ABSTRACT

MOORE, GEROMY GEORGE. Global Population Structure and Aflatoxin Chemotype Diversity in *Aspergillus* Section *Flavi*. (Under the direction of Ignazio Carbone.)

The genus *Aspergillus* is comprised of over 250 species that are grouped into sections based on similar phenotypic characters. Section *Flavi* includes species that have great economic impact on a global scale since they exist all over the world. Beneficial species such as *Aspergillus oryzae* are integral to fermentation of soy sauce and sake, but this research focuses on those species in Section *Flavi* that negatively affect global economies. Among these are species that are pathogenic on important agricultural commodities, but that also have the potential to produce carcinogenic substances known as aflatoxins. These aflatoxins are synthesized in a biosynthetic pathway that involves at least 25 clustered genes. Two of the most common species that produce aflatoxins are *Aspergillus flavus* Link and *Aspergillus parasiticus* Speare. Since the first publicized outbreak of aflatoxins with Turkey X disease over 40 years ago, researchers have sought to control contamination of food by aflatoxin producing Aspergilli. Pre-harvest biocontrol using E.P.A. approved biopesticides is currently one of the preferred methods of control using atoxigenic *A. flavus* strains to competitively exclude indigenous aflatoxin producers. An important consideration for the success of these atoxigenic strains as biocontrol is that these fungi are considered asexual and incapable of recombining with indigenous strains. Recent discoveries of mating type genes in these fungi, and subsequent discoveries of sexual states in *A. flavus* and *A. parasiticus,* call into question the stability of the biocontrol strains since the potential for them to outcross and regain aflatoxigenicity exists.

The purpose of this research was to conduct a molecular examination of the structure of populations of different species within Section *Flavi*, most importantly *A. flavus* and *A. parasiticus*, from geographically isolated peanut fields representing five continents. DNA was isolated from sample strains, oligonucleotide primers were used to amplify and sequence across regions within the aflatoxin gene cluster as well as non-cluster loci, and then statistical analyses were performed for comparisons within populations and between species and localities. We found evidence of recombination in most of the populations, with higher rates detected in *A. flavus* compared to *A. parasiticus*. Recombination breakpoint boundaries appear conserved within each species and across geographies. We found that differences in the proportions of *MAT1-1* and *MAT1-2* were correlated with the amount of asexual (i.e. clonal amplification of VCGs) and sexual reproduction in populations. For both *A. flavus* and *A. parasiticus*, when the numbers of *MAT1-1* and *MAT1-2* were significantly different, there was more extensive linkage disequilibrium in the cluster and isolates grouped into specific toxin classes, either the nonaflatoxigenic class in *A. flavus* or the B_1-dominant and G_1-dominant classes in *A. parasiticus*. Our work shows that a combination of ecological factors, asexual/sexual reproduction and balancing selection may influence aflatoxin diversity in these agriculturally important fungi. Finally, we show that climate change may influence gene flow in these toxin-producing species and thereby exacerbate problems in the future by facilitating selection for better adapted, more toxic pathogens. Collectively the results of this research have significant implications for sustainable biocontrol of aflatoxigenic fungi worldwide.

Global Population Structure and Aflatoxin Chemotype Diversity in *Aspergillus* Section *Flavi*

by
Geromy George Moore

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

2010

APPROVED BY:

Dr. Ignazio Carbone Chair of Advisory Committee	Dr. Gary Payne
Dr. Marc Cubeta	Dr. Eric Stone

DEDICATION

To fate and its mysterious ways!

BIOGRAPHY

Geromy George Moore was born in Dallas, Texas. He moved around a lot while growing up, even living abroad during high school. He earned an Associate of Science degree in 1996 from Navarro College in Corsicana, Texas; a Bachelor of Science degree in 1999 from Stephen F. Austin State University in Nacogdoches, Texas; and is ecstatic to be earning his Doctoral Degree for Plant Pathology in 2010 from North Carolina State University. He had the privilege of working with a great fungus named *Aspergillus* during his time at NCSU. His favorite experience during graduate school (aside from actually finishing) is being an authority on a newly named fungus called *Petromyces flavus*. He is a proud "mama's boy" and "grandmama's boy" with very close family ties. He's had the good fortune of living a happy life full of great friends and adventures; hence, his life-philosophy is Live! Laugh! Love! Fate likes to keep him guessing (see previous dedication) which always makes for an interesting journey and he wouldn't have it any other way. It is his sincerest hope that he can now start a new chapter in life that involves a rewarding career and a family of his own!

ACKNOWLEDGEMENTS

I would like to thank the members of my advisory committee, Drs. Ignazio Carbone, Gary Payne, Marc Cubeta, and Eric Stone for their encouragement and guidance. I would especially like to thank Dr. Ignazio Carbone for selecting me as his very first (and favorite) graduate student. His patience and support, as well as his unwavering faith in my success have meant much. Sincerest appreciation goes to members of the lab who have helped me complete a rather hefty project: Judy Jakobek, Jacalyn Elliott, Carolyn Worthington, Rakhi Singh, Jorge Ramirez-Prado, Eric Price, and others. It is truly as a team that science happens! Thanks must also be given to the funding sources for this project: NC Cooperative State Research, Education and Extension Service, and the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service. Thanks to the many professors who shared their passions for plant pathology with us students. I should not forget *Aspergillus*... a most interesting fungal genus to study! Most importantly I need to thank my partner, Jeff, my parents, Richard and Theresa Wallot, my dad, Ralph Moore, and my grandparents, George and Sandra Smith. Their patience and unconditional love have made the stresses of going back to school worth it. My love to you all!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

CHAPTER 1

CHAPTER 2

CHAPTER 3

Introduction

1. The genus *Aspergillus*

1.1 Taxonomy and Morphology

The filamentous fungi that comprise the genus *Aspergillus* are taxonomically placed within the phylum Ascomycota, the class Eurotiomycetes, the order Eurotiales, and the family Trichocomaceae (Alexopoulos, Mims et al. 1996; Geiser, Gueidan et al. 2006). Fungi in the Ascomycota are so grouped because their teleomorph spores are formed in sac-like structures called asci. The grouping of fungi within Eurotiomycetes/Eurotiales is because the asci are housed within an enclosed fruiting body called a cleistothecium. Fungi in the Trichocomaceae share the common feature of ascospore release only after the prototunicate ascus is broken open (Alexopoulos, Mims et al. 1996; Taylor, Spatafora et al. 2006). There exist genera that share the family Trichocomaceae, e.g. *Penicillium*, which may be confused with Aspergilli, but there are certain morphological characters that are unique to *Aspergillus* such as L- or T-shaped foot cells that anchor the conidiophore. The cell wall of foot cells is thicker, and adds support for aerial hyphae that become the conidiophore stipes (Alexopoulos, Mims et al. 1996; Klich 2002). An aseptate conidiophore stipe is also considered unique to this genus (Pitt and Hocking 1985). The vesicles (swellings) that form at the tips of conidiophores are unique, and though phialides are not found only with Aspergilli, their simultaneous formation on the vesicle is a character only observed with *Aspergillus* species (Klich 2002). Conidia form as chains from each phialide, held together by a connective tissue (Raper and Fennell 1965). All teleomorphs within this genus produce a cleistothecium containing prototunicate, evanescent asci, with each ascus containing eight ascospores (Alexopoulos, Mims et al. 1996; Geiser, Klich et al. 2007). Morphology is also important to distinguish between species within this expansive genus. On a macro-scale,

characters such as conidial color and ascocarp formation are used to identify a species; as well, micro-scale characters such as vesicle shape and spore ornamentation are integral for correct identification of species (Geiser, Klich et al. 2007; Bennett 2010). A physiological character that aids in distinguishing certain species is secondary metabolite, e.g. mycotoxin, production (Cabañes and Bragulat 2008; Cary, Ehrlich et al. 2009).

Within the genus *Aspergillus* there are six subgenera; within each subgenus there may be more than one section, and currently there are 18 sections named (Raper and Fennell 1965; Gams, Christensen et al. 1985; Klich 2002). The groupings can be seen in Table I.1. Each section may include multiple species; there are now more than 250 recognized species, both anamorphic and teleomorphic, encompassed within this genus (Geiser, Klich et al. 2007; Geiser, Samson et al. 2008). Historically, these subgenera, sections, and species have been separated mostly upon morphology, physiology and teleomorph associations (Balajee 2008; Geiser, Samson et al. 2008; Cary, Ehrlich et al. 2009); but recently, phylogenetic inference has been included as an additional means of accurately and quickly grouping, or separating, the vast number of *Aspergillus* species (Pitt and Samson 2007; Balajee 2008). Stephen Peterson's recent phylogenetic work argues that the species concept should not be based on morphology alone, and suggests that there is necessity to change the numbers of subgenera and sections (Table I.1) considering monophyly (Peterson 2000; Peterson, Varga et al. 2008). Taxa that fail to exhibit monophyletic association within their original section/subgenus are moved to other/new groupings, and some species are removed from the

Table I.1 Grouping of subgenera and sections within the genus *Aspergillus*

Historical Subgenera and Sections[a]	Phylogenetic Subgenera and Sections[b]
Subgenus *Aspergillus*	**Subgenus *Aspergillus***
Section *Aspergillus*	Section *Aspergillus*
Section *Restricti*	Section *Restricti*
Subgenus *Circumdati*	**Subgenus *Candidi***
Section *Candidi*	Section *Candidi*
Section *Circumdati*	**Subgenus *Circumdati***
Section *Cremei*	Section *Circumdati*
Section *Flavi*	Section *Cremei*
Section *Nigri*	Section *Flavi*
Section *Sparsi*	Section *Nigri*
Section *Wentii*	**Subgenus *Fumigati***
Subgenus *Clavati*	Section *Cervini*
Section *Clavati*	Section *Clavati*
Subgenus *Fumigati*	Section *Fumigati*
Section *Cervini*	**Subgenus *Nidulantes***
Section *Fumigati*	Section *Bispori*
Subgenus *Nidulantes*	Section *Nidulantes*
Section *Flavipedes*	Section *Ochraceorosei*
Section *Nidulantes*	Section *Raperi*
Section *Terrei*	Section *Silvati*
Section *Usti*	Section *Sparsi*
Section *Versicolores*	Section *Usti*
Subgenus *Ornati*	**Subgenus *Ornati***
	Section *Ornati*
	Subgenus *Terrei*
	Section *Flavipedes*
	Section *Terrei*
	Subgenus *Warcupi*
	Section *Warcupi*
	Section *Zonati*

[a] Historical groupings taken from *Identification of Common Aspergillus species* (Klich 2002).

[b] Suggested groupings by Stephen Peterson's phylogenetic analyses (Peterson, Varga et al. 2008).

genus altogether (Peterson, Varga et al. 2008). A synthesis of classification (Balajee 2008) based on both character and molecular methods would seem the most beneficial to the *Aspergillus* research community (Samson, Hong et al. 2006; Geiser, Klich et al. 2007; Pitt and Samson 2007), but clinical researchers may continue to choose the molecular approach to attempt more rapid identifications (Balajee 2008; Manikandan, Dóczi et al. 2008).

1.2 Timeline and Sexuality

Pietro Antonio Micheli, an Italian botanist and priest, first characterized the genus *Aspergillus* in 1729. He observed that the conidiophores resembled an aspergillum… a device used by Catholic priests to sprinkle holy water. Micheli included his characterization of *Aspergillus* in his work titled 'Nova Plantarum Genera' in 1729 (Raper and Fennell 1965; Varga and Samson 2008). Since then there have been three monographs chronicling Aspergilli: Thom and Church (1926), Thom and Raper (1945), and Raper and Fennell's 'The Genus *Aspergillus*' (1965) considered "the last complete monograph" of the genus (Geiser, Klich et al. 2007; Geiser, Samson et al. 2008). At the time of Micheli's characterization, *Aspergillus* was labeled as an asexual fungus, and this belief held fast until the early 1800s when a mycologist named Johann H. Link discovered a cleistothecium-producing species on a herbarium specimen and named it *Eurotium herbariorum*. This was later associated as the teleomorph to *A. glaucus* by Anton DeBary (Raper and Fennell 1965; Bennett 2010), and would have been the first indication that Aspergilli were pleomorphic by producing both asexual (mitotic) and sexual (meiotic) spores (Reynolds 1994). For the most part, *Aspergillus* is considered asexual (Figure I.1), possibly due to previous failed attempts at mating, and the lack of evidence that sex was occurring in nature

(Raper and Fennell 1965; Pitt and Samson 2007). The first investigation of parasexuality in this genus was for *A. nidulans* in the 1950s; parasexuality (Fig. I.1) being a means of genetic recombination in an asexual fungus (Pontecorvo 1956). In the 1970s, Ken Papa further explored the idea of parasexual reproduction, and performed several experiments to illustrate it, as a means of genetic recombination in *A. flavus* and *A. parasiticus* (Papa 1973; Papa 1978; Pál, van Diepeningen et al. 2008). Then, with the identification, and subsequent use of mating-type genes and their diagnostic PCR tests, the number of sexual states (Fig. I.1) being discovered and characterized is increasing (Paoletti, Rydholm et al. 2005; Pál, van Diepeningen et al. 2008; Ramirez-Prado, Moore et al. 2008). The most recent discoveries are *Neosartorya fumigata* (O'Gorman, Fuller et al. 2009), *Petromyces flavus* and *Petromyces parasiticus* (Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009; Horn, Ramirez-Prado et al. 2009). Today we use *Aspergillus* for the anamorph state, and most of the names for the teleomorph states are based on how their cleistothecia appear and/or form (Table I.2) (Geiser, Klich et al. 2007). A relatively small number of sexual stage associations (~33%) have been assigned for the genus, and most of those teleomorphs are considered homothallic (Geiser 2009; Bennett 2010).

Table I.2 Associations with naming *Aspergillus* teleomorphs based on their cleistothecia.

Teleomorph Genus	Notes on Cleistothecia
Neosartorya	Cleistothecial wall composed of flattened hyphae
Emericella	Dark cleistothecium surrounded by stromatic layer of hülle cells; red ascospores
Fennellia	Cleistothecial wall composed of thick-walled cells
Petromyces	Multiple cleistothecia enclosed in dark sclerotial stroma
Neopetromyces	1-2 cleistothecia enclosed in pale sclerotial stroma
Eurotium	Bright red, orange, and yellow cleistothecia

Information taken from two sources: (Klich 2002; Geiser 2009).

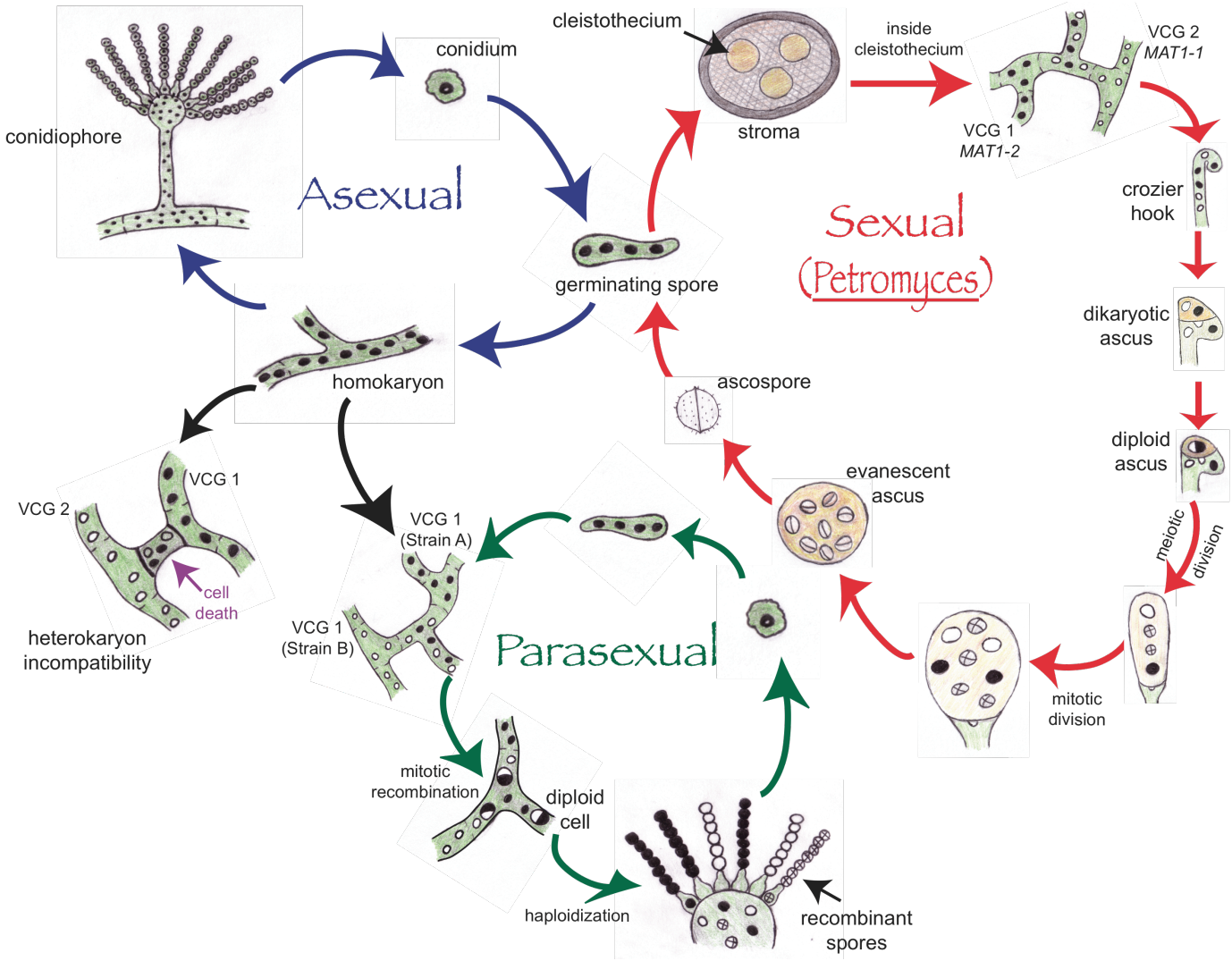


Figure I.1 Life cycle for *Aspergillus* based on asexual, parasexual, and sexual reproduction.

Aspergillus is the only fungal genus that offers evidence of heterothallism as ancestral to homothallism, because some researchers believe that if the homothallic mating system in Aspergilli were ancient then it would have been observed long ago (Geiser, Timberlake et al. 1996; Geiser 2009). Heterothallic fungi will have only one of two idiomorph types at a single mating (*MAT*) locus (*MAT1-1* or *MAT1-2*) and facilitates a self-sterile condition, while homothallic fungi will have both mating types in a single genome facilitating a self-fertile condition. In homothallic fungi the *MAT* idiomorphs mostly reside on different chromosomes, but in *Petromyces alliaceus* both genes are found on the same chromosome (Metzenberg and Glass 1990; Pál, van Diepeningen et al. 2008; Ramirez-Prado, Moore et al. 2008; Kwon-Chung and Sugui 2009).

1.3 Life cycle and Importance

Aspergillus species are frequently regarded as saprobes that colonize non-living tissue (de Vries 2008; Bennett 2010). Their life cycle can be simply explained in three steps: sporulation for conidia or ascospores, dissemination of those spores (mostly by wind), and spore germination to colonize a carbon source (Austwick 1965; Scheidegger and Payne 2003; Bennett 2010). A nutrient recession may cause overwintering as sclerotia/stromata or mycelia until conditions are again favorable for the fungus (Scheidegger and Payne 2003; Abbas, Wilkinson et al. 2009). *Aspergillus* is one of the most economically important fungal genera (Geiser, Klich et al. 2007; Geiser, Samson et al. 2008), and estimates of the global economic impact are reported to be in the hundreds of billions of dollars. For example, in 1997 the impact on the U.S. alone was quoted at approximately $45 billion from *Aspergillus*

species representing the good, the bad, and the ugly (May and Adams 1997). *A. (Emericella) nidulans* is of benefit as a model organism for *Aspergillus* studies. It has offered much insight about the genus through parasexuality and regulation of metabolic pathways, and continues to be studied extensively (May and Adams 1997; Bennett 2010). Some beneficial Aspergilli are used for enzyme production such as citric acid with *A. niger*, and centuries-old fermentation practices such as for soy sauce with *A. oryzae* (Machida, Yamada et al. 2008; van Dijck 2008; Bennett 2010). Both species have the potential to produce harmful toxins and must be monitored/tested before distribution to consumers (Blumenthal 2004), but their overall safety and productivity make them invaluable to industry (van Dijck 2008). The more serious incarnations of the Aspergilli are as pathogens and producers of mycotoxins (Klich 2007). *A. fumigatus* is a known human pathogen affiliated with aspergillosis infections in immunocompromised individuals (Latgé 1999). *A. flavus* is not only a plant pathogen and mycotoxin producer, it is also gaining notoriety as a pathogen of animals and humans (Scheidegger and Payne 2003; Geiser, Gueidan et al. 2006; Klich 2007; Krishnan, Manavathu et al. 2009; Bennett 2010).

1.4 Pathogenicity

Any existence or incarnation of *Aspergillus*, other than saprobic, is not essential to its life cycle, but under certain conditions this fungus can become pathogenic and cause problems in living host tissue (Austwick 1965; de Vries 2008). In 2003, Scheidegger and Payne reported on pathogenicity of *A. flavus* in plant hosts. They explained the predominant condition, favorable for pathogenicity, is weakened host defense. For plants this may be the result of

high temperature and drought stress which might cause wounds in plant tissue that allow the fungus to colonize and invade the plant host (Scheidegger and Payne 2003; Klich 2007). This scenario holds for animal/human hosts as well… that promotion of infection by pathogenic *Aspergillus* coincides with a weakened immune system (Krishnan, Manavathu et al. 2009; Bennett 2010). Another potential scenario for fungal invasion of plant hosts is through tissue damage by insects (Blankenship, Cole et al. 1984; Klich 2007). Host specificity is not well defined for pathogenic *Aspergillus*, and infection occurrences have been reported in fields, hospitals, and even marine environments such as with *A. sydowii* on sea coral (May and Adams 1997; Baker and Bennett 2008; Toledo-Hernandez, Zuluaga-Montero et al. 2008).

1.5 Toxigenicity

Pathogenicity is not the only potential hazard posed by Aspergilli since some species produce mycotoxins. Mycotoxins are extrolites, or secondary metabolites, whose function is still undetermined, but hypotheses for their purpose relate to fitness for the fungus and include: competition/protection (Fox and Howlett 2008), increased conidiation (Calvo, Wilson et al. 2002; Wilkinson, Ramaswamy et al. 2004), natural insecticide (Beard and Walton 1969), pheromones/quorum sensing (Rasmussen, Skindersoe et al. 2005; Amaike and Keller 2009), and virulence factors (Ehrlich 2006). Although their purpose is yet to be determined, their economic impact is without question. In 2003 the Council for Agricultural Science and Technology estimated an economic impact of $932 million by mycotoxins in the United States with estimates of aflatoxin contamination costs ranging from $4-40 million (CAST 2003; Robens 2005; Klich 2007). Ochratoxin, aflatoxin, and various acid compounds are

among the multitude of mycotoxins synthesized by one or more species of *Aspergillus* (Varga and Samson 2008; Kabak and Dobson 2009). A comprehensive list of *Aspergillus* species and the respective mycotoxins can be found at:

http://www.aspergillus.org.uk/indexhome.htm?secure/metabolites/list_by_secmet.php~main (www.aspergillus.org.uk 2009) Ochratoxin A (OTA) is a nephrotoxic mycotoxin, classified as a human renal carcinogen, produced by species such as *A. alliaceus* (Müller, Rosner et al. 2003; Cabañes and Bragulat 2008). Cyclopiazonic acid (CPA) is a mycotoxin that is serious at high concentrations, and it is known that the genes responsible for its synthesis are clustered and adjacent to the aflatoxin gene cluster (Sorenson, Tucker et al. 1984; Chang, Horn et al. 2009; Georgianna and Payne 2009). Sterigmatocystin (ST) and *O*-methyl-sterigmatocystin (OMST) are manufactured in the aflatoxin cluster as precursors to aflatoxin (AF) (Bhatnagar, McCormick et al. 1987; Keller, Kantz et al. 1994; Carbone, Ramirez-Prado et al. 2007). Aflatoxins (AFs) are carcinogens that may be either B-type (B_1 or B_2) or G-type (G_1 or G_2), and type B_1 is considered the most potent AF (Cullen and Newberne 1994). The B/G designation is based on how they fluoresce under long wave ultraviolet light-- B AFs fluoresce blue and G AFs fluoresce green (Klich 2007; Abbas, Wilkinson et al. 2009). AFs may persist in animals raised for human consumption such as cattle or fish, or as derivatives of AFs such as M-1 that accumulate in the milk of cows, fed AF-contaminated feed (El-Sayed and Khalil 2009; Prandini, Tansini et al. 2009). Although reports indicate that the genes responsible for mycotoxin/secondary metabolite production are clustered, the gene order and/or the genes required may vary considerably (Cary and Ehrlich 2006; Ehrlich 2008; Cary, Ehrlich et al. 2009).

2. Aflatoxigenicity

2.1 Global Impact of Aflatoxigenicity in Agriculture

Aflatoxins have earned much attention, and are currently under strict regulation for food/feed commodities, since their first publicized outbreak more than 40 years ago (Blount 1961; Horn 2003; Reddy, Abbas et al. 2009). From Turkey X disease in 1960s UK (Blount 1961), to outbreaks among immune-suppressed people in Africa (2004) and pet deaths from contaminated food in Taiwan (2008), society is aware of the risks posed by AFs and many countries have set strict regulations for AF concentrations in foods (Wu, Liu et al. 2008; Reddy, Abbas et al. 2009). It is reported that 4.5 billion people in developing countries may suffer from exposure to aflatoxins through the foods they eat (CDC 2010). These health risks bring to light the global impact of aflatoxigenicity, and have sparked over 20 years of research to reduce AF contamination through molecular studies, biological control studies, and population studies (Varga and Samson 2008).

2.2 Molecular Studies of Aflatoxigenic *Aspergillus*

Research began, and continues, with molecular investigations of the genes and regulatory mechanisms involved in AF biosynthesis in an attempt to understand, and potentially halt, AF production (Georgianna and Payne 2009). The advances in genomic studies of Aspergilli are increasing due to the sequenced complete genomes of eight species, and potentially more to come (Carbone, Ramirez-Prado et al. 2007; Geiser, Samson et al. 2008; Andersen and Nielsen 2009). Genome sequences have allowed for creation of genetic markers for many informative loci, especially the AF cluster genes (Chang, Skory et al. 1992; Geiser, Samson et al. 2008). The first AF biosynthesis gene to be investigated was *nor-1* (*aflB*) from

A. parasiticus in 1992 (Chang, Skory et al. 1992) Historically the AF cluster has been reported to encompass 25 genes in the 70-kb sub-telomeric region of chromosome 3R (Fig. 1 from Yu, Chang et al. 2004), but recent research found five additional transcripts of hypothetical function that may contribute to AF synthesis (Fig.1 from Cleveland, Yu et al. 2009). These transcripts were discovered after the creation of expressed sequence tags (ESTs) which have offered more resolution to the genomic condition of important species of *Aspergillus* (Yu, Chang et al. 2004; Cleveland, Yu et al. 2009), and led to the subsequent creation of microarrays that allow for better understanding of genes and gene expression across all eight chromosomes (Abbas, Wilkinson et al. 2009; Cleveland, Yu et al. 2009).

Genomic research is accelerating what is known about potential mechanisms for the regulation of extrolite, e.g. aflatoxin, biosynthesis. Regulation of AF production occurs internally with transcriptional regulators *aflR* and *aflS*, and externally with global regulators of secondary metabolism such as *laeA* and *veA* (Amaike and Keller 2009; Georgianna and Payne 2009). The protein encoded by the *aflR* gene is considered integral for transcriptional activation of the genes with which it is clustered. In the older nomenclature the *aflS* gene was known as *aflJ*, and is reported to influence AF production through interaction with *aflR* (Chang 2003; Yu, Chang et al. 2004). An important aspect of genomic studies is gene "knock out" which allows for understanding the importance of a gene's function or role in an organism by removing a particular gene and noting any phenotypic changes that may result (Pickford 2005). Nancy Keller's lab has shown that deletion of *laeA* results in the inability to produce AF, and they hypothesized that *laeA* is a global regulator of secondary metabolism

after showing loss of multiple extrolites in *A. nidulans* (Bok and Keller 2004). The light-activated *veA* gene, important in sclerotial formation, was reported to regulate the production of CPA, aflatrem, and AF in response to light (Kato, Brooks et al. 2003).

2.3 Biological Control of Aflatoxigenic *Aspergillus*

Biological control involves use of naturally-occurring antagonists to control pests with no need for chemicals that may further pollute the environment (DeBach and Rosen 1991). For biological control of AF-producing Aspergilli, studies have been implemented such as introducing nonalfatoxigenic (AF-) strains into fields to out-compete the toxin producers (Dorner 2004). Two *A. flavus* strains are currently registered through the Environmental Protection Agency as pre-harvest biopesticides: NRRL 18543 or AF36, and NRRL 21882 (active component in afla-guard®) (EPA 2003; EPA 2004), and it has been reported that these AF- individuals are more aggressive colonizers than those that are aflatoxigenic (Horn and Dorner 2009). AF36 was discovered in a cotton field in Arizona by Peter Cotty, and is currently used in and around Arizona on a variety of crops (Cotty and Bhatnagar 1994). A nonsense mutation in the *pksA* (*aflC*) gene renders AF36 unable to synthesize AF (Ehrlich and Cotty 2004). The application of AF36 involves colonized wheat seeds that have been sterilized, and the rate of application is 10 lbs/acre once a year (EPA 2003). The NRRL 21882 strain was isolated from a Georgia peanut field by Joe Dorner, and it is used exclusively on peanuts in Georgia (Dorner 2004). This strain lacks the entire AF gene cluster, which prevents synthesis of AF (Chang, Horn et al. 2005). The application of NRRL 21882 involves colonized hulled barley that is applied at a rate of 20 lbs/acre once a year (EPA 2004).

Fields are currently being inundated with these biocontrol strains, and the basis for their success depends on them being asexual and incapable of recombining with the indigenous populations of toxin-producers. In 1997, Milgroom and Fry reported that understanding the ways by which these fungi reproduce is integral to disease management (Milgroom 1997), so research is now looking into population structure to observe evidence of recombination among species of *Aspergillus*. Population genetic studies are important for understanding an organism's biology through genetic diversity resulting from recombination (Pildain, Vaamonde et al. 2004; Rydholm, Szakacs et al. 2006). David Geiser was one of the first researchers to examine recombination in Aspergilli and report that recombination could impact toxicity of the biocontrol strains (Geiser, Pitt et al. 1998; Geiser, Dorner et al. 2000), and in a clinical setting recombination could thwart efforts to control diseases caused by species of *Aspergillus* (Geiser, Klich et al. 2007).

2.4 Population Studies and Recombination in Aflatoxigenic *Aspergillus*

Recombination seeks to generate genetic variation through the sharing of traits that enhance fitness and the removal of mutations that hinder fitness (Pál, van Diepeningen et al. 2008). Factors that might hinder, or promote, recombination and population diversity include: geography, ecology, genetic compatibility systems, and balancing selection. The influence of geography relates to isolation of populations and can only be broken through the migration of individuals that may introduce novel alleles/phenotypes (Church and Taylor 2002). Ecological influences for Aspergilli are based on niche, temperature, and moisture (Abbas, Wilkinson et al. 2009). Genetic compatibility or self/non-self recognition encompasses

vegetative compatibility groups (VCGs) and mating type association (Micali and Smith 2005). VCG testing involves heterokaryon incompatibility that results in the prevention of hyphal anastomosis between individuals exhibiting allelic differences. The gene types involved in vegetative/heterokaryon [in]compatibility are: *het* genes involved with physical (hyphal) interactions, and *vic* genes involved with non-physical (chemical) interactions (Micali and Smith 2005; van Dijck 2008). Studies in *Saccharomyces cerevisiae* and *Podospora anserina* suggest that prions play a part in vegetative compatibility (Wickner, Taylor et al. 1999; Saupe, Clave et al. 2000; Dos Reis, Coulary-Salin et al. 2002), but to date no research has shown that prions exist in Aspergilli. In *Neurospora crassa*, mating type association disregards heterokaryon incompatibility since the idiomorphs do have allelic difference, but mating cannot occur without individuals of opposite mating type (Micali and Smith 2005). Balancing selection is a form of natural selection seeking to maintain genetic polymorphisms in a population, and it is evident when increased levels of genetic variation exist between alleles or haplotypes within a species (Hedrick 2007). According to the literature, evidence of balancing selection is observable in mating-type loci (May, Shaw et al. 1999), trichothecene chemotype evolution in species of *Fusarium* (Ward, Bielawski et al. 2002), and the major histocompatibility complex in humans (Takahata, Satta et al. 1992).

3. Agricultural Significance of Section *Flavi*

Section *Flavi* is in the subgenus *Circumdati* and currently includes over 20 species (Table I.3) of anamorphs and at least four teleomorphs (Rigo, Varga et al. 2002; Peterson, Varga et al. 2008; Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009). This section is named

for its most infamous species, *A. flavus* (Raper and Fennell 1965; Horn 2003), which is capable of occupying many niches and is capable of pathogenic associations with many hosts (Cleveland, Yu et al. 2009; Krishnan, Manavathu et al. 2009). The first reported teleomorph from section *Flavi* was *Petromyces alliaceus* (an.= *A. alliaceus*), characterized in the 1970s (Malloch and Cain 1972). It was not until 2008 that other teleomorphs were found within section *Flavi* for *A. flavus* (*P. flavus*) and *A. parasiticus* (*P. parasiticus*) by Bruce Horn in collaboration with Ignazio Carbone's lab (Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009). It appears species in section *Flavi* exhibit a heterothallic mating system with two *MAT* idiomorphs (*MAT1-1* and *MAT1-2*), whereby *MAT1-1* encodes an alpha-box motif and *MAT1-2* encodes a high mobility group (HMG) gene (Ramirez-Prado, Moore et al. 2008; Horn, Ramirez-Prado et al. 2009).

Table I.3 Species that comprise section *Flavi* and mycotoxins, if any, produced

Anamorphs	*Petromyces* Teleomorphs	Mycotoxins
A. albertensis	*P. albertensis*	Ochratoxins (A,B)
A. alliaceus	*P. alliaceus*	Ochratoxins (A,B)
A. arachidicola		Aflatoxins (B,G), Kojic acid
A. avenaceus		
A. bombycis		Aflatoxin (B1)
A. caelatus		
A. clavatoflavus		
A. coremiiformis		
A. flavofurcatis		
A. flavus	*P. flavus*	Aflatoxins (B,G,M), Aflatrem, Cyclopiazonic acid, O-methyl-sterigmatocystin, Sterigmatocystin, Kojic acid
A. lanosus		Ochratoxin A
A. leporis		
A. minisclerotigenes		Aflavtoxins (B,G), Aflatrem, Cyclopiazonic acid, Kojic acid
A. nomius	*P. nomius* [a]	Aflatoxins (B,G)
A. oryzae		Citrinin, Cyclopiazonic acid, Kojic acid
A. parasiticus	*P. parasiticus*	Aflatoxins (B,G,M), O-methyl-sterigmatocystin, Sterigmatocystin
A. parvisclerotigenus		Aflatoxins (B), Cyclopiazonic acid
A. pseudotamarii		Aflatoxin (B1), Cyclopiazonic acid
A. sojae		
A. subolivaceus		
A. tamarii		Fumigaclavine A
A. terricola		
A. thomii		
A. toxicarius		
A. zonatus		

List compiled from various sources (Malloch and Cain 1972; Klich 2002; Rigo, Varga et al. 2002; Frisvad, Skouboe et al. 2005; Geiser, Klich et al. 2007; Pildain, Frisvad et al. 2008; Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009).

[a] *P. nomius* research results in draft phase by Bruce Horn and therefore unpublished.

Species within section *Flavi* are predominantly soil-inhabiting saprophytes, but *A. flavus* is fairly common being found in non-soil environments as well (Horn 2003; Abbas, Wilkinson et al. 2009). Interestingly, most of the *Aspergillus* species that produce aflatoxins are from section *Flavi* (Horn 2003; Carbone, Ramirez-Prado et al. 2007).

Two important species in section *Flavi* are *A. flavus* Link and *A. parasiticus* Speare (Abbas, Wilkinson et al. 2009). Both induce rot of important agricultural commodities such as corn, cotton and peanuts, both prefer dry, hot weather conditions, and both are often found together in soils (Blankenship, Cole et al. 1984). But even more serious than their prolific colonization of food crops is that *A. flavus* and *A. parasiticus* produce mycotoxins (Scheidegger and Payne 2003; Klich 2007; Horn and Dorner 2009). *A. flavus* has two recognized morphotypes that are differentiated based on sclerotial size and mycotoxins produced. L- (large) strain *A. flavus* forms sclerotia greater than 400 µm in diameter and may produce B AFs and CPA, whereas, S- (small) strain *A. flavus* forms high quantities of sclerotia less than 400 µm in diameter that may produce B AFs, G AFs, and CPA (Horn 2003). Some researchers refer to the *A. flavus* S-strain as species *A. parvisclerotigenus* if they are flavus-like and only produce B AFs (Frisvad, Skouboe et al. 2005), or *A. minisclerotigenes* if they are parasiticus-like and produce both B and G AFs (Pildain, Frisvad et al. 2008). *A. flavus* L-strains (and some S-strains) fail to produce G AFs due to defects in, or complete absence of, a required gene known as *cypA* that encodes cytochrome P-450 (Ehrlich, Chang et al. 2004). Also, it is reported that in a nature one could observe up to 40% of an *A. flavus* L-strain population as AF- (Ehrlich 2008), because of defective or absent genes that are integral to AF synthesis (Chang, Horn et al. 2005; Yin, Lou et al. 2009). *A. parasiticus* is capable of synthesizing both B and G AFs, and OMST (Carbone, Ramirez-Prado et al. 2007). Reports of atoxigenic *A. parasiticus* or *A. flavus* S-strains are rare (Ehrlich 2008). Both *A. flavus* L-strains and *A. parasiticus* are capable of reproducing asexually, parasexually (Papa 1973; Papa 1978), or sexually as heterothallic

(Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009; Horn, Ramirez-Prado et al. 2009). There is evidence to suggest that the S-strain morphotype is actually a hybrid between *A. flavus* and *A. parasiticus*, and this hypothesis is currently being explored further in the Carbone lab (C Worthington, unpublished data). Hybridization is considered to promote evolutionary advantage, termed "hybrid vigor", due to the increased ability of hybrids to adapt to niche changes (Dodge 1942). It exists widely in the plant kingdom, and is gaining attention within pathogenic fungi like *Cryptococcus neoformans* (Lin, Nielsen et al. 2008). Successful inter-specific crosses for Aspergilli could challenge us to rethink how we classify species within this genus.

4. Current Research Challenges and Considerations for *Aspergillus*

It seems as we gain more knowledge about *Aspergillus* more questions arise. What makes one isolate more fertile than another? Are Aspergilli truly haploid? How can the same isolate test both AF+ and AF-? Research should further explore fertility issues, ploidy, and nuclear conditions that give rise to different chemotypes for the same isolate. Exploring these frontiers of scientific research may allow us to better understand these fungi.

4.1 Variable fertility

For a heterothallic system, one might expect that simply introducing an isolate to another of complementary mating type would be enough to facilitate a successful cross. But quite to the contrary, mating for species of *Aspergillus* is hit or miss. In the laboratory, certain isolates exhibit greater fecundity than others despite using the same substrate and environmental conditions (Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009; Horn, Ramirez-Prado et al. 2009).

Possible reasons for infertility could be mutations in either *MAT* idiomorph, false complementation by an amplified but silent *MAT* allele or minority allele due to difference in the number of nuclei in a cell, or defective signaling pathways. In nature we have no evidence of sex, much less the possible conditions that may promote sex, in these fungi. It has been hypothesized that fertility for *Aspergillus* is on the decline from an evolutionary standpoint (Kwon-Chung and Sugui 2009). The high incidences of clonal expansion in populations of *Aspergillus* certainly offer support to this hypothesis. These ideas should be explored further since they could contribute to the prevention of advantageous recombination in Aspergilli.

4.2 Ploidy in *Aspergillus*

Aspergillus has long been considered a haploid fungus meaning that its nuclei contain a single set of chromosomes. Diploid conditions have been observed in the teleomorphic Aspergilli, with heterokaryosis, which is considered unstable causing the fungus to revert back to a more stable haploid state (Käfer 1961), but stable diploids have been observed in *Aspergillus* (Upshall 1981). Ploidy has even been shown to affect virulence in *A. nidulans* (Purnell and Martin 1973). Evidence of diploid condition may be observed in *A. flavus* DNA sequences with double peaks at polymorphic sites, though it could also arise from a dikaryon (I Carbone, unpublished data). Aneuploids, having an abnormal chromosome number, have been observed in filamentous fungi such as *Penicillium* and *Aspergillus* and are reportedly evident through unusual development or phenotypic variation (Bridge, Hudson et al. 1987). In *Aspergillus*, individuals from a single strain may exhibit different vegetative growth,

colony color, and even metabolite production, believed to be the result of aneuploidy (Schroeder 1969; Upshall 1971). This may be the reason an *A. flavus* isolate could produce AF in one test, but be AF- in another. Kistler and Miao hypothesized that incidence of chromosomal aberrations, such as aneuploidy, increases in fungi that undergo little meiosis in their life cycles. Interestingly, meiosis is considered integral to the removal of traits that negatively affect fitness (Kistler and Miao 1992). Aneuploidy has also been reported to result from fungicide use against *Candida albicans*, increasing its fitness through aberrant chromosome conditions (Selmecki, Dulmage et al. 2009).

4.3 Multinucleate *Aspergillus*

In addition to variability in chromosome number is the possibility that a spore may contain more than one nucleus; anywhere from one to twelve nuclei have been observed in a single *Aspergillus* spore, the larger the spore the more likely it will be multinucleate (Raper and Fennell 1965; Foudin, Papa et al. 1981). In *A. flavus* conidia, one to four nuclei have been observed, and in a single hyphal cell 10-15 nuclei have been observed (I Carbone, unpublished data). It has been reported that in *Podospora anserina*, and potentially other filamentous ascomycetes, there is potential to retain distinct nuclear identity during mating, when the nuclei of both parents exist in the same cytoplasm, resulting in uniparental or biparental progeny (Nelson 1996). Perhaps this is why certain isolates can exhibit both mating type genes, or why an isolate can exhibit both an AF+ and an AF- phenotype. This would mean that a dominant nucleus could express and another nucleus would be silent in the cytoplasm upon germination. The act of silencing or suppressing genes in fungi is

referred to as quelling, a term first coined through studies of the phenomenon in *Neurospora crassa*, during the vegetative phase of a fungus, or meiotic silencing which occurs during the sexual phase. The underlying mechanism of quelling in *N. crassa* was reported to be post-transcriptional gene silencing through RNA interference, and meiotic silencing involves inhibition of gene expression only during prophase I of meiosis for any unpaired homologue (Shiu, Raju et al. 2001; Raju 2009). For the AF cluster, a study involved attempts to ectopically insert a functional cluster into an *A. flavus* isolate which lacked the cluster, but the nonaflatoxigenic phenotype remained unchanged for the isolate lacking the AF cluster (Smith, Woloshuk et al. 2007). Also possible is that asynchronous division of nuclei could have them residing in the same cytoplasm yet exhibiting independent behavior. In nutrient-deprived media *A. nidulans* has been reported to undergo asynchronous nuclear division (Gladfelter, Hungerbuehler et al. 2006). There may be implications of heterokaryosis in a multinucleate species such as *Aspergillus*. Studies have shown, in colonies of *A. nidulans*, that parental strains exhibiting different conidial colors may form a heterokaryon of yet a third color. Upon culturing the conidia of this heterokaryon a sector may eventually form that shows haploid conidia of one parental phenotype, and diploid conidia of the other parental phenotype (Garber, Bryan et al. 1961). *A. nidulans* is uninucleate so imagine the possible heterokaryotic phenotypes possible in multinucleate species like *A. flavus*. Isolates having identical nuclear condition would be observed with single peaks along nucleotide sequences of AF cluster genes. Evidence that a few isolates may be heterokaryotic exists because, at polymorphic sites, there are two possible peaks. Though one peak may be

stronger than the other, at least one peak agrees with the consensus and phenomena such as this could prove the existence of multiple genomes in a single individual (I Carbone, unpublished data).

4.4 Effect of Serial Culturing on *Aspergillus*

It has been reported that mycotoxin production has been halted in fungal species that have undergone serial culturing. Other than *Aspergillus*, this phenomenon has been observed in *Claviceps* spp. and *Fusarium* spp., and one proposed hypothesis is based on lack of competition though no real evidence exists to support this (Horn and Dorner 2002). This is yet another potential explanation for an isolate being an AF producer and non-producer at different times.

4.5 Nomenclatural Challenges with *Aspergillus*

Nomenclature is another challenge the *Aspergillus* research community may have to face. Historically, dual nomenclature for holomorphic *Aspergillus* had been ignored being that more species were asexual (Raper and Fennell 1965; Pitt and Samson 2007). With the increasing number of teleomorph associations, it may not take long to uncover the sexuality of most, if not all, of the Aspergilli. Scientists should not ignore the possibility that the abundance of "species" in this genus may all be hybrid relations of each other (I Carbone, unpublished data). It is possible that, at least for genus *Aspergillus*, the definition of species will require more effort and time to refine (Bennett 2010). Figure 1 in (Samson and Varga 2009) lists the multi-facet approach to naming new species in *Aspergillus* which involves at least seven considerations. Morphology is not a true delineator of species (Harrington and Rizzo 1999).

Humans, for example, are all *Homo sapiens*, and yet we exhibit morphological differences based on many different environmental and genetic factors (Hendry 2009). According to the biological species concept, a true species is reproductively limited from another species, or at the extreme, non-reproductive with another species (Hendry 2009). Researchers are challenging this concept with inter-specific mating between *Aspergillus* "species". Early success with inter-specific crosses has been shown between *A. flavus* and *A. parasiticus* (C Worthington, unpublished data). There are other fungal concepts of species based on ecology, cohesion, monophyly, genealogy, and phylogeny, but Harrington and Rizzo choose to define species, based on a modification of a definition put forth by Nixon and Wheeler in 1990, as "… the smallest aggregation of populations with a common lineage that share unique, diagnosable phenotypic characters" (Nixon and Wheeler 1990; Harrington and Rizzo 1999). At the core of the species concept is the idea of an "individual". Researchers have no real answer for what constitutes an individual in fungi. Perhaps it is a single nucleus in a cloned cell, or a dikaryotic cell with nuclei from the parents (Burnett 2003). Can a multinucleate fungal cell containing multiple genomes be considered an individual, especially if an individual is capable of exhibiting distinct phenotypic variation?

5.0 *Aspergillus* Biocontrol Revisited

In order for an agent of biological control to be effective there must be a thorough investigation of the agent's properties, both positive and negative (DeBach and Rosen 1991). Granted, the knowledge of sex in agriculturally important species came after the commercial distribution of biocontrol strains, but the idea now should be to re-evaluate their stability

given their potential for sex. There is evidence that in a single generation of sexual recombination, between a toxin-producer and either biocontrol strain, that the progeny can exhibit the toxic phenotype (R Olarte, unpublished data). No known studies have been performed which look at the soil densities of *Aspergillus* populations in fields where the biocontrol strains have been utilized. Though methods to detect nonaflatoxigenic isolates are being tested (Horn and Dorner 2009), the persistence of the biocontrol, or nonaflatoxigenic phenotype, should be considered in light of genetic recombination and sex among these fungi.

Pre-harvest biological controls are the preferred methods, so research involving AF-strains, plant breeding, and transgenic approaches continues strongly (Wu, Liu et al. 2008; Rajasekaran, De Lucca et al. 2009). As well, studies of other biological means of control involve inhibition of *Aspergillus* growth during storage with antagonistic bacteria, fungi and yeasts; detoxification of foods through enzymatic degradation by bacteria and fungi; and sorbent compounds to bind toxins, thereby preventing them from moving out of the gastrointestinal tract (Kabak and Dobson 2009). Studies using bacteria and essential oils (Reddy, Abbas et al. 2009), offer the existence of other means to control/prevent potential AF contamination without the use of chemicals. Some researchers have tried, with mixed success, to prevent AF production by introducing mycovirues into toxigenic fungi (van Diepeningen, Varga et al. 2008; Schmidt 2009). The approach to harnessing the beneficial applications of species within *Aspergillus*, while effectively suppressing the negative, is going to require the continued contributions of various disciplines of fungal and agricultural research (Andersen and Nielsen 2009; Cleveland, Yu et al. 2009).

