCAAATAACCAAGGCTCTTGAGAAGCATACGAAATTTTTCCAGTGACATTAACAGATCCACTTTCTAATTCGAGCTCTTTCATAATTATG
TGCAACAAAGTTGTCTTTCCGCTTCCTACGGATCCAACAACAGCAGCTACACTACCAGCTGCCACATCCATGTCAATGTTCTCTAAATTGTTCTCGT
ATTGTGACTTCAACCATTTAGCAGACGCCTTTTTAATATGAATGCTTCCCTTATCAGATCCAGTTAACTCTTGTAAGTCTTGTTTACCGTTTTGTTT
ACCATTCAGTGAATTTTTAAGATTTTCAATGTTAAGTTCATCATATTGCAAAAATTTCTCGATTCTCTTCATGGATATCCACATTTCGGAGCCTTGA
GAAACGGCCATGGGTAAGAAGTAAGTAAGGCTGCCCATGAGACGATAATATGAAGTAACTGTAAACGCATATGAAGCTGTTAGGGGATTTCCTGTCA
ATACATATGTCACCAGGCATATGGCAACAGCTATATCATTTAAAGTCATGGCTATGGAAATTAGAATTCCACGGATTACGGATGTTTTCCTAATAAA
TTTTATTTCATTCCCCCTCACATACTCCACTAACTTTGCAAATGGGTATTCCCAAGTATACATTTTGATTACTTGAATTCCAGATATGATTTCATTC
ATAAGTCTAACCCTTTCATCGGTTCGAGTAGCAGTTTTCAGCCTAAACTGCGAAGTCTTCTTGCCCAACCATGACTGTAGCGGTATGGACAGTAGCA
AAAAGAATGTTCCAAGAAGTCCTGTCCATCCCACAACTATATATAAAAATACCATCACAATAAGCGCTTGTATAGGAGCGATATATAAGTGATGTAG
GTGAATGGGAGCTGTGTCAAATCTGCCTACATCATTAGAAAGTAAATTTACCATCTGTCCAATGGTAGTTTCGGCTAAAGCAGATTTGCTGAGTCTT
AGAGCTTTCCTATATATTAGAGAACACGATGCTATTCTCATCTTCATACCAATTTGAAACAGTCCTAAACAAAATCTATGTACAAGGACCACACTGA
TAAATGAAGTCGCAACTATTAAGCCTGAGTACAGGTAAACTTCGTTAATATTTTCTTTGGGATCTTTTTCATAAACGATTAGTAACTTTGATATTAT
AAACGGTTGAGCCATTCTTACTGCTTCACATAATATAACAAGAATGTCCTGACCGAATATTTCCCAAAAAAACGTTCCCATTAACACTTTCCAAAAT
GCAGGATTCTTATTTTTAGCTACGTGTTTGATCCACCTTTCCTCCAGCCTTTGTCCTAGCTTGCTAGAATCGTGCTCACTTCGATGTCTATACATAT
CATCTTCGGTTAGTTCTCGTTTGTAACCTTTGACAAAGTAAGGAAGTAAGTAACAGAAAAATAAACGAGAAATAAAATTTGACCTATCTTTTGGATT
CAGCTTTTTCTCTGGAACCCCACCAACTTTAACTTCCTCCATTTTATAATTTTAGCTGCAAGTCAGACAAATCAGCAAACAGTGTTTTTTTAATTGA
AGCACCTAAACTAAACTCGAGTTAAATCACTTATTATAATATACAAATAATTGACGTCTGAAGTATCTGTGTGTATGTGTGGGATGGAGAGTACTTC
GAGATACTAAAAACGGTGGATATTTTGTTATAACGGTGGTTATCAAATGCA
>Dvv ABC-C T45163_c0_seq1
ATATAAAATTTATTAAACTAGATAATTATCTATATAATAATAACATTAAAAATCAAGCTTCTTGCACTAGCCTATAAAAATAAGTGCTAGAATCAGC
CATCAGTACGTCGGGTTTGTCGAATTCCACCACTTCTCCTTGATGCATCACCAACACCCGATCCGAGTCGAAGATGGTCTGTATCCTATGAGCGATG
GTTATGACTGTACTACTACGGAAAGCGGTACGTATAGTGTTTTGAATAAGTCTATCAGTCTCTTCGTCAACGTTTGCTGTTGCTTCGTCAATGCACA
GTATTTTAGCATTGTGAAGAACAGCCCTGGCTAAACAAACCAATTGTTTTTGACCGACGGAAAAATTCGAACCACTTCCGTCAATCTTGTAATCTAG
TCCTCCCAAGGAATTAATAACATCAGTCAAGTTTACTCTTCCCAGAGCGTTCCATATTTCGTCGTCCTTGAACTCTTCCAGAGGATCCAGGTTTTCT
TTTAAAGTCCCGCTGAACAGGAAAGGATCTTGCGGTATGCAGAACATCCTAGATCTTAATGTTGCAAGCGATATTTTCTTGATATCCACTGCGTCGA
TAGTGATCACTCCTGCAGTTATATCTACTAAATTAAACAATACAGAAATGATGGAGCTCTTTCCAGCGCCAGTTCTACCAACAACCCCAAGTTTCTC
GTACGGCCTGGTATCGAACGATAGATTCATCAGTGATGGTGGCAAATGTTTCCTATAATGTAAGGTGACATGTTCAAAAGAGATAACACCTTGACTG
GGCCAAGCGAATGGAGGATCCATAACAAAATATTTACTTTCCATTGGAATGGATTCTACGTATTGGTTGATGCGTTCGACGGCCACCATTTCCCTCT
CTGTCTCCGTGAACGAGTTGACGACGCCGCTGAGTGATCCAGTGATTCCCAGAGCGTACGAAATGGCCAATCCGATAAAACCTGGATCCGCGATATC
ATACTGATGCTGAATGACGGCAATAAAACTAACACCTGTAATTATGATCACTCCTATAAACTGTAATCTAAGACCAAGCCATCGAGCTGCCGCTTGA
CTGGCGAATTGAGCTTTCAGATTAGCTTCTAGATGTTGATTGTTGTCGTGTTTGAATTTGTGGATCACTCTCATCGCTCTAATTGTAGTCAACCCTT
GCAAAGTTTCATTAAAATGGTTATATATGGGAGAAAGCGTTACACTAGATATCCTTTTTAGTTCCCTTGACGTTATCCTGTACTGATTTAATAAATA
TGTATATATGGGTACTAACGGAATGAGGAAGACACATATCCAAGGCAAACCGTACATCGTTATAAAAAGAGAACCTAGTAATCCGAATAACTGCGCA

AGAAATATATTAAGAATAAACGGCAACGAATCATCAACAGTGTAGGTATCGCTAGAAAACCTATTAATGATTCTTCCTATTGGTGTTACATCGAAAA
ATGTACATCTAGCTCGTAACACGGTCTTCAATAACACCTTATGAAACTTTGTAGCTGCCAGAATTCCACCATAAGCAAATATAAACGATCTTATTAG
CGTGAATAGCGTATTGACACACGCCAACTCGACGTATACTTTTAAAAACTCTGACATGTTATCATCGTCTTCTACGTAATCGATATAGTCAGTCGCC
GTCGCTAAGTATACTGTTAAGTTGGTAGAATTGGTAACGGGGCTGGTTACACCATTGGCTAACCACCAGTCAGTGAAATTCCGTGACACTTGCATTA
AAGTCATCGAAATCAGTATTGAAATGCTGACAAAGTGGCTGATGCCTTTCCAGTAAGAAATTATCACAGTAAAATCTAGATTACCCCGTTCGCTGAC
TTCTTTGAACAGTTCGATATCCTTATCCTTTTCTTGGGCGTTGTCCACGCACAAACTATTCTCTAGAGTGCACTCAGACGATGTATATTCAGATTGA
ATACTATCCTCCAGTTCTAAATCTACGGGTAAAGTGTCGTCTATGTTGCGGAGTACATCGACAGGTTTGCCAATTGACTTGATAGCTCCGTTTTCTA
ACAACAAAATCCTATCCGCGTGCAGCAGATATTTTACGTGGTGGGTGCATAGTAGTCTAGTTTTTCCTTGGAGAAGTCCAATGATGCAATGATTGAA
AATGTGTTTGCCTACTTTGACGTCAACTGCTGATAGAATATCGTCCAAAAGATATACCTGCTTCTCTTGGTAGATCGCTCTAGCTAGGGCTATTCTA
GCCTTTTGACCCCCGGAAAGAGTACCACCGCTATCTCCTACTTCTGTTTTGTCTCCTGATGGTAAGATCTGCATATCTTCGATCAGGCCACAAGCAA
ATATGATATCTTTGTATTTCTTCTCTTCGTACGGTTTTCCGAACAAAATGTTGTCCCTAACTGTTCCTCTCTGTAACCATGGTTGTTGTGTTACAAG
TCCAAATCCTGTTTCTACTTGACGTACAGCAAGACTGCCGGATTCTACAGCGAGTTCACCTAATATTGAGGAAAGCAGGGTAGATTTGCCACAACCA
ACTGTACCCATAATGCCAATAAATTCACCTTTCTTTATCTTCAAAGTTATATCCCTGAGCTTGAAAATAATGGTTTCTCTTCTTTGGACACTGCTTT
CGCCTTGTCCCCTTTTGGATAACGTTCTTTTTCCTTTCCCTTTTCCTTTAGGCACTTGTACAACGGGAGTGGCTTTCTCCCAGCTGAAACTGGCATT
GGAGATGGATATTTCTGTATTTTGGTCTGTATTATCTTCTAGTTCTTTGTGGTCGTAGATTTCATTTAGGTTCATATATGGAAGATCTAGTAACTTC
TGAATCCTCTTCAACGACACCCAGGCTTCAGTCAAACCATTCAGTACCCAAGGAAAAGCGTTAAGAGGGCTCACCAGCATGTTCAGAAGGGCTATGC
TTGTAAACACAGTGGCAGCTGTCAATTTGTTTCCCATCAGAACGTAAGTAACGAACGTCAAAATGGAGATGACCACTGGAGTAGTAGCCCAGAAGTA
TACGCAGAGGGCGTCCAGGTATTTACGACCTTTCAGGTACTTCAGTTCGGCGTCCCTTTGCCTAGTAATAAGTCTTATAAAGTGTTGTTCCCAAACG
TACAATTTGATTGCCTTAATGCCCCTGAGTATTTCGGAAGTCAACTTCACTCTGCCGTCCTTTTGTTCCATCATCTTGGTACTAAGATCGCCGATTT
TGTTCGCTATGCATTTATTTATGGGAATGAGGATTATGCTAAACATAACACCAGCTAAAAACGCCAAGCCCACTTGGTTGTACAGAAGGTAAAGGGA
TACAAATAGCTGGAATGGTATACTCCATACTGCGTGGAAACTTGGACAAGAATTAACAATCCGGTCTGTATCTGTAGACATGTAGTTCATTATCTCC
CCCACAGATAATTTAGACATTAAAACCGTAGATCTCACTGATAAAGTTTTCCTGTATATAGTATTGATTATAGCTCCCCTCATTTTGAGACCGATCA
CAGACATCTTGAAATTAAAGTGGGAATCACAGAGAGCAGACATAGAAGTGCTAACCACCAACAGCACAGCATACAGATAGCCCCATTTGATATCTAT
AGATCTCTTTTCTATAAAGTCTACTAATTTGTTCAACAAAATTGGTCCAGCAAAATTACTACAATCCGCTACCAACCTCAAGATACCAACGCTGTAG
AATTGCCAGGCGTAGCATCTATGAAGCGCCGCTAACAAACTAGAACTTTTGCGTGGCTGAATAACTATTACTTCTCGGGAAGAATCCGTATCGTGTT
CAATCCTCAGGTAATTTTCTAGTCTATTACTATTGTTGTCGCTACTTATTGAGTCCGGTAGATCGTACAAATCGTCAGATGTAACTATTTTATCTTC
AACTCCTTTGTTTACTAGAGACTTTACCCAGGAAAACGATAAACGGGACAACCAGTTAGTGTCTTCCATAGCAACCCCCAGATCGTTTGGATCTCCC
TCTTCACTGAATCTGCCGTAGGCATTGGCGTTTATCAAGGGCTGGCGCTCTGTGATCTCTGTATACCTCTCAGGAAAATTTAATGTGGTGGTGCTAC
CGCCTGGTATCAAACTAAGGGCATAAATGATCTGCAAGACGGTGTAATACAAAGTAATACCTAGCTCCATTTTGGTATTCTCTGTTGGTCTTTGCGA
ATTTTCCAAGTAGGCACTTCTCAACGATACGATGGTCATCGCGAAAACCATACACCAGACGATACCCATTGCAGTAGGACCTCTTTGACTTTTACCT
AACCTCGATCTTAGTCCTACAGTGTAGAGGAAATGGGAACACCAAGAAATGGCTTGGACTGCGCAAACCAGAAAGGAGGCGTTTTCTCTGTTCGGAG
TTTTCGTCATAATAGTATTGTATGCGTGCATTATTGGATAGAACATCAGTAGGAGACTTACTACATTTCTGGTTATGATACAGTATCTTTGAAGCTT
TCCTCGAGAAACAAAACCAGTTTGACGCCCAAAGTAATATGAAGATAGTACTGCAAAAAGGGAAAGTACAGGGATCTGGAAACATAGTTGCTGGAAG
CACTGGCCAAAATCGTGTCTTTGCTGTGACCATATTTGGAACCCTTCGGGACCACAGAGGCTGGTCCAGTTCCAATTACGGCCTATGACATCACTTG
TCATTTTAAATCATCTGGCACCATCTGCAGGTAGCTGGAAGTCCCAAGCAATGGCTTTTCTTTGTTTCTTAATTTCAGTTATAACTACTATGTGTCA
CATTATTACCTCAATACTTATTATTATTAGTTAATAATTAATTAGTAAAAGTAAAAGTAAAACAAAACAGCTGAGCTGATTGATGT
>Dvv ABC-C T49618_c0_seq2*
GCCCTCTCGGCGGCCCTGCTAAATAGCACAACACAGAGAAGATTGAAGAACTTATGGTCCAGGAGTGAATGCTAATAGGCTGAAGAAGATGATACTA
AAATTTTACTTAAAAACATTTTGATCTCGATAATAAATTAAAAATAAAAGTATATAGTAAAAAACAAGAAACTAACATTTAGCTATGTTCAATATTA
AGTGTCTTGGATATTGAAGTGTTAAAAACTGTTAATTTGGAGGAAATTTCTACAACTGATTAATTTATTTAAGTGAAACGCACAAACAATCAGTGCC
CCTATCAAACTACTACCTAAAACTGTCTGGATGAAGATGACTCATAGCTTCGTAACTATATTTAGCTAACTGATGAAGAGTTTCAGCCATAGTTTTT
CCTGTCTGCTGAACCATCCCATAAAAGATACCATCTTTATCTTGTAATAGTACGTGAGGGTGGTCAAACTCCTTCAAAGTTCCAGCGTCCATTACTA
GCACCTTATCTGAATCCATTACAGTATTCAGTCTGTGAGCTATAGTTAGAACCGTGCATTCTGCGAATTTGGTTCTTATGGTTTGCTGGATCAGAGC
ATCTGTTTGTGGGTCGACGTTAGCAGTGGCTTCGTCGAGGACTAAAAGTTTGTTCTTTCTGAGGATGGCTCTTGCTAGACAGACAAGCTGGCGTTGG
CCTACGCTGAAGTTGGTGCCACCTTCGGACATTTTGGAGTTGAGACCGGCTACCAAGTTTTCGACGGCTTCTTTTAGTTCTACTTCTTCTAAAACGC
TCCAAAGTTCAGCATCTGAAAATTCATCAAATGGATCTAAATTCTTCCTCATAGTTCCCGAAAACAGAACAGGCTCTTGTGGGATAATAGAGATCTT
AGTTCTCAAATCATGTAATCCAATCTGTGAAATGTTAACTTCATCTATCAAAATATCGCCGATAGTTTCAGTTAATTGGAAGAGAGCATTAATAAGT
GAAGATTTCCCTGCACCTGTCCTTCCGACGATACCGATTTTTTCTTTTGGTTTCACCACGAAGCTCAAATCCTTCAAAACGAAGGGATCCTGTGGGA
AGTACTTTAGGTAAAGGTCCCTAAATTCCAACTTGCCCTTCTCTGGCCAAGAAGGTTTTGGTTTCTTGTCAGGCTGAGATTCCAACGCGCCTTCTTT
TTCTATATCATTGTACTGGAGTACTCTTTCTACGGATGTCATTTGGTTCTCCAATTCGGTGGACTGCCTCATTCCCCATTGGAACATTCCAGTCAGT
CCAATAGCTTGGGTAATAGCTAAGCCAACGTTTCCTCCGAATTTTTCGTTACCGATAACTAGAAAACTGAATGTAACAAGAGTAACGTAAACCATAC
ACAGTAAGTCCAACCAGTACCCGAAAGCTCTGGACGTTGATATAAAAGCATACCAAGCTGAACTATGCAAATCTTGATGGTTATCAAATTCCTTTTT
CAAAATTTCTTCAGCTTGAAAAGCCCTGATTGTTGTTAATCCCAAAAGAGAGGCATTTAAATGGGAAAATACAGGACTTCTGGTGGTTCCTTCAATT
CTCTTTATGTTTCTACTAGTTCTTAAGTAAAAAATTCTGAAGAAAAAGAAGATTATTCCAATAGCTACAGTCGGTATCATTAAGAAAGGAGTTACTA
CTGCTACAACTACTATAATACCCAAAAGAGCTAAACCGATCTGTAGGCAATCTATTAAAGCTGTCGGCAAAAGTTCGTCAATGGATCCCATGTCTTT
GGAAAACCGGTTTAATATTCTTCCGGAAGCATTAGTATTGAAGAACCTCATGGTGGCTCTGGTAATGCTGACGAACATATTGTCATGGAGTCTGGTA
GAGGCGCGCATACACACAGCAAAGAATGTGAACGATCTGACTAACGTTATAATTACCAACAACACAATTAGACCCGAGTAAATGTATATCGAGGTTT
CTCTTGAAAAGTGCCAAAAGTTATTGGTGATGGTCTCATTTGATGAAAGTCCTGTTGATTGTGTTGAGTTTGCGGCACTGTCTTGTTGTTCTAAATT
TACCCAATAAGATAAGTAGTAATCTGAGGCACTGGCAAACATTTGTGATAACAGGAATAAAAAGAAGAATATAAAAATAACACAACAATTTCCTCCC
GCTCTGAAATAGGCTCTGTATACATATCCACCAATCGTACCGGTGCCTTTTTGTTCCTGGACTTCTTTGGGTTCTTCCACTTCTACTGAAGCAACAC
TTTTTATAGAGAGCTGACGCATGACTTCAGATTCTTTAGTTGTTTCCTCTTCCTCCGTAGATTCCCCGAGTGATTTTGTAAAATCTAATCCACTCGC
TTGTAATTCTCTAAAAGACCCTGTTTTAGGAATGCCATCTTCCAAATACAATATATTGTCCAATTCCTGTAAGTACTGTATTTGATGAGTAACTAAT
ATCACAGTTTTATTTTTTAAGTAACCCGTAATGCACTGTTCAAATATTTCTTTCCCGACGTGAGTGTCCACAGCGGATAAAGGATCGTCTAGCAAAT
AAACGTCGGCTTGCTTGTAGACGGCTCTTGCCAGATTGATTC
>Dvv ABC-C T47333_c0_seq11
CAAAGATAAAATGGAACAACTATAATAAAGATATTGTTTGCACATTAATATTCTTATTTTCATATTTCTTACATTGAAACAGGGGCAATCTCAAATC
TCGCTTTTGAGTCGATTGTTTTGCACTCAAGTGTAGTTTAAACTTTTGTAATACTTCCAATCGTTACGGAAGGAATAACTTTACAATACACTGTTCA
ACTTCTGTCTATTGTATCGTCATTTTCAACGATAAGTAAAACAAAAAATTTTAACTATCTATTCCCCCTTTTTCCCCTTTATTCCGCCGTACTTCTT
ATTATAATCTTCCTCTGCCACTGCCTTCAACATTTCTTCCATAGCTGGTCCCGTCTCTCTAAGCATTCTACTAAAATATCCGTCTGGGTTTTGTAGA
AGTTTAAAAGGATGGTCAAATTCTACTGCTTCTCCGGCATCCATTACCAGCACGCGGTCTGAATCCATAACAGTATTAAGTCGATGGGCGATGGTAA
GTACAGTACAGTCTCTAAACCTCTCACGGATTGTTTTCTGGATTAGGGCATCAGTGCCTGGGTCTACATTGGCAGTAGCTTCATCCATAACAAGAAC

CTTATTGTTTCGTACTATAGCTCTTGCAAGACAAAGGAGTTGTCTTTGACCTGCACTAAAGTTTGAACCTCCTTCGCTGACAGTTTGATTTAAATCA
GTAATGGCATGTTTTAGTTCAACGCTTTCTATTGCCTTCCAAAGTATTTCATCGCTGTATTTTTCAAATGGATCTAAATTGTATCTAACTGATTCCG
AAAATAAAACGGGCTCTTGAGGAATTATAGATACATTCAGTCTTAAATCTTTTAAGCCTATTTCTGCCGTATTTAAATTATCTATGGAAATCGTTCC
CTCGATTGGGGCTAACCTGAACAAAGATGCGATCAATGTTGATTTTCCTGCTCCGGTTCTCCCTACTATACCAACCTTTTCTCCAGAATTGACTTCA
ATATTCAAATTCTTTAAAGCAGGAGGTACATCAAGGGAATATCTAAGGAAAGTATTTTTAAAAATAATTTTTCCTTTCTCTGGCCAATCTCTAGGAG
GTTTGGTTAAAGATTCGAATGGTCCTTCCTTCTCTAACTTTGTGTATTGCAACACCCTTTCTACACTTATCATATTGCTAATTGCCTCCGCAGTTTG
TCGCACTCCATATTGTAACATTCCGGTTAAAATTAAGCTTTGGGATAAAACCAAACCTACATTTCCTCCCGCAGTGCTATCGACACTGAAAATTAAG
AATTGATAGGTAAGTAGTCCCAAAAACAGCGTACTGATAATGTCCAAATAAAGTCCAAAGGCTGTACTACTAACAAGAAAACAGACATATGTACCAG
TATGCTGATCTTGCAAAACGTCAAACTCCTTTATGACCATTTCTTCAGAATTTGAAGCTCTTATAGTTGATATTCCTTGTAACGTAGCAGAAACATG
TGAAAAAACTGGGGCTTTTGCTACTCCTTCCCATCGTTTAATGGATTGTGCACTAGCTAAATAAATATTACGGCAATAATAGAATAAGAATCCAATC
AGTATAGTAGGAATAATCATCCAAGGGGTTACTATAAATACCAAAACTAAAACGCCTACTAAGACCGTGAATATCTGAATTGTGTCAATACCAGTCT
TTGGCAACATTTCATCTATTGCTCCCATATCTTTTGAAAATCTGTTCAATATACGACCTGAAGGATTTGTATCAAAGAATCTCATAGGAGCTTGAAG
AATATTACTGAACATTTTATTATGTAATGTTCTTGAAGATGACATACATACTATGAAGTACAAGAAAGACCTCAGGGTTAATAAGATCATACTGGAA
AACATCAAAATGGTATAAATGATTATGTATCTGTTCTGATCATAAACTTGCTCAGCTTCTATTCTTACTGTAGTAGTGTCGATCACAGAGGATATGG
TTAGCATGGATTCCATGACAGAAATATTCGTGGGTAACGTCGGTAGTGTTAATAAAACCGGATCTGCAGAAGTCGTTAAGGGTAGCAGAGTAGTTAG
TGTGTTTGTCAAATTATTTCGATTTGTTTCACTAGTATTTTCAACATTTGCGGCCAACATATCAGCGTAGTATTTTCCTTCGACATTGGTCCAGTGT
GTCAGCCACAAGTCAGTTATACTGGTAAAAGCTTGCCCAGCAATGAATGCTACCGCTGTGAATGTTAAAAAGAAAACACCTGCACCAGATCGGAAAT
ATTCGACAAATATTGAGAATGGGACAGTGCCTTGTTCAACTTCCTCTTCTCCTCGTTCGTTGGATTTATCTTCTGGCTCCTCGGAAACCGATAATGA
AGTATTAGATTGATAATGCTCTGAACGAAGTGTTTTCTCACTGATGGTGCTTTGTGATCTTTCTACTTTTGATCTCTCTTTTTTAACTTCATTTTCA
TCTTCTTCTGGTTCAGGTTCTTGTTGTAAGACATTAAGCTCATCCTGAGATAATGCACTGAAATTTGCTGCTTTTTCGATTCTGCCATTGTTTAATA
TAACAATAATATCAGCTTTCTTCATATACTGAAGCTGATGAGTCACTAAAATTCTCGTCTTATTTTTTAAATATTTCACCATACATTCATCAAATAG
ATGTCTTCCTACTTTTGTATCCACTGCAGAGAGTGGATCGTCAAATAAATAAATATCAGCATCTGAATATACTGCTCTTGCTAAGTTAACTCTGGCT
TTTTGGCCACCACTTAATGCTGCACCGCGTTCACCTACTAAAGTTTTATCGCCATGCGGTAACTCCTTCAGGTCTCTTTCTAGTGAACACACACGAA
TGACTTCTTGGTATCTCTTTTTATCGTAAGGTCTACCGAACAAAATATTCTCCCGTACATTAGAAACAAATAACCATGGTTCTTGACTAGCATATGC
AACATTTCCATAAACATTCACCTTTCCATTCGTAGGGGACAACTCTTTGAGCAAAAGGTGTAATGTACTACTTTTCCCAGATCCTACATTTCCAACA
ACGCAACACAGAGTTCCTGGCGTTAAGTGTAAGTTCAAATCTACTAGTGTTGGTGTTATTGGATTCTCTGACCAGCTTGCTGTGACATTAAACAACT
CTATTTCGCCTGGTTTCCCTGCTAATGTATCATTTTCTCCAAGTTTTTTCTTCGGCGTGGTCTCTTCCAGAAGAAGGAATTCTTCTATTCTTTTTAT
CGATACTTTAGCTTCAGCATATGTAGATAAGGCTCTTGGGAAGAACATCGCAAAAATCATTTGAACAGTATTTACAAGTTGAGCCATTGAAAATACA
ATATCCCCAGATATTCTTCCTTCCATTAGTACAAACGTTATTACCACAATAAATAATGCAAAACGTTCTGTAAAAACAGACATTGCCGATAGCATAC
CATGAATGTAAGATGATTTTGTGATGCTATCGATTTCCAACCTCCTAGCTAGTTCTACAACTTTCTCAAAAGGTTTTTCCCATGCATACATTTTAAT
TACTTGAATTCCCGTCGTTATTTCACTCATGAGTTTTATTCTTGTGTCTGTCCTAACAGCAATTTTTCCCCGTAATTGTCCTTGCAGTCTTGATAGG
TAACTTTGAAGAGGTAGTGCTTCCACTATAATTGCACCCATTCCTGTAAGGGCAGCCACAAGATTGCATGTTCTGTACATAATGTAAAATGATATCG
CACTGTTAATGGGCATTACCCAAATGTAATGTAGGAATACTGCAGCGTAATCAAACCTCTGCAAGTCGTTGGATAAAAGATTGACTAGTTGCCCTGC
TGCCGTTTTACCAAGGGAAGTGTGGTTTAATTTTAACAGTTTTCTATACAATAAGGAACCCACTGCAATTCTCACTCTCATACCCACTCGTTGAGTA
TTTAAAAATGTATTATGAAGAAGGATAACATTTAAAAATGCCATAATGACAACTCCAGACGCTAAGAGCCACCCAAATTCATTAGGCTTGATTCCAT
GCTTTTTGTCGAAAAAATTAATATATTCTGCGAGTACTATGGGTTGTAACATCCTAATAACTATAAACTGCAAAAATATCATCACACCAGATGCTGT
ATACATTTTTAAATATGTTTTCACAATTGCTGACTTTAAACTGGGTTTCTTTTTATTTTGGCTTTGATCACATTTTTTAATTTGTTCTTCCCAATTC
TTTTGAAGTTGATTTCCCAAAGATTCAGACATATCAGGTTTTGTTGCATTGTATATGTCTTTAAGTTCAACATCGTTTTTATATCCATATTTAAAGA
ACGGAAGAAACCACAAATTCAGTCTGGTAACAGGTTTCCTTACTCTAGTGGAACGCCTTAAAATGGGTTCCTCAGTTTCAGTTTGAGCCCCAGGTAC
ACTATTCGAACTTTCCATATTGTTTGAAATTTCGGATTCATCACCCAGATAGCATGCAATATTCCTGGGACATTCTCTCAATGTCCTGATGTAAATT
TATTTCCAGAGCATGTATCCACTCTCGACCAACCAGTGAGTAGCCATCACCATCCACCACATACAACTTTAGAGTGTGCTTTAATTCCTGGTGCTCA
ACCTCCACTGAAACTAAACCTAAAACATTCAAATTATTTTTGCAGTATGTAGTTAATTTAACATCCGATTTTTGTAATTGTAATGCAGGAAAATATT
TTTCAAATATATTTTTATTTATGATTGTAACAGCAGCGCCAGAATCAATTTCAAAATTTAAAATATTACCATTCACTTTTAAATTTATAAAAAATTT
ATCTTTGCAATTTCCAGCATTTACTTCTAAGACCTCTTGAGCAGAGTTTACTGAATTTACTTGCCTGTGTGAATTTTGTTGTGCCTTTAAACAAACT
TTTTGAATGTGTCCGACGTTTTTACAAAAGTTACAGATTAAATGCTTTTTATTACATTTATCGGCTAAATGCGAGGTGTCTCCGCATCTATAACAAG
CCCTATTACGGTTATTTGACATTGTAAAAGTATTATTAGGAGAGCTACTATGATTATTTACTGTATTATTATCATTAAATTTACTAGTGTTTGTATT
TTTATGAAAACTTTTTGCTTTCGCATTTAAAACATTTATACTAGCCTGATTGTAATTATTGTTGTGAGAAAATTGATTACTATCCTTTTCTGAAGTT
TCCATGCTTGCTGCAATTTCTACGGCTCGGTCCAAGTCCAATCCCTTTGTTTCTAATAGTCTGGCTTGAATCCTCTTTGACTGTAAACCAAAAACAA
ACTGATTTCGTATAGCACTTTTTAAATATGTGGAAAAGTTGCAGTTTATGGCAAGTTTCTGTAGAGAATGTAGGTATTCTTGTATACTTTCGCCTTC
TGCTTGTCTTTTTGATTGAAACCTAAATATTTCAGCAATTTCAAGTGGTGCAGGGTTGTAAAAATTATCCATTAACTTAACTGTATCCTCATATGAC
TTGTCTTCAGGGACCTCTGGAGCTAGTTTATCGCACAGTGTATCGTATGCTTCCGACCCCATGTAATGTAGTAAATAGGGCAATTTCATTTCCTCTG
GAATTTTAAAAACTTTGTATGCTCCTTCCAGCCGTTTTACCCACCTGGACCATTTTGTAGCTGTCTGGTTAAAAGGTTCAACGGAAAACTGGAATGT
AGAAACTTGTGTAGACATTACTGGAGCTGTAGCAGGTGCAACTATCGCAGTTGTAGTCGAGGAACTCGTGTTTGTACCGGCTGTAGTTGGAACGTTG
TAGTTGATGTAGTGTCTGTAATATTTGTATCCATTATCCTCGTCGCCAAATGTCAAGTCCTTTGTACTTAAAGGACGTAAGTATTTGTAATGAAATA
ATACTCGATGTAAAATACTTTATTTTAAAGAATTATGCACATGGTTTCCATTTTCTATCGCAGTATTAAGTATGTGACAGCTGTCGTTCAGATGACA
GTAGTACCAACCTACATTCATGGTTAAGGGAGTTTTACACGAACATAACACCTCCTCCCTTGACCATGAATTACAAAAGTTAACTAGAAGTAAAAGT
ACTT

>Dvv ABC-C D7536

TTTGTTTTTGATTTACAACAATTTATCAATATAGTTTATTTCTTTGTTCTGCAGCTGCAATGTATCAAAAGATCACAAATTTGTTTTCACTTTAATA
AGATATTACAAACGTATAACATTTTCATATATTTATATAAAATAGACATCTTAATTTAAGGAACTAATCTTTGTTATCTTCAATTGCTAAATTTTGT
GGAACTTCACTGGTTAACGRTCCCTTCTTCTTAATGTAGTCATTCTTTGCTACATTACTTAATAGTTCCGCCATGGCGGGTCCAGTTTCCTGGACCA
TCTTGCTAAAGTATCCTTCCGGATTTTGTAGAAGCTGATAGGGATGGTCAAACTCCATTGCCTGTCCGGCATCCATGACTAAGACTCTGTCTGAATC
CATAATGGTGTTTAATCTGTGGGCAATTGTAAGTACTGTACAGTCTTGGAATCTCTCACGAATAGTTTTTTGGATGAGGCCGTCAGTCTGTGGATCT
ACGTTAGCGGTAGCTTCATCCATAACGAGAACTTTGTTATTTCGTATAATAGCTCTAGCGAGACATATAAGTTGTCTTTGGCCAGCACTGAAGTTAG
ATCCTCCTTCACTGACAACTTGATTTAAGTCACTAATAGCACCTTTGAGTTCAACGTCTTCTAACGCTTTCCATAAAACCTCATCACCATGTTTCTC
AAAAGGATCTAAATTATACCTGAGTGAAGCGGAGAACAAAATTGGTTCTTGGGGTATGATTGAGATATTTAATCTTAGATCATTCAAACCTATTTCG
GCAGTATCAACATCATCTATTGATATAGTACCCTCTATCGGTGCTAATCTGAAAAGTGACGCAATGAGGGTTGACTTTCCTGCACCAGTTCGTCCAA
CAATTCCAATCTTCTCGCCAGAGTTTACTTCTATATTTAAATCTTTCAATACAGGTGGAAGTTCTGGAGCATATTTTAAGTACGTATTCTTAAATAT
AATTCTTCCTTTTTCGGGCCAATTCCTTGGTGGCTTCTTGCCAGGTAGCGTTTCGAAGGGCCCTTCCTTATCCAGTTTAGTATATTGCAAAACTCTT
TCTACGCTAATCATATTACTCGCCACTTCTGCACTTTGTCGCACTCCATATTGTACCATGCCCGTTAATATGAGACTTTGAGATATGACCAAGCCCA
CGTTTGCACTTAAAGTATTTTCTGTTCTAAAAAGAAGAAATTGGTAAGTCACAATCGCCAAGAAAAAGCAGCTCATGACATCCAAATAAAATCCAAA

AGCGGTACTGCTGACTATAAATAAGTACCACGTGCTGGTGTGCTGATCTTGTAGGATATCAAATTCTGTAGTTATCATTTTTTCAGCGTTTGAGGCT
CTAATTGTTGTTATACCGTACAGTGAAGCCGATATGTGGGAAAACACCGGAGCTCTTGACACTCCTTCCAGTCTTTTTAACGATTGTGCGCTTGTTA
AGTAAACTACACGGAAGAAGTAGTAAAGAGGAGCAAGTATTACTGCTGGAATGATCATCCAAGGAGTCACTATGAACACCATAGCTAAGATACCAAT
CATAACCATAAAAATTTGAATAGCGTCAATTTGACAGCGTGGTAAAAGTTCGTCTACTGCACCCATATCTTTTGAGAACCTGTTTAAAATTCGCCCT
GATGGATTTGTATCGAAAAATCTCATCGGTGCTTGAAGAACATTGTTAAACATTTTATTATGCAAAACTTTGGAAGCAGTCATACATATCTGGTAAT
AAAGGAATGATCTCCACGTAAGAAATACAACAGAAGCTAATATAATTCCAGTGTACACATAAATATAATATTCTGTAGACGGAATGTCATTCAGTTC
TTCTAAAGTAGCATTTCTGATGHGGTCTGTTAAATATCGTTTCTGACTTAAACTTAGAAGTGAACTATCTGCAGATGGTGTTGAATCTGTTAATAAC
GTAGCAGAAGTTGACATCTCATCAGATAATCCAGTAGACATTAGTAATGGTCTTTCTAAGGTTTTATTGGAAGTTGCATTGTGTAGGGAATCACTAA
GATGGTATCTTTTTGCCTCGTTATTAGTCCAATGTGTAAGCCAAAGATCTGATGCATTCGTAATCATTTGAGCGACGAGGAATATAAAAATTGTAAA
GAATAAGAATAAAACTCCACCGCCTGATTTCCAATATTCCACATAAGTAGCATTTGAAATGGCACCTTTTTCGATAAGTTCGTCTTCTTCGTTCGGA
TCATCCGAGAATACAGAACTAGCCAAAGAACTCATAGACTGGAAATGAGGTACTGTCGAATTTTTACGAATCTTATCGCCAGTCTCTGGTACCTTTT
CTTTCTCTTTATCATCGGCCTCAGGTTCTTGTTGAAGAGCATTCAAGTCATTTTCCGACAATTCGTTAAATTTGGCTATTTTCTCAATTTGACCATT
GTTTATAATAATTATCAAATCAGCTTTCTTCATAAACTGGAGCTGATGTGTGACCAAAATTCTAGTTTTTCCATATAAATATTTCGATATACATTCA
TCAAAAAGATGTCTTGCTACCTTTGTGTCGACCGCCGATAGCGGATCGTCGAATAAGTAAATATCTGCTTCAGTGTATACTGCTCTGGCAAGATTAA
TTCTTGCTCGTTGTCCTCCACTTAACGATGTTCCTCTTTCACCTACTAAGCTTCTATCACCAAAGGGAAATTGTTTAAAATCTCTTTCTAGAGAGCA
CACCTTAACTACGTCATGGTACCTGTTTTTCAAGAATGGTTTTCCAAAAAGTATATTTTCTTTCACGTTTGAAACAAAAAGCCAAGGTTCTTGGGAA
GCATAGGAAATTTTCCCGTTAATCTTCATTTGCCCCCTACTTGCCGGTAATTCCCTCAGAAGTAACTGCAGAAGACTACTTTTTCCACAGCCTACGT
TACCAACGACGCAACATAAAGTTCCCGGTTGAATGTGTAAATTGAGATCCATCAGAGTGGGTGCAATTGGTTTCGATGCCCAACTGGCTGAAATATT
AGTCAATTGAATTTCTCCTAATTCGTCTGGGTTGGAGATTTTTTCCGGCTCCGGTATTTTTTCGTTTTCATCTAATAAAAGAAAGTCTTCTAACCTT
GTAATAGAAACTTTGGCTTCAGAATAAAATGACAAAGCTCTTGGGAAGAATATCGCCATGATCAATTGAACAGTGTTGAATAACTGAGCTAAAGAAA
AAACTACATCTCCCGTTAATCTTTGTCCAACTAGTACAAAAGTAATGACTGTGATGTATAAAATCATACGTTCTGTGAATATGCCCATTGCTGATAG
GACGCCATAGCAATAAGAGGTAGTTGTTATAATATCTATTTCGTATTTTCTAGAGAGCTCTACGACTTTTTCGAATGGTTTTTCCCATGCATACATT
TTGATGACTTGTATTCCAGAAGTAATTTCGCTCATCAGTTTTACTCTTTTGTCTGTCTTTAGAGCAATTTTGTATCTTAGTTTTCCTTGCCATTTCG
AGAAACTACCTTGTAAAGGAAGAGACTCCAGAGTAATGAAAACCATTCCAGTAACGGCAGCGATGATGCCAACATATCGGTATAAAATATAAAACGA
TATCCCTGCAGTAAGTGGCATAATCCAGGCATAATGTATAAATTGGGCGGCTAAGTCAAATCTTTGGACGTCATTAGATAAGAGGTTAACCAGTTGC
CCGGATGCTGTTTGACCCAGGGAATTGTGGCTTAACTTCAAGAGTTTTCTATAAATTAACGAACACACTGCTATACGCACTCGCATTCCTACCCTTT
GGGTATCTAAAGTACAACTATGCGTTATGATAACATTAAGAAAGGCCATTCCGATTACACCAGTTGCCAGCCACCAACCAATGTAAGCTTCTTGATT
ACTGTCAAAATAGTTTATGAATTCGGCTAATACTATCGGCTGAAGCATTCTAATAACTACAAATTGAAAACCTAATGCTATGCCAGATCTGGAATAA
GATTTCCAAAATGTCTTTAATATGGCATTTTTTAGGCTAGGCTTCTTTTGCTTTTTGTCTTCATGATTTCTAATTTCATCTTCCCAATTTTTTTGCA
ACGCATCCCCAAGAGCGGCTGATAAATCACCTTGTGTGGTATTATATACATCTTTTATTGTGATGTCGTGTTTGTATCCAAATTTAAAAAACGGCAG
AAACCACCAGTAAAATAGTTTTGAAATAAGATTAGCTTTTTGTTCCGGTGATACTTTAACGTATTTTTTGGTTGAATCCATCTCGAACTTAAGTACT
GTTAGTTAAGTAAGGAAATATCACAAATGTTTTTGACCAAAAACAAAGTACATATTTGCACTCGGCCCTTTGATTTTCCGACTACACGCCTATGAAG
AGAACTAGGCCTACTACAATAACAGTTACTGTAAATGCACTTCTTCACTTTTGGTACACTTTTCAACAAAGATTTATTTCATAAATTATCAAACATT
TTTATAACCTATAGTATAGCACAGTCGGTGTTTTAATCAGCTAGTTATCACACTCGCACTTATACTTATCACTTATCCACACTTTATAATCTACAAC
CTAATACAAACAGCTCCGCTCCAACTACTATCGACTTAACATGTTCGTCT
>Dvv ABC-C D20002
TTATGATATGATTGTGTAGATATTAGTGACAATATGGATATAGGGTACAAATTGAAGGCAGAAAATCCCAAACAAAATGCAAATATATTTTCAAAGA
TATTTTTCGGGTGGATGATACCACTAATAAGGAGAGGAACTAAACAGAATTTAGAAATAAATGACTTATACAAAACTTTAAAAAAAGATCAATCAAA
AAGACTAACAGATGCTTTGGAAAAAAAATTGGCAAAAACAGGTTGATAAAGCGAAGAAGAAGGGAGGAAAACCCAGTGTACTTATGGCTATATCAAGG
ACGTTCGCTTTTGAATTTATGATGTACGGAATACTATGGGCCATTCAAAATGTAGTTCTAATGTCTCTAAAACCAATATTGATTGCACAATTGATCG
AGTTATTCACTGACGATACTTCTACCAGATTTAGAGAAATGTATTTCTTCAGTACATCATTAATTTTAGTGTCACTTTTAATTGTATTCTTCTTCCA
CCACACGAATTTTGGTCTGCAAGCAATTGGAATGAGGATAAGAGTTGCGACTTCATCCTTAATCTACAGAAAGATTACTCGGTTAAATCAAAAATCT
TTGGGCGAGACTGCTACAGGACAAATTGTCAATCTACTATCAAACGATGTCCAGCGATTTGATATGGTGGTGGTACCGTTACACGCGTTGTGGGTGA
TGCCGCTTCAGGTGGCCATACTTATGTTTATAATCTGGAATCAAGTTGGCATTTCGTCTTTAGCAGGTGTTATTTCGATGGCAATCATTGCTTTGCC
CGTACAAGGTTATATGGCAAAGCTCATGGGTATGCTTCGACAAAAGGTTTCCGGAAAAACCGACACCAGAGTGAAACTAATGAATGAAGTCATTGGT
GGTATACAAGTGATAAAGATGTATGCCTGGGAGAAACCATTCGAAAAAGTCATAAAACAAGCTAGATCATCGGAGATTGGTGATATAACAAAGGCAT
CCTACTTGAGAGGTGTATTTTCTAGCTTCATCGTATTTTTGGATAGAGTGGCTCTATTTTTCACTGTTATGACCTACGTTTTACTAGGTAACGTTAT
TTCCGCTGATATAGTTTTCTCATTGGCCCAAACCTTCAATATTCTACAAACAGCCATGGCCATTTGGTACCCAGTTGCAATTAGTGTTGGAGCGGAG
GCTCTAGTTTCAGCTAAGAGAATTCAAAACTTTTTAATTATGGAGGAAAGAGAAGAAGCGTCTATCGAAAAAATAGATAAGCCCGGTATAATACTAT
CTAACGTGTTTGCTTCTTGGACAACTAAGGGACGTACTTTACAAGATATTTCATTCCAAGTTCCACCAGGTACTTTGTGCGCTGTAGTAGGACCAGT
AGGTGCTGGAAAAAGCTCGTTGTTACAGCTCCTTTTGGGAGAGCTGCCACTAAAACATGGAAGAGTTCAACTGGGAGGAGAAGTTTCCTACAGTTCG
CAAGAACCTTGGCTTTTCCAATCAACTGTACGGAACAACATTCTCTTCGGAAGTCCCTACGAAAAACGCTGGTATGAAAAAGTGGTTAAAGTGTGTG
CATTGGAACGCGATATGGAACAGTTTCCACAAGGAGATAAAACTATTGTTGGAGAAAAGGGTGTTTCACTTAGTGGAGGACAAAGGGCTAGAATAAA
TTTAGCACGAGCAATTTACAGACAAGCTGATGTTTATTTAATGGACGACCCGTTGTCTGCTGTTGACACACACGTTGGACGACATTTGTTTGATCAG
TGTATTCTTCACCATTTGAGAGGTAAAACTAGAATCCTAGTAACGCATCAGCTACAATATCTGAAGAAAGCTGGTTTAATCGTTGTACTTAATGACG
GCAAGATTGAGGCTCAAGGCACCTTTGAAGAGCTAATGGATAGTAAAATGGACTTTACAAAACTCTTAGTTGCTGCCGATGAAACTGGTGAGAAACA
TGACAAACAAGAGGATGCTGATCCAGAACCTGTAGATTTTACAAGAAAGTTTTCTAGTACTAGAAGATTTTCGGTATTGTCCGACGCCAGTGGCGAA
ATATCGTTGAGTATGCGCTCAGTAGATATGACTTCGGATAAAAATGGTGAAGAGGAAGAAAGTGCTGGTGACGGAAAACCATTTAAAGACTATTTGT
TTGCCACTAAGAATATTTGCTTCGTTATTTTTGTATGCGTGCTGATGATAATGGCGCAGGCTTTTGTTGTAGGTGTAGATTTATGGTTAACATTCTG
GACCTCTCAAGAAGCAATACGGCACGAAAATGGAACAATAGTTGAATCAACATCTCCAACTGTAGAAATAATTCCATTACATGGAGACAACTTCAGT
TATTCCTATAATTATAATTACAGTTACATACCAGAGAATAACACAAACTCATTCAATAACACGAAATCATTCAATATCAATGATATATTTGATACTG
TTAACGTCGATGGCCAGTTGAAAAAAATAATCAAAACGAATTGGGCGTTGTACTTCTACAGTGGTCTAATAGGACTGGCTATCGTTTTTACGCTAAC
TAGGTCACTGCTGTTCTTTAAGGGTTGCATGATGGCTTCAGTTAACTTACACAGCAGTATGTTCCATATGTTACTTAAGGCTCCGATGAGGTTTTTT
GATACGAATCCCAGCGGAAGGATCCTCAATAGATTTTCCAAAGACATGGGTGCTATAGATGAACTTTTACCTATGGGTTTCTTAGATACGATGCAGA
TCATGTTAGCATTGTGTGGAATTCTGGTAAATATTACTGTTTCTAACGCCTACATAGTTATAGCCATTGCCATTCTTGCAGCAGTTTTCCTAAAATT
CCGATCATGGTATATATCCTCAGCTAGAGTGCTTAAGCACTTAGAGGGAATAACTAAGTCTCCTGTTTTTTCTCATATTAATGCCACACTAAATGGA
ATCATCACAATACGAGCTTCAAATGCACAAGATGTCCTAATAGAAGAATTTGACGAAAATCAGGACGCAAATACCTCAGCATGGTACTTGACAATTG
CCTGTATGAATTCCTTTGGATTATGGTTGGACTTCTTGGCTATTATATTCTTAGCAATTGTTACATTCTGTTTTGTGATTTTAAGAAAATTCACCGA
TGTCAATGGCAGTTTAGTAGGTCTAGCTGTTTCCCAATGTTCAGCTCTTACTGGCATGTTACAATTTGGAATGAGACAAACTGCAGAAATTATTAAC
CAGTTAACCAGTGTCGAAAGAGTTATGCAATATACTAAACTTGATACAGAAGGTCCATTTGATACACCAGAAGAAAATCGTCCGAGAGGAGTATGGC
CAAAACGGGGTCAAATTGAATTTAGGAATCTATCTTTAAAATATGTTGAGAACGATCCACCAGTACTCAGAAATCTTAACTTTATGATAACGCCAGG

ACAAAAGATTGGTATAGTTGGCCGAACAGGGGCAGGAAAATCCTCTCTTATCTCAGCTCTCTTCCGTCTGGCTCCACTTGAAGGTGCCATTTATATA
GATGGAGTAAATACCAAAAATCTAGGATTAACTGATTTAAGAAGAAAAGTTTCCATCATTCCCCAGGAACCAGTACTGTTTTCCGCTTCTTTGAGAT
ACAATTTGGATCCCTTTAACGAGTTTGATGATGACAAAATTTGGGATGCTTTAGAACAAGTGGAACTCAGGGACAGCGTAGACTCCTTGGACTTCCA
TGTAGCTGAAGGCGGTGGTAACTTCAGTTTAGGCCAAAGACAATTAGTATGTTTAGCAAGAGCTGTTTTGAAGAACAATAAAGTTTTAGTTTTAGAT
GAAGCTACAGCAAATGTTGATCCTAGGACTGACGCACTAATCCAAGCGACAATTCGAAAAAGATTCAAAGACTGCACGGTTCTAACGATTGCCCATA
GGCTGAACACCATCATGGACTCCGATAAAGTGCTGGTAATGAGTTTCGGAAACATGATCGAATTCGACCATCCTCATAAACTCCTTCAAATTCCTGA
CGGACACTTCCACAAGATGTTACTAGAAACTGGACCAGTGATGTCGGCGCAACTCAAAGATGTGGCCATGAGGGCATACCAACAAGAGTGAAGACAA
GCATAGTATAGGAATCTTTATTTATTTCTTTAAAATACCTTTATTTTCTAATAACACGTCACACTACATAAACGTTTTTTTCGGATAAATGTTAAAT
ATAAAAAAGTATCTTAAAAAGGATTTTACTTAATATGAAGATAACGTTGAAAGCGTTGAAATTTTTCTTTTATAGTAATTATGTTTTATAAATACGG
TTTACTACATATTTTATAGTTTTATTCTATTTTCTATTTCCTCGGTAAATTAAGCTTTAATAACTCTAATTGCGTCATATAAGTACTCTCTCTCTCC
TACAGCGTTACAGCCCCAGTCGGGCCATGTATCCGCCTCCAAGTATCTCTGTCCCTAGCTGCTCTCCTCCAGTTATACACCTTTAAGATCTCTTTAG
CATCGGTTCTTCGTCTATCCACCTCCTCCTTTGTCTTCATACTGGTTGACGTCCCACCATAGTTCCGCTAAGCGTTCTATTAGGTGTTCCGTCATTG
TAGGTTATGTTATGAGGTGACGTGCTTAGCGCAATCTTTGTATTTTTGCATAATTTGCTATAGAGTGTCATATAAGTACTCTAAGATAATTACGTAG
ACTACATTTGGTAATGTAATTGGCATAATAACTAAACTACTTACTACCTTGACATTAAATTTTCGTCTTCTTCAAGTACCGTCTCCGCGACAGAAGT
CGGCAACCATCATACCTATTCGAATTTTAAAGATGGCTTCTCTAAAATGTTTATTTAATGTATGATTCTCTCACCTTAGATCCTTTCCTGCATAATC
AGTTGAAGCACGTACCTCTCCACATGTAATATGTTCGTGATATTACAATTTTCTTATTTTGATTGTATTTAACATTTTTTCATTTATTTTTCGTCAT
TTCGTCGTTGTAATCTGCCGAGCGTTGTAGTTTCTCATCCGTTTGTGTTAGAGTAATCCTTGTCGATCCCGAAATGCACCAGGGAGGTTTTGCTTTC
TAGGTCATCCAATTTTTTAACTTTCTTGTGTGGACCAAATAGCTAATCCCTTTTTTGAAGCTCCTCGATCTGTTTGTATTTTCTTCAAATATTTTTT
TTCTCTTGCCTGCTTTCTCCTCTTCCTGATTTCGTGGATAATTTTCCTGTTTCCCTAATTTTTATCTTTGTTCTTACAAAAATACTGGATTTGATGT
AGGAGGTATAGTTGAAATACACTAAATGAGGGTAATTAAAACRRAACGTTTTCGGATTRGAAAATCCAACATCACTATTTCTGAAAAGC
>Dvv ABC-C D22628
GGCGGAATTGTAATTTTCAGTGAAAGTCACTAACGATACAGGTTTAATCTTCTTATATGTGATACGCCGAGTATAGTGAAATATGTGCATTTTCGTT
TCGTTTTTAGGGTCCATCTATAGGATGTTTAGGATATTCTAATGGAATGTACGGAATATTCATAAAAAATAAATTGTTAGGCACTTGGGATTGTTGA
AGTACCTAAAGTAGAAAATGGATTCGACTAAGAAGCATGTAAAAGTGTCACCGGAAGAGAGTGCTAGTCCGTTTTCAGTACTTTTTTATTGGTGGAT
TCTGCCTTTTTTGAAGTATGGATATCAAAATGATATTGGGATGAAAGATGTATATAATACCACACAAGCAGATCAATCAGGACCATTAGGGGATGAA
TTACAAGAAAATTGGGAGCGAGAAATTTTAAGTTATATTGACAAGCAGAAAAACAAACCTAGTTTAAAAAATGCAATATTTAGGACATTTTGGAAAT
CTTTTATCCTTTCCGGTGCTGCAATATTTGTACAATTTATAATTATTAAGACACTCCAGCCAGTAGTTTTGGCCAAATACATTAATTTCTTTGATAC
AAACAACAAACCCTACCTTGGTTGGATATGGGGATGTGGAGTGGTTTTGTTGGCATTGGCCAATGTGGTTCTATACCATAGCACAATGTTAGCAACC
CAAAGGATTGGTATGAGAATTCGTACTGCGGTATCTTCGTTGACTTACAGAAAGCTTTTAAAATTAAACCACAAATCTCTAGGAGAAACTGCGGCAG
GGCAACTGGTTAATCTTATGTCCAATGATGTTCAAAGATTCGATGTGTGTGCCGCTTCCATACATTTTATTTGGATCATGCCTATATACGCGGTACT
TACATTCTATATCTTGTATATCTATGTGGGAATCATTGCTGCTGTGACTGGAATGGCGTTTATTAGTTTGGAATCAATTCCATTACAAGGAACAATA
TCCAGATGGCAGGGTGTATTGAGGTATAAAATTGCCCTAAGAACTGATAAGAGAATTAAACTGATGAGTGAACTAACATCTGGAATACAAGTAATCA
AAATGTACGCATGGGAGAAACCCTTTGAAAAAATTGTAGAAATGTCAAGAAAATACGAAATAGACGTGATAGCCAAGACCTCATATTTGTATGGAAT
TTTATCAGCTACTAGTGTATTTACTGAACGACTGATCCTTTACGTAACTTTAATACCATTTGTGTTACTTGGACATAGATTGACAGGAGGCATTGCA
TTTTCCCTTGCTAACCTTTTTAATAACATACAACTTGTCATGGCGATTAATTTTCCAAGAGCACTATCATCTTATAATGAAGCAAACGTGTCTATAG
CAAGGTTAGAAAAATTCCTTCTATTAGAAGAAGTTGAAGAAGAAACAGTAGTAAACGAAAAGCACGGTGATCATGTGGGATGTATTAATTTAAACAA
TGTTACTGCCAGTTGGTCCCCTAAATCGATCGTACCAACATTGATTGATATTGATCTACATTTAAGATGTGGAACCTTATGTTGTGTGGTGGGTAAT
GTAGGAAGTGGAAAAAGTAGTCTCTTGCAGTTATTTCTAAGAGAATTACCAATAACGTCAGGTTACATGAATATTGCAGGAAAGATCTCATATGCTT
CTCAAGAGCCATGGTTGTTTGTATCAAACGTGAAAGATAATATTTTATTTGGAAAATCATTTAACAAAAAGAGGTATCAAGATGTTATAAAAGTGTG
TTCGCTAGAAAGAGATTTAAAACAGTTACCGTACGGCGATAAAACATTGGTGGGAGAAAGAGGCAAATCGTTGAGTGGAGGACAAAAAGCAAGAATT
AATCTAGCAAGAGCTGTCTACACAGAAGCTGATATTTATTTGTTTGACGACCCCCTTTCTGCCGTCGATACAAAAGTTGGAAAGCATCTATTTGATG
AATGTATAACGAAATATTTAAATGGAAAAACAAGAATTTTGGTAACACACCAACTTCAATACATGAAGAAATCAGATCTGATCGTCATAATAAATAA
TGGCAAAATAGACAAAGTGGCAAAATTTCATGACCTAACTGAGCACGAGTTGAATCTCCTCCAACAAACCCCAGAAATTGATGATAAAGATAAAGAG
AAAATGCCGACTATATTAGAGACGAAAATTCCAAAAGTTTCATCGACTGCAACACTTCAATCAGCATCTTCATTAGCAAGTTCTATACCAACTGAAG
AACCAGGAGAAACAGGCGAATTAATAGAAAAAGGAGATTTATCGACTTCTCTTTATTGGGAATATTTCCGATCTGGGACTGGCATTGGTTTCCTGCT
ATTTACAGGTTTTATGTTTATTTTTTCACAAATCATAACAAATGCTTCAGATTTGTGGTTATCTCATTGGACTAATGTTGAAGCTAGAAGATATGTC
TCAACTTTAAACTTATCTACTGAATTCATTTCTACAACGATTTCTTCTATAAACAACCATACCGTGACAGATATGCTCAAGTCTGTGGCTTTTACAA
ACACAAATCCTCAGGAAATTATCACAACTGCCGATCCTTCAGTTAGTGCTTCAAAGTTGTCTTCAAATATGTCATCACTAGAGTTGGAAAATTTTAA
AAGTGAAGGATACTATATATGGGTATATTCAGCTTTGATCATAGCTGTAGTAATATTACAGATATGGAGATGCTTTTTGTACTATCAAGTGTGCATG
AGCTCATCTAAAGCATTACACAATAAAATGTTTCATAATATACTCCAAGCACCGATGAGATTTTTTGATACCAATCCGTCAGGTAGAATACTAAATC
GTTTCTCCAAAGATATGGGAGCTGTTGACGAATTACTACCAAGCTGCCAAATTGATGCGATTCAAATTTTGATGGTTTTTGTTGGTATATTGGTAAT
GGTTGTTATAGTAAATCCTTGGATGATCCTTACAACCATAATTATCGGTCCGGTTTTGTTTATTCTTCGAAAAATGTACTTAAAAACCGCTCAATCA
GTGAAGAGACTTGAAGGAATCTCAAAAGCGCCTGTTTTCTCGCATATATCAGCATCACTGTTTGGAATTACCACTATTCGGGCATCCAACGCTGAGA
AAATGGTCACTACAGAATTTGATATTTTGCAAGATCAACACTCCAGTACATGGTTTTTGTTTTTGGTCAGCGGAAGAGTATTTGGATTCTATCTGGA
TGTTATTTGCTGCATATTTTTGGCTATTGTTACAATTCAGTTTTTACTTTTTAGAGACGAGAATACGTTGAGTGGAAACGTAGGTTTAGCCATTTCC
CACAGCTACATCTTAACTGGTATGGTACAAATGGGTATACGCCAGAGCGTGGAAGTAGCCAGTCATATGATAAGCGTCGAACGAATTTTACAGTATA
CCAAATTAGAAAAGGATGGTGTATTTGAATCACTGCCTGCTAAAAAACCTCCCCGAGACTGGCCTAATAAGGGGAAGATTATTTTTAAGAATACTTT
CTTAAGATACGCGCTCAACATGACACCCTCCCTAAGAGATTTAAGTATAGACATTAAGTCTGGAGAAAAGGTTGGCATTGTAGGTAGAACTGGTGCC
GGAAAATCCACCCTAATTGCTTCATTGTTCCGCTTAGCACCGGTCGATGGTGAAATTATTATCGATGATATAGAAACTGGAGGAATAGGACTACATG
ATTTGAGGACAAATATTTCTATTATTCCTCAAGATCCCGTTTTGTTTTCAGCTTCGGTCCGTTACAATTTAGATCCCTTCGAAAAGCATAGCGATGA
AATTCTATGGAAAGCGTTGGAAAATGTGGAACTTAAAGGGGTTGTTACAGATCTAAATCAACCTGTTAGTGAAGGAGGTTCCAATTTTAGCGCAGGA
CAAAGACAACTCATATGTCTTGCAAGAGCCATAGTACGAAATAATAAAATTCTTGTCATGGATGAAGCTACAGCAAACGTGGATCCGCAAACTGATG
CTCTGATCCAAAAAACAATTCGTGAGAGATTTAGAGATTGTACAGTACTAACTGTTGCACATCGATTGAACACTATTATGGATTCTGATCGAGTTCT
TGTCATGGATGCCGGCCAAGCTGTAGAATTCGATCACGCCTACCGACTATTACAGAATTCAGATGGGTATTTTGCTAAGTTTCTGAAAGATGCTGGA
CCACCAATGGCAAATAAGTTAAGAGATATTGCTAAGGAGGATTATTATAAGAAACACACGTCGCCAGAAGTTATTGTCGATGACGATTAAGTCGAAT
TTATTAAAATCTTTTTTTATTGATACAAGC
>Dvv ABC-C D14070
CCTAGAAATAAAGAGATAAATATGCAACTGCTATAACAATAATTTATAAATACTTAATCAATCAAATACTACTATCGTTATTCAATCTTGTTGTTAT
AAGCAGCCTTGGCGATACTCTTTAACTGCTGAGTCATTCCAGCCCCAGTCTCCAAGACCATTCGATGGAAGTGTCCATCTGGTATTTCTAGCAGATT
GTGGGGATGGTCGAATTCTATCATGCTTCCAAAGCTCATTACCAGTACCTTGTCTGAGTCCATTATGGTATTTAACCGGTGAGCGATGGTAAGCACA
GTGCAATCTTTAAACCGTTTTCTGATTGTTGCTTGTATCAATGCATCCGTCCTAGGATCAACATTAGCCGTAGCTTCATCCAATACCAATATTTTGT

TGTTTTTTAGAACTGCCCTGGCTAAACATACTAACTGTCTTTGACCAAGACTAAAGTTACCACCTCCTTCTGCAACTTGGAAGTCTAACGAATCAAT
ACTTTCCCTTAAATCTACCTGTTCCAAAGCTTTCCAAATTTTCTCATCATCGAATTCATTAAATGGATCTAGATTATATCTTAATGTTGCTGAAAAT
AATACTGGTTCCTGAGGTATTATTGAGATTTTCTTCCTTAAATCAGTTAGTCCCAAATTTTTAGTGTCCACTCCATCTACGAAAATAGAGCCTTTAG
TCGGAGCAAGACGAAAGAGAGCTGCAATAAGCGATGATTTTCCCGCACCTGTTCTGCCTACTATGCCAACCTTTTGCCCAGGTTTTATAACGAAGTT
GAGATTGTTTAAAACTGGAGGATCATTCACCACATATATTAAAGATAAGTTTCTGAATTCTAGAAGACCTTCCTTGGGCCAGACACCCCTTAGGTATT
TTTTCTTTTGGTGTTTCAAATGGACCTTCCGTTTCTATTTTTGTATATTGCAGAGTCCGTTCTACACTGGTAAGTTGATTGATAAAGTCTGCTAACT
GTCTCACTCCATACTGTGACAATCCAATTAGTACTGAGCCTTGGGAAACCGCTAGACCTACTAGACTACCATCTACATTGGTAAATTTAGTTAAAAT
GATAAAGCAGCCAATCACACATGCCATGAATATGATGGATACAATATCCAGCCACAATCCAAAGGAGCTCATGCATCCTATTGTTAGGTACCATGCG
GACGTGTGGGCATCTTGATTTTCATCGAATTCATCTTTTAATATATTTTGTGCGTTAGCAGCTCTTATTGTTGTTATTCCATTTAGAGTTGAATTTA
CATAAGAAAATACTGGAGATTTTGTTATTCCTTCTATGTGTTTAAGAACTTTGGCAGTTGCTACATACCAATCTTTAAACTTTAAAAATATGTACCC
CATTGGAACCAGTAGTATAAGTATATAATAATTTGAAACAGTGATGCTAACTAAAATTCCAGCCATTACCACAAAAATTTGACAGCTGTCTAACAAA
ACTTTTGGCAGCACTTCATCTATAGCACCCATGTCTTTAGAAAATCGGTTAAGAATCCTTCCGCTAGGATTTGTATCAAAAAATCTCATCGGAGCTT
TAAGTAACGTGTGGAATATGCTTTTGTGCAAATTAGCAGATGACAACATGCAACATTTGTAGAAAAGAAAAGATCTGATTATCGTTAAAATAATAGC
TAGTGCCATTAGAGCACTGTAAATATATATAGCGTAGTCTGTTTTAATGATCTTTTGTATTTGTCCATGAAGATCAATTTTGTCAAATATGTCATCA
AAACTGATCTGCTTTAAAATTTGTGAGCCATTATTCTCAATATTGTAAACCAACCGATGGGTTGCATCGATTTTAGTTGTCTGATTTATTGGATATA
ATTCTACTTCATTAATATCCGCTGATTGCAATGAAATAACTGTACCGTTTGAATGACGTATTTGCTCTTGAGAAGTCCAGTAAGTTACCCATAGGTC
AACCAAAACATTAATACCCTGAGCAATAACGAGCAGAATGGTTGTAAAAGTTACAAAACAAAGACTTTTGGTAGACAGAATGTACTGTTTAAAAGTT
GTGCCAATATTAGTATTTCCTGACTCTTCTTCAATTACTTCAGGATCATCATACATACTTTCCGAAAAATCAGACAATGCCGATACTGCGGATGATT
TTCTTCTACTACGCGCCGAAGCACGTCCACTTTTCCCTAATTTTCCTTCCGAGTCTGCGTCTTCTTCTTTGTCGTCTGTTTCATCAGCGGCAACTAA
TAGTTTGGTGAAATCTAGATTATATTCCATCAGTTCATTAAAGGTACCCTGGGCTTCAATTTTACCATCATTAACAACAACAATTAAATTTGCTTTC
TTTAAGTACTGAATTTGATGGGTTACTAGGATCCTCGTCTTTCCTCTTAAGTGATTTACAATACAGTTTTCAAATAAATGTTTTCCGACATGAGTGT
CCACAGCCGACAAAGGATCATCCAATAGATAAACATCAGCATTCCTGTAGATCGCTCTTGCCAAATTTATCCTTGCTCTTTGGCCTCCACTAAGAGA
TACACCCTTCTCTCCAACGATGGTTCTATCTTTATGGGGAAATTGTTCGAAATCTCTTTCGAGTGCACAAACTTTGACGACCTTATCGTACCAGGTT
TTTTCGTATGGCTTTCCAAAGAGAATATTTTTTTCGTACTGATGATTGGAACAGCCAAGGTTCTTGTGAACTGTATGATATTTCTCCTCCTATCGAAA
CTTTACCGGATTTAGTAGGGAGTTCTCCAAGAAGAAGCTGTAACAACGAACTTTTTCCAGCTCCAACAGGTCCCACAACAGCACATAACGTACCTGA
TGGTATCTGAAGTGAAATATCTTGCAAGGTACTTGTAGTAGGTGTCCATGAAGCGTTCACTTTGGAAAGTATAACCCCCGATTTTCCTAAATCGTCT
ATTTTAGATTCTTCTTTCTCTTCCATAACAAGAAACTCCTGCACTCTTTTGACTGAAACCAAAGCTTCTGCACCCTGGCTAACGGCCCAAGGCCAAT
ATATAGCCATAGCCATTTGTAAGATGTTAAATGTTTGAGCCAAAGAGAAAACTATATCAGCGGTTATTAAGTTACCTAACAATATATAACAAACAAT
AGTAAAAAACAGAGCCATTCTATCCAAGAACACCATGCAACTGGAATATACCCCTCTAAGATAGGAAGCAGCTGTAATATCTTTGACTTCAGTGCCT
CTAGCCAGTTTAATCACAGCCTCGAATGGTTTCTCCCAAGCATACATCTTTATAACCTGTATTCCTGATACCACTTCGCTCATGACCTTCACCCTTT
TGTCTGTTTCTTGGAGATTCTTTCTCGAAATTTACCTTGCAGTCTTGCCAGGTAACCTTGCACTGGTAAACTAACGACTGCCATAGCAACGACTCC
TGCTACAGATGAAATTCCCACTTGTTGCCACATAATAAACAAGAGCACTGCTACCTGAATTGGCATAACCCATAGAGCATGTAGTGGTACTACGACT
AAATCGAAACGCATAACATCGTTTGAAAGCAAATTTGCTATTTGTCCTGCAGCGGTTTGGCCTAACGCCTTCTGATTCAGCTTCGTGATCTTTCGAT
ATACTAATGATGATATGGCCACCCTTATTTTCATTCCCAGTGCTTGCAGTCCAAACGCAGTGTGGTGATAAATAAATACAATTAAAACGTTTACTCC
AATAAAAGCGGCACTATAAATGTAAATATCTGTGACTACATCATCGGAGTAGTTAATAAACAGTGCAACCAACTGGGCAATTAGAAGTGGTTTCAAA
GACGTTAGTACCACATTTTGAATACACCAAAGAATTCCATACATTATCAGCTCAAGACGAAATGTTTTCCATAGCGCCCACATTAAACTTGGTTTTT
TACTTTTTTCTTTGGCCAGCAAAACCTGTCTATCCCAGTGCTTTTCCAAGGCATCTCCCAGTCTTTTGATTTGTCGGCATCTAAAGCTTCGTACAG
ATCTACCACTTCCAGAGTTTGTTTATTTCCTTTTTTAATTATGGGTACGAGCCATGCAAACAATAACTTCGACAAGAAGTTGGCCCGTCTTTTGGGA
TTCTCTGTCTTAAGTTTAAACACCACATCCATGATTACCAAAGAAATCGAAACAAACTCATTACACTAACTTATCACTATTTATTGGAGACTATGTA
ATACTCGGTTACGACCGT
>Dvv ABC-C T49513_c0_seq6
CTAGTTTAATCACATTCATTCTTCAATCAGCAAATTCTTATCAAACGTATTTTTCGCAACTGTTTTCAATTGCTCACACGTACCAGGACCAGTGTCC
TCCACCATCTTAGAGAAGATACCGTTGGAATTCTGAAGTAAAAGATGTGGGTGGTCAAATTCTGCTGCTGTTCCCTGATCCATAACCAATACTTTGT
CAGAATCCATGATGGTGTTCAGTCTATGAGCTACGGTTAGGACAGTACAATTTGCAAATTTGGTTCGTATGGTCTTCTGAATGAGAGCGTCAGTCTG
AGGGTCGACGTTTGCTGTAGCTTCATCCAACATGAGGATCCTATTATTCCTTATGATAGCACGAGCTAAACAAATCAGTTGTCTTTGCCCGACACTG
TAATTGGAACCTCTATCCATAACTCTATTTTCTAATCTGTTAATAATGTTATTGGGATCTTTCAGCTCCACTTCGTCTAATACTTTGTACAGTTGTT
CGTCGGAGTATTCTTCAAATGGATCTAAGTTGTACCGAAGCGTACCCGAAAATAAAACAGGGTCTTGAGGAATGATGGAAATTTTCGATCGGAGGTC
CCTCAAGTAAATGTCTCTCGTATCGATATCGTCAATTGCAATTTTGCCCTCTACGTTAGCTAGTCTAAATAGAGCAGCGATTAAAGATGATTTGCCG
GCACCCGTTCTACCAACTATTCCAACTTTTTCCGTTGCCTTAACATTTAGAGACAAATCTTTTATAATCGTAGGACCACCTTCGAAGTATTGGAGCT
TAACATTTTGAAAATCAATCTTTCCTGAACTGGGCCAATTCTTGGGAAGTGGCTTTGGATTTTTCGGTTGAGGTTCAGGTGGAAGTTGTTCGTATTC
CAGAACTCTTTCGACCGACATCATCTGGTTAGATATGTCAGCTGTCATTCGCATAGCCATTTGAATCATTCCGGTTAAGGTGCTAGCTTGAGACACT
GCCAGACCCACAGTTCCTCCCGTTAGTCCCATAGCTTTATCGAACGTAAGTATGCTGAATATTAGTACTGCTATAAATATATTTGAACACATACTTA
AACTGATTCCAAAAGCTGCACCGCACGTGAAATTCATGAATCCGGCACTGATATGGTAATTTTGAATTTTATCGAATTCTGCCTGCAGAATTTTCTC
TGTTCTTAGTGCCCTGATTGTTGGTAATCCTTGTAAAGTTGCACTTAGGTGAGTAAACACTGGACTTTTCAGTATTCCTTCCACTCTCTTTAAATTT
TTCGAGGTTTTTAAATAAATATTTCTTAAGAAAAATGCAGCTATTGCCAAAATTAAGGCCATTGCAATTGCGTATGGATTGGCAACTCCTACTAGAA
TGAGAGATCCAAATATTTGCAAGATCATCTGACCTCCATCCAAAAGTACTCTTGGTATACTCTCATCAACACTTCCAATATCCTTAGCAAATCTGTT
GAGCACTCTGCCGCTAGGATTAGTATCGAAAAATCTCATTTAGCATAAATGACACCTTTGAACAAAATTTGGTGAAGATTGCGTGAACAAGACATA
ACCAAAGAATAAAACGAAAATGATCTTCCTAAAGTCATCAAAATTGAAATGGATATAAGTATGCCATAGATGTAAAGACATTCCTCAGTAGTCCAAG
CAGGAAATTGGATCCAAGCTTGTGGTTCGAGTTTACAGTGGTATTTAATGTATTTAATGCTGCATCACTTAAAGTAGTAGAAGGAATAGTTGAAGG
AAAATCAGCTAGGACGACAGTTGAATTTATATTATTGGAAGCAATTGTGGTAGAAATATCCTCTAGTAAGGTAGATGAACTATTCGATTCTAATAAG
ATACTAGCACTGGTTCTTTCATCTTCAATATTTGTCCAGATACTTAAGAAATAATCTACCATGCTACTGGAAACTTGTACTAATACGTAAAGAGATA
AAATGAGAAAGACGTTGAAACAATTGGCTCCAGTCCTAAAGTATCTTAAAAGTAAGGATCCTAAAACTTTGCCCTTCGATGATTCTTCTTGCATATC
TCGCATCTCCAATTTTTTGGTGTTCTCTTCATCTTCACTATCTTCGTCACTATCTTCCATTAAACCTTCTACGATTGTCTTTTCACTTATGATACTT
GTCATCGACGATTTTCTACTCTTTCTTGATTGTCTAGAAACTTGTCTGGCAAATTCAATTTGCTTAACTTTTTCCTCTTCTGTTATTTCTGGTTCTT
TCGTTAAAAGTTGAGCGTATAAATTGTCACTATTAGCAAGATCATTAAATGAACCTTCGCCTTCGATTCTTCCATTGTTTAGAATAATTATGTGACT
AGCATTTTTCAAGTAATGTACTTGATGTGTAACTAAAACTCTTGTCTTATTTTTTAAAAATCCATTAATACATTCATCGTAAAGTACTTTAGACACG
TGTATATCCACAGCTGACAAGGGATCGTCCAACAAATATATATCTTTTTCTCTATACACAGCTCTAGCTAAATTGATTCTCGCTTTTTGGCCACCGC
TCAAAGAAGCTCCTTTATCTCCAACAACAGTCATATCCCGATCTTTGAACTGCTCCAAATCTCTTTCCAGAGCACAAACACGAATGACTTCCTTGTA
TCGTTTTCTATCATAGTCAGCTCCAAACAAAATATTTTGCCGAATGGTAGAGGAAAATATCCAAGGTTCTTGTGAAGCGTAGGAGTAAGTTCCGTTG
GCATTTATGGATCCTTTGTCTATATCCAGCTCACCTAATAACGTCTGAAGAAGTGAGCTTTTACCGGCACCAACTGGTCCAATAACACCTAGTAGCT
TTCCCTTGTTGAGCTCTAAAGTAATATCATTAAGTACACAATCGCTCAATGAGGCATCCCATTTAGCATAGATGTTTTTGAGAGATAACATTGATTT

TGAGTTTTCATTCGGATCTGAGCTTATCTTATTTTGATCGTACTTGTATTCATCATTTTCAAGGAATTTTTGAAGACGGGCGATTGACGTGAACATC
TCAGCGATTTCTGCAATTCCTCTTGTCAAAATACCAGTTAAAGTCATAGATAGCATGTTGAAGTATGAAATAAGTACAAATACCTTCGCTGCCGTTA
TTTCTCCACCAGTTAATGTAATAGTAACAAGCGTAGCAAATAGAGCCAATTTGGTCGTAACCATTCCCAAAGTCATAAAAAGAGCCCTGATGTATGA
ATTTTTTTTAATAATCTTAATTTCTGCCTTTCTTGCCAATCCAATTACTTTTTGGAAAGGAATTTCCCAAGCATACATTTTGATAACTTGAATACCA
AGAATGACTTCGTCCATCAGCCTGATTCTTTCATCTGTTCTCAGAGCAGTCATCTTCCTGAATTTTGTTGCCAGTTTAGCGGATGTTGCTTGTAAAG
GAGACATTATAAGTATGACTAATATTCCAGCTATTCCTGGCAGTTGTGCTCTTTGATATAAAAAGTAAATAACGGGTATTGTGGAAACTGGTCCAAT
CCACAGCTGATGGATCATCATTGATGCGAAGTCAAACCTTGCGACGTCGTTGGAGAGGAGATTTACAACCTTTCCTGCAGGTGTGTCTCCCAGAGCA
GTTCTACTGAGTCGAGTTGCTTTTCGGTATATTAAAGCACACACTGAAGATCGAACTTTCATCCCCGCATGGGCTGCATTAGTCATGTACTGATTCA
TTAAAATAGAAGAGGTCAAAGACATTATCACCATACCAGATCCATAAAAGTATGCTTCCAATTTTGATTTTTTAGATCCAGGAGTGTAATAGTCCAA
AACTCCACCTAACATAGATGGTTGAAACGATCGTATACCTATATCAGATACTGCAGCTATAATTCCTATGAGGAGATATTCTAACCAGAAGGTTCTG
ACGACGGCTTTAAATAGACTAGGTTTTTCTTTTTTGTTTTGTTTTACCAGTTCTAATTCTCTATTCCAATTTTTTTCTAATCTATCTCCTAAATATT
CGCTTCTATCACTTCTTAGTGTATTATATACATCCTCTACGCCTATCTGTTTTTTCATACCAGTTCTAAATGTTGGTATTGTAAACCCAAAAAACAT
AATTGATAAGATCCCAGCAGTTTCCCTGGGATTGGGATTGTTTTGCACCTCCGTCGAATCCATCGCTGATTACAAAATATATTTATTTCACACAAA
ATTCTGTAGGTATAATAAATTTAATGTGTTAATACCTAAAATTCTTCTTTTTCATATTCAGTGGAGGACGAGCTTTGGTATGATATGATAAGCGTGT
AACCACTAACAATACGACGACATAAATATACACACAGTTTATATATAAAACTTATATTTGATTTCATTTTATCTTCCTTTATCTGCATTCTTCGGTA
TGCAATGTATTGAAATAATGAAAATCATTAGACAGAGATAATTGTGATAAGATAATCATTCTCGTAGATGAATAGATCATTTTTTGTAAACTAAAAA
TTTGTATTTTTAAGCATTGCT
>Dvv ABC-C D18126+D19164
GCGTTGTGGAAATTTCTTATTTATCTTGAGGTGGTGAACGGTGCGCAAGGTTTCTTCGTAGAATGGATATAGGCTTTAAGTTAAAACATGAAAATCC
AAAAAAGCAAGCAGGTCTAGTTAAAAAACTATTCTTTGGATGGCTAGTAAAACTAGTTAAACAAGGAACAAAGAAACAACTGGAAATATCAGATTTG
TATGAGCCATTAGATAAGGATAATTCCAAAACTTTAGGCGATTGTTTAGAGAGACACTGGAAGAACGAAATTTTAAAAAGCCAAATTAAGAAGACGA
GCCCAAGTCTTTTAAAAGCCATAGTAAAAGCGTTTTATTTCGAATTTTTATTATATGGAATAGCCTGGTTTGTTTTAAATGTTTTGCTAAGATGTTC
CCAACCGATAATCTTATTTCATTTCATAGCTCTATTTTCTGGAGAAAACAGAGAGGAAAATCAAGGAGATATGTATATTTACGGCGGTCTCCTGATA
TTAGTTTCAGTCCTTTCCATATTTTTTATGCACCATTTACAAATTGGGTTAGCATCAATTGGCATGAGAGTCAGAGTTGCTTGTTCCTCACTGGTGT
ACCGAAAAATTACCAAACTCAGTCACCAAACCCTTGGACAAACTGCAGTGGGACAGGTGGTTAACTTAATGTCCAATGACGTTCATCGTTTTGATTT
AGTGCTGTTACCATTGCACGCATTTTGGGCAATACCTTTTCAGTTTGTCATTTTATCTTATTTTATATGGCAACAAGTCCAAATAGCTTCTTTAGCT
GGTTTGGTTTCTATGGTGATTATAAGCTTGCCACTACAAGGTTATCTGGGGAAATTAATGGGCACCCTTCGAGCCAATATAGCTAAGAAAACAGACA
ATAGAGTCAAACTGATGAGTGAAATCATATCAGGCATACAGGTTATAAAAATGTATGCTTGGGAAAAGCCGTTTGAGAAAGTAATAGAAATAGCAAG
GAAAAATGAGATAAGATGCGTTACACTGACTTCTTATCTGCGAGGAATATTCGCCAGTTGTATGGTCTTTCTGGAGAGAATGTCTTTGTGCTTCACC
TTAATTTGTTACGTTTTGTTGGGTAATAACATAACAGCTGAGAAAGTATTTTCTTTAGCTCAGGCTTTCAACATTCTGCAATTATCCATGGCGATTT
GGTACCCTTTAGCAGTTAGTCATGGAGCGGAGGCGTTAATATCAATTAAACGACTTAAAGCGTTTCTTACGTTAGAAGAAAAGGAGGTTAGTCGTAT
CAAAGGACTGTCAACACCGGGCGTTGTTATGTCAAATGTAAGTAGTTCGTGGTGCGATGCTGGCGAAACGTTGCAAGATATTTCATTAAATATTCCA
CCAGGATTTTTGTGTGTAGTGATTGGACCACTTGGTGCTGGAAAAAGTTCTTTATTACAGTTGCTGTTAGGAGAACTAGCTATCAAAACTGGAACGG
TTCTAATGGGAGGAGAAATATCATATTGTTCCCAAGAACCTTGGCTTTTCCAGTCAACAATACGCAATAACATTCTCTTCGGCAGACCTTTCGACAA
ACAGCTGTATGAAAAAGTCGTCAAAGTATGCGCATTAGAACGAGACTTCCAACAATTTCCAGAGCGGGATGAAACTGTCGTTGGAGAAAGAGGTGTT
TCACTGAGCGGAGGTCAAAGAGCAAGAATTAATCTTGCTAGAGCTATTTATAGGCAAGCTGATGTTTATTTGTTAGATGATCCGTTATCAGCTGTAG
ATACACATGTTGGAAAACATTTGTTTAATCAATGTATTGTGAAGCACTTACGAGGAAAAACAAGAATATTGGTGACACATCAATTACAGTATTTGAA
GAAAGCCAACCTCATAGTTGTTCTTAATGAAGGTAAAATTGAAACAGTAGGAACATTTGAACAGTTATCGAGAAGCAAGTTGGATTTTGCCAAAATA
ATAGTGGACAGTGTAGAACCAGGCGACAAACACGAAGAGACAACCGAATCATCGGATCTTACAAATACTGTGTCAAATCCAAGAAAAGCTAGTGTAA
CATCCACCAAAAGTGATTTATCAGAAAGCTTGGAATATTTTGAAGAAAATAACTTCAGCGACGAAGTAATTGAACATGATCCTCAGCATACGGCGTG
GAAAGAGTATTTTCAAGCAACCAGACGTACTGGACTTTTAATAGTAGTTTTCACCATGTTGGTCCTTGCTCAAACAATTTGTTCTGGAACTGATTTA
TGGGTGGCTTTTTGGACAAATCAAGAAGTAATCAGGCATAGTACTTCAAGACCTTTAGAAAGTGATGTATCTTCAGATGAACCTATTCTGTATGAAG
CTACCCCGTTTGATAGCAGCCTGAATTTAACTCACAACTACGAATACAAAGTAGAAGAACCAGTTGTTACAAGCGATGGCATTTTCGATTATGTTTA
TATAAACAACCACGTGTACCATTTGGTTAAAACAGGCTACGCCATTGGTTTTTATGGATTTTTGATAGTAGGAGTTATAATTTTAACCTTATTCAGA
TCGATGATGTTCGTTAAAGTTTGTATGATAGCTTCCGTTAATATACACTCGAAGATGTTTAGTACGTTATTAAGAGCACCCATGAGGTTTTTTGACA
CCAATGCAAGTGGGAGGATTCTGAATAGATTCTCCAAAGATATGGGATCCATAGATGAAATATTACCCAGAGTACTACTAGAATCAGTGCAGATTTT
TTTGGTTCTGATTGGTATATTGGTGAATGTATCGATTTCAAGTCCATACGCCATTATAGCTATGTTACTTTTGGGAATTTGCTTCCTAAAATTAAGA
TCTTGTTACCTATCCATAGCGATGTCGTTGAAACATATAGAAGGAAAAGTCAAATCGCCGATGTTTTCTCACGTCAACTCTTCATTGCATGGAATGG
CAACCATTAGAGCTTCAAATACAGAAGATATACTTATTAAGGAGTTTGACGAGCACCAAAATGTTCACACGTCGGCCTGGTATTTAACAATAGTATG
CATTTCAGCTTTTGGTCTTTGGATGGATATTGTATGCATCATATTTTTGACTTGTGTCATTTTAACCTTCATCTTTTTGCAAAACTTTTTTAAAGTA
AATAGCAGCTTGGTAGGCTTAGCAATTTCTCAAAGTATGACTCTAACTGGCATGTTACAATATGGCATGAAGCAAACAGCCGAGGTGATCAATCAAC
TAACAGCAGTGGAGAGAGTGTTGCAGTATACTCATATTAACACTGAAGGACCTTTTGAATCACCGATAGAAACTCGTCCTGTCGAACCTTGGCCTAA
ATTTGGACGCGTGGACTATAACCATGTTTATTTAAAGTATTCTGAAGATGATCCTCCAGTATTGAGAAACGTTCAGTTTACAATTTTACCTGGACAA
AAGGTTGGTATTGTTGGACGCACCGGCGCTGGAAAATCGTCTTTGGTTGCAGCTCTGTTCCGTTTGGTAAATTTCCAAGGTACCATATCAATAGACG
GAATAGATACAAAACGTATGGGACTTACATATCTAAGGAAGAAAATATCAATAATTCCACAGGAACCAGTATTGTTCTCTGCTACTCTCAGGAATAA
TCTAGATCCTTTCGAAGAATTCACTGACGAACAAATATGGAAGGCACTGGAACAGGTTGAATTAAAGGACATATCACATTCATTAGATTTGATGATA
TGTGAAGGAGGAAGTAATTTAGTTTGGGCCAGAGACAATTAATATGTTTGGCTAGGGCAGTTTTAAGAAACAATAGAATTTTGGTGTTAGATGAGG
CAACAGCTAATGTGGACCACAGGACTGATTCTCTAATCCAAACAACTATTCGTAACAAATTTAAAGACTATACTGTAATAACTATCGCTCATAGACT
AAATACTATTATGGATTATGACAAAGTTATCGTAATGAGTTATGGAAGAGTGGTAGAATTTGGTCATCCGCATCAATTGTTACAACTACCAGATGGA
CACTTCCACAGAATGGTTCTGGAAACGGGACCTGTGATGTCTTTGCAACTTAAAGATGTGGCCATGGTGGCTTATTCAAACCTCCTGAAACAGGAAG
CAGCTTCTATCCGTTGAG
>Dvv ABC-C C222633_3.0+T22098_c0_seq1*
CTCCTCTGTCCAACACCATCACTTTATCACAGCCCAGCACAGAATGCAATCTGTGCGCTATGGTTAATATTGTGCAGTCAAAAAGTTTTCCTTAAT
AATCTTATGTAAAAGTTTATCAGTTTCGTGATCCATGTTGGCAGTGGCTTCATCCATGACTACAATCTTGTTCTTCCGAAGCAAAGCTCTGGCCAAG
CAGAGAAGCTGCTTTTGGCCAGAACTGAAAGTTAGTAAGTTACTGTTCGTTGGTTGTTCTAAGTTTGTGATTGAATCGTTGATTCCAACTGTTTGTA
AAGCATTCCAGAGGTCTTTATCTTCAAATTCATTCAGAGGATCCAAATTTGATCGAATAGTTCCGGAAAATATCGTGGGATCTTGGGGGATAACAAC
AATATGTTTTCTTAAGAAATCTAGAGACAAAGTTTTAATATCAACTCCATCAATAAGGATCTTGCCTTGAACCTCATAAAGTCTAAAAAGCGTAGAT
ATTATTGAAGATTTTCCTGCACCTGTCCTTCCTACTATACCAATCTTCTCTGTAGGTTTAATTTCAAAATTTAGGTCTTTTAAAACATAATGAGTAT
TGCCTTCATAAGACAGGTACACGTTTTGATATGATATGCTTCCAAGTGAAGGCCAGTTGTTGATTTCAGCTCCTTCTTTATTCTCTGTTTTGATGTT
GGTGTACTCCAGGATCCTTTCCAAAGACGTCATGTAATTTTCTAGCTCAGTAAACTGCCGTACGCACCACGTTATGCTAATGGAAAGGGCGGAGACT
TGAGTGATAGTGAGACCAATGTCACCAGCTGTATTATCTTTGTTAAAGAATAAAAATTTGATCACTACAAAGACGATGAATAATGAACATATCAAGT

