MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 269–280 doi: 10.1111/j.1365-294X.2009.04467.x

Genetic isolation among sympatric vegetative compatibility groups of the aflatoxin-producing fungus *Aspergillus flavus*

L. C. GRUBISHA and P. J. COTTY

United States Department of Agriculture, Agriculture Research Service, School of Plant Sciences, The University of Arizona, PO Box 210036, Tucson, AZ 85721-0036, USA

Abstract

***Aspergillus flavus*, a fungal pathogen of animals and both wild and economically important plants, is most recognized for producing aflatoxin, a cancer-causing secondary metabolite that contaminates food and animal feed globally. *Aspergillus flavus* has two self/nonself recognition systems, a sexual compatibility system and a vegetative incompatibility system, and both play a role in directing gene flow in populations. *Aspergillus flavus* reproduces clonally in wild and agricultural settings, but whether a cryptic sexual stage exists in nature is currently unknown. We investigated the distribution of genetic variation in 243 samples collected over 4 years from three common vegetative compatibility groups (VCGs) in Arizona and Texas from cotton using 24 microsatellite loci and the mating type locus (*MAT*) to assess population structure and potential gene flow among *A. flavus* VCGs in sympatric populations. All isolates within a VCG had the same mating type with OD02 having *MAT1-2* and both CG136 and MR17 having *MAT1-1*. Our results support the hypothesis that these three *A. flavus* VCGs are genetically isolated. We found high levels of genetic differentiation and no evidence of gene flow between VCGs, including VCGs of opposite mating-type. Our results suggest that these VCGs diverged before domestication of agricultural hosts (>10 000 yr BP).**

Keywords: Ascomycota, asexual, Eurotiales, microsatellite loci, parasexual recombination

Received 19 February 2009; revision received 21 October 2009; accepted 10 November 2009

Introduction

Self/nonself recognition systems have evolved throughout eucaryotic lineages. These recognition systems range in function from insect social organization (Queller & Strassmann 2002) to promoting outbreeding (e.g. self-incompatibility in plants) and immune defence (e.g. major histocompatibility complex in vertebrates; reviewed by Boehm 2006). Filamentous fungi have two self/nonself recognition systems that mediate gene flow in populations: a sexual compatibility system (discussed below) and a vegetative incompatibility system.

Vegetative incompatibility limits hyphal fusion and subsequent gene flow between individuals belonging to different vegetative compatibility groups (VCGs; Leslie 1993). Vegetative incompatibility may have evolved as a way to limit transmission of deleterious viruses and plasmids (Caten 1972; Debets *et al.* 1994; Biella *et al.* 2002), or parasitic nuclei (Hartl *et al.* 1975). Vegetative incompatibility is regulated by multiple unlinked loci (*vic* loci, Leslie 1993; Glass & Kaneko 2003) that range in number from six in *Cryphonectria parasitica* (Cortesi & Milgroom1998), to at least eight in *Aspergillus nidulans* (Anwar *et al.* 1993), nine in *Podospora anserine* (Saupe 2000) and 11 loci in *Neurospora crassa* (Glass *et al.* 2000). Allelic polymorphism at *vic* loci is thought to be maintained by balancing selection, as shown at the *het-c* locus in *N. crassa* (Wu *et al.* 1998). Isolates within the same VCG form stable fusions of vegetative hyphae, whereas hyphal fusion between isolates of different VCGs usually results in programmed cell death (Glass & Dementhon 2006).

Aspergillus flavus Link (Ascomycota, Eurotiales) is a haploid, filamentous fungus that has a vegetative

Correspondence: Peter J. Cotty, Fax: +1 520 626 5944;
E-mail: pjcotty@email.arizona.edu

incompatibility system (Papa 1986), however the number of *vic* loci in *A. flavus* is currently unknown. Like many filamentous fungi, *A. flavus* populations are highly diverse in terms of numbers of VCGs which vary markedly in relative abundance so that a few VCGs comprise most isolates whereas many VCGs contain very small numbers of isolates (Bayman & Cotty 1991; Novas & Cabral 2002; Pildain *et al.* 2004; Barros *et al.* 2006).

Aspergillus flavus is a widely distributed opportunistic fungal pathogen of both animals and plants. Among *Aspergillus* species, *A. flavus* is second only to *Aspergillus fumigatus* in causing human disease (Hedayati *et al.* 2007). However, *A. flavus* is best known as the primary cause of aflatoxin contamination of diverse crops including maize, cottonseed, almond, pistachio and peanut and causes substantial and recurrent worldwide economic loss (Robens & Cardwell 2003; Wu 2004). Aflatoxin is a naturally occurring carcinogenic fungal metabolite produced by some species in *Aspergillus* section *Flavi* (Klich 2007). In regions where food surveillance is inadequate, disease, impaired development, reduced immune function and death are associated with human consumption of aflatoxin-contaminated food (Wild & Turner 2002; Cardwell & Henry 2004; Gong *et al.* 2004; Azziz-Baumgartner *et al.* 2005; Probst *et al.* 2007). Attempts to eliminate aflatoxin contamination through preharvest fungicide treatments have been ineffective, leading to development of biological control strategies (reviewed by Palumbo *et al.* 2008).

One biocontrol method that has been very successful in reducing levels of aflatoxin contamination is use of indigenous nonaflatoxin-producing VCGs of *A. flavus* to competitively exclude aflatoxin-producing VCGs (Cotty *et al.* 2008). Presently, two commercially available nonaflatoxin-producing VCGs have been developed and registered in the United States as biopesticides (Dorner & Lamb 2006; Cotty *et al.* 2008). A concern with this approach is that potential recombination between nonaflatoxin- and aflatoxin-producing VCGs may result in an aflatoxin-producing VCG with increased competitive ability (Geiser *et al.* 1998). Geiser *et al.* (1998) inferred recombination between *A. flavus* isolates, without a priori assignment to specific VCGs, based on conflict among gene tree topologies. In this case, it is not clear if putative recombination events are the result of historic or contemporary gene flow, or to what extent discordance among gene genealogies is attributable to incomplete lineage sorting. They assessed concordance among gene trees using the partition homogeneity test (Farris *et al.* 1995). Since Geiser *et al.* (1998) the reliability of this test has been called into question (Yoder *et al.* 2001; Barker & Lutzoni 2002; Ramirez 2006) and no evidence for recombination among VCGs was found in a study of six *A. flavus* VCGs, including a nonaflatoxigenic VCG used in biocontrol (Ehrlich *et al.* 2007).

Evidence for recombination has been found in many fungal species that are only recognized by an asexual stage (Burt *et al.* 1996; Taylor *et al.* 1999a, b; Grünig *et al.* 2004; Douhan *et al.* 2007). When sexual reproductive structures are absent in nature, it is difficult to distinguish between signatures of recombination due to (i) ancient sexual reproduction, (ii) an unidentified, cryptic sexual stage, or (iii) parasexuality (Burt *et al.* 1996; Debeaupuis *et al.* 1997; Taylor *et al.* 1999a, b). In fungi, recombination may occur in the absence of meiosis by a parasexual cycle characterized by heterokaryon formation, nuclear fusion and mitotic recombination (Pontecorvo 1956). Sexual reproduction in filamentous fungi occurs in self-compatible (homothallic) individuals or between self-incompatible (heterothallic) individuals (Debuchy & Turgeon 2006). Heterothallic individuals need a partner of alternate mating-type to complete the sexual cycle, while homothallic individuals are capable of self-fertilization and may be capable of outcrossing (Pál *et al.* 2007). In the phylum Ascomycota sexual compatibility is controlled by the mating-type locus, *MAT* (Debuchy & Turgeon 2006). Genome sequences of asexual *Aspergillus* species, including *A. flavus*, have allowed identification of an apparently intact and potentially functional mating-type locus that has two alleles (*MAT1-1* and *MAT1-2*) (Galagan *et al.* 2005; Paoletti *et al.* 2005; Ramirez-Prado *et al.* 2008), referred to as idiomorphs because they are not related by structure or common descent (Metzenberg & Glass 1990). Sexual structures were recently produced under laboratory conditions for *A. fumigatus* (O'Gorman *et al.* 2009), *Aspergillus parasiticus* (Horn *et al.* 2009a) and *A. flavus* (Horn *et al.* 2009b) which suggest that for these species ephemeral, cryptic sexual stages may exist in nature.