With regard to the *A. flavus* biocontrol strain NRRL 21882 not having an AF cluster, or any partial cluster *A. flavus* isolate (Chang, Horn et al. 2005), we may infer that the genes are lost from the cluster but not necessarily from the genome (Moore, Singh et al. 2009). Random cluster genes have been amplified and sequenced that should be absent, if the genome is in fact smaller, for isolates lacking the cluster genes (I Carbone, unpublished data). A helpful task would be to sequence the genome for NRRL 21882 and compare it to NRRL 3357, which has the full AF gene cluster. If the genomes are similar in size then it may speak for translocation of cluster genes from the AF cluster (Hodges, Kelkar et al. 2000); however, recent microarray work offers evidence that the genes are not translocated suggesting they are found in other nuclei as minority copies (I Carbone, unpublished data).

6. Research Objectives

Three research objectives will be explored in this document. The first objective is to look at molecular variation in a single *A. flavus* population using a multilocus approach (Moore, Singh et al. 2009), and compare to previous work on an *A. parasiticus* population sampled from the same field (Carbone, Jakobek et al. 2007). To expand upon this objective is to seek evidence of recombination, balancing selection, variation in cluster size, and rates of recombination for each soil co-inhabiting species. The second research objective involves examination of molecular variation for worldwide, geographically-isolated populations of *A. flavus* and *A. parasiticus* using a multilocus approach, identifying informative loci for detection of recombination, and comparing recombination rates on a global scale. Also explored will be the distribution of mating types in each geographic population sample of *A. flavus* and *A. parasiticus*, and the respective distribution of aflatoxin concentrations for each

isolate. The third objective further explores section *Flavi* within a phylogeographic context. There are other species in section *Flavi* that produce mycotoxins, and warrant further investigations, such as *A. nomius*. The goal is to quantify the magnitude of recombination, selection, subdivision, and migration among worldwide populations. The loci of interest will include both cluster and non-cluster loci, and will represent more than one chromosomal location. Also investigated will be multi-locus sequence typing for VCG clone-correction to replace traditional auxotrophic mutant studies, and to look into ecological influences on these populations based upon cultivar, soil type, temperature, and precipitation.

7. Concluding Remarks

This project is based on natural populations that may be undergoing recombination. The evidence may present itself, but the timescale is still unknown. Experimental crosses will further explore the occurrence of recombination in a single generation. All the research in this document is seeking to strengthen supportive evidence that identifying and developing effective biocontrol strains should be rigorous and tailored to local population conditions.

LITERATURE CITED

Abbas, H. K., J. R. Wilkinson, et al. (2009). "Ecology of *Aspergillus flavus*, regulation of aflatoxin production, and management strategies to reduce aflatoxin contamination of corn." Toxin Reviews **28**(2-3): 142-153.

Alexopoulos, C. J., C. W. Mims, et al. (1996). Introductory mycology. New York, Wiley.

Amaike, S. and N. P. Keller (2009). "Distinct roles for VeA and LaeA in development and pathogenesis of *Aspergillus flavus*." Eukaryotic Cell: EC.00088-09.

Andersen, M. R. and J. Nielsen (2009). "Current status of systems biology in Aspergilli." Fungal Genetics and Biology **46**: S180-S190.

Austwick, P. K. C. (1965). Pathogenicity. The genus Aspergillus. K. B. Raper and D. I. Fennell. Baltimore, Williams & Wilkins**:** 82-126.

Baker, S. E. and J. W. Bennett (2008). An Overview of the Genus Aspergillus. The Aspergilli : genomics, medical aspects, biotechnology, and research methods. Mycology series. S. A. Osmani and G. H. Goldman. Boca Raton, CRC Press**:** 3-14.

Balajee, S. A. (2008). DNA sequence based methods for species identification in the genus Aspergillus. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 261-275.

Beard, R. L. and G. S. Walton (1969). "Kojic Acid as an Insecticidal Mycotoxin." Journal of Invertebrate Pathology **14**(1): 53-&.

Bennett, J. W. (2010). An Overview of the Genus Aspergillus. Aspergillus : molecular biology and genomics. M. Machida and K. Gomi. Wymondham, Caister Academic**:** 1-17.

Bhatnagar, D., S. P. McCormick, et al. (1987). "Identification of O-Methylsterigmatocystin as an Aflatoxin-B1 and Aflatoxin-G1 Precursor in *Aspergillus-Parasiticus*." Applied and Environmental Microbiology **53**(5): 1028-1033.

Blankenship, P. D., R. J. Cole, et al. (1984). "Effect of Geocarposphere Temperature on Pre-Harvest Colonization of Drought-Stressed Peanuts by *Aspergillus-Flavus* and Subsequent Aflatoxin Contamination." Mycopathologia **85**(1-2): 69-74.

Blount, W. P. (1961). "Turkey "X" Disease." Journal of the British Turkey Federation **1**: 52,55,61.

Blumenthal, C. Z. (2004). "Production of toxic metabolites in *Aspergillus niger*, *Aspergillus oryzae*, and *Trichoderma reesei*: justification of mycotoxin testing in food grade enzyme preparations derived from the three fungi." Regulatory Toxicology and Pharmacology **39**(2): 214-228.

Bok, J. W. and N. P. Keller (2004). "LaeA, a regulator of secondary metabolism in Aspergillus spp." Eukaryot Cell **3**(2): 527-35.

Bridge, P. D., L. Hudson, et al. (1987). "Investigation of Variation in Phenotype and DNA Content Between Single-Conidium Isolates of Single Penicillium Strains." Journal of General Microbiology **133**: 995-1004.

Burnett, J. H. (2003). Fungal populations and species. New York, Oxford University Press.

Cabañes, F. J. and M. R. Bragulat (2008). Ochratoxin A in profiling and speciation. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 57-70.

Calvo, A. M., R. A. Wilson, et al. (2002). "Relationship between secondary metabolism and fungal development." Microbiology and Molecular Biology Reviews **66**(3): 447-+.

Carbone, I., J. L. Jakobek, et al. (2007). "Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus*." Mol Ecol **16**(20): 4401-17.

Carbone, I., J. H. Ramirez-Prado, et al. (2007). "Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster." BMC Evol Biol **7**: 111.

Cary, J. W. and K. C. Ehrlich (2006). "Aflatoxigenicity in Aspergillus: molecular genetics, phylogenetic relationships and evolutionary implications." Mycopathologia **162**(3): 167-177.

Cary, J. W., K. C. Ehrlich, et al. (2009). "Characterization of the *Aspergillus ochraceoroseus* aflatoxin/sterigmatocystin biosynthetic gene cluster." Mycologia **101**(3): 352-362.

CAST (2003). Mycotoxins: Risks in Plant, Animal, and Human Systems. Report 139. Ames, IA, Council for Agricultural Science and Technology.

CDC. (2010). "Health Studies: Aflatoxin." Environmental Haxards and Health Effects Retrieved February 27, 2010, from http://www.cdc.gov/nceh/hsb/aflatoxin/.

Chang, P. K. (2003). "The *Aspergillus parasiticus* protein AFLJ interacts with the aflatoxin pathway-specific regulator AFLR." Mol Genet Genomics **268**(6): 711-9.

Chang, P. K., B. W. Horn, et al. (2005). "Sequence breakpoints in the aflatoxin biosynthesis gene cluster and flanking regions in nonaflatoxigenic *Aspergillus flavus* isolates." Fungal Genetics and Biology **42**(11): 914-923.

Chang, P. K., B. W. Horn, et al. (2009). "Clustered genes involved in cyclopiazonic acid production are next to the aflatoxin biosynthesis gene cluster in *Aspergillus flavus*." Fungal Genet Biol **46**(2): 176-82.

Chang, P. K., C. D. Skory, et al. (1992). "Cloning of a Gene Associated with Aflatoxin-B1 Biosynthesis in *Aspergillus-Parasiticus*." Current Genetics **21**(3): 231-233.

Church, S. A. and D. R. Taylor (2002). "The evolution of reproductive isolation in spatially structured populations." Evolution **56**(9): 1859-62.

Cleveland, T. E., J. J. Yu, et al. (2009). "Potential of *Aspergillus flavus* genomics for applications in biotechnology." Trends in Biotechnology **27**(3): 151-157.

Cotty, P. J. and D. Bhatnagar (1994). "Variability among Atoxigenic *Aspergillus-Flavus* Strains in Ability to Prevent Aflatoxin Contamination and Production of Aflatoxin Biosynthetic-Pathway Enzymes." Applied and Environmental Microbiology **60**(7): 2248-2251.

Cullen, J. M. and P. M. Newberne (1994). Acute hepatotoxicity of aflatoxins. The Toxicology of aflatoxins : human health, veterinary, and agricultural significance. D. L. Eaton and J. D. Groopman. San Diego, Academic Press**:** 3-26.

de Vries, R. P. (2008). A genomic look at physiology and extracellular enzymes of Aspergillus in relation to utilization of plant matter. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 87-106.

DeBach, P. and D. Rosen (1991). Biological control by natural enemies. Cambridge, [England] ; New York, Cambridge University Press.

Dodge, B. O. (1942). "Heterocaryotic vigor in Neurospora." Bull. Torrey Botan. Club **69**: 75-91.

Dorner, J. W. (2004). "Biological Control of Aflatoxin Contamination of Crops." Journal of Toxicology -- Toxin Reviews **23**(2/3): 425-450.

Dos Reis, S., B. Coulary-Salin, et al. (2002). "The HET-s prion protein of the filamentous fungus *Podospora anserina* aggregates in vitro into amyloid-like fibrils." J Biol Chem **277**(8): 5703-6.

Ehrlich, K. (2006). "Evolution of the aflatoxin gene cluster." Mycotoxin Research **22**(1): 9-15.

Ehrlich, K. C. (2008). Genetic diversity in *Aspergillus flavus* and its implications for agriculture. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 233-247.

Ehrlich, K. C., P. K. Chang, et al. (2004). "Aflatoxin biosynthesis cluster gene cypA is required for G aflatoxin formation." Appl Environ Microbiol **70**(11): 6518-24.

Ehrlich, K. C. and P. J. Cotty (2004). "An isolate of *Aspergillus flavus* used to reduce aflatoxin contamination in cottonseed has a defective polyketide synthase gene." Applied Microbiology and Biotechnology **65**(4): 473-478.

El-Sayed, Y. S. and R. H. Khalil (2009). "Toxicity, biochemical effects and residue of aflatoxin B-1 in marine water-reared sea bass (*Dicentrarchus labrax* L.)." Food and Chemical Toxicology **47**(7): 1606-1609.

EPA. (2003, November 25, 2009). "Aspergillus flavus AF36 Biopesticides Registration Action Document." Biopesticides Registration Action Document Retrieved 26 February, 2010, from http://epa.gov/opp00001/biopesticides/ingredients/tech_docs/brad_006456.pdf.

EPA. (2003, October 15, 2008). "*Aspergillus flavus* strain AF36 (006456) Fact Sheet." Retrieved 21 November, 2009, from http://www.epa.gov/pesticides/biopesticides/ingredients/factsheets/factsheet_006456.htm.

EPA. (2004, October 15, 2008). "*Aspergillus flavus* NRRL 21882 (006500) Fact Sheet." Retrieved 21 November, 2009, from http://www.epa.gov/pesticides/biopesticides/ingredients/factsheets/factsheet_006500.htm.

EPA. (2004, November 25, 2009). "Aspergillus flavus NRRL 21882 Biopesticides Registration Action Document." Retrieved 26 February, 2010, from http://www.epa.gov/pesticides/biopesticides/ingredients/tech_docs/brad_006500.pdf.

Foudin, L. L., K. E. Papa, et al. (1981). "Nuclear Behavior During Conidiogenesis in *Aspergillus flavus*." Canadian Journal of Botany **59**(11): 2116-2120.

Fox, E. M. and B. J. Howlett (2008). "Secondary metabolism: regulation and role in fungal biology." Current Opinion in Microbiology **11**(6): 481-487.

Frisvad, J. C., P. Skouboe, et al. (2005). "Taxonomic comparison of three different groups of aflatoxin producers and a new efficient producer of aflatoxin B-1, sterigmatocystin and 3-O-methylsterigmatocystin, *Aspergillus rambellii* sp nov." Systematic and Applied Microbiology **28**(5): 442-453.

Gams, W., M. Christensen, et al. (1985). Infrageneric taxa of Aspergillus. Advances in penicillium and aspergillus systematics. NATO ASI series. Series A, Life sciences. R. A. Samson, J. I. Pitt and D. North Atlantic Treaty Organization. Scientific Affairs. New York, Plenum Press. **102:** 55-62.

Garber, E. D., G. W. Bryan, et al. (1961). "Evidence for Parthenogenesis in *Aspergillus nidulans*." The American Naturalist **95**(884): 309-313.

Geiser, D. M. (2009). "Sexual structures in Aspergillus: morphology, importance and genomics." Med Mycol **47 Suppl 1**: S21-6.

Geiser, D. M., J. W. Dorner, et al. (2000). "The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*." Fungal Genet Biol **31**(3): 169-79.

Geiser, D. M., C. Gueidan, et al. (2006). "Eurotiomycetes: Eurotiomycetidae and Chaetothyriomycetidae." Mycologia **98**(6): 1053-64.

Geiser, D. M., M. A. Klich, et al. (2007). "The current status of species recognition and identification in Aspergillus." Studies in Mycology(59): 1-10.

Geiser, D. M., J. I. Pitt, et al. (1998). "Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*." Proc Natl Acad Sci U S A **95**(1): 388-93.

Geiser, D. M., R. A. Samson, et al. (2008). A review of molecular phylogenetics in Aspergillus, and prospects for a robust genus-wide phylogeny. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 17-32.

Geiser, D. M., W. E. Timberlake, et al. (1996). "Loss of meiosis in Aspergillus." Molecular Biology and Evolution **13**(6): 809-817.

Georgianna, D. R. and G. A. Payne (2009). "Genetic regulation of aflatoxin biosynthesis: From gene to genome." Fungal Genetics and Biology **46**(2): 113-125.

Gladfelter, A. S., A. K. Hungerbuehler, et al. (2006). "Asynchronous nuclear division cycles in multinucleated cells." J Cell Biol **172**(3): 347-62.

Harrington, T. C. and D. M. Rizzo (1999). Defining Species in the Fungi. Structure and dynamics of fungal populations. Population and community biology series. J. J. Worrall. Dordrecht ; Boston, Kluwer Academic Publishers**:** 43-71.

Hedrick, P. W. (2007). "Balancing selection." Current Biology **17**(7): R230-R231.

Hendry, A. P. (2009). "Evolutionary biology: Speciation." Nature **458**(7235): 162-164.

Hodges, R. L., H. S. Kelkar, et al. (2000). "Characterization of an echinocandin B-producing strain blocked for sterigmatocystin biosynthesis reveals a translocation in the stcW gene of the aflatoxin biosynthetic pathway." Journal of Industrial Microbiology & Biotechnology **25**(6): 333-341.

Horn, B. W. (2003). "Ecology and population biology of aflatoxigenic fungi in soil." Journal of Toxicology-Toxin Reviews **22**(2-3): 351-379.

Horn, B. W. and J. W. Dorner (2002). "Effect of competition and adverse culture conditions on aflatoxin production by *Aspergillus flavus* through successive generations." Mycologia **94**(5): 741-751.

Horn, B. W. and J. W. Dorner (2009). "Effect of nontoxigenic *Aspergillus flavus* and *A. parasiticus* on aflatoxin contamination of wounded peanut seeds inoculated with agricultural soil containing natural fungal populations." Biocontrol Science and Technology **19**(3): 249-262.

Horn, B. W., G. G. Moore, et al. (2009). "Sexual reproduction in *Aspergillus flavus*." Mycologia **101**(3): 423-9.

Horn, B. W., J. H. Ramirez-Prado, et al. (2009). "Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus parasiticus*." Fungal Genet Biol **46**(2): 169-75.

Horn, B. W., J. H. Ramirez-Prado, et al. (2009). "The sexual state of *Aspergillus parasiticus*." Mycologia **101**(2): 275-80.

Kabak, B. and A. D. W. Dobson (2009). "Biological Strategies To Counteract the Effects of Mycotoxins." Journal of Food Protection **72**(9): 2006-2016.

Käfer, E. (1961). "The Processes of Spontaneous Recombination in Vegetative Nuclei of *Aspergillus nidulans*." Genetics **46**(12): 1581-1609.

Kato, N., W. Brooks, et al. (2003). "The expression of sterigmatocystin and penicillin genes in *Aspergillus nidulans* is controlled by veA, a gene required for sexual development." Eukaryot Cell **2**(6): 1178-86.

Keller, N. P., N. J. Kantz, et al. (1994). "*Aspergillus-Nidulans* Vera Is Required for Production of the Mycotoxin Sterigmatocystin." Applied and Environmental Microbiology **60**(5): 1444-1450.

Kistler, H. C. and V. P. Miao (1992). "New modes of genetic change in filamentous fungi." Annu Rev Phytopathol **30**: 131-53.

Klich, M. A. (2002). Identification of common Aspergillus species. Utrecht, Netherlands, Centraalbureau voor Schimmelcultures.

Klich, M. A. (2007). "*Aspergillus flavus*: the major producer of aflatoxin." Molecular Plant Pathology **8**(6): 713-722.

Krishnan, S., E. K. Manavathu, et al. (2009). "*Aspergillus flavus*: an emerging non-fumigatus Aspergillus species of significance." Mycoses **52**(3): 206-222.

Kwon-Chung, K. J. and J. A. Sugui (2009). "Sexual Reproduction in Aspergillus species of medical or economical importance: why so fastidious?" Trends in Microbiology **17**(11): 481-487.

Latgé, J. P. (1999). "*Aspergillus fumigatus* and aspergillosis." Clinical Microbiology Reviews **12**(2): 310-350.

Lin, X., K. Nielsen, et al. (2008). "Impact of Mating Type, Serotype, and Ploidy on the Virulence of *Cryptococcus neoformans*." Infect. Immun. **76**(7): 2923-2938.

Machida, M., O. Yamada, et al. (2008). "Genomics of *Aspergillus oryzae*: Learning from the History of Koji Mold and Exploration of Its Future." DNA Research **15**(4): 173-183.

Malloch, D. and R. F. Cain (1972). "Trichocomataceae - Ascomycetes with Aspergillus, Paecilomyces, and Penicillium Imperfect States." Canadian Journal of Botany-Revue Canadienne De Botanique **50**(12): 2613-2628.

Manikandan, P., I. Dóczi, et al. (2008). Aspergillus species in human keratomycosis. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 293-328.

May, G., F. Shaw, et al. (1999). "The signature of balancing selection: Fungal mating compatibility gene evolution." Proceedings of the National Academy of Sciences of the United States of America **96**(16): 9172-9177.

May, G. S. and T. H. Adams (1997). "The importance of fungi to man." Genome Research **7**(11): 1041-1044.

Metzenberg, R. L. and N. L. Glass (1990). "Mating Type and Mating Strategies in Neurospora." Bioessays **12**(2): 53-59.

Micali, O. C. and M. L. Smith (2005). Biological Concepts of Vegetative Self and Nonself Recognition in Fungi. Evolutionary genetics of fungi. J. Xu. Wymondham, U.K., Horizon Bioscience**:** 63-86.

Milgroom, M. G., Fry, W.E. (1997). "Contributions of Population Genetics to Plant Disease Epidemiology and Management." Advances in Botanical Research **24**: 1-30.

Moore, G. G., R. Singh, et al. (2009). "Recombination and lineage-specific gene loss in the aflatoxin gene cluster of Aspergillus flavus." Mol Ecol **18**(23): 4870-87.

Müller, G., H. Rosner, et al. (2003). "Effects of the mycotoxin ochratoxin A and some of its metabolites on the human cell line THP-1." Toxicology **184**(1): 69-82.

Nelson, M. A. (1996). "Mating systems in ascomycetes: A romp in the sac." Trends in Genetics **12**(2): 69-74.

Nixon, K. C. and Q. D. Wheeler (1990). "An Amplification of the Phylogenetic Species Concept." Cladistics **6**(3): 211-223.

O'Gorman, C. M., H. T. Fuller, et al. (2009). "Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*." Nature **457**(7228): 471-474.

Pál, K., A. D. van Diepeningen, et al. (2008). Sexual genes in the asexual filamentous fungus *Aspergillus niger* and related Aspergilli. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 107-131.

Paoletti, M., C. Rydholm, et al. (2005). "Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*." Curr Biol **15**(13): 1242-8.

Papa, K. E. (1973). "Parasexual Cycle in *Aspergillus flavus*." Mycologia **65**(5): 1201-1205.

Papa, K. E. (1978). "Parasexual Cycle in *Aspergillus parasiticus*." Mycologia **70**(4): 766-773.

Peterson, S. W. (2000). Phylogenetic relationships in Aspergillus based upon rDNA sequence analysis. Integration of modern taxonomic methods for Penicillium and Aspergillus classification. J. I. Pitt and R. A. Samson. Amsterdam; Abingdon, Harwood Academic; Marston**:** 323-356.

Peterson, S. W., J. Varga, et al. (2008). Phylogeny and subgeneric taxonomy of Aspergillus. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 33-56.

Pickford, A. (2005). Gene silencing as a tool for the identification of gene function in fungi. Genes and genomics, 5. D. K. Arora and R. M. Berka. San Diego, Elsevier**:** 93-116.

Pildain, M. B., G. Vaamonde, et al. (2004). "Analysis of population structure of *Aspergillus flavus* from peanut based on vegetative compatibility, geographic origin, mycotoxin and sclerotia production." Int J Food Microbiol **93**(1): 31-40.

Pildain, M.B., J.C. Frisvad, et al. (2008). "Two novel aflatoxin-producing *Aspergillus* species from Argentinean peanuts." Int J Syst Evol Microbiol **58**: 725-735.

Pitt, J. I. and A. D. Hocking (1985). "Interfaces Among Genera Related to Aspergillus and Penicillium." Mycologia **77**(5): 810-824.

Pitt, J. I. and R. A. Samson (2007). "Nomenclatural considerations in naming species of Aspergillus and its teleomorphs." Studies in Mycology(59): 67-70.

Pontecorvo, G. (1956). "Parasexual Cycle in Fungi." Annual Review of Microbiology **10**: 393-400.

Prandini, A., G. Tansini, et al. (2009). "On the occurrence of aflatoxin M-1 in milk and dairy products." Food and Chemical Toxicology **47**(5): 984-991.

Purnell, D. M. and G. M. Martin (1973). "Heterozygous Diploid Strains of *Aspergillus nidulans* - Enhanced Virulence for Mice in Comparison to A Phototropic Haploid Strain." Mycopathologia Et Mycologia Applicata **49**(4): 307-319.

Rajasekaran, K., A. J. De Lucca, et al. (2009). "Aflatoxin control through transgenic approaches." Toxin Reviews **28**(2-3): 89-101.

Raju, N. (2009). "Neurospora as a model fungus for studies in cytogenetics and sexual biology at Stanford." Journal of Biosciences **34**(1): 139-159.

Ramirez-Prado, J. H., G. G. Moore, et al. (2008). "Characterization and population analysis of the mating-type genes in *Aspergillus flavus* and *Aspergillus parasiticus*." Fungal Genet Biol **45**(9): 1292-9.

Raper, K. B. and D. I. Fennell (1965). The genus Aspergillus. Baltimore, Williams & Wilkins.

Rasmussen, T. B., M. E. Skindersoe, et al. (2005). "Identity and effects of quorum-sensing inhibitors produced by Penicillium species." Microbiology-Sgm **151**: 1325-1340.

Ratnayake, R., H. Sumithra, et al. (2009). "Effect of GRAS compounds on Aspergillus rot of wood-apple (*Feronia limonia*)." Phytoparasitica.

Reddy, K. R. N., H. K. Abbas, et al. (2009). "Mycotoxin contamination of commercially important agricultural commodities." Toxin Reviews **28**(2-3): 154-168.

Reynolds, D. R. (1994). Implications of the holomorph concept for ascomycete systematics. Ascomycete systematics : problems and perspectives in the nineties. D. L. Hawksworth and North Atlantic Treaty Organization. Scientific Affairs Division. New York, Plenum Press**:** 13-19.

Rigo, K., J. Varga, et al. (2002). "Evolutionary relationships within Aspergillus section Flavi based on sequences of the intergenic transcribed spacer regions and the 5.8S rRNA gene." The Journal of General and Applied Microbiology **48**(1): 9-16.

Robens, J., Cardwell, K.F. (2005). The costs of mycotoxin management in the United States. Aflatoxin and Food Safety. H. K. Abbas. Boca Raton, FL, Taylor and Francis**:** 1-12.

Rydholm, C., G. Szakacs, et al. (2006). "Low genetic variation and no detectable population structure in *Aspergillus fumigatus* compared to closely related Neosartorya species." Eukaryot Cell **5**(4): 650-7.

Samson, R. A., S. B. Hong, et al. (2006). Old and new concepts of species differentiation in Aspergillus, Taylor & Francis Ltd.

Samson, R. A. and J. Varga (2009). "What is a species in Aspergillus?" Med Mycol **47 Suppl 1**: S13-20.

Saupe, S. J., C. Clave, et al. (2000). "Vegetative incompatibility in filamentous fungi: Podospora and Neurospora provide some clues." Curr Opin Microbiol **3**(6): 608-12.

Scheidegger, K. and G. Payne (2003). "Unlocking the Secrets Behind Secondary Metabolism: A Review of *Aspergillus flavus* from Pathogenicity to Functional Genomics." Journal of Toxicology -- Toxin Reviews **22**(2/3): 423-459.

Schmidt, F. R. (2009). "The RNA interference-virus interplay: tools of nature for gene modulation, morphogenesis, evolution and a possible mean for aflatoxin control." Applied Microbiology and Biotechnology **83**(4): 611-615.

Schroeder, H. (1969). "Factors Influencing Development of Aflatoxins in Some Field Crops." Journal of Stored Products Research **5**(3): 187-&.

Selmecki, A. M., K. Dulmage, et al. (2009). "Acquisition of Aneuploidy Provides Increased Fitness during the Evolution of Antifungal Drug Resistance." PLoS Genet **5**(10): e1000705.

Shiu, P. K., N. B. Raju, et al. (2001). "Meiotic silencing by unpaired DNA." Cell **107**(7): 905-16.

Smith, C. A., C. P. Woloshuk, et al. (2007). "Silencing of the aflatoxin gene cluster in a diploid strain of *Aspergillus flavus* is suppressed by ectopic aflR expression." Genetics **176**(4): 2077-2086.

Sorenson, W. G., J. D. Tucker, et al. (1984). "Mutagenicity of tetramic mycotoxin cyclopiazonic acid." Appl. Environ. Microbiol. **47**(6): 1355-1357.

Takahata, N., Y. Satta, et al. (1992). "Polymorphism and Balancing Selection at Major Histocompatibility Complex Loci." Genetics **130**(4): 925-938.

Taylor, J. W., J. Spatafora, et al. (2006, 2009). "Ascomycota. Sac Fungi." The Tree of Life Web Project Retrieved 27 February, 2010, from http://tolweb.org/Ascomycota.

Toledo-Hernandez, C., A. Zuluaga-Montero, et al. (2008). "Fungi in healthy and diseased sea fans (*Gorgonia ventalina*): is *Aspergillus sydowii* always the pathogen?" Coral Reefs **27**(3): 707-714.

Upshall, A. (1971). "Phenotypic specificity of aneuploid states in *Aspergillus nidulans*." Genetics Research **18**(02): 167-171.

Upshall, A. (1981). "Naturally Occurring Diploid Isolates of *Aspergillus nidulans*." J Gen Microbiol **122**(1): 7-10.

van Diepeningen, A. D., J. Varga, et al. (2008). Mycoviruses in the Aspergilli. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 133-176.

van Dijck, P. W. M. (2008). The importance of Aspergilli and regulatory aspects of Aspergillus nomenclature in biotechnology. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 249-257.

Varga, J. and R. A. Samson (2008). Aspergillus in the genomic era. Wageningen, Netherlands, Wageningen Academic Publishers.

Ward, T. J., J. P. Bielawski, et al. (2002). "Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic Fusarium." Proceedings of the National Academy of Sciences of the United States of America **99**(14): 9278-9283.

Wickner, R. B., K. L. Taylor, et al. (1999). "Prions in Saccharomyces and Podospora spp.: protein-based inheritance." Microbiol Mol Biol Rev **63**(4): 844-61, table of contents.

Wilkinson, H. H., A. Ramaswamy, et al. (2004). "Increased conidiation associated with progression along the sterigmatocystin biosynthetic pathway." Mycologia **96**(6): 1190-1198.

Wu, F., Y. Liu, et al. (2008). "Cost-Effectiveness of Aflatoxin Control Methods: Economic Incentives." Toxin Reviews **27**(3-4): 203-225.

www.aspergillus.org.uk. (2009). "The Aspergillus Website." Retrieved November 22, 2009, from http://www.aspergillus.org.uk/indexhome.htm?secure/metabolites/list_by_secmet.php~main.

Yin, Y., T. Lou, et al. (2009). "Molecular characterization of toxigenic and atoxigenic *Aspergillus flavus* isolates, collected from peanut fields in China." J Appl Microbiol.

Yu, J., P. K. Chang, et al. (2004). "Clustered pathway genes in aflatoxin biosynthesis." Appl Environ Microbiol **70**(3): 1253-62.

Chapter 1

Recombination and Lineage-Specific Gene Loss in the

aflatoxin gene cluster of *Aspergillus flavus*

2009. Molecular Ecology 48(23): 4870-4887

Recombination and Lineage-Specific Gene Loss in the

aflatoxin gene cluster of *Aspergillus flavus*

G. G. Moore[1], R. Singh[1], B.W. Horn[2], I. Carbone[1]

[1] Center for Integrated Fungal Research, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695
[2] National Peanut Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Dawson, GA, 39842

ABSTRACT

Aflatoxins produced by *Aspergillus flavus* are potent carcinogens that contaminate agricultural crops. Recent efforts to reduce aflatoxin concentrations in crops have focused on biological control using nonaflatoxigenic *A. flavus* strains AF36 (= NRRL 18543) and NRRL 21882 (the active component of afla-guard®). However, the evolutionary potential of these strains to remain nonaflatoxigenic in nature is unknown. To elucidate the underlying population processes that influence aflatoxigenicity, we examined patterns of linkage disequilibrium (LD) spanning 21 regions in the aflatoxin gene cluster of *A. flavus.* We show that recombination events are unevenly distributed across the cluster in *A. flavus.* Six distinct LD blocks separate late pathway genes *aflE*, *aflM, aflN, aflG, aflL, aflI* and *aflO*, and there is no discernable evidence of recombination among early pathway genes *aflA, aflB, aflC, aflD, aflR* and *aflS.* The discordance in phylogenies inferred for the *aflW/aflX* intergenic region and two non-cluster regions, tryptophan synthase and acetamidase, is indicative of trans-species evolution in the cluster. Additionally, polymorphisms in *aflW/aflX* divide *A. flavus*

strains into two distinct clades, each harboring only one of the two approved biocontrol strains. The clade with AF36 includes both aflatoxigenic and nonaflatoxigenic strains, whereas the clade with NRRL 21882 comprises only nonaflatoxigenic strains and includes all strains of *A. flavus* missing the entire gene cluster or with partial gene clusters. Our detection of LD blocks in partial clusters indicates that recombination may have played an important role in cluster disassembly, and multilocus coalescent analyses of cluster and non-cluster regions indicate lineage-specific gene loss in *A. flavus*. These results have important implications in assessing the stability of biocontrol strains in nature.

INTRODUCTION

Species in *Aspergillus* section *Flavi* commonly infect agricultural staples such as corn, cottonseed, peanuts, and tree nuts, and produce an array of mycotoxins (Horn 2007). The most common mycotoxins are the aflatoxins, carcinogenic polyketides that pose a serious health risk to animals and humans (Eaton & Groopman 1994). *Aspergillus flavus* and *A. parasiticus* are the most abundant aflatoxin-producing species in agricultural soils and have the potential to cause considerable contamination of crops (Horn 2005). *A. flavus* is the dominant aflatoxin-producing species in the majority of crops, whereas *A. parasiticus* is more frequently associated with peanuts (Horn *et al.* 1994; Horn 2005). The two species also differ in the types of aflatoxins produced. Aflatoxin biosynthesis involves more than 25 enzymes encoded by genes that are clustered together in a 75-kb telomeric region on chromosome 3 (Carbone *et al.* 2007b). Four structurally related aflatoxins occur in nature, depending upon the presence of polyketide dihydro- (B_1 and G_1) or tetrahydro- (B_2 and G_2)

bisfuran rings. Of these aflatoxins, B_1 is the most toxic and carcinogenic (Eaton & Groopman 1994). While both *A. flavus* and *A. parasiticus* produce the B aflatoxins (B_1 and B_2), *A. parasiticus* also produces the G aflatoxins (G_1 and G_2) or rarely *O*-methylsterigmatocystin (OMST), which is the precursor of aflatoxin B_1 (Yu *et al.* 2004). Because of the high toxicity of aflatoxins to humans and animals, federally and internationally mandated laws require the destruction, decontamination, or reprocessing of crops when the aflatoxin content exceeds specified levels, resulting in huge yield losses worldwide (van Edmond & Jonker 2005). The economic loss is worse when crops are exposed to drought and elevated temperatures, conditions that favor crop invasion and fungal growth by aflatoxigenic fungi (Cole *et al.* 1985).

Individuals within *A. flavus* populations vary widely in their ability to produce aflatoxins, ranging from those that are nonaflatoxigenic to those that are potent producers of aflatoxins (Horn & Dorner 1999). *A. flavus* is characterized by two morphotypes: the typical L strain with large sclerotia >400 μm in diameter and the S strain with abundant small sclerotia <400 μm (Cotty 1989). Nonaflatoxigenic strains are fairly common in *A. flavus* L strain (Chang *et al.* 2005; Horn & Dorner 1999), but are rare in *A. flavus* S strain and *A. parasiticus* (Horn & Dorner 1999). In addition to aflatoxins, *A. flavus* produces another unrelated mycotoxin, cyclopiazonic acid (CPA), an indol-tetramic acid that targets the liver, kidneys and gastrointestinal tract in animals (Burdock & Flamm 2000). Aflatoxins and CPA often co-contaminate agricultural products and several of the disease symptoms in animals can be attributed to CPA (Cole 1986). CPA biosynthesis involves approximately 3-5

enzymes encoded in a 50-kb mini-cluster adjacent to the aflatoxin gene cluster (Chang *et al.* 2009). The inability to produce aflatoxins and/or CPA in some strains of *A. flavus* is the result of various deletions in these gene clusters (Chang *et al.* 2005, 2009). Populations of *A. flavus* show a high level of variation in mycotoxin production, with individuals producing both aflatoxins and CPA, aflatoxins alone, CPA alone, or neither mycotoxin (Horn & Dorner 1999; Horn *et al.* 1996).

Two nonaflatoxigenic strains of *A. flavus* (AF36 and NRRL 21882) are currently registered through the US Environmental Protection Agency (EPA) and applied to fields as biocontrol agents to out-compete indigenous aflatoxigenic strains (Dorner 2005). AF36 was originally isolated from a cotton field in Arizona (EPA 2003) and approved for use on cotton in Arizona and Texas. Ehrlich and Cotty (2004) reported that loss of aflatoxigenicity in AF36 is the result of a nonsense mutation in *pksA* (*aflC*), a critical early pathway gene in aflatoxin biosynthesis. AF36 otherwise has a full aflatoxin gene cluster and a functional CPA cluster. NRRL 21882 (the active component of afla-guard®) was isolated from a peanut seed in Georgia (EPA 2004) and is currently approved for commercial use on peanuts and corn in the United States. Chang *et al.* (2005) reported that NRRL 21882 is missing the entire aflatoxin and CPA gene clusters. The application of highly competitive nonaflatoxigenic strains of *A. flavus* to fields has been shown to be effective in reducing aflatoxin contamination in peanuts (Dorner 2009b), corn (Dorner 2009a) and cottonseed (Cotty 1994). Although the two biocontrol strains have been approved by EPA and are used widely in the US, there is uncertainty as to whether they are capable of reacquiring toxigenicity through genetic exchange and recombination with indigenous aflatoxigenic strains.

The possibility that recombination in *A. flavus* could potentially influence the stability of biocontrol strains was first suggested by Geiser *et al* (1998), who reported a population structure in *A. flavus* indicative of recombination based on a lack of congruence of five gene genealogies. However, the method of recombination, how often recombination occurs, or when recombination occurred in the history of the species could not be determined. Similar evidence for recombination based on incongruence of gene genealogies for four genes was shown for *A. nomius* (Peterson *et al.* 2001), another aflatoxin-producing species. More recently, Carbone *et al.* (2007a) sequenced 21 regions across the aflatoxin gene cluster for a population of *A. parasiticus* sampled from a single peanut field in Georgia. Detection of significant linkage disequilibrium (LD) over the evolutionary history of *A. parasiticus* resolved five recombination blocks across the gene cluster, while coalescent analysis suggested that some recombination had occurred within the last one million years. Furthermore, recombination alone separated some aflatoxin cluster haplotypes (Carbone *et al.* 2007a).

Although genealogical approaches have provided indirect evidence of recombination, our discovery of a sexual state for both *A. flavus* (Horn *et al.* 2009a) and *A. parasiticus* (Horn *et al.* 2009b, c) has provided the first direct evidence that sexual recombination may be an important mechanism for generating diversity in aflatoxin production in these agriculturally important species. In *A. parasiticus*, sexual reproduction in nature, as suggested by a 1:1 distribution of *MAT* genes in populations (Ramirez-Prado *et al.* 2008), increases diversity in mycotoxin profiles and creates new vegetative compatibility groups (VCGs) (BW Horn, unpublished data), whereas balancing selection acts to maintain the ratio of aflatoxins G_1/B_1

(Carbone *et al.* 2007a). In the present study, we show that recombination and balancing selection in *A. flavus* also work in parity to maintain nonaflatoxigenicity and to influence cluster disassembly in a lineage-specific fashion. These underlying evolutionary processes are important considerations when developing biocontrol strategies.

MATERIALS AND METHODS

Population Sample

Our initial sample of 79 single-spore *A. flavus* isolates (L strains) was from a single peanut field (Herod, Terrell County, Georgia, USA) sampled in 1992 (Horn & Greene 1995). This sample was consolidated to a subset of 44 isolates (see Table 1.1) based on vegetative compatibility in which isolates capable of forming stable hyphal fusions are assigned to the same vegetative compatibility group (VCG) (Leslie 1993). VCGs 6-30 contain two or more isolates and VCGs 31-63 each contain a single isolate. We expanded the sample to include two additional strains, if available, within each VCG. Two VCGs (24 and 32) are nonaflatoxigenic; VCG 24 (NRRL 21882; IC252-254) is missing all the genes in the aflatoxin and CPA clusters and the telomeric repeat sequence (TTAGGG) is adjacent to *hxtA* (Chang *et al.* 2005), whereas VCG 32 (IC277) has complete clusters. The production of aflatoxins B_1 and B_2 and CPA was determined in a previous study (Horn *et al.* 1996). Also examined were *A. flavus* L strains representative of deletion patterns A (IC309=NPL MS5-6=NRRL 35735), B (IC310=NPL NC3-6=NRRL 35736), C (IC311=NPL AL3-9=NRRL 35737), D (IC312=NPL GA4-4=NRRL 35738), E (IC313=NPL TX13-5=NRRL 35739), F (IC314=NPL NC7-8=NRRL 35740), and G (IC315=NPL AL1-4=NRRL 35741) reported by

Chang *et al.* (2005) in which NRRL 21882 is pattern H; mycotoxin production was determined by Horn and Dorner (1999). To examine trans-species evolution we included the 24 isolates of *A. parasiticus* sampled from the same Georgia field (Horn & Greene 1995) and analyzed in a previous study (Carbone *et al.* 2007a). Also included were the biocontrol strain AF36 (=NRRL 18543) and the genome strains *A. flavus* NRRL 3357 (http://www.aspergillusflavus.org/) and *A. oryzae* NRRL 5590 (Machida *et al.* 2005); *A. nomius* NRRL 13137 (ex type) was the outgroup species for *Aspergillus* section *Flavi.*

Table 1.1 Strain designations, haplotypes, VCGs, and mean concentrations of mycotoxins in *Aspergillus flavus**

IC	NRRL	*MAT*	VCG	Haplotype	B_1	B_2	Total aflatoxins	CPA
IC229	29459	2	6	H2	39.2(22.4)	0.8(0.5)	40(22.9)	22.5(3.8)
IC230	29460	2	6	H2	15.8(13.3)	0.3(0.2)	16.1(13.5)	26.9(3.2)
IC231	29461	2	6	H2	33.6(13.3)	0.7(0.2)	34.3(13.5)	23.9(2.5)
IC234	29464	2	14	H14	85.6(5.8)	1.4(0.3)	87(5.7)	60.7(8.4)
IC235	29465	2	14	H14	104.2(16.7)	1.7(0.1)	105.9(16.8)	65.6(3.1)
IC236	29466	2	14	H14	105.5(15.3)	1.6(0.7)	107.1(16)	53.2(7.9)
IC244	29473	1	17	H15	104.3(10.3)	1.7(0.1)	106(10.5)	64.9(4.2)
IC245	29474	2	23	H10	96.3(6.2)	1.9(0.6)	98.3(6.8)	48.5(3.9)
IC246	29475	2	23	H10	100(8.6)	2.2(0.2)	102.2(8.8)	54.4(5.7)
IC247	29476	2	23	H10	94.2(12.5)	2(0.3)	96.2(12.8)	48.2(0.6)
IC252	29481	2	24	—	0(0)	0(0)	0(0)	0(0)
IC253	29482	2	24	—	0(0)	0(0)	0(0)	0(0)
IC254	29483	2	24	—	0(0)	0(0)	0(0)	0(0)
IC258	29487	2	25	H15	168.3(28.7)	4.4(0.6)	172.7(29.2)	155.7(51.9)
IC259	29488	2	25	H15	134.3(5)	3.8(0.8)	138.1(5.8)	92.9(2.8)
IC260	29489	2	25	H15	210.6(35.5)	6(0.8)	216.6(36.3)	115.6(25.6)

Table 1.1 Continued

IC	NRRL	*MAT*	VCG	Haplotype	B_1	B_2	Total aflatoxins	CPA
IC263	29492	2	26	H26	0.8(0.1)	0(0)	0.8(0.1)	22(3.4)
IC264	29493	2	26	H26	0.5(0.1)	0(0)	0.5(0.1)	25.2(9.5)
IC265	29494	2	26	H26	0.1(0.1)	0(0)	0.1(0.1)	22.4(3)
IC267	29496	2	27	H6	109(41.4)	2.5(1.3)	111.4(42.7)	93.1(12.9)
IC268	29497	2	27	H6	141.2(50)	3.8(2)	145(52)	117.2(18.3)
IC269	29498	2	27	H6	79(37)	2.5(1.3)	81.5(38.3)	79.9(4.3)
IC270	29499	2	28	H20	72.5(7.2)	3.3(0.7)	75.8(6.6)	44.9(2.6)
IC271	29500	2	28	H20	89.4(6.5)	2.1(1)	91.5(7.5)	53.9(3.1)
IC272	29501	2	29	H22	64.2(6.7)	1.2(0.2)	65.5(6.7)	85.8(12.3)
IC273	29502	2	29	H22	53.7(10.7)	0.4(0.1)	54.1(10.8)	74.8(6.6)
IC274	29503	2	30	H19	4.8(1.5)	0.1(0.1)	4.9(1.5)	53.4(8.2)
IC275	29504	2	30	H19	2.4(1.3)	0.1(0)	2.5(1.3)	55.5(2.5)
IC276	29505	2	31	H11	0.4(0.1)	0(0)	0.4(0.1)	55.3(6.4)
IC277	29506	2	32	H1	0(0)	0(0)	0(0)	99.6(20.7)
IC278	29507	1	33	H15	99.4(46.2)	1.8(1.1)	101.2(47)	76.8(4.8)
IC279	29508	2	34	H2	12.9(1.1)	0.1(0)	13(1.2)	53.2(2.9)
IC280	29509	2	35	H7	36.9(7.8)	1.3(0.5)	38.1(8)	46(7.3)
IC281	29510	2	36	H14	154(8.6)	2.6(0.5)	156.6(8.6)	85(3.8)
IC282	29511	1	37	H19	164.6(26.8)	5.5(1)	170(27.8)	46.4(3.2)
IC283	29512	1	38	H24	9.6(1.4)	0.1(0)	9.8(1.4)	10.2(1.1)
IC284	29513	2	39	H22	24.6(3.3)	0.4(0)	25(3.3)	126(6.5)
IC285	29514	2	40	H26	185.3(47.1)	1.5(0.8)	186.8(47.8)	18.9(3.5)
IC286	29515	2	41	H2	24.7(4)	0.7(0.1)	25.4(4.1)	20.3(1.1)

Table 1.1 Continued

IC	NRRL	*MAT*	VCG	Haplotype	B_1	B_2	Total aflatoxins	CPA
IC287	29516	2	42	H18	104.1(27.5)	2.3(1.2)	106.4(28.7)	75(0.9)
IC288	29517	2	43	H5	97.4(18.3)	1.8(0.6)	99.2(18.8)	87.7(4.4)
IC289	29518	1	44	H4	40.2(26.1)	0.7(0.4)	40.9(26.5)	85.9(14.2)
IC290	29519	2	45	H7	15.3(4.5)	0.3(0.1)	15.6(4.6)	48.2(3.4)
IC291	29520	1	46	H25	47.6(13.9)	0.3(0.1)	47.9(14.1)	43.7(5)
IC292	29521	2	47	H21	105.6(24.5)	2.8(0.4)	108.3(25)	97.5(14.6)
IC293	29522	1	48	H8	54.2(12.7)	1(0.2)	55.2(12.9)	161.5(20.1)
IC294	29523	2	49	H23	19.6(5)	0.1(0.1)	19.7(5)	76.7(2.7)
IC295	29524	2	50	H19	164.6(17.4)	4.7(0.8)	169.3(18.2)	60.8(2.8)
IC296	29525	1	51	H12	0.3(0.1)	0(0)	0.3(0.1)	58.6(5.5)
IC297	29526	1	52	H9	25.7(10.9)	0.8(0.4)	26.5(11.3)	39.4(5.7)
IC298	29527	1	53	H22	18.7(4.3)	0.2(0)	18.8(4.3)	44.6(8.1)
IC299	29528	2	54	H17	36.2(19.2)	0.8(0.5)	37(19.7)	125.7(12.1)
IC300	29529	1	55	H14	51.4(10.9)	0.6(0.2)	52(11.1)	52.3(4.4)
IC301	29530	1	56	H22	16.8(3.3)	0.1(0)	17(3.3)	100.3(17.6)
IC302	29531	2	57	H17	83.2(16.1)	1.7(0.6)	84.9(16.6)	55.6(3.1)
IC303	29532	1	58	H16	94.9(15)	1.8(0.2)	96.7(15.1)	46.1(2.2)
IC304	29533	2	59	H15	69.6(12.2)	1.5(0.7)	71.1(12.9)	114.2(5.3)
IC305	29534	1	60	H27	37.4(10.5)	0.5(0.2)	37.9(10.7)	17.5(1.6)
IC306	29535	2	61	H2	11.5(0.7)	0.2(0.1)	11.7(0.8)	12.9(1.9)
IC307	29536	2	62	H13	139.8(5.1)	3.1(0.1)	142.9(5.1)	77.2(1.5)
IC308	29537	1	63	H3	39.8(6.5)	0.6(0.1)	40.4(6.6)	183.5(18.3)

*Isolates were sampled from a Georgia peanut field in 1992 (Horn & Greene 1995). Culture collection designations: IC (I Carbone), NRRL (Agricultural Research Service Culture Collection, Peoria, Illinois). Mating-type designations are from Ramirez-Prado *et al.* (2008) and vegetative compatibility groups are based on Horn and Greene (1995). Concentration units are μg/mL. Mean concentrations and ratios (n = 3) are shown; standard deviations are indicated in parentheses. Strains were grown on yeast extract–sucrose broth in vials at 30 °C and cultures were analyzed for mycotoxins with high-performance liquid chromatography (Horn *et al.* 1996; Horn & Dorner 1999). Isolates in VCG 24 (IC252, IC253 and IC254) are missing aflatoxin cluster genes precluding inference of multilocus haplotypes.