CAATAAAAAATCCAAAAGCTCTAGCAGAACACATGGATGTAAAACTGGCTGAAGTATATAAATCTAGGTGTCTATTATATTCTTCTATTAAAATTCC
TTGCACTTTTGAAGCTCTCACCGTGGTGACTCCTTCTAAGGTTGCGTTGAGGTGTCCAATCATAGGACTTCGACTTGCTGCCTCCAGACGTTTTAAG
CTCCTTCCTGCAGGTAAGTACATTTTTCTCATTAGAAAAGCCAACGCAAAAATCAAAGCGATAATGAAAGCACCAGAAGGGTTGACTAATGTAACCA
AGAAAACTATTCCACCGACGTTAAATATCGTCCTGCAGCATTGATCCAAAACAGCAGAAAGATACTCATCGACACTCAGCATATCTTGTGATAACCG
GTTCAATATATTTCCTATCAAGTGCGTATCGAAAAACATCATCACTGAGTTGATGATCTTCTTGATCATCGCCTTGTGTAAATTAACAGAGGCCCTC
CTGCAAAACTCCAACAATGCGTAAGTAGTGAGCAGTTCCATTCCTGTACCCAACATAACAAATACCGAGTACAACCTAAACGTGTCTTGTTCGATGG
ATTTGGCCAATGACAATGAGGGAAATAAGGTATCATTAGAAAGCGGATTGGTGGAAATGTTTTTCTCGATGTTTAGCACAATTTGCTTTTTGTCCAC
CCACTTGGTTAGTAGCCTCTCTGCATAACTTACTGCAAACTGAGTGGATCCTCTCATAGAAATATTAAAAAGAACGAATAAGAATCCTCCTCCGTAT
ATAAAGTACTGTTTGTAAACTCTCCAATCCACAGATCCTTGTTTCTTTACTTCACTGTAGATCTTTTTCGTCGACTTTTCGTCTTTTTCAAGCAATT
TTATATCATCTGTCACTTCGTTATCGAAATCTGCAGAGTTATCTTCTTTAATTTGTTGTTTTTCTTGTTTTTGTAAGTCCACACCATGATCTACAAC
TTCTTTTGCTTCTCTTATAATCTGCTCGCTTGGCTTACCTACGTCTTTTACTGTACCATCCTCCAATATTATCACATAATCAGCTTGTTGTATATGA
CTAGGATTTTGACTAACCAAAACACACGATTTGTTTTTAAAAACTTCTTAATGCATTCGTTGAATATATAAACTTGAACTGATGCATCTAATGCAG
TGAGACAGTCATCCAAAAGATAGATATCACTAGACCTGTAGATAGCTCTCGCTAAATTGATTCTTGCTTGCTGACCTTTGCTTAGGTTGAGTCCTCT
ATCGGTTAATATAGTTTCGTCTCCGTGTTCAAAAAGACTTAGGTCAAATGTTAGGGCACAGGCTGTTATAACTTCATTATACCTTTTGGCTTGATAA
GGTTCTCCAAAAAGTATATTCTGCCTTATCGTCGAAGGAAACAACCACGGATCTTGTGACGCATATGAAATCCTACCCTGTGTGAGTACTGATCCGC
TTTCTAATGGTATATCCTTTAAAATAGTTTTAAATAATGTACTTTTACCACATCCAATAGGACCTGTAATCACAGTTAGACCAGGAGATATTTGAAA
TGAAACACCTTTTAAAATTTCATGGCCTCTGACTTTAACAGTAGCTTCCTTCAATTTAATAAAAGGTGGAGATTTAGGGTCTTCTTGTGTAATTTCT
TCGACAAGTTCCTCAGATTTCAAGGCCTTGACTAATCTCTTGTACGAAGATACAAACTCACTGACGTTTCCCATACCCATAGGCATCACAATAGCTA
AAAAATGTCTGAGATCTCTGAAGTTAGATAGAATATAGAATATTATGGTGGTGTCTGTAGATACGCCTGCCCATATGCAAGCCATGATTAGCATGTA
AAAGCCAAAATTTAAAAACATTACTCCTGTGATTATAATAACTCTTCTGCAGTAGAATCCCAGAACAAGTTTTCTCATTTCCTCTTTTCTTGCTTTA
GTAACTTTTTGGGTAAAAAACAGTTCCCAGGTGTACATTTTTATTATTCTTATCGTAGACAATATTTCCTGGCTAAGTTGTAGCCTTTCGTCAGACT
TTTGGCACGTTTGCAGCCTAAACTTCGTTACCCACGTACTGATATAAACTTGAAAAGGAAAAACTGATATTAAAATTCCAATGCCTACGAATGATAC
GACTCCAACTTTATTGAACAGCAAATAGCAAATGACCAAAACTTGCGTTATAGATACCCAAGCATCGTTTACACTAAAAATTACTGTCACAAATGCT
TGGACGTCCCTGGTGATCAGCGTAACTATATTTCCCAGATTTGTGTCGGTTGTACCAGCTGATGAAAGTTTTAGAGCTTTCCTATATAAAAGGGATG
AAAACGCAGTTCTCATCTCAATGGCAAACATCTGAATCCATATCATATAATTGTGGTTGTAAAATAAGGCCACAAAGTGGATGACCAAAATGCCGCA
TCCATAAAAATAGGCGTCATTCCTCGTTAAAACGGTTTGGCCGGGAGCAAAGTAGGCAACGAAATGGGCGATCACATATGGGCGTATAATACTATTA
ATAGTATTAAAAACGAAATCAATGAAACCTCTAAACAAATATCTTTTGCCGTAACATCTCCACATAACCCGAGAAACCGATGGCTTCTTTGAGCGAT
TTTTTTCACTTAACCATTCCCGTTCTAATAGATCACCACATTTTTTGGCTCGGTGTTGATTTAATACTTCATAGATATCGTCTTCATGGAGGTCGAG
CTTAAAAGCTTTTTTGAATAAACTTCCTACATATGCGAAAGTAAGTAGAGAAAATATATTAGCGGTCTCCCTTGGATTGGTTTTTTTCTGTTCTTTT
CTGCAGTGATCCATTTTGAAGTTGCACAAACAATGTAACAATACGAACAATACGAAAT
>Dvv ABC-C S21020Locus_43217_0+T47823_c0_seq1*
CTTAAATGCGCAGCCATATTTTTACCGGTATTCCTAACTAAGCCATAAAATATAGTCTTTTTATTTTGTAATAACGTATATGGATGATCGAACTCAA
CAGCTTCTCCGGCGTTCATCACTAACACTTTGTTAGAGTCCATAACCGTGTTTAGTCTGTGAGCGATCGTCAGCACTGTGCATTCGGCGAACTTTTC
CCGGATAGTTTTTTGGATAATAGCGTCCGTCATCGGATCAACGTTAGCGGTCGCTTCATCTAGAACTAAGATTTTGTTGTTTCTGATGATAGCTCTC
GCTAAGCAGAGTAGCTGCCGTTGACCTACACCGAAGTTGGATCCACCCTCCGCCATCTTGCCGTATAGATTGTGTCTCATCTCTTCCACTGCATGTT
TCAATTCCACCTCGTCTAAGGCGGCCCAAAGAACCTTATCACTGTAATCATCGAATGGATCCAAATTTTTCCTAAGCGTCCCCGAGAAAAGCACAGG
TTCTTGTGGAATAATAGATATCTTAGACCTCAACACCTTCAGCTCAACAGTATTTATATCGACATTATCTATGATTAAGCTTCCTTCATTTATTGCC
AGTCTGAATAACGCTTGGATCAAAGATGACTTCCCAGCACCCGTTCTTCCAACGATGCCAATTTTTTCTCTCGGTTTTATCTCGAAATTAAGATTTT
TAAGTACACAAGGAGCTTCCAAGGAGTACCTCAAACTCATATTTTTGAACGTTAGGCTTCCTGAACTTGGCCAATCTGCTGGAGGTTGTTTTCTCTC
GTTATCTGATTCTATGGGTAGGTCGGTATATTCTTGTGTCCTTTCTACTGAAGTCATTTGGTTTTCTAGTTCGCTCCATTCTCGCATTACCCATTGG
AACATTCCTGATAATGTAATTGCCTGAGTCAGTGAAAGACCGACATTTCCTCCCAATGAATCACTTTGGGAAAAAACTATACTGATGACGACCAGAC
TAATATAAAAGGCACAGTGCAAGTCCAACCAGAATCCAAAGCCCCTACTGGCAGTTAGGAACATGAAATATGAAGAGGTGTAGGCATCTTGGAAATT
ATCAAACTCTTGGGTGAGAATTTGCTCTGCTCCAAATGCTCTTATTGTTGGAAGTCCTTGTAGGGAGGCTGCTAAATGGGTAAAAATTGGACTTCGA
GTGACAGCTTCGATCCGTTTAATATTTCTACTGGATACAAGAAATATTTGCTTAATTTTGTAAAAAATACAAACAATAACAGCCGTGGGTATCAATA
TCCAAGGATTCAAAAGACTAATAACTATTGTTGTAGCCACTACAGTAAGGCCGACCCATGTTGAGTCTGCAGTCGCTAGCGGTAGCAATTCATCTAA
AATACCAATATCCTTGGAAAATCTGTTTAAAATCCTACCAGATGGATTACTATTAAAGAACAGCATCGGTGAGCTAATAATTTTTCTAAACATTGCG
TTATGGACATTCGTGGAAGCAGTCAAAAAATAGCTAAAAGATGCGAGGGATCTAGCAACAGCCAGAATTACTGTTGCGATCACAAGGAAGCTGTATA
TATGCAACGGACTAATTTGGTCCATCACGTGGGGCCAGTAACTTTCAGTGGATTGGTCCTTATATTTCAATATTGTATCTAAGTTTGAGGAATTTGA
GGTTTCATTTATGAAAACCGCTTGTTTTCTTCTCCACTGTTCTATATTTACCCATAACGTTAGAAATATGTCGGATAAAGAAACCAAAACTTGTGTT
CCAACGAAACTTAGTGCCAATATCAAAGTTTTGAATAAATTTCCTCCCGCCTTTAGATAATTTCCATAAACTCTCCAAGACACTCCACCAGATCCAA
TCGTTTCCTTTTGTTGGTCAATAGATATGTAATCTTCTTCTTCGTAGTTGATCGAACTAATAGATTCAGATTTTGACAGTTTCCTGTCATGCTCCAA
TTGATCCATTTCAGTTGCGGTGGCTAATAATTTCGTAAAGGCATTATCACCATTTTTCAAGTTTTCGAAAGTACCAGAAGCTCTAATTCGACCATCA
TCGATTAGGTAGATAAACTTTGCTCTATACAAATATTGTAACTGATGTGTTACTAATACCACACACTTTTTCCTGAGATAGTTGCAAATGCATCTGT
TAAATATTTGTTTGCCAACTTGTGTATCAACTGCAGATAAGGGATCATCAAGAAGATATATGTCTGCTTCCTTGTAAATAGCTCTGGCTAGATTGAT
TCTAGCTCTTTGTCCACCACTTAAAGACGTACCTCTTTCCCCAATAAGAGTTCTATCACCATGTGGAAACAATTTTAGGTCTTTTTCTAGAGCACAC
ACTCTTAACACTTCATAATATTTATTTTCATTGTAAGTTTCTCCAAAAAGGATATTTTGGCGAACGCTGGAAGCGAATACCCAAGGTTCTTGTGAAG
CAAAGGAAACAGTTCCAGTTACTTTGACTGTGCCTTCTAATGGTTTTAGCTCTTTCAAAATAACTTGTAGAAGGGTTGATTTTCCTGCACCCACTGG
ACCTACTAAGGCTACAAGTTGATTTGATCTGGCTTCAAAGTTGATTTGTTCTATTGTGTTTTCTGGATGTTTTTTGATCCATTTTACAGCTGCATTT
TTTATTTTAATACCTACAGGCCGTACTTCGATGACATCTGGCTTGTTGTTTTGTAAAATCTCGTTGTTGTTGATTTGTTTGTGCATTGGTGTATATT
TTACATCTAGTTCTTCAAATAATAAGAACTTCTGTATCCTGGATATGGATACTAATGTCTCAGCAAACTGGGTCATAGCTTGAGGAAACTGCAAAGT
GATTATGGCTCTCATGTATGTATAATATGACGTTACTGTAAAGGCATATGATGCTGTGAGAACGTTTCCTGATAATATGTATGTTAAAATGCATACA
AGAATAGCCGCTCGTGATAGTGCGATGTTGAAGGTTATAGTTAGAGCCTTTATTATTGAGGTATGTCTTATTTGGTCAATTTCCTTCTTTCTGATGA
ATTCAATCAATTTGGCAAATGGTTTTTCCCATGTGTACATTTTAATAACCTGAATTCCTGAGATAATTTCGTTCATCAGTCTTATTCTTTCGTCAGT
CTTTAAAGCTGTTTTTAATCTAAATTGCGAAGTCAATTTCGCCATATAAACTTGTGCAGGAACAAAGCACAATAAGAACATACTTCCAGACAAACCA
GTTATCCCGGTTTCCTTATAAAGCAAAATAGCACAAGCGATTCCAACACATGGTGATAACCACAAATAATGAATATGTTGTCCACTGTAGTCGAATC
TTCCCACATCATTGGACAACAAATTGACCATCTGACCAACGGTGGTTTGCGACAAAGCTGTTTTGCTTAGTTTTAGGGCTTTTCTGTATATTAGAGC
ACATGCCGCTACTCTTATTTTCATTCCTAAAACCATAACTTTTAACTGATAGTGGTGTATGCTCACTACTTGAGCAAATGTGACAATAATTAAAGTA
ATGGCACAAACCACAATATTATCGCTTTTGGCGCCATCTTGGAAGTACGAGACCAGCTTTGATATCAAGAATGGTTGAGCTATCCGCAATAGCTCGA
ACCAAGCGTTGAACACTGCATATATTGCTAACTCCCATTTGAACACTGAACCTATAGCCAGTAGAAGAGAAGGCCGTTTCTTTTTCAGTTGTTTCTT
CCATGCTTGTTCTAGTTTCGTACCAAGATATAAAGATTTTTGTTTTTCGACAGTGACGTACATATCGTCTTCGGTGACTTCATTTTTGAGTCCTCTA
AAGAGGTAGACTGGTAACCAACAAAAAAATAATTTTGACACGCAGTTTGTTGTTTCGATTGGATGTCGTTTTCTTTTCTTCTGTTCTAAAAATGACA

CTTCCTCCATCACTTGTTACACAATCTTTTAAGTATTTACAATTATTTCCAGTACGAATAGTCACTAAAGTACACAACACAAAGTATTAACTCACTT
TGTTAAAAAATGTCACAGTTAACTTTCACACTAACGATTACTCTTTTATTATATATATGTACGTTTTAACATTTTAAGATAAGTAATTCCACAAAAA
ATACGAAGTTCAAATAACTGCCACTTATCAATTTCGTTTTAATATTGTAAATAGAAGAGTGCGGACGAGATTTCTATTTATTTAATGAAATATTGTA
TTTTGCAAGATGTAGGCTATCACTCTTTCATTGTTCTAGTGTCGTGTATCCCGTCGTTACACGAGGATAGTGGAAAAATTTGTTTATTTTTTATTTA
TCTAAATTTTAGTGTACTTTGTAGTGT
>Dvv ABC-C T51687_c0_seq1
CAGTCAAAACGCGAGGCACGTGGTGCGAAAACAGTCATCTCGTATGTGACCGAATATTGTCTGTATAAAACCCTATATTTTATATTTCCACCTGAAA
CTAACAATCCTAACTACAGTCCACACCCTCTGAAAGTATATTCATACTCTCACATATGTACAAGTTATCGTACAGCAAAGAGTGATCAAATTCTTAC
GAGCCGAATGATAAAATATAAATTTTTTGGGACTGTCTTAACCTTCCATAGCTAACACCCGGTCGGTGAGACCGTTGATCTCTTTAGTACCTTTAAG
GCCAAAAGATATCCGTATCACATATACAGTGATAGAGCATTTACGCGTTTTGAGTTATCAGTGTTTTGGTCTGAGAGACCGGTGTTAGAGTCTGCGT
TACTTATATTGAAATATGTATAAAAAATGTAAGTGTAAAACAGGATTTAGCTGCCATAAATGTAAAAATTCGTAAAAATTCAGTTTCACCAAAGACT
TTTTTAATTTCCTCAGATCCACGTTACTTCCTACGCCTTTAGAAATATTTAGTTCATTTTCTTGATTTTGAAGACCTTCAGAGAATCTATATTTTTC
TGTAAAAATACTGGTAAATGTCAGTTTTGAAAATACATGCCATAATTTTGACATGTCGGTCTCTGAGACCCCTGTTATCTTCTATGTGTAATTAGTT
AAGACCATTTAGATAAGATGCTCATCCACTGTAACTCAAATTGGCGACAAAAATTAATCTTACTTTTGTTACTGTCAGTAACCATGTTATTGGGCCA
GTAGAAAGATTGTGACACTTACCTTTTAAATACATACAACTTAAACTTGACAATGTCAGATGATCGATGACAAATGACTGATGAGAAATGATCGATG
ACAGATGACTGATTACAGGTGACGTTTAGTGATCTGCCACCTGTCATTTTAGCCAATTCAAGTGACCCTAGACTACGATTAATTAAAAAACTAGCAC
GATTAGTCAACATTGGCAACTTAATTGATAAGTTAAGACGTTCTATTGCTTTTTTGGAAAATTTGCTTTTAAGCTAAACCGGCGTCTCTCGACATAC
TGGCAAATATTGAATTCTCATTGGCTAGCAGTCGCGCAGGGCTGTCGAATTCTGCTATCCTGCCTTTATCTAGCACGATGACTCTGTCTGAGTCCAT
GATCGTGTTTAATCTATGTGCGATGGTAAGTACTGTGCATTCTTTGAACTCCGTTCTGATTGTCTTTTGAATGAGATCGTCCGTTTCTAAATCTACT
GCTGCTGTGGCTTCGTCCAAGATGAGAATTTTCGTTTTTCGGAGTAGAGCTCTGGCGAGGCAGATTAGCTGTCTTTGGCCGACTGACAGATTTTCAC
CGCCTTCGGTGACTTCGTGATTAAGGCCAGCTGGTAGTCCTTTCACGAATTCTTTAAGATGCGCATGTTCCAAGGACCTCCATACGTCTTCATCGCT
GTGTTTGTTGAATGGGTCCAAATTGATACGTAGGGTTCCAGAAAACAGTACGGCATCTTGAGGGATAATAGTAAGTCTAGAACGAAGAGTGTGTAAA
CCAAGTTCGTCGATGCGGATTCCATCGATTAGGATGTTTCCTTGCGCAGCTTCGATTATTCTGAATAAAGCCAACGTTAGACTCGATTTTCCTGCTC
CAGTTCTTCCTACAATACCAACCTTCTCTCCACCCTTGATTTCGAAGTCCAGATCGTGCAACACGAGATCCAAACCGGGTCTATATCTGACCGAATA
GTTTTTGAATTGGACTGTACCTTCTTCCGGCCAGGCTGGGGACGTCTGCTTGCTAGGGATTTCCCAAGCGGCTTCCTGCGGGGCTTCTCCGTACTCT
TTAATTCTTTCGACGGCAACGATGTTGGTTTCCACATCAGACGTCATTCTGACCAGCCAGTTGAGTGTTTGTGTAATCTGTAGCGAATAGGTCACAG
AAAGACCGACCAATCCTGGCGCCTGATCTTTCTTGAGAACAGCGAACAGAGCAGCGAAAAATATAATTAAGTTTCCAATCATTTCGAGCCTGACTGC
CAACCATCTATTCGATATAATTCCGGGATAATAACATATTTGATTCATATCCACTTTGTTTTCCGATTCCACAATAAATCTTTCCTGTTGATTGTAC
GCTCTGATAGCATGTGCTCCAGAGACACTTTCTCCAAAGTGCGAATAAATCGGGGAACGTGACACAGATTCTAACCTCTTTACTTGTCTAGAGGTAG
CCACGTAGAACCTTTGCATGAAGTAGTAAATGAAACATAAGGGGATTATCACTAAAATAAACACTGGTGTGGAATAGCTGATTACAGCTAGAGTACC
TACGACCGAATAGAAACAAGTGATCCAGCCTCTCATTGTCATGGGCAACACGTTGTCCAACGTGTCAATATCTTTGCTGAATCTATTCAGAATCCTT
CCCAAAGGAGTCACGTCGAAGAATGTAGTACATGGAGCTTTTAGTACGTTACCGAGTATCAAGTTATGCAAAGCTCTTGCTGAGTTTAAAGATCCCA
CGTACAGCGCCAACGAGGCAAACAATATCGTTACAACTTGTCCAAGACCTAGAGCACCATACACCCCGAGATACATATCTCGTCTTGCAGTGTCAAC
GGTATCGTTGACGACGATCTTTTGATCATCTGACCAAAGCCCCAACCAAACGTTGGAGCCGACACTGAAACCTTGGTACACTAAATTAAAAATCAAG
GTGGCCAACATAAACACCAATCCGATAGACTTCAAGTAATGCTTGTAGACGGCCCAGCTGACGTTGCCCGTTTCTGCTTTCTCGGTTTCTATTAGCT
TCTCGCCTTTCTTTGGACCTTTCTTACCGTCATCTATGGATAGTTTGTGGTTGCTTTTGTCCACGCTGTTGAGGCGTTGGATAGATCCGTTGCCGAT
GTGGTCGGAGCCTGTTTCTGATTGAGATTCTGAGACGCGAGACCTGTGCCTGACTAGTTGTCTTGCTACTTCTTGAGATAATGGTGTGTCAGCGAGT
TGGTCTTTGAGTTCATCGAGTTCTGCTTCGGTTTCGGCTTCTTCTGTAATATGTTGGAGAAGGAATTCGGCGAATGCTCCTTTTCTATCCAGGAGTT
CCTGATATGTTCCCGTCTCGGAAATTCTACCGTCTTTCAGAACGATGATTTTTTCAGTTTGTGGCAAATACGTGATTCCGTGTGTGACGAGGATTTT
GGTTTTGCTTTTTAGCAATCCATGGGGACCGATCACCTTTTCGAATATGTGTTTTCCAACGTGACTATCTACAGCACTCAACGGATCATCCAAAAAG
TATATATCAGCATTCGAGTAGACTGCTCTGGCGACGCTAACACGTTGTTTTTGACCTCCAGATAAATTGATACCTTTTTCGCCGATTTCCGTCTGAT
CACCAGCCGGCAACATCTCCAGATCTGGATTCAGGGCACAAGCTTCCACCACTGCGTCATACAGCGTTTTGTCGTAGGGTTTTCCGAAGAGGATGTT
ATCTCGAAGGGTGGCGTTCTGGATCCAAGCTTGCTGGGAGACGTAAGCGACTTTTCCATATGAGTTCACACGGCCTGAGAGTTTGTCCATTTCTCCT
AGGAACGCTGAGATCAAACTGGATTTTCCTGAACCGACTGTTCCGACGATGGCCGTTAGAGTTTTTTTCTTGATGTTAATGTTGATGTCCTTCAAGA
TTGGTCCTTCGCCCCAGGAGAATGTGCCGTTTTCGATCGTAAGTGGCTCTTCAGAGGGTTCGTGCAAAACATTATCAGGATCAAGTTCTTCAGCATT
CATGAATTTGTTTATTCTTTTAACTGACACCCAAGTTTGTACTAAATTTGACAACATCATGGGCAACATGCTGAGTGGAAATCGTATGATATTGAAT
AGAGAGATAGAAACATAGGCTTTATTTGCGTCTAATACATGTTTTTCATCCACCAATACGTATGTAGCGAAAGTGACTAGAGAGACCAGGAATGGAG
CACAAGACCAGATGAACGACGTACCAGCATTCAAATAAGCAGCTTGTTTCAACACTTTAATTTCTTTATTTCTGATTTTCAATACTTGATCCTCAAA
ACTGTGTTCCCAAGCGTACAGCTTCAAAACTTTAATCCCGTTCAAAATTTCGTTCATCAGCTTGACACGTTCATCCTTGTTCTTCATCTGTTTCACT
TGTAGTTTTTTGATCTTGTTAGCAATGAATCCATTGACGGGAATCAGTATTATCATCACAGCAAGTCCGGCGAGTACAGAAGGTCCTAGTTCCTTCC
AGAGGAAATATAAGGACAGGCAGATCTGTAGTGGTGCAGACCAAATCATGTTGAGGTAACCAATTAGATCCATGAATTTTTGAGCATCAACAGCCAT
CAGATTAACGATTTCTCCTACTGTGCTTTCTTTTCGGGCACTGTTTGATATCTTCAATGCTTTTCTGTAAATCGTTGATACCAAGACAGTACGGATC
CTCATTCCAACCACAAACATCCTGTTGAAGTACTGAGCTAGGAAAAGGGTTTGAATGGTGGCTGTGATGAAAAGTAGAACAGCGTAGAGGTAACCCT
TCCACATTTCTTGGTTGAATCGAGTGTAGGAAATAAGGAAACCTAAAATCTGGGGGCTCACAAAAGTGAGCAGATCTTGTATCAACTTCAACAGCGC
TCCAAAGAAATAGATAGGTCCGAAACATTTGAACAGTGCCGGTAAGATGGAGGCTTGTTTCTTCTTCTTGTTATTAACGAAATCGACGCTGGCCGAA
TCTGATTTAAATTTGGCATGGGCAACTTGTGTGGGGGAGGTGCTTTCGGCTTTTTTTAAAGTTTGACTCCAATATTTTTCAAAAACGGGTACCAACT
CCCTCGAACTATCTTCGACATTGATATCCCACAGGTCTTTGGTTTCTAAAGGTTTTCTGAAACCTTTCCAAGCCAGCGGGTCGAACCATGCGAAGAG
TAGACGTGATAAGAAGCTGGCACTCTCTTCTGGTGACGGATTCTGTGATTGTGGATACTTTCTTTCCAAAGGTGCCTTGTCCGCGAAACAGTTAAGA
AGAAACATCAACACGACCAGAGGAAAATATATCAGGTAGCTGGTGTAAAAATAGTAATAATCGGGCAATATCTCGTTCCTTTGGCTTCTTCTGATTT
CGGTCCTGAACTGAGGAATGGCACAAAGGACGACGAAGAACCAGAACAAGAATAACAGTCCGGACGTTTGTAGACCGTGCTTTCTATTGTAGAAAAT
AAAAACTCCGGTAAGGGCAAATGTTAGTATTTTGATCAACGGCGAGTAAATATCCACATCATTCACTTCGCTGCCTCTGTTACCTGCCGTTTTGAAT
GAGTTTACTATATCACTAATTGTTAAAACTATTAGAGTGGCTGTTATTGCTAATTTTGATAAATTTAACCAGTTCCATGGTATATTGCGTCGTTTGC
TGTTGAATATGTAAAATACTTCTAGGCCGGCGAAGACCCATAGAAAAATACACGGTATCCATACTAGTACAGTTTTCTCGAAACACGGTGTGAACTC
TGGGTCGTTTGTGTTCCATGTTAGGTTTGTGTCCCAAAAAGGAGATCCACAAAATGAACTCAACGCTGCATAAGCGTCTTCTTCAATAGTTGCATTT
TGTTGTGAGGAGTTGCTAGTTAATTTTGGTAATTCGGTTACAGTTGTCGGGAGGGTGGTTAATTCTGTTAAGTTTGCCATATTTTAAAATACGAGGG
TACAGGTAGCTGCAGTAAAATATTCCCTTTTTAATATCATCAATAGAATGCAATTTCAACAAACACCTTGAAAACCATCCACAATGCACATTTACTT
GAACACTTTATGTAACATTAATTTAATAAAAGCGAAACACCGCTTGATTAACTGTAAATGGAACAGCGAGAGATCGTAATAATTTTATTCACAGCAT
TTCAGTAAACTGTAATTTAAAACTGTAAATCTGTAAAACCTCGCGCACGAAGTATTTTGTATTCCTCGAATCACCGAATCGCCGACCTGACCGACCT
TACCTC

>Dvv ABC-C T17573_c0_seq1
GAAAATCTTTATTTTTAGAGATACAAAAATATTTGCTACATCTAAACACTACCACTCAATTTGACTTTTGTTTTTCCTCACCATCACCCAGATGAGA
ATTTCTCATCATTTCTGAAAATAAACTGGTTTTGTTGTCTAACAATACTTTGGGTGAATCAAATTCAATAACGTTACCTTTATCTAAAACTAATATT
TTATCACAATCTAAAATAGCGTCTAATCTATGAGCTATAATAAGTACAGTACAAGAACTAAAGTTTTCTTCTATAATTTTCTGGGCTACTTGTTCAG
TTTCAGGATCCATATTAGCTGTGGCCTCGTCCAAAACCATAATTTTATTTTTCCTTATTATAGCTCTGGCTAAACAAATAAGCTGTCTTTGTCCGGT
ACTAAAGTTGACATCGTCAACATCGACTTCGAGTTTTGGAACTACGCTGTCTAGTTGAACTTTGTGAAGAGTTTTCCAAATTTCCTCGTCAGAAAAT
TCTTTTAGAGGATCTATATTACTTCGTACAGTTCCTGAGAACATAATGGGATCCTGTGGAATAATGGATATATGCTGTCTTAGGAATTTGAGAGAGA
GTTGTTTCAGCTCTACCCCATCGACGAATATCTTACCTTGGTAATTGTATAATCGATATAATGTGGATATAATTGAGGATTTTCCGGCACCAGTCCT
ACCTACAATACCAATCTTTGTTCCTGCCTTAACAGTAAACGAAATATCTTTCAAAACTTTTTCGTGAGAATTCGTGTATGTCATAGACACGTCTTGA
TAAACGATTTCTCCATTAGTTGGCCATTTTATATTATCTGTTCCAGTATTACTTTCACCTTCGATTGTTGTATACTCCATGGCTCTTTCAACGGACG
TCATGCTGTTTTCCAATTCTGACCAAGCCGCTAAGCCCATGTGTACAATTCCACTTAGCATACCCGCTTGATTTAGAGTAAGGCCAACGTCTCCAGA
ATTTGTTCCTGTATCAAAAAATAAAAATCTGACTATCACAGTGGTTGTAAATGCTATTGAAGATATTTCCATCAAAAACGAGAAAGCGGCCTTAGCG
CATATTGATGTGTAGAAGGCACTACTAAACAAATCTTGATGTCTATCGAATTCATTCTTTAAAATATCTTGTGCTCTTATAAGCTCTTATCGTCGTCA
AACCTTCCATGGTTGAGTTTAAATGGCCTACTAGTGGACTGCGAGTAGCTGCTTCAAGTCGTTTTAAACTTCTTGATGTTCTAATATAAATACTTCG
GAGCAATAATGAAAAAACGGCTAATACTACCGCTGGTATAATAAATTTCCAGTTTATTGTTGCTATTAAACCTACGACACCAACCAAATGAAACAAG
GCATCAACCATGTGCGACATCATCATCGACAAATGTTCATCGACAATACTCAAGTCCTGAGAAAATCTGTTCAATATATTTCCAATGAAGAAACTGT
CGAAAAATGCCATCGTGGAGTGCACGATCGATGTAACCATCTTTTTATGGAGGTTCACAGACGCTCTAAATCCTACCCTTAAGATGATGTAACGTTT
TAGCAGTTCAACGAATGAGTAGCCGATGACCAAAATTGTGTACAAATTTAAGGACTTGGTGGCCTTTACTTCCAGCCGTCCAATATTCCGGATGACT
TCAGCTGATATATTGTACGTGCTATTATAGTATATTGCTGTTTGGTTTGTTAAGTCAATGTCTTCGAAGTTTATTGAAGTATTTTTCAGATGTTTCT
CTTTTAAACCTTGAATGGTCGACTTTTCGTTGATCCAATTGGTTAACATTTTCTGTGACGTGCTTTCCGTAAGAGTGGCTCCGGCAAAAGTAAACAA
TAATAAGATTGCGAAAATAAATCCTCCACCCATTTTTATATACTGGATATACAAATCAAAGTCAACGGAACCTTGCTTTTTGTTTTCATGATAAACT
TGTCTTTTCCGTAACTGCGGTTTTTCAAGAAGTTCAGTTTTTTCGTCAATAACTTTACTTTCTTCTGTCTCCATCACCACTTCTTCAATTTCTTCGT
CTTCTAAAGCTTCCAAAAGAATATCTTCTGTGGCGTTTGCTTGATCTCCAATATATTTAATAGCTCCATCGTGTAAAATAACTAACTTATCCGCCGC
ATGAATATGTTTTGCGTTGTGCGTAACGAGCACCACACATTTATCTTTCAGAAGTCCTTGAATACACTGGGTGAAGATTTGTTCTTGGACTCTAGTA
TCTAAAGCAGTTAACGCGTCGTCAATTAAGTAAATATCACTGTCTCTGTAAATAGCTCTTGCTAGGTTGATCCTTGCTTGTTGTCCTTTACTTAGAT
TCATACCTCTGTCAGCAACAATTGTGTCGTCACCCTTTTCTAAAATCTTAAAATCGTATTCCAAAGCGCAAGCCGCCACAACCTGTTGGTACTTCTT
GAAATCATATTTTTCTCCAAACAAAATATTTTGTTTGATCGATGAAGGGAATAGCCAAGGATCTTGTGAGGCATAGCTTTTCCGTCCCCTTGTTCTG
ACCTCCCCTCCGTCAAGAATCGGGTAGTCACGTAACACAACCTTTATCAGAGAGCTCTTGCCACATCCTAACTGACCGGTAAGAACATTGAGACCTT
TTTCCAACTTTAAATCTACATTTTTCAATATATCCTTGTTCCTTAAATTTACGCTGACGTTTCTCAAATCAATTTGTGGATCATCATCAGGTTTGTC
GATAACGTCAGGTAGCTCTTCGAGCTCCAAGATGCGATTGATTCTTTTAAGAGAAGCTGAAAGCTCTCCAATTCTTGTAAATCCCATGGAAAACGCC
ATCGCCATAGTGAATCGTAAAGTTCCAAATATTCTCATTATGTAAAATATATCTTCAGCGCTCATATCCTCATGTATATAAAGGTATATCATCAGCA
TAGCGTAAAAAGCAAACTTTCCAACCAAGCTACTCATTATCATCAACGATATTTTGGCATAGGCTGATTTGAGTAAAATTTTAATTCTTTAAATCT
AGCTTCTCCTATTTTATCGGCAAAAACTTGTTCCCAGGTATACATCTTAATCGTTTTGATGGTTGATAAACTTTCTTGTGTTCTTTGTAATCGTTCG
TCAGTTCTTCTTGACATCTTTAGTCTTAGATTTTTTATTATTTTCGCAACGTAGCCTTGTACTGGTATAACGAGTACGAGCATTAGAATAGCTACTA
CAGAAGCTACTCCTGTCCTTTGGTAAATTAAATAACACATTACAAAAGTTTGTAATACAGATATCCATAAGTCGTTAAACATCCATATAGAGTGTTC
AAACTGTTGAACATCTTTGTTAATGACTGTGACGATATTTCCCAAACTCGTGTCTTCTAAAGCTTTTGGACTTAACTTCAATGCTTTTCGATAAATA
AGAGAAGAAAACGAAATGCGGATTTGAAGGGCTAACTCTGTTAAGTAGATGTGGTAATTTTGAAAGTAGATTGTGTGAAAAGCTTTAACTCCAATCA
TGATTCCAGCAAAATACATAGCATCGTGTATTGTCATTTTGGTTTGTCCAGGTTTGAAGTATCCTACCAAATGTGAAAGAGCGTCAGGTTCCAACGT
ACTGGTCAAAAGCCTTCCCGACATATGAAATAGTCCTAGTAATATATACCTCAATCCATACATATGCCATATTATTTTGAATATAGATACCTTCTTG
TTTGGATCTTTAGTTTTAGATCGACTATTGTACTGGTGTTCTAGTTTGTCTGTACATTTTTTTGATCGACACTTCTTAATAACGTCATATAGGTCGT
CATCCTCGAGATCCTGTTTAAATCCTCTTGTAAATAATTTTCCGGTATATCCAAATGTGAGAAGAGAGAATATATTTGCATTTTCCCTGGGATGGGG
CGGCCGTTGCACTTTTTCACAATGATCCATTTTGATCAACTAACCACATTGAGAAACTGGATTTATAAAACGGATGAGATGTTATTGTTGGA
>Dvv ABC-C T48952_c0_seq2+T48952_c0_seq1*
CCTGAATGTTCAATCATTTTGTTAAAGATACCACCTTGTTTTGTTCTCAAGGCTTGTGGAGTATCAAACTCTATAATTTCACCATGATCAACAACCA
TTATCCTATCACAATCCAACACAGTATTTAATTTGTGAGCAATGGTTGGCACTGTACATTGTGAAAAATGAGTTTTAATTGTCCTTTGTAAGAGACT
GCAAGTATCTGGATCCATGTTGGACGTAGCTTCATCCAAAACAATAATTTTATTTTTGCTGACCAACGCTCTGGCCAAGCATATCAATTGGCGTTGT
CCAGAACTATAACGTGAACCATTCTCAACAATGACATCATCTAAACAGGTAAATAAATTTTTTAAATGGACTTTTTCTATTGCGGACCAAATTTCTG
CATCCGTATGGGTTCCTGATGGATCCATATTTGCTCTTATTGAACCTGAAAATAATATCGGATCTTGTGGAATGATACCTATGCTGGATCGTAAATA
TTTCAAACATAAATTCTTAATATCTTCATTGTCTATTAAAATACGACCCTCAATATCATAAAGCCTGAACAATACTGAAATAATAGATGATTTTCCA
GCCCCTGTACGTCCAACAATTCCAATTTTTCTTTGCTTTTAATGGTAAAACTAATATTTTTAATACTCTTTGTTCCGATGAGTTGTAAACTAAAG
AGACGTTTTCGTACTTTATTTCTCCAAACGATGGCCAATTTTTCAACGTACGTCCTTGCAATTTATCTTCAGATTCAATGTCGGTATATTCTAAAAC
TCTTTCTATACCTGTCATGTTATTTTCTATTTCGGTATATTGCCTAATAACATATTGTAGAAAACCCGTAAGTAACATAGCCTGAGTTATTGCTAAT
CCAACGTCACCAGCTTGAGCATCTGTCTTAAAACAAACAAACTTTATTACTATGGCAGCAAAAAAATGTAGTTGATACCATATCCATGATAAAACCGT
AAAACCTAATGGCGCACGTCATCATAAAAGAAGTGGATGTATAATGATCTAGGTGTTGATCAAATTCTGAAATCAGTAATGGCTGCTTCTCAAATGC
TCGAACGGCAGTTATTCCTTCAAGCGTTGCATTTAGATATCCAATCATTGGACTTCTAGTTGCAGCATCCAGTCTTTTTATGCCTCTCCCCGTTGGT
AGATAAAACTTCTGTATAAAATATAACTTGACAAATAAAAATATTGAAGGAATGATAAAGAGTTTGTTGACGGAAATAATTAAGTACAATATTCCCA
GTAATCCTAGAGTTTGCCTAAATATCTCGTAAGTATTCAAAGGAATAACTTCGTCCATAGTCATCAGATCTTTCGAAAACCTGTTCACAACATTTCC
CACAAAGTGTTTGTCGAAGAAGCTCATAAAAGCGTTTATGACGGATTTCAACATTTTGCGATGTAGATTTCGAGACGCTCTTATAGCAAAGAAGAAG
TTCATGTATATTCTAGTAAAGGTGAGTACTACCATCACTACTGTCAAGAATGTATACATCACAAGGTAGTTGTCTCTTTTATTTATGATGTTGATAT
ACTCTTCAGTGTCAGTTTGATTGGAGAGCGTTAAATTCGTAATATTAGGTTCCAGATTAACCCACTTGCTTACAAGTTTTTCTGAGTACGACAGAGC
TGCTTGGGCAGCAATAAAAAATAATTGCTAGAATAACTAACATGATAATGCCACCAGCGTATTTGTAATATCTCAAATAAACTTTAAGGCTTACAGTA
CCACTCTTCTTTTCCTCATCGTACAAATTCTTATTCTCATTTTTATCTCCCCCAATTAGCTGAGTTCTCTCATTCGCTTCGTCCTCAAGTATATCTT
CGTCTGTAACATCAATATCGTCGTCAAAGTAGTTCATTTCAACATCATCGATAAAATAAGTTATTCTTTTGTCCAAAGTTTCTTTTTGTTTTTCCAA
AGACAGCGTTCGACCATCTTCAACAACTAGCGTATTATTTCCGTACAACAATTTGATGTGGTTGATGTTGTTGGACACCATGATAACCACTTTGTCT
TTTAAGAAATCCATAATACAGTTTCTAAATATATAATAATTTACTTTTGAGTCTAAACCAGCTAGACACTCATCTATTAAGTATATATCACTAATAC
GATATATCGCTCTTGCCAGATTGATTCTGGATTGCTGCCCTTTGCTGAAATTGACACCTCTATCTCCTACTATAGTTTTGTCGCCTTTTGGTAGTTT
CTTTAAATCTATATTAAGAGCGCACAATTTAAGTACTTCATTGTACCGTTTTTCGTTATAAGGTTCTCCAAAGAGTATGTTCTGTCGTAAAGTGGAT
GGGAAAAGCCAGGGTTCTTCTGACGCGTAGGAGATGGTTCCATCAATTGCCATTTGCCCACTGCTTACTGGATATTCACCCAGGAGAACTTTGATAA
TTGAGCTTTTACCACTACCAATGTTGCCAGTGATGAGATTGACTCCTTTTTCGGCCGAAAAAGAGATGTCATTAAGAATAGTTTTTTCTCCGACTTT
GGCACTCACATGCCTCATGTAGACTCTCGGTGAAATAGTCTGTGATCGTCTGTCTACGACTTCCTCCAGTTTCAAAAAGTGTGACAATCTTTTAAGG
GAAGCTCTTAAGTCCGAACATTGAGCTATACCAATTGGGATCGATACTGTTATGAACGATTTGAGACTTTGGTAGCAAGTTTGAATAAAATAGACTG

TCTCTGCGTCTGTAAAATGACCAGACCATATATAGGTCATGATCATAAAGAAAAAGGAGAGGCTCGTTGCTGTGGATCCAACAATTAGCACCAAAGA
TTTTAGATAGAAGACCGGAGATAAGTTTTGCAGTTCTAACTTTCTTAATTCGTTAAGTTTTGTCGTAAAGAAGTTTTCCCAAGTGTACATTTTAATA
GTTTTTATATTTCTTATAGTTTCTGTTGTCAGTTGCAGTCTTTCGTCAGTCCTCTTAGCAGATTGGACCCTTTTTGAGCTTACTACTTTTCCTACAA
ACAATTGTAATGGGATAGTCAACATGTAAAAGCCAATTCCTCCAAACACAGACCATTGTATTCTATTAAATATAACTCCTGTGATGATTATTGTTTG
GATGACACCGATCCACATATCGTTTACAAATATCAATGCCATTTCGAATGCAAACACATCCTTCGTTATCAACGTTACAATTTTTCCAACAGTTACG
TGATTCCAAGCATTGGGCCCCAACTTCACTGCTTTCCTGTAAACTAGTGCAGCTACAGAAGTCCTCACTCTTATACCTACTTCTGTAGTAAATTGTT
GATAATTGTGGTTATAAATAACACTAAATACGTTTAAACCTACAACTAGGGAAGCATATATATACAAATCTTTTGTTGTCACATCTGTTTGGTTGGG
AGCGTAGTACGCTACCACTTTACTAAGTGCTCTTGGTTGAATAAACACTAAAACTGTTTTCACAACCAATTGTATTAATCCCAGAATTAAATACTGC
ATACCAAAGCAGCTGATTAAGCACCTTACCAAGGAGGCATCTTCTCCTTTCTTTTTCCGTTGCCTCATCCAAGCATCTTCTAGATTATTTCCCAGCT
GTTCTGATGCCAGCCATTTTGGGATTTCATATATATCACTATCATCAAGTCCTTTTTTGCGTGCTTTTTTAAACAGGGGGTAGTTATAGAAAAATGT
TAGAAGACGAAAAATGGAAGCTTTGGTTCGGGGATTTGTTTCTTTGGCATTTCTTTTTGACGAGTCCATATTGGATAAGGACAGTCAAAAACCATAA
ACGCCTATCTCGAAAAATTATTAGTAGTTGTAAATGTCTCGAGTTGTTTCCTTATCTGTCTTATCGGCGATAATCCACTCATGATATTACCGATATA
TTTGTGTTGATTTGTCTAATAGATATACCGGGGCAGGTGTAGGTATTCATACCTGTTAGGTGGATCTTTCTGTAT
>Dvv ABC-C T44708_c0_seq1
CGCCCAACGTGGGGCGCCATTTGACTATCACAGCTAACTAAAATATTATTTTACCATAGCATTTCAATTTTCTGTTGTAGTTTACCATGACAAAAAA
AATGAGGGATAATACTGGTATATCGTTGTCGGTAGAATATAACAATACCTTCTATAGCCGTAAAATATTACTTAAATTAATACACGTTAACTGGACA
TATCACCCTGTGCTACCATTTTACTAAAGTGTCCTATTTTATTTTCCAGCAACTCCTTTGGAGTACCAAACTCTTTTATCTGTCCTCGATCCAGTAC
CAACACCTTATCACATTCTAATATAGACTGCAATCTATGAGCGATTGTGAGGACAGTACTATCAGAAAAGTTATGTTTAATAGTTTCTTGTATTAGA
TTTTCTGTTTCATGATCCATAGTAGATGTAGCTTCGTCTAAAATAACTATCTTCGATTTAATTAATATAGCTCGTGCTAAGCAAACCATTTGCTTTT
GACCAAGGCTGAAATTAGAAGCATAACTGCTAACTTTGATATCTAAATTTGTTATAGAAGTCTTTAGATTGGCTTTTTCTAAAGCTTCCCAAAGATC
TTTTTCTTGAAACTCACCAAATGGATCTAAATTAGATCGGATAGTACCAGAAAAGAGAATAGGATCTTGCGGGATGATTGCTAGATGTTTTCGTAAA
TGTTTTAAGGGTAGTAGCTTTATATCAACTCCATCGATGGTTATTTTACCGTCTACTTCATAGAGTCTAAATATTGTGGATATTATGGAAGACTTCC
CAGCCCCTGTTCGGCCTACTATTCCAATTTTTTCTTTCGCCTGAACTTGAAAACTCAAATTGTTCAAAATTCTTTGAGTAGAATTATAGGATAAACT
AACATTTTGGAAACTTATGGTTCCGTCTTTAGGCCAATGTTTTGGTACGCTGCCTTGTGTTGTTTCAGACTTTAATCCGGTATATTCCAGGGCTCGT
TCAGTAGCTGTCATTAAAGTTTCCAAATCAGCCCAACATCTCACTGCTCCTTGTATTAGAGCACCTAAATATATCACTTGGGTCAAGGCTAAACCAA
TGCTTCCAGCTGTGGCATCTGTTTCAAAAAGAACGAAACTACAAATGACCAGCGTCACCATAACACCAGCCAACATGCTCATATAGAATCCCATTCC
GAACTTGAAACACAGTAGACTGTAGTGGGCCGAAGTGAAAACGTCTTGATGCCTCTCGTACTCATCGATGATCATATTTTCTGCTTTATAGGCTCTA
ACAGTGGTAACACCTTCCAGAGCAGCATTTAAATGACCTATCATTGGACTTCTTGTTGAAGCTTCAAGTCTTCGCAAATTTCTGGTTATTGGAAGAT
ATAGCTTTAAAAATAATAACATTACTCCAAACGTTACTGATATATAAAAGAAAAAATATGGATTTACTGTCATAAGTAAAGTAGCAGCTGCACCAAC
ATTGATTGCTACCTCTAGACAATTTAACAACTGAAAGGGTAAATTCTCATCAACATTAATCACATCCTGGGAAAACCTGTTTAATATATTACCAATA
AAGTGTGTATCGAAGAAAGCCATCACTGAATGTAGTACATTCTTTATCATCGCTTTATGTATATTTATCGATGCCCTTTTGCAAAAGTCTAGAAAAG
CGAAGGTTTTGAGTAGTTCTAGAACTAAGGAGAGTATTACCATCACGGTATACATTTGGAATGTAGAAGCAGCTTGAACTTCAGCTTGTGTTAGGTT
TGTAGATAGTTGGGTTATATTTTGGCCAATAATGTTAGATATGTTGGCTTTTATGTTAAGAACTTTCTGTTTATCATCGCTCCACACAGATATCAAC
TTTTCGGAGTAACTTTCAGCAGCTTGCTTTGAACCAATTAGCAACAAAGCAAATGTAAACAGGGCAGCTCCTCCCCCATAAGATATATATTTAATAT
AGACAGATTTATCTACGGCACCTTTCTTCATTTCTTCTTCATATATATTTTTATAGTTACTTTGCTCGGTTTCCAGAAATCTGTCAACCTCCATAGA
CTCCATTAATCCATTTTCTTTGATTAAAAATTCTTTCGTTTTATATATGGTATCTGCGAAATTATTTACTTCCTTTTCAGTGTTTTCCTGAGGGTTA
TAGCTCCCTATAATGTGTCCTTTATCCAGTACATAAACAGCATCTGCTTGTGTTAGCGTAGATTTAGATTGGGTAACCAAAATGCATATCTTTTCCT
TTAAAAACTCTTTAATGCACTTTTTGAAGATGAACTGATGAACGTGTGCATCTAAAGCTGAAAGACAGTCATCCAAGAGGTAAATGTCAGCTTGTCT
GTATATTGCTCTAGACAGATTGACTCTGGCTTGTTGGCCTTTGCTTAAGTTTTGGCCATTATCTGTCAGGATCGTTTCGTCTCCCTTTTCAAATAAG
TTTAAGTCATATTCTAGTGCGCAAACTTTTAATACTTCTTGGTACCTCTTTTCATTGTACTTCTCACCAAATAAAATATTTTGCTTGATAGTACTTG
GAAAACACCATGGATGTTGGGAGGCATATGATACAGTTCCGCTACAAACAACTTCACCACTACTTAAAGGGTACTCTTTTAATATTGCTTTTAGTAG
GGTACTTTTACTAGATCCTACTTTACCGGTAACCAACGTTAGTCCAGAATCAACCTTAAAACTGATTTTTGACAATAGTTCTTCTTTTCCTACACAA
ACTGTAATGTTTTTTAACTCTACATATGGCTTAACATTGTTTTGTTTACTTTGATGATCTTGATCTAGTTCTTCGGCTATCAGTACTCTATTGACTC
TTTTAATTGTGACAGACAATTCTGCAAATCTTCCTACAGCTTCTGGAAGAAGATATCCAAAAAATATTACGATGTTGTCATAATGGGCAGTTATGAA
GAAGATTATAGAGGCATCTGGCGGTTGTTCCATCCAAATATAACCCATAATCATTAAGTAGAAGCCGATTTTGGTTGAGATAGCACCCAATAGCATT
AGTATCATACGCAGATAGAAAACTATAACCATTTTTGCTACTTCTCTCGTTCTAGCCAAGTTAATTTTTTGGTAAAAGTAATCTTCCCAAGTATACA
TTTTAATGATTTTTGTTGCAGATAGTACTTCCTGACTAACTTGTATTCTTTCGTCGGTCATCTCACTTGTTTCGTTGCGCATTTTAGTCACCATTTT
ACCAAGAAAAAATTCTATGGGAATAATGCAAAAGTGCATACCCACTGCAAAAAACGTGAGTACACCAAGGCGAGCATATAGTAAATAACATGTGACA
CTAATAACTACAGTAAATACCCAGATATCATTAAAAGCAAACATAGATCTTCTGATTTGCTGGACATCTTTGGTTATCAACGTCACAATATTTCCCA
AACTTGCTTTGGTCATGGCCGCAGGGCTGAGCTTGAGGACTTTTCTGAACAAAAGCGAACACAAAGCTGTTCTGACTTGTATACCCATTATAAACAG
GAAGAGATCGTAGTTGAAGAAGTATATTTTCTGCAAAAAGTTGGAGAATAGTAAGATGGCAGCGTTGAAATAAGCGTCGGATTTGGATAGTTTTGTT
TGTCCTGGTTGAAAGTAGGAAACCAAGTTTGTGATGGCGTTGGGTTCAAAAATGCTTCTGAATAATTCACAAAACAATTGAGTTATTCCCCAAAGAA
AATACTTCCATCCAAAAACTTTACACAAGGTTAGAAAGAGAGAAGGATATTTTCCATCTTTTTTTCCCTTTTTCCATTTCTTTTCAATTTGTTCGGC
ACAATGTTTGGATGCACAATTTTTTGAAACTTCATAGAGGTCCTCTTCTTCTAAATCTTTCTGCACGCCTCGTTTTAAGAGATCTGTTATGTAGCTA
AATGTAACCAAGGACCAAATATTGGCCGAATCTCTAGGATTTAATCTTCTTTTCTTAATTTTTGAGGAATCCATAATTTCTAATTTTTATTGAATAA
TAAAAAATAAATTGTTGAAGATTAGTCGAAAATTCTTTTGACTTAGTATAGTTCCAACAATCCCACCATTATGCCAACTAATCATCCTAAGACTTCA
CCATCACCAATTACTGTCACATTATGCAGTTAAAAAATAGCAATTTGCCAAAATTTCCAGAATTGTCCATACTAATAACGATTCGGTTACTTGGTCT
CACTCCTACATTCAATTAAATAATGCTGTGTTAATAGCGGAAGTAATTTTCGAAATTATAAGGTTACAGATGCGGATTAAGATTTCGTGTTCTGGGT
TTAGGTTCTCTATTAGACGATGTCGCATATTTAAATTTTAAGTAACTAATATAATATATTTACCCGTATAAAGACAATACAAGC
>Dvv ABC-C T41801_c0_seq1
ATTTAATTATATTCACGTAGATCTTAAGAAAATTACATATGAACTGGTCCTGCTTTCACTTGTTACCACCTTTTGTATTTGATACAGAACTGAATGC
TTAGGAATGATGTTGTGTAAAGTGCTTATACGCCATGTCTTCCAAATGGGCTGCGGTCGCATCTCCAGTTTGCTGCACTAGATCATAGAATTTTCCA
CGCTTAAGTATAAGTCGATATGGATGATCACATTCCACTAAATATCCGTTGTCCATAAGTAATACTTTATCGGTCTGCATGATTGTTTGTAGCCGAT
GGGCTATGGTAATAAGGGTAAAGTTGGCAAACCTCATCTTGAGAGTGCGCTGAATTAGTTCATCGGTTCTTTGATCTACGTTAGCTGTAGCTTCATC
TAATATTGCTAGCTCATTTTTTCTTAGTATTGCTCTGGCCAAACATATTAGTTGTTTCTCTCCTGTACTAAAATTGTTTCCACCTGACGTCACCACT
GAGTCCAAGGATGTTATGCAGTCCTTTAATTCTACCTGTTTAATGACTGTCCATAAATTATCATCTGTGAACTCCCCAAACGGATCTAAATTATACC
TCACAGTATTGGTAAATAAAACAGGATCCTGTGGTATTATTGATACCCGTTTTCTTAATATTTCCAATGGCATCGTCTTTGTATCCACTCCATTGAT
TTGTACTGAGCCACTAAATTCGGAAAGTCGAAGTAGAACGTTAATCAAAGAAGATTTGCCAGCTCCTGTCCTTCCAACAACTCCCACCTTCTCTCCT
GAAGCTATACTAACATTTATGCCGTTTAAAACTAAAGATCCTTTGTTGTAAAACATCGACACATTTTTAAATACAACCTCAGCTCCGGAATAACCCG
GTTGAAAGTTATATACACCATTTTTTTTGGTTTCAGAAAATTCTTCAGGTTCAATTTTACTAAATTCTATCATTCTTTCTATACTCATGAGATACTG
TACGAAGTCCAAGGACTTTTGTGTAGCAAACTGAAGGGCACCCATTAGCTGAAATACTTGTAAAATGGCCAGACCTGCGACACCTGGAGTAATTTTG

TGATCATATTCGTTGAGAAGAATAAGTGCAAATATCACTCCAAATAAAAAGAATGCACAAACTGTATCCACAATTACAGAAAATGCAGCGCCACATA
GAACATTTAAACGAAGAACACTTGTGTGGTTATTCAATAATATAATAAATTCTTTTTGTAGTTTTTCTTGATTTTTGGTTGCTCTTATTGTAGTTAT
TCCACTTAAAGTGTTGTTAATGTAAGTATATATTGGAGATTTGGCTGTTATTTCCAAGTGTTTTAATTCACTTCCAATATATTGATACCATTTTAAA
AGCTTGGCATAAATGAAAACAATCAAAACGAAAAAAGACATATAATAATCCGAAACTATTACTAGAATAATTGATCCTATACCCATAAAATATACCT
GAGAAATATCTTTAAGATTGTCTGCAAATCTCACATCGACTGCTGTAGTATCTGAAGAAAATCGATTCAAAATTCTCCCACTTGGATGAGTATAAAA
GAAGCTCAAAGGAGCGCTGATAACTGCTTTGTACATGTTTTTGTACATATTTTTGGATGCCGTATTAAAAAACTTTGTTAATAAAATATTAGACAAT
ATTAACATCAAGAATGCGCTCGATACGATGGAACAATATATAGTAAAATTAGTATTGAATTCTGTAGTCCACTGAATATCTCTCATATGATGAACAT
TGGAGTTTTCTGTTACATTTTTTGATTCAGCGTGTGTGGAATAAGATTGTTTTGTCCAATATAACAACCAATAATCAGTCAATCCTAAGAGAAGTTG
TGCACTGAATACAGTTAAAAATAGGGGAATAATGCAAGAACCAGACTTAATGTATTCAAACAGTGCTCTCCTTGGAGTTATTTTTGTTCTATTTTCA
TTACAAACATAGTTCAATTTCTGGTCTGATTCATCTTTAAGGTTCAATTTTGACGTCACAGTAGTTTCCTTTGTTTCATCGGCTTTATTTAATTTTT
GTATCCATTCAAAAATTTCAGGTTTTTTAAATTTTATCTCTGACAAACTATCAAATGATTCTAAATTGCCCTTGTCTACAATAATAATTTGATCAGC
TTCTTGTAAATACTGCAAGTGGTGTGTTACTAATAATCTAGTTTTGTCTCTTAAATAACTTTTTACACATTCATCGAAAATTCTTTTTCCTACTTCA
GAATCCACAGCAGACAACGGGTCATCCAAAGGTAAATATCAGCATCTCTATACACCGCTCTTGCTAGATTTATCCGAGCTTTTTGTCCCCCACTCA
ATAGAACACCTTGTTCACCTACACAAGACAGGTCCCCATTTCTTAGCTGGTCAAAATCAGGCAATAAACTACAAGCAGAGACTACTTGTTGGTACTT
ATTATAATCATAATCATTGCCAAAGAGAATGTTGTTTTTAAGTGAAGAGGTAAATAACCAAGATTCTTGTCCACTGTAAGACACTGAACCATTCATC
TGTATGTTACCAGAAGATAGCTTAAGCTCTCCTAGAAGTAACTGCAATAGAGTGGATTTTCCAGAACCTACAGGTCCAATGATTACGGTAAGCGATC
TTGATGGTATGTTTATGTTATTGGATGATATAACAAATGATGACTTTGTAGACCAAGATGCTTTGATTTGATCCAACTGAATACCAGACACATCTGC
ATTAATTTTGCTATCGATATCAACTTCATTCATAAGTAACAACTCTGAAATTCTAACAAACACATTAGTAATTTCAAACTGTAGTAGCATTGCTTGG
GGCATAAATGCCAAGACTCCTTTTTTCAATATCACCAGAAATTGAATGCAAGAGAATATTATTTCGGGGAATAGGATTCCATTGACATACAACATTG
TTAGGACTGAAATTGCACTGGCTAGTTTTAAAACAGCGAAGTGGTAATAGCCTATTTGAAACAATAAATATTTAGTTATATTTCTAAGTTCTTCTTT
TCTATAATCGTTGATGATTTTGGTGAATGAGAGTTCCCATACATACATTTTGATAGCTTTGATTCCAGATATTATTTCTTTCATCATGTTTACCCTG
CTATCTGCAATTTTTGATGCTTTAAGACGATAAACTGAAGAGATTTTACTTCCTAAAACCATGATTGGTAACAGGATAGTAAAGAATAAGGCACCAC
TAAATGCCGCTGTACCCAGTAAATACCACATACTCGCATATAATGTTAGGATTTGTAGTGGTGTCAATATCAGATATCCCAAATATGTTGTGTACTG
AACAATCCTATCTAAATCATCAGCTAAAATATTTATGATCTTTCCTACGGTTGCTTCTTCCATTGACTGAAGATGCAAAACCAAGAGCTTGTTATAT
ATTAAAGAAGACAAACCTGCCTTCATCCTAATTGCAATTTTTCGTAGCTGGTAGTGGAAAATATAATCTGAATAATAGTTAATTACGACAAGGACAA
CAACCATGCCTCCTAACATTATTCTTTCTCTAAATCGGTACTCTTCTGTAAAGGTTCGGTTCAAAATTGTTAATAATACACATGTAATTACCCTCAA
AATAACAAAGTTGACTCCCGTGAATGTGTAAATTATTGCTAACTGTTTCCGAAAAATTCCCATTACAACTTTCTTTAAATTCGGCTTTTCTTTGTTT
TCTTTTGATCTTTTCACTTCTGCAGCCCATTTTTCCTCCAACTGCTTTAAAAGAATTTTGGCATCATCCTCAGCATTAACAGCAAAATAAGAAAGTT
TCAATGTACCATTAAGGATTTTTCTAAAAAAATGGAATAGTCCATAAAAAGAATATACGTGATAATAAACATGCTTCATACTCGGGGCTAATTTCTTT
TGTTTCAGTTGCTACGTCCATTACTTTAAAGTTATTTACACACTTTAAAACTGTGGTAATGTTTTGATAATAAAGCTAATTACTTATTGTAAATTAC
TAAAAAAGTATAATGGT
>Dvv ABC-D D11014
GTGATGGGATCGCCTATACTTGTAACATTACCAATGGCCGGCCTTAAACTCGGCAAACCTGCCGAACCACGTTGACTAGAGTGACGTACGGTGTTGC
CAAGTTCTTGAACAAATACATAATAATTACCAAAATTCACTAGATTGCCGTATAGATGTTGCCAAATCCTTGACTGAACACGTAAGAATTTTAAAGT
TGACATTAAAAATAGAAAAAAAAAATGAAAATAAAACTTTTTTTAAAATGCTCAGAGTCCTTCATTTACCGGTATTGTATAGTCCGAAAAATTGTTT
GGAATTATCACAGCACTATCACCAGGATGAAACTGAAAGAGTAAAGGCAAGATTTGTAGTTTTAAATGATTCGACAGGACGTAGATGAATACTATGA
ACTCTTCTCCTGAATATTTGAAATCTACAGTAAATGCTTGTCAGTTAGCTAATCTCTAAGTAGTTATGAATCCTCTCCAAGTAATTTGTTTAGTTCG
TCCAGTCTTTTCGATCGTTCTTCGTTGTTTTCTGCTTTTAGAAGGTCCTCCTTTTCTTTCTTCAAGGTAAGTCTGCTGGTGTGGTTCAACTGACTAA
ATTCCCATGATCCTGTACCGTCGAATTGAAGGATGTGCGTATGGAATTTCCATAGTGTAGGTCTATGTGTGATGGTGAGTAGAGTAATTCCCATGTC
TATAGCACTTTGGTATATGAAGCTTTCAACATCAATGGAGACTGCAGAAGTGCATTCGTCCAAGAGAGCGTATTTCGGTTTGTGATAAAACAATCGC
GCAATAGCCATCCTCTGTTTTTCACCACCGGATAAGATGTCAGTCCAATCTTTCATTTCATAGAATCCGTCCCTCTCGACGATATGATCTAGATGTA
CCATAGTCATTATTTTCAGTAAATTTTCTTCGGTAATACCTTTCTTTACCATATCCGTGTAAGTATCTGGGTATATTACTTGGTCTCTTAAATTACC
TATTACCATATAAGGCCTCTGAGGAATATAAAACATTGAATTTTTAGGAGTATGAAGCTCTCCACCGTAAATAGGCCATAATCCACTCAATATTCTG
AACAAACTGGACTTTCCACAGCCATTTGGACCCGTAATTAGCAAGTGTTGTCCCGGAGTCAATTCAAGACTTAGAGACGGACACACTATATCGCAAT
TTGGAGTTACTATAGGAACATTCTTTAGGATGATCTCGTTATTGGTTGAGTAGATGATTTTACCTTTGGCTAAAGGTTGATCACCACGGAACTCAAT
CTCGAAATCAGCAGACTTCTCTTTCTTAGCGACGAGAGTCTTATGGTAAATTCCGTTGCTGGCTTCCTCTAAAACCTCGAACATGTTGGCCACCCTG
GCCGTATGACCAGCTAACTCGACGATGTTCTTGTAGCTGCTCATGAGCCGTTCTACAGCATCTGAGCCAGTTATCAACAGGTTCTTGGAGGTAGTGA
AGTAGTGCGTTCGTTCTGATACGCTATCTTCTATTTGGTCGTCGGCTGTGTCTTCAGCGATGAGAGATTCTGATTTGGATGAAAAATCTGGTATGGA
AAGAAGACTTTTGTTGGAGATTTTGCGTCGATTGCCGGCTGCTAGGAGAATAGGCAGAGATACGACTATTATACCTGCTCCAGACCACACATATTTC
ATCAAGAATTGCTCCAACATGATGAACCATAATTTCACACCGAACATGTGTTCCAAGTGTTTAGATAAAACTCTAAAAGCTTGTCTAAGTTGACTTT
CTTCGACCTGATGGCCTCCGTAGAAAGCGATTTCTTCGGCATTGCTGACGATTCTTCCGTGGACGTGTCGTAGATATCCTTTCTTCTCCGCTTCTTG
GGCTACCAAGTGGCCGAACCTTGGAGACACAAACCTTAAAAGTAAAGCTGAGAACATCACCACTCCGCAAATAATCACAGGTCCTATCACAAGATTC
GAATGGCGACTCTTCACCAAATTGGCTAAAGCTATAACCATGAGAAGTATATCGAAACATGGCTTCGTTATCTGTCCGTAAAGGTGCGACACGGTGT
TGGCGACAGTTTCTATGTCATCTGTCAGCCGCTGCGCGCAATTGTCTAATCTGCCGTCTAGCACAGTCACCCTGTAGTAACTCTGGTTCTTGAAGTA
TAGTTTGTACGAGTGGTCTACCAGTCTGGTGCGGAAACTCAAGGCTATCCGACTTTCAAGATATCTTATCATGCTGTTAATAAACGTAGCTGGAATC
GCTACGGCAAACCATTTTCCTAGCTGCTTAACGAAGTTCTGGGGGTCTTTCATCACAATATACTTAACTATAGCACCTTCTAAATTAGCGACATAAA
TGCTCAAAAACGTCCTCAAAAACAAAAACGTCGTATGCCCGCTGAGTAAAACCGTTTCAGTACAAATGAAACTGGGTATCATTATCCTAACTAACTT
AATAAATTGTAAGATAAAGGCTAAATTGAAGTTCGGTATACTATTTTTAAGCCTTCCTTTGAGTTTCTTGTTCTTCACTAGGCCATTCTTACGTTGC
AAGTCTTTTTGAACTATGTGGTTGTTGTTGAGATTGTCCTTTGGTTTGTGTATCAAGGAGTCCACGAAGGGGTAGCCGACTTTGTAGGCGTAGGTGC
AAAGAACGGCTGCGATGAATGCTCCGGAGAAGACGGATTTGTTCTGCTTGAAGGTGCTTTCACCCTGTTCCAGGAACTTGGAGATTACTGTTGGCAT
GTTGTTTGGTCTGGTTTGATCTTGGTGAGTATTTTAAATTACTAAAGCTATTTGGTCAAATTGCGTTGGTGGTTTGTTTAAGCCTTAGGTTCCTTGT
TCCATGGTTTCAGATCTTATTGAATATAAACAGCACCTAAACTACACTATAAAGTGAAGAACCTCACTCACTTTGTGTCTGTATTCGACCGAAGGTT
TTAGCAATGGTAATTCTGCAAGGAATAATTCGAATAATAAAATACAAACACGGTCAGTATCGATTTTGGCAAGGTCGACAATAATTACACCAAACAC
TGAAATCTCGAGATTACCTGAGAACGTCGATGTATAAACCGAGAATGAAAACAAAAGTGTCTGTACAAACCAAAATACTATCTAAACAACAAACTGA
CCGGAGGAGACAGTATTCTGACTGAAGAATGTTACTCTGCTCGTTTTGTGTGAATAGAATGAATAAAGCAAAACAAAAAATTGGAACGAGGTAGCGGA
GCGGGTATCTCACTCGGACTCGGTCG
>Dvv ABC-D D11628
GTTCCAGTTTGTTTAACTATAAGGTATACAGCCAATCCAGGGAACTACCCCTTTTTAGTTTAAGATATCTTTATTTCTGAAAAGTACGGTCGACGGA
AAAGTTTTGTAACGAACTCGTGAATTAAAATGGCGATTAACAATCTCAAATATTAACGCGCTCGCGCCGCTCACGCGTTAAAATATCTCAATTGTTA
ATCGTCCTTATTAATACACTTGTTGGTTAAATATCTATTATTTAATTGTAAACATTTAATTACTTACAGATGTCATCAACAATGACGATGTATATTA
CTGAAACTTCACTGATGAGTGTGGATACTTTTAATCAATGATTTTATACTTTTTATTACTTAAAACATAAATATGCGGCAGTTTTTTGTGTATTTGA

AAACTTACAAAACGGTAGAAAATGAGGAATTTTGCGAAAATGTTCTTTTCTGATAATCGAAGGACAAAATAATCAAGATGTGTATAAAGTGTTCTTA
TTTTATGTAAAAAAATATACACTTTCACACACTGGTGATATACACTAAGGTATCAATTGCCCTTGTGGTATATAGCTTTGTGTAGTTTGTTTCGACA
TTTGTTTCTGGATTCTACATTTGTTTATTGAGAAATTCTCCTAAAAATATATTTCATAATTCGGACACTATGATCCGAATTGTTCATCACAATTATC
AATTGGTTTAAAACTATATCCTCCTCGCCCGTCTAAATGTAATACGTATTCGTGATGCTGCCAAAGTGATTTTCTATGTGATACTGTAAGTAACGTG
ATGCCAACATCTCTACAATACTTATACATACTGCCCTCTACATCGACAGATACAGCGCTTGTGCATTCGTCCAAAATGGCGAATTGAGGTTGGTGGT
AGAATAATCTGGCCATAGCTATCCTTTGCTTCTCTCCACCACTCAAAACATCTAACCAGTCCGCTACAGCATCTAGACCACCCTCTCTTTCCAAAAT
GTATCCGAGTTGCACCCTTTGCAAATACTCCTCCAACTTGGCATCTGTCGTTCCTCTTCGCGCTGCTTCTGCCCCAGAATGCGGGTATGTCAGTTGA
TCCCTGAGGCATCCTAACGTCATGTACGGTCTCTGTGGGATATAGAATAATTTCCCCCTAGGTGGTTTCGTGAGCTCACCACCGAACAGAGGCCAAA
GCTCTCCTAAAATTCTAAACAAGGAAGACTTTCCGGCCCCATTGGGACCACATACCAACACATTCATGCCGGAGTTAATTTCAAAAGTGATTTCGTT
AATGAGAACGTCTCCATTGGGGGTTATTAAAGGAACTTTGTGGAATTTAATGATGTTGTTTCTGAAAATTAAYTTTCCGCCGTTGATCTGTAAAGAT
TCTGATCCGGAAACCATTGTTCGTTGGTATTTTCCAGAGTTAAGTTCACTCAACACTGTCTTTAATTGTGTAACTCTAGCTGTGAAACCTGCTAGGC
GGGTTAGATCCCTTCCTGCTAGGACCAATCGACCAATTGCTTCGGCTAACTTTACTAACATCCTACCGTACGTGTAATACAATCTTGATCTTTCGTT
TGTTCCAAGAAGCCTAAAGCTGTGTCCACTTGTCATAAATGGCAACGACACGACCCAAAATCCTACTACTCCGGCAAAATATTTAGCAACAATGTTG
TCCACGACGCCCATTGCGACTCTAAACCTTAAGAACTTCCTTAAGTGATTTAAAAGCTTGTTGTAACTGGCCATGAGAGTTGCCTTTTCTCTAGAGT
TACCGTTGTAGAAGGCTACTTCTTCTGAATGGGTGATTAGACGTGAATTGATATGCCGAAATTCACCTTCTAATTTTTGTTCTCCCGCTGTTAATCT
TGCTGTTGGTTTTCTCAAATTCGTTAATAATACGCCAGAAACTAATAGATACACCAGCATAATTCCAGGAGTGCCCCCTCCCAATGTTGTCGATAAC
TTATAAACATAAATACAAATATCCAGCAAAGGTTTGGCAGTGTTGCAATACAAATCCGTAACTCCTTCACAAAACTTGTCTATGTCGGTCGTCAACA
GTTGATCTGCGTTGGATATCCTATTGTCCAGATTTGATATTCTGTAGTAAGTGTAATTCTTCAGATATTCCTCATAGAGATGTCGTGTCATGTTAGT
CCGTAGTTGGATTTTCAGAGCACCTATTGAGTATTTCAGGACATTGTTTACAACTGAGATTATTGGAATAGCTAAGACGTAGTAGAGCAGGCGTTGC
TTGAAGAGAGTTGGGTCCATTGCGATGATTGAACTTTCGATTTAGTACCATGATTGATCAGCCACAAATCGCACATAGATCGGGATACTAAACTTA
AGGCGATTAAGAAGAATAAGCCACTTTCTTGGCTGGTCCATCCTGGGCAGGCGATTCCAAATAGTTGCGATAATTGGGAGAAGAATTTCTTGTCTAC
TTGTGCTTTGATTTTAGTTTCTCCTTTTTCCGATATTAAATATTTCACCTCTTCTTCTACGGCTTTTTGTACTTTTTTCTTGTTAAGTTTTCCCTGT
TTATTCCTATATTTCAAAATAATAAGCAGAATTGTTCCTAAGGCTCCTGCTCCAGCAAGAGCCTTGTGTTTGTTTAAAACTTTGCTATAATTAGGAG
CCATCACTAACACATTTTTATTAATAAACTTATAAAAAAAATTAGCGTAAAATCTACCGTTCACTAAAATCGAAACGGATTAATTACCCAACTAACT
TCTCCGCCGCTCGCGCTTGAATAAAAATAAAATGAAGTAATTTTAAAAGTTT

>Dvv ABC-E D2830

CTTTAATTATACTCAAGGGATACAATTTATTAAAGTAGTAAAAAGTAGCTCAATATCCTTTGCCTGCAATCGATTATTGACACGTACAATTAATAAT
TATAACGAACGTTTTCCTGTCTAACTGGATACAAAAACCAACGAATTAATCCTCCAAGAAGAAATACTGCCCGGCCCTCTTCTGCTCCACGTCCTTC
ACCGACTCAAGCTTGTTGATCCTAGGTCGGAAATTGTTAGGATCACGTCTGAAAGTTATGCCGAGGAGTTCCAGGAACCGGTTCATTCCGGCCAATA
AGGTTTGAGGGGCGTGGGCTGTCGTCTTCACTGAGGGAGACCCTTCGAATACGATGACACGATCGGCCAAGTAGGTCGCCATTATGAAATCGTGCTC
TACTACGAAACCGGTCTTCTTCGCGTGAAGTATAAACCGTTTTATGACCTTAGCTGCTACCAAACGCTGTTCAGAATCCAAATAGGCACTAGGTTCA
TCAACAAGGTATACATCAGCGGGTTTTCCAAGACAAAGCGTCATTGCCACTCTCTGCAATTCTCCTCCAGATAAATTCTGTACTTCTTGATCGATGA
TGTCTTCTATCTTCAAAGGCTTCATTACATCTGCTATAAATTGTGGATGAATATAGGCATCCCTAATTTTCTCATGAAGCAACTGTCTTACCAATCC
TGTGGATTTAGGACTGATCTTTTGGGGCTTATAACTTATATGCAACTGGGGTAATTCACCTGATCCACTATCAGGCTCCAAATTTCCAGCCAACATT
CTAATAAAAGTAGTCTTTCCGGTACCATTCTCTCCCAACAGCACCAAAATTTCAGAATCAGAAAACTGACCCTGGGCAACCTTCAACTCAAAACTAC
CCATCGTTTTAGTCATAGTTGGGTACTCGTAGTGGTTCATACGTTTAATTTCTTCTTCTGTAGCCGACTCTGCTACTTTAAAAACTAAAGATTCATC
TCGGAAACGGAGATTTTCGGTTGGGACAAAACCGTCCAGGAAAATGTTGATACCTTCACGGACAGAGAAAGGCATAGTTACTACACCATAGGCACCA
GGCACTCCATAAAGACAGCAAATGAAGTCTGACAGGTAGTCAAGAACACTCAAGTCGTGTTCCACCACAATAATAAACTTGCTGGGATCTATCAGGG
ATCTGATAGTTCTGGCAGCATTTAAGCGCTGTTTGACATCTAGATATGATGAAGGCTCATCGAACATGAATATGTCTCCATTCTGAATACATACCAT
GGCACAGGCAAATCTTTGGAGTTCTCCTCCAGATAATGCGGCGATTTCACGTTCTTTTATGTGCAGCAGATCCAACATCCCACAAATTTGATTCATA
TTGTCTAGTTCATTTTTTCTGTCTAGAAGTTGTCCCACAGTGCCTTTAACTGCTTTGGGAATTTGATCAACATACTGGGGCTTTATAAGAGCCTTCA
AATCATCCTCTAGAATCTTAGTGAAGTAATTTTGCAATTCGCTACCTCTGAAATGGCTTAGAATTTCTGTCCAATCTGGAGGATCCATATATCTTCC
TAGATTGGGCTTTTGTTTTCCTGCAAGAATCTTCAGGGCAGTAGATTTACCAATACCGTTAGTGCCTACGAGTCCAAGTACCTCACCAGGACGAGGT
ATAGGTAACCTATGGAGCTTGAATGAGTTCTTACCATATCTATGAGTAGTTTCTTTTTGCAAATTACTTGGTAAATTAATAATAGCAATGGCTTCAA
AGGGGCATTTCTTAACACAGATACCACAACCAATGCACAATTCCTCAGATATGGTAGCAATTTTACTGTTTGGTACAACTTCAATGCACAGTTTTCC
AAGCCGGACCACTGGACAAGATTTCTTGCATTCCTGCCTACATCTTTTGGGTTTGCATTTATCTGCATTGACAATGGCTATACGGGTGAGTTTGTCA
GACTCTTCCATTCCTTTTCGTTTGGACATTTTGCTAAAATCTGTCCTGTTTTTTTCCTTTAAACAGAAATGAAGACGGCAAACGTCACCTTTTAATA
ATTTGCCATGAGGAAAGGAACAA

>Dvv ABC-F D2701

GGTTAACATATTGTTACCAATTTTAATATAATAATTTAGAAATGTCGAAAAAACGTGGGGCCAAGAAGGGAAAAAATTTGGATGATGATTTTGAAGA
AACCTCCAGTATTATTAGCGAAAAAGAAAAAATAAATTCGAAAACTCCCAAAAATAAGACTACTAAAAAAGGGAAAAAAGGAAAGGATGACTGGAGT
GATGACGAGGAAATTGAACAAGGCAAAGAAGTTAATCAATCTGAACCAGCGGAAAGTAAACCAGTAGCCAAGAAAAAGGGCAAAAAGGGCAAAGGCA
ATAAGAATGATGACTGGTCAGACAAGGAAGAAGCGGATATAAAACTTTCAGACAGTGAAACTGAGTCAATGCCAGTAGCAGTGAAAAAATCCGCTAA
GAAAAATAAGAAAAAGAAAGATGATTGGTCAGATAAAGAGGTTGACATACAATTATCTGAAAGTGATATTGAAGAGGCACCTAAAATTGTTAAAAAG
TCAGCTAAGAAAAAATAAGAAAAAAGATGATTGGTCTGATAAAGAAGATATAGAAATAGAAACTAAATTATCTGATAGTGAAGAAGAACTAGCACCAG
CTATTGTTAAAAAGTCTGCTAAGAAAAATAAAAAGAATAAAAATCAGGTACGTGATGAATCTCCAGATATTGAGGAAATGGAAGCACACGATGAAGG
TTTTAAGGAGGATGACGTAAGCGAGGAAGAAATTGTAAAACCTGTTAAACAAGACAAAAAGAAAAGCAAAGGCAAAGAATCAAAGAAGGATGAAAGT
AAAGTTGATGTAGAAGTTAAAGAAGAAATTAAGGAATCTAAAGAAGATTCCCCTATAGTAGAAGACAAAGTAATAGAAAATGGCAAACCTGAAAAAC
AAAACAAGGAGCTAGAAAATAAAAAGGATAATGATGTAGAAGAATTAACAGAAAAAATAAGTACTACGCAAATAAGTGATGATAAATTAGATGAATC
TAAAGAAAAGAAACTCACGCATAAAGAAAAGAAAAAAATGAAAAAACTACAGGAATACGAAAAACAAATGGAAACCATGTTGAAAAAGGGTGGACAG
GGGCATTCTGAATTAGATAGTAATTTTACTGTATCACAAACTCAAAAGACTGCTGGACAATTGGCTGCATTTGAGAATGCAGTAGATATAAAAGTTG
AAAATTTTAGTATTTCTGCTAAAGGTAATGATTTATTTGTGAATGCCAATTTATTGATAGCCCAGGGCAGACATTATGGTTTAGTTGGACCTAATGG
CCATGGTAAAACAACACTGCTGAGACACGTAGCACAACGTGCATTTGATATTCCCCCAAATATCGATATTCTCTATTGTGAACAAGAAGTCGTGGCA
GATGATAATACTGCTGTTGAAACCGTCCTAGCAGCCGATGTCAAACGTAATGATTTATTAGCCGAATGTAAAAAATTAGAAGCCGCTGCTAATAGTG
GAGATTTAGAAATTCAGGAGAGGCTTAACGAAGTTTACAGTGAACTTAAAGCTATAGGGGCCGATTCTGCTGAACCTAGGGCTAGGAGAATTTTAGC
TGGTTTAGGATTCGATAAGGAAATGCAGGATCGTGCTACCAAAAATTTCTCTGGAGGTTGGAGAATGAGAGTATCTTTAGCAAGAGCATTGTATATT
GAACCAACGTTACTTTTGCTTGATGAACCGACTAACCATTTAGATTTAAATGCTGTAATTTGGCTTGACAATTACTTGCAAGCTTGGAAAAAGACTC
TGTTGATCGTTTCCCACGACCAGTCTTTCTTAGACAACGTCTGCAATGAAATTATTCATTTGGATAACAAGAAATTATACTATTATAAAGGAAATTA
TTCCATGTTTAAGAAGATGCATGTGCAAAAGAAGAGGGAGATGATCAAAGAGTATGAGAAGCAAGAAAAGAGGATAAAGGAATTGAAGTCGTCAGGT
TCTTCCAAAAAACAGGCGGAAAAGAAACAAAAAGAAGCCCTTACTAGAAAGCAGGAGAAGAATAGAACTAAGATTCAGAAACAAGAAGAGGATACAA
CTCCCACAGAACTATTACAAAGACCTAAAGACTATTTGGTCAAATTTCGCTTCCCAGAACCTCCACCTTTGCAGCCACCTGTTTTAGGATTACACAA

