MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 57–70 doi: 10.1111/j.1365-294X.2011.05364.x

Evidence for genetic differentiation and variable recombination rates among Dutch populations of the opportunistic human pathogen *Aspergillus fumigatus*

CORNÉ H. W. KLAASSEN,*[1] JOHN G. GIBBONS,†[1] NATALIE D. FEDOROVA,‡ JACQUES F. MEIS*§ and ANTONIS ROKAS†

*Department of Medical Microbiology and Infectious Diseases, Canisius Wilhelmina Hospital, Weg Door Jonkerbos 100, 6532 SZ Nijmegen, The Netherlands, †Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA, ‡J Craig Venter Institute, Rockville, MD, USA, §Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, The Netherlands

Abstract

As the frequency of antifungal drug resistance continues to increase, understanding the genetic structure of fungal populations, where resistant isolates have emerged and spread, is of major importance. *Aspergillus fumigatus* is an ubiquitously distributed fungus and the primary causative agent of invasive aspergillosis (IA), a potentially lethal infection in immunocompromised individuals. In the last few years, an increasing number of *A. fumigatus* isolates has evolved resistance to triazoles, the primary drugs for treating IA infections. In most isolates, this multiple-triazole-resistance (MTR) phenotype is caused by mutations in the *cyp51A* gene, which encodes the protein targeted by the triazoles. We investigated the genetic differentiation and reproductive mode of *A. fumigatus* in the Netherlands, the country where the MTR phenotype probably originated, to determine their role in facilitating the emergence and distribution of resistance genotypes. Using 20 genome-wide neutral markers, we genotyped 255 Dutch isolates including 25 isolates with the MTR phenotype. In contrast to previous reports, our results show that Dutch *A. fumigatus* genotypes are genetically differentiated into five distinct populations. Four of the five populations show significant linkage disequilibrium, indicative of an asexual reproductive mode, whereas the fifth population is in linkage equilibrium, indicative of a sexual reproductive mode. Notably, the observed genetic differentiation among Dutch isolates does not correlate with geography, although all isolates with the MTR phenotype nest within a single, predominantly asexual, population. These results suggest that both reproductive mode and genetic differentiation contribute to the structure of Dutch *A. fumigatus* populations and are probably shaping the evolutionary dynamics of drug resistance in this potentially deadly pathogen.

Keywords: fungi, microbial biology, molecular evolution, population genetics, azole resistance, mutations, *Aspergillus fumigatus*

Received 26 January 2011; revision received 28 September 2011; accepted 10 October 2011

Introduction

Invasive fungal infections, which have become a major health issue over the past three decades, are difficult to diagnose and treat. Identifying the genetic structure (i.e. the genetic differentiation and mode of reproduction) of fungal pathogens is essential for the development of better therapeutics against fungal infections. The majority of medically important fungi are haploid organisms that reproduce mostly asexually while occasionally engaging in sexual or parasexual reproduction (Taylor

Correspondence: Corné H. W. Klaassen, Fax: +31 243657516; E-mail: c.klaassen@cwz.nl

[1]These authors contributed equally to this work.

et al. 1999b; Heitman 2006; Sun & Heitman 2011), resulting in populations that are predominantly nonrecombining. This prevalence of 'asexual' reproduction among human pathogenic fungi is surprising, not only because recombination can be very advantageous in stressful environments (Goddard *et al.* 2005; Zeyl *et al.* 2005), such as inside a human host, but also because several of these pathogens have closely related nonpathogenic sexual relatives (Nielsen & Heitman 2007; Butler 2010).

