ABSTRACT

BLANCO-MENESES, MONICA. Population Biology and Detection of the Tobacco Blue Mold Pathogen, *Peronospora tabacina*. (Under the direction of Jean Beagle Ristaino.)

*Peronospora tabacina* Adam. is the causal agent of blue mold or downy mildew of tobacco. The pathogen is a fungus-like organism and is a member of the Oomycota. *P. tabacina* is an obligate parasite restricted to species in the genus *Nicotiana*. Identification of the pathogen is difficult since symptoms and signs generally are found 6-12 days post inoculation if artificial or post infection if natural. The spread of the pathogen occurs through aerial long distance dispersal of inoculum and severe epidemics occur yearly in tobacco growing areas of the world. The first objective of this work was to develop a real-time Taq Man assay for the detection and quantification of *P. tabacina*. Optimization of the assay was established at a final concentration of 450nM of primers and 125nM of probe. The assay was useful for detection of the pathogen down to a lower limit of 1 fg of DNA. The pathogen could be detected after 4 days post inoculation. The real-time PCR assay was useful for the specific detection of *P. tabacina* in field samples, artificially inoculated leaves, roots, and systemically infected tobacco seedlings and could be used as a tool for regulatory agencies interested in the detection of the pathogen. A second objective was to examine the genetic structure of the pathogen in North America, Central America, the Caribbean and Europe and determine the direction of migration of the pathogen. The intergenic spacer *Igs2* region of the nuclear ribosomal DNA (rDNA)

and the Ras-related protein (*Ypt1*) gene, and the mitochondrial cytochrome c oxidase subunit 2 (*cox*2 gene) were sequenced. Populations of *P. tabacina* were characterized by high nuclear diversity, low population division and a possible mixed sexual and asexual reproductive system. Subpopulations from CCAM and EULE had the highest estimates for nucleotide diversity and mean mutation rate. Neutrality tests were significant and negative for all the subpopulations and the equilibrium model of neutral evolution was rejected. Large population size, the mechanism of dispersal, the parameters of mutation rate and genetic diversity found for the whole population indicate that this pathogen is a high evolutionary risk plant pathogen. Isolation with Migration (IM) model was used to study genetic diversity in the U.S./Central America and the Caribbean (CCAM) and the European subpopulations. Results support migration from the CCAM region, Florida and Texas into north states further in the U.S. including North Carolina. These data validate previous migration reports of the pathogen by the North American Plant Disease Forecasting Center at NCSU. In Europe estimates for the migration of the pathogen from North Central to Western Europe and both these regions to Lebanon support migration reports for the first introductions of the pathogen into Europe. Mitochondrial sequences of *P tabacina* and the *Hyaloperonspora* parasitica genome were generated using bioinforrmatics approaches and PCR methodology. One quarter of the mitochondrial genome of *P. tabacina* has been annotated and compared with that of *Phytophthora infestans* and *Hyaloperonospora parasitica*. Similarities in

direction, arrangement and number of genes and regions have been found. *H. parasitica* mitochondrial genome exhibited higher similarities with *P. ramorum* and *P. sojae* genomes than with *P. infestans*. Results from this research will be useful in understanding the evolutionary history, epidemiology and population genetics of this important plant pathogen.

Population Biology and Detection of the Tobacco Blue Mold Pathogen,
*Peronospora tabacina*

by
Monica Blanco-Meneses

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

Plant Pathology

Raleigh, North Carolina

2009

APPROVED BY:

| | |
|---|---|
| _____ | _____ |
| Dr. Ignazio Carbone | Dr. Kelly Ivors |
| _____ | _____ |
| Dr. Jenny Xiang | Dr. Jean Beagle Ristaino<br>Chair of Advisory Committee |

## DEDICATION

To dedicate a thesis not only represents to devote the final result, the final paper to somebody. A thesis represents many years of hard work and sacrifice attempting to achieve a specific goal. But this time this effort is also full of happy and amazing moments.

In my case, my husband and I moved from our beloved country of Costa Rica. In the beginning it was very hard, and it took us some time to get used to our new surroundings. After two years we were blessed with the birth of our son, Rodrigo, who changed our lives forever.

My two Rodrigos have been the support that I needed to finish my work. To big Rodrigo, my eternal support, my partner, my love, my soul mate, my friend…the one that has been with me for 11 years. Every moment that we have spent together has shown me that together we can reach the stars and the moon… the infinite. But even more important is to know that the time that we have spent together, the love and how close we feel to each other is what really counts! To little Rodrigo, my beautiful baby, a part of me, the one that makes me feels alive every day, my eternal inspiration…he has shown me that it is possible to realize everything even if you don't know anything. That knowledge is something that is cultivated and grows every day. That the best way to reach something is trying thousand times even if it is tiring or devastating. That life is given to you as an empty book and it is your job to fill it

with amazing and unforgettable moments. Also Sophie, our little Cocker Spaniel that makes me feel like the most important person in the world everyday.

To my parents and parents-in-law that have also contributed greatly to my work mainly with support, good ideas, positive energy and all the time that they have spent with us giving us their love.

I love you all…

## BIOGRAPHY

Mónica Blanco- Meneses was born to Liliana Meneses and Gerardo Blanco on an afternoon of July 22nd 1973 in San Jose, Costa Rica. She is the oldest of four brothers and sisters, Laura, Carlos and Diego. She attended a catholic school, the Maria Inmaculada Primary and Middle School and finished in 1990. Before starting college, she spent one year in Portugal with the American Field Services (AFS). With this experience, besides learning a new language, she understood how the world works and that when you are lonely you need to take advantage of every situation that is presented to you.

She went back to Costa Rica in 1991 and started at the Universidad de Costa Rica. During three years she studied Pharmacy but later she changed to Biology at Universidad Nacional were she worked on **"**Micropropagation of different medicinal plants using *in vitro* methods and genetic transformation". She received her degree in 1998. Later she started her Master in Biotechnology in 1999 at Universidad de Costa Rica and finished in 2003 working on "Resistant wild potatoes as source for novel genes mediating resistance against fungal, viral and nematode diseases" under the supervision of Dr. Roberto Valverde. Later she moved to North Carolina State University were she started her PhD in Plant Pathology graduating in May 2009.

## ACKNOWLEDGEMENTS

I would like to express my most sincere thanks to all those people who in one way or another helped me to successfully complete my dissertation. To the members of my advisory committee: Dr. Jean B. Ristaino, Dr. Ignazio Carbone, Dr. Kelly Ivors, and Dr. Jenny Xiang for their advice and support throughout my Ph.D. program. A special thank you to Dr. Ristaino for giving me the opportunity to study at NCSU, to work in her lab and to achieve this important step of my professional career. I also thank Dr. Ignazio Carbone and Geromy Moore (Dept. Plant Pathology) for his advice and help with the data analysis using the SNAP Workbench program.

I would also like to thank the organizations that funded my thesis work: the Tobacco Foundation of North Carolina and the Tobacco Commission of North Carolina. The people that supported my work through fellowships: the Lorillard Tobacco Co., the Phillip Morris Co. and the Fred Bond Scholarship.

I also appreciate all the help provided by former and current members of Dr. Ristaino's lab: Josie, Katrina, Cruz, Bo, Joslyn, Josh, Logan, Jill, Marion, and specially Dr. Julia Hu (Chia-Hui Hu) for her encouragement, support, friendship and fun time together.

At the department; Dr Ritchie's lab., especially Tammy; at Dr Opperman's lab., especially to Angela and at Dr Lommel's lab., to Tim Sit.

Also all the tobacco producers and scientist in the U.S., Central America, the Caribbean and Europe that helped me with the collection of blue mold samples.

## TABLE OF CONTENTS

## LIST OF TABLES

### CHAPTER 1



### CHAPTER 2

**CHAPTER 3**

## LIST OF FIGURES

### CHAPTER 1

**CHAPTER 2**

**CHAPTER 3**

# CHAPTER 1

# DETECTION AND QUANTIFICATION OF *PERONOSPORA TABACINA* USING A REAL-TIME PCR ASSAY


## ABSTRACT

*Peronospora tabacina* is an obligate plant pathogen that causes blue mold of tobacco. The disease is extremely difficult to diagnose before the appearance of symptoms and can be easily spread in nonsymptomatic tobacco seedlings. We developed a real-time PCR assay for *P. tabacina* that uses 5' fluorogenic exonuclease (TaqMan®) chemistry to detect and quantify pathogen DNA from diseased tissue. The primers and probe were designed using 5.8S rDNA sequences from 12 fungal and Oomycete tobacco pathogens and 24 *Peronospora* species. The PtabBM TaqMan assay was optimized and performed with a final concentration of 450nM primers and 125nM probe. The real-time TaqMan assay was assessed for sensitivity and the lower detection limit was 1 fg of DNA. The assay was specific for *P. tabacina*. None of the DNA from other tobacco pathogens, nonpathogens, or the host were amplified. The PtabBM TaqMan assay was useful for detection of *P. tabacina* in field samples, artificially inoculated leaves, roots, and systemically infected tobacco seedlings. The assay was used to quantify host resistance and it was possible to detect the pathogen 4 days post inoculation in both medium

resistant and susceptible tobacco varieties. The real-time PCR assay for *P. tabacina* will be a valuable tool for the detection of the pathogen and of use to regulatory agencies interested in preventing the spread of *P. tabacina*.

## INTRODUCTION

*Peronospora tabacina* Adam. is the causal agent of blue mold or downy mildew of tobacco. The pathogen is a fungus-like organism and is a member of the Oomycota. *P. tabacina* is an obligate parasite restricted to the genus *Nicotiana* spp., and has also been reported to infect peppers, tomatoes, eggplants, and certain weeds, but these are not preferred hosts (Lucas 1980, Mandryk 1971, Lucas 1975, McGrath and Miller 1958). The airborne sporangia, which can be transported over long distances, are the main source of inoculum for disease (LaMondia and Aylor, 2001). Optimal conditions for sporangium germination are high relative humidity and temperatures ranging from 15 to 20°C but it is also known that this pathogen can withstand extreme temperatures and low humidity (Moss and Main 1988). The sporangiospores germinate and infect through the cuticle or stomates of the host plant (Svircev *et al*. 1989, Trigiano *et al*. 1985, Moss and Main 1988).

Blue mold was first reported on tobacco in the U.S. in 1921 (Smith and McKenney 1921). Before 1979, blue mold was largely confined to early-season seed or transplant beds and the disease was sporadic and dependent on cool and wet

conditions (Lucas 1975). During the 1979 and 1980 tobacco growing seasons, field epidemics of blue mold occurred and were reported across large areas of the United States and Canada causing major crop losses estimated at 252 and 95 million dollars, respectively (Moss and Main 1988). Since 1979, blue mold has become a devastating disease of tobacco fields each year and causes epidemics in both seed beds, transplants and production fields in humid and temperate regions of the southeastern and eastern U.S., Canada, Central America and countries bordering the Caribbean basin (Aylor and Taylor 1983, Davis *et al*. 1990, Davis and Monahan 1991, Kucharek *et al*. 1996, Main 1995, Main *et al*. 2001, Moss and Main 1988, Todd 1981). Blue mold is also a problem in South America (Lea 1999, Johnson 1989) and southwestern and southeastern Europe, the Middle East and North Africa (Ristaino *et al*. 2007, Rufty 1989, Delon and Schiltz 1989).

*P. tabacina* is not believed to overwinter in the temperate zones of the United States. It is assumed that inoculum is introduced each year into the U.S. from tobacco crops in Mexico and the Caribbean, or from wild tobacco in the southwestern U.S. The role of infected debris and oospores in overwintering of the disease is not clearly understood (Heist *et al*. 2002, Spurr and Todd 1982). The pathogen can also be dispersed via infected transplants. In some cases, transplants that appear healthy may actually be infected. The latent period is 10-12 days after inoculation. Commercialization and sale of tobacco transplants from state to state is

common and can increase the risk of disease spread and introduction of the blue mold pathogen into new areas.

Fungicides have been used to control *P. tabacina* for many years. Metalaxyl (N-(2,6-dimethylphenyl)-N- (methoxyacetyl) alanine methyl ester), a systemic fungicide, was the primary chemical used for the control of this pathogen. Resistance in populations of *P. tabacina* to metalaxyl was reported in 1980 in Mexico and later in the U. S. (Wiglesworth *et al*. 1988, Bruck *et al*. 1982). Alternative protectant-type fungicides are now available for disease control including Dithane DF (active ingredient, Mancozeb), Acrobat 50WP (active ingredient, Dimethomorph), Forum (active ingredient, Dimethomorph), Aliette (active ingredient, Fosetyl present as aluminium salt), Quadris (active ingredient, Azoxystrobin), and Actigard (active ingredient, 1,2,3-benzothiadiazole-7-thiocarboxylic acid-S-methyl-ester ) (Ivors and Mila 2007).

Blue mold is fairly easy to diagnose after the symptoms and signs have begun. The detection is typically based on symptomology and the presence of diagnostic clusters of sporangiospores on the underside of infected leaves (Lucas 1975, Main 1995, Main *et al*. 2001). In severe situations, *P. tabacina* may also cause systemic infections. Systemic infection was first reported in Italy by Gigante (1962) and in Australia by Mandryk (1966) (Caiazzo *et al*. 2006). Factors affecting the development of systemic infections are not completely understood (Moss and Main 1989). Systemic symptoms of blue mold usually affect entire tobacco

plantations, and can cause high economic losses. It is possible for the pathogen to remain latent until environmental conditions become favorable and plants finally express symptoms (Caiazzo *et al*. 2006). Systemic infection results in leaf curl, chlorosis, stunting, and death of the terminal bud and the abnormal emission of bottom shoots. The presence of internal necrosis is diagnostic and affects the cambium and phloem tissues. Generally there is no sporulation on the leaves, which makes identification very difficult (Caiazzo *et al*. 2006, Moss and Main 1989, Mandryk 1966).

Detection of the pathogen in symptomless plants is critical for growers and extension agents. Since sporangiospores can spread rapidly from one field to another also, chemical control application prior to sporulation is more effective than after sporulation has occurred. Molecular diagnostic tools can provide rapid and reliable detection for early disease development and provide an increased confidence level in the identification, even when the pathogen can not be isolated. A few molecular diagnostic methods have been developed for detection of *P. tabacina*. Ristaino *et al*. (2007) and Tsay *et al*. (2006) used the ITS and 5.8S rDNA sequences of *P. tabacina* to develop specific PCR primers for the rapid detection of the pathogen. Wiglesworth *et al*. (1994) used RAPD probes for the identification of the pathogen. PCR methods can be used directly on infected plant material which is useful since *P. tabacina* is not culturable. Diagnosis is possible in a single day upon receipt of the sample with PCR but symptom expression can take from 8 to 10 days.

Real-time PCR assays have been successfully applied for the detection of many other plant pathogens (Atallah *et al*. 2004, Bohn *et al*. 1999, Brouwer *et al*. 2003, Cullen *et al*. 2007, Hayden *et al*. 2006, Hogg *et al*. 2007, Hukkanen *et al*. 2006, Lievens *et al*. 2006, Qi and Yang 2002, Silvar *et al*. 2005). Compared to the conventional PCR, real-time PCR has significant advantages. This technique eliminates the requirement for post-amplification processing steps and significantly reduces time and labor, greatly increasing the efficiency of PCR testing and is an automated diagnostic system suitable for large-scale analysis. Furthermore, health risks for operators and environmental contamination are reduced (Schena *et al*. 2004).

Real-time PCR is based on the labeling of primers, probes or amplicons with fluorogenic molecules and allows detection of the target fragment to be monitored while the amplification is in progress (Schaad and Frederick 2002). In 5' fluorogenic real-time PCR (TaqMan), a sequence-specific oligonucleotide probe labeled with a fluorescent reporter and a quencher generates fluorescence at a rate directly proportional to the amount of product amplified in the reaction. The method is now being applied to a range of organisms in many different research applications, including detection and quantification of oomycete plant pathogens (Bohm *et al*. 1999, Vandemark and Baker 2003, Ippolito *et al*. 2004, Hayden *et al*. 2004, Schena *et al*. 2004, Silvar *et al*. 2005, Hayden *et al*. 2006, Lievens *et al*. 2006, Schena *et al*. 2006, Tooley *et al*. 2006, Bilodeau *et al*. 2007, Cullen *et al*. 2007, Kox *et al*. 2007).

For *Peronospora* species, real-time PCR has been used in studies for the detection of *P. lamii* (Belbahri *et al*. 2005), *P. sparsa* (Hukkanen *et al*. 2006) and *P. parasitica* (Brouwer *et al*. 2003) using SYBR green assays but has not been developed for *P. tabacina*.

With the advent of quantitative real-time polymerase chain reaction (QPCR), it is possible to accurately quantify a specific pathogen within a host plant. By comparing the amplification pattern of unknown samples against that of a range of quantitative standards, the quantity of target DNA in the unknown sample can be estimated. Where fluorescent probes are employed, QPCR assays are sensitive, highly specific, and can be performed in a relatively short amount of time (Hogg *et al*. 2007, Atallah *et al*. 2004, Hayden *et al*. 2004, Hayden *et al*. 2006, Schena *et al*. 2006, Qi and Yang 2002, Heid *et al*. 1996). The successful design of primers and probes is closely related to the specificity and the region of DNA used. The nuclear encoded ribosomal RNA genes (rDNA) provide attractive targets since they are highly stable, possess conserved as well as variable sequences, and can be amplified and sequenced with universal primers. They also occur in multiple copies (up to 200 copies per haploid genome) arranged in tandem repeats with each repeat consisting of the 18S small subunit (SSU), the 5.8S, and the 28S large subunit (LSU) genes separated by internal transcribed spacer (ITS) regions (ITS1 and ITS2). Another part of rDNA is the spacer between the LSU and SSU genes, known as intergenic spacer (IGS) or the non-transcribed spacer (NTS). The rDNA-conserved

region, allow probes and primers from one species to be used to detect the equivalent ribosomal genes from other species, genera, families, or even kingdoms (Henson *et al.* 1993, Schena *et al*. 2004). These repetitive sequences have proven useful in applied studies as targets for the sensitive detection or identification of plant and animal pathogens by polymerase chain reaction (PCR) (Judelson and Randall 1998).

The objectives of this study were (i) to design, develop and test specific PCR primers and a TaqMan probe for detection of *P. tabacina* based on the absence of shared polymorphisms with other tobacco pathogens and oomycete species; (ii) to validate the real-time PCR assay using DNA isolated from pure cultures from the laboratory, fresh and frozen diseased tobacco leaf samples from the field and artificially inoculated tobacco plant tissues and (iii) to establish a quantitative assay for *P. tabacina* detection.

## MATERIALS AND METHODS

### *Peronospora tabacina* isolates

Seventy-three isolates of *P. tabacina* collected from 11 U.S. states and 11 countries and stored in cryogenic conditions were used in this study (Table 1.1). The collection is housed in the Department of Plant Pathology, at North Carolina State University (NCSU) and is located in the Phytotron facility in a BS Level 2

containment laboratory. The *P. tabacina* collection consists of 260 isolates collected in the U.S. and 10 international countries. New samples are obtained each year via collaborators from the North American Plant Disease Forecast Center and sent by mail to NCSU following APHIS permits and guidelines. Tobacco leaves with lesions were sent inside plastic bags with a wet towel to retain humidity and maintain sporulation. International samples were sent inside a kit containing: eppendorf tubes with DMSO 10% solution, labels, plastic pipettes, gloves and ice packets. In the laboratory, the new isolates of *P. tabacina* were inoculated on tobacco seedlings to increase the amount of sporangia for subsequent cryostorage. Inoculation was performed on one of two susceptible varieties of either *Nicotiana repanda* or *Nicotiana tabacum* cv. *xanthi*. Sporangiospores were collected from the leaves and stored at -4°C on 10% DMSO for one day and later transferred to liquid nitrogen at temperatures of -210°C.

Other important foliar and soil-borne fungal pathogens of tobacco were grown on potato dextrose agar (PDA) (Table 1.2). Mycelia from cultures in Petri dishes were transferred to potato dextrose broth and grown at 25ºC for 3 to 5 days. The potato dextrose broth was then filtered and the mycelia were collected. Roughly 100 mg of mycelia were placed in a sterile 1.5-ml microcentrifuge tube for DNA extraction as described below and stored at -20ºC for future use.

**DNA extraction**

DNA was extracted using the hexadecyltrimethylammonium bromide (CTAB) procedure with some modifications in all of the experiments to optimize the TaqMan assay (Trout *et al*. 1997, White *et al*. 1990, Wangsomboondee and Ristaino 2002). DNA was extracted from sporangiospore solutions using fresh sporangiospores, fresh or dried tobacco leaf lesions (10 mg tissue), or sporangiospores washed from fresh lesions (by vortexing 10 mg tissue in water and removing tissue). If sporangiospores were no longer viable, and inoculation on the tobacco plants was not possible, DNA extraction was performed from the cryogenic spore solution directly. Frozen sporangiospores and mycelia were placed in sterile 1.5-ml microcentrifuge tubes containing 0.2 g of glass beads to which 150 µl of extraction buffer (0.35 M sorbitol, 0.1 M Tris, 0.005 M EDTA, pH 7.5, and 0.02 M sodium bisulfite) was added and each sample was macerated using a Konte pestle attached to a drill. Nuclei lyses buffer (150 µl; 0.2 M Tris, 0.05 M EDTA, pH 7.5, 2.0 M NaCl, 2% CTAB, and 60 µl of 5% sarkosyl [N-Laurylsarcosine]) was added and tubes were vortexed and incubated at 65ºC for 15 to 30 min. After incubation, one volume (300 µl) of chloroform: isoamyl alcohol (24:1) was added to each tube and tubes were centrifuged for 15 min at 13,000 x *g* at room temperature. The aqueous phase was removed to a new tube and chloroform extraction was repeated. DNA was precipitated overnight at –20ºC in 0.1 volume of 3M sodium acetate (pH 8.0) and 2 volumes of cold 100% ethanol. The supernatant was discarded; pellets were washed

with 70% ethanol and dried at room temperature. DNA was suspended in Tris-EDTA (TE; 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). A 1:10 dilution of DNA was used for the molecular experiments.

**Growth of tobacco plants**

*N. repanda* seeds were planted in soil consisting of 1/3 peat-lite and 2/3 gravel, covered lightly with sand, and misted with a nutrient solution in a growth chamber at 25ºC under a photoperiod of 12 h of light and 12 h of dark. Seedlings were moved to the greenhouse at the two-leaf stage at temperatures of 25°C and natural light conditions until inoculation. For inoculation, the tobacco seedlings were moved to growth chambers and placed in separate plastic boxes to prevent cross contamination among isolates. In some experiments, detached leaves were placed inside inverted Petri dishes with 30% water agar to preserve humidity (close to 95%).

**Inoculation of tobacco with *P. tabacina***

Isolates of *P. tabacina* were thawed at room temperature, and washed 4 times with 1 ml of autoclaved distilled water by centrifugation at 12000 rpm to eliminate the DMSO residue. Plants were moved from the greenhouse to the growth chamber. Growth chambers were set at temperatures of 22°C with a 10 h

photoperiod of light reduced to 18ºC with a 13 h photoperiod of dark and 1 h of light interruption in the middle of the dark period. Light was provided by a combination of incandescent and fluorescent lights at an intensity of 200 µmol/$m^2$/s (Thomas and Downs 1991, Ristaino *et al*. 2007). Plants were inoculated by either a) spraying a solution of $10^8$ sporangia/ml directly on the underside of the tobacco leaf, b) transferring sporangiospores from the washed suspension onto the leaf surface using a cotton swab or c) by pipetting one drop (10 µl) of sporangiospores onto the leaf. After 10-12 days signs of sporulation were observed and the leaf lesions were collected to perform the DNA extraction and PCR reactions.

**Primers and probe design**

Specific primers and a probe to identify *P. tabacina* were designed using the amplicon generated by the primers ITS4 and ITS5 of the ribosomal (rDNA) cluster and Primer Express 1.5 software (Applied Biosystems, Foster City, California). GenBank assession numbers are shown in Table 1.2. The following criteria for the probe design were used: melting temperature ($T_m$) 8-10°C higher than the melting temperature of the primers, 15 to 30 bp in length and the total number of G's or C's in the last five nucleotides at the 3' end of the primer not exceeding two. The mismatching nucleotides were positioned as close as possible to the middle of the

probe rather than at the ends while avoiding positions with secondary structures (Heid *et al*. 1996).

A total of 37 sequences from tobacco pathogens, including 7 sequences of *P. tabacina*, from the rDNA region located between primers: ITS4 (5'-TCCTCCGCTTATTGATATGC-3') and ITS5 (5'-GGAAGTAAAAGTCGTAACAAGG-3') were used (Table 1.2). A total of 18 of these sequences of other tobacco pathogens and 6 of *P. tabacina* were reported by Ristaino *et al*. (2007) in a previous study. An additional 27 sequences from other Oomycetes spp. were obtained from the GenBank database at the National Center of Biotechnology Information (NCBI) (Table 1.2) and compared using the BLAST database. The use of the BLAST searching program minimizes the likelihood of nonspecific detection of very similar isolates. A total of 64 sequences representing the major Oomycete species including: *Phytophthora* spp., *Phytium* spp. and *Peronospora*, *Hyaloperonospora* and *Pseudoperonospora* spp. (Table 1.2) were aligned using BioEdit Sequence Alignment Editor Version 7.0.5.3.

The primers and probe were designed inside the 5.8 rDNA (Figure 1.1). The probe was labeled with the 6-carboxyfluorescein fluorescent dye (6-FAM) and the MGB (Minor Groove Binder) (Kutyavin *et al*. 2000). The amplicon generated by the two primers: Ptab2F (5'-GCTGCGAACTGCGATACG-3') and Ptab2R (5'-CCGAAAG TGCAATATGCGTTCAAAA-3') and the PtabBM (5'-CTGAATTCCGCAATTCGT-3')

probe, consists of a fragment of 86 bases. All reactions were performed in 0.2 ml optical grade plates (Applied Biosystems, Foster City, CA) in an ABI PRISM 7000 Sequence Detection System (Applied Biosystems, Foster City, CA).

All real-time PCR reactions were performed in a total reaction volume of 20 µl containing 1 µl of genomic DNA (at a concentration of 10 ng/µl in specific tests), 10 µl of 2X TaqMan Universal Master Mix (Applied Biosystems, Foster City, CA), 1 µl of the 20X Assay Mix (Applied Biosystems, Foster City, CA) with the two primers Ptab2F and Ptab2R (18 µM) and the probe PtabBM (5 µM) and 8 µl of RNase-free water (Fisher Scientific, Pittsburgh, PA). The thermal cycling parameters were 50ºC for 2 min, 95ºC for 10 min followed by 40 cycles at 95ºC for 15 s and 60ºC for 1 min. The threshold line or level of detection was set at fluorescence (ΔRn) of 0.2. A cycle threshold (Ct), the PCR cycle number at which a statistically significant increase in the reporter fluorescent can first be detected, was scored at between 16 and 35 for positive results and >40 for the negative ones.

To control for successful DNA extraction, an additional universal primer set and TaqMan probe, Eukaryotic 18S rRNA Endogenous Control, labeled as VIC dye-MGB (Applied Biosystems, Foster City, CA) was used. The primer produces an amplicon of 187 bases. The 2X TaqMan Gene Expression Master Mix (Applied Biosystems, Foster City, CA) was used to improve the amplification of both targets in

a multiplex real-time PCR reaction of the more abundant target detected by the universal primer and the less abundant target detected by the Ptab primers and probe.

To perform the PtabBM TaqMan assay for quantification of *P. tabacina* the multiplex reactions were performed on separate wells on the plate. One well was used to test the PtabBM primer/probe and another well with the same DNA sample was used for the universal primer/probe.

Each series of amplifications using the TaqMan assay included a duplicate of each sample and two no template negative controls to test for contamination; DNase and RNase-free water as well as DNA from a healthy tobacco leaf. Since the negative controls are used to assess contamination, they were assessed at a higher Ct threshold than the positive controls.

It was important to determine if the detection of DNA from the pathogen could be affected by cross contamination. To determine whether the presence of tobacco DNA and coextracted host compounds affected the amplification of target DNA, DNA extracted from inoculated tobacco leaves with *P. tabacina* was compared with DNA extracted from pure sporangia. The amount of extracted plant DNA used in this experiment was 10 mg, the same as when field plants were sampled.

**Analysis of the data**

The Sequence Detection Software System Version 1.2.3 (Applied Biosystems, Foster City, CA) was used to analyze the data. This software works directly with the ABI PRISM 7000 Sequence Detection System (Applied Biosystems, Foster City, CA). The software was used to define plate standards, and negative and unknown values for the samples; it also allowed the programming of the thermal cycler reaction settings and once the run was completed, created the amplification plots where Ct values and Rn (measure of reporter signal) values could be compared. It was also used to generate standard curves and final values for Ct, quantification, standard deviation of the Ct and mean values for the resulting reactions.

**Optimization of the PtabBM TaqMan assay**

The concentration of the primers and probe were optimized. The primer and probe concentrations recommended by Applied Biosystems (Foster City, CA) of 900 nM of primers and 250 nM of probe were used and then 3 dilutions of primer and probe were tested including 450 nM, 250 nM and 125 nM of primers and 125 nM, 62.5 nM, and 31.2 nM of probe. Six isolates of *P. tabacina* (PennB, CT991, FR184, 123347, PE05 and GA995 (Table 1.1) and a non template water control were tested in duplicate on a plate). All samples were replicated 2 times on a plate and the

experiment was repeated twice. The detection Ct value and the estimated DNA concentration were analyzed as well as the quality of the curve for each of the reactions tested for this assay.

**Sensitivity of the PtabBM TaqMan assay**

To calculate amplification sensitivity of the assay a 10-fold serial dilution of DNA was quantified and tested. The 450 nM of primers and 125 nM of probe concentration were used. A 10-fold serial dilution of DNA was run at first and in each of the real-time PCR assays conducted in order to compare DNA concentrations of known and unknown samples and calculate DNA concentrations. Three different isolates of *P. tabacina* were used for the dilutions (PennB, BU061 and Waynesville1). The DNA concentration used in the reaction was determined with a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Philadelphia, PA). The serial dilution of DNA consisted of an initial concentration of 10 ng/µl diluted in a series of 1:10 from 10 ng/µl to 10 ag/µl (10 ng/µl, 1 ng/µl, 100 pg/µl, 10 pg/µl, 1 pg/µl, 100 fg/µl, 10 fg/µl, 1 fg/µl, 100 ag/µl and 10 ag/µl). A nontemplate negative control and DNA from a healthy tobacco leaf were included. All dilutions were done in duplicate on the plate and the experiment was repeated twice. There were 120 total reactions (10-fold dilution x 3 isolate x 2 replications x 2 experiments).

The results of the real-time PCR reactions were analyzed by plotting the log of template concentration of DNA against Ct values. PCR efficiency was calculated with the formula $E = (10^{(-1/slope)} - 1) \times 100$, where E is the amplification efficiency and the slope is derived from the plot of log of template concentration versus Ct. A slope of 3.32 translates into 100% efficiency of amplification.

**Conventional PCR assay**

Seventy three isolates of *P. tabacina* (Table 1.1) used to test the TaqMan assay were also tested using the conventional PCR with primers PTAB and ITS4 primers (Ristaino *et al*. 2007). Each PCR reaction consisted of the following: 1 µl of genomic DNA (1:100 dilution of original DNA extract in TE buffer, about 10 ng/µl) was added to a 49-µl master reaction mixture containing 5 µl of 10X PCR buffer (100 mM Tris-HCl, 15 mM $MgCl_2$, and 500 mM KCl, pH 8.3), 35.25 µl of sterile distilled $H_2O$, 1.8 µl of 10 mM $MgCl_2$, 2 µl of 2.50 mM dNTPs, 2 µl of 10 µM PTAB primer (5'-ATCTTTTTGCTGGCTGGCTA-3'), 2 µl of 10 µM ITS4 primer (5'-TCCTCCGCTTATTGATATGC-3'), 0.25 µl of BSA (Roche, Nutley, NJ) and 0.2 µl of Taq polymerase (Invitrogen, Carlsbad, California). Thermal cycling parameters were initial denaturation at 96ºC for 2 min, followed by 35 cycles consisting of denaturation at 96ºC for 1 min, annealing at 55ºC for 1 min, and extension at 72ºC for 2 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min.

Amplified products were separated on 1.6% agarose gels containing ethidium bromide at 0.5 µg/ml with 1X Tris-borate-EDTA (TBE) running buffer. DNA bands were visualized using UV light on a transluminator (BioRad, Hercules, CA.).

**Specificity of the PtabBM TaqMan assay**

The specificity of the TaqMan assay was tested using DNA from 6 isolates of *P. tabacina* (WI-07, OH-AD-04, OH-DL-04, PennB, BU061 and Waynesville1) (Table 1.1), DNA from 12 different tobacco pathogens, and DNA from 5 other Oomycete spp. including *Pseudoperonospora humuli*, *Hyaloperonospora parasitica*, *Pseudoperonospora cubensis*, *Phytophthora infestans* and *P. ramorum* (Table 1.2). DNA from a non-inoculated tobacco leaf and water were used as negative controls. Primer probe concentrations were 450 nM and 125 nM, respectively, and the 10-fold dilution using *P. tabacina* isolate PennB was used to compare DNA concentration of the samples tested. A total of 70 real-time reactions were performed (23 isolates, 10 serial dilutions and 2 controls x 2 replications). All the reactions were performed in duplicate and the experiments were done twice.

DNA extracted from the 73 isolates of *P. tabacina* (Table 1.1) was tested using the TaqMan assay. A multiplex reaction was used with a 450 nM of primers and 125 nM of probe concentration (concentration used for all of the experiments). A

10-fold dilution using BU061 isolate of *P. tabacina* were used to determine DNA concentrations of the samples tested. All the reactions were performed in duplicate and the experiment was repeated twice. A total of 170 real-time reactions were performed for the detection of the pathogen.

**Detection of *P. tabacina* in infected field samples, artificially inoculated leaves, seedlings and roots**

The NaOH/Tris DNA extraction method was compared with the CTAB DNA extraction method to test the quality and reliability of both methods on extraction of *P. tabacina* from infected tissue. For this experiment 3 samples were used (OH-AD-04, OH-CL-04 and for WI-07), also a chlorotic lesion, a sporulating lesion and a solution of sporangiophores were extracted. For the NaOH/Tris method a small amount of sporangiospores or leaf tissue were placed in a 1.5-ml microcentrifuge tube. A volume of 90 µl of 0.5M NaOH was added and maceration was performed with a clean Konte pestle until the samples were liquefied. A total volume of 3 µl of the solution was transferred immediately to a new tube and 300 µl of 100mM Tris buffer, pH 8.0 were added (Wangsomboondee and Ristaino 2002). The tube was vortexed and placed on ice. The samples were used directly on the PCR reaction or stored for short time at -20°C.

Field samples of tobacco leaves with signs and symptoms of blue mold were collected. Real-time PCR was done with samples from 2 fields in 2007 (2 samples from each field) and 4 fields in 2008 (3 samples from each field) using fresh leaf lesions (Table 1.1). CTAB DNA extraction was performed on 10 mg of tissue from 3 areas including: the lesion area, an area adjacent to the lesion area, and 4-5 cm from the lesion for each of the samples. A 10-fold dilution with sporangiospores extracted DNA from the PennB isolate of *P. tabacina* was used to compare DNA concentration of the samples tested and a nontemplate water and healthy tobacco leaf control were used in all experiments. All the reactions were performed in duplicate. A final number of 96 real-time reactions were performed for this assay.

To determine whether the real-time PCR assay could be used for detection of the pathogen in leaves with nonsporulating chlorotic lesions, 10 detached leaves were artificially inoculated with the isolate WI-07. *N. repanda* leaves were inoculated with a $1X10^8$ spores/ml using a cotton swab and incubated in Petri dishes for 6 days. When chlorotic nonsporulating lesions were observed, 10 mg of tissue were used for CTAB DNA extraction and the real-time assay was performed for each of the samples.

Late stage lesions with visible sporulation (sporangia), chlorosis and necrosis were sampled. Twenty five seedlings growing under greenhouse conditions and 25 seedlings of *in vitro* cultured *N. repanda* were artificially inoculated and stored in

plastic boxes or in tissue culture containers, respectively. After 12 days, 10 mg of tissue from the sporulating lesions were removed and DNA was extracted with the CTAB method, and the real-time assay was performed for each of the samples as described previously.

Lesions from air dried leaves from Ohio fields sent in 2004 that were stored in the freezer at -20°C were tested. Ten visible lesions (2 mg each) from different leaf samples (OH-AD-04, OH-CL-04, OH-HL-04, OH-GA-04, OH-JK-04, OH-BR-06, OH-CL-06, OH-GA-06, OH-JK-06 and OH-LA-06; Table 1.1) were tested and CTAB DNA extraction and real-time PCR was performed as described previously.

A total of 188 real-time PCR reactions were performed in the test on chlorotic, sporulating and air dry tobacco leaves. A 450 nM of primers and 125 nM of probe concentration were used and a 10-fold dilution with PennB *P. tabacina* isolate were used to compare DNA concentration of the samples tested. Each single reaction was performed in duplicate. Controls included no templates samples and healthy tobacco leaf tissue. All experiments were done twice.

Tobacco plants were inoculated to induced systemic infection by injection with a sterile needle of $1X10^8$ spores/ml solution of *P. tabacina* into the stem of 10 susceptible *Nicotiana repanda* plants. After 20 days and no signs of sporulation on the leaves, tissue was collected from the leaves (0.1 g), stems (0.15 g) and roots (0.01 g) of each tobacco plant and CTAB DNA extraction and real-time PCR were

performed as described previously. Each single reaction was performed in duplicate. A total of 72 reactions were performed.

Twelve tobacco plants were inoculated and left for 30 days under humid conditions, until the aerial part of the plant was consumed by the pathogen. Roots systems from the 12 plants were removed from the soil, and washed with distilled water. DNA was extracted and real-time PCR was performed on 0.01 g of root tissue for each of the plants.

**PtabBM TaqMan assay to quantify host resistance**

In order to quantify the colonization of pathogen and DNA concentration in tobacco leaf tissue during early stages of infection, the ratio of pathogen DNA to host DNA was determined in leaf lesions. The PtabBM TaqMan assay was tested on 8 tobacco genotypes that differ in resistance to blue mold. Samples included burley cultivars “NC2002” and “NC2000” and flue-cured cultivar “Chemical mutant” (resistance obtained after treating seeds of flue-cured tobacco cultivar Virginia Gold with triethylene iminotriazine) (Rufty and Main 1989), which have a moderate level of resistance to the blue mold; burley “TN90” cultivar which is considered tolerant to the pathogen; “Xanthy” and flue-cured cultivars “Hicks” and “NC55” which are susceptible and burley “KY14” cultivar which is very susceptible to *P. tabacina* (Ivors and Shoemaker 2005). Tobacco cultivars were provided by Dr. Ramsey Lewis from

the Crop Science Department at North Carolina State University. Tobacco discs (wet weight: 0.1 g and 2 cm diameter) were inoculated using a cotton swab with a solution of $1X10^8$ sporangia/ml of isolate VA-08 (Virginia 2008) of *P. tabacina* as previously described. Three discs were placed in separated inverted Petri dishes with 30% water agar. The discs were inoculated and then sampled at 0, 2, 4, 6, 8 and 10 days after inoculation and a total of 144 CTAB DNA extractions were performed from each single leaf disk, from each cultivar (n=8), from the 6 sample times. Two un-inoculated controls leaves for each tobacco cultivar, a no template control, and 2 noninoculated negative controls of *N. repanda* were also used.

To compare the highly susceptible and most resistant varieties, DNA was extracted from the 3 discs from “Chemical Mutant” (36 samples) and “KY14” (36 samples) cultivars and from the negative controls. The PtabBM TaqMan assay was used to detect the degree of the pathogen colonization over time in tobacco tissue. For each round of real-time PtabBM TaqMan detection was performed using the universal primer in a different reaction plate, in order to ensure reporting by both fluorogenic probes and no interference on the quantification of the DNA. A concentration of 900 nM of primers and 250 nM of probe was used and a 10-fold dilution with PennB was used to compare DNA concentration of the samples tested. Each single reaction was performed in duplicate and experiments were repeated two times. DNA concentrations for “Chemical mutant” and “KY14” cultivars were

compared between each treatment and between different sample times by analysis of variance using the Proc GLM procedure of SAS 9.1 (SAS Institute, Inc., Cary, N.C.). Alpha levels <0.05 were used to denote statistical significance.

## RESULTS

### Primers and probe design

We developed a real-time PCR assay for *P. tabacina* that uses 5' fluorogenic exonuclease (TaqMan®) chemistry to detect and quantify pathogen DNA from diseased tissue. The primers and probe were designed using 5.8S rDNA sequences from 12 fungal and Oomycete tobacco pathogens and 24 *Peronospora* species. The amplicon generated by the two primers: Ptab2F and Ptab2R and the PtabBM probe, consists of a fragment of 86 bases (Fig. 1.1)

Primers and probe were designed using sequences from GenBank to minimize the likelihood of nonspecific detection. DNA sequences were compared using 20 *Peronospora* spp., two *Pseudoperonospora* spp., two *Hyaloperonospora* spp. and three *Phytophthora* spp. The PtabBM probe had four mismatches when compared with the 27 sequences from *Peronospora* species and Oomycetes from GenBank (Table 1.2). When the DNA sequences were examined from other tobacco pathogens including four Pythium spp., two *Phytophthora* spp., one *Alternaria* sp.,

one *Cercospora* sp., one *Thielaviopsis sp*., one *Rhizoctonia* sp., one *Sclerotium* sp. and one *Sclerotinia* sp., the Ptab2F and Ptab2R primer sequences were 100% identical except for the *Alternaria* sp., *Cercospora* sp., *Thielaviopsis sp*., *Sclerotium* sp. and *Sclerotinia* sp. sequences; again supporting that the specificity is located in the PtabBM probe. The primer and probe sequences are shown in Appendix 1.

The PtabBM assay was tested to determine if the presence of DNA from tobacco and coextracted compounds would interfere with the efficiency of the assay. The amount of DNA detected from the pathogen, the Ct values and the levels of fluorescence were very similar over the range of DNA concentrations tested in the presence or absence of tobacco DNA (data not shown).

To control for successful DNA extraction and presence of DNA in the sample universal primers were used. The addition of universal primers and probe interfered with the quantification of the DNA from *P. tabacina* but not with the detection of the pathogen. Lower levels of DNA were detected from the same sample when the universal primers and probe were used in a multiplex reaction. For example, when *P. tabacina* sample 123368 was tested, the resulting DNA concentration using the PtabBM assay was 414.72 ng/µl; when the multiplex reaction was performed using the Universal Master Mix (recommended to perform the PtabBM single reaction) the DNA concentration was reduced to 272.07 ng/µl and when the multiplex reaction was performed using the Gene Expression Master Mix (recommended to perform

the multiplex reaction using both primers/probe assays) the concentration was even lower, 48.49 ng/µl. The use of a multiplex reaction, where two sets of primers and probes are used, reduced the amount of master mix that can be utilized by the PtabBM assay itself and the DNA concentration reported by the assay was reduced as well. When the multiplex reactions are used to diagnose the *P. tabacina* pathogen the assay is not affected significantly but when quantification assays are performed it is recommended to use both sets of primers/probe in separate reactions.

No false positive or false negative results were observed in any of the repetitions of the assay.

**Optimization of the PtabBM TaqMan assay**

Four different concentrations of primer/probe including 900nM primers/250nM probe, 450nM primers/125nM probe, 225nM primers/62.5nM probe and 125nM primers/31.2 probe were tested (Fig. 1.2). Positive reactions were detected sooner at the higher concentration of primer/probe (900nM primers/250nM probe) and the slope was higher than when the other concentrations of primer and probe were tested. The Ct value was 16.17 for the 900nM primers/250nM probe and a Ct value of 20.3 for the lowest (125nM primers/31.2 probe) concentration was obtained (Figure 1.2, plot A).

**Sensitivity of the PtabBM TaqMan assay**

The detection limit for *P. tabacina* was improved by the use of the two highest primers/probe concentrations. The 900 nM primers/250 nM probe concentration values were useful for detection of DNA concentrations between 10 ng (Ct value of 16.17) and 10 fg (Ct value of 37.92) (Figure 1.2, plot B). The lower primer probe concentrations of 450 nM primers/125 nM probe were useful to detect values between 10 ng (Ct value of 17.32) and 100 fg (Ct value of 36.2) of DNA (Figure 1.2, plot B). The 450nM primers/125nM probe concentration were chosen to perform most of the subsequent assays since the quality, detection and reliability of the reactions were optimum and a reduction of the cost of the detection assay was achieved by diluting reagents. The highest concentration of primer and probe detected lower concentrations of pathogen DNA (10 ng to 10 fg) than the lower concentrations of primer and probe (10 ng to 100 fg).