TACAAGATTTGCCTATCCTGGACAAAAACCTCTATTTGTAGATACAGATTTTGGTATTGATATGAGCAGTAGAGTCGCAATCGTAGGTCCAAACGGT
GTTGGCAAATCAACATTCTTGAAGTTATTAACAGGAGATCTATCTCCTGATAAAGGAGAAAATAGGAAAAACCATAGATTGCGAATTGGTAGATTCG
ATCAGCACTCTGGTGAACACTTAACGGCTGAAGAAACACCTAGTGAATATCTGATGCGTCTTTTCGACCTTCCCTATGAAAAGGCTCGTAAACAATT
AGGAACATTCGGATTAGCTAGTCACGCACATACCATTAAGATGAAAGATTTATCCGGTGGTCAAAAAGCAAGAGTGGCCTTGGCAGAACTCTGTTTG
AATGCTCCCGATGTATTAATTTTGGATGAACCGACGAACAATTTAGATATAGAATCCATAGATGCCTTAGCTGAAGCAATTAACGAATATACCGGAG
GAGTCATAATTGTGTCTCACGACGAGCGACTTATCAGGGAAACCAATTGCTCTTTATATGTCATAGAAGATCAAACCATCAACGAATTAGAAGGAGA
CTTCGATGATTACAGGAAGGAATTATTGGAAAGTTTGGGAGAAGTCATCAACAGTCCCAGTATAGCAGCCAATGCAGCGGTTGCCCAATAATAAATA
TAATTTTGTCCTTGAAATATGATTTTTTTAAATTCTGCGATCATGATTCGACCTTGTTTTCCAAAAACTTAAAACATATTAAGGCTTGTCACGCACA
GGGAGCTATACTATAATATACACACACCTTATATGGAATCATCTGATATTTCTTACTATTTCGTACGAATTTAAAAAAACTAACACACGAAGTCAAA
CAATATTTATTTTAAAGAAATTTAATAACAAATACATAACAATTAAATAAAACAATAAAGTATTTAACAAACAAATTGAAAGTAGTTGTTTTAAAGT
CAATAGCATACAAAATTTCAATGTAAAACGCATAATGTTTAAATTGTTTTTATTTTTTTTTGTATTTTGTGGTAGTCAGTGTTATCACCTCTAGTTC
TTATGCAAGATTCAGTTTGCGTGGGCACGCTCCTAATCAAATTATCAACATTTTGTTGTGGTAGGTTGCTCCATCTTTTAAGAGCAGCTCGTACTAG
CCGTGCGGTGTTTTGTGGATTATCCCGGCCAGCTCTAATTTTTCTTTTAAGCATATCCCACAAATACTCTATAGGATAAAGCTCGGGTGAGCAAGCA
GGCCACTCCAAACAAGGGATACCTTCTGCTTCAATGATGTCTGTAGTCACTCTAATGGTATGTAGAGGTCCATTATCATGCAATAAAATTAAATTTT
CTCCTGTTGCACCTCTCCAAAGCCTGACTACAGGTTATCAAAATACCTGTGAACAGTTAAAGTTGATTGGATGAAAATTAAAGGAGTTTTTTTAACG
GATCACAATTCCTCTCCAGAACATTACACTTCCCCTTGTATATTTGTGAACAGATCTGACAGTTTTCGTTCTTGCTTGTCTTCCTCGACCTCTAAGT
ACACGATTTCGTGGGTCATCTGATTTTCCATCTGATTACATTAAATTCTGACTGTTTTTGAAAATAGCACATTTTGTCAATTCCCAATGTTCCAGTT
TTGGTGGTGAAGACACCAATTTAGGCGATCAATCTTGTGCTGCCTGGATAACTCGAGAACCYATAAYTGTCACCTGCTGTATACTTCTTSGTACGAA
CTCTTCTTCTTATCCT
>Dvv ABC-F D802
TCAATTTGTTTATTACTCTGTTTTAAATACTTTTTATCACAATGATGTTCAAGAGGACCTACCAATCACGATTAAAAACTGTCGGTAAATTTTAGCT
AATGATGTTAAATGCTAAGAAATGTGCATAATTTTCTAGTGTAAATTAAAAAAAGCCTAATTCTAATATTTTTTTAATAATCAATCCTTTGAAGCCC
ATAATATAACAATATGCAAATATGCCTCACCGAAAAAAGACTTTAGTTAAATAATCAAAAACTTAATAGCATTACAAAAGATATACATCATCTTCCT
AATTAGATGACTAACGGTTTTGAGTGTTATAAAAAAGCTAGCATGTATGGCTTCCAAAGACCGATTTTAGTCGAATTTTATATAAAGCACATAAATA
AATGTATATTTAGACAATAATAAACAATCTAATTATTAGTATAGATCAATGAAATACTTGGATTTGAATATCACTTTTGTATTCGTAACATGATTTA
ATGAAACAACTTATATATGAAACACGTTTATTAAATAAATAAAACCAAGGCCGGCCAATCATTTAATTTCGACATACCATACAATATGCTTAAATAA
TTTTATGTTAAAAATGAATCTAACGTGATAAAAGAATTACAGTTTCTTCAATAGCTCCTAATAATCTGAAGTATTGCTTTAAATGTTTATCATTATT
TCCGTTTTGCCGCATCCTTCAAAATTTTAGTTTTCAAGTGATCTTTGTATGAAAGAATGTCGCCTTGCCACTTCGTTACTGTTCCTTTTTCGCATAC
CCAAATTTCTTCCGCAACCTGACTGATTAATCTAAAATCGTGACTAACAAGCACCATTCCTCCTTCAAATTCATTAATAGCATCTGCAAGAGCATCG
ATCGTTTCCATATCCAAGTGATTCGTAGGTTCATCAAGAAGAAGCATATGAGGTACTTGCCATGCTAGCCAAGCGAATACTACCCTGCACCGTTGAC
CGTCAGATAATTGTCTAATAGGGCATACCTGTTGACGCCCTGTAAGACCATACCTTCCTATGATTTTTCTCATTTCTTCTTTTTCTTTAATATCAGG
GAATTCTTTCATCATATATTCAAGCGGCGAGAGGTCCAAGTCTAGCAGTTCGTGTAGATGCTGATGGTACCGGGCAATACGTAAATGAGAGTTTTTC
CTAATCATACCTTCTGTTGGAGTAAGATCACCATATAATAATTTCAAGAGAGTACTTTTTCCTGCTCCATTGGGTCCCACCAACGCTAACCTAGTGT
CCAGATCTATACCGAATTCCAAATTTTTATAAATCAAAGGTGTAGAATCGTTATATCTAAAACTGACATTTTGTACCATAATTACTGGTGGAGGTAT
AGTACCACAACTGGGGAAATAAAAAGTGACAATTTTGTCACTAGTCACTTTTTCTGTTAGACCCTGAGCTACCATTTTGGCTAGCGTTTTTTCTTTG
GACTGCGCTTGTCTCGCAAGCTTTGCTGATCCGTGACCAAATCGTGCAATATAGTTCTTCATATGGTTGATTTGATCTTGTTCCCAATTATACTGTT
TCATTTGATTTTCCAAAAGTTCCATTCTAGTCTTAACAAAGGCGTCATAATTGCCGGTATAATATTTGAGCCTTTTTTTATTTATGTGTATTATGTT
TGTGCATACTCCGTTTAAGAAATCTTGCGAATGTGATATCAACACTAATATTCTCTTATAATTTCTTAGTTCCTCTTCTAACCATACACAGGCATCC
AAATCCAAATGGTTGGTAGGTTCATCTAATAAGAGAAGATGTGGCTTGACATAAAGAGCTCTAGCTAGTGCAATACGCATTCTCCAACCTCCACTGA
AATCTTTAGTCTTCTTATTCTGCATTTCTCTGGTAAAACCAAGACCATGTAGAATATTAGCAGCTCTTGCTTCAGCTGTATCAGCTGCCATATCATC
CAATCTCTCGTAGATATCCATTAGTTGTTCCTGAGATTCGTCATCTTCACATGCAACTAGCTCCTCTGCCAACTTTTCCAACCTTACTCTTTCTTCA
TCTACCTCCATAACACATTCAAGGGCTGTTTTATCAGAAGCAGGCATCTCCCTAGTTAAATGGAAAATATCTATATGATCTGGAATGGGCACTTCTC
TGTTACCAAGTACTGCCAAAATTGTGGATTTTCCACTTCCATTTAAACCTAGAAGACCATATCGTCTACCACAGTTCAATTCCAATAAAGCATCTTG
TAACATTTCACAACCATGGAAAGTAATGGAAAAAGTATCTATTTTCACATCTCTAGATTTTGGATGCACTGCTAAAGAACCAGTACATGCACGAGCC
TCTGCATTCAAACGGGCATCCGCTTCAAGTTTTGCACATAAGGCTTCTTCAGCACTTAATTCACTCGAGCCATTTGTTGCACCATTACTGAGACCAT
TTGTTTTAAGAGTGTCATTTTGCTCTTCTCCTTTTTGTTTTTCATTCTTTTTTCCTGCTTGTCTGGCTTTTGCTGCATCTTTTTTACGTTGTTGTTC
ACGTTTCTTAGCGTCGGAGGGCATTGTTATTAATTAATTTGAATACTTTATGCTCGATAATACCTAAAAATAAAATCGGCTCAAAATTACATGTGTC
CTCTAATTGAATGGACACAAAAGAGGTCTATCGTCTATGGCTG
>Dvv ABC-F D9935
GTTTTAGGTTAGGGTAACTTGTCACATAAATTTTTATAAGAAAATCTGGATTTTACTATTCTAAACAATGGGGACGTGTAGTGAGTACATTAAAAAT
GTTTTTCCTGCGATAGATGAAGAATCTAAACAATATGTAGAAGGTGTATTATTAAATGGTGCGGATGATTTTGAGGATAGTGAAGAAGTATATGATG
CAGTAGGAGAAGTTCTAAAAGAAATTTCGAACGACAAGTCCGAAGATGATATTAGAAATATCTGCAATGATCTTCTATGCATGTTAAAACCGGATAA
AGGAGAAAAGGCTACTAACGGGGCTATGAAAGTTTTAAATGCGCCAGTACATTTGGGGTCAATGGTAGACAATACGGATACAAATATTGATGATGTT
AAAAGTATATGGCTTATACAACGGGATGATTCTCTGAAAGTCGATGCAAGAAAACTAGAAAAAGCTGAAGCTAAGCTACAAGAAAAATTAGACAAGA
GAACGAAAGAAATAAAGGTTATAGCACCGCCCAAACTACAGACCGCTACAGCATCGCAAGTTACCAGCAAAAAGGACAGCAAACTAGAAGCCAAAGG
TACCAATAGGACACAGGACATAAGAATAGAAAACTTCGATGTTGCGTATGGAGATAGGGTTTTGCTTCAAGGAGCTGATCTAACTCTTGCTAGTGGT
AGAAGATATGGTCTTGTAGGTAGAAACGGTTTAGGAAAAAGTACTCTCCTAAGGATGATTTCTGGAAGCCAATTAAGAATACCCTCCCACATATCAA
TTCTTCATGTGGAACAAGAAGTTGTAGGTGATGATACTGTAGCCTTAGATAGTGTACTAGAATGTGATACTGTGCGCGAAGAATTGTTAAAAAAGGA
AAAAGAAATTAGTGCTGCAATTAATAGTGGATCGGTTGATCCACAACTAAACAGTCAATTGACAGAAGTGTATAATCAATTACAAAACATAGAAGCC
GATAAAGCACCTGCCAGGGCTTCTATTATCTTAAATGGTTTAGGTTTCACGTCCGAAATGCAACAGAACGCCACTAAGACATTTTCTGGAGGGTGGA
GGATGAGATTGGCATTAGCTCGTGCTCTTTTCTCTAGACCTGATCTCTTACTTCTTGATGAACCTACTAACATGTTGGACATTAAAGCGATCATCTG
GTTAGAAAACTATTTACAAAACTGGCCTACGACGTTACTGGTAGTTTCTCACGATCGCAATTTCTTAGACACGGTGCCTACTGATATTCTACACCTT
CACTCTCAAAGGATAGAGGCCTACAGGGGAAATTATGAGCAATTTGAGAAAACAAAGACGGAAAAATTGAAGAACCAACAACGAGAATATGAGGCGC
AAATGCAGCAAAGACAGCACGTCCAAGAGTTCATCGATAGGTTTAGGTATAACGCAAACAGAGCCGCATTGGTGCAGTCTAAAATAAAAATGTTGGA
AAAACTTCCTGAATTGAAGCCTATTGTGAAAGAAACAGAAGTTGTATTAAGATTACCAGAGACAGAACCATTGTCGCCTCCCATTTTACAACTAGAC
GAAATATTGTTCCGGTATAACTCCGAACGAGTTATATTTAGCAACGTCAACCTCGGAGCTACCATGGACTCCAGAATATGTATCGTCGGTGACAATG
GCGCTGGTAAGACTACCCTTCTCAAAATCATCATGGGAATATTATCTCCGACATCTGGTATGAGAAACGTGCACCGGAATCTTAAATTTGGATATTT
TAGTCAGCATCACGTGGATCAGTTAGATATGAACGTCAATTCCGTAGAACTTTTACAACAAACATATCCCGGTAAACCAATCGAAGAATATAGAAGG
CAACTGGGCAGTTTTGGTGTTTCCGGTGATTTAGCTCTCCAAACAGTATCGAGTCTCTCTGGAGGTCAAAAATCGAGAGTAGCCTTCGCTACGATGT
GCATGGGCCGACCTAACTTCCTAGTTCTCGACGAACCTACGAATCACCTGGACATAGAAACGATAGAAGCCCTCGGCAAAGCCTTAAAGAAGTACAC
GGGAGGTGTAATTTTAGTATCTCACGACGAACGTCTGATAAGAATGGTCTGCAGCGAATTGTGGGTCTGCGGAAACGGATCGGTCAAAAGCGTGGAG

GGAGGTTTCGACGAATACCGCAAGATCGTCGAGCAGGAGTTGGAGGCTGCGGCGCAGTCGAAATAGCATTTCTAAACAAATTATTTATACGGTAATA
AAAAAAATAAAAGCACTGTGATTGAGGCATGTTGAAGCTGAGATTGACTGTTCCGG
>Dvv ABC-G D9811
GCGCACTATTTGCGGACTTTGCATACCTATATTATTAATATCTATATGGAATCATAAGATTCGATTCCAGCAATAAAACTGCTGGTAAATAACTTTT
CCTTGTATTTTGCTAATTAGCCCAGAGTATTAATGAACATTAGTAAAACACATTTTTAATTACAACTGAACACTGAACAAAATCAGGAAACATCATT
TTGAACTAAAGAAATTTGAAAATATACACTGACTTACGCCGCCACAGGTATAAAAAACCATTTTTGTTAAATATTGGGATGTTCCTTGCAGGCAAAA
TGTTTGCATTTATCGCGTTTCGATGTGAATTTTTCTGTCGGATAATCAACTGCAAATTTTACATCGGCCACGTTACCTACCATTTTCAGTTGGCTCT
GTATCTTCAGAATCTTCTGAAAATTTTGTTAGTACCATATCCAAGGCTGTTTGAATGTGTTTATAAAAACCTTGACGCGGTCGGGATTGTATAAAGG
GTGTGACTAGTTCAAGGGCCAAATTTTGAAGAAAGAGCCTCCTTGGAAAAGATTTCCTTTTTCGTAGTCTTCATACTTATCTATGTATARGGCAAAT
GCATTTAGCCGCCTATTACTCCAAAATTACTCAGAATTTCACATAATAGACGGTATTTACTTAATATATAATTGGAACATGGAAGGCCAAGGGCGGT
CGTCTTTGACCCCAAATTCAGAAAATATTTTTTATTCGTGAAATATTGGGACCGTGCAGGTTCGGCAAAGCGACACCTGGTTTCTACGCTCAGCAAC
ATTATTCGCACTTTTAATTATATTGGCCAATCATATTGGTCCTGGTTACTGGATAATTGTCAAGGTCATAGTCCAAAAAATAATAAGGAAAAATACG
ATTCAGGTTATGTTATTAAAACGTAAACTATATGTAATTATATGTATTAAATAAAATTAGTTATTAAAATGCAGTACTACAAGCAAAATACAATTAC
TTAAATTTACCTTTATATAATAATTGCATATCATATCAATATTGTGGAGCAATATATAATTTTTCTTCTTCAATGACAGTAGGTATGAAATATACGT
CAATTTGACAATTTCAATTGACAATATGAATTATTTAAGAAAGTTGCAATATTTCTCCGCTATTCGTGCACGATCGTTTCTCGTATCTCCTCCAGGT
ACTTGCACACCTTGAATATTGGTTAAACAATATTGTCCTACTGAGATAAAAATTTGTTTATATATTGGATAATATTGTGACTTAAGGTCTTTTGCCT
CGGCATATAGCAGCAGATAAAAAAGTGTACAAAATAGTATGTGCCCACATAAGCGCAAATAAAAACACAAAAATGATGATGACAAAGATGATTAACA
ATTTTTTGGTAACAAGTAAATTTCTATAAAACTCATGAAACTGACTGGGGATCCACTGTAAAAGGATTAGTATGTGATCATTTTTCTTTTGTGTTAT
ACCTTTCGGTTTTTCATAACGTAGTGGTAACTGTTGTGAAAGATTGGTCGCAGTTGCTTTTATTTTTTGTTACTGTCAACAATTTATCAAGTAAAAC
TCTCTGTCAGTCAAAGTGACTAACACTTGTATGGATATACTGCTGTTTACCTCGCGCGAAATAGCGCGCCCGATAAACAAATCCGACAGCTGATGGA
GTTACCACGCTTTACTTCGTTCTCCAAAGGATATACTACTTTTTCTTATAAGAATATAATATAGGCCACCAGTAATACTTCGCATATCATATTTTGC
CACAAATAAAATTGGAACATGTTAAGCTTAGCAAAAGTATTAAGTCCAAATTAGAAAAAAATGGAGTTACTACCCTTTATCTCTTCGATACAGATAT
AGATCTTCTGTCATATTTTCAGTAGTTTGTTTTCTTTTCCTCGATAAAATCATATTATCTTTTACACAGACATAATCTTAAGCTTTGGAATTTCAAA
ATAAATTTTTACAAAATAAGTTATCAGTATCAATACAATACATCAGTAAATATCAGCAAAAAATCAACATTAATTCAGAAAAAAATCAAAATTCACT
CAATTTAAAAAGTGCACAACAGGATTTATAAAAGACCTCGTTAGATACCAAGCCTTATTGAAATATATTCTGATATTTATTAGTCATCCATGGAGCT
TAAACATTATGAAAAATAAAGCTGAGTTTAACAATTATGAAAGGCATATGAAATTCTCTTACTGGGCATACAACCCTGGGCAGACAAGTCTTTGTCT
TTTGGTGTCATCTACAGTTAAACTTCAATTCCGAACGACTAAATACAGATGACCTACAGATTCAAATTTTTGGGTTTTGTAATAATAGGACCTACAT
CATTTTTTGTAGTATAGATTGGTACTATATAGACTTCTGGACTATATAGGTGGTTCAAATCTATTACTTATCAAAAATAAAGCATCAGATATGGTTT
TGCAGCTATTTCCCACAAGTATTTCATTATTAATTTCACGTAGGTACACCAAACATGCCGACATTAATATTTATATCTATAACAAACAGACAGAACC
CGTACAAAGATTCCAGTATTTTGATTGCTCGATAACAAGCGATCTTGCCGTGAGGTCGAAATTCATACAAAGCATACTCATAGGTATAGAGTATGCT
AGGCCCGCTTTTCAAACAATGAAAAAATTCCTAATAAACTCTACCATATCGTTGGATATTATGTACCAATGTGTAATGGTTATTCCATACGAATGGT
TATTCCACGAAGTGGAAACATGGACTTTCGGAATTATACTAGTATGAGGCGTATGGTGTGGGTTGTTTGACGAATGTTAAAGATCTCTTGGACAGAA
GAGCACGTGATCAACAACAAGGTGCTGACAAGAATGGGGACTGAGAGAGAACTCATAAATATTAAAAAAAAACAGAAACGGGAAGAAGAGGTCCAGG
AAGAAGAAAATACTCTTGTCCGAAGAACGTAAGAGACTCAACAGGCAAGGACATACATTATATACTAAGAACAGCTCAAAATAGAGAGCAATTTGCT
GTAGTTATGTAATAAAGTAATTACCTTCAGTAATGGAGAAGGCACATTAAAATGAAGATTTCATTGAAAGTAGATATGTGGACCATCTTGTAGCAAA
TACACACTTCTTAATCTTCTTGATCCAGTAAATTAACAATTGTATTGAATATAGTATTCCGATAATAAGTTATTGTAAACAAAATAACTCACTAACT
CGTTTGTCGACGCTCGTCGTTCTTTTTTTTGCAAAATCACAGCCACAAGATGAAGCTAGCAAATAATAAAGAAATAAAGTGAAAAATAAACTTAGCA
TTTTACTGTGACCTAAAACGTGTATCTACTACTATACACTGCGAGGCATAAGAAAATTATGCTCATTTTCATAGTTCATTTTTTCGTGAACAGATTG
ACGGATCTGTATCTGTATCTGATCTAATCATCTACAAGAACCTGCCTAATACTGCCTTTGAGAAGCCACTGCCCTGAAGAGCGACAATGTTGTCAAG
AAAATTGCCCCATAGGGTTGTATGGCCCTATAATTGTACACTGCATCACATAATATTTGAAATACGCTGTAGAGTAAATGGACTGGGCATACGAAAA
ATCTGTATTCTGTACACACCATTATTTTAAGTAAAGACAATATGCGCAATACATAAAAACATAACAGCTTTCAAGTATATAATAAAAAATACGTATA
ATTTAACAATAAAGATTGTTGTTTCACAGTTTCTGCGTGGTTTATCATCAAATCTTTTGCATGTTAAAGAATCATCGCTTATTTAATCTCAACCATA
AGGTCAAATACGTCACAGCATGCATAACACACCAAACAATAACAATCAATATAAAATTTGAAACTAAATCTACATCTAAAATATCCATTTCCGAGAG
GAACTTGTGAGGATCTTTATAATGACAGTATTCCTCCTCTATTGGGCAAACCAAGATCTCTCGGTTAAAGCCGTAAATACTATAAACGACGCTTTGG
AATGCTGCTCTAAAGTAACTGATGAAGAAGATCGGTTTGAACATGGTAGAGGTGTCTATTGCTCTGATGCAAAAACCGAAGACTGATAAGAAGCAGG
CTAACACTGGGGCGATGAATACAGCTACCTTTACTGGTGTTGTGGCTCCTATAAAATATCCCATGGATTGGCCACAGAGAGAACACGCTGTTACAAA
AATGACAAACAAGAATAACCTAAAGTCTAATGGCTGGTTCGTTAACCAATAAGATACTGCTATGTACATCCATGTACATATAACTTGGAAAGGAATT
TCAACTAATATCACAGACAATAAATATGGAGTTAGCTTATACCAACGGTTAAAATGCTCTCTAGAGAGTATCTTCATTTCTAGAGGAAATGATAAAG
TTACGGGCATTTTACCAGTATATACTGTGAGTAAAAGAGTACCATACAAATATACATAATTTGCTAAAACTGTATCCGCGGCACTTCCAACGTTTCT
ATACAAATAGCCAAAGATCAGACCAATAACAATGTGGGCCAATACTCTGTTCAGCGATGGACCGTAACATCTCTTTATTATCAGCAGATTCCTCTTA
TACAATAACAAAAACTGCATAATGATTGCTGCGGGCACTGCTTTCGTTTCGTCGGAATAGAGACCTACTTTTTTTTCCTCGAAGAGATCTCTCCCTT
TTAAGTCATCGTTTTCACCACTTTCCATACTTTTATATTGATACGATTTGACACTATCTGCCAATTTGTTAATATTTGCATTGTGATCTCCGATTGA
TACTTCTATTAAGTAATCAGCTGGATTGTGGTATGGGGGACATTTTAGATTGAGTTTTTCGAAATATGGTACTACATCGCTAACCCTTCCATCATAT
ATGCATTTTCCTTCAGAGAGAGCATAGAGTTTATCGAACATTTCGAATATTAACGCTGAAGGTTGATGAACAGTGGCTATTACGGTTCTTCCTTCTG
AGGCGAGGTTTTTGAACAGAGACACGCATTGAGTACACGACAAACTGTCCAGGCCCGTTGTTGGTTCGTCGAGAAACAATATCGGTGGATTACTAAG
TAGCTCCAGTGCCACGGCCAACCGTTTCCTTTGTCCACCAGATAATTGTGCTGTTAGTGTATGTTGGCATTCTTCTAAACCCAGCATTTTCAGAATC
CCTGCAATCTGTTTGAATTTATAATCGTTAGAAACACGGTAACCTAGTTTTAGGTGAGCAGCAAACGTCATGGCCTCCTTCGCTGTCAAAGCCATTC
TAAGTTCTTCATCTTGTGGAATATAAGCTGATAATTTTCTATACCGGGGGCTCTGGTCTCGAACGACATCGTTTAATTTTACTGTTCCGGTGCTTCC
TTTGGTTATATATCCTGCTAACACATTAAGAAGAGTGCTCTTTCCTGCACCAGAAGGTCCCATTATTACAGATAGTTCACCTGATTTGAATTGACCG
GAGACCCCATGAAGAATTTTCTTCGTTTCTTTCTTGAACTTCGTAATGGACCATGACGTGGAGTAGAAAGTGATATCCTCAAAATTCAAGTCTACTG
CCGTCTTCTTGGGGAATACACTGCTCATATTTTCCTGTGTACTTTTAAAACGGTAATTGTTGTTGTAGTTGATGGCTAAGCTCTTGGTGTCTGTCAT
GTTGACATTTTCTGTCAACGGTATTGACGATGGTGCAGTGTAGTTGTGTGATGGGTTGATTTTGATTTTGTCATCGTCGATATAACGCGTAATTTCC
TCAGTTCCAATCATCCTGAAGATGGTAACATTGAACTTCACACAAAAAAAAAATTCACAAAACGTTTGAAATGGATACGTATCGGAGAAGAACACAA
CAATACTCGCGAGCGCGAGTTCCGGCAGATTCCAATATTTCAACCGTCACACAACAACTGCACGTTGAGACCAAACGAATGAGTCTG
>Dvv ABC-G D3712
GTCAWCACGCYCAAATGGATGACGTCACTAGTACGATATATATGTCAAAAAATCATAATTTAAAAATCGAATAACTTTTGTATATTACATTTTTTTT
CTAATTCTGTAAATAAAGCAGTTTACAAGGGGGTCAAAATATAGTCAACAACCCCGCACACTCACCGGGGCCGATCGACCGGCTCCGTCCCATCATA
CAGCCGACCGACCATAATAACCCCTACTTATATTTAATTTAGAATGTATGCACTGACCCCCAGCCCCAGCAAATGGATTTAATCACCTTCGTAGGAA
AGCAACCCTGACATCCTCTCAGCAACCCCCGTTCCACGCTTCGCCCGACCGACCGCAGCAYGCCTCTCCACACAATCACAGCAATCAACTGTGAAAA
ATTCAGCTCCAGCAAACTCAAAGATACTTTTCAGCGCTGCTTCTTGGACTATTAATCTTACAGTAGCGGAATGCTTAAAAACAGGCAAAAAAATGGGT
TCTGTAGTCACTTATTGGTTAGATCATTATTTGATATATTATGACAACTGACATATAATAGAGCAAAATACATCGTAGCCTGTAGAACCAAAGCAAT

GAAAGGTTTAAAGAATATTGTTTCCTACGTAAAATAATGAATCAATCAATTTTACATTTTTTTAGACGCTAAACATCATATAAATATAAAGAAAGT
TTACAAACAATAAATTCTGTATAAATCATAATACGTTTAATTTTAGTACGTCATAGAGAATATAAATAGAAGCAGTACTAGGAAAATAGTACGTTTA
AAGATGATGAAACGACGGTAACTACTCATCGGCCATTTTTTATTGGAATATCCTTCATACCAGTATCTAGAGCTGAATAAATAAAAGAACACAGGTT
TAATTCTAGGCATCACGCTGTTAATGAGTCAGAAGATAAAAAAATAGCTGCTAAGTACAATTCTGGCAAGAACAGCTGTATTTTTGAAATTCAGTAT
GAGAAATTGATTGTATATAACATAATATTACATTATTATTAAAGATACCAATGTACAACCTGCGTCCTTTCTATATAAAACCACGATTAATAACAAT
ATTTAACATTATTAAATATATCCTAAACTTAATGGTTTGATTTAAGTTTCCACTTCAAGAAGAGGTAAGCGGAAACTCTTAAAAAGAAGAATATCAC
AACTAGAGCCAAAATGTCGAACCAGTAACTCGATTTATCCATGTCCAACTCTTCGAGAGTTGTGAGTGCCTTTTTAAAGTGGCAATATTGAGCGAAA
CAATCCAGGTTTGGTCTGTTGTATCCGTATGTGGCAAGCGCAGTCCCTTCGAATCCGTATCTGATGTAGGACAGATACGTTATCCAGCGAAGGTAGA
TGGGAATAGCATCGAATGATACAAAGAATCCGGAGAACAACAGGAATGGTACTGACATTACTGGAGCTAAGAATACACCGTTCTGAACATTCATAGC
AGCTCCTACTACGAGTCCAACACTTTGGGCCACGAATGATACTAACAAGGAAGACAGTAAAAACATTCCAAATCTAGCCGGATCCAAGGGTTGGGAA
GTCATGAAATATACTATCGTTACGTAAAGTATACAGAAGATGGTCTGGAATGGCATATCTGATATCGTTATTGCCAGATAATAAGACCTCAGGGAAT
ACCATCGATTAAAATGTTCCTTTAGTAAAACTGGCATTTCAAGTGGGAAGGACAAAATGGTTATAGTCATGGAAGTGTACATCAAGAACAACATGTT
GAAAAACAGGAAACCAAGATTACTAAGTACTTTGGACCCATCGTTTCCAATTTTGAAGTATAAGGCACCAATTAGGAATCCAACCAAAATATGGGCG
AATAAACGTAGGTACATCAAAGTCCAATCTCTTCTACTGAACAATAACGCACGTTTTAAAATAATGAAAAATTGTTGGAATTCCGAATTACCATATC
TGGGCTGCTTAACTACAACGGAAGTATTTAATAAGGCACTATCAACATTAGCTTTTTCGACATCTTTAGAATCTATTCCAGTATCTTCGTTAACGTT
GTTTGATTTTCTATTGCCATTCACGGTACTGCCCTCTTTGGACAATTCATCAGAAATAAGGGTTTTCAAGTTATCTTTGGCGTATTGGTAAGCGTTA
TTCAAACCGTCGCTCATTTTGAAACCTACTATTTGGCTAGCATCTCGGATGTCATTCTTACCATTTTCTATAGCGTTAACTAACGTTCTCGTGTGAT
CTCCATATTCCCCACATGCTACTTCTATAATATAGGAGGCTGGATTGTGGTATGAAGGACACTGTAACTGCAGACTTGCCAAAAAGGGCACCAAGAA
TGTCGTAGAACCTTGGTAAACACATTGACCATCAGCGAGTGTATAAAGTTGGTCGAACATTTCAAACAGTCTAGCTGAAGGTTGGTGAATAGTACAT
ATTATAGTTCTTCCACCTTTAGCTAAGGTTTTCAACAGTGATATACATTGAAAACATGATGAAGAATCCAATCCACTAGTAGGTTCATCGAAAAACA
TAACTGGAGGATTACTAACTAATTCTAAGGCGATTGACAATCTCTTCTTTTGGCCCCCCGACAACCCGCTTGTCATTGTTTTCTTATGATCCAGCAG
TCCAAGTGTGTCTAGGATTTCGTTAATGATATCTTCTCGATCTTTTGGACTTTTGGCGCCAATTTTAAGTGCAGCTGCGACAGCCATAGCTTCATCC
ACTCTTAGATTAGCATGTAGTTGGTTATCCTGCATAATATATGCCGATAATTTTCGGAATTGACTAAGATCTCTATCACTGCCGTTCATTAACACAT
GGCCTTTTACCCCTATGGTTTTGTATCCAGTGAGTATATTGAGTAAGGTGGATTTTCCAGCACCAGATGGTCCCATAATAGCGCATAGTTCACCGGA
ACGTAGCATTCCACTGACATTTTTCAAAATCTTTTTCTCTTTGCTCCCTTGCTTAACCACATATTCCAAGTCCGAAAAAGCCAAATCAACTTTGGGT
CTCTCAGGTAAATGAACTAACTTTTTTGGTTGATTAGATACGATTTGCAACTTCACATCAGACTTATTTTTTAAGAGTCCTTCGGTGGCCATCTTAA
AAAGATTTAAAACAAATAATTTAGAGGACAAAGAAGCGACTACCGTTTAAATAAAATGTTCTAATCCACTAAAACACCACTAGAAAAGTTATCACCA
ACTAATAGTGTACCTAATCTTATCACACATACACTGCTAGCTCATACCGAAAACAATGATAATATTGAATTGATATTGGGTATTTATTAATCACCGA
TCTTCGAGTAACGAGGACATCAACAATAAAATTAACGTTCGAACAAATGAGGTTTTGATGATGATTATTTGCACTTGGAACTTCGACTTGACACTTA
GTAAGTGGGAACAGGTTTTGTTTTGAAATGTTTTTGTATTTGTCAAAAGTGTTATTCTTCCATTGACATGTACGTCAAAAAC
>Dvv ABC-G D14042
TGTATTCGTTAAGCCAGCCACTATAACTATCTCTGTCTGTATCTGTAGAGCAGGGAAGGCGAATAAAAGTTCCACAGCCCTTAGGCCACAACTGAGC
CATTGTGTATCCTTCTCAGGGGCGTAACAGAGGCCCCCGCGGCGTGAAGCTTGCGGTGGGGCCCCAATCGAGCAGGGACCCGATCTTGTCGTCTGAT
ATTATGTAAAAATCTGTTATAGGTAAAGAAATGATATTACTTTGTGTGTTTAAATCTAAAAACGTAAATTTCTAAATAGGCGTATTCGGCAAGTGAA
TCATCTCATATCTAGAAACTTAATCCGTTCCGGTCTCACGAACTTTTATGGGTGACGACAATAAACTATAATGATTTGTACGTAACTATTGAAAGAT
TTGAGAATTGCAATTTGCCGAATTTTAGAACAATATTTCTAAATCTAGAAATGGGGTTTCTCTTTATTATTTACCTACGTTTAGTATTTCGATACAA
TTATCCAGAAGCGTAACAGAGGCCCCCGTAGCAGGAAGGGGGCTATACCAATAAAGATGAAAAATGTAAGTCGTCATAAACAATGTATGAATACACC
AATAGCTCAGTAAATGACAGTTTTGGAGTGAAATTATCAGCTTCACGACGGACGTATTAGCTTGACAATTTGGATTTAGGTTCTAGTTACACTCCAC
TTCAAAGTTGGATTTATGCCGTAAATTGTTGGATTTATGTTGTAAAGTGGATTTATGCTTTTACTTCGGGGGGTGACACCCCTGCTTGTTAAGCTTG
ACACTAATTATAAGCAACTTTTGTTCTATAGAGTTTCTTTATCTACGTTAATACTTTTCAAGTCATTTGTGAGTGAAAATGTTTATTTTTCAACAAA
AAGAAACAACGTTTTCAAACGGGTTTTTTGCAAATAACTCAAAAATAAGTATTTATCGAATTAAATATTTGTAGCAAAAATGTAGCTTGTAAGATAC
AAAAAAATGGTGTACTTATTCATGAAGTCTATCGACACAGTAAAAGCACAGTTGAAGCTCATGAAAAATTATTCTTATTCGTCTAATTCCAAATCGA
ATATTTCAACCTGAAATAACCAAAAAATGAAGCAATTTTCGGGAAAATCTCAAAATTTGTTTTAAATGTTTCAAAAAAAAGTTTTTGTTTTATTTTT
TATAAAAGTTTCTAGTACCAAAACTATAAGAGTCACACTCAAAATAAAGTTAGTTTGGTAAAAAAAAATCGTGAAAATATGTCCCTATTTCGTACCTC
AAATAAAATTTATCATTAGCGCTTCACCATTTACTTTAAATATTTATTGTTTATATGATCTGTAAGTTTGACCGGTTCGAAGTTCTTATTTTTGAAA
AAATTTGGTTTTATAGTAAAAAAAGTTTTTATTATTTTGGAAAAATGCCTTTTTTTTTTCAAAATAACTTAAAAAGTATTAGTGATACGAAAAATCT
CAAAGAGTAAAGGAATGTAGGTCTTGCTTTTATAAATATTTTGGCTTTATTTTGTTTTTCTGTAAAACAAAAATTGGTTGGCTGTTTAAAATTTGCA
TATACTCACGATTAATGACTCGTTCAAGCCCTTTTAACTAGAACCCTTTTAAAAATAAGCACTTTGAACCGGTGAAACTTACAGGTCATATAAAAAA
GTTAAGTAATTTGTAAGGCGATAACGATTAGTTTCATTTAGGGTGCTAAATAGGGCAGATTTTCACAATTTCTTACCAAAAAAATGGGATTAATTGT
ATTTTGAGCGTACCTGGCTTAGTTTTGATGCTACAAACTTTTGTAAAAAAAAGCTTCTTTTAAACACTTTCTTTACTACTAAACTTTTCTTTTGCTG
AGTGGATAGACTTCATTTTTGTTTAATTTTTTTATAAGCTAGATACATTTTTGCTAAAAATATTCTTTTTTTTGACAAGGTTTTCCTAGCCATTTTC
ACATTCCTAAGCCATTTTAAGGTTTCTTTTTATTAAAAGAAGATTCAAAAGTATGTTTTTCACAGATAAACAATAAATAACACTTATTAATCTTTGA
ACAAACATTTATTTAATGATCAGGGAATTAAAAACACCACTTTGGTCATAAAAATATTCAGTGAAATGACTTTTGGTACTTTGTTATTTAAAAATTG
TTTAAATATGGTTTTTACATGGCTTAAAAGAATTGTAAGTGTATTAAATATTAAATTACTAAATATTCGCCATATCTTGATATAAAAAATTTCTACA
TAAGCTTCAAAATCCCATGTATGGTTCAGATAATGTAGCAGTGACAGAATTTATGATTTTAAACTAGACAATATGGAGATTAAATGTCTAAAACTGA
AAGTTTATTTAACACGTTAATGAATTAGGAAATCATTCCCTCACATTGTTTCCTACTTCTTTGTAAAGTACTTGTCAACTCACATTTCTGATCTCTG
TTACACCATCTGCTTGATTGGCGCACCAAAATATAACGAACTGTAACAAATATGTATACCATATTCCTACATAATTATAACATATAAGATCACCAGC
TCCTTTGAAATACTTGTGTTACTACCTACCCTAACAAATAATAAATTGATCTAAAATAATATAAATACATTGATTTCGTATAAAACTGTGACATTTT
TGGTTTACATGAAACTAGATGGGCAAAAATCTCGACATATCACTGGAATTACTATCTTACACTCTTATCCCAGGAAATTTCAAGAAGAATAAGTATA
CTCTATCTTTTGAGTCATCGTTACAGACGTTGGAAATTATCCATAATCCAACATCAGCGAAAACAAGTTTAAAAAGCCCAACAAAAGGTTTGTTTTG
ACATAATTGTCCAATATGCTGACAAACAATTGTCTAATATACTGACAAACCAAATTGCACTGACAAAAGCAATAAATGTGTTACTTAAAATATGAAG
ATTGTTACGTTCTACACTCTGTAGACAAATATTTTGGAATTAAACATAGATTTTTACGATTTTTTATACTATTTCTAAAACTCTACATCAAAGGTTC
CAAAATTAAAATAATGTATTACAGTGGGAAACAGTAACCCCATCTACAATTCTAACTATATTAGCTATAGTCAATCTGTAATCATAAGTTAATCTGT
TTCTAGAGATCAGATAAATATCACATTAGGCTACTAATAACTTCAACTTTTTTACAAATGTGTTTTTACCTCATCAAAATTCTTACTAATTCTTTTG
AGCTTAGCACCAGCTAGACAGTCTTTAATCGATAGATGTAGCTGACCCACTAGTATTCCTTACAGTATTTCCACTAAATTATCACTTTGTATATAAT
TTTAAATTATATACTGCTTATTTTGTAATAAACACAACAATATATGTCTCAAAAAACAGAGTATACCCACTTGATGCAATACAACAATTTAAACAGC
TAAATATTGCCCTTAATTGCTAATGATGTGTTGTCGAAAAGGTACGAATTATTGAAGAAAACATTTAGGAATACAAAATTGTAATATTAAAAATGCA
GTTATAAACTTTATTTCCATTGCTATGGTACGGAATACTGTTCCAGTACATACAGAAATAAACAAGACTTAATTGGTCAGTACAATATATAGTTACT
TTATTGCCCTAGAATATTTAAACTTTCGAAACAAGTTTTTACGTTATTATAGATCTAGAATATGTCTAGTTTCATTCAATAACTTTTCTTTATCGCT
ACAACAGCCCTGTTGGTTGTATATCGTTCTGAGGCGGTAATTTAGGTTATGTGAGGTTAACTCTGCAATAATTTCAACAGAGGCAATAAAGTATCTA
TTTGACAAGTGCCTAGTAACCTCAAAGTCCCATCCTATTACATAAATTCTCTATAAAAAGAAAATTTTACCGATTTTGCATCAATTTAATGCGAAGC

ATAAAGTACGCTATCACCCTAAGAAAGATGAACAGTCCTCCCAGCACCCCGACGTCGATGAAATAGCTGACCATGTCGTCTTTCATGGACATCTGCT
CCAAGAACTTCTTGGGGTACTTAAAGTGGCAGTACTCGATGTTGCACGTCAGCTTCGGTCTATCCAAGCCGTAGATGGCGATCATACACGCTTCGAA
GCCGTATTTCAGGTAACTCAGGTATGGCAGCCATCTCAGGTAGAAGGGAATGTCGTTCAGGTTTGTGAAGAAGCCTGAGAAGAGCACCATGGGAATG
GTGGAGATTGGTCCGAGGAATACACCTCCCTCAATATTAAACGCAGCACCTATGAGCAATCCGAAGCTTTGAGATACAAGGGCAGTTAACACTGTAA
CTAACAATATCATCCCAAATCTAGTGACATCTAACGGTTGGGATGTTATAAAGTAAACACCTATTATATAGCACAACGTCATAACCACCTGAAACGG
TATGTCCGCCAATGTTTTAGCCATATAGTACGCTTTGAGCGAGTACCAATAGTTCAGGTGCTCCCTTACGGTAACTGACATTTCCAAAGGAAATGTT
AAAATTGTGGGCATCATAGCTGTGTACATCATAAACATGACGCAGAAGAAGAGGCACCCTGCGTTACTGGTAACTTTGGCAGCATCTTGTCCTATAT
CATAGTAAATTAGGCCTATAAGACATCCTATCACAAAGTGGGACACCAGTCTCATCCTAGTTAGCGTTTTGTCCATTAGAATCATGTACATACTTCT
TTTGAGCAATATTGTGAACTGTTGGAGACCTGTCGTTGAAAAACCATTCTTTTCCGTTGGCGACAAATTCTCCGAAGAATCTAAAAGTGACGTTGTG
CAAGTGACTGGCGTTGTGGGGGTGGTAGGCTTCACAGAACCATTTGGCACGGATATTACATCTCCCGAGGAGTTCTTCCCGTTGGCTTCTTTAGCAA
TATCATTGGAAACTATCTTACTACTTCTATGATCTGGCGTGGCGAATTTGGTGCAACGACCAGCGTTTACTGCCACCACCAGTTTCTGAACGTAGTC
TCCATGCTCTCCACATGCTACCTCCATAACGTAATCAGCAGGATTATGATAACTGGGACAATTCAAGCCCATGCTTGACAAAAATGGCACCAATCCT
AGCACAGGCCCTCTATAAATACACTGTCCTTCCGCCATCATGTAAAGATGATCGAACATTTCAAATAACCTAGCTGATGGTTGATGTATAGTGCATA
TTATTGTTCGACCTCCTCGTGCGAGTGACTTGAGTAGGCAGAGGCACTGGAAACATGATGAACTGTCAAGACCACTCGTTGGTTCGTCGAAAAACAT
TACTGGTGGGTTATTTACTAATTCCAAACCTATTGACAAACGTTTCCGTTGACCCCCCGAGAGATTGGAGGAGTTTATATCGATGCAAGTTTGAAGT
CCTAAATTTTCAATAATTTCGTTGATCACAACTTTCTTCTCGGACAGTGTGACAGTCTTTCCTAATTTGAGATTCGCTGATACCATCATGGCCTCTT
TGACGGTGAGATGTGGGGAAAGACAGTCGTCCTGCATGATGTAACAGGACATCTTCCTGAAGCGTCGCAGGTTGCGCTCTTTGCCGTTGATCATAAC
CTGCCCACTCAGATTTGACGTTTTGTAGCCAGCTAAAATATTCATAATGGTACTCTTTCCGGCCCCTGATGGTCCCATAATGGCCGTCAACTCTCCC
GACTTGCATTTTCCACTGATGCATTTCAATATCGTTTTATACCCTCTTTTCCTTCCTTCGGATACCGAGTAGGAGAGATCTGTGAAGGTAATGTCCA
CGGGCGGCCTTTTGGGTAGGTGTGTGAGGGCAATCATTGGCCTTTTATGGTTGCTCGGACTGCTGTTGGGGACCTTACGGAGCGCACCCGCTCCGCT
GATGGTTATTGGTGAGGAACCATTGTATATGTTGGTTTGGGAGCCACTTAAAACAGAAGCATTATTAAGAGATTCGGTTTCTTTCAGAAGCGGGGGA
GCGTGAACAGTCTTGGTGTAAGCAACAGCTTGGTTTGTGTCTTCGGGATCAAGAGGAACATTTATCTTTATACATTTCGAATCGTCGGCCATCTTTT
GCGCTTTGTCATAGATACCTCACACGGTAGCAACGTCGATGATGTTCGTCGGTTTTTCTCATCAATTTTAACTAGGATAATTAGGAGTCATTTTTCA
TCAAATAATAAGGACACTAACCCCCAGCACTAGCACTCAACATCTACCAAATAATTCGACCAAATCACAGAAAATAACAAAAATTAGCTAAGAGTTC
AAGAAATTAACATTGTTAGAACAACTGTCAACAATCAGCTGTTCTGAACATCTATGACGTTT
>Dvv ABC-G D10897
GCTTGAGTCAACCACAATTCCGATAATCGAAATCTCATTTCGACAGGGTAAATGCTGTCGAAGTGATTTCTATTCACTATTACGATTGTAGAGATTC
CGAAGTAGTTATGGAATACGGATTTGGACTATTTATATTCGACTTGGCCACTTCTATTCGATTTTACTGACTTCTATCATGGAAATGAGTTACAGAA
TAGGCGTGATTGACTCCTCTAAGAAATGCAACCCTAAATGTAACATTTGAATGGAGTACCTACATACCATAACGGATAATATCATAACGTAAGTATG
AAATTATATTTCTTATTTCACTTTTAATTCCACTAGGTACTACCAAAATTGTATGAGAGTTCATGATTTTTACAGAATTGGCTCATTAATAGGTATA
ACCACTGTACTGAGTGACGATCTCATACCTTTATCTTTTAATAAATCGTAATCTTTGGAAAATTTTAAAAATATTATTGTAAATTTTCTCTTAGTTC
CTTTGATTGTCACAAAAAATGATGTTTTCTAAAGAAAAAAAAATATTTTTTTGAGTAAAATTTAAATAAATGTCCTTCTAACAACACTGTTTACCAA
AACCGAGGCTTCAATCACCTACTTCTAATAAAAGTTATTAAACCATAAACATGCTCTATCGCACAGATGCCAGTTTGTACCTCAGCACTACGTACGC
CGCAATTCTCAAAACAAACAAGAAAAGCAGTAAAGCAATGATGTCATTATCGAACTGATCCGACCTGACATCTACTATTTCCAGGAAAGTCTTCGGA
TACTTGTAGAAGCAAAATTTGTCTTCGGGACATTGAAGAATACCCCTGTCCATTCCGTACACCGCCTGGACTACTCCTTCCAAGCCGTACCTTAAGT
AAGAAGTAAAACTGCCGTAGTACATTATGGTCGGCAGATCACGTAGACGAACGCCAAAACCGGAGAACATCATCATGGGTACCATCAGGGTTGGTCC
AACAAAAGTTCCGTTCACAACATCAAAATATGCGCCAATAGTTAGACCAACTCCTTGAGCTACTAGAACTACCAACATGCAAGTGACGAAGAATATC
AGGAATCTGGCTTTATCGTTTGGTTGATCTGTCATATAGTATATTATGAGTGAAAACAGGAAGCAGCCAATAACCGAAACCGGAATATCCACTATGG
TTACCGATGTGTAATACATTTTTAATGAATACCAACGGTTGAAGTGTTCTTTTATTAAAATCGACATTTCCTGGGGAAATGTTAAAATCGTCAACAT
CATTGTGGACATCATATGGTGCAATAATATCGCAAACAATAAATTAAAATTATCCAAAACTTTAGATCCATCACTTCCAGCTTTCCAATAGAGGCTG
CCCAACATCAGTCCAACTATTATGTTTGTTCCAATACGAAGGTAGGTAAGCATCTTATCTCTGTAGGCTTTAATAAATCCCCGTCTCATAAGAACTC
CTAGTTGATGGAATTGTGATGTGGGAAAAGAGGATHCATCACTGCTARCAACTCGTTTAATAGGTTTGGGACCTGGGTAGTTACAAATGGTGCTTTT
CATCTTTTCAGGCTCGTTAAAATATTGGTAAGATTGTCCGTTCTGAATTCTTTCCCTCATATATTCAGGTTTCTCAGGTCCGTGCTCAATACATGCT
AGTTCTATTACGTAGTCAGCAGGATTGTGATACTGCGGACAGGGAAACCCACAGTCTTGCAAAAATGGAACCAGTTTATTCGTACTACCTTGATAAA
GGCAGTACCCATTTCCTACTACGTACACCTGATCAAACATTGCAAATAAAGTAGCTGAAGGCTGATGAATGGTACATATGATTGTTTTTCCTTGCAT
AGTTAGTTTCTTAAGTAGTGTTATACATGTTGAACAAGACGAACTATCCAATCCCGTGGTTGGTTCATCCAAGAACATGACCATTGGATTATTTACT
AACTCTAAAGCTATTGACAACCTCTTTTTCTGTCCACCGGAAAGACCAGCTGCTCTCACTTTTTTTGTATTTTCAAGATTTAGTGTTTTTAAAATAT
TGTTAATTATACTCCGTTTATTACTGTTGCTGACACTGGAAGGTAGTTTCAAATCTGCAGCTACCCACATATTTTCGTCCACAGTGAGTAAAGGCTG
CAGACGGTCATCTTGCGTAATGTAACATGACGTTTTCCTGAATTCCTTTAAACTTCTTGGTTGGCCGTTAACGTATACTGAACCTGTTACGCCACGT
ATCCTATACCCACTCAAAACATCCAATAATGTACTTTTGCCTGCACCCGATGGCCCCATTATGGCTATCAGCTGCCCGGGCTGGAAACGTCCTCCGA
CTTGATGCAAAATTTCTTTTGTAGTCCTTTTGGTAAATATATTCCCTTGCGTCGCACGAAAACTAATATTTTGAAACTCCAAGTCCACCGTATTTTG
GCACATTTTGCCATTAGTACCATCGTCACAAATCAAAGTCTTTTCTTGGTTAAGATTCACCTTTACGGAAAAATTCACTGAATCGCCGTTGACCCCG
TTCGATTTGACAACGGCGGCGGCGTCTTTAGATTGCCCTACAAAAGTACTCTCGTGCAGAACTGCCTCCATGGTGCGGATTGAAATTCGCGGACTGT
CAGCGACTGAACCAGTAAGTACTCTCGAAGATTGTGATTGCCTATCTATTCGTATTTGTG
>Dvv ABC-G D22358
GTGATGTTGGGTTAATCCATAATATTTTGAGTCCACGAATCTACAGCTAAAACATTTCCAATTAATAATGAAACACCCTATATTGTTTTGTTTTATT
TCATTATCTTTTATTTTGTAGTAATATTGATGATTTTTGTAATTTCAGTAAACTGTATTTGTTATTGTTTTTTTCCTAATCTTTGTAAGCTTTGTCC
ATAAAATTGTAAAATTTTTAGTGATAATAAAGCATATTTCTATTCTATTATATAGAAGCGTTTTTGCTTCTTTTGTTGAAATTTTTTTTAGCACTAT
GCCATATCAGTGGTGGAAAAAGGCCCGCGACATACTTTGATATGGGGCACCTGTGGAGCTAGCTACTGAAAGGACTATGAAACAAAATCAGCTCTTC
AATTTTTCCATAGAATTTTTTAAAGTTAGGTGAAAATGCAATGCTTCCACTCAACCCCGTATAATTCCATCCAAGCAAAAAATTTTGTTACCGGAGT
AATAGCAAACGGTACCAATACACGTACCCACAAACCGATGTAAGAATCTTGAAAATCGGAGCACAAATAACAAAGTTACAAGTTGTCAAACTTAATC
AAAAACCCTTAGTGGCGGAACCTTGTGCCGATGAGTGTATATAAAAATAATTCTGACTACAATATTTTACAAAGTTATAACAATGATCCTTTAAAAA
TATCACTTAAATTAATAGACAGTTACCCAAGTTCCTAATATCCTATATAATAATATCACTTAAAATATAATAAGATTAACTAAAGACATAAACCAAC
TTTCGCTTATCATCGCGCTAGTCCAAACTGTGATTTAACCAATCTTCCAATATATTGTAGAGCTCTAAAGGTTTTGTTAGGAGTTAATCTTTGTCGT
AATAGATAATAACTACCACCCCTGAAGAGGATTAAAATAAAAATTAGCACYAAAACGTCAACCCAATAAATAGTGTTTTCCATGCCCATGTCCCTTA
CGAATAAGTTGAGGTCTGTGTATATACAGTATTCTTCAGTATCCGGACAGCTTAGCTTTTCTCTATCTTTTAACATGGCGTGAATCAACCCTTCGAG
CGAGTACCTGAGGTAGCTAAAGTGCATGAAGAATTTTATTATAGATGGAATAGTGTCGTAACCAGATCCGAATCCGTAGACAGCTAGTAGCATGAAA
GGAACAGCAGATACTGGTCCCATAAACACAGCATTCACAACTTTCAACTGAGCAGAAATCAGAAGTCCAAGACTTTCGGAAATTACACTGGTCAAAA
TGCAAATGAAAAAGAACATTGACATCCTCCGAAACTCTAACGGTTGGTCAGTTAAAAGGTAAACACACGATACATACACAACTCCAAGACAAAGTTG
TACAGGCACGGTCGAAAAAGACAATGCTGAGAAATATGCACTAAGTCGGTACCATTTGTTAAAATGTTCTCTCTTCACTAACTGAATTTCTTGAGGA
AATTGCAGTAGAATTGGCATCATTGGAATATACATGAAAAATATAATGCAGCAGTAATAAAAACCGAAATTAAAGATGGTTTTGGACCCATCCTGAC

CCATTTTATAATATAGTGATCCTATTAGAAGAGCAAGGATTATATATAGGACAGTCCTCATTATTAAGTAAGTTTTATCCCTCCACATCTGCATCCA
TAACCGGGTGGTTAAGATCCAGAATTGGCTAAACCATGAAGACTCGTGATTAAAATCTGAAGAACTTAAATCCGCGTAAACTTCTTCTCTTATTGAT
ATTGATGAAATACTGCTTCGATTTTCTAGACTATCGTTGGCAAGTTCGACTTCTTGTAGTTGTGGTAAGTTACCGTTAGTTCTATAAATATTACGAC
CATTGTCTATCACAGATACCATCCGTTCTTGGAAATTACCATATTCGTTACAGCAAACTTCTATAATAAAATCTGCTGGGTTATAGGTTTTAGGACA
TTCAATTCCTACTTTACTTAAATAGGATACTACTTCGGGTCCGTAACCTTGATACACGCATTGTCCGTCAGATAAAATGTAGACGTTGTCAAATTGT
GAAAATATTTTGGCACTTGGTGTATGTATAGAGCAAATCACAGTTCGACCACCTTCGGCGATCTTTTTCAGTAATGAAATACACTGGGAACATGACA
GATCATCTAAACCCGTAGTTGGCTCGTCCAGAAATAGAACAGGAGGATTGTTTAATAGTTCCAAAGCTATCGATAACCTTTTTCGTTCTCCACCGGA
TAACCTACTTGTACCAGTCCTTTTTGCTTTGTCTAATCGTAGTAAACTTAAAATTTCATCTATCGCTGAAGATTTATCGCTTAGACTCAACGTGTTT
CCTAACTTTAAATTTGCTGCTATCATCATCGCTTCTTCCACAGTCAACAATGGTTGGATCATATCCTCTTGCATAATATACCTAGACATTTTCCTAA
ATTGTTTTAAATTTCTAGTCGCTCCATTTATCAAGATTTGCCCTGTCGCTCCCTGTGTTTTATATCCTGCTAGTATATTTAACAATGTGCTCTTTCC
AGCTCCAGAAGGACCCATAATTGCTGTCAATTGTCCAGCCTGGAAACTTCCATTAACGCTTCTTAAAATTAATTTGGATCCTTTTCTTCCTTGGGGC
ACAGTGTAACTGACATCGTGAAATTCGATATCGACAGGAGGTCTTTTCGGCAAATGCCTCAGGGATGTGTTCCTAGAATGCAATTCCATTTTAATAT
ATCACTAGCACTAAGTCATCTAACAAAAAACGTCAAAAACCGTACAGAAATCGAGAGGTAAACAAACGTAACGAAGTATTTAAAAGTATCGACTAGA
CTAGCACACGCGTTGCGCAGTTCACGTTTCAGTACAACATACGTCTATGCACATCTGTT
>Dvv ABC-G D23081
CGAGTATTTACGTTAGGATGCAATGTCAACATTCAGGTGCCACAACATTGCAATACGCTGCATCCTTTATGAAGAAAAATCAGCAAAATTTTATTTG
ATTTTGAGTGATTTGGTTCATAAATCCTTTCAATAAAATAAATTATTTATTTTGGCGAGTTTCCTCGTTTTAATATATTTTTCTGTGATAATTCGGG
ATGGAGCTTCAACAATTCGCGTCGAATAACTTAAGAATTAGTGATAAAATTCTAGATGTAGAATTCCAGGATATTTCGTTCGTTACCAAGGAGAAAA
ATGGTTATAAAAAAATTATTGATGGAGTCTCAGGAAAATTTCATTGTGGACACCTTACAGCCATTATGGGTCCCTCAGGAGCTGGCAAAACCTCGTT
GCTTAATATTCTCACAGGATACCAGGTCACTGGAACTACTGGTACCATAAAATGCAATAGCTCTAGTAGAAAGCAAAAAGGCGTCTTACAATACAAA
AAAGAATCATGTTACATCCTCCAAGATGATAGCCTACCAAATTTGTTTACAGTTGAGGAGTGTATGATGATTGCCAGTAAACTTAAAATTGCCAATA
TGGCCAAAAAAGCAAGGGAATTTTTGATTAACGAAATCCTAACCAATCTAAGCTTATTAAAGGCAAAGAATACTAGATGCCAATCTCTATCTGGTGG
TCAAAAGAAAAGGTTATCAATAGCTTTGGAATTAGTTGATAATCCACCGATTCTATTTTTGGACGAACCAACTACAGGTCTTGACAGCGCATCCACC
ACCCAATGTGTAGATTTGTTAAAAAAATTAGCCAATGGAGGAAGAATTGTAATCTGCACTATTCATCAACCCAACACCCAAACTTATGAAATGTTCG
ATCAGGTTTATATGTTGGCCAAAGGTCGATGTGTTTACCAAGGGCCAAGTACCAATACTGTTCCTTTTTTAGCTTCAGTAGGGTTGCACTGTCCTCA
ATACCACAACCCTGCGGATTATATTATGGAAGTCGTTAGTGGAGAATATGGTGATCATATAGATCAATTAGCAGTAGCAGCACAGGATAAAAAGTGG
CAAAATATACCTACAATTAAACTTAGTGATACTCCAGCTATAGACAGTAAAGATAATAATATCATATACTCCGACGAAAACGTTACTCTCTCTAAAT
CACCTTCAGAATGGAAAAGGTTTTTTATATTATTACAACGAAGTTCTGTTCAACTTTATAGAGATTGGACGATATCTCAACTAAAACTGGTACTTCA
TTTGTTAGTAGGATTATTTTTAGGTATAACTTTCCAAAATTGTGGTAGGGATGCCACAAAAGTGATAAGTAATTTAGGATTTTTACAAGTCGGTATA
GTTTATTTGGCTTATACGTCAATGATGCCTGCAGTATTAAAATTTCCTACCGAATTAGTAATATTGAAGAAAGAAAGTTTCAATAATTGGTACAAAC
TGACAACGTATTACGCTGCCTTTTTGGTTTTCGACATACCTCAACAGATGTTATTTAGCACAGTTTACTGTATTGGATGCTATTTTGTCAGTGACCA
ACCATTGGAGGTTGAYCGGTTTTTCAGTGTACTTTTTGTTCTAGTGCTAGCATCATTGTCCTCGTCAGGATTTGGACTGATATTGGGAACAATCACT
AATCCAATTAACGGAGTATTTTTTGGAGCAGTGGGATTATGCTTCTTTATCTCAGTTGGAGGATTTTTCATAATGTTCACGCATATGTCGAATGTAA
TGTACTTATTTTCGTATATTTCGTATATCAGTTTCTCAGTAGAAGGAGTGATGCAAGCTATATATGGATACGGCAGAGGCCAGCTTCATTGTCCAGA
AGAAGCAGAGTTCTGTCAGTACGTTTCATCAGAGGTGCTTCTCGAAGATATCGGAATGAGTAAACCTAACTATTGGATAGACATTATCTACCTTACT
TGCACATTCCTAACTTTCAGAACCATCGCCTTTGTAACACTTAAAAGAAAATTAGCTAATCCGTAATAGCTAGAAAAAGGAGTTATATTTTTAAGCT
TATGTTAAGTAGGTCGGAATATTGGGAAAAATACACAATATAATTATGTAGGACTGTGTCTTATGGTATGGTTGGAAAAATGAGCTGCTTATATAAC
ATTAAAAAATATTACCAATTAACAACATAACTAAATAATTTAAGTGGTAGTTGGGATAATCCGCTATATGCAGCGCGGATAATCCGGAGTGTGGATT
ATTCGCTCCATGAGTTTTCCATTACTCAAATGAATTTTTGACGTTTTATTGATAGAGCGTCACCGTAAATCACTCGACGGAAACTCTATAATACCGC
ATTTACACTCAGTCCTGTTGGACAGGGAATTGGGCAGGGAAGTGGGCAGCGTGAATGTGTAAATAGCTCACTTACGCTGCTGCTGCTCATACTACTG
CYATTGCACGGCGCTCCCTCAAAATTCGGTTTTAGAGATGGACGTCAAAGGACTGGCAGCGCGAGCGCGAGTGGGTATTCACATGCAACTACTCTCT
CTCTCACTTGCCGCCGGCAAAGCGGCAAGTGAGTAAAGGAGCAAGACGATGCTGTCGTCACATTTCATTCTTGAGATTCG
>Dvv ABC-G D13051
TTTTTTTTTAACTTCAAACATATTTACTATACACATAAAAATTATAAATAGTACACTTACGCTGAGAGTATATTATTATAAATATTAGTAGTACATC
TAATACTGATACAAATTCGGTACTAAACTTTTGGTCCCAGAGTAGTTCAGTACTAAAAGTACTGAACATGGGTATTTGAATCTCAGACATTTATTAT
CTAATATACCTACTAATAACCATAAAAACCCGTATAATTAATATTATTGTTTGGCATAATGTGGCATAAAATGTATGAAATATAGTTACTATTGAAA
ATAGATTTATTATAAAAAGTAAAATATCCATTATTTTGTTAACATAATATTTATGCTGACAATAACTAATAAAATTTAACAAAAATACGAAAAGGAG
TACAGTCAATGCCCCAAGACAATGTTTTATTAGTATAGTACAAACGGAATTCGGTCAGAATCTGCAGCCATTATTTAAGCTTTTAATCCCTGAGTCA
TATTACCGAAAACAGCGATGAATGGAAATATAAAGGTTATGAATTATAAACTCACAATAAGTTTTACGAACGTAATTGTTCAACAGACGAGGATTTA
CGAAGAAAAAATAACGTTTAGTCTGCTAATCACTTTACCCGAATTGAAAACATTGCGTTCTTCACACAATGTTGCCAAACTGCATTTTTTTATTTTC
GGTTTAAATTATCTAGCCGAATTACGAGTCTACTATACACCTTCTTACACCATCACGGCAATCACAATTTCATTCCTTATTAACTTATAGGAAAATA
AATAGACCTACAGTTTATAAGCCCATTCATGTTTCTTTCTGCTTGACAATTTTCGCCGCATAGTACACCAACTTGTTTCTTAGTTCGGAAGTCATTC
GGTACTTCAATGATATGTACGCTATGATTCTAAACAGGACAGTGTAACCCAATATGTAAGCAAACTGATGGAGTGGTACATTTTGGCTCATTCCCAT
GTCTTCTAATAACTTTTCTGGATTCTTGTAGTGACAATATATGTCGTTGCAATCCATTTCTGCTCTGTCGTTTAAAAGCGTCGATGATATTCCTACC
AGGCCATATCGTACATAACTTAAAGACATGAATGCTTTCATCACTGGTTCGATGCCATCTTTATAACCCATTCCGTACACGGACAAAGCCAAGAGAA
TAGCTAACCAGCATGGAGCAACCACAGATCCGCTCAGTATACCACAATTGGATCCAATTGCATATCCTAATCCCTGAGAACAAAATCCGACGTTGAT
GGCGATCATAGTAAACATTAAGAATCGTTGCCATTCTATTGGTTGACCCGTCATGAAAAATACTATTGATATGAATATCATTCCATAACCAACCAAC
ATTGGAATGTTTACTATCGTACTAGCAAGAAAGTACGGCTTTAAGCTAAACCATCGATTAAAATACTCTCGTTTTAACAGTTTGACTTCTATCGGAA
ACAGTAACACTGGCACCATGACATGAGTATACACGAAAAACACGTTAATCGTTACACAATACTTAAATATTGGTAGAACTTGACTGGCTTCATTTCC
AATCTGGTAAAATATTCCACCGACTAACGTTGCTGAAATGACGTGATGGAAAAATTGTATGATCCACATGGATTTATTTCTTTTCATCTGTAAAGCC
ATGCGACACAAAAGCACGGTGAATTGTGTCCAGAAAGACGTTGGGTATTCTATGTCGTGAATATGCATACCTGTTTGTGTGGTTTCCTGGTAAATTT
CATATATTCCTAATGTCTTATGACTTTGTAGAGGTTTTAACTTTTTATCTTTCATATTTATTTTACCGTTTTGAATTTGGTTTTGCAGTATTGGAAT
ATTGTCTTGATTGGTTTGGATTACTTCTATAATAAAATCAGCTGGTGTTGAGGTTTCTGGACATACACAATTTACCGACGACATAAAAGGTACTAAT
TGATTTGGCGACCCGTTATAAACACAATATCCATTAGCCATAATATATACTTGATCAAAAATTTGGAAGAGCGATGCTGGCGGCTGATGGATTGTAC
ATATCACAGTTCTTTCGAGCCTAGTTATTTTCTTAAGAAGATCTATACATTGTTTAATAGATACATTATCCAATCCCGTTGTGGGCTCATCTAAAAA
TATTACAGGGGGGTTATTAACTAATTCTAAAGCTACACTCAGTCGTTTTCTTTGTCCACCGGAAAGGTATTCAGTTTTGGTATCGTAACATTTTTCT
AGACCAAGAAGTTGGATCACTTCATGTACCACAGCCACTTTTTGTGTATGCCCTATCGACGCGCTCAATTTAAGATTAGCTGCTACCATCATTGATT
CTCTGACACTCAACCTCGGTTGCACTAGATCTTCTTGCATGATGTACGACGAGAGCTTGGTGAAGATTCTCATATCTCGAGGCTTTCCGTTGATTTT
TACACTACCTTTAACTCCCGCCGTCATATATCCCGCCAAAATATTAAGTAACGTTGATTTTCCTGCGCCAGAGGGACCAAGAATGGCAGTTAGTTCT
CCAGATCTAAACTTTCCATTGATCGATTTTAATAATTGCCTCCATCCCCCTTTATAACTGGAATCAGAAATCGAGTAAGTTAAATCACAGAATTCAA
CATCAATGGGTGGTTTTTTAGCGATTGTATTTAACTGAGTTAAAGGTTTAAATTTCGGCAACAAGACCTCCATTGCACTACCACTGGCTTCACTCAT

GTTTGTCAATTAAACTACTGCCGTTTTGACTCGAACACGCAACTCAACAATAAATCCTATGGTTCTGGGAAACCGAACTGATTTGCTGCAATCGACC
AACTAAATAGTTTCTTAGTTCGCGCGGAGGCAGTTTATAAATCAACTGGCAAGCTGTAAGGAACGTTTCAGAACTACGTATAGAGCAGAAGGTCGAT
AAATGCCGCGTGCTATCCGTTGGTAATGGAAGCCCATTCGTTTCGACGGTTTTTTG
>Dvv ABC-G T38769_c0_seq1
ACTCCTATCGTATAGCTGGCATTTTCCATGTCTTTGTCTTTAATAAACTGCTCAGGATATACGTAGAGACAGAAATCCTTGGACTCGCAGGGTAGTT
TTCCTCTGTCATAGCCGAATATCGTTAGTACCAACCCTTCAAGCGCATATCTGGGGAAAGAAATGTCAAACATCCATTTAAGTTGGTGCGGACAATC
ACGAAGTTGAACGAAGTATCCGGAAAACATGATGAATGGTAGAAGGAAGAAAGGACCTATTACTACGCCGTTCGTTACACTCAAACAGGAGCCCACA
ACCAATGCCCAACTTTGGGAGATAACCGAGACCAAAATGCATATAAACAAAAACGAAGATACTCTACTCCATTCCATAGGTTGTTGTGTACATAAAT
AGGTTATTGTGGCATACAAAACGGTAGCCAAAATTTGGACGGTTATATCTGAAATAGATACTGCTGCATAATACGATCGTATGGCATACCATTTATT
AAAGTGTTCCTTAATTATTATAGGTAATTCTGAAGGAAATGTGGTAACTACACAGTTGAACGCCGTCATCATAAGGAACATAACCGTAAAAAACATA
AAACTAAAATTGTCATTAATATTAGCGGCATCCAGCCCAACACCGAAGTACAAAATTCCTAAAAATAGCGCTATGGCAGAATGAGTCCCCAATCTGA
AGTACGTCAAAGTTCTGTCTCTAGAGAGCATTAGAAAAGATCTCTTCGTCAGCACCACAAACTGGTTTATCGTTGAAGTTGGATAAGAATTTAAGCC
GCACTCATCCCAAGTATATTCGATTTCGTTTACGATTGTTCTTTTTGTTGCTGCAAGAAAATCTTCTCCAATATTGGGTTCTTTTATTGAATTTAGT
TTCATTGTACTAGCACTTAATTTCGCATAACTTTCATTGATTCCATTATTCGACTTAGTTGAGAGAGCATGTGTAAAATCCCCATAATCTCCTGAAG
CTACTTCCAATAAGAAGTCAGCTGGATTATGTTGAGTAGGGCAAATCAAATCAAAATCTTCAAAATATGGTAAGAGATTATTCGGTGATCCTTGGTA
TATACATTTACCACTTGATACAACATACAGGTGATCGAACAATTTAAATATCGTGGCACTGGGTTGATGAATCGTACAAATAATGGTTTTACCACTT
TCCGACATTTCTTTCAGCAATTTAACAAGACGAATTGAATTTAAACTATCTAATCCACTAGTAGGTTCATCAAAAAACATAACTTGTGGATCTCTAA
GCAGCTCTAATGCTACAGAAAGACGTTTCTTCTGTCCCCCAGATAGGGCGTCTGTCCTGGTGTGTTTTATTTCCCAAAGAGTGACGGATTTCAAAAT
GGATGCTATCTTTTTATCTTTGTCAGTCTTAGAGGCCGTAAGCATTAAATCAGCAGCCACTCTCATGGATTCTAATACCGTTAGTAAAGGTTGGAGA
TCATCATTTTGCATTATATAACACGATTGTTTCCTAAAAGCAACATTGTTTCTTGGTTCATCATTGACTAAAACTTCACCCAGTATGCCTGTAGTCG
CATAGCCTGATAAAATATTCATTAGACATGTCTTTCCAGCACCTGAAGGTCCCATTATAGCCGTCAATTGCCCTCCTGGAAAATTTCCTGTGATATC
ATCCAGCAGTTTCCTTTTCGTTCTTGAAAATAATATTCCGTCATAAACACTGTAAGATATATGTCTAAATCCTACATTTATTAGAGACGCTGGATCA
TTAACTTGATTTTCTGGTAGTTCCAAAGAATGCAAACGGAAGTCTTCGTCATGTATCTTATTGCTCTCTTGAGGTGGATCCACAGTTCCTAAGATAT
CACAAACTGTATTTTCACCCATTTTAGTTCTGTTGCATCGTTTTTATTGAATATTCTAGCACAATCCACAAACTGTTCAAATATACATCGACACAAT
CATTGAATTTGTCTATCCAACCGGCAGGTAC
>Dvv ABC-G D13829
AAAAAGACATTCGTATTTTGTGTTACGTGAAACGTCTGATATACATGTATGTTGTATTTTTTGTAATCGTTTTATACATATACTCGTATATTTTATC
TGTGAATTTTTGTCTATTATGGCTGTTTACAATCGTGGAGGCAGGGAAGAAGAATCCTCAGCCTCAGAATCGTTATCGTTAGCAGCTAGGCTAGGTT
TAAGAGTAGTTAGAGATTATGGGACGGATTTTCGCTCAGTGGGAAGAGTTATTCCAATTGAGGAGCGGATGAAATTTACCTGGAACGATATAAATGT
ATTCGCAATGACCAAAACCATCAAAAATAAATTTTGTTGTTTCGGATTTATGTCTCATGGTTATGATGAACACATATTAAAAAGTGTTGACGGAGTA
GCGTGCTCAGGAGAGTTATTAGCCATATTGGGCTCAAGTGGGTCTGGAAAGACCACTCTTTTAAATGCTCTAACATATCTACCAATGGATGGAATGG
TCGTTTCAGGAATGCGGTGTATAAATGGTGTGCCCGTGGATGGTCAGAGGTTAAGAAACGTGTCAGCTTTTGTTCAGCAGGATGATTGTTTTATAAC
AACCTTGACAGTTCGAGAACATTTGGTTTTTCAGGCTCTGGTAAGAATGGAACGCGGTGTGTCATATAATCAAAGACTGATAAGAGTTGATGAAGTA
TTGTCTGAATTGCTGCTTAAAAAATGCGAAAATACCATAATTGGCTCTACAGATAAAGGAGGCATCTCAGGAGGAGAGAGAAAGCGGCTGACGTTCG
CCTCAGAAATGCTGACAAATCCGCATATAATGTTCTGTGACGAGCCGACCTCTGGCTTGGATTCTTTCATGGCGCTACAAGTCGTCCAAGCTCTAAA
AGCCATGGCTCAAAATGGGAGGACCGTTATTTGTACCATTCACCAACCTAGCTCCGAGTTATACGTCTTGTTTGACAAAATTATGTTGATGACTGAA
GGAATTACTTGTTTTCTGGGAACCAGAGAGGATGCTGATGGGTTCTTTATAACCATGGGGGCCGGTTGTCCGAGAAACTATAATCCTGCAGATTATT
TCGTAAAATTGCTATCAGTTATTCCAGATAGAGAAGAATCCTGTAGACAAGCCATTGCTTTAATAGCGGACAAATTTCACAATTCCAATTTGGGTAG
GAGGCTAGCCGCAGACTCTGCTTATATAAGAGCTCAGGAAGAAGTCGACAACGCTGTATGGCTTTCTAATAACTTGAGACCATACAGAAATTCATGC
TGGGCACAGTTTCGCGCGGTTCTGTGGAGATCGTGGATAAGCATGTTAAAAGATCCGCTGATAATCAAAGTTAGGTTTTTACAAACTGTGATAACAT
CTTTACTTATAGGAATGATTTATTATGGACAAGAACTTAACGAAGAAGGAGTAATGAACATTAATGGTGTCCTTTTTATATTTCTAACCAATATGAC
TTTCCAAAATGTTTACGCTGTAATCCATGTATTTACCGCAGAGTTACCAGTGTTCCTTCGAGAGCACAGAAGTGGCATGTATCGCACAGATGTATAT
TTTATCAGTAAAACATTAGCAGAACTACCCTTCTTTATAATAATACCTGTAGCGTTCACCACAGTATGTTATTATCTAATTGGTCTAAACGGAACAC
TAACGAAATATTTTATAACCTGTGGAATTGTCATTCTAGTTGCCAATGCTGCCTTAAGTTATGGTTATTTAGTTTCTTGTATTTCAAGAAATACATC
TATGGCCTTGACGCTGGGTGCACCATTAGTCATACCTTTCTTGCTTTTCGGTGGATATTTCATGAACCTGGGTTCTCTACCACACTACCTCAAGTGG
TTGTCTTATTTTTCATGGTTTAGATTCGGAAACGAAGCCCTTATGATAAATCAATGGAAAATATAACTGATATTGATTGTTCAACCAACAGCACGA
TCTGTCCCAAAAATGGACACGTTATTTTGGAAACATATAATTTCGCAGAGGAAAACTTTACTGTGGATATTATAGCACTGTGTATGCTTATACTTGG
ATTCCGATTCTTTGCATACTTAGCCCTGTTAAACAAAACATGTTCGTGTTAATAAAAACAGATCGTTTTTTTTTATATTATATTTATTAACGTATTC
CTTAGCATTGGGAAACGAAATTTTTATAATTTTTTTATGAAATTTGTAGCTTAGATATATTGACTGTCTATGTTTATATAAGCATTTTAAAATAAAC
GATTTTATACAT
>Dvv ABC-G T49457_c0_seq1
GTCGGTAGGGCAAAAAATTCAGTACACCTAGAACGGCCTCGGTGACCTCCATTCACCATTAAAGCATTTCCGCTTCCCAATGAAACACCCTGTATAA
TATTAAACTTATCATTTATATCAGAAATAAATAGATAATAATAAAATATTTCAAATATTACAGATCTTATGTACAAAATATGTCTAAATGTTAACAA
TTTAATATTTATTGGTAATGCTAGAGATATAAACAACATGTTTCAGCTATAATATCTATAAAACTGTTTTAGTGTTTTTCAAACAAACTCATAGCTT
TTTCTGTGCATTGGCACATTAACTTTTATATTTTTTATAGTGCCCTTTCTTCTTCTATAACTTTACAATCTTCTTTACAATTTCTTTAGTTAAAATT
CGTATTACAGCAACCTATTAACTATTTAAAATTTCATGGTTTATTTGCATCGTTCTTAGTGTCTTGTCTACGCCTTCGTCTGCTTCTGCACAAATTG
AGAGCGAACATGAAGCAAGAGATCACGAAGAAGATACAGTAAAATACGCCCATTGCTATAACTGGGTTACCATAGTCGAAGGTTAAGTTATCAGACT
TTAAGTAACCAAAGCTTTTTAGACTATTTGTTCCATTTGGAGGTGGACATGGCAATTGGACTATAATATCCTGGTGTTGAATCTGTTTATTTCTGCA
AAGGGTGGTTATTGTCCCTTGCAAACTCTGTATGGCTTCTGTAGAAAGTTCTCTATTTAGTAAATATGGTATCAGCCACTCGCTCGGGCTTATTTTT
TCAATCCATTGTAAATACATTGGAATATCTTTGAGGTGCAATGAATATCCAGCACTCAGGAAGAAAGCTGAGAGTACTGTACCGCAGAATATTGTGG
CTGTGTTGCTCAAAGGGACTGTGTAGATAAATGCCATCAAAAATATTTGGATGCAGCTAAGATATAACAACATTACCCCTATATAAATATAGAAACC
ATCGTAATTATTCAAATGCTGCATATATAAACCAGTCATAGAGTAACTAGGGACTACATAGATCAGCCATACAAACAATGATGGAAATATATTGATG
ATGGATTTTGTTATGATATAGATGACTCTACCATACAGTCCATCTTTAATATCTCTCTCAATCACCTTTCTATTTCTTCTAACTTCATTGACAGTCA
TCGCTAGCAATAATGGCCAATGTACAATGCACATGACGCTGTAGTGGTACCCGTACCTGTCGTTTAATATAAGTTGATGATCGGTGGAAGGGATATC
CCAAAATATGGCGCCAAGAATAAGGGATAGACTGGCTGAAAGGACTATCACTGTGAGCCAGCTGAGGAAAGTCGCGGGTTGGGTGTAGATCATGGAT
TTCGTAAACAAAGCAAATGCAGCTACAAAACAATTACAGTTTCGTACAGTGAGCGGCAAAGATGAAGGTGGTCCTGGGTCGCTTAAGGGTTCCTGTT
TTTGTCTAAATATTTCAGCCAACTGTTCAATTCGTTGTGATGATTCTAACATAGCCTCCGCTGCGAGATCGTCCAAGGTTACGAGATCAAGGTAGTA
ATCCGAGGGATTCTTAAAAGCCGGACAGGGATACTCAACTAGAGCAAAATATGGCAGCATGTCCCGTCGTCTTCCACTATACATAGTCCTCCCCGCC
GACACCAACAAAATCCTAGACAGCATAGTAAATATCTCGTAAGTGGGGGGATGCATCGTCAATATCACTATCCGCCCCAAAGTACTACCTGCCCCAC
CGCTCGCCCATTGTTTTAAATATTCGACTAAAAAAAAGGTATCGAAAATGTCCATGCTTTTGGTGGGTTGGTCTAAGAGTACGATATCAGTGTCGAG
GATAAGGTGACACGCTACGTTCAGTCTTCTTCTCTCTGAGATTGTCATCATAGATACCTTGGTGTTTCGGACTTGCTCTAAGCCTAGATCGTCTATT

AAAACGTTTATCCTATCCATAGATTCAATTTTTAAATACCCTAATTTCTCTGTTGGTTTCTTAAGATCATAATGTAATCTTAAAGTCTGTACAACTG
TCATATCCTTACATAAATTCAAGTCATTTTGGACATAAGCTACTCTAGATTTCAATGTATGCGCTCTGACGCTATGTCCATTTAAAAATATTGTTCC
TAAAGCTGGTGACGAGAACCCAGCAATGATGTCCAACAAGGCGGTACCTTCTTCCTCGGAAGTCGTCATAATGGCCATTATTTCGCCCCCTCGAATT
TCAAACGATATTTCGTTAAGAAGAAGGTGTTGACGGGATTTGGTGCAAGACTGGGAGCTTGGGAAAATGCCTAATCCATGTGCCTGTAAATGTGGAA
ACATAAACGTGGTTCCAGATTCATCATCTTGATTGAGATAATCGGTTCGAAGAATAGATGGTGGTCCTCTTCTTGAATGAGGCACAGAGAAATGACT
TTCAGTTTTGGACATTGGTCTTCCCATATGACCCATGTCACTATGGTCCACGACGTCATGATTGTGGTGTCCATAATCGGCAACACTGTTTCTTCGA
GGATCTATTAGATGGCGCTCACTATTATGACAGGATCTATTATAACCTAATGCCAAAAGATTGGCTTCCGAAATGCTCTTATTTCTATTCCATTGAC
TGTTTGAATGTCGTATATCACCTTGGTGCATGGCTGCCAGTGGGATCCCAGGACTTGTTTGCTCAGCGTGGTACACATTTTGCATTCTAAAGTCCGG
ATCCGATCGGTATTTTCGATTTTGCTTATAGTTGTGGCTCGTTTCATTGTCAGAACTATCATATCCAGAACTTCCATCTCTTGGTCCATGACTTCCT
GCTTTCGAACGCCGAATGTTATGCGGTCTAGTGGACGAGTTTCGTCGAAAATGTTGTGGAGTACCAGACCGATTAGAACCATTATGGTTCGGTGGGA
ACACCCTCGGAAGGCCGAATTTTAGGTAGGTGTACATGTTGGAGCCCAGGGCTGACTTAGGACCATAACGAGGATGAGAAAGGATAGACTGCACCGT
TGTATCTCGAAGCTGAAAATTGCCATAAGGAAGTGGACTCTTATCTGTGGATCCCAAAGCACTATCAGTGAAATCTGAATTTAGATTTTGCCTATAT
ATCGACCAAGCGTGCAGATCTTCACTGGTAGTTCACCCGAAAAGGCTCTTGATTCCGGGTTAGAAGGTACCGAGTATTTCCTCTCCATTTCGAATC
CTTGCGAGCCCCCTCCTTGGGAGGGGGTGCGACTACGCATCATGGTCTATGTTATCTCAGACCTTTCATTCCCATCTCTGTAAACCCTATTAACTGA
ACAAAAATAAGTAAATATTGGAGTTACAGTTAGTCCAAGTACTAGAATCAACAACTTCGTACACTAACCCGGTACCACCAACAATCATCCAATGAAC
CGGTTAAACACAAATAACACTACACAAACGACAAATGGAGATGTTTTGGTACTACTTTAGTACCTTTAATTAGCTTATGTATGCGTGCACCAACACA
AAAACCGATATGATTTAACAAAGTGAAGTCTGGTTGTCTGATGTTATTTTGGAATCTACTACGCAAACACACAGTACTACCAGTACCACAGAATG

>Dvv ABC-G T36869_c0_seq1

AACTTTATTTAAACCTTAATCTATTCCCTAAACTTTGCCTTAATGAAAGAAGGCAAAGGAATCAGATACAAAAACAAATTAAAAACGATCATTCCCA
GTGCAAACCCACAGGTGAGTCCAACATTGAAATCTTCTTCCAGAACTCCGCTAAAAATGACCTCCGTGGGATCTCTGCTGTATCGTTCCGTGAGGAA
GGCTTGTCCACTGGCGTATCTGCAGTAAGCATTGTTGGCTCCGTTCAAGATAGAAGTTTCGAAGCTGATTTGAGTGCAGTTGTGCAGGAGATCAAAA
GGAAGGGCTTTGCTGAGGGCAGAGTGGGTAAACACTTTGCGATTTAGATAGGCAGCAGCGTAGCGTGCTTGAGTCCCGTAGGTGGCGTAATACAGCC
ATTCTGGTAAGCCTTTCATTGACCTTAATATCCCGCTGCCAAGAGTAATGCACACCACTGTCAGATAAATGCTAAATATAGCGGCTTTCATGTAATC
CTTCACAATCATTAAAATCGCCATGGTTTGCTGTTCTGCGAAGATGTAACAAGCCCATAATATTAGAGTGAAGTACACGAAATCGACAGGATTGTCG
AATGATATAATTAACGGGAAAATTATAACTGCAGACAAAACTGTCGAAAGAAATGAGAAAGGCAGCGAAACCAAATTGTAGGTTAGCAGAAACAACG
TGCCTCCATATAATCCTTCTTGAGTATCTTGGTAGTACCTCGTCCTATATATAGGATAGATTAATATGGTGTTTATTATTCCAACGAAGTAAACACA
GCACAGACAGTTCAAAATTAAGCCATTTCTTGATATGAATGTATGTTGCCAATCCTTCATTTCTCTATAGAACAGCCACATCAAGAAGAAATATAAA
GGTAGCGCAAAAGTACGCAAAAACGTTTGTCGCAGTCCAGCTCTTTTAAAACTCACAGTTGCCGCTAAAAGTCGCACGTAAATCGTCCACCCAGTTG
AAAACCTTCCAGGTCTTCCATGCATGAATGGAACCTTCTCGTGTTGTCCATGCTCGTGGTTCGGATTGGGCGAGATCATGCTTGGACTTTTCTGGAA
AATAACTCCTTCGTTTTTAAACTTCTCCACTAATGCTGCTATTTGGTAATTGCTTTCTACGAATCGTTCTCTGGATCTTCTGTCCACTGTAGAGAGA
CAAAGATAGTACATCAAAGGATTTTCCAGTTGTGGACAAGGAAAGCCAATAACGTTGAAGTACTCGAGCATTTGCTTGGTACCTCCAGTGTAGACGA
CATCTCCAAGACACAAGTACAACACCCGATCCAAGAACGGGAAAACATCAGATCTAGGCTTCTCCATCGTCAAGATTATAGCAGTTCCATACTTTTT
AGCTGCATTTGAAAGTATTGAAATTATTAGGTATGTGTTTAGTGGGTCCAGATCCCAGGTGGGTTCGTCTAGTAATAGTACCACTGGGTTTTTGATT
AGTTGGACACCAATCATCAGTCTTCTGTATTCACTTTTTGTCAAATCTTCAACGCATTTTTTAGCGACTTGTGACAATGCTAAATCTGCTATCACTT
GCTTTACTTTAGACATTTTTAAATAGCCAGTGAACTTAGTCGGAGTATAATACAACGTCTGTTCAACATTTAGTCCTGGTATAAAATCACATTTATG
TGTAACATATGCACATTTCTGCTGGAATAAACACATACTGACTGGATGGTTGTTCAGATATATCTGTCCTCTTATAGGGCCTTGTGCTCGTCTTGAA
ATAACATCAAGTAAGGCTTTCTTCCCACTACCTTTTGATCCCAGCACTGCTAGGACCTCACCAGAATGTACCAACATGGACACATCTTTTAAGATTA
CTGCAGTTTTAATGTTGCCAACCATTCTTTGAAGAACAGAGCGCTTTTCTACCTGTCCAGAATAAAAAACATTGCACAATTCCAATATATAGCCACC
AGACAACATCTTCAAAGATTAATATATAAATAAATAAAAGAAATGATATTATTGAACGAATCTGTCAAAACACAACTAAAGCAGTACCATACATGTA
GTACCCATCCTGAAACTGAAGTTATTAAAACGAGTACCGCTCGCGCGATAGTCAAGCCACGTAGACCAGAACCACATTCGCCTCCCCAACTTAACAA
TTACGATCTAAATTTGACTGGTCAACTGTAATTGCCCAATAATATCGTTAATTACGATACAAAATGGTGTCTACCTGTACAGGTAAAAA