One of the most important fungal pathogens of humans is *Aspergillus fumigatus*, the leading cause of a rapidly progressing, frequently deadly, systemic infection called invasive aspergillosis (IA) (Denning 1998). Although many fungal species, including several human pathogens, show genetic differentiation (Milgroom 1996; Taylor *et al.* 2006; Hittinger *et al.* 2010), several studies have reported that *A. fumigatus* lacks genetic differentiation (Debeaupuis *et al.* 1997; Pringle *et al.* 2005; Rydholm *et al.* 2006); for example, an early restriction fragment length polymorphism analysis of hundreds of clinical and environmental isolates (Debeaupuis *et al.* 1997) and a multilocus sequence-based phylogenetic analysis of clinical and environmental isolates from five continents (Rydholm *et al.* 2006) both found no evidence of genetic differentiation. This absence of genetic differentiation in *A. fumigatus* is consistent with the species' cosmopolitan distribution, its high abundance and ease of aerial dispersal. Alternatively, the observed absence of differentiation could be due to the use of markers that are less informative, or, to sampling design; for example, a different multilocus sequence-based phylogenetic analysis of isolates from around the world identified two globally distributed and genetically differentiated lineages within *A. fumigatus* (Pringle *et al.* 2005), arguing that additional studies are required prior to concluding that *A. fumigatus* is not genetically differentiated.

Early evolutionary analyses also suggested that *A. fumigatus* had lost the ability to reproduce sexually (Geiser *et al.* 1996). However, the presence of intact meiosis-related genes in the *A. fumigatus* genome (Galagan *et al.* 2005; Nierman *et al.* 2005; Rokas & Galagan 2008), the presence of mating-type loci at near-equal frequencies (Paoletti *et al.* 2005) and the demonstration that certain isolates can undergo sexual reproduction in the laboratory (O'Gorman *et al.* 2009) suggest that natural *A. fumigatus* populations are likely to reproduce both asexually and sexually. This inference is consistent with the detection of historical recombination in several population genetic studies of *A. fumigatus* populations (Varga & Toth 2003; Paoletti *et al.* 2005; Pringle *et al.* 2005).

Understanding the genetic structure of *A. fumigatus* is important for human affairs, because recent reports indicate an increase in the frequency of multi-triazole-resistant (MTR) *A. fumigatus* isolates worldwide. Triazole drugs are the primary and most effective therapy against IA infections (Meis & Verweij 2001; Herbrecht *et al.* 2002). In the majority of isolates, MTR resistance is attributed to two mutations in the *cyp51A* gene that encodes 14α-sterol demethylase, the triazole target (Mellado *et al.* 2007; Verweij *et al.* 2007; Snelders *et al.* 2008). The two mutations, known as the TR/L98H allele, involve (i) a duplication in the *cyp51A* promoter and (ii) a non-synonymous point mutation in the *cyp51A* coding region that results in a L98H amino acid change in the protein product. *A. fumigatus* is not transmitted host-to-host, so the rapid spread of the TR/L98H allele raises the possibility that it might have originated outside the clinical environment (Snelders *et al.* 2008; Verweij *et al.* 2009). Notably, genetic analysis of a collection of MTR isolates shows that all isolates with the TR/L98H allele are confined within a single clade and are less variable than nonresistant isolates (Snelders *et al.* 2008), consistent with a single and recent origin.

The recent spread of the TR/L98H allele in *A. fumigatus* presents a unique opportunity to investigate the role of genetic differentiation and reproductive mode in shaping the evolution of drug resistance in this potentially deadly pathogen. Our hypothesis was that sexual reproduction is facilitating the emergence and spread of the TR/L98H allele across the entire *A. fumigatus* population. To test this hypothesis, we studied the genetic structure of *A. fumigatus* in the Netherlands, the country where MTR first emerged. Using 20 neutral markers dispersed across the *A. fumigatus* genome, we genotyped and analysed 255 clinical and environmental Dutch isolates, including 25 MTR isolates with the TR/L98H allele. Contrary to our expectation, we found that Dutch *A. fumigatus* genotypes cluster into five distinct populations that differed in levels of genetic diversity and recombination patterns, which allowed us to infer that four of the five populations were predominantly reproducing asexually. Importantly, the analysis showed that all isolates with the TR/L98H allele nested within one of the four asexually reproducing populations. These results suggest that genetic differentiation and reproductive mode are influencing the dynamics of drug-resistance patterns in natural *A. fumigatus* populations.