Questions concerning reproductive mode, dispersal abilities and levels of genetic diversity within VCGs in *A. flavus* populations are crucial for understanding the evolutionary history of *A. flavus*, as a human or plant pathogen, and for management of disease (McDonald & Linde 2002; Hedayati *et al.* 2007). Furthermore, knowledge of the genetic structure of *A. flavus* populations is important for development of successful biocontrol strategies for reducing aflatoxin levels in agriculture. Relevance of the ability to produce a sexual stage of *A. flavus* from isolates of different VCGs and mating-type under laboratory conditions (Horn *et al.* 2009b) to what occurs in nature needs to be examined through population studies (Tibayrenc *et al.* 1991). In this study, we tested the hypothesis that *A. flavus* VCGs are genetically isolated. The objectives were to determine if: (i) there is gene flow among *A. flavus* VCGs and (ii) *A. flavus* VCGs have clonal or recombining population

structures. We genotyped 243 *A. flavus* isolates from three common VCGs in sympatric populations from cottonseed in Arizona and Texas with 24 short tandem repeat loci (STR; microsatellite loci) located throughout the eight chromosomes of *A. flavus* (Grubisha & Cotty 2009).

Materials and methods

Aspergillus flavus *isolate collection, identification and storage*

Samples used in this study were obtained as part of a large-scale project to investigate diversity of *Aspergillus flavus* VCGs from cotton in Arizona and Texas. Cottonseed was sampled from commercial cotton gins from 1999, 2000 and 2001 in south Texas or 2000, 2001 and 2002 in central and western Arizona (Fig. 1, Table 1). This allowed random sampling of relatively large areas (~25–300 hectares per field) representing numerous fields (Table 1). Cottonseed was separated from lint at a gin stand and then subsampled. For Texas sampling, seed was collected after ginning in an overhead hopper and transferred to trucks for transport to the Valley CO-OP Oil Mill in Harlingen, Texas. Upon receipt at the oil mill (July through September) in 1999–2001, cottonseed was subsampled with a vacuum probe. Six cores (3–5 kg) of seed were taken from each truck and immediately split and subdivided resulting in a single 1–1.5 kg sample. In Arizona, cottonseed was ginned into seed piles (up to 100 tons per pile) and sampled with a 50″ corkscrew Trier. Ten probes (1.5 kg each) were taken from each pile that was composed of seed from an individual field. Seed was mixed and split as for samples from Texas. In Texas during 1999, 2000 and 2001, a total of 19, 17 and 12 trucks were sampled, respectively. In Arizona during 2000, 2001 and 2002, a total of 28, 39 and 46 seed piles were sampled, respectively. Seed was stored (up to 3 weeks) dry at room temperature until fungal isolations were complete.

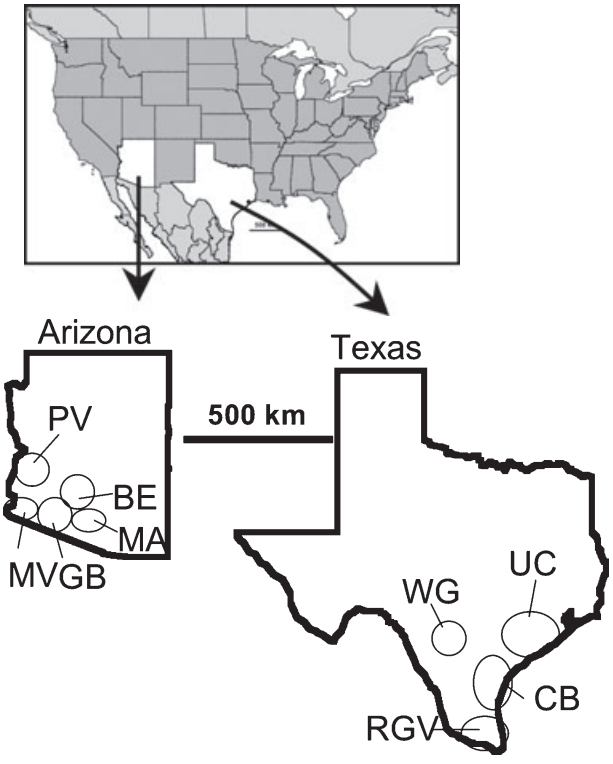

Fig. 1 Collection sites of cottonseed in Texas and Arizona used to isolate *Aspergillus flavus*. Site abbreviations are listed in Table 1.

Cottonseed was washed with 0.01% Tween 80 and *A. flavus* was recovered from washings through dilution plate technique on modified rose Bengal (Cotty 1994). Isolate collections representative of the area for each year were screened for three VCGs (OD02, CG136 and MR17) as previously described (Bayman & Cotty 1991). Briefly, nitrate nonutilizing auxotrophs were selected on agar medium containing chlorate. Isolate auxotrophs were paired with tester auxotrophs (*cnx*$^-$ and *niaD*$^-$) and isolate auxotrophs complemented by a tester strain were classified as belonging to the VCG defined by that tester pair. Plugs of sporulating cultures (3 mm) were submerged in 5 mL of sterile distilled water in 15-mL vials for long-term storage (8°C). *Aspergillus flavus* VCGs are further categorized into L- or S-strains by the size of sclerotia (L-strain sclerotia average diameter >400 μm, S-strain sclerotia <400 μm in diameter), which are haploid, asexual dormant progagules (Cotty 1989). The aflatoxin-producing L-strain VCGs OD02, MR17, and CG136 were selected for this study because they were known to be relatively frequent in the target regions of both Arizona and Texas and, as a result, relatively large sample sizes for each VCG could be obtained (Table 1).

DNA extraction, polymerase chain reaction and genotyping

Mycelial cultures were grown and harvested as described by Elias & Cotty (1996). Isolates were cultured with potato dextrose broth (70 mL in 250-mL flask, 48 h, 160 r.p.m., 32°C). DNA isolation, multiplex polymerase chain reaction (PCR) and genotyping using 24 previously characterized STR loci from eight chromosomes of 243 isolates followed Grubisha & Cotty (2009). Reproducibility of genotyping results was assessed by comparing three to five independent PCR and genotyping runs from at least 10% of the isolates within each VCG for all loci.