DNA isolation, PCR amplification and sequencing

Sequences of oligonucleotide primers (Table 1.2) and procedures for isolation of DNA and for amplifications were described previously (Carbone *et al.* 2007a). Approximately 1-2 kb was sequenced per isolate for each of 21 regions within the *A. flavus* aflatoxin gene cluster positioned on chromosome 3 and for two non-cluster regions, tryptophan synthase (*trpC)* and acetamidase (*amdS)*, located on chromosomes 4 and 6, respectively. We focused on the same regions that were previously examined in *A. parasiticus* (Carbone *et al.* 2007a). This was important for comparing patterns of LD and rates of recombination in the gene cluster between the two species and for reconstructing trans-species evolution. We used the primers *aflC-F* (TTAGATCGGTCCCTTTACTTTC) and *aflC-R* (GGTGTCAGTCCTTGTCTCTGTA) to amplify and sequence a coding portion of the *aflC* gene to determine whether nonaflatoxigenic strains contained a nonsense mutation, as previously reported (Ehrlich & Cotty 2004).

Table 1.2 Oligonucleotide primer sets used for amplification and sequencing of 21 cluster regions in *Aspergillus flavus*

Target Region	Primer Name	Primer Sequence	Intergenic/Gene[1] (bp)
aflT/aflC	*aflTF-XIR*	AACTGGTCCAACCGGAGTAC	390/0
	aflCR-XIR	CGGCAGATACAGTCATGGAC	
aflC/aflD	*aflCF-XIR*	GGTTCGAGCCAACTTGTGAT	340/0[2]
	aflDR-XIR	GACGTTGGAGAAAAGCTTCAAT	
aflD/aflA	*aflDF-XIR*	AACAACTGTCGCAGACAGTGT	265/0
	aflAR-XIR	TGAGATCGAGCATGGAGGTA	

Table 1.2 Continued

Target Region	Primer Name	Primer Sequence	Intergenic/Gene[1] (bp)
aflA/aflB	*aflAF-XIR*	AATTGCCGTCTTCAGCTTTC	534/200 (*aflA*)
	aflBR-XIR	GTGTTGGATGCCACGTCTAG	
aflB/aflR	*aflBF-XIR*	GAGGGGAGATTGAGCCTTATC	183/0
	aflRR-XIR	CCAGTCGCTGGTGAACCTAT	
aflR/aflS	*aflRF-XIR*	CCTGGCTGAAGGAAGACTCT	529/192 (*aflR*)
	aflSR-XIR	CTGGCGAGGGCTAATACTTG	
aflS/aflH	*aflSF-XIR*	GGTCAGTCTGAGCGATCTCTC	311/8 (*aflS*)
	aflHR-XIR	AGCAGACGTAGTGGACGTGT	
aflH/aflJ	*aflHF-XIR*	ACCTTCTTGCTCCTTGGTTC	523/63 (*aflH*)
	aflJR-XIR	CCGTAGCGCGTAGCTAATGTA	
aflJ/aflE	*aflJF-XIR*	AAGGCTCCTGAGACTCGCTA	0/456 (*aflJ*)
	aflER-XIR	GTTCGACTTGATCTTTTGCG	
aflE/aflM	*aflEF-XIR*	GAACCATTTGACGTCGGATT	336/59 (*aflE*)
	aflMR-XIR	GTTGGCCTTGATCTGTTGAA	
aflM/aflN	*aflMF-XIR*	GCTTGGCTCTCTCCTTTGAA	546/506 (*hypE*)[3]
	aflNR-XIR	GCTGCTGAGGGAGTTGAAAC	
aflN/aflG	*aflNF-XIR*	TATTACGCCAGCATACGATGA	428/5 (*aflN*)[2]
	aflGR-XIR	CGGTTGATCCTAGTCAAGCTT	
aflG/aflL	*aflGF-XIR*	ATAGCTCATTGGGTGCGATT	454/163 (*aflG*)
	aflLR-XIR	AGGAACCGTACAAGTACGACAA	
aflL/aflI	*aflLF-XIR*	AGAGAACTGGCTCGCCATAG	450/103 (*hypB*)[4]
	aflIR-XIR	GCCACTGGTAGTGTCCATCA	
aflI/aflO	*aflIF-XIR*	GCACTATCCTGGCACAATGT	358/278 (*aflO*)
	aflOR-XIR	TTCTTCGAGTCACAGCCTATTC	
aflO/aflP	*aflOF-XIR*	CTGCTCAATCGCATAACCAC	350/12 (*aflO*)
	aflPR-XIR	CAGCGCTTCATGAATCAGAT	

Table 1.2 Continued

Target Region	Primer Name	Primer Sequence	Intergenic/Gene[1] (bp)
aflP/aflQ	*aflPF-XIR*	GTGGAGGATGGAGTCCTCTCT	349/20 (*aflP*)
	aflQR-XIR	CCAGCAGCTTCTCCAGAATC	
aflQ/aflK	*aflQF-XIR*	GGTCAGCTGCTTCTTCATCTC	284/0
	aflKR-XIR	CGTAGTCGAATGACTGTCCC	
aflK/aflV	*aflKF-XIR*	CTATGGGCAAAGCAGATGATT	409/0
	aflVR-XIR	CAGAAGAGGCGAAAATGTCA	
aflV/aflW	*aflVF-XIR*	GGGACTGAAATATGCGGTTT	331/133 (*aflV*)
	aflWR-XIR	GGAGCAAAGGGGTAGGTGTAG	
aflW/aflX	*aflWF-XIR*	GCACACGGTGTGGAAAGATA	529/296 (*aflW*)
	aflXR-XIR	GACTAGTGCACGATGTGCAAC	

[1]The ratio of the intergenic region to gene (exons plus introns).
[2]*hypC* (GenBank accession XM-002379909) and *hypD* (GenBank accession XM-002379897) do not overlap with our sequenced regions.
[3]506 bp of the total length overlap with a hypothetical gene, *hypE* (GenBank accession XM-002379899). Of these 506 bp only 414 bp are coding and 92 bp are noncoding (intron).
[4]103 bp of the total length overlap with *hypB* (GenBank accession XM-002379894), and 61 bp of an intron.

Molecular sequence variation

DNA sequences for each locus were aligned and manually adjusted using Sequencher Version 4.5 (Gene Codes Corporation, Ann Arbor, MI). The multiple sequence alignment for each region was trimmed and exported as NEXUS files for further analysis in SNAP Workbench (Price & Carbone 2005). We calculated π, the mean number of pairwise nucleotide differences per site (Nei 1987), for each of the 21 multiple sequence alignments and for *trpC* and *amdS* using the program SITES version 1.1 (Hey & Wakeley 1997).

Alignment files for the 21 regions were then combined using SNAP Combine (Aylor *et al.* 2006) to create concatenated sequences for each isolate that span the cluster. The multiple sequence alignment was collapsed into haplotypes with the options of recoding insertions/deletions (indels) and excluding all variation that violates an infinite sites mutation model. The latter was important for coalescent analyses (described below) that assume an infinite sites model. The real positions of single nucleotide polymorphisms (SNPs) in the concatenated multiple sequence alignment were based on the physical mappings of the 21 regions within the 75-kb aflatoxin gene cluster of *A. flavus*, using AF36 (AY510455) as a reference sequence.

Linkage disequilibrium and compatibility analyses

We first examined LD across the 21 regions for the subset of 44 isolates representing distinct VCGs. After inferring LD blocks in this subsample, we examined the expanded VCG sample to determine if block boundaries were conserved within VCGs. We performed LD and compatibility analyses using SNAP Clade and Matrix (Bowden *et al.* 2008). Recombination blocks were based on informative SNPs and recoded indels that are infinite sites compatible; uninformative variation was included to identify non-recombining regions possibly under the influence of balancing selection (Carbone *et al.* 2007a). For example, previous research showed evidence of trans-species evolution in *hypE,* a protein-encoding gene of unknown function between *aflM* and *aflN* in the aflatoxin gene cluster of *A. parasiticus* (Carbone *et al.* 2007a). The size of recombination blocks for the concatenated multiple DNA sequence alignment depended on the number of contiguous pairs of sites that

were both strongly correlated ($0.8 < r^2 < 1$) and significantly linked ($P < 0.01$). The actual sizes of the blocks were determined by performing LD analysis on complete cluster sequences for *A. flavus* AF13 (L strain, AY510451), AF36 (L strain, AY510455) and NRRL 3357 (L strain; http://www.aspergillusflavus.org/), and for *A. oryzae* NRRL 5590 (Machida *et al.* 2005).

We further examined patterns of incompatibility among pairs of sites using the four-gamete test (Hudson & Kaplan 1985) implemented in SNAP Clade. This analysis may reveal one or more blocks with distinct evolutionary histories that need to be examined separately in coalescent analyses that assume no recombination (described below). Recombination hotspots, defined as regions in the cluster that are more prone to recombination, were deduced by mapping the minimum number of recombination events (R_h) calculated using the program RecMin (Myers & Griffiths 2003). An estimate of the population recombination rate, per base pair, within the cluster was obtained using Hey and Wakeley's γ estimator (Hey & Wakeley 1997) as implemented in the SITES program. A minimal ancestral recombination graph (ARG) for all SNPs spanning the cluster was inferred using the branch and bound method implemented in Beagle (Lyngsø *et al.* 2005); if there were too many recombination events for branch and bound to go to completion we used the program Kwarg (http://www.stats.ox.ac.uk/~lyngsoe/section26/), which is a heuristic implementation of the Beagle algorithm. The ARG was useful in examining the ancestral history of haplotypes and provided a rough estimate of haplotype age; in this case we define older haplotypes as requiring more mutations and recombination events before coalescing to a recent common ancestor in the ARG.

LD and compatibility analyses were performed for the 21 sequenced regions in complete clusters and for homologous regions in partial gene clusters. The latter was important to determine if recombination block boundaries are conserved among complete and partial clusters. This would indicate that complete and partial clusters share a history of recombination and possibly that partial clusters are derivatives of full cluster ancestors.

Telomeric regions and G + C content

Several studies report on the importance of recombination and G + C content in the maintenance and disassembly of gene clusters (Howlett *et al.* 2007; Patron *et al.* 2007). We examined G + C content distribution along the right arm of chromosome 3 (200 kb) for the sequenced genome strain of *A. flavus* NRRL 3357 to identify regions or islands that undergo an abrupt change in G + C content (Zhang & Zhang 2004). In bacteria, sharp transitions in G + C content have been associated with horizontal transfer of genomic islands (Zhang & Zhang 2004). We used a windowless method to display the cumulative GC profile (or *z'* curve), which is reported to have better resolution than window-based methods in identifying putative boundaries of genomic islands (Zhang & Zhang 2004). The *z'* curve displays variations in G + C content along a chromosome based on the cumulative count of G and C bases. A *z'* curve that is approximately a straight line indicates a constant G + C content; a jump (positive slope) in the *z'* curve indicates a decrease in G + C content, whereas a drop (negative slope) in the *z'* curve indicates an increase in G + C content.

Trans-species evolution

To reconstruct patterns of trans-species evolution we expanded the *A. flavus* sample to include 24 *A. parasiticus* isolates examined previously (Carbone *et al.* 2007a), two representative nonaflatoxigenic strains with complete clusters (*A. flavus* AF36 and *A. oryzae* NRRL 5590), and eight nonaflatoxigenic *A. flavus* strains that represent distinct deletion patterns in the aflatoxin cluster (Chang *et al.* 2005). Because we are interested in detecting the targets of balancing selection in *A. flavus*, we initially restricted our analysis to regions in the *A. flavus* cluster harboring uninformative polymorphisms, which are easily identifiable as strongly correlated pairs of sites in the LD plot with no significant linkage. This is important since recombination will uncouple polymorphisms, thus making it difficult to identify chemotype-specific haplotypes. The two non-cluster genes, *trpC* and *amdS,* are orthologous and provided independent estimates of the species phylogeny. Phylogenies were inferred using unweighted parsimony in PAUP* 4.0 (Swofford 1998) and rooted by specifying *A. nomius* as the outgroup (Peterson 2008). If parsimony searches yielded more than one equally parsimonious tree, support for branches was assessed using 500 bootstrap samples. We also inferred a strict consensus tree and resolved polytomies by reconstructing ARGs. The ARG for the aflatoxin cluster region under balancing selection was examined more closely to determine if the effects of mutation and recombination result in significantly different toxin profiles, which would be expected if the region is important in maintaining chemotype-specific differences in *A. flavus.* Recombination between chemotype-specific lineages throughout the cluster is the basis for our detection of recombination blocks, as was shown in *A. parasiticus* (Carbone *et al.* 2007a). We examined the ARGs for evidence of

recombination between chemotype-specific lineages, or in the case of *A. flavus*, between aflatoxigenic and nonaflatoxigenic lineages.

Neutrality tests and coalescent analyses

Departures from neutrality and population-size constancy were tested separately for each of the 21 cluster regions using Tajima's *D* (Tajima 1989), Fu and Li's *D** and *F** (Fu & Li 1993) and Fu's F_S (Fu 1997; Fu & Li 1993; Tajima 1989). Significance thresholds were Bonferroni-corrected by dividing by the total number of neutrality tests ($n = 84$). Significant values of Tajima's *D* and Fu & Li's *D** and *F** and non-significant values of Fu's F_S (Fu 1997) are indicative of balancing selection. One effect of balancing selection is to preserve polymorphisms in two or more lineages over an extended period, resulting in strong genetic differentiation and positive values of *D* and *Fs*, which should not be interpreted as population subdivision or reproductive isolation. Negative values of *D* and F_S would indicate population growth or a selective sweep.

To reconstruct the ancestral history of complete and partial aflatoxin gene clusters, we examined molecular sequence variation in conserved cluster regions as well as in adjacent non-cluster genes. This is particularly relevant because the genes required for the biosynthesis of CPA are immediately adjacent to the aflatoxin cluster (Chang *et al.* 2009). Specifically, we wanted to ascertain whether gene loss was occurring only for genes in the aflatoxin cluster or included genes in the CPA cluster and possibly other telomeric regions. Identifying conserved genes that flank complete and partial clusters would indicate a process of gene loss in complete clusters giving rise to partial clusters. This was investigated using

two approaches. First, we focused on cluster genes and reconstructed the ancestral history of complete and partial clusters by simulating mutational histories using the coalescent. Second, we inferred rooted gene genealogies for two adjacent non-cluster regions that are conserved in complete and partial aflatoxin clusters. This was important to further corroborate inferences of gene gain or loss in the cluster. We determined the largest non-recombining partition for complete and partial aflatoxin gene clusters in *A. flavus* using SNAP CladeEx (Bowden *et al.* 2008). As in *A. parasiticus*, the recovery of *A. flavus* isolates belonging to the same VCG across the US (Horn & Dorner 1999) supports a population model with migration and random mating (Bayman & Cotty 1993). Assuming panmixis and constant population size, coalescent analysis was performed using Genetree version 9.0 (http://www.stats.ox.ac.uk/~griff/software.html). This involved generating all possible rooted gene genealogies for the largest non-recombining partition in the cluster and then calculating the relative probabilities of all rooted genealogies by performing one to ten million simulations of the coalescent assuming Watterson's (1975) estimate of θ, panmixis, and constant population size.

RESULTS

Molecular sequence variation

We sequenced a total of 8351 bp for 21 regions of the aflatoxin gene cluster, 323 bp for the 5'-untranslated region of *trpC*, and 333 bp (285 bp exon and 48 bp intron) for *amdS* in our subsample of 44 *A. flavus* strains. For the additional strains within VCGs, we examined only 14 of the 21 regions. Examination of the additional strains excluded most of the early

pathway regions between *aflT* to *aflJ* (see Fig. 1.1), which had low nucleotide diversities per site (π = 0.00041 - 0.00496) comparable to non-cluster genes, *trpC* (0.00458) and *amdS* (0.00342). The late pathway regions from *aflE* to *aflW* had nearly 10-fold higher nucleotide diversities per site (0.00253 – 0.0489). Similar nucleotide diversity estimates were obtained when alignments for late pathway genes included aflatoxin deletion pattern strains. We also sequenced 700 bp of *aflC* for four nonaflatoxigenic isolates (IC277, IC309, IC310 and IC312). Only one isolate (IC309) showed the same nonsense mutation as in AF36; five other polymorphisms encoded three synonymous and two replacement substitutions, and were present in both aflatoxin- and nonaflatoxin-producing strains. All sequences have been deposited in GenBank under Accession numbers FJ877157-FJ878623, GQ456072-GQ456075 and GQ479324-GQ479333. A total of 27 haplotypes were inferred for the 21 aflatoxin-cluster regions (Tables 1.1, 1.2). The 14 regions from *aflS* to *aflJ* and *aflE to aflW* examined in additional representative strains within VCGs yielded 24 distinct haplotypes, indicating that 89% of all haplotype-specific variation in the aflatoxin cluster of *A. flavus* resides in the late pathway genes.

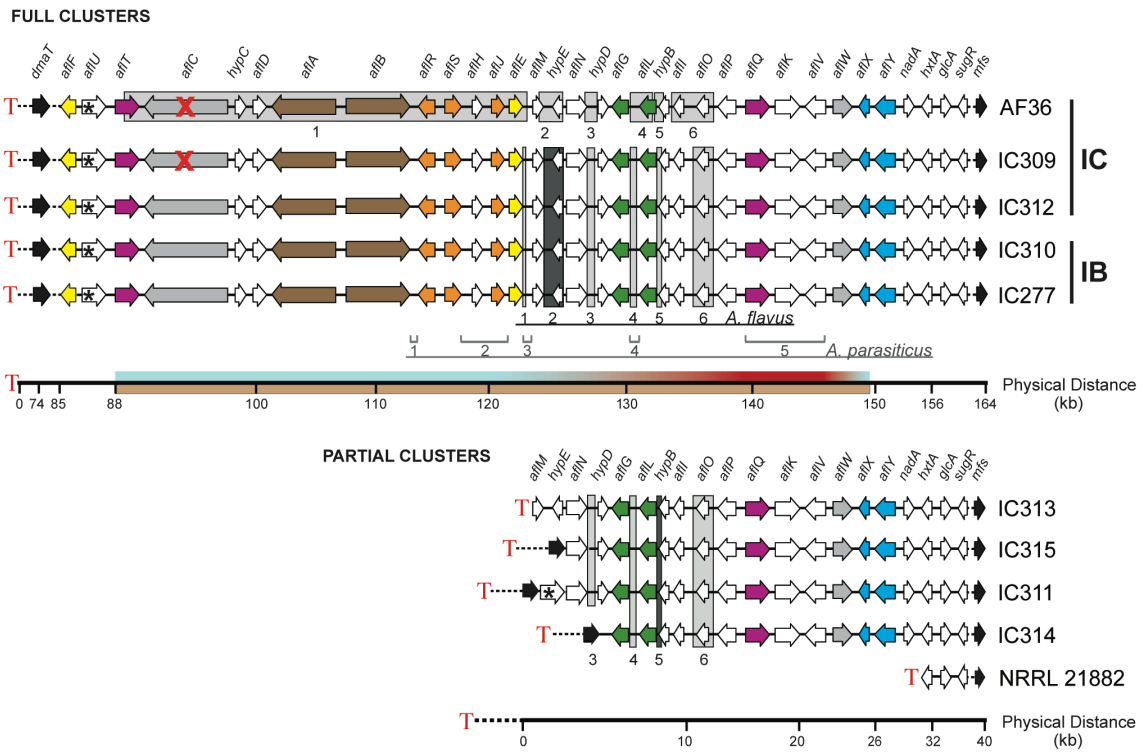


Figure 1.1 Schematic representation of full and partial aflatoxin gene clusters in *A. flavus*. The aflatoxin gene cluster starting with *aflF* is approximately 85 kb from the telomeric end of chromosome 3R, which is indicated with a 'T'. A total of 18 putative transcripts have been identified in this telomeric region (Chang *et al.* 2009). One of these transcripts, encoding for dimethylallyl tryptophan synthase (*dmaT*) and involved in CPA biosynthesis, is approximately 10 kb from *aflF* and is present in all complete *A. flavus* gene clusters, as well as in three of the partial cluster deletion strains (IC311, IC314, and IC315). As reported by Chang *et al* (2005), only one of these partial cluster strains (IC311) synthesizes CPA; IC314 and IC315 are reported to be missing two other genes important in CPA biosynthesis: monoamine oxidase (*maoA*) and the hybrid polyketide and non-ribosomal peptide synthase gene (*pks–nrps*). The nonsense mutation in *aflC* for AF36 and IC309 is designated with a 'X'. Genes shown in the same color arise from gene duplication (Carbone *et al.* 2007b). The six LD blocks in *A. flavus* from the concatenation of 21 cluster regions are boxed and numbered; the position of the five *A. parasiticus* blocks is shown below the *A. flavus* blocks and is from a previous study (Carbone *et al.* 2007a). The real size of LD blocks in *A. flavus* is shown in AF36 (e.g. block 1 spans *aflT-aflE*). LD blocks that have a different evolutionary history (i.e. incompatible partitions) are shown in a darker shade of gray (e.g. block 2 in complete clusters and block 5 in partial clusters). The upper scale bar shows real distance (kb) from the telomere across the CPA (*dmaT* only) and aflatoxin gene clusters, and is based on the genome strain of *A. flavus* NRRL 3357. The dashed segments between the telomere and *dmaT* in full and partial cluster strains denote variation in the length of the subtelomeric region. The scale bar for the partial cluster strains shows only real distance between contiguous aflatoxin cluster genes. A recombination hotspot in full clusters is shown in red above the scale bar and recombination coldspots are shown in blue. Full cluster strains belonging to groups IC and IB (see Fig. 1.4) are indicated; all partial cluster strains are in IB.

Linkage disequilibrium and compatibility analyses

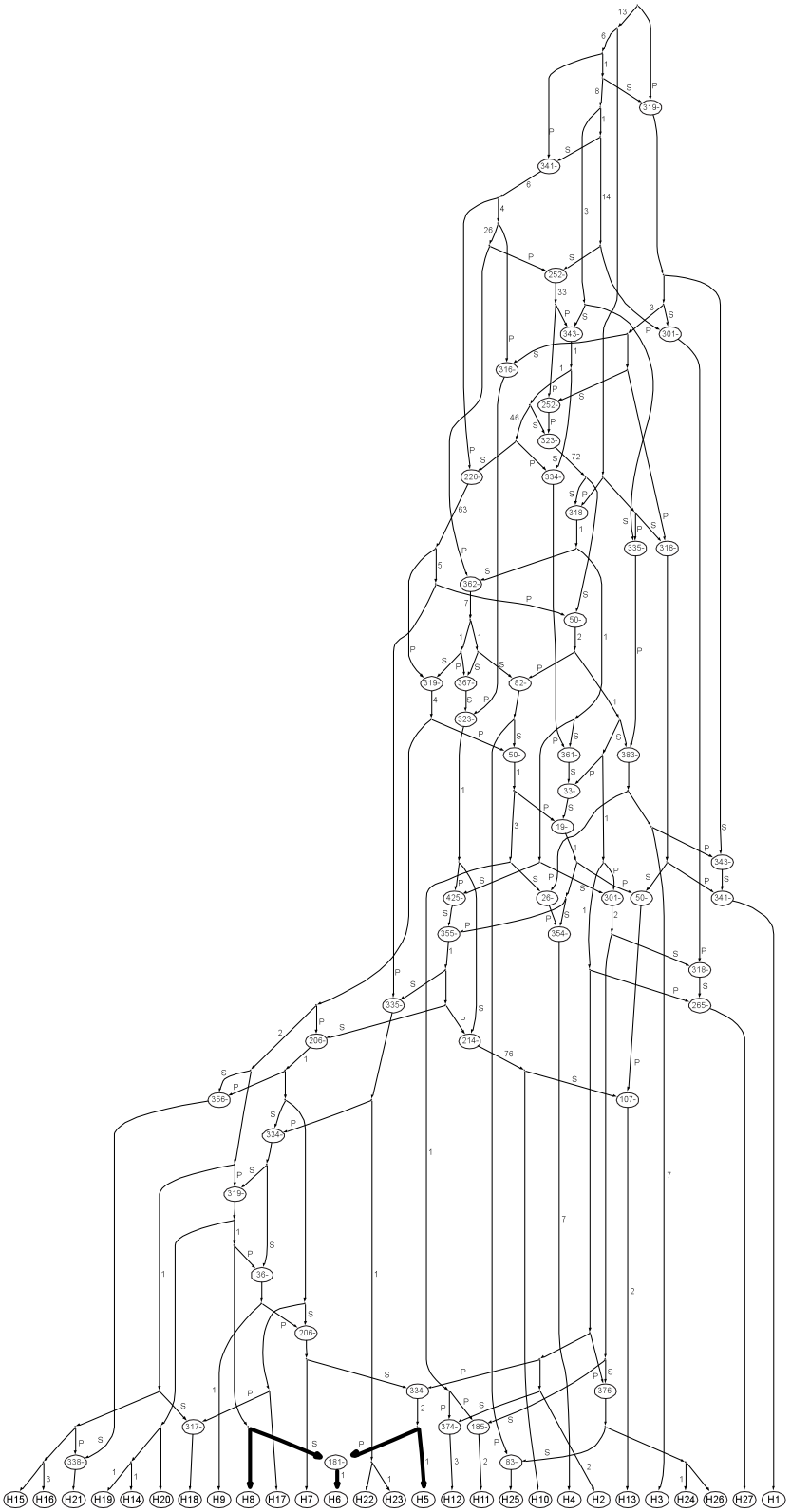
Linkage was examined for 453 polymorphisms distributed across the 21 regions of the aflatoxin gene cluster in *A. flavus* (Table 1.2). We identified six distinct blocks spanning the intergenic regions of late-pathway genes: *aflE/aflM, aflM/aflN (hypE)*, *aflN/aflG, aflG/aflL*, *aflL/aflI (hypB)* and *aflI/aflO* (Table 1.2; Fig. S1.1, Supporting Information). The early pathway regions from *aflT* to *aflE* (not inclusive) accounted for only 8% (36/453) of the total nucleotide sequence variation. Only one 3'-untranscribed region, *aflW/aflX,* showed a large number of fixed polymorphisms for a nonaflatoxigenic haplotype H1 (IC277); all other strains shared the consensus sequence (Fig. S1.1). In Figure S1.2 (Supporting Information), the six putative blocks identified from the 21 sequenced regions are shaded in the LD plot of complete cluster sequences in which only the lower triangular portion of the LD plot, showing correlations among pairs of sites, can be used for inference of block structure since significance of linkage cannot be assessed with only four strains. The upper triangular portion shows compatibility relationships for all pairs of sites. As expected, blocks are larger in size and may span several intergenic regions and genes when based on full cluster sequences (Fig. S1.2). For example, block 1 includes all the early pathway genes from *aflT* to *aflE,* blocks 2-5 each have one boundary in a gene and the other in an intergenic region, and block 6 extends from the middle of *hypB* in *aflL/aflI* to the *aflO/aflP* intergenic region (Fig. S1.2). There is rapid decay in LD with physical distance between *aflP* and *aflW*, as revealed in the LD plot for the 21 concatenated regions (Fig. S1.1), and for the four complete *A. flavus* clusters (Fig. S1.2). LD analysis for homologous regions in full and partial clusters could only be performed for regions spanning blocks 3-6; only block 5 was incompatible

with the other blocks (Fig. S1.3; Table S1.1, Supporting Information).

Compatibility analysis using RecMin revealed a minimum of 24 recombination events across the 21 concatenated regions. Recombination events were unevenly distributed, with 3, 6, 13 and 2 inferred breaks spanning *aflT-aflE, aflE-aflP, aflP-aflW,* and *aflW-aflX* regions, respectively (Fig. S1.1). This results in an overall population recombination rate of 0.000394 per generation per base pair for intergenic regions. Based on significance of linkage and compatibility analysis with CladeEx, the six blocks could be grouped into two distinct evolutionary histories: block 2 (shaded in Fig. S1.1) and combined blocks 1, 3, 4, 5, and 6. LD breaks down completely between *aflP* and *aflW* but is restored from *aflW* to *sugR* (Figs. S1.1, S1.2).

The ARG indicates extensive recombination in the ancestral history of the aflatoxin cluster in *A. flavus* (Fig. 1.2). Although many recombination and mutational events are required to interconnect all haplotypes to a most recent common ancestor in the past, it is clear from the ARG that many haplotypes are very closely related and recently evolved, often arising from only single crossovers and one to three mutations. For example, proceeding from left to right in the ARG: clades (H15, H16), (H19, H14, H20), (H22, H23) and (H24, H26) are separated by mutation alone; haplotypes H21, H18, and H25 arise from a single crossover; and H6, H5, H12 and H11 have evolved by both mutation and recombination.

Figure 1.2 One possible minimal ancestral recom-bination graph (ARG) inferred using the Kwarg heuristic method for all polymorphisms in the 21 regions of the aflatoxin gene cluster in *A. flavus.* The ARG shows possible mutation and recombination paths giving rise to the sampled haplotypes. The direction of paths is from the top of the ARG (past) to the bottom (present); moving backwards in time one of three events (mutation, coalescence or recombination) is possible. The paths leading to the recombination nodes (ovals) are labeled with a P (prefix) of S (suffix), indicating the 5' and 3' segments of the recombinant sequence, respectively; the number labels on the paths indicate the number of polymorphisms. The number in the ovals indicates the variable position immediately to the left of the recombination breakpoint. The paths leading to the putative parental haplotypes (H8, H5) of the recombinant (H6) are thickened. The ARG was rooted with haplotype H1 (IC277).


H15 H16 H21 H19 H14 H20 H18 H9 H8 H17 H7 H6 H22 H23 H5 H12 H11 H25 H10 H4 H2 H13 H3 H24 H26 H27 H1

Telomeric regions and G + C content

The distribution of G + C content along 200 kb of the right arm of chromosome 3 is shown in Figure 1.3. There is a strong bias for higher G + C content as the distance from the telomere increases and G + C content reaches a maximum with the first gene in the aflatoxin cluster, *aflF,* shown in the *z'* curve as a sharp downward peak. The downward peak is followed by a sharp decrease in G + C content seen as an upward slope in the *z'* curve until reaching *aflE,* after which G + C content stabilizes and follows a sinusoidal *z'* curve ending with *aflY*. The end of the cluster marks another GC-rich region. Mapping the distribution of recombination blocks and hotspots below the *z'* curve (shown as a highlighted bar in Fig. 1.3) shows that the early pathway genes (*aflF* to *aflE*), with reduced sequence diversity and significant LD, have higher G + C content than the late pathway regions (*aflE* to *aflY*), which harbor the six LD blocks and a recombination hotspot. There is a localized increase in G + C content in the region of the recombination hotspot from *aflI* to *aflW*. Similarly, the G + C content of the genes involved in CPA biosynthesis (*maoA*, *dmaT, pks–nrps*) is relatively constant until after the *maoA* gene where it reaches a maximum with *aflF*.

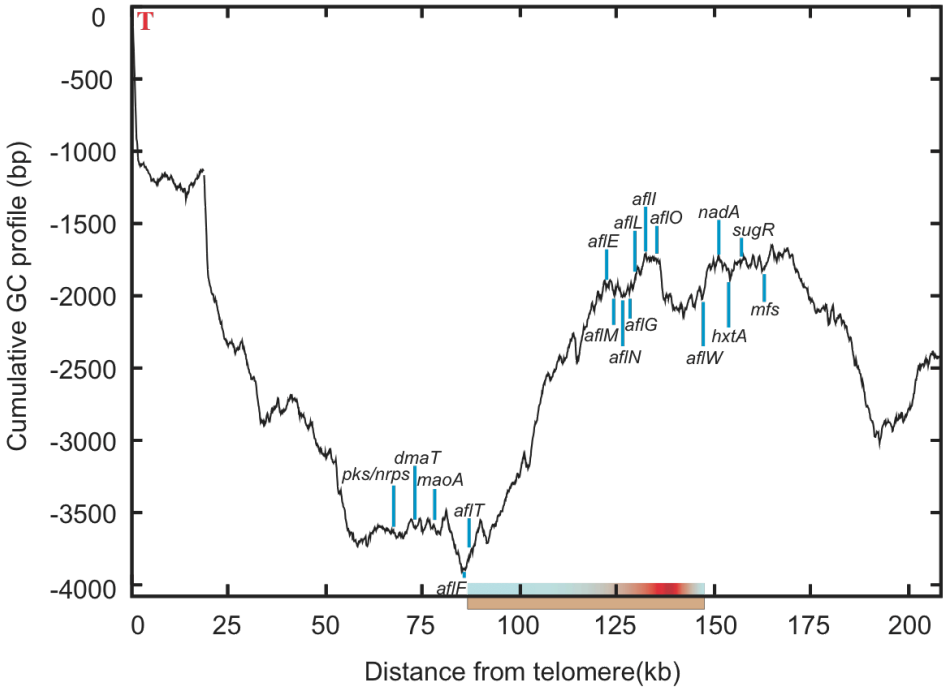
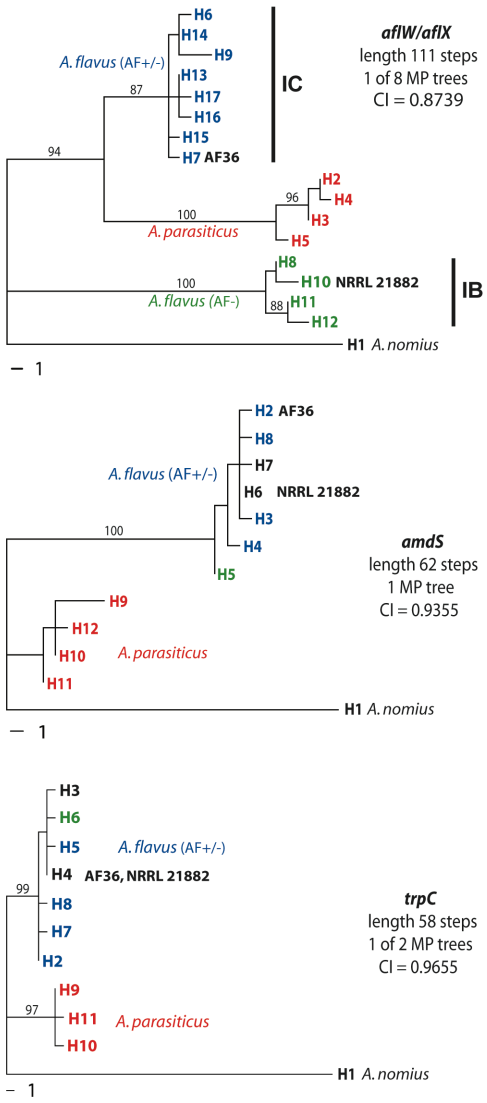


Figure 1.3 The cumulative GC profile (or *z'* curve) for 200 kb of chromosome 3R in *A. flavus* NRRL 3357. The AT-rich telomere is indicated with a 'T'. On the y-axis, -4000 bp indicates the lowest point in the *z'* curve at a distance of 85 kb from the telomere, and signifies that approximately 4000 G + C bases are in excess of the average G + C content calculated for this sub-region. G + C content increases with distance from the telomere and reaches a maximum with the first gene in the aflatoxin cluster, *aflF* at ~85 kb, and then a minimum at about 133 kb with *aflI*. The higher G + C content at ~85 kb also coincides with the 3' end of the CPA cluster, after *maoA*; the other maximum at 190 kb marks the end of the aflatoxin and sugar clusters. The scale bar shows the distribution of recombination hotspots and coldspots in the aflatoxin cluster; the recombination hotspot is localized in a region of relatively higher G + C content at about 140 kb in the late pathway genes.

Trans-species evolution

A total of 17, 11 and 12 haplotypes were inferred for the *aflW/aflX, trpC* and *amdS* regions, respectively, for the expanded *A. flavus* and *A. parasiticus* sample. The haplotype distribution in these regions is shown in Figure 1.4 (Table S1.2, Supporting Information).

Figure 1.4 One most parsimonious (MP) phylogeny inferred for *aflW/aflX* and two non-cluster loci (*amdS* and *trpC)*. Taxa include *A. flavus* and *A. parasiticus* sampled from Georgia, the two biocontrol *A. flavus* strains AF36 and NRRL 21882, the genome strains of *A. flavus* NRRL 3357 and *A. oryzae* NRRL 5590, eight nonaflatoxigenic *A. flavus* strains that represent distinct deletion patterns in the aflatoxin cluster, and *A. nomius* NRRL 13137 (GenBank Accessions AY510454, GQ479323, and GQ479322 for *aflW/aflX, trpC* and *amdS,* respectively) as the outgroup species. Each phylogeny is labeled with its respective gene name, tree length, number of MP trees, and consistency index (CI). The scale bar below each tree represents a single character state change. For each phylogeny, the distribution of isolates among haplotypes and whether isolates are aflatoxin producers (+) or nonproducers (-) are shown in Table S1.2 (Supporting Information). The *aflW/aflX* phylogeny shows two distinct *A. flavus* aflatoxin gene clades: one clade (IC, shown in blue) includes aflatoxigenic and nonaflatoxigenic strains (AF+/-), including AF36, and shares a more recent common ancestor with *A. parasiticus* (shown in red)*;* a second clade (IB, shown in green) is completely nonaflatoxigenic (AF-) and includes complete/partial cluster *A. flavus* strains as well as NRRL 21882. With the exception of *A. nomius,* haplotypes shown in black for *amdS* and *trpC* include strains from both *A. flavus* lineages. The AF36 and NRRL 21882 biocontrol strains show divergent origins only in *aflW/aflX*. As expected, the non-cluster genes are orthologous and show no evidence of trans-speciation. Bootstrap values are shown for branches with greater than 80% support.



Phylogenetic analysis of the *aflW/aflX* region revealed two distinct *A. flavus* lineages (Fig. 1.4). One lineage containing both aflatoxigenic and nonaflatoxigenic strains, including AF36, shares a more recent common ancestor with *A. parasiticus* than a second distinct *A. flavus* lineage, which is completely nonaflatoxigenic and includes complete/partial cluster

strains and NRRL 21882. These two *A. flavus* aflatoxin gene lineages, or clades, correspond to groups IC and IB, respectively (Geiser *et al.* 2000). As expected, these patterns of trans-speciation were not observed in *trpC* and *amdS* phylogenies. The partitioning of *A. flavus* into clades IB and IC in *aflW/aflX* was incongruent with the monophyly of *A. flavus* in the *trpC* and *amdS* organismal genealogies. Figure S1.4 (Table S1.2) shows the strict consensus phylogenies and ARGs inferred using branch and bound (beagle) and heuristic (kwarg) search methods. Although we observed three distinct clades in *aflW/aflX* when we assume no recombination, the ARGs indicate a history of genetic exchange and recombination. For example, in the branch and bound and heuristic ARGs for *aflW/aflX* (Fig. S1.4, Table S1.2), haplotype H9 arises from a single crossover with putative parents haplotype H8, belonging to the strictly nonaflatoxigenic clade, and haplotype H14, which is a member of the *A. flavus* clade sharing a recent common ancestor with *A. parasiticus* (paths to parents determined by lack of mutations and additional recombination events are shown in bold in both ARGs in Fig. S1.4). Also, *aflW/aflX* shows evidence of an *A. parasiticus* G_1-dominant (high G_1/B_1 ratio) haplotype H5 acting as one of the parents in a recombination event in the ancestor of all *A. parasiticus* haplotypes represented in H2, H3, and H4. The *A. parasiticus* OMST-producing haplotypes H9 and H10 show a recent history of recombination with *A. flavus* in the ARGs for *amdS* and *trpC*, respectively.

Neutrality tests and coalescent analyses

All neutrality tests were non-significant after Bonferroni correction. Negative values of *D* and F_S were obtained for early pathway genes spanning *aflT* to *aflE* (13 of 18 tests) while

positive values of D and F_S were observed among late pathway genes from *aflE* to *aflX* (18 of 24 tests). Collectively, this reversal in values indicates that recombining regions in the cluster are under strong balancing selection for maintaining chemotype-specific differences – in this case aflatoxigenicity versus nonaflatoxigenicity – while regions of low recombination spanning the early pathway genes are swept to fixation by selection.

The largest non-recombining partition for complete and partial clusters included the intergenic regions: *aflI/aflO, aflO/aflP, aflP/aflQ, aflQ/aflK, aflK/aflV, aflV/aflW* and *aflW/aflX*. This corresponds to a total concatenated sequence length of 2605 bp, 184 polymorphic sites, and 20 haplotypes (Fig. 1.5; Table S1.3, Supporting Information). The *aflG/aflL* region could not be amplified in IC314 and was excluded in the concatenation. The *aflL/aflI* region was incompatible with the largest non-recombining partition and excluded from the analysis. The genealogy with the highest root probability (Fig. 1.5; Table S1.3) shows two distinct evolutionary lineages (IB and IC) as observed in the *aflW/aflX* region (Fig. 1.4). The strong genetic differentiation observed in the cluster is the result of balancing selection acting on a nonaflatoxigenic lineage (IB, haplotypes in bold type in Fig. 1.5 and Table S1.3) that includes all the aflatoxin cluster deletion strains from the sample we examined. The rooted genealogy shows that partial clusters are recent descendents of full clusters. For example, IC311 (H13) and IC315 (H14) from Alabama share a recent common ancestor and descended from a full cluster ancestor with a mutation configuration similar to IC277 (H17). Similarly, partial deletion strain IC314 (H10, which includes *A. oryzae* NRRL 5590) and full cluster strain IC310 (H8) from North Carolina share a recent common ancestor (Fig. 1.5).

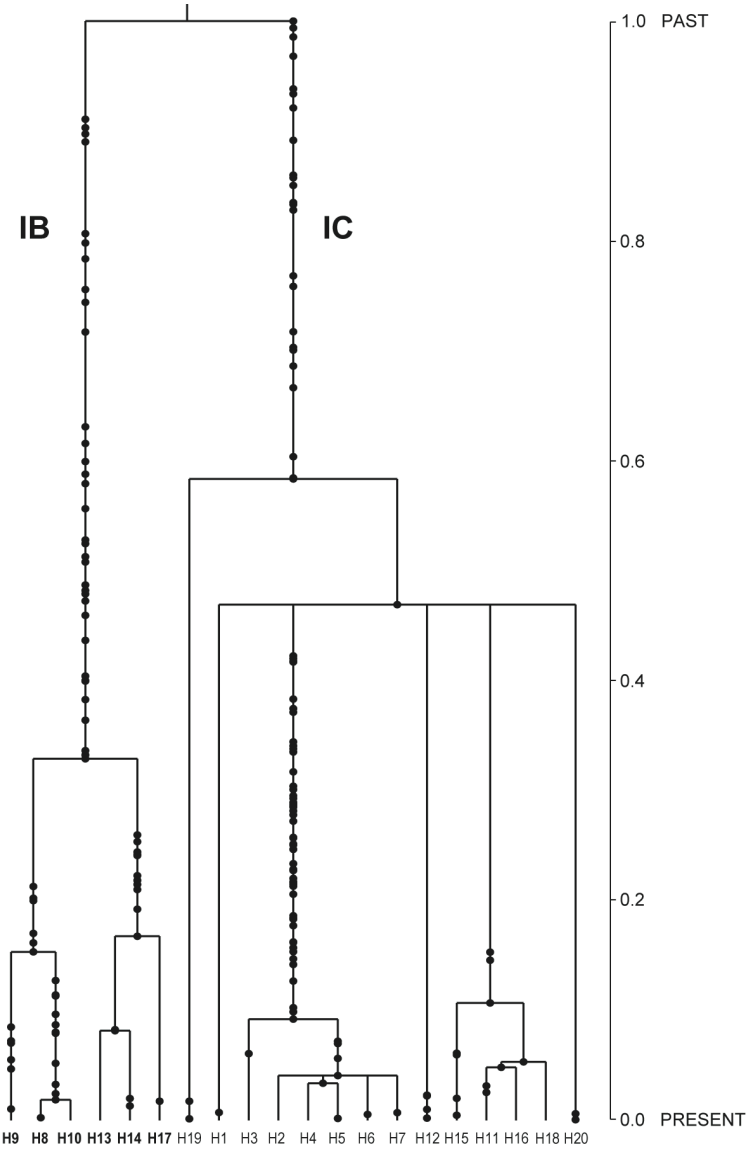


Figure 1.5 The rooted coalescent-based gene genealogy for the largest non-recombining partition in full and partial gene clusters. Mutations are shown as dots along the branches. The estimated mean time to the most recent common ancestor (TMRCA) in coalescent time units is 2.07 (SD = 0.02). The scale at the right shows time rescaled to a TMRCA of 1.0. All partial cluster strains examined are in a distinct nonaflatoxigenic evolutionary clade (IB), which comprises haplotypes shown in bold. The other distinct clade (IC) includes AF+/- strains. The gene genealogy with the highest root probability was selected based on maximum-likelihood estimates of all possible rooted trees using genetree (likelihood of best tree = 1.16×10^{-80}, SD = 3.64×10^{-77}, $\theta = 40.72$). Convergence to the best tree was achieved within six independent runs of 5 million simulations each and one 10 million simulation, assuming panmixia and constant population size. The distribution of strains among haplotypes is shown in Table S1.3 (Supporting Information).

Similar patterns of descent among full and partial aflatoxin clusters were observed in regions flanking the aflatoxin gene cluster (Fig. 1.6; Table S1.4, Supporting Information). In the *dmaT* gene genealogy, there was no clear separation of partial and full aflatoxin cluster strains; however, the genealogy does reveal two distinct evolutionary lineages similar to the balancing selection observed in the aflatoxin late pathway genes. In the *mfs* gene genealogy, the full-cluster nonaflatoxigenic haplotype H8 (IC277) is ancestral to partial cluster haplotype H9 (IC315) and to haplotype H1, which includes a mixture of partial (IC313, IC314) and full cluster strains (*A. oryzae* NRRL 5590*,* IC310). In rooting genealogies, haplotypes with no mutations along the edges of the genealogy (branches) occupy interior positions and serve as a root from which all mutation paths to descendant haplotypes can be retraced (Griffiths & Tavaré 1994). The immediate descendents of H1 are partial cluster haplotypes H6 (IC311) and H11 (IC252 and NRRL 21882). There is no signature of balancing selection in *mfs* positioned at the 3' distal end of the cluster.