> DvvABC-G T79525

GGTTTCTGTACAAGCAAACCACAAAAAGTTTCATACGAGGTAAATTTACACAGTCCGCAAAGACAATATTTGTTTTCTAGTATATCAAGTGTATTGT
TTAATATTCAAGATAAAATAATAGAAAGTCGCGTCTGGCGTTGTTTTTAATTTAAAGTTTTCAATTGTTCTACAACTGCAAATAAAATGCTTGATGT
TGAGGAGACCCCATTGAATTTATTTCAAAGCCATGGATCGTCTGGATCGTATAACGAATTTGACATATATGCCTCAGCTAGTTTGGCGAAGAAAAGG
CAACGTACCTACTCACGATGGTCACCTATTGAAGAGGGAGTAACTTTAGTTTGGGAAAATCTTTCAGTATACTCGTCAACGACGAAAAATGGCCAGA
TGCAGCACAAACAGATAATTAACGGAGTTACTGGAGCTGTTAAAGCAGGCAGTTTGGTAGCTATAATGGGATCAAGCGGTGCTGGCAAAAGTACTTT
GATGACTGCTTTAGGATATAGAACTGAAGGTTCAATATTAACTGAAGGCAACATTCTAATAAACGGTCGACAAATTGGGGACTATATGAAATATTTA
AGCGGATTTATGCACCAAGAGGATATGTTGCTATCTTATCTTACAGTAAGGGAACACATGAATATAATGGCCAATTTAAAGCTTGATAGAAGACTAT
CGGGAAATGATAAAAAACAGCTAATATACGACATTCTCAGACAACTTGGACTAATGAAGTGCATTGATCTTAAAATTGGAGGTATAGATCAGGCCAA
ATCATTATCTGGTGGAGAGAAGAAGCGATTGGCATTTGCCACAGAGTTGCTTACAGATCCACCTCTTTTATTCTGTGACGAACCTACAACTGGTCTA
GACTCGTATTCTGCTCAGAAATTGGTGGTAATTATGAACCAAATGGCCATTACCGGAAAAACGATTTTGTGTACCATACATCAACCTTCTTCAGATA
TATTTGCCATGTTTAGTCAACTAATTTTAGTTGCTGATGGACGAATAGCTTATATGGGTTCTACAAACAATGCTCTTGACTTCTTTGAGGAGGATGGG
TTATGTATGTCCAACATCTTACAACCCTGCTGATTTTTATATTAAGACGTTATCCACAACTCCGGGTTATGAAGATAACTGTAGACAAACAGTTAAA
AGAATATGCGACCAATTTGCTGTTAGTGATGAAGCAAAAGAAGTCGAAATAGTAGTACAATATGAACTTCATATGGGCAGAGTTGCAACACAAAGAA
AATTTGAATTGCGTCAAAATTTCAAAGAGATTCGTTGGCTGTCAAAATTGTTTTGGTTAACGTACAGATGGATTTTGGAAATTTATAGAAACCCGTC
ATTGGAGGCAATGAAAATTGCACAAAGAATGCTGATTGGTTTCATAGTCGGGTTCTGTTATCTTGGGACAGATGCCTTGACACAGAATGGTGTACAG
TCTGTGACAGGAATAATCTTTATGTTTGTTAGCGAAAATACATTTAATCCCATGTACAGTGTACTGCATCAATTTCCTTCTTACTTACCCCTATTTT
TAAGGGAATACAAGTCTGGGTTATATCATCCTGCGACTTACTATTTATCAAGAATTTTATCACTGCTGCCAGGATTCATATTGGAACCATTAATATT
TGTAATAACTGCTTACTGGTTAGCTGGACTACGTACTAGTGGATACGCATTTCTAATGACATTATCGGTCATTGTTCTGACTATGAACGTTTCATCC
GCGTGTAGTGTCATGTTTTCAAATGCTTTTGATTCTGTTCCTACGGCTTTAGCCTATTTGGTACCCCTTTGACTACGCCCATATGATTACGTCTGGTG
TCTTCGTAAAGTTGAGTACCCTTCCACTTGTATTTTCTTGGATAAAATATCTCTCCTGGTTAATGTATTCGACGGAGGCGTTGTACATATTGCAATG
GAAAGATGTTACTAATATAACTTGTGTCAATCCACCAGAAGGTTTCTCTTGTATGAATAATGGTCAAGACGTTTTGGCAAGATATGACTTTAAAGAA
ACTGACTTTCAATTTGACATCTTTGGAATGCTTGTATTGCTATTATGTTTTTATATTTTGGGATACTTTTTCTTATGGAGACGTGTAAGATATCATT
AA

>Dvv ABC-H D20789

GCAAACGTGGCTACTTTGTAGAAAAATCAACATATTCACTCGCAAGCAACTCGAAAACTATTAACCYGGTAAAAGAACTTTATAGAACAAAAGTTGC
TCGGAATTAGTCATTCTATCCATTTCCGGACTTAATTCGAACATATATTTTTCACCCACAAAAGGGGGTTAAACTCACCCCTAATGCAAAAGCACAC
ATCGGCACAATATCACTTCTTTTCTTTGACCTGTTAGCCATGTGTATGCCAAATTCCATGTCAATCCTAGAGTTCTTTAAAACTTAGAGGTTTTGCA
ATATTTTACCTTTAAAGAACGTAATATTAGTAAATACCAACCGCTAGCACTACAATAGTTTCTGTATTTTTCAATAACTTTTGTCGTTGAAAATATA
CAAATATGTAAATATATAATTTATTTTCTACAATCACTCGAAATTGTACCTATGGTAAAATATGTAATGCCAAACTTGATTAAATTATTTTTGTTTT

TTCCTATAAAAATATTTACTGACTATCGCTTTTTCTTTATCAAATACACACTCAAAACTCCGAAGAATACTGTCCACATAAACCCTACACCCATCCC
GCTGTATACTTGAAAGTTGTCAATCGACCAACCTCTTTTGGATACGTTTCTCAATGATTCGATGGCCATCGTAAATGGAAGGCATCTGGAGAAGATT
CTTAAAGCTGGAGGCATTCCCTCTGTTGGCCACATCAATCCACTTAACATCATCATTGGAAGAAATATTCCTGTCAGAACAGTATTGGCCATGGAAT
GATCAGTGCTGATAACAGATACCCAAAAACCATACGCCATTCCACAAATACCTTGAAGATAGACCATCACGTACATTAACCACAAACTTCCTGAATA
TTCTTGTTGGTAAATCAAATAGACGACAACCAATAGCTCTGCCGTTTGAATTATGCAAATCGAAGCCTGCAACACCAAATGCGTAATGGTAATCTCT
AAAGATGTTACTCCTGCAACAATGGATCTGTCCCACACACCGTCGTGTCGGTCGGTAATGATTATCTGGGATGTCATAATTGCACCCATAAAGAACA
TGATGGTAATAAGACTTCCCGGTATCATAAAAACTGTGTACGGCTCATCTCCATCTCCATAAATGAAATTTATGTTAACAGGCAAGTCGCCAAAGCC
AGGAACAAAATCACAATCATCAAATAAGCTGTTTTGGAATTTGGTGTATAAATCAATGAGCTTGTATTTTAAGGTAGCTCCTATTTGGCGGTTTGAC
ATATCCATCCAAACTTTAATCTCACTCAAAGACAAAATATCTTTCTCGATATCCTTCCCTTTATCGATTCTTTCTTCTAAAAATGATGTAAAATTCG
AAGACATGTACATAGCTCCTACTATCTTGCCATGCAAGACTGCATCTTTGGCATCCTCGAGAGTCTCAAAATGAACCTTTTCTATCATGGGATGGTC
CAGATACGATAAAAACCGACAGCTGATATTCCTAAGCTGACAGGCTCTATCGTCTGTAGCAGTTGCCGTACCGTTAAATGAAAAGCCTGGGCATGTA
ACGCTCATCGCTTCGTCGTTAACGATTCCCAAGGGTATGCTTCTAATGTCTCCACCTACGGCTCCCATAAATACCCCTACTTGTAGAATTGGAAATG
TCATTAAAAATATAATACCCGTGATGTTCCTATAAAATTGTTTAAATTTTTATCAAACAATGCTTTCATTCGACTTTTATTTAAGGCATTTCTTGC
CTTCAGTATTTTCTTTTTGGCTAAAATATCAGTGGAACCATGACCGATTTCAAATGTAGAAACCGACGTTGCTACGCTAGTGGTGGATCCTGTGGTG
TCATTAGCTAACATAGAATTATTCTGATCATCTACAACTCGATGTGAGGTTAGTTCTTGTAATCGTCCTTCTTCCTGTCTTTTGCTGAGCAACAAAA
ATACGTCTTCCAAAGTCTCACTATTAAACAATGTCAATAACCTAGTTGGTGACTCCTCTGCTAAAAGTTTTCCCTCTCGCATTAGACCAATCTTATT
TGCCTGTCTACATTCCTCTATGTAGTGCGTAGTAATTATCACAGACGTATTATCCTTTTTAGTAATATCTACTAGATGCTTCCATATACGATCTCTT
AACACTGGATCCACTCCAACCGTGGGTTCGTCCATTATTAAAAGTTCAGGTTTATGGACCAAAGAAGCTGCTAGGGAAACTCTCCTTTGTTGACCTC
CACTACAATTTTTTAAATACCTATCATCTGGAGGCAACTCTAAAAGAGTATGAAGATTTCGATAACGTTTTGCTATTAAACTATCTTCCATCGAAAA
GATTCTTCCAAAATAGTAGATCGCATCTTTGACTGTAAACTCGCCAACAAGAGCAATGTCCTGTGGCATATATCCAACTCTGGGTCCAGGTACTCCG
CTGCCAGCTTCTCCCGGTTTTCCACCCAAAACCCAAATTTCACCTCCATCTATTTTCTTTCGGCCTACTATGCTGCTTAATAATGTAGTCTTTCCGC
AACCACTAGCTCCCAAAAGACCGTATATACTACCCCTTTCTACTTTCATACATATTCCAGTCAAGACTTCTTTTTGTCCATATTTTTTCACTACTCT
ATCAACAAACACAGCATAGTCGTCGTGCATTTTCGATTTTGAAGAAAGTGAAATTATGCGCCGTACTCCGCCAAACTGTTAGAACATATCACATAAC
TAATAAATTTAACAAGCAAATTACACAGGGGGGTTAAAATGCACTTTTAATTGAGCAGTTGTGTTTTACAAAACAA
>Dvv ABC-H D5118
TAAATTAAAAAGTTAATACTTTAATCATAAAAAAATCAAAACCTCGTGTTCCTATTCACCCCACCAAGTGGGGTGAATGGGATCACCTGTATTAAAA
TCTATTGTGTTCCTATTCGTACATTTGTAGGGTCAATAAAAATAAAAAGTGTCATAACGTGTAAAGTAACATAAACGAATAAATATAAAGTTTGTAT
TGTTGGAAAAATATATTGTTGAAATGAAAAATATTTTTGAATATTTTCTAATTTGGCATTTTTCTCTGAAAGACTATTAACCAGTCTCTTTAATAGT
CTAGGAAACTTGTCTTTAGCTATTGTGGATTCATTCCTGCCTGGTACTGCTTTTCATTCCTCAACTATTTTTCGCCACATAATCTTAATGGGCCTTG
TTGAGTCTGTACATGCATTTAAACTTTTTGTACAACAGTTCTGATGATTGCAGAAAACCGTATAAATATTTATTTCAAATTTACGGAATCCACATTT
GGTCGTTCGAGCCAGATGCTGATAATTAGTGCCCATAATAATTGGAAAAATTATATTCATTAATACTACGCATGGATGAGAACCCTGGTAGAAGTTT
GTTTTGGATTGGAGAGGATAACTAAACGTAGAAACACGTTTTGCAAGGCATGGTTCTTTCTATTCTCTTTGTGTGGATCAATATAATATGTATAATT
TGTGGTTAGATGAAATATTGTTTCGTTTGGTCTTACTCTATGGAGATATAAGATAAAAAAAAGTAACTTCTAATATGTGGGGTGAATGGCAACACGG
TAAATATTATGTTCCTATCCACCCCACATATCTCAAATTCATGACGTCAACAGTAATAATTGTTGTCAACAAAAAAGGTTTACAGATAAATCAAAAT
ACCATCATGTTATGTTGTTGTAAAAGTTAGAATAAACAAAAAAACAATACTCTTCAATCTATTTTGACAGACATTTTGTACTTACTGAGAAAGTTTT
TTAGGATGAGAAATATCCTCAACATTAACGACTTTGTTTGCACGCCGAGAACATAGGAACCAGTGACGCCAACAACACAAACCTAAATTTTCTAAAG
CCATATTACGTTAGCCAACTTGAAAAAAATTAAATAGCCCAACGAAAGTGATATATATGTTTAGGAAGATAAGTAATTGTTTAGTGAAGTTTTGTTC
CTATTCACCCCATTTTACGGTACTTGCTAGGTTTAAAGCTAAAACAATTACCCTTAAATTAAGAGGCCACATCAAATGACTATTTTACAATTGTTTT
TTACTTAATTTGTTATAATTTAATTGTCAAACACCAGTTTGTTAATTTTTAAAAAGGAACTACATATCTATTTTCATTTAGAATTAAATATCTCTAA
TACATAGTTTAGGAGATAGAATCATTTACATAATTAAAGTTTTCAGCCTGGTACCCTAGAAATATCGAATCTACGCACCAAAAAACTTAAAAACACC
ATGGTACCAAAAGGACATAAAAGCGCATACTCAACCTTCTTGGTTCCTACACTCCTAATAATGAAATACTATTTGTTTTATAATATATTATTCTTGA
CCTCATAACGCAACGGTACCTTTGTGATTATTTCAATATCAAATTTCAGCTTTAATTTTAGTATTCCTCAAAAACCTATCTACCTAACGCCATAAAA
TACCATAAAGACAGTACGCGAAATGAGTAAAAAATATACAAAAATTAAATTAATACTAGTACTTACAGTATTTAAGGGGGCAAGTTTATAACATAGA
TTTACTGTCCATCACTTAAACCAATTTTTGGCTGTATTTACTGGATTCGCAAAAACACCCGTCCAAACAATTAAATAGAGTTTTTAGATCTGATCTG
TTTATTAAAAAGTAGAATTATACATACAAAATTTCCCACGCGGTAAACAATTTACTAAATGGTTTTCAACCCCTCTTAGATCTTAGCACCAGTAAAC
TAATCGTTAAAAATAGTACTATCCATACTATCGTAGCTATAAATCCGTAATAGACGTCGGGTTCTCCTATGGACCAGCCCCTGGTCATCATGGATCT
CAAAGAGGTCGTGGCCAGGGTCAGAGGCAGGAACGTGGATATATATCGCAGCACCGTTGGCATTCCCTCTATCGGCCAGATGACACCGCTCAACAGA
AGAGTTGGGTAAAAGCTACCCAGAGCCAGCTGGATTGCATTTCGTTCGAGCTCACAGATGGCAGAGATGACGAATCCGAAGCACATACCACACAATC
CTTGAAGAATCGTAATAACTACGACAAGGAAGATGTCTCCTTTGCATTCAACTTGGAAAACGATAATCATGAAGATAAGTACGAGAGTTGTTTGGCC
GCACATTACCACGAACTGAGTGATGACGTGGGAAAATAGAATTTCGCTTGGTGTTACACCGGCTACCCAAGATCTGTCCAATAATCCTTCCATTCTT
TCTATAATTAAAGCAGAGGATGTCAGAGCCACAGCCAAGAAAAACACGATCGTTAATATTACACCAGGTGCTACGAAGTCTGTAAAGGACGGCTGGT
TAGAGCCATAGATAGGTTCCTTAAACGATATAGGTATCTCGGCTAATTCTTCATTATATTCACAATCCCTAAATATATCTTTCGTGAAATTTTGGAA
CGCCAGCTGAAGATCTCGTTGAAGTATAATACCTATTTGTTGATTTGACATATCCAGCCAAACTCTCACTTCGCTTTGATCCAACGTTTCTTCGTCC
GCGTCTTTCCCTAAGGCCATCCTAGCTACCAGAGCGTCGGTAAAGTTCTCGGTAAAGTACAAGGCACCCCACGCGTTGCCCGTTCTAACCGCCTCTT
TGGCTTCTTCGGGCGTCGCATAATATTGTTTGATGATCGTTGTCGTATTTATTGTGTCTACGTATCGACAACTTAGGTTGCCGAACTTACATCCTTT
TTCATACTCGCATTCCTGATAAGTCAAATTCGTAAAGTTCTTTTCGTGATTTACTATGGCTAGTTTGAGTCCTTTGGGATCTCCCCCGATGGCCAAA
CAGAAGAGGATAACTTGCATCACAGGTAGAGCGAAGATGAATAACATCACCCCGACATTCCTCCACATTCTGAGAAAATTCTTTTGGAGGAGGGCTC
GAAGTTTTCCCGTCGTTGTGAAATCTGAACAGCTACTACAGTCGTCGCATGCTTCTTTGATACCATGAGTTGCCGAAGAAGGTGGTTTTCCCATATC
TAAATGTCCGTTGCTATCAGGAACTAAAACTTCTTTGCTTTGATGGAAATTCAAGCCGACCACTCCACTCTCCTCCGTCACGGAGACGGACTCTTTC
TTGGACCAATTAAGTGTAGCTAAACTAATGTTATTAGCCATATTTTGGTCATTAGCTGCCGCACCTACTTGAACTTGCTTCCTAGATAGTTTTAAAA
ATACGTCTTCTAAGGAGGTACATTGATACATCTGTAGCAAAACTTGTGGGGATTCTTCTGCCAGGAGCTTACCGCTCCTCATAAGGCCAATAGTATG
TGCTTGCCTAGCTTCTTCAATATAATGTGTTGTAATTATGACTGTCTTGTTGCCGTCCTTTGTGATTTGTACAAGGTGATTCCATATACTTTGCCTA
AGCAAAGGATCCACACCAACAGTGGGTTCGTCAAGGATTAAAAGTTCTGGATCGTGCATTAGAGCCACCGCAAACGACACTCTTCGTTGTTGACCAC
CACTAAGATTCTTAACCATACGATTTTGACTAGGTAAATCCAAGAAGTTAAGTAAAAACTGCAGTCTTTCGTAAATTTCTTTTGACTCCATACCGAA
GATCCATCCAAAGTACATCATCGTTTCTTTGATTGTGAATTCTCCGTAGAGAGCTATTTCCTGAGGCATGTAACCCACACGCTTACCCGGTACTCCC
GAACCTTTTGTACCTGGTTTTCCTCCCAGTACCCATATCTCTCCAGTGTTGAGTCTCCGCCTTCCCACGATGCAGGACAAGAGAGTCGTTTTCCCAC
AACCTGATGCGCCCAGCAGACCATATATTGTTCCCTTGCCCACAGTCATGTTTAAATCACTCAACACGTGGTTGGGTTTCTTCTTCGAGCCATAGTG
CTTGTACGCGTGCCTTACACTCACTGCATTCTGTCTTCGGTTCCATACTGTGGACTGTTGATGTACTAGAGGCGGCCTTTCCACATTTACACCCACC
AAATCGGGCCCAGGAACTTGTGGCTCCACCCGATCACCCATTGTCGTACCCTCAATGCTTAATGCTCTCCTATAGATTCCTAACAATCTACCCAAAA
CATTCACTAAATTATTCCACGAAAACATCTACAAACGATTAACATTCCACTACGTTCCTCTGGCACATGACACTGATCACTAACTCACCAACACAGA
TAATTAGGAATATTAAACCAAAAAGTAATTATTATCAGTTTACAACAAAAGGTAATTGATATCACTTGTTTTCTTGGACACTTATCATTATTGCGAA

TATATTAATGGTGGTCTATTTGATTATTATTATAATTGCTTTTAAAATACCATTAGCTAACGTTATTTCATGGACTAAACACTAAACTTTGGAGTTG
TAACTGGGATAAAAATAGCTCGGACTCGACCACGTTCAGTGTTAGCGTGTCAAGATCAGCGAAAATTAATTTTGTTTTTATAATAGATTCATATATG
CATACGCATTAGCGAAATTAGACCAGTGATAGGAACATCAATTTTATCCGAGAATGAGTCGCAAAATGCATATGATTAGCTAGCTATACAGGGTTTG
G
>Dvv ABC-H D18290
ATAGCTAGAATGCAGGGAGGAAGAGTCCCCATAGAATTAATAAATATTGAAGCTGGCACAAAAAGCATAAGAGGACGGTCACGAAGAAAATGGATGG
AATCGGTAAAAAGAAGACTTGGCATTGTTGGGAGTACAAAACTAGAAAACAAGATAGGAAGACAGGAAGGAAACCATTGAAGCCTTCGGCCTCTAGG
GTCTGTTTGAGCTGATGTTCATGTATATGGTAGCAAGTCAGTCAATAACAAACAAATTCTATGATGTCAATACAAAACCGTCGTTTGTTTCAGAATA
AGCGTTCACAAATCATAAGATGGAAAGGTAGAGAAGAAAAACGAAACTTGTACAATTAGGTACCTACGTACTGCGTACAAGAAAAGTGAAACAATAT
TACAATAATTTAAAGTATTTAATAACTGTTTATAATATTCCTTATTTCTTTTCTCATCTTTTTGTTTTCGTCAAGTAGATGTTCAGCGCTAAGAGAA
ATAATATCCAGAAAGACAGTATTCCTATTGAGTGAAGGGTATGTAAGTTGGTAAGAGTCCATCCTTTCCTGATGACATTTCGTAATGCTGATGTTGG
AATTGCCAATGGCAACCACKCTATGACGAAATATATATATTTGGGTAATGCTTCCGCGGGCCACAAGATACCGCTAACTAGTACAAGTGGATTGAAG
ATTCCTGTAACTATTGTGTTTGCTGTTACATGATTATCGCTTAGAACTGATATTGCAAATCCTAACATGATTCCAGCTAAACCTTGTAGGTATAAAA
GAGCGAAAATTAGCCATAAAGACCCGATATACGGTTGTCCAAAAATTACAAATACAAACAACAAGAAGTCCAACGATATAATAATCATAAACACGTT
TTGATACAGAAGATGTGACAAGCTTAGCTCCAGAGCTGTAACACCGGCAACGATGGAGCGGTCCCAAATACCTTCACATTTTTCATCCAAGATTATC
ACGCAGGTCATCAAAACTCCCATAAAAAATATCATCGAGGTTATTATTCCAGGAGTGATAAAAATTACATATTCTTCATTCTTATTTCCGTAAAAGC
TTTCTTCTACTCTTAAAGGTATTTGTACAGCTTTCGATAAATATCTGCAGTCATCGTAGATCTTTTCGTGAAATAGCTGGTACTTATCAGCAATTCT
TTCTTTAATAATTCCGCCTACCTGAATAGATGTCATATCCATATAGACCTGCATTTCTTCATCTGGCAAATCGTACTTATCTACTCCATAATTAACT
CTAGCTTCAATTAGTTCCGACAAATTCTGATGCATATGTAGGAACCCCCTTATTTTTCCTGACTCCAAAGCTGACAAGGCTTCATATTTTGAATCAA
AGCGAATCATGTAAAATAATGGTACATCCAAATAGTCCTCAAACAGGCAACTGGCACTATTAATATGGCAACTGAATTTATCGTACGGCACAGCTGA
TTCGTTTCTGGGAAATGTTTTGCAGAACTCTTGAGATCTTTCGTTGTTGACTATACCGAATGGATTTCCTTTTGTTTCTCTGCCAACTGCTGTGAGA
AAAAACAAAACCTGAAGTGCGGGGAACACAACGAGAAATATTGTAGATATTAAATTTGATTTGAGCTTTTTATAATTTTTAAACAACAATGACTTGA
ATCTGATCATTTCGAAAATGGATTGTTTTTTTCTTTCAGTTCTCGTTTGTTGAGGTGCCTCACTAATTACATCCAATGCTGCATTTGTTCGTTCAGT
AGATTCATTTATGCCTGTTTGAGATGACTGAAAAGCCAAATTATCTACTCCAGAAGTTTGTTCTGTTTCTTCTTGTTTAGAACTTAGAATAACAAAG
ACTTCTTCTAGCAGATCACAGCCGTACCTTTCTAAGAGCGTCTCCGGGTGTTCTTCTGTTAATAACTTTCCCTTTCTCATTAGAGCAATCTTATCAG
CTTGTCTAGTTTCTTCTATATAATGTGTTGTTATTATCACAGTCGTTTTGTCCTCTTCAGTTATTTTCGTTAAATACTTCCATATTCTATTTCTAAG
AATTGAATCTGTACCAACAGTAGGTTCATCCAATATAAGTAACTCTGGTTTATGAACTAACGTTACAGCAAAGGACACTCTTCTCTGTTCTCCTCCG
CTACAATCTCTTACACATCTGTTATCCGTTGGAAGGTCTAATAATTCAATCAGTTCATCACAACGACTTTTTATTATGTCATCCTTTAAATTGTAGA
TTCTTCCAAAGAAATAAATAGAATCCCTTGCACTTAATTCTCCAACTAATGATATATCCTGTGGCATATACCCCACTCTATCTGCTAGGACTACACT
ACCATCCTGTTTTTCTCCCAGTACATATATATCCCCTTTATTTACATTACTCCTTCCAATTATACACCTTAATAAAGTTGTTTTACCACAACCACTG
GATCCCAAAAGTCCGTATATGGTACCTTTTTGAACGTTTAAACATAATTTGTCCAGCACTTTGTTTGAACCAAACGATTTTTCTACATCCTTGACAT
ATAGAGCAAGATGGTTATTCATTGTGTTCATGATAGAACACAAACT

>Dvv ABC-H
D11818CTAGAACAAAATTAAGTACAAAATTAAAAACTCCTAATGGAATATACACCCCCTAACCCTTCTTGAATTTAAGTAACAGAATACTGATCGT
CAAAAATATAAAGATCCATATGGTCAACGATATAAAACCGTTGTAAACTGTGGGATTACTGATGGGCCATCCTCTTGCTAGTATACATCTCAAACTC
TCTGTGGCTTGGGTCAAAGGTAGTACCACCGACATCCACGATATGTACTTATGCATGGCTTCTATGGGCCAAATGATGCCACACAACATTACGATCG
GTAAGAAACTGCCCATAGCTAGGTAGGTAGCAGATCTTTCATTTTCACAACTGCAGGCCACCACGAAACCAAAACACATTCCACATATACCGGCCAG
TACAGTCAAGGTTGTCACTGTCGTCCAATCTCCTCGTTGCGTTAAACCAAAGAGAATGAAAGCGACAAGAAGTACTGCGATGGACTGTCCCAACATG
ACGACGAACTGGGTGATAATTTGCGAAAACAACAGCTCCGTTCCGCTGATGCCGTTGACCAAAGATCTTTCCAAGATTCCTTCGTTCCTTTCCATGA
TCATTGAACCGGACGTGAGGGCAACAGCCATGAAGAAGATGATAGTCAAGACTACTCCAGGTGCTGCAAAATCGGTAAAATCTGGGTCTTGGTATCC
ATAAATTGGTTCATTGTATCGAATAGGAACGCCAACACTTTTTTCATTATATCCACAACTTTTCACAAATTCAGATATAAAGGTGAGGAATCCATAC
AGCATATCTCTTGTAAGGAAAGTTGCGATGTTTTCATTTGATTTATCTTCATAGACACTGATAGTAGAAGCCATTATATCTTCTGGAGGAGTATCTC
TCGAATTTTCAATTCTACTCCAAAGCGCGTCGGTGAAATTGTGTGGAACGACTACAACACCCCAGGACTTTCCTCGTTCCACTCTATGTCTCGCTTC
TTCTTCCGTTTTCATAAATTCCCACGTTATAGAATAACTCTTAGCCACTTCGTTCAAGTAATTACAACTTATTTTTGTACCATTACAATTTAGAGGT
TGCTGATGACAAGTCTCATCAGGGAAATTAATCTCGTTATTGACAACTGATACTTTTATTCCTTGAGGGTCGTGTCCAATAGTCCAACAGAACAAAA
TGGTCTGACTTATTGGTAAACCGATAATGAACATCATCATGGGAATATTTCTCCACATCCACAGGAAATTTTTCCATATCAAGGCGCGCATGTGATG
CGATTGGACAAATTTAAAGTAATCTTTGAAGGACACTTCCGGTTCCTCGTCTGGTGGTAGTTCTGGAATAATGTGTTCTGCCGTGGGTTCAGGAGCA
ATAGACACTCTTCTTCCTGTCCTACTAGTAACCGAAACACTGTCTCCGAATTCCCCAGAAATCTCTCCCAAGTCGTCGTCCAACACAGCAGCAGGAT
TAACAGCTCCACCTAGTTCAGGCACCGTTATCGTCTCGACGACACTTTTTGCTATACTGGAGCGTCTTCGTTTTCCCATGTTTTGCATCACGCTGAG
TTTCAAGAAGACGTCTTCCAAACTATCCACCCCGAATTGTGTTATCAGTCGTTCTGGGGATTCCTCGGCCAAGAAGTAGCCTCCACGCATAAGACCA
ATCAAATGAGCCTGCCTAGTCTCATCAATATAATGCGTGGTAATGATCACAGTCGTCCTACCATATTTAGTTATTTCAACCAAGTGGTTCCAAATAT
TCGTTCGAAGCAAAGGATCAACACCAACGGTGGGCTCGTCCAAAATTAGTAACTCAGGTTCGTGAAGAAGCGTAGCTGCTAAGGAAACTCGCCTCTG
TTGCCCTCCGCTGAGGGTTTTGACCTGCCTATCTGCGTCCGGTAACATCAAGAAGTTGATGAAAAAATCAACTTTGGCTTCTACTTCGTCTGTGGTC
ATGCGGGATATCCAGCCGAAGTATTTTAGAGTTTCTCGGATTGTGAATTCGCCGTTCAGTGCTGTCTCCTGTGGCATATATCCTACTCGTGGACCAG
GAACTCCACTGCCTCTCGAACCAGGCGTACCACCCAGTACCCATAGTTCTCCGGAATTAAGTCTTTTTCTACCTACAATACAATTGAGCAAAGTCGT
TTTACCACAGCCACTAGCACCCAAAAGTCCGTATATACATCCTTTGGGCACTGTCATGTTAAGATTGTCCAATATAACATATGGATTACTTTTGGTG
CCATATTTTTTCACAGCCCTCCTGACACAAACCGCTTGTTGTCGTCTGGAATTCAATGTAGAAGCTTGAGTTTTGAACTGCCTTCTGCGTAGTTCGA
TGTCACTGGGACGTTCCATATTGGCAAGAATCGTCCCTTAATAAAATAAACTTTTTGGACGTAATCCAACTTTTAGAGGATATCAATGGCACTTGAC
ACTAACGCGGTTTTTTTTTCAAATCACGCCGACATCTGAACGAGTCAAAGAATTGAAAAAAATCACCTCGATCGATGTAGACCGCACGATTTGGCGG
ACTGTTGTGTTTTGTTTCGCGATGCGTGGCGTTCTTC

Appendix C
Sequences of RNAi-related genes in *Megacopta cribraria*

siRNAi pathway

Dcr-2

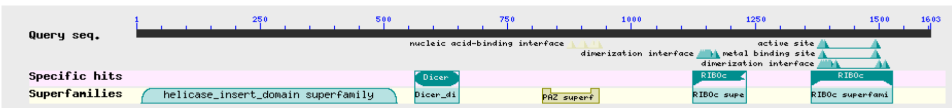
cDNA

>Contig 7670, length 5375
GTTAGAATGCTTTGTGTCTCAGTCGTAGGTTATGACTGTTACCGGTGTTTATTTTTGGTTGGCATGTTCAAAGGAATTTTTTCTTTTTAATAAACAG
TTTTATGATCATTTACTTTTATTTTATTTTAAAGTAAACCATTTTAATTAAGTAATTCGTTAACATATAAAAACAACCATTTTCAAATACCATTTCT
ACTTTTTTTCTAGTTAAAACTAGAGCAGGTTTTAAAAAATGGAAAACCAAATAAATGATTTTCAACCCAGAGAATACCAGGTTGAATTATTTGAAAA
AGCGTTGAAAGAAAACACCATAATTTTCTTGCCAACTGGCTCTGGTAAAACATTCATTGCTGTTATGCTAATAAAGCATAAATCCAATGATTTAAAA
TTACCTTATTCAGAAGGAGGTAAAAGAACTTTTTTTATTGTGGACACGGTGGCATTGGTGTCCCAACAGGCTCAAAGCATAAGACAACAAACTCCTT
TAACTGTTGGGGAGTACAGTGGTGATTTGGGTGTAGATTATTGGGATATTGACAAGTGGCTTGAAGAGCTTGAAAACCATCAGGTTTTAGTTATGAC
TTCTGCAATTTTCCTTGATATGATATTACATAAATTTATTTCTTTAACTCGAGTTAATCTTCTGGTGATTGATGAGTGTCATCATGCTGTTAATAAG
CATCCCATGAGGCAGATTATGGAACAGTTTGAAGGAGTTCAAAACGATAGGCCTCACATTTTGGGTTTAACAGCTGCTTTATTAAATTCTAATGTGA
AACCTAATCGAATTGAGAAAGAAATTCATGACTTAGAAATCACTTTACATGCAAAAATAGCTAGGGCTAAAAATGAAGCTATGATTCAAGTATATTC
AACAAGTCCTAAAGAAGAAATCATTTATTTTGATTTACCTCATGAAGATGAATGTTGGACAGTCATTAATTCATCATTTGAACAATTATACTCGATT
TTGGACTTGATTGTTCTTCAGGATAAATCTGTTAAACCCATGATTGCTCCTTTAGGTGCAGAACAAATATTTAATAAAACTGATCCCATTAATGAGA
TAAGAAATATCTTAAAAGAAATTTCGATTATTACTTTGGAATTAGGTGGGTATGCATCTTCATATGCCCTGTTATCTCGAATAGTACACTTTGAACG
TATGCGGATTTCTACTAATGACCCCGTTTTAATTACTGTATTAGGGCAGGTCATAACCCAATTTCTTCTAATAAGAAAGAAGCTTGATTTAATGATG
AATGTAGGAAATGAAAAAATTTTTAAATTTGCATCTCCAAAATTAAGAGAACTCTTCAAAATAATGAAGGATTATGATGGTAGTAATATAGCTCTTG
TTTTTGTTGAAAGAAGAACCACAGCAAAAGTTTTATACTATATTTTTGGTGATTTAGCTGAATTAACGGAAGAATATAAGAATATAAAACCGGATTT
CATTGTTGGATTTAATAATAATCCTGCATGCTTAACTAGAGAGCATATTATTGAAAGGAAAAAAAATCATAAAACCTTAAAGAAGTTTAATACAGGA
GAAACAAATGTTTTGTTTGCCTCAGATGTTATCGAGGAAGGCATTGATGTAAGAAATTGCAATTTTGTTATTAAATTTGATCAACCAAAAACAGTCA
GATCTTACATTCAATCAAAAGGTCGTGCCAGACATCGAGATAGTAAATTCATTGTCTTAGCTCCCCAGTCCAAATCAGTGGCATTTAATATAACATT
TTCAAATTTCAAAGAAGCAGAAGGCACTTTAAAGGAGATATTGTGCTATGGAAAAGATTCTTTGCGTAGAGGTGGTACAGAGAAAGAAGCAGAAGAA
TTTCTTTATACCAAGGAAGTAGAACCATTTTGTGTTGGTAACTTGAATGGAGCCAAAGTAACCGTAGATTCTGCTGTTTCTTTGTTAAACAGATATT
GTACATCACTTCCATCTGATAAATTTACGAAATTGACTGTATTTTATTGGAAAAAAAATGTAGTTATTAATAAAGAAAGTAAAGTTATTGGTTTTCT
TCAATTGCCTGTAAATTCTCCCCTTAAAGGATTCATAAAGGGAAAACCTTGTGATACTTTAAATTCAGCCAAAGGAGCTGTTTCGTTAGAAGCATGC
AGGATGTTGTATGAACTTGGTGAACTTGATGACTACTTATTACCTGTAGGCAATGACATGAATAGGGAAGTTGAAAACTCCCTACATTTTCCCCTCT
GGGAATTATTTAAAGAAGAAAAATCTTCGCTTAAAACAGGAACAAAAAAATGTCGTCGTATTTATGATAAAAATATTCCAAATTGGATCTGTGAAAG
TTTACCAGTGGAAAATAAACCAGTTTATATGCATGTATTAAACATAAGACCAAATTATTCTAAACCAGACAGCCATAGATTGTTAGCATTTTATAAT
TTATTGCTTTCAAACAAAACTTTTTCAGTTATTACTTCGAAGAAATGGCCCAAGATCTGCAATTTTCCATTATTTATGAATGTTGGGGAGGTCAGTA
TTGTCATCCAACAAAATGCTCGAAAGGTTACATTGAGCAAAGAGCAAATCAAATCTATATTTGATTTTCATTGTGTATTATTTACGCAAGCGTGGAC
TTATAAAAAAAATTTTTTGGTTAGAGATTATGAAAACAAAGAGAATTCCTATTTGATAGTGCCAACTATCGTTGATGGAGATGATGTTCAAATAGAT
TGGGAAACTATCAACAATCATAAATTTATTCCAGCTGTTAAGCCTGTACCAATTGAATCTAGAGAAAATATTGTAGTTGATAAGGAAAAATATTTCA
ATAAAATTGTTATCCCCTGGTATCGCCCTAAAGGTGACACATTGAAATATGTTGTTACAAGAGTTTGTGAAGAGGAAACTCCCCTCTCTGAATTTCC
TTCAAGGAAATATAATTCTTATTCTGATTATTTTCTTGAACGTTATAATCAATCAATTTTGAATAAAAATCAGAAATTGATTGAAGTCAGGGCTATT
TCAGGAATGTCCAATCATATACTACCAAGAGGAAGGTATAAAATTGGTAAAAAAAAGGAGACGAGATGAAGAAGACCTAGAAGAAACACTGGTTCCAG
AATTGTGTTTCCTTTTAGAATTTCCTGCTGTTTATATGCTAAAAGTTACTTTATTACCATCACTTTTGCACCGAATCACTATGCTGCTTAATGCTGA
GGAATTGAGACAAAAAATAGCTGCTGAAACAGGTTTAGGGGCTGTGAAACTCCCTAAAGGTGTTTATTGGAATGCTATAGAAGTTGATAATAATGTT
TTGGCAACTGATGAAAATTCAAATGCACGGAGTACTATTGAAGTTAGCAACATGTTGACTGAAAAAATGGTAGAATTACTACCTAAAGAGTATGGAT
GGAAAAAGGAAGAAGAACCAATGGATATAGAAAGAAACATTAGAAATATTGATTTGCTTTCAGTTTTGCATTATGAAAAATTTGTTAAAATGGAGTT
TAATCCTATTGATAAGTCTAGCAATATAAATCTTCCTCCCTGTATCAATGCACCAGATATTCCTATCCTAAACAAAAAGAATTCACTTTATGGTCCT
GAGCTGTCTGAAATATTAAAGGCACTTCTTGCAAGTTCTGCCAATGATATTATGAATTATGAGAGACTGGAAACATTGGGAGATTCTTATCTTAAGT
TTCTAGTCTCACTTATATTATTTGTTTATTTTGAAAATAAAGATGAAGGGAAGTTGAGTAATATCAAAGGAAAGATAATTGGAAATCGAAATCTATT
TTATTGTGGGAAGAATATTAATCTTGGAAATTTACTCAAGGTTTATGATTTTGTTTCCGATGATTGGGATGTACCCGGATTTGTAATTTTGGGAAGC
CTTAAAGAAATAATTAAAAATGCAAATATCCCTCCTTCCATACTTTACCAAATAAGTCTCACTAGTGAAGAAAGAGATTCATCTATTTTAAATGATG
AAACTTTAATTAAAGTTCAAGAAATATTTTGTAATGCTGTCCAAGCAGACGAGAATACACATAACAGTTTAGGTATAATTGGTCAGCATAGTATTAC
TGATAAAGTTATTGCAGATTCAATTGAGGCCCTTGTTGGGGTGTATCTGGAGGCATATGGAATTGAAGGAGCTTTTAAGTTGAGTAAGTGGTTAAAA
ATTCTTCCTGAAAACCTAGACGACATAAATAAAGTATTATATGAACCTCCGCCAACTGCCCAAGTATTGAATGGTGGAAATATTAATAATGTATTGA
TAATGCCTGATATTATTGAGTCTATTTTGGGGTACAAATTTAAGGATAGGAGTTATTTGTTACAAGCTTTAACTCATTGTTCTTTCGACAATGGTCT
TACTGACTGTTATCAAAGGTTGGAATTTCTTGGAGATGCCATATTAGATTTTTTGATCACATCATACATATATTCTAAATGTAAAAACATTACTCCT
GGTGAACTGACAGACTTGCGATCTGCCTTGGTAAATAATGTTACATTCGCTTGCCTCACTGTACGTTATGGTTTTCATAAGTTTATGTTAATAAAGG
CCTGTAAACTGACTAATATTATTAACAGATTTGTTGAACACCAAGACAAAAGAGGCCACAAAATTGGAGCAGAGATCTTATTTCTTTTAAATGAAAA
TGATGTGGAAGCTGCGGAAATTGTTGATGTCCCTAAAGCACTTGGTGATTTATTTGAATCTCTTGCTGCTGCTATTTATTTGGATAGTGGAAAATCT
TTGACTACTGTTTGGAATACATTCTATAAGTTAATGCATAATGAGATTGAGGAATTTATGAAAAACATTCCAAAGAATAGCATTAGATTGTTGTATG
AAAAATTCCCTAATCCTCTGCCGTCCTATGGGAAACCTGAAATTTTGAATGAAGGTCGGGCTGTGATGGTTCCATTGACCATTCATCAAAATAATGA
AATGAGATTTGTTGGTGTTGGAGAAAATAAAAGTCGAGCGAAACTTGCTGCAGCAAAAGTAGCTCTTAGATATTTCCATTTCAAAGATCCATTGTGA
TTTAATGTTGTTTTATTATACTTGTGAATATAATTTTATATAGATATAATTGTTTGATGATTATAAATTATTTTTAAAATATATAATTAATAATTAA
AAATTATATAAAATAATAATACAGTAAAACCTCGTATCGACTGCCCACTTGTGAATAAGATAAATGATTCCATTCCAGTGAGAGATAATTTTCTTTC
CTATACTTTTTGCCTGAACCTACAAATATTATTTCTTAAATGAGGTTTTTTCGCATATGCAGTTTCCAGAATTATAAGTTTGACTATTCCTTTTGTG
TATTTTTTAAAATTTATCTTTACTTTAAAAAAAAAAACAAA

Protein 1603(AA)
>MENQINDFQPREYQVELFEKALKENTIIFLPTGSGKTFIAVMLIKHKSNDLKLPYSEGGKRTFFIVDTVALVSQQAQSIRQQTPLTVGEYSGDLGV
DYWDIDKWLEELENHQVLVMTSAIFLDMILHKFISLTRVNLLVIDECHHAVNKHPMRQIMEQFEGVQNDRPHILGLTAALLNSNVKPNRIEKEIHDL
EITLHAKIARAKNEAMIQVYSTSPKEEIIYFDLPHEDECWTVINSSFEQLYSILDLIVLQDKSVKPMIAPLGAEQIFNKTDPINEIRNILKEISIIT

LELGGYASSYALLSRIVHFERMRISTNDPVLITVLGQVITQFLLIRKKLDLMMNVGNEKIFKFASPKLRELFKIMKDYDGSNIALVFVERRTTAKVL
YYIFGDLAELTEEYKNIKPDFIVGFNNNPACLTREHIIERKKNHKTLKKFNTGETNVLFASDVIEEGIDVRNCNFVIKFDQPKTVRSYIQSKGRARH
RDSKFIVLAPQSKSVAFNITFSNFKEAEGTLKEILCYGKDSLRRGGTEKEAEEFLYTKEVEPFCVGNLNGAKVTVDSAVSLLNRYCTSLPSDKFTKL
TVFYWKKNVVINKESKVIGFLQLPVNSPLKGFIKGKPCDTLNSAKGAVSLEACRMLYELGELDDYLLPVGNDMNREVENSLHFPLWELFKEEKSSLK
TGTKKCRRIYDKNIPNWICESLPVENKPVYMHVLNIRPNYSKPDSHRLLAFYNLLLSNKTFSVITSKKWPKICNFPLFMNVGEVSIVIQQNARKVTL
SKEQIKSIFDFHCVLFTQAWTYKKNFLVRDYENKENSYLIVPTIVDGDDVQIDWETINNHKFIPAVKPVPIESRENIVVDKEKYFNKIVIPWYRPKG
DTLKYVVTRVCEEETPLSEFPSRKYNSYSDYFLERYNQSILNKNQKLIEVRAISGMSNHILPRGRYKIGKKRRRDEEDLEETLVPELCFLLEFPAVY
MLKVTLLPSLLHRITMLLNAEELRQKIAAETGLGAVKLPKGVYWNAIEVDNNVLATDENSNARSTIEVSNMLTEKMVELLPKEYGWKKEEEPMDIER
NIRNIDLLSVLHYEKFVKMEFNPIDKSSNINLPPCINAPDIPILNKKNSLYGPELSEILKALLASSANDIMNYERLETLGDSYLKFLVSLILFVYFE
NKDEGKLSNIKGKIIGNRNLFYCGKNINLGNLLKVYDFVSDDWDVPGFVILGSLKEIIKNANIPPSILYQISLTSEERDSSILNDETLIKVQEIFCN
AVQADENTHNSLGIIGQHSITDKVIADSIEALVGVYLEAYGIEGAFKLSKWLKILPENLDDINKVLYEPPPTAQVLNGGNINNVLIMPDIIESILGY
KFKDRSYLLQALTHCSFDNGLTDCYQRLEFLGDAILDFLITSYIYSKCKNITPGELTDLRSALVNNVTFACLTVRYGFHKFMLIKACKLTNIINRFV
EHQDKRGHKIGAEILFLLNENDVEAAEIVDVPKALGDLFESLAAAIYLDSGKSLTTVWNTFYKLMHNEIEEFMKNIPKNSIRLLYEKFPNPLPSYGK
PEILNEGRAVMVPLTIHQNNEMRFVGVGENKSRAKLAAAKVALRYFHFKDPL

Graphical representation



Comparison with *Cylas puncticollis Dcr-2*

```
Query  20     DEMTPRNYQVALMEIALEKNTIIYLPTGSGKTFIAVLVLKRLSQSLSKPYSEDGKLSFIL  79
              ++  PR YQV L E AL++NTII+LPTGSGKTFIAV+++K  S  L  PYSE GK +F +
Sbjct  6      NDFQPREYQVELFEKALKENTIIFLPTGSGKTFIAVMLIKHKSNDLKLPYSEGGKRTFFI  65

Query  80     VNTVALVDQHTKCIEQRTSFKVGSYSGDMNLDGWSKEKWLNEFDKYQVIVMTVQILANLV  139
              V+TVALV Q  + I Q+T   VG YSGD+ +D W  +KWL E + +QV+VMT  I  +++
Sbjct  66     VDTVALVSQQAQSIRQQTPLTVGEYSGDLGVDYWDIDKWLEELENHQVLVMTSAIFLDMI  125

Query  140    NQGFIDLNKVNLLIFDECHRGVNDHSMRQLCKSLENV-NEPPRIIGLTATLLNGNCKPGR  198
                 FI L +VNLL+ DECH  VN H MRQ+ +  E V N+ P I+GLTA LLN N KP R
Sbjct  126    LHKFISLTRVNLLVIDECHHAVNKHPMRQIMEQFEGVQNDRPHILGLTAALLNSNVKPNR  185

Query  199    VMDMVKDLEITYHGKVATVNGLDQVIGYSTNPWENIVEFKTHTLLPIENKVIGVLN---E  255
              +   + DLEIT H K+A       +  YST+P E I+ F     LP E++   V+N   E
Sbjct  186    IEKEIHDLEITLHAKIARAKNEAMIQVYSTSPKEEIIYFD----LPHEDECWTVINSSFE  241

Query  256    LIVTIGCVKIVSSYNDPPQIDLKPV------NKNKGLKQLQNLIRDIIFQIESLGSYGGD  309
               + +I  + ++   +  P I   P+      NK   + +++N++++I      LG Y
Sbjct  242    QLYSILDLIVLQDKSVKPMI--APLGAEQIFNKTDPINEIRNILKEISIITLELGGYASS  299

Query  310    KSLLAHLIQIERLRKHCDDVTLYKVLNYVGVTFMYVRGLFSQAMKDYDEFNKILLFSSEK  369
               +LL+ ++  ER+R   +D  L  VL  V   F+ +R      M   +E  KI  F+S K
Sbjct  300    YALLSRIVHFERMRISTNDPVLITVLGQVITQFLLIRKKLDLMMNVGNE--KIFKFASPK  357

Query  370    VKKLITLLENYPTMSKEELCAIIFTKRRFTAKVVYYILDSLSKSSPKFEHIKTDFVVGNN  429
              +++L  ++++Y   +     A++F +RR TAKV+YYI   L++ + ++++IK DF+VG N
Sbjct  358    LRELFKIMKDYDGSN----IALVFVERRTTAKVLYYIFGDLAELTEEYKNIKPDFIVGFN  413

Query  430    SNPYNDTRECSYIRKKTRKILNSFGNKEINVLVASNVLEEGVDITMCSLVVKFEHPEDYR  489
              +NP   TRE    RKK  K L  F   E NVL AS+V+EEG+D+  C+ V+KF+ P+  R
Sbjct  414    NNPACLTREHIIERKKNHKTLKKFNTGETNVLFASDVIEEGIDVRNCNFVIKFDQPKTVR  473

Query  490    SYIQSKGRARNKSSLYMMFVDDKDKDTFKTRYQEYQEVESTLNEFLI-GKNLERSEPSQR  548
              SYIQSKGRAR++ S +++         F   +  ++E E TL E L  GK+  R   +++
Sbjct  474    SYIQSKGRARHRDSKFIVLAPQSKSVAFNITFSNFKEAEGTLKEILCYGKDSLRRGGTEK  533

Query  549    DIEEMYNEDCLAPYYVNGPNSARVTSTSAISLLCSYCLSLPSDKYTVYTPEVFYKSFLDS  608
              + EE      + P+ V   N A+VT  SA+SLL  YC SLPSDK+T  T   + K+ + +
Sbjct  534    EAEEFLYTKEVEPFCVGNLNGAKVTVDSAVSLLNRYCTSLPSDKFTKLTVFYWKKNVVIN  593

Query  609    SGLELIKVVIRLPTICPI-DIVEGIPMPSIKLAKRAAALQACKLLHQCGELDDNMLPKKA  667
                 ++I   ++LP   P+   ++G P  ++  AK A +L+AC++L++ GELDD +LP
Sbjct  594    KESKVIG-FLQLPVNSPLKGFIKGKPCDTLNSAKGAVSLEACRMLYELGELDDYLLPVGN  652

Query  668    GISEE-DVSFLFKHYPAEKEPDA----GTNKRKRLHKKVIPQYIRGPI-TSTKPVYLHII  721
               ++ E + S  F  +   KE  +    GT K +R++ K IP +I   +    KPVY+H++
Sbjct  653    DMNREVENSLHFPLWELFKEEKSSLKTGTKKCRRIYDKNIPNWICESLPVENKPVYMHVL  712

Query  722    EVTPAFKREENINRATMYDMYNSSLCYGIITPERVPEICDFPVFVSLGTLNISLKENYST  781
               + P + + ++      Y++  S+  + +IT ++ P+IC+FP+F+++G ++I +++N
Sbjct  713    NIRPNYSKPDSHRLLAFYNLLLSNKTFSVITSKKWPKICNFPLFMNVGEVSIVIQQNARK  772

Query  782    VNLTAENIDDIKAFNVLVYSDILTCLREFLIFDNSKDAESILIVP--VDKEHGCIDFSIL  839
              V L+ E I  I  F+ ++++   T  + FL+ D      S LIVP  VD +   ID+  +
Sbjct  773    VTLSKEQIKSIFDFHCVLFTQAWTYKKNFLVRDYENKENSYLIVPTIVDGDDVQIDWETI  832

Query  840    KEHKTVKSVMEPSQEEKLNVEVNLDTYLNKIVAPWYRNTG---FYIVTEVTFTKTARSVF  896
                HK + +V     E + N+ V+ + Y NKIV PWYR  G    Y+VT V   +T  S F
Sbjct  833    NNHKFIPAVKPVPIESRENIVVDKEKYFNKIVIPWYRPKGDTLKYVVTRVCEEETPLSEF  892

Query  897    PNESFGTYEQYFNDKHNLQLVNPNQPLLFVKSLSKKLNCFKPKG---AASRRKRDEMFEE  953
              P+  + +Y  YF +++N  ++N NQ L+ V+++S   N   P+G      +R+RDE  E+
Sbjct  893    PSRKYNSYSDYFLERYNQSILNKNQKLIEVRAISGMSNHILPRGRYKIGKKRRRDE--ED  950

Query  954    LDIHLIPELVIKQEFPASLWIQASLLPTILSRLSFLFRLEYLRVQIATEAGLGQKDVPNK  1013
              L+  L+PEL    EFPA   ++ +LLP++L R++ L   E LR +IA E GLG   +P
Sbjct  951    LEETLVPELCFLLEFPAVYMLKVTLLPSLLHRITMLLNAEELRQKIAAETGLGAVKLPKG  1010

Query  1014   A---SLELDLHLL  1023
                  ++E+D ++L
Sbjct  1011   VYWNAIEVDNNVL  1023
```

Ago-2

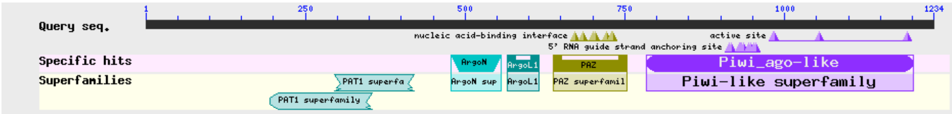
cDNA

Contig7506, Length-4208

GTAACCTCCCTATGTTATAGATTATATATTTTTAAAAAAAAGATAAGCATGGAATAAACGTCAAGGCTTATTGAAAACAAATTATTGCGTTTTTATT
TTTAGTGAAAAAAAAATCTCAAATTTTCATTTAACTACCTTGCTGGACACTGCATTAAATATGTCGCGAAATTAATTTATATTTAAAGTGCAAGTGA
CTGGTATTACACTGCATGACGAAAGTTCTGTGATTTCCCAGTACCAGAAAAGTCTTGAGACTTATTGAGTTAAACTTTCCATTTTTTTCAGTGTGTT
TTTATGAAAAGTGATTCGACTTTAAGTTATTATTGCTTTTTCTTGTTTGATCTAATTAAATTGTGTTGTTATTATTCATTGTAAAATTGGTTATTCG
CTCTTCTAAATGGGAAGAAAAGGAAAGCGTGGAGGTGGAGGTGGAGGTAGAGGTGGAGGTGGTTCTGGGGGCGCGGGTGGTGGAGATGCCCCAGGAA
GTTCCAAGCAACAACAACCCTTAAGGCAAGAAACCTCCCAGGAAATCCCTTCTCGTGGTGAAGGGCCAGTGATGCACCAGCAGGAGGATTTTCCATC
TCTTGGTGACAATAGTAGAAGACAAGGTGGCAGAGGGCGAGGGAGAGGTAGAGGGGGCAGCGGGGACAATATAGAACAGATAAGACAATTGCCTCCT
AATCAACCAGAGTATTCCCAAGAAACAAAGGAACAGCCTCGAGATAAGGGAAATGTTGGACGAGGAAGAGGAAGAGGGCAAGGAAGCCTTCAAATGC
AACAACCACAGCAGCCTGTTCAAATGCAACAACCACAGCAGCCTGTTCAAATGCAACAACCACAGCAGCCTGTTCAAATGCAACAACCACAGCAGCC
TGTTCAAATGCAACAACCACAGCAGCCTGTTCAAATGCAACAACCACAGCAATACATAGCTCCGCACCCTGAGCTACTTCAACAAATTAATATGCAG
CGACAAAGGGAATACCAGCTCATTCAGTCACTTCAGGCCATGGGACCACAAATAGCAAACATTCAACACTTACATCCACAGCAACAATCAACTCTTC
AAGAGTCATATAATGCTTGCCGAAATGAGTTATTAGCATGTCAGCAGCTATTACAACAACTTAGCATACAACAACAAGTAGTTGGTCCTCAGGGTGC
TCAATCAGGACCACAACCCACTCAGCAGCAAAGGGGCCCACCGCAACGTCAGGGTCAGGGGAGTGCTCCATCAGGTCAGGATGCTAGATCAGTTGCT
CAAAGGCCTCCATCAAGCAGTGCACAAGATAAGGAAGCTCGCCAGCCACCTCCTGCTAAAGCTCCATCTCAGGGTCCCAGTAGGCCTCAATCTCGTC
AAGGCATACAACAAGTAGTTGGTCCTCAGGGTGCTCAATCAGGACCACAACCCACTCAGCAGCAAAGTGGCCCACCGCAACATCAGGGTCAGGGGAG
TGCTCCATCAGGTCAGGATGCTAGATCAGTTGCTCAAAGGCCTCCACCAAGCAGTGCACAAGATAAGGAAACTCGCCAGCCACCGCCTGCTAAAACT
CCATCTCAGGGTCCCAGTAGGCCTCAATCTCATCAAGGTAAAAGTGAACCACCCTCCTCTTCAACTGGTACAGGCCAAGTTCCCAGAAGACTTGATA
GACCCAAGAAAACTAAAGCAGGAAGGAAAATCATTGTAGAAACAAACCATCTTGAGCTTAAGTTAAAAAATCCAAAAATGGTGGTACACCACTATGA
TGTCACAGTAAATCCGGAAAAGCCTTTTAGATTTTATCGACGAGCCTTGGATGCGGTGAGGGAAATTTGTTATCCTAGTAGATTTCCTGCATTTGAT
GGTAAGAAAAATCTCTACAGTTACCCGGAGCTTCCTCTAAGATCAAAGGAATTGGTAGAAACAGTTAACATCTATGATGAAGAAAAAGATCAAACCA
AAGAAATAACTGTCACTATTAAATATGCCAACGCTGTAGATGTGTCCATCATTGGTGAATACTTGAGGTGCGGTACTACATTAAACACTCCCCAAGA
AGCCATTCAAGCTCTTGATGTTGTCCTAAGGCAACCTTCTGCTCGAAGGTTAGTGACGGTTGGAAGGTCCTTTTTTTCTCCTCCTCCTGGCAGAATT
ATTTCACTTGGAGATGGTCTCGATTTATGGTATGGCTTCTTCCAGTCTGCAATTATTGGTTGGAAACCTTTCCTCAATATTGACGTTGCTCATAAAG
GATTCCCTACCCCGGAAAACTGTGTTGCAGCCCTAAAGAATATTATCCGTTCTAGAGATGGGGAACTGGAACATCCAAACTTTCAGTTACAAGGATA
TCAACGGGAATCTTTTCTTTCATACATCAAGAATCTTAAAGTAGTGTATGCCGTTCCCCATGGTAGAGGCAAAAGAACTTATAAAGTTATTGATTTA
TCGAAATGTCCCAGAGATAACAAATTTAAGTCCACAGATAAAACAACTCAACAAACATCTGAGATGACTGTAGAGAAATATTTTCTTAGTCAGTATA
GATATAAACTTAATTATCCTCACTTGCCATGTTTGGCGGTTGGTTCGAAAGATCGGCCAATATCTCTTCCATTGGAGTTATGTACTATTGCAGAGGG
ACAAGTCACAATGAAGAAAATGAATGAAATACAAACTCGTAATATGGTACAACAAGCGGCTGTGGATACTGATCGACGAAAAGAGAAGATTATGGAC
TCGATAAAAAGGATCAGCTTTAATAATGACCCCTGTGTGAAAGAGTTTGGACTTTCAGTTTCTGAAAGCTTCACGAGAGTTGATGCTAGAGTCTTAA
TGCCTCCCAAACTTGATTATAGTGAGAAGATGCAGGTGGATGTTAAACAAGGTGTGTGGAGGAATGAAAAATTCCTTCAAGGGGCTGAATTATTAAC
TTGGCGAATTTTGAACACTAATCGTTATATTCGTGATAATGACATAAGAGGGCTTGCTGAAGGCTTAATCAAACATGGCAAGACCTGTGGAATGAAT
ATTTCTAACAACTATGGTACTGCACTGAAGGATAAACTTTCAGATATAAAGGATTATTTCACACAATGTCAAAAGGATGGTGTCAGTCTGGTTATTA
TAATTTTGCCAGATAGAGGAGCAGTTACTTATGGTGGAGTCAAAAAGTGCGCTGAACTTCAAGTAGGAATTCTGACCCAGTGTTTGAAATCATTCAC
AGTCCAAAAGAAGCTTTCCTCTGCTACGTTTGTCAATATATTGCAAAAAGTCAATGCAAAATTAAATGGTGTGAACCACAGTATTACCCAAAATATT
TGGCCTTCATATTTCCATAAGCCGGTAATAGTGGTTGGTGCAGATGTGACCCACCCGTCTCCTGATCAAGTCAATATACCTTCCATAGCTGCGGTTG
CTGCTTCCCATGATGTTAAAGCATTTAAGTACAACATGATCTGGAGGCTCCAACCCCCACGGGAGGAAGTAATAAAGGATCTTCAAGCCATAATGAA
GGATCATTTGTTAAAGTTTTATTTAAGTACTAAGCAAAAGCCGTTAGCCATTCTCTTCTATAGAGATGGAGTTTCTGAGGGCCAGTTCAAAATAGTC
AGGGATATTGAACTAAATGCTATCCGAAGGGCCTGTTCTTCTCTTGATGCTGATTATAAGCCTCCAATTACTTTTTTTGTCGTACAGAAGAGGCATC
ATACAAGATTTTTCCCTACTTCACCGAAGGATGCTGAAGGAAGGAACAAGAATGTGCCTGCAGGAACAGTCGTTGACACAGAAATTACCCATCCCAC
TGAATTAGATTTCTATCTCGTTAGTCACGCTAGTATACAGGGAACGGCCCGGCCAACAAAATATCATTTACTTTGGGATGATGCAGAGTTATCAGAA
CAAGAATTGGAAGAGATAACTTTTTATCTTTGTCATCTTTTTACCAGGTGCACTAGGTCAGTTTCATATCCTGCACCCACCTATTACGCTCATCTTG
CTGCTTTTCGAGCACGCTCATATATGGAGACAGAACATATACATTTAGAAAGGCTAGAAGAAGAGCAGAAACGAAGAATAATTAAAGATACATTTAT
TAAGAAAAATCCTATGTTCTTTGTTTAAGAAAAAAATTGATTTGTGACTATGTGATAAAAGTTATATTTTATCTTTTGTAATTAAAGAATAGGTCATA
TTTTTAATTATTATGTTTGTTATAAATGTAATCAAT

Protein 1234(AA)
MGRKGKRGGGGGGRGGGGSGGAGGGDAPGSSKQQQPLRQETSQEIPSRGEGPVMHQQEDFPSLGDNSRRQGGRGRGRGRGGSGDNIEQIRQLPPNQP
EYSQETKEQPRDKGNVGRGRGRGQGSLQMQQPQQPVQMQQPQQPVQMQQPQQPVQMQQPQQPVQMQQPQQPVQMQQPQQYIAPHPELLQQINMQRQR
EYQLIQSLQAMGPQIANIQHLHPQQQSTLQESYNACRNELLACQQLLQQLSIQQQVVGPQGAQSGPQPTQQQRGPPQRQGQGSAPSGQDARSVAQRP
PSSSAQDKEARQPPPAKAPSQGPSRPQSRQGIQQVVGPQGAQSGPQPTQQQSGPPQHQGQGSAPSGQDARSVAQRPPPSSAQDKETRQPPPAKTPSQ
GPSRPQSHQGKSEPPSSSTGTGQVPRRLDRPKKTKAGRKIIVETNHLELKLKNPKMVVHHYDVTVNPEKPFRFYRRALDAVREICYPSRFPAFDGKK
NLYSYPELPLRSKELVETVNIYDEEKDQTKEITVTIKYANAVDVSIIGEYLRCGTTLNTPQEAIQALDVVLRQPSARRLVTVGRSFFSPPPGRIISL
GDGLDLWYGFFQSAIIGWKPFLNIDVAHKGFPTPENCVAALKNIIRSRDGELEHPNFQLQGYQRESFLSYIKNLKVVYAVPHGRGKRTYKVIDLSKC
PRDNKFKSTDKTTQQTSEMTVEKYFLSQYRYKLNYPHLPCLAVGSKDRPISLPLELCTIAEGQVTMKKMNEIQTRNMVQQAAVDTDRRKEKIMDSIK
RISFNNDPCVKEFGLSVSESFTRVDARVLMPPKLDYSEKMQVDVKQGVWRNEKFLQGAELLTWRILNTNRYIRDNDIRGLAEGLIKHGKTCGMNISN
NYGTALKDKLSDIKDYFTQCQKDGVSLVIIILPDRGAVTYGGVKKCAELQVGILTQCLKSFTVQKKLSSATFVNILQKVNAKLNGVNHSITQNIWPS
YFHKPVIVVGADVTHPSPDQVNIPSIAAVAASHDVKAFKYNMIWRLQPPREEVIKDLQAIMKDHLLKFYLSTKQKPLAILFYRDGVSEGQFKIVRDI
ELNAIRRACSSLDADYKPPITFFVVQKRHHTRFFPTSPKDAEGRNKNVPAGTVVDTEITHPTELDFYLVSHASIQGTARPTKYHLLWDDAELSEQEL
EEITFYLCHLFTRCTRSVSYPAPTYYAHLAAFRARSYMETEHIHLERLEEEQKRRIIKDTFIKKNPMFFV

Graphical representation



Comparison with *Cylas puncticollis Ago-2*

```
Query  22    GSQPQPQVSSSQPQPQPQAKVPAPQTKSSPSTEDLTKRVSEMQV----SKPKTLADDSGL  77
             GS P  Q + S  Q  P +     +T+  P  +  ++  S  Q     S+P + +  +G
Sbjct  352   GSAPSGQDARSVAQRPPPSSAQDKETRQPPPAKTPSQGPSRPQSHQGKSEPPSSSTGTGQ  411

Query  78    VSKVQIKPGTQ--GRRIKVESNHLKLSLGTLKEV-YHYDVNIVPDRPKKFLRAVVEKFRQ  134
             V +   +P     GR+I VE+NHL+L L   K V +HYDV + P++P +F R  ++  R+
Sbjct  412   VPRRLDRPKKTKAGRKIIVETNHLELKLKNPKMVVHHYDVTVNPEKPFRFYRRALDAVRE  471

Query  135   TRFPNKYPAYDGNKNMFTSFLLSGNLYDVLEEEVTIVDEE-GREKRFKIEVKWANTIDLT  193
               +P+++PA+DG KN+++   L     +++E  V I DEE  + K   + +K+AN +D++
Sbjct  472   ICYPSRFPAFDGKKNLYSYPELPLRSKELVET-VNIYDEEKDQTKEITVTIKYANAVDVS  530

Query  194   PLRN---LHSSPATPQLALQAVDIILRSAPVTTCLPVGRSFFVKPEGRIIDLGQGMEMYH  250
              +       ++  TPQ A+QA+D++LR       + VGRSFF  P GRII LG G+++++
Sbjct  531   IIGEYLRCGTTLNTPQEAIQALDVVLRQPSARRLVTVGRSFFSPPPGRIISLGDGLDLWY  590

Query  251   GFYQSAIRGWTPLLNVDVAHKAFPKSQNVIDTLVDLLSRPRQEFRRNDLQYLNPQQQEIL  310
             GF+QSAI GW P LN+DVAHK FP  +N +  L +++     E    + Q    Q++  L
Sbjct  591   GFFQSAIIGWKPFLNIDVAHKGFPTPENCVAALKNIIRSRDGELEHPNFQLQGYQRESFL  650

Query  311   SHFIKTLRVNYEIPGHPGTRRGYRVNGLGQPASQARFQHENRN------MTVAEYFRSVK  364
             S +IK L+V Y +P H   +R Y+V  L +     +F+  ++       MTV +YF S
Sbjct  651   S-YIKNLKVVYAVP-HGRGKRTYKVIDLSKCPRDNKFKSTDKTTQQTSEMTVEKYFLSQY  708

Query  365   NYALRYPDLPTLWVGSVQRTDKILLPLELCTIVEGQALNRKMTDIQTSNMIRQAATSTDV  424
              Y L YP LP L VGS  R   I LPLELCTI EGQ   +KM +IQT NM++QAA  TD
Sbjct  709   RYKLNYPHLPCLAVGSKDRP--ISLPLELCTIAEGQVTMKKMNEIQTRNMVQQAAVDTDR  766

Query  425   RKQKIMTGVQQASYNQSPCVREFGFSVGNEFQKLDARVLQPPQIQYN-RSNVRVSKGVWR  483
             RK+KIM  +++ S+N  PCV+EFG SV   F ++DARVL PP++ Y+ +  V V +GVWR
Sbjct  767   RKEKIMDSIKRISFNNDPCVKEFGLSVSESFTRVDARVLMPPKLDYSEKMQVDVKQGVWR  826

Query  484   NDVFLKGAVIKKWTIASVVSRYGPRPDDLKRMADMFHRAARDVGMIFETPAQEPFTTIGP  543
             N+ FL+GA +  W I +  +RY  R +D++ +A+   +  +  GM
Sbjct  827   NEKFLQGAELLTWRILNT-NRYI-RDNDIRGLAEGLIKHGKTCGMNISNNYGTAL-----  879

Query  544   RQDFQTISQYFNVQKEKQLDVIFVVVPDSGP-QYSYVKRAAEINVGCLTQCVKSTTVARK  602
             +     I  YF   ++  + ++ +++PD G   Y  VK+ AE+ VG LTQC+KS TV +K
Sbjct  880   KDKLSDIKDYFTQCQKDGVSLVIIILPDRGAVTYGGVKKCAELQVGILTQCLKSFTVQKK  939

Query  603   MNAQTALNILLKVNSKLNGINHTIADR--PPIMKRPVMIMGADVTHPGPDAQNIPSVAAV  660
             +++ T +NIL KVN+KLNG+NH+I     P    +PV+++GADVTHP PD  NIPS+AAV
Sbjct  940   LSSATFVNILQKVNAKLNGVNHSITQNIWPSYFHKPVIVVGADVTHPSPDQVNIPSIAAV  999

Query  661   TASHDPKAFQYNICWRLQSPRLEIIEDLQAIVVEQLKFFYAQTKCKPEAIVFFRDGVSEG  720
              ASHD KAF+YN+ WRLQ PR E+I+DLQAI+ + L  FY  TK KP AI+F+RDGVSEG
Sbjct  1000  AASHDVKAFKYNMIWRLQPPREEVIKDLQAIMKDHLLKFYLSTKQKPLAILFYRDGVSEG  1059

Query  721   QFQQVRDSEIRAIRAACKTLQQDGYEPKITFVVVQKRHHTRLFPTNPRDSEDKNCNVPAG  780
             QF+ VRD E+ AIR AC +L  D Y+P ITF VVQKRHHTR FPT+P+D+E +N NVPAG
Sbjct  1060  QFKIVRDIELNAIRRACSSLDAD-YKPPITFFVVQKRHHTRFFPTSPKDAEGRNKNVPAG  1118

Query  781   TCVDTDITHPFMQDFYLVSHASIQGVAKPTKYCTLWDDNLLDNDAVEQLAYYLCHMFSRC  840
             T VDT+ITHP   DFYLVSHASIQG A+PTKY  LWDD  L    +E++ +YLCH+F+RC
Sbjct  1119  TVVDTEITHPTELDFYLVSHASIQGTARPTKYHLLWDDAELSEQELEEITFYLCHLFTRC  1178

Query  841   NRSVSYPAPTYYAHLAAARAKVYIENDNLDMRALQREFQRFQIRDEIRKDLPMFFV  896
              RSVSYPAPTYYAHLAA RA+ Y+E +++ +  L+ E +R  I+D   K  PMFFV
Sbjct  1179  TRSVSYPAPTYYAHLAAFRARSYMETEHIHLERLEEEQKRRIIKDTFIKKNPMFFV  1234
```

miRNAi pathway *Megacopta cribraria*

Dcr-1

cDNA

Contig11725, Length-4144

TATCACAGTTTCAACAAAAATATATATAATTAACATAAAACATTTTTTTTAATTGGTACAGATTGATTTTTCAATAAATTTAAAATATCATATTTGT

AATTTAACTTTGACAACTAATAAATCTGAATGGAATGAAACCCATTATTTGAATGGCTACCTCTAAAAAGGTTCTATAACTGATTTTACAGACAATT

GAATCTTTTTAGTAGGAAGCTATTCATTATCAACTAGTGAGTGAGATGTTAAATAGTTCCCAAATGCAAAATGTCTTTCTTACAATCTATATCAATA

ACAATTAAACAAGGAAGAAATAAATATTATTGGTTCAGAATTAAAGAGTATCTTTCATGCCTAAGAATAGGCTAATGAAAATAAATTTACTATAAAC

AATTTGGTGTCAAGTTTTTCCTTTTTAGCTAAAATATAAATATAATAATGAAAACTTCTTCTTTTAATTCACTCTTCAAAAAATAAAACTGTAGAGC

ACTTTAAATGTAACTGCACCATGAAGATTAATACATTCATTGATATAATGCAATTGTGTTTTATTGATTTTTTAAGTCTTGACAAACTCTGAGAGCT

TTAAGTAACGACTGACTAGGATAATCTGTCAACTTTCTGGCCAGAAGCCCTATCTTTAATTGTTTTAATGCACATTTTGCTGCTGTGCATTTGGCTA

TTCGATAATTTCTCCCAATACCTTTAAAAGTTCCTTTACCAAAGACCTCTACTGTAACTCTAACTCTTCTTCCATCTGCTAACTTTTCAGGTTTACT

GAATTTTGCAGTCTCCGGTTCCAATTCAAGTAGTTCTCGAATTGGAGATTTTGGAACTTTTGTACTAAATTGTTCTATTTCTTTCTTCATCATCCTG

TAATAAACCTTCCAGACAGTGTCTAGTGACATCCCGCTATCAAGAAATATTGCTCCAGCAACAGATTCAAAAACATCGCCTAGAGCTTTTGGCACTT

CAACATCTTCAGCCTCCTCACCAACTAAATAATACTCTTCACTAATTAGATGACTATTTTCTTCCTGAATTGTGACAAATCTTTGAACGACTTCTGC

CAGGCCAGGAGAAAGATGTCGGAAATACTTATGGAGTCCATATTTAACAGCAAGTGAAGCGAAGATTGTGTTATTAACGAGGGCAGACCTTAGGTCT

GTAAGTGCACCTGGACTATGTGCACGTTTGTCTTCAAACAGGTGGCGTGTTATAAGATAGTCAAGAACAGCATCACCCAAAAATTCTAATCGCTGAT

AGCAATCAGTAACTCTATTTGGAGAATAAGAAGCATGACTCATTGCTTGTAATAAGTAAGATTTATCACGGAAAGAATAATGTAGGGATGATTCTAA

AGCCTCAAAACCAGCCATGAGAAATGCTAGTTCCCCTTCAGGATCATCTACATGCTGGAGCAATGGTGAGAGAGGATCTGATAGGGTAGTGAACTTG

ACCACAGTATTCTCGTTAGTTTTTATTCTCCCAACAGTCTTCATCGATGAACTTGTTCTTTTTGATACATCTATTTCCTCTCGTGGCAGGACTTTAA

TACCAAGCCAAGACATGAACAACAATGCACCTCTAGGACCACATTCTATTAGGTAAGCTCCTATCAAAGCTTCCACACAATCAGCAATACTCTTATC

AGGAATTGAATGCTGAGTAACAAGATTGTAGGGAAGAGTAGTTTGTGGCTCTGAATCTAATAATCTTTGTTCTGCATCTCGGGGTACTAAGTAACAA

GGAGGCAACCAGTTATCATGAGGTTCAAATTTACTTGCTATCATACTTTCACCAAACATTTTCTTCTGTCCTAACCTATAGAGCTTCAAATTACTGA

CCTGTTTTGAACGTAGATGACTGAGCTTTCCTTCATGAATACTCTCGTGAGTACAATATAGGTAAGTTGTAATTGCATATTTAAGGAACGAATCACC

AATAGTTTCCAACCGTTCAAGATTAATCCCATCATTGGCATTTGACATGGTCAAAGCTTGAAGAAGGATGGATGGACTAGGACCTGGATGGCCCTCC

AGTTTAGGCTGATGGTCAAAACTAAATTTACTGTTCTCTTCAACAACAATGTCAGAAATTCTTTCCGTTGGGTATGGGTATGATTCAAATTTCTTTT

TTTGATACTTAGGAATCAAACTTTCAGGCATACCATAACCTCTTAATTTAAAACCTAGTTCTTTTGTTTCACTTTGAAGATTTTCATCCGCTTTGAG

CTCCTCTTTATTACGAATAAAAGGTTCTTCGGATTTAATTAATATTTTTTTAACCATTTCTACATTTTCGATGGTTGCTTTTTGAAAATCATGAGTC

TCTTCCTCAACTTTATTCTTTACTTGAGAAATGAAATCATCAGCATCCCATTTCCAGACATCAAAGTCCTTACTTTTATCTTTATCAATTTCCCTGT

TGAGCTCATGTTCATTATCTTCAACTGCTTCAGCCAAATAGTCACTTCTGAAAGTGATTTTCAATTTTCCAGTTCCTCCATCAGATGAATCTGAAGT

TTCTGATTCTACATATAGGTCGCTTGAATCATCAGATTCACAGTCTGAACTGTAATAATTAAAAGAATGGTTTTCACCTACATCCCAACTGGTAGGA

GACCCATAACGCACCAATGAAAGGTCTGCAGGTAAATCTTTCTGAAATTCCTCATCAAGCTGAGCCATATCATTGGACCAAGTCCCTATTTCCATCC

AGTCACCTTTTTCTTTTTCATCACTTTTATCAATTTCAGTTTCCTCGATAGGTTCATTTTCTGATTCTTCAGCTTCTGATTCTTTTTTTTTGCTGGC

TAAATCATTATTACTCTTTTCATTCAAGAGTCCATTTTTGGTGCCTACTTTTCTATTTCCATCATAGACTTTATTGGCATCATTTTCTTTATTTTCA

ATATCTTGTTTTGTTCTTTTTAAGACATCTGCTAAACTCCAACCAAAATCTAAAGGAGGCCATTCAAAATCTTTAGGAAGATTAACAAGCCCAAGAC

CTATTTCATCAGCAACTTTGGTTCGAATTTCATCTGCCAGAAGGAGAGCATTAAGCCTATATAAAATGCATGGGAGACAAACAGCTTGTCTCCAGAG

AGAAGCTGGAACAACATGAACTTCACATAACTCCCCAACTAAAATTTGTTTTTGTTCCAAGTTTTCCCTTTTGGCTCTCTTTGTTTCTTCACTAGAT

GTGGGCAAAGCAACACCTTTCCGATTTACATACCTTGGTGTAAGGAAATTTAGTCTAGCAGAAGTATGATCAACATCAAGAAGAGGTTGTGTAGTAT

TTTGTATTTGTATATCGTACTTTTTCAGATAGTAGTCCTCAAAGGTATTATAGAAACTACCAGGAAATCCTGATTTTGGATTGAGATGGTAACATAT

TTCAGCAACATAGAAATACTGAGGTTGATCTTGATTTCTATACCACGGCATAACAACAGCATCATTATACTTCTCTGGATCGAAAACAAAGTTTTTT

CGTTCTTCATCAGATACTAAATGAGGATTATCATCTTTTCTTTTATAGATTACTTCAAGGAAGTCCCAATCTACTTGACTTAATCCATTATCATCTT

TCTTAGTTGGCACAATGAAAAATGAGTTCTCCTTAGCTTTAGGATCAAAACTCATTAAATATTTTTGAAGTCGAAGAACTGAGGTGAAAATAAAATT

CAAAAATGTTGCTATCGTTTCTACATTCCTTTCAGATAACCTTATTGAGCTAGAGCATAGTTTTAAGCTGACTTGCACTTCTCCTGATCTAGTAAAT

ATTGGAAATGGACAGACCTTAGGAATTTCTTTCATGGTTAGAATTCCAAAAGCTTGAGGTGCAAGTTCAGGAGGATGAAGTTTTCTACCTCTGGTGT

TTTGCTCCTCTGGAAGTGGACATGTAAGCACCAAGTCAATATTATACAAATAGCATTCGATTCCAGCAGCAGGGCGACATTCATTCAAGCAATTAGC

AATTCTTTTGTAATAATATTGTCTGCTTTTAGTTGTACCAGGTCTAGGTAAATTCGGGTCCCATGGCAAGGGAGAAAGATCTTCTTCTTCAAATCCA

GAATTAAATGCTTCTTTACCAATAGGCTCCAAGTTATCATCAAGTTCTCCTTTTTCATGCAATAAACGG

Protein 1090(AA)

MKEIPKVCPFPIFTRSGEVQVSLKLCSSSIRLSERNVETIATFLNFIFTSVLRLQKYLMSFDPKAKENSFFIVPTKKDDNGLSQVDWDFLEVIYKRK

DDNPHLVSDEERKNFVFDPEKYNDAVVMPWYRNQDQPQYFYVAEICYHLNPKSGFPGSFYNTFEDYYLKKYDIQIQNTTQPLLDVDHTSARLNFLTP

RYVNRKGVALPTSSEETKRAKRENLEQKQILVGELCEVHVVPASLWRQAVCLPCILYRLNALLLADEIRTKVADEIGLGLVNLPKDFEWPPLDFGWS

LADVLKRTKQDIENKENDANKVYDGNRKVGTKNGLLNEKSNNDLASKKKESEAEESENEPIEETEIDKSDEKEKGDWMEIGTWSNDMAQLDEEFQKD

LPADLSLVRYGSPTSWDVGENHSFNYYSSDCESDDSSDLYVESETSDSSDGGTGKLKITFRSDYLAEAVEDNEHELNREIDKDKSKDFDVWKWDADD

FISQVKNKVEEETHDFQKATIENVEMVKKILIKSEEPFIRNKEELKADENLQSETKELGFKLRGYGMPESLIPKYQKKKFESYPYPTERISDIVVEE

NSKFSFDHQPKLEGHPGPSPSILLQALTMSNANDGINLERLETIGDSFLKYAITTYLYCTHESIHEGKLSHLRSKQVSNLKLYRLGQKKMFGESMIA

SKFEPHDNWLPPCYLVPRDAEQRLLDSEPQTTLPYNLVTQHSIPDKSIADCVEALIGAYLIECGPRGALLFMSWLGIKVLPREEIDVSKRTSSSMKT

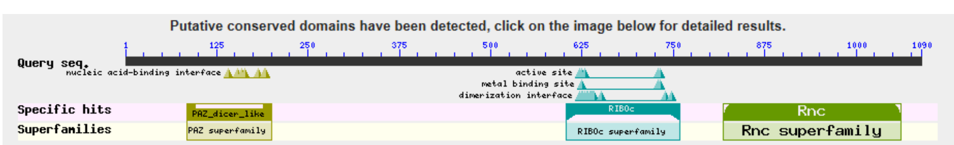
VGRIKTNENTVVKFTTLSDPLSPLLQHVDDPEGELAFLMAGFEALESSLHYSFRDKSYLLQAMSHASYSPNRVTDCYQRLEFLGDAVLDYLITRHLF

EDKRAHSPGALTDLRSALVNNTIFASLAVKYGLHKYFRHLSPGLAEVVQRFVTIQEENSHLISEEYYLVGEEAEDVEVPKALGDVFESVAGAIFLDS

GMSLDTVWKVYYRMMKKEIEQFSTKVPKSPIRELLELEPETAKFSKPEKLADGRRVRVTVEVFGKGTFKGIGRNYRIAKCTAAKCALKQLKIGLLAR

KLTDYPSQSLLKALRVCQDLKNQ

Graphical representation



Comparison with *Cylas puncticollis DCR-1*

```
Query  395   REIPKISAFPIFTRSGEVEVELRLCAKDVIVNEEQTHKIQEFVNYTFTTVLRLQKYLMLF  454
             +EIPK+  FPIFTRSGEV+V L+LC+  + ++E     I  F+N+ FT+VLRLQKYLM F
Sbjct  2     KEIPKVCPFPIFTRSGEVQVSLKLCSSSIRLSERNVETIATFLNFIFTSVLRLQKYLMSF  61

Query  455   KPELSDKNYLIVPTIQSEHGID-VDWEFIDIIYKNLHFVPRVIPDEDRKDYVFNTDDYKD  513
              P+  + ++ IVPT + ++G+  VDW+F+++IYK     P ++ DE+RK++VF+ + Y D
Sbjct  62    DPKAKENSFFIVPTKKDDNGLSQVDWDFLEVIYKRKDDNPHLVSDEERKNFVFDPEKYND  121

Query  514   AVVMPWYRNQDQPQYFYVAEICSFLNPTSAFPGSEYKSFEQYYLIKYAIQIQNLEQYLLD  573
             AVVMPWYRNQDQPQYFYVAEIC  LNP S FPGS Y +FE YYL KY IQIQN  Q LLD
Sbjct  122   AVVMPWYRNQDQPQYFYVAEICYHLNPKSGFPGSFYNTFEDYYLKKYDIQIQNTTQPLLD  181

Query  574   VDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILVPELCAIHPFSASLWR  633
             VDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILV ELC +H   ASLWR
Sbjct  182   VDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILVGELCEVHVVPASLWR  241

Query  634   KAVALPCVLYRINALLLADQIRTTVASDLNLGKTKLDDDFKWTPLNFGWSLSDVLKKSRD  693
             +AV LPC+LYR+NALLLAD+IRT VA ++ LG    L  DF+W PL+FGWSL+DVLK+++ 
Sbjct  242   QAVCLPCILYRLNALLLADEIRTKVADEIGLGLVNLPKDFEWPPLDFGWSLADVLKRTK-  300

Query  694   EEMKKQEEKLKSLREKTTELKITN-LVSDASEED---------PVNSDADISEAEEEEGN  743
             ++++ +E     + +   ++   N L+++ S  D            S+ +  E  E + +
Sbjct  301   QDIENKENDANKVYDGNRKVGTKNGLLNEKSNNDLASKKKESEAEESENEPIEETEIDKS  360

Query  744   DEGKGGDNWVVGTWSNEMAQ---SSQGDSTA---LVRYASPTSWLHGNSYDEDF------  791
             DE + GD   +GTWSN+MAQ     Q D  A   LVRY SPTSW  G ++  ++      
Sbjct  361   DEKEKGDWMEIGTWSNDMAQLDEEFQKDLPADLSLVRYGSPTSWDVGENHSFNYYSSDCE  420

Query  792   ---------SDDMSDNSDFDSEESTHEWGRLRIEFTGDHQAEALDDDD---SKEDELYFV  839
                        + SD+SD  +       G+L+I F  D+ AEA++D++   ++E +    
Sbjct  421   SDDSSDLYVESETSDSSDGGT-------GKLKITFRSDYLAEAVEDNEHELNREIDKDKS  473

Query  840   EDVNAWTVDDN------SQETERLRHEFNNACLRNKEHILSSGILIKENEEFEKKSQNTP  893
             +D + W  D +        + E   H+F  A + N E +    ILIK  E F +  +   
Sbjct  474   KDFDVWKWDADDFISQVKNKVEEETHDFQKATIENVEMV--KKILIKSEEPFIRNKEELK  531

Query  894   ADDIPPVCPDNFDFD-------SLMIGDINIEQSSSAPQEPFVNVYDIS-EHIQNFSFDE  945
             AD+          F          +I     ++  S P  P   + DI  E    FSFD 
Sbjct  532   ADENLQSETKELGFKLRGYGMPESLIPKYQKKKFESYPY-PTERISDIVVEENSKFSFDH  590

Query  946   QPNLDTHPGPSPNVLLQALTMSNANDGINLERQETVGDSFLKYAITTYLYKTYENIHEGK  1005
             QP L+ HPGPSP++LLQALTMSNANDGINLER ET+GDSFLKYAITTYLY T+E+IHEGK
Sbjct  591   QPKLEGHPGPSPSILLQALTMSNANDGINLERLETIGDSFLKYAITTYLYCTHESIHEGK  650

Query  1006  LSHLRSKHVSNLNLYKLGKLKNLAEYMVATKFDPHDNWLPPCFYVPKQLEDALIDAQFPA  1065
             LSHLRSK VSNL LY+LG+ K   E M+A+KF+PHDNWLPPC+ VP+  E  L+D++ P 
Sbjct  651   LSHLRSKQVSNLKLYRLGQKKMFGESMIASKFEPHDNWLPPCYLVPRDAEQRLLDSE-PQ  709

Query  1066  NCWTVADMAATRNMTIDEVCTLVRERGIYSLPNIIPYNLVTQHSIPDKSIADCVEALIGA  1125
                                               +PYNLVTQHSIPDKSIADCVEALIGA
Sbjct  710   TT--------------------------------LPYNLVTQHSIPDKSIADCVEALIGA  737

Query  1126  YLIECGPRGALLFMAWLGIKVLPKNE-DGS----------------------YGELDLPK  1162
             YLIECGPRGALLFM+WLGIKVLP+ E D S                      +  L  P 
Sbjct  738   YLIECGPRGALLFMSWLGIKVLPREEIDVSKRTSSSMKTVGRIKTNENTVVKFTTLSDPL  797

Query  1163  SPLLRNIPNPEEELEKLLDGYDTFEKHIGYKFRDRSYLLQAFTHASYSPNRLTDCYQRLE  1222
             SPLL+++ +PE EL  L+ G++  E  + Y FRD+SYLLQA +HASYSPNR+TDCYQRLE
Sbjct  798   SPLLQHVDDPEGELAFLMAGFEALESSLHYSFRDKSYLLQAMSHASYSPNRVTDCYQRLE  857

Query  1223  FLGDALLDFIITKALFEDSRMHSPGALTDLRSALVNNTIFASLAVKHGFHKYFKHLSPGL  1282
             FLGDA+LD++IT+ LFED R HSPGALTDLRSALVNNTIFASLAVK+G HKYF+HLSPGL
Sbjct  858   FLGDAVLDYLITRHLFEDKRAHSPGALTDLRSALVNNTIFASLAVKYGLHKYFRHLSPGL  917

Query  1283  NEVIERFVRLQEESGHTIVDEHYLVVDTECEEVEDVEVPKALGDVFESVAGAIYLDSGMS  1342
              EV++RFV +QEE+ H I +E+YLV     EE EDVEVPKALGDVFESVAGAI+LDSGMS
Sbjct  918   AEVVQRFVTIQEENSHLISEEYYLV----GEEAEDVEVPKALGDVFESVAGAIFLDSGMS  973

Query  1343  LDAVWKVYHRMMKTEIEQFSNKVPKSPIRELLELEPETAKFGRPEKLADGRRVRVTVQVF  1402
             LD VWKVY+RMMK EIEQFS KVPKSPIRELLELEPETAKF +PEKLADGRRVRVTV+VF
Sbjct  974   LDTVWKVYYRMMKKEIEQFSTKVPKSPIRELLELEPETAKFSKPEKLADGRRVRVTVEVF  1033

Query  1403  GKGTFKGIGRNYRIAKCTAAKCALKHLKRRGLLRNKND  1440
             GKGTFKGIGRNYRIAKCTAAKCALK LK   L R   D
Sbjct  1034  GKGTFKGIGRNYRIAKCTAAKCALKQLKIGLLARKLTD  1071
```