Table 1 Collection information for 243 *Aspergillus flavus* isolates used in this study listed by location and number of isolates for vegetative compatibility groups OD02, MR17 and CG136. Sampling occurred over 3 years, 1999–2001 in Texas and 2000–2002 in Arizona

Commercial cotton gin region (no. fields)	GPS	OD02			MR17			CG136			Total
Texas		1999	2000	2001	1999	2000	2001	1999	2000	2001	
Upper Coast, UC (22)	29.05 N 96.20 W	3	0	2	5	2	3	2	5	9	31
Coastal Bend, CB (19)	27.75 N 97.48 W	10	1	2	3	1	2	8	2	2	31
Rio Grande Valley, RGV* (14)	26.22 N 97.65 W	NA	6	7	NA	1	2	NA	4	2	22
Wintergarden, WG (5)	28.95 N 99.63 W	2	2	1	1	0	0	3	0	0	9
Arizona		2000	2001	2002	2000	2001	2002	2000	2001	2002	
Buckeye, BE (30)	33.50 N 112.45 W	5	1	2	6	6	7	8	5	5	45
Gila Bend, GB (14)	33.05 N 113.05 W	1	4	0	5	3	0	7	5	0	25
Maricopa, MA (16)	32.90 N 112.02 W	3	2	0	3	3	5	4	2	1	23
Mohawk Valley, MV (26)	32.76 N 114.07 W	1	2	5	3	9	8	2	5	4	39
Parker Valley, PV (13)	33.74 N 114.46 W	0	1	3	0	2	7	0	3	2	18

GPS, general location for sampled fields; NA, not applicable.
*Not sampled in 1999.

Allelic and genotypic diversity

Allele frequencies were calculated using GENALEX version 6.1 (Peakall & Smouse 2006). Haplotypes (multilocus, haploid genotypes) were identified, and Nei's (1987) genetic diversity corrected for sample size and evenness was calculated using GENODIVE version 2.0b11 (Meirmans & Van Tienderen 2004). Evenness is an indicator of how evenly genotypes are divided over the population.

Population structure

To identify population composition, admixture and genetic structure due to VCG or geography (Arizona or Texas) within a VCG, haplotypes were assigned to K populations using the Bayesian clustering program Structure version 2.2.3 (Pritchard *et al.* 2000). Using the admixture model and default parameters, Markov chain Monte Carlo (MCMC) simulations were run for K = 1–10, with 12 replicates for each K. Each simulation was run with a burn-in length of 100 000 MCMC generations followed by 1 million MCMC iterations. The ad-hoc statistic ΔK, based on the rate of change in the log probability of data between successive runs of K, was calculated following Evanno *et al.* (2005). Principal coordinate analysis (PCO) of the covariance matrix of the Φ_{PT} distance matrix was conducted in GENALEX. To determine how genetic variation was partitioned among VCGs, hierarchical analysis of molecular variance (AMOVA) was estimated using Arlequin version 3.11 (Excoffier *et al.* 2005). Arizona and Texas were considered 'regions' and 'populations within regions' were the sampled gin locations as designated in Table 1. For each year of sampling, Structure, PCO and AMOVA analyses were performed twice (i) all loci were used, and (ii) hypervariable loci were removed even if a locus was hypervariable in only one VCG.

Estimates of divergence time between VCGs

Divergence time between pairs of VCGs was estimated following Zhivotovsky (2001) and used the approach by Munkacsi *et al.* (2008). Briefly, divergence time in terms of generations, T_D, was estimated as: $T_D = (\bar{D}_1/2\bar{w}) - (\bar{V}_0/\bar{w})$, where $\bar{D}_1$ is the average squared difference in allele size between pairs of alleles sampled over all loci (Goldstein *et al.* 1995); $\overline{V}_0$ is the estimate of average variance in the number of repeats over all loci in the ancestral population; and $\bar{w}$ is the average mutation rate across loci. As we do not know the variance in the ancestral population, we estimated conservative earlier and later boundaries by setting $\overline{V}_0 = 0$ for the earlier limit, and assuming $\overline{V}_0$ is equal to the variance in the extant populations for the later limit (Zhivotovsky 2001; Munkacsi *et al.* 2008). Evolutionary error for T_D was estimated following Zhivotovsky (2001). The average mutation rate, $\bar{w}$, was estimated using Equation 1 of Thuillet *et al.* (2005). We assumed one generation per year.

Linkage disequilibrium and detection of potential recombination

Comparisons of linkage disequilibrium (LD) between pairs of polymorphic loci were calculated separately for each VCG and population (all sampling locations combined within either Texas or Arizona), by year sampled using MULTILOCUS version 1.3b (Agapow & Burt 2001). Significance was determined by 1000 permutations

followed by sequential Bonferroni correction for multiple comparisons (Rice 1989). Multilocus LD was assessed using the index of association statistic (I_A; Brown *et al.* 1980; Maynard Smith *et al.* 1993). The I_A standardized for the number of loci was estimated as $\bar{r}_d$ using MULTILOCUS and significance was determined by generating 1000 randomizations and comparing the observed value of $\bar{r}_d$ to that expected under the null hypothesis of panmixis (Agapow & Burt 2001). The most variable loci were removed within each VCG for multilocus LD tests because hypervariability at a repeated locus may be due to selection (Taylor *et al.* 1999b; Verstrepen *et al.* 2005; Levdansky *et al.* 2007) or may bias estimates by supporting random mating (Taylor *et al.* 1999a). The remaining loci (reduced data set, Table 2) were placed into linkage groups. When linkage groups are defined, the randomization process is performed such that all loci within a linkage group are permuted together (Agapow & Burt 2001). A second set of analyses were performed on two to three subsets of the data by randomly selecting one locus with low polymorphism (loci with 2–3 alleles)/linkage group. Reproductive mode was assessed with a phylogenetic approach using the parsimony tree length permutation test (PTLPT; Burt *et al.* 1996). A population with clonal reproduction would produce a single most parsimonious tree that is significantly shorter than the distribution of tree lengths produced by artificially recombined

Table 2 Descriptive statistics and results from statistical tests for reproductive mode in *Aspergillus flavus* vegetative compatibility groups OD02, MR17 and CG136 in Arizona (AZ) and Texas (TX) excluding the most variable loci within each VCG (loci AF10, AF48 and AF64 for OD02, loci AF8, AF10, AF11, AF31 and AF42 for MR17, and locus AF11 for CG136)