Materials and methods

Isolate collection

We analysed 255 clinical and environmental isolates from the Netherlands (Fig. 1A). We included 201 isolates from a nation-wide survey conducted in 2005 (Klaassen *et al.* 2009, 2010), 12 isolates from the clinical

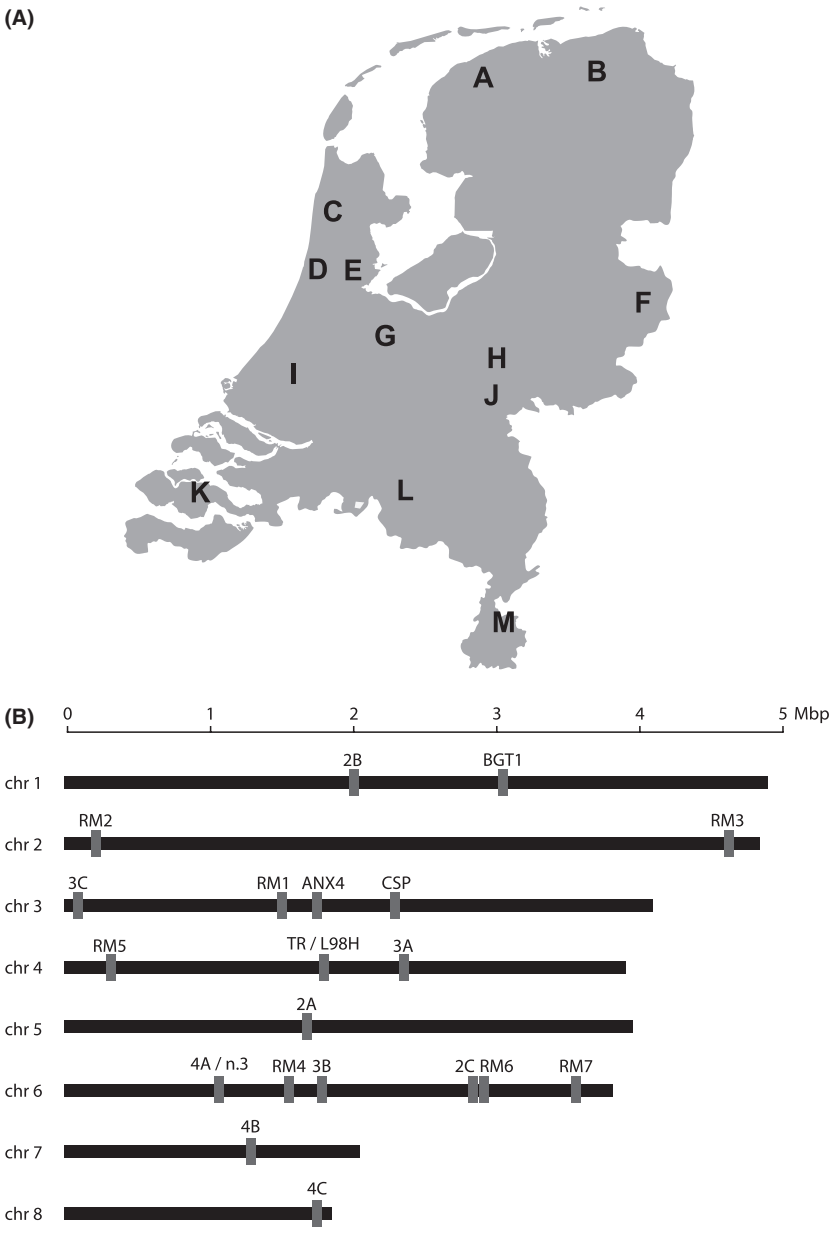

Fig. 1 (A) Sampling location of the 156 genotypes from the Netherlands and (B) the chromosomal location of the TR/L98H (MTR) locus and the 20 markers used in the present study. In panel A, city and number of isolates and genotypes collected per city, respectively (in parentheses), are as follows: A: Leeuwarden (8, 3), B: Groningen (1, 1), C: Alkmaar (10, 9), D: Haarlem (7, 5), E: Amsterdam (19, 10), F: Enschede (15, 9), G: Utrecht (21, 7), H: Arnhem (6, 3), I: Rotterdam (45, 29), J: Nijmegen (102, 68), K: Goes (7, 5), L: Veldhoven (9, 4) and M: Heerlen (5, 3). The chromosome length scale (in million base pairs or Mbp) is shown on top of panel B.

samples collection of the Canisius Wilhelmina Hospital in the city of Nijmegen and 42 isolates from samples of cultivated garden soils from the greater Nijmegen area from a survey conducted in 2008. In all isolates, we determined the presence/absence of the TR/L98H allele using a recently developed real-time PCR screening method (Klaassen *et al.* 2010). Sixteen of the 213 clinical isolates and nine of the 42 environmental isolates contained the TR/L98H allele.