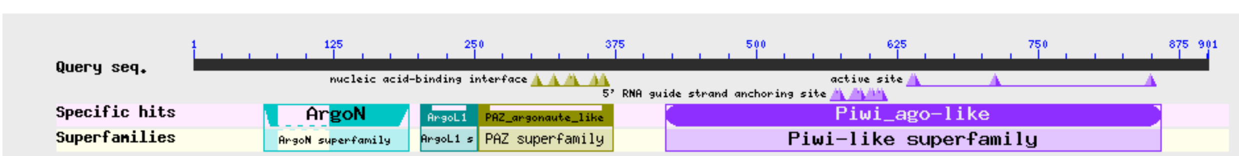
Ago-1

cDNA

Contig 7709, Length-4070
CTAGCAGCCATGTTGATAAACTTTGAGACCAGTTTATTTTAGGGCTTATAATTCCTCTATAGTTTTCTTCCGTCGTTATAGTGAAATAGAAACATGT
ATCCTGTTGGACAACCTCCTCCTGTGGGGCCACCAGGAGCACCTGCAGGAGGTGGAGCACCACCAGGCCCTGTTGGCGTGCCGCCTGGGGCAATAGC
TTTAGCTATACAGCAGCCACCACACCAGCCGCCCGATCTCCCGATGTTTACCTGCCCCCGAAGGCCAAATCTCGGCCGAGAAGGTCGGCCCATCGTG
CTTCGTGCCAATCATTTCCAGATAACTATGCCTCGTGGTTTTGTACATCATTATGAAATAAGCATACAACCTGACAAATGTCCTCGTAAAGTAAACC
GGGAAATTATTGAGACTATGGTTCATGCCTATTCGAAAATTTTTGGCAACTTGAAACCAGTTTTTGATGGAAGAAGTAATCTCTACACTCGTGATCC
TCTCCCAATTGGCAATGATCGGATGGAACTTGAGGTGACACTACCAGGTGAAGGTAAAGACCGTATATTTCGAGTTGCTATAAAATGGATGGCGCAG
GTCTCTCTATTTGCTCTAGAAGAAGCGTTAGAAGGAAGGACTAGGCAGATTCCTTATGATGCTATCATGGCCCTTGATGTAGTGATGCGTCATTTAC
CTTCAATGACATATACCCCTGTAGGAAGGTCATTCTTTTCTAGTCCAGATGGCTATTACCATCCTCTCGGTGGTGGTCGTGAAGTATGGTTTGGGTT
TCATCAGTCAGTTAGACCCTCGCAGTGGAAGATGATGTTAAATATTGATGTTTCGGCCACTGCATTTTATAAAGCACAACCTGTTATTGAGTTTATG
TGCGAAGTATTAGACATTCGGGACATAAACGAACAACGAAAGCCTTTAACAGACTCTCAGAGGGTCAAATTTACAAAAGAAATAAAAGGTTTAAAAA
TAGAAATAACACATTGTGGTTCAATGAGAAGAAAATACAGGGTATGTAATGTGACTCGAAGACCTGCTCAAATGCAGTCATTCCCTTTACAGTTGGA
GAATGGACAAACTGTAGAATGTACAGTGGCTAAGTACTTCCTTGACAAATACAAAATGAAGTTGAGATATCCACATTTACCATGCCTTCAAGTCGGT
CAGGAACATAAACATACATATCTTCCATTAGAGGTTTGCAATATAGTTGCTGGTCAGAGGTGCATTAAAAAGCTTACTGACATGCAGACCTCCACGA
TGATCAAAGCAACTGCTAGGTCAGCTCCTGACAGAGAAAGGGAAATTAATAACCTTGTGAGGCGGGCTGATTTCAATAATGATGCCTATGTCCAAGA
GTTTGGACTAGCCATCTCAAATAATATGATGGAAGTTCGAGGGAGAGTCCTCCCTCCACCAAAGCTGCAGTATGGTGGGCGAGCACCGAGCTTGCCG
AGTCAGCTCGAATACCAAGGAATGTTGTCAGCAAAACAGCAAGCTCTTCCAAATCAAGGAGTTTGGGATATGCGGGGAAAACAATTCTTCACAGGTG
TCGAAATTAGGGTTTGGGCGATTGCTTGCTTTGCACCCCAAAGGACTGTGAGGGAAGATGCCCTAAGGAATTTTACGCAGCAATTGCAGAAAATAAG
TAATGATGCTGGAATGCCTATTATTGGGCAGCCATGTTTCTGTAAATATGCAACTGGGCCTGATCAAGTAGAACCTATGTTCCGTTACCTGAAGTCT
TCCTTCCAAGCATTACAGCTTGTTGTTGTGGTACTTCCAGGGAAAACTCCAGTTTATGCGGAAGTGAAGAGAGTTGGAGATACAGTGTTGGGAATGG
CAACTCAATGCGTTCAGGCCAAAAATGTAAATAAGACTTCCCCTCAAACACTATCCAATCTCTGCCTCAAAATCAATGTTAAACTAGGCGGTATCAA
CAGCATTCTTGTCCCTAGTATCAGGCCAAAGGTATTCAATGAACCTGTCATATTCCTTGGGGCTGATGTGACTCACCCACCAGCTGGAGACAACAAG
AAACCTTCGATCGCAGCTGTTGTTGGATCCATGGATGCACATCCTAGCAGGTATGCTGCCACTGTTAGGGTCCAACAGCATCGCCAAGAAATAATCC
AAGAGCTCTCTTCTATGGTCAGGGAACTTCTCATAATGTTTTATAAGAGCACTGGAGGGTACAAGCCTCACAGGATCATTCTTTACAGAGATGGTGT
TTCGGAAGGACAATTTTTACATGTATTGCAACATGAATTGACAGCAATAAGAGAAGCCTGTATCAAATTAGAGGGTGATTACAAACCTGGGATTACT
TTTATTGTGGTGCAGAAGAGGCATCATACGAGGTTGTTTTGTGCTGATAAGAAAGAACAGAGTGGAAAATCAGGGAATATACCCGCTGGCACAACTG
TCGATGTTGGAATCACCCATCCAACTGAATTTGACTTTTACCTTTGCAGTCATCAGGGTATACAGGGAACAAGTCGACCTAGTCATTACCATGTCCT
ATGGGATGATAATCACTTTGATTCAGATGAATTACAGTGCCTTACTTATCAGTTGTGCCATACTTATGTGAGGTGTACACGCTCTGTATCCATACCT
GCTCCTGCTTATTACGCTCATCTTGTTGCTTTTAGAGCAAGGTATCATCTGGTTGAAAAGGAGCATGACAGCGGTGAAGGGTCTCACCAGTCAGGTT
GCAGCGAGGATCGTACTCCCGGTGCAATGGCTCGAGCTATTACTGTCCACGCGGACACAAAGAAAGTAATGTACTTTGCATAACCAAAGAAACGAAA
AAACTTCCTGCCACCATTTGATTGGAATTATCCGATACAGCCAGCTCTCGTTTCTAGCTTACGTAATATCTATTGTTTTTTAAAGAGTGGTTGTAAC
TGTGTCGGTTTTATATCGTAAGATTTATTATGTTAAGGTTAATATTTTACAGTTAAAAATAACAAAAATTAAAGCAGTGATAGTTATATTTTCTTAA
GAGTCTAATGAAGACATTTTCCCGGAATTATAATTTTAACCCGAGAAGATTTCACTAGATGGAAGAAAACTAATTGAGGAACGCTGCCGTAATTATT
TTAAGTCTTCATTAAGCTGGTGCACAAGTTAAAGCATGTAAATAAGGAAGAGTAGCCTGTAAGAGAGATTTTTGGAAAATTCAATTAAATGTAAGTT
GTAAGCTTTTTTTCTAAAAAAATATATGTGTTTATATATTTATTTTTAATTCCTCTCTATGTTCTCTCTTAGGCCTCTTATTTCCAAGTCAGCTGTT
AACGTGTAGCTTTAAACATTCTCATTAATATATATTTATATATATATACCAAATGGTTGTGTTAGTCATGTCTCCTGGCACTGTACTAAAGTCAGCG
TTAAGTTAGTTTGAGGTGATATAGAGTGCTCATTTAGTTGTACTTAAAAAAAAAATTTTTTTTAACTGTGGAATTAAACAATAACAGGTAATAGATT
TCCTTTATTTTACTGAATTGGATATTATCAAATAAACATGTTAAATTTTAATGAATGTTTACAAAGAAAAAAAAAAAAAAAAATTAAAGAATTATTTA
TTGTATGATGTTCTTGTTATTGAGAGATGGAACTATTATTTTGTTATTTAGATTTTTTAATTTTTGAAGTTATGCTGCAGATAATTACTGCAGAAAA
AAAACAAACTATGTAATGTTTTTATTGTAATATGAAGGATTAGGGAGTACGAAGTCCAGTAATCAATGCTTCATTTGGTGAACTAATTGTATCTTCC
ACAAACAAGGTTCAAAATTTATTTCTGTATGTATATATTAAACAATTAAAATTTACTTTGAAACCCAAAGTAATGTAAGATAAAGTATCATTTAAAA
AAATTTTTTTTTTAATTAGTAAGACCATGATTATGCGAACTAATTGATATGAAAGCAATGCTTCAAAAGTCAATTTGTTTGGATAATTGAATTCATT
GTGAAATAAATGTGCTTTAGTTAATTTTAGATTTATAAAAAAAAAAAAAAAAAGAAAAAAAAAAATAAATCACACCATATGATACTGTAGCAAGC

PROTEIN 901(AA)
MYPVGQPPPVGPPGAPAGGGAPPGPVGVPPGAIALAIQQPPHQPPDLPMFTCPRRPNLGREGRPIVLRANHFQITMPRGFVHHYEISIQPDKCPRKV
NREIIETMVHAYSKIFGNLKPVFDGRSNLYTRDPLPIGNDRMELEVTLPGEGKDRIFRVAIKWMAQVSLFALEEALEGRTRQIPYDAIMALDVVMRH
LPSMTYTPVGRSFFSSPDGYYHPLGGGREVWFGFHQSVRPSQWKMMLNIDVSATAFYKAQPVIEFMCEVLDIRDINEQRKPLTDSQRVKFTKEIKGL
KIEITHCGSMRRKYRVCNVTRRPAQMQSFPLQLENGQTVECTVAKYFLDKYKMKLRYPHLPCLQVGQEHKHTYLPLEVCNIVAGQRCIKKLTDMQTS
TMIKATARSAPDREREINNLVRRADFNNDAYVQEFGLAISNNMMEVRGRVLPPPKLQYGGRAPSLPSQLEYQGMLSAKQQALPNQGVWDMRGKQFFT
GVEIRVWAIACFAPQRTVREDALRNFTQQLQKISNDAGMPIIGQPCFCKYATGPDQVEPMFRYLKSSFQALQLVVVVLPGKTPVYAEVKRVGDTVLG
MATQCVQAKNVNKTSPQTLSNLCLKINVKLGGINSILVPSIRPKVFNEPVIFLGADVTHPPAGDNKKPSIAAVVGSMDAHPSRYAATVRVQQHRQEI
IQELSSMVRELLIMFYKSTGGYKPHRIILYRDGVSEGQFLHVLQHELTAIREACIKLEGDYKPGITFIVVQKRHHTRLFCADKKEQSGKSGNIPAGT
TVDVGITHPTEFDFYLCSHQGIQGTSRPSHYHVLWDDNHFDSDELQCLTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLVEKEHDSGEGSHQS
GCSEDRTPGAMARAITVHADTKKVMYFA

Graphical representation



Comparison with *Cylas puncticollis Ago-1*

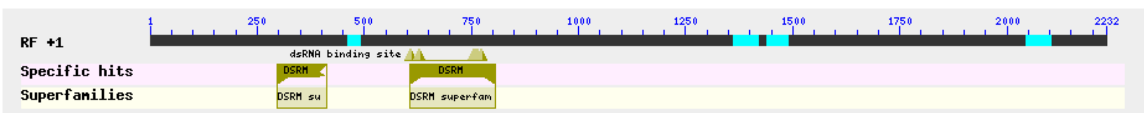
```
Query  33   SLTTVSPPAEPP---MFQCPRRPNLGREGRPIGLKANHFQISMPRGFVHHYDINIQPDKC  89
            +L    PP +PP   MF CPRRPNLGREGRPI L+ANHFQI+MPRGFVHHY+I+IQPDKC
Sbjct  34   ALAIQQPPHQPPDLPMFTCPRRPNLGREGRPIVLRANHFQITMPRGFVHHYEISIQPDKC  93

Query  90   PRKVNREIIETMVKSYGKIFGNLKPVFDGRNNLYTRDPLPIGNSREELEVTLPGEGKDRL  149
            PRKVNREIIETMV +Y KIFGNLKPVFDGR+NLYTRDPLPIGN R ELEVTLPGEGKDR+
Sbjct  94   PRKVNREIIETMVHAYSKIFGNLKPVFDGRSNLYTRDPLPIGNDRMELEVTLPGEGKDRI  153

Query  150  FRVSIKWVAQVSLYGLEEALEGRTRQIPYEAILALDVVMRHLPSMSYTPVGRSFFSSPEG  209
            FRV+IKW+AQVSL+ LEEALEGRTRQIPY+AI+ALDVVMRHLPSM+YTPVGRSFFSSP+G
Sbjct  154  FRVAIKWMAQVSLFALEEALEGRTRQIPYDAIMALDVVMRHLPSMTYTPVGRSFFSSPDG  213

Query  210  YYHPLGGGREVWFGFHQSVRPSQWKMMLNIDVSATAFYKAQPVIEFMCEVLDIRDINEQR  269
            YYHPLGGGREVWFGFHQSVRPSQWKMMLNIDVSATAFYKAQPVIEFMCEVLDIRDINEQR
Sbjct  214  YYHPLGGGREVWFGFHQSVRPSQWKMMLNIDVSATAFYKAQPVIEFMCEVLDIRDINEQR  273

Query  270  KPLTDSQRVKFTKEIKGLKIEITHCGAMRRKYRVCNVTRRPAQMQSFPLQLENGQTVECT  329
            KPLTDSQRVKFTKEIKGLKIEITHCG+MRRKYRVCNVTRRPAQMQSFPLQLENGQTVECT
Sbjct  274  KPLTDSQRVKFTKEIKGLKIEITHCGSMRRKYRVCNVTRRPAQMQSFPLQLENGQTVECT  333

Query  330  VAKYFLDKYKMKLRYPHLPCLQVGQEHKHTYLPLEVCNIVAGQRCIKKLTDMQTSTMIKA  389
            VAKYFLDKYKMKLRYPHLPCLQVGQEHKHTYLPLEVCNIVAGQRCIKKLTDMQTSTMIKA
Sbjct  334  VAKYFLDKYKMKLRYPHLPCLQVGQEHKHTYLPLEVCNIVAGQRCIKKLTDMQTSTMIKA  393

Query  390  TARSAPDREREINNLVRRADFNNDEYVQEFGLTISNNMMEVRGRVLPPPKLQYGGRVASL  449
            TARSAPDREREINNLVRRADFNND YVQEFGL ISNNMMEVRGRVLPPPKLQYGGR  SL
Sbjct  394  TARSAPDREREINNLVRRADFNNDAYVQEFGLAISNNMMEVRGRVLPPPKLQYGGRAPSL  453

Query  450  SGQ--------SKQQACPNQGVWDMRGKQFFTGVEIRVWAIACFAPQRTVREDALRNFTQ  501
              Q        +KQQA PNQGVWDMRGKQFFTGVEIRVWAIACFAPQRTVREDALRNFTQ
Sbjct  454  PSQLEYQGMLSAKQQALPNQGVWDMRGKQFFTGVEIRVWAIACFAPQRTVREDALRNFTQ  513

Query  502  QLQKISNDAGMPIIGQPCFCKYATGPDQVEPMFRYLKTTFQSLQLVVVVLPGKTPVYAEV  561
            QLQKISNDAGMPIIGQPCFCKYATGPDQVEPMFRYLK++FQ+LQLVVVVLPGKTPVYAEV
Sbjct  514  QLQKISNDAGMPIIGQPCFCKYATGPDQVEPMFRYLKSSFQALQLVVVVLPGKTPVYAEV  573

Query  562  KRVGDTVLGMATQCVQAKNVNKTSPQTLSNLCLKINVKLGGINSILVPSIRPKIFNEPVI  621
            KRVGDTVLGMATQCVQAKNVNKTSPQTLSNLCLKINVKLGGINSILVPSIRPK+FNEPVI
Sbjct  574  KRVGDTVLGMATQCVQAKNVNKTSPQTLSNLCLKINVKLGGINSILVPSIRPKVFNEPVI  633

Query  622  FLGADVTHPPAGDNKKPSIAAVVGSMDAHPSRYAATVRVQQHRQEIIQELSSMVRELLIM  681
            FLGADVTHPPAGDNKKPSIAAVVGSMDAHPSRYAATVRVQQHRQEIIQELSSMVRELLIM
Sbjct  634  FLGADVTHPPAGDNKKPSIAAVVGSMDAHPSRYAATVRVQQHRQEIIQELSSMVRELLIM  693

Query  682  FYKSTGGYKPHRIILYRDGVSEGQFLQLLQHELTAIREACIKLEADYKPGITFIVVQKRH  741
            FYKSTGGYKPHRIILYRDGVSEGQFL +LQHELTAIREACIKLE DYKPGITFIVVQKRH
Sbjct  694  FYKSTGGYKPHRIILYRDGVSEGQFLHVLQHELTAIREACIKLEGDYKPGITFIVVQKRH  753

Query  742  HTRLFCADKKEQSGKSGNIPAGTTVDVGITHPTEFDFYLCSHQGIQGTSRPSHYHVLWDD  801
            HTRLFCADKKEQSGKSGNIPAGTTVDVGITHPTEFDFYLCSHQGIQGTSRPSHYHVLWDD
Sbjct  754  HTRLFCADKKEQSGKSGNIPAGTTVDVGITHPTEFDFYLCSHQGIQGTSRPSHYHVLWDD  813

Query  802  SHLDSDELQCLTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLVEKEHDSGEGSHQS  861
            +H DSDELQCLTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLVEKEHDSGEGSHQS
Sbjct  814  NHFDSDELQCLTYQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHLVEKEHDSGEGSHQS  873

Query  862  GSSEDRTPGAMARAITVHADTKKVMYFA  889
            G SEDRTPGAMARAITVHADTKKVMYFA
Sbjct  874  GCSEDRTPGAMARAITVHADTKKVMYFA  901
```

Loquacious
cDNA
Contig 1167, Length-2231
CAATTGCAGCTGTTTATTCCGCCATATTTCTGAAGTTGTACAGTATTTTGGTACCGGAGCTCGGATATTTCTAAACCTATTACTGTACAAATACTTG
AAATATTGATGGCCGAAGCTACTGTGAAGATTGAAGATAATTTGAGTGGGCATGTGGGACTATCAAGGCATACACCAACATTACGTGGGGGACGAAG
TCGGATGAGAAATATTCCATCTTGTAAATTTTCCCCACCTGAGCCTGTTCCCCTGGAAGAGGCTGCAAAAAGGGAGTTGCAAACTATGCCCACCAAA
ACCCCAGTTTCTATGCTCCAAGAGCTTCTCTCGAGAAGGGGTACAACACCAAAATATGAACTGGTTCAAGTTGAAGGAGCAGTTCATGAGCCGACCT

TTAGGTACACAGTTTCTGTTGGTGAAACCGTAGGATCCAACAGATTTGCTGCAATGGGTACTGGTAGATCTAAAAAGGAAGCCAAGCACGCTGCAGC
TAAAGCAATTTTAGATAAGATCACAGGTAATACTGAACATCATCTTGAAATTTCTAGTCCTACTTCAACAGCTGATAGTTTAAGTCCAAATATTGGA
ACTGTTGATGAACGAATGGTTGGAAACCCAATAGGCCTTCTGCAGGAACTTTGTATGACAAGGAGATGGCCTCCACCTGCCTACGAGACAGAAAATG
AGGAAGGTCTTCCCCATGAAAGACAGTTTACCATCTCCTGCCTTGTTTTCCGCCACAAAGAAATAGGGACTGGCAAATCCAAAAAAATTGCCAAGAG
GTTAGCCGCGCATAAAATGTGGATGAGGCTAAGAGACTTACCCTGTGAAAATAATACTATCAATTATGGCTTAGATGATGATGATGAAATTGGCCAA
CGGATTCCGCAAAGTGAGTTTGACTCATTGAAGATGATCAAAGTCCCTTCTCTCTCCTCAGCGCCATACAACATCAAAGTGTCACAATTCCATAAGA
ACCTCAAATCATCTTGTGGGCCTAAACTTGAAGAGTTGCAGATAATTTCATTAAAAGAGCCTGGTTTTGATGCAGTTCATTTCCTGCAGGAAATAGC
CACAGAGCAACAGTTTGAAGTAACTTTTGTTGATGTCGAGGAGAGGTCTGTTTCAGGAAAATGCCAGAGTTTGGTCCAGCTGTCTACATTACCTGTT
GCTGTATGTTACGGTACAGGAATGTCCTCATCAGAGGCCCAAATGTGTGCAGCCTACAATGCACTTCAGTACCTTAAGCTTATGACGAAGAAATGAA
ATGTTAAACCAAATATCGCGAACTGTGTACCTACTATTCTCCTTGTTGCCATTTTTCCCTTATTCTCCCTGGCTATTTCAGCTCTTTATGAAATCAA
GTGATAACTGCGAATCAGAATATCACTTACTTAAATTATTTTATTTAGTAGTATAGTAATAATTGAAAATTAGAATGATTGATAACTTATATTAATC
ATATACCTGTGATTTAAATTAATCTTTTTATAGTTCTCGTGGAACGTCCCAGAGTGCTGAAACTAATAATAGTTAGATGCCAGTAAGTTTTGCTGCC
TTAACCAACGCGTTTATTCACAAATTTTTATTCCAAAAGACATTCTATTTAAGAGAGTCAAAAATAAAACATTTACTTTCCATCATTTAAAAAAGTG
TCATTACTTATTTAGGATTTAAGTCATTTTAATCAACCTCATAAAATATTACAACTCACACAAATCTGTTATAAAAAGTGGCTTGATTGATCCCTCC
ACAATTATAGTATTGAATCATTTAATTATTTCTCTAGAACCTTCACATCATGATTAAAAGTTAAGATTGTATAATGGCATGACTCTATATTATCAGT
TCATGTATTTGAAATTCTACTGTGAATTCAAGGTATACCTAATTTATATTCCCAACAGCTGATGTGTTATGTTACTTTTATTAATGACAAAACTTTG
TTTTGATTAAGACAAGCAGCGGCTAGTGCAGGTATTTGCCTTATTTTGTAAACAAACTCATAGTGTTCTTAATGAAAAAAAATTAACATTTTGCAGT
AATCATTTTTTAATGTGGATTTTTTGTTTCTTTTTTAAAATGTTTAAAATCTTTTACTTTCATTTCTTATACAGAAAATTAAATTCTTTTTTTTTGT
GTTATTGATCATAGGTTTTGGAAGTTAGCCATTGTTTATTTTAAATCGATCTGGTATATTATTTAATGAGATTGTTATTGTTTTTTGTTTGTTTGTG

PROTEIN 384(AA)
MAEATVKIEDNLSGHVGLSRHTPTLRGGRSRMRNIPSCKFSPPEPVPLEEAAKRELQTMPTKTPVSMLQELLSRRGTTPKYELVQVEGAVHEPTFRY
TVSVGETVGSNRFAAMGTGRSKKEAKHAAAKAILDKITGNTEHHLEISSPTSTADSLSPNIGTVDERMVGNPIGLLQELCMTRRWPPPAYETENEEG
LPHERQFTISCLVFRHKEIGTGKSKKIAKRLAAHKMWMRLRDLPCENNTINYGLDDDDEIGQRIPQSEFDSLKMIKVPSLSSAPYNIKVSQFHKNLK
SSCGPKLEELQIISLKEPGFDAVHFLQEIATEQQFEVTFVDVEERSVSGKCQSLVQLSTLPVAVCYGTGMSSSEAQMCAAYNALQYLKLMTKK

Graphical representation



Comparison with *Cylas puncticollis* Loquacious

```
Query  19   GGPIHLRRGKNPKHSMASTITLAEEAKLVYSQEMSGINSKTPVSVLQELLSRRGSTPKYE  78
            GG    +R    + K S     + L E AK     +E+  + +KTPVS+LQELLSRRG+TPKYE
Sbjct  27   GGRSRMRNIPSCKFSPPEPVPLEEAAK----RELQTMPTKTPVSMLQELLSRRGTTPKYE  82

Query  79   LVQIEGAIHEPVFRYRVFL-----SNDLVATGTGRSKKDAKHAAAKNLLDLIVGKQTPEQ  133
            LVQ+EGA+HEP FRY V +      SN   A GTGRSKK+AKHAAAK +LD I G    E
Sbjct  83   LVQVEGAVHEPTFRYTVSVGETVGSNRFAAMGTGRSKKEAKHAAAKAILDKITGNT--EH  140

Query  134  ANQTNGTPGSTDITAQVVSPFDDKVMGNPIGWLQEMCMSRRWPPPLYEMEQEEGLPHERQ  193
              + +     + D   +   +    D++++GNPIG LQE+CM+RRWPPP YE E EEGLPHERQ
Sbjct  141  HLEISSPTSTADSLSPNIGTVDERMVGNPIGLLQELCMTRRWPPPAYETENEEGLPHERQ  200

Query  194  FTIACQVLKFREVGTGKSKKLAKRVAAHRMWQALQDLPMEGTNVQTFESDEELTGNAVDI  253
            FTI+C V + +E+GTGKSKK+AKR+AAH+MW  L+DLP E    +     D++  G  +
Sbjct  201  FTISCLVFRHKEIGTGKSKKIAKRLAAHKMWMRLRDLPCENNTINYGLDDDDEIGQRIPQ  260

Query  254  SVRYVGLKDSKITKLSTQQSQ-KVSQFHKGLKQSAGKKLNELQSLNMKE--FNFVQFLQE  310
            S   +  LK  K+  LS+     KVSQFHK LK S G KL ELQ +++KE  F+ V FLQE
Sbjct  261  S-EFDSLKMIKVPSLSSAPYNIKVSQFHKNLKSSCGPKLEELQIISLKEPGFDAVHFLQE  319

Query  311  IASEQNFEVTYVDVEEKSLSGRCQCLVQLSTLPVAVCFGTGKTPKDARANAALNALEYLK  370
            IA+EQ FEVT+VDVEE+S+SG+CQ LVQLSTLPVAVC+GTG +  +A+A  AA NAL+YLK
Sbjct  320  IATEQQFEVTFVDVEERSVSGKCQSLVQLSTLPVAVCYGTGMSSSEAQMCAAYNALQYLK  379

Query  371  IMTKK  375
            +MTKK
Sbjct  380  LMTKK  384
```

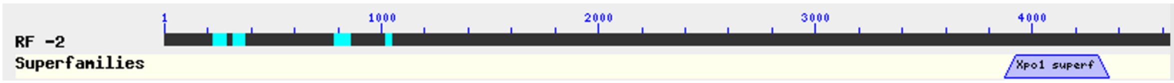
Exportin-5
cDNA

Contig 4070, Length-8269

CAGCAAACAAAAATAACTTCAATGGGTTCAAACTGAATTAGTCAAATTGAAGAAAAATTAAATGCACTCAATTCCTTACAAAGAAATTAATAATTTA
ATAAACATAATATACCTGAATGCAATAAACAGTCATTTTCTTTCTTTTGTAAAAATAAAAACAAGATAATGTTTTATCCATTACGTTTAACCTATTA
TAACAATTCCTTTTGATTGATAATGACCAATAGTAATACATATATACATAGTAACTCAGACTTATATAATACATACACAAGTATTTATTAACACTCT
AATCTAACATTCCAAATGTTACATATATTACAACAAAACAATAATAGAAATATATATACATATGCCTATACACAACGAACATTTACATGTTGTATTA
AGAGGGGTACAGGGTTTCCCAGTCCAGTATACATTGCCCAAAACATAACTTTGAGACTTTTAAGCAAATATGTTTTGCATAAAAACAAAATTCATAG
GGAAGCTTATTTTTTATGGAAAGAGTATTCGTATTTTATGAAATATTAGGATATGTACACTGTTTTTCGTTTAACTATAAATATCAGTTAAATACAA
AATAATAAAACCCAAGGGAAGTACTCTGACAACTATGTACTTCATTGGCCCCTACATTGTGAGAATGCTTTCAGTCAATTTCAAAACAAGAACCGTG
ATTTAAAAAATCTGGGTTTTATACGACTTGCTTTGAGAGGAAATAAATTTCATATTACAGGATAAAAAAAATTACGAAGTTTGAACAAAAACTTTTG
AATTTAACAAAGAATGAATAATTAATAAAAATTAATCAATAATAATTAGATTTATAAATTATGATATACTGATTTAAAAACCATTTAAATTTTGGAC
TAATATTCCAAGCAGAACTAAAGTTTATTGTGATTTTAGTTTTTGTAATACATTCAAACTGGAATGTAAATCATAACAATAAAAAACTCTTAGAACT
AAAAACAAAAAATAAGTTAATTTTGATCACATTTTGAATTTAAAATAAACCACAGAAGATAATAAAAGATAGGAAAGGAAGTTAACAGCATCACAAC
ACTGGAGAGACTAGATGAATAGGATAAGTGAATTGGATGGCATCTGATGCTACGTATTAAACAGTTTCACAAGACCATTTTCTTCTCCATTAATCAA
AAGATCTGTATGTAGTTTTGATTTAGAATTTTCAATTTTAGGCAAGTCCCGAATAGCAACCTTCTTCCGGAAAAGCTGACCAACATTTCTTCCAATC
AAAGGGCTAGTTACTTTTTTAAACATTTCTTTCTTAGCTTTATCAACTTTAGATCCTTTTTGAATTTCTTTCATCATCCTATCATCTAATTTTTGTA
AATCAATCATATTAACATCTGGGATCTGTTTCATCACTTGAAGGATGGGAGGAGCCATAGGGCGAAGAATCCCATACATAAGTGTTCCAAGAACAAG
TAAAGAGCCTTGGTTTGCATCATGCTGACCATGACACTGCAGACCAAATAAGACGGAAGTTAGAGCATGACATGCCAGGTCAGAAGTCATTTTCCCA
GTAGAAGCGAGATGAAACATCACACCACCCAGGAGTGAAGTACCTTTGAGGCTGGCAGTACTGTCTCCCCAATATAAAGCTCTTAACAAACAAAGCA
AAATAGGTTGAGAAGTCGGATCAAAAGCAAGGACTCTCATGCCAAGTTCACTTATAACATCAGATTGTGGCTTGACATCCTGGCCTTCTTCCTCCAT
ATTTCCTCCACTGGCGGTACCTGTATCAGTTCCTCCATAGAGAACAACTTTCAGTACATCTAAATACTCCCTTGTCAATGTTCTGTTCAGCATATCT
TCTAACATCTCCTGAGTATCTGTATTTTCTTCATGCTCATCAACATTACCATTTTCACCACTGATTTCAGATAAATGCTGCCATCTAGCACTGAGCC
TAGTAAACATATATGGAGTAAAATGAGCAAGAACTGGAATCAAAACAGAATCATAACATGGTGGAGGACAGCTAGTAACAAACGATTTCAAAAACAC
CCGAACAAATGGTCTCATCCGGTGATCTGGAACTAACTCCAGGTTTGAAAATACAGTAGCTATAAGAGAATGCGCTAGGTTCGGTACATGATATATA
TCATGGCTCAAGGCATGAAATGCATGCGCTAAGGTATGGCAACAATCCTCATTCACATTTGATATAAAGTTCTGCATACGTTGAAGGGGTGTTGTTT
CACCTTTTGACTCATTATTCATAACTTCACTTGTACTTCCAATCCCCAACAAATTCTGTCTCTCCAATTCTAGCATTGTATATGCACCTCGGTATCC
CTCAGATAAATTGGTCAGTGCAGCAGGAGAAGACAAACCATTCAGAACACAAGAAAGGCTAAAGAGTCCTGGAAGAAGTGGAAGAACATGCGGCGTG
GCAGGATTCCTATAGATTGGATTACCAGCCTCTGTGGTACAAACCAAAAAACCACCCCTTGCAGCAAGTTCTGGATCACTAGGCCACATGACACGCT
TCATAACGGCTAAAATAACACCCAAGCACCACATGATCTGGGATCTGTTTTGACCATTGATATCATCAGAACTAGGCTCGACAGGTGGCTTATCCAA
TCCTACATACGACATAAAGGCTGGAGTGGAAAGGAATGCTGTTGAAGCCATATTGGTCCATGGTCCGATGACAGGGTTCATTACTTCACCTATGAAC
CTGCATTCACGTTCATAGTCACATAAGTGATTACTGATCAGGAGCAGTGCTTCCTGGAGACAAAGAGCCTCCATAGTTGACAAGGTGTGAGGTTTGG
CCTGGAGACTGACAACAATGCCATGTATCCTAGAAAAGACTGGTAGGAGAAGGAGAGGATACTTCTGGGCAATTTTCACCATCAAGGAAGCAGCATG
GCGACGAAGATTCTTAACGCTTCTTGACCTTTGGTTTTTTGGCTCTCCAGGGAGTGTAAAAACAAGAGTTGCAAATATTTTATCCAAACTTTTGGT
AAATAAATAGCAGTAGTTTCAGGAGGAGCCATGCTGAGAAAGACGAACAAGGCCGATATACAAGACAACATAGCTGAGAGAATGAGAGGATCTGTTG
GCTCTAATGCTAAGCAAAGAACCAAAAGTTTAAGACCATTATTAACACCAGGTCTCTCCTTGGCCATAACAAGCTTCCCGAGTACTGCATCAAGAGC
CAATGATAAAGCTTCCCATTCTAAGTAAGCAACAGAAGTTAAAGTGCAAGGCTCCATGTTATTTTGAGATACCGTCTTTTGCATTTGGACTTCTAAC
CACCTCTCAACATAAGAATAAGCAACCAATGGTGAAACTAAAGTGGCCTGTTTAAATGTATCAAGAAGCTCTACACGATGTTTATGAAAAAAAGCAT
TGTACTCATCTTCTGCATCAAAATCTTGAACAGCATAAGCTTCTGGAGTATTAGGTTCCAAATTAGATTTGGATGATGGAAATGGAACCTTCATAAT
TTTTGGTCCAGCGGCCTCAATCCATTTAGGTACAAAACTAAGAAATACTTCATCTTTGGAAATTTGGTCATGCTTCATCAAAGCTAGCCAAAGTGAA
TTAGCATAACAAACTAGACTAAGAGAAGGGTGACGGGTAAATGTAACTATTGCTTCAAGGTACATTTGGAAATTAGGAGGGCGTGATTGACTATCTT
TAGTCCACAATGAGGACAACTGTGTTCCCAAGCCAGTCAAAACACGAGTAAGTTTTTTTAAAAACTGGTAGTTCTTATTTTCAGAAGTATCAGCTGA
TTGAAAAATACATCTCATAGCATCTTCACTGAATAATATAAGAAGGGGCCTTCGTTCATCAAGTTTACCTTTCCTGGAAACAATCTGAAGAAGACAT
TCTGCAGAGCCATCTTGGAAATTTGCATCATTTAATAAAACACAAAATATTTGTAATAACCTTCCCATCTTCAGCCATTATGTGAGTCATAGAAACCC
ATTCCACAAAACCACTTAATGTTAACAGAACAACCTGCACAACACGAGCATGAGAATTCGCCAGCATAACATTTCCTAATTCAATGTTTTTACGATA
TTCTTCATAGTGTTTTCCAATAAGACCTCGAAAGAATATAAATAAATCTGCCATGTTTGTTGTTAATGCCTGATATAAATCTTTTCTGCGTTGGTTT
GACTCTAATGTCTGTAGCACAGCAACATCTTCAACAAGTCTGAGTAAAATTAGCAAAACCATTTCTGTTTGAATATTCCCTTGTGAAGCACACTGAT
TCAGTTCTGCCAATAATGTGTTCCACTGTTGTGGCCACTCCCTCTTGATCATTTCTACTACCACTCGAGACAGAGCATCTCTCATGTGAGCCTTATC
TTCTAAAGTGTTGTCCATACCACGTTCAACCAATTTCATGGCATTTTCCTTAATAAATAGTTTCTCAGCTTGAGTCATATTGTACCATCGGTATTTT
ATGCAATGCTCCATTAATTGCAATCCGAAATGCCTTACAAAATAAGAATAATCTGCTCTCTGAGCAAGATATAAACCACATTGTACACATAATGGGC
TCTTCTCTTTAAATACTTCACATGCATTGTAAGCTTCTAATCTTTCAGTCTGTGATGCAAGTGGATTCATCATAAGCTCCACAGCAGCAGCCAGTTG
TGTAGCAACTGCACTGATCTCTACATCCATGGTTAATTTAATGTTTTCTAACTGTATAAAAAAAAAGGAAGGAATTATACAAATTGTGTGTTGCTTT
AAAAATAACACAGATGACACTGTTTCAGACTAGTCAATGTCCATAGGAGCATGTGAGGACCATGCCTTCACTCTTTTAAATGAAACAGTGAAGTTTA
GGAACTGCTTATAAAACCTTAATTTGTCACATACAGCTTAACATAACAATATATATGGAAATATTTGAATATTCTGGTGTATTGCCACAAGTTACAT
CCTTAGATGAAGATAATTCTTACGTAAATGAAACGAATCCTAGGTCATGCTTGATATGTATGCGGTCTAACCTTTCATATTTGACTGGACTAATCTC
AACTGCTTCTCTGTTTTGTTG

Protein 1889 (AA)
MDVEISAVATQLAAAVELMMNPLASQTERLEAYNACEVFKEKSPLCVQCGLYLAQRADYSYFVRHFGLQLMEHCIKYRWYNMTQAEKLFIKENAMKL
VERGMDNTLEDKAHMRDALSRVVVEMIKREWPQQWNTLLAELNQCASQGNIQTEMVLLILLRLVEDVAVLQTLESNQRRKDLYQALTTNMADLFIFF
RGLIGKHYEEYRKNIELGNVMLANSHARVVQVVLLTLSGFVEWVSMTHIMAEDGKLLQIFCVLLNDANFQDGSAECLLQIVSRKGKLDERRPLLILF
SEDAMRCIFQSADTSENKNYQFLKKLTRVLTGLGTQLSSLWTKDSQSRPPNFQMYLEAIVTFTRHPSLSLVCYANSLWLALMKHDQISKDEVFLSFV
PKWIEAAGPKIMKVPFPSSKSNLEPNTPEAYAVQDFDAEDEYNAFFHKHRVELLDTFKQATLVSPLVAYSYVERWLEVQMQKTVSQNNMEPCTLTSV
AYLEWEALSLALDAVLGKLVMAKERPGVNNGLKLLVLCLALEPTDPLILSAMLSCISALFVFLSMAPPETTAIYLPKVLDKIFATLVFTLPGEPKNQ
RSRSVKNLRRHAASLMVKIAQKYPLLLLPVFSRIHGIVVSLQAKPHTLSTMEALCLQEALLLISNHLCDYERECRFIGEVMNPVIGPWTNMASTAFL
STPAFMSYVGLDKPPVEPSSDDINGQNRSQIMWCLGVILAVMKRVMWPSDPELAARGGFLVCTTEAGNPIYRNPATPHVLPLLPGLFSLSCVLNGLS
SPAALTNLSEGYRGAYTMLELERQNLLGIGSTSEVMNNESKGETTPLQRMQNFISNVNEDCCHTLAHAFHALSHDIYHVPNLAHSLIATVFSNLELV
PDHRMRPFVRVFLKSFVTSCPPPCYDSVLIPVLAHFTPYMFTRLSARWQHLSEISGENGNVDEHEENTDTQEMLEDMLNRTLTREYLDVLKVVLYGG
TDTGTASGGNMEEEGQDVKPQSDVISELGMRVLAFDPTSQPILLCLLRALYWGDSTASLKGTSLLGGVMFHLASTGKMTSDLACHALTSVLFGLQCH
GQHDANQGSLLVLGTLMYGILRPMAPPILQVMKQIPDVNMIDLQKLDDRMMKEIQKGSKVDKAKKEMFKKVTSPLIGRNVGQLFRKKVAIRDLPKIE
NSKSKLHTDLLINGEENGLVKLFNT

Graphical representation



Comparison with *Cylas puncticollis* Exportin-5

```
Query  9     KKLIQVFGGLTIQICTLWGKDPSFQPHNFSMFLEAILAYSSHSSLFLAHLANPLWNSMLK  68
             KKL +V  GL  Q+ +LW KD   +P NF M+LEAI+ ++ H SL L   AN LW +++K
Sbjct  315   KKLTRVLTGLGTQLSSLWTKDSQSRPPNFQMYLEAIVTFTRHPSLSLVCYANSLWLALMK  374

Query  69    QEHVARDLTFLSFLPKWVQCTAPKIIKFNYPAGKSPTLGEDISETVAYAKIDFDSEEEFS  128
              + +++D  FLSF+PKW++   PKI+K  +P+ KS  L  +  E  AYA  DFD+E+E++
Sbjct  375   HDQISKDEVFLSFVPKWIEAAGPKIMKVPFPSSKS-NLEPNTPE--AYAVQDFDAEDEYN  431

Query  129   AYFYRCRSDFLDSFRQATLVAPLITFNYVEQWLLKCLQ---VPNTTFGLSVSDPVYQEWE  185
             A+F++ R + LD+F+QATLV+PL+ ++YVE+WL   +Q     N     +++   Y EWE
Sbjct  432   AFFHKHRVELLDTFKQATLVSPLVAYSYVERWLEVQMQKTVSQNNMEPCTLTSVAYLEWE  491

Query  186   ALANFLESVLSRVLQAHERPSIPSGLRLLEMCLRYQPADPLIMSTLLTCISALFVFLSMS  245
             AL+  L++VL +++ A ERP + +GL+LL +CL  +P DPLI+S +L+CISALFVFLSM
Sbjct  492   ALSLALDAVLGKLVMAKERPGVNNGLKLLVLCLALEPTDPLILSAMLSCISALFVFLSM-  550

Query  246   TGQMAPAANSIAVSGAALLPQVLEKIFSTLVYNPNNQSKEVRSRPLKNVRRHAASLMVKI  305
                 AP   +     A  LP+VL+KIF+TLV+    + K  RSR +KN+RRHAASLMVKI
Sbjct  551   ----APPETT-----AIYLPKVLDKIFATLVFTLPGEPKNQRSRSVKNLRRHAASLMVKI  601

Query  306   GNKYPLLLLPVFDQIRTTVENLSKPDGPAQLSTLEKVTLREALLLISNHFADYERQSAFV  365
               KYPLLLLPVF +I   V +L     P  LST+E + L+EALLLISNH  DYER+  F+
Sbjct  602   AQKYPLLLLPVFSRIHGIVVSLQAK--PHTLSTMEALCLQEALLLISNHLCDYERECRFI  659

Query  366   GEVLREASAQFISIVNSGALKGAASFISFVGLDKPSVPANTKDICGHNRSNLVFCVNLVL  425
             GEV+      + ++ ++  L   A F+S+VGLDKP V  ++ DI G NRS +++C+ ++L
Sbjct  660   GEVMNPVIGPWTNMASTAFLSTPA-FMSYVGLDKPPVEPSSDDINGQNRSQIMWCLGVIL  718

Query  426   GAIKRCSWPDDPERATRGGFVVSLTESGNPVCRNPAAPHVIPLLPHMLSLICIFNELFTN  485
               +KR  WP DPE A RGGF+V  TE+GNP+ RNPA PHV+PLLP + SL C+ N L +
Sbjct  719   AVMKRVMWPSDPELAARGGFLVCTTEAGNPIYRNPATPHVLPLLPGLFSLSCVLNGLSSP  778

Query  486   EAQNLIHQSYKGCLAMPETEKSNLLGLIGHSSD--NGDNLIIQSPLERMQRFLTSLHESF  543
              A   + + Y+G   M E E+ NLLG IG +S+  N ++    +PL+RMQ F+++++E
Sbjct  779   AALTNLSEGYRGAYTMLELERQNLLG-IGSTSEVMNNESKGETTPLQRMQNFISNVNEDC  837

Query  544   YHLIGSMGPSLGRDLYNIPDLALAVINSVLAYLQYVPDYRIRPIIRVFLKSFIFACPVPY  603
              H +     +L  D+Y++P+LA ++I +V + L+ VPD+R+RP +RVFLKSF+ +CP P
Sbjct  838   CHTLAHAFHALSHDIYHVPNLAHSLIATVFSNLELVPDHRMRPFVRVFLKSFVTSCPPPC  897

Query  604   YEIVILPVLGHITPIIFNRLHSNWQRVIEFR----NREVQEDNADTQEVLEDILTRALTR  659
             Y+ V++PVL H TP +F RL + WQ + E      N +  E+N DTQE+LED+L R LTR
Sbjct  898   YDSVLIPVLAHFTPYMFTRLSARWQHLSEISGENGNVDEHEENTDTQEMLEDMLNRTLTR  957

Query  660   EYLDLLKVALVGGSLTPESTSETMETEELSNSPTPPPVRSNLTSEVISDLGLILLRNTKT  719
             EYLD+LKV L GG+ T  ++   ME E     P          S+VIS+LG+ +L    T
Sbjct  958   EYLDVLKVVLYGGTDTGTASGGNMEEEGQDVKPQ---------SDVISELGMRVLAFDPT  1008

Query  720   CQPIVLAVLGALSWIDSNASLKATLLTGPIVRQLSQDNSLNGEMATHIMAAILNALMLHG  779
              QPI+L +L AL W DS ASLK T L G ++  L+    +  ++A H + ++L  L  HG
Sbjct  1009  SQPILLCLLRALYWGDSTASLKGTSLLGGVMFHLASTGKMTSDLACHALTSVLFGLQCHG  1068

Query  780   QHEANQGSLLTLGAQIYELLRPSFVEVLGVMQQIPGVNPVDLQKLDERISGSTSKGNKVE  839
             QH+ANQGSLL LG  +Y +LRP    +L VM+QIP VN +DLQKLD+R+     KG+KV+
Sbjct  1069  QHDANQGSLLVLGTLMYGILRPMAPPILQVMKQIPDVNMIDLQKLDDRMMKEIQKGSKVD  1128

Query  840   KVKKDLFKKITANLIGRSVGQLFKKEVKIQDLPQLVIPKKAEKSEVIGDL  889
             K KK++FKK+T+ LIGR+VGQLF+K+V I+DLP++    +  KS++  DL
Sbjct  1129  KAKKEMFKKVTSPLIGRNVGQLFRKKVAIRDLPKI----ENSKSKLHTDL  1174
```

piRNAi pathway *Megacopta cribraria*

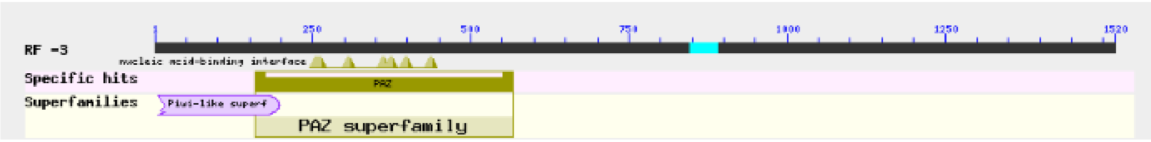
AGO-3

cDNA

Contig 1655, Length-3033
CCAGCACCAAACAGAATATCAATAGGTGGAATATTACGACCTTCAAGAACTAAATAGGCATCATCTAGACTCATGCCCCAATCAGATAATACTTTGG
TAGCTTCAGGAGAAACTCTGATATTATTAATAAACTGTTTTATAGCACGATGTCTTTGGTTAGGGGAGATACTAATACACTGAGTAATATCTTTCAT
AACTCTAGCATCAGCTCGCATAGTGTCAGTTATTCCAGTGAGAAAGCAAAGTTCTGGAACTAATGAAACTAACTGCCTTGTTTCCTTTCCTTTAATC
CGTGTTTTTACTCTGCTTTTCAAAAGTGGTTGATTTTCATCACTTATTTTAATATTATACTGAAGTCTGTAGTAGTCAACATACCTGACTTTTTCAC
CATTGCTCTTTTCAAAGGAGTCGTTAGGACTTTCATCAAATGATACATCATCAATTCTATAGACCTTATTATTATATTTAGTCAGAACAGATTGTCC
AAGGATTAACTGGACAAATGCCTCTTTCCAATGTGAATGATTTGAGTTCTGAATATCATCCATCAACTGTAATACTGTTTGATTCCTAAGAACTCTG
TGTGAAGCATCACAACATAACATAAGGCCACCTTCAAATTCATCTATTGCTGTAATATATCCAGGCCAAACTTCCAATTTATGCATTGGAATGGGCA
TAGCTGATTTAGGGTCAAAAAAATTTTTTCCATGCAGCGCTAAACCCAAAATGACCATAATCCGCCTAAAAAGCACATTATAAAGATGAACCGCTTC
TGGATCTCCAAACTTTCGTTCCTTGGTATATGTAATTTTAACATTAACAACACTGTTATCATGGGCTAAAGTTGTTTTTAATAATGTTATTTTTTTT
GGTAGCTGGATTGGTAAATACAGAACCATGCCATCAAAGAGCCTGGTTTCACCAATTACATGCTTTAGACTATTCAATACACCAAATTTTAGACTTC
TGGATTCTAAAGGAGGATCATATTGCACATCATATTGGAATAGACCTTTTCCTTTTTCAACCTTTAGTTTAACATAGTTAACCCCAACAGATATCTC
TCTTCCAAATGATCCTTTTTTTATAACAGGTTGCCTTTCTTCTTCAGGACTATAACGAAGTGACATTCCACCCATCTCCTGAGAAATATGTTCAGTC
ACTGGATCATCTACTTCCCTTTTTGAAACCAATGAAGCTAGTAAAGCTCCTCTTCCAGATATTTTAGGGACACATCCTTCATCCTTTAATTCCTTTG
AAAGGATAGCACTTCGGCCAACAAGAGTTCTTCTCATTCCTTGAGATGAAGCAGATCTGTGAGAAGCACCTTCTTCCTGATACTCAGACGGTTCATG
GCCTTCATGTGACTGGGTTCCCGAACTTGTCTGAGGTACGGATTGAAGCAGTTTTCTTAAGGCTTCACCCCTTCCCAATCGTTTTCCATCCATTTTT
ATGTGACACTACTTATCTAAAAATTACAAATACCACATACATCCACAGATCAAGACCTTAAAAG

Protein 483(AA)
MDGKRLGRGEALRKLLQSVPQTSSGTQSHEGHEPSEYQEEGASHRSASSQGMRRTLVGRSAILSKELKDEGCVPKISGRGALLASLVSKREVDDPVT
EHISQEMGGMSLRYSPEEERQPVIKKGSFGREISVGVNYVKLKVEKGKGLFQYDVQYDPPLESRSLKFGVLNSLKHVIGETRLFDGMVLYLPIQLPK
KITLLKTTLAHDNSVVNVKITYTKERKFGDPEAVHLYNVLFRRIMVILGLALHGKNFFDPKSAMPIPMHKLEVWPGYITAIDEFEGGLMLCCDASHR
VLRNQTVLQLMDDIQNSNHSHWKEAFVQLILGQSVLTKYNNKVYRIDDVSFDESPNDSFEKSNGEKVRYVDYYRLQYNIKISDENQPLLKSRVKTRI
KGKETRQLVSLVPELCFLTGITDTMRADARVMKDITQCISISPNQRHRAIKQFINNIRVSPEATKVLSDWGMSLDDAYLVLEGRNIPPIDILFGA

Graphical representation



Comparison with *Cylas puncticollis* AGO-3

```
Query  88   SKREVDDPVTEHISQEMGGMSLRYSPEEERQPVIKKGSFGREISVGVNYVKLKVEKGKGL  147
            S++   DP  E +S+++  +S+    EEE+ P + KG  G  + +  NY++L VEKG+G+
Sbjct  76   SEKPKADPDLE-VSKQLQQVSIS---EEEKVPALYKGESGSPLKLSANYIRLDVEKGRGV  131

Query  148  FQYDVQYDPPLESRSLKFGVLNSLKHVIGETRLFDG-MVLYLPIQLPKKITLLKTTLAHD  206
            F+Y+V+++P L++++ +  ++N +   +   ++FDG   LYLP ++ ++  + K TL
Sbjct  132  FEYEVRFNPELDAKNERIKMVNQMMRDMESVKVFDGGHCLYLPQKIKEEQKVYKGTLPGG  191

Query  207  NSVVNVKITYTKERKFGDPEAVHLYNVLFRRIMVILGLALHGKNFFDPKSAMPIPMHKLE  266
            +  V+V I Y ++++FGD E +HLYN+LF+RIM IL     G+N+F+P     IP HKLE
Sbjct  192  DQEVSVTIIYKRQKRFGDRECLHLYNILFKRIMHILLYTQMGRNYFNPAHKHLIPQHKLE  251

Query  267  VWPGYITAIDEFEGGLMLCCDASHRVLRNQTVLQLMDDIQ-NSNHSHWKEAFVQLILGQS  325
            V+PG+   +DE EGGLMLC D  HRVLR Q   +L+ +I+  S+   +KE  ++ I+G
Sbjct  252  VYPGFAVTVDELEGGLMLCLDTQHRVLRTQNAYELLTEIRCASDPRTFKENAMKNIIGSC  311

Query  326  VLTKYNNKVYRIDDVSFDESPNDSFEKSNGEKVRYVDYYRLQYNIKISDENQPLLKSRVK  385
            V T+YNNK Y IDD+ +D +P+D+F   +G  + +V+YY+ QYNI I+D NQPLL  R
Sbjct  312  VFTRYNNKTYIIDDILWDMTPSDTFPTRDGNTITFVEYYKQQYNININDVNQPLLLHRRS  371

Query  386  TRIKGKETRQ--LVSLVPELCFLTGITDTMRADARVMKDITQCISISPNQRHRAIKQFIN  443
             ++ GK  ++  ++ LVPEL FLTG+TDTMR D RVMKD+ Q   ++P+QR +A++ +++
Sbjct  372  VKVSGKAEKEDRMICLVPELSFLTGLTDTMRNDFRVMKDVAQYTRVTPHQRMQALRVYLD  431

Query  444  NIRVSPEATKVLSDWGMSLDDAYLVLEGRNIPPIDILFG  482
            N+R S +A +VLS WG+S+ +  + L+GR +    I FG
Sbjct  432  NVRSSEKAQQVLSQWGLSVAEGNIDLQGRQLEHEKIRFG  470
```

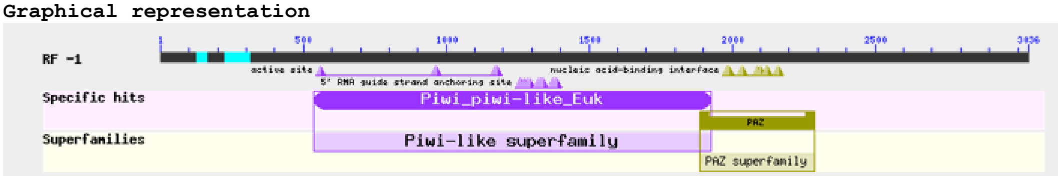
Aubergine/PIWI

cDNA

Contig 1655, Length-3033
GTCAAAATAATCATATTACAACTTAACCTTATTATGTTAGTTAAATGAGTTAATGAAACAAATAATATTCATTCATTATAAATAATTCAATTGGTAG
AGGGAAAATAATCCCAATCAAATAAACAAATACGATGATAAACTTTTTTTGTAAACATAATAATAACCTAGAAAATTAAATAAAATACTATAATTTA
TAGTATTAAACAAAAAGAAAACCACTGATAATAATAATAATTTATATAAAATAATCTCTCAAATTATTTCACAATTTAATAATAATTACAATCCAAA
AAATTTTAATTTTCTGAAAAGTATTACAAAATTTATTTTTTTTAAATCACAATAAAAATTAATCAGCAGCCTAAGTTATAAAAATAAGACTATATCA
CAATGAAGTTGCAGAATATTAAAAGAAAAAAAAGAAGAAGAAGAACCTCTGTCAACTTCCTTTGTGAATTATTTACAATTAAAAGCAGGTCCTCAAA
GAAAAAACAAACATCTTGTTAGCACTTCAAGAGGTGCAAGGATGGTAACCTGGCCCACTAAAAATGCCAACTTGTGAGCATATTGAACCTGAGCTGG
TACCCTTACAGTGGTACTGCAATTGTAATACAGATGACACATTTTATATGTTATTCTCTGAATAATATTGGCAGGAGTTGCCATTGTATCATATATC
ACATGGTAATTGGTAGGTGTGACAGTACCTTGCCGAACTGATTGTGAAACCAAAAAGAAATCGTATTTCTCAGGAGATGTAATTATATCATCAACAA
CAGTTCCGGGTTGTGGATTGTATGCATTTCTACCTTCAAGAAGAAAGATCCTGGTTTGAATGCGTTTTGAAACAAGAACAATCATTAATTTGATCTT
ATTCCCTGGATAAGATGAATTAAGTCTACTCTTTATTGCTGAAACCTCATGTTCCAAAACATAACCAACATTTGATTCACCAACACCATCTCTATAC
AGAATTATTCTAGGTGGCAAAGCACCTTTATTGACTTCTTTAAATTTGTAGACAGCAATGACAATATGCGCAGCAATGTCATTAGATAGTTCCTCAC
CAAAATTGTGCTTACTTACAGCTGAATACCACTTCGTGAAATGGTAGTCTAGTGATGCAACCATTGCTCCCACAGATTTGTCTTTCACTTTTGAGTC
ATGACATACGTCAAATCCAACAACCATTGATACATTACAATGATCAGGGACAATCTTAAAAGGTATAGGTGTAAACCAAGGGGACCCACCAAGTTTA
CAGTTGATCTGAATGATTATTTTTGTAACTACTGAGGGAAGATTTTTATGTGATAGTGTTTTAGTAGCTACCACTTGAGATGCAACTCCCCGATCAC
AACAGCATTTCTTCTTAATTGCACCATATATATCAGATTTACTGCCGTTTATAATGCACATGACAAGATCAATTTTGGACATGACAGATAACTGTTT
CTCAATTGCCTCAACATATCCCAATGGTCTCTCATCGTTAATAAATATTACCTCAGGCCGATGTACTTCATACTTCAAGGATTTGGCAATTCGTGAT
ATTTCTTGAAGGAATATATCAAACACATTCTTTCCTGTCTTGTCTTTAGCATGAAATTTGTGAATGACAACAAACCAATTTTGCAAGGGCATATTGT
TGAACATGATTCTACCCCTCATTTGACGAGTCCAGTCACCATCATGTCCCTCAATAGGGCTTTTATCTTTTGCCACAAATTGGTTCATATATAACGG
TTCAGGATCCAGTCTTCTGGCCTTAATTCTTTCAAGGTCTGAGGCAAATTTAAAATTCCAAATGCTTAATTTACTCTGCACCTCTTTACTCTCCTTA
AGTCGATCCATGAATCTATAGTAAGATGCAATACGGGCAGTAGGGCCAATACGGGTTTTTTCAGCAACTGCCTTCATTAGGTTGAAATTTGTCCTCA
TTTCATCACTTAGACCAGTCATTTGACATAGTTCAGGAACAAGGACAATGTTTTGATTCCTGCCACCTCTTATGTCTCTAGCTCTTGGCCTCGAAAT
AAGAAGTGGTTGCTTTGGATCTTTAATATTAGCATTGTATTTCTGACCATAATATTCAATATAAGAAATATTAGTATTTCCTTTCTTGAATGTGTCA
AGAGGTGACATATCAAAAGCGATATCATCAACACGGTAAGGTTTATTGTTGTAGAAAGTGAATACCACACAACCCAAAACATGTCTGTTAAAATGAG
TCATCCAATCATGACGATGTACCCTCTTAAGATCATTCAAAAAATCTAGAACAGTATCAGTTCTTAAAACTTTGTGAGAAACATCGATTAACATCAT
TACATTATACTCATATTGGTTGATTGAATTATTGTAACCAGGCCATATTTCCAGCGCAAAGCTCTCCAGGGTTGTAGTGCGGGTAGGATCAAAATAC
TCGCGCCCAACCAGTTGAAATTTCATAGCATCCATACAATGTCGGAGTATTATATTGTACAACTGAATGTACTGAAAGTCACCAACTGCCAACTCGT
TAGTAAGTTTTATTGTTATCCTGATTTTAGCTTGATCAGACTCACGTATCGAAAAGACCTCAAATTCATCTGGTCGCATGCGCTGAATAGTAAAAAG
CATTGTGCCATCAAACATGTACGCCTCACCCAAACGATCCTTATGATCTCTTAACATAGCTTTACGAATTTTAGTCCTCTCTTCCTCAGGACTAAAA
TCAACTCTGTATTGATAAAGTGCCCAATCGGTTTTCTTAATCAGTCTAAAATGGTTGCTAACAACTTCAATCCCTTTGCCATCTGTGCCTTGTTTGA
CTTCTAGATTATCAGGTCTGGTTTTGATGGGAGGAACCAACTCAAAACGACCACGCATTGCACCCCTTCTTAATCCTTGACCTGGCTTAACTTCTCC
CAAAGACAGTCCACGGAAAGCACCAGCCACTTGACTCAGAGTTATTCTCCCTTCTCCAGCTTCTTGTGGTGAATACTGCTGTGGAGGCGGTGACTTC
ACATATCCTCTAGCTCGACCAGAGGC

Protein 796(AA)
MRGRFELVPPIKTRPDNLEVKQGTDGKGIEVVSNHFRLIKKTDWALYQYRVDFSPEEERTKIRKAMLRDHKDRLGEAYMFDGTMLFTIQRMRPDEFE
VFSIRESDQAKIRITIKLTNELAVGDFQYIQLYNIILRHCMDAMKFQLVGREYFDPTRTTTLESFALEIWPGYNNSINQYEYNVMMLIDVSHKVLRT
DTVLDFLNDLKRVHRHDWMTHFNRHVLGCVVFTFYNNKPYRVDDIAFDMSPLDTFKKGNTNISYIEYYGQKYNANIKDPKQPLLISRPRARDIRGGR
NQNIVLVPELCQMTGLSDEMRTNFNLMKAVAEKTRIGPTARIASYYRFMDRLKESKEVQSKLSIWNFKFASDLERIKARRLDPEPLYMNQFVAKDKS
PIEGHDGDWTRQMRGRIMFNNMPLQNWFVVIHKFHAKDKTGKNVFDIFLQEISRIAKSLKYEVHRPEVIFINDERPLGYVEAIEKQLSVMSKIDLVM
CIINGSKSDIYGAIKKKCCCDRGVASQVVATKTLSHKNLPSVVTKIIIQINCKLGGSPWFTPIPFKIVPDHCNVSMVVGFDVCHDSKVKDKSVGAMV
ASLDYHFTKWYSAVSKHNFGEELSNDIAAHIVIAVYKFKEVNKGALPPRIILYRDGVGESNVGYVLEHEVSAIKSRLNSSYPGNKIKLMIVLVSKRI
QTRIFLLEGRNAYNPQPGTVVDDIITSPEKYDFFLVSQSVRQGTVTPTNYHVIYDTMATPANIIQRITYKMCHLYYNCSTTVRVPAQVQYAHKLAFL
VGQVTILAPLEVLTRCLFFL

Graphical representation



Comparison with *Cylas puncticollis* Aubergine

```
Query  130  VRGRTNRNEIISTRPSHIQSKKGFDGTAIRLRANYFRFIPKGHWGLNQYRVDFSPDLDST  189
            +RGR      I TRP +++ K+G DG  I + +N+FR I K  W L QYRVDFSP+ + T
Sbjct  1    MRGRFELVPPIKTRPDNLEVKQGTDGKGIEVVSNHFRLIKKTDWALYQYRVDFSPEEERT  60

Query  190  STRKYLVRTGMKNKNVSGYLFDGTVLYTPNRIQPEPLQ-FVVDTDDGTHTTVTIRLVGEV  248
              RK ++R   K++    Y+FDGT+L+T  R++P+  + F +   D     +TI+L  E+
Sbjct  61   KIRKAMLRDH-KDRLGEAYMFDGTMLFTIQRMRPDEFEVFSIRESDQAKIRITIKLTNEL  119

Query  249  KWGDYHYLQLFNIIMRKCLESMQLQLMNRNYYDPKLKISVPEHNLELWPGFFTSIRQYEK  308
              GD+ Y+QL+NII+R C+++M+ QL+ R Y+DP    ++    LE+WPG+  SI QYE
Sbjct  120  AVGDFQYIQLYNIILRHCMDAMKFQLVGREYFDPTRTTTLESFALEIWPGYNNSINQYEY  179

Query  309  DIMMNTDLAFKTVRTDNVYDLF--LECLQGRNPQNEFRSRIIGSVVLTYYNNKTYRIDDV  366
            ++MM  D++ K +RTD V D    L+ +   +    F   ++G VV T+YNNK YR+DD+
Sbjct  180  NVMMLIDVSHKVLRTDTVLDFLNDLKRVHRHDWMTHFNRHVLGCVVFTFYNNKPYRVDDI  239

Query  367  DFKQTPASTFQMKDGSQISYAQFLQNRWRVSVNTMDQPMLISRASMREVRLGMPETIILV  426
             F   +P  TF+ K  + ISY ++   ++  ++    QP+LISR   R++R G  + I+LV
Sbjct  240  AFDMSPLDTFK-KGNTNISYIEYYGQKYNANIKDPKQPLLISRPRARDIRGGRNQNIVLV  298

Query  427  PSTCIMTGLTDKQRENFHLMKALGEYTRVNPQNRMQKLREFSQRLQNCPEAMNEIRRWDL  486
            P  C MTGL+D+ R NF+LMKA+ E TR+ P  R+     F  RL+   E  +++  W+
Sbjct  299  PELCQMTGLSDEMRTNFNLMKAVAEKTRIGPTARIASYYRFMDRLKESKEVQSKLSIWNF  358

Query  487  DVARDIIELEGRVLPEESLLLRNGAMITGGP----EADFTKQLRSAPMYT-IAQADKLAV  541
              A D+  ++ R L  E L +         P    + D+T+Q+R   M+  +   +   V
Sbjct  359  KFASDLERIKARRLDPEPLYMNQFVAKDKSPIEGHDGDWTRQMRGRIMFNNMPLQNWFVV  418

Query  542  ICPGRFRQGTNE-----FINCLVKAGRGMHFNLGNPKILEMQDDRPQSYLDILEHVITSL  596
            I     +  T +     F+  + +  + + + +  P+++ + D+RP  Y++ +E  ++ +
Sbjct  419  IHKFHAKDKTGKNVFDIFLQEISRIAKSLKYEVHRPEVIFINDERPLGYVEAIEKQLSVM  478

Query  597  -QPRMVLVVLPNNSADRYNAVKKKCYVDRAMPCQVVVGRTLQNKGVMSIATKVAIQMNCK  655
             +  +V+ ++  + +D Y A+KKKC  DR +  QVV  +TL +K + S+ TK+ IQ+NCK
Sbjct  479  SKIDLVMCIINGSKSDIYGAIKKKCCCDRGVASQVVATKTLSHKNLPSVVTKIIIQINCK  538

Query  656  MGGAPWGTALP--------KATMVVGYDVCRDTANRGKSFAGMVASMDTACSQYFSLTIE  707
            +GG+PW T +P          +MVVG+DVC D+  + KS   MVAS+D   ++++S   +
Sbjct  539  LGGSPWFTPIPFKIVPDHCNVSMVVGFDVCHDSKVKDKSVGAMVASLDYHFTKWYSAVSK  598

Query  708  HEQEQELSSNIATFMLLACKNYQERNK-VIPERIVIFRDGVGDGQIQYVKEHEVEQLKKK  766
            H   +ELS++IA  +++A   ++E NK  +P RI+++RDGVG+  + YV EHEV  +K +
Sbjct  599  HNFGEELSNDIAAHIVIAVYKFKEVNKGALPPRIILYRDGVGESNVGYVLEHEVSAIKSR  658

Query  767  LNAEIYQTQQLKMAFIIVSKRINTKIFRAQGVGPRNDYNPPPGTVVDDVITWPERYDFYI  826
            LN+  Y   ++K+  ++VSKRI T+IF  +G   RN YNP PGTVVDD+IT PE+YDF++
Sbjct  659  LNSS-YPGNKIKLMIVLVSKRIQTRIFLLEG---RNAYNPQPGTVVDDIITSPEKYDFFL  714

Query  827  VSQCVRQGTVAPTSYNVIEDSLGIDANKLQRFTFKLCHMYYNWSGTVRVPAPCQYAHKLA  886
            VSQ VRQGTV PT+Y+VI D++    AN +QR T+K+CH+YYN S TVRVPA  QYAHKLA
Sbjct  715  VSQSVRQGTVTPTNYHVIYDTMATPANIIQRITYKMCHLYYNCSTTVRVPAQVQYAHKLA  774

Query  887  FLTAQSLHRSANPALKNTLYYL  908
            FL  Q    +    L   L++L
Sbjct  775  FLVGQVTILAPLEVLTRCLFFL  796
```

Antiviral RNAi in *Megacopta cribraria*

Ars2

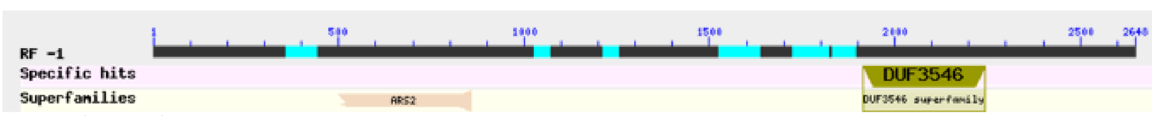
cDNA
Contig 4509, Length-2647
GGAAATCATTTTAAAACCTAATTGTCAAACTTACATCACTGATTCATATTCAAAGATATCAAATACTTATAGCAATTAAACAATGATGCTAAATTAA
ACTAATATTTGTTTTCCTTTAAAATACAAGCACTGTTATTAGTATCCACAGAGGAATGTTAACACAATATTTTCTCCTTTATACAAACAAAATCCAT
AATACAATTTATGAATGAAGATAATGATAAAAAGATATTTAACTCATATTTATTACATAAATAAATAGGAATATTTTAAAAATGCTATTAAATTATA
TAATTTCGTCAGGTTCACAAGGAGCGTCCAAGTCCCTATAAGTTATAATAGGCCTATAGTCCATACCCCTTCCCCTTCCAAAGCCACCACGTCCACG

```
GTTGAAACCACCATGTCCACCAAAATGGCCAAAGCCACGTTTGAAACCACCACCATATTGATAATGATAAGGAGATTCAGGCTCCTTTTTGCCTCCA
CGGTTACCAGGGTGTTCAGCTAACTGAGGACGTTTAGAGTCTCGAAGATAGTTATTAAAATAATCAACCTCTTTTTTTACTTCTTCTAATTCTTGAG
CAAATTTATTATAAATATGCTTTCGGACGAATTCCGGTCCTTTAAATTTCTTCCCTGACAAAGGACATAACCACTTGTCCTTTGCTAATTCCTGTGT
ATTAGCTTGTATGAATTTCTCTACTTCTGTATCTCTATCCTTTGCACCTAAAGAAATTAATTCTTGCTCAGTTACTATTGGTACTGGGGTATACAAA
GGATCCAACTTCGTTTTATAACCATCAATATATTCCTGTATTTCTTGAGGTGATACCTCTGTACTTGAAGGTAATCCCCTAAGATGCATTATGCCAC
ACCTATTAGGCATTTCATCTTCAGATGTATATTGGGAGGTATTGTAGTAATCGACAGAATGAACTATCCTCAAATAGAATAACAATTTATCTAAAAT
CTTAATAGCACCAGCAGTTCTAGATTTTGATTCATCGGTAACCACACCTTGGCCTAATAGTTCTTCCTCTTCAGCAGGAGCTTCTTCAACAAGATAT
TCAGTAATGCCTTTAAACACTGGATTTCTTGATTCCAATCCAAAGGTTTCATTTGTATCAGATAAACTTAAACCAGCTCTGGAATCCAACTCTTGGA
TGATTTTAGCAGCTAATTTCAAATCGGCGAGAACAACCTGTTTATGTCCCGAGAGACCACTAACAGATCTTACTCGACGAGACAAATCCCGATTTAC
TATTGCTCCCATTTCACATTCTTTTACCCTAATATTATTCAAATTCCAACAGATTTCTTTGATATTGACATCACGTCGAAATGTGACCCATCCCCTT
CTATACCAGTGTTGTTCAGCTTGAGGTTCGGAAAGAGAAACTCTGAGAAATCCTGGATACTTCTTACAAATAGATTCTATTTCAGCTTTAGTAACCC
TTGGAGAAATTGTTCGCAGGAAAATAGATGAAGTTTTGTGCAGTGCTCTTGGTGTCAATGGAACCACTTCATCATGATCCTTGTCAGAATCTTGGTC
ATCCTTTGATTTCTTAGGTTCTTCTTTTTTTTCTTTCTCAGGTGAGGATTCTTTTTCATCAGGGGAAGACTTTTTTCTGGAGAGCCTTTTGCTGCA
GAAATATCTTCTTTATCCGTGTCACTCTCTGGATCAACTCCAGGAGGTCTTTCATCTTCAGAGAGTTCGTGCACTTCATTCTTATCTTCCTCTTCTT
TTCGGTCATCATCTTCTATAACAATCTCATCTTTACTTTCTTCATGTTCCTTAACTTTTTCTTCATCCTCTGATTTGTTATTTTCTTTGGGCTTTTC
ATCGATTTTATCCTTTTCTTTTTGCTCTTCCTTTTGACCTTCTGGCTGATCAAGTATGGTCAAATCAAGATCTGTTCCTCCTTCAAGCCTAATTACA
ACAGCATCAAGAAGCTTCACAACATCATCTCCCTGCTCAGTGTCTACTCTTATATTCTCCATGCGGTCTGAATTCATCATTTCTAGGAAAACGTCAC
ACCTTTTCTTTAAGGCAGTATTAAGTTCTTCTTTTCTTTTCACTGATTCTTCAGGATGATATTTTAACTTGAACCATTCTTCCTCTTTGTGTGCTAC
AAAAAACTCATTCAGCTGCTGGCGTCTGAATTCAAGTTTATATTCAGCATATTTAGATATAGCCTCTTCATCAGTAATATTATCATCTTGTGATTGA
AGGAAAGCTTTGAATGTCATAATTGGAGGCTGTGTATCCATTTCACCAAGTCTTGAATTTCCAGGATATGGAGGTTCAGGACCCCAACTATTGTAAG
GTCCATAACCCGGGTCATTATATCTCATCCTCCTGATGTCATCTGAATCATATCTCATTCTCTTCGCAGGAGGTATATCTCCTTGGTATTGATATCT
ATCTCGATAGTCATTCCTCATTCTTGGCCGTCCCCAATCATCATCTTTGCGTTTATCTGTTCTGTATGAATCGCTTCTTTCTCCTTGAAATTTATCT
CTCCTCTTCTTGTCATATTCGTCCTCACTATCGCCCATTTCAAACACTTTCTCTTAATCGTGCCTTACTAGAAAGCTACAAGCCGCGGCAGAAGCCA
ACGTTTCCCAAAACAGAAAACTATAAAC
```

Protein 757(AA)

```
MGDSEDEYDKKRRDKFQGERSDSYRTDKRKDDDWGRPRMRNDYRDRYQYQGDIPPAKRMRYDSDDIRRMRYNDPGYGPYNSWGPEPPYPGNSRLGEM
DTQPPIMTFKAFLQSQDDNITDEEAISKYAEYKLEFRRQQLNEFFVAHKEEEWFKLKYHPEESVKRKEELNTALKKRCDVFLEMMNSDRMENIRVDT
EQGDDVVKLLDAVVIRLEGGTDLDLTILDQPEGQKEEQKEKDKIDEKPKENNKSEDEEKVKEHEESKDEIVIEDDDRKEEEDKNEVHELSEDERPPG
VDPESDTDKEDISAAKGSPEKKSSPDEKESSPEKEKKEEPKKSKDDQDSDKDHDEVVPLTPRALHKTSSIFLRTISPRVTKAEIESICKKYPGFLRV
SLSEPQAEQHWYRRGWVTFRRDVNIKEICWNLNNIRVKECEMGAIVNRDLSRRVRSVSGLSGHKQVVLADLKLAAKIIQELDSRAGLSLSDTNETFG
LESRNPVFKGITEYLVEEAPAEEEELLGQGVVTDESKSRTAGAIKILDKLLFYLRIVHSVDYYNTSQYTSEDEMPNRCGIMHLRGLPSSTEVSPQEI
QEYIDGYKTKLDPLYTPVPIVTEQELISLGAKDRDTEVEKFIQANTQELAKDKWLCPLSGKKFKGPEFVRKHIYNKFAQELEEVKKEVDYFNNYLRD
SKRPQLAEHPGNRGGKKEPESPYHYQYGGGFKRGFGHFGGHGGFNRGRGGFGRGRGMDYRPIITYRDLDAPCEPDEII
```

Graphical representation



Comparison with *Cylas puncticollis* Ars2

```
Query  1    MGDSDDEYDRKRRDKFRGERAEASYRGAERVVRPRDEWVERESWSRPRPR-EYRPRDRGY  59
            MGDS+DEYD+KRRDKF+GER++ SYR  +R         + + W RPR R +YR  DR
Sbjct  1    MGDSEDEYDKKRRDKFQGERSD-SYRTDKR---------KDDDWGRPRMRNDYR--DRYQ  48

Query  60   SPSMEPTAKRMRHD--------YYGEGY--YNHYTPYHQPPHREQTMSQMEGQQPPMMSF  109
                 P AKRMR+D        Y    GY  YN + P  +PP+    +       QPP+M+F
Sbjct  49   YQGDIPPAKRMRYDSDDIRRMRYNDPGYGPYNSWGP--EPPYPGNSRLGEMDTQPPIMTF  106

Query  110  KAFLATQDDNISDSEAIEKYNDYKLEFQRQQLNEFFVAHKDDEWFRLKYHPEDSVRRKQE  169
            KAFL +QDDNI+D EAI KY +YKLEF+RQQLNEFFVAHK++EWF+LKYHPE+SV+RK+E
Sbjct  107  KAFLQSQDDNITDEEAISKYAEYKLEFRRQQLNEFFVAHKEEEWFKLKYHPEESVKRKEE  166

Query  170  QAAALKKRLEVFIDLLNSGKIESTTVDCARTTDLLKLLDTVVIKLEGGTDEDLAALEQEY  229
               ALKKR +VF++++NS ++E+   VD  +   D++KLLD VVI+LEGGTD DL   L+Q
Sbjct  167  LNTALKKRCDVFLEMMNSDRMENIRVDTEQGDDVVKLLDAVVIRLEGGTDLDLTILDQP-  225

Query  230  EEMEQKYREQKEKEAAAERAREARMAETASQKEEGE  265
               E +  EQKEK+   E+ +E   +E   + +E E
Sbjct  226  ---EGQKEEQKEKDKIDEKPKENNKSEDEEKVKEHE  258
```

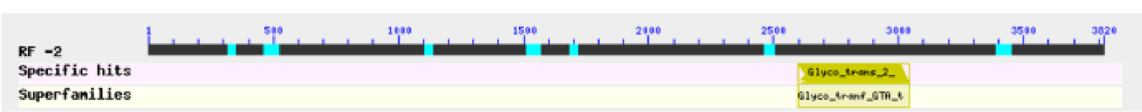
Egghead

cDNA

Contig 3764, Length-3819
GAAAGGGCACAAAACCAAACTTAATTTAAAGATAAGTTTTGTAGTTAAATTAAATTCGAGTGAAATAATTTAATACTAGATGAAGAATTTACAAAAG
GCAAAATTTAATGAGTCCAAAAATGTTTTTGTTTAAGAATTATGAAATGCTTCATAAGTTTATATTTGTATGTTCAAGAAAATAAAGGAAAGAAAGT
GAAATTCACAAAGACAAAATTTCCACATATAAATAAATAGTTAATATTGTATTTTGCAAATAAAATATAAATACATTTAAAAAAAATTAATCTCCCT
TTTTATTAATTTATGAATACCAAATAGGATTAGATCAATTAATGACTATATTATTAAGGGATTTTAAGAAAAACATTGGTCATCCATGGACACCAGT
AATATGAATTTCTTTTAATGGATTTGGTATGTTGAAAAGATGAACCACTTAATTATGTGGTTGCCACTGTACAATAATTTTTTAAATCATTCTTTCA
GAACTAATAATAATTAATTCAAGATAAAAAAAGTTAACCTAAATGAATTTTAAATGTTCTTTGAGCACATTAGAAGTAATTTTTATTCAAAGAAAGA
GCCCTGTATTTTCCATGACTTCATTATTATAATCAAACATGATTGCAATTAATTATCAAAGTAATAAGTGAACAAATTTTGTTTTATAATATTGGAT
AAATCATTATAAGTTTAATAAATAAGATAAGAAATGGAAATTACACAGCCAAAATTTTTCATCCTACATGTATAAATCAAATATAACTAGCAAAATA
TAACCATGAAAAACAAATTAATACAATCATTCTTAGTAGATTAATTCTCGTGTATTTATAAATATATATTTATATAAATTAAGTCCTCTTGTATGTA
CTTACAAAAGATAGTAAATCACAAATGGTATTAATTGAAAAATCTTAAGAGCAGAACCAAATAAAATTTTTTGTACCAGGCCTATTACAAAATTAAA
TTACAAAATATAGATTATTCAATAAAATTATAATTTCTATATCAAATAATTAAATCTAAATAAGATTTCTAATGAATAACTAATTAAAATTCACTAA
AAGAGAAAAAGTAATTGGTTCCAAATTTATATTATATTCAATTCAACATACTTAAATTATTCATCAACTAAGCAACTTTCCTTCATCCTGACAGAAC
TTTTTCAATGAGGTTTTCTCTCACATTATAGGAAAAGAATTAAAAGTTGAAAATGGGAAGAGTGGTGAAGAGCTTTTTCCCTTGTCTGATATGGAGT
TAAAAACAGTATCTGCATGGAGTCACAATAACGGAAAATTAAGTTCACAAACAACTTTGGGAGAAAATAATACAAGTCCTAAACCATCCTTGCTAGA
GGAATCATCTTACAAACGGGATAATCAAACTCATATATCTTAATAATAGGACTATCTTTATATTTTTCATCTCCCTAATAACATTAATTTGATTTTT
TTTTCTTTGAACTTTAAGTATGAGCCCTGAAAAAAGATAATTAAGATATGCATCCAAAATTATTTACAGATAACGTGACAGATATTAGATAAGCCTT
ATGCAATATATTTAACTAGTTAGACAAAATGTTGGAGTTGAAAACCAATATTTATAAACAGTCAAAGAAAAAGAGGGCGGAGGGGGGAGGAATCCTT
CACACCAATTAATTATTCTTACCTTAATTGTATTGTAATTAAGAATATCAGAAAGAACTATAATTTCCCATTTGAATTTAAAATAAAACCAATATGT
TTGAAAATGGACTTGGTAATTGAAACATAGTTGGAATACATGGCCAAGGATAAGTGATAACCAAGGATAGACCGAGGCTGAACTGGATTCAGCCCTG
AAAGGAACCAAAGTTCAGGGAAAAGAAGTTTTTCCAAAACATGAATATACTTTACACACAAACATATATAAAACAAAAAATAGGAGAAAATATAAAT
GTGAATCTTTTTAAAACAAAAATTTTACATTATATTTATACTTCAAAAGTATTTTTTACCCATCTGCCAAAAAAAAAAATTTTTGAAGGGAGAGAAA
TTCTCATGGGTATGTGTTCAACTTATTACCTAGTTCTGTCGGGTTACACACATAAAGAAGAAATTTATAAAGTTAAGAATTGTATTACCTTAAACGA
TAATGAAGCTCATTAAGCAAATTAGAATCGTCCAAGTTTAATTTTAATACAAAACATTAATAAATTGGTGGAAGAAAAGAACCTATACTATCATGGC
AGGTGTGATGTCCTTATTAACTACATAAAATTTATGTTTTTTCCCAAAGAGACCCCACAGTACAGCAATATTTTCAATTACAACATTCATAGGAATT
GTAGCCAATGCACCAAAAATACAAAGACTAAATTTCACGATTCCAAAACGATATAAAGAAAATGATTTTATAACACCAAAAATATACATGTAAATAC
TTACAGCGCCAATAAAAGCACATAAGAAATCAATTATTGGAGGGCAAGGGATGGGACATATAGCTGCAAGAATTATATTTGAGGTTGATAGCGGCAT
TGTAACCCAAGAATAACACGAAATGCTTAGAAATATCTTATTTCTAAATGGAATTTTTTTTGAGCGTACAACTAAAAGAATTCCTTGAAGCCATCGT
TTTCTTTGCTGTAAGAAGTCACAAAGGGTAAATGGAGATTTTTCCCACATTTCACCTTCTATAAAGTTGAATGAATAACCCATTGTGTATGCTTTCA
TTGCAAAAAAACAATCCTCAGCAACAGAGCCGTCTAATCCATTATCAAAAGTAACATCACGTTCAGCCTTAACCTGAGTTACGACATAGGATCCTTT
CCAACTAAGCAGAGGCTTATGGAACATTAAAAACTGATAACGAAGCTTTCCCATGTCATCAGCTACCCTGAACGTATCGGCCAAAGTAGTAATCCAG
TTCACGACTTCTTCATTGGCATATGTAATGAGCCCTTGTCCAAATGGATACTTCCCATCAATAACGAAATTCAAAATGCCTCTAATAGAATTTTCAG
TCAGAAGGGTTTCCTCATCAAGATGCACTATCCAATCGCTGTCAGAAAGGGTGTTCACACCCTCTTCCAAACAATACTGCAGAGCACGAGCTTTAAA
CAAGGCTCCTGATTTTGTTTTATAATTAGCAGGAATTACCGTTTCAATAATCCTCCGATCTTTAATAAGATGAATTGGCTTATCAGTGACAACTTCA
ATAATGAAATGTTCTAGCCCTGCATTCAAACATTTTGATAAATTACGACTTACATTGTTTTTCACTAATTGGGGATAATCACCACGAGTAACTATTC
TGATACAGATGAAAGGGGCTAAAAGCGGACTGCCCTTCAGAATAACTTTGTCAGGAAATGCATTGTAAAATGTAAGTCCAGCAAAATTTAAGAGAAT
TTGAGGCAGTGGTAAGAAAGCCAACAATCTCAAAACATAGAGAAGAGTTGTTAAAACAAAGCCATATTCCTCCCAAGGATCTATTATCACATCATCA
CTGGAAAACTTTACACCTCCAGTAACAAGTTCGAAAACAATAATAACAGTAAACAATAATACACAGTGAAGAACATGTTTGGTTTTACTCCTTAACA
TATCTGAGCCCTTCTAAACTTTTAGTTAAGAACCTATTAAATGTTATCTACCCACGAAAAGCTTCATCACAATTGATCACAAACTATTTGCTGGAAG
TGAAAATTTAGGTAAACAAAAATGTTCATGAACATATTGCAAAATCTTAAGCTCGTATTTGAAACATTATAATAAAAAATTTCGCGGTAGGGAGTAA
CCACCAAAATCAAGAAGCATGACTGCTTCGTT