VCG population	*n*	*H*	*L*	N_A (range)	*D*	*E*	$\bar{r}_d$ (*P*-value)	PTLPT *P*-value
OD02								
1999 TX	15	9	0.50	2.1 (2–3)	0.89	0.64	0.03 (0.18)	13 (11–15) 0.498
2000 TX AZ	19	16	0.65	2.4 (2–5)	0.98	0.90	0.03 (0.09)	27 (24–31) 0.260
2000 TX	9	8	0.55	1.9 (2–3)	0.97	0.92	0.05 (0.09)	16 (14–19) 0.278
2000 AZ	10	9	0.45	2.2 (2–4)	0.98	0.93	0.02 (0.19)	14 (11–15) 0.944
2001 TX AZ	22	17	0.40	3.1 (2–5)	0.98	0.89	0.01 (0.33)	20 (19–24) 0.063
2001 TX	12	9	0.40	2.3 (2–3)	0.96	0.89	0.01 (0.42)	12 (11–13) 0.676
2001 AZ	10	10	0.30	2.7 (2–4)	1.0	1.0	−0.10 (0.98)	10 (10–12) 0.304
2002 AZ	10	6	0.30	2.0 (2)	0.84	0.69	−0.05 (0.89)	NA
MR17								
1999 TX	9	7	0.16	2.3 (2–3)	0.92	0.77	−0.16 (1.0)	6 (4–6) 1.0
2000 TX AZ	21	11	0.37	2.9 (2–5)	0.88	0.57	0.01 (0.84)	17 (16–17) 1.0
2000 TX	4	4	0.26	2.8 (2–4)	1.0	1.0	−0.33 (1.0)	NA
2000 AZ	17	8	0.32	2.2 (2–3)	8.2	0.56	−0.11 (0.91)	NA
2001 TX AZ	30	16	0.42	2.4 (2–4)	0.87	0.40	−0.02 (0.81)	20 (16–21) 0.984
2001 TX	7	5	0.26	2.2 (2–3)	0.86	0.75	−0.02 (0.97)	6 (6–7) 0.579
2001 AZ	23	12	0.37	2.3 (2–4)	0.87	0.51	−0.02 (0.61)	13 (10–13) 1.0
2002 AZ	27	14	0.47	2.2 (2–4)	0.87	0.43	−0.06 (0.99)	15 (12–16) 0.913
CG136								
1999 TX	13	11	0.57	2.5 (2–5)	0.97	0.90	−0.01 (0.59)	**19 (19–22) 0.034**
2000 TX AZ	32	26	0.57	3.1 (2–7)	0.98	0.79	−0.01 (0.64)	36 (33–40) 0.341
2000 TX	11	9	0.35	2.5 (2–4)	0.96	0.89	−0.02 (0.55)	NA
2000 AZ	21	18	0.52	2.6 (2–6)	0.99	0.91	0.0004 (0.38)	27 (24–30) 0.556
2001 TX AZ	33	21	0.57	2.8 (2–6)	0.96	0.64	−0.02 (0.85)	28 (25–32) 0.385
2001 TX	13	10	0.43	2.2 (2–4)	0.95	0.81	−0.04 (0.82)	16 (15–17) 0.879
2001 AZ	20	14	0.39	2.7 (2–5)	0.96	0.84	−0.03 (0.72)	17 (15–18) 0.879
2002 AZ	12	10	0.26	2.5 (2–4)	0.97	0.90	−0.002 (0.39)	12 (9–12) 1.0

n, number of samples; *H*, number of haplotypes; *L*, proportion of polymorphic loci; N_A, mean number of alleles across all polymorphic loci, range of number of alleles in parentheses; *D*, gene diversity, according to Nei (1987) as $(n/n-1)(1-\sum p_i^2)$ (Meirmans & Van Tienderen 2004); *E*, evenness is an indicator of how evenly genotypes are divided over the population. All genotypes have equal frequencies when $E = 1$ (Meirmans & Van Tienderen 2004)

$\bar{r}_d$, index of association standardized for the number of loci. Significance was determined with 1000 randomizations by comparing the observed value of $\bar{r}_d$ to that expected under the null hypothesis of panmixis (Agapow & Burt 2001).

PTLPT, parsimony tree length permutation test. Significance was determined by comparing the observed tree length to tree lengths of 1000 artificially recombined data sets (Agapow & Burt 2001). NA, only one parsimony informative character was detected; therefore all 1000 trees are the length of the observed tree. Test where $P < 0.05$ is in bold.

data sets whereas a population with a recombining population structure would not (Burt *et al.* 1996; Taylor *et al.* 1999a). The null hypothesis of recombination was tested by comparing the tree length of the observed data to the distribution of tree lengths of 1000 randomizations generated by PAUP version 4.0b10 (Swofford 1999) from an input file prepared in MULTILOCUS. The PTLPT was performed on data sets as described for I_A-tests. Finally, to estimate the frequency of the mating-type genes present in VCGs, portions of *MAT1-1* and *MAT1-2* were multiplex-PCR-amplified from all 243 isolates using primers M1F and M1R (*MAT1-1*) and M2F and M2R (*MAT1-2*) following Ramirez-Prado *et al.* (2008). Because some loci may be preferentially amplified in multiplex-PCR reactions, PCR-amplification of each locus was performed separately on 24 isolates for each VCG.

Results

Genotyping

Final primer combinations used for multiplex-PCR were free of PCR artefacts and only generated a single peak in the expected size range for each locus. In all cases, reproducibility tests confirmed genotype data. Null alleles were not detected except that locus AF18 did not amplify in VCG OD02.

Allelic and genotypic diversity

Allelic and genotypic diversity were analysed with all loci (full set, Table S1) and with the most polymorphic loci removed (reduced set) for each VCG (Table 2). Within VCGs the following loci were highly polymorphic and removed to form the 'reduced set': in VCG OD02 loci AF10, AF48 and AF64; in VCG MR17 loci AF8, AF10, AF11, AF31 and AF42; and in CG136 locus AF11. Allelic polymorphism for each locus had been previously verified to represent length variation in microsatellite repeats by sequencing representative alleles including the shortest and longest alleles within each of the VCGs OD02, MR17 and CG136 using primers designed to PCR amplify >650 bp to sequence the entire locus used for genotyping (Grubisha & Cotty 2009). Overall allelic diversity was low and consistent with clonal reproduction (Table 2). All loci were polymorphic between at least two VCGs, although some loci were fixed for different alleles across the three VCGs. The 27-page table with allele frequencies by year is available upon request from the second author. In each population, isolates with redundant haplotypes were excluded to derive clone-corrected data sets (Table 2). Haplotypes were VCG specific with no evidence of haplotypes occurring in multiple VCGs, even when the most polymorphic loci were removed.

Population structure

Analyses for Structure and PCO were conducted for each year (1999, 2000, 2001 and 2002) separately and yielded essentially identical results therefore results from 2001 will be presented below. Structure, PCO and AMOVA results when hypervariable loci (loci AF8, AF10, AF11, AF31, AF42, AF48 and AF64) were removed (Figs 2 and 3, Table 3, Figs S1 and S2) were identical to the results from analyses with all loci (data not shown). Population composition by geographic location and VCG was determined with the program Structure (Pritchard *et al.* 2000). Although the highest estimate of the log probability only weakly indicated $K = 3$ [mean ln $P(D) = -164.33$] as the most likely number of genetic clusters, the rate of change in log probability among 12 successive runs for each K (Evanno *et al.* 2005), the ad-hoc statistic ΔK, clearly indicated $K = 3$ ($\Delta K = 879.32$) as the most likely number of genetic clusters (Fig. 2). The three clusters correlate with the three VCGs OD02, MR17, and CG136 and there is no evidence of admix-

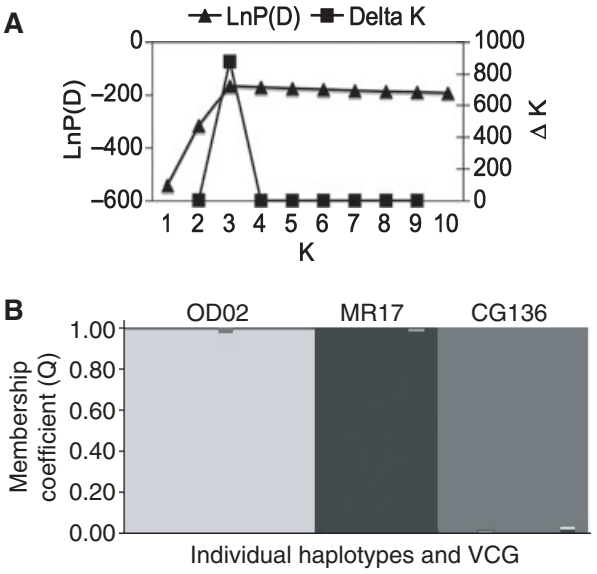

Fig. 2 Results from Structure (Pritchard *et al.* 2000) analysis, based on 12 simulations for each K ($K = 1$–10) for 34 clone-corrected haplotypes of *Aspergillus flavus* VCGs OD02, MR17 and CG136 collected in 2001 from Arizona and Texas. (A) Plot of the posterior probability [ln $P(D)$] averaged across 12 simulations for each K given by Structure output, and calculation of ΔK following Evanno *et al.* (2005). (B) Graphical presentation of the Structure output for one of 12 iterations for $K = 3$. Each haplotype is represented as a vertical bar along the x-axis. The estimated proportion of membership (Q) of a haplotype in each of $K = 3$ inferred genetic clusters is denoted by shade.