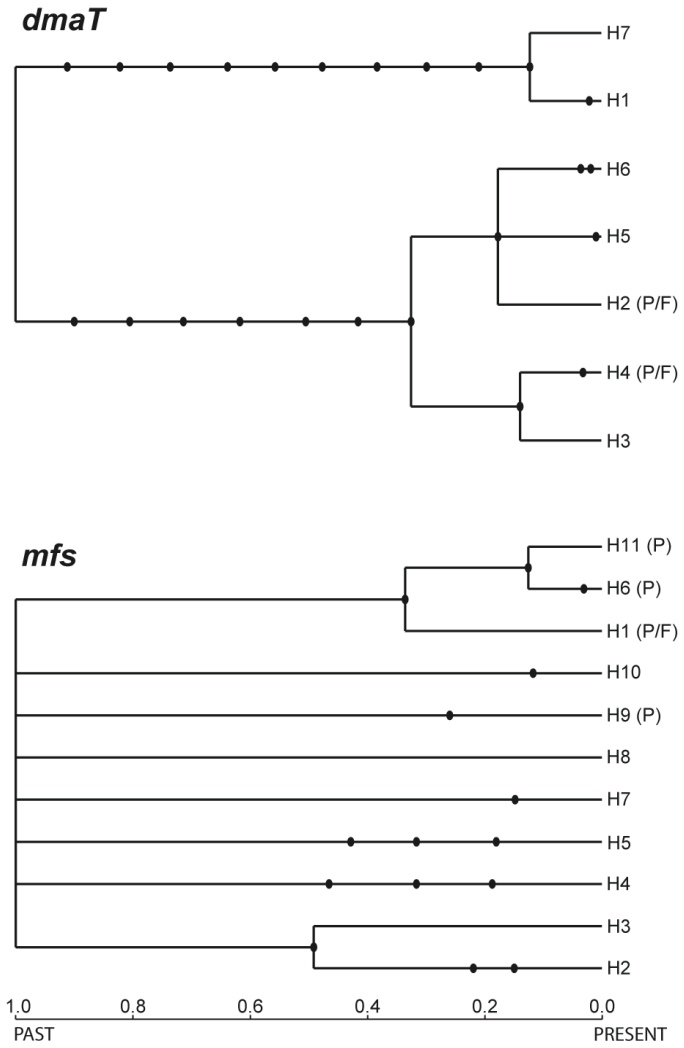


Figure 1.6 Coalescent-based gene genealogies for *dmaT* and *mfs*. Mutations are indicated with dots along branches and time is scaled to the most recent common ancestor (TMRCA) of 1.0 for each locus. The mean TMRCA for *dmaT* is 3.32 (SD = 0.69) and for *mfs* is 1.67 (SD = 0.48). The haplotype structure for each genealogy is shown in Table S1.4 (Supporting Information). Haplotypes that contain partial (P) or full (F) aflatoxin cluster strains are indicated in the genealogies. The *dmaT* rooted genealogy (likelihood of best tree = 1.0265×10^{-13}, SD = 4.4358×10^{-12}, $\theta = 5.22$) indicates two divergent lineages similar to, but topologically discordant with, the aflatoxin cluster-based genealogy in Fig. 1.5. In the *mfs* rooted genealogy (likelihood of best tree = 5.9185×10^{-13}, SD = 6.5180×10^{-12}, $\theta = 3.69$), haplotype H8 (IC277) is the most recent common ancestor of the sample, and partial cluster strains in haplotypes H6 and H11 are the most recently evolved. At least three independent runs of 1-10 million simulations each were examined to ensure convergence to the best tree.

DISCUSSION

The early portion of the cluster from *aflT* to *aflE* (~ 35 kb) showed 10-fold lower nucleotide diversity than the late pathway segment from *aflE* to *aflP* (~ 13 kb). The reduced sequence diversity could be the result of functional constraints on pathway genes such as *aflC*, selection acting on adjacent noncluster genes such as *dmaT,* or differences in gene regulation or chromatin organization. We cannot rule out genetic hitchhiking effects due to the proximity of the CPA and aflatoxin gene clusters independent of whether they are co-regulated or not. In the absence of recombination we would predict adjacent genes or clusters to be dragged through the selection process. With recombination between *maoA* and *aflF*, as well as in the late aflatoxin pathway genes starting with *aflE/aflM,* we would predict that hitchhiking would break down at these junctions. This is consistent with the observed variation in mycotoxin production, with individuals producing both aflatoxins and CPA, aflatoxins alone, CPA alone, or neither mycotoxin. Only 25% of nonaflatoxigenic strains examined in this study had the nonsense mutation in *aflC*. In the study by Criseo *et al* (2008), gene loss and nucleotide variation in the early pathway genes *aflD* and/or *aflR* comprised about 20% of the 134-nonaflatoxigenic *A. flavus* strains that they examined. This suggests that the majority of nonaflatoxigenic strains have early pathway genes. The abrupt transition in G + C content in this region (Fig. 1.3) indicates the potential for the entire segment from *aflT* to *aflE* to be exchanged or lost as a single block during the shuffling process (Fig. 1.1). Only 12% (3/24) of recombination events fall in the early part of the pathway, which further supports the tight linkage of these genes.

LD analysis for homologous regions in full and partial clusters revealed that polymorphisms in deletion strains align with blocks 3-6 (Fig. S1.3; Table S1.1), indicating a shared history of recombination in complete and partial clusters; however block 1, which spans *aflT* to *aflE,* is the largest block and is missing in all partial clusters (Chang *et al.* 2005; Yin *et al.* 2009). Two recent studies examining nonaflatoxigenic *A. flavus* strains isolated from peanut fields in China (Yin *et al.* 2009) and from food and feed commodities in Italy (Criseo *et al.* 2008) report a high frequency of cluster deletions, which could account for more than 60% of the nonaflatoxigenic trait (Criseo *et al.* 2008). All deletion patterns reported so far are consistent with the block structure we report. In all cases, missing genes in partial clusters are flanked by recombination blocks in complete clusters; for example, loss of *aflN* in deletion pattern D in Yin *et al* (2009) is flanked by *A. flavus* blocks 2 and 3 (Fig. 1.1). Similar deletion patterns have been reported in *A. oryzae* groups 2 and 3 (Kusumoto *et al.* 2000) and observed in *A. oryzae* IC902 (=RIB 430) (I Carbone, unpublished data). Collectively, these patterns are consistent with the hypothesis of directed stepwise deletions, which postulates that gene loss begins at the *aflF* end of the aflatoxin cluster and ends in the late pathway segment (Kusumoto *et al.* 2000). Because the position of recombination block boundaries coincides with deletion patterns, it might be possible to predict the existence of other partial aflatoxin clusters by using these boundaries as a guide.

The recombination blocks we identified in *A. flavus* coincide with the sequence breakpoints in deletion variants reported previously (Chang *et al.* 2005). For example, deletion strain IC313 is missing all the early pathway genes upstream of *aflM,* consistent with a double-strand break (DSB) in the *aflE/aflM* region (block 1). Similarly, strain IC315

could arise from a DSB in the *aflM/aflN* region (block 2) and IC314 from a DSB in the *aflN/aflG* region (block 3) (Fig. 1.1). Most DSBs, which presumably give rise to deletion variants, appear to be located in intergenic regions containing transcriptional promoters at the junction of recombination blocks 1, 2 and 3. While DSBs can occur in genes, as observed in Figure S1.2 for blocks 2 (*aflM*), 3 (*aflN*), 4 and 5 (*aflL*), most are located in intergenic promoter-containing regions (Baudat & Nicolas 1997). There are no reports of deletion variants associated with blocks 4, 5 and 6, but it is possible that variants exist as these blocks are in a region of increasing G + C content.

The clear transition in G + C content between the CPA and aflatoxin gene clusters (Fig. 1.3) suggests a possible recombination hotspot between the two gene clusters. This might explain why gene loss in the aflatoxin gene cluster sometimes involves only aflatoxin cluster genes, as seen in IC311, which makes CPA but not aflatoxins as a result of a deletion of the *aflT-aflM* region (Fig. 1.1). A recombination hotspot between the two gene clusters might also explain the observed diversity in mycotoxin production in populations, from which it is possible to isolate strains that are AF+/CPA+, AF+/CPA-, AF-/CPA+, and AF-/CPA- (Chang *et al.* 2009; Horn & Dorner 1999).

Although the mechanism that gives rise to deletion variants in *A. flavus* is unclear, crossing over between full and partial clusters during sexual reproduction is one possibility. Alternatively, intra-genomic translocations may move cluster genes to other chromosomes. This may involve translocations of entire telomeric/subtelomeric regions between chromosomes (Smith *et al.* 2007) or translocations within a single gene such as *stcW* (=*aflW*)

of the sterigmatocystin gene cluster which disrupts the open reading frame (Hodges *et al.* 2000). In addition, movement of cluster genes may be due to duplication followed by gene loss (Carbone *et al.* 2007b; Powell *et al.* 2008); a partial cluster duplication of genes from *aflR* to *aflM* and *aflO* has been reported for several *A. parasiticus* strains (Chang & Yu 2002). Several late pathway genes, including the *aflW/aflX* region, have been amplified and sequenced in NRRL 21882, and *aflA, aflB, aflR, aflS* genes also have been sequenced in other representative strains within the same VCG 24 (I Carbone, unpublished data). Although these genes are present, they are not associated with the aflatoxin gene cluster on the telomere end of chromosome 3R. Whether these cluster genes are the result of a translocation to another chromosomal location in a full cluster ancestor or a duplication of the original cluster to another genomic location followed by gene loss needs to be investigated further.

In previous work, we showed that gene modularity arising from duplications of a single gene is a common feature across many genes in aflatoxin biosynthesis, but such modularity is clearly more pronounced in the *aflT-aflE* region of the cluster (Carbone *et al.* 2007b). We now provide evidence that LD blocks and recombination hotspots in *A. flavus* may be responsible for the observed modularity, such that regions in the cluster with low recombination rates and reduced sequence diversity show greater modularity than regions that harbor distinct LD blocks and recombination hotspots. In Figure 1.1, the 3' terminal end of the cluster starting with *aflW,* which flanks the recombination hotspot, shows significant LD and a return to modularity (Carbone *et al.* 2007b). In *A. parasiticus,* block 5 spans *aflQ* to *aflV* (Carbone *et al.* 2007a), which in turn spans with the recombination hotspot in

A. flavus (Figs. 1.1, S1.1). At least two blocks in *A. parasiticus* span two or more cluster genes (blocks 2 and 5); these blocks are potentially larger and more encompassing of the cluster but this cannot be determined without the sequencing of full gene clusters in *A. parasiticus.* Only two small LD regions (blocks 3 and 4) were detected in the span of *aflE* to *aflP* in *A. parasiticus*; the same spanned region in *A. flavus* harbors six distinct LD blocks (Fig. 1.1). The reduced recombination in *aflE* to *aflP* and *aflQ* to *aflW* simplifies the ARG in *A. parasiticus* (Fig. 2B in Carbone *et al.* 2007a) compared to the ARG in *A. flavus* (Fig. 1.2). Although recombination rates are approximately 10-fold lower in *A. parasiticus* than *A. flavus* (0.000394 versus 0.0011), the existence of common blocks (*A. flavus* block 1 with *A. parasiticus* block 3, and *A. flavus* block 4 with *A. parasiticus* block 4) indicates shared common ancestry in the history of these species. The different rates of recombination in *A. flavus* and *A. parasiticus* may be the result of a larger effective mating population in *A. flavus* than in *A. parasiticus*. VCG diversity in the single peanut field in Georgia was higher for *A. flavus* than *A. parasiticus,* with 44 and 17 VCGs, respectively (Horn & Greene 1995). Because mating in *A. flavus* and *A. parasiticus* is between parents of different VCGs (Horn *et al.* 2009a; Horn *et al.* 2009b, c) and results in progeny that belong to new VCGs (BW Horn, unpublished data), ongoing sexual reproduction would continue to yield disproportionate numbers of VCGs and as a consequence, different rates of recombination in these fungi.

In *A. parasiticus*, trans-speciation and balancing selection acting on a G_1-dominant lineage was observed in *aflM/aflN* (*hypE*), which is a relatively large nonrecombining region of the *A. parasiticus* cluster. Similarly, in *A. flavus,* the nonrecombining *aflW/aflX* shows a strong signature of balancing selection, which is acting on a nonaflatoxigenic lineage (Fig.

S1.1). *AflW* is a flavin-dependent oxidase and *aflX* is an NAD(P)H-dependent oxidoreductase that may be involved in electron transfer to the flavin (Cary *et al.* 2006). Positive selection for these genes may explain the retention of a basal cluster even in nonaflatoxin-producing strains of *A. flavus*. All six LD blocks in *A. flavus* show two highly divergent haplotypes; however, within each block there is no specific association of haplotype with either an aflatoxigenic or a nonaflatoxigenic trait. Only the *aflW/aflX* region in our population study, and possibly the genes and intergenic regions from *aflW* to *hxtA* (Fig. S1.2), show two distinct haplotypes in which one of the two haplotypes is a nonaflatoxigenic strain (IC277 in Fig. S1.1, and *A. oryzae* NRRL 5590 in Fig. S1.2). Geiser *et al.* (1998) also showed disproportionately more polymorphisms in *omtA* (=*aflP*) than in four non-cluster genes within *A. flavus* group I. Moreover, *aflP,* which is located in the recombination hotspot, has been reported to harbor a greater number of phylogenetically informative sites than other genes in the cluster (Chang *et al.* 2006). Both distinct *A. flavus* lineages that we report in this paper are in group I of Geiser *et al.* (1998). The existence of a strictly nonaflatoxigenic clade (IB) in the *aflW/aflX* region (Fig. 1.4) suggests that selection is acting to maintain this phenotype in the population. The strains examined with deletion patterns B (IC310), C (IC311), E (IC313), F (IC314) and G (IC315) share the same nonaflatoxigenic clade in *aflW/aflX.* The specific genes under balancing selection for the nonaflatoxigenic phenotype are not obvious when examining other cluster regions, as gene conversion or double crossovers will shuffle highly divergent sequences among aflatoxigenic and nonaflatoxigenic strains. This process results in alternating haplotype patterns among blocks such that one common haplotype in one block transitions to a different one in the

adjacent block (see haplotype polymorphism map at bottom of Fig. S1.1).

If selection were favoring the nonaflatoxigenic phenotype, then *aflC* (*pksA*) and *aflM* (*ver-1*) would seem to be the ideal candidates to target. The aflatoxin precursor, versicolorin A, has been shown to accumulate in strains that are nonaflatoxigenic but have a functional *aflC,* and both *aflX* and *aflY* are integral in the conversion of versicolorin A to demethylsterigmatocystin (Cary *et al.* 2006; Ehrlich *et al.* 2005) as are *aflM* and *aflN* (Yu *et al.* 2004). Versicolorin A also accumulates in toxigenic strains, leading us to believe that it is the rate-limiting step before the split into B_1 or B_2 production (D. Bhatnagar, personal communication). Although we identified a distinct nonaflatoxigenic clade, the actual target of selection may be ecological adaptiveness rather than the loss of aflatoxin production per se. *A. flavus* strains vary considerably in factors influencing aggressiveness in crop invasion, such as the production of plant-degrading enzymes (Brown *et al.* 1992; Cotty *et al.* 1990; Shieh *et al.* 1997). There appears to be a positive correlation between regions with a high frequency of drought and *A. flavus* soil populations with high aflatoxigenicity (Dorner & Horn 2007; Horn & Dorner 1999). Since crops are susceptible to *A. flavus* under drought-stress conditions, this suggests that nonaflatoxigenic *A. flavus* strains may be less aggressive than aflatoxigenic strains in invading crops. Alternatively, nonaflatoxigenic strains might be better adapted to other as yet undefined ecological niches. This warrants further investigation.

Balancing selection in the nonaflatoxigenic lineage was not evident in the 5' end of the cluster corresponding to the early pathway genes, which suggests that these genes are more prone to loss. Further upstream, a phylogenetic signature of balancing selection was

observed in *dmaT* (Fig. 1.6), which encodes a critical enzyme in CPA biosynthesis (Chang *et al.* 2009); however, the *dmaT* genealogy is incongruent with the aflatoxin cluster genealogy and as expected, loss of genes in the 5' end of the aflatoxin cluster does not always result in a concurrent loss of genes in the CPA cluster. With the exception of IC313 and NRRL 21882, all deletion strains examined contain an intact *dmaT* gene as well as several additional genes between the telomere and *dmaT*. It appears that *dmaT* is functional for at least one deletion strain (IC311) and one nonaflatoxigenic strain with a full cluster (IC277).

Comparative genome hybridization of *A. oryzae* NRRL 5590 and *A. flavus* IC277 to the *A. flavus* NRRL 3357 reference genome reveals deletions of approximately 75 and 45 kb in size, respectively, from the telomere ends of these strains (I Carbone, unpublished data). This indicates that gene loss is also occurring in NRRL 5590 and IC277 full cluster strains, at the distal end of chromosome 3R. Deletions in the telomere of IC277 do not extend to the CPA and aflatoxin clusters, whereas in NRRL 5590, the breakpoint is in the CPA gene cluster. One possible explanation is that gene clusters slow down the deletion process, perhaps through the actions of balancing selection in these clusters. The only exception is NRRL 21882 in which all telomeric genes are missing and the telomeric sequence is immediately adjacent to *hxtA*.

Results from coalescent simulations indicate that aflatoxin cluster deletions in *A. flavus* lineage IB are recent evolutionary events and that deletion strains are descendants of full cluster ancestors (Fig. 1.5). This has important implications for biological control using nonaflatoxigenic *A. flavus* strains. While our population study cannot exclude the possibility of a nonaflatoxigenic strain with a full gene cluster reacquiring toxicity genes, there is no

evidence that deletion strains can reassemble a full cluster via crossing over or gene conversion in the cluster. A single cross over between a partial cluster strain and a full cluster strain in the recombination hotspot region would give rise to progeny with full clusters that have the *aflW/aflX* region of partial clusters and vice-versa. Alternatively, progeny from either AF36 or NRRL 21882 can become aflatoxigenic by acquiring chromosome 3 from an aflatoxigenic parent, through the independent assortment of chromosomes during meiosis. We are currently investigating both possibilities in laboratory crosses.

Coalescent simulations of nonaflatoxin cluster *dmaT* and *mfs* gene genealogies show that the full cluster nonaflatoxigenic strains occupy interior positions in the genealogies and are therefore older than derived deletion strains (Fig. 1.6). Haplotypes that include deletion variants in interior positions such as H2 in *dmaT* and H1 in *mfs* also include full cluster strains (Fig. 1.6), which suggest a lack of resolution. These non-cluster gene genealogies further support our observation of lineage-specific gene loss in the *aflW/aflX* region (Fig. 1.4) and late pathway genes (Fig. 1.5). The existence of common LD blocks in full and partial cluster strains indicates that the process of crossing over in full clusters may be driving gene loss. Because only one or two mutations separate *A. flavus* partial deletion (IC314 [H10]) and full cluster (IC310 [H8]) strains (Fig. 1.5), it is possible that sexual recombination may be accelerating gene loss. The close similarity between full and partial cluster strains argues against the hypothesis that gene loss (partial or entire deletion) is due to genetic drift, when the aflatoxin cluster is no longer needed, long after the recombination events have ceased. In contrast, the shared ancestry and geographic proximity among IC310 and IC314 strains

indicate that cluster disassembly is recent and may be influenced by local population dynamics or environmental conditions. Future work will examine recombination, gene loss and heritability of toxin production in *A. flavus* offspring from sexual matings.

LITERATURE CITED

Aylor DL, Price EW, Carbone I (2006) SNAP: combine and map modules for multilocus population genetic analysis. *Bioinformatics* **22**, 1399-1401.

Baudat F, Nicolas A (1997) Clustering of meiotic double-strand breaks on yeast chromosome III. *Proceedings of the National Academy of Sciences, USA* **94**, 5213-5218.

Bayman P, Cotty PJ (1993) Genetic diversity in *Aspergillus flavus*: association with aflatoxin production and morphology. *Canadian Journal of Botany* **71**, 23-31.

Bowden LC, Price EW, Carbone I (2008) SNAP Clade and Matrix, Version 2. Distributed over the internet, http://www.cals.ncsu.edu/plantpath/faculty/carbone/home.html, Department of Plant Pathology, North Carolina State University.

Brown RL, Cleveland TE, Cotty PJ, Mellon JE (1992) Spread of *Aspergillus flavus* in cotton bolls, decay of intercarpellary membranes, and production of fungal pectinases. *Phytopathology* **82**, 462-467.

Burdock GA, Flamm WG (2000) Safety assessment of the mycotoxin cyclopiazonic acid. *International Journal of Toxicology* **19**, 195–218.

Carbone I, Jakobek JL, Ramirez-Prado JH, Horn BW (2007a) Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus*. *Molecular Ecology* **16**, 4401–4417.

Carbone I, Ramirez-Prado JH, Jakobek JL, Horn BW (2007b) Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. *BMC Evolutionary Biology* **7**, 111.

Cary JW, Ehrlich KC, Bland JM, Montalbano BG (2006) The aflatoxin biosynthesis cluster gene, *aflX*, encodes an oxidoreductase involved in conversion of versicolorin A to demethylsterigmatocystin. *Applied and Environmental Microbiology* **72**, 1096-1101.

Chang P-K, Ehrlich KC, Hua SS (2006) Cladal relatedness among *Aspergillus oryzae* isolates and *Aspergillus flavus* S and L morphotype isolates. *International Journal of Food Microbiology* **108**, 172-177.

Chang P-K, Horn BW, Dorner JW (2005) Sequence breakpoints in the aflatoxin biosynthesis gene cluster and flanking regions in nonaflatoxigenic *Aspergillus flavus* isolates. *Fungal Genetics and Biology* **42**, 914-923.

Chang P-K, Horn BW, Dorner JW (2009) Clustered genes involved in cyclopiazonic acid production are next to the aflatoxin biosynthesis gene cluster in *Aspergillus flavus*. *Fungal Genetics and Biology* **46**, 176-182.

Chang P-K, Yu J (2002) Characterization of a partial duplication of the aflatoxin gene cluster in *Aspergillus parasiticus* ATCC 56775. *Applied Microbiology and Biotechnology* **58**, 632-636.

Cole RJ (1986) Etiology of turkey "X" disease in retrospect: a case for the involvement of cyclopiazonic acid. *Mycotoxin Research* **2**, 3–7.

Cole RJ, Sanders TH, Hill RA, Blankenship PD (1985) Mean geocarposphere temperatures that induce preharvest aflatoxin contamination of peanuts under drought stress. *Mycopathologia* **91**, 41-46.

Cotty PJ (1989) Virulence and cultural characteristics of two *Aspergillus flavus* strains pathogenic on cotton. *Phytopathology* **79**, 808-814.

Cotty PJ (1994) Influence of field application of an atoxigenic strain of *Aspergillus flavus* on the populations of *A. flavus* infecting cotton bolls and on the aflatoxin content of cottonseed. *Phytopathology* **84**, 1270-1277.

Cotty PJ, Cleveland TE, Brown RL, Mellon JE (1990) Variation in polygalacturonase production among *Aspergillus flavus* isolates. *Applied and Environmental Microbiology* **56**, 3885-3887.

Criseo G, Racco C, Romeo O (2008) High genetic variability in nonaflatoxigenic *A. flavus* strains by using Quadruplex PCR-based assay. *International Journal of Food Microbiology* **125**, 341-343.

Dorner JW (2005) Biological control of aflatoxin crop contamination. In: *Aflatoxin and Food Safety* (ed. Abbas HK), pp. 333–352. Taylor and Francis, Boca Raton, Florida.

Dorner JW (2009a) Biological control of aflatoxin contamination in corn using a nontoxigenic strain of *Aspergillus flavus*. *Journal of Food Protection* **72**, 801-804.

Dorner JW (2009b) Development of biocontrol technology to manage aflatoxin contamination in peanuts. *Peanut Science* **36**, 60-67.

Dorner JW, Horn BW (2007) Separate and combined applications of nontoxigenic *Aspergillus flavus* and *A. parasiticus* for biocontrol of aflatoxin in peanuts. *Mycopathologia* **163**, 215-223.

Eaton DL, Groopman JD (1994) *The Toxicology of Aflatoxins: Human Health, Veterinary, and Agricultural Significance.* Academic Press, San Diego.

Ehrlich KC, Cotty PJ (2004) An isolate of *Aspergillus flavus* used to reduce aflatoxin contamination in cottonseed has a defective polyketide synthase gene. *Applied Microbiology and Biotechnology* **65**, 473-478.

Ehrlich KC, Montalbano B, Boue SM, Bhatnagar D (2005) An aflatoxin biosynthesis cluster gene encodes a novel oxidase required for conversion of versicolorin A to sterigmatocystin. *Applied and Environmental Microbiology* **71**, 8963-8965.

EPA (2003) *Aspergillus flavus* AF36. July 3, 2003 Biopesticide Registration Action Document. http://www.epa.gov/pesticides/biopesticides/ingredients/tech_docs/brad_006456.pdf.

EPA (2004) *Aspergillus flavus* NRRL 21882. March 24, 2004. Biopesticide Registration Action Document. http://www.epa.gov/oppbppd1/biopesticides/ingredients/tech_docs/brad_006500.pdf.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.

Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics* **133**, 693-709.

Geiser DM, Dorner JW, Horn BW, Taylor JW (2000) The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*. *Fungal Genetics and Biology* **31**, 169-179.

Geiser DM, Pitt JI, Taylor JW (1998) Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. *Proceedings of the National Academy of Sciences, USA* **95**, 388-393.

Griffiths RC, Tavaré S (1994) Ancestral inference in population genetics. *Statistical Science* **9**, 307-319.

Hey J, Wakeley J (1997) A coalescent estimator of the population recombination rate. *Genetics* **145**, 833-846.

Hodges RL, Kelkar HS, Xuei X*, et al.* (2000) Characterization of an echinocandin B-producing strain blocked for sterigmatocystin biosynthesis reveals a translocation in the *stcW* gene of the aflatoxin biosynthetic pathway. *Journal of Industrial Microbiology & Biotechnology* **25**, 333-341.

Horn BW (2005) Ecology and population biology of aflatoxigenic fungi in soil. In: *Aflatoxin and Food Safety* (ed. Abbas HK), pp. 95-116. CRC Press, Boca Raton.

Horn BW (2007) Biodiversity of *Aspergillus* section *Flavi* in the United States: a review. *Food Additives and Contaminants* **24**, 1088–1101.

Horn BW, Dorner JW (1999) Regional differences in production of aflatoxin B_1 and cyclopiazonic acid by soil isolates of *Aspergillus flavus* along a transect within the United States. *Applied and Environmental Microbiology* **65**, 1444-1449.

Horn BW, Dorner JW, Greene RL, Blankenship PD, Cole RJ (1994) Effect of *Aspergillus parasiticus* soil inoculum on invasion of peanut seeds. *Mycopathologia* **125**, 179-191.

Horn BW, Greene RL (1995) Vegetative compatibility within populations of *Aspergillus flavus, Aspergillus parasiticus,* and *A. tamarii* from a peanut field. *Mycologia* **87**, 324-332.

Horn BW, Greene RL, Sobolev VS*, et al.* (1996) Association of morphology and mycotoxin production with vegetative compatibility groups in *Aspergillus flavus, A. parasiticus,* and *A. tamarii*. *Mycologia* **88**, 574–587.

Horn BW, Moore GG, Carbone I (2009a) Sexual reproduction in *Aspergillus flavus*. *Mycologia* **101**, 423-429.

Horn BW, Ramirez-Prado JH, Carbone I (2009b) Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus parasiticus*. *Fungal Genetics and Biology* **46**, 169-175.

Horn BW, Ramirez-Prado JH, Carbone I (2009c) The sexual state of *Aspergillus parasiticus*. *Mycologia* **101**, 275-280.

Howlett BJ, Idnurm A, Heitman J (2007) Fungal pathogenesis: gene clusters unveiled as secrets within the *Ustilago maydis* code. *Current Biology* **17**, R87-90.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**, 147-164.

Kusumoto K, Nogata Y, Ohta H (2000) Directed deletions in the aflatoxin biosynthesis gene homolog cluster of *Aspergillus oryzae*. *Current Genetics* **37**, 104-111.

Leslie JF (1993) Fungal vegetative compatibility. *Annual Review of Phytopathology* **31**, 127-150.

Lyngsø RB, Song YS, Hein J (2005) Minimum recombination histories by branch and bound. *Proceedings of the 5th International Workshop on Algorithms in Bioinformatics (Lecture Notes in Bioinformatics 3692)*, 239-250.

Machida M, Asai K, Sano M*, et al.* (2005) Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* **438**, 1157-1161.

Myers SR, Griffiths RC (2003) Bounds on the minimum number of recombination events in a sample history. *Genetics* **163**, 375-394.

Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.

Patron NJ, Waller RF, Cozijnsen AJ*, et al.* (2007) Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. *BMC Evolutionary Biology* **7**, 174.

Peterson SW (2008) Phylogenetic analysis of *Aspergillus* species using DNA sequences from four loci. *Mycologia* **100**, 205–226.

Peterson SW, Ito Y, Horn BW, Goto T (2001) *Aspergillus bombycis,* a new aflatoxigenic species and genetic variation in its sibling species, *A. nomius*. *Mycologia* **93**, 689-703.

Powell AJ, Conant GC, Brown DE, Carbone I, Dean RA (2008) Altered patterns of gene duplication and differential gene gain and loss in fungal pathogens. *BMC Genomics* **9**, 147.

Price EW, Carbone I (2005) SNAP: workbench management tool for evolutionary population genetic analysis. *Bioinformatics* **21**, 402-404.

Ramirez-Prado JH, Moore GG, Horn BW, Carbone I (2008) Characterization and population analysis of the mating-type genes in *Aspergillus flavus* and *Aspergillus parasiticus*. *Fungal Genetics and Biology* **45**, 1292-1299.

Shieh MT, Brown RL, Whitehead MP*, et al.* (1997) Molecular genetic evidence for the involvement of a specific polygalacturonase, P2c, in the invasion and spread of *Aspergillus flavus* in cotton bolls. *Applied and Environmental Microbiology* **63**, 3548-3552.

Smith CA, Woloshuk CP, Robertson D, Payne GA (2007) Silencing of the aflatoxin gene cluster in a diploid strain of *Aspergillus flavus* is suppressed by ectopic *aflR* expression. *Genetics* **176**, 2077-2086.

Swofford DL (1998) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0.* Sinauer Associates, Sunderland, Massachusetts.

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

van Edmond HP, Jonker MA (2005) Worldwide regulations on aflatoxins. In: *Aflatoxin and Food Safety* (ed. Abbas HK), pp. 77–93 Taylor and Francis, Boca Raton, Florida.

Watterson GA (1975) On the number of segregating sites in genetic models without recombination. *Theoretical Population Biology* **7**, 256-276.

Yin Y, Lou T, Yan L, Michailides TJ, Ma Z (2009) Molecular characterization of toxigenic and atoxigenic *Aspergillus flavus* isolates, collected from peanut fields in China. *Journal of Applied Microbiology*, doi: 10.1111/j.1365-2672.2009.04356.x.

Yu J, Chang P-K, Ehrlich KC*, et al.* (2004) Clustered pathway genes in aflatoxin biosynthesis. *Applied and Environmental Microbiology* **70**, 1253-1262.

Zhang R, Zhang CT (2004) A systematic method to identify genomic islands and its applications in analyzing the genomes of *Corynebacterium glutamicum* and *Vibrio vulnificus* CMCP6 chromosome I. *Bioinformatics* **20**, 612-622.

ACKNOWLEDGEMENTS

We thank two anonymous reviewers for their insightful comments and suggestions that helped us improve this manuscript. This work is funded by the North Carolina Cooperative State Research, Education, and Extension Service, grant nos 2006-34500-17032, 2008-34500-19396, and by the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service, grant no. 2005-35319-16126 to I. C.

Figure S1.1 LD plot for 453 polymorphisms in 21 regions of the aflatoxin gene cluster in *A. flavus*. The scale bar at the top shows the physical distance of polymorphisms in units of base pairs. Below the scale bar is a schematic of the aflatoxin cluster where gene modules are color-coded based on correlations among gene duplicates across five *Aspergillus* genomes (Carbone *et al.* 2007b). The upper triangular matrix represents the *P* values calculated using Fisher's Exact test; the lower triangular matrix represents r^2, the coefficient of determination between the allelic states at pairs of sites. Colored shading in the LD plot indicates statistical significance in linkage and strength of associations. The physical location of blocks with respect to genes in the cluster is shown above and below the plots. A mapping of SNPs and indels among the 27 haplotypes is shown below the LD plot. Haplotypes are based on 437 base substitutions and 16 recoded indels; four indels and six SNPs violated infinite sites and were excluded. In the haplotype map a dot indicates a match to the consensus base; haplotype H1 transitions from one common consensus sequence in one block to a different one in the adjacent block. The physical positions of recombination breaks, inferred from the RecMin program, are shown with dashed vertical lines below the matrix. These dashed lines indicate numerous recombination events (recombination hotspot) among haplotypes in the region between *aflP* and *aflW*. There are six distinct blocks based on *P* and r^2; block 2 (shaded) is the only incompatible partition in the cluster. The uninformative (with respect to LD) *aflW/aflX* intergenic region harbors many polymorphisms for a nonaflatoxigenic lineage represented by haplotype H1.

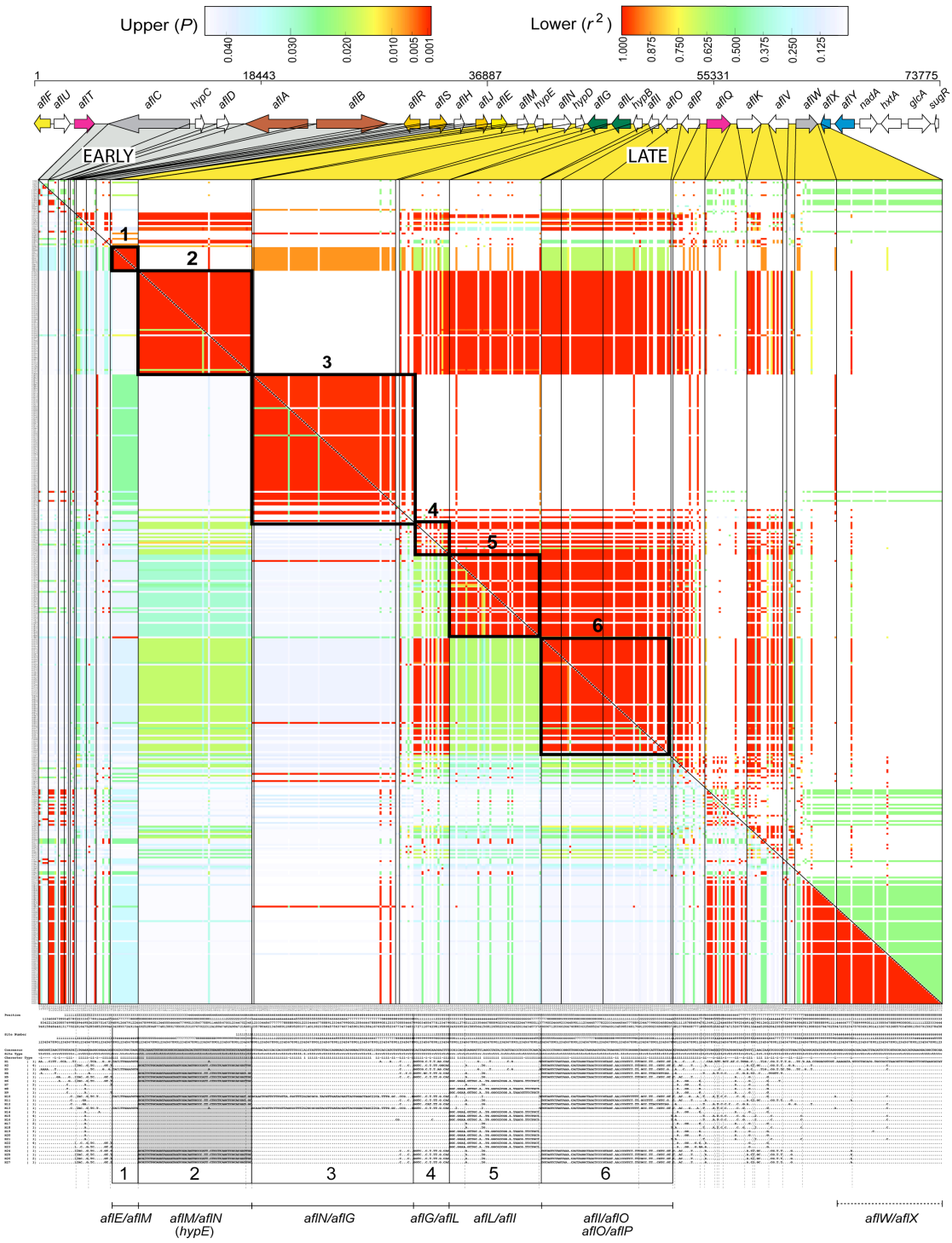

Upper (P)
0.040
0.030
0.020
0.010
0.005
0.001
Lower (r 2)
1.000
0.875
0.750
0.625
0.500
0.375
0.250
0.125
1
18443
36887
55331
73775
aflF
aflU
aflT
aflC
hypC
aflD
aflA
aflB
aflR
aflS
aflH
aflJ
aflE
aflM
hypE
aflN
hypD
aflG
aflL
hypB
aflI
aflO
aflP
aflQ
aflK
aflV
aflW
aflX
aflY
nadA
hxtA
glcA
sugR
EARLY
LATE
1
2
3
4
5
6
aflE/aflM
aflM/aflN
(hypE)
aflN/aflG
aflG/aflL
aflL/aflI
aflI/aflO
aflO/aflP
aflW/aflX

Figure S1.2 LD plot for 2188 polymorphisms across four full aflatoxin cluster sequences: *A. flavus* AF13, AF36, NRRL 3357, and *A. oryzae* NRRL 5590. The upper triangular matrix is a site compatibility matrix in which incompatible sites are indicated with a black square; the lower triangular matrix represents r^2, the coefficient of determination, between the allelic states at pairs of sites. Recombination blocks inferred from the 21 regions in Fig.1 are shaded and each block is extended (outer box) to show the putative real size of the LD block (approximately 35752, 683, 611, 1464, 761 and 3020 bp for blocks 1-6, respectively). This full cluster analysis was important for determining whether block boundaries fall in intergenic regions or exons (labeled in the figure). Extended block 1, which spans the early pathway genes, is compatible with itself but incompatible with the late pathway genes. The two regions showing trans-species polymorphisms are *hypE* in which we have evidence of balancing selection acting on G_1-dominant (high G_1/B_1 ratio) strains in *A. parasiticus* (Carbone *et al.* 2007a) and *aflW/aflX* in which there is evidence of balancing selection acting on a nonaflatoxigenic phenotype in *A. flavus*.

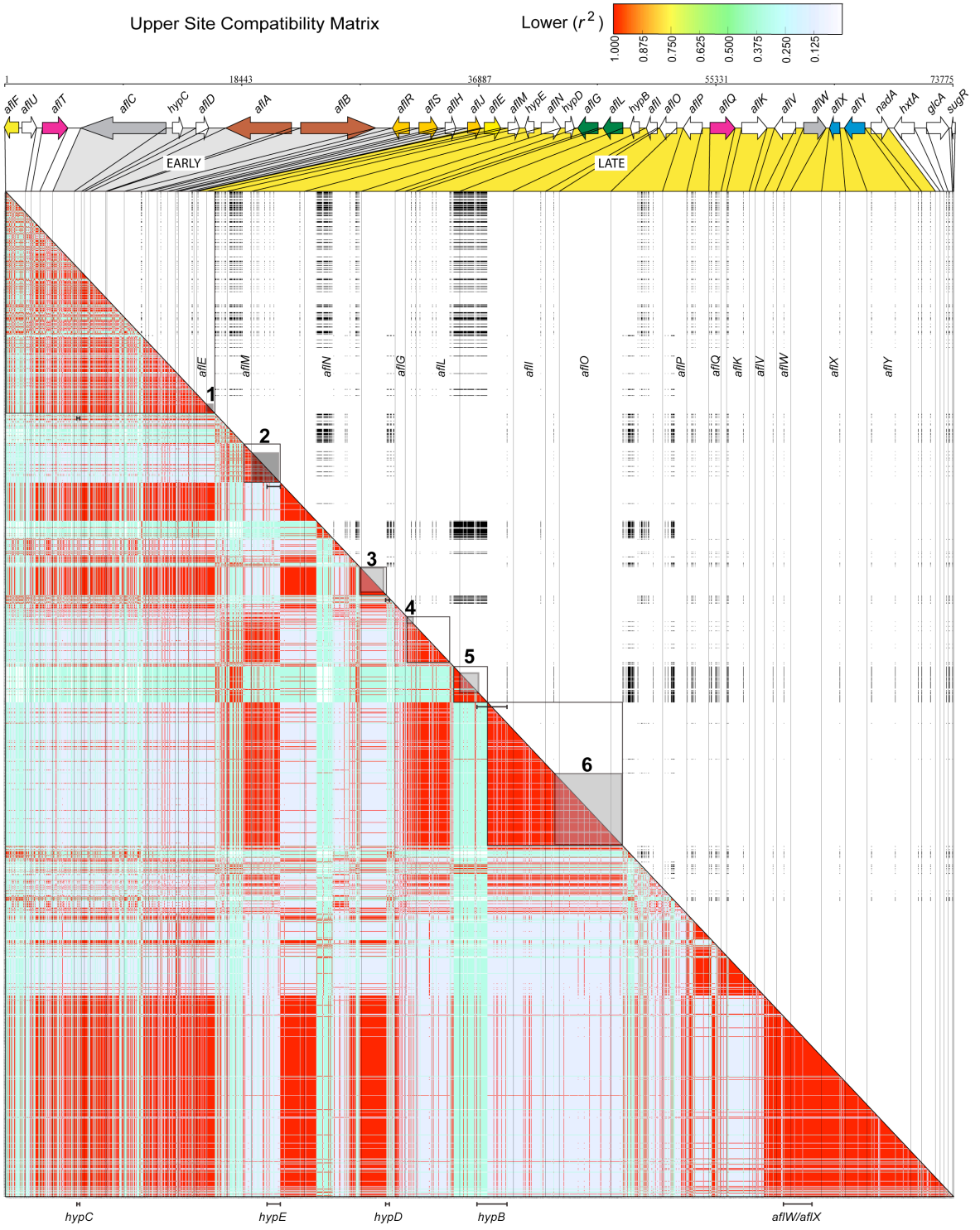

Upper Site Compatibility Matrix
Lower (r^2)
1.000
0.875
0.750
0.625
0.500
0.375
0.250
0.125
1
18443
36887
55331
73775
aflF aflU aflT aflC hypC aflD aflA aflB aflR aflS aflH aflJ aflE aflM hypE aflN hypD aflG aflL hypB aflI aflO aflP aflQ aflK aflV aflW aflX aflY nadA hxtA glcA sugR
EARLY
LATE
aflE aflM aflN aflG aflL aflI aflO aflP aflQ aflK aflV aflW aflX aflY
1
2
3
4
5
6
hypC
hypE
hypD
hypB
aflW/aflX

Figure S1.3 LD plot based on 228 polymorphisms spanning intergenic regions from *aflN* to *aflP* for the 43 Georgia isolates (excluding IC252, which is missing the entire cluster) plus three partial deletion strains (IC311, 313, and 315) and three nonaflatoxigenic full cluster strains (IC309, IC310, and IC312). The partial cluster haplotypes H5 (IC313) and H7 (IC311, IC315) are highlighted in the polymorphism map below the LD plot. The nonaflatoxigenic strain IC277 (H6) shares a recent common ancestor with H7 across blocks 3-6; however, H5 diverges from the consensus haplotype in block 5, giving rise to a partition (shaded) that is incompatible with blocks 3, 4 and 6. The distribution of strains among haplotypes is shown in Table S1.1.


Upper (P)
0.040
0.030
0.020
0.010
0.005
0.001
Lower (r 2)
1.000
0.875
0.750
0.625
0.500
0.375
0.250
0.125
1
18443
36887
55331
73775
aflF
aflU
aflT
aflC
hypC
aflD
aflA
aflB
aflR
aflS
aflH
aflJ
aflE
aflM
hypE
aflN
hypD
aflG
aflL
hypB
aflI
aflO
aflP
aflQ
aflK
aflV
aflW
aflX
aflY
nadA
hxtA
glcA
sugR
3
4
5
6
Position
Site Number
Consensus
Site Type
Character Type
3
4
5
6
aflN/aflG
aflG/aflL
aflL/aflI
aflI/aflO
aflO/aflP

Table S1.1 Haplotypes and strain designations for Figure S1.3

Haplotype	Strain
H1	IC229, IC279, IC286, IC306, IC312
H2	IC245
H3	IC310
H4	IC307
H5	IC313
H6	IC277
H7	IC311, IC315
H8	IC303
H9	IC234, IC244, IC258, IC267, IC270, IC272, IC274, IC278, IC281, IC282, IC284, IC292, IC293, IC294, IC295, IC298, IC300, IC301, IC304
H10	IC280, IC287, IC290, IC299, IC302
H11	IC297
H12	IC289
H13	IC309
H14	IC296
H15	IC308
H16	IC263, IC283, IC285, IC288, IC291
H17	IC305
H18	IC276

Table S1.2 Haplotypes and strain designations for *A. flavus* and *A. parasiticus*[1] phylogenies in Figures 1.4 and S1.4

aflW/aflX

Haplotype	Strain[2]
H1	*A. nomius*(+)
H2	IC2(+), IC11(+), IC15(+), IC25(+), IC38(+), IC56(+), IC67(+), IC69(+), IC75(+), IC77(+), IC78(+), IC98(+), IC115(+), IC123(+)
H3	IC59(+), IC61(+), IC65(OMST), IC70(+), IC71(OMST), IC72(+), IC74(+), IC76(OMST)
H4	IC43(+)
H5	IC73(+)
H6	IC296(+)
H7	**AF36(-)**, IC309(-), IC244(+), IC258(+), IC278(+), IC287(+), IC292(+), IC303(+), IC304(+), IC899(+)
H8	IC277(-), IC311(-), IC315(-)
H9	IC308(+)
H10	**NRRL 21882(-)**, IC313(-)
H11	IC310(-), IC314(-)
H12	NRRL 5590(-)
H13	IC234(+), IC267(+), IC270(+), IC272(+), IC280(+), IC281(+), IC284(+), IC288(+), IC289(+), IC290(+), IC293(+), IC294(+), IC297(+), IC298(+), IC299(+), IC300(+), IC301(+), IC302(+)
H14	IC312(-), IC229(+), IC279(+), IC286(+), IC306(+)
H15	IC245(+), IC307(+)
H16	IC263(+), IC276(+), IC283(+), IC285(+), IC291(+), IC305(+)
H17	IC274(+), IC282(+), IC295(+)

Table S1.2 Continued

amdS

Haplotype	Strain[2]
H1	*A. nomius*(+)
H2	**AF36(-)**, IC309(-), IC312(-), IC267(+), IC270(+), IC278(+), IC281(+), IC295(+), IC298(+), IC303(+), IC305(+), IC899(+)
H3	IC297(+)
H4	IC258(+), IC308(+)
H5	IC313(-), IC310(-)
H6	**NRRL 21882(-)**, IC244(+)
H7	NRRL 5590(-), IC277(-), IC311(-), IC314(-), IC315(-), IC229(+), IC234(+), IC245(+), IC263(+), IC272(+), IC274(+), IC276(+), IC279(+), IC280(+), IC282(+), IC283(+), IC284(+), IC285(+), IC286(+), IC287(+), IC289(+), IC290(+), IC291(+), IC292(+), IC293(+), IC294(+), IC296(+), IC299(+), IC300(+), IC301(+), IC302(+), IC304(+), IC306(+),IC307(+)
H8	IC288(+)
H9	IC59(+), IC65(OMST), IC70(+), IC71(OMST), IC72(+), IC74(+), IC76(OMST)
H10	IC2(+), IC11(+), IC15(+), IC38(+), IC56(+), IC69(+), IC75(+), IC77(+), IC78(+), IC98(+), IC115(+), IC123(+)
H11	IC25(+), IC43(+), IC61(+), IC67(+)
H12	IC73(+)

Table S1.2 Continued

trpC

Haplotype	Strain[2]
H1	*A. nomius*(+)
H2	IC244(+), IC245(+), IC278(+), IC280(+), IC281(+), IC285(+), IC288(+), IC289(+), IC292(+), IC293(+), IC303(+), IC304(+), IC306(+), IC307(+), IC899(+)
H3	NRRL 5590(-), IC270(+), IC302(+)
H4	**AF36(-)**, **NRRL 21882(-)**, IC277(-), IC309(-), IC310(-), IC311(-), IC315(-), IC296(+), IC308(+)
H5	IC234(+), IC258(+), IC267(+), IC282(+), IC290(+), IC294(+), IC297(+), IC300(+), IC301(+)
H6	IC313(-), IC314(-)
H7	IC312(-), IC229(+), IC274(+), IC279(+), IC283(+), IC284(+), IC286(+), IC287(+), IC291(+), IC295(+), IC298(+)
H8	IC263(+), IC272(+), IC276(+), IC299(+), IC305(+)
H9	IC2(+), IC11(+), IC15(+), IC25(+), IC38(+), IC43(+), IC56(+), IC59(+), IC61(+), IC67(+), IC69(+), IC72(+), IC74(+), IC75(+), IC77(+), IC78(+), IC98(+), IC115(+), IC123(+)
H10	IC65(OMST), IC70(+), IC71(OMST), IC76(OMST)
H11	IC73(+)

[1]IC2 - IC123 are *A. parasiticus* strains
[2]Aflatoxin producers (+) or nonproducers (-)

Figure S1.4 The corresponding strict consensus phylogenies (assuming no recombination) and minimal ARGs (assuming recombination) inferred for the phylogenies in Figure 1.4. The ARGs are based on branch and bound (beagle) and heuristic (kwarg) methods and rooted with *A. nomius.* The interpretation of ARGs is described in Figure 1.2. The haplotype color designations are the same as described in Figure 1.4. The recombination nodes (blue ovals) in the ARG are not associated with haplotype color. The ARGs show evidence of recombination between the two distinct *A. flavus* lineages, as well as interspecific recombination between *A. flavus* and *A. parasiticus.* For *aflW/aflX*, branch and bound (beagle) and heuristic (kwarg) methods estimate a minimum of 16 and 20 recombination events, respectively; the ARGs for *amdS* and *trpC* are topologically congruent and show 5 and 4 recombinations, respectively. In the ARGs, the thickened lines show the paths leading to 1) a recombinant H9 between the two distinct *A. flavus* evolutionary lineages in *aflW/aflX* and 2) a putative hybridization between *A. flavus* and *A. parasiticus* giving rise to an *A. parasiticus* haplotype H9 in *amdS* and an *A. flavus* haplotype H5 in *trpC*. The distribution of strains among haplotypes is shown in Table S1.2.