Protein 457(AA)
MLRSKTKHVLHCVLLFTVIIVFELVTGGVKFSSDDVIIDPWEEYGFVLTTLLYVLRLLAFLPLPQILLNFAGLTFYNAFPDKVILKGSPLLAPFICI
RIVTRGDYPQLVKNNVSRNLSKCLNAGLEHFIIEVVTDKPIHLIKDRRIIETVIPANYKTKSGALFKARALQYCLEEGVNTLSDSDWIVHLDEETLL
TENSIRGILNFVIDGKYPFGQGLITYANEEVVNWITTLADTFRVADDMGKLRYQFLMFHKPLLSWKGSYVVTQVKAERDVTFDNGLDGSVAEDCFFA
MKAYTMGYSFNFIEGEMWEKSPFTLCDFLQQRKRWLQGILLVVRSKKIPFRNKIFLSISCYSWVTMPLSTSNIILAAICPIPCPPIIDFLCAFIGAV
SIYMYIFGVIKSFSLYRFGIVKFSLCIFGALATIPMNVVIENIAVLWGLFGKKHKFYVVNKDITPAMIV

Graphical representation



Comparison with *Cylas puncticollis* Egghead

```
Query  1    MLSGKTKHYLHCTLFLTVILFFEIFSGGLRVFNGGVFVPASEINPWVEYGFFGACVLYAL  60
            ML  KTKH LHC L  TVI+ FE+ +GG++  +  V      I+PW EYGF    +LY L
Sbjct  1    MLRSKTKHVLHCVLLFTVIIVFELVTGGVKFSSDDVI-----IDPWEEYGFVLTTLLYVL  55

Query  61   RLVTFLPLPQVIFNFVGLIFYNAFPEKVSLKGSPLLAPFICIRVVTRGDFPQLVKNNVNR  120
            RL+ FLPLPQ++ NF GL FYNAFP+KV LKGSPLLAPFICIR+VTRGD+PQLVKNNV+R
Sbjct  56   RLLAFLPLPQILLNFAGLTFYNAFPDKVILKGSPLLAPFICIRIVTRGDYPQLVKNNVSR  115

Query  121  NMNKCLDTGLENFLIEVVSDKPIGLEQHRRVRELVVPSDYRTPTGALFKARALQYCLEEK  180
            N++KCL+ GLE+F+IEVV+DKPI L + RR+ E V+P++Y+T +GALFKARALQYCLEE
Sbjct  116  NLSKCLNAGLEHFIIEVVTDKPIHLIKDRRIIETVIPANYKTKSGALFKARALQYCLEEG  175

Query  181  INILSDNDWVVHLDEETVLTENSIRGILNFVMDGKHHFGQGLITYANEEVVNWITTLADS  240
            +N LSD+DW+VHLDEET+LTENSIRGILNFV+DGK+ FGQGLITYANEEVVNWITTLAD+
Sbjct  176  VNTLSDSDWIVHLDEETLLTENSIRGILNFVIDGKYPFGQGLITYANEEVVNWITTLADT  235

Query  241  FRVTDDMGKLKLQFLWFHKPLFSWKGSFVVTQVAAEKDVSFDNGLDGSIAEDCFFAMKAY  300
            FRV DDMGKL+ QFL FHKPL SWKGS+VVTQV AE+DV+FDNGLDGS+AEDCFFAMKAY
Sbjct  236  FRVADDMGKLRYQFLMFHKPLLSWKGSYVVTQVKAERDVTFDNGLDGSVAEDCFFAMKAY  295

Query  301  SMGYSFNFIEGDMWEKSPFTLWDFIQQRKRWIQGILLVVHSRSIPIRNKLLLACSCYSWV  360
            +MGYSFNFIEG+MWEKSPFTL DF+QQRKRW+QGILLVV S+ IP RNK+ L+ SCYSWV
Sbjct  296  TMGYSFNFIEGEMWEKSPFTLCDFLQQRKRWLQGILLVVRSKKIPFRNKIFLSISCYSWV  355

Query  361  TLPLSTSNIFLATKFPIPCPALVDFLVAFIGAVNIYMFIFGVIKSFNVYRFGLARF  416
            T+PLSTSNI LA   PIPCP ++DFL AFIGAV+IYM+IFGVIKSF++YRFG+ +F
Sbjct  356  TMPLSTSNIILAAICPIPCPPIIDFLCAFIGAVSIYMYIFGVIKSFSLYRFGIVKF  411
```

Nucleases* RNAi in *Megacopta cribraria

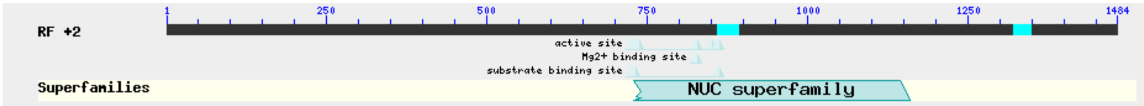
dsRNAse

cDNA

Contig 3484, Length-1491
CAGTATGGCTGTACTCCATTTCTTCTGCTTGGCGGGCTTCTTTGTCGCCTCTCATGGACGAGCAGTAAGCAGGGTCGACCCTGTCGGAGGTGCTTGT
GTCTTGGACGTGAACGAAGACTTACCCAAAAAGAACGAACCTCTGCTGCTGAAGAGAGGAAAATCAGGATCGCTGGAATTTGCGTATCCAGAACTGG
AAGGAAGACGAGGTGTCATTGCGCTAAGAGAAGGAGAGGAACTGGTGATCTCATGCCCTGGTAGACGTAATTACCTGCAAAAAACCAAAAATGACAT
TGCCAAAACTTCCTGCGCAGAAGGAAACAGACTAACACTAGACGATAAGAAATACAATTCAGATGAGTTTGAATGTAACAGCAGAGTCGCTGCCTCA
AGTAAACCCACAGAGCAGAAGTGTGCTGGTGGTAAAGGAATCATAGTACAACTTGGATTTAATGTTGGAGAATGGCTGCCTCTGATTGACTCCTGCC
ACAATGTCGATGAGAGCGTCACTTACTATTCAAATCACACCATCCACGGTGCCATTCTCGGCGGAAAGGTGAGGAGGACTACAGGGCGACCGAAGTT
CGATAGAGGAGACGCCATCTTATTCAAAGGGTTTAACCCGGAAATCGCTTATACTCAGAGGAATCAGCAAAGAATTCTCTCACGTGATCTTGGAAAA
GAAACAGCAGAGGGTCTCCTCGACACCAAGAAGTCACTATTCATGGCGAGAGGTCACCTGGCACCGGATGCCGATTTCCTCTTCAGCTCCCACCAGT
TCCTCACTTATTTCTATGTTAACGTGGCGCCCCAATGGCAGTCCATAAACGCCGGAAATTGGCTGAAAGTGGAGGAAAACACCAGGAGATTTGCTAG
AAGGCTTGCAGCTGACCTCGATGTAGTTACTGGTACCTACGGAGTCCTTAAACTGAAGGGTACCGATGGAGAGAAGGAAATTAGATTGCAAGAATCA
AGACTTCCCGCCCCAGAACTATTCTGGAAAGTGGTCAGGAATACGCAGGATGACTCTTGTATAGCCTTTGTCTCGACCAACAACCCATTCCTGGAAT
CCAAACCAAAACCAATCTGCACAGACATCTGTAAAATGGCAAATTGGCCTACATTTCAGGATGATCTGTCCAAAGGATACGTTTATTGCTGTACCTA
CCAAGAGCTGAAGAAGGTGGTCGCCGAGATGCCAGATGTGACTTGTAACTCCGTTCTCAAAGGAAAATAAAAACAACTTTGTTTACCACCAATAAAC
ATATATAATAAATAATAAACAGTAAAAGTACTCCTGCAATGCATTTTTAATGTGGAAGCCATAATCTTATATATATATATATATATATATACATTAA
AAGCATTTATATATGGTAAAATGCTTTACCATTCAACAAAACGTCACAAATTCAATAAATTGTTATTTTTAAGGATTAAATTATTGATGTAATTTCA
TAAGAATCATAATAATAAAATTCGTTGG

PROTEIN 409 (AA)
MAVLHFFCLAGFFVASHGRAVSRVDPVGGACVLDVNEDLPKKNEPLLLKRGKSGSLEFAYPELEGRRGVIALREGEELVISCPGRRNYLQKTKNDIA
KTSCAEGNRLTLDDKKYNSDEFECNSRVAASSKPTEQKCAGGKGIIVQLGFNVGEWLPLIDSCHNVDESVTYYSNHTIHGAILGGKVRRTTGRPKFD
RGDAILFKGFNPEIAYTQRNQQRILSRDLGKETAEGLLDTKKSLFMARGHLAPDADFLFSSHQFLTYFYVNVAPQWQSINAGNWLKVEENTRRFARR
LAADLDVVTGTYGVLKLKGTDGEKEIRLQESRLPAPELFWKVVRNTQDDSCIAFVSTNNPFLESKPKPICTDICKMANWPTFQDDLSKGYVYCCTYQ
ELKKVVAEMPDVTCNSVLKGK

Graphical representation



Comparison with *Cylas puncticollis* dsRNAse

```
Query  45   SGTSDFIYPK--PFEKTLKFSSGQSLDFVCPERTLILGATKTNDSVLQGTCVKDDTFKIN  102
            SG+ +F YP+        +    G+ L    CP R    L   KT + + + +C + +    ++
Sbjct  53   SGSLEFAYPELEGRRGVIALREGEELVISCPGRRNYL--QKTKNDIAKTSCAEGNRLTLD  110

Query  103  GGEVLWNQISCSNYPYRVAKRTGLSCENG-GTDIEIGFEVGDGRFLRSLSVCFNESSQIA  161
              +    ++  C++      +K T    C  G G  +++GF VG+  +L   +  C N    +
Sbjct  111  DKKYNSDEFECNSRVAASSKPTEQKCAGGKGIIVQLGFNVGE--WLPLIDSCHNVDESVT  168

Query  162  LYSFHNMTTAI--NQRVLATPRPSWLQGSGLY--TISTITNYYVRNKQRSTINTLLGLDI  217
             YS H +   AI    +      T RP + +G   +      +     Y   RN+QR   ++   LG +
Sbjct  169  YYSNHTIHGAILGGKVRRTTGRPKFDRGDAILFKGFNPEIAYTQRNQQR-ILSRDLGKE-  226

Query  218  NSTKYIEDNSNYYLARGHLSARSDHYYAAQQNATFFLINVAPQWQTFNSLNWNQVEIDVR  277
             +     ++    + ++ARGHL+    +D    +++ Q   T+F +NVAPQWQ+ N+ NW +VE + R
Sbjct  227  TAEGLLDTKKSLFMARGHLAPDADFLFSSHQFLTYFYVNVAPQWQSINAGNWLKVEENTR  286

Query  278  DYAEERGINLQVWTGIYGVTILPHEQTGRPVELYLYVDDNNNKALPVPEIYWKVAYDPIT  337
             +A      +L  V TG YGV   L          + + L                LP PE++WKV    +
Sbjct  287  RFARRLAADLDVVTGTYGVLKLKGTDGEKEIRL-------QESRLPAPELFWKVVRNTQD  339

Query  338  EKGLAMIGVNNPYLANYTE-ICNDVSDRITWLHWKKDDQARGYSYACAVDVFRRVVTSLP  396
            +   +A +   NNP+L +    + IC D+        W   + +DD ++GY Y C        ++VV   +P
Sbjct  340  DSCIAFVSTNNPFLESKPKPICTDICKMANWPTF-QDDLSKGYVYCCTYQELKKVVAEMP  398

Query  397  DIPVRGLL  404
            D+    +L
Sbjct  399  DVTCNSVL  406
```

Exosome

cDNA

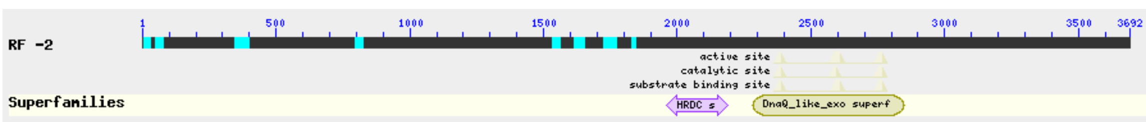
Contig 2692, Length-3677

TTTTTTAATTTTTTTTTTTTTTTTTTTTTTTTCAATCAGAATTAAACAAAGTAAAAAATAAACTAAAACTGCTATAAAAAATCTTCAGTAATTCAACA
AAATTTTAAAAATATATTAATAAAAAAACATATGACTGGCTCAGAGGTAAAACAGTCCAAGGTCTACTTATAACTAATTCAATTTGAAAAACTTTTC
TTCTGAGTGGCGTACTGTACACATAGTATAGTTGTAGAAATACTTTATTTTTATTTTCATAGATTAGTTACTATACATTATTTTATACAATGCGAAT
TGTACATATTTTACACACATTTTTTTTTTTTTCAAGAGAATATTGGCAGTTGTTTCCTTAGATTTTTTTAAAGGGACAATCTAAGTCAAAGTTTTTT
GAAATTTTTATCAATATCTGCTACAGAATCTTAATTTTACTTATGAAAAAATAAAAAACCAGAAAAAAAATATCAGTTTCTATTTTCCCAAAACAAAT
ATTCATAAAGTAAGACTTTATATTATCTGACCCATAAAATTCAATCCTTAATTATAGAAGAAAAAACAAGAGAATCTATGAATACTAACTGACAGGG
TATGAGATGAGAGTGTTCATATCAGTAAATGAGTTATCAGTGATCTTCAGAAATCGAGTTAAAAAGTTACAAATTAAGAAACCAATATCTGACCAAG
GTCTTTATTGAAATCCCATAAATATCCTGGGTTTCAATAAATTGATTAGAAAACAAATGCTGGATGGATCAAGAATTAAATAATGAAACATTACCAT
AATTATTCGAAAAACAGATTTTAATTATAATTTACAATACTATTTAATTTGATTATAATTTATGATAAAAGATCACTAATTTCTCTAAAAACTTTCT
ATTTAAAAAAATAAAATTAAACATAGTCTAATATGTATAAACAATAAAATCATAAATTGTTTTGAATTGCTTTTGACAAGCAAAATTATGGTAATTA
CATATTTTATACCATATATGGCAACATCTATAATAATTTTAAGATCAGAAAATTAAAACTTCAAGGCTTAACATTAAACAAAAAATTTAGTTATGAT
CAATGAGCTAAGGATAACAAAATTCAGCTAGGAAATCAGTTTTAAAATAAGGCAATGGGCCTTATTTCACATTATTTTGTCCTCCTCGATGCTTTTT
ATTCACTTTCCTGTCTTGAAAGCCTTTCCTCTTTCTTTTATTTTGAAACTTATTATTAACTCCTTTCTTCTTTTGATGAGGGGTAGCTGCACCACTA
TGAAATTTGTTGTAATCGACTTTACTGTAATCATAGTTAGGAGGCTCACTAGGTTTAGACTTTTCATTTCCTTCTACATGCTCTCGCCTATTTTTCC
CAATATTTCTGTTAAAATTCTTTCCATTGTTTTTAGGTGATATCTTAACTTCCCCTCCAGCAACACCAAAACTTGATTTCAATGGAATTTTTGGTTC
CTTCTCCTTTTTCACTGATATTTCACCCTTTTCTATTCTCGCTTTTTTAAGATACCCACGGTTAGTGGCAATTGGTTTTGGTTCTGGTTCCTCTGGC
TTTTCAGGTACTTTGCTGGCTGCATTTAATTCCAACGCATTTGTTGCACTAGAAAGCCTTTCCTCTTCTTCCTTTTCTCTCCTTTCTCTCTCTTCTT
TTTCGAGTTGTATTGCATACGGAAGAACGTTCAAATAACGCTCATATGGGCTAACAAACCTCAGCTTTACAGTACTGCTTTCTTCAGAATGATGACT
AGATTTATCCTTATGATTGCTTTCAAAAGCAAAGATAATAGGAGCACAATTTTGTTCGGGCACTGATGGAATTTCTGGTAATCCTCCTCCAAGCAAT
GTCGGTAGGTCATCTCTGAACTCGAACCCAGAGTTGTCTTGAGGAAGAATTAGCCTTAAGTCTCTTTCAACACCCGTGGCACTTTTGGAAGTTTCTT
CCACTAAACTTTTTACTAAAGGCTGTTCTTTAGCTTTAGAACAATCTGATGGAGTTCTAATAAACAAGTTCGAACAAGTGGAGGTATAGGGTTGCA
ACAAGCGAGTATCCCCTGTACTTCACGGGGAAGTGTTCTACATATCTGCACCATCATATGATTTGGAAGAACATAACCACAACTTTCATCTTCCTGC
CGAGCAATACGATCTCTCCACCTGTAGATCTCTTTCAAGCGTATAATTGCCTATTGTCAAAACATAGCTTCAGTCGTCGGTAAATAGCCATGTGAC

TATCCTCTTCCAAACTAGGTTTGCAATACACCTTTTTACATGTTTCTTTACTCCTATCAAACACAGCTTCTAAAAGATTTGTCTTCCCATTGGCATG
TTGTATTAAGTCATTACGCATATGTTGATATATATGTAATAGATAGTGGGTATCTTCTCGCGCGTATTTTAAAGCACTTTCTGGAAGTGGTCTTTTT
CTCCAATCAAAGAGCTGATATTCTTTGTTTGCTTCTACTTTACAATATTTATAAAGAAGATGAGCAAGACTGAGTCTAGCAAATCCTAACAACTTGC
TCGCTTGATAGGTATCAAACATGTTGACAACATAGATTGATAAGTCTCTCTGAAGCCATTCAATGTCACTATCGGCTCCATGAAACACCTTGAGAAT
TGAAGGCTTTGTAAAAACTTCATTAAGAATATGTAACTTATCCCTCAGTTTTAAAGCATCAATTACATAATCTTTTTCCCAAGTGGAAATTTGCATA
ACACAAGTAAATCCCATAAATGTCCTATATGAATGATGCTCTAAATCCACAGCGAATTCTTGAACCTCTGATAAATCTTGTAACAAGCCATTAAGCT
GTTCTTCTGTTTCAATGTAAATAAAAGGAGTATCGGCAAGAGGGAGAGGTGTTTTTACTTCTACATTAGTTAGCAAACTTTCATCTGGTTTCCAAAG
AGTCAGTTCATAATCATATGGGTGTGAAAAGCAATCTTGCCCAAGATCGTTCAACTCCAACATTATTGACAATGGTTTCAAGGAATTAGGCTTTTCT
TTAATGCGAGGTATCCATGGAGAATTGCTATTGTCAATCTTATCTTTAAAAAGGGTTTGGGGTCGTACAGTCCCTGATGAAGCACACAAAGTTATTT
CTTCTTTTGAAGGATAAGAATTAGTTTGTTTTTCAAACCAACCTCCACTTGGCACAGCTTGGGGAGCTACATTAGTAGCTGGTTTGGGAGGCAAAGG
ATTTTTTTTCCGGCCATCCAATTCTAACAAGTTATTATTCACCCTATCAAGACATCGATCATTAGAATCAATAACAGATTCAAAGCCCTGTTCAAGC
TTCCGGATACCATTCCTCTTCATTAAACATTCCATCAATCTTTGTAATTTGCTGCCCTCCTGCTTCATAATGTTAATAAACTCGGGATATGTTCGGT
AATACTTAAACTCCTGTGAAGTAGGTAAACCATTGGAGGCTTTGACTCCTTCAACAACAAAAGGAAATGCATTCTGCAAAAAATGCTTGTAGATTAGG
GTAACCTGGGATTAAATCTCCTGCACTGTTTTCTGCATTTGTAGAACATTCATTTTCACCGTCCATTTTCAGCCAAATGTATTTAAATAATCTGTAG
ATGTGATGTTACATGAAGGTGGATCTTTAAAAAAATATAAATAAGCAGTTCCCCAGCTCCACTGATCAAAAAATATAAACTTAACCTCAGAAAAACC
TCAAC

PROTEIN 809 (AA)
MDGENECSTNAENSAGDLIPGYPNLQAFLQNAFPFVVEGVKASNGLPTSQEFKYYRTYPEFINIMKQEGSKLQRLMECLMKRNGIRKLEQGFESVID
SNDRCLDRVNNNLLELDGRKKNPLPPKPATNVAPQAVPSGGWFEKQTNSYPSKEEITLCASSGTVRPQTLFKDKIDNSNSPWIPRIKEKPNSLKPLS
IMLELNDLGQDCFSHPYDYELTLWKPDESLLTNVEVKTPLPLADTPFIYIETEEQLNGLLQDLSEVQEFAVDLEHHSYRTFMGFTCVMQISTWEKDY
VIDALKLRDKLHILNEVFTKPSILKVFHGADSDIEWLQRDLSIYVVNMFDTYQASKLLGFARLSLAHLLYKYCKVEANKEYQLFDWRKRPLPESALK
YAREDTHYLLHIYQHMRNDLIQHANGKTNLLEAVFDRSKETCKKVYCKPSLEEDSHMAIYRRLKLCFDNRQLYALKEIYRWRDRIARQEDESCGYVL
PNHMMVQICRTLPREVQGILACCNPIPPLVRTCLLELHQIVLKAKEQPLVKSLVEETSKSATGVERDLRLILPQDNSGFEFRDDLPTLLGGGLPEIP
SVPEQNCAPIIFAFESNHKDKSSHHSEESSTVKLRFVSPYERYLNVLPYAIQLEKEERERREKEEEERLSSATNALELNAASKVPEKPEEPEPKPIA
TNRGYLKKARIEKGEISVKKEKEPKIPLKSSFGVAGGEVKISPKNNGKNFNRNIGKNRREHVEGNEKSKPSEPPNYDYSKVDYNKFHSGAATPHQKK
KGVNNKFQNKRKRKGFQDRKVNKKHRGGQNNVK

Graphical representation



Comparison with *Cylas puncticollis* Exosome

```
Query  1    MNSILRKNEVKGNMRNNVLEKKTELLIEACDVILERVTNNIDEVNGIRKNDVEPMVFQTV  60
            M  ++++N ++       LE+  E +I++ D  L+RV NN+ E++G +KN + P     V
Sbjct  76   MECLMKRNGIRK------LEQGFESVIDSNDRCLDRVNNNLLELDGRKKNPLPPKPATNV  129

Query  61   SVQLPVNGSWNRVSKPVMSATSAITSSQTTASGSSTQTIRLITAENIVRPQKFFKDKIDN  120
            + Q   +G W                 + T S  S + I L  +   VRPQ  FKDKIDN
Sbjct  130  APQAVPSGGW---------------FEKQTNSYPSKEEITLCASSGTVRPQTLFKDKIDN  174

Query  121  SRQRPWEPRIKEKPNSIKPLAIFLEETENGQE-YSHPYECELDKFSPPEDQLRKSDVVKP  179
            S   PW PRIKEKPNS+KPL+I LE  + GQ+ +SHPY+ EL   + P E  L   +V  P
Sbjct  175  SNS-PWIPRIKEKPNSLKPLSIMLELNDLGQDCFSHPYDYELTLWKPDESLLTNVEVKTP  233

Query  180  KLLKDTPLVEITQEEQLNDLVEDLRSAKEIGIDLEHHSYRSFMGITCLLQISTLEKDYLI  239
              L DTP + I  EEQLN L++DL   +E  +DLEHHSYR+FMG TC++QIST EKDY+I
Sbjct  234  LPLADTPFIYIETEEQLNGLLQDLSEVQEFAVDLEHHSYRTFMGFTCVMQISTWEKDYVI  293

Query  240  DTLVLRDKLYILNEIFTDPKIIKIFHGAKHDIQWLQRDLSLYVVNMFDTHFAAKQLEYPG  299
            D L LRDKL+ILNE+FT P I+K+FHGA  DI+WLQRDLS+YVVNMFDT+ A+K L +
Sbjct  294  DALKLRDKLHILNEVFTKPSILKVFHGADSDIEWLQRDLSIYVVNMFDTYQASKLLGFAR  353

Query  300  LSLAYLMQKFCNFAPNKQFQLADWRIRPLPHELKTYAREDTHYLIYIYQNMKNELIDKYN  359
            LSLA+L+ K+C    NK++QL DWR RPLP     YAREDTHYL++IYQ+M+N+LI   N
Sbjct  354  LSLAHLLYKYCKVEANKEYQLFDWRKRPLPESALKYAREDTHYLLHIYQHMRNDLIQHAN  413

Query  360  GQPNLLKSVIQMSTDLCKTRYVKPVFTSDSYLNLYRRCSRMFHNRQLFALQELYRWRDEI  419
            G+ NLL++V   S + CK  Y KP    DS++ +YRR    F NRQL+AL+E+YRWRD I
Sbjct  414  GKTNLLEAVFDRSKETCKKVYCKPSLEEDSHMAIYRRLKLCFDNRQLYALKEIYRWRDRI  473

Query  420  SREEDESTGYVLPNNMLLEISERLPREMQGILACCNPTPPLVKANLLDLHKIMLKALEQP  479
            +R+EDES GYVLPN+M+++I   LPRE+QGILACCNP PPLV+  LL+LH+I+LKA EQP
Sbjct  474  ARQEDESCGYVLPNHMMVQICRTLPREVQGILACCNPIPPLVRTCLLELHQIVLKAKEQP  533

Query  480  FESPILKEDTRSRGITTKISKIDVNNPLHCPHDLTKREEVCDNLPTILSGDSINLDSIFT  539
                +++E ++S       + ++ +  L  P D     E  D+LPT+L G    + S
Sbjct  534  LVKSLVEETSKS------ATGVERDLRLILPQD-NSGFEFRDDLPTLLGGGLPEIPS---  583

Query  540  HYIEEDSSVSSIFDFPVNAANETRLPNKLEGV--LKFLSPYARYKLIKPF  587
              + E +    IF F  N  +++   ++      L+F+SPY RY  + P+
Sbjct  584  --VPEQNCAPIIFAFESNHKDKSSHHSEESSTVKLRFVSPYERYLNVLPY  631
```

Polymerase

cDNA

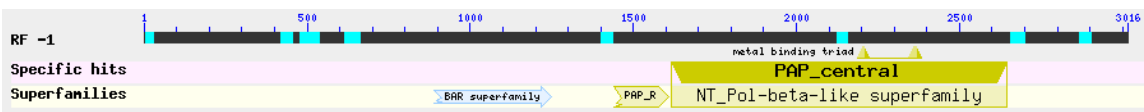
Contig 7462, Length-2376

AAAAAAAAAAAAAAAAAAAAAAAAGTAGTAAAGAACAATGAGTTAGGATTGATAAAAATAGAGTCTATATATTGTTCTTAATATTCAGCGGTAATAAT
GAGGAAGAAATCTTCTCAAATTATCACCTGCCACCATGAGCCCTCCAAGGTATAGCTAGAGCAGGTCTTAATGGGAAATGATGCAAGAAAGTATTAA
ATTTAAACTGCATTAAAAGCCCTTCGTTTCACAAAAAAATATAACAGGTGCACGACAGCTCTAATCGACAATTGTAAAGAGTTCAATAAAACGAACA
GGAAATTTAAAGGTTTCTTACACCTTGACCCAAGCTTAAAAATAAATTAAGACAACGACTATTAGAGAGAATATATTTCTTGATACAGTGGTGTCGA
TATATATATACAAAAGTAAGTCACTATATCAGTAAGAGGCGCTGGTTGGTTCACCTTAACAAAGTCTTCATTACCCTTTAACTTGGAAACAAAAACA
AAACAAAACAAATACAAAGCAATAACAATTCACTCACAACATTAACTTATAAAAAAGAATTTTTAAATGAATTTTAAGTATCCAAATGATAAGGGTT
CTGAATTTTGTTGAGATACTACTCCAAATTGAACTCTCAAAAAAAAAAAAAAAAAAAATTAAATAATTAAAAAAGAAATAATTAAGTATGGAATATTT
ATATCCACTATGTCCATTGCCTTCTGTGATTGGTTCAGAACCCCAACGTACTTACCGTGTCACAACAGGCATAAAGGCGAGAGCTAACTCATTCACC
TCATCGAGGTACAGCACGCACTTCATTACTATATATTACAAAATATATAGTTCCACAATGGTGATATCACCAAGGTTAATCTTAAAATGTATATTTA
TCATAAGCAAATAACCTCTTCCTGAATGGTGTCACTTTTTTTAGACTCTAAATTAGATTTCGAATCAATTATATCGTGGTGTCGATTGTGCTCATCA
GCGGAGAGAGAGGAATTAGAGTCTTCACCACAACTAACGACAATCCTGTTCATTGGAGGATCTCCGTCCATTTCAGGAAGGTGTGGATGAAGCTTCT
GCTTCTTCTGGATCAAGTCGAGGGAAGAGTCCGAATGACGCTTCTTGTTGACATCTCCGCTCTGCTGCTGTGAGACACCGTTCTTGTGAATTGGAGA
GCTACTGTTCCGAGCAGCCGCAACCTTTTTCTCTCTCTTCAATATTTGAGCTGGCAGGTAAGCGCTAAGCTGTTTTCTCTTAACATGCCTAGCTTCA
ATTTTCATGGAATCCTTGATCATCTTGGTTGCGGATGCTTGTCGGTGAACTGAGGTAGTGAAAGATTGTATGTCGTATGTGAGGTCAACATTCAAAT
TTTCCAACTTTCTGAATACAATGCCGATGAACCACATGGAAGTGGGTCCCTGTGGTTCACCAGGAGGAGTCGGAGGCCCAGCAAAGCACTCAGGGTT
TACGTGGGCCAGGCTGATATACTGGTTCCTTTCCAGGTTACCAACAAGGTGACGGACTTTAGACTCGATAAGCCCACACCACTCTAACTGTTCTTCC
TGGCTACCAGATTTTGCCAATAATACAATAAAGTGCTTGTATTTTAAAAAGAAATTAGGTGCTTCGAACAACTTATCCCAAGTTGCTTTGCCATTGA
AAATTTCATCAGTTATGGAAAGGCCATTATTAAACGCCTCTTGAAGTATAGTTCGAGTAGATGCGGAGACATTGAATGTCGAATTCTGTTGGGGGTA
AACAGGGGTAATTATTGGCATCAGGTGGTACCTATCTGAAACAGTCACCCTCGGGTCCCAAACAGGGAATCCAAGATTGACTTGGTCTGGCTGCTTC

AGAAGAACAGGCTGTGGCCATTTCCACTGTGAAAAAACTAAAAAAAAATTTATGGACTAACGTAGCAGCGACTGCATTCGGGTACAACTGGCATGTCC
TAGCGACAAGCATCGCCCAACTGACACCTCCTAAATATCCCATCACATTGCTGTAAACTCCATGCTTTTTTGCCCAAAGCTTTATCGTCCTTAAGGT
CAGACGAAAGTTTTCTATGTCAGGTACCAATCGGAGGATTTCATCAGTGACCCTGCAACCATTGAGGCTTCTTACACATTTCTGGTCCAAATTCTTC
AGCAGCAGGTCATCCTTGAGGTCCATTGAATCAGGAATCTCTTTGAGAGCGAGTCGAGCAAAAAGCATATCAATCTCAATTCCATCAAAGTTCATCT
TTATTACAGGTACAAATGCTTCTTCAACAGCCCTTAAATCAGTAACCTCTGAATTGAGTTTCAGTTTTTCAAAGAACGAGGTAAAATAATCTGATCG
ATCAATATGCCTAGGAGCAACACAGAGAGCGTCAATATCTGCACCTTTGTGATGAACACCTAGGCGGTAGGAACCAAAGGTGTATATTTTCCCACCA
ACAGAACTAGCAACACTCTCAGGCATGTTGCGACTTATACTAACTTCCTTTATCCATTCTTTCACCAAACTATTCAATTTACTTAAAATCTCCATCC
TATGGTTTAACTCCTCCTCCGTCTCCATGACATTGTATGGCTTTAGGGCTTCTGTGAGTTCATTGGTTTTTTGGTAATCGATGGGCTTTGGTCCAGC
AAGTGTGATAGCCGAGGTCAACCCAAGAGTGCGCTGTTGTATTTGCTGCTGTGGTGCAGGTTGCGCTGAGCCCTGCGACGGCCACATACCTCTTCTT
GTTACAGCGGCCAGGGAGCATGATCAATGGATAAGGTTCATTAGAGAAGTATTTGGATATAGTGGGTCATCGTGTTCTTCAATTCCCCATTTGTTAT
GGACTACCACTTGAGAACAAAGACAAATCAACGAAAGAAAAAAAATATTATTTTTTTAAATAGTCAAGTATACTATTTTTTTTTTAAAAACACGTG
CTGTTACAAATAAACACAGCTTCACAAAGGGGAAACTAATCCATTTAAGGTTTCACTTAGTCTTAGGAACTTAGTATTAAGTTGTTGACAAGTCCTA
CTGCTTAC

PROTEIN 610 (AA)
MWPSQGSAQPAPQQQIQQRTLGLTSAITLAGPKPIDYQKTNELTEALKPYNVMETEEELNHRMEILSKLNSLVKEWIKEVSISRNMPESVASSVGGK
IYTFGSYRLGVHHKGADIDALCVAPRHIDRSDYFTSFFEKLKLNSEVTDLRAVEEAFVPVIKMNFDGIEIDMLFARLALKEIPDSMDLKDDLLLKNL
DQKCVRSLNGCRVTDEILRLVPDIENFRLTLRTIKLWAKKHGVYSNVMGYLGGVSWAMLVARTCQLYPNAVAATLVHKFFLVFSQWKWPQPVLLKQP
DQVNLGFPVWDPRVTVSDRYHLMPIITPVYPQQNSTFNVSASTRTILQEAFNNGLSITDEIFNGKATWDKLFEAPNFFLKYKHFIVLLAKSGSQEEQ
LEWCGLIESKVRHLVGNLERNQYISLAHVNPECFAGPPTPPGEPQGPTSMWFIGIVFRKLENLNVDLTYDIQSFTTSVHRQASATKMIKDSMKIEAR
HVKRKQLSAYLPAQILKREKKVAAARNSSSPIHKNGVSQQQSGDVNKKRHSDSSLDLIQKKQKLHPHLPEMDGDPPMNRIVVSCGEDSNSSLSADEH
NRHHDIIDSKSNLESKKSDTIQEEVICL

Graphical representation



Comparison with *Cylas puncticollis* Polymerase

```
Query  25   TLGMTSAISAAPPKPIDLVKTKELEQTLIPFDVFETEQELNHRMLILGKLYSLVKQWIKE  84
            TLG+TSAI+ A PKPID  KT EL + L P++V ETE+ELNHRM IL KL SLVK+WIKE
Sbjct  20   TLGLTSAITLAGPKPIDYQKTNELTEALKPYNVMETEEELNHRMEILSKLNSLVKEWIKE  79

Query  85   LSIKCNMPESVAENVGGKVYTFGSYRLGVHNKGADIDALCVAPKHICRSDFFTSFYELLK  144
            +SI  NMPESVA +VGGK+YTFGSYRLGVH+KGADIDALCVAP+HI RSD+FTSF+E LK
Sbjct  80   VSISRNMPESVASSVGGKIYTFGSYRLGVHHKGADIDALCVAPRHIDRSDYFTSFFEKLK  139

Query  145  KQPEVTELRAVEEAFVPVIKMSFDGIEIDLLFARLLQKEIPDSMDLRDDNLLKNLDQKCV  204
               EVT+LRAVEEAFVPVIKM+FDGIEID+LFARL  KEIPDSMDL+DD LLKNLDQKCV
Sbjct  140  LNSEVTDLRAVEEAFVPVIKMNFDGIEIDMLFARLALKEIPDSMDLKDDLLLKNLDQKCV  199

Query  205  RSLNGCRVTDEILRLVPNIETFRLALRAIKLWAKRHGIYSNVLGYLGGVSWAMLVARTCQ  264
            RSLNGCRVTDEILRLVP+IE FRL LR IKLWAK+HG+YSNV+GYLGGVSWAMLVARTCQ
Sbjct  200  RSLNGCRVTDEILRLVPDIENFRLTLRTIKLWAKKHGVYSNVMGYLGGVSWAMLVARTCQ  259

Query  265  LYPNAAPATLVHKFFLIFSKWQWPQPVLLKQPTNASLGFTVWDPRVNIQDRYHLMPIITP  324
            LYPNA  ATLVHKFFL+FS+W+WPQPVLLKQP    LGF VWDPRV + DRYHLMPIITP
Sbjct  260  LYPNAVAATLVHKFFLVFSQWKWPQPVLLKQPDQVNLGFPVWDPRVTVSDRYHLMPIITP  319

Query  325  AYPQQNSTFNVSNSTRNIIMEEFERGLKITDQIMLNKSTWERLFVQPPFFMKYRHFIVLI  384
             YPQQNSTFNVS STR I+ E F  GL ITD+I   K+TW++LF  P FF+KY+HFIVL+
Sbjct  320  VYPQQNSTFNVSASTRTILQEAFNNGLSITDEIFNGKATWDKLFEAPNFFLKYKHFIVLL  379

Query  385  ISAETSKDLLEWSGLVESKIRLLISTLERNGAITLAHINPENFML-PEPLRKPKFQYSMW  443
              + + ++ LEW GL+ESK+R L+  LERN  I+LAH+NPE F   P P  +P+   SMW
Sbjct  380  AKSGSQEEQLEWCGLIESKVRHLVGNLERNQYISLAHVNPECFAGPPTPPGEPQGPTSMW  439

Query  444  FIGLEFRKMEGLNVDLTQDIQQFTETVNKHAMHIQMFKDTMKLEARHVKRKMLPNYLPSG  503
            FIG+ FRK+E LNVDLT DIQ FT +V++ A   +M KD+MK+EARHVKRK L  YLP+
Sbjct  440  FIGIVFRKLENLNVDLTYDIQSFTTSVHRQASATKMIKDSMKIEARHVKRKQLSAYLPAQ  499

Query  504  VLKREKKYLATFTE--------CKEQKEG---RKRQSESSDSETTVKKTRLS---NEMNG  549
            +LKREKK  A             +Q+ G   +KR S+SS  +   KK +L     EM+G
Sbjct  500  ILKREKKVAAARNSSSPIHKNGVSQQQSGDVNKKRHSDSS-LDLIQKKQKLHPHLPEMDG  558

Query  550  --------TNFDDSSNNSLNNDD  564
                     +  + SN+SL+ D+
Sbjct  559  DPPMNRIVVSCGEDSNSSLSADE  581
```

dsRNA uptake in *Megacopta cribraria*

Hermansky(CG4966)

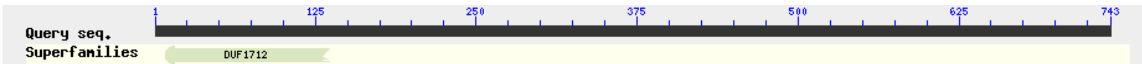
cDNA

Contig 5140, Length-2217
ATTTAGTTTATCTCTTGTCTACATTTCAATTAAAATACCCTGAATACCATTGACTAGTTAAAAACTGGCATCAATTTATTCATTCAATAGCAACATT
AACATCCGATGTTAACATCGGAAAGTAAATCATAACAAAATCTCAAATTTGCATAAACTTTATAAACTTAATACAACAGGCTATAAACAATTTATAG
AAATGTCTTTTAAACCTACTTTAACATTAGTAACTTTTAAAAATTTTACAAAGGTCACAAATTGTATTATAATAAGGTCTTGGTTTTTTCCTATGTG
AAGTTTCACACCATTCAGGATTAAGAGAAGGCAGACCTATACTGTAATACCTTATAGGAATTTTTTTTGTGGTAAGATCATCAGGGGTCATAGCTGA
TCTTACATTACTGGAAGCTTCCGGTATTATGAAATTGTTAGGAACTACTGCCTCAATTCTCGATAAATGGCGAGCTTTTACAGCATCACGGACAACA
TCAGGAATATCACCTTGATCTTTTCCTTCCAGATTTTCATTTCCTTCTTCCTCTAGCACCGTAAAAATTCTAGATACATCTCTTTTTAAACCACACA
TTGAAATGTATGAGTCAGAAGTTTTTAAACAAGTGTTCTTATTTTGATTTTGAATCATAGAATCCTCCAGTTGGTTCCTCCAAATGGCGACATCTTT
TAAGGTATTATGTAGCTCATTATATTCATCTTTCAGTCTTTTTATTTCAGCTGATGAAATATAAACTTTCATGAGTTGCACACCAAATGGAAGATGA
AAAGGAGTAGATACAATTTCTGCAGGTAACTTAATTCTATATGGGTCAGTCACAACAAAGAGTTTGGTTAAGGATTTAGCCAGTTGTGTAGCAACAA
CCCTATTTTGATAAAGCAGTACACCTCCAAGGATTCCCTTTATTTCTTGAAATCCTTGTAAAACTTGCCTGGCTTCTAAATACAGAGTACTTGCACT
CTTAGGAATTTTGAATAATGGAATATTTCCAAATTTATTAGTGGCATATTGCAATATCGGTAAATAAATCTGGAGCAGTTCAGACAATTTTTCAGAA
ATATCATCATCTAATAATGATTCTTGTAATGTTTCAATACTGTTATGAAACAACCTTATTAACTTATATAAAATCTTAACTCTTGTTTCCAATACTA
TATCAGGAATATTTCTGTCAGTTCCTACACACAAAGCATATTTACCTAATTTTTTCAAAGCAAATTTGCCACTTTGTAGAGAAATAATATTTGGAGG
AGAAAATGATGTAGAACAAAATTGGGTAACACCCATTAATTGACTACACAAGGCATGTCTCTGTTCATCGGTGACCCAAGAGGGGTAGAAGTAAAGA

ATAGAATTTTGTGGATCATCTTCTTCTCCTTTACATAATTCACAATCATAAATAAAAACTATCAACAATTCCTTTGCCATTGAGCTGATGTAAAATT
TCTATCAGTATTACATAACAGACATATCCAACAGTTGTTTGACTGTTTATTTTGAG

PROTEIN 405 (AA)
MAKELLIVFIYDCELCKGEEDDPQNSILYFYPSWVTDEQRHALCSQLMGVTQFCSTSFSPPNIISLQSGKFALKKLGKYALCVGTDRNIPDIVLETR
VKILYKLIRLFHNSIETLQESLLDDDISEKLSELLQIYLPILQYATNKFGNIPLFKIPKSASTLYLEARQVLQGFQEIKGILGGVLLYQNRVVATQL
AKSLTKLFVVTDPYRIKLPAEIVSTPFHLPFGVQLMKVYISSAEIKRLKDEYNELHNTLKDVAIWRNQLEDSMIQNQNKNTCLKTSDSYISMCGLKR
DVSRIFTVLEEEGNENLEGKDQGDIPDVVRDAVKARHLSRIEAVVPNNFIIPEASSNVRSAMTPDDLTTKKIPIRYYSIGLPSLNPEWCETSHRKKP
RPYYNTICDLCKIFKSY

Graphical representation



Comparison with *Cylas puncticollis* Hermansky(CG4966)

```
Query  1    MAKETTIIFVYDTDKLQNEEDDPASAILYFHPTWVSDQQKTALCGQIVGTISCLKSIFSM  60
            MAKE  I+F+YD +   + EEDDP ++ILYF+P+WV+D+Q+ ALC Q++G      + FS
Sbjct  1    MAKELLIVFIYDCELCKGEEDDPQNSILYFYPSWVTDEQRHALCSQLMGVTQFCSTSFSP  60

Query  61   PKVVCLQTGKFYIIENGQYILAVGTDRNIADWLLEHRGTLLYSLVTLFNKDFDMLGRIYK  120
            P ++ LQ+GKF + + G+Y L VGTDRNI D +LE R  +LY L+ LF+   + L
Sbjct  61   PNIISLQSGKFALKKLGKYALCVGTDRNIPDIVLETRVKILYKLIRLFHNSIETLQESLL  120

Query  121  DEALKAKLYHLFETYLRMLFYGGSIFSHVPVLCLPKSASAVFTEAVTFLEACREYSHVLG  180
            D+ +  KL  L + YL +L Y  + F ++P+  +PKSAS ++ EA   L+  +E   +LG
Sbjct  121  DDDISEKLSELLQIYLPILQYATNKFGNIPLFKIPKSASTLYLEARQVLQGFQEIKGILG  180

Query  181  GTLLYHNKVVATQLPTDITKRIVIADPYRIKSPAEPAETAFKLPLGVQLLQVYITN----  236
            G LLY N+VVATQL   +TK  V+ DPYRIK PAE   T F LP GVQL++VYI++
Sbjct  181  GVLLYQNRVVATQLAKSLTKLFVVTDPYRIKLPAEIVSTPFHLPFGVQLMKVYISSAEIK  240

Query  237  --KDFCDLLETSMRNKCIFQYLNNKSIKRAHSKS----SSEAPI-VSAMKRDQSLLFTA  288
              KD  + L  ++++  I++     S+ +  +K+    +S++ I +  +KRD S +FT
Sbjct  241  RLKDEYNELHNTLKDVAIWRNQLEDSMIQNQNKNTCLKTSDSYISMCGLKRDVSRIFTV  299
```

FBX011

cDNA

Contig 3711, Length-6097
AAATGGGTATTTTGCCAGACGTAAAGGCTCAGTACATACATTATTGACAGAATACAACATCCTATAATAGGTAAAACACAAACACCTCTCACAGTAG
TATCCATTTTACCTTCTCAAAACAAAATATCATGTGATCACGTGACGTCATTATTAGAGTTTTTCTTCCGCCTTTAGTAATGGCTGCTCGGTGATTG
GGTTCTTCTCTCGTTGTCGTATTGTTTTTACTCTGCTCGTGACTTGTTATTAAAATGCCGAGCGCTTCATATGGTTCGTCCAGATATTATCCAAGAA
GATCAAGGCGTAAGGGTAATAATAGGATAGCAGTTTCTAACAGGAACAATAATGCTGAGACTTGTGCATCAGGTAGTGGCTCAAATGGAACTTCTTC
ACCTGTAATGAGTTCACAACCTCCTCATAACCCTCCTTATGACCTCCGTAGAAAGTCACCTCCATACCCAATTTTGGATGCCGGTTTTCCTAATCTA
CGTAAAAGACCTCGTAGATCATGCTCAGTGAGTACAGATGTATGTACACCACCTGCTGCTCATTATTTGCAATATGAGCTACCTGACGAGGTGTTGT
TAACTATTTTTGGTTACTTGTTGGAAAAAGACTTATGCCGTCTTTCTCAAGTCTGTAAGCGTTTTCAAACTATTTCAAATGATAATGGTCTATGGAA
ACGTTTGTATCATAATATTTTTGAATATGATGTCCCATTATTTAATCCAGAACCATGCAAATTTGAATTTATACCTCCAAATGATTGTGAATATGCA
AACCCTTGGAAAGAAAGTTTTAGGCAATTATATAGGGGAATCCATGTTCGTCCTGGGCAAGGATCACGTTATAAAGCAAGAGGTAGATCTCTTGCCC
ATTTTGATACGGTTCAAACTGCCTTAGACCATGCTGATGAAAACCAAACCCCTGGTGTACCAGCTATTGTTTTTCTTCATGCTGGTACCTATAGAGG
AGAATTTTTGGTTATAGATTCTGATGTTGCTCTTATTGGAGCTGCAGCTGGCAATGTTGCTGAATCTGTCGTTTTAGAAAGACAATCCGAATCTACA
TGATGTTTGTTGAAGGTGCCAAATCTGCATATGCCGGTCATCTCACTTTAAAATTCACTCCTGATGTTGCCTCCAGTGTTCCTCACCACAAACATTA
TTGTTTGGAAATTGGAGAAAACTGCAGCCCCACAATTGACCATTGTATTATTAGAAGCACTTCAGTTGTTGGAGCTGCTGTTTGTGTTAGTGGTGCT
GGAGCTGCTCCCACTATCAAGCATTGTGATATTAGCGACTGTGAAAATGTTGGTTTGTATGTTACTGATTATGCTCATGGAACATACGAAGACAATG
AAATTTGCAGGAATGCTCTTGCTGGTATTTGGGTGAAAAATTATGCAAATCCTATAATGAGGAGAAATCATATCCATCATGGAAGAGATGTTGGCAT
TTTCACATTTGATAATGGTTTGGGTTTTTTCGAGGCAAATGACATTCACAATAATAGAATTGCTGGGTTTGAAGTTAAAGCTGGAGCAAATCCTACT
GTAGTCCAGTGTGAAATTCATCATGGTCAAACTGGTGGAATATATGTTCATGAAAATGGTCTTGGACAATTTATTGATAACAGAATTCACTCTAATA
ATTTTGCCGGAGTTTGGATTACTTCTAACAGTAATCCTACCATAAGACGTAATGAAATATATAATGGTCAGCAAGGCGGTGTTTATATTTTTGGAGA
AGGACGTGGTTTGATTGAGCACAACAATATTTATGGAAACGCCCTGGCAGGAATTCAAATACGCACCAACAGTGATCCAATTGTTCGTCACAACAAA
ATACATCATGGACAACATGGTGGCATTTATGTTCATGAAAAAGGAGAAGGCTTAATTGAAGAAAATGAAGTATATGCTAATACCTTGGCAGGTGTTT
GGATTACTACAGGTTCTACTCCAGTTCTCCGCCGTAATAGGATTCATTCTGGAAAACAAGTTGGCGTTTATTTTTATGATAACGGCCATGGGAAGTT

AGAAGATAATGATATATTCAATCATTTGTATTCAGGAGTTCAGATTAGGACTGGTAGCAATCCTATTATTCGTGGTAATAAAATTTGGGGAGGTCAG
AATGGTGGTGTTTTAGTCTATAATGGGGGTCTTGGCCTATTGGAACAGAATGAAATTTTTGATAATGCCATGGCAGGTGTTTGGATTAAAACAGATT
CTAACCCAACTCTTAAGCGGAATAAAATTTTTGATGGAAGGGATGGAGGCATTTGTATATTTAATGGCGGCAAAGGTATTTTAGAAGAAAATGACAT
ATTTAGAAATGCTCAAGCTGGTGTATTAATTTCTACTCAAAGTCATCCAGTTCTAAGACACAATCGAATATTTGATGGATTAGCTGCTGGAATTGAA
ATAACTTTGAATGCCACCGCTACTTTAGAAAGTAATCAAATTTTCAATAATAGATTTGGTGGACTTTGTTTGGCTACTGGTGTTAACCCTATTGTGA
GATGCAATAAAATATTTAGCAACCAAGATGCCGTGGAAAAGGCTGTGCGAAATGGGCAGTGTCTTTATAAAATATCAAGTTACACTTCTTTTCCCAT
GCATGATTTTTATCGTTGCCAGACATGTAATACTACTGGTAGAAATGCAATTTGTGTAAATTGTATTAAAACTTGTCATGCAGGTCATGAAGTTGAA
TTTGTTAGACATGATAGGTTTTTTGTGATTGTGGAGCTGGCACTCTATGTAACCAGTGTCAACTACAAGGAGAGCCTGCACAAGATACTGATACTT
TATATGATTCCGCTGAACCAATGGAATCTCATACTCTTATGGTTAATTGAACTTTGTGGTGTTACCATTATTTATTGAAACTGTACACTTTCTATTG
TTCAGGATTATCGTTGTATTATAGAACTAGGTAAAATGTGTGTTTTTAAGAGTCAACAAGAGCTATAGCATATTCACACCACACACATCACTGAGGT
AGTAAATAGAAATTATGTATATAAATATGTTGAAAGTGATATATTTGTATTTATCTTTCCACGATTATTTTCATTATTCTACTTTTTAAAATTATCT
TATTGGTGGTAGAAAGAAACGTTTTCAATATTCATCCTTTTTGGGAGTATTCTTCTTCCAGGAGGTTTCTAAATAAGAAGATCGAAACTCCTAAAAA
GATGATTTTTTGTTGTTTTTTTTTGTTTATTGTTATCAAGTCTTTTTTCTTATAATTTGAGAGGAAACTTCTTTCATTCTGATTAATTAAAAAAATT
CTGATATTGAAAATAATTGTTATTTTAACTATTAATAACTGGTGCATTTTGTTTACTTGATTATAATTGTGTTAATTCACATTGAATTTAAATTTTA
CGAAGCAATCTGAAAATTTAAGCTATTATATCTATTATTGTTTTTGATTTAGTTAAGTTTTTTGCCCCCTTTATATTCAAACTATTGTATAATTGAT
TGTTTAAATTGTTATTTCAATGTTGGAAACATCATTGTTGCTTGATATTCCATTTCTAAGTATATAAAATTATTTGAATATTAGGGGAATGATCAGT
TATTTTATAGAATGTATGGTGTAAACACAACTTTTTTTTATTTACTCAGACAAAATTTAACAAATTTTTTTGCATCTTATAAAAAACCTCTGATACA
TATGTATGTGTAAAGGATGTCTAATAATTAATTTTAAAACTAACCTCTTCGATTTTGTGTATAATGAAAAAATAAAATTATTTTTTTATTTAAAATT
TTGTATTGTATTGTGTATATATATTGTTGGTTTTACTGAACTGTTGAGTATGATGAGTGAATTATTTGTTTGTTTTTTTTTCTATTTACTTTCGTTT
TCTTGCCATTTATTCTTAGACTTTGAAAGACTGTTAGTATTTTAAGTATTTTGTAATTGATGGTACATTATACTTGATGGTGTATTAAGTAAGAGTT
TGTGTGCGTGTATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCGTGTGAAGCTGTAGTAAACTCTTGTAACTTACCAATTTATAGCTTAT
GACTATTTAATGTAACTTCTTTTTCATTGTCTTAATATAAATATCGACTGTTTTTGAATAACTGATATAGCCCTTCAGCTGGTTAGATGTTAGAATT
TAAACCCATTGATAAGTATATGATTATGTTTTTTATGTCAAAACTTTTTGCTAATTATTTTTGATTGGTGGTTGCATCCCTGAGATATATGCTAAAT
GTTTGGCTTAGCAAAGAATTATGGCTGCTTATCCATTTTCAGACTTTTATAATGTATACAATTTGTTTCTTTTCCAAACAATATGTTTTTAAACTCC
AATTAATTTATAGTTTAATTTTTGTCTAAAAAGTCCCTGAAATAAGGCTTCTATAAACTAGTGTATAACATGAGCGAATGTTTTATGACTGATAAGT
GTGTTATAACTTTGGCTGTATAATAATCATATTTTTTAAATAATTGTAAAAATTGCTTTGTAAGTAGTGATATGGAAAGTTGAGTTAATTCATAAGT
TCTAAAACTCAGCTTTTCCCCCCTTATGTTTAAAATTAAAGGTGAGAAACTAGTGTAATTTTAATTTGTTTATTTTTTTCCCCCTATATACATTTTA
GTTAACTTAGGAAGAGGGCGCTGTGTAATCTTCCACATCAATAACAAATTTTAAGGAATTAATCAAGTAGATATCAGTTATTCGTATTGATAATAAT
TTAATGGACTCCATTATTAAATAAATAAGTTGCACAATTCTTTGGAGATCATAAACTAGATACAGTTACTTGGAATTATGAAAATCAGTTTTTTTTT
GCAGTTTATAAGAATGCTTAGGTGAAACTACTTATTTACACCATTATGTAAAACACTCGATTTGGTTCAAGCAGTTCGTAGAAAATTTAGGAATTTT
CCTACTATATTACATTTTCATCATATTTGATTTAAATGAGTTTGCTTTTGAAATTCATTATGCATTATATGCAGGTTTTACTAATTATACTAATTGA
ATCATTAATGAATAATTTTTTTGTGAATTAATTACCTTTTCAAAATTAGTTAAGTATTATGATGAATTAGTTAAGTTAACTGTAGATTTTAATATTG
TGATTTTTTTTTTAATCATACTTGATTTTGCTTCTCCAATTTAAGACTGTATGTACCTTATGAATATATGGAAATGCAGATATCTATTTAATTGTAA
AGTTGATTTTTCTCTTTACTTATGGAGTCCTTATATTTAAAAATACTGTCTTCCTATTAAAAGCCACATCATTATAATGCTACAACAATTTTGTGAA
TAGTTACTGCTTTATATTTTTCCTCCCACTTCATTTGTTTGTTTAACATGTTTTATTTTGGTTAAAATTTTAAACTGTAATTATTGTAAAAAATGAA
CCCAGTTATATAACATTTCCTCATCTTTTTTGGAGAACATTTAAAACTCTTTCCTATAACTGTTATGTTGTGGGATAAATAGTTGTGCTGGATGTGC
CTCATGTTTTTGATTAAATTTCATCATTACTTAAAAGTTTCTTGTCTTTGGACTTGGTGAAGTCCAGTATCACCACTATAAAAGTGCCTTATTCAAA
AACAATATGACACTGGTTTTAAAATAATTAAGACATTTATTTAATTGGGAAATGAATTGAGTTCATTAATATTGTTTTATAATCATTTTTAGATGT
TTCTGGTATTATTAATTAAAATATATATAACTACTCAGAAAACTGAACAAGAGCTTTATTCTTTCTATCTGTTACAAAATTCCAAAAATACACGAAA
TGTATTTTTCTTACTTATTACTGACCCAATAAAAATATAGACTAATATACTGCAGTTGCATAGTAAACCAATTTTTTGCTCCTTGTGTTTTAATTTG
TAAAATATCCTAAATTTGGCTTTCAGCTCTTTGTTATTTTAATTCTAGTTTCTTTATTGGTTGTGATTTACTAATGAAGGTATGTTATATGTAGGGA
AAGAATTTTTATAATGAAAATTTTCTAACCTTATGGCTTGAACTTCGATTTTTTCATATAGTAATCATTTACATTCTTCTTTTTTATCTGTGTTTCT
ATATTTCTGATAGTTTTTTGCATAATTGCAAG

PROTEIN 871 (AA)
MPSASYGSSRYYPRRSRRKGNNRIAVSNRNNNAETCASGSGSNGTSSPVMSSQPPHNPPYDLRRKSPPYPILDAGFPNLRKRPRRSCSVSTDVCTPP
AAHYLQYELPDEVLLTIFGYLLEKDLCRLSQVCKRFQTISNDNGLWKRLYHNIFEYDVPLFNPEPCKFEFIPPNDCEYANPWKESFRQLYRGIHVRP
GQGSRYKARGRSLAHFDTVQTALDHADENQTPGVPAIVFLHAGTYRGEFLVIDSDVALIGAAAGNVAESVVLERQSESTMMFVEGAKSAYAGHLTLK
FTPDVASSVPHHKHYCLEIGENCSPTIDHCIIRSTSVVGAAVCVSGAGAAPTIKHCDISDCENVGLYVTDYAHGTYEDNEICRNALAGIWVKNYANP
IMRRNHIHHGRDVGIFTFDNGLGFFEANDIHNNRIAGFEVKAGANPTVVQCEIHHGQTGGIYVHENGLGQFIDNRIHSNNFAGVWITSNSNPTIRRN
EIYNGQQGGVYIFGEGRGLIEHNNIYGNALAGIQIRTNSDPIVRHNKIHHGQHGGIYVHEKGEGLIEENEVYANTLAGVWITTGSTPVLRRNRIHSG
KQVGVYFYDNGHGKLEDNDIFNHLYSGVQIRTGSNPIIRGNKIWGGQNGGVLVYNGGLGLLEQNEIFDNAMAGVWIKTDSNPTLKRNKIFDGRDGGI
CIFNGGKGILEENDIFRNAQAGVLISTQSHPVLRHNRIFDGLAAGIEITLNATATLESNQIFNNRFGGLCLATGVNPIVRCNKIFSNQDAVEKAVRN
GQCLYKISSYTSFPMHDFYRCQTCNTTGRNAICVNCIKTCHAGHEVEFVRHDRFFCDCGAGTLCNQCQLQGEPAQDTDTLYDSAEPMESHTLMVN

Graphical representation

Comparison with *Cylas puncticollis* FBX011

```
Query  86   HQSPYDLRRKSPSHNHGGGAPVADGGPGPSTSAATASGSPASPTTPCNLPAAQGAYSSAM  145
            H  PYDLRRKSP +      P+ D G                       P
Sbjct  56   HNPPYDLRRKSPPY------PILDAG----------------------FPNL--------  79

Query  146  LPARKRPRRTCSASYENCTNTAAHYLQYELPDEVLLTIFNYLLEQDLCRVSQVCKRFQVI  205
               RKRPRR+CS S + CT  AAHYLQYELPDEVLLTIF YLLE+DLCR+SQVCKRFQ I
Sbjct  80   ---RKRPRRSCSVSTDVCTPPAAHYLQYELPDEVLLTIFGYLLEKDLCRLSQVCKRFQTI  136

Query  206  ANDTELWKRLYQNVYEYDLPLFNPSPCRFEFVSSEESELANPWKESFRQLYRGIHVRPGY  265
            +ND  LWKRLY N++EYD+PLFNP PC+FEF+    + E ANPWKESFRQLYRGIHVRPG
Sbjct  137  SNDNGLWKRLYHNIFEYDVPLFNPEPCKFEFIPPNDCEYANPWKESFRQLYRGIHVRPGQ  196

Query  266  -QELSFKGRNLVYFNTVQGALDYADERSGSGSGTGLAVSGGGSHSSENAGQGALIFLHAG  324
                  +GR+L +F+TVQ ALD+ADE                   ++  G  A++FLHAG
Sbjct  197  GSRYKARGRSLAHFDTVQTALDHADE-------------------NQTPGVPAIVFLHAG  237

Query  325  TYRGEFLVIDSDIALIGAAPGNVAESVILERESESTVMFVEGAKNAYAGHLTLKFSPEPA  384
            TYRGEFLVIDSD+ALIGAA GNVAESV+LER+SEST+MFVEGAK+AYAGHLTLKF+P+ A
Sbjct  238  TYRGEFLVIDSDVALIGAAAGNVAESVVLERQSESTMMFVEGAKSAYAGHLTLKFTPDVA  297

Query  385  SATPHHKHYCLEVGENCSPTIDHCIIRSTSVVGAAVCVSGVGANPVIRHCDISDCENVGL  444
            S+ PHHKHYCLE+GENCSPTIDHCIIRSTSVVGAAVCVSG GA P I+HCDISDCENVGL
Sbjct  298  SSVPHHKHYCLEIGENCSPTIDHCIIRSTSVVGAAVCVSGAGAAPTIKHCDISDCENVGL  357

Query  445  YVTDYAQGTYEDNEISRNALAGIWVKNYANPIMRRNHIHHGRDVGIFTFDNGLGYFEAND  504
            YVTDYA GTYEDNEI RNALAGIWVKNYANPIMRRNHIHHGRDVGIFTFDNGLG+FEAND
Sbjct  358  YVTDYAHGTYEDNEICRNALAGIWVKNYANPIMRRNHIHHGRDVGIFTFDNGLGFFEAND  417

Query  505  IHNNRIAGFEVKAGANPTVVHCEIHHGQTGGIYVHENGLGQFIDNKIHSNNFAGVWITSN  564
            IHNNRIAGFEVKAGANPTVV CEIHHGQTGGIYVHENGLGQFIDN+IHSNNFAGVWITSN
Sbjct  418  IHNNRIAGFEVKAGANPTVVQCEIHHGQTGGIYVHENGLGQFIDNRIHSNNFAGVWITSN  477

Query  565  SNPTIRRNEIYNGHQGGVYIFGEGRGLIEHNNIYGNALAGIQIRTNSDPIVRHNKIHHGQ  624
            SNPTIRRNEIYNG QGGVYIFGEGRGLIEHNNIYGNALAGIQIRTNSDPIVRHNKIHHGQ
Sbjct  478  SNPTIRRNEIYNGQQGGVYIFGEGRGLIEHNNIYGNALAGIQIRTNSDPIVRHNKIHHGQ  537

Query  625  HGGIYVHEKGQGLIEENEVYANTLAGVWITTGSTPVLRRNRIHSGKQVGVYFYDNGHGKL  684
            HGGIYVHEKG+GLIEENEVYANTLAGVWITTGSTPVLRRNRIHSGKQVGVYFYDNGHGKL
Sbjct  538  HGGIYVHEKGEGLIEENEVYANTLAGVWITTGSTPVLRRNRIHSGKQVGVYFYDNGHGKL  597

Query  685  EDNDIFNHLYSGVQIRTGSNPVIRGNKIWGGQNGGVLVYNGGLGLLEQNEIFDNAMAGVW  744
            EDNDIFNHLYSGVQIRTGSNP+IRGNKIWGGQNGGVLVYNGGLGLLEQNEIFDNAMAGVW
Sbjct  598  EDNDIFNHLYSGVQIRTGSNPIIRGNKIWGGQNGGVLVYNGGLGLLEQNEIFDNAMAGVW  657

Query  745  IKTDSNPTLKRNKIFDGRDGGICIFNGGKGILEENDIFRNAQAGVLISTQSHPILRRNRI  804
            IKTDSNPTLKRNKIFDGRDGGICIFNGGKGILEENDIFRNAQAGVLISTQSHP+LR NRI
Sbjct  658  IKTDSNPTLKRNKIFDGRDGGICIFNGGKGILEENDIFRNAQAGVLISTQSHPVLRHNRI  717

Query  805  FDGLAAGVEITNNATATLEFNQIFNNRFGGLCLASGVQPIVRGNKIFSNQDAVEKAVANG  864
            FDGLAAG+EIT NATATLE NQIFNNRFGGLCLA+GV PIVR NKIFSNQDAVEKAV NG
Sbjct  718  FDGLAAGIEITLNATATLESNQIFNNRFGGLCLATGVNPIVRCNKIFSNQDAVEKAVRNG  777

Query  865  QCLYKISSYTSFPMHDFYRCQTCNTTDRNAICVNCIKTCHAGHDVEFIRHDRFFCDCGAG  924
            QCLYKISSYTSFPMHDFYRCQTCNTT RNAICVNCIKTCHAGH+VEF+RHDRFFCDCGAG
Sbjct  778  QCLYKISSYTSFPMHDFYRCQTCNTTGRNAICVNCIKTCHAGHEVEFVRHDRFFCDCGAG  837

Query  925  TLTNQCQLQGEPTQDTDTLYDSAAPMESHTLMVN  958
            TL NQCQLQGEP QDTDTLYDSA PMESHTLMVN
Sbjct  838  TLCNQCQLQGEPAQDTDTLYDSAEPMESHTLMVN  871
```

Auxiliary factors RISC in *Megacopta cribraria*

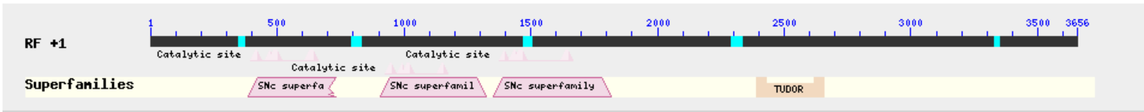
Tudor-SN

cDNA

Contig 2299, Length-3663
TACAAATTACTGTATTACAGTAAAATGAATTTATTAATGATGTATATAATTTATTAGAAAGGCTTGATTAAAAGAGCTAAGCGATGTGACCATAAAA
TTCCTCTTTCACATTTCAGTTTAGTTTCTAGCCACATCATCACCACCTTTCGCCCCTCAGCACGTTTGAAGTCAGTTCGATGAGGTGTTCGTGAATA
TGAAAACGTGTTGTATCCACTTCGTTACTGAAGCTGAAGCATCCCATTTCTAGGAAGTACTTTCTTCTGTGTTTTTTAATTCATTGATCAGGATATA
TTCGTCTGACGTTTTATTAAATATTTCCTATTAAAAAGTACGATGAGTGCTGCCCCACAACCTCAAAATGAACAGCCACCCCCAACTTTGAACAGGG
GGGTTGTTAAACAAGTACTATCCGGAGATACTGTGGTTATTAGAGGTGTGCCTAAAGGGGGCCCCCCACCAGAGAAGACCTTAAGTTTAACTTCCCT
AATAGCCCCAAAACTTGCCAAAAGACCTCAAGCTCCTGGCTCAAATTCTGATGAAGATGAACCTTATGCATGGGAAGCTCGTGAGTTCTTAAGAAAG
AAACTCATCGGTCAAACTGTGCTTTTTACTACTCCGAAAGAGAAAGCAGCCAATCAAACTCGTGATTATGGAACTTTGTTCATCGACCTCCCAAATG
GTGTTAGAGAAAACATCAATGAATTACTGGTCAAAGAAGGATTGGTCAGTGTGAGGACTGTTCCTGGAAGCAAAGATCCTGTCTTACAACATTTGAC
TGAGTTACAGGAACAAGCGAAAGCTGCAAAAAAAGGGAAATGGGCGGATGATGCCAGCAGCCATGTGAGACAAATCAGATGGTCTGTCGAAGCTGAT
AAAATGATGAATTTTGTTGATAAAGCTGAAGGGAAGCCAATAAAGGCAGTCATTGAGCACGTTCGTGATGGATCAACCATTAGAGCTTTCTTATTAC
CTGACTTCACCTACATTACGTTGATGATGTCTGGAATCAGATGCCCTGGTTTTAAACTTGACAGTGAAGGTAAACCTGATCTTACAGTCAAAGTTGA
ATTTGCCGATGAAGCAAAATATTTCACTGAAGCAAGACTTCTTCACCAAGAAGTTGAAATTGTGTTGGAGTCTGTCAATAATAACAATTTTGTTGGT
ACTATCCTGCATCCTAAAGGTAATATTGCTGAAGAGTTATTAAAAGTAGGCCATGCTCATTGTGTTGATTGGTCCATTGCAAAAATGAAGAAAGCTG
ATGCTGACAGATTGCGACAAGCTGAAAAAATGGCTAAAGAAAAGAGATTAAGGATCTGGAAGGACTGGACTCCTCCACCTCAGATAAGTGATCGTGA
ATTTACTGGTACTGTTGTGGAAGTTGTAAATGGAGATGCCTTAATGGTGAAACTTCCTAATGGTACTGTGAAGAAAATATTTTTAGCCAGCATCAGG
CCCCCAAGAGATACGAGTGCAGCTGCCGAAGAGAGAGCTGCTTCTGAAAGGCCAAGAACTGGAAGAGTGCGACCATTATATGATATTCCTTGGATGT
TTGAAGCGAGAGAATTTCTGAGGAAGAAGCTCATTGGAAAGAAAGTCAATGTTACTGTTGACTATATCCAACCTGCGAGGGAGACCCTTCCAGAAAA
AATATGTTGTACCGTCACTATTTCAAATGTTAATGTAGCCGAAGCCATGGTGAGCAAAGGTCTGGCTACAGTTGTGCGATATAGACAGAATGATGAT
CAGAGAGCTTCAGCCTATGATGCACTTCTTGCTGCTGAAAACAAGGCAGCCAAAAGCCTGAAAGGTGTACATGCTAAAAAGGATATACCTATTCACA
GAGTTAATGATATCACAGCAGACCCCTCAAAAGCCAAAAACTTACTTCCTCATATAATGAGGTATTCCAGGATAGAAGCTCTGGTAGAATTTGTCGT
TAACGGGTCGAGAATGAGATTGTACGTCCCTAAGGATTCTCATCTCATTACTTTTCTTCTATCAGGAATTGAAGTCCGCTGTGATAAAAATTCTGCT
GTTAACACAGAAGAAGCTGCATTACAGTTCACTCGAGAGAGGTGCATGCAAAGAGAGGTTGAGATCCAAGTTGAAGGTATTGACAGAAATGGGAACT
TCATAGGTTGGTTGTGGGTTGATAATGTTAATTTATCCGTTGCTCTGGTGGAAGCTGGGCTTGCAAGGATACACTCCTCTGGTGAAAACTCACGATA
CTCCAAGGATTTGTATAATGCAAAGAATACTGCAAAAGCTAATGGAGTTTGGAAATATGAAGAAAAGGAGGAAAAACCAGATACAATAGAAGAAGAT
AAAACGATTGATAGAAAAACAGATTATAAAGAAATTTTGGTTGTTGAAGTTTCTGAAGTAGGATTTTATTGTCAGTATGATTCGAGTAAAAGCAAGT
TGAACTCACTTCAAGCCAAAATGGCTCAAGAAATGGGCGCTAATCCTCCCCTTCCTGGTGCATATACTCCTAAGAAAGGTGATGTTTGTGCGGCCAA
GTACCCTTTGGATAACATGTGGTATCGAGCCAAAGTGGAGAAATTAGCGGGTTCTGGAAAAGTGAGCGTATTTTACATTGATTATGGTAACAGAGCT
GAAGTTCAAAATGTACAATGTGCTCCTCTGCCCCCTGGTCTTGACATTAATACGGACAAACCTTATGCTCACTACTTCGCACTTGCTTTTGTCTCAG
TTCCCAATGAGCCTGAAGATAAAAAAGCTGCCGAAAATGAATTCAGGGATTGTGTTGTTGACAGAAAATTCTTGATGAATTATGAATATGAACTGAA
TGGTGTTAGACATGCTACCCTTCTTGACCCAGAGATGGACAAAGAGAAAGATAAACAAGACGTTGGTAGGTCCCTCGTTGCTGATGGATTTTTACTT
GTTGCCAAACGGAGGGAAAAGAGATTTGAAAAAATTATGAGCGATTACTTGGCAGCGGAAGATGAAGCTAAGAAAAAACATTTAGGTATATGGCAAT
ATGGTGACATAAGAGAAGATGATGATAAAGAATTCGGAATGGGCCGCTAATCATCTTTTAACAATCAATAGGAAGCAAAAAGTACGAGATTTCAACA
GAACACATTCCCAAACCAAACAATGAAATGTCCTCTTTCATTATAATTTCTTGTGGGAAACCGTGTTAGAACTGTCAACATCATATTTCTTCAAGG
AAGAAAATAAGAAGGGATTTTGTCCAGTGGTATCCGTCTAAAAGCGTAATCTATTCCTATAAATAAATGTAACTAAATTGCCTTTCAGTCTTGACTG
ATATACTGTATATCTTTTAAAAGACACAGACAGTTATAATTGCTTAGGTAATTAATAGTAACCAGTTTTATGATATTTACATAAATATTTATATGTA
TAGTTATTTATACAAGAACCAAAAATGACTTCCAATCGTAATTCTAAGCATTGCTGGGTAAATTCATAGAGACTGATATATTATTTAATTGTTATAA
ATTCTTTATTTTCATTTTGATATAAACATAAGCATACATACATTATGTAGTGCTCATCGAATCTAATTATTGCAATGTAAAATTTGTTATTCGGTTA
AATTATATTACGGAGTGTGGTTGTGTTGATAGAATAAATATTGTTGGAAATGATAAAATTTATTATTA

PROTEIN 907 (AA)
MSAAPQPQNEQPPPTLNRGVVKQVLSGDTVVIRGVPKGGPPPEKTLSLTSLIAPKLAKRPQAPGSNSDEDEPYAWEAREFLRKKLIGQTVLFTTPKE
KAANQTRDYGTLFIDLPNGVRENINELLVKEGLVSVRTVPGSKDPVLQHLTELQEQAKAAKKGKWADDASSHVRQIRWSVEADKMMNFVDKAEGKPI
KAVIEHVRDGSTIRAFLLPDFTYITLMMSGIRCPGFKLDSEGKPDLTVKVEFADEAKYFTEARLLHQEVEIVLESVNNNNFVGTILHPKGNIAEELL
KVGHAHCVDWSIAKMKKADADRLRQAEKMAKEKRLRIWKDWTPPPQISDREFTGTVVEVVNGDALMVKLPNGTVKKIFLASIRPPRDTSAAAEERAA
SERPRTGRVRPLYDIPWMFEAREFLRKKLIGKKVNVTVDYIQPARETLPEKICCTVTISNVNVAEAMVSKGLATVVRYRQNDDQRASAYDALLAAEN
KAAKSLKGVHAKKDIPIHRVNDITADPSKAKNLLPHIMRYSRIEALVEFVVNGSRMRLYVPKDSHLITFLLSGIEVRCDKNSAVNTEEAALQFTRER
CMQREVEIQVEGIDRNGNFIGWLWVDNVNLSVALVEAGLARIHSSGENSRYSKDLYNAKNTAKANGVWKYEEKEEKPDTIEEDKTIDRKTDYKEILV
VEVSEVGFYCQYDSSKSKLNSLQAKMAQEMGANPPLPGAYTPKKGDVCAAKYPLDNMWYRAKVEKLAGSGKVSVFYIDYGNRAEVQNVQCAPLPPGL
DINTDKPYAHYFALAFVSVPNEPEDKKAAENEFRDCVVDRKFLMNYEYELNGVRHATLLDPEMDKEKDKQDVGRSLVADGFLLVAKRREKRFEKIMS
DYLAAEDEAKKKHLGIWQYGDIREDDDKEFGMGR

Graphical representation



Comparison with *Cylas puncticollis Tudor-SN*

```
Query  27   PQIFKRQFSVVFVFRQRNVRHDVFGGERLIGRIGGG  62
            PQI  R+F+   V        +V  G+ L+ ++  G
Sbjct  336  PQISDREFTGTVV--------EVVNGDALMVKLPNG  363
```

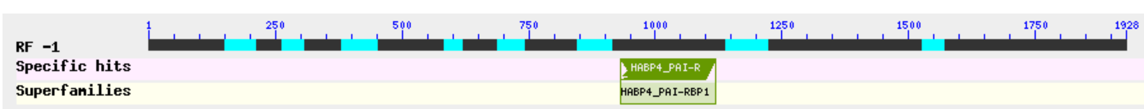
Vasa Intronic (VIG)

cDNA

Contig 2606, Length-2579
CAATTATCATACTAAAACAAACAAAACAAGAATATAATAGTTTATGATCATAATAATGTTAATCAATAACTGTGAATACAACCATCTCAGAGATTAT
TGAAAGAGATAATTTATCAATTATAAAGAAAACTGTGTCCTGAAGTTAATAGATAGTATTAAAATGAAATAAAAAAAGTAATATCAAAATGAAATGG
AATGTCAAAATAAAAAACTCCATTACAAAAAAGTTGATAATTCATTAAAAATATGGATTTAAACACTTCGATTACTATTTGCTAAATAATAAGATTA
AAGAAGAAAATAAAATTCCTTACATTCTACTTCTTTTTAAACAAAGTAAGTTTTATTGCCAGAACTCAATAAATTACACTGAATAAAACATAATGAA
TATTCTAATCCGAAATGGTTTAAATTAAAAAATCAAAAGTAAAATCATTTAATGAAATAAAAAGCTATTTACTCAAGCAAGAGAAGGGAAATCCCTT
TCATCATCAACCTTTGGAGCAGGGATCAAAGTGCGGTTATTAGCTGATGCGCTCTGACGTGGTGAAAACTGCCCTCGCCCAGATTGATTGGGAGGTC
CCCATCCGCGCCCGAAGCCACGGCCTCTTCCTCCGCCCAAACTTCTCGAATCTTTATAATGTATATCGATCTCTAACATTTTTTGCTTACCAACTCT
CTGTGGATATTCACCCAATTCATATTCTTCATAATCATCATCATCAGTTTCCTTCTTTTTCTCAAGTGCGTACATTTTTTTCCATTGAGAAAGGTCC
TCCCCCTCTCCAGCTTTGCGTATATTGTAAGTAGGAGTTTTTCGAGGGGTTTTCAATGCCTTCCACTCATCTAAAGTCAATTCTCTTCTCTCTTCAT
CTTCTTCAACATTAGTCTCCACATTATCATCTTCTTCTTCATTCAAAATTTCAGGGCATTCAACTGGCGGTTCATCTTCGACTTTTTCAGAAGACCA
AGGCTCTTCTGCAGCGAATGAAGTTTGAAGATTATCAATATCATCACGAGGGTTGCCCCAATTATGAGCACCAGATCCTTCACGTTTGTCAATTGAT
TTTATTCCTGTTTTATCTGAGCCTGACTGCCTATCGTATTCTCTTTTCCCTCTAGTGTCGAAAGTAGACCTCGGAAAACCACCCCTGTGACCACCAC
GGCCACCACCCCGGACAAACACACCTCTTGGCAAACCCCTCCTATTATTTCTTCCTCTATCAGCTGAATCCCCTTCAAAGGCAGCACCTGTGTTAAT
GAGATCGCTCTTATCACCTTCTTCAAAATAATCACCAAATTGCTGGGATTCAGGTACATCTGCTGGATAGTGTTTGTTGAAATTTTCAACAATAGTT
ACATGTTTCTTAGTGAAATCAGCATCTCTTGGTTTGTTCTTAATAATTACATCATCTTTATGACTTATGACTTCATTATGTTTGGCTACGGAATCTT
TTCCAGCTTTTGAAGACTTTAATTTTGCAGAATCTTTAGTTTTAGTAACTTTATCCAAAAGTTTGGACTTCTTTCTAGCTTCTTTTTCCAATTCATG
ACGCTTCAACAATTCAAGAGGGTCGATTTCTTCTTCTATACCACTTAAAGCAAATCTGTTCTTAACAATGCCAATCCCATACTCCGTAGCTGACATC
GTTGCAGCCAGTTTTGTCCCAATCTTTACTGGATTAAGTTACAGGATAACAATCAGTTATTCAAAAGAGTAACTAATCACAAAACTCAGACAGTTCT
TGAGTCTTGGGTTTTCCTGAGACAAACATTGATCCAATCATTGTGAAAATCACTGGCATTAAAGAAGAGAAAAGCGCAGAAGATAAGCATAAGGTCG
GCCAACTCTCAGAACTGCATGAAGACGCCGGTCGCAGCCTCTTAGCAGTAAACAATTGTTACTATAATCTGTTGAAGAGCTAAA

PROTEIN 395 (AA)
MSATEYGIGIVKNRFALSGIEEEIDPLELLKRHELEKEARKKSKLLDKVTKTKDSAKLKSSKAGKDSVAKHNEVISHKDDVIIKNKPRDADFTKKHV
TIVENFNKHYPADVPESQQFGDYFEEGDKSDLINTGAAFEGDSADRGRNNRRGLPRGVFVRGGGRGGHRGGFPRSTFDTRGKREYDRQSGSDKTGIK
SIDKREGSGAHNWGNPRDDIDNLQTSFAAEEPWSSEKVEDEPPVECPEILNEEEDDNVETNVEEDEERRELTLDEWKALKTPRKTPTYNIRKAGEGE
DLSQWKKMYALEKKKETDDDDYEEYELGEYPQRVGKQKMLEIDIHYKDSRSLGGGRGRGFGRGWGPPNQSGRGQFSPRQSASANNRTLIPAPKVDDE
RDFPSLA

Graphical representation



Comparison with *Cylas puncticollis Vasa Intronic (VIG)*

```
Query  5    YGIGVA-NRYALF-LDDESDPLETLSIKEQEKELKKKAKVAEKENKGKNEQPKAKPAANV  62
            YGIG+  NR+AL  +++E DPLE L   E EKE +KK+K+ +K  K K+   K K +
Sbjct  6    YGIGIVKNRFALSGIEEEIDPLELLKRHELEKEARKKSKLLDKVTKTKD-SAKLKSSKAG  64

Query  63   QKKPIKETTINKAQENKREDNKPNQRAAVDGKPERTFTKFNNENREERNNRRNREDRPYN  122
            +    K   +   +++    NKP             +  FTK +    E  N        P +
Sbjct  65   KDSVAKHNEVISHKDDVIIKNKPR---------DADFTKKHVTIVENFNKHY-----PAD  110

Query  123  GPAENRDRDRE---NRPRRENNSENFENRNRDRGERGERR------EGRAIGNRPGGPRG  173
             P   +  D     ++    N    FE  + DRG R  RR         R  G
Sbjct  111  VPESQQFGDYFEEGDKSDLINTGAAFEGDSADRG-RNNRRGLPRGVFVRGGGRGGHRGGF  169

Query  174  PRKTFDDRRGKREFDRQSGSDKTGVKPIEKRDGAGAHNWGSHKDIIEAETERPSEADQSW  233
            PR TFD  RGKRE+DRQSGSDKTG+K I+KR+G+GAHNWG+ +D I+   +    A++ W
Sbjct  170  PRSTFD-TRGKREYDRQSGSDKTGIKSIDKREGSGAHNWGNPRDDID-NLQTSFAAEEPW  227

Query  234  GETEKIE---PIE---VEKKEEQPEVEA-VPAEEEPKELTLDEWKAQRAGRAKPQYNIRK  286
              +EK+E   P+E   +  +EE   VE  V  +EE +ELTLDEWKA +  R  P YNIRK
Sbjct  228  S-SEKVEDEPPVECPEILNEEEDDNVETNVEEDEERRELTLDEWKALKTPRKTPTYNIRK  286

Query  287  AGEGEDPSQWKKMFELKKKEKEEESEEEEFDASEYPQRAGRQKLCHSIDIQFNDNRRPGG  346
            AGEGED SQWKKM+ L+KK++ ++ + EE++  EYPQR G+QK+   IDI + D+
Sbjct  287  AGEGEDLSQWKKMYALEKKKETDDDDYEEYELGEYPQRVGKQKMLE-IDIHYKDS-----  340

Query  347  GRGRGQRSGVRGGRGGSQRGGGPPRDGGERPRFREEQ---GDEKGTPRAPKVDDERDFPS  403
                  RS   G   G  RG GPP   G R +F   Q    + +    APKVDDERDFPS
Sbjct  341  ------RSLGGGRGRGFGRGWGPPNQSG-RGQFSPRQSASANNRTLIPAPKVDDERDFPS  393

Query  404  LV  405
            L
Sbjct  394  LA  395
```