ture of haplotypes among these VCGs (Fig. 2). Structure results did not find genetic clusters within VCGs. Independent analyses for each year also indicated $K = 3$ as the most likely number of clusters both from calculation of ΔK and ln $P(D)$ estimates (data not shown). Output from Structure analyses for 1999, 2000 and 2002 is presented in Fig. S1, which show no evidence of admixture among the three VCGs. When individual genotypes were plotted using PCO, the first two axes accounted for 90% of the genetic variation (56% first axis, 34% second axis). The amount of variation explained by the third axis was inconsequential (3.72%). Analyses for each year (1999, 2000 and 2002) yielded similar results and are presented in Fig. S2. Three clusters were detected, each corresponding to a specific VCG (Fig. 3). Given the high level of similarity of the results from Structure and PCO analyses from each year, AMOVA was conducted only on 2001 isolates and over 87% ($P < 0.0001$) of the variation was partitioned among VCGs (Table 3).

Estimated time of divergence between VCGs

Earlier and later boundaries of divergence time were estimated to be between 18 000 and 42 000 (±4500) yr BP for OD02 and CG136, 21 400–46 000 (±9700) yr BP for MR17 and CG136, and 30 000–63 500 (±12 000) yr BP for MR17 and OD02, from samples collected in 2001 when hypervariable loci were removed. The average mutation rate, $\bar{w}$, was estimated to be 2.42×10^{-4}, which is within the rates reported for fungi and other organisms (Estoup & Angers 1998; Munkacsi *et al.* 2008). The divergence dates estimated in this study are based on the assumption of one generation per year.

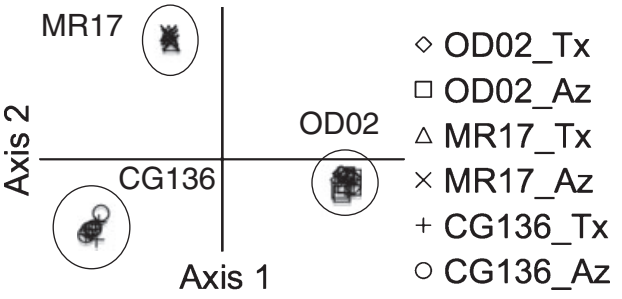

Fig. 3 Results from the principal coordinates analysis indicating three genetic clusters representing the VCGs OD02, MR17 and CG136 in a two-dimensional plot; TX, Texas; AZ, Arizona. The first two axes accounted for 90% of the genetic variation (56% first axis, 34% second axis).

Linkage disequilibrium and detection of potential recombination

Four statistical tests were performed to test for random mating. LD was not detected in all pairwise comparisons between polymorphic loci either before or after sequential Bonferroni correction (Rice 1989) except for three pairs of loci (AF53 with AF43 and AF55; and AF43 with AF55) in VCG MR17 Texas 2000 samples. In this case, significant LD is more likely to be a result of clonal reproduction or small sample size and not linkage as the loci in question are located on different chromosomes.

In I_A-tests of multilocus LD, the null hypothesis of random association of alleles between loci within VCGs could not be rejected for both sets of analyses (i) after hypervariable loci were removed within each VCG and remaining loci put into linkage groups (Table 2), or (ii) when two or three different combinations of one low polymorphic locus/linkage group were analysed (data not shown).

In the first analysis, as described for I_A-tests, the PTLPT test also failed to reject random mating, except for CG136 Texas 1999 (Table 2). However, in four cases only one of the variable characters was parsimony informative and thus the tree lengths resulting from the 1000 randomizations were all equal to the length of the observed tree (Table 2). In the second analyses in VCGs OD02 and CG136, all PTLPT tests failed to reject random mating, except that for OD02 Arizona 2002, CG136 Texas 2000 and CG136 Arizona 2001 only one of the variable characters was parsimony informative. In VCG MR17, the null hypothesis could not be rejected only for Texas 2001, all other tests had only one parsimony informative character. MR17 had the lowest number of loci and smallest sample sizes for this particular test.

Table 3 Analysis of molecular variance results for *Aspergillus flavus* VCGs OD02, MR17 and CG136 in Arizona and Texas from 2001

Source of variation	d.f.	Sum of squares	Covariance component	% of molecular variance	Fixation indices	*P*-value*
Among VCGs	2	210.44	5.62 V_a	87.45	$F_{CT} = 0.88$	<0.0001
Among populations within VCGs	22	18.13	0.02 V_b	0.25	$F_{SC} = 0.02$	0.37
Within populations	31	24.48	0.79 V_c	12.30	$F_{ST} = 0.88$	<0.0001

*Significance was based on 10 100 permutations.

The second analysis for both the I_A and PTLPT was based on the fewest number of loci (as low as only two loci), as not all populations had a single low polymorphic locus in each of eight linkage groups, and the smallest clone-corrected sample size.

All isolates within a VCG had the same mating-type as identified by PCR product size with primers M1F/M1R (*MAT1-1*, 390 bp) and M2F/M2R (*MAT1-2*, 270 bp; Ramirez-Prado *et al.* 2008). All isolates of VCG OD02 contain *MAT1-2* and all isolates of VCGs MR17 and CG136 contain *MAT1-1*. Results from PCR-amplification of each locus separately were identical to the multiplex-PCR results.

Discussion

In this study, we examined the population structure of three sympatric *Aspergillus flavus* VCGs in Arizona and Texas. These results support the hypothesis that *A. flavus* VCGs OD02, CG136 and MR17 are genetically isolated lineages. We found no evidence of gene flow among VCGs, including VCGs of opposite mating-type. While this study was not designed to test whether sexual recombination between VCGs of opposite mating-type resulted in a new VCG, such a scenario would result when the *vic* loci are shuffled during a recombination event between sexually compatible VCGs. However, in subsequent generations, recovery of multilocus parental *vic* haplotypes would be expected and this would result in introgression between the parental VCGs. In addition, if the *vic* loci are not shuffled, then recombinant individuals should retain the VCG of the original isolates and that would have been detected in this study. Both introgression and recombinant haplotypes were not observed in the current work and, as such, these results suggest that VCG OD02 (*MAT1-2*) and CG136 and MR17 (both *MAT1-1*) are not sexually compatible.

Close association between VCG and haplotypes has been found in other plant pathogenic fungi including *Colletotrichum coccodes* (Heilmann *et al.* 2006) and *Verticillium dahliae* (Collado-Romero *et al.* 2008). However, incongruence between VCG and isolate haplotype were found in *Cercospora kikuchii* (Cai & Schneider 2008) and *Sclerotinia sclerotiorum* (Atallah *et al.* 2004). Chulze *et al.* (2000) found that within one multi-isolate VCG, members were no more closely related to each other than to the population as a whole in *Fusarium verticillioides*. Possible explanations for the lack of association between VCG and haplotype may include recombination (discussed above) or mutations at *vic* loci.