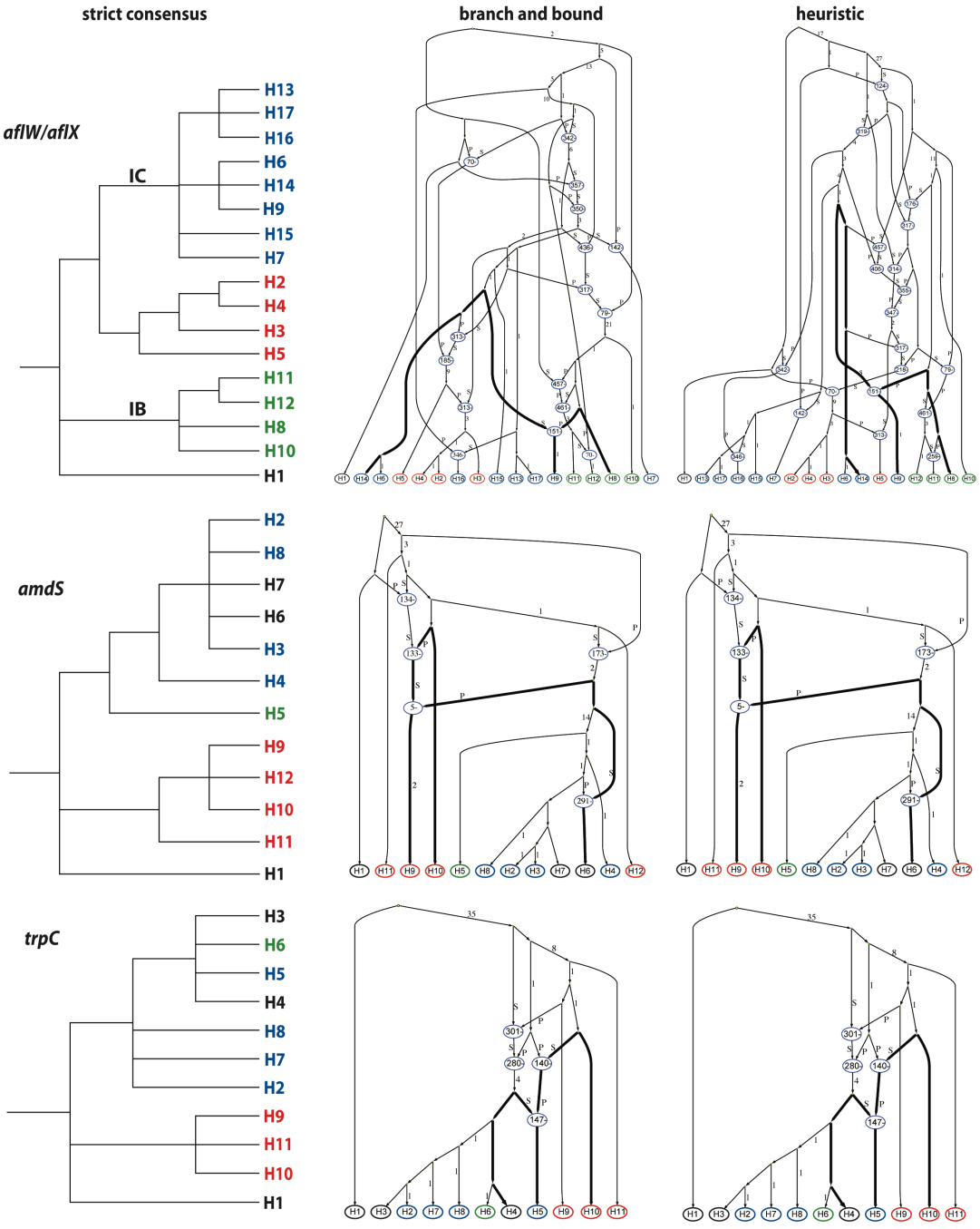

strict consensus
branch and bound
heuristic
aflW/aflX
IC
IB
H13
H17
H16
H6
H14
H9
H15
H7
H2
H4
H3
H5
H11
H12
H8
H10
H1
amdS
H2
H8
H7
H6
H3
H4
H5
H9
H12
H10
H11
H1
trpC
H3
H6
H5
H4
H8
H7
H2
H9
H11
H10
H1

Table S1.3 Haplotypes and strain designations for multilocus gene genealogy in Figure 1.5

Haplotype	Strain
H1	AF36, IC309
H2	IC234, IC270, IC280, IC281, IC290, IC293, IC297, IC299, IC300, IC302
H3	IC244, IC258, IC278, IC287, IC303, IC304, IC292
H4	IC272, IC284, IC298, IC301
H5	IC294
H6	IC267
H7	IC274, IC282, IC295
H8	**IC310**
H9	**IC313**
H10	**IC314, NRRL 5590**
H11	IC305, IC276, IC263, IC283, IC285, IC291
H12	IC289
H13	**IC311**
H14	**IC315**
H15	IC245, IC307
H16	IC229, IC279, IC286, IC306, IC312
H17	**IC277**
H18	IC288
H19	IC308
H20	IC296

Table S1.4 Haplotypes and strain designations for gene genealogies in Figure 1.6

dmaT

Haplotype	Strain
H1	NRRL 5590
H2	AF36, IC229, IC234, IC244, IC245, IC258, IC263, IC267, IC272, IC274, IC276, IC278, IC279, IC280, IC281, IC282, IC283, IC284, IC285, IC286, IC287, IC288, IC289, IC290, IC291, IC292, IC293, IC294, IC295, IC296, IC297, IC298, IC299, IC300, IC301, IC302, IC303, IC304, IC305, IC306, IC307, IC309, IC312, IC314, IC315, NRRL 21882
H3	IC277
H4	IC308, IC311
H5	IC270
H6	IC252
H7	IC310

Table S1.4 Continued

mfs

Haplotype	Strain
H1	NRRL 5590, IC310, IC313, IC314
H2	IC229, IC279, IC286, IC306, IC312
H3	IC245, IC285, IC291, IC305, IC307
H4	IC270, IC302
H5	IC296
H6	IC311
H7	AF36, IC244, IC258, IC278, IC287, IC292, IC303, IC304, IC308, IC309
H8	IC277
H9	IC315
H10	IC204, IC234, IC274, IC276, IC263, IC267, IC272, IC280, IC281, IC282, IC283, IC284, IC288, IC289, IC290, IC293, IC294, IC295, IC297, IC298, IC299, IC300, IC301
H11	IC252, NRRL 21882

Chapter 2

Global Recombination and Aflatoxigenicity in Peanut-Growing Regions

Global Recombination and Aflatoxigenicity in Peanut-Growing Regions

Geromy G. Moore[1], Bruce W. Horn[2], Jacalyn L. Elliott[1], Sofia N. Chulze[3], German Barros[3], Manjunath K. Naik[4], Graeme Wright[5], Kerstin Hell[6], Ignazio Carbone[1]

[1] Department of Plant Pathology, North Carolina State University, Raleigh, NC, 27606
[2] National Peanut Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Dawson, GA 39842
[3] Departamento de Microbiología e Inmunología, Universidad Nacional de Río Cuarto, Córdoba, Argentina
[4] Department of Plant Pathology, College of Agriculture, Karnataka, India
[5] Deaprtment of Primary Industries, Kingaroy, Queensland, Australia
[6] International Institute of Tropical Agriculture, Cotonou, Republic of Benin

---

ABSTRACT

Mycotoxins, and especially the aflatoxins, are a serious problem in agriculture, with aflatoxin B_1 being the most carcinogenic known natural compound. The worldwide costs associated with aflatoxin monitoring and crop losses are in the hundreds of millions of dollars. Aflatoxins also account for considerable health risks, even in countries where food contamination is regulated. *Aspergillus flavus* and *A. parasiticus* are the most common agents of aflatoxin contamination of oil-rich seed and grain crops. Sexual reproduction in these species occurs between individuals belonging to different vegetative compatibility groups (VCGs), and suggests that the vegetative compatibility system is not a barrier to gene flow and that recombination could result in progeny differing in toxicity from the parents. To study this we examined natural genetic variation in *A. flavus* and *A. parasiticus* sampled from single peanut fields in the United States (Georgia), Africa (Benin), Argentina (Córdoba),

Australia (Queensland) and India (Karnataka). Analysis of molecular sequence variation across multiple intergenic regions in the aflatoxin gene clusters of *A. flavus* and *A. parasiticus* revealed significant linkage disequilibrium organized into distinct blocks that are conserved across different localities, which indicates that genetic recombination is a global occurrence. To determine whether sexual reproduction gives rise to recombination blocks, we tested the null hypothesis of an equal number of *MAT1-1* and *MAT1-2* in populations sampled from each locality/species. All samples were clone-corrected using multilocus sequence typing which associates one-to-one with VCG. We found that differences in the proportions of *MAT1-1* and *MAT1-2* were correlated with the amount of asexual (i.e. clonal amplification of VCGs) and sexual reproduction in populations. For both *A. flavus* and *A. parasiticus*, when the numbers of *MAT1-1* and *MAT1-2* were significantly different, there was more extensive LD in the cluster and isolates grouped into specific toxin classes, either the nonaflatoxigenic class in *A. flavus* or the B_1-dominant and G_1-dominant classes in *A. parasiticus*. *Aspergillus flavus* S strains sampled from Córdoba share toxin classes with sympatric *A. parasiticus*, whereas *A. flavus* S strains sampled from Georgia and Queensland do not share the same chemotype classes as their sympatric *A. parasiticus* populations. Our ability to detect and estimate recombination in the aflatoxin gene cluster is dependent on the frequency of distinct toxin classes in population samples. Also, comparing quantile plots for population aflatoxin concentrations with distribution of *MAT1-1* and *MAT1-2* shows a correlation, with a more continuous distribution of toxicity exhibited among populations with a higher component of sexual reproduction. The potential for these fungi to have sex,

coupled with our evidence of recombination within natural populations, should be enough to warrant re-evaluation of biocontrol methods. Current biological control through competitive exclusion by atoxigenic *A. flavus* may eventually cause more harm than good as its efficacy will be determined by a population's recombination potential, which is related to the distribution of mating types in that population. The existence of sympatric species should also be taken into consideration when devising biocontrol strategies. *Aspergillus flavus* S strains appear to have isolates with toxin profiles similar to *A. flavus*, *A. parasiticus* or both. The ability to detect recombination, along with understanding the factors that contribute to recombination in populations may help us better control aflatoxin contamination throughout our world. Our work shows that a combination of ecological factors, asexual/sexual reproduction and balancing selection may influence aflatoxin diversity in these agriculturally important fungi.

INTRODUCTION

Aspergillus flavus Link and *Aspergillus parasiticus* Speare are economically significant fungal species. This is not only because they are prolific colonizers of important food crops, or pathogens of plants and animals, but also because they both are capable of producing mycotoxins such as aflatoxins (Scheidegger and Payne 2003; Klich 2007; Horn and Dorner 2009). In peanut fields, *A. flavus* and *A. parasiticus* are often found together in the soil and they prefer drier, hotter climates since the subsequent stress on their plant hosts enhances their pathogenic success (Blankenship, Cole et al. 1984; Scheidegger and Payne 2003). Several molecular studies provide evidence of recombination in field populations of *A. flavus*

(Geiser, Pitt et al. 1998; Moore, Singh et al. 2009) and *A. parasiticus* (Carbone, Jakobek et al. 2007). Linkage disequilibrium, of DNA sequence variation in the aflatoxin gene cluster, analyses also revealed a history of recombination within both species (Carbone, Jakobek et al. 2007; Moore, Singh et al. 2009). Both species have a heterothallic mating system (Ramirez-Prado, Moore et al. 2008) and the sexual states belong in the genus *Petromyces* (Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009; Horn, Ramirez-Prado et al. 2009). Gene flow is possible through independent assortment and crossovers during sexual reproduction (Milgroom 1996) or via parasexuality in heterokaryons, which are formed by the fusion of vegetatively compatible strains (Papa 1973; Papa 1978; Leslie 1993). Vegetative incompatibility among strains gives rise to vegetative compatibility groups (VCGs), which limit further gene flow and may lead to genetic isolation (Grubisha and Cotty 2009). Sexual reproduction in these species occurs between individuals belonging to different VCGs (Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009), which suggests that the vegetative compatibility system is not a barrier to gene flow and that recombination could result in progeny differing in toxicity from the parents.

A. flavus has two recognized morphotypes that are differentiated based on sclerotial size and mycotoxins produced. L- (large) strain *A. flavus* forms sclerotia greater than 400 μm in diameter and may produce B AFs and CPA, whereas, S- (small) strain *A. flavus* forms high quantities of sclerotia less than 400 μm in diameter that may produce B AFs, G AFs, and CPA (Horn 2003). Some researchers refer to the *A. flavus* S-strain as species *A. parvisclerotigenus* if it produces only B AFs (Frisvad, Skouboe et al. 2005), or *A. minisclerotigenes* if it produces both B and G AFs (Pildain, Frisvad et al. 2008). *A. flavus* L-

strains (and some S-strains) fail to produce G AFs due to defects in, or complete absence of, a required gene known as *cypA* that encodes cytochrome P-450 (Ehrlich, Chang et al. 2004).

In previous work we examined 21 intergenic regions in the aflatoxin gene cluster of *A. flavus* and *A. parasiticus* sampled from a single field population in Georgia (Carbone, Jakobek et al. 2007; Moore, Singh et al. 2009). Genealogical analysis of non-recombining regions in the cluster in *A. parasiticus* (*hypE*) and *A. flavus* (*aflW/aflX*) revealed trans-species polymorphisms and balancing selection acting on specific aflatoxin chemotypes. Our ability to detect and estimate recombination in the aflatoxin cluster depends on the frequency of two or more distinct allelic classes in the population (see Fig. 2.1). In the hypothetical example in Figure 2.1, recombination between blue and green alleles in population A results in the formation of distinct LD blocks, whereas the same recombination events in population B are not detectable. The reduced recombination activity in population B may be the result of a selective sweep for an advantageous allele (in blue) or a population bottleneck that greatly reduces genetic variation. The increased recombination activity as evidenced by the detection of LD blocks and recombination hot spots in the late pathway genes of the aflatoxin cluster in *A. flavus* and *A. parasiticus* has resulted in patterns of recombination similar to population A (Fig. 2.1). In this manuscript we show that recombination is a global occurrence for both *A. flavus* and *A. parasiticus*, and that sex increases the diversity of aflatoxin chemotypes and total aflatoxin load in a population. This knowledge is integral for improving biocontrol strategies worldwide.

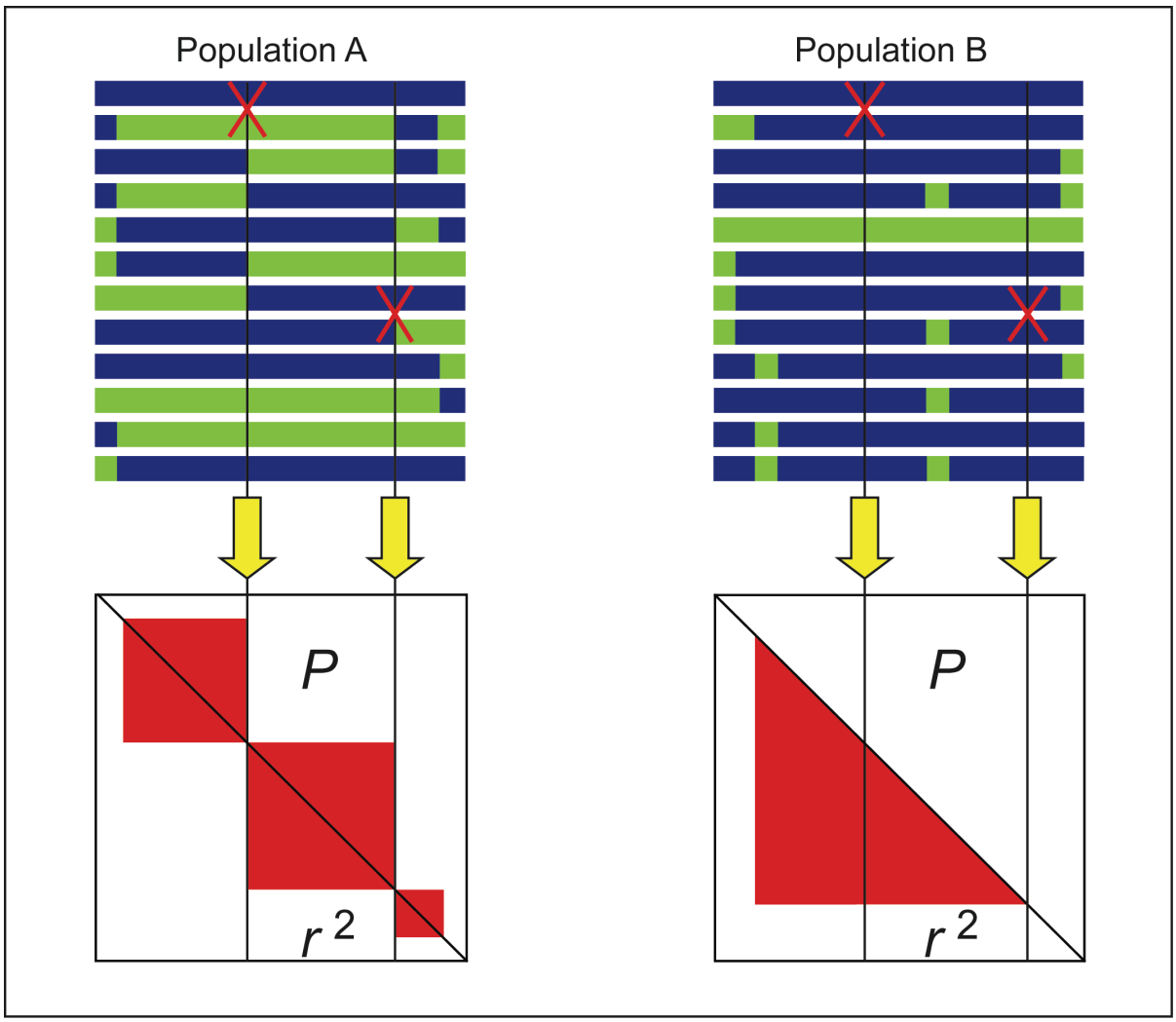


Figure 2.1 Schematic representation for detection of recombination in which two populations are compared. Individuals are represented as rows of blue and green blocks to indicate possible genetic differences. Recombination breakpoints are shown as a red X. Below each population is a linkage disequilibrium (LD) plot in which the upper triangular matrix represents the *P* values calculated using Fisher's Exact test; the lower triangular matrix represents r^2, the coefficient of determination between allelic states at pairs of sites. Colored shading in the LD plot indicates statistical significance in linkage (upper) and strength of associations (lower).

MATERIALS AND METHODS

Sampling and aflatoxin analyses

Population samples for the different section *Flavi* species were from soils collected in different geographic regions representing five continents: Georgia (N. America), Argentina (South America), Queensland (Australia), Karnataka (India), and Benin (Africa). Aspergilli isolated from the soils were identified, and single-spore cultures were grown, at the National Research Peanut Laboratory (NPRL) by Dr. Bruce Horn. The sample size for each population is shown in Table 2.1. Concentrations of B and G aflatoxins for each isolate were determined at the NPRL facility using high performance liquid and thin-layer chromatography, as described previously (Horn, Greene et al. 1996; Horn and Dorner 1999).

Table 2.1 Population sample sizes for each geographic region

Location	*Aspergillus flavus* L	*Aspergillus parasiticus*	*Aspergillus flavus* S
Georgia/US	79	76	25
Córdoba/AR	80	80	4
Queensland/AU	80	80	80
Littoral/BE	80	n/a[1]	44
Karnataka/IN	80	n/a[1]	n/a[2]

[1] *A. parasiticus* isolates were not sampled in this region.
[2] *A. flavus* S isolates were not sampled in this region.

Quantile plots were generated for *A. flavus* L ($B_1 + B_2$), *A. parasiticus* (G_1/B_1), and *A. flavus* S (G_1/B_1 or $B_1 + B_2$) for each geographic location. For *A. flavus* L we determined the aflatoxin midpoint concentration from quantile plots and the proportion of high B_1 producing

strains ($B_1 + B_2 > 100$ µg/mL). *Aspergillus parasiticus, A. flavus* L and *A. flavus* S were further grouped into distinct chemotype classes. In *A. parasiticus* the three classes are B_1 dominant ($G_1/B_1 \leq 0.5$), equivocal ($0.5 < G_1/B_1 < 2.0$) and G_1 dominant ($G_1/B_1 \geq 2$). In *A. flavus,* we counted the number of nonaflatoxigenic isolates that fall in IB, the lineage undergoing lineage-specific gene loss (see Chapter 1). In *A. flavus* S, we examined both the *A. parasiticus* and *A. flavus* chemotype classes.

DNA isolation and multilocus sequence typing (MLST)

Fungal isolates were grown on potato dextrose broth for 3-5 days at 30° C. Mycelial pellets for each isolate were harvested and freeze-dried, and DNA was isolated from a single pellet as previously described (Carbone, Jakobek et al. 2007). PCR amplification and DNA sequencing of target loci were performed using oligonucleotide primers and thermal cycling conditions, also described previously (Carbone, Jakobek et al. 2007). The loci of interest included the blocks of recombination identified in the Georgia populations of *A. flavus* and *A. parasiticus* (Carbone, Jakobek et al. 2007; Moore, Singh et al. 2009). For *A. flavus* the cluster regions investigated were: *aflE/aflM, aflM/aflN (hypE), aflN/aflG, aflG/aflL, aflL/aflI, aflI/aflO.* For *A. parasiticus* the cluster regions of interest were: *aflB/aflR, aflS/aflH, aflH/aflJ, aflJ/aflE, aflE/aflM, aflG/aflL, aflK/aflV*. Non-cluster loci for both species included acetamidase (*amdS*), tryptophan synthase (*trpC*), the mating type genes (*MAT1-1* or *MAT1-2*), and a major facilitator superfamily gene (*mfs*). Mating types were determined for all isolates in the sample using multiplex-PCR (Ramirez-Prado, Moore et al. 2008). All isolates were clone-corrected by multilocus sequence typing (MLST)

using DNA sequence variation in *aflM/aflN*, *aflW/aflX*, *amdS*, and *trpC*. This MLST uniquely types approximately 84 and 59 percent of the VCG diversity in *A. flavus* and *A. parasiticus,* respectively; VCGs were determined for only the single peanut field in Georgia sampled in 1992 (Horn and Greene 1995).

Mating type distribution

We tested the null hypothesis of no significant difference in the number of *MAT1-1* and *MAT1-2* isolates for each species and geographic region using a two-tail binomial test. The test was performed on two genetic scales: the uncorrected sample and the clone-corrected distribution, as determined by MLST. Differences in the proportions of *MAT1-1* and *MAT1-2* were correlated with the amount of asexual (i.e. clonal amplification of VCGs) and sexual reproduction in populations. The influence of asexual and sexual reproduction on recombination in the aflatoxin cluster and overall toxin diversity was examined by reconstructing patterns of LD in the aflatoxin cluster.

Linkage disequilibrium and genealogical inference

Our previous work showed that recombination hotspots and coldspots in the aflatoxin cluster (see LD analysis for full clusters in Chapter 1) are conserved among geographically isolated strains, which suggests that recombination is highly localized (Muers 2010). We therefore determined the extent of recombination in the aflatoxin cluster by focusing on the regions separating six LD blocks (*aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*, *aflI/aflO*) in *A. flavus* L populations that were originally identified in the Georgia sample (Chapter 1).

Similarly we examined regions that separated five LD blocks (*aflB/aflR*, *aflS/aflH*, *aflH/aflJ*, *aflJ/aflE*, *aflE/aflM*, *aflG/aglL*, *aflK/aflV*) in the *A. parasiticus* Georgia sample for detecting recombination in *A. parasiticus* isolates sampled from other geographic regions. For *A. flavus* S populations we examined either the six *A. flavus* blocks, if the strains produced only B aflatoxins (Georgia and Australia), or the five *A. parasiticus* blocks if the strains produced Bs and Gs (Australia, Argentina and Benin). To quantify the amount of recombination we counted the number of phylogenetically informative sites, deduced the minimum number of recombination events (R_h) calculated using the program RecMin (Myers and Griffiths 2003), and estimated the population recombination rate using Hey and Wakeley's γ estimator (Hey and Wakeley 1997) as implemented in the SITES program.

To determine if there is a signature of balancing selection in the aflatoxin cluster of *A. flavus* S we inferred a gene genealogy for each geographic region using GENETREE version 9.0 (http://www.stats.ox.ac.uk/~griff/software.html) assuming panmixis and constant population size. This involved generating all possible rooted gene genealogies for the largest non-recombining partition in the combined blocks and then calculating the relative probabilities of all rooted genealogies by performing 1–10 million simulations of the coalescent, assuming Watterson's (1975) estimate of θ. The different chemotype classes were superimposed on the genealogies to determine if clades separated by many fixed polymorphisms delimit distinct aflatoxin chemotypes classes: G_1 or B_1 dominant in *A. parasiticus* (Carbone, Jakobek et al. 2007) or nonaflatoxigenic belonging to lineage IB in *A. flavus* (Moore, Singh et al. 2009).

Ecological Effect on Aflatoxigenicity

We determined if mating type and aflatoxin distributions are associated with mean annual temperature and precipitation across the different geographical regions using the climate data compiled by WorldClim (http://www.worldclim.org/). The climate data are based on compilations of monthly measurements taken at weather stations all over the world, mostly between the years of 1950-2000. The program uses the latitude/longitude geographical coordinate system and includes all major continents except Antarctica. From this, global climate surfaces were created for monthly precipitation as well as minimum, mean, and maximum temperature offering a high '1 km' resolution for general use (Hijmans, Cameron et al. 2005).

RESULTS

Sampling and aflatoxin analyses

We examined 788 isolates, including *A. flavus* L-strain, *A. parasiticus* and *A. flavus* S sampled from five continents (Table 2.1). *Aspergillus flavus* L was sampled in all regions; *A. parasiticus* was not sampled in Benin and India and *A. flavus* S was not sampled in India. Across all *A. flavus* samples the total concentrations of B aflatoxins ranged from zero to approximately 200 μg/mL, with only a few outliers in Georgia and Australia having concentrations greater than 200 μg/mL (Fig. S2.1, Supporting Information). The number of isolates having a high concentration of B aflatoxins (>100 μg/mL) was highly skewed among localities with Australia harboring the most toxigenic (36/80) followed by Georgia (35), Benin (21), India (9), and Argentina (6) (Table 2.2; Tables S2.1-S2.5, Supporting

Information). Consistent with this trend, Australia also has the highest midpoint aflatoxin concentration of approximately 80 μg/mL, and Argentina has the lowest (0 μg/mL). Nearly half of the Argentina population was nonaflatoxigenic (Fig. S2.1; Table S2.1, Supporting Information); most of these (35/48) are from a lineage IB (Table 2.2) that is nonaflatoxigenic, and undergoing lineage-specific gene loss (Moore, Singh et al. 2009). The other four populations have smaller numbers of nonaflatoxigenic isolates belonging to lineage IB.

In *A. parasiticus*, the G_1/B_1 ratio for most isolates was between 0.5 μg/mL and 2 μg/mL (Fig. S2.2, Supporting Information), indicating equivocal concentrations of G_1 and B_1. The Argentina sample has more G_1- and B_1-dominant isolates (41 and 27, respectively) than G_1/B_1 equivalent (12) isolates (Table 2.3). This trend is reversed in Georgia and Australia resulting in a more continuous distribution of these three chemotype classes. In Figure S2.2 the distribution of G_1/B_1 for the Argentina population is discontinuous and clearly partitioned into three chemotypes classes: B_1 dominant, $G_1 = B_1$ and G_1 dominant. Only *A. parasiticus* sampled in Georgia and Australia have isolates that accumulate *O*-methylsterigmatocystin (OMST) (Supplemental Tables S2.6-S2.8).

Isolates of *A. flavus* S sampled from four localities can be partitioned into chemotype classes that are consistent with either *A. flavus* or *A. parasiticus* (Table 2.4). Figure S2.3 (Supporting Information) illustrates the population aflatoxin concentrations as well as LD if recombination was detected. The toxin profiles for all *A. flavus* S isolates are found in Tables S2.9-S2.12 (Supporting Information).

Table 2.2 Aflatoxin, recombination and mating type distribution in populations of *A. flavus* with different proportions of asexual and sexual reproduction

Region	AF midpoint[a] (μg/mL)	High B AF[b]	Lineage IB[c]	IF Sites[d]	Gamma[e] (γ/bp)	Rh[f]	Genetic Scale[g]	Mating-type Frequency[h] *MAT1-1*	Mating-type Frequency[h] *MAT1-2*	*P*-value[i]
Asexual > Sexual										
Córdoba/AR	0	5/80	35/80	0.0913	0.002429	7	Uncorrected	85 (67)	15 (12)	2.049E-10
							Haplotype corrected	76 (19)	24 (6)	0.0146
Karnataka/IN	30	7/80	4/80	0.0909	0.001608	5	Uncorrected	63 (50)	37 (30)	0.0330
							Haplotype corrected	69 (25)	31 (11)	0.0288
Littoral/BE	40	17/80	2/80	0.0856	0.000238	2	Uncorrected	61 (49)	39 (31)	0.0567
							Haplotype corrected	64 (42)	36 (23)	0.0248
Sexual > Asexual										
Georgia/US	60	27/79	1/79	0.0896	0.001119	6	Uncorrected	19 (15)	81 (63)	3.749E-08
							Haplotype corrected	38 (15)	62 (24)	0.1996
Queensland/AU	80	29/80	1/80	0.0848	0.001029	5	Uncorrected	37 (29)	63 (50)	0.0238
							Haplotype corrected	37 (13)	63 (22)	0.1755

[a] AF concentration ($B_1 + B_2$) midpoint, in μg/mL, for quantile plots in Figure S2.1.
[b] Number of isolates having AF (B_1+B_2) > 100 μg/mL out of the total isolates in sample.
[c] Number of AF- isolates that group with Geiser's group IB, out of the total number of full-cluster isolates, based on phylogenetic inference for *aflW/aflX* region.
[d] Number of informative (IF) variable sites out of the total number of variable sites.
[e] Gamma is rate of recombination for six genomic loci used in LD analysis: *aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*, *aflI/aflO*.
[f] Minimum number of recombination events predicted by RecMin.
[g] Samples either uncorrected, or haplotype corrected based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*.
[h] Numbers in parentheses refer to number of isolates (uncorrected), or haplotypes (corrected), examined for each genetic scale.
[i] Probability of a binomial test (two-tailed). Test was performed on the raw data.

Table 2.3 Aflatoxin, recombination and mating type distribution of *A. parasiticus* with different proportions of asexual and sexual reproduction

Region	$B_1 >> G_1$[a]	$G_1 = B_1$[b]	$G_1 >> B_1$[c]	Gamma[d] (γ/bp)	IN Sites[e]	Rh[f]	Genetic Scale[g]	Mating-type Frequency[h] *MAT1-1*	Mating-type Frequency[h] *MAT1-2*	*P*-value[i]
Asexual > Sexual										
Córdoba/AR	27	12	41	n/e[3]	0.00016	0	Uncorrected	98 (78)	2 (2)	-1.998E-15
							Haplotype corrected	88 (15)	12 (2)	0.0023
Sexual > Asexual										
Georgia/US[1]	9	59	4	0.000706	0.020905	4	Uncorrected	81 (61)	19 (14)	3.808E-08
							Haplotype corrected	72 (13)	28 (5)	0.0963
Queensland/AU[2]	4	32	43	0.009947	0.018466	12	Uncorrected	38 (29)	62 (48)	0.0395
							Haplotype corrected	41 (12)	59 (17)	0.4582

[a] Number of isolates that are B_1 dominant ($G_1/B_1 \leq 0.5$).
[b] Number of isolates that have equivalent amounts of G_1 and B_1 ($0.5 < G_1/B_1 < 2.0$).
[c] Number of isolates that are G_1 dominant ($G_1/B_1 \geq 2.0$).
[d] Gamma is rate of recombination for five genomic loci used in LD analysis: *aflB/af/R*, *aflS/aflH*, *aflH/aflJ*, *aflE/aflM*, *aflG/aflL*, *aflK/aflV*.
[e] Number of informative (IN) sites out of the total number of variable sites.
[f] Minimum number of recombination events predicted by RecMin.
[g] Samples either uncorrected, or haplotype corrected based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*.
[h] Numbers in parentheses refer to number of isolates (uncorrected), or haplotypes (corrected), examined for each genetic scale.
[i] Probability of a binomial test (two-tailed). Test was performed on the raw data.
[1] Four isolates produce only *O*-methyl-sterigmatocystin (OMST) at $80 < \text{OMST} < 250$ μg/mL.
[2] One isolate produces only *O*-methyl-sterigmatocystin (OMST) at 119 μg/mL.
[3] No estimate determinable from gamma analysis.

Table 2.4 Aflatoxin, recombination and mating type distribution *A. flavus* S

Region	Aflatoxin Profile				IN Sites[e]	Gamma[f]	Rh[g]	Genetic Scale[h]	Mating-type Frequency[i]		*P*-value[j]
	B only[a]	$B_1 >> G_1$[b]	$G_1 = B_1$[c]	$G_1 >> B_1$[d]		(γ/bp)			*MAT1-1*	*MAT1-2*	
Georgia/US	25	0	0	0	0.003761	n/e[2]	0	Uncorrected	62 (13)	38 (8)	0.3833
								Haplotype corrected	60 (6)	40 (4)	0.7539
Queensland/AU[1]	50	0	19	8	0.042038	0.002895	5	Uncorrected	29 (15)	71 (37)	0.0032
								Haplotype corrected	38 (10)	62 (16)	1
Córdoba/AR	2	0	0	2	0.094964	n/e[2]	0	Uncorrected	50 (2)	50 (2)	1.375
								Haplotype corrected	67 (2)	33 (1)	1
Littoral/BE	2	0	24	18	0.127387	0.019272	26	Uncorrected	41 (14)	59 (20)	0.3915
								Haplotype corrected	41 (7)	59 (10)	0.629

[a] Number of isolates that produce B (B_1+B_2) AFs only.
[b] Number of isolates that are B_1 dominant ($G_1/B_1 \leq 0.5$).
[c] Number of isolates that have equivalent amounts of G_1 and B_1 ($0.5 < G_1/B_1 < 2.0$).
[d] Number of isolates that are G_1 dominant ($G_1/B_1 \geq 2.0$).
[e] Number of informative (IN) sites out of the total number of variable sites.
[f] Gamma is rate of recombination for six genomic loci used in LD analysis: *aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*.
[g] Minimum number of recombination events predicted by RecMin.
[h] Samples either uncorrected, or haplotype corrected based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*.
[i] Numbers in parentheses refer to number of isolates (uncorrected), or haplotypes (corrected), examined for each genetic scale.
[j] Probability of a binomial test (two-tailed). Test was performed on the raw data.
[1] Three isolates are nonaflatoxigenic making neither B nor G aflatoxins.
[2] No estimate determinable from gamma analysis.

Mating type distribution

There was a significant disparity in the number of *MAT1-1* and *MAT1-2* isolated in the Argentina, India and Benin *A. flavus* L populations, with *MAT1-1* being the dominant idiomorph, for both the uncorrected and MLST corrected samples (Table 2.2). Approximately 60% (48/80) of the *A. flavus* isolates sampled in Argentina are nonaflatoxigenic (Fig. S2.1, Table S2.1), and 35 out of the 48 nonaflatoxigenic isolates (73%) are from lineage IB (Moore, Singh et al. 2009). By comparison the India and Benin population samples contained only four and two isolates belonging to lineage IB, respectively. Even after clone-correction the number of *MAT1-1* individuals is significantly higher than *MAT1-2* indicating a predominantly clonal population structure in Argentina, India and Benin. This trend is significantly reversed in *A. flavus* L population samples from Georgia and Australia. In these populations *MAT1-2* is more abundant than *MAT1-1* in the uncorrected samples, but this difference is not significant after clone-correction (Table 2.2), indicating a predominantly recombining population structure. Because these fungi are heterothallic there is independent segregation of mating type genes during meiosis in the transient diploid formed during sexual reproduction. This would result in equal proportions of *MAT1-1* and *MAT1-2* provided there is no post zygotic mechanism affecting the fitness of a particular mating type.

Mating type ratios are also skewed in favor of *MAT1-1* in *A. parasiticus* populations sampled from Argentina and Georgia, whereas *MAT1-2* predominates in Australia (Table 2.3). In Argentina, 98% (78/80) of the isolates were *MAT1-1* indicating strong clonality. By contrast clone-correction of the Georgia and Australia samples showed that differences in

MAT1-1 and *MAT1-2* were not significant (P = 0.0963), indicating a predominantly sexual population structure. Similarly, mating type ratios showed no significant deviation from one in *A. flavus* S population samples (Table 2.4).

Linkage disequilibrium and genealogical inference

Patterns of LD in the aflatoxin cluster were conserved across all populations but there were differences in the size of blocks and recombination rates. The most significant linkage was observed for the Argentina population, which also had the highest recombination rate (γ = 0.002429) and minimum number of inferred recombination events (R_h = 7). Ranking the other clonal populations by decreasing linkage and gamma/R_h was as follows: India (γ = 0.001608/R_h = 5) and Benin (γ = 0.000238/R_h = 2). The predominantly sexual populations in Georgia and Queensland revealed almost identical LD plots and were matched in their toxin profiles and concentrations, recombination rate estimates (γ = 0.001119 and 0.001029, respectively) and minimum number of recombination events (R_h = 6 and 5, respectively).

Sympatric populations of *A. parasiticus*, *A. flavus* L and S were sampled only from Argentina, Georgia and Australia. No recombination was detected in the Argentina *A. parasiticus* population (Fig. S2.2, Table 2.3). In the predominantly sexual *A. flavus* L populations, recombination rates in the aflatoxin cluster were ten-fold lower in Georgia (γ = 0.000706; R_h = 4) than in Argentina (γ = 0.009947; R_h = 12). In all cases, populations with higher recombination rates have elevated numbers of informative sites. The *MAT1-1/MAT1-2* of 1 was not significantly different from *A. flavus* S sampled from Georgia, Australia, Argentina and Benin; however, patterns of LD in the aflatoxin cluster indicate reduced

recombination activity (Fig. S2.3). Multilocus gene genealogies reveal a signature of balancing selection acting on G_1 dominant (G_1/B_1) and B aflatoxins (Fig. S2.4) in the Australia, Argentina, and Benin population samples. When isolates produced only B aflatoxins, as in the Georgia population, mutations rates were consistent with neutrality.

Ecological Effect on Aflatoxigenicity

All five sampling areas had a similar mean annual temperature but could be split into wet (Benin and India) and dry (Georgia, Argentina and Australia) areas according to annual precipitation data (Fig. 2.2).

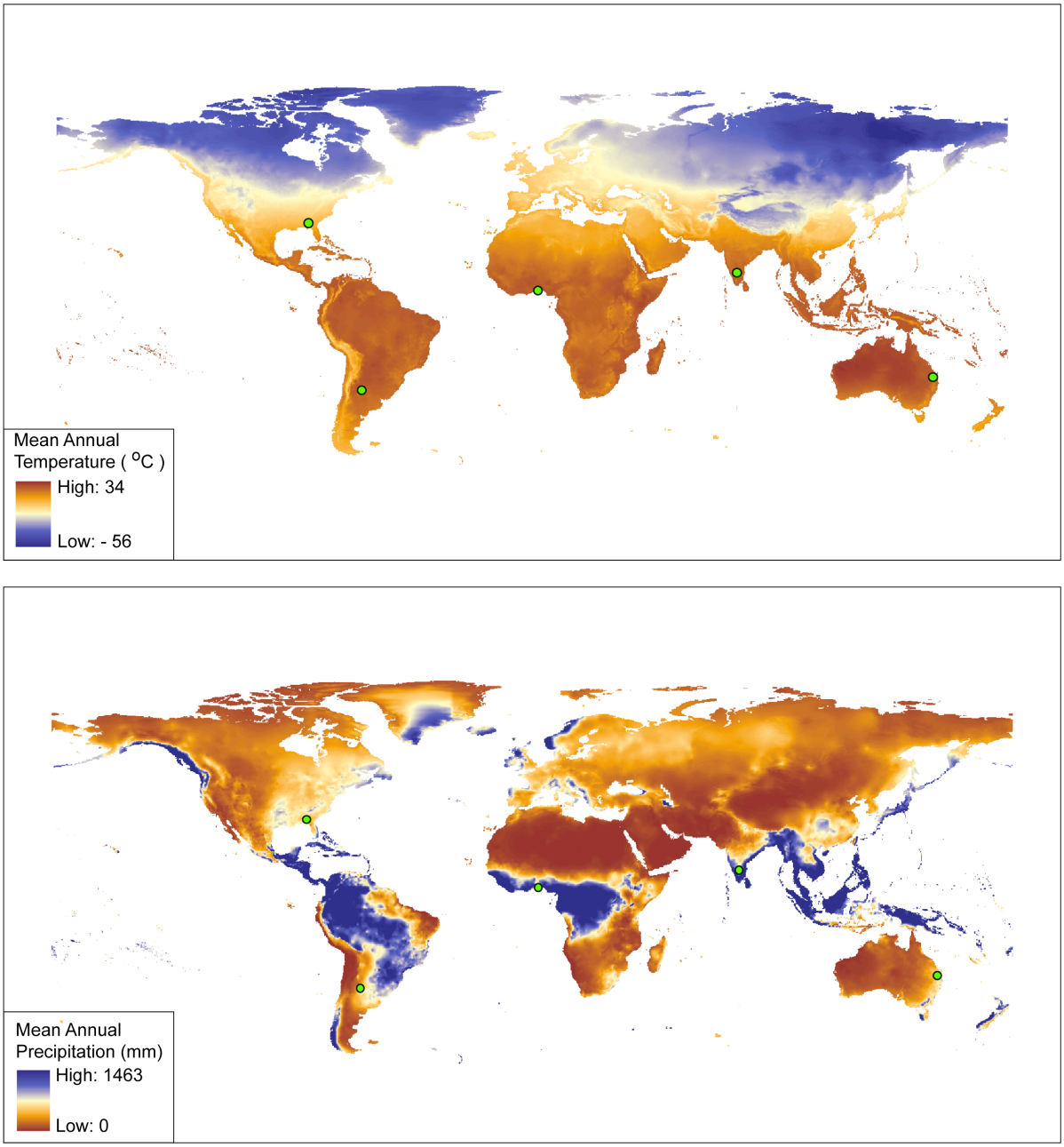


Figure 2.2 Climate maps illustrating mean annual temperature in °C (upper) and precipitation in mm (lower) on a global scale. The green dots mark the locations of fields sampled for this study. For the temperature gradient, blue is cooler and red is warmer. For precipitation, blue is higher quantity and red is lower.

DISCUSSION

There are studies that report how climate directly affects fungal community structure (Cotty and Jaime-Garcia 2007). *Aspergillus flavus* L was sampled in all regions, but *A. parasiticus* and *A. flavus* S were absent or uncommon in certain regions (Schroeder and Boller 1973; Cotty and P 1997; Cotty and Cardwell 1999). Worldclim data shows that Georgia, Australia and Argentina have significantly less precipitation than Benin and India. This difference correlates with species richness in these areas; soil-inhabiting *A. parasiticus* was not sampled in the wetter regions. Both species prefer dry, hot weather conditions, but perhaps *A. flavus* has a better adaption to wetter soils. Another possibility is that *A. parasiticus*, being mostly limited to soil environments, is hybridizing with *A. flavus* in an attempt to make a niche change. This might explain why there are many *A. flavus* S strains that produce B and G AFs in the absence of any *A. parasiticus* isolates, but where or when this out-crossing between *A. flavus* and *A. parasiticus* occurred in the Benin field is unknown. Yet another possible reason for absence of *A. parasiticus* in certain fields is the soil composition. It has been reported that *A. flavus* population density is higher in clay-rich soils (Joffe and Lisker 1969; Griffin and Garren 1974), and both Benin and India have clay-rich soils (Table S2.13). In heavy (clay) soils that may have higher quantity of organic matter (Griffin and Garren 1974), if *A. flavus* has an advantage in obtaining nutrients over *A. parasiticus* then it might competitively exclude *A. parasiticus* in a field.

Skewed mating type distributions in different geographic regions could depend on the ratio between sexual and asexual reproduction (Leslie and Klein 1996) or ecological differences. The *MAT* genes, *MAT1-1* and *MAT1-2* encode putative transcription factors

regulating pheromone and pheromone receptor genes as well as other genes not involved directly in the mating process (Hornok, Waalwijk et al. 2007). A higher frequency of one mating type may be the result of increased fitness on a function other than mating. In some heterothallic organisms, such as in the oomycete *Phytophthora cinnamomi* a single mating type (A2) is sampled at a higher frequency worldwide without any host preference while the other mating type (A1) is rare and recovered only from certain hosts (Galindo and Zentmyer 1964; Duan, Riley et al. 2008).

Linkage disequilibrium analyses revealed that recombination blocks are conserved among populations. The higher recombination activity in *A. flavus* L populations with a greater proportion of asexual reproduction (Argentina and India) may be the result of balancing selection acting on the cluster in lineage IB isolates which may favor higher rates of allelic shuffling. The three LD blocks detected in Benin are consistent with reduced recombination activity in this population ($\gamma = 0.0002$; $R_h = 2$). The lower rate may not necessarily be the result of lower recombination rates *per se* but a lack of individuals with divergent DNA sequences to allow us to track recombination events when they occur (Figure 2.1). In *A. flavus* L divergent sequences arise from balancing selection acting on a specific nonaflatoxigenic phenotype belonging to lineage IB (Moore, Singh et al. 2009). As seen in Table 2.2, when the number of isolates in lineage IB increases from two (Benin) to 35 (Argentina), there is a ten-fold increase in the rate of recombination and a three times increase in the minimum number of recombination events. Similar population dynamics are observed in *A. parasiticus* populations.

The lack of recombination in *A. parasiticus* from Argentina suggests a recent selective sweep of the *MAT1-1* mating type acting on B_1 and G_1 dominant chemotypes, whereas the higher recombination activity in Georgia and Queensland may be the result of balancing selection acting on B_1 and G_1 dominant strains. For the sympatric *A. parasiticus* and *A. flavus* populations in Georgia, it appears that mating type distributions are skewed in opposite directions such that *A. parasiticus* has a higher frequency of *MAT1-1* and *A. flavus* has a higher frequency of *MAT1-2.* This may arise from interspecific balancing selection acting on *MAT1-1* and *MAT1-2* to facilitate inter-specific matings. Alternatively, there may be a higher frequency of female sterility in *MAT1-2* isolates in *A. parasiticus* and *MAT1-1* sterility in *A. flavus,* but this cannot be known for certain without further mating studies. In both species, fertile crosses are only possible between parents of different VCGs suggesting that sexual reproduction serves as a mechanism to purge strains of deleterious mutations arising from DNA methylation.