FXMR1

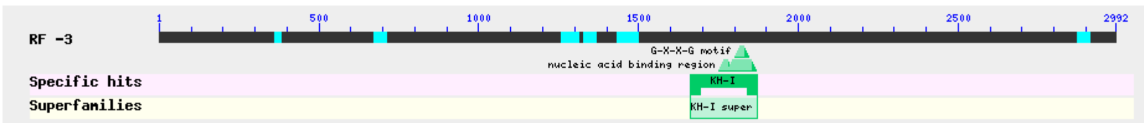
cDNA

Contig 1649, Length-4242
ATCAAATATAACATTGTTACTGTAAGATGTCAGTGACCTCGAAAAGGACATAAGTCATGCCAGTTGTCCTCTGCTTTTATTGGACAACTAGAAACTA
TAAGATTTCACAGTCGAATTTTCATACATAAATACTTATTAAACTCTTTATTTTCAATACCACAGAGGTAAAAAATTTAATTAGTATTTGTATGTTC
CTAAACCCATAATTAACCTTACCAAACACATCTCAGTCAAAGCAAAAACATCTTGAATGTATTTTTAATCCATAATTCTCAATTAATCATGTAATTT
ATGAAAAACTTACAAAGCATGTTGATTAAAAAAAAAACTTCCAGAATAGCACATTATTTTTCATAACATAAAAAAAACAAAAAAAAATTAAATTTTC
TTTCCTCAAAACATCATTAGTAGCTCAAAAACAAATTAAATTTGCATATATCTGGTAAGTACAGTACTAAGCCCCTCCAAATGAATTACTTGATATA
AAGAAAAATAAAATTACAAAAAGCAAAGATGCTCTGGATTTAAAATCCTTTATTTGTAGATAGCATTATGTGAATATGGTGTTAGGGGTAATAATAA
GCCAATTAAACACATAGAAGAACACAAACACAAGAATTGTAGAATAAGACAGAAACAGAAAAGATATCTGTTTGTATAATCTTTCTCAAAAAATTTT
CAGAATCAGGGAGAAAAAAAAAAGAAAGTCAATGTTTATCCACAACGTATTATAAATAAATACGAGGAGGGGTTTCTAGTCAAACATAAAGTTCAAC
TGAGATATTAGATACAATTAAGAGATTTTTGTTTCCAACAGTTAAGAGTGAAAAGAAATATAGAAAAAAAAGGGAAATGTGTAAAAGAGAAGTTTGCC
AGGATGGCGAATCCTAGTTCCTTTGCCCACTGCCTCCAGCCAAAGGCACAGCACAAAACACGGTAGCAAGTTATTCGGACGAGGGGAGAAGTGGAAA
TATCGGTTTGGTGTCCTTATTTGGCAAAAGGCAGTGAGCCGCGGTGACTAGATCCTCCCCCTCCTTTTCTCCTCTTTAGGTTGAAGTTCCGTTGTAC
AACGGCTCCTTCACCTTCTCGGTGGTCCTTTTCTGCCTATTATTTTTACGCTGACCTTTCGGCCCGTCATTGCTATTCACAGCTTCTTTGGAATTCT
CGACAGGTGCAATGTTGATGGCAGCAAGACTATGATGGCTACCTGAGTCCCCAGTTTCGAAGTCTTGTGTAGGTGGCTGGTCTTCATCGTGCCTAAC
TCGAGGATTGTTGGCCCTTGGAGGACCTCCCCTGTCTCCCCGATCCCTCCTTTGCATGTAATTACCACCTCCACCTCGGCCTCCTCCACTGTAACCT
CGCCCACCTTCTTTTACCCTCTCGTCCGTGTCAGGTGTTTGATGGCGGGATCCTTGGTACCTGGAGTCATGTCGGCCTGGACCTCCTCGACCCCTTC
CTCCCCGCCCCCTTGGAGGTCCACCACGACCTCGTCCTCCTCCAACCTCCATGTCAGAGTTGTACCCCCTTTCATTGCGCCTTTGAATGGGAAATCC
CATCATAGAACCCATGTTAGTCCCTTGCATGTTCCTCACCTGACTCTCTATGGCTTCTTTGTTCTGCCTCAGTTTGTCAACTTCCTTGAGGTGAGCC
AGGTGGTACTCCAATAGGATCTTGGCATTGGTGATTGACTCAATAGTGCCAACAAACATGAAAGGAACAAGCCCTTCTTCTCTAGGTACAGATGGGG
ATGGTTCATTGTCGCCCTCAATTTTAACCCGCACGACTCCACTTTTATCAACTATTTCTTGGATTATGCGCCCATTTTTTCCTATAACTTTCCCAAC
AAGGCTTCTCTCCACCTGCATGGACTCCTCACTGTACTCAAGCATAGTCCTTGCTTTTTTCACAGCTTCACTACTTTCACCATAAATCTTAAAGGTA
CATCCTTCTTCTAATTCAATATTAGTGATCCCATCCAATTTCCTTGCTTGTTGAATGTTCACACCATGAGCTCCAATCGCAAGACCCATAAGGTCCT
CTTTGACTGAAAACTCATCAGAGAACCCTCCAGTGTTTTGAAGCTTGGTGGACTCCAACTGTTTCTGTGCCTCTTCTGTCCTCTTAAGCAGAAGGAG
TTTTTGATGTAAATTCCTAAAATGCATTTCCTGAAGCATTGCTGCTCTTTTCTTTGAACCTTCACACCTTGAGATGAAGCTTAATACTCCACTCTCG
GTATGTATCGGCAGATTGAGGCAGTTATGGCTTTTTGGAAATCTTTATGCACATTTTCCTTTTTGGCACTGCAGTCGTAAAGGTATTCCCTCAGTTC
TTCAGGAACTTCGATTTCAAATTTGGTAAATGTTTGTGCAGTGATAGGGGGATTACTATTTTTGTGCCTTAACCTGTCCGATGGAACAATTTCCGTA
TATGTAGTTTCCCATCCAAGGTACTCCAAAACCAAGAAGTCACCTTTCATCATCTTAATGACAGCCCGCCACCAACCACAGGACTCGTGGTCATTGG
AACGGGAAAGGACCTCAACCTCTTCATTTTCTTTAAACTCTGTTTTGGGGGCATCCTTTGGAGGTGGTAATCGCACAAGGTTAAAGGTAAACCTGGA

TTCGGCCTGCCATTCATTTTCAAAGGCAACTGATACTTCTTTTTCAAATACATCAGTCACATAAGCCTTATAATAAGCACCATTTTCGCCGCATACT
TCCACCTCTAAGTCATCCATCTTGCTATAGCTCTATTGCTATCAACGGGACCCGCACAAGTTACTACAACCAATAGGATCTTCTCATTTTAGCCATT
GGCAGCACTGTAGCAGTAACCGATCTAGCAGGCTGTGTTAATTACTTCTTTCTTCGTAATATTTTATCTTTTTTAATTATTATGCCTTAAAATACTT
AGATACCAGCCTCCTCTTCTAAAATGTCCACTGGTAGCGATTCTGACGAATACTATGATGCTGAAGATCACACTCCTAAC

PROTEIN 563 (AA)
MDDLEVEVCGENGAYYKAYVTDVFEKEVSVAFENEWQAESRFTFNLVRLPPPKDAPKTEFKENEEVEVLSRSNDHESCGWWRAVIKMMKGDFLVLEY
LGWETTYTEIVPSDRLRHKNSNPPITAQTFTKFEIEVPEELREYLYDCSAKKENVHKDFQKAITASICRYIPESGVLSFISRCEGSKKRAAMLQEMH
FRNLHQKLLLLKRTEEAQKQLESTKLQNTGGFSDEFSVKEDLMGLAIGAHGVNIQQARKLDGITNIELEEGCTFKIYGESSEAVKKARTMLEYSEES
MQVERSLVGKVIGKNGRIIQEIVDKSGVVRVKIEGDNEPSPSVPREEGLVPFMFVGTIESITNAKILLEYHLAHLKEVDKLRQNKEAIESQVRNMQG
TNMGSMMGFPIQRRNERGYNSDMEVGGGRGRGGPPRGRGGRGRGGPGRHDSRYQGSRHQTPDTDERVKEGGRGYSGGGRGGGGNYMQRRDRGDRGGP
PRANNPRVRHDEDQPPTQDFETGDSGSHHSLAAINIAPVENSKEAVNSNDGPKGQRKNNRQKRTTEKVKEPLYNGTST

Graphical representation



Comparison with *Cylas puncticollis FXMR1*

```
Query  27    MEDLAVEVVGENGALYKGSILDVHDDGVLVHFEDEWQPDSKFPFEQVRLPPNPDQ-KVEF  85
             M+DL VEV GENGA YK  + DV +  V V FE+EWQ +S+F F  VRLPP  D  K EF
Sbjct  1     MDDLEVEVCGENGAYYKAYVTDVFEKEVSVAFENEWQAESRFTFNLVRLPPPKDAPKTEF  60

Query  86    TENMEVEVFSRSNNQEGYGWWKARIKMMKGDFYVLEYLGWDSTYTEIVSIDRLRMKNSNP  145
              EN EVEV SRSN+ E  GWW+A IKMMKGDF VLEYLGW++TYTEIV  DRLR KNSNP
Sbjct  61    KENEEVEVLSRSNDHESCGWWRAVIKMMKGDFLVLEYLGWETTYTEIVPSDRLRHKNSNP  120

Query  146   PIERGMFHKFEIEVPEDVRDY---C---KKDTYKEFQKAIGAGRIFYNPDKGVLVVISRC  199
             PI    F KFEIEVPE++R+Y   C   K++ +K+FQKAI A    Y P+ GVL  ISRC
Sbjct  121   PITAQTFTKFEIEVPEELREYLYDCSAKKENVHKDFQKAITASICRYIPESGVLSFISRC  180

Query  200   ESSRKRASLLQDMHFRSLQQKVLLMKRTEEAARQLESTKLATIGGFSDEFHVKEDLMGLA  259
             E S+KRA++LQ+MHFR+L QK+LL+KRTEEA +QLESTKL   GGFSDEF VKEDLMGLA
Sbjct  181   EGSKKRAAMLQEMHFRNLHQKLLLLKRTEEAQKQLESTKLQNTGGFSDEFSVKEDLMGLA  240

Query  260   IGAHGVNIQQARKVEGITNIELEENSCTFKIYGETDEAVKKARSMLEYSEESLQVPRALV  319
             IGAHGVNIQQARK++GITNIELEE  CTFKIYGE+ EAVKKAR+MLEYSEES+QV R+LV
Sbjct  241   IGAHGVNIQQARKLDGITNIELEE-GCTFKIYGESSEAVKKARTMLEYSEESMQVERSLV  299

Query  320   GKVIGKNGRIIQEIVDKSGVVRVKIEGDNEPQPTFPREEGQVPFVFVGTVESISNAKVLL  379
             GKVIGKNGRIIQEIVDKSGVVRVKIEGDNEP P+ PREEG VPF+FVGT+ESI+NAK+LL
Sbjct  300   GKVIGKNGRIIQEIVDKSGVVRVKIEGDNEPSPSVPREEGLVPFMFVGTIESITNAKILL  359

Query  380   QYHLAHLKEVEQLRQEKLEIDQQLRSIHHGSNLGSMQSLSMSGRRNDRGYNSDMDGRSSG  439
             +YHLAHLKEV++LRQ K  I+ Q+R++  G+N+GSM    +  RRN+RGYNSDM+
Sbjct  360   EYHLAHLKEVDKLRQNKEAIESQVRNM-QGTNMGSMMGFPIQ-RRNERGYNSDMEVGGGR  417

Query  440   RGGSMRGRGGGRGRGGPRQNDRYNSGSRHQTPDTADERVQDGGGPRRYGGYRGGRGGGGG  499
               G      GGRGRGGP ++D    GSRHQTPDT DERV++GG     GG  GG
Sbjct  418   GRGGPPRGRGGRGRGGPGRHDSRYQGSRHQTPDT-DERVKEGGRGYSGGGRGGGGNYMQR  476

Query  500   RGRGDARRGGEDRRRN----NDDEETVLDSQEVSSADRVNNHSDQA---PATDNGVDNQP  552
             R RGD  RGG  R  N    +D+++    +Q+  + D  ++HS  A      +N  +
Sbjct  477   RDRGD--RGGPPRANNPRVRHDEDQP--PTQDFETGDSGSHHSLAAINIAPVENSKEAVN  532

Query  553   RGGGPKNRYNRGPPPRRGGGVDKHQQHQQQQQQKEALVNGTSA  595
                GPK            G    ++Q +  ++ KE L NGTS
Sbjct  533   SNDGPK------------GQRKNNRQKRTTEKVKEPLYNGTST  563
```

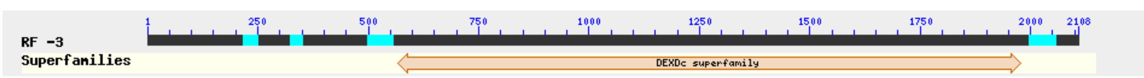
p68 RNA helicase

cDNA

Contig 211, Length-1773
TTTTTTTTTATTCTAAAAATTTGTATTCAAAAATCTTTCTCTAAAATTTAAAGTTAACAGAACATCTCTCCATTAAGCATTACTTAATACTGTGCAA
CAGAAATATGAAAATCATAGGAGGTAAGTTAAAAAAATGCCCTCCAATAGAAAAAATAAAATAAAAGAGCCTCATTTATATAAACCTATTATCACAA
AGAAATAACGAATTAAGTTCACAATAAAAAAAGTTTAATAACTAACATTAATTATAAAATAGATAAATTATGTTCTTCTTATTTACATTTTATTAGC
AATATATATAGAAGAAAAAAATTAACACTAAAAAAAAAAAAAGGGTGGGGGAGGGGGAGGATCACATAGATTGAAGAGTCAAATAGGAATTATTTGTA
CATAGGCTCAACAATTTTTTTTTTGTATACAATCATAATTAAGGAATAAAAGTTCTTAACTTTATAGGAATTAATCAAAATCAAATCTTTTATAATT
TGTAGACTTTTGACCATAACCACCATTTCTTTGGAAATTATTCCCTCCATTGAACCGGCTGCCACCACCTCCAAATTTATCATTTCTTTTGAATGTT
CCATTGCGGAAATTCCTAGAGTATCCATTTCCTCCACCAGAGAATCTGGTCATAGCCTGCAGCTCTGGAGAAATGCTCTGTTTGGCTTCTGCTAATA
CAGCAACAAGATCCCTAGCTTGCTTTTGATTGGCACTTGTAAAAAAGGTATAAGCTGTTCCTGATGCAGTGGACCTTCCAGTCCTACCAATCCTATG
GATGTAATCTTCAGACGAATTGGGAAAGTCATAATTGATGACATACATCACATCCTCAACATCTAGACCTCGGGCGGCAACATCTGTTGCAACGAGG
ATCAGAGATTTTCCTCGTCTAAAATCTCTTAAAACATAATCTCTTTCATTTTGGGATTTGTCACCATGAATACAAACAGCAGGCCAACCAACTTTTT
TGAGAGCATTGGCAATAGCTTCAACCTTCTTTTTAGTTTCCACAAAAATAATAGTTCTTTCCAGTTTCTGATTACCAATTTCTTCTAGCAAGTTGTT
AAGTTTGGATTGTTTATCACCTTCTTGGCATACCTCCACGATTTGTTGAATATTATGATTTGCTGCTAGTTCTAGTGATCCTATATTAACTTGAACA
TAATCTTGCAAAAAGTCTTCAGCAAGGGCTTTGACTTCTTTGGGCCAAGTTGCAGACCACATCAGCACCTGACGGTCAGGCCTGATTTGCTGCATTA
TTTTTCTAATTTGTGGCTCAAAACCCATGTCTAACATCCTGTCGGCCTCATCAAGTACCAAATAAGTACATCTTCTCAGGTTAGTAGTGCCTCTAGC
CAAGAAATCAATCAAGCGCCCAGGAGTAGCAATAACTATTTCCACTCCTCTTTCCAGTTCTCTAGCTTGGGGTCCTTTTGGAGCACCACCAAAAATA
GCAACCGATTTTTGACCACATTGACCACCAAAATGTGATGCAACTGTCTGAATTTGCTGAGCTAACTCACGAGTGGGTGCAAGAACGAGGGTTATAG
GTCCATCTCCACGCTGCACTCTAGGCTGATTTCTGATGTGGATTACAGCAGGAAGCATGTAAGCTAAAGTTTTCCCACTGCCAGTTTGAGCAATTCC
AACCATATTTTTTCCACTCAAAGCTATTGGCCAACCTTGAGCTTGAATAGAAGTGGGAAAGTCAAATCCCATCTTCTTAATGTTGGATGCAACATAA
TCAGGGAAGTTGCCTTCTTCAAAAAACTGAATTGGATTAAAGATGGCATCTGAAGCCCCTTTAACTGTAATATCCTTACTTTCTCTGTAACGCTTAA
CTTCAGTGTATGGTCTGTTTAATACATTAGGATGTGGACAATAGAAGTCTTTCTTGAAATCTGACAGTTCATGATTCTCCCAAGCAACTGCTTGCAG
CCCACCCCCGAAGTTACCTCTGTCATCATTGTTCACTCTCCCTCGGAAACCTGGCTTATCATTCTTATAACCATTTCCTCCTCCAAATCCTCCTTTT
CTGTTTGAATTATTGAATTTCATTTTGGACGAATTAATTGCAGTAGTACAATTCCTAAGGAGAGATGTTC

PROTEIN 533 (AA)
MKFNNSNRKGGFGGGNGYKNDKPGFRGRVNNDDRGNFGGGLQAVAWENHELSDFKKDFYCPHPNVLNRPYTEVKRYRESKDITVKGASDAIFNPIQF
FEEGNFPDYVASNIKKMGFDFPTSIQAQGWPIALSGKNMVGIAQTGSGKTLAYMLPAVIHIRNQPRVQRGDGPITLVLAPTRELAQQIQTVASHFGG
QCGQKSVAIFGGAPKGPQARELERGVEIVIATPGRLIDFLARGTTNLRRCTYLVLDEADRMLDMGFEPQIRKIMQQIRPDRQVLMWSATWPKEVKAL
AEDFLQDYVQVNIGSLELAANHNIQQIVEVCQEGDKQSKLNNLLEEIGNQKLERTIIFVETKKKVEAIANALKKVGWPAVCIHGDKSQNERDYVLRD
FRRGKSLILVATDVAARGLDVEDVMYVINYDFPNSSEDYIHRIGRTGRSTASGTAYTFFTSANQKQARDLVAVLAEAKQSISPELQAMTRFSGGGNG
YSRNFRNGTFKRNDKFGGGGSRFNGGNNFQRNGGYGQKSTNYKRFDFD

Graphical representation



Comparison with *Cylas puncticollis p68 RNA helicase*

```
Query  28   KSGYEGGSYRNSRAGFNNRLRTVEWSSKQLRPFKKDFYVPHPSICSRSTFEVDQFRSAKE  87
            K G+ G    + R  F   L+ V W + +L  FKKDFY PHP++ +R   EV ++R +K+
Sbjct  22   KPGFRGRVNNDDRGNFGGGLQAVAWENHELSDFKKDFYCPHPNVLNRPYTEVKRYRESKD  81

Query  88   ITVEGEAP---KPIQSFHEANFPDYVMDEIVSQGYEFPTAIQAQGWPIAMSGNDMVGVAH  144
            ITV+G +     PIQ F E NFPDYV   I   G++FPT+IQAQGWPIA+SG +MVG+A
Sbjct  82   ITVKGASDAIFNPIQFFEEGNFPDYVASNIKKMGFDFPTSIQAQGWPIALSGKNMVGIAQ  141

Query  145  TGSGKTLAYILPAIVHINNQPEVQRGDGPIVLVLAPTRELAQQIQQVANDFGRRSKIRNS  204
            TGSGKTLAY+LPA++HI NQP VQRGDGPI LVLAPTRELAQQIQ VA+ FG +   ++
Sbjct  142  TGSGKTLAYMLPAVIHIRNQPRVQRGDGPITLVLAPTRELAQQIQTVASHFGGQCGQKSV  201

Query  205  CVFGGAPKGPQARDLERGVEICIATPGRLIDFLEKGTTNLERCTYLVLDEADRMLDMGFE  264
             +FGGAPKGPQAR+LERGVEI IATPGRLIDFL +GTTNL RCTYLVLDEADRMLDMGFE
Sbjct  202  AIFGGAPKGPQARELERGVEIVIATPGRLIDFLARGTTNLRRCTYLVLDEADRMLDMGFE  261

Query  265  PQIRKIIGQIRPDKQTLMWSATWPKEVKKLAQDFMNNPIQLNVGSLQLSANHNILQIVDV  324
            PQIRKI+ QIRPD+Q LMWSATWPKEVK LA+DF+ + +Q+N+GSL+L+ANHNI QIV+V
Sbjct  262  PQIRKIMQQIRPDRQVLMWSATWPKEVKALAEDFLQDYVQVNIGSLELAANHNIQQIVEV  321

Query  325  CQEHEKESKLNNLLQEIGTNGEPDAKIIIFVETKKKVEGITKTIRRLGWPAVCMHGDKSQ  384
            CQE +K+SKLNNLL+EIG   +   + IIFVETKKKVE I    ++++GWPAVC+HGDKSQ
Sbjct  322  CQEGDKQSKLNNLLEEIGN--QKLERTIIFVETKKKVEAIANALKKVGWPAVCIHGDKSQ  379

Query  385  QERDYVLREFRSGKSTILVATDVAARGLDVDGIKYVINYDYPNSSEDYIHRIGRTGRSDT  444
             ERDYVLR+FR GKS ILVATDVAARGLDV+ + YVINYD+PNSSEDYIHRIGRTGRS
Sbjct  380  NERDYVLRDFRRGKSLILVATDVAARGLDVEDVMYVINYDFPNSSEDYIHRIGRTGRSTA  439

Query  445  TGTSYAFFTPSNIRQAKDLVSVLKEANQVVNPKLSDMASKSSGFGGFQKGGRWG  498
            +GT+Y FFT +N +QA+DLV+VL EA Q ++P+L  M   S G  G+ +  R G
Sbjct  440  SGTAYTFFTSANQKQARDLVAVLAEAKQSISPELQAMTRFSGGGNGYSRNFRNG  493
```

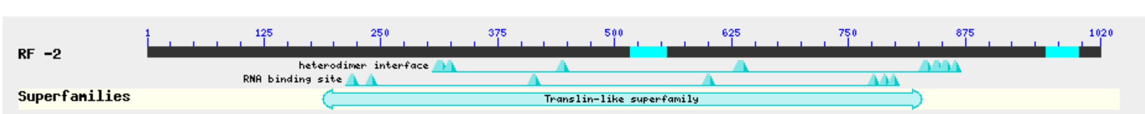
Translin associated factor X

cDNA

Contig 5300, Length-1018
CTTAAATTGATTTATTTAGCTTCTACAAAAATAAAAATATTTATACATAATTATAAGAAAATTATAAATAAATTAACACTTTATTAATAATAGTTTA
AAAAAACAATAAAATATCTTCATTTAACCAGCATAAGTTTCAAACCCTAGAGCATTTTCTGAAGTGAAGCCATAATCCAGGGCTTCTGATCCACGAA
TGTGACTTGCATAACGAACTAATTCCATTTTATTTAAAGACTGAAGGGCTACAACCAATTTTCGTGAGAAATCTCTCCCTTGCATGTTTTGACTCAA
GCTTGTTCCAGACAAATTTAGCAATAGCATTCCAGTGTATAAATTTTTAACAAAATTACAATCTGCATTCCCTGCATTTAAATTACGAGCACTAAAA
CTATTAATACAATGCCTCATTAGTTCACCAGTTAGGTCCTGGATTCCAAGAATATAATCAAAAAGAGGAATGAAGAGTGAAATGGATTCTTTAGATT
GGTCAACCTGATTCATAAAATTTTCATGGAATGTTAATTCTTTTTCTAATAGTTTTAGATCAAATAAATCTCCTTTTGTAATAAAATGAAAAAAAGT
CACAGCTTCAATATATTCCTGCAAACCTGCTGTATAAGCCCTGCAATATTGATAAGGATCATCATCCTTTAGTTCAAATGCTATATTTTTCAAATGA
GTATTTCTAACTGTTTTAAGATTTTCTGTTATTTTGGCCAAGTCTTGCTCATAGTTTGAAGATCTTAATAAAGCATGCAAAGAAAATATAATACGTT
TACTGGCAATAGCAACATCTCTGCTTAATTTTATAATGCGCTCATACTTGTCTTGTTTAAACTCTAATTCTTTGGAATACTTTTCAAATGCTTCAAT
TATTGGTGTTTTTGGATTGTTATTCTTTTTAGTTTGAGATGGTTTATCATATTTTCTGTTATTTTCGTCATAACTTGCCGAGAGCTTTCTCTTCATT
TTGAACACCTTGAATACTCATCGTACATAATTTTCACTCTTATGTAAC

PROTEIN 282 (AA)
MKRKLSASYDENNRKYDKPSQTKKNNNPKTPIIEAFEKYSKELEFKQDKYERIIKLSRDVAIASKRIIFSLHALLRSSNYEQDLAKITENLKTVRNT
HLKNIAFELKDDDPYQYCRAYTAGLQEYIEAVTFFHFITKGDLFDLKLLEKELTFHENFMNQVDQSKESISLFIPLFDYILGIQDLTGELMRHCINS
FSARNLNAGNADCNFVKNLYTGMLLLNLSGTSLSQNMQGRDFSRKLVVALQSLNKMELVRYASHIRGSEALDYGFTSENALGFETYAG

Graphical representation



Comparison with *Cylas puncticollis Translin associated factor X*

```
Query  31   PIIEAFEKYSKELEFKQDKYERIIKLSRDVAIASKRIIFSLHALLRSSNYEQD--LAKIT  88
            P+I+ F++++ EL+ K D+YER++K+SRD+ I +KRIIF LH+      ++D  L +
Sbjct  33   PVIQMFQEFTLELDDKHDRYERVVKISRDITIEAKRIIFLLHSTNTDIESKRDSVLEEAF  92

Query  89   ENLKTVRNTHLKNIAFELKDDDPYQYCRAYTAGLQEYIEAVTFFHFITKGDLFDLKLLEK  148
            E L T+   + KNIA EL+D D Y Y +A+T G+QE+IEA+ F+H+   G++     L K
Sbjct  93   ERLSTLYKNNFKNIAMELRDKDHYLYHKAFTCGMQEFIEALCFYHYTKNGNITSWLELNK  152

Query  149  ELTFHENFMNQVDQSKESISLFIPLFDYILGIQDLTGELMRHCINSFSARNLNAGNADCN  208
                   F    DQ +E ++L    +D+ILGI D TGELMR CIN  S  N+N     C
Sbjct  153  -------FFQYKDQDEEDLNLLFTQYDFILGIADFTGELMRKCINVLSVGNVNECFKLCK  205

Query  209  FVKNLYTGMLLLNLSGTSLSQNMQGRDFSRKLVVALQSLNKMELVRYASHIRGSE  263
            FV+N++TG + L+ +G         ++ S+K  V  QSL KMELV Y   IRGSE
Sbjct  206  FVRNIHTGFVGLSFAG--------NKELSKKAYVLRQSLAKMELVCYNIKIRGSE  252
```

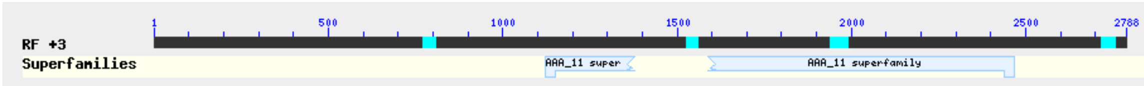
Armitage

cDNA

Contig 5309, Length-2219
AGTATGTATGGGTATTTACTGAATTTAGTATTGGTAGAATATTTGATTTAGAAGTGATTGATCCAACCTTAGCTGAAAATGAACCAGACTTTTCCTC
TTCACAGTCACTTATCCCCAGGAAGGATGTTACAGCCACGCAATTGTTGGACCGGGCATCTGAGTTATCATGCCAGGACGGAGATCTTTTTCCCCTG
CTGCATTTGTACCAGTGAAACTTGGACAGTTTCCATTGTCAGAAAAAATTCTTATAGCTGCTCTACCAGATGTGAAAATGAACCTTGATGACTTGAT
GATTCGTGTTGAGAAGGTGTTGCCATGTTTAAAGGAAGATTTATCCCCACGTACTTATTTGGCTCGTATGCACGGACTACTGTACCTTGAGGAAGTG
GCTATGTTGAAGCAGGTCGGTGAGCTTCAGATGGATAGAGCTTCTTTCAGAAAGGAAAAAGACTGCCTTATATTAGACGTCCCACCACTATCTCCTT
ATACAAGCAAATTAGTCGCTGGTGATATGTTAATAGCTTCATTACCTTGGGCTGCAGATGAAAGAAAATATGAAGGTGTAATCCATCGTGTAACTCA
AACTGAACTTTGGCTCCAATTTGCAAAGGATTTTCATGATTCTTATGGTGAAGGGATACAATATTCAGTTTCATTTGTGACTTCAAGGGCATCATTG
CGTAGAATGCATCAAGCTATTAACCTTGCAGCTAAACATTTAGGATTTGGTTGGCTTTTCCCGACTGGAGTAACATCTAGATTACCTCAAGTCATCG
TTGAAGAAGAGATAGAAGAAAGTGAATCCGTAAGAATTGCAAGGAAAGAAAACTATTCAAAGAAGTTACACTGTGTACCAACTCCTGAAAGTAAGGA
TTTACTATACAGCCCTGATCCCAGGAAAATTTCCGTTGCTGGTCAGGCAAATCTTAGTATTGAACACAAATTTCAGGCTGCTTCTTACAAATCAGGA
TACCATGGTAGAGGCAGAAAATATAACAACAGGGACATTCAATGGAACAGGCTAGGAAGTGGTTATGGGATGCATTGTACTGTTGACAATGAAGTTC
GTGTAAACATTGGCCTTGAAAGAATAAAGAAAATTAGATGGTTTAACAAAGGCCTTAATAGGCAACAAAAAGAGGCTGTAAATAATGTCCTTTTAGG
AGAAGCAAGGCCTTTGCCTTACGTCATATTTGGTCCTCCTGGTACTGGAAAAACTGTTACAGTTGTTGAAACAATTTTACAAATACATGCTCTGATA
CCAGAAAGCAGATTGTTAGTTGCCACGCCTTCCAACTCTGCTGCAGATCTTATTACTGAAAGATTACTGGATGCAGGTGATTTAGAACAAGGAGATT
TATTGAGGATGGTTGGCTATCATTATTTAGAACAAGGAAGAATCGCTGCATCTATTGTTCCATATGCTGCTGTACCAGATGTTAAAGCTATAAATGT
GGCTGGACTCTCAGGTGCATCTCATGAAGGGGTGCAGATGTGTGGACGAGAACTCCTTGGCCAACATCGAGTTACTGTAGGGACTCTAGGTTGCCTT
GGTCTCCTTTACAATATGGGTTTTCCTAGAGGCCATTTCACCCATGTGATAGTAGATGAAGCTGGCCAAGCAACAGAACCTGAGTTGCTTATACCTA
TGGTTTTTTTGCACATGGAATATGGGCAAGTGGTTTTAGCTGGTGATCCTCTACAGCTTGGGCCTGTAGTTACTTCACGTCTTGCTTCAAGATGTGG
ACTTCAGGACTCACTTCTTGCTAGATTTCTTAATCGATTTCCATACTCGAGAGATCCCCATGGTTTTCCAGATAGTTCAGGCTATGATCCACGTTTA
GTCACCAAGCTGGTGAACAACTATCGATCTCTTCCTACCATACTGGAATTACCCAGCATGCTTTTTTATGACAATGATCTAATTCCAAATGTCTCGG
AGGATTCTAGTGAAGAAGCAGCTTTATTAAGAGCTTTGGCTTCATTACTTCCTTGTCGAACCTTTGGGGGATATGCTCCTCCTCTTCTCTTCCATGG
TGTTCGGGGTACAAACTGTCAGGAAACTGAGTCCCATTCTTGGTACAACCCTCAAGAAGTGTTCCAAGCATTTGTTTATTTGAATCTATTATATAAA
GCTGGTCTTCGCCCTGATCAAGTTGGAATAATTACTCCATATCAACTCCAGTCTAACAAAATTAGATTCATGTTGGAACGTATTAATGTGGAGCCCC
CAAAAGTTGGTTCAGTGGAAGAATTTCAAGGGCAAGAAAAAATGGCTATAATTGTTAGTGTTGTGAGGAGTAGCCCTGATTTGATCAGTTATGATAT
GCAAAGAGCACTTGGATTTGTTGCTAATGCTCGGAGGTTAAATGTTGCATTATCACGTGCAAGAGGTATCTTAATTATTTTAGGTAATCCTCATTTA
CTTTACCTTGACATGCACTGGAGAAGTGTCCTCAGGCATTGTGTTAAGAAGAAGTTTTATACTGGCTGTGATCTTCCTCCAGAATTTACTTTAGAGA
GTACTGTTGACGACCCTTCATTAAACTTTGAAGACTAATAAACCAAGTAATCTTTTGACTTTTCTACTGACGGATCAGCTGGGTTTTGACACAAGAA
GATTTTATACTGGCTGTGATTTTCTCAAGATGCTAATCATTTATAGTTACTGATATTATTTGTCATATTATATAAAAGTGTAATTAAAAAAAATTAT
TTTTGAATATGTTCCTAAATATTTTGTATTTTCAGTTGTTTTTTAAATATAAGTTTAATTTCAGTTTTAAA

PROTEIN 797 (AA)
MPGRRSFSPAAFVPVKLGQFPLSEKILIAALPDVKMNLDDLMIRVEKVLPCLKEDLSPRTYLARMHGLLYLEEVAMLKQVGELQMDRASFRKEKDCL
ILDVPPLSPYTSKLVAGDMLIASLPWAADERKYEGVIHRVTQTELWLQFAKDFHDSYGEGIQYSVSFVTSRASLRRMHQAINLAAKHLGFGWLFPTG
VTSRLPQVIVEEEIEESESVRIARKENYSKKLHCVPTPESKDLLYSPDPRKISVAGQANLSIEHKFQAASYKSGYHGRGRKYNNRDIQWNRLGSGYG
MHCTVDNEVRVNIGLERIKKIRWFNKGLNRQQKEAVNNVLLGEARPLPYVIFGPPGTGKTVTVVETILQIHALIPESRLLVATPSNSAADLITERLL
DAGDLEQGDLLRMVGYHYLEQGRIAASIVPYAAVPDVKAINVAGLSGASHEGVQMCGRELLGQHRVTVGTLGCLGLLYNMGFPRGHFTHVIVDEAGQ
ATEPELLIPMVFLHMEYGQVVLAGDPLQLGPVVTSRLASRCGLQDSLLARFLNRFPYSRDPHGFPDSSGYDPRLVTKLVNNYRSLPTILELPSMLFY
DNDLIPNVSEDSSEEAALLRALASLLPCRTFGGYAPPLLFHGVRGTNCQETESHSWYNPQEVFQAFVYLNLLYKAGLRPDQVGIITPYQLQSNKIRF
MLERINVEPPKVGSVEEFQGQEKMAIIVSVVRSSPDLISYDMQRALGFVANARRLNVALSRARGILIIILGNPHLLYLDMHWRSVLRHCVKKKFYTGC
DLPPEFTLESTVDDPSLNFED

Graphical representation



Comparison with *Cylas puncticollis* Armitage

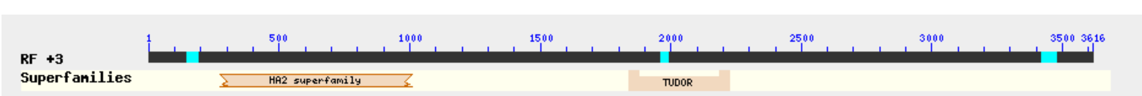
```
Query  414   GQRN-STVRFKSVKIPDYPVPKKLLDLIVKYHNQIDKINLIEELKVTKRSLFNNLTSTNY  472
             G+R+ S   F  VK+  +P+ +K+L +      +++  +L+  ++     L  +L+   Y
Sbjct  3     GRRSFSPAAFVPVKLGQFPLSEKIL-IAALPDVKMNLDDLMIRVEKVLPCLKEDLSPRTY  61

Query  473   EDKFHTLLHLDEIANLICIRNYDQDLACFIRNGEFLMLEIENLSERRPSIVLGDRIIASD  532
               + H LL+L+E+A L  +     D A F +  + L+L++  LS     +V GD +IAS
Sbjct  62    LARMHGLLYLEEVAMLKQVGELQMDRASFRKEKDCLILDVPPLSPYTSKLVAGDMLIASL  121

Query  533   PLRIAKEDYEGNVFKVGAHHVYLKFSALFHDKYN-GEDYSIRVIPGRSTYRRQHHAIYMV  591
             P    +  YEG + +V    ++L+F+  FHD Y  G  YS+  +  R++ RR H AI +
Sbjct  122   PWAADERKYEGVIHRVTQTELWLQFAKDFHDSYGEGIQYSVSFVTSRASLRRMHQAINLA  181

Query  592   ARNLGQDWLFP----SKIVERDVQVRFTYDHYSRVMR--NIKKRLR--------------  631
             A++LG  WLFP    S++ +  V+         R+ R  N  K+L
Sbjct  182   AKHLGFGWLFPTGVTSRLPQVIVEEEIEESESVRIARKENYSKKLHCVPTPESKDLLYSP  241

Query  632   -PREL------------------YQIAVEENRRKSEANATEVEETGS-------------  659
              PR++                  Y+       RK      +    GS
Sbjct  242   DPRKISVAGQANLSIEHKFQAASYKSGYHGRGRKYNNRDIQWNRLGSGYGMHCTVDNEVR  301

Query  660   --IDDSSLLKLQWFNQQLNCKQKDAVVNILKGVARPLPYIIFGPPGTGKTVTVIEAILQI  717
               I    + K++WFN+ LN +QK+AV N+L G ARPLPY+IFGPPGTGKTVTV+E ILQI
Sbjct  302   VNIGLERIKKIRWFNKGLNRQQKEAVNNVLLGEARPLPYVIFGPPGTGKTVTVVETILQI  361

Query  718   IRLIPESRLLVTAPSNSAADLIALRLIDSGVLKPGNLVRLVSVNYAVGDNIPARLVPYCA  777
               LIPESRLLV  PSNSAADLI  RL+D+G L+ G+L+R+V  +Y     I A +VPY A
Sbjct  362   HALIPESRLLVATPSNSAADLITERLLDAGDLEQGDLLRMVGYHYLEQGRIAASIVPYAA  421

Query  778   TAGFAKEGTAEANNVLENGMICDCSRSVLGRHKITVSTCSSAGSLYQMGFPRGHFTHIIV  837
                      A  +     G +  C R +LG+H++TV T    G LY MGFPRGHFTH+IV
Sbjct  422   VPDVKAINVAGLSGASHEG-VQMCGRELLGQHRVTVGTLGCLGLLYNMGFPRGHFTHVIV  480

Query  838   DEAGQAAEPEVMIPISFLDKWNGQIILAGDPMQLGPVVISRIAEECGLNESFLERLTNRF  897
             DEAGQA EPE++IP+ FL    GQ++LAGDP+QLGPVV SR+A  CGL +S L R  NRF
Sbjct  481   DEAGQATEPELLIPMVFLHMEYGQVVLAGDPLQLGPVVTSRLASRCGLQDSLLARFLNRF  540

Query  898   PYIRDAEGFPDTEGFDPRLVTKLLYNYRSLDVILELFSSMFYHGDLIPTISDRDSKEAMF  957
             PY RD  GFPD+ G+DPRLVTKL+ NYRSL  ILEL S +FY  DLIP +S+  S+EA
Sbjct  541   PYSRDPHGFPDSSGYDPRLVTKLVNNYRSLPTILELPSMLFYDNDLIPNVSEDSSEEAAL  600

Query  958   LTSLSEILPKRD-DGSVPSIVFHGVKGENYQTNDSPSWYNPHEVAQVFYYVNELYRLGCE  1016
             L +L+ +LP R   G  P ++FHGV+G N Q  +S SWYNP EV Q F Y+N LY+ G
Sbjct  601   LRALASLLPCRTFGGYAPPLLFHGVRGTNCQETESHSWYNPQEVFQAFVYLNLLYKAGLR  660

Query  1017  PTNLGIITPYSKQVKEIKSILREAEFDLPKIGTVEDFQGQEFGVIILSTVRSSRDYIPSD  1076
             P  +GIITPY  Q  +I+ +L     + PK+G+VE+FQGQE   II+S VRSS D I  D
Sbjct  661   PDQVGIITPYQLQSNKIRFMLERINVEPPKVGSVEEFQGQEKMAIIVSVVRSSPDLISYD  720

Query  1077  LEHSLGFVSSPRRLNVAISRSKSLLIIIGNPNLLCYDTYWRTVITYCMEKGAYTGCDL  1134
             ++ +LGFV++ RRLNVA+SR++ +LII+GNP+LL  D +WR+V+ +C++K  YTGCDL
Sbjct  721   MQRALGFVANARRLNVALSRARGILIILGNPHLLYLDMHWRSVLRHCVKKKFYTGCDL  778
```

Homeless (spindle-E)
cDNA

Contig 4827, Length-3647
CAATGAATGTAAACCAGTTCTCAGATTATTTCTCAAGATCGGTCAATGGAGTTCTCATTGAACCTCCTATTATTGAACTTGGCCATGCACCTAAATT
TTCTGTCCAATTTTATTATTTAGACAGCCTGGTTAACAGATTACCTGGAATCAAAATTCCAGAAATTGATATTAGTTCTCCATCTATTTCCTCCTGT
ACATATGAATTGGCTCGAGCTCTAATTGAAGTCTTCCATGTTATTGATGAGCATGAACACCATGATACTCGTTTCATTGGTTCTGTTCTTATATTTT
TACCTGGTATTGCTGAAATTGAGAATATGTATGCTAAACTCATCAATTATTGTGGCTCAAATGAGAAATGGTGGTTATGCCCTCTTCATTCATCAGT
TACTTATGATGAACAAATGAAGGTCTTCCAGCCTGCTCCTAGAGGGCACCGAAAGATCATTCTTGCTACTAACATAGCTGAAAGTTCTATCACCGTT
CCTGATATAAAATATGTGATTGATTTCTGTCTATCTAAACAGCAAATTTTGGAGCCCAATACGAACTATTCCTGCCTTCAGTTAACCTGGATATCTC
GTTCTCAAGGAACGCAACGAGCTGGAAGGGTGGGAAGAGTTATGCCTGGGCGTGTATACAGACTTATGCCTAAGAATTACTACATGGAACTGCCAGA
AGATATTACACCAGAAATTATGAGGTGCCCTCTTGACCAGTTGGTCTTAAAAGCAAAACAGCTCAAGATGGGACCTCCTGCTACACTTTTAGGTTTT
GCCATTGATCCACCTGATTTGTCAAATATTCACAAGACCATTCTTCATTTAAAAGAGGCTGGCGCTCTTCTTTTGACTGCTGATGGTAAATATAATG
CCAATGATGGTGACCTCACATTCATAGGCCAAATAATGGCTTCTCTTCCTCTTGACATACATCTTTCAAAGCTCATAATTCTCGGTCACATGTTTTC
ATGTCTTTCAGATGCTATAATTATGGCTAGTGCTTTGTCGGTGAAAAGTATATTCAGCACTCCATTCAGACAGCAACTTGAAGCTTACAATTCCAAG
CTGACTTGGGCCGACTCTTCTTGCAGTGATCCTATAGCTTACCTTAATGCTTATTCGCTTTGGAGATACAAATCTAAAATGGGGTATTTCAAGAGAT
CGGGTGGGGAGAGTGAGGTTGATTGGTGTCGAAAGTATTTTATTCAAGGAAATTTAATTAAAGAGGTTCTGCGATTAGAGAATGAAATAACTCAACG
CTTGAGAGTACTTGGTGTAGAAGAAATTAAATTGGAAACTGGTGTGTCCTGGTCTGCATCAGAAAAGCCGCTTGTATTAAAACTCATATTAGCCGGT
GCTTTCTACCCAAATTACTTTATCCCCTATGGTTCAGATGAGAAAGACTCAGTAAAACTGTTAGGTGGACGTGATCCATTTAATACTGTCTATTTAA
TGGGAATGCCGAACAATCAACCTGGTCCTTTGTACACTCATGGTATAAGAAGCCATTTTTTGAACTGTGGTTCGAACATAGAAGTATCTTTTGATGG
CAGCAGTAAGGTGTATTTGTCTTTTGGCAGCAGTATTCCCCTTAATACGGAACCTGAAAACATTGGTGTGGTACCTGGAAAAATAAGTATGGCTGTT
TATAGGAGCATAAAACTACGACAACTTCAGATTCCTATAAAAATTCCTGTCCTACTACCTCATGAAGCACGGAAAAGGGCACAAGAAATTTTTGGGG
AAAGGTTGAGACCCACTCTCTTCAATGTAAACAAGAGAAATGCAGTAACTCCGGTTAAGAAGTCAAACATGCCCTCTTTAAGAACAAATATTTTATC
AGTAGTTATAACCCATGTTGAAAATCCCTCCAAGTTTTGGACTTCTATTAATGAACCTGTTAATGCCAACAAAGTGTACTGGATTCATTCTTCATTA
AACAGGCGACTGCAACCATTACCACCTTATTATGGGCCTTACGATGATGGAAGTGAATGTATTGCACCATATAGAGATGGCAATGCACCGATACAAT
ACTTCAGGGCAAGGATCATCAGTAGTATTGAAACTGATAGATTCTCTGGGAAGGTCACCAGAGTACAGATATTTTTTTATGATTATGGTAATCAAGA
ATCAGTTTATACAAAAGAGTTGAAGGATTATCCTCCAGAACTAAGTTCAATTAAAGAAGAACCTGCTCTTGCAATGGAAACAAGTTTGGCTGAGGTT
GGGCCATCAGTAAGCAAAAACCCTCGTGGGGACTGGTCTTTCGAAACGATTAAGGAGTTTCGGAAATTGTTAATAAAGGACCAAGCAGTTGTAAAAA
TATACTCAGTTGTCAATGAGGTTATGGCAGTTACCCTCTACACTCATGAGGCATTAGAGAATTGTACTGACACGGATTTCCCATTCAAAAGGTCCAT
TAATTATTATCTTATTAAACATGGTTTTGCTGAAGCTATTGATGAACCATATCTTTCTCGAGATAATCATTTGCTTCGAGAAGCTTCTGAGAAGATT
CCAGGTTTGATCACAACCACTATGGACAATGACTTGTTCAATGACAATTTATCTGATATAAACTTTACTCCTCCTAGACCTGAAGAATGCCGGAGTA
AGATAACTTTGAAAGGTCCACGCTCACCCTTGGAGATGAATCTATATAGCTTGCCAAAGAAGTGTCAAGCTAAAGAAACTCTAATTGAATGGAGTTC
AGTTAACTCAGTTCTTCTTGACACTTGCCCCATGGATCCGCACAGCAGGTTATTAGTAGCAGCATCAGTTACACAGAATGCCAACAGCAGTCGCCTA
ACTCTCAGAAATACTACCTTAATGCCAAATATCCATGGTTTAGGTTCATTGATCTGTCTCCTCTTTTCTCCTAGGGTAGAACTCCGGGCTGATCCAG
AGAGACGATATTTGACTGGTGCCTTATGTGGTCTTGGCTTTGATCCAGAAACTGGAGTTGCTTTTTACCCAGAAAATGATATAGAAATCACCTTTGA
TACCAAACTAACACTTTTGGATTTAGAAATGATAAACAAATTGAGGTTTTGGATGGATTATATGATGGGTGGAGGAGATGGACCTAATAGTGAAATA
TCCAGACCAGGAATCCTCAAAGCCCAAGAGAAGATCAAATCTTACATTATAGACCTGCTGAGTAAGCGACGTCCTTCAGTGACCAGTTATATAACAT
CATATCCTTATGAATGGGACCAGTTGGAACCAGACGAGCTCCTCGATCCTCAGGGCTCCGTAGTCAGTCTTTACCCTCTGATTTGGGGAATCTCACT
TGTTGGTGAGGGTGCCAGGGACAGTGCAGTTCTATCAAGGGTTGAGACACTGAATATGATTGCTGAAGGGAAAGAACCATTTAAAACAGCTGTACAA
TGTGAACTATGCAACATTTACAATGAAACTTTGCAAGAATTAAAATTACACCTGACAACTACATTACACACTGTGAAGTTGACTGATTTCAAAGAAA
GTATGAAGCCTAGAAGGAGGCATTGATTTTCTCCCTTGTTTTATTTATTTCCAATTTTACTTTTACAGAAATATTTTGTCTATATTCATCTTAAAAG
TTACTATTTTATAGTTTATACAAGA

PROTEIN 1771 (AA)
MNVNQFSDYFSRSVNGVLIEPPIIELGHAPKFSVQFYYLDSLVNRLPGIKIPEIDISSPSISSCTYELARALIEVFHVIDEHEHHDTRFIGSVLIFL
PGIAEIENMYAKLINYCGSNEKWWLCPLHSSVTYDEQMKVFQPAPRGHRKIILATNIAESSITVPDIKYVIDFCLSKQQILEPNTNYSCLQLTWISR
SQGTQRAGRVGRVMPGRVYRLMPKNYYMELPEDITPEIMRCPLDQLVLKAKQLKMGPPATLLGFAIDPPDLSNIHKTILHLKEAGALLLTADGKYNA
NDGDLTFIGQIMASLPLDIHLSKLIILGHMFSCLSDAIIMASALSVKSIFSTPFRQQLEAYNSKLTWADSSCSDPIAYLNAYSLWRYKSKMGYFKRS
GGESEVDWCRKYFIQGNLIKEVLRLENEITQRLRVLGVEEIKLETGVSWSASEKPLVLKLILAGAFYPNYFIPYGSDEKDSVKLLGGRDPFNTVYLM
GMPNNQPGPLYTHGIRSHFLNCGSNIEVSFDGSSKVYLSFGSSIPLNTEPENIGVVPGKISMAVYRSIKLRQLQIPIKIPVLLPHEARKRAQEIFGE
RLRPTLFNVNKRNAVTPVKKSNMPSLRTNILSVVITHVENPSKFWTSINEPVNANKVYWIHSSLNRRLQPLPPYYGPYDDGSECIAPYRDGNAPIQY
FRARIISSIETDRFSGKVTRVQIFFYDYGNQESVYTKELKDYPPELSSIKEEPALAMETSLAEVGPSVSKNPRGDWSFETIKEFRKLLIKDQAVVKI
YSVVNEVMAVTLYTHEALENCTDTDFPFKRSINYYLIKHGFAEAIDEPYLSRDNHLLREASEKIPGLITTTMDNDLFNDNLSDINFTPPRPEECRSK
ITLKGPRSPLEMNLYSLPKKCQAKETLIEWSSVNSVLLDTCPMDPHSRLLVAASVTQNANSSRLTLRNTTLMPNIHGLGSLICLLFSPRVELRADPE
RRYLTGALCGLGFDPETGVAFYPENDIEITFDTKLTLLDLEMINKLRFWMDYMMGGGDGPNSEISRPGILKAQEKIKSYIIDLLSKRRPSVTSYITS
YPYEWDQLEPDELLDPQGSVVSLYPLIWGISLVGEGARDSAVLSRVETLNMIAEGKEPFKTAVQCELCNIYNETLQELKLHLTTTLHTVKLTDFKES
MKPRRRH

Graphical representation



Comparison with *Cylas puncticollis Homeless (spindle-E)*

```
Query  273   IKANEFAYYFRRQTFGQSISAPIISIDKVSQYTKTIFYLD-MINKITQ-KTPKFDIYKPE  330
             +  N+F+ YF R   G  I  PII +     +++    +YLD ++N++    K P+ DI  P
Sbjct  1     MNVNQFSDYFSRSVNGVLIEPPIIELGHAPKFSVQFYYLDSLVNRLPGIKIPEIDISSPS  60

Query  331   ISEEVWAVFVFLISIFDRLEEGEATVNKN-GTVLVFLPGIYEIEEAHRRLMNEETNTNKW  389
             IS   + +   LI +F  ++E E     +  G+VL+FLPGI EIE  + +L+N   +  KW
Sbjct  61    ISSCTYELARALIEVFHVIDEHEHHDTRFIGSVLIFLPGIAEIENMYAKLINYCGSNEKW  120

Query  390   EIIPLHSSLPNDEQGRAFAPRKPGLRKIILSTNIAESSVTVPDSSFVIDFCQTKVMTVHP  449
              + PLHSS+  DEQ + F P   G RKIIL+TNIAESS+TVPD  +VIDFC +K    + P
Sbjct  121   WLCPLHSSVTYDEQMKVFQPAPRGHRKIILATNIAESSITVPDIKYVIDFCLSKQQILEP  180

Query  450   ETKYMSLKLEWASHVNCDQRAGRVGRTCNGRVYRLVNRQFYMSEMEKHSKPEILSAPLER  509
              T Y  L+L W S      QRAGRVGR   GRVYRL+ + +YM E+ +    PEI+  PL++
Sbjct  181   NTNYSCLQLTWISRSQGTQRAGRVGRVMPGRVYRLMPKNYYM-ELPEDITPEIMRCPLDQ  239

Query  510   IVLQAKMLELNETPAQILALAMTPPGLKNIEDTIWQLKEIGGMLRTCRGIQANADGDITF  569
             +VL+AK L++    PA +L  A+ PP L NI  TI  LKE G +L T  G     DGD+TF
Sbjct  240   LVLKAKQLKMG-PPATLLGFAIDPPDLSNIHKTILHLKEAGALLLTADGKYNANDGDLTF  298

Query  570   LGQVMANLPIDVHLSKLIVLGQLFSCLEETIIIAAGCSIQNIFSIPFQKRLDAYRKQLLW  629
             +GQ+MA+LP+D+HLSKLI+LG +FSCL + II+A+  S+++IFS PF+++L+AY  +L W
Sbjct  299   IGQIMASLPLDIHLSKLIILGHMFSCLSDAIIMASALSVKSIFSTPFRQQLEAYNSKLTW  358

Query  630   SDGSCSDLIALLNLYTVWLNLKRENAF--GSNRQELEWCRTNLVSLKGLKEWNLLITEIR  687
             +D SCSD IA LN Y++W    +   F         E++WCR    +    +KE   L  EI
Sbjct  359   ADSSCSDPIAYLNAYSLWRYKSKMGYFKRSGGESEVDWCRKYFIQGNLIKEVLRLENEIT  418

Query  688   QRLERMQIKETDGPGRVILSDIEKPTVLKIIMAGAFYPNFFIKSPHSSQMTEREAVKIVG  747
             QRL  + ++E       V  S  EKP VLK+I+AGAFYPN+FI  P+ S   E+++VK++G
Sbjct  419   QRLRVLGVEEIKLETGVSWSASEKPLVLKLILAGAFYPNYFI--PYGSD--EKDSVKLLG  474

Query  748   GRDPFSTVYFTGMDPRQPGQVYVRPIKKLIRENGEREMDVQVGFDGSSKIYVEFKNV--L  805
             GRDPF+TVY  GM   QPG +Y   I+       G    +++V FDGSSK+Y+ F +   L
Sbjct  475   GRDPFNTVYLMGMPNNQPGPLYTHGIRSHFLNCGS---NIEVSFDGSSKVYLSFGSSIPL  531

Query  806   QREPITVNVDGFQKLTNIPGRIPMSMYEILRKRQLQYPFQLKVLPYNEAWEFAEKHGLKR  865
               EP  + V         +PG+I M++Y  ++ RQLQ P ++ VL  +EA + A++   +R
Sbjct  532   NTEPENIGV---------VPGKISMAVYRSIKLRQLQIPIKIPVLLPHEARKRAQEIFGER  583

Query  866   NMPSGKLPYQERNCYSSVEYSPLPSMDTTYISLNISNHIDAGHFWAQNATEETNVFLQQI  925
               P+       +RN  + V+ S +PS+ T  +S+ I++  +   FW              + I
Sbjct  584   LRPT-LFNVNKRNAVTPVKKSNMPSLRTNILSVVITHVENPSKFWTSINEPVNANKVYWI  642

Query  926   EEALNK--QVLKQVTENIKTDKIYAARYAEDG----QFYRCKVTV------LGGQLN--Q  971
               +LN+  Q L                A Y  DG    Q++R ++          G++   Q
Sbjct  643   HSSLNRRLQPLPPYYGPYDDGSECIAPY-RDGNAPIQYFRARIISSIETDRFSGKVTRVQ  701

Query  972   ILFIDYGNLQEVRSDELY-YVPQRPQCVMAP-LAFECILHGVKPTYRLNPVGMWDGKVNE  1029
             I F DYGN + V + EL  Y P+        P LA E  L  V P+   NP G W  +  +
Sbjct  702   IFFYDYGNQESVYTKELKDYPPELSSIKEEPALAMETSLAEVGPSVSKNPRGDWSFETIK  761

Query  1030  DFKRLTDGILLYGEVYSVVGDVVELILYRSQ------NRDL----SLNQYLLNRQFAEPS  1079
             +F++L          ++YSVV +V+ + LY  +        + D     S+N YL+    FAE
Sbjct  762   EFRKLLIKDQAVVKIYSVVNEVMAVTLYTHEALENCTDTDFPFKRSINYYLIKHGFAEAI  821

Query  1080  TPSFLSKQNHEQRQTVQMAENPDQEASRLNYDKI---ISYSDFVTPQLNHSKCSTIYLKG  1136
                +LS+ NH  R+  +    P    + ++ D      +S  +F  P+      + S I LKG
Sbjct  822   DEPYLSRDNHLLREASEKI--PGLITTTMDNDLFNDNLSDINFTPPRPEECR-SKITLKG  878

Query  1137  PFSPLEVKLYSCTNMGSTKSVFVEGTSVNAVLLDNEPESPHSRLVVAVHVGQSVDGMKLK  1196
             P SPLE+ LYS        K    +E +SVN+VLLD  P  PHSRL+VA  V Q+ +  +L
Sbjct  879   PRSPLEMNLYSLPKKCQAKETLIEWSSVNSVLLDTCPMDPHSRLLVAASVTQNANSSRLT  938

Query  1197  LRQTTLMPSLPGFPMLMTAMFCPTMEPKLTSDNMLVAAILCGLGYNEITGKALYPAHDIC  1256
             LR TTLMP++ G   L+  +F P +E +   +   +    LCGLG++  TG A YP +DI
Sbjct  939   LRNTTLMPNIHGLGSLICLLFSPRVELRADPERRYLTGALCGLGFDPETGVAFYPENDIE  998

Query  1257  LGLDTELKIEELEMVNRLR----YLMNNGIKLLHDIEQNMSNQNELIHTQIKIKENLFDL  1312
             +   DT+L + +LEM+N+LR    Y+M  G      D    +  ++   ++   Q KIK  + DL
Sbjct  999   ITFDTKLTLLDLEMINKLRFWMDYMMGGG-----DGPNSEISRPGILKAQEKIKSYIIDL  1053

Query  1313  IRMERNPIERRNVVPQNTTWGKHAAGLGLLKPSMDNDEQ----DIWSLLWFVKLAGYNPR  1368
             +   +R P      +       W +        L+P    D Q      ++ L+W + L G    R
Sbjct  1054  LS-KRRPSVTSYITSYPYEWDQ-------LEPDELLDPQGSVVSLYPLIWGISLVGEGAR  1105
```

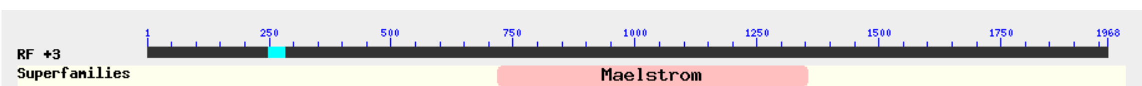
Homeless Maelstrom

Cdna

Contig 9069, Length-1803
GGGGTCGACCCCATGATCTGGAAGACCTGTGTCTTTATTTCATTCAAACAGGAAACGTTTTACGAAAATTAACAATTTTTCTAGCAATCGAGATTAA
TTGAATGTACACAATTTTTTAAGGATTAATGTTACCTCAAATATCACTATACTCGATATTAAAATTTTCAATTTCGAAACGGTATTGAAAGTATCCC
ACAGTTAATGTGACAATTTAGGATGGCGATTTAATATTTTAATCGATTTGTAGAAGTTAACTTTTGAACTAATCGTTTCTTACATTATTCCCTTATC
CGTCAATTGCTAGACTGATCCCGGAATTCAAAATCATGCCACCAAAAAAGAAGCAAAAGGGTGGAAGGAACGGCTTTTATTATTTCATGTTGGATCT
ACAGAAGCAAGAAGCACTAAAGGGGAACCGTTACTCCCTTCAAGAAATTTCAGGATTAGCAAACCCTTTATGGTCGGATATGTCTTTAGAGGAAAGG
AAACCTTTTACTGATTTGGCCCTTAAAAATGCCTCTAAAAACTCTGGAGACTTAACTAAGAAGTTCACTTCTTTAGGAATAAGTTATGCTGAGTTAG
ACGCTGAGGAAAGAGAAATAGAGGATGCCAGAAACATTATGATTGCAACAATTAAGAATACTATTGCAAATCTCGATTTAAATTCTACTCTGAAGAG
TCATATATTCTTTGTGTGTCATGCAAATTATTTCTATAAATCTGAAAAGCAAGTCTATTTCCCAGCTGAAGTTGCTGTTGCTGCATTTTCCCTTGAG
CTAGGTGTTCTTGCAACTTTACATTTTTTTGTTGATCCTGGGAAAATCCCTCTCGGTTTTAAGTATGAATCTTCAGAATGGTCTTCAAGGACCCATC
GCATACCTGTCAGTGACCCATCTTGGAAAGAGGGAATAAAAGATCCATTAGAAATGTATCGAAGAGTTAAAGAATTCATCAAGAAATTTACAGTTGG
GACTGAAGTTCCTCCAATTTATACTTTGTTTGAAACCTCGAACTCACACTATCGTCTTTCTGTTGTAAAATCTGCCTTGAATATGATGTGCGAAGCG
GCATATGAAGATACATCCTTATTTCGAGTATATTGTCTAGCACATTTATTTTTTGAAATAAGAAATAAATGTTCTGCAGAAATACCTAGTGTAGCTA
TTATGAAGAATGAACTTGAAAAAGACGTGTTTTCCTATGCAAGAGATTTAGGATGTTATTATCATGAGGAAAAGGATCTCTCAATACATTGCTCACT
CTCTATTGTAACTAGGTGGGTTTTTATTATTTGTGACCATTGCTGTAAACATCTTGGTATTAAACTGATTCTAGGAAGACATGTCCCAAAAGATACA
GCTTTATTAGATAGGGCACATTATGTGAGAAAGAGTGATGGCTTATCACCGAAACAAAATGTTGCTGAATCCAGCAGTAAACCATTGACAATAATTG
ACCATGGAAGTTTAAAAGAACAAAGAGCTGAAGAGAGAAGATCTCAAGAGATGTACTTAAGAGAAAGTGAAAAAATCAGACCTCCAAAGTACTGTTA
CAGTGATAAAATTGGTGCTAGATTTAGTCAAAATAGTTCCTCACAATTGTCTCCTGAAATAGGAGCAACATCCAATGAAAATGAATGGTGTCTGGCT
AAAGGAAAGTCTTCTGGGCGAGGTCGTGGAATTTTAAGTTATTCATCTCCAGAACATCCTCAATGGCCAATATCTGGACGAGGACGAGCTTTTGCAA
ATTTTCATAATTAAAGATTTTATTTTGTTCCTATAAAAAGCAAAGTGGTTCTAATTTTTAGCGTTAATATATTTTTGTTTTATTAAGGAAAGGTGAC
AGTTCCTATTTTAATGCGATAAGTTGCTATTTTTTTAATTTAGATAATTAAGTTCAATTTCAGGTCTTAAGTAACAATTGTTAAATTTTAAATATGT
AATTGTTTATAAATAAATATACAGGAC

PROTEIN 477 (AA)
MPPKKKQKGGRNGFYYFMLDLQKQEALKGNRYSLQEISGLANPLWSDMSLEERKPFTDLALKNASKNSGDLTKKFTSLGISYAELDAEEREIEDARN
IMIATIKNTIANLDLNSTLKSHIFFVCHANYFYKSEKQVYFPAEVAVAAFSLELGVLATLHFFVDPGKIPLGFKYESSEWSSRTHRIPVSDPSWKEG
IKDPLEMYRRVKEFIKKFTVGTEVPPIYTLFETSNSHYRLSVVKSALNMMCEAAYEDTSLFRVYCLAHLFFEIRNKCSAEIPSVAIMKNELEKDVFS
YARDLGCYYHEEKDLSIHCSLSIVTRWVFIICDHCCKHLGIKLILGRHVPKDTALLDRAHYVRKSDGLSPKQNVAESSSKPLTIIDHGSLKEQRAEE
RRSQEMYLRESEKIRPPKYCYSDKIGARFSQNSSSQLSPEIGATSNENEWCLAKGKSSGRGRGILSYSSPEHPQWPISGRGRAFANFHN

Graphical representation



Comparison with *Cylas puncticollis Maelstrom*

```
Query  1    MAPKKPAKP--NGFFMFASEFKNRH---GRKYASMKEACDAAAPYWAKMNAQEKQKYQDM  55
            M PKK  K   NGF+ F  + + +    G +Y S++E    A P W+ M+ +E++ + D+
Sbjct  1    MPPKKKQKGGRNGFYYFMLDLQKQEALKGNRY-SLQEISGLANPLWSDMSLEERKPFTDL  59

Query  56   AREE--KGSG---KKYTSEGLDVEELTRKEKEQEQFQEKMRQDIRQIIELASMKGSEALA  110
            A +   K SG   KK+TS G+   EL  +E+E E  +  M   I+  I  A++  +  L
Sbjct  60   ALKNASKNSGDLTKKFTSLGISYAELDAEEREIEDARNIMIATIKNTI--ANLDLNSTLK  117

Query  111  SETFFLIHFNIFCFHPPSDQYYPAEVAAVCFNLKDGVKPENVFHEFVKPGPMPLGYSFDA  170
            S  FF+ H N F +      Y+PAEVA   F+L+ GV      H FV PG +PLG+ +++
Sbjct  118  SHIFFVCHANYF-YKSEKQVYFPAEVAVAAFSLELGVLA--TLHFFVDPGKIPLGFKYES  174

Query  171  KQHSEETHKISVP---YDDGENNMEEVFVKLTTFLKKKKPANVTRLPILF-----ANDKY  222
             + S  TH+I V    + +G  +  E++ ++  F+KK      T +P ++     +N  Y
Sbjct  175  SEWSSRTHRIPVSDPSWKEGIKDPLEMYRRVKEFIKKFTVG--TEVPPIYTLFETSNSHY  232

Query  223  Q----KMFQNILDRWSWEYGGEENLFRVYSLQVLLFWLRNKVSGGEVWHTVTFSDREIEK  278
            +    K   N++   ++E   + +LFRVY L  L F +RNK S  E+  +V     E+EK
Sbjct  233  RLSVVKSALNMMCEAAYE---DTSLFRVYCLAHLFFEIRNKCSA-EI-PSVAIMKNELEK  287

Query  279  DVYAYVPDIACHYHSKMSSSLYCSQSNVLRMAYVVCDNCCEDLDIELLPGQHVP  332
            DV++Y  D+ C+YH +   S++CS S V R  +++CD+CC+ L I+L+ G+HVP
Sbjct  288  DVFSYARDLGCYYHEEKDLSIHCSLSIVTRWVFIICDHCCKHLGIKLILGRHVP  341
```

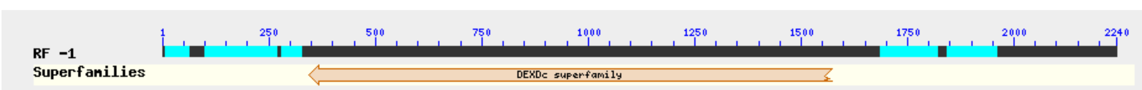
Homeless RNA helicase Belle

cDNA

Contig 3595, Length-2233
TCTGGCTAGGCTGAAGAAGAGCGGGGGAGGGGGCGAGATCGATTGGCTGAGAGAGGCAGGGGAACGGTGAGAGAATTATTGCCCCCACCAATCAGGG
CCGGAGCTGTTGTTTTGGTTCGATGAATATGAACCTCCATAGCCTGAACTGCTGCTACCAACTCCATAGTAATTATTAGAACTGGAATAACCACCAT
AGCCAGATCCTCCAGACCTCGAGATATTTCTTGAAGAATTTCCTCCTCCACTGCCACCACCCTGTTGTCGGTAGTCTCGAGCACCAAATCCACCAGA
AAATCTACCGCCGCCTTTTCCACCCCGACGAGAAGAAAGAGGCTGTTTGCTTTCAGAAGCCATTCTCTGCAACCACGAAGGCATATCTTGATTTGTT
TCTTGAATAAGATCATAAAGGTCTCTGGCTAAATTACGATTTTTATCATTAAAGAATGATGTGGCTAACCCTAAGTTACCCATACGTCCAGTACGCC
CAATTCTATGGACATATTCTTCAATATCATTTGGTAAGTCAAAATTAATGACGTGTTTCACATGAGGAATATCCAAACCTCTAGCAGCAACAGCAGT
TGCCACAAGAACAGGTGTTTGACCTGTTCTAAAACTTCTGAGCGCTTCTTCTCTTTCTTTCTGAGACCTATCACCATGAATGGAGGTCACAGGATAT
CCCGAACCAAACAAAAAGTCTTCAAGAGAGTCTGCGCCTTTTTTAGTTTCAACAAATACCAAGGTTAAAGATACAGCATTCTTCGATGCATTTAGAA
GGTCTAATAGGAAACCCCTTTTGTCATTTTCTTCAACCCAAACAATATTCTGAGTAATGTTTTCTGAGGTAGACCCAACACGACCAATAGCCAAAAA
TATGTAGTCATTTAAAAAATCTAAGGCCAACATCTGAATTTCTTTTGGAAAAGTAGCTGAAAACATCAAAGTTTGCCTATCACCAGTAGGAGGCATT
GAGTGTTTTTCTACAATCCTCCTTATTTGTGGTTCAAAACCCATATCTAACATTCTATCAGCTTCATCAAGTACTAGATACTTGCAATTTTCTAATC
CTATTCTTCCTCTTGCAAGCATGTCAACCAAACGCCCAGGAGTAGCAACCAATAAATGGCAACCACGTTCAAGATCTCTCATTTGATCAGCAACATG
TAAACCACCATACACAACACAAGGCCGAAGTCGAGATCTGTATGCAAACTTGCGAGCTTCATCATAAATTTGGGTAGCTAATTCCCTAGTAGGTGCC
AGAACAAGACCCATTGGATATTGTTTACGTTTACCAGGATATCTCGTGCTCTGAGGTCCACGTTCAAGCATTTGGTTCAAGATGGGTACAAGAAAGG
CAGCAGTTTTACCACTACCAGTTTGTGCACATGCCATCAGATCTCTCTTGGCCATAATTATAGGGATAGCATGCTTCTGAACAGGGGTAGGCTTATC
ATAACGAGCCAAACGAATATTATTTCCAATGATTTCAGTCATCTGTAAATCCTCGAATGAACTAATGTGGGGAGGCACTTGATCACCAGTAGCCTCA
ACTGGTATATCCTCATATTTATTGAAGTTAATGCCAGTGTTAGCATTTTCAAAAAGCTCAAATTCAAGTCTATCATCCTTGGGTAAAAGAGTGGTCC
AGCCGGCAGAAGAACCTTGTGGCTTTTCTTGCCATCTGTCATTGCCACCGGAAGGGGTTCCTCCTCCACCTCTCCTACTACCCCAGTCATCTCTGTC
ACGCTCTCTATCTCTGTCCCTTCCTCCACTCCTGAAAGAAGAAGAACCACCTCTGTCCCCCCACGACCGATCTCCATTTTGGGAGTCAAAATCACCA
TTTCTTCTGCTTCCTCCACGGTAGTTATCACTTCTACCACTACCACCGCTACTGCGAAAATTGCCACTTCCTCGGGAAGAAAAGTCACGACTGTCGC
TGTAAGATGAAAAACTTGAGGGGGGCTGTTCACCACTACGAGATTCTCCAGCTTGTTTATTCCGGAGATGAGGAGGAATGTAGGGGCCGGTTTCTTT
ACGGCTACCCTCCAAGTCCAGACCAGCAAATCGCTGCTCTAGACCTGATCCATTTTGGTTGGGTACATTACTCATATTACTTTAAATCAGTACTCTG
AAATTAAGCCCTTTTAAGTGCTTTTCCAATATAATAAGACACAAAGAAAAGTGGAATGATGTACACGGGATTATACACTAGACCTGTACACTAAAGC
ATCGATA

PROTEIN 678 (AA)
MSNVPNQNGSGLEQRFAGLDLEGSRKETGPYIPPHLRNKQAGESRSGEQPPSSFSSYSDSRDFSSRGSGNFRSSGGSGRSDNYRGGSRRNGDFDSQN
GDRSWGDRGGSSSFRSGGRDRDRERDRDDWGSRRGGGGTPSGGNDRWQEKPQGSSAGWTTLLPKDDRLEFELFENANTGINFNKYEDIPVEATGDQV
PPHISSFEDLQMTEIIGNNIRLARYDKPTPVQKHAIPIIMAKRDLMACAQTGSGKTAAFLVPILNQMLERGPQSTRYPGKRKQYPMGLVLAPTRELA
TQIYDEARKFAYRSRLRPCVVYGGLHVADQMRDLERGCHLLVATPGRLVDMLARGRIGLENCKYLVLDEADRMLDMGFEPQIRRIVEKHSMPPTGDR
QTLMFSATFPKEIQMLALDFLNDYIFLAIGRVGSTSENITQNIVWVEENDKRGFLLDLLNASKNAVSLTLVFVETKKGADSLEDFLFGSGYPVTSIH
GDRSQKEREEALRSFRTGQTPVLVATAVAARGLDIPHVKHVINFDLPNDIEEYVHRIGRTGRMGNLGLATSFFNDKNRNLARDLYDLIQETNQDMPS
WLQRMASESKQPLSSRRGGKGGGRFSGGFGARDYRQQGGGSGGGNSSRNISRSGGSGYGGYSSSNNYYGVGSSSSGYGGSYSSNQNNSSGPDWWGQ

Graphical representation



Comparison with *Cylas puncticollis RNA helicase Belle*

```
Query  1    MSNAPNQNGSGLEQQFAGLDLQ-SRGISDRYVPPHLRNKQQHSS----QPESGSSYDREK  55
            MSN PNQNGSGLEQ+FAGLDL+ SR  +  Y+PPHLRNKQ   S    QP S  S   +
Sbjct  1    MSNVPNQNGSGLEQRFAGLDLEGSRKETGPYIPPHLRNKQAGESRSGEQPPSSFSSYSDS  60

Query  56   GRGESRGGGRNFNRGGGSGGSRDAR---ENRGDFSSFN-TRNRRDNFQNGEVFENREKEE  111
                SRG G NF   GGSG S + R      GDF S N  R+  D   +          +
Sbjct  61   RDFSSRGSG-NFRSSGGSGRSDNYRGGSRRNGDFDSQNGDRSWGDRGGSSSFRSGGRDRD  119

Query  112  WGREARG-GGRFQERERDLPRNDRWQEPEKPKDRWNESSNSREGGNGRWNDNRERRNEND  170
              R+    G R          NDRWQE  KP+        S  G
Sbjct  120  RERDRDDWGSRRGGGGTPSGGNDRWQE--KPQ-------GSSAG----------------  154

Query  171  WTIPLARDERLEMELFGTGNTGINFSKYEDIPVEATGDKVPRHIASFEEVQLTEIIRSNI  230
            WT  L +D+RLE ELF   NTGINF+KYEDIPVEATGD+VP HI+SFE++Q+TEII +NI
Sbjct  155  WTTLLPKDDRLEFELFENANTGINFNKYEDIPVEATGDQVPPHISSFEDLQMTEIIGNNI  214

Query  231  AAARYDTPTPVQKYAIPIILSKRDVMACAQTGSGKTAAFLVPILNQMYEVGPPNIQHGRS  290
              ARYD PTPVQK+AIPII++KRD+MACAQTGSGKTAAFLVPILNQM E GP + ++
Sbjct  215  RLARYDKPTPVQKHAIPIIMAKRDLMACAQTGSGKTAAFLVPILNQMLERGPQSTRYPGK  274

Query  291  RRSKQYPLGLVLAPTRELATQIYDESKKFAYRSRVRPCVVYGGAHIGDQLRDLERGCHLL  350
            R  KQYP+GLVLAPTRELATQIYDE++KFAYRSR+RPCVVYGG H+ DQ+RDLERGCHLL
Sbjct  275  R--KQYPMGLVLAPTRELATQIYDEARKFAYRSRLRPCVVYGGLHVADQMRDLERGCHLL  332

Query  351  VATPGRLLDMIDRGRIGLDYCKYLVLDEADRMLDMGFELQIRRIVEKESMPRTGERQTLM  410
            VATPGRL+DM+ RGRIGL+ CKYLVLDEADRMLDMGFE QIRRIVEK SMP TG+RQTLM
Sbjct  333  VATPGRLVDMLARGRIGLENCKYLVLDEADRMLDMGFEPQIRRIVEKHSMPPTGDRQTLM  392

Query  411  FSATFPHPIQMLARDFLDNYIFLAVGRVGSTSENITQKVMWVEEHDKRSLLLDLLNVNDL  470
            FSATFP  IQMLA DFL++YIFLA+GRVGSTSENITQ ++WVEE+DKR  LLDLLN +
Sbjct  393  FSATFPKEIQMLALDFLNDYIFLAIGRVGSTSENITQNIVWVEENDKRGFLLDLLNASK-  451

Query  471  HQPSAESLTLVFVETKKGADSLEDFLHQEGYPVTSIHGDRTQREREDALKQFRSGNTPIL  530
               +A SLTLVFVETKKGADSLEDFL   GYPVTSIHGDR+Q+ERE+AL+ FR+G TP+L
Sbjct  452  ---NAVSLTLVFVETKKGADSLEDFLFGSGYPVTSIHGDRSQKEREEALRSFRTGQTPVL  508

Query  531  VATAVAARGLDIPHVKHVINFDLPSDIEEYVHRIGRTGRMGNLGLATSFFNDRNRNLASG  590
            VATAVAARGLDIPHVKHVINFDLP+DIEEYVHRIGRTGRMGNLGLATSFFND+NRNLA
Sbjct  509  VATAVAARGLDIPHVKHVINFDLPNDIEEYVHRIGRTGRMGNLGLATSFFNDKNRNLARD  568

Query  591  MLDLLIEAKQEYPSWLESVASDSRMP  616
            + DL+ E  Q+ PSWL+ +AS+S+ P
Sbjct  569  LYDLIQETNQDMPSWLQRMASESKQP  594
```

PRP16, mut6 homolog

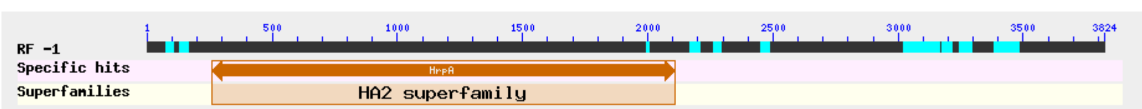
cDNA

Contig 3764, Length-3819
GCCTGAATGGTGTCCTAAGAGGTGTGCTTCCCGGAGCTCCCGGCTCAGCGACTTGATTACTTCTTTTGGACCATCGACCACGCTCTTCCTCTTCCTG
TCGCTCTCGATCTTTCATCTGCTTCTGGGCTGCTTCCATCTCTGCCTCCATCCCCTCCATAGCCTCACGCCTCTTGGAGGAGCCCGAACGCCCTGTC
TCCTTGATAGAAAAGAACATCGGTCCCAGTTCTGCCAACCATACACCATCCACAGCAGTAACACACTGCATGTATTCCTTTGATGTCATAACAAGTT
CGTGATAAACGACATAATCTGGGGTAAAACCCATGCCAAAGAGAGCGGATGTTGGATGTAGATGACAAGGCATACCTGTCCTCAGGTGAACGTATTC
TCCTATCCCTTTCAATCGAGCTCCTTGATGGAAGTAAGCTGAGCAAATACATTTTCTCACAATGTCCCAATCAGTTCCGCAGGAGATAACTTCCATC
TTCTGTTGGTCCAAGATATCCTTCAGCTGTTGTCTGACCTCACGAACCTTTCTCATTGCTTTTGCGTGTATAAAGTGGTCATTGCACCAAGTTGAAG
AATACTTGTTCTGCTTCCATTGGTTGTAAACGTTCAAATAAGTCAAATGATCTGATTCTGGAACCTGAAATTTTTCTCTTACACTGTCTGCTTCATC
CTCTCTTCCTTTTGGTCGGTAGAAAATAGTAGGAACTGAGAGCATCGAAACGATAATTAGAATCTCAGCCGTACAGCCCATTTGATTTGAGACAATC

AACATCTGACATTGTGGTGGATCCAAAGGAAATTCGGCCATCTGCCTTCCGAGAGCGGTCAATGAACCTGTATGATCCAAAGCTCCGAGGATCCATA
ATTGGTATAAAGAGTTAAGAATGTTGTCCTGCGGGGGTGGATCCATGAAATGAAACTGTAAAAGATCTTGAACTCCCAAAGATTTGAGCAGTAGCAC
TGTATTGGCAAGGTTAGTTCTTTGTATTTCAGGTACTGTTGATGTAAGGAGTTCTTCCTTATATTGTCTTTCCGTGTATAATCTGTAAGCGTGACCG
GGTCCAGTCCTCCCTGCCCTGCCGGACCTCTGATTGGCATTGGCTTGGGAAATTGGGTAGATCTGGAGTGCGTCCATACCTATCCGAGGGTTATAAA
CCTTCAGCTTGCAATACCCGGAGTCAACGACATACATGATTCCGTCGACAGTAAGAGATGTCTCGGCAATGTTGGTGGCAACTACACACTTACGTAC
TCCATCAGGCGAAGCTTGAAATATCTTAGCTTGTAGATCTGAGGGTAACTGCGAATAGATTGGCAATATGGAGAGAGGCGGAGCAGCGTCTATTTCT
CCAAGCCTTTCTCCTAACACCTCGCATGTCACTTCAATATCTTCCTGACCTGGCATGAAGATAAGTATATCACCTTTGGTTGGTTGAAGATGTATCT
GGAGTGCCTGCTTTACAGCTGCTTCCACGTAGTCCTCAACAGGATTCCTAGAAAAGAGTACTTCCACAGGAAACGTCCTTCCTGGTATTGTGTACAC
TGGAACATGTCCAAAGAACATTGAGAACTTGGAGGAGTCCATTGTAGCAGAGGTAACTATCAGTTTTAAGTCCTGCCTTCTAGCCACAACCTGTCTT
AAAAGACCAAAAAGTACATCAGTATTCAGTGATCTTTCGTGAGCTTCGTCCATAATTATGGCAGAGTAATTGTCGAGGTCTGGTTCTCGTAAAGATT
CCCTTAGCAATATACCATCAGTCATATACTTTATAACAGTGTCCTCTGAAGTACAATCTTCAAATCGAATGGCATATCCGACAGTCTTTCCTAACGG
TGAAGCCATCTCGTCAGACACTCTCTTGGCGACAGACATGGCAGCGACTCTCCTCGGTTGCGTACATCCAATCATCCCTGACCTAGAATAACCATCT
TCGTGTAGGTACTGGGTCAGTTGTGTCGTTTTACCAGAACCAGTTTCACCGACTACGACCACAACACTATTTTCCCTGATAATATTCAACAGCTCTT
CTCTAACAGCAAAGACAGGCAAGTACTGTCGCTGCTGTTGAATAGTTTTCTTCTTTGCGAATTCACTAGCAGCGGGTGTCATGTCTTTCATGTGTTC
AGCGAATTTCTGACTAGTTTTGTAATCGGCCTCAACTTCTCCATCTATTTTGTCATCTTCCTTCTCATCCTTTTTAGGTACACCCATGATGTTACCG
ATAGTAGTTCCTGACAGCTCCCAATGCTTCTTCTGAGCCCGTTTCCTCTCTTTTTGTTCCCTATAGGCTCGGACCAATAATGATCCTTTTCTTGAGA
CGAGTGCCATATCTGATGTAGGGTCCTTGACTGGTATAACAGGCTCCGGTTGTTTGGTAAAGACTATACGACCGTCCAAGAATGGAGGGACAATATT
GTGAACCAGAAGATGTACTCTAGCTTCTGCTTCCTCATCGAAATCTTCGTCCAAGTCGAGGTTCTGAACTACTCCTGAGGTCAACATCCTGTTCCTT
TCCCACAGCTCATTATCCTTATTGATCTGCCTTTGCTGAGCTGACATTCTTTTCTTCTTAGAGCGCTCTAACTGCTTCTCCCTTTTGTTGATATATT
CCTGGCTGACAGAGCCGAATGGGTCTCTACCCTCCTCAAAACCTTGATCCAGACCATACCATTCCCGGTCTAGCCTTTTTTGTTCCTCTTCCCACAG
GATGCGGTCCTTCTCAGATCCCCAAGGAACTTCTTCTTCTTTACCGACTAACGGAGTTGCTCCAGTTCTTGCTCTCTCCGGTGCCCAACTATTAAAC
CTATGCGCAGGAGTAGGCCTCCTCATGGACCAATCGTCTTCCCTTGGTTCAGCCTTCGGAGTTGGTTGTTCCCAAGAAGATTTTTTTGCATCAGGAG
TAATATCATCTTCATCCCACCCGCATCTGGAAGGTGTGTCTTTGACAGAAAACATTGGTGTACGAGGTTCGTCATGAAACCTAGGAGATTCTGACCA
AGATCTCTGTGATCTGCCGTAGCTGCGCTCACTCCTTCGATCACTCCTGTCATCTCTTCTATCCCTCTCTCTGTCCCTATCACCTCGGCTCCGATCT
CTGTCCCTATTTCTTTCCCTGTCTCTGCTCCTTTCTTTGTATCTGTCACGATCCCTGGACTTACCCTGCCTCTCATCTTTTCTTTTGTCCTCTTCA
TTCTCATCATTGTTCTTTGAATCAGCTGTAACATGAATTCCTCTTTTCTTTTCCTCTCTTTTTCTCCTTTCATCCATTCTCTCTTTCGCCTCCCTTG
TAACTCCACCTGTATGGGTAGGAGTCTCATCGAACTTTTCTCTATAATGCCTATCATTATGTTTTTCTGGTTTAGCAAACGTGCACTTCTCCTCCTT
GCCTTCCTTCAGCTTCTTCTCTGTTTCTTTTTTATCTTCATTTTCTCTTTCCTTCCTTCGAAGAGCTGCAAGCCTATCCAATCCCAACAAAGATGTC
TGGGGGACTTTAAACTCAAAATTCCCTTTAGGTTCCGGTTTCTTTTTTATAATAAGTCCACCTTTTTGTTCTTCAACACCTTCAAGTCTATAAACAC
CGCTATCATCTTCCACTTCCATCGCTAACAAAACTTTGGAACACAATTGTAATATATATGCGCTTCGTAAATTACATTACAATCGTATATTTTCACT
GAACGTAATTAAGAAACTAATTACAAACACATAACATCATGGTAATTAAGATATTAACACTTCTGTTTTTGTTTAGGAAGGTAACATTTTAATAAAC
CCTCATAAGCATAGTAAAATGTACAAAGAGATCCTTCGAAT