Studies of association of VCG and haplotype in *A. flavus* have also found conflicting results. Barros *et al.* (2007) found high genetic diversity among 31 *A. flavus* VCGs from a peanut growing region in Argentina using amplified fragment length polymorphisms. They also found variation within 11 VCGs that had either two or three isolates per VCG. However, haplotypes within these VCGs did not always cluster together in a dendrogram. In contrast, Ehrlich *et al.* (2007) used DNA sequences to infer relationships among six VCGs and found haplotypes from all isolates within a VCG were similar. McAlpin *et al.* (2002) were able to assign 75 *A. flavus* isolates correctly to 44 VCGs using a DNA probe. While nearly identical DNA fingerprints were found in VCGs with multiple isolates, all single isolate VCG DNA fingerprints were unique. Two multiple isolate VCGs had much higher levels of variability with 83% or 87% similarity among isolates within the respective VCG. Bayman & Cotty (1993) found cladistic analysis of DNA polymorphisms from random amplified polymorphic DNAs (RAPDs) was consistent with VCGs. In a study of *Aspergillus fumigatus* and *A. flavus*, Fedorova *et al.* (2009) found gene content was almost identical within VCGs but differed by up to 2% between VCGs using complete genome hybridization.

Other studies have used various molecular methods such as DNA sequences (Geiser *et al.* 1998, 2000), RAPDs (Tran-Dinh *et al.* 1999) and DNA amplification techniques (Baird *et al.* 2006) to examine genetic diversity within *A. flavus* populations and to assess relationships among isolates based on ability to produce toxin, sclerotia size, and geography. While these studies found high levels of genetic diversity within *A. flavus*, isolates were not assigned to a VCG and thus genetic diversity within VCGs could not be assessed, nor could relationships among VCGs.

Estimates of divergence time between pairs of VCGs suggest that these *A. flavus* VCGs diverged well before agriculture. These findings contrast patterns found for fungal plant pathogens with divergence patterns that coincide with host domestication including *Magnaportha oryzae* on rice (Couch *et al.* 2005) and *Mycosphaerella graminicola* on wheat (Stukenbrock *et al.* 2007). Populations of *Ustilago maydis* in North America, Mexico and South America do not predate maize domestication (Munkacsi *et al.* 2008). *Aspergillus flavus* is not restricted to agricultural plants but is a native component of the mycota in warm regions, where it associates with plant debris and fruit of leguminous trees such as mesquite (*Prosopis* spp.) in uncultivated settings (Boyd & Cotty 2001). Desert wild cotton, *Gossypium thurberi*, and wild cotton, a variety of *Gossypium hirsutum*, are potential hosts and sources of inoculum. Both species are native to Texas and Arizona in regions where current cotton agricultural fields are found. Determining the evolutionary history of *A. flavus*, including identifying the centre of

origin, may require sampling populations from a variety of domesticated and wild hosts from diverse geographical locations (Stukenbrock & McDonald 2008).

Reproductive mode was assessed within *A. flavus* VCGs because there was no evidence supporting genetic exchange among VCGs. *Aspergillus flavus* VCGs reproduce clonally through copious production of asexual propagules (conidia and sclerotia) that is reflected in the observed loci with low allelic variation and multiple identical haplotypes within each VCG. A signature of strictly clonal reproduction is association among pairs of loci (LD) due to lack of recombination. However, statistical tests of pairwise LD between loci failed to find LD within VCGs except for MR17 Texas 2000. Likewise, statistical tests of multilocus LD (I_A, PTLPT) failed to reject random mating, with few exceptions. The primary exceptions were the PTLPT analyses when only one low polymorphic locus per linkage group (between two and eight loci total) was analysed that had only one parsimony informative character, especially for VCG MR17. It is possible these VCGs have some type of mixed mating system (clonal and sexual or parasexual) and MR17 may have a stronger clonal component than OD02 or CG136. However, in addition to non-random mating, multilocus LD may also be a result of genetic drift, linkage, population admixture, population expansion or selection (Milgroom 1996). Examination of fine-scale population genetics within VCGs OD02, MR17 and CG136 is currently underway, but is beyond the scope of this study.

It is unclear if the potential recombining population structure observed within these VCGs may be attributed to (i) a cryptic sexual stage, (ii) a parasexual cycle, or (iii) ancient recombination, loss of the sexual cycle and subsequent mutation. All isolates within a VCG had the same mating-type locus, a finding similar to that of other studies of *A. flavus* (Ramirez-Prado *et al.* 2008; Fedorova *et al.* 2009). Meiotic recombination between individuals with alternative mating-type would require that the second mating-type locus within each VCG is present but was undetected or in a very low frequency. Mating-type bias might occur if there is a selective advantage for one mating-type in the population. Brasier (1988) found in the fungus that causes Dutch elm disease, *Ophisotoma novo-ulmi*, that the epidemic front was characterized by few VCGs with one mating-type that was possibly due to frequency-dependent selection exerted by a virus, which was absent in these VCGs. Highly skewed frequencies of opposing mating-type genes in populations have also been correlated with range expansion as in the chestnut blight fungus, *Cryphonectria parasitica* (Milgroom *et al.* 2008), genetic bottleneck associated with introductions of plant pathogens, e.g. *Phytophthora infestans*, or geography, as in the human pathogenic fungus, *Cryptococcus neoformans* (Heitman 2006). VCGs of *A. flavus* with both *MAT* loci are currently not known. However, both *MAT* loci were found in a single VCG in the closely related species, *Aspergillus parasiticus* (Horn *et al.* 2009a). Isolates of alternative mating-type within this VCG failed to form sexual reproductive structures (Horn *et al.* 2009a), although the sample size of one VCG with two isolates with each mating-type gene was too small to be conclusive. In asexual pathogenic fungi, accumulation of mutations in haploid individuals and subsequent mitotic recombination as a result of a parasexual cycle would provide a mechanism for evolution of novel genotypes with specific adaptive advantage. While parasexual recombination has been demonstrated in *A. flavus* under laboratory conditions (Papa 1973), it has not been considered important in natural populations. Lastly, recombining structure may also result from ancestral meiotic recombination and subsequent mutation. Contemporary meiosis might be obviated by either extinction of one mating type locus or loss of sexual compatibility. Future studies should address these hypotheses.

Acknowledgements

This work was supported by Agricultural Research Service, United States Department of Agriculture CRIS project 5347-42000-019-00D. We thank Elizabeth Arnold, Sarah Bergemann, Bruce McDonald and anonymous referees for their comments on the manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Departure of Agriculture.

References

Agapow P-M, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, **1**, 101–102.

Anwar MM, Croft JH, Dales RB (1993) Analysis of heterokaryon incompatibility between heterokaryon-compatibility (h-c) groups R and GL provides evidence that at least eight *het* loci control somatic incompatibility in *Aspergillus nidulans*. *Journal General Microbiology*, **139**, 1599–1603.

Atallah Zk, Larget B, Chen X, Johnson DA (2004) High genetic diversity, phenotypic uniformity, and evidence of outcrossing in *Sclerotinia sclerotiorum* in the Columbia Basin of Washington State. *Phytopathology*, **94**, 737–742.

Azziz-Baumgartner E, Lindblade K, Gieseker K *et al.* (2005) Case-control study of an acute aflatoxicosis outbreak—Kenya 2004. *Environmental Health Perspectives*, **113**, 1779–1783.