In predominantly clonal *A. flavus* L populations, recombination rates are higher in the cluster when balancing selection is present. For example, in Argentina recombination and selection act to maintain a high proportion of nonaflatoxigenic IB isolates. In contrast, recombination is relaxed in Benin, which has only two nonaflatoxigenic IB isolates. In sexual *A. flavus* L populations toxin concentrations are higher and very few isolates are nonaflatoxigenic. For example, only a single nonaflatoxigenic isolate belonging to lineage IB was sampled in Georgia and Australia. In these populations the target of balancing selection is more difficult to discern because recombination shuffles distinct alleles among isolates. Sexual reproduction also appears to have a diversifying effect on aflatoxigenicity

such that the more actively recombining populations exhibit a continuous (diverse) gradient between the lowest and highest aflatoxin concentrations, as well as giving rise to some of the highest aflatoxin producers. A similar phenomenon was observed in the *A. parasiticus* populations. For example, in Argentina there was no evidence of recombination in the cluster and isolates could be clearly partitioned into the G_1 and B_1 dominant chemotypes. Although these two chemotype classes were also found in Georgia and Australia, a greater proportion of the total population were G_1 and B_1 equivocal. This indicates that balancing selection is still acting in these gene clusters to maintain G_1 and B_1 dominant chemotypes in the population, which is also supported by the high heritability of G_1/B_1 observed in experimental mating populations of *A. parasiticus* (I Carbone, unpublished data).

The high proportion of chemotype-specific isolates of *A. flavus* L and *A. parasiticus* in Argentina is notable. In *A. flavus* a large number of nonaflatoxigenic isolates in lineage IB belonging to different haplotypes (Table S2.1) gives rise to significant linkage in the cluster (Fig. S2.1). Increased recombination activity may be the result of more sequence diversity in the population to track recombination when it occurs, or it may be in response to strong selection to maintain the nonaflatoxigenic phenotype. In the sympatric *A. parasiticus* population, recombination in the cluster was not detected and the G_1 and B_1 chemotype classes were distinct. In previous studies in Georgia we predicted that balancing selection was acting on G_1/B_1 in *A. parasiticus* and on the nonaflatoxigenic trait in *A. flavus*. Both of these inferences were based on single isolates in a single field population and we argued that recombination shuffled chemotype allelic classes making it difficult to identify the phenotypic targets of balancing selection. From our examination of the Argentina population

we now have supporting evidence that in the absence of sexual reproduction these specific chemotypes are sampled at high frequency. The strong skew in the mating type distributions, favoring *MAT1-1* in *A. parasiticus* and *A. flavus* L in Argentina, indicates strong clonality or a selective sweep. It's possible that *MAT1-1* isolates may be more fit or there is a high frequency of female sterility among *MAT1-2* isolates (Leslie and Klein 1996; Mekwatanakarn, P et al. 1999). At the other extreme is Australia, which is predominantly sexual and exhibits the least linkage in the cluster and the most continuous distribution of aflatoxin concentrations. This population harbors the greatest number of G_1 and B_1 equivalent isolates, but also maintains a significant number of unique chemotypes (G-dominant, B-dominant, OMST), which suggests that these chemotypes are maintained by a mechanism of negative frequency dependent selection. In Argentina these chemotype classes are found at a higher frequency indicating a greater fitness of the chemotypes or *MAT1-1*.

The results of this work may be useful for designing effective biocontrol in these different localities. A major contribution to this effort is our observation that the amount of recombination in a population is determined by the frequency of *MAT1-1* and *MAT1-2* mating-types genes. Populations that have a 1:1 distribution of *MAT1-1* and *MAT1-2* will exhibit greater rates of recombination as well as greater diversity of aflatoxins because the likelihood of strains of complementary idiomorphs finding one another and mating increases (Zhan, Kema et al. 2002). For example, in the Argentina populations we observe reduced diversity of aflatoxins and little recombination. This is a mostly clonal population for both *A. flavus* and *A. parasiticus* where *MAT1-1* greatly outnumbers *MAT1-2* even after clone correction. A biocontrol isolate that is *MAT1-1* might do well in these fields, as the potential

to recombine with an indigenous toxin producer is low. At the other end of the spectrum is Australia. Even with clone correction the distribution of *MAT1-1* and *MAT1-2* isolates, for both species, is 1:1. Under such circumstances the potential of a biocontrol strain recombining with a toxin producer increases. The implications of introducing a biocontrol strain at high density to the overall population genetic structure of these species is unknown. It is clear that sexual reproduction is a mechanism for transferring genetic material between VCGs and possibly between sympatric species. In the next chapter we will explore these possibilities.

LITERATURE CITED

Blankenship, P. D., R. J. Cole, et al. (1984). "Effect of Geocarposphere Temperature on Pre-Harvest Colonization of Drought-Stressed Peanuts by *Aspergillus-Flavus* and Subsequent Aflatoxin Contamination." Mycopathologia **85**(1-2): 69-74.

Carbone, I., J. L. Jakobek, et al. (2007). "Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus*." Molecular Ecology **16**(20): 4401–4417.

Cotty and J. P (1997). Aflatoxin-producing potential of communities of Aspergillus section Flavi from cotton producing areas in the United States. Oxford, ROYAUME-UNI, Elsevier.

Cotty, P. J. and K. F. Cardwell (1999). "Divergence of West African and North American Communities of Aspergillus Section Flavi." Appl. Environ. Microbiol. **65**(5): 2264-2266.

Cotty, P. J. and R. Jaime-Garcia (2007). "Influences of climate on aflatoxin producing fungi and aflatoxin contamination." International Journal of Food Microbiology **119**(1-2): 109-115.

Duan, C. H., M. B. Riley, et al. (2008). "Characterization of *Phytophthora cinnamomi* populations from ornamental plants in South Carolina, USA." Archives Of Phytopathology And Plant Protection **41**(1): 14 - 30.

Ehrlich, K. C., P. K. Chang, et al. (2004). "Aflatoxin biosynthesis cluster gene cypA is required for G aflatoxin formation." Appl Environ Microbiol **70**(11): 6518-24.

Frisvad, J. C., P. Skouboe, et al. (2005). "Taxonomic comparison of three different groups of aflatoxin producers and a new efficient producer of aflatoxin B-1, sterigmatocystin and 3-O-methylsterigmatocystin, *Aspergillus rambellii* sp nov." Systematic and Applied Microbiology **28**(5): 442-453.

Galindo, J. and G. A. Zentmyer (1964). "Mating Types in *Phytophthora cinnamomi*." Phytopathology **54**: 238.

Geiser, D. M., J. I. Pitt, et al. (1998). "Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*." Proc Natl Acad Sci U S A **95**(1): 388-93.

Griffin, G. J. and K. H. Garren (1974). "Population levels of *Aspergillus flavus* and *Aspergillus niger* group in Virginia peanut field soils." Phytopathology **64**(3): 322-325.

Grubisha, L. C. and P. J. Cotty (2009). "Twenty-four microsatellite markers for the aflatoxin-producing fungus *Aspergillus flavus*." Molecular Ecology Resources **9**(1): 264-267.

Hey, J. and J. Wakeley (1997). "A coalescent estimator of the population recombination rate." Genetics **145**(3): 833-46.

Hijmans, R., S. Cameron, et al. (2005). "Very high resolution interpolated climate surfaces for global land areas." International Journal of Climatology **25**(15): 1965-1978.

Horn, B. W. (2003). "Ecology and population biology of aflatoxigenic fungi in soil." Journal of Toxicology-Toxin Reviews **22**(2-3): 351-379.

Horn, B. W. and J. W. Dorner (1999). "Regional differences in production of aflatoxin B_1 and cyclopiazonic acid by soil isolates of *Aspergillus flavus* along a transect within the United States." Appl Environ Microbiol **65**(4): 1444-9.

Horn, B. W. and J. W. Dorner (2009). "Effect of nontoxigenic *Aspergillus flavus* and *A. parasiticus* on aflatoxin contamination of wounded peanut seeds inoculated with agricultural soil containing natural fungal populations." Biocontrol Science and Technology **19**(3): 249-262.

Horn, B. W. and R. L. Greene (1995). "Vegetative compatibility within populations of *Aspergillus flavus, Aspergillus parasiticus,* and *A. tamarii* from a peanut field." Mycologia **87**(3): 324-332.

Horn, B. W., R. L. Greene, et al. (1996). "Association of morphology and mycotoxin production with vegetative compatibility groups in *Aspergillus flavus, A. parasiticus,* and *A. tamarii*." Mycologia **88**(4): 574–587.

Horn, B. W., G. G. Moore, et al. (2009). "Sexual reproduction in *Aspergillus flavus*." Mycologia **101**(3): 423-429.

Horn, B. W., J. H. Ramirez-Prado, et al. (2009). "Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus parasiticus*." Fungal Genet Biol **46**(2): 169-175.

Horn, B. W., J. H. Ramirez-Prado, et al. (2009). "The sexual state of *Aspergillus parasiticus*." Mycologia **101**(2): 275-280.

Hornok, L., C. Waalwijk, et al. (2007). "Genetic factors affecting sexual reproduction in toxigenic Fusarium species." Int J Food Microbiol **119**(1-2): 54-8.

Joffe, A. Z. and N. Lisker (1969). "Mycoflora of fresh and subsequently stored groundnut kernels on various soil types." Israel Journal of Botany **18**(2): 77-&.

Klich, M. A. (2007). "*Aspergillus flavus*: the major producer of aflatoxin." Molecular Plant Pathology **8**(6): 713-722.

Leslie, J. F. (1993). "Fungal vegetative compatibility." Annual Review of Phytopathology **31**: 127-50.

Leslie, J. F. and K. K. Klein (1996). "Female fertility and mating type effects on effective population size and evolution in filamentous fungi." Genetics **144**(2): 557-67.

Mekwatanakarn, P, et al. (1999). Sexually fertile *Magnaporthe grisea* rice pathogens in Thailand. St. Paul, MN, ETATS-UNIS, American Phytopathological Society.

Milgroom, M. G. (1996). "Recombination and the multilocus structure of fungal populations." Annu Rev Phytopathol **34**: 457-77.

Moore, G. G., R. Singh, et al. (2009). "Recombination and lineage-specific gene loss in the aflatoxin gene cluster of *Aspergillus flavus*." Molecular Ecology **18**: 4870–4887.

Muers, M. (2010). "Recombination: The gene that hits the hot spot." Nat Rev Genet **11**(2): 92-92.

Myers, S. R. and R. C. Griffiths (2003). "Bounds on the minimum number of recombination events in a sample history." Genetics **163**(1): 375-394.

Papa, K. E. (1973). "Parasexual Cycle in *Aspergillus flavus*." Mycologia **65**(5): 1201-1205.

Papa, K. E. (1978). "Parasexual Cycle in *Aspergillus parasiticus*." Mycologia **70**(4): 766-773.

Ramirez-Prado, J. H., G. G. Moore, et al. (2008). "Characterization and population analysis of the mating-type genes in *Aspergillus flavus* and *Aspergillus parasiticus*." Fungal Genet Biol **45**(9): 1292-9.

Scheidegger, K. and G. Payne (2003). "Unlocking the Secrets Behind Secondary Metabolism: A Review of *Aspergillus flavus* from Pathogenicity to Functional Genomics." Journal of Toxicology -- Toxin Reviews **22**(2/3): 423-459.

Schroeder, H. W. and R. A. Boller (1973). "Aflatoxin production of species and strains of the *Aspergillus flavus* group isolated from field crops." Appl Microbiol **25**(6): 885-9.

Watterson, G. A. (1975). "On the number of segregating sites in genetic models without recombination." Theoretical Population Biology **7**: 256-276.

Zhan, J., G. H. Kema, et al. (2002). "Distribution of mating type alleles in the wheat pathogen *Mycosphaerella graminicola* over spatial scales from lesions to continents." Fungal Genet Biol **36**(2): 128-36.

Figure S2.1 Graphical representations of population aflatoxin concentrations and recombination in populations of *A. flavus* L with different proportions of asexual and sexual reproduction. Quantile plots were created which show the number of individuals along the x-axis and the total B ($B_1 + B_2$) aflatoxin (AF) concentrations in μg/mL. To the right of each quantile plot is an LD plot in which the upper triangular matrix represents the *P* values calculated using Fisher's Exact test; the lower triangular matrix represents r^2, the coefficient of determination between allelic states at pairs of sites. Colored shading in the LD plot indicates statistical significance in linkage (upper) and strength of associations (lower). The LDs are based on six genomic regions within the AF cluster: *aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*, *aflI/aflO*. Six blocks of recombination first recovered in (Moore, Singh et al. 2009) are outlined and numbered in the Georgia LD plot. The populations are grouped based on proportion of reproductive methods as either asexual > sexual, or as sexual > asexual. Tables S2.1-S2.5 (Supporting Information) list the isolates included in each quantile plot and LD analysis.

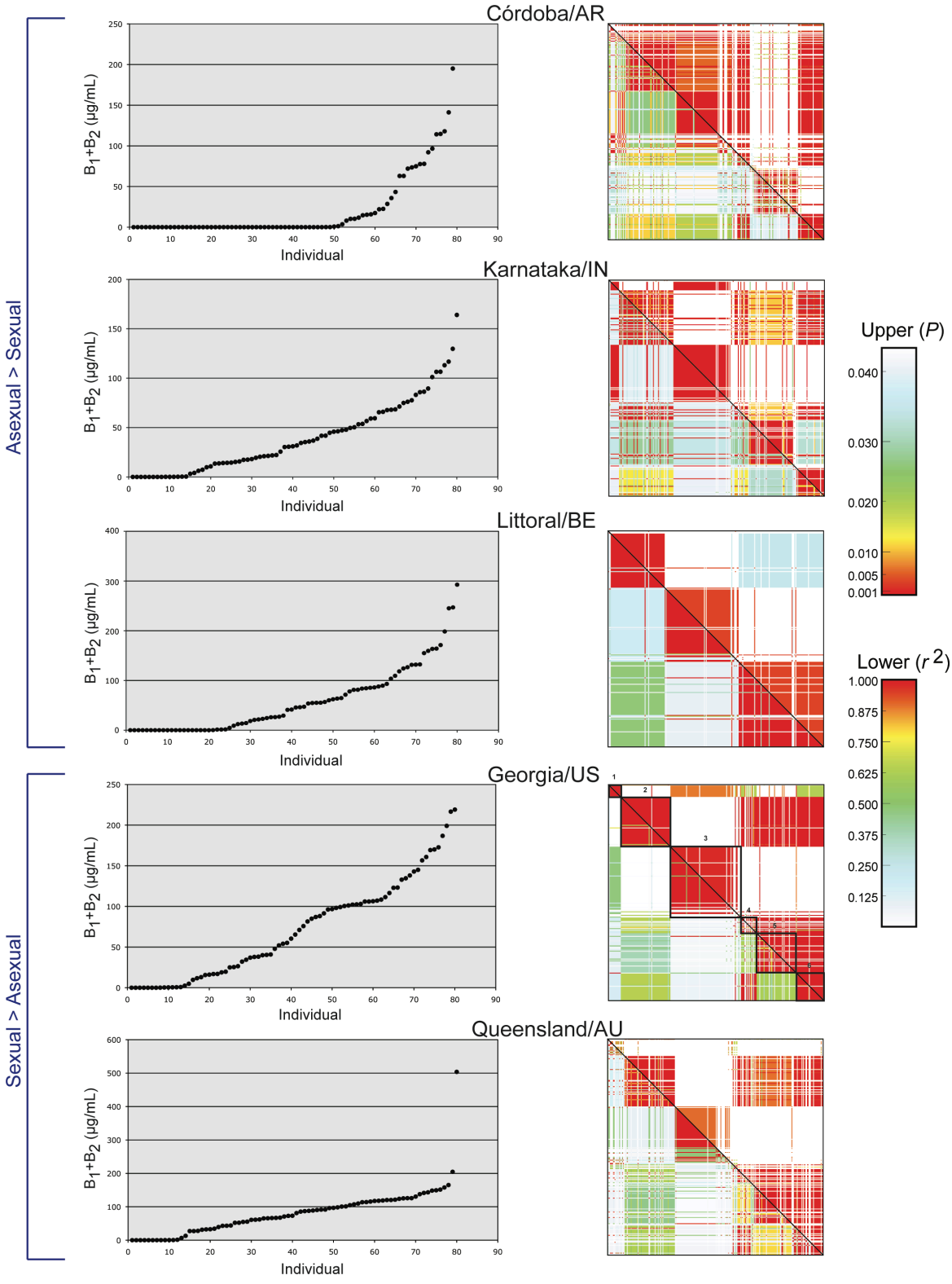

Aspergillus flavus L
Córdoba/AR
Karnataka/IN
Littoral/BE
Georgia/US
Queensland/AU
Asexual > Sexual
Sexual > Asexual
B_1+B_2 (µg/mL)
Individual
Upper (*P*)
0.040
0.030
0.020
0.010
0.005
0.001
Lower (r^2)
1.000
0.875
0.750
0.625
0.500
0.375
0.250
0.125

Table S2.1 *Aspergillus flavus* L isolates from Córdoba, Argentina

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
396	1	0.0 (0)	0.0 (0)	0.0	H1
397[1]	1	0.0 (0)	0.0 (0)	0.0	H2
398[1]	1	0.0 (0)	0.0 (0)	0.0	H2
399	1	15.1 (1)	0.2 (0)	15.3	H3
400	1	0.0 (0)	0.0 (0)	0.0	H4
401[1]	1	0.0 (0)	0.0 (0)	0.0	H2
402[1]	1	0.0 (0)	0.0 (0)	0.0	H10
403	1	0.0 (0)	0.0 (0)	0.0	H1
404[1]	1	0.0 (0)	0.0 (0)	0.0	H2
405	2	0.0 (0)	0.0 (0)	0.0	H5
406	1	62.3 (21)	0.7 (0.3)	63.0	H6
407	1	0.0 (0)	0.0 (0)	0.0	H4
408[1]	1	0.0 (0)	0.0 (0)	0.0	H2
409	2	112.2 (17)	2 (0.3)	114.2	H13
410	1	22.4 (0.8)	0.1 (0)	22.5	H19
411	1	76.3 (9)	1.3 (0.1)	77.6	H15
412	1	0.0 (0)	0.0 (0)	0.0	H4
413	1	43 (3)	0.3 (0)	43.3	H17
414[1]	1	0.0 (0)	0.0 (0)	0.0	H2
415	2	138.4 (31)	2.8 (0.8)	141.2	H13
416	1	70.8 (6)	1.2 (0.1)	72.0	H15
417	2	14.5 (5)	0.1 (0)	14.6	H20
418[1]	1	0.0 (0)	0.0 (0)	0.0	H2
419[1]	1	0.0 (0)	0.0 (0)	0.0	H2
420	1	94.8 (19)	1.8 (0.3)	96.6	H16
421	1	17.3 (4)	0.1 (0)	17.4	H19
422	1	190.1 (22)	5 (0.5)	195.1	H12
423[1]	1	0.0 (0)	0.0 (0)	0.0	H2
424	2	115.5 (29)	2.4 (0.8)	117.9	H13
425[1]	1	0.0 (0)	0.0 (0)	0.0	H10
426	1	72.2 (16)	1.3 (0.2)	73.5	H15
427	2	10.1 (0.8)	0.1 (0)	10.2	H20
428[1]	1	0.0 (0)	0.0 (0)	0.0	H10
429	1	0.0 (0)	0.0 (0)	0.0	H4
430	1	0.7 (0.2)	0.0 (0)	0.7	H24
431[1]	1	0.0 (0)	0.0 (0)	0.0	H2
432[1]	1	0.0 (0)	0.0 (0)	0.0	H10
433[1]	1	0.0 (0)	0.0 (0)	0.0	H10
434[1]	1	0.0 (0)	0.0 (0)	0.0	H2
435[1]	1	0.0 (0)	0.0 (0)	0.0	H10
436	1	90.5 (9)	1.6 (0.2)	92.1	H18
437[1]	1	0.0 (0)	0.0 (0)	0.0	H10
438	1	3.6 (0.7)	0.0 (0)	3.6	H1

Table S2.1 Continued

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
439[1]	1	0.0 (0)	0.0 (0)	0.0	H10
440	2	15.7 (6)	0.1 (0)	15.8	H21
441	1	0.0 (0)	0.0 (0)	0.0	H1
442[1]	1	0.0 (0)	0.0 (0)	0.0	H2
443	1	112.9 (24)	1.8 (0.2)	114.7	H16
444	1	0.0 (0)	0.0 (0)	0.0	H4
445	1	1.2 (0.3)	0.0 (0)	1.2	H2
446	1	76.7 (18)	1.2 (0.3)	77.9	H15
447[1]	1	0.0 (0)	0.0 (0)	0.0	H10
448[1]	1	0.0 (0)	0.0 (0)	0.0	H10
449[1]	1	0.0 (0)	0.0 (0)	0.0	H10
450	1	73.7 (9)	1.2 (0.2)	74.9	H15
451	1	0.1 (0.1)	0.0 (0)	0.1	H23
452[1]	1	0.0 (0)	0.0 (0)	0.0	H2
453[1]	1	0.0 (0)	0.0 (0)	0.0	H2
454[1]	1	0.1 (0)	0.0 (0)	0.1	H22
455	1	0.0 (0)	0.0 (0)	0.0	H24
456[1]	1	0.0 (0)	0.0 (0)	0.0	H2
457	2	28.5 (5)	0.1 (0)	28.6	H14
458[1]	1	0.0 (0)	0.0 (0)	0.0	H8
459	2	21.8 (2)	0.2 (0)	22.0	H17
460[1]	1	0.0 (0)	0.0 (0)	0.0	H2
461[1]	1	0.0 (0)	0.0 (0)	0.0	H2
462	1	35.7 (3)	0.1 (0)	35.8	H7
463[1]	1	0.0 (0)	0.0 (0)	0.0	H10
464[1]	1	0.0 (0)	0.0 (0)	0.0	H10
465[1]	1	0.0 (0)	0.0 (0)	0.0	H10
466[1]	1	0.0 (0)	0.0 (0)	0.0	H11
467	1	0.0 (0)	0.0 (0)	0.0	H9
468	2	7.9 (0.9)	0.1 (0)	8.0	H21
469[1]	1	0.0 (0)	0.0 (0)	0.0	H2
470	2	11.6 (3)	0.1 (0)	11.7	H20
471	2	10.2 (4)	0.1 (0)	10.3	H20
472	1	62.1 (4)	0.9 (0)	63.0	H15
474	1	0.0 (0)	0.0 (0)	0.0	H1
475	1	76.4 (10)	1.1 (0.2)	77.5	H15

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.1.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*.
[1] AF- isolates that group with Geiser's IB clade.

Table S2.2 *Aspergillus flavus* L isolates from Karnataka, India

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
1229	2	20.2 (7)	0.2 (0.1)	20.4	H1
1230	2	116.1 (72)	0.6 (0.4)	116.7	H2
1231	2	21.7 (3)	0.5 (0.1)	22.2	-
1232	2	125.9 (13)	3.8 (0.4)	129.7	-
1233	1	100.4 (12)	0.7 (0.1)	101.1	H19
1234	1	0.3 (0)	0.0 (0)	0.30	-
1235	1	19 (2)	0.4 (0)	19.4	-
1236	1	21 (9)	0.3 (0.2)	21.3	-
1237	2	80.1 (8)	3.0 (0.2)	83.1	H33
1238	1	64.1 (9)	1.4 (0.4)	65.5	-
1239	1	0.0 (0)	0.0 (0)	0.0	H9
1240	2	58.6 (27)	0.8 (0.3)	59.4	-
1241	1	9.7 (3)	0.1 (0)	9.8	H10
1242	2	70.3 (9)	1.0 (0.1)	71.3	-
1243	1	75.2 (6)	0.9 (0.1)	76.1	-
1244	1	45 (23)	1.2 (0.8)	46.2	-
1245	2	66.4 (12)	1.3 (0.1)	67.7	H25
1246	2	66.6 (7)	1.4 (0.2)	68.0	-
1247	2	0.0 (0)	0.0 (0)	0.0	-
1248	2	84.5 (5)	1.3 (0.2)	85.8	-
1249	2	162.9 (26)	1.1 (0.1)	164	H2
1250	1	58 (9)	1.1 (0.2)	59.1	H26
1251	1	17.2 (6)	0.1 (0)	17.3	H17
1252	1	13.4 (2)	0.3 (0)	13.7	H28
1253	1	34.7 (21)	0.6 (0.4)	35.3	H27
1254	2	104.9 (6)	1.5 (0.2)	106.4	H16
1255	1	35.7 (3)	1.1 (0.2)	36.8	H30
1256	2	30.4 (9)	0.6 (0.2)	31.0	-
1257	1	0.1 (0)	0.0 (0)	0.10	H12
1258	1	10.6 (0.8)	0.1 (0)	10.7	H18
1259	1	33.3 (5)	1.1 (0.2)	34.4	-
1260	2	112.5 (11)	0.6 (0.1)	113.1	H2
1261	1	0.0 (0)	0.0 (0)	0.0	-
1262	1	56.1 (12)	0.5 (0.1)	56.6	H21
1263	2	15 (2)	0.2 (0)	15.2	-
1264	2	15.6 (5)	0.1 (0)	15.7	H20
1265	1	6.2 (2)	0.0 (0)	6.20	H22
1266	1	13.4 (3)	0.1 (0)	13.5	H3
1267	1	83.3 (7)	3.0 (0.3)	86.3	-
1268	1	63.4 (11)	2.5 (0.4)	65.9	H31
1269	1	0.1 (0.1)	0.0 (0)	0.10	H34
1270	2	20.1 (12)	0.9 (0.5)	21.0	H24
1271	1	3.2 (2)	0.0 (0)	3.20	H14

Table S2.2 Continued

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
1272	1	0.1 (0)	0.0 (0)	0.10	H7
1273	1	49.4 (18)	0.9 (0.5)	50.3	-
1274[1]	2	0.0 (0)	0.0 (0)	0.0	H4
1275	1	47 (9)	0.2 (0)	47.2	H20
1276	1	46.2 (14)	1.6 (0.5)	47.8	H32
1277	1	18 (6)	0.1 (0)	18.1	H23
1278	1	21.2 (2)	0.6 (0.1)	21.8	-
1279	2	49 (18)	0.6 (0.2)	49.6	H13
1280	1	105 (18)	1.5 (0.2)	106.5	H11
1281	1	52.8 (41)	0.5 (0.4)	53.3	H22
1282	1	0.4 (0.2)	0.0 (0)	0.40	H12
1283	1	41.4 (23)	0.3 (0.2)	41.7	-
1284	2	38 (10)	0.7 (0.2)	38.7	-
1285	2	41.1 (9)	0.7 (0.1)	41.8	-
1286	1	31.5 (8)	0.4 (0.1)	31.9	-
1287	1	76.8 (26)	0.8 (0.3)	77.6	-
1288	1	4.0 (0.5)	0.1 (0)	4.10	-
1289	2	17.4 (8)	0.1 (0.1)	17.5	-
1290	1	88.7 (32)	0.8 (0.3)	89.5	H12
1291[1]	1	0.0 (0)	0.0 (0)	0.0	H5
1292	1	53.1 (13)	0.6 (0.1)	53.7	-
1293[1]	2	0.0 (0)	0.0 (0)	0.0	H4
1294	2	35.1 (14)	0.8 (0.3)	35.9	-
1295	2	43.9 (23)	0.8 (0.6)	44.7	H29
1296	1	0.2 (0.1)	0.0 (0)	0.20	H12
1297	1	25.1 (9)	0.5 (0.2)	25.6	H27
1298	1	14.2 (5)	0.3 (0.1)	14.5	-
1299	1	44.9 (16)	0.9 (0.3)	45.8	-
1300	2	30 (4)	0.6 (0)	30.6	-
1301	2	7.3 (4)	0.0 (0)	7.30	-
1302	1	14 (4)	0.3 (0.1)	14.3	-
1303[1]	1	0.0 (0)	0.0 (0)	0.0	H6
1304	1	0.0 (0)	0.0 (0)	0.0	H9
1305	2	67.5 (15)	0.9 (0.3)	68.4	H15
1306	1	29.6 (6)	0.7 (0.1)	30.3	H8
1307	1	73.1 (15)	1.8 (0.4)	74.9	H1
1308	2	13.8 (2)	0.3 (0.1)	14.1	-

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.1.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[1] AF- isolates that group with Geiser's IB clade.

Table S2.3 *Aspergillus flavus* L isolates from Littoral, Benin

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
1027	1	83 (52)	2.3 (1)	85.3	H1
1028	1	59.6 (25)	0.5 (0.2)	60.1	H2
1029	2	46 (5)	1.2 (0)	47.2	H50
1030	1	86 (14)	1.8 (0.4)	87.8	H54
1031	2	88.5 (47)	4.6 (1)	93.1	H59
1032	1	42.4 (19)	3.2 (0.5)	45.6	H60
1033	2	152.8 (13)	6.7 (1)	159.5	H16
1034	2	53.3 (11)	1.4 (0.3)	54.7	H51
1035	1	52.8 (11)	0.9 (0.1)	53.7	H32
1036	2	28.8 (18)	0.5 (0.2)	29.3	H34
1037	1	45.4 (12)	0.8 (0.3)	46.2	H17
1038	2	22.4 (8)	0.4 (0.1)	22.8	H28
1039	2	0.0 (0)	0.0 (0)	0.0	H6
1040	2	81.5 (37)	2.2 (1)	83.7	H56
1041	1	63.8 (16)	0.8 (0.1)	64.6	H17
1042	2	0.0 (0)	0.0 (0)	0.0	H38
1043	1	0.0 (0)	0.0 (0)	0.0	H10
1044	1	0.0 (0)	0.0 (0)	0.0	H25
1045	1	0.0 (0)	0.0 (0)	0.0	H8
1046	2	13.5 (5)	0.8 (0.2)	14.3	H1
1047	2	82.3 (18)	2 (0.6)	84.3	H58
1048	1	0.0 (0)	0.0 (0)	0.0	H25
1049	1	79.4 (7)	1.9 (0.1)	81.3	H26
1050	1	8.7 (4)	0.3 (0)	9.00	H5
1051	2	11.7 (2)	0.5 (0.1)	12.2	H1
1052	2	85.5 (15)	0.5 (0.1)	86.0	H19
1053	2	0.0 (0)	0.0 (0)	0.0	H6
1054[2]	2	0.0 (0)	0.0 (0)	0.0	H39
1055	1	0.0 (0)	0.0 (0)	0.0	H13
1056	1	21.5 (4)	0.1 (0)	21.6	H37
1057	2	17.9 (1)	0.5 (0.1)	18.4	H55
1058	2	13.2 (3)	0.1 (0)	13.3	H47
1059	1	102.4 (13)	1.2 (0.1)	103.6	H17
1060	1	61.8 (10)	0.4 (0)	62.2	H20
1061[2]	1	0.0 (0)	0.0 (0)	0.0	H3
1062	1	235 (83)	10.1 (5)	245.1	H57
1063	1	0.0 (0)	0.0 (0)	0.0	H36
1064	1	153.2 (20)	1.8 (0.4)	155.0	H33
1065	2	4.7 (1)	0.0 (0)	4.80	H2
1066	1	69.9 (16)	1.5 (0.6)	71.4	H29
1067	1	0.0 (0)	0.0 (0)	0.0	H15
1068	1	0.0 (0)	0.0 (0)	0.0	H6
1069	2	20.4 (5)	0.3 (0)	20.7	H51

Table S2.3 Continued

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
1070	1	0.0 (0)	0.0 (0)	0.0	H6
1071	1	0.0 (0)	0.0 (0)	0.0	H12
1072	2	55.7 (13)	1.0 (0.2)	56.7	H2
1073	2	54.4 (36)	0.8 (0.4)	55.2	H2
1074	2	124 (23)	2.6 (0.4)	126.6	H4
1075	1	126.7 (102)	5.1 (3)	131.8	H52
1076	1	87.3 (37)	2.2 (0.7)	89.5	H53
1077	1	113.7 (24)	4.6 (0.6)	118.3	H42
1078	1	106.6 (40)	2.8 (1)	109.4	H40
1079	2	130.5 (67)	1.8 (0.9)	132.3	H22
1080	1	38.5 (12)	3 (0.3)	41.5	H41
1081	1	165.5 (49)	5.7 (1)	171.2	H17
1082	1	0.0 (0)	0.0 (0)	0.0	H27
1083	2	40 (6)	0.9 (0.2)	40.9	H14
1084	1	1.3 (0.3)	0.0 (0)	1.30	H9
1085	1	1.2 (0.2)	0.0 (0)	1.20	H11
1086	1	237.3 (29)	9.6 (0.9)	246.9	H45
1087	1	25.8 (10)	0.2 (0.1)	26.0	H30
1088	1	126.5 (28)	4.9 (0.9)	131.4	H43
1089	2	0.0 (0)	0.0 (0)	0.0	H2
1090	2	76.4 (12)	1.5 (0.3)	77.9	H23
1091	1	0.2 (0.1)	0.0 (0)	0.2	H38
1092	1	157.7 (31)	6.4 (1)	164.1	H44
1093	1	26.9 (4)	0.2 (0)	27.1	H34
1094	1	0.0 (0)	0.0 (0)	0.0	H31
1095	2	121.9 (8)	2.2 (0.3)	124.1	H48
1096	1	54.5 (24)	0.6 (0.3)	55.1	H24
1097	2	285.5 (57)	7.1 (2)	292.6	H18
1098	1	192.1 (67)	6.4 (2)	198.5	H21
1099	2	0.0 (0)	0.0 (0)	0.0	H38
1100[1]	2	0.1 (0)	26.2 (7)	26.3	H49
1101	1	0.0 (0)	0.0 (0)	0.0	H34
1102[1]	2	0.1 (0)	24.3 (3)	24.4	H49
1103	1	62.2 (5)	1.4 (0.2)	63.6	H35
1104	1	79.3 (16)	1.8 (0.4)	81.1	H35
1105	1	157.9 (27)	5.6 (0.6)	163.5	H46
1106	1	1.6 (0.6)	0.0 (0)	1.60	H7

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.1.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*.
[1] Isolates that produce more B_2 than B_1.
[2] AF- isolates that group with Geiser's IB clade.

Table S2.4 *Aspergillus flavus* L isolates from Georgia, United States

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
229	2	39.2 (22)	0.8 (0.5)	40.0	H1
230	2	15.8 (13)	0.3 (0.2)	16.1	H1
231	2	33.7 (13)	0.7 (0.2)	34.4	H1
232	2	16.6 (11)	0.4 (0.2)	17.0	H1
233	2	31.4 (7)	0.7 (0.1)	32.1	H1
234	2	85.6 (6)	1.4 (0.3)	87.0	H2
235	2	104 (17)	1.8 (0.1)	105.8	H2
236	2	106 (15)	1.6 (0.7)	107.6	H2
237	2	130.4 (13)	2.5 (0.9)	132.9	H2
238	2	98.6 (18)	1.9 (0.6)	100.5	H2
239	2	86.7 (24)	1.4 (0.4)	88.1	H2
240	2	101.2 (20)	1.9 (0.3)	103.1	H2
241	2	59.5 (15)	0.9 (0.7)	60.4	H2
242	2	115.1 (11)	1.4 (0.5)	116.5	H2
244	1	104.3 (10)	1.7 (0.1)	106.0	H6
245	2	96.3 (6)	1.9 (0.6)	98.2	H28
246	2	100 (9)	2.2 (0.1)	102.2	H28
247	2	94.2 (13)	2.0 (0.3)	96.2	H28
248	2	131.8 (5)	2.8 (0.6)	134.6	H28
249	2	119.8 (13)	3.1 (0.5)	122.9	H28
250	2	120.1 (17)	3.1 (0.5)	123.2	H28
251	2	157 (20)	3.8 (0.7)	160.8	H28
252	2	0.0 (0)	0.0 (0)	0.0	-
253	2	0.0 (0)	0.0 (0)	0.0	H29
254	2	0.0 (0)	0.0 (0)	0.0	-
255	2	0.0 (0)	0.0 (0)	0.0	-
256	2	0.0 (0)	0.0 (0)	0.0	-
257	2	0.0 (0)	0.0 (0)	0.0	-
258	2	168.3 (29)	4.4 (0.6)	172.7	H7
259	2	134.3 (5)	3.8 (0.8)	138.1	H7
260	2	210.6 (36)	6 (0.8)	216.6	H7
261	2	193.5 (19)	5.6 (0.9)	199.1	H7
262	2	213 (22)	6.1 (1)	219.1	H7
263	2	0.8 (0.1)	0.0 (0)	0.8	H30
264	2	0.5 (0.1)	0.0 (0)	0.5	H30
265	2	0.1 (0.1)	0.0 (0)	0.1	H30
267	2	109 (41)	2.5 (1)	111.5	H34
268	2	141.2 (50)	3.8 (2)	145.0	H34
269	2	79 (37)	2.5 (1)	81.5	H34
270	2	72.5 (7)	3.3 (0.7)	75.8	H17
271	2	89.4 (7)	2.1 (1)	91.5	H17
272	2	64.2 (7)	1.2 (0.2)	65.4	H22
273	2	53.7 (11)	0.4 (0.1)	54.1	H22

Table S2.4 Continued

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
274	2	4.8 (2)	0.1 (0.1)	4.9	H12
275	2	2.4 (1)	0.1 (0)	2.5	H12
276	2	0.4 (0.1)	0.0 (0)	0.4	H24
277[1]	2	0.0 (0)	0.0 (0)	0.0	H3
278	1	99.4 (46)	1.8 (1)	101.2	H9
279	2	12.9 (1)	0.1 (0)	13.0	H37
280	2	36.9 (8)	1.3 (0.5)	38.2	H21
281	2	154 (9)	2.6 (0.5)	156.6	H18
282	1	164.6 (27)	5.5 (1)	170.1	H14
283	1	9.6 (1)	0.1 (0)	9.7	H32
284	2	24.6 (3)	0.4 (0)	25.0	H23
285	2	185.3 (47)	1.5 (0.8)	186.8	H31
286	2	24.7 (4)	0.7 (0.1)	25.4	H1
287	2	104.1 (28)	2.3 (1)	106.4	H8
288	2	97.4 (18)	1.8 (0.6)	99.2	H35
289	1	40.2 (26)	0.7 (0.4)	40.9	H26
290	2	15.3 (5)	0.3 (0.1)	15.6	H2
291	1	47.6 (14)	0.3 (0.1)	47.9	H27
292	2	105.6 (25)	2.8 (0.4)	108.4	H11
293	1	54.2 (13)	1 (0.2)	55.2	H21
294	2	19.6 (5)	0.1 (0)	19.7	H2
295	2	164.6 (17)	4.7 (0.8)	169.3	H13
296	1	0.3 (0.1)	0.0 (0)	0.3	H25
297	1	25.7 (11)	0.8 (0.4)	26.5	H15
298	1	18.7 (4)	0.2 (0)	18.9	H19
299	2	36.2 (19)	0.8 (0.5)	37.0	H22
300	1	51.4 (11)	0.6 (0.2)	52.0	H2
301	1	16.8 (3)	0.1 (0)	16.9	H16
302	2	83.2 (16)	1.7 (0.6)	84.9	H20
303	1	94.9 (15)	1.8 (0.2)	96.7	H10
304	2	69.6 (12)	1.5 (0.7)	71.1	H11
305	1	37.4 (11)	0.5 (0.2)	37.9	H33
306	2	11.5 (0.7)	0.2 (0.1)	11.7	H36
307	2	139.8 (5)	3.1 (0.1)	142.9	H4
308	1	39.8 (7)	0.6 (0.1)	40.4	H5

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.1.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[1] AF- isolates that group with Geiser's IB clade.

Table S2.5 *Aspergillus flavus* L isolates from Queensland, Australia

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
640[1]	2	496.2 (368)	7.5 (6)	503.7	H1
641	1	155 (23)	2.4 (0.3)	157.8	-
642[1]	2	162.6 (14)	2.5 (0.1)	165.1	H28
643	2	0.1 (0)	0.0 (0)	0.10	H3
644	1	59.9 (6)	1.0 (0.1)	60.9	-
645	1	64.2 (6)	1.0 (0.1)	65.2	-
646	2	64.5 (19)	1.1 (0.3)	65.6	H17
647	2	51.4 (19)	0.9 (0.2)	52.3	-
648	1	55.2 (9)	0.2 (0.1)	55.4	H5
649	2	119 (20)	2.4 (0.6)	121.4	-
650	1	13 (2)	0.0 (0)	13.0	H10
651	2	112.5 (8)	1.5 (0.2)	114.0	H29
652[1]	1	101.9 (19)	2.8 (0.3)	104.7	H25
653	2	6.6 (1)	0.1 (0)	6.70	-
654[1]	2	72.8 (6)	0.4 (0.1)	73.2	-
655	2	84.1 (22)	1.4 (0.5)	85.5	H17
656	2	147.2 (21)	2.2 (0.5)	149.4	H26
657	2	42.7 (10)	0.5 (0.2)	43.2	H23
658	2	149.1 (24)	2.4 (0.7)	151.5	H26
659	2	0.0 (0)	0.0 (0)	0.0	H21
660	2	146.3 (32)	2.1 (0.6)	148.4	H29
661	2	0.4 (0.1)	0.0 (0)	0.40	H15
662	1	99.5 (12)	1.1 (0.1)	100.6	H11
663	2	0.3 (0)	0.0 (0)	0.30	H15
664	2	27.2 (4)	0.1 (0)	27.3	H7
665	2	31.2 (7)	0.1 (0)	31.3	-
666	2	32.3 (4)	0.1 (0)	32.4	H8
667	2	27.8 (12)	0.1 (0)	27.9	H8
668	1	60.7 (14)	0.6 (0.1)	61.3	-
669	2	114 (9)	1.8 (0.2)	115.8	-
670	1	43.8 (13)	0.2 (0.1)	44.0	H9
671	1	0.0 (0)	0.0 (0)	0.0	H6
672	2	200.7 (161)	3.9 (4)	204.6	H31
673[4]	2	0.0 (0)	0.0 (0)	0.0	H4
674	2	71 (25)	1.7 (0.5)	72.7	H24
675	1	0.2 (0)	0.0 (0)	0.20	H27
676	2	124.3 (16)	1.5 (0.2)	125.8	H29
677	2	53.8 (5)	0.9 (0.2)	54.7	H12
678	2	42.2 (14)	0.7 (0.2)	42.9	H18
679	2	66.1 (16)	0.6 (0.2)	66.7	H16
680	2	88.8 (17)	1.6 (0.4)	90.4	H18
681	2	105 (2)	1.9 (0.2)	106.9	-
682	1	96.7 (30)	1.6 (0.7)	98.3	H18

Table S2.5 Continued

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
683	2	141.4 (16)	1.8 (0)	143.2	H30
684	1	90 (27)	1.5 (0.6)	91.5	H32
685	1	117.2 (34)	1.6 (0.5)	118.8	H20
686	2	0.0 (0)	0.0 (0)	0.0	H14
687	2	116 (3)	1.3 (0.1)	117.3	-
688	1	122 (11)	1.8 (0.1)	123.8	H20
689	1	95.4 (7)	0.7 (0)	96.1	-
690	2	32.6 (4)	0.2 (0)	32.8	-
691	2	66 (11)	0.3 (0.1)	66.3	-
692	1	61.1 (22)	1.0 (0.4)	62.1	-
693	1	122 (21)	3.4 (0.4)	125.4	-
694	1	69.8 (7)	1.1 (0.1)	70.9	-
695	1	0.1 (0)	0.0 (0)	0.10	H22
696	2	0.0 (0)	0.0 (0)	0.0	H14
697	2	135.8 (18)	1.9 (0.5)	137.7	H29
698	2	0.0 (0)	0.0 (0)	0.0	H13
699	2	119 (39)	1.4 (0.6)	120.3	-
700	2	129 (29)	1.6 (0.3)	130.6	-
701	2	34 (7)	0.4 (0.1)	34.4	H25
702	1	123.2 (34)	2 (0.6)	125.2	H20
703	2	116.6 (19)	1.4 (0.4)	118.0	H29
704	1	139.3 (33)	2.2 (0.6)	141.5	H19
705	1	119 (29)	1.0 (0.3)	120.0	-
706	2	113 (16)	1.0 (0.1)	114.0	-
707	2	39.2 (15)	0.1 (0)	39.3	-
708	1	89.7 (17)	2 (0.4)	91.7	H21
709[2,3]	1	4.1 (0.9)	23.2 (5)	27.3	H5
710	2	99.2 (29)	2 (0.6)	101.2	-
711	2	80.7 (14)	0.5 (0.1)	81.2	H12
712	1	67.2 (17)	0.2 (0.1)	67.4	H5
713	2	109 (48)	1.2 (0.7)	110.2	-
714	1	86.1 (14)	0.9 (0.1)	87.0	-
715	2	86.7 (27)	1.1 (0.5)	87.8	-
716	2	95.1 (29)	1.7 (0.6)	96.8	-
717	1	87.7 (35)	1.1 (0.6)	88.8	-
718	1	51.6 (11)	0.5 (0.1)	52.1	-
719[2]	2	1.3 (0.5)	0.0 (0)	1.3	H2

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.1.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[1] Isolates that also produce $G_1 < 0.5$ μg/mL.
[2] Isolates that also produce OMST < 40 μg/mL.
[3] Isolate produces more B_2 than B_1.
[4] AF- isolates that group with Geiser's IB clade.

Figure S2.2 Graphical representations of population aflatoxin concentrations and recombination in populations of *A. parasiticus* with different proportions of asexual and sexual reproduction. Quantile plots were created which show the number of individuals along the x-axis and the G_1/B_1 aflatoxin (AF) ratio. To the right of each quantile plot is an LD plot (if recombination was detected) in which the upper triangular matrix represents the *P* values calculated using Fisher's Exact test; the lower triangular matrix represents r^2, the coefficient of determination between allelic states at pairs of sites. Colored shading in the LD plot indicates statistical significance in linkage (upper) and strength of associations (lower). The LDs are based on seven genomic regions within the AF cluster: *aflB/aflR*, *aflS/aflH*, *aflH/aflJ*, *aflJ/aflE*, *aflE/aflM*, *aflG/aflL*, *aflK/aflV*. Five blocks of recombination first recovered in (Carbone, Jakobek et al. 2007) are outlined and numbered in the Georgia LD plot. The populations are grouped based on proportion of reproductive methods as either asexual > sexual, or as sexual > asexual. Tables S2.6-S2.8 (Supporting Information) list the isolates included in each quantile plot and LD analysis.