Molecular typing

We genotyped all 255 isolates using 20 markers (nine microsatellite, one indel and 10 sequence/PCR-typing markers), distributed across the eight *A. fumigatus* chromosomes (Fig. 1B, Table S1, Supporting information). Typing of the nine microsatellite markers, which are STR*Af*2A, STR*Af*2B, STR*Af*2C, STR*Af*3A, STR*Af*3B, STR*Af*3C, STR*Af*4A, STR*Af*4B and STR*Af*4C, was performed as described previously (de Valk *et al.* 2005). We also scored the presence/absence of a one-base-pair deletion in the flanking region of microsatellite marker 4A as a separate indel marker; this is the n.3 marker. We sequenced the BGT1 and ANXC4 loci and identified marker alleles using the multilocus sequence-typing scheme described by Bain *et al.* (2007). We sequence-typed the CSP (cell surface protein) locus as previously described (Balajee *et al.* 2007; Klaassen *et al.* 2009). We

PCR-typed the mating-type (MAT) locus using the 'one common and one specific primer' strategy, which yields amplicons of different length for each of the two MAT alleles (Paoletti *et al.* 2005). Using a similar approach, we PCR-typed the alleles of six putative HET (for heterokaryon or vegetative incompatibility) loci, which are thought to regulate self-/non-self-recognition during filamentous growth (Fedorova *et al.* 2008, 2009). For the MAT and HET loci, which together will be further referred to as recombination markers RM1–RM7 (Fig. 1B, Table S1, Supporting information), the A allele corresponds to the allele present in strain Af293 (Nierman *et al.* 2005) and the B allele corresponds to the allele present in strain A1163 (Fedorova *et al.* 2008), whereas the C allele corresponds to a recently identified third allelic variant (C. H. W. Klaassen and J. F. Meis, unpublished observations). We scored alleles that failed to yield a PCR product as nulls and confirmed negative results using a different set of PCR amplification primers.

Neutral evolution analysis

Markers undergoing positive selection are poorly suited for the study of population structure (Avise 2000). Therefore, we tested whether the coding sequences containing our markers are probably undergoing positive selection by estimating the ω ratio of the nonsynonymous substitution rate (d_N) to the synonymous substitution rate (d_S) for each gene using the CODEML module from the PAML software, version 4.4 (Yang 2007). We first identified and aligned orthologues between the transcriptomes of *A. fumigatus*, *Neosartorya fischeri* and *A. clavatus* as described previously (Rokas *et al.* 2007; Fedorova *et al.* 2008; Rokas 2009). To test for positive selection in each coding gene in our marker set, we first evaluated the log likelihood of the null M7 model. Under M7, ω values at different codon positions in a gene follow a beta distribution, where ω is constrained to fall between zero and one. We then measured the difference (Δ*L*) between the log likelihood of the M7 model and that of the alternative M8 model, which, in addition to the zero to one beta distribution for ω values, also allows for a subset of codon sites to have ω values above one (Yang 2006; Scannell *et al.* 2011). We excluded all genes showing rates of synonymous substitutions larger than 2, because in these cases, substitution saturation is likely to reduce the power and reliability of the performed comparisons. Finally, for the BGT1 and ANXC4 markers, for which typing was carried out by sequencing, we also evaluated whether the *A. fumigatus* population departed from neutrality by calculating Tajima's *D* (Tajima 1989), as implemented in the DnaSP software, version 5.10.01 (Librado & Rozas 2009). All tests were performed at $P = 0.01$ significance.