A standard curve was constructed by plotting known concentrations of *P. tabacina* DNA against the Ct values obtained from real-time PCR. The 10-fold dilution series was done with 10 ng of DNA from *P. tabacina* isolate PennB. In order to validate the reproducibility of the assay, PCR amplifications were conducted based on three replicate serial dilutions of DNA extracted from different preparations. Highly reproducible Ct values with very small standard deviations were observed and a linear response from 10 ng down to 10 fg of DNA (Ct values

between 19.0 and 37.0) with the detection limit of 1 fg and a late Ct value of 39 (Figure 1.3 plot A) were found. The regression equation for the DNA standard curve (Figure 1.3, plot B) was y= -2.91 log(*x*) + 22.50 with an $R^2$ value of 0.99; with an efficiency value of 120%. Each 10-fold difference in initial DNA amounts was represented by approximately three-cycle differences in Ct value.

**Specificity of the PtabBM TaqMan assay**

The detection of *P. tabacina* pathogen was highly specific using the primers Ptab2F and Ptab2R and the PtabBM probe. When the assay was tested against DNA from 12 other tobacco pathogens and 5 other Oomycetes, no cross-amplification by the PtabBM TaqMan assay was observed (Figure 1.4).

Only *P. tabacina* DNA showed a Ct value in the optimal range and the other species were not detected before 40 cycles. Isolates of *P. tabacina* showed Ct values ranging from 15.0 to 35.0 representing DNA concentration from 1805 ng/µl to $7.53X10^{-6}$ ng/µl. Controls with no template were negative showing Ct values of more than 40 cycles for all of the assays.

**Detection of *P. tabacina* in infected field samples and artificially inoculated leaves, seedlings and roots**

Pathogen DNA from both sporangiospore solutions and infected leaves were detected with the CTAB-extracted DNA by the PtabBM TaqMan assay (Figure 1.5). The real-time PCR assay worked equally well with fresh or dried leaf lesions. Sporangiospores washed from infected leaves were also detected at latter Ct values with the assay. A simple NaOH/Tris extraction method used in a previous study (Wangsoomboondee and Ristaino, 2002) was also tested and worked well with the real-time assay for detection of *P. tabacina* in infected leaves (data not shown) (Wang *et al*. 1993).

The assay was also positive for detection of *P. tabacina* extracted from DNA from field samples and Ct values between 15.2 and 35.4 and DNA concentration from $5X10^{-5}$ ng/µl to 2760 ng/µl were found. In plants with early stage infection with chlorotic symptoms only, Ct values were between 30 and 33 and DNA concentration ranged from $4.6X10^{-6}$ ng/µl to $1.42X10^{-3}$ ng/µl were observed, while late stage lesions with presence of sporulation, chlorosis and necrosis Ct values between 24 and 32 and DNA concentrations from $3.8X10^{-5}$ ng/µl to 5.67 ng/µl were observed (Table 1.3). Tests on leaves with dried lesions that were stored for long periods (4 years or less) were also positive with Ct values between 18 and 36.4 and DNA concentration from $5.5X10^{-5}$ ng/µl to 68.0 ng/µl were observed.

Plants inoculated with *P. tabacina* to mimic the systemic infection that occurs naturally in tobacco transplants were also positive. Ct values between 15.3 and 16.8 and DNA concentration from 1863 ng/µl to 2260 ng/µl were found in older leaves and stems (1 cm adjacent to the inoculation point) (Table 1.3). Detection of the pathogen in the newer leaves of the systemically infected tobacco plants was negative.

Roots taken 1 month after inoculation from plants with aerial symptoms of blue mold were also positive in 10 of 10 samples and the Ct values were extremely high and ranged from 18.25 and 25.57 and DNA concentrations ranged from 12.60 ng/µl to 155 ng/µl. Highest DNA concentrations were found in fresh lesions from field samples, systemically infected seedlings and roots.

**PtabBM TaqMan assay to quantify host resistance**

Different tobacco cultivars were inoculated with *P. tabacina* and phenotypic characteristics as well as the amount of DNA from the pathogen present in the tobacco seedlings was determined. Two resistant burley cultivars “NC2000” and “NC2002” exhibited a hypersensitive response 6 days after the inoculation with the pathogen. When these visible results were compared with the DNA concentrations from the PtabBM assay, lower DNA concentrations and less colonization by the pathogen occurred at this initial stage. Ten days after inoculation the pathogen was

sporulating on both varieties. The resistant “Chemical mutant” had the lowest DNA concentration of *P. tabacina*. There was no sporulation on the discs and necrotic lesions were observed (data not shown). In contrast, the susceptible cultivar KY14 had higher concentrations of DNA of the pathogen and sporulation occurred at 8 days from the inoculation.

Colonization of the cultivars “Chemical mutant” which has a medium level of resistance and “KY14” which is very susceptible were analyzed over time. Final concentrations of DNA in leaf discs were significantly different between the cultivars ($P<0.0001$) (Figure 1.6). The susceptible cultivar “KY14” had higher pathogen DNA concentrations after 4 days and the pathogen increased exponentially up to 10 days after inoculation. The medium resistant cultivar “Chemical mutant” had lower pathogen DNA concentrations, maybe due to the hypersensitive response that occurred over time. At 10 days post inoculation, there were no visible symptoms of the disease in “Chemical mutant” but the DNA from the pathogen was detected in lesions (Figure 1.6).

## DISCUSSION

*Peronospora tabacina* is an extremely difficult obligate parasite for identification. The identification of the pathogen is only feasible when the appearance of the symptoms of the disease start, which include chlorosis, necrosis

and sporulation. We developed a real-time TaqMan based method for the detection and quantification of *P. tabacina* in symptomatic and asymptomatic tissue. This method enables the rapid detection of DNA before the appearance of the first signs and symptoms of the disease. The PtabBM TaqMan assay is based on a multiplex PCR reaction with two universal primers, two specific primers for the pathogen and two probes to monitor the identification, quantification and quality of the DNA extracted from infected tissue.

The detection of *P. tabacina* using the TaqMan assay was extremely high and the method was improved by the use of primers and probe generated inside a conserved region of the 5.8S rDNA. The nuclear encoded ribosomal DNA genes (rDNA) provide an attractive target since this region is highly stable, possess conserved as well as variable sequences, and can be amplified and sequenced with universal primers. The rDNA is a multiple copy region of genomic DNA. *P. tabacina*'s sporangiospores have from 4-35 nuclei per sporangium (Trigiano *et al*. 1987, Trigiano *et al*. 1985). The multinuclear characteristic of *P. tabacina* increases the amount of available DNA that can be detected.

The assay was optimized using reduced concentrations of primer/probe than the concentrations recommended by the manufacturer and the ability of the assay to detect and quantify *P. tabacina* DNA was not compromised. The sensitivity of the assay was not affected by the use of half of the manufacturer recommended

primers/probe concentration. The possibility to use half of the recommended concentrations can also reduce costs when the assay has to be performed with high numbers of samples.

It was possible to detect *P. tabacina* DNA at concentrations as low as 10 fg. Higher detection limits have been reported for a range of other plant pathogens with primers developed on multicopy genes such as the ITS region of the rDNA, as well (Atallah and Stevenson 2006, Schena *et al*. 2006, Hayden *et al*. 2006, Silvar *et al*. 2005, Hayden *et al*. 2004). This lower detection limit demonstrated the high sensitivity of the assay as a diagnostic tool for *P. tabacina*.

Using serial dilutions of target DNA, linear responses and high correlation coefficients between the amount of DNA and cycle thresholds for the target pathogen was achieved. Furthermore, detection limits and correlation coefficients were not affected by the presence of tobacco and coextracted plant DNA extracts. These results indicate that the method developed herein is appropriate for both qualitative and quantitative analyses of pathogen biomass in infected tissue.

The PtabBM TaqMan assay was tested for primers and probe specificity. The method proved useful and detected 62 isolates of *P. tabacina* tested previously using the PCR detection system from Ristaino *et al*. (2007) and additional isolates of the pathogen. Also the assay did not detect many other different foliar and soilborne tobacco pathogens, including: *Alternaria alternate*, *Cercospora nicotianae*,

*Phytophthora glovera*, *P. parasitica*, *Pythium aphanidermatum*, *P. dissotocum*, *P. myriotylum*, *P. ultimum*, *Rhizoctonia solani*, *Scletotinia sclerotiorum*, *Sclerotium rolfsii*, and *Thielaviopsis basicola*, and five different Oomycetes including: *Pseudoperonospora humuli*, *Hyaloperonospora parasitica*, *Phytophthora infestans*, *P. ramorum* and *P. citricola* using this method. The PtabBM TaqMan primers and probe were species specific and the assay showed an increase in fluorescence exclusively for isolates of the target pathogen and did not cross-react with any of the other *Peronospora* spp. or any other tobacco pathogens. All the *P. tabacina* samples tested showed Ct values between 15.0 and 35.0 which is considered optimal for pathogen diagnostic assays.

The PtabBM TaqMan method was capable of detection of the presence of *P. tabacina* in field samples and artificially inoculated leaves, seedlings and roots. The detection limits were highest for leaves with sporulation, since the internal mycelia in colonized tissue as well as aerial sporangiophores and spores are detected. DNA detection and quantification during the development of early symptoms (chlorosis or necrosis of leaves) and from DNA extracted from stored dried leaf samples was possible.

Detection of the pathogen in systemically infected tobacco plants could be a very useful tool for producers, industry and extension agents. The detection assay allows the detection of the pathogen even before the occurrence of field symptoms

which will enable control strategies to be deployed more effectively.

The detection of the pathogen on different plant organs was also possible. The detection of the pathogen in leaves, stem and roots of infected tobacco plants was successfully accomplished. The detection of *P. tabacina* in the root system a month post inoculation indicates that the pathogen is able to survive in root material for long periods of time. There are two possible explanations for this result: the presence of vegetative hyphae inside the root tissue (Heist *et al*. 2002) or the presence of oospores or other kinds of survival structures in the roots. It is possible that mycelia colonized the roots and remains viable once the plant dies. However further experiments will be necessary to determine the role of root inoculum in the disease cycle. On the other hand, oospores have been only reported on *in vitro* plants with sporulation induced under laboratory conditions (Heist *et al*. 2002). In tobacco fields the presence of oospores has not been reported since 1997 in the U.S. and only at very low levels of (Spurr and Todd 1982, LaMondia and Aylor 2001). The role of oospores in the survival of *P. tabacina* are unknown and further work is necessary including microscopic assays, storage of roots over longer periods of time and other stressful conditions including variable temperature and humidity.

The assay was used to quantify host resistance. The PtabBM TaqMan assay was useful for detection of *P. tabacina* 4 to 5 days post inoculation. This result provides a wide range in time for the detection of the pathogen under natural

conditions. The response of medium resistance varieties to the pathogen was expected, with lower DNA concentration ranges of the pathogen than in the susceptible varieties. The presence of resistance factors did not suppress the total colonization of the pathogen inside the plant, and even though there was no sporulation, mycelial colonization was detected by the assay.

The real-time method presented here not only is a sensitive and specific method for the diagnosis of *P. tabacina* in planta but also gives an accurate, reliable and highly efficient procedure for quantification of the target pathogen DNA present in infected tissue. The ability to quantify the amount of DNA present in a symptomatic or non symptomatic lesion using the PtabBM TaqMan assay will contribute to the improved diagnostics of this important plant pathogen.

The disadvantages of the method include cost, the need for well-equipped laboratories, and the inability to determine if the target organism is viable. Also, the risk of obtaining false positives and negatives can never be fully discounted due to the inability of fully assessing the biodiversity of close relatives (Bilodeau *et al*. 2007).

The real-time PCR protocol developed here can be a useful tool to detect *P. tabacina* rapidly and sensitively from pure cultures and plant tissue samples. The assay has great potential as a tool for identification and detection in a wide range of applications from pathogen surveys to statutory testing. This method offers

advantages over conventional PCR procedures and will provide a useful and rapid tool in nationwide efforts to detect the blue mold pathogen, *P. tabacina*. The need for such a test has been increased over the last few years with the spread of the pathogen from one tobacco producing area in the U.S. to another via shipments of nursery stock (Personal communication with Dr. Kenneth Seebold) or the spread of the pathogen on wind currents in the American and European continents. Also, there is a great interest in world markets for tobacco leaf, transplants, and seeds but there is concern about the potential movement of *P. tabacina* with these materials. The sensitivity, accuracy and rapidity of this PCR-based technique are valuable for the detection of possible infective material and could be valuable to national regulatory agencies to help prevent the spread of *P. tabacina* and subsequent disease.

## REFERENCES CITED


1. Atallah, Z. K., and Stevenson, W. R. 2004. A methodology to detect and quantify five pathogens causing potato tuber decay using real-time quantitative polymerase chain reaction. Phytopathology 96: 1037-1045.

2. Aylor, D., and Taylor, G. 1983. Escape of *Peronospora tabacina* spores from a field of diseased tobacco plants. Phytopathology 73(4): 525-529.

3. Belbahri, L., Calmin, G., Pawlowski, J., and Lefort, F. 2005. Phylogenetic analysis and real-time PCR detection of a presumably undescribed *Peronospora* species on sweet basil and sage. Mycol. Res. 109(11): 1276-1287.

4. Bilodeau, G. J., Levesque, C. A., de Cock, A. W. A. M., Duchaine, C., Briere, S., Uribe, P., Martin, F. N., and Hamelin R.C. 2007. Molecular detection of Phytophthora ramorum by Real-Time Polymerase Chain Reaction using TaqMan, SYBR Green, and Molecular Beacons. Phytopathology 97: 632-642.

5. Bohm, J., Hahn, A., Schubert, R., Bahnweg, G., Adler, N., Nechwatal, J., Oehlmann, R., and Obwald W. 1999. Real-time quantitative PCR: DNA determination in isolated spores of the mycorrhizal fungus *Glomus mosseae* and monitoring of *Phytophthora infestans* and *Phytophthora citricola* in their respective host plants. Journal of Phytopathology 147: 409-416.

6. Brouwer, M., Lievens, B., Van Hemelrijck, W., Van den Ackerveken, G., Cammue, B. P. A., and Thomma B. P. H. J. 2003. Quantification of disease progression of several microbial pathogens of *Arabidopsis thaliana* using real-time fluorescence PCR. FEMS Microbiology Letters 228: 241-248.

7. Bruck, R. I., Gooding, G. V., and Main C. E. 1982. Evidence of resistance to metalaxyl in isolates of *Peronospora hyoscyami*. Plant Dis. 66 (1): 44-45.

8. Caiazzo, R., Tarantino, P., Porrone, F., and Lahoz, E. 2006. Detection and early diagnosis of *Peronospora tabacina* Adam in tobacco plant with systemic infection. J. Phytopathology 154: 432-435.

9. Cullen, D. W., Toth, I. K., Boonham, N., Walsh, K., Barker, I., and Lees, A. K. 2007. Development and validation of conventional and quantitative polymerase chain reaction assays for the detection of storage rot potato pathogens, *Phytophthora erythroseptica*, *Pythium ultimum* and *Phoma foveata*. J. Phytopathology 155: 309–315.

10. Davis, J. M., Main C. E., and Nesmith, W. C. 1990. Aerobiological aspects of the Kentucky blue mold epidemic of 1958. Pages 55-71 in: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

11. Davis, J. M., and Monahan and J. F. 1991. Climatology of air parcel trajectories related to the atmospheric transport of *Peronospora tabacina*. Plant Dis. 75:706-711.

12. Delon, R., and Schiltz, P. 1989. Spread and control of blue mold in Europe, North Africa and the Middle East. Pages 19-42 in: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

13. Gigante R. 1962. Osservazione sulla *Peronospora tabacina* Adam sistemica, effettuate nel 1961. Boll. Staz. Patol. Veget. 20: 21-52.

14. Hayden, K., Ivors, K., Wilkinson C., and Garbelotto M. 2006. TaqMan chemistry for *Phytophthora ramorum* detection and quantification, with a comparison of diagnostic methods. Phytopathology 96: 846-854.

15. Hayden, K., Rizzo, D., Tse, J., and Garbelotto, M. 2004. Detection and quantification of *Phytophthora ramorum* form California forest using a real-time polymerase chain reaction assay. Phytopathology 94: 1075-1083.

16. Heid, C. A., Stevens, J., Livak K., and Williams, M. 1996. Real-time quantitative PCR. Genome Res. 6: 986-994.

17. Heist, E. P., Nesmith, W. C., and Schardl C. L. 2002. Interaction of *Peronospora tabacina* with roots of *Nicotiana* spp. in gnobiotic associations. Phytopathology 92: 400-405.

18. Henson, J. M. 1993. The polymerase chain reaction and plant disease diagnosis. Ann. Rev. of Phytopathol. 31: 81-109.

19. Hogg, A.C. Johnston, R. H., and Dyer, A. T. 2007. Applying real-time quantitative PCR to Fusarium crown rot of wheat. Plant Dis. 91: 1021-1028.

20. Hukkanen A, Pietikäinen L, Kärenlampi S, Kokko H. 2006. Quantification of downy mildew (*Peronospora sparsa*) in Rubus species using real-time PCR. Eur. J. Plant Pathol. 116:225-235.

21. Ippolito, A., Shena, L., Soleti-Ligorio, V., Yaseen, T., and Nigro, F. 2004. Real-time detection of *Phytophthora nicotianae* and *P. citrophthora* in citrus roots and soils. Eur. J. Plant Pathol. 110: 833-843.

22. Ivors, K., and Mina, M. 2007. Burley and flue-cured tobacco: Blue mold. Web-link: http://www.ces.ncsu.edu/depts/pp/bluemold/Blue_mold_disease_note_07.php

23. Ivors, K., and Shoemaker, P. B. 2005. Chapter 9. Disease management. Pages: 85-105 in: 2005 Burley Tobacco Information. North Carolina Cooperative Extension Service. College of Agriculture and Life Science. North Carolina State University, Raleigh, NC.

24. Johnson, G. I. 1989. *Peronospora hyoscyami* de Bary: taxonomic history, strains and host range. Pages 1-18 in: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

25. Judelson, H., and Randall, T. 1998. Families of repeated DNA in the oomycete *Phytophthora infestans* and their distribution within the genus. Genome 41: 605-615.

26. Kox, L. F. F., van Brouwershaven, I. R., van de Vossenberg, B. T. L. H., van den Beld, H. E., Bonants, P. J. M., and de Gruyter, J. 2007. Diagnostic values and utility of immunological, morphological and molecular methods for in planta detection of *Phytophthora ramorum*. Phytopathology 97: 1117-1129.

27. Kucharek, T. A., Young, T. R., and Thomas, W. D. 1996. Occurrences of and some factors affecting blue mold of tobacco in Florida from 1921 to 1995. Proc. Soil Crop Sci. Soc. Fla. 55:81-85.

28. Kutyavin, I., Afonina, I., Mills, A., Gorn, V., Lukhtanov, E., Belousov, E., Singer, M., Walburger, D., Lokhov, S., Gall, A., Dempcy, R., Reed, M., Meyer, R., Hedgpeth, J. 2000. 3'-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. Nucl. Acids Res. 28(2): 655-661.

29. LaMondia, J. A., and Aylor D. E. 2001. Epidemiology and management of a periodically introduced pathogen. Biological Invasions 3: 273-282.

30. Lea H. W. 1999. Resistance of tobacco to pandemic blue mold (*Peronospora hyoscyami* de Bary syn *P. tabacina* Adam): a historical overview. Australian J. of Exp. Agric. 39: 115-118.

31. Lievens, B., Brouwer, M., Vanachter, A. C.R.C., Cammue, B. P.A., and Thomma B.P.H.J. 2006. Real-time PCR for detection and quantification of fungal and oomycete tomato pathogens in plant and soil samples. Plant Science 171: 155–165.

32. Lucas, G. B. 1975. Chapter 20. Blue Mold. Pages: 235-266 in: Diseases of Tobacco. Biological Consulting Associates, Raleigh, NC.

33. Lucas, G. B. 1980. The war against blue mold. Science. 210: 147-153.

34. Main, C. E. 1995. Tobacco blue mold epidemic of 1995 in the United States. Coresta Meeting Report, Oxford, England

35. Main, C. E., Keever, T., Holmes, G., and Davis, J. 2001. Forecasting long-range transport of downy mildew spores and plant disease epidemics. APS Net Feature, April-May 2001. Web link: http://www.apsnet.org/online/feature/forecast/

36. Mandryk, M. 1971. Resistance of Solanaceous and non-Solanaceous species to *Peronospora tabacina* as shown by necrotrophic reactions. Aust. J. of Exp. Agric. and Animal Husbandry 11: 94-98.

37. Mandryk, M. 1966. Stem infection of tobacco plants with three strains of *Peronospora tabacina* Adam. Aust. J. Agric. Res. 17: 39-47.

38. McGrath, H., and Miller, P. R. 1958. Blue mold of tobacco. Plant Dis. Rep. S. 250:1-35.

39. Moss, M. A., and Main, C. E. 1988. The effect of temperature on sporulation and viability of isolates of *Peronospora tabacina* collected in the United States. Phytopathology 78:110-114.

40. Moss, M. A., and Main, C. E. 1989. Factors affecting systemic infection of tobacco by *Peronospora tabacina*. Phytopathology 79: 865-868.

41. Qi, M., and Yang Y. 2002. Quantification of *Magnaporthe grisea* during infection of rice plants using real-time polymerase chain reaction and northern blot/phosphoimaging analyses. Phytopathology 92: 870-876.

42. Ristaino J.B., Johnson A., Blanco-Meneses M., and Liu B. 2007. Identification of the tobacco blue mold pathogen, *Peronospora tabacina*, by polymerase chain reaction. Plant Dis. 91: 685-691.

43. Rufty, R.C. 1989. Genetics of host resistance to tobacco blue mold. Pages: 141–164 in: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

44. Schena, L., Hughes, K. J. D., and Cooke D. E. L. 2006. Detection and quantification of *Phytophthora ramorum*, *P. kernoviae*, *P. citricola* and *P. quercina* in symptomatic leaves by multiplex real-time PCR. Molecular Plant Pathology 7(5): 365-379.

45. Schena, L., Nigro, F., Ippolito, F., and Gallitelli D. 2004. Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. Euro. J. of Plant Pathol. 110: 893-908.

46. Schaad, N. W., and Frederick, R. D. 2002. Real-time PCR and its application for rapid plant disease diagnostics. Can. J. Plant Pathol. 24: 250-258.

47. Silvar, C., Díaz, J., and Merino F. 2005. Real-Time Polymerase Chain Reaction quantification of *Phytophthora capsici* in different pepper genotypes. Phytopathology 95: 1423-1429.

48. Smith, E. F. and McKenny, R. E. B. 1921. Dangerous tobacco disease appears in the United States. U. S. Dept. Agr. Cir. 174: 3-6.

49. Svircev, A.M., McKeen, W. E., and Smith, R.J. 1989. Host-parasite relations: morphology and ultrastructure, Pages 43-104, in: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

50. Spurr, H. W., and Todd, F. A. 1982. Oospores in blue mold diseased North Carolina burley and flue-cured tobacco. Tobacco Science. 26: 44-46.

51. Thomas, J. F., and Downs, R. J. 1991. Phytotron Procedural Manual for Controlled- Environmental Research at the Southeastern Plant Environment Laboratory. North Carolina State University Technical Bulletin.

52. Todd, F.A. 1981. The blue mold story. Page: 109 in: Report: 29th Tobacco Workers Conference. F. A. Todd, compiler. Lexington, KY.

53. Tooley, P. W., Martin, F. N., Carras, M. M., and Frederick, R. D. 2006. Real-time fluorescent polymerase chain reaction detection of *Phytophthora ramorum* and *Phytophthora pseudosyringae* using mitochondrial gene regions. Phytopathology 96: 336-345.

54. Trigiano, R. N., Van Dyke, C. G., Spurr, H.W., Jr., and Main, C. E. 1985. Ultrastructure of sporangiophore and sporangium ontogeny of *Peronospora tabacina*. Tob. Sci. 29:116-121.

55. Trigiano, R. N. and Spurr, H.W. 1987. The development of the multinucleate condition of *Peronospora tabacina* sporangia. Mycologia 79(3): 353-357.

56. Trout, C. L., Ristaino, J. B., Madritch, M., and Wangsomboondee, T. 1997. Rapid detection of *Phytophthora infestans* in late blight infected tissue of potato and tomato using PCR. Plant Dis. 81:1042-1048.

57. Tsay, J. G., Chu, C., Chuang, Y. H., and Chen, R. S. 2006. Specific detection of *Peronospora tabacina* by PCR-amplified rDNA sequences. Plant Pathology Journal 5(3): 378-382.

58. Vandermark, G. J., and Baker, B. M. 2003. Quantifying *Phytophthora medicaginis* in susceptible and resistant alfalfa with a real-time fluorescent PCR assay. Journal of Phytopathology 151: 577-583.

59. Wang, H., Qi, M., and Cutler A. 1993. A simple method of preparing plant samples for PCR. Nucleic Acids Res. 21: 41-53-4154.

60. Wangsomboondee T., and Ristaino, J. B. 2002. Optimization of samples size and DNA extraction methods to improve PCR detection of different propagules of *Phytophthora infestans*. Plant Dis. 86: 247-253.

61. White, T. J., Bruns, T., Lee S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315-322 in: PCR Protocols: A Guide to Methods and Applications. M. A. Innis, D. H. Gelfand, J. J. Sninshy, and T. J. White, eds. Academic Press, New York.

62. Wiglesworth, M. D., Nesmith, W. C., Schardl, C. L., Li, D., and Siegel, M. R. 1994. Use of specific repetitive sequences in *Peronospora tabacina* for the early detection of the tobacco blue mold pathogen. Phytopathology 84:425-430.

63. Wiglesworth, M.D., Reuveni, M., Nesmith, W. C., Siegel, M. R., Kuc, J., Juarez, J. 1988. Resistance of *Peronospora tabacina* to metalaxyl in Texas and Mexico. Plant Disease 72: 964-967.

**Table 1.1** *Peronospora tabacina* isolates collected between 1945 and 2007 from U.S. and international sources and tested by real time PCR.

| Isolate | Year | Source | Collector and location |
|---|---|---|---|
| **United States samples** | | | |
| CT991 | 1999 | Connecticut | J. LaMondia, University of Massachussets |
| CT021 | 2002 | Connecticut | J. LaMondia, University of Massachussets |
| FL002 | 2000 | Florida | T. Kucharek, University of Florida |
| FL08 | 2008 | Florida | A. Gevens, University of Florida |
| GA992 | 1999 | Georgia | P. Bertrand, University of Georgia |
| GA996 | 1999 | Georgia | A. Csinos, University of Georgia |
| GA995 | 1999 | Georgia | A. Csinos, University of Georgia |
| GA08 | 2008 | Georgia | J. Jacobs, University of Georgia |
| KY79 | 1979 | Kentucky | B. Nesmith, University of Kentucky |
| KY206 | 2006 | Kentucky | K. Seebold, University of Kentucky |
| KY062 | 2006 | Kentucky | K. Seebold, University of Kentucky |
| KY071 | 2007 | Kentucky | K. Seebold, University of Kentucky |
| BPT | 1984 | Maryland | Clayton, Texas A&M University |
| Hadley | 2004 | Maryland | P. Shoemaker, North Carolina State Univesity |
| BPTS | 1945 | North Carolina | no information |
| RPT83 | 1983 | North Carolina | C. Main, North Carolina State University |
| Bertie | 1983 | North Carolina | R. Rufty, North Carolina State University |
| PT87W | 1987 | North Carolina | C. Main, North Carolina State University |
| Waynesville1 | 1998 | North Carolina | J. Ristaino, North Carolina State University |
| NC998 | 1999 | North Carolina | P. Shoemaker, North Carolina State University |
| NC9911 | 1999 | North Carolina | P. Shoemaker, North Carolina State University |
| NC997 | 1999 | North Carolina | T. Melton, North Carolina State University |
| NC994 | 1999 | North Carolina | G. Pate, North Carolina State University |
| NC0212 | 2002 | North Carolina | P. Shoemaker, North Carolina State University |
| AV041 | 2004 | North Carolina | F. Bolick, North Carolina State University |
| BU041 | 2004 | North Carolina | S. Holloway, North Carolina State University |
| BU061 | 2006 | North Carolina | K. Ivors, North Carolina State University |
| PennB | 1998 | Pennsylvania | J. Yocum, Pennsylvania State University |
| Colum | 1998 | Pennsylvania | J. Yocum, Pennsylvania State University |
| PE05 | 2005 | Pennsylvania | B. Maksymowicz, University of Kentucky |
| PE07 | 2007 | Pennsylvania | B. Maksymowicz, University of Kentucky |
| PE08 | 2008 | Pennsylvania | B. Green, Pennsylvania |
| BPTP | 1945 | Texas | Clayton, Texas A&M University |
| TX84 | 1984 | Texas | B. Nesmith, University of Kentucky |
| TX062 | 2006 | Texas | M. Black, Texas A&M University |
| Mosel | 1998 | Virginia | J. Ristaino, North Carolina State Univesity |
| VA011 | 2001 | Virginia | Avery, Virginia Tech |
| VA08 | 2008 | Virginia | R. Wells, Virginia Tech |
| OH-AD-04 | 2004 | Ohio | R. Stevenson, Ohio State University |
| OH-CL-04 | 2004 | Ohio | R. Stevenson, Ohio State University |
| OH-HL-04 | 2004 | Ohio | R. Stevenson, Ohio State University |
| OH-GA-04 | 2004 | Ohio | R. Stevenson, Ohio State University |
| OH-JK-04 | 2004 | Ohio | R. Stevenson, Ohio State University |
| OH-BR-06 | 2006 | Ohio | R. Stevenson, Ohio State University |
| OH-CL-06 | 2006 | Ohio | R. Stevenson, Ohio State University |
| OH-GA-06 | 2006 | Ohio | R. Stevenson, Ohio State University |
| OH-JK-06 | 2006 | Ohio | R. Stevenson, Ohio State University |
| OH-LA-06 | 2006 | Ohio | R. Stevenson, Ohio State University |
| WI-07 | 2007 | Wisconsin | M. Drilias, University of Wisconsin |

Table 1.1 continued

| Isolate | Year | Source | Collector and location |
|---|---|---|---|
| **International samples** | | | |
| 123046 | 1988 | Bulgaria | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| DR022 | 2002 | Dominican Rep. | I. Abreu, FERQUIDO |
| DR051 | 2005 | Dominican Rep. | A. Ubiera, FERQUIDO |
| DR054 | 2005 | Dominican Rep. | A. Ubiera, FERQUIDO |
| DR062 | 2006 | Dominican Rep. | M. Peralta, Tabadom Holding Inc. |
| DR064 | 2006 | Dominican Rep. | M. Peralta, Tabadom Holding Inc. |
| DR066 | 2006 | Dominican Rep. | M. Peralta, Tabadom Holding Inc. |
| 123392 | 1989 | Guatemala | Peterson, USDA, Ft. Detrick, MD [a] |
| 123364 | 1987 | France | R. Delon, Institut du Tabac-Altadis |
| FR184 | 2005 | France | J. L.Verrier, Institut du Tabac-Altadis |
| FR194 | 2005 | France | J. L.Verrier, Institut du Tabac-Altadis |
| FR178 | 2005 | France | J. L.Verrier, Institut du Tabac-Altadis |
| FRy1 | 1980s | France | J. L.Verrier, Institut du Tabac-Altadis |
| 123350 | 1963 | Germany | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123368 | 1960 | Hungary | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123340 | 1963 | Lebanon | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| NIC063 | 2006 | Nicaragua | N. Placencia, Segovia Cigars |
| NIC065 | 2006 | Nicaragua | N. Placencia, Segovia Cigars |
| 123347 | 1989 | Mexico | P. Shoemaker, North Carolina State University |
| 123296 | 1989 | Mexico | P. Shoemaker, North Carolina State University |
| Mex | 1999 | Mexico | R. Rufty, North Carolina State University |
| MX02 | 2002 | Mexico | V. Nikolaeva, Amapa, Nayarit |
| 123349 | 1989 | Poland | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123344 | 1989 | Poland | G. L. Peterson, USDA, Ft. Detrick, MD [a] |

[a] Samples belong to a collection located at the USDA, Ft. Detrick, MD and donated to the Plant Pathology Department at NCSU in 2004.

**Table 1. 2** Isolates of other tobacco pathogens, *Peronospora tabacina* and other related Oomycetes used to develop primers and probe and test the specificity of the PtabBM real time PCR assay.

| Species | GenBank accession no. [b] | Source [c] | Host | Collector [d] |
|---|---|---|---|---|
| **Tobacco pathogens** | | | | |
| *Alternaria alternata* [a] | DQ059568, DQ323699, AY751456 | N. Carolina, USA | *Nicotiana tabacum* | H. Spurr |
| *Cercospora nicotianae* [a] | DQ059569, DQ059571 | N. Carolina, USA | *Nicotiana tabacum* | M. Daub |
| *Peronospora tabacina* [a] | DQ067896, DQ067897, DQ067898 DQ067899, DQ067900, DQ665672, AY198289 | N. Carolina, USA GenBank | *Nicotiana tabacum* | J. Ristaino, M. Thines and H. Voglmayr |
| *Phytophthora glovera* [a] | DQ059570, AF279126, AF279127, AF279128 | Brazil | *Nicotiana tabacum* | D. Shew |
| *Phytophthora parasitica* [a] | DQ059571 | N. Carolina, USA | *Nicotiana tabacum* | D. Shew |
| *Pythium aphanidermatum* [a] | DQ059572, DQ298521, DQ298523, AF271227, AB160845, AF271227 | N. Carolina, USA | *Nicotiana tabacum* | J. Ristaino |
| *Pythium dissotocum* [a] | DQ059573, AF271228, AF330184 | N. Carolina, USA | *Nicotiana tabacum* | D. Shew |
| *Pythium myriotylum* [a] | DQ059574, DQ102701, DQ222438 | N. Carolina, USA | *Nicotiana tabacum* | W. Gutierrez |
| *Phythium ultimum* [a] | DQ059575, AF271225, DQ211527, AY986952 | N. Carolina, USA | *Nicotiana tabacum* | W. Gutierrez |
| *Rhizoctonia solani* [a] | DQ059576 | N. Carolina, USA | *Nicotiana tabacum* | J. Ristaino |
| *Sclerotinia sclerotiorum* [a] | DQ059577 | N. Carolina, USA | *Nicotiana tabacum* | D. Shew |
| *Sclerotium rolfsii* [a] | DQ059578 | N. Carolina, USA | *Nicotiana tabacum* | J. Ristaino |
| *Thielaviopsis basicola* [a] | DQ059579 | N. Carolina, USA | *Nicotiana tabacum* | D. Shew |
| **Other Oomycete species** | | | | |
| *Hyaloperonospora parasitica* [a] | AY210988.1 | Virginia, USA | *Arabidopsis thaliana* | J. McDowell |
| *Hyaloperonospora niessleana* | AY531465.1 | GenBank | *Alliaria petiolata* | Goeker et al. |
| *Peronospora arabis-alpinae* | AY531466.1 | GenBank | Brassicaceae | Goeker et al. |
| *Peronospora arborescens* | DQ885384.1 | GenBank | *Papaver alpinum* | Landa et al. |

Table 1.2 continued

| Species | GenBank accession no. [b] | Source [c] | Host | Collector [d] |
|---|---|---|---|---|
| **Other Oomycete species** | | | | |
| *Peronospora arenariae* | AY198280 | GenBank | *Moehringia trinervia* | H. Voglmayr |
| *Peronospora arthurii* | AY198284 | GenBank | *Oenothera biennis agg.* | H. Voglmayr |
| *Peronospora astragalina* | AY608608.1 | GenBank | *Astragalus membranaceus* | Choi et al. |
| *Peronospora boni-henrici* | AY198286 | GenBank | *Chenopodium bonus-henricus* L. | H. Voglmayr |
| *Peronospora campestris* | AY608609.1 | GenBank | Amaranthaceae | Choi et al. |
| *Peronospora chenopodii-polyspermi* | AY198291 | GenBank | Chenopodiaceae | H. Voglmayr |
| *Peronospora claytoniae* | AY198281 | GenBank | Minerlettuce | H. Voglmayr |
| *Peronospora corydalis* | AY211015.1 | GenBank | Fumariaceae | Choi et al. |
| *Peronospora cristata* | DQ885375.1 | GenBank | *Papaver somniferum* | Landa et al. |
| *Peronospora destructor* | DQ885385.1 | GenBank | *Allium* spp. | Landa et al. |
| *Peronospora erysimi* | AY531461.1 | GenBank | *Erysimum* herbaceous plants | Goeker et al. |
| *Peronospora farinosa* | AF528558 | GenBank | *Beta* spp. | Choi et al. |
| *Peronospora holostei* | AY198283 | GenBank | Chickweed | H. Voglmayr |
| *Peronospora iberidis* | AY531461.1 | GenBank | Brassicales | Goeker et al. |
| *Peronospora lepidii-sativi* | AY531463.1 | GenBank | Pepperweed | Goeker et al. |
| *Peronospora manshurica* | AY211019.1 | GenBank | *Glycine max* | Choi et al. |
| *Peronospora polygoni* | AY198280 | GenBank | *Polygonum aviculare* | H. Voglmayr |
| *Peronospora sparsa* | AY608610.1 | GenBank | *Prunus laurocerasus* | Choi et al. |
| *Phytophthora citricola* | AM235209 | GenBank | Citrus | Zea-Bonilla et al. |
| *Phytophthora infestans* [a] | EF173472 | Ecuador | *Solanum tuberosum* | J. Ristaino |
| *Phytophthora ramorum* [a] | DQ873514 | Georgia, USA | *Lithocarpus densiflora* , *Quercus* spp. | J. Ristaino |
| *Pseudoperonospora cubensis* [a] | EF050035.1 | N. Carolina, USA | Cucurbits | L. Kanetis |
| *Pseudoperonospora humuli* [a] | EF126356.1 | Oregon, USA | *Humulus lupulus* | G. Holmes |

[a] Species used to test the *P. tabacina* TaqMan assay

[b] GenBank accession numbers used to design primers and probe for the TaqMan assay.

[c] Source indicates state and country where the samples were collected or GenBank when the accession numbers were used.

[d] Indicates the name of sample collector or the author for the accession numbers from GenBank. Harvey Spurr, North Carolina State University; Margaret Daub, N.C.S.U.; David Shew, N.C.S.U.; Jean Ristaino, N.C.S.U.; Walter Gutierrez, N.C.S.U.; John Yocum, Pennsylvania State University; John McDowell, Virginia Tech; Gerald Holmes, N.C.S.U.; David Gent, Oregon State University.

**Table 1. 3** Ct values and lowest and highest concentrations of DNA of *P. tabacina* for field samples and different tissue samples inoculated with *P. tabacina*.

| **Sample** | **Ct value** | | **DNA concentration (ng/µl) / 10 mg of tissue** | |
|---|---|---|---|---|
| | lower | higher | lower | higher |
| Field leaf sample | 15.2 | 35.4 | $5X10^{-5}$ | 2760 |
| Chlorotic leaf | 30 | 33 | $4.6X10^{-6}$ | $1.42X10^{-3}$ |
| Sporulating leaf | 24 | 32 | $3.8X10^{-5}$ | 5.67 |
| Dry leaf (2 mg) | 18 | 36.4 | $5.5X10^{-5}$ | 68 |
| Systemic seedlings | 15.3 | 16.8 | 1863 | 2260 |
| Roots (15 mg) | 18.25 | 25.57 | 12.6 | 155 |
| No template controls | | >40 | none | none |

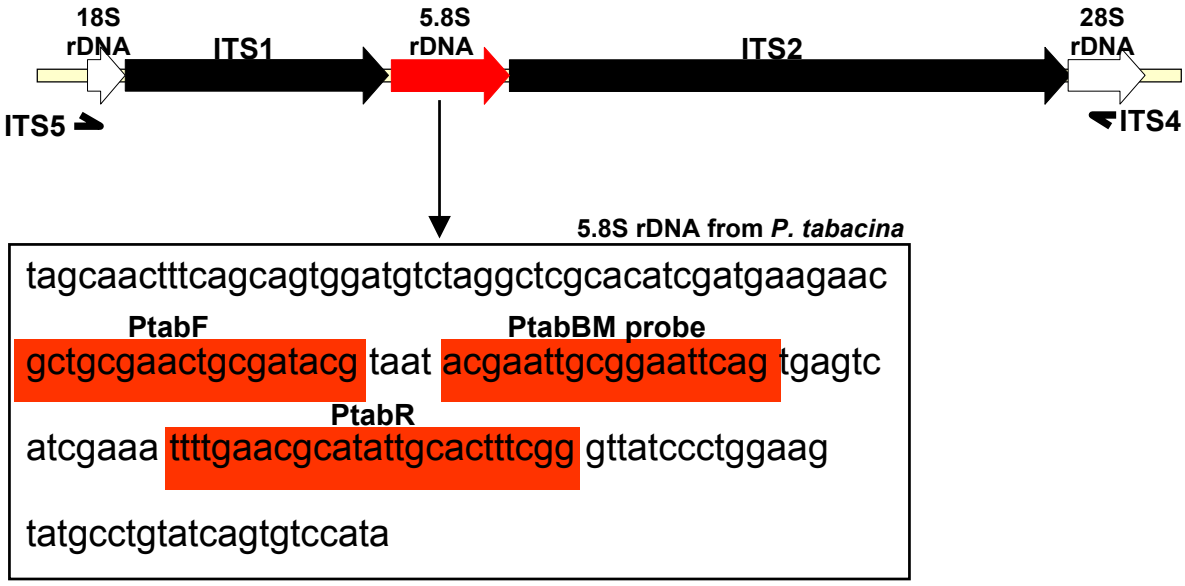



**Figure 1.1** Diagram of the internal transcribed spacer regions and 5.8S rDNA of *P. tabacina* indicating the location of designed primers and probe for the PtabBM TaqMan assay.

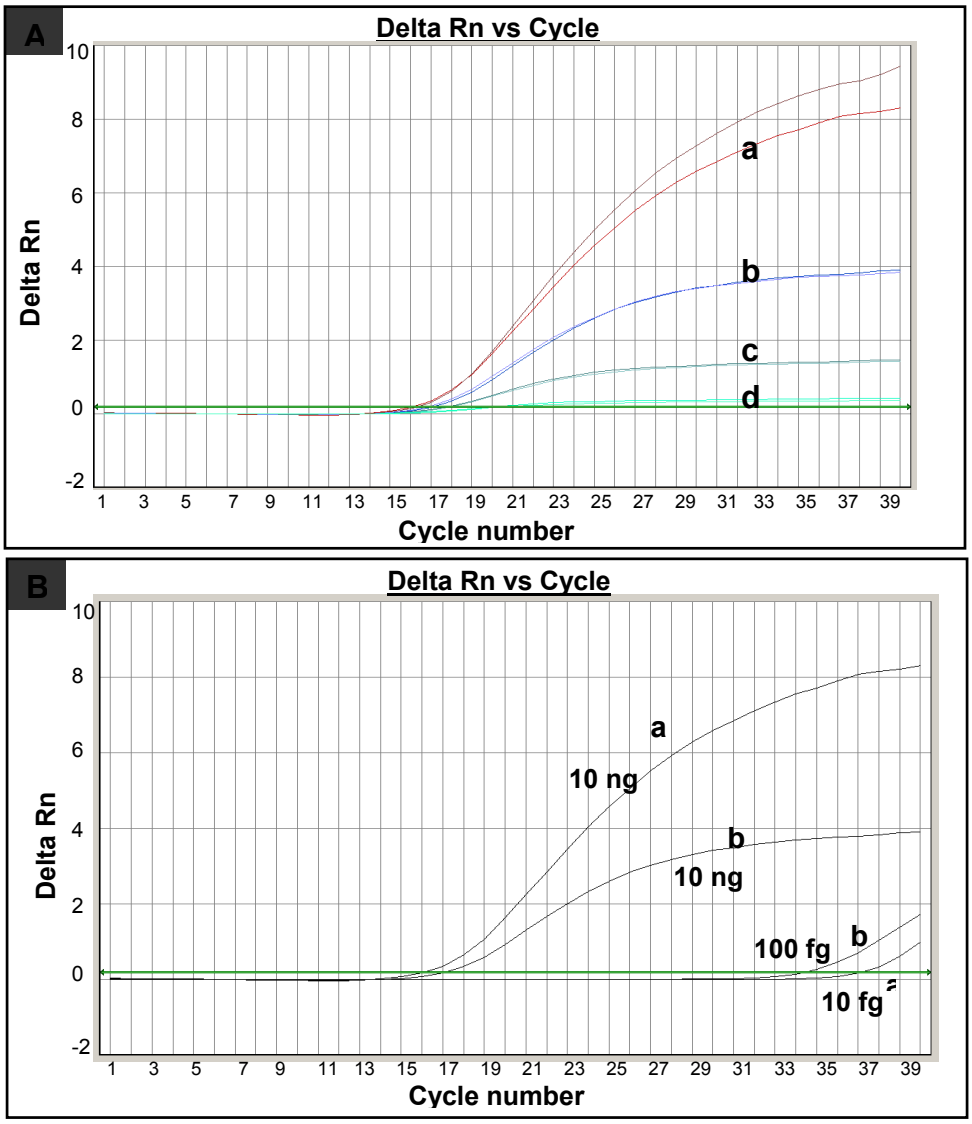



**Figure 1. 2** Plots of the sensitivity of the primers and probe for detection of *P tabacina*; A) The concentrations of: a. 900nM primers / 250nM probe, b. 450nM primers / 125nM probe, c. 225nM primers / 62.5nM probe and d. 112nM primers / 31.2nM probe. The Ct value for the duplicated samples of *P. tabacina* is similar for all of the concentrations used but the quality of the curve for fluorescence was affected by the different concentrations; B) DNA detection limit at the two highest concentrations of primer and probe is shown for: a. 900nM primers / 250nM probe and b. 450nM primers / 125nM probe.