Aspergillus parasiticus

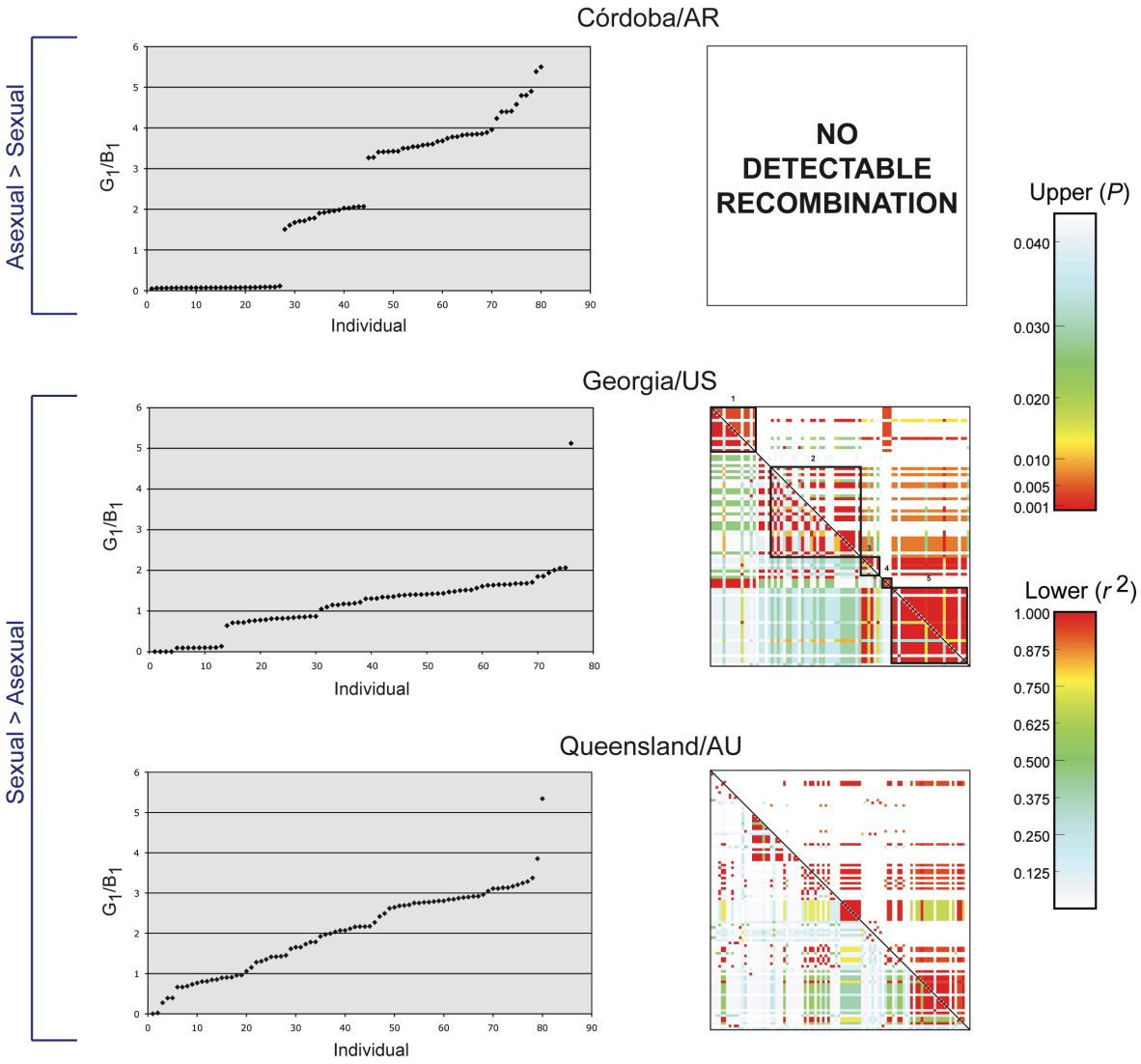


Table S2.6 *Aspergillus parasiticus* isolates from Córdoba, Argentina

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
480	1	23.2 (5)	379 (56)	0.061	H1
481	1	244.4 (33)	128.5 (20)	1.902	-
482	1	251.4 (26)	128.2 (8)	1.961	-
483	1	214.8 (32)	121.5 (12)	1.768	-
484	1	376.1 (9)	98 (0.5)	3.838	H12
485	1	16.1 (8)	218.1 (79)	0.074	H1
486	1	293.5 (4)	54.5 (0.2)	5.385	-
487	1	337.2 (22)	99 (9)	3.406	H12
488	1	161.2 (48)	35.2 (11)	4.580	-
489	1	340.4 (32)	104.2 (7)	3.267	H12
490	1	45.5 (11)	23.4 (5)	1.944	H2
491	1	374.7 (29)	94.7 (3)	3.957	H12
492	1	33.4 (4)	373.8 (58)	0.089	-
493	1	314.2 (27)	87.5 (9)	3.591	-
494	2	215.6 (21)	65.8 (13)	3.277	H5
495	1	20.4 (5)	273.5 (52)	0.075	H1
496	1	225 (12)	111 (8)	2.027	H12
497	1	350.1 (50)	73 (11)	4.796	H12
498	1	315 (41)	64.3 (14)	4.899	-
499	1	28.6 (6)	255.1 (51)	0.112	H1
500	1	26.6 (4)	298.9 (49)	0.089	H1
501	1	344.3 (17)	78.3 (4)	4.397	-
502	1	341.1 (25)	77.3 (3)	4.413	H12
503	1	333.8 (36)	69.5 (9)	4.803	-
504	1	221.5 (12)	115.4 (6)	1.919	H12
505	1	197.2 (32)	122.7 (23)	1.607	H12
506	1	276 (11)	77.2 (0.9)	3.575	H8
507	1	260 (21)	155.3 (20)	1.674	H7
508	1	310.4 (26)	80.5 (9)	3.856	H12
509	1	288.1 (23)	81.3 (4)	3.544	H12
510	1	345.9 (15)	98.9 (3)	3.498	H12
511	1	22 (4)	305.8 (41)	0.072	H1
512	1	368.7 (19)	181.7 (13)	2.029	H12
513	1	371.4 (26)	187.3 (12)	1.983	H16
514	1	360.4 (33)	202.1 (16)	1.783	H12
515	1	341.9 (42)	166.6 (25)	2.052	-
516	1	354.8 (51)	94.8 (12)	3.743	H12
517	1	160 (43)	37.8 (8)	4.233	H3
518	1	139.6 (14)	36.9 (6)	3.783	H3
519	1	19.1 (4)	272.5 (29)	0.070	H11
520	1	342.1 (24)	88.9 (12)	3.848	H12
521	1	349.2 (36)	89.8 (12)	3.889	H12
522	1	284.4 (28)	77.6 (7)	3.665	H12

Table S2.6 Continued

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
523	1	19.2 (2)	264.9 (17)	0.073	H1
524	1	24.8 (6)	356.8 (45)	0.069	H17
525	1	25.1 (4)	321.7 (52)	0.078	H1
526	2	1.1 (0.1)	0.2 (0)	5.500	H6
527	1	205.2 (15)	120.2 (11)	1.707	-
528	1	26.1 (5)	348.6 (69)	0.075	H1
529	1	276.6 (1)	78.2 (3)	3.537	H12
530	1	251.8 (29)	73.8 (5)	3.412	H13
531	1	20.7 (2)	313.7 (33)	0.066	H1
532	1	290.2 (39)	66 (9)	4.397	H12
533	1	292.5 (12)	85.6 (5)	3.417	H12
534	1	30.6 (9)	20.3 (8)	1.507	H14
535	1	20.5 (3)	313.7 (17)	0.065	H15
536	1	18.1 (0.5)	297.5 (13)	0.061	H9
537	1	8.9 (0.3)	193.7 (4)	0.046	H1
538	1	20.5 (3)	291.8 (45)	0.070	H1
539	1	20.9 (2)	309.2 (24)	0.068	H1
540	1	306.3 (29)	89.4 (16)	3.426	H12
541	1	311 (36)	86.3 (6)	3.604	H12
542	1	330.2 (30)	89.7 (5)	3.681	H12
543	1	325.9 (8)	95.1 (5)	3.427	H12
544	1	229.6 (37)	133.9 (19)	1.715	H15
545	1	339.3 (21)	96.8 (3)	3.505	H12
546	1	22.5 (3)	315.1 (38)	0.071	H1
547	1	18.1 (3)	232.3 (46)	0.078	H1
548	1	34.6 (5)	386.7 (50)	0.089	H15
549	1	385.1 (23)	100.4 (7)	3.836	H12
550	1	24 (5)	297.4 (44)	0.081	-
551	1	26.3 (5)	376.3 (43)	0.070	H10
552	1	29.1 (2)	337.7 (26)	0.086	H1
553	1	266.2 (5)	129 (5)	2.064	H4
554	1	233 (11)	112.7 (6)	2.067	H12
555	1	336.9 (16)	88.2 (4)	3.820	H12
556	1	29.3 (2)	390.7 (24)	0.075	H16
557	1	280.3 (26)	74.2 (3)	3.778	-
558	1	20.4 (4)	274.8 (52)	0.074	-
559	1	29.1 (5)	415.4 (58)	0.070	-

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.2.

[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.

[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.

Table S2.7 *Aspergillus parasiticus* isolates from Georgia, United States

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
1	1	94.5 (6)	111.1 (7)	0.851	H1
2	1	72.5 (6)	94.9 (9)	0.764	H1
3	1	71.2 (15)	91.7 (9)	0.776	-
4	1	91.9 (16)	105.8 (19)	0.869	-
5	1	60.4 (3)	94.1 (0.7)	0.642	H1
6	1	54.4 (4)	75.7 (8)	0.719	-
7	1	53.4 (4)	75.3 (3)	0.709	H1
8	1	75.7 (9)	91.6 (11)	0.826	H1
9	1	59.3 (10)	78.8 (7)	0.753	-
10	1	72.2 (13)	88 (11)	0.821	H1
11	1	79.8 (8)	91.7 (3)	0.870	H1
12	1	81.9 (36)	104 (29)	0.788	H1
13	1	68 (18)	83.9 (11)	0.811	H1
14	1	79.1 (12)	94.3 (13)	0.839	H1
15	1	74.5 (12)	91.1 (15)	0.818	H1
16	1	57.1 (18)	79.5 (7)	0.718	-
17	1	98.1 (46)	114.9 (51)	0.854	H1
18	1	168.1 (21)	119.9 (10)	1.402	H16
19	1	102.6 (35)	69.3 (20)	1.481	H6
20	1	172.8 (9)	105.9 (4)	1.632	-
21	1	195.9 (36)	130.2 (25)	1.505	H16
22	1	149.4 (16)	105.7 (2)	1.413	H16
23	1	169.9 (15)	103.1 (5)	1.648	H16
24	1	161.1 (13)	123 (6)	1.310	H16
25	1	158 (37)	112.3 (24)	1.407	H16
26	1	202.5 (35)	133.3 (14)	1.519	H16
27	1	220.7 (31)	141.2 (33)	1.563	H16
28	1	218.9 (18)	155.7 (1)	1.406	-
29	1	200.4 (7)	140.9 (14)	1.422	H16
30	1	170.2 (63)	130.8 (38)	1.301	-
32	1	217.8 (23)	132 (13)	1.650	H1
33	2	117.4 (26)	86.6 (12)	1.356	H1
34	1	227.9 (5)	136.1 (9)	1.675	H1
35	1	212.6 (7)	114.5 (2)	1.857	H12
36	1	223.6 (44)	135.2 (25)	1.654	H1
37	1	288.1 (127)	176.1 (85)	1.636	H1
38	2	95.8 (7)	91.4 (8)	1.048	H1
39	1	231 (39)	137.4 (14)	1.681	H1
40	1	195.5 (2)	141.4 (15)	1.383	H1
41	1	230.8 (31)	143.6 (10)	1.607	-
42	1	262.6 (21)	127.2 (3)	2.065	H1
43	1	36.2 (11)	376.9 (70)	0.096	-
44	1	15.9 (1)	120.3 (14)	0.132	H15

Table S2.7 Continued

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
45	1	26.5 (6)	270.8 (30)	0.098	-
46	1	28.9 (6)	310.2 (66)	0.093	H15
47	1	30.7 (3)	319.7 (10)	0.096	H5
48	1	25.2 (2)	271.3 (17)	0.093	H15
49	1	28.4 (4)	273.7 (20)	0.104	H15
50	1	21.8 (4)	217.5 (22)	0.100	H15
51	1	20.3 (3)	203.9 (20)	0.100	H15
52	1	158.8 (52)	144.5 (52)	1.099	H1
53	2	196.1 (16)	161.2 (10)	1.217	H1
54	1	197.4 (11)	168.1 (12)	1.174	H4
55	1	198.7 (28)	166.8 (31)	1.191	H1
56	1	203.1 (13)	176.6 (12)	1.150	H1
57	2	243.5 (13)	145.2 (15)	1.677	-
58	2	326.2 (17)	158.9 (21)	2.053	H13
59	2	229.3 (23)	124.1 (11)	1.848	H13
60	2	231.3 (8)	135.5 (7)	1.707	H13
61	1	190.9 (41)	142.2 (34)	1.343	H14
62	1	178.2 (9)	132.5 (10)	1.345	H14
63	1	195 (23)	140.2 (21)	1.391	H14
64	1	169.7 (25)	130 (26)	1.305	H3
65[2]	2	0.0 (0)	0.0 (0)	0.0	H9
66	2	0.0 (0)	0.0 (0)	0.0	H9
67[2]	2	143.2 (8)	97.5 (7)	1.469	H8
68	1	161.2 (17)	137 (17)	1.177	H7
69	2	85.1 (62)	59.3 (44)	1.435	H1
70	1	474.6 (43)	236.6 (32)	2.006	H11
71[1]	2	0.0 (0)	0.0 (0)	0.0	H9
72	2	253.4 (34)	167.3 (15)	1.515	H13
73	1	377.6 (22)	73.7 (4)	5.124	H2
74	1	189.2 (9)	97.6 (5)	1.939	H13
75	1	203 (6)	142 (2)	1.430	H1
76[1]	2	0.0 (0)	0.0 (0)	0.0	H10

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.2.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[1] Isolate produces OMST < 200 μg/mL.
[2] Isolate produces OMST > 200 μg/mL.

Table S2.8 *Aspergillus parasiticus* isolates from Queensland, Australia

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
800	2	149 (42)	51.1 (15)	2.916	H1
801	2	141.1 (46)	46.3 (13)	3.048	H1
802	2	117.8 (62)	42.8 (24)	2.752	-
803	2	162.3 (26)	56.5 (6)	2.873	-
804	2	110.5 (64)	39.4 (23)	2.805	H1
805	2	111.7 (18)	41.5 (7)	2.692	H1
806	2	267.1 (80)	50 (14)	5.342	H7
807	2	134.6 (22)	175.5 (37)	0.767	H20
808	2	162 (32)	181.3 (30)	0.894	H14
809[1]	2	0.0 (0)	0.0 (0)	0.0	H19
810	2	74 (3)	107.5 (10)	0.688	-
811	1	64.2 (6)	20.5 (2)	3.132	H10
812	1	68 (17)	20.7 (3)	3.285	-
813	1	51.3 (22)	18.5 (5)	2.773	H8
814	2	86.3 (10)	129.8 (9)	0.665	H21
815	2	53 (6)	17.9 (2)	2.961	-
816	2	165.6 (58)	84.1 (40)	1.969	H26
817	1	211 (23)	103.7 (13)	2.035	-
818	2	187.6 (4)	86.6 (6)	2.166	-
819	1	220.3 (63)	102.1 (39)	2.158	-
820	2	174.8 (13)	80.4 (7)	2.174	-
821	2	173.5 (16)	87.1 (5)	1.992	-
822	2	169.1 (30)	74.5 (19)	2.270	H25
823	2	121.3 (12)	165.6 (9)	0.732	-
824	1	185.5 (28)	112.1 (12)	1.655	H23
825	1	142.9 (49)	362.9 (131)	0.394	H18
826	1	166.4 (9)	117.7 (6)	1.414	-
827	2	0.011 (0)	0.012 (0)	0.917	-
828	2	196.1 (39)	109.8 (23)	1.786	H17
829	1	357.7 (47)	279.3 (40)	1.281	-
830	2	174.5 (41)	98 (19)	1.781	-
831	1	112.5 (24)	29.2 (5)	3.853	-
832	2	7.8 (1)	372.3 (37)	0.021	H11
833	2	16.1 (3)	58.6 (8)	0.275	-
834	1	43.2 (12)	14.9 (4)	2.899	-
835	1	225.1 (9)	195.4 (22)	1.152	H25
836	1	59.4 (15)	21.3 (3)	2.789	H12
837	2	124 (14)	43.6 (4)	2.844	H2
838	1	165.5 (41)	76.4 (21)	2.166	-
839	2	218.4 (43)	103.3 (23)	2.114	H3
840	2	183 (19)	58.8 (5)	3.112	H1
841	2	182.3 (46)	94.8 (27)	1.923	-
842	1	159.1 (28)	49 (8)	3.247	-

Table S2.8 Continued

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
843	1	209 (10)	160.2 (8)	1.305	-
844	1	216.2 (31)	148.9 (22)	1.452	H22
845	1	58.6 (18)	18.7 (3)	3.134	-
846	2	107.1 (28)	37.1 (10)	2.887	-
847	1	151.1 (12)	73 (17)	2.070	-
848	2	338.6 (147)	122.2 (50)	2.771	H11
849	1	55.5 (9)	52.5 (6)	1.057	-
850	1	39.7 (13)	41.5 (12)	0.957	-
851	1	209.1 (10)	129.5 (9)	1.615	H16
852	2	166.2 (9)	52.5 (2)	3.166	-
853	2	84.7 (9)	127.6 (9)	0.664	H21
854	2	163.4 (18)	67.6 (5)	2.417	H27
855	2	141.8 (16)	56.9 (9)	2.492	-
856	2	159.3 (8)	59.4 (0.9)	2.682	-
857	2	152 (25)	54.2 (13)	2.804	-
858	1	125.4 (18)	88.4 (14)	1.419	-
859	2	151.8 (13)	57.4 (6)	2.645	-
860	2	43.8 (15)	15 (4)	2.920	H9
861	2	51.3 (1)	16 (1)	3.206	-
862	1	164.1 (32)	421.3 (61)	0.390	-
863	2	123.9 (11)	47.3 (2)	2.619	H6
864	1	201 (11)	208.8 (24)	0.963	H14
865	1	69.1 (45)	81.9 (37)	0.844	-
866	2	84.7 (11)	105.8 (13)	0.801	-
867	2	91 (8)	112.8 (9)	0.807	H4
868	1	179.5 (17)	132.8 (8)	1.352	H24
869	1	193.2 (18)	116.8 (22)	1.654	-
870	1	26.3 (8)	12.7 (3)	2.071	-
871		185.1 (42)	107 (29)	1.730	-
872	1	201.8 (18)	237.2 (15)	0.851	H15
873	2	165.4 (9)	115.8 (3)	1.428	-
874	2	423.6 (33)	153.9 (11)	2.752	-
875	1	195.3 (9)	216.7 (13)	0.901	H5
876	2	52.9 (7)	17 (3)	3.112	H13
877	2	68.5 (7)	20.3 (2)	3.374	-
878	2	383.8 (26)	141.8 (13)	2.707	-
879	2	437.3 (24)	154.1 (10)	2.838	-

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.2.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[l] Isolate produces OMST < 200 μg/mL.

Figure S2.3 Graphical representations of population aflatoxin concentrations and recombination in populations of *A. flavus* S. Quantile plots were created which show the number of individuals along the x-axis, and either the total B ($B_1 + B_2$) aflatoxin (AF) concentrations in μg/mL or the G_1/B_1 AF ratio. To the right of each quantile plot is an LD plot (if recombination was detected) in which the upper triangular matrix represents the *P* values calculated using Fisher's Exact test; the lower triangular matrix represents r^2, the coefficient of determination between allelic states at pairs of sites. Colored shading in the LD plot indicates statistical significance in linkage (upper) and strength of associations (lower). The LDs are based on six genomic regions within the AF cluster: *aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*, *aflI/aflO*. Tables S2.9-S2.12 (Supporting Information) list the isolates included in each quantile plot and LD analysis.

Aspergillus flavus S

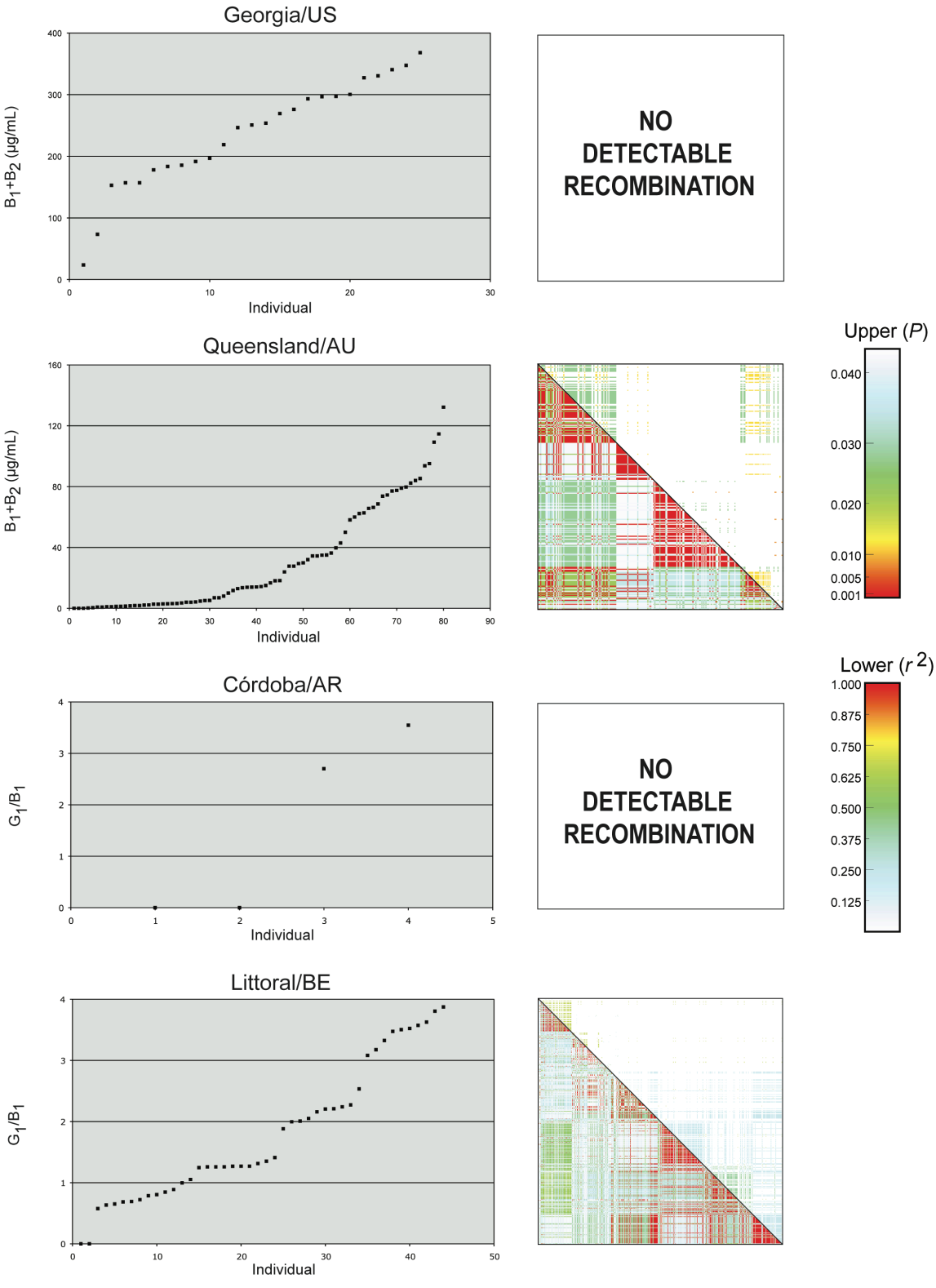


Table S2.9 *Aspergillus flavus* S isolates from Georgia, USA

IC Strain Number[a]	*MAT*	B_1 (μg/mL)[b]	B_2 (μg/mL)[b]	Total B (μg/mL)	MLST[c]
1154	2	148	4.92	152.9	H1
1155	2	71.4	2.08	73.48	H4
1156	2	291	9.39	300.4	H6
1157	2	177	8.46	185.5	H1
1158	2	151	5.57	156.6	-
1160	1	267	9.08	276.1	H6
1161	1	186	5.54	191.5	H6
1162	1	23.3	0.45	23.75	H2
1163	2	174	4.44	178.4	H1
1164	1	244	9.69	253.7	H6
1165	1	178	5.1	183.1	H6
1167	1	356	11.8	367.8	H8
1168	1	242	9.09	251.1	H7
1169	1	317	10.7	327.7	H6
1171	1	287	10	297	H6
1173	2	238	8.21	246.2	-
1174	1	336	11.9	344.9	H1
1175	1	289	8.14	297.1	H6
1176	1	331	9.77	340.8	H5
1177	2	191	6.01	197	H3
1178	1	319	10.9	329.9	H6

[a] Shaded isolate numbers part of a subset for Genetree analysis in Figure S2.4.

[b] AF concentration is based on average of one measurement per isolate. No standard deviation.

[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.

Table S2.10 *Aspergillus flavus* S isolates from Queensland, Australia

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
720[1]	1	11.1 (2)	4.2 (0.4)	2.643	H1
722[1]	2	1.3 (0.4)	0.8 (0.2)	1.625	-
723[2]	2	0.0 (0)	76.2 (0.7)	0.0	H17
725[2]	2	0.0 (0)	108 (46)	0.0	H17
727[2]	2	0.0 (0)	59.4 (13)	0.0	H17
728[2]	2	0.0 (0)	82.3 (14)	0.0	H17
729[2]	2	0.0 (0)	61.5 (2)	0.0	H17
731[1]	1	1.3 (0.1)	1.0 (0.2)	1.3	H22
732[2]	1	0.0 (0)	29.6 (1.2)	0.0	H17
733[1,2]	1	2.0 (0.2)	1.3 (0.1)	1.538	H23
734[1]	1	2.1 (0.5)	1.2 (0.3)	1.75	H10
735[1,2]	1	1.6 (0.1)	1.5 (0.2)	1.6	H23
736[2]	2	0.0 (0)	82.9 (17)	0.0	H17
737[2]	2	0.0 (0)	84.3 (20)	0.0	H17
738	2	0.0 (0)	62 (9)	0.0	-
739	2	0.0 (0)	65 (4)	0.0	-
741[2]	2	0.0 (0)	67.9 (16)	0.0	H21
742[1]	1	5.9 (1)	3.2 (0.5)	1.844	H25
743[2]	2	0.0 (0)	23.8 (4)	0.0	H2
744[1]	1	4.7 (1)	3.1 (0.5)	1.516	H24
748[2]	2	0.0 (0)	78.9 (9)	0.0	H17
749[2]	2	0.0 (0)	73.7 (22)	0.0	H17
751[2]	2	0.0 (0)	14.9 (2)	0.0	H18
753	2	0.0 (0)	30.4 (4)	0.0	H11
755[2]	2	0.0 (0)	9.9 (2)	0.0	H19
758[2]	2	0.0 (0)	6.9 (0.7)	0.0	H13
760	2	0.0 (0)	78.3 (18)	0.0	H11
762[2]	2	0.0 (0)	49.5 (8)	0.0	H17
768[2]	2	0.0 (0)	13.9 (7)	0.0	H19
770[2]	2	0.0 (0)	57.6 (7)	0.0	H14
774[1]	2	8.4 (7)	3.5 (2)	2.4	-
775[1]	2	5.9 (3)	3.0 (1)	1.967	-
776[1]	2	2.8 (0.1)	1.6 (0.3)	1.75	-
777[2]	2	0.0 (0)	12.8 (3)	0.0	H16
778[1,2]	1	2.6 (0.3)	1.4 (0.1)	1.857	H23
779[2]	2	0.0 (0)	113 (12)	0.0	H17
780[2]	2	0.0 (0)	76.6 (20)	0.0	H15
781	2	0.0 (0)	13.4 (3)	0.0	-
782[1]	1	7.8 (2)	4.0 (0.6)	1.95	-
784	2	0.0 (0)	13.7 (3)	0.0	-
785[2]	2	0.0 (0)	34.7 (23)	0.0	H3
786[2]	2	0.0 (0)	31.8 (16)	0.0	H9
787[2]	2	0.0 (0)	27.6 (5)	0.0	H4

Table S2.10 Continued

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
788[1,2]	2	0.3 (0)	0.4 (0)	0.75	H7
790[2]	1	0.0 (0)	16.5 (2)	0.0	H20
791[2]	1	0.0 (0)	13.8 (5)	0.0	H6
792[2]	2	0.0 (0)	73 (9)	0.0	H8
793	1	0.0 (0)	0.0 (0)	0.0	H12
796[2]	1	0.0 (0)	130 (16)	0.0	H17
797[2]	2	0.0 (0)	65.6 (8)	0.0	H17
798[2]	2	0.0 (0)	14.2 (3)	0.0	H5
799[2]	1	0.0 (0)	5.1 (2)	0.0	H10

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.3.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[1] Isolates that also produce OMST < 0.5 μg/mL.
[2] Isolates that share haplotypes with Geiser's group II strains based on *amdS* and *trpC*.

Table S2.11 *Aspergillus flavus* S isolates from Córdoba, Argentina

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
476	2	0.0 (0)	157.9 (17)	0.0	H2
477	1	6.35 (1)	2.35 (0.1)	2.702	H3
478[1]	1	7.06 (0.8)	1.99 (0.3)	3.548	H1
479	2	0.0 (0)	80 (17)	0.0	H2

[a] Shaded isolate numbers part of a subset for Genetree analysis in Figure S2.4.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*.
[1] Isolates that share haplotypes with Geiser's group II strains based on *amdS* and *trpC*.

Table S2.12 *Aspergillus flavus* S isolates from Littoral, Benin

IC Strain Number[a]	*MAT*	G_1 (μg/mL)[b]	B_1 (μg/mL)[b]	G_1/B_1	MLST[c]
1112	1	129.6 (9)	33.5 (5)	3.874	H1
1113	1	13.4 (0.7)	10.7 (0.4)	1.258	H3
1114	2	96.7 (3)	42.5 (3)	2.272	-
1115	2	20.3 (2)	6.59 (0.6)	3.082	-
1117	2	25.7 (7)	12.8 (3)	2.008	H9
1118	1	17.7 (3)	13.9 (2)	1.269	H11
1119	1	18.4 (2)	13 (0.6)	1.411	H13
1120	1	18.9 (0.9)	15 (0.6)	1.26	H3
1121	1	12.5 (1)	9.3 (0.7)	1.351	H3
1122	2	26 (3)	20.6 (2)	1.258	-
1123	2	33.8 (3)	17 (0.9)	1.996	-
1124	2	11.3 (3)	15.7 (3)	0.722	-
1129	1	9.7 (0.8)	14 (3)	0.692	-
1130	1	22.9 (2)	27 (2)	0.847	-
1131	1	11.9 (3)	18.8 (4)	0.634	-
1132	1	8.2 (2)	8.3 (1)	0.997	-
1133	1	14.3 (2)	11.5 (1)	1.247	H14
1134	2	15.5 (3)	12.2 (2)	1.27	H10
1135	1	17.9 (2)	14.1 (2)	1.27	H3
1136	2	20 (2)	5.6 (0.7)	3.574	-
1140	2	17.8 (1)	4.7 (0.3)	3.806	H4
1141	1	16.5 (2)	12.5 (1)	1.314	H12
1142	2	116 (3)	45.8 (2)	2.534	H8
1143	2	15.6 (0.3)	4.7 (0.2)	3.326	-
1144	2	30.3 (3)	14.8 (2)	2.051	H5
1145	2	12.3 (2)	15.6 (1)	0.787	H16
1146	2	33.4 (3)	15.1 (2)	2.209	H7
1147	2	9 (0.5)	13.9 (2)	0.652	H17
1148	2	36 (5)	16.4 (2)	2.205	H7
1149	2	39.9 (8)	18.5 (3)	2.159	H6
1150	2	47.3 (7)	21.1 (3)	2.241	H7
1151	2	10 (1)	14.6 (2)	0.685	H17
1152[1]	2	0.0 (0)	16.6 (1)	0.0	H2
1153[1]	1	0.0 (0)	46.1 (19)	0.0	H15

[a] Shaded isolate numbers part of a subset for LD analysis in Figure S2.3.
[b] AF concentration is based on average of three replicate measurements per isolate. Number in parentheses is standard deviation.
[c] Haplotypes based on four genomic loci: *aflM/aflN*, *aflW/aflX*, *amdS*, *trpC*. Isolates not labeled with haplotype number were not included in MLST analysis.
[1] Isolates designated X strains by Bruce Horn (unpublished).

Table S2.13 Climate, soil, and peanut cultivar for each geographic region

	Georgia/US	Córdoba/AR	Queensland/AU	Littoral/BE	Karnataka/IN
Climate	Temperate	Temperate	Semi-arid subtropical	Semi-arid subtropical	Semi-arid subtropical
Soil	Tifton (sandy loam)	Franco Slimy	Red Kraznozem (sandy loam)	Feralitic Clay	Rampur Series (alfisol/clay)
Peanut Cv.	"Florunner"	"Gran Oleico"	"Streeton"	"Chinese"	"Spanish"

Figure S2.4 *A. flavus* S coalescent-based gene genealogies for six genomic regions within the aflatoxin (AF) gene cluster: *aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*, *aflI/aflO*. The haplotype structure for each genealogy is shown in Table S14 (Supporting Information). Runs of one million simulations were examined to indicate the best tree. Mutations are indicated as dots along branches and time is scaled to the most recent common ancestor (TMRCA) of 1.0 for each genealogy. Haplotypes that exhibit distinct aflatoxin (AF) chemotype profiles are indicated/grouped to the right of the genealogies. All genealogies appear to show either grouping or separation of haplotypes based on AF chemotype differences.

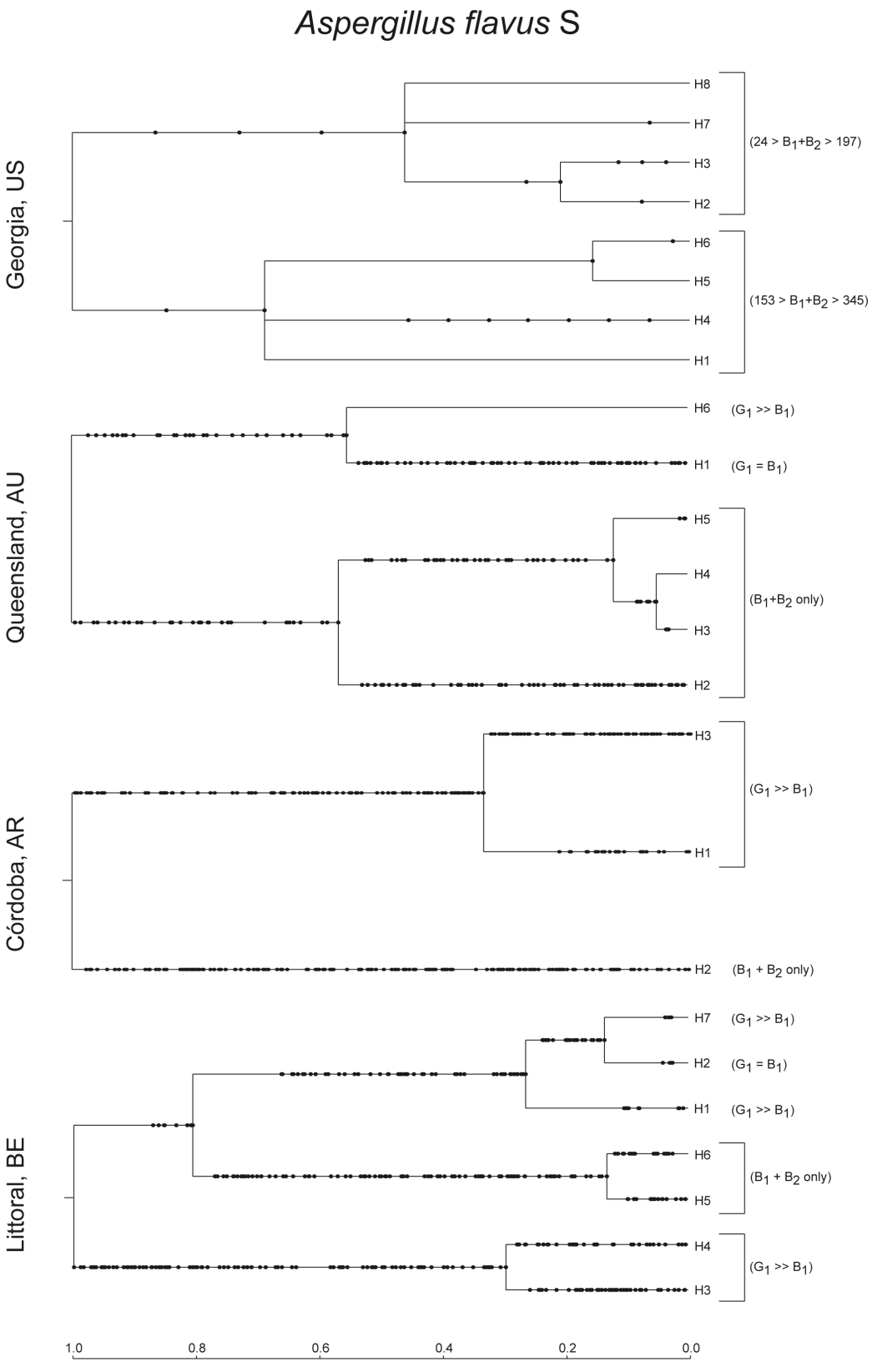

Aspergillus flavus S
Georgia, US
H8
H7
H3
H2
(24 > B1+B2 > 197)
H6
H5
H4
H1
(153 > B1+B2 > 345)
Queensland, AU
H6
(G1 >> B1)
H1
(G1 = B1)
H5
H4
H3
H2
(B1+B2 only)
Córdoba, AR
H3
(G1 >> B1)
H1
H2
(B1 + B2 only)
Littoral, BE
H7
(G1 >> B1)
H2
(G1 = B1)
H1
(G1 >> B1)
H6
H5
(B1 + B2 only)
H4
H3
(G1 >> B1)
1.0
0.8
0.6
0.4
0.2
0.0

Table S2.14 Haplotypes and strain designations for multilocus gene genealogies in Figure S2.4

Haplotype[a]	Georgia, US isolates
H1	IC1154, IC1157, IC1169, IC1174, IC1175, IC1176, IC1178
H2	IC1155
H3	IC1162
H4	IC1160
H5	IC1161, IC1164, IC1168, IC1171
H6	IC1165
H7	IC1163
H8	IC1177
	Queensland, AU isolates
H1	IC720
H2	IC760
H3	IC723, IC779, IC797
H4	IC755, IC768
H5	IC790
H6	IC744
	Córdoba, AR isolates
H1	IC477
H2	IC476, IC479
H3	IC478
	Littoral, BE isolates
H1	IC1112
H2	IC1113, IC1120
H3	IC1150
H4	IC1140
H5	IC1152
H6	IC1153
H7	IC1142

[a] Haplotypes based on combining six genomic loci: *aflE/aflM*, *aflM/aflN*, *aflN/aflG*, *aflG/aflL*, *aflL/aflI*

Chapter 3

Phylogeography and Adaptive Evolution in *Aspergillus* section *Flavi*

Phylogeography and Adaptive Evolution in *Aspergillus* section *Flavi*

Geromy G. Moore[1], Bruce W. Horn[2], Ignazio Carbone[1]

[1] Department of Plant Pathology, North Carolina State University, Raleigh, NC, 27606
[2] National Peanut Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Dawson, GA 39842

ABSTRACT

Aflatoxins produced by several species in *Aspergillus* section *Flavi* are an enormous problem in agriculture and a continuous threat to human health. In Chapter 2 we reported that the amount of precipitation appears to influence the frequency and distribution of aflatoxin and nonaflatoxin producing species more than temperature. To provide further insights into the biology and global population structure of species in section *Flavi* we sampled a single field population from major peanut-growing regions in the U.S., Australia, Argentina, India and Africa. We inferred interspecific gene genealogies for two cluster (*aflM/alfN* and *aflW/aflX*) and four non-cluster regions (*amdS, trpC, mfs*, and *MAT*), and extended phylogenies with six phenotypic categories (geography, species, precipitation, temperature, aflatoxin chemotype profile, and mating type). With the exception of *A. flavus* S populations in Australia and Benin we found no evidence of geographic differentiation. Populations of *A. flavus* S in Australia and Benin were genetically distinct from all other section *Flavi* species in the noncluster regions including the mating type genes, which suggests reduced gene flow and genetic isolation. Patterns of trans-speciation in the two cluster regions indicate that *A. flavus*

S in Benin sharing chemotype profiles with *A. parasiticus* in Australia suggests a potential hybrid origin for these S strains. Mean annual temperature and precipitation were tightly correlated with variation in the *MAT1-1* and *MAT1-2* genes of *A. flavus* S in Australia and Benin as well as with geographic specific lineages of *A. tamarii, A. caelatus* and *A. alliaceus,* which suggests a process of climate-mediated selection potentially acting on different products controlled by mating type genes, thereby driving adaptive evolution. There was no evidence of genetic differentiation based on temperature or precipitation for either *A. flavus* L or *A. parasiticus.*

INTRODUCTION

Aspergillus section *Flavi* contains the most serious mycotoxin-producing fungi in the genus, being named for one of its well publicized toxigenic species, *Aspergillus flavus* Link (Raper and Fennell 1965; Horn 2003). Being a prolific colonizer of many taxonomic groups, *A. flavus* is found around the world and occupies different ecological niches (Klich 2007; Cleveland, Yu et al. 2009). Most of the other species in section *Flavi* are soil-inhabiting saprobes such as *A. parasiticus*, another producer of mycotoxins (Horn 2003; Abbas, Wilkinson et al. 2009). Aflatoxins are the most studied of the mycotoxins in this genus (Horn 2003; Bennett 2010), in fact the term "mycotoxin" was created to describe aflatoxins after the Turkey S disease outbreak in the 1960s (Bennett and Klich 2003). In developing countries the risk to humans by carcinogenic aflatoxins is very high and chronic exposure is estimated to be in the billions (CDC 2010). Strict regulations are being set the world over which means the need for effective control of aflatoxin-producing fungi is crucial

(Klich 2007). Recent discoveries of the sexual states in *A. flavus* and *A. parasiticus* suggest that barriers to genetic recombination, such as heterokaryon incompatibility, are not impassable (Micali and Smith 2005; Ramirez-Prado, Moore et al. 2008; Horn, Moore et al. 2009; Horn, Ramirez-Prado et al. 2009; Horn, Ramirez-Prado et al. 2009) and gene flow could potentially create new strains or species that are better adapted to particular niches (Milgroom 1996). In the case of the two approved biocontrol strains, which are both nonaflatoxigenic *A. flavus*, gene flow may restore toxigenicity (Geiser, Pitt et al. 1998; Moore, Singh et al. 2009). Moreover, gene flow between species may result in hybrids that are super toxin producers, better pathogens, or more fit under adverse and changing environments.

Balancing selection seeks to maintain genetic polymorphisms in an individual and is evidenced when genetic variation increases between alleles or haplotypes within a species (Hedrick 2007). Evidence of balancing selection has been observed in the mating type loci (May, Shaw et al. 1999; Ramirez-Prado, Moore et al. 2008), and balancing selection has been reported to maintain chemotype specific alleles in *A. parasiticus* (Carbone, Jakobek et al. (2007)), and *A. flavus* (Moore, Singh et al. (2009)). In the aflatoxin gene cluster, selection maintains the nonaflatoxigenic phenotype in *A. flavus*, and G_1 or B_1 dominance in *A. parasiticus* (Carbone, Jakobek et al. 2007; Moore, Singh et al. 2009). Because balancing selection could be acting on chemotype or morphological differences (Nei 2007; Moore, Singh et al. 2009), species delimitation using these regions is tenuous and a more holistic approach is warranted (see Figure 1 in Samson and Varga) (2009). For phylogenetic inference of species trees, multiple single-copy neutral loci such as *amdS* and *trpC* should be

targeted, as seen in Figure 1.4 of Chapter 2 (Yelton, Hamer et al. 1984; Michielse, Ram et al. 2004).

Climate change can result in fragmentations and bottlenecks eliminating or reducing the population sizes of endemic species and selecting for more fit taxa (Opdam and Wascher 2004; Ali and Roossinck 2008). Climate has been reported to influence the aflatoxin producing ability of *A. flavus*, alter the numbers of aflatoxigenic fungi in the environment, and change fungal population structure (Cotty and Jaime-Garcia 2007). Given the diverse habitats and differences in mean annual temperature and precipitation in peanut growing regions worldwide (Chapter 2), inferences of genetic isolation should be examined with the possibility of climatic conditions driving genetic differentiation and adaptation. In this chapter we examine the evolutionary history of populations of *Aspergillus* section *Flavi* from geographically isolated regions using a bottom-up micro-evolutionary approach. We tested whether there are significant associations of haplotypes and clades with different phenotypic categories and found that climate is the most important isolating and diversifying force, acting primarily on the products of mating type genes. Further exploration of the role of mating type in adaptive evolution in *Aspergillus* section *Flavi* will allow us to better understand, and possibly improve our biocontrol strategies of these agriculturally important species in different parts of the world.

MATERIALS AND METHODS

Sampling, DNA isolation and target loci

Species in *Aspergillus* section *Flavi* were sampled from geographically isolated fields spanning five continents (Table 3.1). DNA was extracted from freeze-dried mycelia and target loci for phylogenetic inference were amplified and sequenced, as described previously

Table 3.1 Species and total individual counts for each geographic region

	Argentina	Australia	Benin	India	Japan	USA
A. alliaceus	0	9	0	0	0	0
A. caelatus	80	0	0	0	0	31
A. flavus L	80	80	80	80	0	104
A. flavus S	4	80	44	0	0	26
A. nomius	0	0	0	0	0	33
A. oryzae	0	0	0	0	51	1
A. parasiticus	80	80	1	0	0	182
A. sojae	0	0	0	0	1	0
A. tamarii	0	6	80	56	0	35
Total	244	255	205	136	52	412

(Moore et al 2009). DNA sequence variation was assayed in six genomic regions: *aflM/aflN* and *aflW/aflX* in the aflatoxin gene cluster and a non-cluster locus (*mfs*) adjacent to the cluster on chromosome 3; *amdS* and the *MAT* loci on chromosome 6; and *trpC* on chromosome 4. Table 3.2 lists the number of individuals in each species examined per locus.

Table 3.2 Species and total individual counts for each genomic region

	aflM/aflN	*aflW/aflX*	*amdS*	*trpC*	*mfs*	*MAT1-1*	*MAT1-2*
A. alliaceus	4	5	1	1	1	2	7
A. caelatus	n/a	40	40	55	47	56	19
A. flavus L	342	347	352	351	351	192	160
A. flavus S	89	89	89	89	89	39	50
A. nomius	1	19	7	9	0	15	24
A. oryzae	30	48	34	34	35	32	7
A. parasiticus	244	244	244	245	244	209	70
A. sojae	1	1	1	1	1	0	0
A. tamarii	n/a	33	86	66	1	36	61
Total	711	826	854	851	769	581	398

Phylogenetic Inference

Multiple sequence alignments for each locus were collapsed into haplotypes using SNAP Map (Aylor, Price et al 2006). Collapsing was performed with the options of recoding indels and excluding infinite sites violations. Because our variation spans the population species interface collapsing in this fashion allowed us to take full advantage of indels arising with species and exclude only those indels that violate position homology between species. All alignments were rooted with *A. nomius* type strain NRRL 13137, whenever possible, as an outgroup species, except for the *MAT1-2* locus where we used a homothallic strain of *A. alliaceus* as an outgroup species. Both species are considered appropriate outgroups according to Peterson, Varga et al. (2008). For the *mfs* locus we used *A. parasiticus* as the outgroup species. We used PAUP* 4.0 to infer phylogenies (Swofford 1998), and if parsimony searches inferred more than one equally parsimonious tree bootstrap values were calculated for branch length support. Bootstrap consensus trees were based on 1000 replicates. A single most parsimonious tree for each locus was saved in Newick format and imported in GTMiner v.1.25 (Brown, Powell et al. 2008). Haplotypes in each phylogeny were extended with phenotypic information that was also color-coded to facilitate detection of associations. Phenotypes included species, locality, mating type (*MAT1-1*, *MAT1-2*), and mean chemotype concentrations (B_1, B_2, G_1, G_2, G_1/B_1, OMST). Any apparent association between haplotype and phenotype was further examined for evidence of selection.

RESULTS

With the exception of *A. flavus* S isolates in Australia and Benin there was no evidence of geographic effect on population structure (Figures S1 and S2, Supporting Information). Phylogenies inferred for the non-cluster loci for the most part resolve the different species as monophyletic groups (Fig. S2). Shared haplotypes between species from different geographic regions suggests recent descent from a common ancestor prior to geographic isolation, or alternatively, extensive gene flow between localities. In *amdS*, there was strong bootstrap support for the *A. caelatus* clade (H27, H30-H34) and for the *A. tamarii* clade (H2, H29, H35-H41), as well as strong support for these species sharing a recent common ancestry (Fig. S3). Strong monophyly was also observed for *A. nomius* isolates (H4, H54, H55, H1), but there was poor support for other species. For example, haplotypes H12, H20, and H25 (Fig. S3) included a mix of *A. flavus* L and S, and *A. oryzae* isolates from different localities. The *trpC* phylogeny (Fig. S3) is mostly concordant with *amdS* but offers less resolution on a population scale making it difficult to separate population and species-specific variation. For example, *A. parasiticus*, *A. caelatus* and *A. alliaceus* share a recent common ancestor. The *mfs* phylogeny (Fig. S4) shows three major clades: *A. parasiticus* (bottom), *A. caelatus* (middle), and *A. flavus* (top). This is concordant with clades inferred from *amdS* and *trpC.*

Distinct patterns of trans-speciation were observed in phylogenies for the cluster loci, *aflM/aflN* and *aflW/aflX* (Fig. S5). In these phylogenies, species (Fig. S2) are not monophyletic as observed in the non-cluster loci. For example, in Argentina the IB clade includes *A. flavus* L and *A. caelatus* isolates, and in Australia, the IB clade includes *A. flavus* L and *A. alliaceus*. This discordance in *aflM/aflN* (*hypE*) arises from balancing selection

acting on G_1 dominant strains in *A. parasiticus* such that they cluster more closely with *A. flavus* than with other *A. parasiticus* isolates. A similar pattern of trans-speciation is observed in *aflW/aflX*, whereby a nonaflatoxigenic lineage in *A. flavus* is well differentiated from a second *A. flavus* lineage, which includes both aflatoxigenic and nonaflatoxigenic isolates (Fig. S6). We also observed trans-speciation within each mating type gene phylogeny for the same two species, *A. tamarii* and *A. nomius* (Fig. S2).