Baird RE, Trigiano RN, Windham G *et al.* (2006) Comparison of aflatoxigenic and nonaflatoxigenic isolates of *Aspergillus flavus* using DNA amplification fingerprinting techniques. *Mycopathologia*, **161**, 93–99.

Barker FK, Lutzoni FM (2002) The utility of the incongruence length difference test. *Systematic Biology*, **51**, 625–637.

Barros G, Torres A, Rodriguez M, Chulze S (2006) Genetic diversity within *Aspergillus flavus* strains isolated from the

peanut-cropped soils in Argentina. *Soil Biology & Biochemistry*, **38**, 145–152.

Barros GG, Chiotta ML, Reynoso MM, Torres AM, Chulze SN (2007) Molecular characterization of *Aspergillus* section *Flavi* isolates collected from peanut fields in Argentina using AFLPs. *Journal of Applied Microbiology*, **103**, 900–909.

Bayman P, Cotty PJ (1991) Vegetative compatibility and genetic diversity in the *Aspergillus flavus* population of a single field. *Canadian Journal of Botany*, **69**, 1707–1711.

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

Biella S, Smith ML, Aist JR, Cortesi P, Milgroom MG (2002) Programmed cell death correlates with virus transmission in a filamentous fungus. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **269**, 2269–2276.

Boehm T (2006) Quality control in self/nonself discrimination. *Cell*, **125**, 845–858.

Boyd ML, Cotty PJ (2001) *Aspergillus flavus* and aflatoxin contamination of leguminous trees of the Sonoran Desert in Arizona. *Phytopathology*, **91**, 913–919.

Brasier CM (1988) Rapid changes in genetic structure of epidemic populations of *Ophiostoma ulmi*. *Nature*, **332**, 538–541.

Brown AHD, Feldman MW, Nevo E (1980) Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics*, **96**, 523–536.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proceedings of the National Academy of Sciences, USA*, **93**, 770–773.

Cai G, Schneider RW (2008) Population structure of *Cercospora kikuchii*, the causal agent of Cercospora leaf bight and purple seed stain in soybean. *Phytopathology*, **98**, 823–829.

Cardwell KF, Henry SH (2004) Risk of exposure to and mitigation of effect of aflatoxin on human health: a West African example. *Journal of Toxicology: Toxin Reviews*, **23**, 217–247.

Caten CE (1972) Vegetative incompatibility and cytoplasmic infection in fungi. *Journal of General Microbiology*, **72**, 215–229.

Chulze SN, Ramirez ML, Torres AM, Leslie JF (2000) Genetic variation in *Fusarium* section *Liseola* from no-till maize in Argentina. *Applied & Environmental Microbiology*, **66**, 5312–5315.

Collado-Romero M, Mercado-Blanco J, Olivares-García C, Jiménez-Díaz RM (2008) Phylogenetic analysis of *Verticillium dahliae* vegetative compatibility groups. *Phytopathology*, **98**, 1019–1028.

Cortesi P, Milgroom MG (1998) Genetics of vegetative incompatibility in *Cryphonectria parasitica*. *Applied and Environmental Microbiology*, **64**, 2988–2994.

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

Cotty PJ (1994) Comparison of four media for the isolation of *Aspergillus flavus* group fungi. *Mycopathologia*, **125**, 157–162.

Cotty PJ, Probst C, Jaime-Garcia R (2008) Etiology and management of aflatoxin contamination. In: *Mycotoxins: Detection Methods, Management, Public Health, and Agricultural Trade* (eds Leslie JF, Bandyopadhyay R, Visconti A), pp. 287–299. CAB International, Oxfordshire.

Couch BC, Fudal I, Lebrun M-H *et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, **170**, 613–630.

Debeaupuis JP, Sarfati J, Chazalet V, Latgé JP (1997) Genetic diversity among clinical and environmental isolates of *Aspergillus fumigatus*. *Infection and Immunity*, **65**, 3080–3085.

Debets F, Yang X, Griffiths AJF (1994) Vegetative incompatibility in *Neurospora*: its effect on horizontal transfer of mitochondrial plasmids and senescence in natural populations. *Current Genetics*, **26**, 113–119.

Debuchy R, Turgeon BG (2006) Mating-type structure, evolution and function in euascomycetes. In: *The Mycota I: Growth, Differentiation and Sexualit* (eds Kües U, Kischer R), pp. 293–323. Springer-Verlag, Berlin, Germany.

Dorner JW, Lamb MC (2006) Development and commercial use of afla-Guard®, an aflatoxin biocontrol agent. *Mycotoxin Research*, **22**, 33–38.

Douhan GW, Martin DP, Rizzo DM (2007) Using the putative asexual fungus *Cenococcum geophilum* as a model to test how species concepts influence recombination analyses using sequence data from multiple loci. *Current Genetics*, **52**, 191–201.

Ehrlich KC, Montalbano BG, Cotty PJ (2007) Analysis of single nucleotide polymorphisms in three genes shows evidence for genetic isolation of certain *Aspergillus flavus* vegetative compatibility groups. *FEMS Microbiology Letters*, **268**, 231–236.

Elias KS, Cotty PJ (1996) Incidence and stability of infection by double-stranded RNA genetic elements in *Aspergillus* section *Flavi* and effects on aflatoxigenicity. *Canadian Journal of Botany*, **74**, 716–725.

Estoup A, Angers B (1998) Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In: *Advances in Molecular Ecology* (ed. Carvalho G), pp. 55–86. NATO Press, Amsterdam.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Farris J, Källersjö S, Kluge AG, Bult C (1995) Testing significance of incongruence. *Cladistics*, **10**, 315–319.

Fedorova ND, Harris S, Chen D *et al.* (2009) Using aCGH to study intraspecific genetic variability in two pathogenic molds, *Aspergillus fumigatus* and *Aspergillus flavus*. *Medical Mycology*, **47**, S34–S41.

Galagan JE, Calvo SE, Cuomo C *et al.* (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature*, **438**, 1105–1115.

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

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.

Glass NL, Dementhon K (2006) Non-self recognition and programmed cell death in filamentous fungi. *Current Opinion in Microbiology*, **9**, 553–558.

Glass NL, Kaneko I (2003) Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi. *Eukaryotic Cell*, **2**, 1–8.

Glass NL, Jacobson DJ, Shiu PKT (2000) The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi. *Annual Review of Genetics*, **34**, 165–186.

Goldstein DB, Linares AR, Cavelli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463–471.

Gong YY, Hounsa A, Egal S *et al.* (2004) Post-weaning exposure to aflatoxin results in impaired child growth: a longitudinal study in Benin, West Africa. *Environmental Health Perspectives*, **112**, 1334–1338.

Grubisha LC, Cotty PJ (2009) Twenty-four microsatellite markers for the aflatoxin-producing fungus *Aspergillus flavus*. *Molecular Ecology Resources*, **9**, 264–267.

Grünig CR, McDonald BA, Sieber TN, Rogers SO, Holdenrieder O (2004) Evidence for subdivision of the root-endophyte *Phialocephala fortinii* into cryptic species and recombination within species. *Fungal Genetics and Biology*, **41**, 676–687.

Hartl D, Dempster ER, Brown SW (1975) Adaptive significance of vegetative incompatibility in *Neurospora crassa*. *Genetics*, **81**, 553–569.

Hedayati MT, Pasqualott AC, Warn PA, Bowyer P, Denning DW (2007) *Aspergillus flavus*: human pathogen, allergen and mycotoxin producer. *Microbiology*, **153**, 1677–1692.