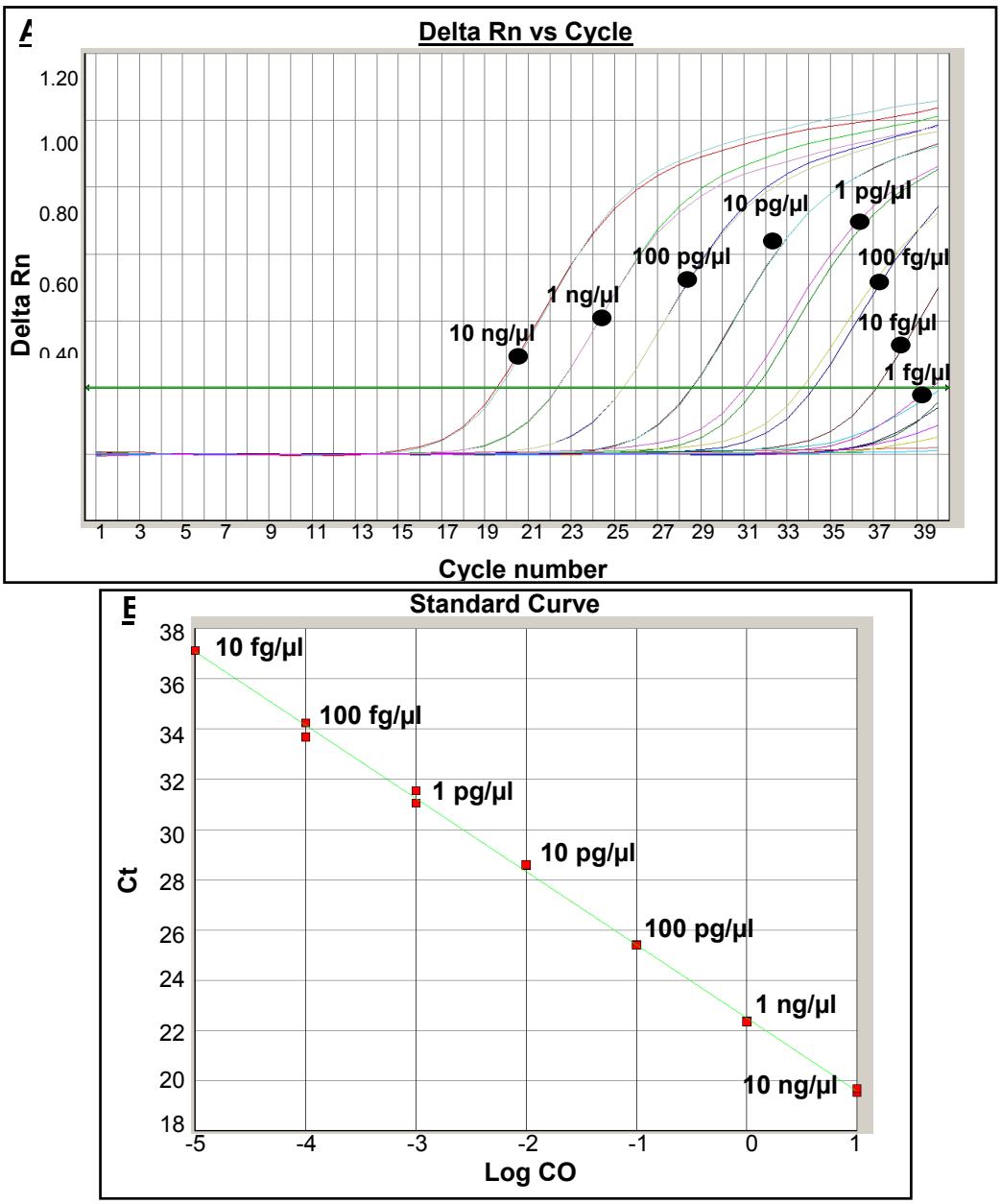



**Figure 1. 3** A 10-fold dilution series showing the sensitivity of the PtabBM real-time assay. A) Plot of Delta Rn vs Cycle number. The highest DNA concentration detected was at 10 ng/µl and a Ct:19.4 and the lowest DNA concentration detected was at 1 fg/µl and a Ct: 39.8 using a concentration of 900nM primers / 250nM probe. B) Plot of Ct value vs log of concentration rate of *P. tabacina* DNA. The standard curve with a Slope: -2.89, Intercept: 22.52 and $R^2$: 0.99.

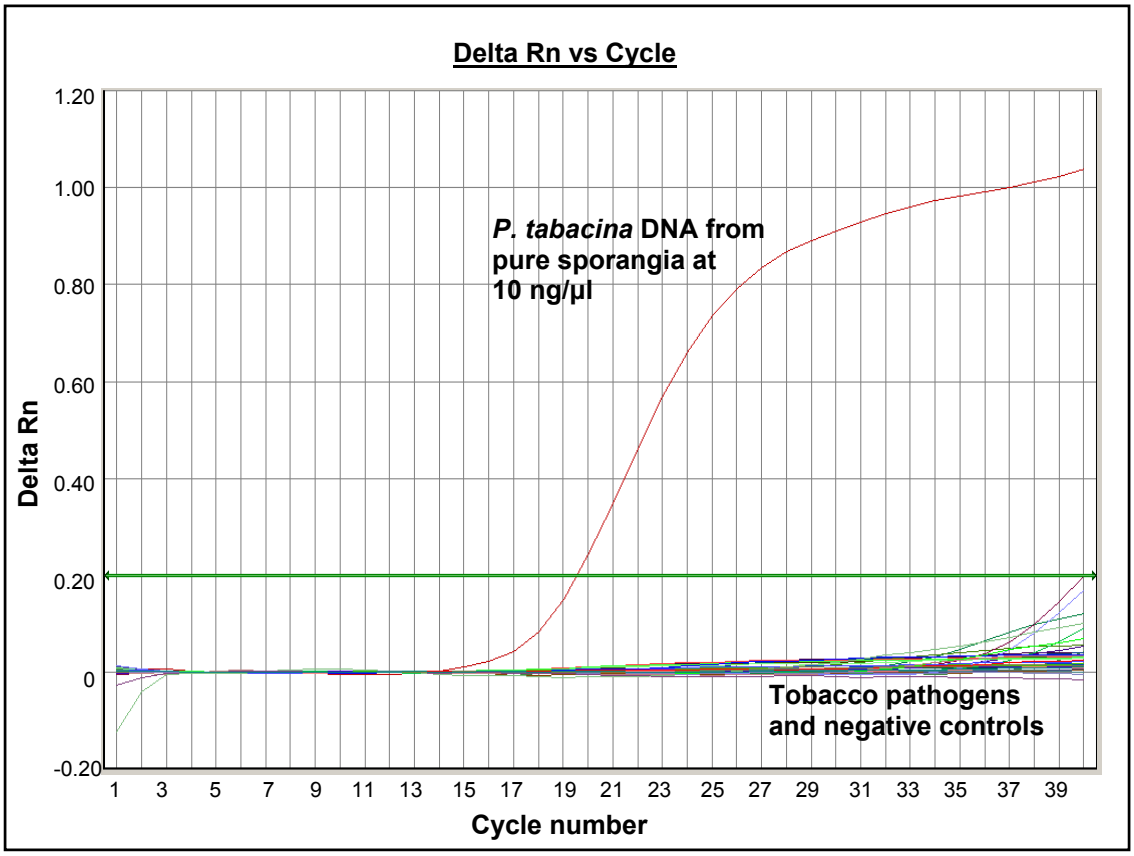



**Figure 1. 4** Plot of the Delta Rn vs Cycle number of the PtabBM real-time PCR assay for detection of DNA from *P. tabacina* sporangiospores (10 ng/µl), and DNA from 12 other tobacco pathogens including four Pythium spp., two *Phytophthora* spp., one *Alternaria* sp., one *Cercospora* sp., one *Thielaviopsis sp*., one *Rhizoctonia* sp., one *Sclerotium* sp. and one *Sclerotinia* sp., and DNA from 5 other Oomycete spp. including *Pseudoperonospora humuli*, *Hyaloperonospora parasitica*, *Pseudoperonospora cubensis*, *Phytophthora infestans* and *P. ramorum* tobacco pathogens and negative controls.

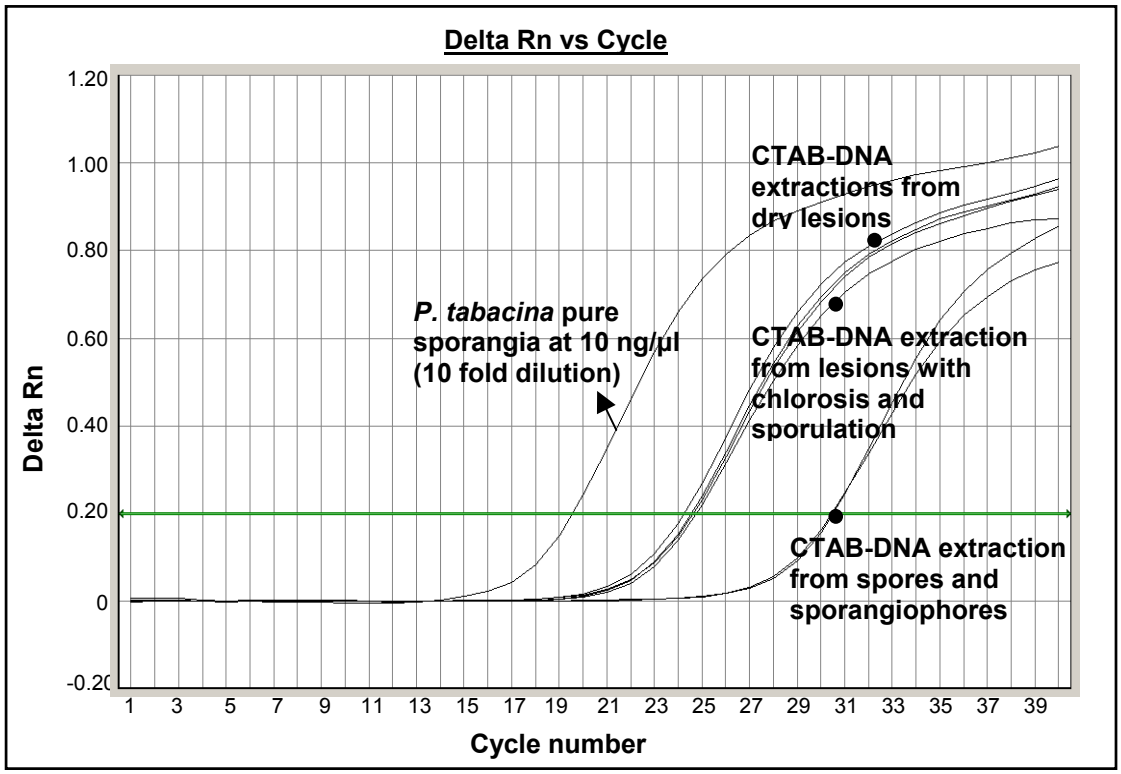



**Figure 1. 5** Plot of Delta Rn vs Cycle number. Detection of *P. tabacina* by real-time PtabBM PCR from a standard sporangiospore suspension (10 ng/µl) or fresh tobacco leaf lesions with sporulation, dried leaf lesions, and sporangiospores washed from fresh lesions. Each sample is represented by two lines (duplicates).

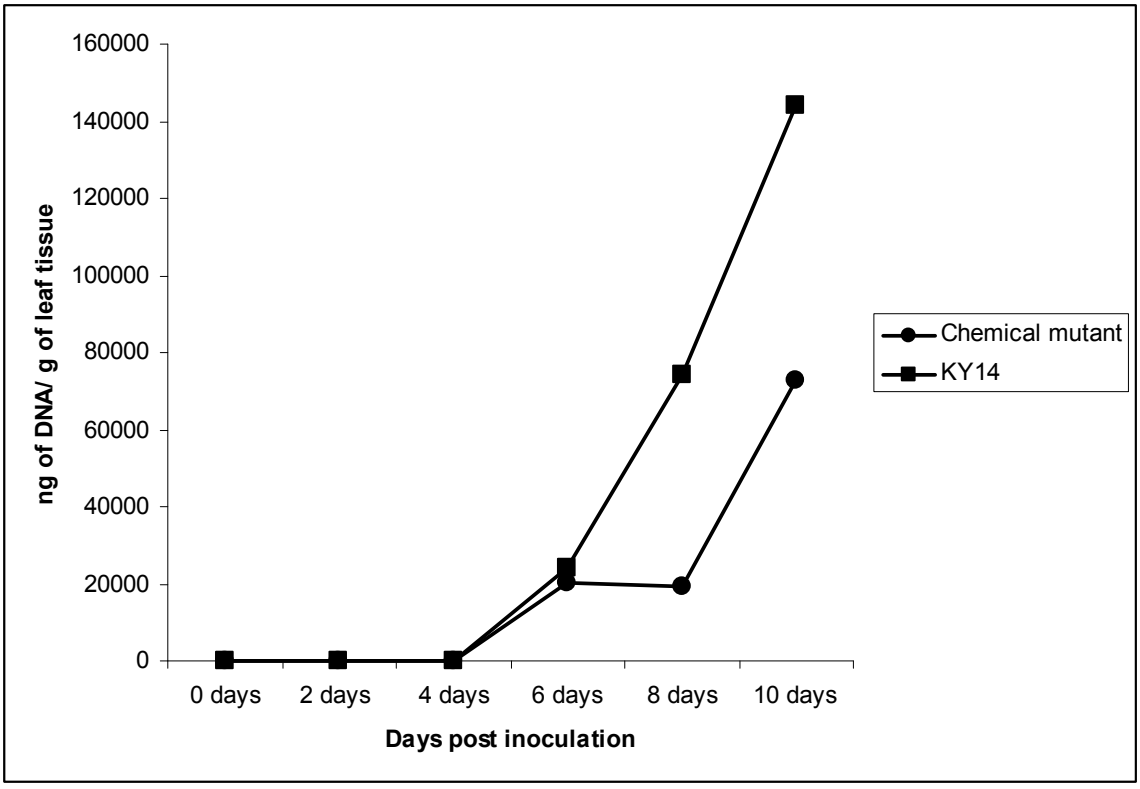



**Figure 1. 6** Detection of DNA by real-time PCR from *P. tabacina* in "KY14" (susceptible burley variety) and "Chemical mutant" (medium resistant flue-cured variety) 2, 4, 6, 8 and 10 days after inoculation.

# CHAPTER 2

# GENETIC STRUCTURE OF POPULATIONS OF THE TOBACCO BLUE MOLD PATHOGEN, PERONOSPORA TABACINA IN NORTH AMERICA, THE CARIBBEAN AND CENTRAL AMERICA, AND EUROPE


## ABSTRACT

Tobacco blue mold, caused by the oomycete pathogen *Peronospora tabacina*, causes yearly epidemics in tobacco (*Nicotiana tabacum*) in North Carolina and other tobacco growing regions in the US and Europe. The genetic structure of *P. tabacina* populations in North America, Central America, the Caribbean and Europe was examined in order to understand genetic diversity, gene flow and migration within and between the regions. The intergenic spacer *Igs2* region of the nuclear ribosomal DNA (rDNA) and the Ras-related protein (*Ypt1*) gene, and the mitochondrial cytochrome c oxidase subunit 2 (*cox*2 gene) were amplified, cloned and sequenced from fifty four isolates from the U.S., Central America-Caribbean-Mexico (CCAM) and Europe and the Middle East (EULE). Nucleotide sequences from the three genes showed high genetic variability across all populations and 93 haplotypes for the *Igs*2 region, 93 haplotypes for the *Ypt*1 region and 82 haplotypes for the *cox*2 gene were observed. Nucleotide diversity and the mean mutation

parameter per site (Watterson's theta) were higher in EULE and CCAM, populations and lower in U.S. populations. Neutrality tests were significant for all populations and the equilibrium model of neutral evolution was rejected, indicating an excess of recent mutations or rare alleles. $S_{nn}$ Hudson's tests were performed to examine population subdivision and gene flow among populations. The U.S. and CCAM populations were pooled into a two genetically subdivided populations: North and South. An isolation-with-migration analysis (IM) indicated migration from South to North supporting the hypothesis of long-distance migration of *P. tabacina* from the Caribbean regions, Florida and Texas into other states further north. Within the European populations, the Hudson's test statistics were also significant and three subdivided populations: North-Central EU, Western EU and Lebanon were observed. An isolation with migration model documented migration from North Central Europe into Western Europe and Lebanon and migration from Western Europe into Lebanon. These data support migrations reported from first disease introductions into Europe. *P. tabacina* could be considered a high evolutionary risk plant pathogen due to its large effective population size and mechanism of dispersal. Intraspore genetic diversity, genome instability due to transposable elements, gene conversion, mitotic and/or sexual recombination may play a role in the high levels of nucleotide variability observed in this pathogen.

## INTRODUCTION

*Peronospora tabacina* Adam (syn. *P. hyoscyami* de Bary), causal agent of tobacco blue mold is an Oomycete pathogen and obligate parasite of tobacco (Nicotiana spp.) (Johnson 1989, Lucas 1980). *P. tabacina* can reproduce both asexually and sexually. The asexual cycle is the predominant phase of the life cycle. Sporangia can be transported by wind currents over long distances and infect tobacco leaf surfaces. Under cool and humid conditions sporangiophores emerge through stomata and produce sporangia which are then released and reinfect other plants. The polycyclic pathogen can complete an infection cycle in 7-10 days. Hyphae, sporangiophores and sporangia are coenocytic and multinucleated and can contain 4-35 nuclei in one sporangiospore (Spurr *et al*. 1990, Trigiano *et al*. 1985).

The sexual cycle occurs less commonly and is usually initiated internally in the leaf mesophyll tissue, within stem tissue or in the roots. The pathogen is believed to be heterothallic and compatible hyphae can give rise to oogonia and antheridia that form oospores. The oospore is a thick-walled resistance structure that can survive adverse environmental conditions for longer periods of time than sporangia (Heist *et al*. 2002, Spurr *et al*. 1990, Svircev *et at*. 1989).

Blue mold of tobacco (*Nicotiana tabacum* L.) was first reported in Australia in 1890 (Tyron, 1890) and is believed to be endemic to Australia and the Americas. (Lucas 1980). In the United States, blue mold was recorded for cultivated tobacco in

1921. In the following years, blue mold was reported on tobacco in Canada, Brazil and Argentina (1939), Chile (1953), Cuba (1957) and Mexico (1964). The pathogen was unknown in Europe until 1957, when it was found in England on tobacco imported from Australia for fungicide trials. By 1959, the pathogen had spread to the Netherlands, Belgium and Germany. In the1960's, the pathogen spread into France, Italy, Yugoslavia and Poland, Spain, Morocco, Syria, and Lebanon (Johnson 1989, Wolf *et al*. 1933).

Prior to 1979, the disease was mostly confined to seed beds of tobacco in the U.S. In 1979, the U.S. and Canadian growers estimated over $250 million in losses due to field epidemics of the disease (Nesmith 1984, Lucas, 1980). The downy mildew of tobacco now causes epidemics on seed beds, transplants and production fields in the humid farming zones of the south eastern and eastern U.S., Canada, Central America and countries bordering the Caribbean basin (Aylor and Taylor 1983, Davis *et al*. 1990, Davis and Monahan 1991, Kucharek *et al*. 1996, Main 1995, Main *et al*. 2001, Moss and Main 1988, Todd 1981). Blue mold pathogen is also a problem in South America (Lea 1999, Johnson 1989) and southwestern and southeastern Europe, the Middle East and North Africa (Ristaino *et al*. 2007, Rufty 1989, Delon and Schiltz 1989).

Several pathways of migration of *P. tabacina* may be operating in the United States and Europe. First, inoculum produced on tobacco in the Caribbean basin

could be disseminated unidirectionally northward to spring and summer-grown tobacco in the more temperate zones. In this case, inoculum arrives anew each year from suspected source regions in the Caribbean or Central America. Similarly in the Mediterranean region and Northern Europe, putative sources of inoculum are from North Africa or the Middle East (Main and Spurr 1990).

A second possible source of inoculum includes survival of the pathogen on fiddleleaf tobacco (*Nicotiana repanda* Willd.). This is an annual herb native to central and southern Texas and adjacent regions of Mexico (Reuveni *et al*. 1988, Wolf 1947). It has also been postulated using atmospheric trajectory analysis that Texas could be the origin for blue mold spores in some years in North Carolina (Davis and Main 1990).

A third possible source of inoculum could include survival of the pathogen as oospores on roots, stems or host debris in the field (Heist *et al*. 2002, Heist *et al*. 2001, Spurr and Todd 1982, Lucas 1980). The significance of oospore production in the epidemiology of blue mold is unclear since oospores are difficult to germinate and have not been recently observed in tobacco leaf tissue. Sexual reproduction could lead to greater genetic variation in populations than strictly asexual reproduction (LaMondia and Aylor 2001, Person *et al*. 1955).

The pathogen can also survive and spread in infected tobacco seedling beds or in transplants in the greenhouses. Since the symptoms and signs of the disease

are not visible immediately after infection and systemic infections can occur, infected seedlings could also be an important inoculum source involved in spread of the pathogen between tobacco growing regions.

The current North American Plant Disease Forecast Center, located at North Carolina State University was initially established as the North American Blue Mold Warning System after the blue mold epidemics in 1979. The center reports and predicts blue mold outbreaks in the U.S. (Nesmith 1990, Davis and Main 1986, Nesmith 1984). The North American Blue Mold Warning System was modeled after the European Blue Mold Warning Service developed by CORESTA (Centre for Cooperation in Scientific Research Relative to Tobacco) for the Euro-Mediterranean zone in 1961 (Ledez 1990). Both centers use atmospheric satellite weather data to examine, integrate and predict the movement of the pathogen from one place to another, in time and geographically. However, validation of the migration patterns predicted by the forecasting system with genetic data from potential source and sink populations of *P. tabacina*, has not been done.

Little is known about the population genetics of *P. tabacina,* mainly due to its obligate nature and difficulty working with the parasite. Some molecular markers have been used to study the genetic diversity in *P. tabacina.* Edreva (1998) and Micales *et al*. (1986), used isoenzymes electrophoresis to study isolates from France and Bulgaria and the U.S. and found little genetic variability. Sukno *et al*. (2002)

used restriction fragment length polymorphisms (RFLPs) markers to analyze 34 isolates in a U.S. population and differentiated 10 haplotypes. Her work also suggested low levels of genetic variability among U.S. populations of *P. tabacina*.

The mitochondrial gene, *cox*2 has been used extensively for phylogenetic studies within the Peronosporomycetes and Oomycetes (Hudspeth *et al*. 2003 and Cook *et al*. 2001, respectively), but not for population genetics.

Phenotypic changes in the population structure of *P. tabacina* suggest that the pathogen may be more genetically diverse than previous studies suggest. Metalaxyl was introduced for control of *P. tabacina* in the seedbeds and fields in 1978. Resistance was soon recorded in Nicaragua in 1981 and in Honduras, Mexico, Cuba (Johnson 1989) and the U.S. between 1982 and 1986 (Shoemaker *et al*. 2000, Wiglesworth *et al*. 1988, Bruck *et al*. 1982). Some populations have shifted from heat sensitive to heat tolerant stains. There is a recent report from Nicaragua in 2006 of the occurrence of strains that can grow at temperatures between 30-33° C (Personal communication with Mr. Nestor Placencia). Systemic infections have been reported and it is possible that the pathogen may remain latent in tobacco tissue until environmental conditions become favorable for symptom expression and sporulation (Caiazzo *et al*. 2006). Changes in pathogen aggressiveness have also been reported with greater disease severity among isolates from Texas than Kentucky (Reuveni, 1988). Pathotypes have also been reported on various tobacco

differentials (Johnson, 1989).

We examined the global genetic structure of populations of *P tabacina* in order to test several hypotheses including: the long distance aerial dispersal of the pathogen and migration from the Caribbean region into the U.S.; and migration from the Mediterranean region into Europe. The objectives of this study were to analyze the population structure of *P. tabacina* from geographically different populations using mitochondrial and nuclear genes in order to 1) estimate genetic variability in pathogen populations and perform neutrality tests, 2) estimate the contribution of gene flow and directional migration to the observed genetic structure and to 3) generate coalescent trees to identify possible ancestral populations of *P. tabacina*.

## MATERIALS AND METHODS

### *P. tabacina* sampling strategy

A collection of 260 isolates from *Peronospora tabacina* was sampled nonrandomly. The collection is housed in the Department of Plant Pathology at North Carolina State University (NCSU) and is located in the Phytotron facility in a BS Level 2 containment laboratory. Fifty four isolates of *P. tabacina* from the US, Central America, the Caribbean, Europe and the Middle East were studied. Thirty isolates from the U.S. (Connecticut, Florida, Georgia, Kentucky, Maryland, North

Carolina, Pennsylvania, Texas and Virginia), 13 isolates from the Caribbean (Dominican Republic), Central America (Guatemala and Nicaragua) and Mexico and 11 isolates from Europe (Bulgaria, France, Germany, Hungary and Poland) and the Middle East (Lebanon) (Table 2.1) were used .

The nuclear encoded ribosomal DNA genes (rDNA) exist as a multiple-copy gene family comprised of highly similar DNA sequences. Previous studies on Oomycetes have shown the internal transcribed spacer (ITS) regions of the rDNA to be useful for the differentiation at the species level (Voglmayr 2003, Petersen and Rosendahl 2000, Cooke *et al*. 2000). This region is useful for species specific identification of *P tabacina* but is not useful for differentiation within populations of *P. tabacina* (Ristaino *et al*. 2007, Wiglesworth 1994). The identity of each isolate used in our study was confirmed as *P. tabacina* by using PCR of ITS DNA using methods described in Ristaino *et al*. 2007 and a diagnostic Real time PCR reaction (Chapter 1).

**DNA extraction**

Thirty five isolates of *P. tabacina* were inoculated onto fresh tobacco leaves (*Nicotiana repanda)* produced *in vitro* to increase the amount of sporangiospores and to collect sporangiospores free of external contaminant's (Table 2.1).

Total genomic DNA was extracted from sporangiospores and sporangiophores collected from the leaves using the standard cetyltrimethylammonium bromide (CTAB) protocol (Trout *et al*. 1997) with some modifications (Wangsomboondee and Ristaino 2002). If sporangiospores were no longer viable, and inoculation on the tobacco plants was not possible, DNA extraction was performed from the spores stored in cryogenic solutions. The spores were rinsed 4 times with autoclaved distilled water and centrifuged between each rinse. Sporangiospores and mycelia were placed in sterile 1.5-ml microcentrifuge tubes containing 0.2 g of glass beads to which 150 µl of extraction buffer (0.35 M sorbitol, 0.1 M Tris, 0.005 M EDTA, pH 7.5, and 0.02 M sodium bisulfite) was added. Samples were macerated using a Konte pestle attached to a drill. Nuclei lyses buffer (150 µl; 0.2 M Tris, 0.05 M EDTA, pH 7.5, 2.0 M NaCl, 2% CTAB, and 60 µl of 5% sarkosyl [N-Laurylsarcosine]) was added and tubes were vortexed and incubated at 65ºC for 15 to 30 min. After incubation, one volume (300 µl) of chloroform: isoamyl alcohol (24:1) was added to each tube and tubes were centrifuged for 15 min at 13,000 x *g* at room temperature. The aqueous phase was removed to a new tube and chloroform extraction was repeated. DNA was precipitated overnight at –20ºC in 0.1 volume of 3M sodium acetate (pH 8.0) and 2 volumes of cold 100% ethanol. The supernatant was discarded; pellets were washed with 70% ethanol and dried at room temperature. DNA was suspended in Tris-EDTA

(TE; 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). DNA was diluted 1:10 (100-200 ng/µl) for further use. Each DNA extraction was checked for the presence of *P. tabacina* DNA using the specific PCR primer PTAB (Ristaino *et al*. 2007) and a real time diagnostic primer PtabBM TaqMan (Blanco-Meneses *et al*., Chapter 1).

**DNA amplification, TA cloning and sequencing**

Initially, 66 different primer pairs from previous published studies of Oomycete pathogens were tested for amplification of *P. tabacina* (Appendix 2.1). Two different thermal cycling parameters were used including either 1) an initial denaturation at 96ºC for 2 min, followed by 35 cycles consisting of denaturation at 96ºC for 1 min, annealing at the specific primers temperature for 1 min, and extension at 72ºC for 2 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min (Ristaino *et al*. 2007) or 2) an initial denaturation at 95ºC for 2 min, followed by 35 cycles consisting of denaturation at 94ºC for 30 s, annealing at the specific primers temperature for 30 s, and extension at 72ºC for 1 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min (Schena and Cooke 2006). Choice of thermal cycling parameters was determined considering the size of the target DNA to be amplified and temperature requirements for each set of primers to anneal with the template DNA. PCR products were analyzed on 1.6% agarose gels. Thirty six of the pair primer amplified the target DNA in *P. tabacina* (Appendix 2.1). Fifteen

different amplified targets generated by the primer combinations were sequenced using 15 isolates of *P. tabacina* and sequences from two isolates were submitted to Gene Bank (Appendix 2.1).

Two nuclear regions including the intergenic spacer *Igs2* (primers Igs3F: 5'-GYGCGAAGGWKTGCTG-3' and Igs7R: 5'-ATATCCTCCATACGWAAGAAGACG-3') region of the nuclear ribosomal DNA (rDNA), and the Ras-related protein (*Ypt1*) gene (primers Ypt1F: 5'-CGACCATYGGYGTKGACTTT-3' and Ypt5R: 5'-GCAGC TTGTTSACGTTCTCR-3'), and the mitochondrial cytochrome c oxidase subunit 2, *cox*2 gene (primers FM78: 5'-ACAAATTTCACTACATTGTCC-3' and FM82: 5'-TTGGCAATTAGGTTTTCAAGATCC-3') (Appendix 2.1) were chosen for subsequent amplification, cloning and sequencing. Primer pair sequences and thermocycling protocols are shown in Appendix 2.1. Oomycetes are diploid organisms and can show disomic inheritance, such that individuals can be heterozygous or homozygous at some loci (Tooley and Therrien 1991). The presence of heterozygous coincidence peaks in the forward and reverse sequence chromatograms were examined for the three gene regions (Figure 2.1).

Cloning was performed using a PGEM-T easy vector TA cloning kit (Promega, Madison, Wisconsin). Reactions were incubated overnight to obtain a maximum number of transformants. Transformation efficiency was higher than $1x10^8$ colonies (JM109 Competent Cells (Promega, Madison, Wisconsin). The cells were

plated on LB (Luria-Bertani) medium (Difco, Franklin Lakes, New Jersey) on Petri plates supplemented with 20 µg/µl – 100 µg/µl ampicillim (Sigma, St. Louis, Missouri) and 100 µl of X-Gal/IPTG 1000X Stock Solution (Bioline, Taunton, Massachusetts) and stored overnight at 37°C. Ten white colonies were selected and the QIAprep Spin Miniprep Kit (QIAGEN, Valencia, California) was used for the purification of the plasmid DNA on 96-plate plates using a Centrifuge 5810R (Eppendorf, Hamburg, Germany). DNA was stored at -20°C for further use and long term storage. Each of the colonies was also stored in sterile glycerol (15%) at -70°C. Ten clones of each isolate (n=54) were randomly selected for sequencing and sequence from 6 of 10 clones were used in the analysis. A total of 324 clones were analyzed.

Sequencing reactions for six clones from each of the 54 isolates were prepared by using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit and analyzed on an ABI PRISM 377 automated sequencer (Applied Biosystem, Foster City, California) at the North Carolina State University Genome Research Laboratory.

**Sequence analysis**

Multiple DNA sequences were aligned and edited manually using either BioEdit Sequence Alignment Editor Version 7.0.5.3 (Hall 1999) or Clustal X

(Thompson *et al*. 1994) from Snap Workbench version 2.0 (Price and Carbone 2005, Aylor *et al*. 2006). All polymorphisms were rechecked from the chromatograms.

A large number of single nucleotide polymorphism (SNPs) were found within isolates. Five isolates of *P. tabacina* isolates (123046, 123344, BU061, TX062, FR184; Table 2.1) were used and amplification of the target DNA, cloning and sequencing was repeated. Sequence data from six clones from each isolate was analyzed for the presence in both the forward and the reverse sequences of the same polymorphisms to exclude the possibility of sequencing error due to the *Taq* DNA Polymerase (Invitrogen, Carlsbad, CA). Repeatable sequence polymorphisms were observed.

**Statistical analysis**

The total population was divided in 4 subpopulations for analysis including: 1) U.S. (all isolates from U.S. pooled); 2) CCAM (isolates from Mexico, Guatemala, Dominican Republic and Nicaragua); 3) EULE (isolates from Bulgaria, France, Germany, Hungary, Lebanon and Poland) and 4) U.S. vs. CCAM (clones from isolates from each U.S. state vs. clones from each isolate from CCAM countries) (Table 2.2).

Statistical analyses were performed using the SNAP (Suite of Nucleotide Analysis Programs) Workbench version 2.0 (Price and Carbon 2005, Aylor *et al*.

2006). The SNAP Workbench is a Java program that manages and coordinates a series of analysis programs for making inferences on population processes. SNAP Workbench allows the user to customize the implementation of complex console programs and functions for the purpose of automating and enhancing data exploration. Population processes can be examined and integrate both summary-statistic methods and coalescent-based population genetic models.

**Population structure**

Each gene or region was analyzed separately for each subpopulation for each of the tests. Cloned sequences were formatted using the Phylogeny Inference Package (PHYLIP) format (Felsenstein 1989).

Sequence components can combine (usually through recombination or gene conversion) from more than one source to form a new hybrid sequence through the process of reticulated evolution. When sequences have evolved in a reticulate fashion, different parts of the sequences can be aligned in relatively high or low phylogenetic compatibility. Two sites in a set of aligned nucleotide sequences are defined to be compatible if there exist a possible evolutionary history of the sequences in which all nucleotide changes at both sites can be inferred to have occurred only once. Sites are incompatible when one or both sites must have experienced repeated mutation or been involved in recombination or gene transfer

(Jakobsen *et al*. 1997, Jakobsen and Easteal 1996). Initially, sequences were collapsed into haplotypes removing indels and excluding infinite site violations using SNAP Map (Aylor *et al*. 2006) in SNAP Workbench version 2.0 (Price and Carbon 2005; Aylor *et al*. 2006). A haplotype is a specific sequence of alleles or SNPs that is a useful way for grouping individuals according to genotype (Aylor *et al*. 2006). Incompatible sites were removed from the sequences using CladeEx (Bowden *et al*. 2008) in the SNAP Workbench platform.

After incompatible sites were removed, the sequences were collapsed with phenotypes into haplotypes removing insertion and deletions (indels) and excluding infinite-sites violations using the SNAP Map (Aylor *et al*. 2006). Base substitutions were categorized as phylogenetically informative or uninformative and transitions vs. transversions and each of the polymorphic sites received a specific site number on the consensus sequence.

**Nucleotide diversity, neutrality tests and population subdivision**

DnaSP version 4.50.3 (Rozas *et al*. 2003) and the Neutrality Test Program (Li and Fu 2002) were used to analyze nucleotide diversity estimates of the DNA sequences. Analysis was performed on the total population, separate regional subpopulations (U.S., CCAM and EULE subpopulations) and the states in the U.S. subpopulation. Each of the two nuclear regions and the mitochondrial gene were

analyzed separately. For each analysis the population mutation parameter per nucleotide site θ using Watterson's (1975) estimate $\theta_w$, based on the number of segregating sites, s, and the average pairwise nucleotide diversity, π (Tajima 1983) were estimated. Nucleotide diversity is the average number of nucleotide differences per site between two sequences (Nei 1987). The Watterson's estimate is the number of segregating sites observed in a sample from a population which is subject to recurring new mutations but not subject to recombination (Watterson 1975). Neutrality test statistics were calculated with Tajima's *D* (Tajima 1989), Fu and Li's *D** and *F** (Fu and Li 1993), and Fu's Fs (Fu 1997) and tested using 5,000 simulations to accept or reject the hypothesis that mutations in these regions follow the neutral model of molecular evolution (Kimura 1980). These neutrality tests assume a constant population size, no recombination and no migration and that each isolate is taken from a single randomly mating population and random sampling, that is, the population evolves according to the Wright-Fisher model. Departures of neutrality may be caused by evolutionary forces such as population subdivision, population shrinkage, over-dominance selection, population growth, genetic hitchhiking, and background selection (Fu 1997). Carbone *et al*. (2004) and Carbone *et al*. (2007) examined population growth and background selection by comparing Tajima's D, Fu and Li's D* and F* and Fu's *Fs* tests. If only Tajima's D, Fu and Li's D* and F* are significant and of Fu's Fs is nonsignificant; background

selection is assumed (Fu 1997). If only Fu's *Fs* is significant there is evidence for population growth or genetic hitchhiking. Nonsignificant test can be also informative; when D and Fs tend to be uniformly negative there is an excess of rare variants, and indication of genetic hitchhiking / selective sweep or population growth, and tend to be uniformly positive when there is an excess of intermediate-frequency alleles, indicating a model under population subdivision or balancing selection. The neutrality tests provide an inference on population processes and enable detection of deviations in population size constancy that may result in departures of neutrality (Carbone *et al*. 2007).

Nonparametric tests for population subdivision were analyzed using Hudson's test, the nearest neighbor statistics $S_{nn}$ (Hudson *et al*. 1992a, Hudson *et al*. 1992b, Hudson 2000) on the Snap Workbench platform. This statistic uses the infinite sites model to examine how often the nearest sequence neighbors are from the same populations. The $S_{nn}$ statistic allows pooling different populations into genetically subdivided groups of regions. To calculate a significance value for the $S_{nn}$ statistic a permutation test of 1000-pseudoreplicates was performed. When $S_{nn} = 1$, the two or more populations are highly genetically differentiated, whereas an $S_{nn}$ near 0.5 indicates that the two or more populations are part of one panmictic population.

At first, analysis was preformed for the U.S. and the CCAM subpopulations. The three gene regions were analyzed separately and population pooling was

performed when the non-significant values were supported by the three regions. The $S_{nn}$ test was repeated a second and third time until it was impossible to pool the populations further. A second analysis was performed for the EULE subpopulation and the same procedure for pooling was performed.

**Migration and Coalescent analysis**

Once it was demonstrated, using the $S_{nn}$ approach that the different populations of *P. tabacina* were subdivided and the boundaries of the subpopulations were determined, the observed patterns of genetic divergence were further explored to determine whether there was evidence for constant migration among populations and shared ancestral polymorphism.

The MDIV program was used to test for equilibrium migration vs. shared ancestral polymorphisms between two subdivided populations (Nielsen and Wakeley 2001). This approach uses either an infinite sites or finite sites model without recombination and implements both likelihood and Bayesian methods using a Markov chain Monte Carlo coalescent simulation for jointly estimating the population mean mutation rate (θ), divergence time (T), migration rate (M), and time since the most recent common ancestor (TMRCA) between two subdivided populations. This approach assumes that populations descended from one panmictic population that may (or may not) have been followed by migration.

If MDIV showed evidence of equilibrium migration to estimate population divergence time, population mean mutation rate and direction of migration the Isolation with migration (IM) program (Nielsen and Wakeley 2001, Hey and Nielsen 2004) within Snap Workbench platform was used. Both MDIV and IM assume neutrality and no recombination. Under the null hypothesis of isolation, an expected migration near 0 would be expected. To run the IM program the following parameters were used, one million numbers of steps for the burn in duration, 12 and 24 hours were used between outputs and a 20 value was used for the maximum migration rate and the maximum time of population splitting. Multiple chains (total of 6 different runs for each two populations compared) were run using different starting values for the parameters. If all chains give similar results, it was assumed as good evidence to support the data.

To place migrations events in time the tree with the highest root probability was inferred from the coalescent using Genetree (Version 9.0) (Griffiths and Tavaré 1994, Bahlo and Griffiths 2000). This program provides a maximum-likelihood estimate of the tree with the highest root probability, and estimate of the population mean mutation rate (θ), and the time to the most recent common ancestor (TMRCA), in coalescent time units of $N_e$ generations for the entire tree, where $N_e$ is the effective population size (Carbone and Kohn 2001b). Coalescent simulations were performed assuming an infinite-site model, constant population size and populations

subdivision. Gene genealogies for each subpopulation were inferred using one million simulations of the coalescent. Additionally we performed three independent runs of one million simulations using a different starting random number seed for each run to endure convergence.

## RESULTS

### DNA sequence variability

A total of 435 nucleotides were sequenced for the intergenic spacer (*Igs*2) region of the rDNA. A total of 493 nucleotides were sequenced and analyzed for the Ras-related protein (*Ypt*1) gene (primers Ypt1F/Ypt5R) and a total of 631 nucleotides for the mitochondrial genome region that covers cytochrome c oxidase subunit 2, *cox*2 gene. The number of segregating nucleotide sites were 116 nucleotides for the *Igs*2 region (26.67%), 126 nucleotides for the *Ypt*1 region (25.56%) and 107 nucleotides for the *cox*2 gene (16.96%) (Table 2.2). A total of 324 clones for the *Igs2* region and *cox*2 gene and 306 clones for the *Ypt1* region were sequenced and analyzed including 6 clones from each of the 54 isolates from the populations (Table 2.1).

Summary statistics describing the nucleotide diversity in the pooled and individual populations for both nuclear and mitochondrial loci are presented in Table

2.2. Nucleotide diversity (π) estimates for the total population were $8.11X10^{-3}$ for the *Igs2* region, $10.28X10^{-3}$ for the *Ypt1* region and $7.66X10^{-3}$ for the *cox*2 gene. The mean mutation parameter per site or Watterson's estimate, $\theta_w$, for the total population was 0.157 for the *Igs2* region, 0.159 for the *Ypt1* region and 0.157 for the *cox*2 gene.

When three subpopulations (U.S., CCAM and EULE) were compared, nucleotide diversity at the *Igs*2 region and the *Ypt*1 region were higher for the EULE ($53.04x10^{-3}$ and $71.09 x10^{-3}$) and CCAM ($42.4 x10^{-3}$ and $64.17 x 10^{-3}$ subpopulations than U.S. subpopulations ($20.48 x10^{-3}$ and $15.68 x10^{-3}$) (Table 2.2). However nucleotide diversity for the mitochondrial *cox*2 gene was highest for the CCAM ($36.83X10^{-3}$) subpopulation. Similarly, $\theta_w$ was higher for the EULE and CCAM populations than U.S. populations for both nuclear regions. Waterson' $\theta_w$ for the *cox*2 gene was higher for CCAM than U.S. and EULE subpopulations.

Within North Carolina, when NC-Burley and NC-Flue-cured subpopulations were compared the nucleotide diversity was higher in the NC-Burley for the *Ypt*1 region ($103.88X10^{3}$) but was higher for the NC-Flue-cured population for the *Igs*2 region ($91.87X10^{3}$) and the *cox*2 gene ($1.09X10^{-3}$) (Table 2.2). Little difference in $\theta_w$ was observed for NC-Burley and NC-Flue-cured subpopulations for any gene analyzed (Table 2.2).

Nucleotide diversity (π) was also estimated for the U.S. populations

separately. Since the population sampled from North Carolina was larger than the populations from other states, we examined the effect of population size on the nucleotide diversity estimates (Appendix 2.2). Populations from North Carolina (n=72) had lower nucleotide diversity estimates for the nuclear gene regions than populations from many other states. Waterson's $\theta_w$ for the North Carolina subpopulation was intermediate in value for the nuclear gene regions analyzed, but highest for the *cox*2 gene when compared with isolates from the other tobacco states (Appendix 2.2). Interestingly, the highest mitochondrial diversity occurred in Texas populations, which are the putative overwintering source in the U.S.

**Tests of neutrality**

Four neutrality tests were performed to determine whether the data departed from an equilibrium model of neutral evolution. Highly significant and negative values were observed for all the neutrality tests for the three subpopulations indicating that the neutral model of evolution could be rejected (Table 2.2). For the total population, Fu's Fs was not significant for the *lgs2* and *cox2* genes indicating background selection. The negative values for all of the tests indicate an excess of recent mutations or rare alleles in the populations, an indication of genetic hitchhiking, selective sweep or population growth.

**Population genetic structure**

Snap Map was used to collapse sample sequences into haplotypes. There were a high number of low frequency haplotypes in the populations. There were 93 haplotypes identified for the nuclear genes *Igs2* and *Ypt1*, and 82 haplotypes identified for the *cox*2 gene (Table 2.2, Appendix 2.3, 2.4 and 2.5).

For the *Igs2* region, one high frequency haplotype (H90) dominated the populations. Ninety-three different haplotypes were identified of which 15 were frequent haplotypes (present in 2 or more clones) and 78 were less frequent haplotypes (presented in only 1 clone). The total haplotypes were differentiated by 116 single nucleotide polymorphism (SNPs), 104 transitions and 12 transversion substitutions. Eight transition substitutions were informative for the data analysis and were located at sites 9, 32, 59, 62, 71, 83, 108 and 112 and only one transversion substitutions were informative and located at the 69 site of the consensus for the *Igs2* region (Appendix 2.3).