PROTEIN 1203 (AA)
MEVEDDSGVYRLEGVEEQKGGLIIKKKPEPKGNFEFKVPQTSLLGLDRLAALRRKERENEDKKETEKKLKEGKEEKCTFAKPEKHNDRHYREKFDET
PTHTGGVTREAKERMDERRKREEKKRGIHVTADSKNNDENEEDKRKDERQGKSRDRDRYKERSRDRERNRDRDRSRGDRDRERDRRDDRSDRRSERS
YGRSQRSWSESPRFHDEPRTPMFSVKDTPSRCGWDEDDITPDAKKSSWEQPTPKAEPREDDWSMRRPTPAHRFNSWAPERARTGATPLVGKEEEVPW
GSEKDRILWEEEQKRLDREWYGLDQGFEEGRDPFGSVSQEYINKREKQLERSKKKRMSAQQRQINKDNELWERNRMLTSGVVQNLDLDEDFDEEAEA
RVHLLVHNIVPPFLDGRIVFTKQPEPVIPVKDPTSDMALVSRKGSLLVRAYREQKERKRAQKKHWELSGTTIGNIMGVPKKDEKEDDKIDGEVEADY
KTSQKFAEHMKDMTPAASEFAKKKTIQQQRQYLPVFAVREELLNIIRENSVVVVVGETGSGKTTQLTQYLHEDGYSRSGMIGCTQPRRVAAMSVAKR
VSDEMASPLGKTVGYAIRFEDCTSEDTVIKYMTDGILLRESLREPDLDNYSAIIMDEAHERSLNTDVLFGLLRQVVARRQDLKLIVTSATMDSSKFS
MFFGHVPVYTIPGRTFPVEVLFSRNPVEDYVEAAVKQALQIHLQPTKGDILIFMPGQEDIEVTCEVLGERLGEIDAAPPLSILPIYSQLPSDLQAKI
FQASPDGVRKCVVATNIAETSLTVDGIMYVVDSGYCKLKVYNPRIGMDALQIYPISQANANQRSGRAGRTGPGHAYRLYTERQYKEELLTSTVPEIQ
RTNLANTVLLLKSLGVQDLLQFHFMDPPPQDNILNSLYQLWILGALDHTGSLTALGRQMAEFPLDPPQCQMLIVSNQMGCTAEILIIVSMLSVPTIF
YRPKGREDEADSVREKFQVPESDHLTYLNVYNQWKQNKYSSTWCNDHFIHAKAMRKVREVRQQLKDILDQQKMEVISCGTDWDIVRKCICSAYFHQG
ARLKGIGEYVHLRTGMPCHLHPTSALFGMGFTPDYVVYHELVMTSKEYMQCVTAVDGVWLAELGPMFFSIKETGRSGSSKRREAMEGMEAEMEAAQK
QMKDRERQEEEERGRWSKRSNQVAEPGAPGSTPLRTPFR

Graphical representation



Comparison with *Cylas puncticollis PRP16, mut6 homolog*

```
Query  1     MDEEDGVHRLEGTTGNEKGGLVVKKKT----GSTFKVPQVSLLGLDRLAAIKRKEKEEAA  56
             ++++ GV+RLEG    +KGGL++KKK        FKVPQ SLLGLDRLAA++RKE+E
Sbjct  3     VEDDSGVYRLEGVE-EQKGGLIIKKKPEPKGNFEFKVPQTSLLGLDRLAALRRKERENED  61

Query  57    RKMSFSMDDEAAAENSDDHEGRKVPKDGRKFRSPAEETPTYTGGVSKEAKDRLREHMKNK  116
             +K +     E   E     +  K   + R +R   +ETPT+TGGV++EAK+R+ E  K +
Sbjct  62    KKETEKKLKEGKEEKCTFAKPEK--HNDRHYREKFDETPTHTGGVTREAKERMDERRKRE  119

Query  117   YKEKGVY--ASTKDRRDRERKRDRGDRSERRRNKDSSRR---------------------  153
              K++G++  A +K+  + E  + + +R  + R++D  +
Sbjct  120   EKKRGIHVTADSKNNDENEEDKRKDERQGKSRDRDRYKERSRDRERNRDRDRSRGDRDRE  179

Query  154   -----------------------EETPKFRDEPRTPNIKIKDNTARTSWEDEDESVPVKR  190
                                     E+P+F DEPRTP   +KD  +R  W+++D +   K+
Sbjct  180   RDRRDDRSDRRSERSYGRSQRSWSESPRFHDEPRTPMFSVKDTPSRCGWDEDDITPDAKK  239

Query  191   SSWDFPTPTSYKREDPSWSERSFRTKYESSRSFKHERSSRDSKSFRRYSDDTPRATPAHK  250
             SSW+ PTP +  RED  WS R                                R TPAH+
Sbjct  240   SSWEQPTPKAEPRED-DWSMR--------------------------------RPTPAHR  266

Query  251   FNNWMKDRKRTGATPGTGRDKDEPLKWDATVDRENWEEEQKRLDREWYNMDEGYDDENNP  310
             FN+W  +R RTGATP  G++++ P  W +  DR  WEEEQKRLDREWY +D+G+++  +P
Sbjct  267   FNSWAPERARTGATPLVGKEEEVP--WGSEKDRILWEEEQKRLDREWYGLDQGFEEGRDP  324

Query  311   FSSVSDEYTKKKEEQLEQRKKKRMSAQQRQINKDNELWERNRMLTSGVVQSVDLNEDFDE  370
             F SVS EY  K+E+QLE+ KKKRMSAQQRQINKDNELWERNRMLTSGVVQ++DL+EDFDE
Sbjct  325   FGSVSQEYINKREKQLERSKKKRMSAQQRQINKDNELWERNRMLTSGVVQNLDLDEDFDE  384

Query  371   ESIDRVHLLVHNIVPPFLDGRIVFTKQPEPVIPVKDPTSDMALVARKGSHLVRVYREQKE  430
             E+  RVHLLVHNIVPPFLDGRIVFTKQPEPVIPVKDPTSDMALV+RKGS LVR YREQKE
Sbjct  385   EAEARVHLLVHNIVPPFLDGRIVFTKQPEPVIPVKDPTSDMALVSRKGSLLVRAYREQKE  444

Query  431   RRKAQKKHWELGGTKIGNIMGIKKKEDEEDKKYNKEDDTTDYKADHKFAEHMKGSSEAAS  490
             R++AQKKHWEL GT IGNIMG+ KK+++ED K + E +  DYK   KFAEHMK  + AAS
Sbjct  445   RKRAQKKHWELSGTTIGNIMGVPKKDEKEDDKIDGEVEA-DYKTSQKFAEHMKDMTPAAS  503

Query  491   DFARKKSISEQRRYLPVFAVRQELLNVIRENSVVIIVGETGSGKTTQLTQYLHEDGYSKY  550
             +FA+KK+I +QR+YLPVFAVR+ELLN+IRENSVV++VGETGSGKTTQLTQYLHEDGYS+
Sbjct  504   EFAKKKTIQQQRQYLPVFAVREELLNIIRENSVVVVVGETGSGKTTQLTQYLHEDGYSRS  563

Query  551   GMIGCTQPRRVAAMSVAKRVSDEMGTQLGDEVGYAIRFEDCTSENTVIKYMTDGILLRES  610
             GMIGCTQPRRVAAMSVAKRVSDEM + LG  VGYAIRFEDCTSE+TVIKYMTDGILLRES
Sbjct  564   GMIGCTQPRRVAAMSVAKRVSDEMASPLGKTVGYAIRFEDCTSEDTVIKYMTDGILLRES  623

Query  611   LREPDLDHYSAIIMDEAHERSLSTDVLFGLLREIVARRHDLKLIVTSATMDSSKFSMFFG  670
             LREPDLD+YSAIIMDEAHERSL+TDVLFGLLR++VARR DLKLIVTSATMDSSKFSMFFG
Sbjct  624   LREPDLDNYSAIIMDEAHERSLNTDVLFGLLRQVVARRQDLKLIVTSATMDSSKFSMFFG  683

Query  671   NVPTFTIPGRTFPVEVLFSKNPVEDYVDAAVKQALQIHLQPPSGDILIFMPGQEDIEVTC  730
             +VP +TIPGRTFPVEVLFS+NPVEDYV+AAVKQALQIHLQP  GDILIFMPGQEDIEVTC
Sbjct  684   HVPVYTIPGRTFPVEVLFSRNPVEDYVEAAVKQALQIHLQPTKGDILIFMPGQEDIEVTC  743

Query  731   EVLAERLGEIDNAPELSILPIYSQLPSDLQAKIFQRSPEGIRKCVVATNIAETSLTVDGI  790
             EVL ERLGEID AP LSILPIYSQLPSDLQAKIFQ SP+G+RKCVVATNIAETSLTVDGI
Sbjct  744   EVLGERLGEIDAAPPLSILPIYSQLPSDLQAKIFQASPDGVRKCVVATNIAETSLTVDGI  803

Query  791   IFVIDSGYCKLKVYNPRIGMDALQIYPISQANSNQRSGRAGRTGPGQAFRLYTERQYKDE  850
             ++V+DSGYCKLKVYNPRIGMDALQIYPISQAN+NQRSGRAGRTGPG A+RLYTERQYK+E
Sbjct  804   MYVVDSGYCKLKVYNPRIGMDALQIYPISQANANQRSGRAGRTGPGHAYRLYTERQYKEE  863

Query  851   LLVTTVPEIQRTNLANTVLLLKSLGVQDLLQFHFMDPPPQDNILNSLYQLWILGALDHTG  910
             LL +TVPEIQRTNLANTVLLLKSLGVQDLLQFHFMDPPPQDNILNSLYQLWILGALDHTG
Sbjct  864   LLTSTVPEIQRTNLANTVLLLKSLGVQDLLQFHFMDPPPQDNILNSLYQLWILGALDHTG  923

Query  911   VLTKLGRQMAEFPLDPPQCQMLIVSTQMGCTAEILIIVSMLSVPSIFYRPKGREEEADGV  970
              LT LGRQMAEFPLDPPQCQMLIVS QMGCTAEILIIVSMLSVP+IFYRPKGRE+EAD V
Sbjct  924   SLTALGRQMAEFPLDPPQCQMLIVSNQMGCTAEILIIVSMLSVPTIFYRPKGREDEADSV  983

Query  971   REKFQVPESDHLTFLNVYNQWKQNNYSSHWCNEHFIHVKAMRKVREVRQQLKDILVQQKF  1030
             REKFQVPESDHLT+LNVYNQWKQN YSS WCN+HFIH KAMRKVREVRQQLKDIL QQK
Sbjct  984   REKFQVPESDHLTYLNVYNQWKQNKYSSTWCNDHFIHAKAMRKVREVRQQLKDILDQQKM  1043

Query  1031  EVKSCGTDWDIVRKCICSAYFHQAARLKGIGEYVNCRTGMPCYLHPTSALFGLGNTPDYV  1090
             EV SCGTDWDIVRKCICSAYFHQ ARLKGIGEYV+ RTGMPC+LHPTSALFG+G TPDYV
Sbjct  1044  EVISCGTDWDIVRKCICSAYFHQGARLKGIGEYVHLRTGMPCHLHPTSALFGMGFTPDYV  1103

Query  1091  VYHELVMTAREYMQCVTSVDGHWLAELGPMFFSLKETGKSGRAKKKQAAEHLLEMENQMQ  1150
             VYHELVMT++EYMQCVT+VDG WLAELGPMFFS+KETG+SG +K+++A E    ME +M+
Sbjct  1104  VYHELVMTSKEYMQCVTAVDGVWLAELGPMFFSIKETGRSGSSKRREAMEG---MEAEME  1160

Query  1151  VAQEEMKARKEAADKKEAAMYKGQEIV----SAGTTPRRTPAR  1189
              AQ++MK R+   +++     K    V    + G+TP RTP R
Sbjct  1161  AAQKQMKDRERQEEEERGRWSKRSNQVAEPGAPGSTPLRTPFR  1203
```

Gawky

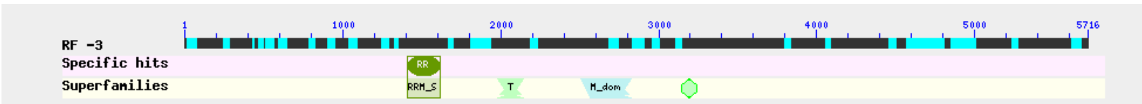
cDNA

Contig 602, Length-3560
CAGATAAATGAAATGTGTATGTGTGTGTGTGTGTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGTTGAGTGTGGATCCCACCTCATTCCCAACA
GGCGATACCAAAACCTATCCACAATTCACAAACCTAATATCTAAAAACATTTGCACCATCATTTATAAAACTACATAGAATAAAAATTATTTTTAAC
TATAACTCTTTCGTCCAAATAAAAGTTCTAATAAATTATGTACATAAACAAAAATAAAAAATTAAATTTCAGGTTCCACTTAGAAAATCTATGGTAG
CAACGAGACTGTATTGTATCAATCTGTAGATTAAAAATACATACTAAAAAGCCACTTGCTTGCTCACAGCGCATTGTTATCTTCAAAATATATACAA
TATATAATCATCAAAAGACTGCAGGTTTAATGCTATATAGTAACAATAGAAAATCTCACTTATATTAAAATCATTATTAAAAAAAAATCCCAAGGAA
AAAACCTAATACCATATAACACATGACTTAAATCGTTATAAAAATAAAAACAATGTGACAAATAAAACAAAATTACTTGCTAAATTATAAATTACCA
GATGACAACCATAAATTTATTACTAAAATTAAATTATTCAGCCAACATATATTATTTTACAATTATGTTTAAAGCATGTTTAAACAAATATAATAAA
ATGAAAAACACATAACCTATTTATTTTTAATTCACTTGTACCTCCAGTTTACAATTTTTCTTTTACCATTAATTTATAAAATAGCTCTTTATAATCT
TTTAATATAAATTATTACTTAGTAATATTGAATTTATAACAAATACTCAAACAGTCTGATTTCGTCCGACAGGAACTACGCATGTCGTTTTACAATT
TAAAACTCATTAAATTCATTCTTTAAATACAATAGTATTATTTATATTTATAATAATAATAATAATAATTATAACATATATGTTCCTGGTACTAATA
ATTGCATATACTTTATATATCATACATTTACAATAAAAAGGTCATGTGGTGTTTCTGAGATCAAAGTATTTGTTGTTTTTTTTTTTTTTTTTTTTTT
TGTTAGTTTTTGTTATAATTTCCTTGATAAGTTTGATTGACGTACCTATCGATCGAGAAGAGAGGGAGAATTCAGAAATTAGTGAAGATTTTCTTAA
ATTGATTCACCGCCTAGAAGATCACCTGGCAAATATGAATTAATGGAGGAAGGGGTATTGCGGTGCTGATCATTATCTCCAGCCCACAAGGAGGCCG
CAGTCGATCCAGCTCCAGTGGCACCCCACAGCTGGGAATTTCCCCAAGCCTCCTTTCCACGCCATGCAGAACTTCCACTGCCAGCACCACCTTGCAC
ACTTAGATGTTGGAGTAATGATAGAACATCGGATTCACTGGGACTTTCAGCAAATATTGTCGTGTTACCAAGGACACAATTATTGAGAGCGCCTTGG
GCCTTGCTGGCTTCTTCTCTAGAAGTATACTTAGCAAGTGCAATACCATGATTAAGGTAGAGATGAAAATTCTGAAGAGGTCCATGTTGCATACAAA
GAGTTTTTAAAGTTGACCCATCAATCTGAGGAGTAAGATTGCGTAATAATAACCATGTTGATCCAGGCTGAGTGCTAGCTGGCTGATTAGGTCCAGG
CCAGCCTTGTCCAGTAGACCAACGACCACTGCTGGACAAAGAGCTCCAACCATTAGCTGTCTGCCCGCCACTTGAACTTCCAGGTTTAGCAGGAAGA
CCTGGAGGAGGCCCTCTTTTAGGGGCCCATAATTCAGAGGAATTAGTTTCCCAAGTAGAAGCACTATTCTTCCCACCGCCATTGTTTGAGGCTGATC
CACCACCACTAACCCCACCACCACTAGAATTGAGCTTCACACTAGCACTAGATGAAGATGGTGGTGGGTTGAAACTCCATGTAGATGATGACAAGCC
AAGAGAAGGGATAATATCAAGAGATGAACTGGTAGTTGGAGAAGCTTTACCAGCTCCAGTATTAGTTGATAAAATATGAGTATCTTTAATTGAAGGT
AAAGAAAGAGGTGATCTCACAACTGAACCCGGAGTAAGGCTCGGGTCATCTTCAATACTTTTCAATGGATTACCCTTCCAAGGCTTTCCAGGTTCAA
ATTCTGGAACAAGGTCAGAAAATGCCTGCGATGGAGGTTGACTTGAATTTGGCCAATCCTTTCCAGAAGAATCATCACCTGAAGAATCAGGCCAACC
AGATTCACGATTCACAGAAGACCAAGTGCCATCAGCTTGACCAAGTAATAAATTGGTATTGGGTGAAGAAGAACCCTGTCCTGACTTAGCAGTAGTC
CCTGGAGCTCTTGAGAATTCACCAGTCCCAATATCAGGGTCTTTTTCCAATGCGGGTAACTTCCACTGATTTAATCTTGACTGCTGACTAGTTCCAC
TGGTCGGTGGATCCTTAAGGGTAAGATCTGAAAAATTTGGATGAAGCGCTGCAATAGGTTCATGAAGAGCGCTCTTAAAAAATTCAGTTGAGGGAGG
CGTGTGCTGTTGATGCTTAACATACACTGCTTGCTGAGCAGCTATTTGATTCTGGAGATTAGATATCTGCTGCTTAGTTTTAGTTATTTGCACAGAA
AGCTGCAGTAAATGGTTGGAATTTGGTTTTCCAAGAGGATTTACTTGCATTACCGTGTGATGTTGCATAAGCTGCTGAAGAGACTTGATTTGTTGAA
GTAATTGATTCAACAAAATGAGTGTTTGAGGCGCTAATGGCTGATTTAAGATCTGAGGGCTTAGGTAGCCAGTTTGAACTGCCATCTGAATTTGTTG
AACCAACATTCTAAGTTGTTGAGCAGAGGGTTGTTGTGTTTGCTGAGTTCTTGAAGGCTGCTGAGCAAAAGGCTGCTGCTGTGGAGGTTGCTGAGCC
AATAGTTTTTGCATTACAGCAGGAGTTATTGTAGCCAATGTAGGTCCACCTCCTGGACCACTTCCACTTCCACCACCACCTCCTCCACTAGGAAACG
GAGCTAAACAGATGGGCTGTTGGGGTGATGGAAATGCATTCATGATTGCAGCATGATCAGCATCTAAATCACCTCTCGGCCACCGATCACTAGCTCC
ACCTCCACTACTTAGGCCTCTTGAGTTGCTGTTGAGCTGCTCCAAAGCATCTTCCAAACTCATGCCAGAATTTCTTAATGCACTCTCTATGTCTTCC
TTCTTGTAACCCATGTCAGACAGAATTCTAAATTGTTTGCTGGTCCAAATAAACTCCTTGGAAACAGGTTTAGAAGGCTGTTTGGAATGGACCCAAG
AAGCAGGATCTAAGTCAGTTCCATCCCAGGCAGGTTTATGTTTAGGTCCTCCTTGCCATGATGATGGAGTCAAACCTTGTTCCATCCATGGAGCAGC
ACTTTGTGTAGGTTTATCATCAACCCATGGACCACCTTCCATCCCTCCCTCCCATCCTCGCTCATTTCTAGGATGATGAGGCCACAAGGGTTGATCA
GGTTTAATAGATGGATTAGCCCTCCCAGGAGGACACTGTAAACCACCTCTTCCAATTCCAGGAGCAGGCATCTCTTTCCAATGAGAAACTTTATTTT
GACCCTGAATGGAAGGTCTTGGATGTTGTGGTCCCCATAGTGATGTCCCATCATCGTAATTAGGCATATTCCTTCGTTGTGCAGGTGGAGAAGGTTC
TTCCCAACCAGTTGATTTTCCACCAAGACCTGTCATATCCTTCATTTGTGGTCCTGAGCTGGTCCACTGATTAACTGGTTGATTTGGTTGAGCTGGT
AATTTTCCTTGGTGATGAGGTATGTGATGATGAGGTCCTGGACCAGGCCACATGGCTTCACTAGTCACACCATTCAATCGACCACTTATTCCTCTTG
GATCACCTCGCATTATAACATCATTCCCACCCATTGACCTATGATCAATAGAAGTTGGTCTCATATGATCTCTAGCATCCATCATCTGTCTCATATC
ATTATGACGCAAATCCCTGGAGTCACGCTGGTCATTCCAATTGGGGCCACCCCAGTCTCCCTCTTTTTTGGTAGCTGGCCAAATAGGAGGTGGCTGA
GCAGGCCACTGAGGTCCACAACCAGCCTGTGTTTGAGGAACACCAGTCCATACATTGGAGGCATCCCCTTCATCATCTTCACCCCATGTTCCTCCAA
TGTTAGTACTTGGAGTATGAACCCATGGAGTCTTTTGATTTTGCTGGCTTACTGGAGGAGGTACACCTCCATTTCGTAAATTGGCTTCCCACAAATC
AGTTCCATTATTAATTGGTAATTTCCATAAAGGTACCGGAGTGGCACCGTTGCTCTCCTTGTTCCCTGGTTCTGGAGAACCTGGTATATCCCAGTTG
CTATCTTGATTGACATTTTGCCCTCCCCATCCATCTTGACTATAGAGAGCTTCCCTCATAGTATTTAACTGCTCCATTTGCTGCTTGGTAGAAGAGT
TTCCACTTGATGAAGGAGTAGTAGCCGAACTAGTAGGAGGAAGCCCTTTACCAACTGCAGTTGCCCATGTCGAACCACTGCCGACCACCTGGTTCCC
TGGTGCTCCATTCTGTTGTGTAGTAGATGTAGTAATAGTATTATTATTTGATGTATTGGTTGAAGAGGAGGCTGGAGCAGTCGTAACTGTAGATGTC
GATACACTTGCTGCAGGAGGAGCCGAAGATTGACTAGGTGGCCCACCTCCTGGTTGCTGTTGTTGGGCATTACCTCCCCCTGGTTGGTTCACTGAAT
TACGAGATCCATTATTGTTACCATTATTGGAATTGTTGTTGCCATTATTTTGAGTCGTACCCCATGGCTGCCCAGCTGAGGTGGGAGCCGGACTTTG
TCCAGTGGATCCCTGTCCTTGTCCTGACGAATTATTTCCTCCACTAGTGCTGCTAGTGCCATTCCACCCTGTACTTCCACTCTGATTCGAAGGAGGG
GACCCCCATCCAGTAGTCCCCGTAGTTAAAGAACTCTCACCTCCCCCCAATAATTTTAAATTTCGTGGGATTCCCCATCGAGTATAATAATCACAGA
TTTTAAAGCTTTTACCTTCTTGATCGGTATCATCGTTGGTTTGTGGAATGTTTGGCAATAAAATTGAACTCTTGCTTGGGTAAACACATTTATCCTT
AACACTAAATGATTCATTAACACTAATAGACTTGATGCTATCTAAGCAAAGAGCACCGATTTCTTGTGTCTCTGCTAGTGGCATTTTGGAAGAACCG
TCAGATTCACTATCACCAATTTCTGATGAAGAAATCTTGGCGGGCGCTTGGGACGGATGGTTGGTTATCACAGTCAGGACGGGGTCTTCTTCTTCTT
CTCTTACCCCACAATCTACCACCATTACGTCCCTTTTCTCAGGCACTTTAAGAGTAGGACTATTGGGCTTAGGAGGCTGCGCCATACCCCTTAAGCT
ATCAGATTTGTCCTCCTCGTCTTTATTTTGTACAAAGGCATTAGTTTTAGAAGAAATCTCATTTGAACTAGAATTGTTTCGAAACATTGGCTCTCAG
TACGTCTTTGGCAGGCAAAATGTAGAAACCAAAGTCTTACAGACTAATTATCTTGGGCTTTTCTCATTGTAGCAGAAATCTTAACTTAACGATGATA
CAAGCTACATGGATATCACAAACACAAATACAAGAGATAATTATTTAGTTTCGGTAAGAATTCAGCCTACCTCTGAACTACTTAAAC

PROTEIN 1452 (AA)
MFRNNSSSNEISSKTNAFVQNKDEEDKSDSLRGMAQPPKPNSPTLKVPEKRDVMVVDCGVREEEEDPVLTVITNHPSQAPAKISSSEIGDSESDGSS
KMPLAETQEIGALCLDSIKSISVNESFSVKDKCVYPSKSSILLPNIPQTNDDTDQEGKSFKICDYYTRWGIPRNLKLLGGGESSLTTGTTGWGSPPS
NQSGSTGWNGTSSTSGGNNSSGQGQGSTGQSPAPTSAGQPWGTTQNNGNNNSNNGNNNGSRNSVNQPGGGNAQQQQPGGGPPSQSSAPPAASVSTST
VTTAPASSSTNTSNNNTITTSTTQQNGAPGNQVVGSGSTWATAVGKGLPPTSSATTPSSSGNSSTKQQMEQLNTMREALYSQDGWGGQNVNQDSNWD

```
IPGSPEPGNKESNGATPVPLWKLPINNGTDLWEANLRNGGVPPPVSQQNQKTPWVHTPSTNIGGTWGEDDEGDASNVWTGVPQTQAGCGPQWPAQPP
PIWPATKKEGDWGGPNWNDQRDSRDLRHNDMRQMMDARDHMRPTSIDHRSMGGNDVIMRGDPRGISGRLNGVTSEAMWPGPGPHHHIPHHQGKLPAQ
PNQPVNQWTSSGPQMKDMTGLGGKSTGWEEPSPPAQRRNMPNYDDGTSLWGPQHPRPSIQGQNKVSHWKEMPAPGIGRGGLQCPPGRANPSIKPDQP
LWPHHPRNERGWEGGMEGGPWVDDKPTQSAAPWMEQGLTPSSWQGGPKHKPAWDGTDLDPASWVHSKQPSKPVSKEFIWTSKQFRILSDMGYKKEDI
ESALRNSGMSLEDALEQLNSNSRGLSSGGGASDRWPRGDLDADHAAIMNAFPSPQQPICLAPFPSGGGGGGSGSGPGGGPTLATITPAVMQKLLAQQ
PPQQQPFAQQPSRTQQTQQPSAQQLRMLVQQIQMAVQTGYLSPQILNQPLAPQTLILLNQLLQQIKSLQQLMQHHTVMQVNPLGKPNSNHLLQLSVQ
ITKTKQQISNLQNQIAAQQAVYVKHQQHTPPSTEFFKSALHEPIAALHPNFSDLTLKDPPTSGTSQQSRLNQWKLPALEKDPDIGTGEFSRAPGTTA
KSGQGSSSPNTNLLLGQADGTWSSVNRESGWPDSSGDDSSGKDWPNSSQPPSQAFSDLVPEFEPGKPWKGNPLKSIEDDPSLTPGSVVRSPLSLPSI
KDTHILSTNTGAGKASPTTSSSLDIIPSLGLSSSTWSFNPPPSSSSASVKLNSSGGGVSGGGSASNNGGGKNSASTWETNSSELWAPKRGPPPGLPA
KPGSSSGGQTANGWSSLSSSGRWSTGQGWPGPNQPASTQPGSTWLLLRNLTPQIDGSTLKTLCMQHGPLQNFHLYLNHGIALAKYTSREEASKAQGA
LNNCVLGNTTIFAESPSESDVLSLLQHLSVQGGAGSGSSAWRGKEAWGNSQLWGATGAGSTAASLWAGDNDQHRNTPSSINSYLPGDLLGGESI
```

Graphical representation



Comparison with *Cylas puncticollis* Gawky

```
Query  302   TSTKQQLEQLNNMREAIFSQDGWGGQHVNQDTIWDIPSSPEPQIKMDGAGGPPP-WKPAV  360
             +STKQQ+EQLN MREA++SQDGWGGQ+VNQD+ WDIP SPEP  K      P P WK  +
Sbjct  354   SSTKQQMEQLNTMREALYSQDGWGGQNVNQDSNWDIPGSPEPGNKESNGATPVPLWKLPI  413

Query  361   NNGTELWEANLRNGGQPPP--QPQQKTPWGHTPSTNIGGTWGEDDDAADSSNVWTGVPSN  418
             NNGT+LWEANLRNGG PPP  Q  QKTPW HTPSTNIGGTWGEDD+  D+SNVWTGVP
Sbjct  414   NNGTDLWEANLRNGGVPPPVSQQNQKTPWVHTPSTNIGGTWGEDDE-GDASNVWTGVPQT  472

Query  419   QQQWGNNGSNAMWGGAAAASAGTGGAAPTSG-WGDPRSADPRAAVAAVSAMDMRP-DMRV  476
             Q      G    W   A       A    G WG P   D R       + D+R  DMR
Sbjct  473   QA-----GCGPQW---PAQPPPIWPATKKEGDWGGPNWNDQR------DSRDLRHNDMR-  517

Query  477   AAAAGNLDPRQLDPREQMR-----HMTGGGD---MRGDPRGITGRLNGAGAEF-------  521
                      + +D R+ MR     H + GG+   MRGDPRGI+GRLNG  +E
Sbjct  518   ---------QMMDARDHMRPTSIDHRSMGGNDVIMRGDPRGISGRLNGVTSEAMWPGPGP  568

Query  522   ---------------------WGQTGPLGGPAGMHHQNKMPGVGPGNGTAWDEPSPPSQR  560
                                  W   +GP             M G+G G  T  W+EPSPP+QR
Sbjct  569   HHHIPHHQGKLPAQPNQPVNQWTSSGP--------QMKDMTGLG-GKSTGWEEPSPPAQR  619

Query  561   RTMPNYDDGTSLWG--NPQQGAGG---------MGGRGGGPGG---PPGMA-PSRVGGGG  605
             R MPNYDDGTSLWG  +P+    G         M   G G GG   PPG A PS
Sbjct  620   RNMPNYDDGTSLWGPQHPRPSIQGQNKVSHWKEMPAPGIGRGGLQCPPGRANPS------  673

Query  606   LKPDGSVWGGGP---AGGAGVGRGNGWDDVAAGPTAASGWDDPT--VGPWPKQKIPGAAG  660
             +KPD  +W   P    G  G   G  W D           D PT    PW +Q +  ++
Sbjct  674   IKPDQPLWPHHPRNERGWEGGMEGGPWVD-----------DKPTQSAAPWMEQGLTPSSW  722

Query  661   GGG---GGLWDTGDLD---WGH-KQSIKPPQLTKEMVWNSKQFRTLVDMGHKKEDVENAL  713
              GG      WD  DLD   W H KQ  KP  ++KE +W SKQFR L DMG+KKED+E+AL
Sbjct  723   QGGPKHKPAWDGTDLDPASWVHSKQPSKP--VSKEFIWTSKQFRILSDMGYKKEDIESAL  780

Query  714   RMREMNFDEALDMLS-----ISRNRGDPGWITRHDDHYDHPQFPGMGAQRGFPNVGPTNP  768
             R   M+ ++AL+ L+      +S   G       R D   DH
Sbjct  781   RNSGMSLEDALEQLNSNSRGLSSGGGASDRWPRGDLDADHAA------------------  822

Query  769   LSNAFPPNNAPNLL------NNMPGTGGGQSNSSLINNISPAIMQKMLTQQGGGGGAAQS  822
             + NAFP    P  L            G+G G       +  I+PA+MQK+L QQ
Sbjct  823   IMNAFPSPQQPICLAPFPSGGGGGGSGSGPGGGPTLATITPAVMQKLLAQQPPQQQPFAQ  882

Query  823   FGGGGGGPTGGRPLQPQSQPSTQQLRMLVQQIQMAVQAGYLNHQILNQPLAPQTLVLLNQ  882
                           Q   QPS QQLRMLVQQIQMAVQ GYL+ QILNQPLAPQTL+LLNQ
Sbjct  883   QPSR---------TQQTQQPSAQQLRMLVQQIQMAVQTGYLSPQILNQPLAPQTLILLNQ  933

Query  883   LLQQIKNLQQLNNQHPIAAASGKGNLTNNVLMNYSVMITKTKQQILNIQNQIQAQQALYV  942
             LLQQIK+LQQL   H +   +  G   +N L+  SV ITKTKQQI N+QNQI AQQA+YV
Sbjct  934   LLQQIKSLQQLMQHHTVMQVNPLGKPNSNHLLQLSVQITKTKQQISNLQNQIAAQQAVYV  993

Query  943   KQQQNNGNIGYDTFKTNTMHDTIHALQGNFGELAITKDSQVN--QQQSRLNQW----INK  996
             K QQ+          F  + +H+ I AL  NF +L + KD  +   QQSRLNQW    + K
Sbjct  994   KHQQHTPP--STEFFKSALHEPIAALHPNFSDLTL-KDPPTSGTSQQSRLNQWKLPALEK  1050

Query  997   DKEE-NGEFSRAPGSSSKP--VATSPNMAPLGLTQPDGPWSTGRTADTGWPDSSGGDSSN  1053
             D +    GEFSRAPG+++K    ++SPN   L L Q DG WS+     ++GWPDSSG DSS
Sbjct  1051  DPDIGTGEFSRAPGTTAKSGQGSSSPNTN-LLLGQADGTWSS-VNRESGWPDSSGDDSSG  1108

Query  1054  DKDAQWATTAQP---SLTDLVPEFEPGKPWKGNQIK-IEDDPSITPGSVVRSPLSIATIK  1109
              KD  W  ++QP   + +DLVPEFEPGKPWKGN +K IEDDPS+TPGSVVRSPLS+ +IK
Sbjct  1109  -KD--WPNSSQPPSQAFSDLVPEFEPGKPWKGNPLKSIEDDPSLTPGSVVRSPLSLPSIK  1165

Query  1110  DNELFSMNT---SKSPPVS---DAMQSLSLSSSTWSFNPTSTSSAFTSSPQIKLPTSKGG  1163
             D  + S NT     SP  S   D + SL LSSSTWSFNP  +SS+ +    +KL +S GG
Sbjct  1166  DTHILSTNTGAGKASPTTSSSLDIIPSLGLSSSTWSFNPPPSSSSAS----VKLNSSGGG  1221

Query  1164  LGDLNPSMAV-------------TSELWGAPKSRGPPPGLSGKGG--GGASLANGWSSNI  1208
             +    +                  +SELW APK RGPPPGL  K G  G   ANGWSS
Sbjct  1222  VSGGGSASNNGGGKNSASTWETNSSELW-APK-RGPPPGLPAKPGSSSGGQTANGWSSLS  1279

Query  1209  AGSVP-----WAGASGSAAQRNSGNWGVSQSQWLLLRNLTAQIDGSTLRTLCMQHGPLQS  1263
             +         W G +  A+ +        S WLLLRNLT QIDGSTL+TLCMQHGPLQ+
Sbjct  1280  SSGRWSTGQGWPGPNQPASTQPG-------STWLLLRNLTPQIDGSTLKTLCMQHGPLQN  1332

Query  1264  FHLHLHQGFALAKYSSREEATKAQTALNNCVLGNTTILAENPTDWDASTLLQSIANQQGA  1323
             FHL+L+ G ALAKY+SREEA+KAQ ALNNCVLGNTTI AE+P++ D  +LLQ ++ Q GA
Sbjct  1333  FHLYLNHGIALAKYTSREEASKAQGALNNCVLGNTTIFAESPSESDVLSLLQHLSVQGGA  1392

Query  1324  --SSAGWRTSSSKPGVTGGDTWSTGWPNNPSGVG-----LWATPSLDSNDPARATPSSLN  1376
                S+ WR         G + W         +G G     LWA       ND  R TPSS+N
Sbjct  1393  GSGSSAWR---------GKEAWGNSQLWGATGAGSTAASLWA----GDNDQHRNTPSSIN  1439

Query  1377  SFLPNDLLGGESM  1389
             S+LP DLLGGES+
Sbjct  1440  SYLPGDLLGGESI  1452
```

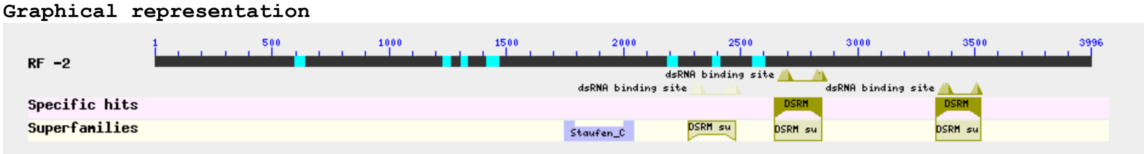
Staufen

cDNA

Contig 522, Length-1684
TTTTTTTTGGTTGGTGATTTATTTATGGTTTCTTTAGAAGAACAATATTACAATACGATAACTAACATTCTCTCCTCAGTAATATTTTTTATTACCA
TTTTTTTGGTTGATAACAACTCCTAAAAATGTATTTATCTATACACATCTATATCATTAAGCCTTATTATATAACTATGAATTAATATGCATCACAC
CATTAACAATACGTCTCAATGCACCTAAAACACAATGCTTCTAACTAATGTATTGATTAACACCCAAAACATCTTTCATTTACATTTATAATTATAA
TATATGCTGAGAATCATAAATAATTAGAATACATTTATTAGGAAAGAGGGGAAAAAAATGACAAAACACTGTGATTAGATATACATATATGAATATA
TGTAGCCAAAGAAAAATATTTAAGGAACTATTGAGCAAAAATGCGTTACGCTTAAGTTTTTTGCATATTTTAATCAAGCGAAACGCATATGCTCTAA
CATAAGCACCTAAAAACACATATAGAAATATATAAACTGAAGCCAAACAAATTAGATGAATTCCTGAACACTGAAAAATTTCAAAAAATTTTTAAAA
GCCAAAACAAATGAGATACAATATAGAATACTATCGTTATACAATATAATAGATAGAAAAACTAACACCAACCATATCCACTAATAGGATTATATAA
TGGACAATCATACATAATAAACAAAAAAAATGTTTAATACTTCAGGAATCTAAATCTAAATAGTGTATGTATTTTTTTTTTCTGCATTCCTACAGCAA
TGATAATGTTTATGTAAGCTTAATAAAGTTTTATTTGGTATCATACTGAAACACAAACACACACTGCGTATGTTCGTGTATATATACATAAAATAAA
TATATGACAAACATACATATTTTTGTACACAAGTATAGACGTAACGCATGTATAATTTTATAATTACTTTATAATTTTAAAAATGAAACTTACATAC
CCATTAACAGCTGTATTTGAAAAGAAAGAGTGATAGTGTTAACAAATACGCCTATTACCTAATAAAAACCCTACCATTATTTTTTTTCCATTTTTTT
TTTCTTTATGTTTGAAACATAATACTTCCAGTTGCGCAATCATACCTAATAATAAATAAACACTTAAAAATGTGTCCCTAGAGAAGATGATGTTATA
AGTGCATTTAGAGACTGCTATCGAATATAAATAGCGAATGTTCCAAAATAATGAAACTTTAATGATAAAAAAAAAAAATTAAATTAAATTAAATAAAC
TTTGAGGCCCACTTGTAACAGTCGGTTTAGCAGAAATAATATTTTAATAATTTTTCTGTATTTTTTTTTTTTTTTAAATTAATAAATCAAAACCGCTC
CAGCTAGAAGTGGCCCACATTTCAAATTTTCTTCTCTAATCAGTCATTCATTTACAAAATATTTTTTAAAAAATATTAAAACAAAAATAACAATAAA
AAAAAAAAGAAGACCGTTTGTGACAATATAACAAATACTTCAGATCTCTGACATCTGATGTACTACGTATTCAAATAATGCTATAAACTCTCAATAT
ATCGATTCTGTTAAATAAATTTTAAAACTAAAGGAATCTTTTTTTATACAGTATTAAATTAAAACAAAACTTTACTTTCCGTTAACATAAGGGGTTT
CTTCCTTGAGATCAATATACAACCCAGACTCAGGTTGCGGTAGGGTCGGTGGCGCAGTATGGGTATTCTCTTTCTTTGAGGGGGTGATGGAATCGAG
GCCAACTTTGGCAAGGGACCTGAGTGCATTCATTGCAGCCTGCTCGTGGGAGAGTTCAATAGTTGCACCCCCTCCATGACAAACTTGAGGTGGATCT
GTGTCAAGGGAGACTAGTGTTAAGAACTCAGAATGGTTTCCCTTGGGAAAATCTGAGAACTGGACTTGAAAGTTCAGAACTGAAGCTAAATGCGCTA
ACTGTTCTGTTGGTCGAACTGAAAAATTTTTGGCCTGGCGTTTAACCCCAGCAAGTGCATCAGCAGTTCGTGAAGACCCTTCATTCAAATATTCCTT
AGCGATAGTTTCCTTAATATTGACACCTGAGCCCGGCTTGTGGCCATTATTTGTAGCTTTATTGTCCTTCATTAGGATGAGGCCAGGAACCAACTGT
CGTCCGGAGACACCACCTTCAGGTTTTTCATCGCTGAACCTAACCTTATGTTGTTGTTGAGAAGGTTGCTGCTGAGACGGTTGGGTCTGAGTTGGTT
GAAGAGGATGACCAGTACGAGATAATGTTGGACTAGAGTAACCTAACTGCTGTAATAGACCTTCCGCAGCATTTCGCTTAGCTACTTTTTTATTTGG
TCCAGAGCCGATATACGAATGATTACCTACAGTTACTTCCATAATGAACTCTCTCCTTCTTGGAAGACCTCGTTCTTCCCGTAAAGTATAAATGGGT
TCTCTTTCCTTCTTAGCTTGTTGAATTTGAATTAATCGTGAAATTGGATTCATTTCATCACTTGTATCCCTCGATTCATCATACTTGGCTTCCTGTG
AAACAGGTGTTTCCTTTATAATATTCCTTGATTTCTTTTTTCCAGTGGTAGCCTTACGCTTAGCTCGAGCGAGAGCAGAAGCGCAACCAGGAGCTGT
AACAGTAGTGGTCGGTGGGAGTGTTCTCAGCTGTTCAAGCATCTTTTCTGCAGCTCTCTTCTTAGAAATCTTCTTTCCATTCCCTTCTCCGGTTGTT
TCAAACTTATCCCCAACACAACACCTTGTAACAAAAGTTTTCATGTGAGGTTGACCTTTTTCAGAAACAACTTCAAAGAGAACAGGGAGGTTCCTTT
TTAATGCTATTTCATGGACCAGAGATATTGGAGACTTAAGTTCTGATTCTTGTTCAGTCAATCCACCACAAATGGTAGAATTGGCTGGGCAAGTGTC
TTCCAAGGGCAAGCTATGGAGGTGCTGTAAAGCCTTAGATGCAGCATCATGTCGTGCAGCTTGCGCTGAGTTACCTTCTCCAGTAAACTCTCGTTCA
CCAACTTTCACAACCACCTTATAGAGAGGCGCCTTGTTTCCAGAGCCATAGAACCTTGGATCAACACCTCGATTGAAACCTCTAAAATTTGCTTTTC
CTCTACCATAAGGAAAAGAAAACTGGTCATTGTAAGGCTGAGTGTACATTCCCCTATGCGAATGATGGTATGTATAATGATGGATTTGTTGAGGTG
ATTCAACTAGCGTATACACAGTAGGTTCCCCTCTCTTCATTGCTAATGCGTTTAATTCAACCGTTGGAGTAATATTACTTTTAGAACCTCGAGATTG
CCTAGAGGATTTAGGTGGAGGGTGTTTGTAGGTAGTACTTTTGAGTGCAATGCCGGCAGCAGCATGTTGAGCTTTTTTTATACTCGCCCCTTCTGCA
GTATACTCCTCATTGCCAAGCTTCAAGGTGACGGTAAACTTCTTTTTATGAGGGGGGCCCATCTTCACTGGTCAAACGGTACTGATGTTCAATCTTAT
TGTACCTGGCAAGTTCATTGACCAGGCACATCGGCGTTTTTTCTTTTCCGAGGAGAGCCTGGCTGCCATCAGCTGTGGCAGGTTCCTCTTGGTTGGA
TTGATCAGCAGCATTACTTTCACTATCTGTCCCAACTGGAACAGGTGCGCTACCTGCCATTACTAAATGGCCTGTAATACATCCATTCTGATCAAAT
AAATCCTTTATCTCATCATAGTTTATCCTTTTTTTATATACTGTGGCTTCACATAATGTGGAAGTAAGGAGAAACAGGATAAGATGATATCCTGTAA
TCCTCTGACCCACACTGTGGGTTTGGCGGAGATTTGGATTAACCATTTACTTCATAAAAGTCGAAACCTTGAGATAAGCGTAGTCTCTATCCGCGGC
AAATATTAACTTCTGGGAATCACAGAGAGAAAAAAGTATCCTGCGTTAAGTTCGCTCGCTGGGGCTTATCTACGCTAAGCGACCGTCAGCTCCCAAC
CGCTCGCATAACCTCCCT

Protein 734 (AA)
MVNPNLRQTHSVGQRITGYHLILFLLTSTLCEATVYKKRINYDEIKDLFDQNGCITGHLVMAGSAPVPVGTDSESNAADQSNQEEPATADGSQALLG
KEKTPMCLVNELARYNKIEHQYRLTSEDGPPHKKKFTVTLKLGNEEYTAEGASIKKAQHAAAGIALKSTTYKHPPPKSSRQSRGSKSNITPTVELNA
LAMKRGEPTVYTLVESPQQIHHYTYHHSHRGMYTQPYNDQFSFPYGRGKANFRGFNRGVDPRFYGSGNKAPLYKVVVKVGEREFTGEGNSAQAARHD
AASKALQHLHSLPLEDTCPANSTICGGLTEQESELKSPISLVHEIALKRNLPVLFEVVSEKGQPHMKTFVTRCCVGDKFETTGEGNGKKISKKRAAE
KMLEQLRTLPPTTTVTAPGCASALARAKRKATTGKKKSRNIIKETPVSQEAKYDESRDTSDEMNPISRLIQIQQAKKEREPIYTLREERGLPRRREF
IMEVTVGNHSYIGSGPNKKVAKRNAAEGLLQQLGYSSPTLSRTGHPLQPTQTQPSQQQPSQQQHKVRFSDEKPEGGVSGRQLVPGLILMKDNKATNN
GHKPGSGVNIKETIAKEYLNEGSSRTADALAGVKRQAKNFSVRPTEQLAHLASVLNFQVQFSDFPKGNHSEFLTLVSLDTDPPQVCHGGGATIELSH
EQAAMNALRSLAKVGLDSITPSKKENTHTAPPTLPQPESGLYIDLKEETPYVNGK

Graphical representation



Comparison with *Cylas puncticollis* Staufen

```
Query  132  KEKTPMCLVNELARYNKIQHQYRLTGEQGPAHKKIFTVTLKLGKEEYEAEGPSIKKAQHS  191
            KEKTPMCLVNELARYNKI+HQYRLT E GP HKK FTVTLKLG EEY AEG SIKKAQH+
Sbjct  98   KEKTPMCLVNELARYNKIEHQYRLTSEDGPPHKKKFTVTLKLGNEEYTAEGASIKKAQHA  157

Query  192  AAAKALTKTEFKHPPQKTIINRPGNRPSNPGVVTPTVELNALAMKRGERTVY-VVESGGA  250
            AA  AL  T +KHPP K+     G++ +    +TPTVELNALAMKRGE TVY +VES
Sbjct  158  AAGIALKSTTYKHPPPKSSRQSRGSKSN----ITPTVELNALAMKRGEPTVYTLVES---  210

Query  251  PPHQGYMNQPSYYPRHNINYPQ---AQPRYGYD-ARRNIRG-HYTYDNRYYGQFRPGAPH  305
             P Q  ++  +Y+  H   Y Q    Q  + Y   + N RG +   D R+YG
Sbjct  211  -PQQ--IHHYTYHHSHRGMYTQPYNDQFSFPYGRGKANFRGFNRGVDPRFYGS-------  260

Query  306  HNPGDPYTVRLRVGEREYPGHGYTVQAARHDAAAKAIEHIKQLGGVEGADSVPANEN--G  363
             N    Y V ++VGERE+ G G + QAARHDAA+KA++H+  L      D+ PAN    G
Sbjct  261  GNKAPLYKVVVKVGEREFTGEGNSAQAARHDAASKALQHLHSL---PLEDTCPANSTICG  317

Query  364  QMSDQQQAQSSLTNDVNSELKSPISLVYEIALKRNLNVTFEVLSEKGPPHMKVFVTQCRV  423
             +++Q+           SELKSPISLV+EIALKRNL V FEV+SEKG PHMK FVT+C V
Sbjct  318  GLTEQE-----------SELKSPISLVHEIALKRNLPVLFEVVSEKGQPHMKTFVTRCCV  366

Query  424  GS-FVAEGEGNGKKISKKRAAEKMLEELSKLPPLPNV------ANISQMKRKRVTNKKKT  476
            G  F   GEGNGKKISKKRAAEKMLE+L  LPP   V      + +++ KRK  T KKK+
Sbjct  367  GDKFETTGEGNGKKISKKRAAEKMLEQLRTLPPTTTVTAPGCASALARAKRKATTGKKKS  426

Query  477  RNLIK-------VNVDKNSETQEEINPISRLIQIQQANKEREPVYTVLEERGAPRRREFV  529
            RN+IK          D++ +T +E+NPISRLIQIQQA KEREP+YT+ EERG PRRREF+
Sbjct  427  RNIIKETPVSQEAKYDESRDTSDEMNPISRLIQIQQAKKEREPIYTLREERGLPRRREFI  486

Query  530  IEASVSGHSCTGVGPNKKVAKRNAADALLTELGYS--TTVPKPQTQTKSEKEPEPSTSSD  587
            +E +V  HS  G GPNKKVAKRNAA+ LL +LGYS  T          ++ +P     S
Sbjct  487  MEVTVGNHSYIGSGPNKKVAKRNAAEGLLQQLGYSSPTLSRTGHPLQPTQTQPSQQQPSQ  546

Query  588  KNRKVTFVEEKAEQPPTQSIGGSGGRQLVPGVLLVTEQNA---------------GFSKP  632
            +  KV F +EK E       GG  GRQLVPG++L+ +  A                 +K
Sbjct  547  QQHKVRFSDEKPE-------GGVSGRQLVPGLILMKDNKATNNGHKPGSGVNIKETIAKE  599

Query  633  KTENTEQNKPVQMQPAPMKNPPPGVRSKDQLLYLAQLMNVHVQFSDFPKANHEMYLTLVS  692
                        +     +     VR  +QL +LA ++N  VQFSDFPK NH  +LTLVS
Sbjct  600  YLNEGSSRTADALAGVKRQAKNFSVRPTEQLAHLASVLNFQVQFSDFPKGNHSEFLTLVS  659

Query  693  LGTNPPQVCHGEGPTTEASHEKAALEALKVLSELGLD  729
            L T+PPQVCHG G T E SHE+AA+ AL+ L+++GLD
Sbjct  660  LDTDPPQVCHGGGATIELSHEQAAMNALRSLAKVGLD  696
```

Clp1 homolog (kinase)

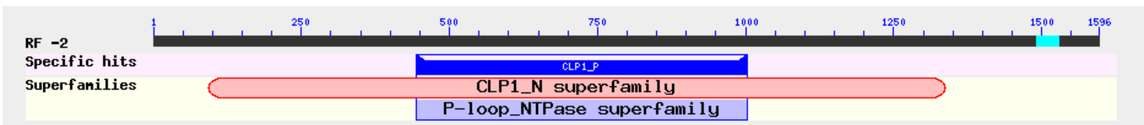
cDNA

```
Contig 10636, Length-1594
AACAACATAAGTTAAATAGTACAAATTTTACTATCAAAAATAATAAAAATACAATATTGTATTAACAAGAAAATAAAAAAATTTTAATGGCTGTCCA
TAAATTGTACGTCTGAAAGAATTAGGAGGGTGTCTGGGAGCGGCCTCGGCTGCGGTGAAAGTACAGTCAGTGTTTGGTTTTCGACATCTACGTTTGT
CACGCATATATAGCCAGCAACATTGGTTGTTATCACATCTTCATCAGTGGATGTAGCAAAACTAACAGCAAGAATGTGATGAAGGAGATTGGGACCA
GGCTGAACTGCAACAAGTTTGGTCATGTTGTCCTCAGCTTTCATACCGAGAGGCATGCAAGAATCTGGCAGAGGGGGAGCCCCTATTTTGTAAACCC
TCAGGTCAGACCATTTGACATCAAAACTGTGAGGGTGAAGTGGGGTCTTTCCACCGTAAAAATATTCCCTAACTCTGGAATCACGAGAATTCATTCT
AAATTTTTTTGTACGTTCTACTACACCACCACTTTTCGGAAGGAGGACCACCTTCACAAAATTAGGCATATCCCTCACAAGTTCATTATAAAGCTTT
```

TCCTGATCCAAAACAAGAATTATATCGACTTCGAAAGCTTGGGCAGCATGTGTTAGATGGCGATAGCCATCTCCCTTCACCCATCCACATGTATTAA
TAATTATTCCAGATACATTAGATTTTCTATTAGTTTGTAGCCTTTCATGTACAACTTCAGCCATTTTATTAACTAAAAGGTTAAATGTTTGGGTATT
TTTCTCCATACTAGTGTGGCCATAATGATAAACCAGGGGGGCATTGTGAGGAAATCCTTCCTCCACATCAGCCGGTCTTTCTACTAAAATAGCACCT
ATAGCTCCTGGAATTGATATTTGTCCTTGACCAACATCTAAGTCAACATAAACGGGTGCTCTCCCCATTCGAACAGCATAATTTAGAAGGATACGAC
AAACGGTAGACTTTCCCACATCAGTTGGGCCAACTACCATTGTAATTGGACCCCGCTTTACTTCACTTTCATTTTCGGCTTTCATCCTGAGTTGTTC
AAGGCAAGCATGGCAATTTAAATAAAACATCATTGGAGTTTCTTTTGCAATGTAACATAAATTTCTTTTTCCTCGAAGTTCTAACGTGCATCCTTGC
CATGTAAAAACTGCCAGCTTAGCTCCTGCAGTAAATGAGTACGGCTTCCCTTTAACTAATTCAGTGCCAAAAATCTCAGCCAAACCAACTTTAAGCT
CTAAGACAATAGTATCTGATTTATTTTCAACTTCAAATCTCAGTTCTGCATCAGGTTCTAGAACAATTTCGTGCACGTCTGGTTCCTTTTCTTCACT
CATCTTTTTTAATACTGTCTTATGATGTCAAGCTCGATTACTACATACCAACTACAGAAGCGGGCGATTTTAACAACAAACCTATTCACTTTAGTAA
GGTAACAATAGGTTATAGAGCATAGAAATAGGGAATACAACGAGCAATAACTAGAATAGTGAAAGGAACCGAAACTTGACAGTTGGGGACATGTTTC
GTCAACTGTAGCAAATAAAGGTAACATTTGATGATCAAAAT

PROTEIN 425 (AA)
MSEEKEPDVHEIVLEPDAELRFEVENKSDTIVLELKVGLAEIFGTELVKGKPYSFTAGAKLAVFTWQGCTLELRGKRNLCYIAKETPMMFYLNCHAC
LEQLRMKAENESEVKRGPITMVVGPTDVGKSTVCRILLNYAVRMGRAPVYVDLDVGQGQISIPGAIGAILVERPADVEEGFPHNAPLVYHYGHTSME
KNTQTFNLLVNKMAEVVHERLQTNRKSNVSGIIINTCGWVKGDGYRHLTHAAQAFEVDIILVLDQEKLYNELVRDMPNFVKVVLLPKSGGVVERTKK
FRMNSRDSRVREYFYGGKTPLHPHSFDVKWSDLRVYKIGAPPLPDSCMPLGMKAEDNMTKLVAVQPGPNLLHHILAVSFATSTDEDVITTNVAGYIC
VTNVDVENQTLTVLSPQPRPLPDTLLILSDVQFMDSH

Graphical representation



Comparison with *Cylas puncticollis* Clp1 homolog (kinase)

```
Query  9    QKNEKVYLTLKSGLAEVFGTELVKGKTYEFVSGAKVAVYTWQGCTIEVKGKTDVIYTAKE  68
             K++ + L LK GLAE+FGTELVKGK Y F +GAK+AV+TWQGCT+E++GK ++ Y AKE
Sbjct  26   NKSDTIVLELKVGLAEIFGTELVKGKPYSFTAGAKLAVFTWQGCTLELRGKRNLCYIAKE  85

Query  69   TPMVIYSNCHAALELMRIEAEKDN--KRGPIAMIVGPCDVGKSTVCRILLNYAVRMGRRP  126
            TPM+ Y NCHA LE +R++AE ++  KRGPI M+VGP DVGKSTVCRILLNYAVRMGR P
Sbjct  86   TPMMFYLNCHACLEQLRMKAENESEVKRGPITMVVGPTDVGKSTVCRILLNYAVRMGRAP  145

Query  127  IYVDLDVGQGSVSIPGTVGALVIERPAGIDEGFSQEAPLVYNFGHKSPGHNTKLYRMIVD  186
            +YVDLDVGQG +SIPG +GA+++ERPA ++EGF   APLVY++GH S   NT+ + ++V+
Sbjct  146  VYVDLDVGQGQISIPGAIGAILVERPADVEEGFPHNAPLVYHYGHTSMEKNTQTFNLLVN  205

Query  187  QLATIVKERLEVNKKTRQSGVIINTCGWIKGEGYKQILQSGKSFEVDVIMVLDQERLYNE  246
            ++A +V ERL+ N+K+  SG+IINTCGW+KG+GY+ +   + ++FEVD+I+VLDQE+LYNE
Sbjct  206  KMAEVVHERLQTNRKSNVSGIIINTCGWVKGDGYRHLTHAAQAFEVDIILVLDQEKLYNE  265

Query  247  LVRDMPNYVKIVFLQKSGGVVERSKQARSEARDQRIREYFYGTPKNSLYPHSFDVKFSEI  306
            LVRDMPN+VK+V L KSGGVVER+K+ R   +RD R+REYFYG  K   L+PHSFDVK+S++
Sbjct  266  LVRDMPNFVKVVLLPKSGGVVERTKKFRMNSRDSRVREYFYGG-KTPLHPHSFDVKWSDL  324

Query  307  KIFKIGAPALPDSCLPLGMKAEDHLTKVVPITPNPGILHHLLGVSFAEKEEDDIILSHVA  366
            +++KIGAP LPDSC+PLGMKAED++TK+V + P P +LHH+L VSFA   ++D+I ++VA
Sbjct  325  RVYKIGAPPLPDSCMPLGMKAEDNMTKLVAVQPGPNLLHHILAVSFATSTDEDVITTNVA  384

Query  367  GFVCVSNVDLERQTITLLSPQPKPLPNNMLVLSEIQFMDSH  407
            G++CV+NVD+E QT+T+LSPQP+PLP+ +L+LS++QFMDSH
Sbjct  385  GYICVTNVDVENQTLTVLSPQPRPLPDTLLILSDVQFMDSH  425
```

Elp-1

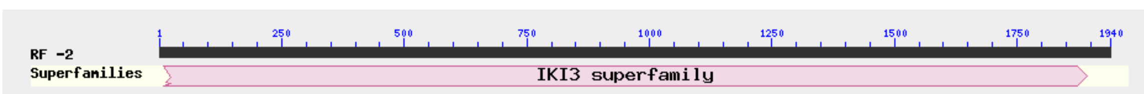
cDNA

Contig 10864, Length-3825
CACAACAAAATACTACTTACCTCTGAATTTATAATTTGTAATATAACTAGTTTGTGTGATGATGTTGTCACAAGTAGATAAGGATCTCTTAAAATAA
ATGAAGTGACTTTGCCAATAATAGATGATAAGCCAAGAGGAGTATTATTCAAATAAAGTCTAGATAATTCTGATAGACAAAAGAGACCTGTAGGATG
TACTTTAATATCATAACAGGGTTCAATCAAGGCTAAACTTTCTGTCACCATCATTGACTTTGTATTATATTTAAATAATTTTCCGCATGAAGTTTGC
ATAATTATTTCTTCACAAATTCTATTAGAAGTAATAATTCTTCCGTTGAAAAAAAAAACATTACTTAATTCAATTTCATCTTTGCTAATATACAATT

GATGAATAAGTGTCTCATTAAATTCATTTGAATAGCAGAAGACTAGGATATCAGCACTGACCCACTGCCAGTGAGAAAGTGATAACATTTCAATTGA
GAAATTGTGTTGTTTTACCAGACCAGGCATAATCAGAGGTTCCGAGATTGGGTCTTTGAAAATTAAAATATCACCACTATACAAAATGACACATAGC
ATATATGGAGAACATTCTTCCCTATTTACTGGAGCAAAGAGAATATGATTCACTGCATCAGAACACTTTAGTTTATATGAAGACATTGGTGGAGGAA
CTACTGCTTCTTTGAATGAGGTAAGCAAAATTGAATCATTGTCTATAACACCAACATAAGATTTTTCTTCACAATAGGAATAATAACTGTCATTAAC
AACAAATGCCCATTCCAGTATCTGTAAAAGTCCTCCCTTACACAGGGCAATCATTCTGTTTGGGCATTCCTTGTCCCAGTAAACTTTAATTGATAAG
GATGTGAATTTAAACTTCTGTTTTAAATACCAATGATAATTTCCAACTGTCCATAACATTACTTCTTCCTCTGAATTATCTTGGTTTTCACAAATTG
TACAAAGAATTGTTGAAGCTTCATTCCACATAACTTCCTTAACTTTGATATTAGATGGTAACGTGAAATCTCCATGCTTCAAACCATTTTTTTCCAA
AAATGAAATAACTTTCTTATTAGGTAACTGTTGGGGAACTGTGATCAGACTTCCACTAGGCTTCCAGCAGATAGGTTCATTTAAACCTTGAATATCT
TCGCTAGTGCTTTGCAAAATTCCTTCAGAATTAAAAATTTTAACCCTTCTAATATTTGTATGAGAGCACCAATAACTAACTGCAAATAAATTGCTAT
CTCCCCTCCAAGTCACATTTGTAGTATGGTCATCATCTTGATGTGGAGAATGAATGACTGGTTTTGATAATGCAGCATTTTTACCTTCAGAACCATG
AAACTGAGTTTCTTTTTTACCCCAACCAACACTAATCAGTTCATTTTCTCCAAGAAATGTGTCGCTTAAATCTATCTCTGCTAAAATAAAAAATTCT
GAATTTATTAAAATTAGTAAGTTTGAGGTGGAAACAACAACAGCAAACTCCTCATCTGGGCTCCAGTGAATCTTTAAAACATCAATATTTAAATGAT
GACAACTTCTGCTATGCTGCACACCTCTATCATCTAAAAGAAATGAATAAATTTCCCCTGTTTCAAAACATATATAAACTGTCTTGTACAAATCCCT
ATACCACATTTCTCTAATTTTAGGAACATTTCCTGAAAATAAAACTCTCAAATCATAATATTCAAGAACTTTTCCATCCGGTAATTCGACGCAATAT
AAAATATTATTAGACAAAACAAAAACATTCTTAGTTTTAGAACACAGATCATTGTCCTGCAAACTTTGACCGCTAATAATTATTGTAGAATCTACAA
AAACCTCACGATGAATTTTAAAATTTAATAACAATTTTAAGTTTCTCATATTTGATTTCAAAAGATCTGGGAACAAAATGATACCAAAAGGTTAC

PROTEIN 639 (AA)
MRNLKLLLNFKIHREVFVDSTIIISGQSLQDNDLCSKTKNVFVLSNNILYCVELPDGKVLEYYDLRVLFSGNVPKIREMWYRDLYKTVYICFETGEI
YSFLLDDRGVQHSRSCHHLNIDVLKIHWSPDEEFAVVVSTSNLLILINSEFFILAEIDLSDTFLGENELISVGWGKKETQFHGSEGKNAALSKPVIH
SPHQDDDHTTNVTWRGDSNLFAVSYWCSHTNIRRVKIFNSEGILQSTSEDIQGLNEPICWKPSGSLITVPQQLPNKKVISFLEKNGLKHGDFTLPSN
IKVKEVMWNEASTILCTICENQDNSEEEVMLWTVGNYHWYLKQKFKFTSLSIKVYWDKECPNRMIALCKGGLLQILEWAFVVNDSYYSYCEEKSYVG
VIDNDSILLTSFKEAVVPPPMSSYKLKCSDAVNHILFAPVNREECSPYMLCVILYSGDILIFKDPISEPLIMPGLVKQHNFSIEMLSLSHWQWVSAD
ILVFCYSNEFNETLIHQLYISKDEIELSNVFFFNGRIITSNRICEEIIMQTSCGKLFKYNTKSMMVTESLALIEPCYDIKVHPTGLFCLSELSRLYL
NNTPLGLSSIIGKVTSFILRDPYLLVTTSSHKLVILQIINSEVSSILL

Graphical representation



Comparison with *Cylas puncticollis* Elp-1

```
Query  57   NIREVRYLQYLSIHQVLFITND----FDLIIYNGNSKEFVVVYSSPIKIESVAWNPSETF  112
            N+ ++R + Y  +++ ++I  +    +  ++ +   +     +   I +  + W+P E F
Sbjct  72   NVPKIREMWYRDLYKTVYICFETGEIYSFLLDDRGVQHSRSCHHLNIDVLKIHWSPDEEF  131

Query  113  LILCDTEENLYGLTFECNIGFKEYRFIETCKSNLHDPVPNS---VYVGWGSHNTQFKGSK  169
             ++  T   L  +  E         F    + +L D         + VGWG   TQF GS+
Sbjct  132  AVVVSTSNLLILINSE---------FFILAEIDLSDTFLGENELISVGWGKKETQFHGSE  182

Query  170  ---KIIDEEIQEIPQQYSPPI--ISWRGNGEQFVINFWKEER--RYLKLFDSTLRPLYQS  222
                  + + +    P Q        ++WRG+    F +++W      R +K+F+S      S
Sbjct  183  GKNAALSKPVIHSPHQDDDHTTNVTWRGDSNLFAVSYWCSHTNIRRVKIFNSEGILQSTS  242

Query  223  ESYPNLLALTSFKQQGNFIACATIHNNQNKIVIFEKNCLVKSEFLVPE-VKGVIRNILYH  281
            E    L      +K  G+ I         N+  I   EKN L   +F +P   +K  ++ ++++
Sbjct  243  EDIQGLNEPICWKPSGSLITVPQQLPNKKVISFLEKNGLKHGDFTLPSNIK--VKEVMWN  300

Query  282  PFLHIL-AIYSLEEAGNGYINIFVNSNAKWYLKQQLLYVTEKQVKCLNWADSLEN---SI  337
                IL   I    ++       + ++     N   WYLKQ+   + T    +K + W      N   ++
Sbjct  301  EASTILCTICENQDNSEEEVMLWTVGNYHWYLKQKFKF-TSLSIK-VYWDKECPNRMIAL  358

Query  338  CKLNILTSSSIERIVFRMVINRSPNNGTVAVVDGSIVHFTFFDSEIVPPPMSSAFLKFDK  397
            CK  +L        +V     +        V V+D    +   T F    +VPPPMSS   LK
Sbjct  359  CKGGLLQILEWAFVVNDSYYSYCEEKSYVGVIDNDSILLTSFKEAVVPPPMSSYKLKCSD  418

Query  398  PINRIQFHP-NLPQC  411
             +N I F P N  +C
Sbjct  419  AVNHILFAPVNREEC  433
```

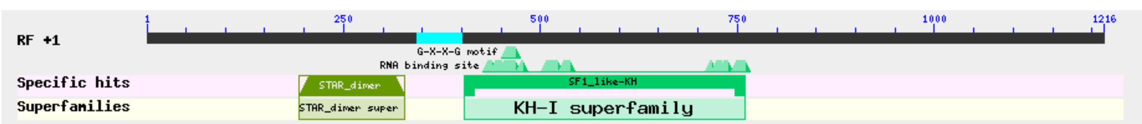
GLD-1 homolog

cDNA

Contig 3261, Length-3219
CCGGAGCCCTGCTGTTGTGTTGTATTTGGGGTCATTCTATTCCGGCCATAACTTATTTTTTGTTTTTACTTGATTCGATTCTAACTGCCTTGTTCCT
CTCCGGCACCGGGACCCCACCGCTTAAATCTTTTACCCCGTATCAAGCTCTCCGCGATCAGGAAGGTCGAAACGATGAGCGATCAAACGGGAAATTC
AACGCAGTCCATAGCAGATTACCTGGCTCAGCTGCTCAAGGATCGCAAGCAAGTGACGGCGTTTCCGAATGTCTTCATGCACGTCGAGCGTCTCATA
GATGAAGAAATCGGGAGGGTGCGGTCATCGCTCTTCCAGGTTGACGGCGTGAAGAAGGAACCTCTGGTACTCCCAGAACCAGAAGGAGAAACGACGA
CTTTGACCGAAAAAGTCTACGTACCTGTTAAAGATCACCCAGACTTTAACTTTGTCGGAAGGATCTTGGGACCAAGAGGAATGACAGCCAAACAGCT
CGAACTGGAAACAGGCTGTAAAATAATGGTGAGAGGAAAAGGCTCAATGAGGGATAAAAAGAAGGAAGAAGCAAGTAGAGGAAAACCAAATTGGGAG
CATTTAAATGAAGATCTGCACGTTCTGATTACAGTAGAAGATACTGAAAATAGAGCTAAAGTCCGCCTTGCCAGGGCAGTTGAAGAAGTTAAAAATC
TCATCATCCCAGCTGATGGAGAAGATGAGCTGAAGAAGAGGCAGTTGATGGAGTTAGCGATCATTAATGGAACTTACAGAGATTCCAACACAAAAGC
TCTTGCTGCTGCAGCTGACGAGGAATGGCGAAGAATGGCCCTAACTGCAGCCGCTGCTGAGACTCAGCGTTTGATGTCGCCAGGAATCCCTGGGATG
GCCACGCCCCTAAGGACACCTGCTGCCAATCCTCTGGGAGCTCCTCTCATTCTGTCCCCTAGGATCCCAGTGCCTACTTCAGCAACTCCCCTCCTGT
CACCACCCGAAGCACACCCTCATCAGCTCTTGTACACTCCCTTTCCAGACTACACCAACTATGCAGCTCTCACGGGCACTCCTCTGATATCAGATTA
CCCAGGTGCAGATCATTCTGAAATGGCCCTCTCCGACGAACTTCCATGTGCCAGCATTACGGTTATTAACTTACCTGAACCAGCTCTCAGTTGGCAG
AACTACGGCCTCGCGTTCACTTCCGTAGTCAATTGATTATTGCATGTCAT

PROTEIN 342(AA)
MSDQTGNSTQSIADYLAQLLKDRKQVTAFPNVFMHVERLIDEEIGRVRSSLFQVDGVKKEPLVLPEPEGETTTLTEKVYVPVKDHPDFNFVGRILGP
RGMTAKQLELETGCKIMVRGKGSMRDKKKEEASRGKPNWEHLNEDLHVLITVEDTENRAKVRLARAVEEVKNLIIPADGEDELKKRQLMELAIINGT
YRDSNTKALAAAADEEWRRMALTAAAAETQRLMSPGIPGMATPLRTPAANPLGAPLILSPRIPVPTSATPLLSPPEAHPHQLLYTPFPDYTNYAALT
GTPLISDYPGADHSEMALSDELPCASITVINLPEPALSWQNYGLAFTSVVN

Graphical representation



Comparison with *Cylas puncticollis* GLD-1 homolog

```
Query  1    MCDQNAAASTQSIADYLAQLLKDRKQLAAFPNVFIHVERLLDEEIAKVRASLFQINGVKK  60
            M DQ    STQSIADYLAQLLKDRKQ+ AFPNVF+HVERL+DEEI +VR+SLFQ++GVKK
Sbjct  1    MSDQTGN-STQSIADYLAQLLKDRKQVTAFPNVFMHVERLIDEEIGRVRSSLFQVDGVKK  59

Query  61   EPLILPEPDGPVTTLTEKVYVPVKEHPDFNFVGRILGPRGMTAKQLEQETGCKIMVRGKG  120
            EPL+LPEP+G  TTLTEKVYVPVK+HPDFNFVGRILGPRGMTAKQLE ETGCKIMVRGKG
Sbjct  60   EPLVLPEPEGETTTLTEKVYVPVKDHPDFNFVGRILGPRGMTAKQLELETGCKIMVRGKG  119

Query  121  SMRDKKKEDQNRGKPNWEHLSDELHVLLTVEDTENRAQIKLQRAVDEVRKLLVPQAEGED  180
            SMRDKKKE+ +RGKPNWEHL+++LHVL+TVEDTENRA+++L RAV+EV+ L++P A+GED
Sbjct  120  SMRDKKKEEASRGKPNWEHLNEDLHVLITVEDTENRAKVRLARAVEEVKNLIIP-ADGED  178

Query  181  ELKKRQLMELAIINGTYRDSSAKAASASACDEEWRRVA--AAAVDAQRLLSPAIPGLAAP  238
            ELKKRQLMELAIINGTYRDS+ KA +A+A DEEWRR+A  AAA + QRL+SP IPG+A P
Sbjct  179  ELKKRQLMELAIINGTYRDSNTKALAAAA-DEEWRRMALTAAAAETQRLMSPGIPGMATP  237

Query  239  MRGPAA-PLGAPLILSPRMSVPTTAASILNGSAPPGSLLSPGDPHGLLYAPYADYTNYAA  297
            +R PAA PLGAPLILSPR+ VPT+A  +L   +PP +      PH LLY P+ DYTNYAA
Sbjct  238  LRTPAANPLGAPLILSPRIPVPTSATPLL---SPPEA-----HPHQLLYTPFPDYTNYAA  289

Query  298  LAATPLLTEYATADHS  313
            L  TPL+++Y  ADHS
Sbjct  290  LTGTPLISDYPGADHS  305
```

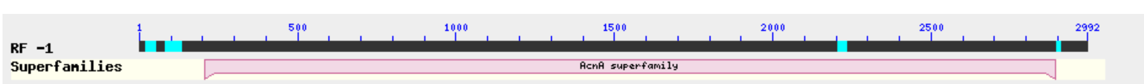
ACO-1 homolog

Cdna

Contig 3707, Length-2990
ATAATTTTCAATCCCTGTTTTATTTTTTAGATCAAGCATTATAAAAATAAAATTCTATAATTATTTTACAGTACAAAACATATTACAGTTAAAACCA
TAACCATGATTAATAATAATATGCACTATATTTATGTATACTATATTTATATACATATATAATTTATCTCATGATATTTATTGTCAATTATATGGCA
CTATTTTGAACTGAGCATCCTTCTGACCATGCAGTTGAGAATTCCTCCATGTTTGAAGTAAAGAAGATCGACTTCAGTATCGAACCTCACAATCACT
TCAAAACTCTTTCCGCTGTCAGTCTGTACAGTAATCCTTTGGTGAGGTCTAGGATCAGCAGGTAAGTTGATAGTGTACCTTTCCTTTCCTGTCAGAC
CGAGGGATTCAGCTGATTCTCCTTCCAGGAACTGGAGTGGTACAATACCCATTCCGACCAAGTTAGAACGATGAATACGTTCATAAGATTCAGCGAT
AACAGCCTTAATGCCAAGTAACAATGGGCCCTTAGCAGCCCAATCCCTCGAAGAGCCAGAACCGTAATCTTTTCCAACGATTGCAATGAGTGGTGTC
CCTTCAGCCATATATCTTTCGGCAGCATCAAAGACATCCATCTCATCTCCAGAGGGGATATGAACAGTTCTTGGAGCAGCTCGGTTGAGAAACTTAT
TCACAAGACGAATGTTAGCAAAAGTCCCTCGAGCCATAACTTCATCATTACCTCTTCTTGAACCATAAGAATTGAATTCACGGGGTGTTAAACCTTT
GGAAGCTAGGTAACGGGCCGCTGGACTGTTTCTAGCGATACTTCCAGCAGGAGATATGTGATCAGTTGTGACAGAGTCACCAAGGAAAAGTAACACC
TTGGCATTTTCGATTTTTTTAGGCTTCGGTAAATCTCTAGTCATTCCATCAAAAAAAGGAGGCCTCCTGATGTAAGTAGACTTGGGATCCCAAGGAT
ACAATTTATCATCGGGAGCATTCAGAGATCTCCAGTTATTGGAACCGTTTTCTATCCTAGCATAGACATCACGGAACATGGCTGGTATAACAGTTTT
CTTTTCAACTTCTTGAATTTCTGAACGTAAAGGCCAGACATCTCGCAGGTAAATGTTTGTACCATCGGGTTTTTTTCCTAAGGGTTCAGTTTCGAAG
TCGATGTCAATTCTGCCAGCAATAGCATAAGCAACAACAAGAAGGGGTGAAGCCAAATAATTAGCTCTGGTATTAGGATGGATGCGCCCCTCAAAGT
TTCTGTTACCACTCAGCACACCACAACAAACAAGTTCATTCTTTTCAATAGCATCGATCATCGGCTCGGGCAAAGGTCCAGAGTTACCGATACATGT
CATGCATCCATAACCGACGGTGGTGAATCCTAACTTGTCGAGGTAAGGGACTACACCAGACTCCTGAAGATAGTGGGTGACAACTCCAGACCCAGGA
GAAAGAGATGTTTTGATGTAAGGTGCGACTTTGAGCCCCGCCTCAACTGCTTTCTTCGCCAAGAGACCTGCTCCTAACATGACGGAAGGATTCGACG
TGTTAGTGCAGGAAGTGATGGCCGCGATGACGACGGAACCATGGCGTATCGTGTACTCCGAGCCCTCGAAGGCGAAGGTACCACTCGCTTCTACGGC
GGCAGGCTTCAGGCCGAACCCTTGGAAACCGACCTTATTTGTCAAAGCCTTTGTAAAGAGTTCCTTGACTTCTGTAACAGGAGCTCTATCTTGAGGC
CTCTTCGGTCCCGATACACAGCTAACCACAGTCGAAAGGTCCAGTTCTATGACTTCTGAATAAACTGGATCTTGTGCAGCATTTGAAAAGTCTCTCA
ACATTCCAACAGCCTGCAGGTACTTCCTGATAGCATTTATTTTGGCAGGATCACGGTTGGTCTGAGAGAGGTAGTAGAGACTGTTTTCATCAACTGG
GAAGAATCCAACAGTCGCACCATATTCTGGACACATGTTAGAAATAGTTGCCCTATCTGCGATAGACAATTCTGCAACACCAGGACCAAAGAATTCC
ACAAACTTACCGACCACTCCAAGTTGTCTCAAGTTCTTAGTGATTGTAAGGACAAGATCAGTTGATGTAGCATATTGATTTAATTTTCCGACCAATT
TGTAACCAATTACTTGAGGAAGGAGCATACTTATGGCCTGACCAAGCATAACTGCTTCTGCCTCAATTCCTCCGACTCCCCAACCAACAACTCCGAG
CCCATTGATCATTGTAGTATGGGAATCAGTACCGACAACACTATCAGGATACAATAGTTGTTCGGTAAATACCACTCTTGCCAAATATTCCAGATTC
ACTTGATGGACAATCCCAGATCCAGGAGGGACGATAAGCATGTTTCGGAATGCCTTCGCTCCCCACTTAAGAAATGTAAATCTTTCCTTATTTCTTT
CAAATTCAATTTCTTCATTTTTTTGCAAGGCCTTATCCACTCTAGCAAAATCAACCTGAACAGAATGATCAATGACGAGATCGGAAGGGCATATAGG
ATTAATTTTGTCTGGGTCACCTCCTAGTGATTTAACAGCATCACGCATAGCAGCAAAGTCCACAACGGCAGGCACACCAGTCAAGTCCTGCAATAGA
ACCCTGGCTGGCTTGAACGCTACTTCTACCGAACTCTCAGTTTTTCCAGTCTTTTCCCAGTTTAAGATGTTCTCTACATCGGTCTGCTTTACTTCAA
AGTTATCACAGTTCCTTACAGCTGATTCCAGAAGAACTCTGATAGAAAATGGGAGTTTTGCATACTGGGGACCAAGTTCTCCGAGGCTGTAAAACTT
ATAAGATACCCCATCCACATCTAAGGTTTTTAGAATTTTATTGTAAGGATTTGAGCCTGACATTCTTAAAGATATCCGAAAGAAAAAAAAAATGAAA
CTACACCACAAGTCGTGAACAGTTGTATACTGAACCCAAACTTTGATAAGGCGATTGAGAGAGGTAACTTTCTCTTATC

PROTEIN 893 (AA)
MSGSNPYNKILKTLDVDGVSYKFYSLGELGPQYAKLPFSIRVLLESAVRNCDNFEVKQTDVENILNWEKTGKTESSVEVAFKPARVLLQDLTGVPAV
VDFAAMRDAVKSLGGDPDKINPICPSDLVIDHSVQVDFARVDKALQKNEEIEFERNKERFTFLKWGAKAFRNMLIVPPGSGIVHQVNLEYLARVVFT
EQLLYPDSVVGTDSHTTMINGLGVVGWGVGGIEAEAVMLGQAISMLLPQVIGYKLVGKLNQYATSTDLVLTITKNLRQLGVVGKFVEFFGPGVAELS
IADRATISNMCPEYGATVGFFPVDENSLYYLSQTNRDPAKINAIRKYLQAVGMLRDFSNAAQDPVYSEVIELDLSTVVSCVSGPKRPQDRAPVTEVK
ELFTKALTNKVGFQGFGLKPAAVEASGTFAFEGSEYTIRHGSVVIAAITSCTNTSNPSVMLGAGLLAKKAVEAGLKVAPYIKTSLSPGSGVVTHYLQ
ESGVVPYLDKLGFTTVGYGCMTCIGNSGPLPEPMIDAIEKNELVCCGVLSGNRNFEGRIHPNTRANYLASPLLVVAYAIAGRIDIDFETEPLGKKPD
GTNIYLRDVWPLRSEIQEVEKKTVIPAMFRDVYARIENGSNNWRSLNAPDDKLYPWDPKSTYIRRPPFFDGMTRDLPKPKKIENAKVLLFLGDSVTT
DHISPAGSIARNSPAARYLASKGLTPREFNSYGSRRGNDEVMARGTFANIRLVNKFLNRAAPRTVHIPSGDEMDVFDAAERYMAEGTPLIAIVGKDY
GSGSSRDWAAKGPLLLGIKAVIAESYERIHRSNLVGMGIVPLQFLEGESAESLGLTGKERYTINLPADPRPHQRITVQTDSGKSFEVIVRFDTEVDL
LYFKHGGILNCMVRRMLSSK

Graphical representation



Comparison with *Cylas puncticollis* ACO-1 homolog

```
Query  1    MAESNPYNKYLKTLTVGSKEYTYYDLSALGEQYNRLPYSIRILLESAVRNCDNFSVKEQD  60
            M+ SNPYNK LKTL V     Y +Y L  LG QY +LP+SIR+LLESAVRNCDNF VK+ D
Sbjct  1    MSGSNPYNKILKTLDVDGVSYKFYSLGELGPQYAKLPFSIRVLLESAVRNCDNFEVKQTD  60

Query  61   VQNVLNWESNQNIEDGVEIPFKPARVILQDFTGVPAVVDFAAMRDAVRDLGGNPEKINPI  120
            V+N+LNWE     E  VE+ FKPARV+LQD TGVPAVVDFAAMRDAV+ LGG+P+KINPI
Sbjct  61   VENILNWEKTGKTESSVEVAFKPARVLLQDLTGVPAVVDFAAMRDAVKSLGGDPDKINPI  120

Query  121  CPADLVIDHSVQVDFARSPDALQKNQDLEFERNFERFVFLKWGAKAFNNMLIVPPGSGIV  180
            CP+DLVIDHSVQVDFAR   ALQKN+++EFERN ERF FLKWGAKAF NMLIVPPGSGIV
Sbjct  121  CPSDLVIDHSVQVDFARVDKALQKNEEIEFERNKERFTFLKWGAKAFRNMLIVPPGSGIV  180

Query  181  HQVNLEYLARVVFTGSNKPILYPDTVVGTDSHTTMINGLGVLGWGVGGIEAEAVMLGQAI  240
            HQVNLEYLARVVFT   + +LYPD+VVGTDSHTTMINGLGV+GWGVGGIEAEAVMLGQAI
Sbjct  181  HQVNLEYLARVVFT---EQLLYPDSVVGTDSHTTMINGLGVVGWGVGGIEAEAVMLGQAI  237

Query  241  TMLLPQVVGYKLYGSLGPYVTSTDLVLTITKHLRQIGVVGKFVEFYGPGVAALSIADRAT  300
            +MLLPQV+GYKL G L  Y TSTDLVLTITK+LRQ+GVVGKFVEF+GPGVA LSIADRAT
Sbjct  238  SMLLPQVIGYKLVGKLNQYATSTDLVLTITKNLRQLGVVGKFVEFFGPGVAELSIADRAT  297

Query  301  ISNMCPEYGATVGYFPADEISLSYLRQTNRSEEQVKLIEAYLLATRQMRNYE-SEENPVF  359
            ISNMCPEYGATVG+FP DE SL YL QTNR   ++  I  YL A   +R++  + ++PV+
Sbjct  298  ISNMCPEYGATVGFFPVDENSLYYLSQTNRDPAKINAIRKYLQAVGMLRDFSNAAQDPVY  357

Query  360  SQSVELDLSTVVSSVSGPKRPNDRVSVSNMKKDFLSCLTNKIGFKGFGIPENKIGTKTKF  419
            S+ +ELDLSTVVS VSGPKRP DR  V+ +K+ F   LTNK+GF+GFG+    +     F
Sbjct  358  SEVIELDLSTVVSCVSGPKRPQDRAPVTEVKELFTKALTNKVGFQGFGLKPAAVEASGTF  417

Query  420  MFDGKQYTIGHGSVIIAAITSCTNTSNPSVMLGAGLLAKKAVEAGLSVEPYIKTSLSPGS  479
             F+G +YTI HGSV+IAAITSCTNTSNPSVMLGAGLLAKKAVEAGL V PYIKTSLSPGS
Sbjct  418  AFEGSEYTIRHGSVVIAAITSCTNTSNPSVMLGAGLLAKKAVEAGLKVAPYIKTSLSPGS  477

Query  480  GVVTYYLRESGVIPALEQLGFNIVGYGCMTCIGNSGGIDENIANAIEQNDLVCCGVLSGN  539
            GVVT+YL+ESGV+P L++LGF  VGYGCMTCIGNSG + E + +AIE+N+LVCCGVLSGN
Sbjct  478  GVVTHYLQESGVVPYLDKLGFTTVGYGCMTCIGNSGPLPEPMIDAIEKNELVCCGVLSGN  537

Query  540  RNFEGRIHPNTRANYLASPLLVIAYAIAGRVDIDFETEPLGVNANGTQIFLRDIWPTRRE  599
            RNFEGRIHPNTRANYLASPLLV+AYAIAGR+DIDFETEPLG   +GT I+LRD+WP R E
Sbjct  538  RNFEGRIHPNTRANYLASPLLVVAYAIAGRIDIDFETEPLGKKPDGTNIYLRDVWPLRSE  597

Query  600  IQSVEQQHVIPAMFKEVYAKIESGSSQWQTMKAPEGKLYPWSNESTYIKKPPFFDGMTRE  659
            IQ VE++ VIPAMF++VYA+IE+GS+ W+++ AP+ KLYPW  +STYI++PPFFDGMTR+
Sbjct  598  IQEVEKKTVIPAMFRDVYARIENGSNNWRSLNAPDDKLYPWDPKSTYIRRPPFFDGMTRD  657

Query  660  LPKSKPIQNARVLLYLGDSVTTDHISPAGSIGRNSPAARFLASKGLTPKDFNSYGSRRGN  719
            LPK K I+NA+VLL+LGDSVTTDHISPAGSI RNSPAAR+LASKGLTP++FNSYGSRRGN
Sbjct  658  LPKPKKIENAKVLLFLGDSVTTDHISPAGSIARNSPAARYLASKGLTPREFNSYGSRRGN  717

Query  720  DAIMARGTFANIRLVNKFMSKSGPKTLYLPNNEEMDIFDCAQRYATSGTPLIIIAGKDYG  779
            D +MARGTFANIRLVNKF++++ P+T+++P+ +EMD+FD A+RY   GTPLI I GKDYG
Sbjct  718  DEVMARGTFANIRLVNKFLNRAAPRTVHIPSGDEMDVFDAAERYMAEGTPLIAIVGKDYG  777

Query  780  SGSSRDWAAKGPFLLGVRAVIAESFERIHRSNLVGMGLVPLQFLPGENAESLGLTGKEIY  839
            SGSSRDWAAKGP LLG++AVIAES+ERIHRSNLVGMG+VPLQFL GE+AESLGLTGKE Y
Sbjct  778  SGSSRDWAAKGPLLLGIKAVIAESYERIHRSNLVGMGIVPLQFLEGESAESLGLTGKERY  837

Query  840  NIDLPENLKPHEHITVSTDTGKSFKVLLRFDTEVDLLFYKHGGILNYMVRKML  892
             I+LP + +PH+ ITV TD+GKSF+V++RFDTEVDLL++KHGGILN MVR+ML
Sbjct  838  TINLPADPRPHQRITVQTDSGKSFEVIVRFDTEVDLLYFKHGGILNCMVRRML  890
```