Heilmann LG, Nitzan N, Johnson DA *et al.* (2006) Genetic variability in the potato pathogen *Colletotrichum coccodes* as determined by amplified fragment length polymorphism and vegetative compatibility group analyses. *Phytopathology*, **96**, 1097–1107.

Heitman J (2006) Sexual reproduction and the evolution of microbial pathogens. *Current Biology*, **16**, R711–R725.

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

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

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

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

Levdansky E, Romano J, Shadkchan Y *et al.* (2007) Coding tandem repeats generate genetic diversity in *Aspergillus fumigatus* genes. *Eukaryotic Cell*, **6**, 1380–1391.

Maynard Smith J, Smith Noel H, O'Rourke M, Spratt BG (1993) How clonal are bacteria? *Proceedings of the National Academy of Sciences, USA*, **90**, 4384–4388.

McAlpin CE, Wicklow CE, Horn BW (2002) DNA fingerprinting analysis of vegetative compatibility groups in *Aspergillus flavus* from a peanut field in Georgia. *Plant Disease*, **86**, 254–258.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349–379.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Metzenberg RL, Glass NL (1990) Mating type and mating strategies in *Neurospora*. *Bioessays*, **12**, 53–59.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, **34**, 457–477.

Milgroom MG, Sotirovski K, Spica D *et al.* (2008) Clonal population structure of the chestnut blight fungus in expanding ranges in southeastern Europe. *Molecular Ecology*, **17**, 4446–4458.

Munkacsi AB, Stoxen S, May G (2008) *Ustilago maydis* populations tracked maize through domestication and cultivation in the Americas. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **275**, 1037–1046.

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

Novas MV, Cabral D (2002) Association of mycotoxin and sclerotial production with compatibility groups in *Aspergillus flavus* from peanut in Argentina. *Plant Disease*, **86**, 215–219.

O'Gorman CM, Fuller HT, Dyer PS (2009) Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Nature*, **457**, 471–474.

Pál K, van Keipeningen AD, Varga J *et al.* (2007) Sexual and vegetative compatibility genes in the aspergilli. *Studies in Mycology*, **59**, 19–30.

Palumbo JD, O'Keeffe TL, Abbas HK (2008) Microbial interactions with mycotoxigenic fungi and mycotoxins. *Toxin Reviews*, **27**, 261–285.

Paoletti M, Rydholm C, Schwier EU *et al.* (2005) Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Current Biology*, **15**, 1242–1248.

Papa KE (1973) The parasexual cycle in *Aspergillus flavus*. *Mycologia*, **65**, 1201–1205.

Papa KE (1986) Heterokaryon incompatibility in *Aspergillus flavus*. *Mycologia*, **78**, 98–101.

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pildain MB, Vaamonde G, Cabral D (2004) Analysis of population structure of *Aspergillus flavus* from peanut based on vegetative compatibility, geographic origin, mycotoxin and sclerotia production. *International Journal of Food Microbiology*, **93**, 31–40.

Pontecorvo G (1956) The parasexual cycle in fungi. *Annual Review of Microbiology*, **10**, 393–400.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Probst C, Njapau H, Cotty PJ (2007) Outbreak of an acute aflatoxicosis in Kenya in 2004: identification of the causal agent. *Applied and Environmental Microbiology*, **73**, 2762–2764.

Queller DC, Strassmann JE (2002) The many selves of social insects. *Science*, **296**, 311–313.

Ramirez MJ (2006) Further problems with the incongruence length difference test: ''hypercongruence'' effect and multiple comparisons. *Cladistics*, **22**, 289–295.

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.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Robens J, Cardwell FK (2003) The costs of mycotoxin management to the USA: management of aflatoxin in the United States. *Journal of Toxicology*, **22**, 139–152.

Saupe SJ (2000) Molecular genetics of heterkaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews*, **64**, 489–502.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75–100.

Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution*, **24**, 398–411.

Swofford DL (1999) *PAUP*. Phylogenetic Analysis Using Parsimony* (version 4.0). Sinauer Associates Inc., Sunderland, MA.

Taylor JW, Geiser DM, Burt A, Koufopanou V (1999a) The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews*, **12**, 126–146.

Taylor JW, Jacobson DJ, Fisher MC (1999b) The evolution of asexual fungi: reproduction, speciation, and classification. *Annual Review of Phytopathology*, **37**, 197–246.

Thuillet AC, Bataillon T, Poirier S, Santoni S, David JL (2005) Estimation of long-term effective population sizes through the history of durum wheat using microsatelite data. *Genetics*, **169**, 1589–1599.

Tibayrenc M, Kjellberg F, Arnaud J *et al.* (1991) Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proceedings of the National Academy of Sciences, USA*, **88**, 5129–5133.

Tran-Dinh N, Pitt JI, Carter DA (1999) Molecular genotype analysis of natural toxigenic and nontoxigenic isolates of *Aspergillus flavus* and *A. parasiticus*. *Mycological Resources*, **103**, 1485–1490.

Verstrepen KJ, Jansen A, Lewitter F, Fink GR (2005) Intragenic tandem repeats generate functional variation. *Nature Genetics*, **37**, 986–990.

Wild CP, Turner PC (2002) The toxicology of aflatoxins as a basis for public health decisions. *Mutagenesis*, **17**, 471–481.

Wu F (2004) Mycotoxin risk assessment for the purpose of setting international regulatory standards. *Environmental Science and Technology*, **38**, 4049–4055.

Wu J, Saupe SJ, Glass NL (1998) Evidence for balancing selection operating at the *het-c* heterokaryon incompatibility locus in a group of filamentous fungi. *Proceedings of the National Academy of Sciences, USA*, **95**, 12398–12403.

Yoder AD, Irwin JA, Payseur BA (2001) Failure of the ILD to determine data combinability for slow loris phylogeny. *Systematic Biology*, **50**, 408–424.

Zhivotovsky LA (2001) Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. *Molecular Biology and Evolution*, **18**, 700–709.

L.G. is a postdoctoral researcher interested in evolutionary biology and ecology of pathogenic and mycorrhizal fungi. Her current research focuses on clarifying population genetics of *Aspergillus flavus* and in the process developing molecular markers useful for assessing genomic variation. P.C.'s laboratory seeks practical solutions to agricultural problems through improved understanding of the biology, pathology and evolution of Aspergilli.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Fig. S1 Results from Structure (Pritchard *et al.* 2000) analysis, based on 12 simulations for each K (K = 1–10) for clonecorrected haplotypes of *Aspergillus flavus* VCGs OD02, MR17 and CG136. Graphical presentation of the Structure output for one of 12 iterations for K = 3 are shown for (A) 1999, (B) 2000 and (C) 2002. K = 3 was shown to be the most likely number of clusters based on ΔK and ln $P(D)$ estimates (data not shown). Each haplotype is represented as a vertical bar along the x-axis. The estimated proportion of membership (Q) of a haplotype in each of K = 3 inferred genetic clusters is denoted by colour.

Fig. S2 Results from principal coordinates analysis. Samples from each year were analysed separately: (A) 1999 Texas, (B) 2000 Texas and Arizona, and (C) 2002 Arizona. Three genetic clusters are identified for each year of sampling, which correspond to the three VCGs OD02, MR17 and CG136. The per cent of variation explained by the first two axes is 95.60% for 1999, 87.96% for 2000 and 99.74% for 2002.

Table S1 Descriptive statistics for *Aspergillus flavus* vegetative compatibility groups OD02, MR17 and CG136 in Arizona (AZ) and Texas (TX) based on all loci

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.