For the *Igs2* region, the predominant haplotype H90 was present in 209 of the 324 clones (64.51%). Greatest numbers of haplotypes were found in the North Carolina (22.58%) and Kentucky (11.83%) subpopulations; the other tobacco producing states had between 2 and 5 (5.38% and 2.15%) haplotypes. Within the CCAM population, 14 of the haplotypes were present in Dominican Republic (15.05%) and 12 in Mexico (12.90%). Within EULE, 11 of the haplotypes were

present in France (11.83%) and between 1 and 6 haplotypes (1.08% and 6.45%) were found in the other EULE countries (Appendix 2.6 and 2.7). North Carolina had the highest number of unique haplotypes (15 of 79, 18.99%), followed by Kentucky (11 of 79 haplotypes, 13.92%), Mexico (10 of 79, 12.66%), Dominican Republic (9 of 79, 11.39%) and France (8 of 79, 10.13%). The number of unique haplotypes for the other regions were lower (between 4 and 0; 5.06% and 0%) (Table S2.6).

For the *Ypt1* region, a total of 93 different haplotypes were found, 17 were frequent and 76 less frequent haplotypes. Similarly one high frequency haplotype was predominant (H1). The total haplotypes were differentiated by 126 SNPs, of which 107 were transitions and 19 were transversion substitutions. Eight transition substitutions were informative for the data analysis and were located on the 9, 25, 26, 57, 58, 72, 98, and 101 and only one transversion substitutions were informative and located at the 96 site of the consensus for the of the *Ypt1* region (Appendix 2.4).

The predominant haplotype named H1 was present in 176 of the 306 clones (57.52%) of the *Ypt1* region and was found in all of the subpopulations studied. Within the USA population, 30 of the 93 haplotypes were present in North Carolina (32.26%), and 6 were found in Texas, Pennsylvania, Connecticut, and Georgia (6.45% each). The other tobacco producing states had between 2 and 4 (2.15 - 4.30%) haplotypes. Within the CCAM population, 16 of the haplotypes were present

in Dominican Republic (17.20%) and 6 of the haplotypes were in Mexico (6.45%). Within EULE, 6 of the haplotypes were present in France, Bulgaria and Poland, respectively (6.45% each) (Appendix 2.6 and 2.7). The highest number of unique haplotypes were found in North Carolina (24 of 76, 31.58%), followed by Dominican Republic (13 of 76, 17.10%). The number of unique haplotypes for the other regions were low (0 - 5; 0-6.58%) (Table S2.6).

The *cox2* gene sequence was the less variable when compared with the other two gene regions. A total of 82 haplotypes were found for this gene, 8 were frequent and 74 less frequent haplotypes. Similarly one high frequency haplotype was predominant (H-1). The total haplotypes were characterized by 107 SNPs, of which 87 were transition and 20 were transversion substitutions. Six transition substitutions were informative for the data analysis and were located on the 16, 31, 33, 39, 46 and 107 and 2 transversion substitutions were informative and located at the 1 and 5 sites of the consensus for the *cox2* gene (Table S2.4).

The predominant haplotype, H1 was present in 235 of the 324 clones (75.53%) of the *cox2* gene and it was present in every geographical location analyzed. Within the U.S population, 16 of the 82 haplotypes were present in North Carolina (19.51%), 7 in Pennsylvania and Texas (8.54%) and 6 in Kentucky (7.32%). The other tobacco producing states had between 1 and 5 (6.10% and 12.19%) haplotypes (Appendix 2.6 and 2.7). Within the CCAM population, 15 of the

haplotypes were present in Dominican Republic (17.20%), 8 in Nicaragua (9.76%) and 5 in Mexico (6.10%). Within EULE, 9 of the haplotypes were present in France (10.98%) and 4 in Germany (4.88%) (Table S2.6). The highest number of unique haplotypes for the *cox2* gene, were found in North Carolina (15 of 74, 20.27%), followed by Dominican Republic (10 of 74, 13.51%) and France (7 of 74 haplotypes, 9.46%). The number of unique haplotypes for the other regions was low (between 0 and 6, 0-8.10%) (Table S2.6).

**Population subdivision**

Hudson's tests $S_{nn}$ were performed to examine population structure and subdivision within and among populations (Hudson *et al*. 1992). Sequences from the U.S. and the CCAM subpopulations were analyzed together and sequences from Europe were analyzed separately. Populations from Mexico, Guatemala, Nicaragua and the Dominican Republic in the CCAM region were not genetically subdivided from each other for the two nuclear and one mitochondrial regions examined and the sequences were pooled together for the CCAM region (Table 2.3, Figure 2.2). The Hudson's test was performed a second time and populations from N. Carolina and Kentucky, Pennsylvania, Maryland and Connecticut were not genetically subdivided so they were pooled together (Table 2.4, Figure 2.2). The $S_{nn}$ test was performed a third time and the populations were classified into the North U.S (N. Carolina,

Kentucky, Georgia, Virginia, Pennsylvania, Maryland and Connecticut) and South U.S./CCAM (Texas, Florida and the CCAM) (Table 2.5, Figure 2.2). The $S_{nn}$ test was perform a fourth time, and the North and South populations were genetically subdivided for the Ypt1/4 region and the *cox*2 gene but not for the *Igs*2 region (Table 2.6).

The EULE population was also analyzed using the Hudson's $S_{nn}$ test. Initially, the population was pooled into 4 different groups; Northern Europe (Poland and Germany), Central Europe (Bulgaria and Hungary), Western Europe (France) and Lebanon (Table 2.7, Figure 2.3). The test was performed again and North and Central Europe were not subdivided so the data was pooled (Table 2.8, Figure 2.3). When the test were performed a third time, North/Central Europe was genetically subdivided from Western Europe for the *Igs*2 and the Ypt1/5 regions but not for the *cox*2 gene. Both populations were not genetically subdivided from Lebanon for all three gene regions (Table 2.9).

**Migration analysis**

The MDIV program estimates the population mean mutation rate (θ), divergence time (T), migration rate (M), and time since the most recent common ancestor (TMRCA) between two subdivided populations (Appendix 2.8). The direction of migration for the populations was examined using the Isolation with

Migration program (Nielsen and Wakeley 2001, Hey and Nielsen 2004). When the U.S./CCAM populations were compared, migration from the South U.S./CCAM to the North U.S. was supported by the two nuclear gene regions and the mitochondrial gene. The migration rates from one population to another were very low. These results were supported by the two nuclear region, *Igs*2 (North $m_1$= 0.67, South $m_2$= 1.08) and *Ypt*1 (North $m_1$= 0.26, South $m_2$= 0.27); and the mitochondrial region (North $m_1$= 0.90, South $m_2$= 1.02) (Figure 2.4). Due to the large number of different haplotypes and the effective population size for the U.S.-CCAM subpopulation coalescent analysis trees were not generated by the programs.

The three subdivided geographic regions were analyzed in the EULE population: NCE, WE and LE. Migration rates were higher between populations in these regions than between populations in the U.S. When direction of migration was tested between NCE and WE, migration from the NCE population ($m_1$=7.29) to the WE population ($m_2$=6.94) was identified. When direction of migration was tested between NCE and LE, migration was identified from NCE ($m_1$=10.89) to LE ($m_2$=7.16) and finally when WE and LE populations were analyzed, migration from WE ($m_1$=9.09) into LE ($m_2$=6.66) was identified. These results indicated that there is migration occurring from the Northern part of Europe (NCE) into France and Lebanon, and from France into Lebanon as well (Figure 2.4).

## DISCUSSION

We analyzed the genetic structure of populations of *Peronospora tabacina* from a global collection of isolates from 11 countries in three geographic regions including the U.S., Central America, the Caribbean and Mexico (CCAM) and Europe and Lebanon (EULE). This work is the first to use DNA sequence variation to study the population structure, gene flow and directions of migrations of an obligate pathogenic Oomycete. We used three primer sets that amplified phylogenetically informative sites in both nuclear and mitochondrial genes. Nucleotide sequences from the three genes showed high genetic variability across all populations which were dominated for a single high frequency haplotype and many rare haplotypes.

There are only two previous studies that examined the genetic structure of populations of *P. tabacina.* Sukno *et al*. 2002 used RFLPs to analyze 35 isolates from 9 tobacco states in the U.S. and found several polymorphisms that allowed differentiation of the populations into at least 10 haplotypes. Seven haplotypes were found in Kentucky while other U.S. regions had fewer haplotypes. Edreva *et al*. (1998) used 3 isoenzymes to examine variation in a collection of isolates of *P. tabacina* from France and Bulgaria. Genetic uniformity among isolates collected from *N. tabacum* was reported in that study but variation in allozyme alleles as a response to host plant was only observed when isolates collected from *N. tabacum* were passed to *N. repanda*. Our data do not support the previous studies which

indicated low levels of genetic variation in populations of the pathogen.

High levels of genetic variation were found among populations of *P. tabacina* when DNA from allelic regions of several loci were cloned and sequenced. Single nucleotide polymorphisms were found in different positions along the DNA, which were resolved using sequencing techniques. Several possible mechanisms could explain the high genetic variation observed in this pathogen. The most common form of dispersal of this pathogen is aerial, long distance dispersal from putative overwintering sites in Texas and the Caribbean. Long-distance dispersal is an important survival strategy for many fungal pathogens including rusts, powdery mildews and downy mildews, enabling them to colonize new territory rapidly or to migrate between summer and winter habitats (Brown and Hovmoller 2002). The production of huge numbers of spores in each polycyclic event is essential for reproduction and survival because the pathogen is completely dependent on living host tissue. *P. tabacina* causes a polycyclic disease and aerial dispersal over long distances and expansive population growth occurs after infection. *P. tabacina* could be considered a high evolutionary risk plant pathogen due to its large effective population size and mechanism of dispersal (McDonald and Linde, 2002). Population size affects the probability that mutants will be present and can influence diversity of genes. Also aplicacion of fungicides in the U.S. subpopulations occurs each year. Fungicides are used in the *P. tabacina* populations for the black shank

pathogen (*Phytophtora tabacum*) mainly but at the same time help to protect the plant against the blue mold. Remaining populations of *P. tabacina* survive everyyear increasing the probability of mutations and variation in the population.

Neutrality tests were significant for all populations and the equilibrium model of neutral evolution was rejected, indicating an excess of recent mutations or rare alleles. Tests of neutrality including Tajima's D, Fu and Li's D*, F* and Fu's Fs showed significant and negative values for the subpopulations. The significant values indicated a departure from neutrality, assuming that the population does not have a constant population size and recombination and migration could be occurring. The negative values indicate an excess of recent mutations in the population. It is interesting to note that the Fu's Fs statistic test was not significant for the analysis of the total population for the *lgs*2 region and *cox*2 gene. This could indicate a process of background selection. In cases where there is strong negative selection on a locus combined with a high mutation rate, the frequent purging of deleterious variants may result in the occasional removal of linked variation, producing a decrease in the level of variation surrounding the locus under selection. The result is the accidental purging of non-deleterious alleles due to such spatial proximity to deleterious alleles (Fu and Li 1993, Fu 1997).

**Migrations from Central America and the Caribbean to North America**

Our data presented here confirm previous data from forecasting models and support the hypothesis of long distance dispersal of this pathogen from the Caribbean region, Florida and Texas into the tobacco growing states further north (Aylor and Taylor, 1983, Davis and Main, 1986). For many years the Tobacco Blue Mold Forecasting System at N.C. State University has been predicting the trajectory of tobacco blue mold epidemics in tobacco growing regions in the U.S., based on weather data and reports of source populations in the Caribbean (Davis and Monahan, 1991; Main et al, 2001). Our work, using genetic analysis of both nuclear and mitochondrial genes provides the first genetic evidence to support the long-distance migration theory of *P. tabacina*.

We used Hudson's statistics to examine population subdivision and gene flow within the CCAM. Subpopulations within this region were not subdivided and gene flow was documented. Interestingly, subpopulations in Texas and Florida were not subdivided from subpopulations in CCAM and were pooled into a South population. The other U.S. subpopulations were genetically subdivided from CCAM subpopulations and divided into two populations: North and South. An isolation-with-migration analysis (IM) indicated migration from South to North supporting the hypothesis of long-distance migration of *P. tabacina* from the Caribbean regions, Florida and Texas into other states further north. These data also support previous

reports of disease occurrence first in the CCAM region (including Cuba) and later as the green tobacco belt is planted further north (Nesmith 1990, Davis and Main 1986, Nesmith 1984). Due to the embargo, we were unable to collect and genotype isolates from Cuba so they are not included among isolates from the CCAM region.

These genetic results agree with previous reports that inoculum of *P. tabacina* is known to survive on wild *Nicotiana* repanda in Texas (Reuveni *et al*. 1988) and overwinter and form oospores in Florida (McGrath and Miller 1958). Both these sources of inoculum can contribute to epidemics in regions further north.

Nucleotide diversity ($\pi$) and the Watterson's estimate ($\theta_w$) was also higher for the CCAM subpopulation for all the gene regions examined. Subpopulations in the Dominican Republic for the CCAM and North Carolina from the U.S. had the highest mean mutation rate (Watterson's estimate), the highest total number of haplotypes and the highest number of unique haplotypes among all populations examined. The CCAM region is clearly a center of diversity for *P. tabacina* and provides the source of inoculum for blue mold epidemics in North Carolina.

High levels of genetic diversity have been reported in other Oomycete pathogens such as *Phytophthora* spp. (Judelson 2002, Chamnanpunt *et al*. 2001, Francis and St. Clair 1997) and this diversity has been attributed to factors such as genetic instability, transposable elements, gene conversion, mitotic recombination and/or dispensable chromosomes (Kamoun 2003). Since nucleotide diversity is the

average number of pairwise nucleotide differences between sequences in a sample and the Watterson's estimate is based on the number of polymorphic sites in a sample, π depends on both the number of polymorphic sites and their frequency, whereas $\theta_w$ is independent of frequencies (Watterson 1975, Nei and Li 1979). A population that has a high number of haplotypes could inflate the estimate for the $\theta_w$, as was observed in our data.

Dispersal of *P. tabacina* occurs every year through the U.S. Our data indicate that gene flow is occurring between Central American and Caribbean regions and migration is occurring to the U.S. CCAM could serve as a reservoir for inoculum since the region has the appropriate weather for the development of the pathogen and tobacco is grown there year round. Tobacco grown in Florida and Texas could also serve as reservoirs for pathogen inoculum. Dispersal of the pathogen on transplant material has been reported in the last 2 years (2007-2008) (Dr. K. Seebold, personal communication), however, we did not analyze isolates specifically from transplant material to differentiate those sources in this study.

**European migrations**

When the EULE subpopulations were analyzed, higher nucleotide diversity and similar mutation rate estimates were observed when compared with the U.S./CCAM subpopulations for the nuclear gene regions examined. Nucleotide

diversity for the EULE subpopulations was lower for the mitochondrial gene than CCAM estimates. Hudson's statistics separated the EULE subpopulation into three genetically subdivided populations including North Central Europe, Western Europe and Lebanon. The isolation with migration model clearly documented gene flow between the regions and the direction of migration was determined. When migration analysis was performed migration rates were high and direction of migrations was established from North Central Europe to France and from North Central Europe to Lebanon. Directional migration from France to Lebanon was also observed.

Historic first introductions of this pathogen into Europe occurred in the UK in 1958 from isolates on the center origin, Australia; and continental Europe in 1959 (Jonhson 1989). In continental Europe, migration occurred first to the Netherlands and Germany and later to the east to France and then Middle East and North Africa. Our model of migration supports the historic disease reports (La Mondia and Aylor 2001, Jonhson 1989, Ledez 1988). The oldest isolates studied in our work were from Germany and Lebanon in 1963 which is close to the time of the original introductions into Europe. Transport of inoculum by winds was believed to involve in movement of the pathogen in these first European epidemics.

CORESTA (2009) developed a blue mold forecasting system, but it has not been updated since 2004. In more recent years including 2004, epidemics started in North Africa, continued in the Middle East, North and Central Europe and finally into

Western Europe (France). Further analysis using more European samples collected more recently from other countries could provide better detail on population subdivision within current European populations of *P. tabacina* and the direction of migration.

**Mechanisms of genetic variation**

*P. tabacina* is coenocytic and, therefore, it is a multinucleate organism. *P. tabacina*'s sporangiospores have from 4-35 nuclei per sporangium (Trigiano *et al*. 1987, Trigiano *et al*. 1985). The high amount of spores produced in each yearly epidemic plus the multikaryon characteristic of *P. tabacina* increases the possibility of high amounts of variation in subsequent populations.

*P. tabacina* reproduces predominately by asexual means. It is unknown how nuclei migrate in the coenocytic hyphal stage of *P. tabacina* or the level of variation among nuclei. It is possible that divergence occurs in mycelium, to form lineages from each nucleus and each will not necessarily diverge in the same way. It is possible that in one asexual event many or all of the nuclei pass to the subsequent generation. It is also possible that developing spores receive one nucleus that then divides to form all the other nuclei. This phenomenon has been observed in mycorrhizal fungi (Schardl and Craven 2003, Sanders 2002, Kuhn *et al*. 2001, Clapp *et al*. 1999, Lloyd-McGilp *et al*. 1996) that also show high levels of genetic variation.

In some mycohrrizal species (*Glomus* spp.) it is believed that recombination can occur between different nuclei that coexist in one single spore (Bever *et al*. 2008, Pawlowska and Taylor 2004, Schardl and Craven 2003, Sanders 2002, Kuhn *et al*. 2001).

Intraspore genetic diversity has been reported in mychorrizal spores. Sanders (1995) first showed that mycorrhizal spores contain variable ribosomal DNA (rDNA) sequences, and many subsequent studies have confirmed this (Lloyd-McGilp *et al*. 1996, Clapp *et al*. 2001, Lanfranco *et al*. 1999, Antoniolli *et al*. 2000, Rodriguez *et al*. 2001). This means that the heterokaryotic organisms have single spores containing more than one genome that coexist together (Sanders 2002). For *P. tabacina*, the presence of genetic variability was observed in different clones from one single isolate and the result was supported for two nuclear regions and the mitochondrial gene. Further work is needed to examine genetic diversity of nuclei within spores and the frequency of nuclear genotypes inherited per spore. It would be necessary to analyze each nuclei separately to confirm if there is intraspore genetic diversity in *P. tabacina*.

It has been suggested that the genome of *P. tabacina* may be comprised of repetitive DNA (Sukno *et al*. 2002) as has been reported in other oomycetes including *Bremia lactuca* (Francis *et al*. 1990) and *Phytophthora infestans* (Judelson and Randall, 1998). The genome of *P. infestans* is greatly expanded with repetitive DNA. Rearrangements in the genome are associated with transposon activity which

results in high levels of genomic variation (Judelson 2007). Transposons have not been reported in *P. tabacina* but they occur in a related *Hyaloperonospora* sp. so their occurrence in *P. tabacina* is likely and needs to be examined (Beynon and Coates 2009).

Another possible explanation for the high variability in *P. tabacina* is the presence of sexual recombination. The presence of oospores in this pathosystem has been reported on *in vitro* plants after inoculation induced under laboratory conditions (Heist *et al*. 2002). In tobacco fields the presence of these structures was last reported in 1997 in the U.S. at very low levels (Reuveni *et al*. 1985, LaMondia and Aylor 2001). Sexual reproduction is known to occur on populations from overwintering cites in Florida, Texas and has been reported in Canada (LaMondia and Aylor 2001, Reuveni *et al*. 1988, McGrath and Miller 1958, Wolf 1936). Presence of oospores has been reported in Florida, Texas, Connecticut, North Carolina, Massachusetts and Pennsylvania (McKeen 1989, Reuveni *et al*. 1988, Spurr and Todd 1981, McGrath and Miller 1958, Wolf 1936). The presence of sexual recombination in the overwintering sites and dispersal of clonal lineages further north is a possibility since a large number of rare alleles were found in the populations examined. The presence of incompatible sites would support conflicting phylogenies and provide evidence of recombination and genetic exchange. However, further work will be necessary to confirm this hypothesis. Sexual

recombination has been reported in many other Oomycetes including *Bremia lactuca* on wild Lettuce species (Lebeda and Blok 1990), *Phytophthora infestans*, *P. capsici*, *P. sojae* and *Plasmopara viticola* (Judelson 2007).

Another explanation for the genetic variation observed could involve non-meiotic recombination of genetic material between cohabitating nuclei within heterokaryons (i.e. parasexuality) (Leslie 1993). Chromosomal exchange is due to mitotic crossing over and takes place during mitotic divisions. The resultant nucleus, because of crossing over, and random loss of choromosomes, is genetically different form the parents (Kamoun 2003).

The nuclear regions and mitochondrial gene regions identified and sequenced in this study have provided a valuable tool to study the genetic structure of populations of *P. tabacina*. Further studies comparing U.S. populations to those in other nearby tobacco growing regions including Cuba would be interesting to determine if the pathogen is introduced from Cuba each season as has been suggested in the forecasting trajectories. The data presented here have validated previous historic hypotheses about the long distance migration of this pathogen into the US and Europe and could be a valuable tool for similar studies with other downy mildews. This analysis has provided a basis for further studies to better understand the genetic processes that allow this pathogen to survive and colonize new areas. Ultimately, understanding the population structure of *P. tabacina* will lead to

improved strategies for the management of the disease.

## REFERENCES CITED


1. Antoniolli, Z. I., Schachtman, D. P., Ophel-Keller, K., and Smith, S. E. 2000. Variation in rDNA ITS sequences in *Glomus mosseae* and Gigaspora margarita spores from a permanent pasture. Mycological Research 104: 708-715.

2. Aylor, D., and Taylor, G. 1983. Escape of *Peronospora tabacina* spores from a field of diseased tobacco plants. Phytopathology 73(4): 525-529.

3. Aylor, D. L., Price, E. W. and Carbone, I. 2006. SNAP: Combine and Map modules for multilocus population genetic analysis. Bioinformatics 22: 1399-1401.

4. Bahlo, M., and R. C. Griffiths, 2000 Inference from gene trees in a subdivided population. Theor. Pop. Biol. 57: 79-95.

5. Bever, J.D., Kang, H.J., Kaonongbua, W., and Wang, M. 2008. Genomic Organization and mechanisms of inheritance in arbuscular mycorrhizal fungi: contrasting the evidence and implications of current theories. Pages: 135-148. *In*: Mycorrhiza. Varma A., and Hock B., ed. Springer Berlin Heidelberg, NY, U.S.A.

6. Beynon, J., and Coates, M. 2009. *Hyaloperonospora parasitica* as a pathogen model. Annual Review of Phytopathology (*in press*).

7. Bowden, L. C., Price, E.W. and Carbone, I. 2008. SNAP Clade and Matrix, Version 2. Distributed over the Internet: **http://www.cals.ncsu.edu/plantpath/faculty/carbone/home.html**

8. Brown, J.K.M., and Hovmoller, M. 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science 297: 537-541.

9. Bruck, R. I., Gooding, G. V., and Main C. E. 1982. Evidence of resistance to metalaxyl in isolates of *Peronospora hyoscyami*. Plant Dis. 66 (1): 44-45.

10. Caiazzo, R., Tarantino, P., Porrone, F., and Lahoz, E. 2006. Detection and early diagnosis of *Peronospora tabacina* Adam in tobacco plant with systemic infection. J. Phytopathology 154: 432-435.

11. Carbone, I., Liu, Y., Hillman, B. I. and Milgroom, M. G. 2004. Recombination and migration of *Cryphonectria hypovirus 1* as inferred from gene genealogies and the coalescent. Genetics 166: 1611-1629.

12. Carbone, I., Jakobek, J. L., Ramirez-Prado, J. H. and Horn, B. W. 2007. Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus.* Molecular Ecology 16: 4401-4417.

13. Carbone, I. and Kohn, L. M. 2001b. Multilocus nested haplotype networks extended with DNA fingerprints show common origin and fine-scale, ongoing genetic divergence in a wild microbial metapopulation. Molecular Ecology 10: 2409-2422.

14. Chamnanpunt, J., Shan, W. X., and Tyler, B. M. 2001. High frequency mitotic gene conversion in genetic hybrids of the oomycete Phytophthora sojae. Proc. Natl. Acad. Sci. 98: 14530-14535.

15. Clapp, J. P., Rodriguez, A., and Dodd, J.C. 2001. Inter and intra-isolate rRNA large subunit variation in *Glomus coronatum* spores. New Phytologist 149: 539-554.

16. Cook, K. L., Hudspeth, D. S. S., and Hudspeth, M. E. S. 2001. A *cox*2 phylogeny of terrestrial and marine parasitic Peronosporomycetes (Oomycetes). Nova Hedwigia. 122:231-243.

17. Cooke, D. E. L., Drenth, A., Duncan, J. M., Wagels, G., and Brasier, C. M. 2000. A molecular phylogeny of *Phytophthora* and related oomycetes. Fungal Genetics and Biology 30: 17-32.

18. CORESTA. 2009. Euro-Mediterranean Blue Mould Information Service. Web link: http://www.coresta.org/.

19. Davis, J. M., Main C. E., and Nesmith, W. C. 1990. Aerobiological aspects of the Kentucky blue mold epidemic of 1958, Pages 55-71. *In*: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

20. Davis, J. M., and Main, C. E. 1986. Applying atmospheric trajectory analysis to problems in epidemiology. Plant Dis. 70:490-497.

21. Davis, J. M., and Monahan and J. F. 1991. Climatology of air parcel trajectories related to the atmospheric transport of *Peronospora tabacina*. Plant Dis. 75:706-711.
22. Delon, R., and Schiltz, P. 1989. Spread and control of blue mold in Europe, North Africa and the Middle East, Pages 19-42. *In*: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

23. Edreva, A., Delon, R., and Coussirat, J. C. 1998. Variability of *Peronospora tabacina* A. – and isoenzyme study. Beiträge zur Tabakforschung International. 18(1): 3-13.

24. Francis, D. M., Hulbert, S. H., and Michelmore, R. W. 1990. Genome size and complexity of the obligate fungal pathogen, Bremia lactucae. Exp. Mycol. 14: 229-309.

25. Felsenstein, J. 1989. PHYLIP-Phylogeny Inference Package (version 3.2). Cladistics. 5: 164-166.

26. Fu, Y. X. and Li, W.H. 1993. Statistical test of neutrality of mutations. Genetics 133: 693-709.

27. Fu, Y. X. 1997. Statistical test of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-925.

28. Gómez-Alpizar, L, I. Carbone, and J. B. Ristaino. 2007. An Andean origin of *Phytophthora infestans* inferred from nuclear and mitochondrial gene genealogies. Proc Nat. Acad. Sci. 104: 3306-3311.

29. Griffiths, R. C., and Tavaré, S. 1994. Ancestral inference in population genetics. Stat. Sci. 9: 307-319.

30. Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

31. Heist, E. P., Nesmith, W. C., and Schardl, C. L. 2001. Cocultures of *Peronospora tabacina* and *Nicotiana* species to study host–pathogen interactions. Phytopathology 91:1224-1230.

32. Heist, E. P., Nesmith, W. C., and Schardl, C. L. 2002. Interactions of *Peronospora tabacina* with roots of *Nicotiana* spp. in gnotobiotic association. Phytopathology 92: 400-405.

33. Hey, J., and Nielsen, R. 2004. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. PNAS 104(8): 2785-2790.

34. Hudson, R. R., Boos, D. D., and Kaplan, N. L. 1992a. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. 9: 138-151.

35. Hudson, R. R., Slatkin, M., and Maddison, W. P. 1992b. Estimation of levels of gene flow from DNA sequence data. Genetics 132: 583-589.

36. Hudson, R. 2000. A new statistic for detecting genetic differentiation. Genetics 155: 2011-2014.

37. Hudspeth, D. S. S., Stenger, D., and Hudspeth, M. E. S. 2003. A *cox*2 phylogenetic hypothesis for the downy mildews and white rusts. Fungal Diversity 13: 47-57.

38. Jakobsen, I. B., and Easteal, S. 1996. A program for calculating and displaying compatibility matrices as an aid in determining reticulate evolution in molecular sequences. Bioinformatics 12(4): 291-295.

39. Jakobsen, I. B., Wilson, S. R., and Eastel S. 1997. The partition matrix: exploring variable phylogenetic signals along nucleotide sequence alignments. Mol. Biol. Evol. 14(5): 474-484.

40. Johnson, G. I. 1989. *Peronospora hyoscyami* de Bary: taxonomic history, strains, and host range. *In*: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN. pp. 55-71.

41. Judelson, H. S and Randall, T. A. 1998. Families of repeated DNA in the Oomycete *Phytophthora infestans* and their distribution within the genus. Genome 41: 605–615.

42. Judelson, H. S. 2007. Sexual Reproduction in Plant Pathogenic Oomycetes: biology and impact on disease. *In*: Sex in Fungi: Molecular Determination and Evolutionary Implications. J. Heitman *et al*., ed. ASM Press, Washington, D.C. pp. 445-458.

43. Kamoun, S. 2003. Molecular Genetics of Pathogenic Oomycetes. Eukaryotic Cell. 2: 191-199.

44. Kimura, M. 1980. A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111-120

45. Kuhn, G., Hijri, M., and Sanders, I. R. 2001. Evidence for the evolution of multiple genomes in arbuscular mycorrhizal fungi. Nature 414: 745-748.

46. Kucharek, T. A., Young, T. R., and Thomas, W. D. 1996. Occurrences of and some factors affecting blue mold of tobacco in Florida from 1921 to 1995. Proc. Soil Crop Sci. Soc. Fla. 55:81-85.

47. LaMondia, J. A., and Aylor D. E. 2001. Epidemiology and management of a periodically introduced pathogen. Biological Invasions. 3: 273-282.

48. Lanfranco, L., Delpero, M., Bonfante, P. 1999. Intrasporal variability of ribosomal sequences in the endomycorrhizal fungus *Gigaspora margarita*. Molecular Ecology 8: 37-45.

49. Lea H. W. 1999. Resistance of tobacco to pandemic blue mold (*Peronospora hyoscyami* de Bary syn *P. tabacina* Adam): a historical overview. Australian Journal of Experimental Agriculture 39: 115-118.

50. Lebeda, A., and Blok, I. 1990. Sexual compatibility types of *Bremia lactucae* isolates originating from *Lactuca serriola*. Neth. J. Pl. Path. 96: 51-54.

51. Ledez, P. 1988. The CORESTA tobacco blue mold warning service for the Euro-Mediterranean zone. In: Blue Mold disease of tobacco, Symposium proceedings, ed. C. E. Main and H. W. Spurr, Raleigh, North Carolina, February 1988. pp. 79-91.

52. Leslie, J.F. 1993. Fungal vegetative compatibility. Annu. Rev. Phytopathol. 31: 127-150.

53. Li, H and Fu, Y. 2002. Neutrality test program. Release 1.1. Web link: http://hgc.sph.uth.tmc.edu/neutrality_test

54. Lloyd-McGilp, S. A., Chambers, S. M., Dodd, J. C., Fitter, A. H., Walker, C., and Young, J. P. W. 1996. Diversity of the ribosomal internal transcribed spacers within and among isolates of *Glomus mosseae* and related fungi. New Phytologist 133: 103-111.

55. Lucas, G. B. 1980. The war against blue mold. Science. 210: 147-153.

56. Main, C. E. 1995. Tobacco blue mold epidemic of 1995 in the United States. Coresta Meeting Report, Oxford.

57. Main, C. E., Keever, T., Holmes, G., and Davis, J. 2001. Forecasting long-range transport of downy mildew spores and plant disease epidemics. APS Net Feature, April-May 2001. Online publication.

58. Main C. E., and Spurr, H. W. 1990. Blue mold – the problem and the phenomenon. In: Blue Mold disease of tobacco, Symposium proceedings, ed. C. E. Main and H. W. Spurr, Raleigh, North Carolina, February 1988. pp. ix-x.

59. McDonald, B. A., and Linde, C. 2002. Pathogen Population Genetics, evolutionary potential, and durable resistance. Annu. Rev. Phytopathol. 40: 349-379

60. McGrath, H., and Miller, P. R. 1958. Blue mold of tobacco. Plant Dis. Rep. Suppl. 250: 1-35.

61. McKeen, W. E. 1989. *Peronospora hyoscyami* de Bary: taxonomic history, strains, and host range. *In*: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN. pp. 55-71.

62. Micales, J. A., Bonde, M. R., and Peterson, G. L. 1986. The use of isoenzyme analysis in fungal taxonomy and genetics. Mycotaxon. 27: 405-449.

63. Moss, M. A., and Main, C. E. 1988. The effect of temperature on sporulation and viability of isolates of *Peronospora tabacina* collected in the United States. Phytopathology 78:110-114.

64. Nei, M., and Li, W. H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. 76(10): 5269-5273.

65. Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

66. Nesmith, W. C. 1984. The North American Blue Mold Warning System. Plant Dis. 68(11): 933-936.

67. Nesmith, W. C. 1990. The North American Blue Mold Warning System – an update. In: Blue Mold disease of tobacco, Symposium proceedings, ed. C. E. Main and H. W. Spurr, Raleigh, North Carolina, February 1988. pp. 73-78.

68. Nielsen, R., and Wakeley, J. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics 158: 885-896.

69. Pawlowska, T., and Taylor, J. 2004. Organization of genetic variation in individual of arbuscular mycorrhizal fungi. Nature. 427: 733-737.

70. Person, L. H., Lucas, G. B., Koch W. J. 1955. A chytrid attacking spores of *Peronospora tabacina*. Plant Disease Reporter. 39 (12): 887-888.

71. Petersen, A. B., and Rosendal, S. 2000. Phylogeny of the Pernosporomycetes (Oomycota) based on partial sequences of the large ribosomal subunit (LSU rDNA). Mycological Research. 102: 1295-1303.

72. Price, E.W., and I. Carbone. 2005. SNAP: workbench management tool for evolutionary population genetic analysis. Bioinformatics 21:402-404.

73. Reuveni, M., Nesmith, W. C., Siegel, M. R., and Keeny, T. M. 1988. Virulence of an endemic isolate of *Peronospora tabacina* from Texas to *Nicotiana tabacum* and *N. repanda*. Plant Dis. 72: 1024-1027.

74. Ristaino J.B., Johnson A., Blanco-Meneses M., and Liu B. 2007. Identification of the tobacco blue mold pathogen, *Peronospora tabacina*, by polymerase chain reaction. Plant Dis. 91: 685-691.

75. Rodriguez, A., Dougall, T., Dodd, J. C., and Clapp, J. P. 2001. The large subunit ribosomal RNA genes of *Entrophospora infrequens* comprise sequences related to two different glomalean families. New Phytologist 152: 159-167.

76. Rozas, J., Sanchez-delBarrio, J. C., Messeguer, X., and Rozas, R. 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

77. Rufty, R.C. 1989. Genetics of host resistance to tobacco blue mold. p. 141–164. *In*: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

78. Sanders, I. R. 2002. Ecology and evolution of multigenomic arbuscular mycorrhizal fungi. The Am. Naturalist 160: S128-S141.

79. Sanders, I. R., Alt, M., Groppe, K., Boller, T., and Wiemken, A. 1995. Identification of ribosomal DNA polymorphism among and within spores of the Glomales: application to studies on the genetic diversity of arbuscular mycorrhizal fungal communities. New Phytologist. 130: 419-427.

80. Schardl, C. L., and Craven, K. D. 2003. Interspecific hybridization in plant-associated fungi and oomycetes: a review. Molecular Ecology 12: 2861-2873.

81. Schena, L., and Cooke, D. E. L. 2006. Assessing the potential of regions of the nuclear and mitochondrial genome to develop a “molecular tool box” for the detection and characterization of *Phytophthora* species. J. Microbiol. Methods. 67: 70-85.

82. Shoemaker, P.B., Milks, D.C., and Main C. E. 2000. Sensitivity of *Peronospora tabacina* isolates to metalaxyl and dimethomorph in North Carolina. In: 2000 CORESTA Congress, Lisbon, Portugal.

83. Spurr, H. W. and Menetrez, M. L. 1990. Pathogen specificity and variation. In: Blue Mold disease of tobacco, Symposium proceedings, ed. C. E. Main and H. W. Spurr, Raleigh, North Carolina, February 1988. pp. 1-9.

84. Spurr, H. W., and Todd, F. A. 1982. Oospores in blue mold diseased North Carolina Burley and Flue-cured tobacco. Tobacco Science. 26: 44-46.

85. Sukno, S. A., Taylor, A. M., and Farman M. L. 2002. Genetic uniformity among isolates of *Peronospora tabacina*, the tobacco blue mold pathogen. Phytopathology. 92: 1236-1244.

86. Svircev, A.M., McKeen, W. E., and Smith, R.J. 1989. Host-parasite relations: morphology and ultrastructure, Pages 43-104. *In*: Blue Mold of Tobacco. W. E. McKeen, ed. American Phytopathological Society Press, St. Paul, MN.

87. Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105**:** 437-460.

88. Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

89. Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc. Acids Res. 22: 4673-4680.

90. Todd, F.A. 1981. The blue mold story. p. 9-25. *In*: Report: 29th Tobacco Workers Conference. F. A. Todd, compiler. Lexington, KY. 109 pp.

91. Tooley, P.W., and Therrien, C.D. 1991. Pages: 204–217 in *Phytophthora*, eds Lucas, J.A., Shattock, R.C., Shaw, D.S. and Cooke, L.R. Cambridge Univ Press, Cambridge, UK.

92. Trigiano, R. N., Van Dyke, C. G., Spurr, H.W., Jr., and Main, C. E. 1985. Ultrastructure of sporangiophore and sporangium ontogeny of *Peronospora tabacina*. Tob. Sci. 29:116-121.

93. Trigiano, R. N. and Spurr, H.W. 1987. The development of the multinucleate condition of *Peronospora tabacina* sporangia. Mycologia 79(3): 353-357.

94. Trout, C. L., Ristaino, J. B., Madritch, M., and Wangsomboondee, T. 1997. Rapid detection of *Phytophthora infestans* in late blight infected tissue of potato and tomato using PCR. Plant Dis. 81:1042-1048.

95. Tyron, H. 1890. Tobacco disease. The Queenslander, Queensland Newspapers, Brisbane. July 5, p. 29.

96. Voglmayr, H. 2003. Phylogenetic relationships of *Peronospora* and related genera based on nuclear ribosomal ITS sequences. Mycol. Res. 107: 1132-1142.

97. Wangsomboondee T., and Ristaino, J. B. 2002. Optimization of samples size and DNA extraction methods to improve PCR detection of different propagules of *Phytophthora infestans*. Plant Dis. 86: 247-253.

98. Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7: 256-276.

99. Wiglesworth, M.D., Reuveni, M., Nesmith, W. C., Siegel, M. R., Kuc, J., Juarez, J. 1988. Resistance of *Peronospora tabacina* to metalaxyl in Texas and Mexico. Plant Dis. 72: 964-967.

100. Wolf, F. A. 1947. Tobacco downy mildew, endemic to Texas and Mexico. Phytopathology. 37: 721-729.

101. Wolf, F. A., Dixon, L. F., McLean, R., and Darkis, F. R. 1933. Downy mildew of tobacco. Plant Dis. Rep. 17:37-39.

102. Wolf, F. A., McLean, R., and Dixon, L. F. 1936. Further studies on downy mildew of tobacco. Phytopathology. 26: 760-777.

**Table 2. 1** *P. tabacina* isolates collected between 1945 and 2006 from the U.S., Caribbean, Central America and Mexico and Europe and Lebanon used for cloning and sequencing in this study.

| Isolate | Year | Source | Collector and location |
|---|---|---|---|
| **United States of America (US)** | | | |
| CT991 [b] | 1999 | Connecticut, Windsor Co. | J. LaMondia, University of Massachussets |
| CT021 | 2002 | Connecticut, Hartford Co. | J. LaMondia, University of Massachussets |
| FL002 | 2000 | Florida | T. Kucharek, University of Florida |
| GA992 [b] | 1999 | Georgia, Colquitt Co. | P. Bertrand, University of Georgia |
| GA996 [b] | 1999 | Georgia, Tifton Co. | A. Csinos, University of Georgia |
| KY79 [b] | 1979 | Kentucky | B. Nesmith, University of Kentucky |
| KY206 [b] | 2006 | Kentucky, Simpson Co. | K. Seebold, University of Kentucky |
| KY062 [b] | 2006 | Kentucky, Greenup Co. | K. Seebold, University of Kentucky |
| BPT | 1984 | Maryland, Beltsville Co. | Clayton, Texas A&M University |
| Hadley [b] | 2004 | Maryland, Hampshire Co. | P. Shoemaker, North Carolina State Univesity |
| BPTS | 1945 | N. Carolina, Bertie Co. | no information |
| RPT83 | 1983 | N. Carolina, Robeson Co. | M. Moss, North Carolina State University |
| Bertie [b] | 1983 | N. Carolina, Bertie Co. | R. Rufty, North Carolina State University |
| PT87W [b] | 1987 | N. Carolina, Bumcombe Co. | no information |
| NC998 [b] | 1999 | N. Carolina, Yancey Co. | P. Shoemaker, North Carolina State University |
| NC9911 [b] | 1999 | N. Carolina, Madison Co. | P. Shoemaker, North Carolina State University |
| NC997 [b] | 1999 | N. Carolina, Jones Co. | T. Melton, North Carolina State University |
| NC994 [b] | 1999 | N. Carolina, Robeson Co. | G. Pate, North Carolina State University |
| NC0212 | 2002 | N. Carolina, Granville Co. | P. Shoemaker, North Carolina State University |
| AV041 [b] | 2004 | N. Carolina, Avery Co. | F. Bolick, North Carolina State University |
| BU041 [b] | 2004 | N. Carolina, Bumcombe Co. | S. Holloway, North Carolina State University |
| BU061 [b] | 2006 | N. Carolina, Bumcombe Co. | K. Ivors, North Carolina State University |
| Penn [b] | 1998 | Pennsylvania, Lancaster Co. | J. Yocum, Pennsylvania State University |
| Colum [b] | 1998 | Pennsylvania, Columbia Co. | J. Yocum, Pennsylvania State University |
| PE05 [b] | 2005 | Pennsylvania, Lancaster Co. | B. Maksymowicz, University of Kentucky |
| BPTP | 1945 | Texas | Clayton, Texas A&M University |
| TX84 [b] | 1984 | Texas | B. Nesmith, University of Kentucky |
| TX062 [b] | 2006 | Texas, Uvalde Co. | M. Black, Texas A&M University |
| Mosel [b] | 1998 | Virginia, Moseley Co. | J. Ristaino, North Carolina State University |
| VA011 | 2001 | Virginia, Dinwiddie Co. | Avery, Virginia Tech |
| **Caribbean, Central America and Mexico (CCAM)** | | | |
| DR022 | 2002 | Dominican Rep., Santiago | I. Abreu, FERQUIDO |
| DR051 [b] | 2005 | Dominican Rep., Valverde | A. Ubiera, FERQUIDO |
| DR054 [b] | 2005 | Dominican Rep., Valverde | A. Ubiera, FERQUIDO |
| DR062 [b] | 2006 | Dominican Rep., Valverde | M. Peralta, Tabadom Holding Inc. |
| DR064 | 2006 | Dominican Rep., Valverde | M. Peralta, Tabadom Holding Inc. |
| DR066 [b] | 2006 | Dominican Rep., Valverde | M. Peralta, Tabadom Holding Inc. |
| 123392 | 1989 | Guatemala | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123347 [b] | 1989 | Mexico, Veracruz | P. Shoemaker, North Carolina State Univesity |
| 123296 | 1989 | Mexico, S. Andres | P. Shoemaker, North Carolina State Univesity |
| Mex [b] | 1999 | Mexico, Veracruz | R. Rufty, North Carolina State Univesity |
| MX02 | 2002 | Mexico, Nayarit | V. Nikolaeva, Amapa, Nayarit |
| NIC063 [b] | 2005 | Nicaragua, Jalapa | N. Placencia, Segovia Cigars |
| NIC065 [b] | 2005 | Nicaragua, Jalapa | N. Placencia, Segovia Cigars |

Table 2.1 continued

| Isolate | Year | Source | Collector and location |
|---|---|---|---|
| **Europe (EU)** | | | |
| 123046 | 1988 | Bulgaria | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123364 | 1987 | France | R. Delon, Institut du Tabac-Altadis |
| FR184 [b] | 2005 | France | J. L. Verrier, Institut du Tabac-Altadis |
| FR194 [b] | 2005 | France | J. L. Verrier, Institut du Tabac-Altadis |
| FR178 [b] | 2005 | France | J. L. Verrier, Institut du Tabac-Altadis |
| FRy1 [b] | 1999 | France | J. L. Verrier, Institut du Tabac-Altadis |
| 123350 | 1963 | Germany | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123368 | 1980 | Hungary | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123340 | 1963 | Lebanon | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123349 [b] | 1989 | Poland | G. L. Peterson, USDA, Ft. Detrick, MD [a] |
| 123344 | 1989 | Poland | G. L. Peterson, USDA, Ft. Detrick, MD [a] |

[a] Samples belong to a collection located at the USDA, Ft. Detrick, MD and donated to the Plant Pathology Department at NCSU in 2004.

[b] DNA extracted from an isolate purified usign an *in vitro* tobacco plant.