All phylogenies showed a significant association between haplotype and climate (temperature and precipitation) for *A. flavus* S isolates in Australia and Benin (Figs. S7, S8). Significant associations between climate and other species were observed only in the *MAT1-1* and *MAT1-2* phylogenies (Fig. S9). It appears that balancing selection on the mating type genes is also climate-mediated, such that haplotypes that contain *A. tamarii* and *A. nomius* isolates in *MAT1-1* and *MAT1-2* phylogenies were all sampled in the southern U.S. (Figs. S1, S2)

DISCUSSION

The results from this expanded global sample of species in *Aspergillus* section *Flavi* confirms and strengthens inferences of population processes in the aflatoxin gene cluster of these agriculturally important species. First, we see widespread evidence of balancing selection acting on G_1 and B_1 dominant chemotypes in both the *aflM/aflN* and *aflW/aflX* regions. Previously this was detected only in the *aflM/aflN (hypE)* locus in the Georgia *A. parasiticus* population (Carbone et al. 2007). Similarly, there is evidence of an evolutionarily distinct nonaflatoxigenic lineage in *A. flavus* that was first inferred from the

Georgia population (Chapter 1). It appears that these patterns of trans-speciation for specific chemotype profiles transcends geographic and species boundaries. Whether these chemotype-specific lineages have increased pathogenicity or enhanced fitness for some other trait or ecological condition is unknown and warrants further study.

Climate has been reported to influence toxin production and population size of *A. flavus* S in the U.S. (Bock, Mackey et al. 2004; Cotty and Jaime-Garcia 2007). Although there are no reports of *A. flavus* S strains in the U.S. producing G aflatoxins (S_B), S strains producing both B and G aflatoxins (S_{BG}) are common elsewhere (Cotty and Cardwell 1999). For example, the Benin population we examined included S_{BG} isolates; however no *A. parasiticus* isolates were sampled in Benin. With regard to populations of aflatoxin producing fungi, species that produce both B and G aflatoxins such as *A. parasiticus* and *A. nomius* are seldom found in certain geographies (Cotty and Cardwell 1999). Egel, Cotty et al. (1994) reported that *A. flavus* S_B are closely related to the *A. flavus* L type strain based on molecular variation, suggesting a possible common ancestor between S_B and *A. flavus* L. Moreover, patterns of trans-speciation in the two cluster regions indicate that S_{BG} isolates in Benin share chemotype profiles with *A. parasiticus* in Australia. Collectively, molecular data in cluster and non-cluster loci indicates a potential hybrid origin for S_{BG} strains. In *Aspergillus* section *Flavi*, the possibility of interspecific hybridization as a driving force for adaptive evolution merits further investigation.

Mean annual temperature and precipitation were tightly correlated with variation in the *MAT1-1* and *MAT1-2* genes of *A. flavus* S in Australia and Benin as well as with geographic specific lineages of *A. parasiticus, A. tamarii, A. caelatus* and *A. alliaceus,* which

suggests a process of climate-mediated selection potentially acting on different products controlled by mating type genes, thereby driving adaptive evolution. There was no evidence of differentiation in *MAT* based on temperature or precipitation for *A. flavus* L sampled from different localities. This may indicate increased plasticity of these species to adapt to different environments. For example, in arboreal systems when a species is undergoing adaptation to a new environment it often receives genetic material (through pollen and seeds) from neighboring trees that are better adapted to local conditions, thereby becoming better equipped to thrive in a new niche (Davis and Shaw 2001). The possibility of climate change driving gene flow in these toxin-producing species has significant implications for sustainable biocontrol of aflatoxigenic fungi.

Future work will examine variation in *MAT1-1* and *MAT1-2* for signatures of adaptive molecular evolution. Whether specific replacement substitutions alter the function of mating type genes will be determined in mating experiments.

LITERATURE CITED

Abbas, H. K., J. R. Wilkinson, et al. (2009). "Ecology of *Aspergillus flavus*, regulation of aflatoxin production, and management strategies to reduce aflatoxin contamination of corn." Toxin Reviews **28**(2-3): 142-153.

Ali, A. and M. J. Roossinck (2008). Genetic Bottlenecks. Plant Virus Evolution**:** 123-131.

Aylor, D.L., E. Price, I. Carbone (2006). "SNAP: Combine and Map modules for multilocus population genetic analysis." Bioinformatics 22(11): 1399-1401.

Bennett, J. W. (2010). An Overview of the Genus Aspergillus. Aspergillus : molecular biology and genomics. M. Machida and K. Gomi. Wymondham, Caister Academic**:** 1-17.

Bennett, J. W. and M. Klich (2003). "Mycotoxins." Clinical Microbiology Reviews **16**(3): 497-+.

Bock, C. H., B. Mackey, et al. (2004). "Population dynamics of *Aspergillus flavus* in the air of an intensively cultivated region of south-west Arizona." Plant Pathology **53**(4): 422-433.

Brasier, C. (2000). "Plant pathology: The rise of the hybrid fungi." Nature **405**(6783): 134-135.

Brown, D.E., A.J. Powell, I. Carbone, R.A. Dean (2008). "GT-Miner: a graph-theoretic data miner, viewer, and model processor." Bioinformation 3(5): 235-237.

Carbone, I., J. L. Jakobek, et al. (2007). "Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus*." Mol Ecol **16**(20): 4401-17.

CDC. (2010). "Health Studies: Aflatoxin." Environmental Haxards and Health Effects Retrieved February 27, 2010, from http://www.cdc.gov/nceh/hsb/aflatoxin/.

Cleveland, T. E., J. J. Yu, et al. (2009). "Potential of *Aspergillus flavus* genomics for applications in biotechnology." Trends in Biotechnology **27**(3): 151-157.

Cotty, P. J. and K. F. Cardwell (1999). "Divergence of West African and North American communities of Aspergillus section Flavi." Appl Environ Microbiol **65**(5): 2264-6.

Cotty, P. J. and R. Jaime-Garcia (2007). "Influences of climate on aflatoxin producing fungi and aflatoxin contamination." Int J Food Microbiol **119**(1-2): 109-15.

Davis, M. B. and R. G. Shaw (2001). "Range Shifts and Adaptive Responses to Quaternary Climate Change." Science **292**(5517): 673-679.

Egel, D. S., P. J. Cotty, et al. (1994). "Relationships among isolates of Aspergillus sect. Flavi that vary in aflatoxin production." Phytopathology **84**: 906-912.

Geiser, D. M., J. I. Pitt, et al. (1998). "Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*." Proc Natl Acad Sci U S A **95**(1): 388-93.

Hedrick, P. W. (2007). "Balancing selection." Current Biology **17**(7): R230-R231.

Horn, B. W. (2003). "Ecology and population biology of aflatoxigenic fungi in soil." Journal of Toxicology-Toxin Reviews **22**(2-3): 351-379.

Horn, B. W., G. G. Moore, et al. (2009). "Sexual reproduction in *Aspergillus flavus*." Mycologia **101**(3): 423-9.

Horn, B. W., J. H. Ramirez-Prado, et al. (2009). "Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus parasiticus*." Fungal Genet Biol **46**(2): 169-75.

Horn, B. W., J. H. Ramirez-Prado, et al. (2009). "The sexual state of *Aspergillus parasiticus*." Mycologia **101**(2): 275-80.

Klich, M. A. (2007). "*Aspergillus flavus*: the major producer of aflatoxin." Molecular Plant Pathology **8**(6): 713-722.

May, G., F. Shaw, et al. (1999). "The signature of balancing selection: Fungal mating compatibility gene evolution." Proceedings of the National Academy of Sciences of the United States of America **96**(16): 9172-9177.

Micali, O. C. and M. L. Smith (2005). Biological Concepts of Vegetative Self and Nonself Recognition in Fungi. Evolutionary genetics of fungi. J. Xu. Wymondham, U.K., Horizon Bioscience**:** 63-86.

Michielse, C. B., A. F. J. Ram, et al. (2004). "The *Aspergillus nidulans* amdS gene as a marker for the identification of multicopy T-DNA integration events in Agrobacterium-mediated transformation of Aspergillus awamori." Current Genetics **45**(6): 399-403.

Milgroom, M. G. (1996). "Recombination and the multilocus structure of fungal populations." Annual Review of Phytopathology **34**: 457-477.

Moore, G. G., R. Singh, et al. (2009). "Recombination and lineage-specific gene loss in the aflatoxin gene cluster of *Aspergillus flavus*." Mol Ecol **18**(23): 4870-4887.

Nei, M. (2007). "The new mutation theory of phenotypic evolution." Proc Natl Acad Sci U S A **104**(30): 12235-42.

Opdam, P. and D. Wascher (2004). "Climate change meets habitat fragmentation: linking landscape and biogeographical scale level in research and conservation." Biological Conservation **117**(3): 285-297.

Peterson, S. W., J. Varga, et al. (2008). Phylogeny and subgeneric taxonomy of Aspergillus. Aspergillus in the genomic era. J. Varga and R. A. Samson. Wageningen, Netherlands, Wageningen Academic Publishers**:** 33-56.

Price, E. W. and I. Carbone (2005). "SNAP: workbench management tool for evolutionary population genetic analysis." Bioinformatics **21**(3): 402-4.

Ramirez-Prado, J. H., G. G. Moore, et al. (2008). "Characterization and population analysis of the mating-type genes in *Aspergillus flavus* and *Aspergillus parasiticus*." Fungal Genet Biol **45**(9): 1292-9.

Raper, K. B. and D. I. Fennell (1965). The genus Aspergillus. Baltimore, Williams & Wilkins.

Samson, R. A. and J. Varga (2009). "What is a species in Aspergillus?" Med Mycol **47 Suppl 1**: S13-20.

Swofford, D. L. (1998). PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, Massachusetts, Sinauer Associates.

Yelton, M. M., J. E. Hamer, et al. (1984). "Transformation of *Aspergillus-nidulans* by using a *trpC* plasmid." Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **81**(5): 1470-1474.

Figure S3.1 Phylogenies for six genomic regions extended with geography. The legend illustrates the corresponding color for each locality. Each dot represents an individual in the sample. The open circles represent haplotypes from which individuals radiate.

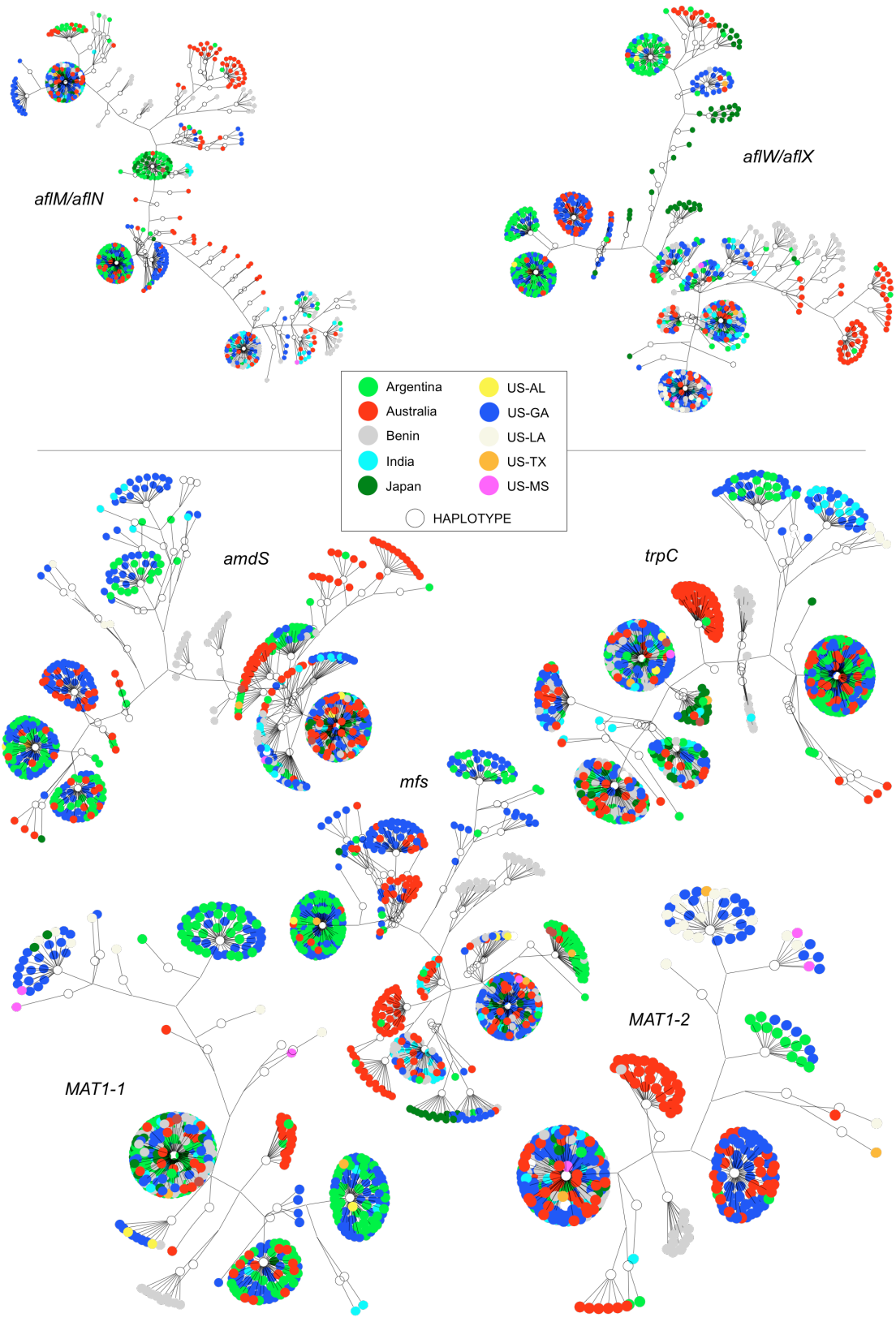
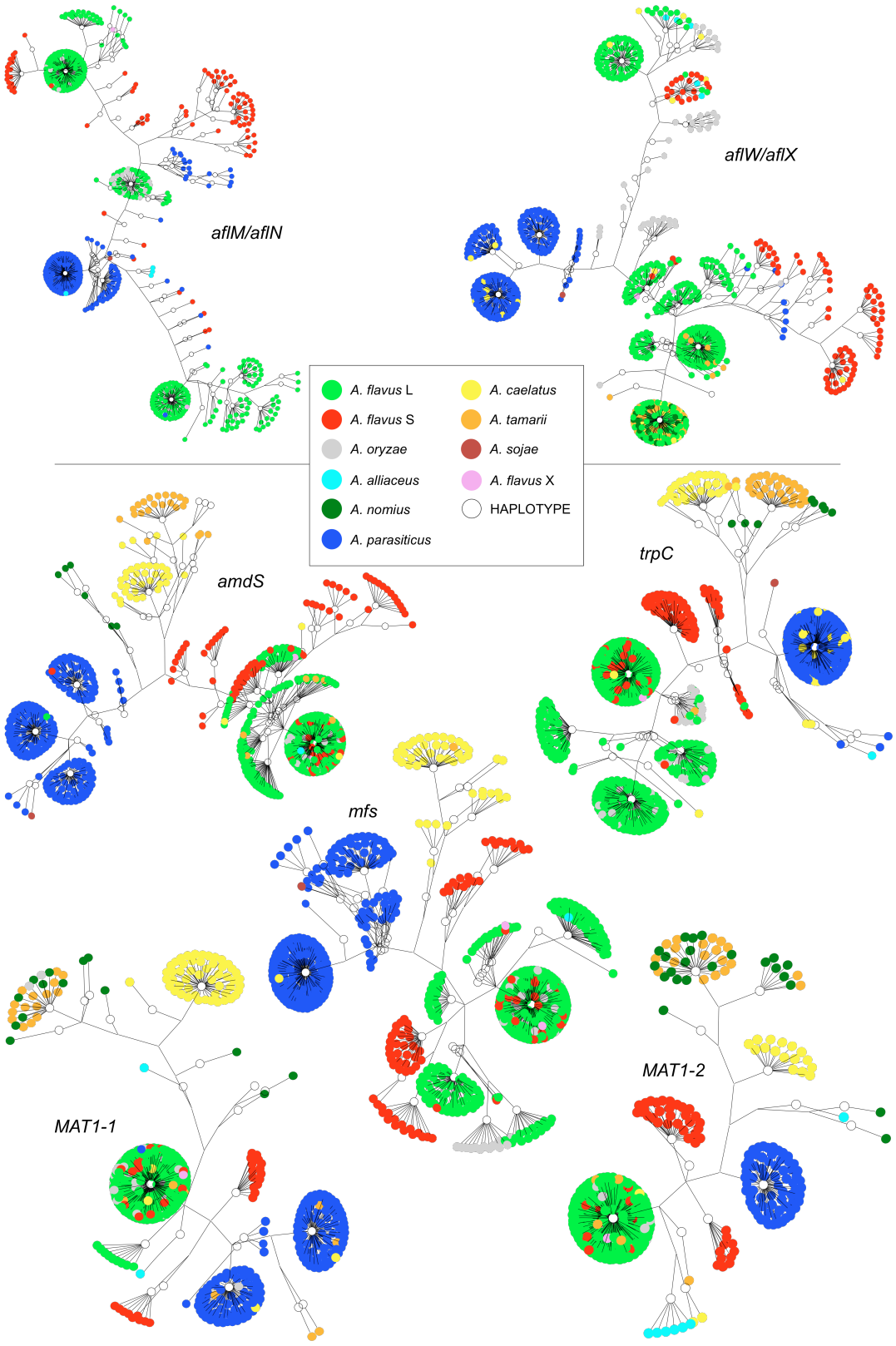

aflM/aflN
aflW/aflX
Argentina
Australia
Benin
India
Japan
US-AL
US-GA
US-LA
US-TX
US-MS
HAPLOTYPE
amdS
trpC
mfs
MAT1-1
MAT1-2

Figure S3.2 Phylogenies for six genomic regions extended with species. The legend illustrates the corresponding color for each species. Each dot represents an individual in the sample. The open circles represent haplotypes from which individuals radiate.


aflM/aflN
aflW/aflX
A. flavus L
A. flavus S
A. oryzae
A. alliaceus
A. nomius
A. parasiticus
A. caelatus
A. tamarii
A. sojae
A. flavus X
HAPLOTYPE
amdS
trpC
mfs
MAT1-1
MAT1-2

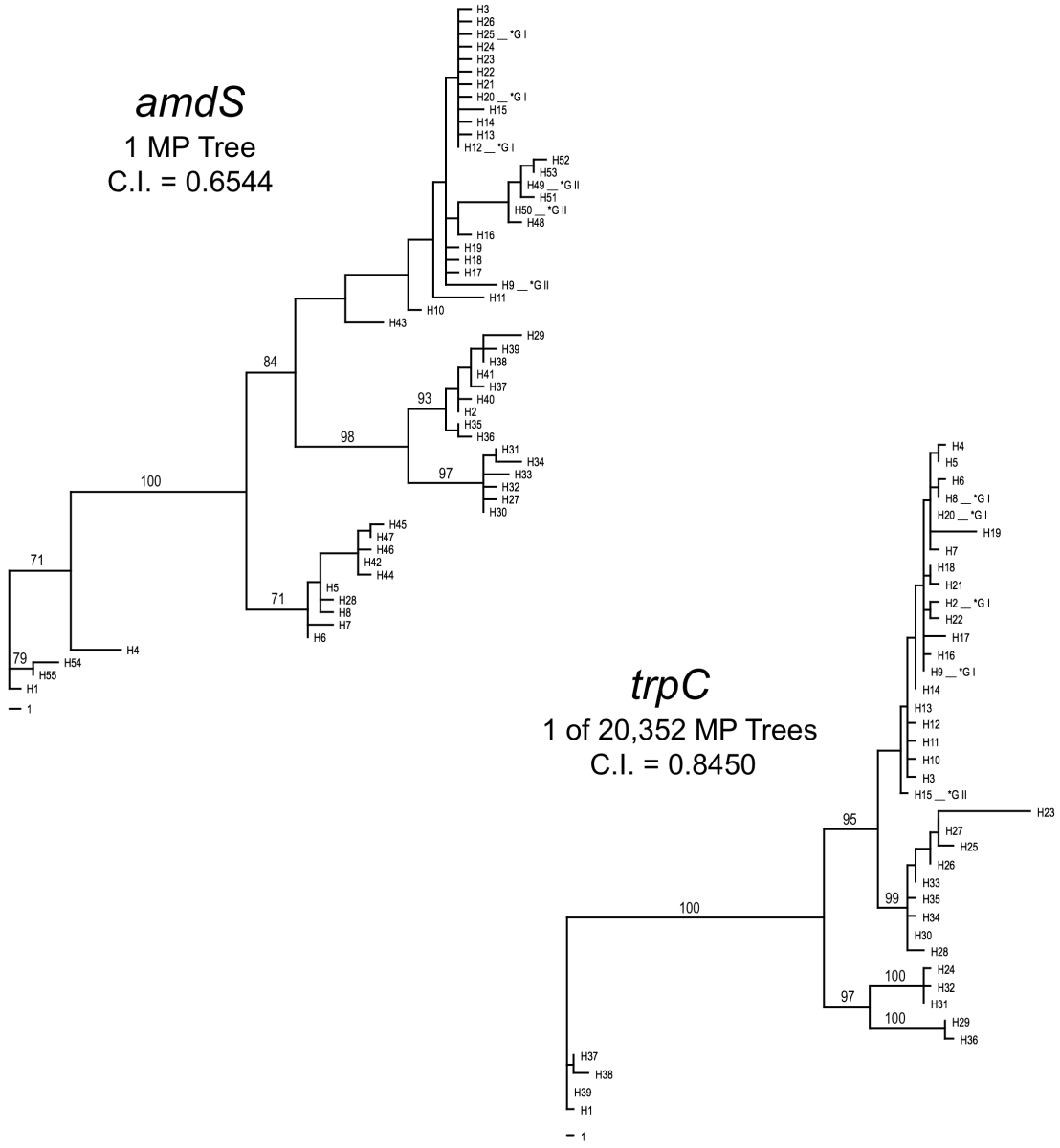


Figure S3.3 Phylogenies for two non-cluster loci: *amdS*, and *trpC*. To the right of each tree is the haplotype number given for each branch. Bootstrap values are found along branches based on 1000 replications. Bootstrap values are shown for branches with >70% support. Additional analyses were performed whereby isolates from Geiser's groups I and II were included to see which isolates share haplotypes with them. The haplotypes are noted with a * G I or * G II. For *amdS*, G I includes *A. flavus* L and S, *A. oryzae*, *A. tamarii*, *A. caelatus*, and *A. alliaceus*. For *trpC*, G I includes *A. flavus* L and S, *A. oryzae*, and *A. caelatus*. For both loci, G II includes only *A. flavus* S strains from Australia and Argentina.

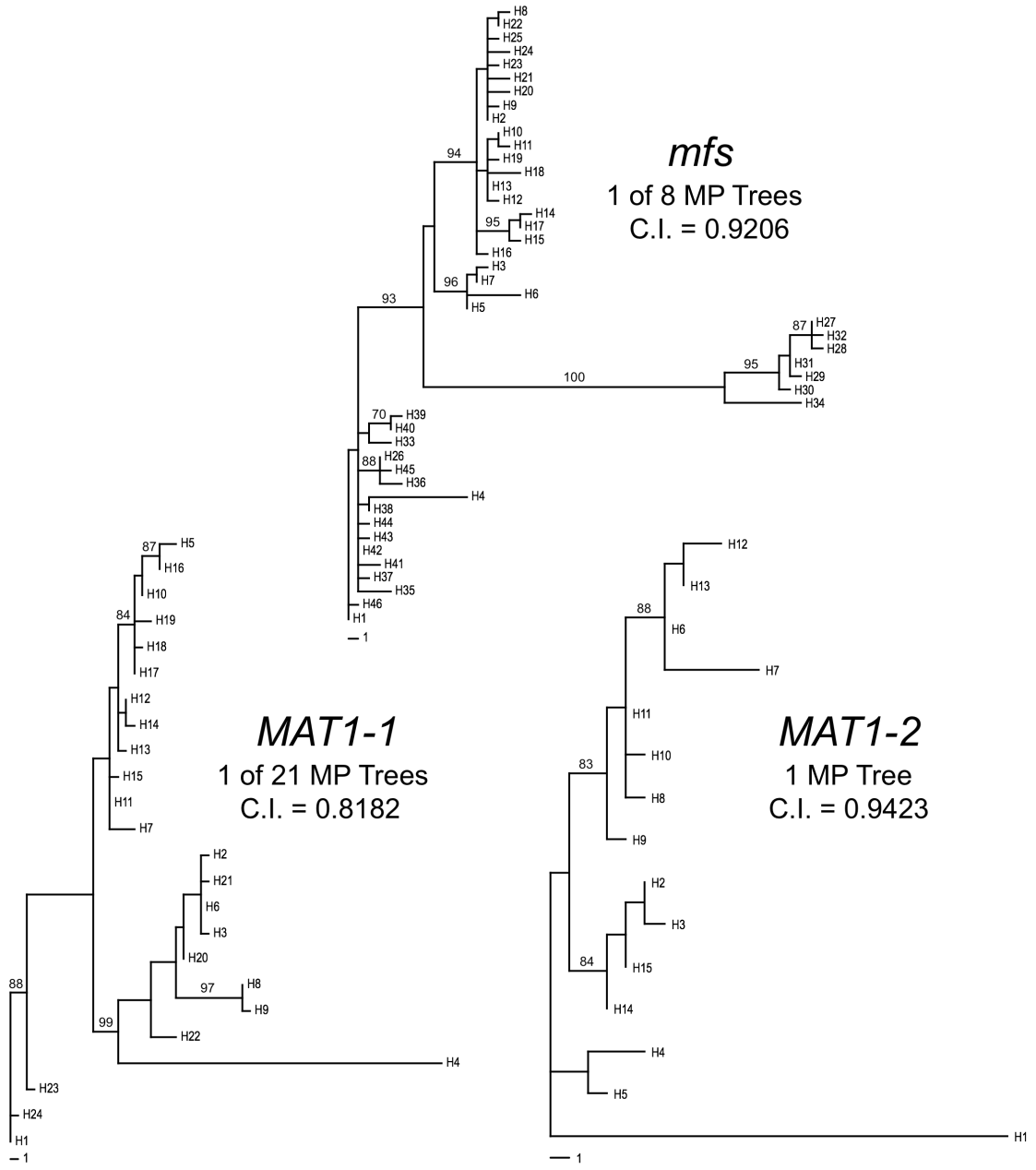


Figure S3.4 Phylogenies for *mfs* and the *MAT* loci: *MAT1-1* and *MAT1-2*. To the right of each tree is the haplotype number given for each branch. Bootstrap values are based on 1000 replications; only branches with >70% support are shown.

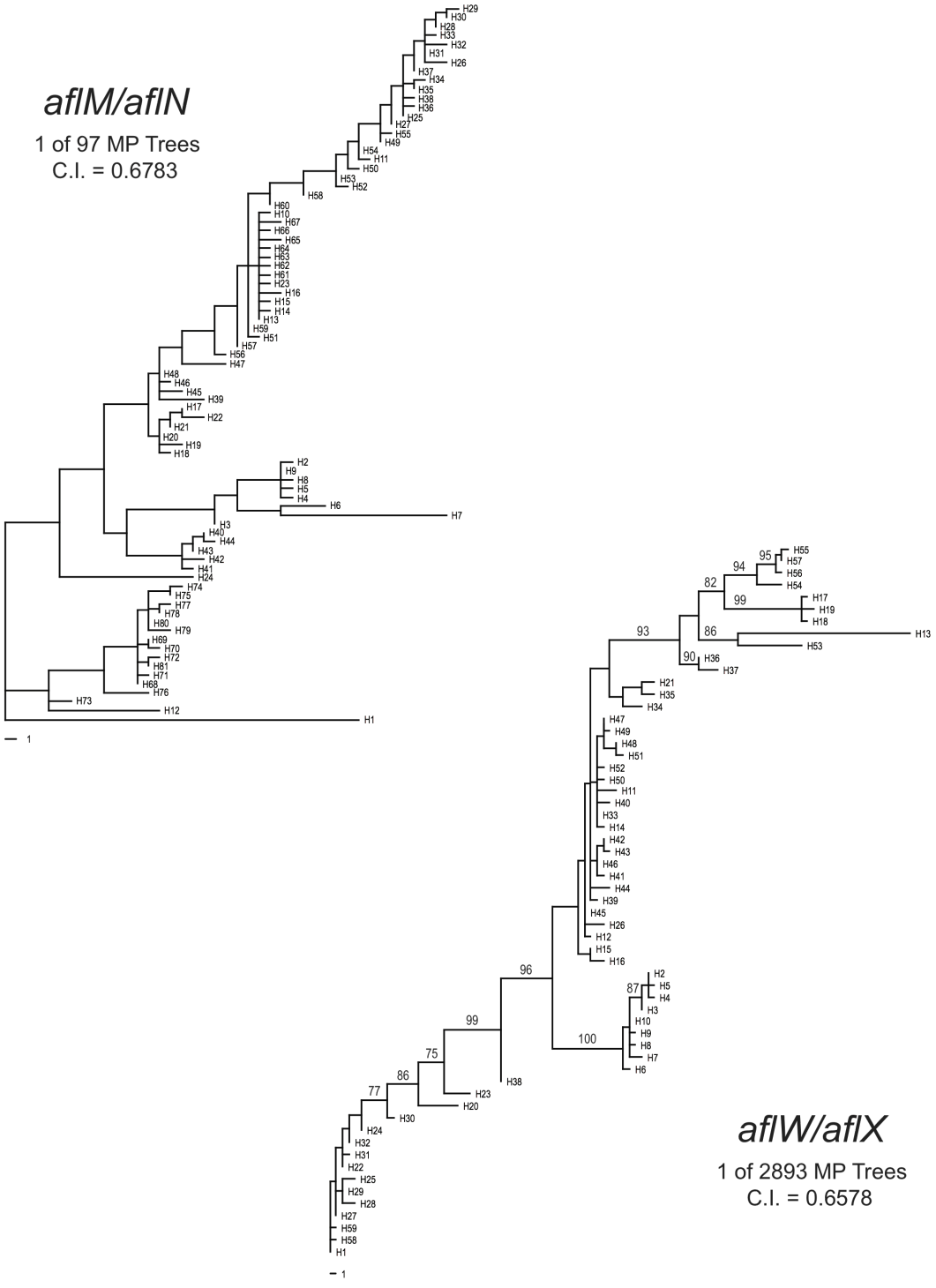


Figure S3.5 Phylogenies for two cluster loci: *aflM/aflN* and *aflW/aflX*. To the right of each tree is the haplotype number given for each branch. Bootstrap values are based on 1000 replications; only branches with >70% support are shown.

Figure S3.6 Phylogenies for six genomic regions extended with chemotype. The legend illustrates the corresponding color for: nonaflatoxigenic (AF-), B only, OMST, $G_1>B_1$, $B_1>G_1$, and $G_1=B_1$. Each dot represents an individual in the sample. The open circles represent haplotypes from which individuals radiate.

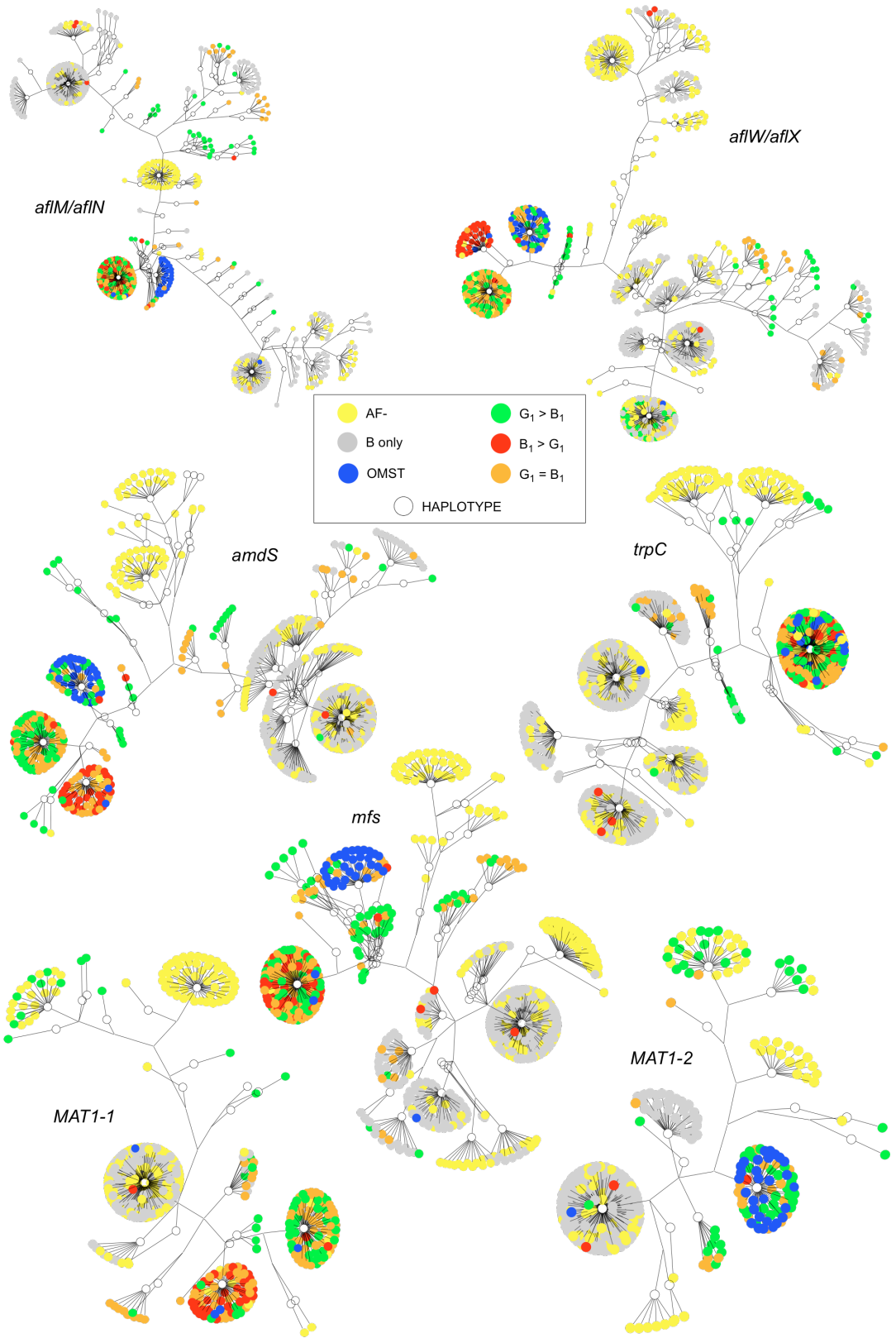

aflM/aflN
aflW/aflX
AF-
B only
OMST
G1 > B1
B1 > G1
G1 = B1
HAPLOTYPE
amdS
trpC
mfs
MAT1-1
MAT1-2

Figure S3.7 Phylogenies for six genomic regions extended with temperature in degrees C. The legend illustrates the corresponding color for each measurement listed. Each dot represents an individual in the sample. The open circles represent haplotypes from which individuals radiate. Temperature estimations were taken from WorldClim map (Fig. 2) in Chapter 2.

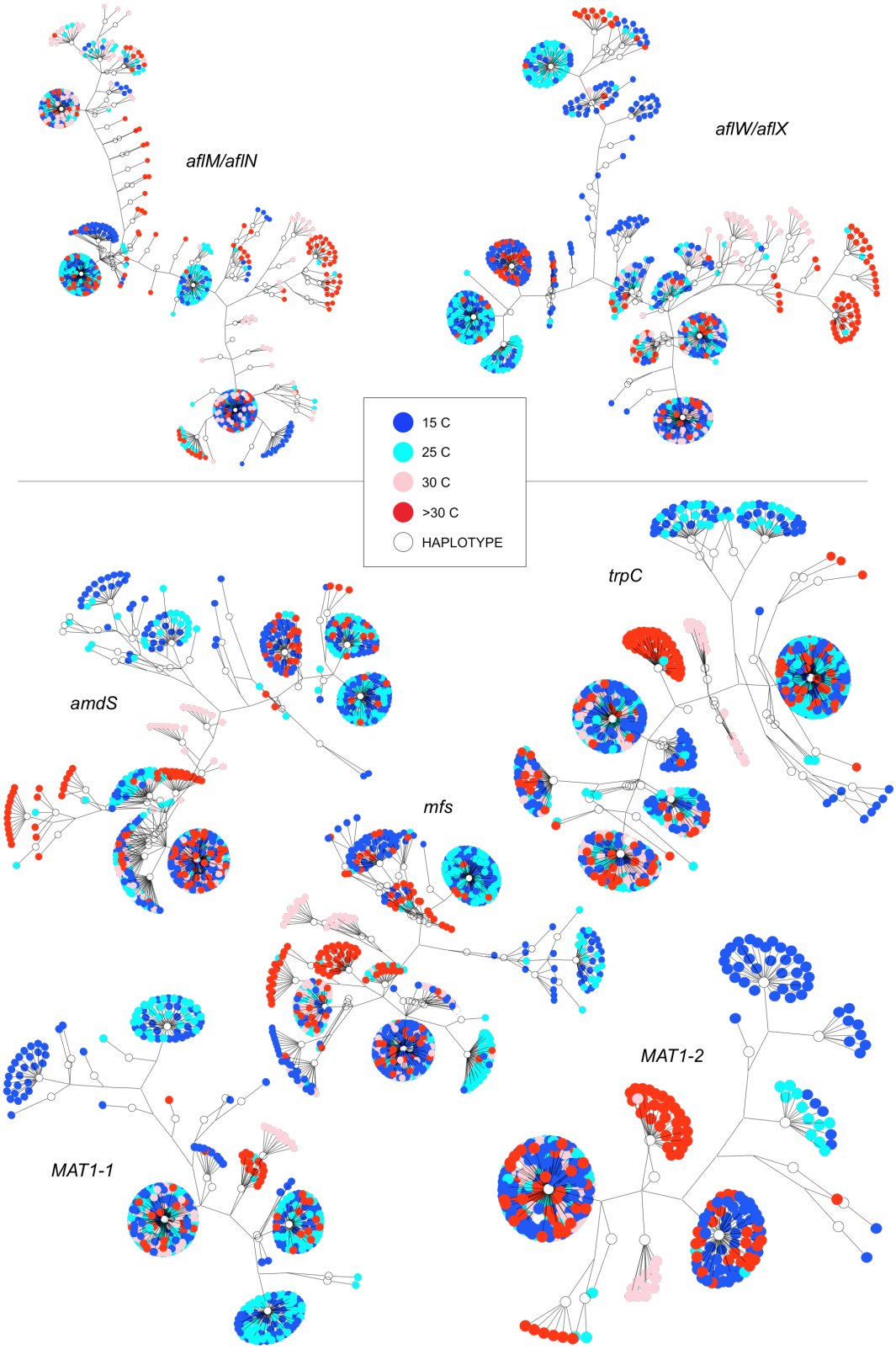

aflM/aflN
aflW/aflX
15 C
25 C
30 C
>30 C
HAPLOTYPE
amdS
trpC
mfs
MAT1-1
MAT1-2

Figure S3.8 Phylogenies for six genomic regions extended with precipitation in mm. The legend illustrates the corresponding color for each measurement listed. Each dot represents an individual in the sample. The open circles represent haplotypes from which individuals radiate. Precipitation estimations were taken from WorldClim map (Fig. 2) in Chapter 2.

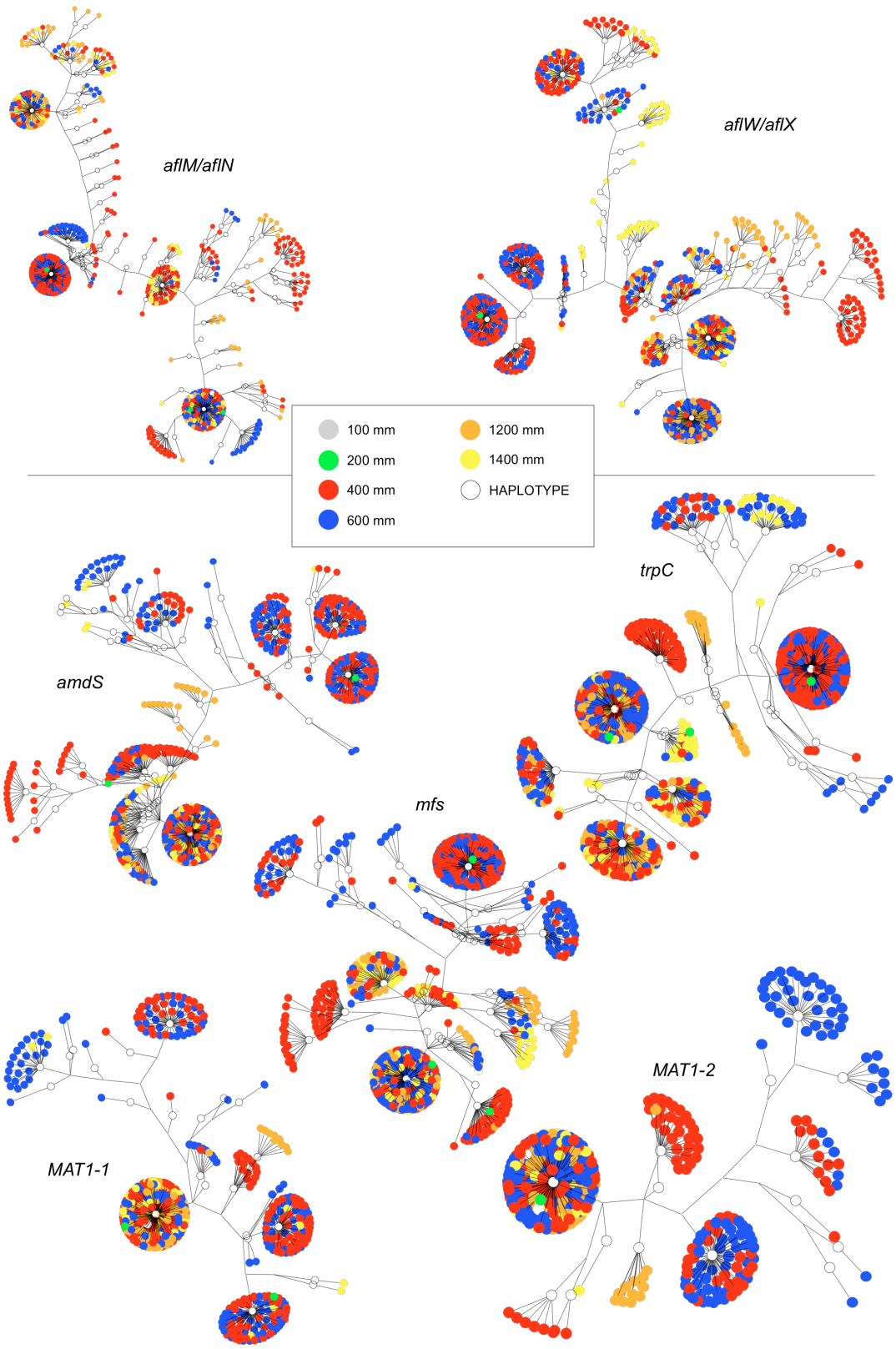
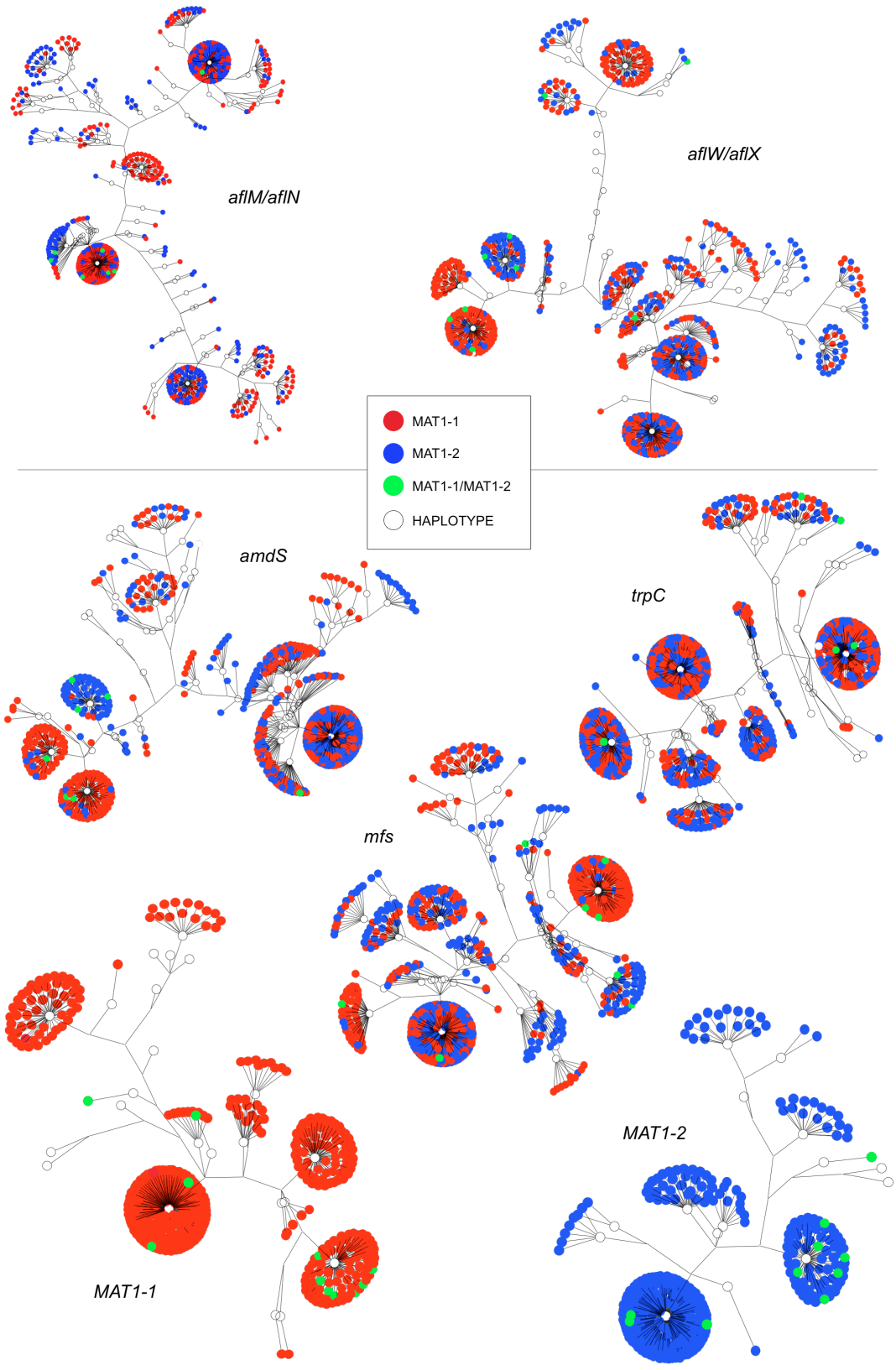

aflM/aflN
aflW/aflX
100 mm
200 mm
400 mm
600 mm
1200 mm
1400 mm
HAPLOTYPE
trpC
amdS
mfs
MAT1-2
MAT1-1

Figure S3.9 Phylogenies for six genomic regions extended with mating type. The legend illustrates the corresponding color for each *MAT* idiomorph. Each dot represents an individual in the sample. The open circles represent haplotypes from which individuals radiate.


aflM/aflN
aflW/aflX
MAT1-1
MAT1-2
MAT1-1/MAT1-2
HAPLOTYPE
amdS
trpC
mfs
MAT1-1
MAT1-2