**Table 2. 2** Population statistics, diversity estimates and neutrality tests based on variation in nuclear and mitochondrial regions for *P. tabacina.*

| | Sample summaries | | | | | Parameter estimates | | Test of Neutrality | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus/Population | *l* | *n* | *s* | *h* | *k* | $\pi$ (SE) $X10^{-3}$ | $\Theta w$ $X10^{-3}$ | Tajima's D Statistic | Fu and Li's D* Statistic | Fu and Li's F* Statistic | Fu's Fs Statistic |
| ***Nuclear*** | | | | | | | | | | | |
| **(*Igs* 2 region)** | | | | | | | | | | | |
| USA | 435 | 180 | 68 | 51 | 1.393 | 20.48 ($1.64X10^{-3}$) | 173.39 | -2.628 *** | -10.307 *** | -7.884 *** | -20.056 *** |
| CCAM | 435 | 78 | 54 | 36 | 2.290 | 42.41 ($3.97X10^{-3}$) | 202.94 | -2.503 *** | -8.323 *** | -6.749 *** | -21.279 *** |
| EULE | 435 | 66 | 27 | 25 | 1.432 | 53.04 ($5.82X10^{-3}$) | 210.12 | -2.273 *** | -5.550 *** | -4.872 *** | -21.239 *** |
| TOTAL | 435 | 324 | 116 | 93 | 0.941 | 8.11 ($7.00X10^{-4}$) | 157.32 | -2.788 *** | -12.642 *** | -8.925 *** | ns |
| NC-BURLEY | 435 | 36 | 25 | 17 | 1.951 | 78.03 ($1.27X10^{-2}$) | 241.15 | -2.214 ** | -4.869 *** | -4.350 *** | -12.529 *** |
| NC-FLUE-CURED | 432 | 36 | 16 | 11 | 1.470 | 91.87 ($1.73X10^{-2}$) | 241.15 | -1.922 ** | -4.340 *** | -3.881 *** | -5.392 * |
| **(*Ypt* 1 region)** | | | | | | | | | | | |
| USA | 495 | 174 | 89 | 58 | 1.395 | 15.68 ($1.74X10^{-3}$) | 174.42 | -2.760 *** | -9.191 *** | -7.165 *** | -20.234 *** |
| CCAM | 493 | 66 | 34 | 28 | 2.182 | 64.17 ($5.97X10^{-3}$) | 210.12 | -2.160 * | -5.385 *** | -4.681 *** | -20.192 *** |
| EULE | 494 | 66 | 27 | 24 | 1.919 | 71.09 ($7.21X10^{-3}$) | 210.12 | -2.012 * | -5.550 *** | -4.757 *** | -20.446 *** |
| TOTAL | 493 | 306 | 126 | 93 | 1.295 | 10.28 ($8.10X10^{-4}$) | 158.75 | -2.769 *** | -11.709 *** | -8.350 *** | -20.051 *** |
| NC-BURLEY | 490 | 36 | 27 | 18 | 2.805 | 103.88 ($1.37X10^{-2}$) | 241.15 | -1.877 ** | -4.156 ** | -3.704 ** | -10.397 ** |
| NC-FLUE-CURED | 489 | 36 | 27 | 18 | 2.481 | 91.89 ($1.52X10^{-2}$) | 241.15 | -2.041 ** | -5.017 *** | -4.381 *** | -10.028 ** |
| ***Mitochondrial*** | | | | | | | | | | | |
| **(*cox* 2 gene)** | | | | | | | | | | | |
| USA | 626 | 180 | 65 | 46 | 0.744 | 1.25 ($1.80X10^{-4}$) | 18.88 | -2.776 *** | -11.691 *** | -8.854 *** | -21.098 *** |
| CCAM | 624 | 78 | 40 | 29 | 1.473 | 31.63 ($6.82X10^{-3}$) | 202.94 | -2.534 *** | -4.453 *** | -4.205 *** | -21.368 *** |
| EULE | 622 | 66 | 23 | 21 | 0.697 | 1.14 ($2.20X10^{-4}$) | 7.92 | -2.554 *** | -6.943 *** | -5.993 *** | -20.799 *** |
| TOTAL | 631 | 324 | 107 | 82 | 0.820 | 7.66 ($1.02X10^{-3}$) | 157.32 | -2.787 *** | -12.295 *** | -8.775 *** | ns |
| NC-BURLEY | 604 | 36 | 9 | 8 | 0.552 | 0.91 ($2.9X10^{-4}$) | 3.59 | -2.099 ** | -3.726 ** | -3.520 *** | -6.436 ** |
| NC-FLUE-CURED | 609 | 36 | 12 | 11 | 0.667 | 1.09 ($3.0X10^{-4}$) | 4.75 | -2.286 *** | -4.786 *** | -4.358 *** | -10.943 *** |

l, # of nucleotide sites; n, sample size (number of clones); s, segregating nucleotide sites; h, haplotypes; K, average number of pairwise nucleotide differences;
$\pi$, average number of base differences per site; SE, standart error; $\Theta w$, population mean mutation rate per site or Watterson's theta estimator
ns, not significant; *, 0.01<P<0.05; **, 0.001<P<0.01; ***, P<0.001

**Table 2. 3** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the U.S. and CCAM subpopulatons of *P. tabacina* pathogen. The total population was used for the test.

Snn U.S.-CCAM - *lgs* 2

| | Texas | Florida | Georgia | N. Carolina | Kentucky | Virginia | Maryland | Pennsylvania | Connecticut | Mexico | Guatemala | Nicaragua |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Texas** | | | | | | | | | | | | |
| **Florida** | ns | | | | | | | | | | | |
| **Georgia** | 0.015 * | ns | | | | | | | | | | |
| **North Carolina** | ns | ns | ns | | | | | | | | | |
| **Kentucky** | 0.021 * | ns | ns | ns | | | | | | | | |
| **Virginia** | 0.000 *** | ns | 0.001 ** | 0.000 *** | 0.001 ** | | | | | | | |
| **Maryland** | ns | ns | ns | ns | ns | 0.000 *** | | | | | | |
| **Pennsylvania** | 0.001 ** | ns | ns | ns | ns | 0.000 *** | ns | | | | | |
| **Connecticut** | ns | ns | ns | ns | ns | 0.000 *** | ns | ns | | | | |
| **Mexico** | ns | ns | ns | ns | ns | 0.000 *** | ns | ns | ns | | | |
| **Guatemala** | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | | |
| **Nicaragua** | ns | ns | ns | ns | ns | 0.000 *** | ns | ns | ns | ns | ns | |
| **Domicam Rep.** | ns | ns | ns | ns | ns | 0.000 *** | ns | ns | ns | ns | ns | ns |

Snn U.S.-CCAM - *Ypt* 1

| | Texas | Florida | Georgia | N. Carolina | Kentucky | Virginia | Maryland | Pennsylvania | Connecticut | Mexico | Guatemala | Nicaragua |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Texas** | | | | | | | | | | | | |
| **Florida** | ns | | | | | | | | | | | |
| **Georgia** | ns | ns | | | | | | | | | | |
| **North Carolina** | 0.013 * | ns | ns | | | | | | | | | |
| **Kentucky** | 0.004 ** | ns | 0.014 * | 0.000 *** | | | | | | | | |
| **Virginia** | ns | ns | ns | ns | 0.015 * | | | | | | | |
| **Maryland** | ns | ns | 0.007 ** | 0.026 * | 0.042 * | ns | | | | | | |
| **Pennsylvania** | 0.004 ** | ns | 0.001 ** | 0.000 *** | 0.002 ** | 0.004 ** | 0.021 * | | | | | |
| **Connecticut** | ns | ns | 0.037 * | ns | 0.038 * | ns | ns | 0.007 ** | | | | |
| **Mexico** | ns | ns | ns | ns | 0.001 ** | ns | 0.006 ** | 0.000 *** | 0.028 * | | | |
| **Guatemala** | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | | |
| **Nicaragua** | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | |
| **Domicam Rep.** | ns | ns | ns | 0.000 *** | ns | ns | ns | ns | ns | ns | ns | ns |

Snn U.S.-CCAM - *cox* 2

| | Texas | Florida | Georgia | N. Carolina | Kentucky | Virginia | Maryland | Pennsylvania | Connecticut | Mexico | Guatemala | Nicaragua |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Texas** | | | | | | | | | | | | |
| **Florida** | ns | | | | | | | | | | | |
| **Georgia** | ns | ns | | | | | | | | | | |
| **North Carolina** | ns | ns | ns | | | | | | | | | |
| **Kentucky** | ns | ns | ns | ns | | | | | | | | |
| **Virginia** | 0.027 * | ns | 0.000 *** | 0.002 ** | 0.000 *** | | | | | | | |
| **Maryland** | ns | ns | ns | ns | ns | ns | | | | | | |
| **Pennsylvania** | ns | ns | ns | 0.000 *** | 0.000 *** | ns | ns | | | | | |
| **Connecticut** | ns | ns | ns | ns | ns | 0.020 * | ns | ns | | | | |
| **Mexico** | 0.001 ** | ns | ns | ns | ns | 0.000 *** | 0.008 ** | 0.000 *** | ns | | | |
| **Guatemala** | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns | | |
| **Nicaragua** | ns | ns | ns | ns | ns | 0.002 ** | ns | ns | ns | ns | ns | |
| **Domicam Rep.** | ns | ns | ns | 0.013 * | ns | 0.032 * | ns | 0.010 * | ns | ns | ns | ns |

**Table 2. 4** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the U.S. and CCAM subpopulatons of *P. tabacina* pathogen. Some of the locations were grouped together for the first time after test analysis. CARMX= Caribbean (Dominican Republic) and Mexico, CAM= Central America (Nicaragua/Guatemala).

**Snn U.S.-CCAM - *Igs* 2- pooled#1**

| | Texas | Florida | Georgia | N. Carolina | Kentucky | Virginia | Maryland | Pennsylvania | Connecticut | CARMX |
|---|---|---|---|---|---|---|---|---|---|---|
| **Texas** | | | | | | | | | | |
| **Florida** | ns | | | | | | | | | |
| **Georgia** | 0.006 ** | ns | | | | | | | | |
| **North Carolina** | ns | ns | 0.047 * | | | | | | | |
| **Kentucky** | ns | ns | 0.023 * | ns | | | | | | |
| **Virginia** | ns | ns | ns | ns | ns | | | | | |
| **Maryland** | ns | ns | ns | ns | ns | ns | | | | |
| **Pennsylvania** | 0.013 * | ns | ns | ns | ns | ns | ns | | | |
| **Connecticut** | ns | ns | 0.002 ** | ns | ns | ns | ns | ns | | |
| **CARMX** | ns | ns | 0.000 *** | ns | ns | 0.012 * | ns | 0.000 *** | ns | |
| **CAM** | ns | ns | 0.000 *** | ns | ns | ns | ns | 0.000 *** | ns | ns |

**Snn U.S.-CCAM - *Ypt* 1- pooled#1**

| | Texas | Florida | Georgia | N. Carolina | Kentucky | Virginia | Maryland | Pennsylvania | Connecticut | CARMX |
|---|---|---|---|---|---|---|---|---|---|---|
| **Texas** | | | | | | | | | | |
| **Florida** | ns | | | | | | | | | |
| **Georgia** | 0.000 *** | 0.059 * | | | | | | | | |
| **North Carolina** | ns | ns | 0.000 *** | | | | | | | |
| **Kentucky** | ns | ns | 0.000 *** | ns | | | | | | |
| **Virginia** | 0.000 *** | ns | 0.017 * | 0.000 *** | 0.000 *** | | | | | |
| **Maryland** | 0.000 *** | ns | 0.018 * | 0.001 ** | 0.000 *** | 0.035 * | | | | |
| **Pennsylvania** | 0.000 *** | ns | 0.001 ** | 0.000 *** | ns | 0.005 ** | ns | | | |
| **Connecticut** | 0.004 ** | ns | 0.000 *** | ns | 0.049 * | 0.002 ** | ns | ns | | |
| **CARMX** | ns | ns | 0.000 *** | 0.000 *** | ns | 0.000 *** | 0.005 ** | 0.005 ** | ns | |
| **CAM** | 0.057 * | ns | 0.000 *** | ns | ns | 0.000 *** | ns | ns | ns | ns |

**Snn U.S.-CCAM - *cox* 2- pooled#1**

| | Texas | Florida | Georgia | N. Carolina | Kentucky | Virginia | Maryland | Pennsylvania | Connecticut | CARMX |
|---|---|---|---|---|---|---|---|---|---|---|
| **Texas** | | | | | | | | | | |
| **Florida** | 0.000 *** | | | | | | | | | |
| **Georgia** | 0.000 *** | ns | | | | | | | | |
| **North Carolina** | ns | 0.001 ** | 0.000 *** | | | | | | | |
| **Kentucky** | ns | 0.000 *** | 0.000 *** | ns | | | | | | |
| **Virginia** | ns | 0.000 *** | 0.000 *** | ns | ns | | | | | |
| **Maryland** | 0.000 *** | ns | ns | 0.002 ** | 0.000 *** | 0.002 ** | | | | |
| **Pennsylvania** | 0.000 *** | 0.054 * | 0.032 * | 0.000 *** | 0.000 *** | ns | ns | | | |
| **Connecticut** | ns | 0.002 ** | ns | ns | ns | ns | 0.010 * | ns | | |
| **CARMX** | ns | 0.001 ** | 0.000 *** | ns | ns | ns | 0.001 ** | 0.000 *** | ns | |
| **CAM** | ns | 0.000 *** | 0.000 *** | ns | ns | ns | 0.000 *** | 0.000 *** | ns | ns |

**Table 2. 5** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the U.S. and CCAM subpopulatons of *P. tabacina* pathogen. Some of the locations were grouped together for the second time after test analysis. NCKY= North Carolina/ Kentucky, PAMDCT= Pennsylvania/Maryland /Connecticut, CCAM= Caribbean, Central America and Mexico.

**$S_{nn}$ U.S.-CCAM - *Igs* 2- pooled#2**

| | Texas | Florida | Georgia | NCKY | Virginia | PAMDCT |
|---|---|---|---|---|---|---|
| **Texas** | | | | | | |
| **Florida** | ns | | | | | |
| **Georgia** | ns | 0.027 * | | | | |
| **NCKY** | ns | ns | ns | | | |
| **Virginia** | ns | 0.025 * | ns | ns | | |
| **PAMDCT** | ns | ns | ns | ns | ns | |
| **CCAM** | ns | ns | ns | ns | ns | ns |

**$S_{nn}$ U.S.-CCAM - *Ypt* 1- pooled#2**

| | Texas | Florida | Georgia | NCKY | Virginia | PAMDCT |
|---|---|---|---|---|---|---|
| **Texas** | | | | | | |
| **Florida** | ns | | | | | |
| **Georgia** | ns | ns | | | | |
| **NCKY** | ns | ns | ns | | | |
| **Virginia** | ns | ns | ns | ns | | |
| **PAMDCT** | ns | ns | ns | ns | ns | |
| **CCAM** | ns | ns | ns | 0.016 * | ns | 0.007 ** |

**$S_{nn}$ U.S.-CCAM - *cox* 2- pooled#2**

| | Texas | Florida | Georgia | NCKY | Virginia | PAMDCT |
|---|---|---|---|---|---|---|
| **Texas** | | | | | | |
| **Florida** | ns | | | | | |
| **Georgia** | ns | ns | | | | |
| **NCKY** | ns | ns | ns | | | |
| **Virginia** | ns | ns | ns | ns | | |
| **PAMDCT** | ns | ns | ns | ns | ns | |
| **CCAM** | ns | ns | ns | 0.004 ** | ns | ns |

**Table 2. 6** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the U.S. and CCAM subpopulatons of *P. tabacina* pathogen. Some of the locations were grouped together for the third time after test analysis. South= Texas/Florida/ CCAM (Dominican Republic, Guatemala, Nicaragua/ Mexico), North= North Carolina/Kentucky/Georgia/Virginia/ Pennsylvania, Maryland/Connecticut.

**$S_{nn}$ EULE-*lgs* 2**

| | Hungary | Germany | Poland | Bulgaria | Lebanon |
|---|---|---|---|---|---|
| **Hungary** | | | | | |
| **Germany** | 0.000 *** | | | | |
| **Poland** | 0.000 *** | ns | | | |
| **Bulgaria** | ns | ns | 0.000 *** | | |
| **Lebanon** | ns | 0.045 * | 0.000 *** | ns | |
| **France** | 0.002 ** | 0.043 * | 0.016 * | 0.054 * | 0.024 * |

**$S_{nn}$ EULE-*Ypt* 1**

| | Hungary | Germany | Poland | Bulgaria | Lebanon |
|---|---|---|---|---|---|
| **Hungary** | | | | | |
| **Germany** | 0.003 ** | | | | |
| **Poland** | 0.001 ** | ns | | | |
| **Bulgaria** | 0.027 * | 0.017 * | 0.001 ** | | |
| **Lebanon** | 0.001 ** | 0.002 ** | 0.000 *** | ns | |
| **France** | 0.012 * | ns | 0.049 * | 0.001 ** | 0.000 *** |

**$S_{nn}$ EULE-*cox* 2**

| | Hungary | Germany | Poland | Bulgaria | Lebanon |
|---|---|---|---|---|---|
| **Hungary** | | | | | |
| **Germany** | ns | | | | |
| **Poland** | ns | ns | | | |
| **Bulgaria** | ns | ns | ns | | |
| **Lebanon** | ns | ns | 0.053 * | ns | |
| **France** | ns | ns | ns | ns | ns |

**Table 2. 7** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the EULE subpopulation of *P. tabacina* pathogen. The total population was used for the test.

**$S_{nn}$ U.S.-CCAM - *Igs* 2- pooled#3**

| | South |
|---|---|
| **South** | |
| **North** | ns |

**$S_{nn}$ U.S.-CCAM - Ypt1/5- pooled#3**

| | South |
|---|---|
| **South** | |
| **North** | 0.000 *** |

**$S_{nn}$ U.S.-CCAM- *cox* 2- pooled#3**

| | South |
|---|---|
| **South** | |
| **North** | 0.000 *** |

**Table 2. 8** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the EULE subpopulation of *P. tabacina* pathogen. Some of the locations were grouped together for the first time after test analysis. North Europe (NE) = Poland/ Germany, Central Europe (CE)=Bulgaria/Hungary, Western Europe=France, LE= Lebanon.

**$S_{nn}$ EULE- *Igs* 2- POOLED#1**

| | NE | CE | WE |
|---|---|---|---|
| **NE** | | | |
| **CE** | ns | | |
| **WE** | 0.007 ** | 0.003 ** | |
| **LE** | ns | ns | ns |

**$S_{nn}$ EULE- *Ypt* 1- POOLED#1**

| | NE | CE | WE |
|---|---|---|---|
| **NE** | | | |
| **CE** | ns | | |
| **WE** | 0.010 * | 0.004 ** | |
| **LE** | ns | ns | 0.000 *** |

**$S_{nn}$ EULE- *cox* 2- POOLED#1**

| | NE | CE | WE |
|---|---|---|---|
| **NE** | | | |
| **CE** | ns | | |
| **WE** | ns | ns | |
| **LE** | ns | ns | ns |

**Table 2. 9** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the EULE subpopulation of *P. tabacina* pathogen. Some of the locations were grouped together for the first time after test analysis. North-Central Europe (NCE) = Poland/ Germany/Bulgaria/Hungary, Western Europe=France, LE= Lebanon.

**$S_{nn}$ EULE- *lgs* 2- POOLED#2**

| | NCE | WE |
|---|---|---|
| **NCE** | | |
| **WE** | 0.001 ** | |
| **LE** | ns | ns |

**$S_{nn}$ EULE- *Ypt* 1- POOLED#2**

| | NCE | WE |
|---|---|---|
| **NCE** | | |
| **WE** | 0.000 *** | |
| **LE** | ns | ns |

**$S_{nn}$ EULE- *cox* 2- POOLED#2**

| | NCE | WE |
|---|---|---|
| **NCE** | | |
| **WE** | ns | |
| **LE** | ns | ns |

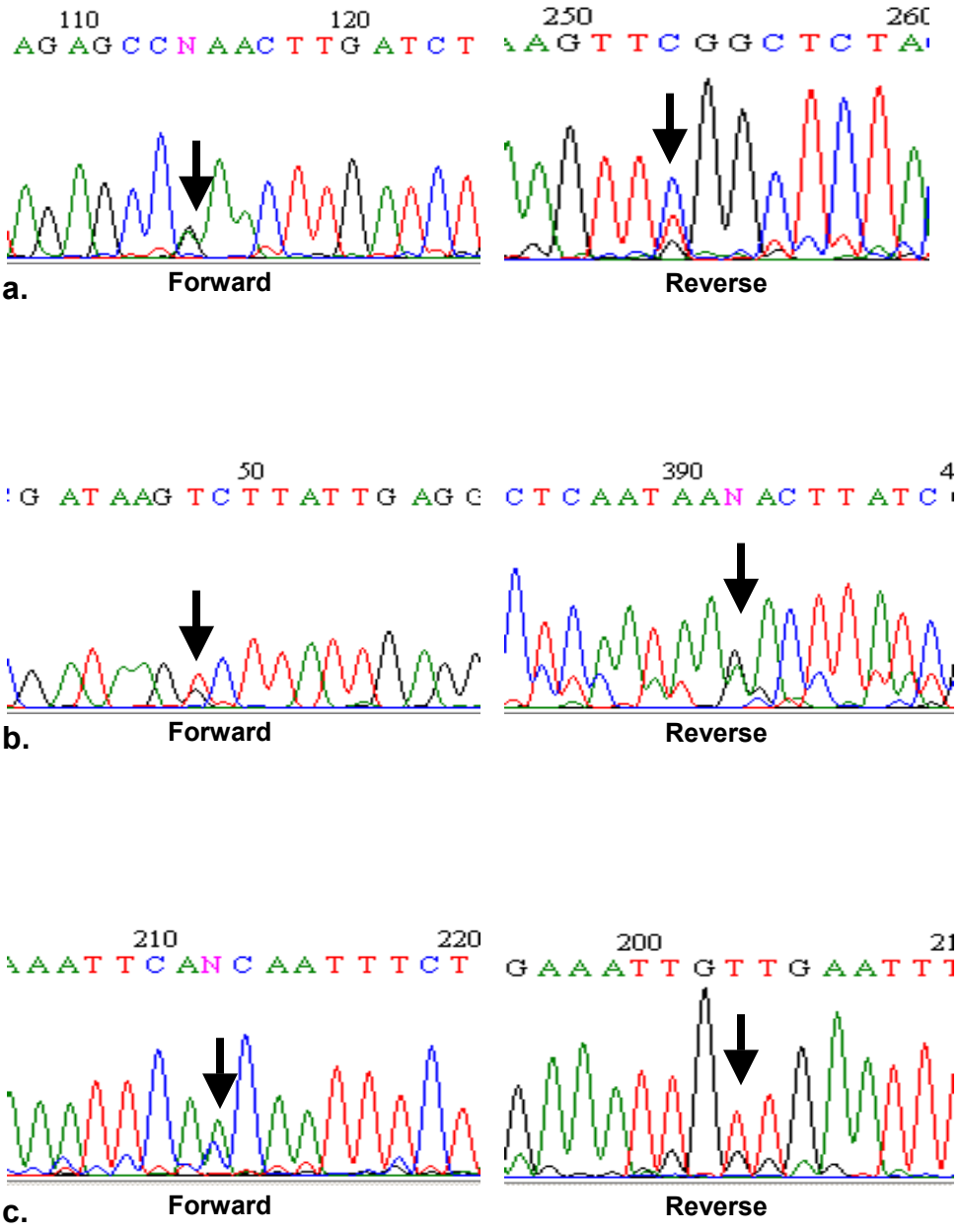



Figure 2.1 Heterozygous coincidence peaks in the forward and reverse chromatograms of a. Nuclear region, *Igs2*, isolate KY98; b. Nuclear region, *Ypt1*, isolate MA0402 and c. Mitochondrial region, *cox2* gene, isolate PE05.

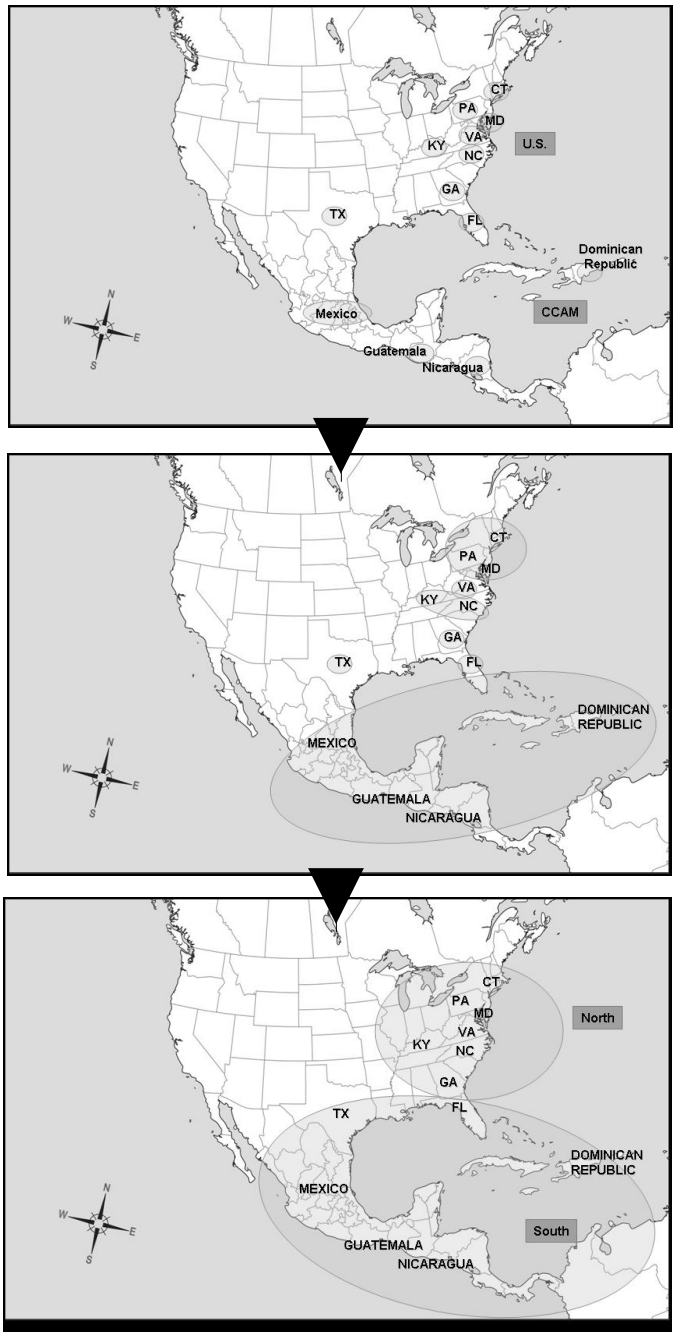



**Figure 2. 2** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the US and CCAM subpopulation of *P. tabacina* pathogen. Locations were grouped into North and South after tests show no genetic subdivision between them.

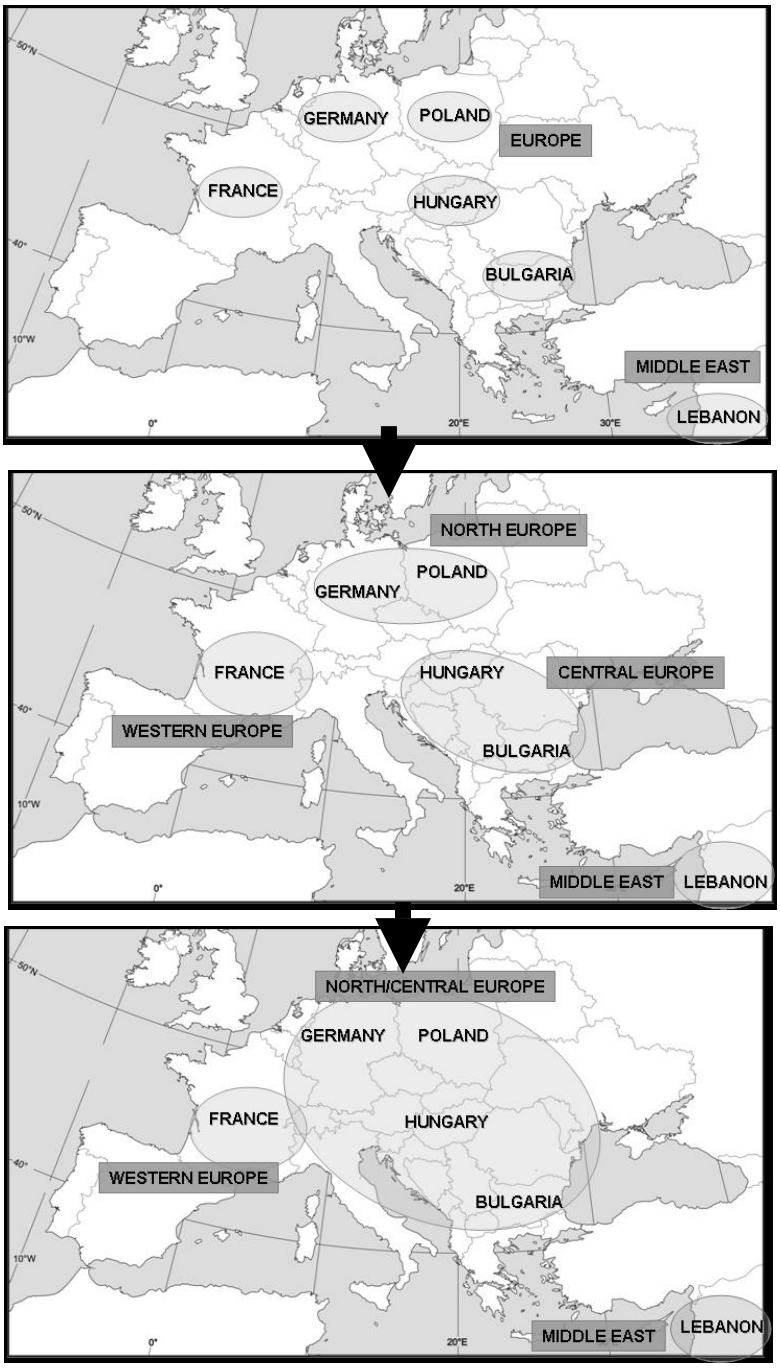
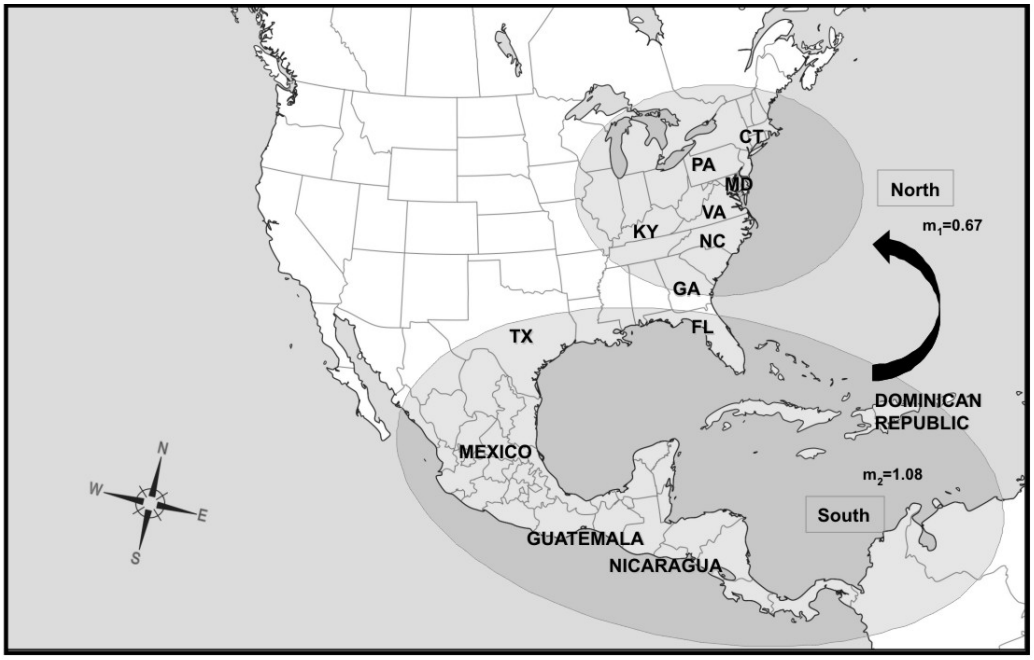



**Figure 2. 3** The Nearest Neighbor Statistic ($S_{nn}$) evaluated on the EULE subpopulation of *P. tabacina* pathogen. Locations were grouped together after tests show no genetic subdivision between them.



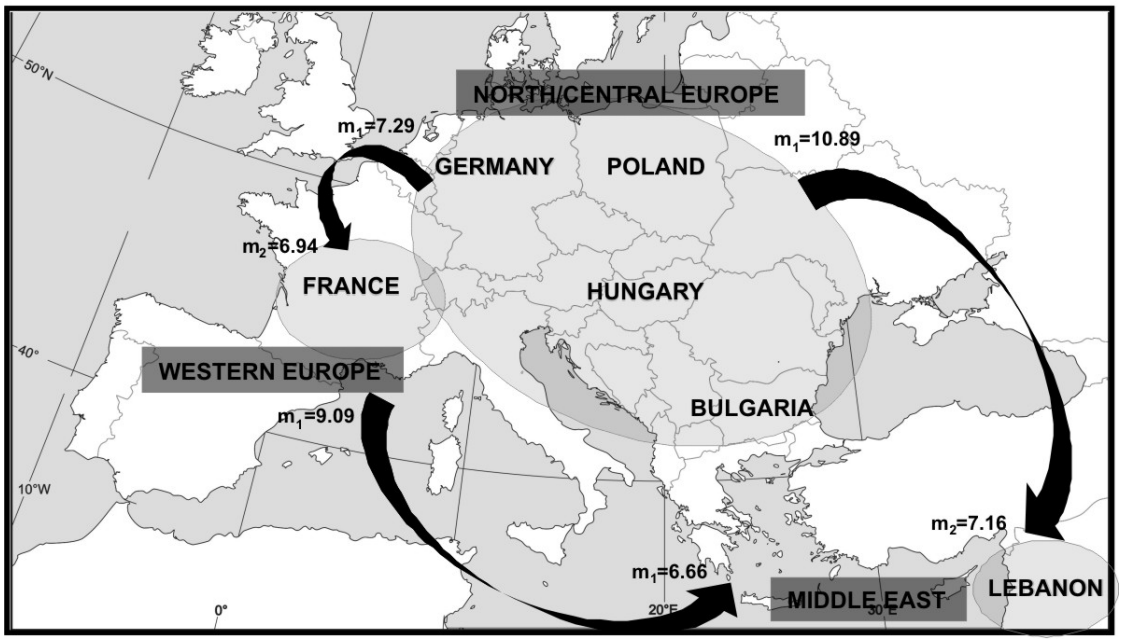




**Figure 2. 4** Isolation with migration (IM) program was used to evaluate the US/CCAM and EULE subpopulation of *P. tabacina* pathogen. Rates of migration are shown in the figures for the *Igs*2 regions but the direction of migration was supported by the two nuclear regions and the mitochondrial gene

# CHAPTER 3

# SEQUENCE AND ANNOTATION OF THE MITOCHONDRIAL GENOMES OF *PERONOSPORA TABACINA* AND *HYALOPERONOSPORA PARASITICA*


## ABSTRACT

Mitochondrial genome sequences are useful for determining evolutionary relationships and tracking migrations of organisms. In this work, portions of the mitochondrial genome of *Peronospora tabacina*, the causal agent of tobacco blue mold and *Hyaloperonospora parasitica* were amplified by PCR and sequenced in order to annotate the genome. *P. infestans* (Ib) mitochondrial proteins were blasted against the whole genome sequence of *Hyaloperonspora parasitica*. Contigs were identified that contained mitochondrial genes of *H. parasitica* and preliminary information on the arrangement and location of these mitochondrial genes were mapped. We also used PCR based approaches. Forty pairs of PCR primers designed from the *Phytophthora infestans* Ib mitochondrial genome were used to amplify and sequence 8 regions of the *P. tabacina* mitochondrial genome. Amplified DNA of *P. tabacina* was used to blast the *P. infestans* (Ib) mitochondrial proteins and the *P tabacina* nucleotides amplified were in the same direction, arrangement and number as those found in the *P. infestans* Ib mitochondrial genome. Specific primers allowed us to sequence approximately one fourth of mitochondrial genome

of *P. tabacina*. The mitochondrial genes identified in the *P tabacina* genome will be useful to study the evolutionary history and population genetics of this important obligate Oomycete plant pathogen.

## INTRODUCTION

Dramatic improvements in our understanding of deep-level evolutionary relationships have occurred in the last few years due to advances in molecular phylogenetics. Different methods have been used to elucidate the diversity of organisms for which data are available, as well as maturing of knowledge about some key features of eukaryotic cells (Voglmayr 2008, Govers and Gijzen 2006).

The sequences from individual genes from mitochondria have been use widely in the last 30 years for evolutionary analysis of organisms. Mitochondrial genomes have been used as a tool for inferring the evolutionary and demographic pasts of both populations and species (Ballard and Rand 2005). The mitochondrial genome is an attractive model to study an organism's evolutionary history because of its mutation rate, uniparental inheritance with rare or no recombination and uniform genetic background due to homoplasmy (Avila-Adame *et al*. 2006, Schena and Cook 2006, Kroon *et al*. 2004, Chesnick *et al*. 2000, Griffith and Shaw 1998). Although the genetic role of mtDNA appears to be universally conserved, the mitochondrial genome exhibits remarkable variation in conformation and size as well

as in actual gene content arrangement and expression within different species (Gray *et al*. 1999).

Over fifteen hundred mitochondrial genomes have been sequenced. The Organelle Genome Megasequencing Program (OGMP) has been investigating mtDNAs from several Stramenopiles. There are 14 complete genomes reported in GenBank which include whole genomes for *Cafeteria roenbergensis* (Accession #NC000946), *Thalassiosira pseudonana* (Ac. #NC007405), *Chrysodidymus synuroideus* (Acc. #NC002174), *Ochromonas danica* (Acc. #NC002571), *Dictyota dichotoma* (Acc. #NC007685), *Fucus vesiculosus* (Acc. #NC007683), *Laminaria digitata* (Acc. #NC004024), *Pilayella littoralis* (Acc. #NC003055), *Desmarestia viridis* (Acc. #NC007684), and *Thraustochytrium aureum* (Acc. #AF288091); and the oomycetes, *Saprolegnia ferax* (Acc. #NC005984), *Phytophthora infestans* (Acc. #NC002387, AY894835, AY898627, AY898628), *P. ramorum* (Acc. #DQ832718) and *P. sojae* (Acc. #DQ832717); and partial genomes for other *Phytophthora* spp. and *Hyaloperonospora parasitica*. The mitochondrial genome size of the oomycete plant pathogens studied so far are larger than 45Mb and all are all diploids (Avila-Adame *et al*. 2006, Govers and Gjizen 2006, Martin 2008, Martin *et al*. 2007, Oudot-Le Secq *et al*. 2006, Paquin *et al*. 1997).

The study of the mitochondrial genomes of Stramenopiles has helped to elucidate the mechanisms and reconstruct the pathways by which evolutionary diversification have occurred in this kingdom. Comparative analysis of complete

mtDNA sequences could provide a genome-level perspective on the presence and arrangement of genes, the function of introns, the distribution and length of spacer sequences and the importance and function of repeated sequences and other relevant information. Studies using mitochondrial genome sequences can help to clarify taxonomic relationships that were previously developed based on morphological features.

Genetic studies of Oomycetes have been done mostly with species of *Phytophthora* a group of plant pathogens that attack a wide range of agricultural and ornamental plants (Gomez-Alpizar *et al*. 2007, Tyler *et al*. 2006, Avila-Adame *et al*. 2006, Klimczak and Prell 1984). Recent phylogenetic studies of species in the Peronosporaceae family have included analysis of the nuclear SSU-rDNA and the LSU-rDNA (Dick *et al*. 1999, Goker *et al*. 2007, Riethmueller *et al*. 2002, Riethmueller *et al*. 1999, Petersen and Rosendahl, 2000). Hudspeth *et al*. 2003 and Cook *et al.* 2001 used the mitochondrial *cox*2 gene to develop phylogenies of downy mildew and white rust taxa and seven marine and three terrestrial representatives of the Peronosporomycetes using neighbor joining, maximum parsimony and maximum likelihood, respectively. The Hudspeth *et al*. (2003) study included three isolates of *Peronosopora tabacina* from Kentucky which were placed according to the consensus tree into the Peronosporomycetidae with the Sclerosporales, Pythiales and Rhipidiales. This group was separated from the Saprolegniomycetidae including the Saprolegniales and the Leptomitales.

The objective of this study was to develop a partial sequence of the mitochondrial genomes of *P. tabacina* and *Hyaloperonospora parasitica,* using the Ib haplotype of *Phytophthora infestans* as a reference genome. These sequences will be useful for further population genetics work and to clarify the extent and types of mutation events that have contributed to genome evolution within these species.

## MATERIALS AND METHODS

### Origin of isolate for sequencing

Isolate Penn07, collected in Pennsylvania during the 2007 epidemic by Mr Greg Seamster from Universal Leaf Co. was used for the analysis. Sporangia from the isolate were collected from tobacco leaf lesions and inoculated onto plants of *Nicotiana repanda.* The sample leaves infected with *P. tabacina* received from the fields were stored under humid conditions at 22°C until sporulation occurred in a Biosecurity Level 2 containment laboratory at the Phytotron facility at North Carolina State University. Spores were then reinoculated by rubbing the lesions onto healthy tobacco plantlets growing under laboratory conditions and free of external contaminants. Further inoculations were performed on approximately 500 tobacco plantlets to collect enough sporangiospores and sporangiophores for the DNA extraction.

**DNA extraction and mitochondrial DNA purification from *P. tabacina***

Total genomic DNA was extracted from sporangiospores and sporangiophores collected from the leaves using a modified cetyltrimethylammonium bromide (CTAB) protocol (Trout *et al*. 1997, White *et al*. 1990) with some modifications (Wangsomboondee and Ristaino 2002). Highest concentrations of DNA were extracted from sporangiospore solutions using the CTAB protocol with some modification according to Avila-Adame *et al.* (2004). Sporangiospores (0.1 g) were placed in sterile 1.5-ml microcentrifuge tubes containing 0.2 g of glass beads to which 150 µl of extraction buffer (0.35 M sorbitol, 0.1 M Tris, 0.005 M EDTA, pH 7.5, and 0.02 M sodium bisulfite) was added and each sample was macerated using a Konte pestle attached to a drill. Twenty tubes containing sporangiosphores were used for each extraction. Extraction buffer (10 ml per 1 g of wet mycelia; 100 mM Tris-HCl, 20 nM EDTA, pH 8.0, 0.7 M NaCl, 1% CTAB, and 1% 2-β-mercaptoethanol) was pre-warmed to 65ºC and added. Tubes were vortexed and incubated at 65ºC for 1.5 hours. After incubation, one volume (15 ml) of chloroform: isoamyl alcohol (24:1) was added to each tube and tubes were centrifuged for 10 min at 10,000 rpm at 15°C on a Sorvall SS-34 rotor (Dupont Instruments, Wisconsin, U.S.). The aqueous phase was removed and 1/10 volume of a solution containing 10% CTAB and 0.7 M NaCl was added. Chloroform extraction was repeated two more times. DNA was precipitated for 30 minutes with a precipitation solution (50 mM Tris-HCl, 10 nM EDTA and 1% CTAB). The supernatant was discarded; pellets

were washed with 70% ethanol and dried at room temperature. DNA was suspended in Tris-EDTA (TE; 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). Twenty tubes with *P. tabacina* sporangia give a total genomic DNA concentration of 50 ng/ul. Each DNA extraction was checked for the presence of *P. tabacina* using the specific PCR primer PTAB (Ristaino *et al*. 2007) and the Real time diagnostic primer PtabBM TaqMan (Blanco-Meneses *et al.*, Chapter 1).

**Annotation of *H. parasitica* genome**

Identification of genes, open reading frames (ORFs), and other biological features in the mitochondrial genomes was facilitated using the previously published sequence of the mitochondrial genome of the Ib mtDNA haplotype of *Phytophthora infestans* available in GenBank (accession number NC002387). Sequence analysis was performed using the software Vector NTI suite 9.0.

The entire mitochondrial of *P. infestans* (mtDNA haplotype Ib) was used to blast the whole genome sequence of *H. parasitica* (http://vmd.vbi.vt.edu/toolkit/), assembles V1.0, V3.0, V6.0 and V7.0, in order to locate contigs containing mitochondrial genes using TBLASTN 2.2.14 program (Altschul *et al*. 1997) (Table 3.1). The start and stop nucleotides of the hits were used to determine the contigs with mitochondrial genes and generate the possible order in *H. parasitica*. Twenty four contigs were selected from the *H. parasitica* total genome and blasted against

the *P. infestans* mitochondrial genome. The 24 contigs gave 6x coverage of the mitochondrial genome. Circular prediction of the arrangement of the mitochondrial genes of *H. parasitica* on the contigs was plotted in Vector NTI (Figure 3.1).

**PCR amplification of mtDNA genes from *P. tabacina* with *P. infestans* primers**

Oomycete mitochondrial genomes contain conserved gene order and arrangement of some genes and noncoding sequences (Lang *et al*. 1999). Forty one primer pairs specific for mitochondrial genes were designed (Table 3.2) using the *Phytophthora infestans* (mtDNA haplotype Ib) mitochondrial genome sequence. It was assumed that since *P. infestans* and *H. parasitica* showed a high degree of coding region similarity, that a high degree of similarity between the *H. parasitica* and the *P. tabacina* mitochondrial genomes would be found. Thus, PCR amplification of the *P. tabacina* mitochondrial genome would be possible using primers made from *P. infestans* genes.

Amplification of regions of the mitochondrial genome of *P. tabacina* using the 40 *P. infestans* primer pairs was performed in separate PCR reactions and each primer was tested twice. Each PCR reaction consisted of the following: 1 µl of *P. tabacina* genomic DNA (1:100 dilution of original DNA extract in TE buffer, about 10 ng/µl), 49-µl master reaction mixture containing 5 µl of 10X PCR buffer (100 mM Tris-HCl, 15 mM $MgCl_2$, and 500 mM KCl, pH 8.3), 35.25 µl of sterile distilled $H_2O$,

1.8 µl of 10 mM $MgCl_2$, 2 µl of 2.50 mM dNTPs, 2 µl each of 10 µM primer pairs (Table 3.2), 0.25 µl of BSA (Roche, Nutley, NJ) and 0.2 µl of Taq polymerase (Invitrogen, Carlsbad, California). The program SCO (Schena and Cook 2006) was used. Thermal cycling parameters were initial denaturation at 96ºC for 2 min, followed by 35 cycles consisting of denaturation at 96ºC for 1 min, annealing at the recommended primer temperature for 1 min, and extension at 72ºC for 2 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min.

Sequences obtained from PCR amplification with specific mtDNA primers including Mt1F/Mt6R, CoxF4N/COXR4N, COX2F/COX2R, and FM78/FM82 (Schena and Cooke, 2008, Kroon et al, 2004, Hudspeth et al, 2000, and Martin and Tooley, 2003) were also used to develop a partial sequence of the mitochondrial genome of *P. tabacina* (Table 3.3). One of two thermocycling protocols were used and are shown in Table 3.3. Amplified products were separated on 1.6% agarose gels containing ethidium bromide at 0.5 µg/ml with 1X Tris-borate-EDTA (TBE) running buffer. DNA bands were visualized using UV light on a transluminator (BioRad, Hercules, CA.).

The mtDNA amplified from *P. tabacina* was sequenced using an ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit and analyzed on an ABI PRISM 377 automated sequencer (Applied Biosystem, Foster City, California) at the North Carolina State University Genome Research Laboratory.

### Annotation of *P. tabacina* mitochondrial genome

DNA sequences were aligned and edited manually using BioEdit Sequence Alignment Editor Version 7.0.5.3. (Hall 1999). All polymorphisms were rechecked from the chromatograms.

Alignments between the mtDNA regions of *P. tabacina* and *P. infestans* mtDNA haplotype Ib (GenBank accession NC002387) were done using BioEdit. Vector NTI suit 9.0 was used to generate the circular and the linear gene predictions of the arrangement of the mtDNA regions from *P. tabacina* (Figure 3.2 and 3.3).

## RESULTS

### Amplification of mtDNA from *P. tabacina* with *P. infestans* primers

Portions of the mitochondrial genome of *P. tabacina* were successfully amplified with seven primer pairs designed from the *P. infestans* mt genome (Table 3.2). Four primer pairs designed from other research (Table 3.3) also amplified *P. tabacina* mtDNA. Amplified DNA was aligned with the *P infestans* mt haplotype Ib (Fig. 3.2). The entire amplified target DNA was used to annotate and align the *P. tabacina* mt genome with the exception of amplified DNA from the F6F/F6R primers since they amplified the same genes amplified by the primer combination Ptmt7F/Ptmt7R.

Approximately 8168 bp of *P. tabacina* mtDNA was amplified from coding regions and 507 bp from noncoding regions using the 8 mtDNA primer combinations and the genes amplified by the primers IgCox and FM 78/82. This represents approximately one third of the mtDNA genome of *P. tabacina*. These regions were designated *P. tabacina* 1-8 and correspond to gene regions amplified by primer pairs (Ptmt2, Ptmt7, cox2-Igscox-cox1 (from table 3.3), Ptmt17, Ptmt22, Ptmt31, Ptmt12 and Ptmt34 respectively) (Table 3.2).

**The mt genome sequence of *Hyaloperonopsora parasitica***

A total of 18 genes involved in electron transport, 2 RNA-encoding genes, 16 ribosomal protein genes and 25 transfer-RNA genes are encoded on both strands in the Ib haplotype of *P. infestans* (Avila-Adame *et al*. 2006). The coding and non coding regions are highly similar between the *P. infestans* and the *H. parasitica* mitochondrial genomes and only some variation in the size of the genes could be observed (Table 3.4). The mtDNA gene order in *P. infestans* is conserved to a high degree in *H. parasitica* (Fig. 3.1). Most of the mitochondrial genes were found on the 24 contigs in the whole genome sequence of *H. parasitica*. The 6x coverage of the regions gave support to the polymorphism found in the mitochondrial genome which differentiates from the ones in *P. infestans*.

**The mt genome of *P. tabacina***

PCR amplification of *P. tabacina* genomic DNA with 40 mitochondrial genome primer pairs from *P. infestans* was done and 8 regions of the mitochondrial genome were sequenced and annotated (Fig. 3.2). Specific genes present in *P. infestans* mt haplotype Ib were also present in *P. tabacina* and their size was determined (Table 3.4). Gene order was conserved in the *P. tabacina* mitochondrial genome within the 8 regions that were sequenced by PCR. Final gene order arrangement and direction is not yet known so the linearized versions of portions of the genome are shown (Fig. 3.3)

An average of 88.30% of the *P. tabacina* mitochondrial gene regions was identical to those of *P. infestans* Ib (Table 3.4). We annotated three regions from the large subunit of the ribosomal DNA (rDNA), nine regions from the transfer RNA (tRNA), part of the region from the small subunit of the rDNA, NADH dehydrogenase subunits 4 and 11, ribosomal proteins S12, S7, S13 and L2, part of the SecY-independent transporter, part or total genes *cox*1, *cox*2 and *cox*3 and the open reading frames 32 and 42.

**DISCUSSION**

Mitochondrial genome sequences are useful for determining evolutionary relationships and tracking migrations of organisms. In this work, portions of the

mitochondrial genome of *P. tabacina*, the causal agent of tobacco blue mold were amplified by PCR and sequenced in order to annotate the genome.

PCR methodology has been used for the amplification, sequencing and annotation of plastomes, specifically chloroplast DNA (Dhingra and Folta 2005). A blast analysis of the *P. infestans* mitochondrial genes was performed to design the primers to amplify the *P. tabacina* mitochondrial genome. *P. infestans* (Ib) mitochondrial was blasted against the entire genome sequence of *H. parasitica*. Contigs in the *H. parasitica* genome were identified that contained mitochondrial genes and the arrangement and location of these mitochondrial genes was determined.

The PCR methodology was more rapid than the traditional CsCl purification but only 7 of the forty one primer pairs from *P. infestans* amplified *P. tabacina* mtDNA. However, the sequences generated were of good quality and when blasted with the *P. infestans* (Ib) mitochondrial proteins, the *P. tabacina* genes were in the same direction and arrangement as those found in the *P. infestans* Ib genome sequence.

Specific primers allowed us to sequence approximately one fourth of the total mitochondrial genome of *P. tabacina* using an average size mtDNA genome of other Oomycete plant pathogens for comparison. The size of the *P. infestans* mitochondrial genome is between 36016 and 37957 bp (Avila-Adame *et al*. 2007). Other mitochondrial genome sizes are: *P. ramorum* 39314 bp (Martin 2008); *P.*

*sojae* 42975 bp. (Tyler *et al*. 2006); and *Saprolegnia ferax* 46930 bp (Grayburn *et al.* 2004).

**Gene content and gene order in *Hyaloperonospora parasitica***

We compared the mitochondrial genome of *P. infestans* to the *Hyaloperonospora parasitica* mitochondrial genome. There was a high degree of similarity between the genomes. But 2 regions covering coding and non coding regions varied in direction between the *H. parasitica* and *P. infestans* mitochondrial genomes. The physical arrangement and direction of 3 of the mitochondrial genomes of *Phytophthora* species (*P. infestans*, *P. sojae* and *P. ramorum*) that have been sequenced were also compared (Martin 2008, Martin *et al*. 2007, Paquin *et al*. 1997). The region that covers cob, nad9 and atp9 were identical for the *H. parasitica* and *P. ramorum and P. sojae* genomes but not for *P. infestans*. The region that covers nad5, nad6, trnR, nad4L, nad1 and nad11 varied between the four different genomes and further analyisis will be necessary to stablish the direction of the region. The next step will be to obtain the maximum colinearity between the *H. parasitica* and *P. tabacina* genomes. We will then use the *H. parasitica* sequences to generate specific mitochondrial primers that will allow the amplification of the rest of the *P. tabacina* mitochondrial genome.

**Gene content and order in *P. tabacina***

There was a high degree of similarity (> 88%) between the mitochondrial genome of *P. tabacina* and *P. infestans*. Both genomes shared similar gene order and arrangement.

Seven of the 41 primer combinations designed from the *P. infestans* mitochondrial genome amplified the *P. tabacina* mitochondrial genome allowing us to obtain approximately a fourth of the entire mitochondrial genome. The regions of DNA that were successfully amplified were spread around the genome and many of them have a gap between genes that will allow the creation of new primers using the flanking regions of the sequences. We plan to use the preliminary sequences of both the *H. parasitica* and the *P. tabacina* mitochondrial genomes to design specific primers to complete the mitochondrial genome sequence of *P. tabacina*.

**Phylogenetics and evolutionary history**

Total mtDNA genome sequences of *P. tabacina* will provide further information for comparison with other Oomycetes members and possibly a more robust phylogeny could be obtained comparing the different organisms. Complete mitochondrial genomes offer the opportunity to evaluate divergence and evolutionary relationships between species.

Complete mitochondrial genomes sequences for three other Oomycete plant pathogens including *P. infestans, P ramorum and P. sojae* have been completed (Paquin *et al*. 1997, Avila-Adame *et al*. 2006, Martin *et al.* 2007, Martin 2008). There are four mitochondrial haplotypes in *P. infestans* (Avila- Adame, *et al*. 2006, Lang and Forget, 1993) that have been used in population genetic studies to characterize isolates of *P. infestans* on a global basis. The evolutionary relationships between the four mitochondrial lineages of *P. infestans* indicated that the type I and type II lineages evolved independently from a common ancestor and the mutations associated with them have been characterized (Avila-Adame *et al*. 2006).

The whole mitochondrial genome sequence of *P. ramorum* was used to identify an additional haplotype not found in previous studies when only *cox*1 gene was used (Martin 2008, Kroon *et al*. 2004, Ivors *et al*. 2006). Tyler *et al*. (2006) published the whole genome sequence of *P. ramorum* and *P. sojae* and Martin (2008) used these sequences to annotate the mitochondrial genomes of these species. In the whole genome sequencing study of *P. sojae* and *P. ramorum*, an ancestral phototrophic organism was proposed as an ancestor for the Stramenopila kingdom (Tyler *et al*. 2006).

**Population genetics**

The mitochondrial *cox2* gene was sequenced from 54 isolates of *P tabacina* from global populations (Chapter 2). One dominant haplotype and many rare haplotypes were identified in that study. We have mapped the SNPs present in this gene. These could be used to either design PCR based RFLPs to separate more common haplotypes, or a SNP chip could be developed for more high throughput haplotyping of isolates.

The analysis of genes and features of the genome of the mitochondrion of *P. tabacina* will be useful to elucidate the phylogenetic position of this pathogen.

The polymorphisms found in this species could be employed to study the origin, migration and population genetics of related downy mildew pathogens. Determining the amount and distribution of genetic variation can contribute to our undestanding of the evolutionary history of the Oomycetes as well as their potential to adapt to changing environments. The identification of specific mitochondrial haplotypes will be useful for tracking migrations of this pathogen.

## REFERENCES CITED


1. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

2. Avila-Adame, C., Gomez-Alpizar, L., Zismann, V., Jones, K.M., Buell, C.R., and Ristaino, J.B. 2006. Mitochondrial genome sequences and molecular evolution of the Irish potato famine pathogen, *Phytophthora infestans*. Curr. Genet. 49:39-46.

3. Ballard, W., and Rand, D. 2005. The population biology of mitochondrial DNA and its phylogenetic implications. Annu. Rev. Ecol. Syst. 36:621-642.

4. Chesnick, J. M., Goff, M., Graham J., Ocampo, C., Lang, F., Seif, E., and Burger, G. 2000. The mitochondrial genome of the stramenopile alga *Chrysodidymus synuroideus*. Complete sequence, gene content and genome organization. Nucleic Acids Res. 28: 2512-2518.

5. Cook, K. L., Hudspeth, D. S. S., and Hudspeth, M. E. S. 2001. A *cox*2 phylogeny of terrestrial and marine parasitic Peronosporomycetes (Oomycetes). Nova Hedwigia. 122:231-243.

6. Dhingra, A., and Folta, K. M. 2005. ASAP: Amplification, sequencing and annotation of plastomes. BMC Genomics 6:176-189.

7. Dick, M. W., Vick, M. C., Gibbings, J. G., Hedderson, T. A., and Lopez-Lastra, C. C. 1999. 18S rDNA for species of *Leptolegnia* and other Peronosporomycetes: justification for the subclass taxa Saprolegniomycetidae and Peronosporomycetidae and division of the Saprolegniaceae *sensu lato* into the families Leptolegniaceae and Saprolegniaceae. Mycological Research 103:1119-1125.

8. Goker, M., Voglmayr, H., Riethmuller, A., and Oberwinkler, F. 2007. How do obligate parasites evolve? A multi-gene phylogenetic analysis of downy mildews. Fungal Gen. and Biol. 44:105-122.

9. Gómez-Alpizar, L, I. Carbone, and J. B. Ristaino. 2007. An Andean origin of *Phytophthora infestans* inferred from nuclear and mitochondrial gene genealogies. Proc Nat. Acad. Sci. 104: 3306-3311.

10. Govers, F., and Gjizen M. 2006. *Phytophthora* genomics: the plant destroyers' genome decoded. MPMI. 19(12): 1295-1301.

11. Gray, M. W., Burger, G., and Lang, F. 1999. Mitochondrial Evolution. Science 283:1476-1481.

12. Grayburn, W. S., Hudspeth, D. S. S., Gane, M. K., and Hudspeth, M. E. S. The mitochondrial genome of *Saprolegnia ferax*: organization, gene content and nucleotide sequence. Mycologia 96(5):981-989.

13. Griffith G., Shaw D. 1998. Applied and environmental microbiology. 64(10): 4007-4014.

14. Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

15. Hudspeth, D. S. S., Stenger, D., and Hudspeth, M. E. S. 2003. A *cox*2 phylogenetic hypothesis for the downy mildews and white rusts. Fungal Diversity 13: 47-57.

16. Ivors, K., Garbelotto, M., Vries, I-D. E., Ruyter-Spira, C., Hekkert, Bt., Rosenzweig, N., and Bonants, P. 2006. Microsatellite markers identify three lineages of *Phytophthora ramorum* is US nuerseries, yet single lineages in US forest and European nursery populations. Mol. Ecol. 15: 1439-1505.

17. Klimczak, L.J., and Prell, H.H. 1984. Isolation and characterization of mitochondrial DNA of the oomycetous fungus *Phytophthora infestans*. Curr. Genet. 8:323-326.

18. Kroon, L. P. N. M., Bakker, F. T., van den Bosch, G. B. M., Bonants, P. J. M., and Flier, W.G. 2004. Phylogenetic analysis of *Phytophthora* species based on mitochondrial and nuclear DNA sequences. Fungal Genetics and Biology. 41: 766-782.

19. Lang, B. F., Gray, M., and Burger, G. 1999. Mitochondrial genome evolution and the origins of eukaryotes. Annu. Rev. Genet. 33:351-397.

20. Lang BF, and Forget L 1993. The mitochondrial genome of *Phytophthora infestans*. In: O'Brien SJ (ed) Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Press, Plain View, New York, pp 3.133-3.135.

21. Martin, F.N. 2008. Mitochondrial haplotype determination in the oomycete plant pathogen *Phytophthora ramorum.* Curr. Genet. 54:23-34.

22. Martin, F., Bensasson, D., Tyler, B., and Boore, J. 2007. Mitochondrial genome sequences and comparative genomics of *Phytophthora ramorum* and *P. sojae*. Curr Genet. 51: 285-296.

23. Oudot-Le Secq, M., Loiseaux-de Goer, S., Stam, W., and Olsen, J. 2006. Complete mitochondrial genomes of the three algae (Heterokonta: Phaeophyceae) *Dictyota dichotoma*, *Fucus vesiculosus*, *Desmarestia viridis*. Curr Genet. 49: 47-58.

24. Paquin, B., Laforest, M. J., Forget, L., Roewer, I., Wang, Z., Longcore, J. and Lang, B. F. 1997. The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. Curr Genet. 31:380-395.

25. Petersen, A. B., and Rosendahl, S. 2000. Phylogeny of the Peronosporomycetes (Oomycota) based on partial sequences of the large ribosomal subunit (LSU rDNA). Mycological Reseach. 104:1295-1303.

26. Riethmuller, A., Volglmayr, H., Goker, M., Weib, M., and Oberwinkler, F. 2002. Phylogenetic relationship of the downy mildews (Peronosporales) and related groups based on nuclear large subunit ribosomal DNA sequences. Mycologia 94:834-849.

27. Riethmuller, A., Weib, M., and Oberwinkler, F. 1999. Phylogenetic studies of Saprolegniomycetidae and related groups based on nuclear large subunit DNA sequences. Can. J. Bot. 77:1790-1800.

28. Ristaino J.B., Johnson A., Blanco-Meneses M., and Liu B. 2007. Identification of the tobacco blue mold pathogen, *Peronospora tabacina*, by polymerase chain reaction. Plant Dis. 91: 685-691.

29. Schena, L., and Cooke, D. 2006. Assessing the potential of regions of the nuclear and mitochondrial genome to develop a “molecular tool box” for the detection and characterization of *Phytophthora* species. Journal of Microbiological Methods. 67: 70-85.

30. Trout, C. L., Ristaino, J. B., Madritch, M., and Wangsomboondee, T. 1997. Rapid detection of *Phytophthora infestans* in late blight infected tissue of potato and tomato using PCR. Plant Dis. 81:1042-1048.

31. Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R., Aerts, A., Damasceno, C. M. B., Dou, D., Dubchak, I., Gijzen, M., Gordon, S., Govers, F., Grunwald, N., Huang, W., Ivors, K., Kamoun, S., Krampis, K., Lamour, K., McDonald, W. H., Medina, M., Meijer, H., Nordberg, E., Ospina- Giraldo, M. D., Morris, P., Putnam, N., Rash, S., Rose, J. K. C, Sakihama, Y., Salamov, A., Savidor, A., Smith, B., Smith, J., Sobral, B. W. S., Terry, A., Torto-Alalibo, T., Win, J., Zhang, H., Grigoriev, I., Rokhsar, D., and Boore, J. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313:1261-1266.

32. Voglmayr, H. 2008. Progress and challenges in systematics of downy mildews and white blister rusts: new insights from genes and morphology. Eur. J. Plant Pathol. 122:3-18.

33. Wangsomboondee T., and Ristaino, J. B. 2002. Optimization of samples size and DNA extraction methods to improve PCR detection of different propagules of *Phytophthora infestans*. Plant Dis. 86: 247-253.

34. White, T. J., Bruns, T., Lee S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315-322 in: PCR Protocols: A Guide to Methods and Applications. M. A. Innis, D. H. Gelfand, J. J. Sninshy, and T. J. White, eds. Academic Press, New York.

**Table 3. 1** Comparision of *P. infestans* against *H. parasitica* genes. Annotation of the mitochondrial genome of *H. parasitica* was performed using the protein sequences from the mitochondrial genome of *P. infestans* Ib as a reference.

| *Phytophthora infestans* Ib mitocondrial genome [a] | | | *Hyaloperonospora parasitica* total genome [b] | | | |
|---|---|---|---|---|---|---|
| Start base | Stop base | Function | Contig | Start base | Stop base | Orientation |
| 835 | 230 | Ribosomal protein L6 | 1816 | 747 | 1295 | + |
| 1223 | 843 | Ribosomal protein S8 | 1816 | 311 | 685 | + |
| 1536 | 1237 | Ribosomal protein S14 | 1816 | 1 | 294 | + |
| 2116 | 1724 | ATP synthase F0 subunit 8 | NONE | | | |
| 2831 | 2178 | Hypothetical protein | 2745 | 330 | 4 | - |
| 3332 | 2928 | Ribosomal protein L16 | 2745 | 921 | 520 | - |
| 4138 | 3335 | Ribosomal protein S3 | 2745 | 1290 | 958 | - |
| 4375 | 4142 | Ribosomal protein S19 | NONE | | | |
| 5185 | 4379 | Ribosomal protein L2 | 585 | 12282 | 13025 | + |
| 5214 | 5627 | Ribosomal protein S13 | 585 | 12253 | 11933 | - |
| 5640 | 6056 | Ribosomal protein S11 | 585 | 11830 | 11414 | - |
| 6264 | 6470 | Hypothetical protein | 585 | 11179 | 11069 | - |
| 7181 | 6435 | SecY-independent transporter protein | 585 | 8835 | 9131 | + |
| 7461 | 9467 | NADH dehydrogenase subunit 11 | 135 | 11649 | 13598 | + |
| 9464 | 10444 | NADH dehydrogenase subunit 1 | 135 | 13652 | 14629 | + |
| 10447 | 10749 | NADH-ubiquinone oxidoreductase chain 4L | 135 | 14771 | 14935 | + |
| 11723 | 11013 | NADH dehydrogenase subunit 6 | 135 | 15819 | 15118 | - |
| 13761 | 11767 | NADH-ubiquinone oxidoreductase chain 5 | 135 | 17803 | 15953 | - |
| 13761 | 11767 | NADH-ubiquinone oxidoreductase chain 5 | 135 | 19163 | 18339 | - |
| 14334 | 13981 | NADH-ubiquinone oxidoreductase chain 3 | 135 | 11040 | 11390 | - |
| 15194 | 14475 | ATP synthase A chain | 135 | 10196 | 10912 | + |
| 16139 | 15222 | Cytochrome c oxidase subunit 3 | 135 | 9269 | 10180 | + |
| 16575 | 16195 | Ribosomal protein S7 | 135 | 8828 | 9208 | + |
| 16978 | 16598 | Ribosomal protein S12 | 135 | 8428 | 8802 | + |
| 17631 | 17305 | Ribosomal protein S10 | 135 | 7756 | 7980 | + |
| 19223 | 17730 | NADH dehydrogenase subunit 2 | 135 | 6621 | 7661 | + |
| 19341 | 20519 | NADH dehydrogenase subunit 7 | 135 | 6060 | 4885 | - |
| 20529 | 20957 | Hypothetical protein | 135 | 4754 | 4368 | - |
| 21098 | 22531 | NADH-ubiquinone oxidoreductase chain 4 | 135 | 4226 | 2796 | - |
| 22713 | 24242 | ATP synthase alpha chain | 135 | 2594 | 1068 | - |
| 25532 | 24381 | Cytochrome b | 585 | 7367 | 6228 | - |
| 26159 | 25593 | NADH-ubiquinone oxidoreductase | 585 | 7981 | 7427 | - |
| 26548 | 26321 | ATP synthase subunit 9 | 585 | 8339 | 8115 | - |
| 28175 | 26697 | Cytochrome c oxidase subunit 1 | 585 | 4047 | 5522 | + |
| 28175 | 26697 | Cytochrome c oxidase subunit 1 | 585 | 2555 | 3142 | + |
| 29153 | 28377 | Cytochrome c oxidase subunit 2 | 585 | 1564 | 2340 | + |
| 29812 | 29573 | Hypothetical protein | NONE | | | |
| 33265 | 32732 | Ribosomal protein L5 | 5351 | 692 | 222 | - |
| 33643 | 33272 | Ribosomal protein L14 | 5351 | 1071 | 703 | - |
| 35089 | 34787 | Hypothetical protein | 3742 | 172 | 53 | - |
| 35553 | 35092 | Ribosomal protein S4 | 3742 | 776 | 318 | - |
| 36153 | 35563 | Ribosomal protein S2 | 3742 | 1147 | 788 | - |

[a] Each gene on the *P. infestans* Ib mitochodrial genome is specify using the counterclockwise direction from the physical arrangment.

[b] V1.0 assembly from *H. parasitica* entire genome was used to blast the *H. parasitica* against the *P. infestans* Ib mitochondrial genome.

**Table 3. 2** Primers sequences and annealing temperature designed for amplification of *P. tabacina* mitocondrial genome.

| Name | Sequence 5' to 3' | Annealing temperature (°C) |
|---|---|---|
| Ptmt-f0 | ATGGTTTAAGTATAAATAATAAAAGC | |
| Ptmt-r0 | AGATATTACGCTTTCTTTAAAGGATG | 55 |
| Ptmt-f1 | CTAAATAATATTCTAAGTGAAAAAGG | |
| Ptmt-r1 | ATATTAGTTATTTATTATAATAAGAG | 48 |
| Ptmt-f2 [a] | TAATACGTGTAGCATAAGTGGTAATC | |
| Ptmt-r2 [a] | AATAATAATGATGGTAGGATTTGAAC | 56 |
| Ptmt-f3 | TTATTTTCTTAATAATTAATAAGATG | |
| Ptmt-r3 | TTAAAATTTATTTTAGTTAAATTTGG | 51 |
| Ptmt-f4 | TTATTTTCTTAATAATTAATAAGATG | |
| Ptmt-r4 | AAATATCATGGAAATTATCAAATAGG | 55 |
| Ptmt-f5 | ATAACTGGCAGATATTTTAATAGATG | |
| Ptmt-r5 | AAGGTGTTATTAATATTAATAATAGG | 52 |
| Ptmt-f6 | TTATTTTGATAATTTAAAATGTTTTG | |
| Ptmt-r6 | CTTCTATTATCAACGTATTTTATCAC | 53 |
| Ptmt-f7 [a] | TTATTTTCTTAATAATTAATAAGATG | |
| Ptmt-r7 [a] | CTCAACTTTATACATTTTTTCAAAAC | 58 |
| Ptmt-f8 | TAAAAAAAAATCAACAATATTATGAC | |
| Ptmt-r8 | TTGCTGGATCTTGAAAACCTAATTGC | 59 |
| Ptmt-f9 | ATAATCAATTTACTTTTTTAGATATG | |
| Ptmt-r9 | AGCACTAAAAATTAAATATAAAGTTC | 51 |
| Ptmt-f10 | TGAATTTTCAAAATATAAATAAATGG | |
| Ptmt-r10 | TAATAAAGGTGTTTCAAATTTTAGAG | 55 |
| Ptmt-f11 | AGCAGCTACTATGATTATTGCGGTAC | |
| Ptmt-r11 | TTAGATATTTACTATAAATAATTACG | 55 |
| Ptmt-f12 [a] | TTTGCTTTTTAATTAAAATTAAATCC | |
| Ptmt-r12 [a] | ATTTGCATGTGTATATCTAATTAACC | 56 |
| Ptmt-f13 | CATTATACACCACATATCGATTTAGC | |
| Ptmt-r13 | AAGTAATATAATAATAAATTACTTAG | 51 |
| Ptmt-f14 | TATACCATTTTTAGGTAAATTTGAAG | |
| Ptmt-r14 | TCTGTTATTGTAGCAGCTACTGCTTC | 58 |
| Ptmt-f15 | TTTCCATTATCTCTAAAATATTCACC | |
| Ptmt-r15 | GGTTTATATCCTAAAATTTTTGAAAG | 56 |
| Ptmt-f16 | TTATTTAATAAAAATAAGAAAAATCG | |
| Ptmt-r16 | CAATGTTTTTAATTATTGGTATTTGG | 55 |
| Ptmt-f17 [a] | AAACATAAATAATGAACCTGCTAATG | |
| Ptmt-r17 [a] | AATCAATTACAAAAAATTAAACAAAC | 56 |
| Ptmt-f18 | GAATTTTTTTTAAAGATATTATTTCG | |
| Ptmt-r18 | TCAATATATTCGAGTTTTATTTTCTG | 55 |
| Ptmt-f19 | AGGTGTCATAGCTCCTACATCTAAAG | |
| Ptmt-r19 | TAATATAAAACCTGAAGCAATTGATC | 58 |
| Ptmt-f20 | ACAAAAAAAATCAATATTATCTATTG | |
| Ptmt-r20 | TGATCCTAAAATAAAATATTTTAAGC | 53 |
| Ptmt-f21 | AAGCCCAGCCTTAAAACAATGTCCTC | |
| Ptmt-r21 | ATGGACTTTTGCTACTTTACGCATAG | 64 |
| Ptmt-f22 [a] | TTAAAACAATGTCCTCAACGTAAAGG | |
| Ptmt-r22 [a] | ACCTGTTAAAGTTAAAAAATTACTAC | 57 |

Table 3.2 continued

| Name | Sequence 5' to 3' | Annealing temperature (°C) |
|---|---|---|
| Ptmt-f23 | CATGGCCTTTTGTTATTTCATTTGGG | |
| Ptmt-r23 | CATAATAAAATAAAAAAAAAAACAAC | 60 |
| Ptmt-f24 | GTTACTATTGAATTAACTTATTCAAG | |
| Ptmt-r24 | TTTTCAATTTACCATTCTAAAGCATC | 55 |
| Ptmt-f25 | TTCATGGGGATTTTGGACTATGATTG | |
| Ptmt-r25 | TAGCCATAAATGTAAAAAAATTTGAG | 62 |
| Ptmt-f26 | TAATTATATTTTAAGTATTATTGATG | |
| Ptmt-r26 | ATTTAAAAAGTTATAAAAATAAAGAC | 48 |
| Ptmt-f27 | ATTTTTATATAATTTAAAAATTTCAG | |
| Ptmt-r27 | TTTTTTTCTTTTTTTGTTAAAGTTAC | 52 |
| Ptmt-f28 | AGTTATTATGCTTTTTTTTTTTTAGC | |
| Ptmt-r28 | GAATATGCGAATATGTTATTAATGAG | 56 |
| Ptmt-f29 | AAATTAATAAATTTAATACTAAAAGG | |
| Ptmt-r29 | AAGTTTTTAATAAGTAAAAATCAATC | 51 |
| Ptmt-f30 | TTTAATTTTTTTTTTAATAAATAAGG | |
| Ptmt-r30 | AAACACGTTACACTTAGGATGTAAAG | 55 |
| Ptmt-f31 [a] | TATTAATTATATTTAAAGCTTCTAAC | |
| Ptmt-r31 [a] | TACACTTAGGATGTAAAGCTGTAAAG | 52 |
| Ptmt-f32 | ATATATCTTTATAGAAATTATCAAAG | |
| Ptmt-r32 | AGTGGTGAAATTGGTAAACACATTAC | 54 |
| Ptmt-f33 | TTTATTTTTATTTAAATTAATATCTG | |
| Ptmt-r33 | ATTATATTAACAAAAATTTAAAAGTC | 49 |
| Ptmt-f34 [a] | ATTCCACGATTACTTTTAATTTCTAC | |
| Ptmt-r34 [a] | TATCTTACCTTGAGGTGTTACAGAAG | 57 |
| Ptmt-f35 | TGTACGTGGTGTAGCAAAAAATCCTG | |
| Ptmt-r35 | TTTTGAAAATTCATTTTTTTTTTACG | 62 |
| Ptmt-f36 | TTTCAAATGTTTTCCAATTTTATTGG | |
| Ptmt-r36 | TAATATTATATTTTCATATTTTACAG | 54 |
| Ptmt-f37 | AATAATTATGAAAATATTTATTTACC | |
| Ptmt-r37 | ATATCCCCTAATAAGTGTATTGGGAC | 54 |
| Ptmt-f38 | CTAATTAACTTAAAATAAGTGTATAG | |
| Ptmt-r38 | CTCAAAAAAAAATTGTTTTATTTGTAG | 51 |
| Ptmt-f39 | TTTATAGATAAAACTTTTTTTTTTCC | |
| Ptmt-r39 | ATCTATATTTTTAATTGTTTGATTTG | 53 |
| Ptmt-f40 | AATAAAAGATACTTTTTTAAATCAAG | |
| Ptmt-r40 | ATTAATTTTAAATTTAAATAAAAAGG | 51 |

The primers were kindly desing by Dr. Amit Dhingra, University of Florida.

[a] Primer combinations used to amplify *P. tabacina* .

Note: *P. tabacina* fragments 1-8 correspond to: Ptmt2 (P.tabacina-1), Ptmt7 (P.tabacina-2), cox2-Igscox-cox1 (from table 3.3) (P.tabacina-3), Ptmt17 P.tabacina-4), Ptmt22 (P.tabacina-5), Ptmt31 (P.tabacina-6), Ptmt12 (P.tabacina-7), and Ptmt34 (P.tabacina-8) primer pairs.

**Table 3. 3** Primers used for amplification and sequence of *P. tabacina* mitochondrial genome.

| Target DNA | Primers | Sequence (5' - 3') | $T_m$ (°C) | Program [a] | Reference |
|---|---|---|---|---|---|
| **MITOCHONDRIAL PRIMERS** | | | | | |
| Mitochondrial genome region between gene Cox2 and gene Cox1 | IgCoxF | AAAAGAGARGGTGTTTTTTAYGGA | 50 | SCO | Schena and Cooke 2007 |
| | IgCoxR | GCAAAAGCACTAAAAATTAAATATAA | | SCO | Schena and Cooke 2007 |
| Mitochondrial genome region that covers cytochrome c oxidase subunit 2 | FM78 | ACAAATTTCACTACATTGTCC | 50 | JBR | Martin and Tooley 2003 |
| | FM82 | TTGGCAATTAGGTTTTCAAGATCC | | JBR | Martin and Tooley 2003 |
| Mitochondrial genome region that covers *rpl14, rpl5 and tRNAs* | F3 | ATGGTAGAGCGTGGGAATCAT | 62 | JBR | Griffith and Shaw 1998 |
| | R3 | AATACCGCCTTTGGGTCCATT | | JBR | Griffith and Shaw 1998 |
| Mitochondrial genome region that covers part of *trnR, nad4L* and part of *nad1* | F6 | GAGTATTAGCCTTCTAAGC | 50 | JBR | Gavino and Fry 2002 |
| | R6 | CTGTTTTACCGAGATATCG | | JBR | Gavino and Fry 2002 |

[a] PCR programs: SCO, an initial denaturation at 96ºC for 2 min, followed by 35 cycles consisting of denaturation at 96ºC for 1 min, annealing at the specific primers temperature for 1 min, and extension at 72ºC for 2 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min (Schena and Cooke 2006).
JBR, an initial denaturation at 95ºC for 2 min, followed by 35 cycles consisting of denaturation at 94ºC for 30 s, annealing at the specific primers temperature for 30 s, and extension at 72ºC for 1 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min (Ristaino *et a*l. 2007).

**Table 3. 4** Size of coding regions and total of base pairs when mitochondrial genomes from *P. infestans, Hyaloperonospora parasitica* and *Peronospora tabacina* are compared.

| Coding regions | *Phytophthora infestans* Ib [a] | | | *Hyaloperonospora parasitica* [b] | | | *Peronospora tabacina* [c] | | |
|---|---|---|---|---|---|---|---|---|---|
| | start | end | bp | start | end | bp | start | end | bp |
| **rnl** | 26 | 2679 | 2653 | 26 | 2671 | 2645 | 26 [d] | 2659 | 2633 |
| **trnN** | 2686 | 2757 | 71 | 2686 | 2757 | 71 | 2686 | 2757 | 71 |
| **trnS1** | 2770 | 2858 | 88 | 2770 | 2855 | 85 | 2770 | 2853 | 83 |
| **trnM1** | 2879 | 2950 | 71 | 2879 | 2950 [d] | 71 | 2879 | 2950 [d] | 71 |
| **trnP** | 2994 | 3068 | 74 | 2994 | 3068 | 74 | 2994 | 3068 | 74 |
| **trnM2** | 3082 | 3153 | 71 | 3082 | 3154 | 72 | 3082 | 3153 | 71 |
| **rpl14** | 3171 | 3542 | 371 | 3171 | 3542 | 371 | 3171 | 3542 [d] | 371 |
| **rpl5** | 3549 | 4082 | 533 | 3549 | 4081 | 532 | 3549 | 4082 | 533 |
| **trnG1** | 4090 | 4161 | 71 | 4090 | 4161 | 71 | 4090 | 4161 [d] | 71 |
| **trnG2** | 4253 | 4324 | 71 | 4253 | 4324 | 71 | | | |
| **trnY** | 4330 | 4413 | 83 | 4330 | 4413 | 83 | | | |
| **rns** | 5199 | 6701 | 1502 | 5199 | 6701 [d] | 1502 | 5199 [d] | 6701 | 1502 |
| **trnW** | 6736 | 6807 | 71 | 6736 [d] | 6807 [d] | 71 | 6736 | 6807 [d] | 71 |
| **orf79** | 7002 | 7241 | 239 | 7002 | 7241 | 239 | | | |
| **cox2** | 7661 | 8437 | 776 | 7661 | 8437 | 776 | 7661 | 8436 | 775 |
| **orf32** | 8453 | 8551 | 98 | 8453 | 8551 | 98 | 8453 | 8552 | 99 |
| **cox1** | 8639 | 10117 | 1478 | 8639 | 10117 | 1478 | | | |
| **atp9** | 10266 | 10493 | 227 | 10266 | 10493 | 227 | | | |
| **nad9** | 10655 | 11221 | 566 | 10655 | 11221 | 566 | | | |
| **cob** | 11282 | 12433 | 1151 | 11282 | 12433 | 1151 | | | |
| **atp1** | 12572 | 14101 | 1529 | 12572 | 14101 | 1529 | | | |
| **trnE** | 14182 | 14253 | 71 | 14182 | 14253 | 71 | | | |
| **nad4** | 14283 | 15758 | 1475 | 14283 | 15758 | 1475 | 14283 | 15758 [d] | 1475 |
| **trnH** | 15781 | 15853 | 72 | 15781 | 15853 | 72 | 15781 | 15853 | 72 |
| **orf142** | 15857 | 16285 | 428 | 15857 | 16285 | 428 | 15857 | 16285 [d] | 428 |
| **nad7** | 16295 | 17473 | 1178 | 16295 | 17473 | 1178 | | | |
| **nad2** | 17591 | 19084 | 1493 | 17591 | 19084 | 1493 | | | |
| **trnF** | 19091 | 19164 | 73 | 19091 | 19164 | 73 | | | |
| **rps10** | 19183 | 19509 | 326 | 19183 | 19509 | 326 | | | |
| **trnR** | 19514 | 19587 | 73 | 19514 | 19586 | 72 | | | |
| **trnQ** | 19598 | 19669 | 71 | 19598 | 19669 | 71 | | | |
| **trnI** | 19671 | 19744 | 73 | 19671 | 19744 | 73 | | | |
| **trnV** | 19747 | 19819 | 72 | 19747 | 19820 | 73 | | | |
| **rps12** | 19836 | 20216 | 380 | 19836 | 20219 | 383 | 19836 [d] | 20219 | 383 |
| **rps7** | 20146 | 20619 | 473 | 20146 | 20622 | 476 | 20146 | 20619 | 473 |
| **cox3** | 20675 | 21592 | 917 | 20675 | 21593 | 918 | 20675 | 21592 [d] | 917 |
| **atp6** | 21620 | 22339 | 719 | 21620 | 22339 | 719 | | | |
| **trnD** | 22370 | 22443 | 73 | 22370 | 22443 | 73 | | | |
| **nad3** | 22480 | 22833 | 353 | 22480 | 22833 | 353 | | | |
| **nad5** | 23053 | 25047 | 1994 | 23053 | 25047 | 1994 | | | |
| **nad6** | 25091 | 25801 | 710 | 25091 | 25801 | 710 | | | |
| **trnR** | 25827 | 25899 | 72 | 25827 | 25899 | 72 | | | |
| **nad4L** | 26065 | 26367 | 302 | 26065 | 26367 | 302 | | | |
| **nad1** | 26370 | 27350 | 980 | 26370 | 27350 | 980 | | | |
| **nad11** | 27347 | 29353 | 2006 | 27347 | 29353 | 2006 | 27347 | 29353 [d] | 2006 |
| **trnL1** | 29438 | 29520 | 82 | 29438 | 29521 | 83 | | | |
| **trnL2** | 29530 | 29613 | 83 | 29530 | 29614 | 84 | | | |
| **ymf16** | 29633 | 30379 | 746 | 29633 | 30376 | 743 | | | |
| **trnC** | 30570 | 30640 | 70 | 30570 | 30640 | 70 | | | |

Table 3.4 continued

| Coding regions | *Phytophthora infestans* Ib [a] | | | *Hyaloperonospora parasitica* [b] | | | *Peronospora tabacina* [c] | | |
|---|---|---|---|---|---|---|---|---|---|
| | start | end | bp | start | end | bp | start | end | bp |
| **trnS2** | 30652 | 30736 | 84 | 30652 | 30736 | 84 | | | |
| **rps11** | 30758 | 31174 | 416 | 30758 | 31177 | 419 | | | |
| **rps13** | 31187 | 31600 | 413 | 31187 | 31606 | 419 | | | |
| **rpl2** | 31629 | 32435 | 806 | 31629 | 32435 | 806 | | | |
| **rps19** | 32439 | 32672 | 233 | 32439 | 32672 | 233 | | | |
| **rps3** | 32676 | 33479 | 803 | 32676 | 33479 | 803 | | | |
| **rpl16** | 33482 | 33886 | 404 | 33482 | 33886 | 404 | | | |
| **trnI2** | 33889 | 33962 | 73 | 33889 | 33962 | 73 | | | |
| **orf217** | 33983 | 34636 | 653 | 33983 | 34636 | 653 | | | |
| **atp8** | 34698 | 35090 | 392 | 34698 | 35090 | 392 | | | |
| **trnK** | 35111 | 35183 | 72 | 35111 | 35183 | 72 | | | |
| **trnA** | 35187 | 35259 | 72 | 35187 | 35259 | 72 | | | |
| **rps14** | 35278 | 35577 | 299 | 35278 | 35577 | 299 | | | |
| **rps8** | 35591 | 35971 | 380 | 35591 | 35971 | 380 | | | |
| **rpl6** | 35979 | 36584 | 605 | 35979 | 36584 | 605 | | | |
| **rps2** | 36591 | 37181 | 590 | 36591 | 37178 | 587 | | | |
| **rps4** | 37191 | 37652 | 461 | 37191 | 37642 | 451 | | | |
| **orf100** | 37655 | 37957 | 302 | 37655 | 37934 | 279 | | | |
| **Total** | | | 34217 | | | 34185 | | | 12179 |

[a] GenBank accession number NC002387

[b] Asseblies V1.0, V3.0, V6.0 and V7.0 from Virginia Tech database

[c] Isolate Penn07

[d] These correspond to start or end points where the entire region was not available. The size of the gene was hypothetically estimated.

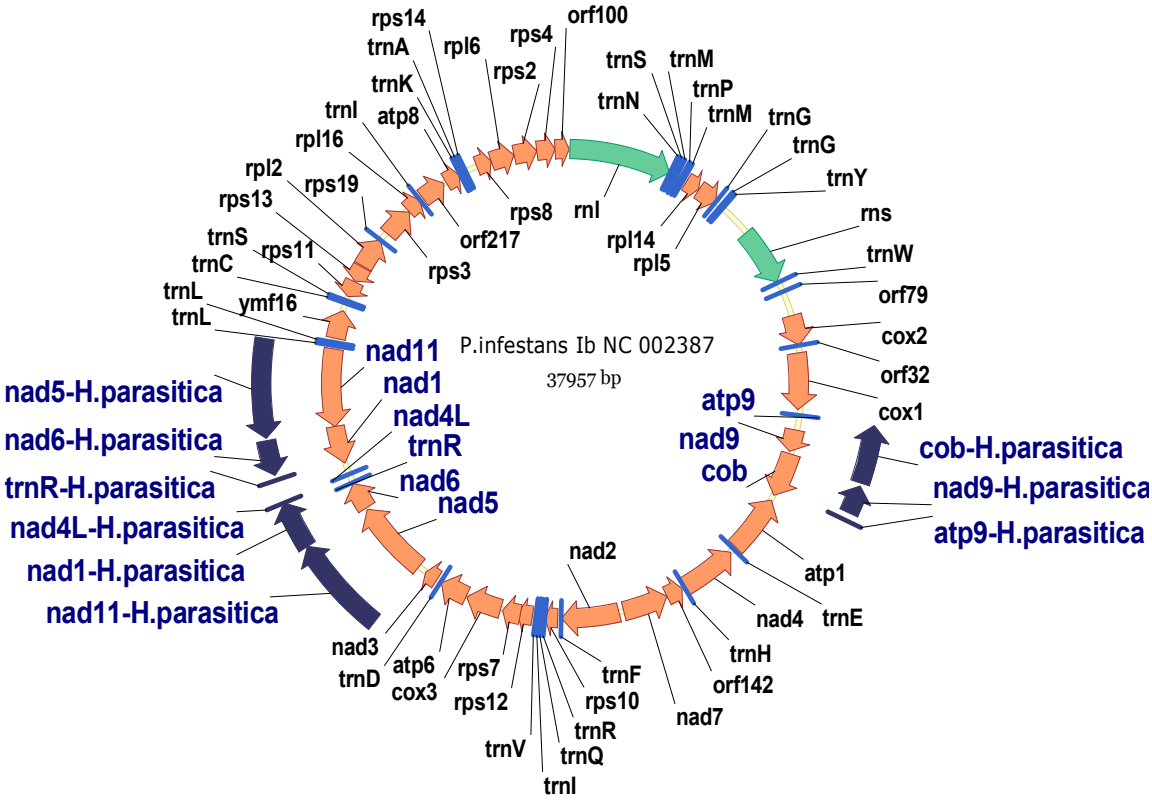



**Figure 3. 1** Physical arrangement of the mitochondrial genes of *Hyaloperosnopora parasitica* on *Phytophthora infestans* mitochondrial Ib genome as a reference. The arrangement of genes is identical for the most of the genome with the exception of the bold arrows outside of the figure. These regions had different direction or were flipped when compared with the *P. infestans* mitochondrial genome. Refer to Table 3.1 and Table 3.4 for gene information.

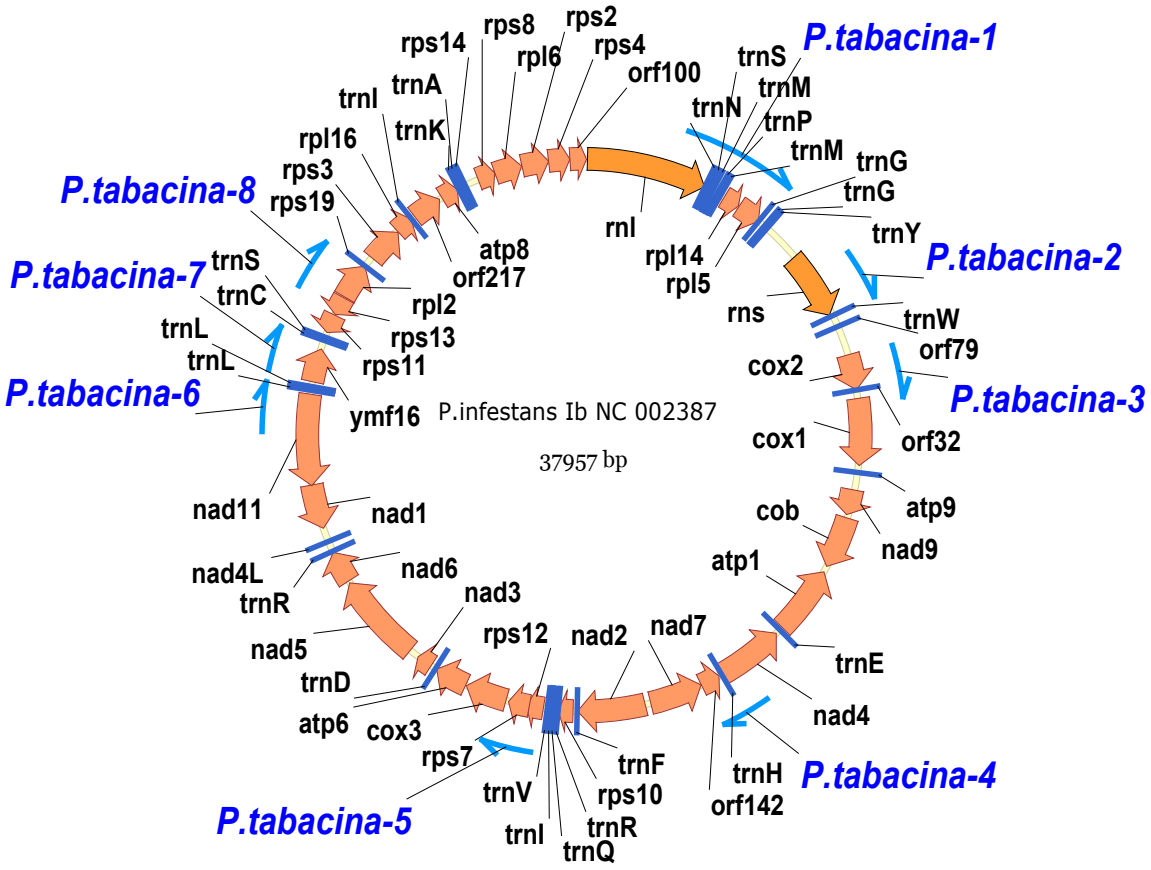
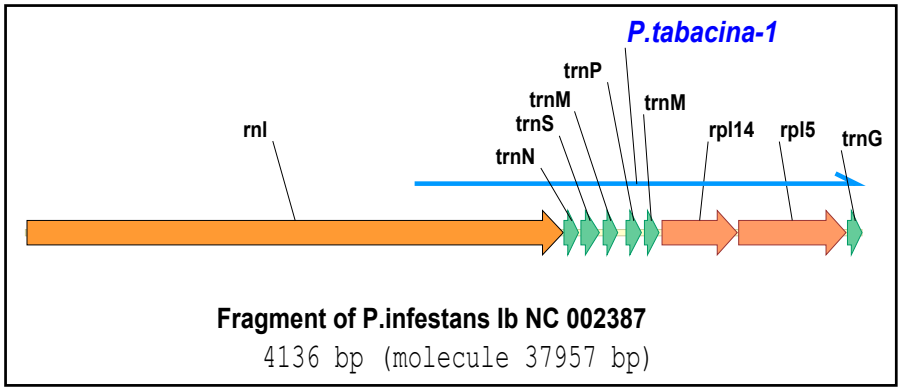


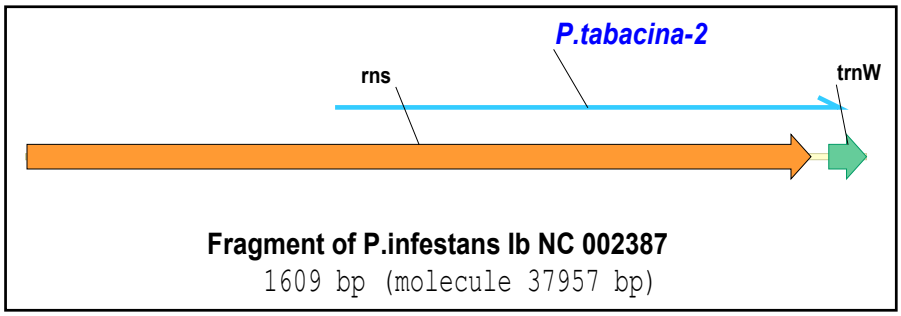

**Figure 3. 2** Schematic diagram of the mitochondrial genome of *P. infestans* mitochondrial Ib genome indicating putative regions of the *P. tabacina* mitochondrial genome that have been sequenced. Genes are indicated by lines and filled boxes. The direction of transcription is indicated by arrows. *P. tabacina* fragments are indicated in the outer circle by blue arrows.

**Figure 3. 3** Linear arrangement of the regions of the *P. tabacina* mitochondrial genome that have been sequenced using the *P. infestans* mitochondrial Ib genome as a reference. Refer to Table 3.4 for genes information.



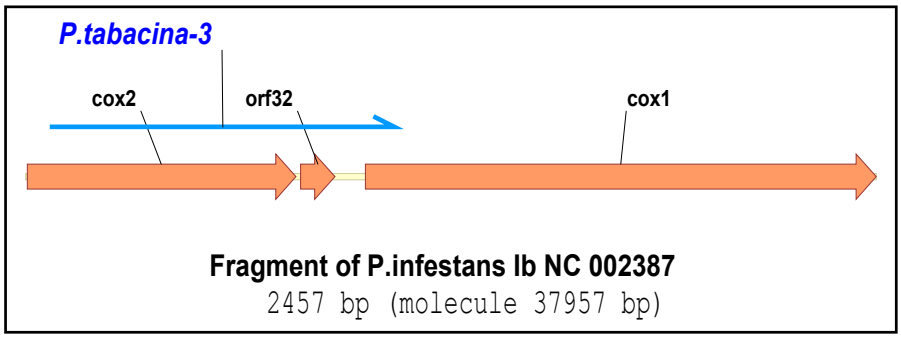

**Fragment of P.infestans Ib NC 002387**

4136 bp (molecule 37957 bp)



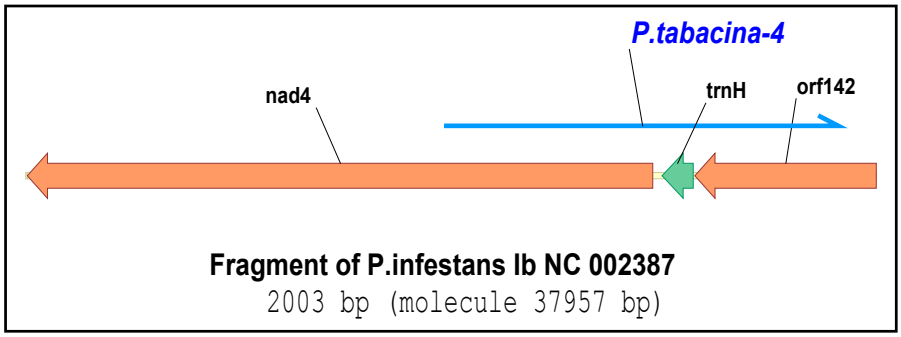

**Fragment of P.infestans Ib NC 002387**

1609 bp (molecule 37957 bp)



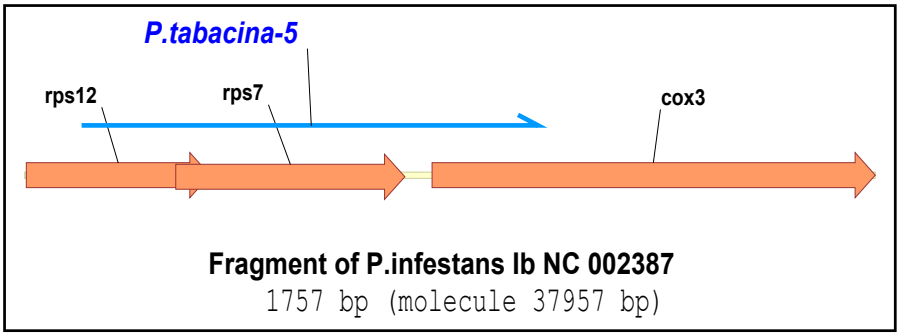

**Fragment of P.infestans Ib NC 002387**

2457 bp (molecule 37957 bp)



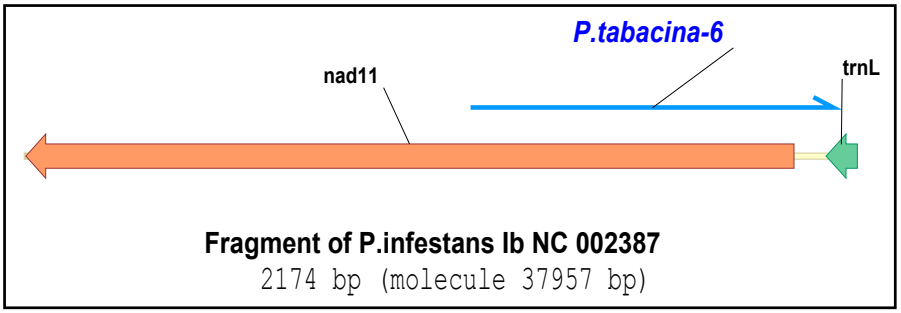

**Fragment of P.infestans Ib NC 002387**

2003 bp (molecule 37957 bp)



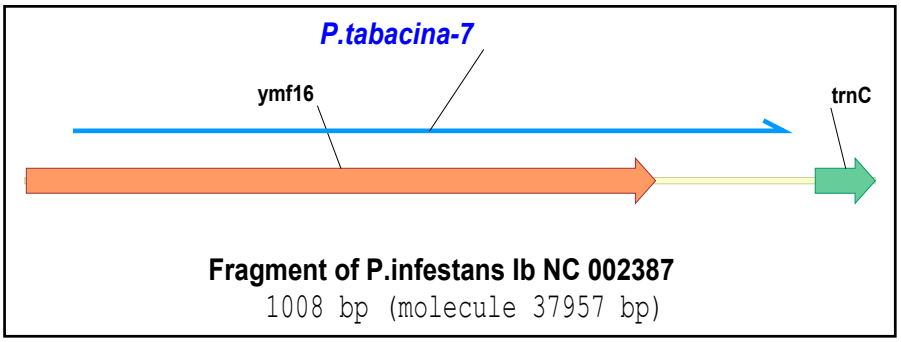

**Fragment of P.infestans Ib NC 002387**

1757 bp (molecule 37957 bp)



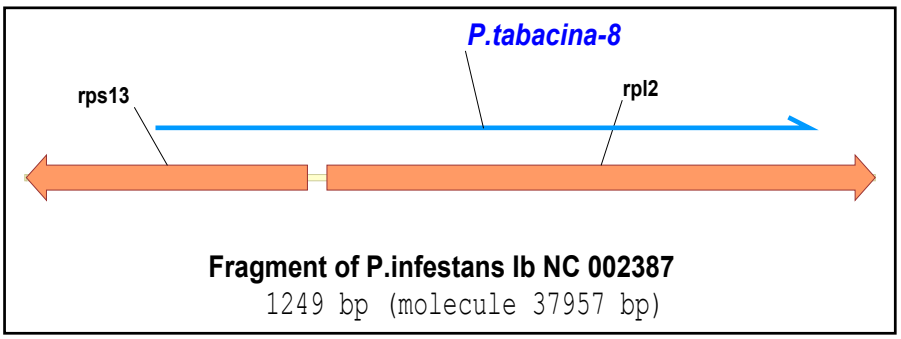

**Fragment of P.infestans Ib NC 002387**

2174 bp (molecule 37957 bp)




**Fragment of P.infestans Ib NC 002387**

1008 bp (molecule 37957 bp)

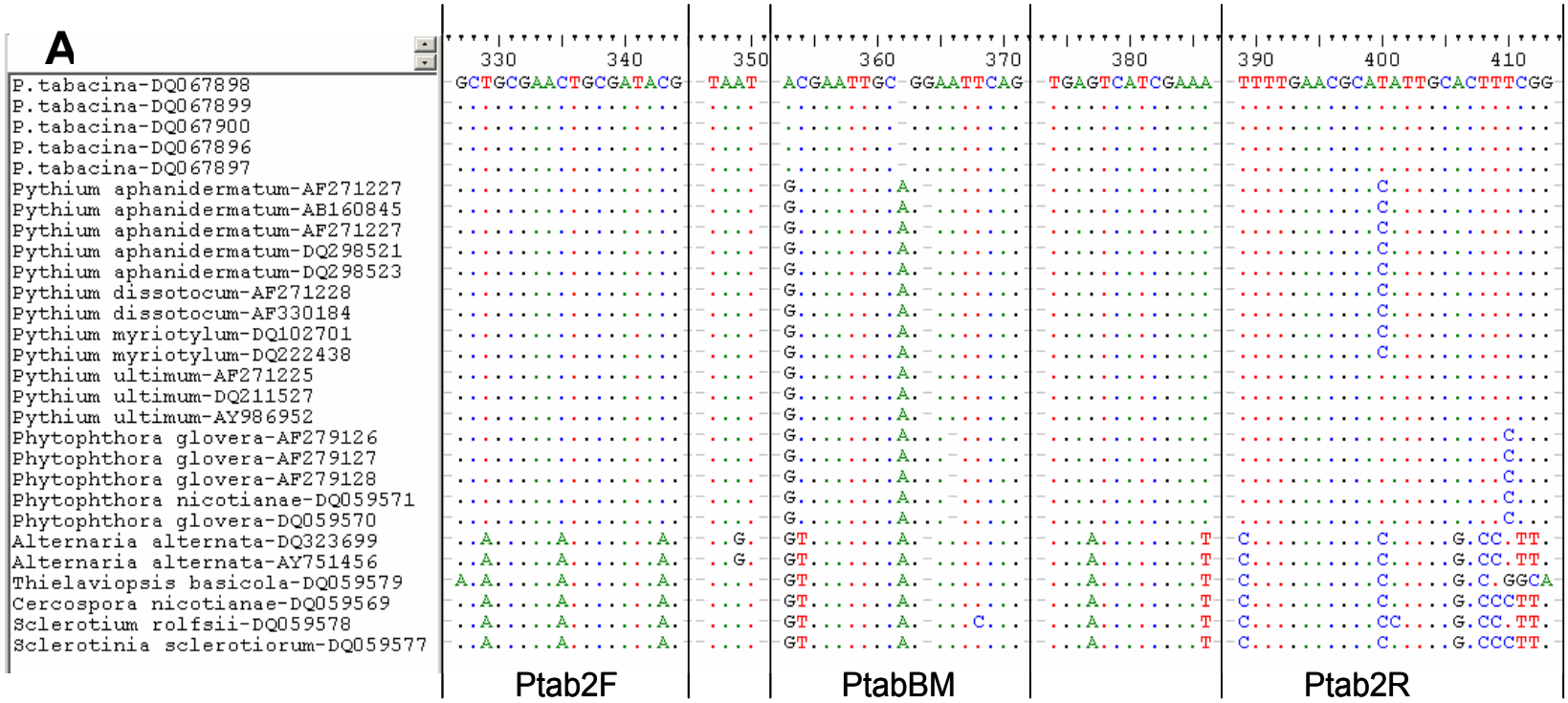



**Fragment of P.infestans Ib NC 002387**

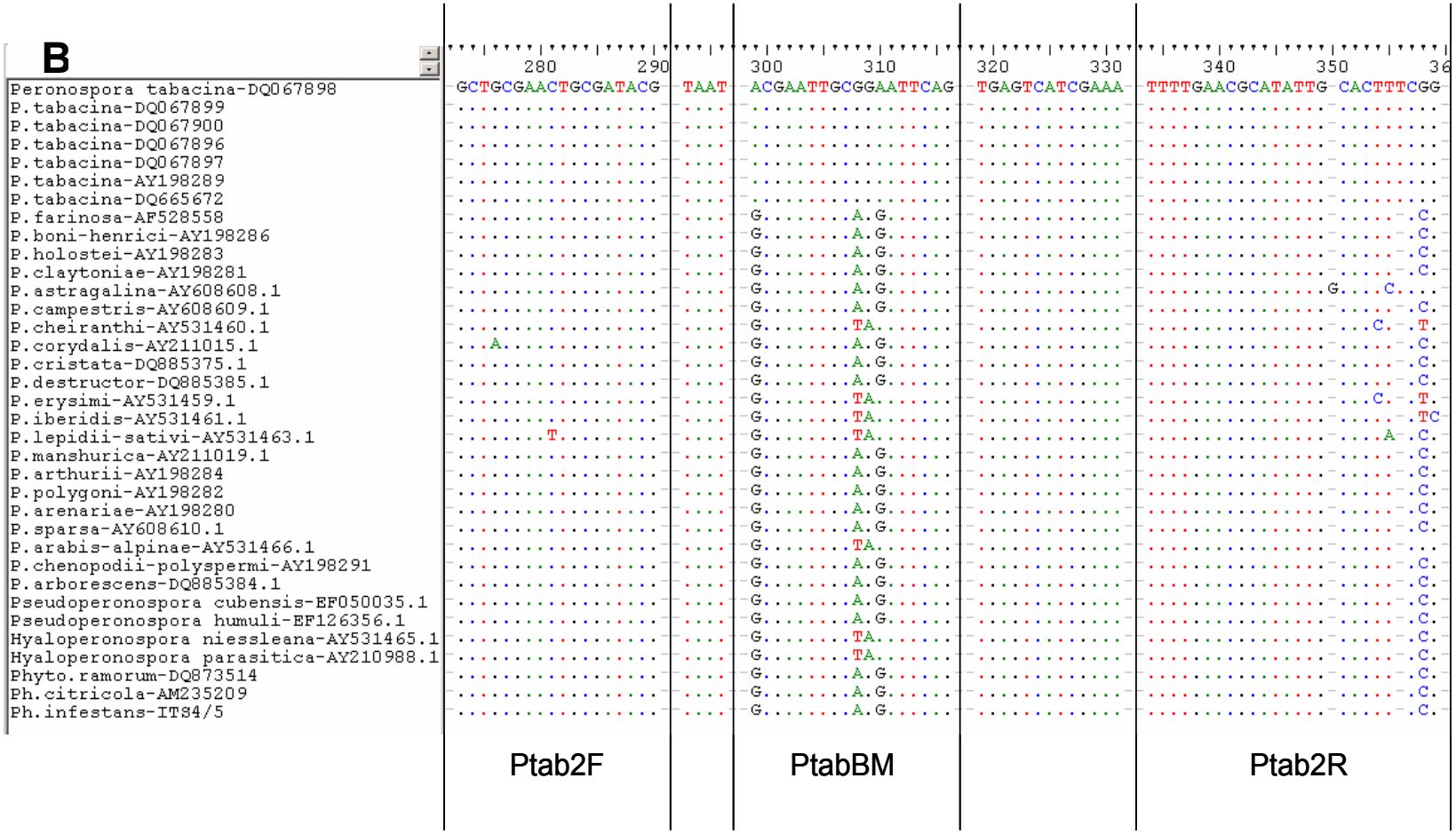
1249 bp (molecule 37957 bp)

# APPENDICES

**Appendix 1. 1** Comparative aligments of A) *P. tabacina* and other tobacco pathogens and B) *P. tabacina* and other Oomycetes species; for the design of the Ptab2F and Ptab2R and probe for a real-time detection assay.



Appendix 1.1 continued



**Appendix 2. 1** Selected primers tested for amplification and sequencing of *P. tabacina* pathogen.

| Target DNA | Primers | Sequence (5' - 3') | $T_m$ (°C) recommended or used | Program [c] | Reference |
|---|---|---|---|---|---|
| **MITOCHONDRIAL PRIMERS** | | | | | |
| Mitochondrial genome region between gene *trnG* (gcc) and gene *trnY* (gua) | Mt13F | ACAGTTTTTCGAATTAAAAACAGAA | 50 | SCO | Schena and Cooke 2007 |
| | Mt15F | TTGCCAAGGTTAATGTTGAGG | | SCO | Schena and Cooke 2007 |
| | Mt3R | GGAGAAAGTAGGATTCGAACCT | | SCO | Schena and Cooke 2007 |
| Mitochondrial genome region between gene trnY (gua) and gene Rns | Mt1F [a] | TGGCTGAGTGGTTAAAGGTG | 50 [b] | SCO | Schena and Cooke 2007 |
| | Mt2F | TGGCAGACTGTAAATTTGTTGAA | | SCO | Schena and Cooke 2007 |
| | Mt5R | TTGCATGTGTTAAGCATACCG | | SCO | Schena and Cooke 2007 |
| | Mt6R [a] | CTCACCCGTTCGCTATGTTT | | SCO | Schena and Cooke 2007 |
| Mitochondrial genome region between gene Cox2 and gene Cox1 | IgCoxF | AAAAGAGARGGTGTTTTTTAYGGA | 50 | SCO | Schena and Cooke 2007 |
| | IgCoxR | GCAAAAGCACTAAAAATTAAATATAA | | SCO | Schena and Cooke 2007 |
| Mitochondrial genome region between gene Atp9 and gene Nad9 | ATPF | TTTATTCTGTTTAATGATGGC | 50 | SCO | Schena and Cooke 2007 |
| | ATPR | CAGCACAAATTTCAGATAATAC | | SCO | Schena and Cooke 2007 |
| Mitochondrial genome region that covers cytochrome c oxidase subunit 1 | COXF4N [a] | GTATTTCTTCTTTATTAGGTGC | 52 [b] | JBR | Kroon et al. 2004 |
| | COXR4N [a] | CGTGAACTAATGTTACATATAC | | JBR | Kroon et al. 2004 |
| Mitochondrial genome region that covers cytochrome c oxidase subunit 2 | COX2F [a] | GGCAAATGGGTTTTCAAGATCC | 50 [b] | JBR | Hudspeth et al. 2000 |
| | COX2R [a] | CCATGATTAATACCACAAATTTCACTAC | | JBR | Hudspeth et al. 2000 |
| Mitochondrial genome region that covers cytochrome c oxidase subunit 2 | FM78 [a] | ACAAATTTCACTACATTGTCC | 50 [b] | JBR | Martin and Tooley 2003 |
| | FM82 [a] | TTGGCAATTAGGTTTTCAAGATCC | | JBR | Martin and Tooley 2003 |
| Mitochondrial genome region that covers part of *nad5* and *nad6* | F1 | GCAATGGGTAAATCGGCTCAA | 56 | JBR | Griffith and Shaw 1998 |
| | R1 | AAACCATAAGGACCACACAT | | JBR | Griffith and Shaw 1998 |
| Mitochondrial genome region that covers part of *atp1, trnE* and part of *nad4* | F2 | TTCCCTTTGTCCTCTACCGAT | 62 | JBR | Griffith and Shaw 1998 |
| | R2 | TTACGGCGGTTTAGCACATACA | | JBR | Griffith and Shaw 1998 |
| Mitochondrial genome region that covers *rpl14, rpl5 and tRNAs* | F3 [a] | ATGGTAGAGCGTGGGAATCAT | 62 [b] | JBR | Griffith and Shaw 1998 |
| | R3 [a] | AATACCGCCTTTGGGTCCATT | | JBR | Griffith and Shaw 1998 |
| Mitochondrial genome region that covers cytochrome c oxidase subunit 1 | F4 | TGGTCATCCAGAGGTTTATGTT | 62 | JBR | Griffith and Shaw 1998 |
| | R4 | CCGATACCGATACCAGCACCAA | | JBR | Griffith and Shaw 1998 |
| Mitochondrial genome region that covers *rpl14, rpl5 and tRNAs* | F5 | CTGTTTTACCGAGATATCG | 60 | JBR | Gavino and Fry 2002 |
| | R5 | GAGTATTAGCCTTCTAAGC | | JBR | Gavino and Fry 2002 |
| Mitochondrial genome region that covers part of *trnR, nad4L* and part of *nad1* | F6 [a] | GAGTATTAGCCTTCTAAGC | 50 [b] | JBR | Gavino and Fry 2002 |
| | R6 [a] | CTGTTTTACCGAGATATCG | | JBR | Gavino and Fry 2002 |

## Appendix 2.1 continued

| Target DNA | Primers | Sequence (5' - 3') | $T_m$ (°C) recommended or used | Program [c] | Reference |
|---|---|---|---|---|---|
| **NUCLEAR PRIMERS** | | | | | |
| Large ribosomal subunit (LSU rDNA) | LR7R-0 [a] | GAAGCTCGTGGCGTGAG | 55 [b] | JBR | Goker et al. 2003 |
| | LR3R [a] | GTCTTGAAACACGGACC | | JBR | Goker et al. 2003 |
| Large ribosomal subunit (LSU rDNA) | LR9 [a] | AGAGCACTGGGCAGAAA | 52 [b] | JBR | Goker et al. 2003 |
| | LR16-0 [a] | TTGCACGTCAGAATCG | | JBR | Goker et al. 2003 |
| Internal transcribed spacer (ITS) | ITS4 [a] | TCCTCCGCTTATTGATATGC | 56 [b] | JBR | Wiglesworth et al. 1994 |
| | ITS5 [a] | GGAAGTAAAAGTCGTAACAAGG | | JBR | Wiglesworth et al. 1994 |
| Intergenic spacer (IGS) region of the rDNA | Igs1F [a] | AAAGTGRKGMGGWGWWGCKGA | 55 [b] | SCO | Schena and Cooke 2007 |
| | Igs2F [a] | AAGTRYMTKAACAACGCTCT | | SCO | Schena and Cooke 2007 |
| | Igs3F [a] | GYGCGAAGGWKTGCTG | | SCO | Schena and Cooke 2007 |
| | Igs6R [a] | CCCAGCRYAAACAACAACAC | | SCO | Schena and Cooke 2007 |
| | Igs7R [a] | ATATCCTCCATACGWAAGAAGACG | | SCO | Schena and Cooke 2007 |
| Ptab | Ptab [a] | ATCTTTTTGCTGGCTGGCTA | 56 [b] | JBR | Ristaino et al. 2007 |
| **RAS-RELATED PROTEIN (Ypt1) GENE** | | | | | |
| Ras-related protein (Ypt1) gene | Ypt1F [a] | CGACCATYGGYGTKGACTTT | 56 [b] | SCO | Schena and Cooke 2007 |
| | Ypt4R [a] | TTSACGTTCTCRCAGGCGTA | | SCO | Schena and Cooke 2007 |
| | Ypt5R [a] | GCAGCTTGTTSACGTTCTCR | | SCO | Schena and Cooke 2007 |
| Intron Ras | IRF | TTGCAGCACAACCCAAGACG | 56 | JBR | Gomez et al. 2007 |
| | IRR | TGCACGTACTATTCGGGGTTC | | JBR | Gomez et al. 2007 |
| Ras | RASF | CGTGTCTGCTTCTCCGTTTCG | 56 | JBR | Gomez et al. 2007 |
| | RASR | CCAGGCTTTCGGCAAATTCC | | JBR | Gomez et al. 2007 |
| Ras-related protein (Ypt1) gene that covers exon 6 (RASexon6) | F | GAGAACGTGAACAAGCTGCTGG | 56 | JBR | This study |
| | R | GCAAAGAAAGCGTCAGCCTG | | JBR | This study |
| Ras-related protein (Ypt1) gene that covers exon 3, 4 and 5 | F | GGAGAGCTACATCTCGACCATCG | 55 | JBR | This study |
| | R | TGTTACCAACCAGCAGCTTG | | JBR | This study |
| Ras-related protein (Ypt1) gene that covers exon 1 and 2 | F | TCGGTTTAGTTATCGCATGCTG | 56 | JBR | This study |
| | R | GGTCGAGATGTAGCTCTCCGTG | | JBR | This study |
| Ras-related protein (Ypt1) gene that covers intron 3 and 4 | F [a] | GGAGAGCTACATCTCGACCATCG | 56 [b] | JBR | This study |
| | R [a] | AACGACTCCTGGTCCGTCAC | | JBR | This study |

## Appendix 2.1 continued

| Target DNA | Primers | Sequence (5' - 3') | $T_m$ (°C) recommended or used | Program [c] | Reference |
|---|---|---|---|---|---|
| **TRANSLATION ELONGATION FACTOR 1 ALPHA GENE** | | | | | |
| Tranlation elongation factor 1α | ELONGF1 | TCACGATCGACATTGCCCTG | 60 | JBR | Kroon et al. 2004 |
| | ELONGR1 | ACGGCTCGAGGATGACCATG | | JBR | Kroon et al. 2004 |
| **BETA-TUBULIN GENE** | | | | | |
| β-tubulin | TUBUF2 [a] | CGGTAACAACTGGGCCAAGG | 60 [b] | JBR | Kroon et al. 2004 |
| | TUBUR1 [a] | CCTGGTACTGCTGGTACTCAG | | JBR | Kroon et al. 2004 |
| β-tubulin | TUBU901 [a] | TACGACATTTGCTTCCG | 56 [b] | JBR | Gomez et al. 2007 |
| | TUBU1401 [a] | CGCTTGAACATCTCCTGG | | JBR | Gomez et al. 2007 |
| β-tubulin | TUBU945 [a] | GCACACCAGGTGGTTC | 56 [b] | JBR | Gomez et al. 2007 |
| | TUBU415 [a] | CGCATCAACGTGTACTACAA | | JBR | Gomez et al. 2007 |
| **NADH DEHYDROGENASE SUBUNIT 1 GENE** | | | | | |
| NADH dehydrogenase subunit 1 | NADHF1 [a] | CTGTGGCTTATTTTACTTTAG | 53 [b] | JBR | Kroon et al. 2004 |
| | NADHR1 [a] | CAGCAGTATACAAAAACCAAC | | JBR | Kroon et al. 2004 |

[a] Primers that were able to amplified *P. tabacina*. Fifteen isolates using each primer combination were sequenced and two were submitted to Gene Bank: *P. tabacina* - Mt 1/Mt 5, XXX; *P. tabacina* - COXF4N/COXR4N, XXX; *P. tabacina* - COX2F/ COX2R, XXX; *P. tabacina* - FM78/ FM82, XXX; *P. tabacina* - F3/ R3, XXX; *P. tabacina* - F6 /R6, XXX; *P. tabacina* - *LR7R-0/ LR9, XXX; P. tabacina* - LR3R/ LR16-0, XXX; *P. tabacina* - Igs3F/ Igs7R, XXX; *P. tabacina* - Ypt1F/ Ypt5R, XXX; *P. tabacina* - Ras3-4F/ Ras3-4R, XXX; - P. tabacina - TUBUF2/TUBUR1, XXX; P. tabacina TUBU901/ TUBU1401, XXX; *P. tabacina* - TUBU945/ TUBU415, XXX; *P. tabacina* - NADHF1/ NADHR1, XXX.

[b] Annealing temperature used for the amplification of *P. tabacina* pathogen

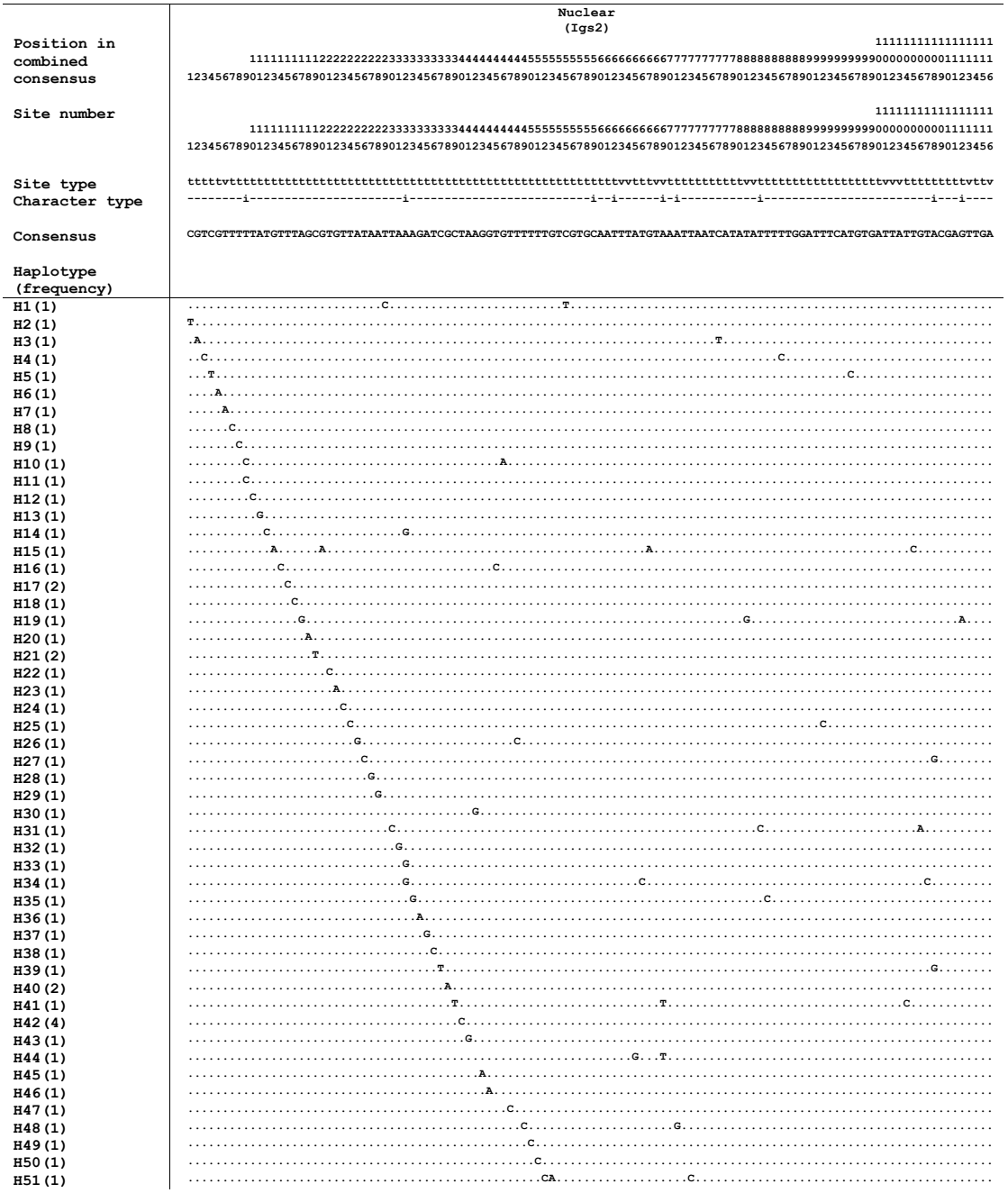
[c] PCR programs: SCO, an initial denaturation at 96ºC for 2 min, followed by 35 cycles consisting of denaturation at 96ºC for 1 min, annealing at the specific primers temperature for 1 min, and extension at 72ºC for 2 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min (Schena and Cooke 2006). JBR, an initial denaturation at 95ºC for 2 min, followed by 35 cycles consisting of denaturation at 94ºC for 30 s, annealing at the specific primers temperature for 30 s, and extension at 72ºC for 1 min. Thermal cycling was followed by a final extension step at 72ºC for 10 min (Ristaino *et a*l. 2007).

**Appendix 2. 2** Population statistics and diversity estimates for the *P. tabacina* U.S. population.

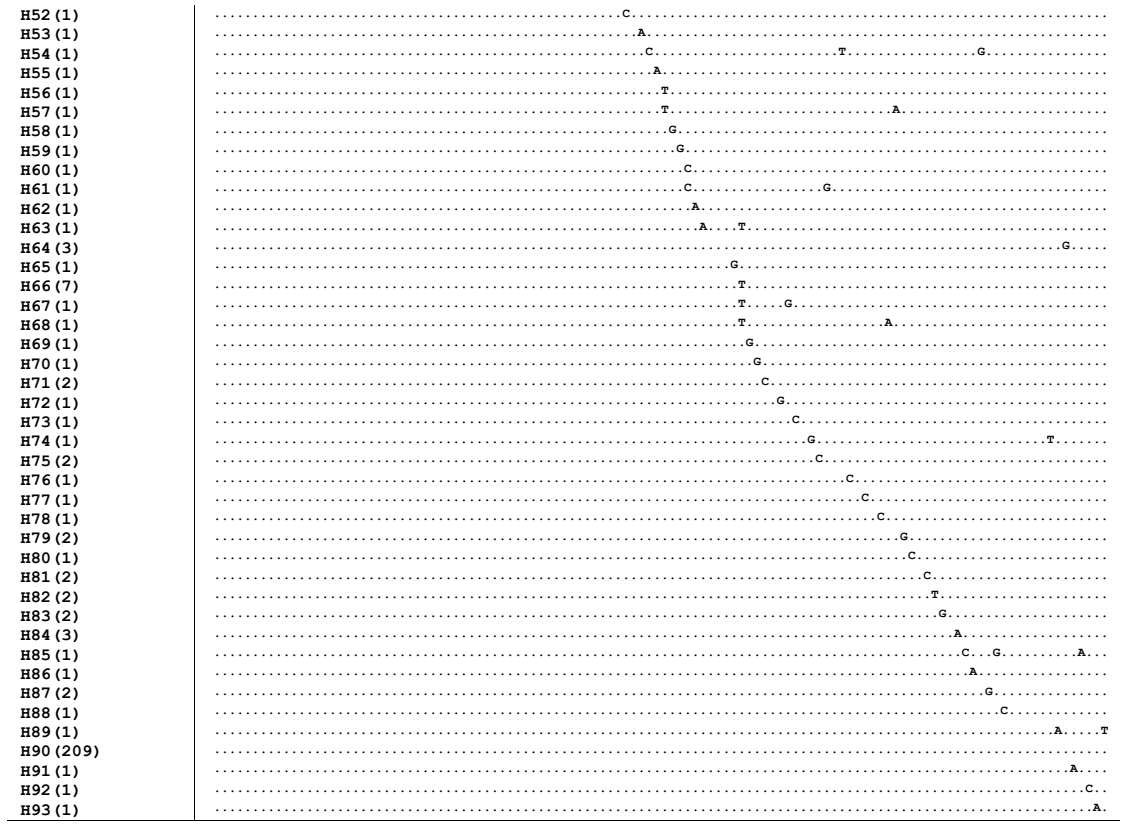
| | Sample summaries | | | | | Parameter estimates | |
|---|---|---|---|---|---|---|---|
| **Locus/Population** | ***l*** | ***n*** | ***s*** | ***h*** | ***k*** | $\pi$ **(SE)** $X10^{-3}$ | $\Theta_W X10^{-3}$ |
| ***Nuclear*** | | | | | | | |
| **(*Igs* 2 region)** | | | | | | | |
| Connecticut | 435 | 12 | 5 | 6 | 1.606 | 321.00 (3.99X10-2) | 331.14 |
| Florida | 435 | 6 | 7 | 6 | 2.533 | 5.82 (1.38X10-3) | 7.05 |
| Georgia | 435 | 12 | 6 | 7 | 1.848 | 4.25 (7.0X10-4) | 4.57 |
| Kentucky | 435 | 18 | 20 | 15 | 3.327 | 166.34 (2.1X10-2) | 290.74 |
| Maryland | 435 | 12 | 10 | 10 | 2.045 | 204.55 (3.4X10-2) | 331.14 |
| N. Carolina | 435 | 72 | 38 | 24 | 5.885 | 31.83 (5.68X10-3) | 206.32 |
| Pennsylvania | 435 | 18 | 2 | 3 | 0.614 | 307.19 (6.4X10-2) | 290.74 |
| Texas | 435 | 18 | 7 | 8 | 1.497 | 213.00 (3.2X10-2) | 290.74 |
| Virginia | 435 | 12 | 5 | 5 | 1.076 | 215.15 (5.9X10-2) | 331.14 |
| **(*Ypt* 1 region)** | | | | | | | |
| Connecticut | 493 | 12 | 16 | 10 | 3.667 | 229.17 (4.5X10-2) | 331.14 |
| Florida | 493 | 6 | 9 | 6 | 4.2 | 466.67 (1.0X10-1) | 437.96 |
| Georgia | 493 | 12 | 12 | 9 | 2.879 | 239.90 (4.4X10-2) | 331.14 |
| Kentucky | 493 | 12 | 8 | 4 | 3.258 | 407.20 (7.7X10-2) | 331.14 |
| Maryland | 493 | 12 | 11 | 7 | 4.273 | 388.43 (6.7X10-2) | 331.14 |
| N. Carolina | 493 | 72 | 47 | 31 | 1.995 | 42.44 (4.9x10-3) | 206.32 |
| Pennsylvania | 493 | 18 | 9 | 8 | 2.124 | 236.02 (4.4X10-2) | 290.74 |
| Texas | 493 | 18 | 12 | 9 | 2.34 | 194.99 (3.0X10-2) | 290.74 |
| Virginia | 493 | 12 | 14 | 6 | 3.545 | 7.21 (1.6X10-3) | 331.14 |
| ***Mitochondrial*** | | | | | | | |
| **(*cox* 2 gene)** | | | | | | | |
| Connecticut | 619 | 12 | 4 | 5 | 0.667 | 1.08 (3.7x10-4) | 2.15 |
| Florida | 619 | 6 | 0 | 1 | 0 | 0.00 | 0.00 |
| Georgia | 619 | 12 | 6 | 5 | 1 | 1.63 (6.1X10-4) | 3.25 |
| Kentucky | 619 | 18 | 9 | 6 | 1 | 1.63 (7.0x10-4) | 4.26 |
| Maryland | 619 | 12 | 3 | 4 | 0.5 | 0.81 (3.4X10-4) | 1.61 |
| N. Carolina | 625 | 72 | 20 | 17 | 0.556 | 0.92 (2.1X10-4) | 6.84 |
| Pennsylvania | 620 | 18 | 8 | 7 | 0.889 | 1.45 (4.7X10-4) | 3.81 |
| Texas | 621 | 18 | 13 | 8 | 1.542 | 2.52 (7.4X10-4) | 6.18 |
| Virginia | 618 | 12 | 2 | 2 | 0.333 | 0.54 (4.4X10-4) | 1.08 |

l, # of nucleotide sites; n, sample size (number of clones); s, segregating nucleotide sites; h, haplotypes; K, average number of pairwise nucleotide differences; $\pi$, average number of base differences per site; SE, standart error; Θw, population mean mutation rate per sequence or Watterson's theta estimator
ns, not significant; *, 0.01<P<0.05; **, 0.001<P<0.01; ***, P<0.001

**Appendix 2. 3** Distribution of haplotypes, base substitutions, and insertions/ deletions events in nuclear gene, *Igs*2 of *P. tabacina.*

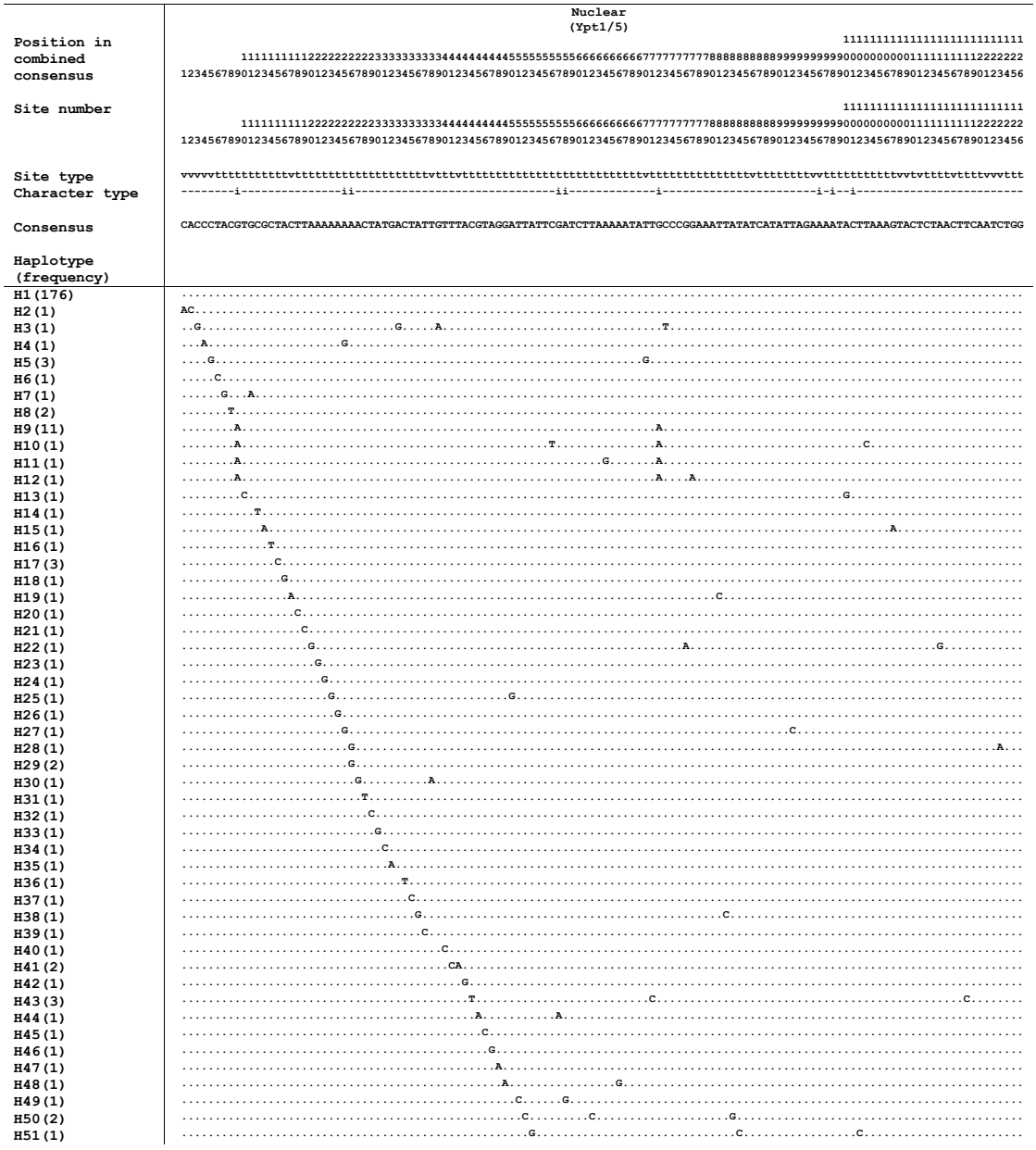


Appendix 2.3 continued

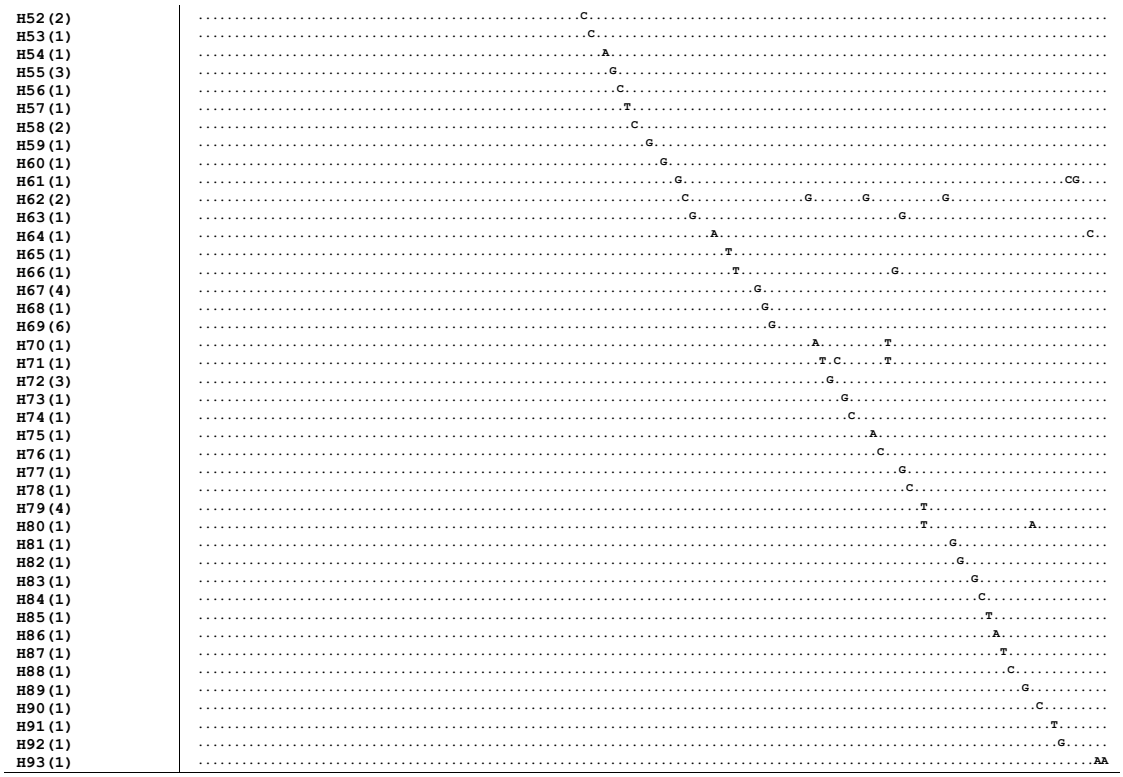



t, transitions; v, transversions; i, phylogenetically informative sites; -, uninformative sites.

**Appendix 2. 4** Distribution of haplotypes, base substitutions, and insertions/ deletions events in nuclear region, *Ypt*1 of *P. tabacina*.

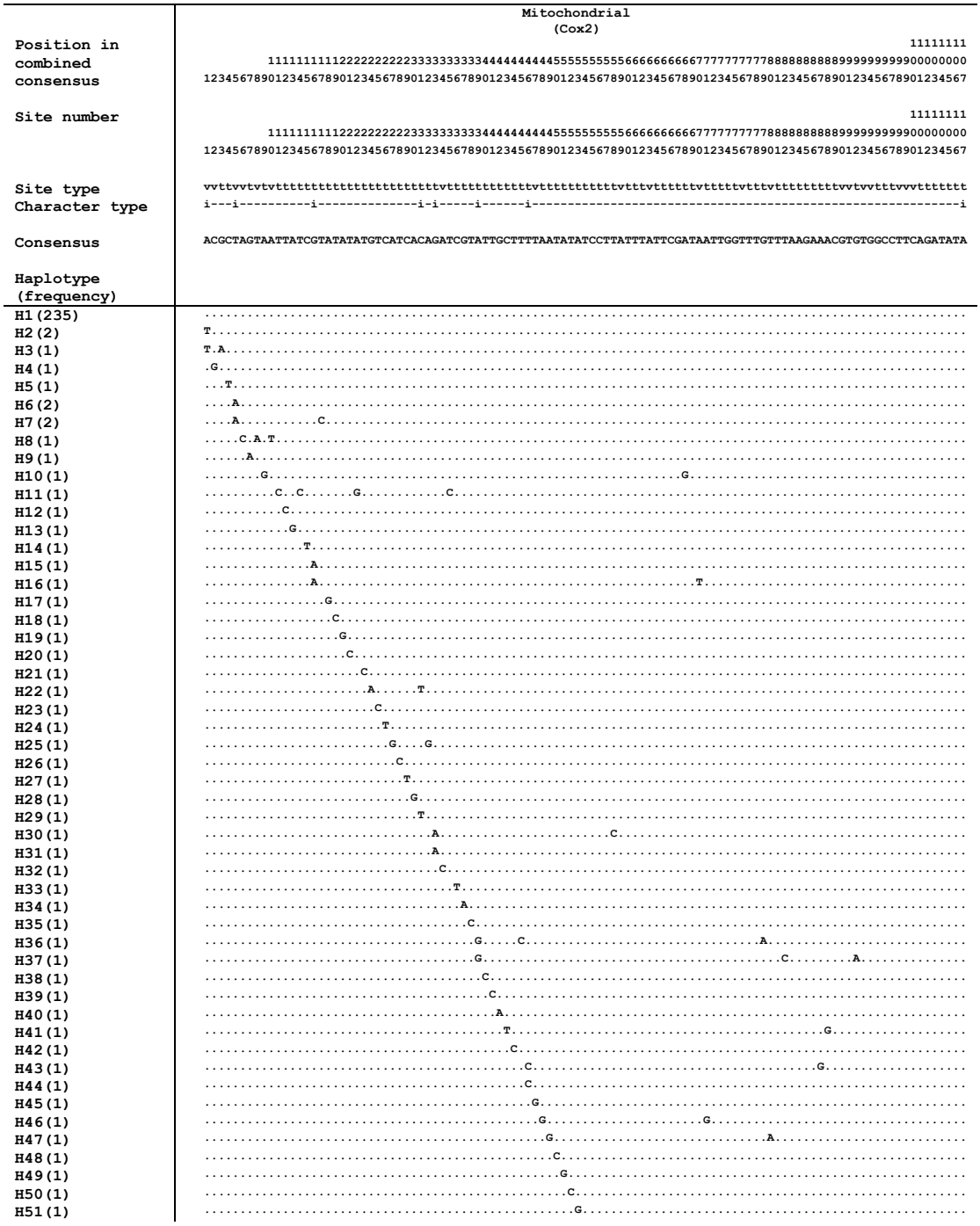


Appendix 2.4 continued.

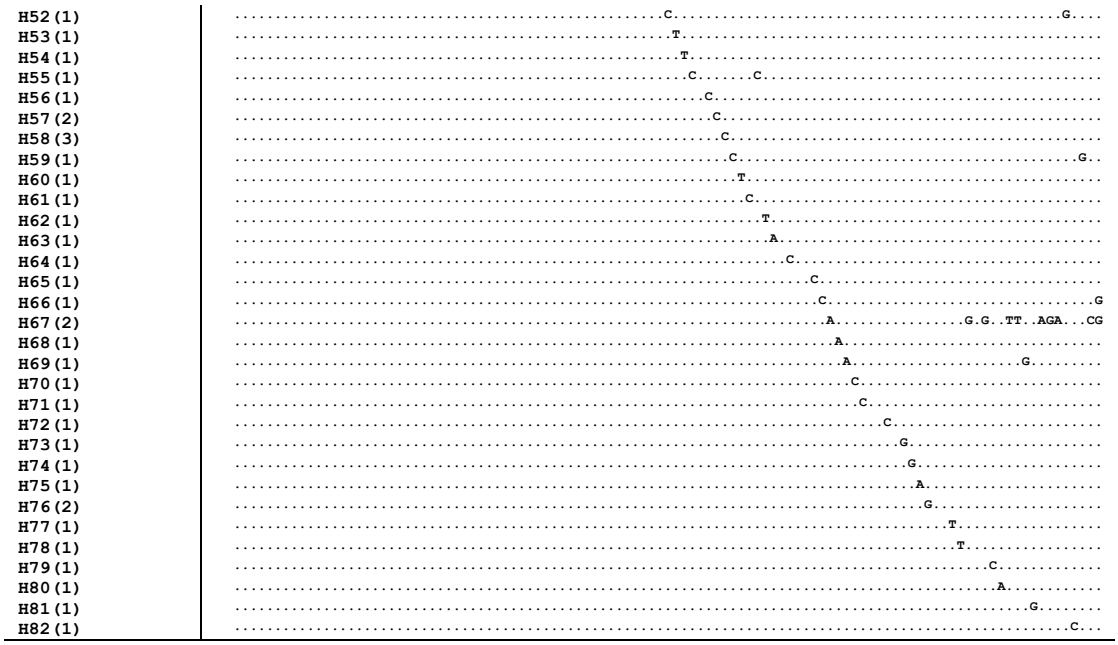



t, transitions; v, transversions; i, phylogenetically informative sites; -, uninformative sites.

**Appendix 2. 5** Distribution of haplotypes, base substitutions, and insertions/ deletions events in mitochondrial gene, *cox*2 of *P. tabacina*.



Appendix 2.5 continued




t, transitions; v, transversions; i, phylogenetically informative sites; -, uninformative sites.

**Appendix 2. 6** Frequency of haplotypes within the nuclear and mitochondrial genes for U.S., CCAM and EULE subpopulations.

| | Frequency of frequent and unique haplotypes | | |
|---|---|---|---|
| | *Igs* 2 | *Ypt* 1 | *cox* 2 |
| **USA** | | | |
| Connecticut | 2 | 6 | 3 |
| Florida | 4 | 3 | 1 |
| Georgia | 4 | 6 | 5 |
| Kentucky | 12 | 2 | 6 |
| Maryland | 5 | 3 | 3 |
| North Carolina | 21 | 30 | 16 |
| Pennsylvania | 2 | 6 | 7 |
| Texas | 5 | 6 | 7 |
| Virginia | 4 | 4 | 2 |
| **CCAM** | | | |
| Dominican Republic | 14 | 16 | 15 |
| Guatemala | 6 | 4 | 2 |
| Mexico | 12 | 6 | 5 |
| Nicaragua | 5 | 2 | 8 |
| **EULE** | | | |
| Bulgaria | 3 | 6 | 3 |
| France | 11 | 6 | 9 |
| Germany | 4 | 2 | 4 |
| Hungary | 1 | 6 | 3 |
| Lebanon | 2 | 1 | 2 |
| Poland | 6 | 6 | 3 |

**Appendix 2. 7** Identity of haplotypes of *P. tabacina* inferred from nuclear regions (*Igs*2 and *Ypt*1) and a mitochondrial gene (*cox*2).

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| **Nuclear** | | | |
| *Igs* 2 region | H90 (209) | C4GA996, C3CT991, C5CT991, C6CT991, C3Penn, C1DR022, C2DR022, C5Colum, C2FR184, C2FRy1, C7FR194, C6FR178, C7GA996, C5NC9911, C9123344, C4PT87W, C7FR178, C1BU04, C1FRy1, C2BU061, C9TX84, C7Bertie, C12VA011, C4Hadley, C5123349, C4123368, C3NC997, C7Mosel, C4123349, C3NC0212, C7123368, C6TX84, C12FRy1, C1NIC063, C6NIC063, C1DR051, C3DR051, C1TX062, C3TX062, C2DR066, C4DR066, C5DR066, C8KY206, C3DR054, C5DR054, C2NIC065, C4NIC065, C7NIC065, C10KY206, C12KY206, C4FL002, C7BPTP, C5Penn, C6Mosel, C5123350, C6PT87W, C4BU04, C1123349, C8123349, C1NC0212, C2123368, C5123368, C11NC997, C8BU061, C2DR064, C7DR064, C2123296, C1123364, C4RPT83, C3BPTS, C2NC994, C1Mosel, C10Colum, C3NC9911, C4NC9911, C2123340, C7PT87W, C8PT87W, C8NC0212, C2KY79, C6KY79, C12NC997, C4TX84, C8TX84, C3CT021, C14Bertie, C10123347, C6BU04, C6NC994, C3DR064, C7123392, C6123340, C3KY79, C10BU061, C8FR178, C4BPTP, C3FR184, C5BPT, C1CT991, C7FR184, C2FR194, C2Penn, C5DR022, C8GA992, C1PE05, C5BPTP, C9Colum, C2PE05, C4DR062, C5FR194, C4Mex, C4BPTS, C6FR184, C1FR194, C4FR194, C7MX02, | All |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| *Igs* 2 region | H90 (209) continued | C4DR022, C6DR022, C2GA996, C3Hadley, C5DR064, C8VA011, C6RPT83, C2RPT83, C4Mosel, C1CT021, C2CT021, C6BPT, C7TX84, C1BPTP, C4CT991, C3FR194, C1GA992, C2GA992, C3GA992, C1AV04, C5AV04, C3NC998, C4NC998, C3NC994, C1123350, C5123340, C7123340, C5123344, C3PE05, C7PE05, C9PE05, C3PT87W, C5BU04, C6CT021, C6VA011, C3NIC063, C4NIC063, C3DR066, C3FL002, C6DR054, C5TX062, C6DR066, C7DR066, C1FL002, C6NIC065, C6NC998, C1BPT, C6BPTP, C2BPTS, C5BPTS, C6BPTS, C1Penn, C4Penn, C6Penn, C2Mosel, C8Colum, C7NC0212, C3123368, C6123368, C7BU061, C11123046, C6DR064, C9KY062, C9VA011, C1123347, C2123347, C3123347, C7123347, C3BPTP, C7BPTS, C2CT991, C1Mex, C2Mex, C7Mex, C4GA992, C7Colum, C1RPT83, C7RPT83, C6GA996, C6123350, C3123340, C8PE05, C1FR178, C4FR178, C3BU04, C7BU04, C12123046, C1123296, C13123296, C2Bertie, C3Bertie, C7CT021, C1NC994 | |
| | H17 (2) | C2PT87W, C6Hadley | U.S. (MD, NC) |
| | H21 (2) | C3DR022, C1TX84 | U.S. (TX)<br>CCAM: (Dom. Republic) |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H42 (4) | C7123046, C8123046, C13123046, C3123392 | CCAM: (Guatemala)<br>EULE: (Bulgaria) |
| | H64 (3) | C1GA996, C2123344, C7123344 | U.S. (GA)<br>EULE: (Poland) |
| | H66 (7) | C6123364, C6KY062, C10NC997, C14NC997, C5NC994, C5BU061, C5VA011 | U.S. (KY, NC, VA)<br>EULE: (France) |
| | H71 (2) | C2VA011, C1KY062 | U.S. (KY,VA) |
| | H75 (2) | C8Mosel, C2FL002 | U.S. (FL, VA) |
| | H79 (2) | C1123392, C3FRy1 | CCAM: (Guatemala)<br>EULE: (France) |
| | H81(2) | C1NC998, C9123347 | U.S. (NC)<br>CCAM: (Mexico) |
| | H82 (2) | C1DR062, C2BPT | U.S. (MD)<br>CCAM: (Dom. Republic) |
| | H83 (2) | C4DR05, C4NC994 | U.S. (NC)<br>CCAM: (Dom. Republic) |
| | H84 (3) | C4DR051, C6DR051, C1Hadley | U.S. (MD)<br>CCAM: (Dom. Republic) |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H87 (2) | C3BPT, C4AV04 | U.S. (MD, NC) |
| | H56 (1) | C4CT021 | U.S. (CT) |
| | H34 (1), H78 (1) | C5FL002, C6FL002 | U.S. (FL) |
| | H73 (1), H76 (1) | C7GA992, C3GA996 | U.S. (GA) |
| | H13 (1), H26 (1), H41 (1), H55 (1), H72 (1), H80 (1), H86 (1), H92 (1), H93 (1) | C9KY206, C1KY79, C5KY062, C11KY206, C12KY062, C3KY206, C3KY062, C8KY79, C5KY79 | U.S. (KY) |
| | H23 (1), H53 (1), H57 (1) | C5Hadley, C7BPT, C2Hadley | U.S. (MD) |
| | H1 (1), H10 (1), H15 (1), H16 (1), H22 (1), H24 (1), H27 (1), H31 (1), H38 (1), H48 (1), H50 (1), H61 (1), H63 (1), H67 (1), H88 (1) | C2AV04, C9NC997, C8RPT83, C6NC9911, C5NC0212, C1NC9911, C1Bertie, C3AV04, C9NC0212, C5NC998, C6Bertie, C7NC9911, C6AV04, C2NC998, C6BU061 | U.S. (NC) |
| | H46 (1) | C4Colum | U.S. (PA) |
| | H45 (1), H47 (1), H58 (1) | C4TX062, C2TX062, C6TX062 | U.S. (TX) |
| | H20 (1), H25 (1), H51 (1), H54 (1), H69 (1), H70 (1), H74 (1), H77 (1), H91 (1) | C2DR051, C8DR064, C9DR062, C7DR051, C2DR054, C2DR062, C5DR062, C6DR062, C1DR054 | CCAM (DR) |
| | H4 (1), H11 (1), H85 (1) | C4123392, C2123392, C10123392 | CCAM (GU) |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H4 (1), H11 (1), H85 (1) | C4123392, C2123392, C10123392 | CCAM (GU) |
| | H3 (1), H5 (1), H6 (1), H12 (1), H30 (1), H32 (1), H33 (1), H36 (1), H49 (1), H89 (1) | C11123296, C4123296, C3Mex, C3123296, C5MX02, C4MX02, C1MX02, C2MX02, C6MX02, C6Mex | CCAM (MX) |
| | H2 (1), H19 (1), H39 (1), H65 (1) | C3NIC065, C5NIC065, C2NIC063, C5NIC063 | CCAM (NIC) |
| | H44 (1) | C3123046 | EULE (BU) |
| | H7 (1), H9 (1), H18 (1), H29 (1), H40 (1), H43 (1), H59 (1), H60 (1) | C10FRy1, C2FR178, C2123364, C1FR184, C4FRy1, C5123364, C4FR184 | EULE (FR) |
| | H8 (1), H62 (1), H68 (1) | C3123350, C9123350, C4123350 | EULE (GE) |
| | H52 (1) | C4123340 | EULE (LE) |
| | H14 (1), H28 (1), H35 (1), H37 (1) | C10123344, C6123344, C3123349, C6123349 | EULE (PO) |
| **Nuclear** | | | |
| *Ypt region* | H1 (176) | C1CT991, C3Hadley, C3CT991, C2GA992, C5123350, C2BPTS, C1123340, C2123340, C4123340, C6Hadley, C2FL002, C2FR178, C2NC997, C6NC994, C7123349, C4123344, C6123344, C4TX84, C2Mosel, C5Mosel, C4123392, C4RPT83, C5RPT83, C5Mex, C8DR022, C3KY062, C5TX84, C6TX84, C2GA996, C4GA996, C5FR178, C4Penn, | All |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| *Ypt region* | H1 (176) continued | C5Hadley, C5KY062, C2Penn, C2PT87W, C2KY062, C8KY062, C1FL002, C12123296, C1BU04, C8TX84, C3GA996, C4MX02, C5Bertie, C7Bertie, C8123350, C1123347, C1Penn, C3Penn, C4Hadley, C1Colum, C5123349, C6123340, C5123364, C6PE05, C5NC994, C3123364, C6123347, C4NC998, C5NC998, C4AV04, C5Colum, C4DR066, C10DR066, C4123364, C4BU061, C5DR066, C6123350, C2Colum, C12CT021, C6MX02, C2123347, C8GA992, C4NC9911, C7Mosel, C7DR066, C3PT87W, C6PT87W, C3PE05, C1FR194, C2Mex, C7CT991, C1NC9911, C2NC0212, C5123340, C5NC0212, C5GA996, C4DR062, C2CT991, C4FL002, C1BPTP, C12BU04, C2TX84, C2AV04, C2DR062, C3123344, C6DR062, C4NC997, C4BPTP, C5BPTP, C7BPTP, C9KY79, C1NC997, C3123340, C1CT021, C8FR184, C5FR184, C4Mex, C8Mex, C9DR066, C5PT87W, C2123392, C9GA992, C6123296, C9123296, C3123347, C3BPTP, C2BPTP, C1NIC065, C2NIC065, C5NIC065, C2KY79, C11CT021, C8123296, C7FR194, C6FR194, C1VA011, C2VA011, C3VA011, C4VA011, C5VA011, C6VA011, C5PE05, C1Mex, C7AV04, C2123350, C10123350, C1123349, C8123046, C9AV04, C1KY79, C2Hadley, C4FR184, | |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| *Ypt region* | H1 (176) continued | C3FR184, C3NC997, C1MX02, C2MX02, C3MX02, C5MX02, C1PE05, C2PE05, C1KY062, C7KY062, C3123368, C6NC997, C2Bertie, C1DR062, C8123364, C9123364, C7123364, C3BPTS, C7FR178, C6FL002, C7KY79, C4BU04, C3Colum, C1NC994, C4NC994, C10123296, C1Hadley, C3Bertie, C4Bertie, C2123349, C1NC0212, C6NC0212 | |
| | H8 (2) | C2NC994, C3Mex | U.S. (NC)<br>CCAM: (Mexico) |
| | H9 (11) | C1FRy1, C2FRy1, C4FRy1, C3FR178, C3FR194, C4FR194, C4FR178, C6FR178, C2123368, C5FRy1, C3FRy1 | EULE: (France, Hungary) |
| | H41 (2) | C8Mosel, C4TX062 | U.S. (TX, VA) |
| | H55 (3) | C1TX062, C5DR022, C8123368 | U.S. (TX)<br>CCAM: (Dom. Republic)<br>EULE: (Hungary) |
| | H58 (2) | C9123347(CACMX), C4Colum(USAPA) | U.S. (PA)<br>CCAM: (Mexico) |
| | H67 (4) | C2DR054, C4DR054, C6DR054, C2FR184 | CCAM: (Mexico)<br>EULE: (France) |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H69 (6) | C2BPT, C4BPT, C6BPT, C8NC994, C13CT021, C5Penn | U.S. (CT, NC, MD, PA) |
| | H2 (1), H30 (1), H70 (1), H71 (1) | C8CT021, C4CT021, C4CT991, C5CT991 | U.S. (CT) |
| | H7 (1), H75(1) | C3FL002, C5FL002 | U.S. (FL) |
| | H15 (1), H34 (1), H84 (1), H89 (1), H93 (1) | C1GA996, C5GA992, C3GA992, C1GA992, C8GA996 | U.S. (GA) |
| | H62 (2) | C3KY79, C5KY79 | U.S. (KY) |
| | H43 (3) | C1BPT, C3BPT, C5BPT | U.S. (MD) |
| | H3 (1), H4 (1), H16 (1), H18 (1), H19 (1), H20 (1), H22(1), H24 (1), H28 (1), H38 (1) H42(1), H45(1), H47(1), H50 (2), H52(2), H56(1), H57(1), H61 (1), H66 (1), H68 (1), H77 (1), H79 (3), H80 (1), H85 (1), H87 (1), H91 (1), H92 (1) | C4NC0212, C5AV04, C7PT87W, C10BPTS, C5BU061, C5NC9911, C7BU061, C1PT87W, C3RPT83, C6AV04, C1RPT83, C1Bertie, C11BU04, C7NC9911, C2BU04, C5BPTS, C6NC9911, C7NC997, C2RPT83, C7NC998, C8NC998, C9NC998, C10NC998, C3BU061 C9BPTS, C7BU04, C1BPTS, C3NC0212 | U.S. (NC) |
| | H14 (1), H46 (1), H65 (1) | C6Penn, C4PE05, C6Colum | U.S. (PA) |
| | H13 (1), H23 (1), H72 (3) | C5TX062, C7TX84, C6TX062, C2TX062, C3TX062 | U.S. (TX) |
| | H25 (1), H27 (1) | C6Mosel, C4Mosel | U.S. (VA) |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H6 (1), H17 (3), H11 (1), H31 (1), H33 (1), H35 (1), H53 (1), H59 (1), H64(1), H76 (1), H78 (1), H81 (1), H82 (1), H90 (1) | C4DR064, C9DR064, C2DR022, C3DR022, C3DR064, C4DR022, C7DR064, C5DR064, C3DR066, C1DR022, C5DR062, C6DR064, C3DR062, C1DR054, C3DR054, C5DR054 | CCAM (DR) |
| | H21 (1), H29 (1), H40 (1) | C5123392, C3123392 | CCAM (GU) |
| | H36 (1), H51 (1) | C4123347, C1123296 | CCAM (MX) |
| | H5 (3) | C3NIC065, C4NIC065, C6NIC065 | CCAM (NIC) |
| | H32 (1), H39 (1), H49 (1), H63 (1), H74 (1) | C2123046, C4123046, C5123046, C7123046, C6123046 | EULE (BU) |
| | H10 (1), H12 (1), H83 (1) | C6FRy1, C1FR184, C5FR194 | EULE (FR) |
| | H73 (1) | C9123350 | EULE (GE) |
| | H26 (1), H37 (1), H60 (1) | C5123368, C1123368, C10123368 | EULE (HU) |
| | H44 (1), H48 (1), H54 (1), H86 (1), H88 (1) | C7123344, C3123349, C1123344, C2123344, C4123349 | EULE (PO) |
| **Mitochondrial** | | | |
| *Cox2 gene* | H1 (235) | C4123340, C1TX84, C5TX84, C1GA992, C3TX84, C5DR054, C3GA992, C2Penn, C4RPT83, C4FR184, C4NC997, C4DR066, C2BU04, C1Colum, C1Penn, C6GA996, | All |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| **Mitochondrial** | | | |
| *Cox2 gene* | H1 (235) continued | C6123296, C5FR178, C1FL002, C5Penn, C3BU061, C4TX84, C4BPT, C5GA996, C7BPT, C6BPTP, C2BPT, C3FR194, C4FR178, C8FRy1, C1PE05, C5123368, C2FRy1, C6FRy1, C3123368, C6KY79, C2123368, C5KY79, C4KY79, C2KY79, C3BPTS, C6KY206, C6123344, C2BPTS, C4123344, C5123344, C1FR194, C3123344, C5BPT, C6BPT, C7KY206, C8BPT, C1NIC065, C8BPTP, C6NC0212, C5NC9911, C4Penn, C7Mex, C5DR064, C6DR064, C4Hadley, C3NC9911, C2NC9911, C1FRy1, C1AV04, C4AV04, C4BU061, C7TX84, C2BU061, C7DR022, C5NC998, C6DR022, C2Mex, C6VA011, C5VA011, C4VA011, C6DR054, C6DR066, C4123046, C6BU04, C2RPT83, C6DR062, C2DR062, C8123340, C1BU04, C5123349, C5123340, C6123340, C1RPT83, C5CT991, C7123347, C6123392, C6KY062, C4Colum, C3DR051, C6CT991, C2123349, C2KY062, C5FL002, C7CT021, C6GA992, C2NC0212, C5BPTP, C1123344, C3KY206, C2123364, C5PT87W, C3AV04, C7BPTS, C5FRy1, C5DR022, C4NC9911, C6NC9911, C1123368, C4NC0212, C1CT991, C2123347, C3123347, C4123349, C6FR178, C9123340, C7CT991, C4123347, C4Mex, | |

Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H1 (235) continued | C8PT87W, C3NC994, C3123349, C4123364, C4MX02, C5MX02, C5Colum, C3Mosel, C3CT021, C3VA011, C2VA011, C5CT021, C5RPT83, C3BPTP, C2Colum, C4CT991, C6Colum, C2NC998, C3NC998, C2123392, C5Mosel, C3BU04, C1MX02, C1123392, C2CT021, C7Bertie, C6Bertie, C5BU04, C6NC997, C3FR184, C2FR184, C1FR184, C4Mosel, C6Mosel, C6MX02, C5123347, C3RPT83, C3123392, C6123347, C2123350, C6123350, C2PT87W, C1NC994, C4NC994, C6NC994, C4DR051, C2123296, C5123296, C5123364, C6123364, C7DR051, C6NIC063, C2FL002, C4FL002, C1DR062, C6FL002, C3FL002, C1KY062, C5NIC063, C6DR051, C3123296, C1123349, C7Hadley, C6Hadley, C5Hadley, C4KY206, C6BPTS, C5BPTS, C4BPTS, C1FR178, C2FR194, C2GA992, C1GA996, C1DR022, C2TX84, C1NC9911, C5NC0212, C2KY206, C5KY206, C1Mosel, C2Mosel, C3Penn, C6BU061, C2AV04, C1NC0212, C7BPTP, C6FR194, C4BPTP, C3NC0212, C6AV04, C1BU061, C6Penn, C5GA992, C5Bertie, C2DR054, C1PT87W, C3DR066, C3DR054, C2123046, C4NC998, C6NC998, C7123392, C6123046, C5123046, C6FR184, C3123350, C1Mex, C2NIC063, C2NC994, C1123296, C1CT021 | |

Appendix 2.7 continued

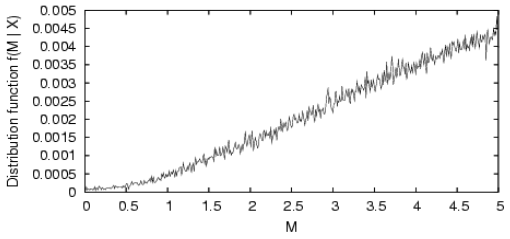
| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H57 (2) | C4DR062, C3Colum | U.S. (PA)<br>CCAM: (Dom. Republic) |
| | H58 (3) | C5Mex, C2FR178, C6Mex | CCAM: (Mexico)<br>EULE: (France) |
| | H76 (2) | C4CT02, C3NIC065 | U.S. (CT)<br>CCAM: (Nicaragua) |
| | H13 (1) | C2CT991 | U.S. (CT) |
| | H4 (1), H35 (1), H43 (1), H51 (1) | C4GA992, C3GA996, C4GA996, C2GA996 | U.S. (GA) |
| | H11 (1), H27 (1), H46 (1), H53 (1), H77 (1) | C5KY062, C4KY062, C1KY79, C3KY79, C3KY062 | U.S. (KY) |
| | H62 (1), H65 (1) | C3Hadley, C2Hadley | U.S. (MD) |
| | H5 (1), H26 (1), H28 (1), H31 (1), H33 (1), H34 (1), H38 (1), H52 (1), H55 (1), H60 (1), H66 (1), H70 (1), H71 (1), H79 (1), H81 (1) | C5NC997, C3NC997, C4PT87W, C2NC997, C3Bertie, C3PT87W, C1NC998, C5NC994, C5AV04, C1NC997, C6RPT83, C4BU04, C4Bertie, C5BU061, C2Bertie | U.S. (NC) |
| | H25 (1), H41 (1), H42 (1), H68 (1), H73 (1) | C5PE05, C2PE05, C6PE05, C3PE05, C4PE05 | U.S. (PA) |

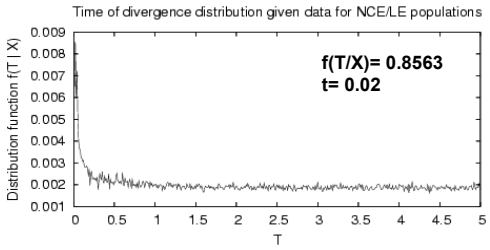
Appendix 2.7 continued

| Locus | Haplotype (frequency) | Isolates | Population |
|---|---|---|---|
| | H8 (1), H10 (1), H59 (1), H69 (1), H74 (1), H78 (1) | C2TX062, C3TX062, C7TX062, C1TX062, C4TX062, C5TX062 | U.S. (TX) |
| | H22 (1) | C7VA011 | U.S. (VA) |
| | H7 (2), H12 (1), H14 (1), H15 (1), H16 (1), H17 (1), H23 (1), H24 (1), H36 (1), H37 (1), H49 (1), H67 (2) | C5DR051, C3DR062, C3DR022, C4DR022, C5DR062, C8DR051, C4DR054, C1DR054, C3DR064, C5DR066, C2DR064, C4DR064 C1DR066, C2DR066 | CCAM (DR) |
| | H50 (1) | C5123392 | CCAM (GU) |
| | H54 (1), H61 (1), H64 (1) | C3MX02, C2MX02, C4123296 | CCAM (MX) |
| | H2 (2), H3 (1), H30 (1), H39 (1), H40 (1), H45 (1) | C2NIC065, C1NIC063, C4NIC063, C3NIC063, C5NIC065, C6NIC065, C4NIC065 | CCAM (NIC) |
| | H20 (1), H48 (1) | C3123046, C1123046 | EULE (BU) |
| | H19 (1), H29 (1), H44 (1), H47 (1), H63 (1), H75 (1), H82 (1) | C4FR194, C1123364, C3FR178, C3FRy1, C5FR194, C3123364, C5FR184 | EULE (FR) |
| | H18 (1), H21 (1), H72 (1) | C5123350, C1123350, C4123350 | EULE (GE) |
| | H9 (1), H80 (1) | C6123368, C4123368 | EULE (HU) |
| | H32 (1), H56 (1) | C2123344, C6123349 | EULE (PO) |

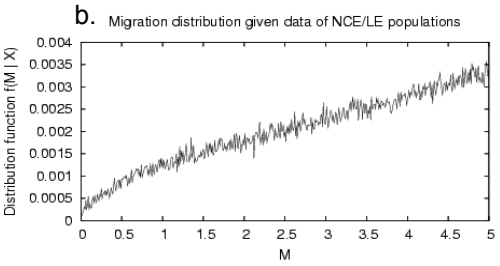
**Appendix 2. 8** Migration and time of divergence of subdivided populations in U.S., CCAM and EULE subpopulations. a. NCE/LE for *Igs*2 region; b. NCE/LE for *Ypt*1 region; c. NCE/LE for *cox*2 gene; d. NCE/WE for *Igs*2 region; e. NCE/WE for *Ypt*1 region; f. NCE/WE for *cox*2 gene; g. WE/LE for *Igs*2 region; h. WE/LE for *Ypt*1 region; i. WE/LE for *cox*2 gene; j. US/CCAM for *Igs*2 region and k. US/CCAM for *Ypt*1 region. Data for the US/CCAM for the *cox*2 gene was insufficient to generate the plots.

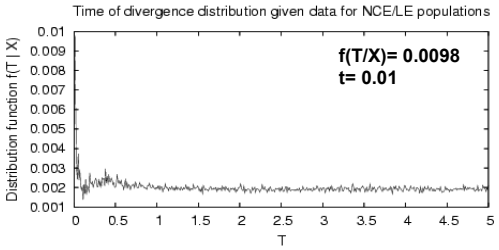
a.



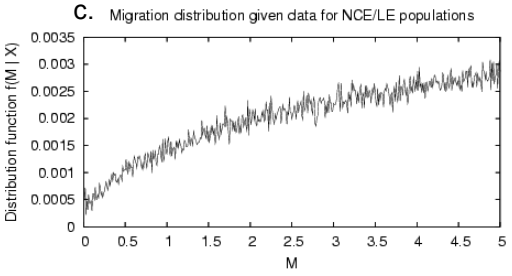



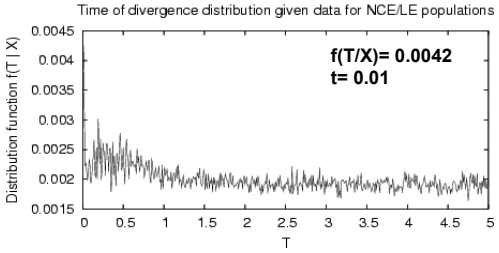

b.



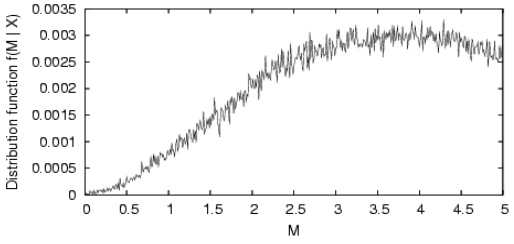



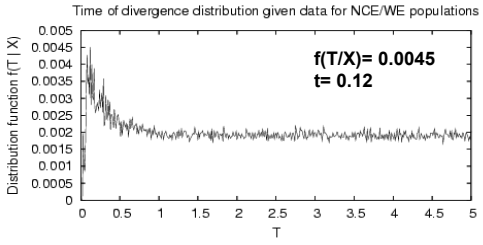

c.



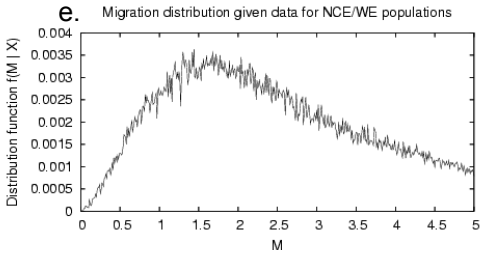

d.

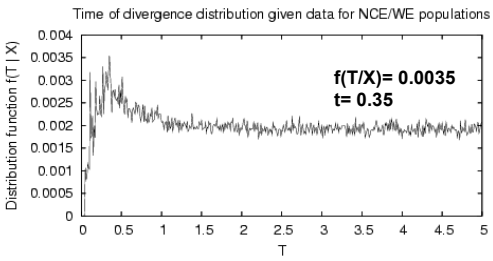


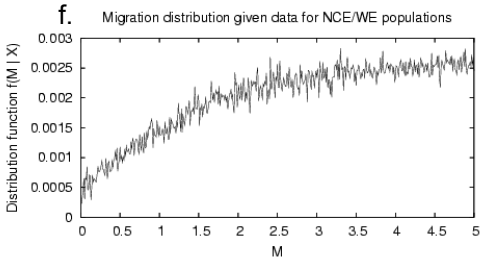
Appendix 2.8 continued

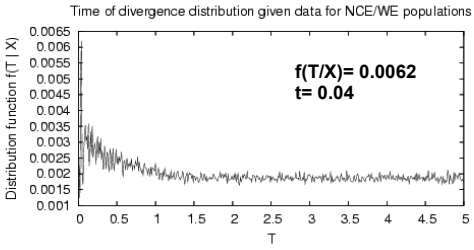
e.



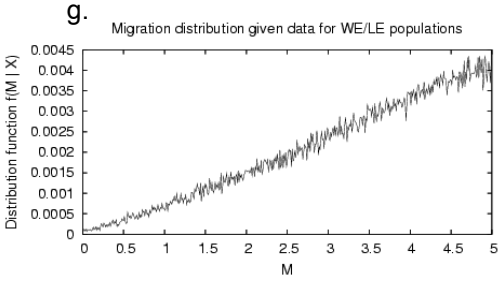



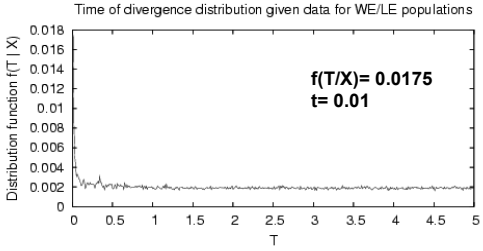

f.



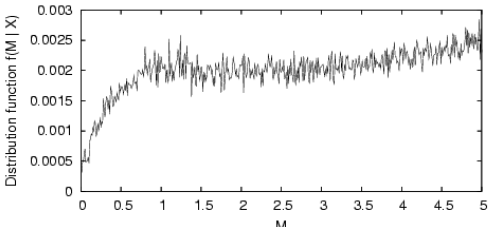



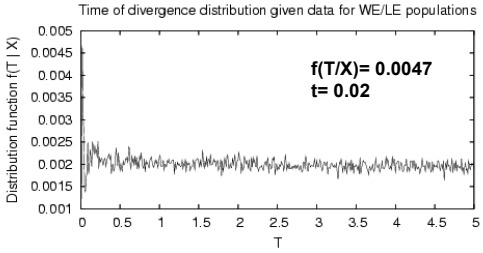

g.



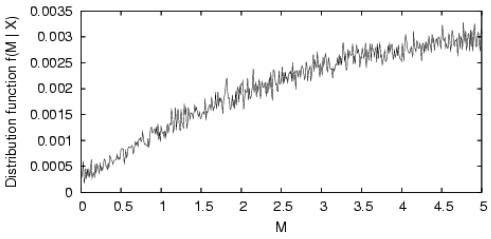



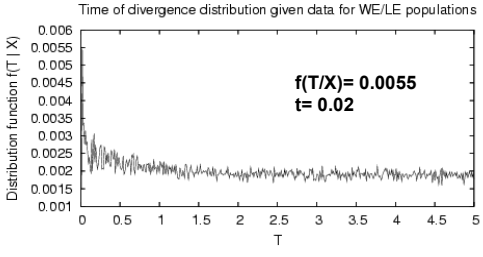

h.



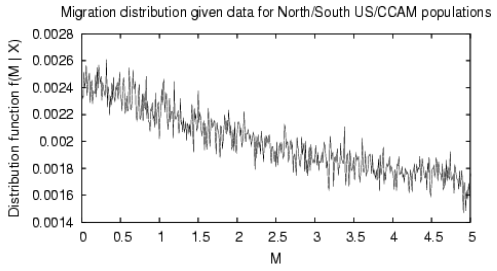

i.

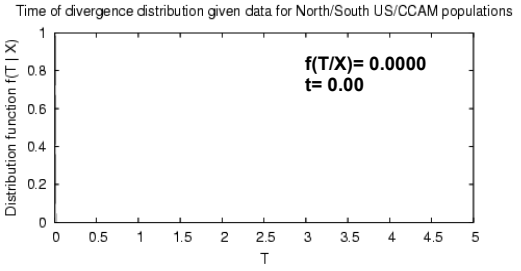


Appendix 2.8 continued

j.




k.