Clonal correction

Inclusion of clonally related genotypes can blur analyses of genetic differentiation, haploid diversity and linkage disequilibrium, because it violates the assumptions of the evolutionary models used in these analyses (Pritchard *et al.* 2000; Jombart *et al.* 2010). Because microsatellite markers have very high mutation rates (Lynch *et al.* 2008), when a large number of microsatellite markers are used, clonally related genotypes might not have identical alleles at all microsatellite loci. Furthermore, for organisms, such as *A. fumigatus*, that are capable of reproducing both sexually and asexually, it is difficult to determine a priori what the best clonal correction threshold might be. To avoid these problems, we generated a series of clonally corrected data sets by eliminating the genotypes of all but one (randomly chosen) isolates with identical alleles for the n.3, ANXC4, BGT1, RM1–7 and CSP markers, and with 0–9 identical microsatellite markers.

Genetic differentiation analysis

We examined the genetic differentiation of Dutch *A. fumigatus* isolates using both model-based and non-model-based approaches for the microsatellite (nine markers), nonmicrosatellite (11 markers) and full (20 markers) data sets, as well as for the series of clonally corrected full (20 marker) data sets.

We examined genetic differentiation using the Bayesian model-based approach implemented in the software STRUCTURE, version 2.3.3 (Pritchard *et al.* 2000). We used 'admixture' and 'allele frequencies are correlated among populations' as our ancestry and frequency models, respectively. We ran 100 replicates of 200 000 Markov Chain Monte Carlo (MCMC) generations for $K = 1$–10, where K = number of populations. In each run, we discarded the first 100 000 generations as burn-in. To identify the optimal K value, we used two different approaches. The first approach makes use of calculating the average log probability (Ln$P(D)$) of each K value (Pritchard *et al.* 2000). Under this approach, the optimal K value is the one showing the highest Ln$P(D)$ score. The second approach is based on the *ad hoc* statistic ΔK, which calculates the rate of change in the log probability of data between successive runs with different K values (Evanno *et al.* 2005). Under this approach, the optimal K value is the one that maximizes ΔK. To evaluate the robustness of population assignment across the 100 STRUCTURE replicates, we compared population assignments for each replicate with the replicate used in this study and calculated average individual membership coefficients using the CLUMPP software, version 1.1.2 (Jakobsson & Rosenberg 2007).

Because natural populations often violate Hardy–Weinberg equilibrium and linkage equilibrium assumptions, inferences drawn solely from model-based methods can be problematic. Therefore, we also analysed our data set using the non-model-based multivariate approach DAPC, as implemented in the ADEGENET software, version 1.3-0 (Jombart 2008; Jombart *et al.* 2010). We predicted the optimal number of clusters (populations) using the *k*-means clustering algorithm, find.clusters', retaining all principal components. We calculated the Bayesian information criterion (BIC) for $K = 1–10$, where K = number of populations. The optimal number of populations was identified as the one for which BIC showed the lowest value and after which BIC increased or decreased by the least amount. We then used DAPC to assign individuals into populations, retaining the number of principal components encompassing 80% of the cumulative variance.

Genetic differentiation by geography analysis

We performed two analyses to test the hypothesis that genotype geography was associated with genetic differentiation. In the first analysis, we grouped genotypes into populations based on their city of origin and based on their STRUCTURE population assignment. We then estimated global and pairwise population differentiation (φ_{PT}) values, a suitable measure of population differentiation analogous to F_{ST} for haploids, and performed AMOVA analysis in each data set using the GENALEX software, version 6.41 (Peakall & Smouse 2006). If genotype geography is significantly associated with genetic differentiation, then we expect to see greater among-population variation, smaller within-population variation and significant population differentiation when genotypes are grouped by the city of origin compared with when genotypes are grouped to the populations they are assigned to by the STRUCTURE software.

In the second analysis, we performed a χ^2goodness-of-fit test between the *observed* number of genotypes from each location and the *expected* number of genotypes from each location because of chance. For each population, we calculated the expected number of genotypes from each location using the equation $N_{LOC}/(N_{TOT} \otimes N_{POP})$, where N_{LOC} is the number of genotypes from location X, N_{POP} is the number of genotypes from population X and N_{TOT} is the total number of genotypes.

Haploid diversity analysis

We calculated Nei's unbiased haploid diversity (uh), a measure that calculates haploid diversity corrected for sample size, independently for the total marker set as well as the microsatellite marker set, for each one of the populations delineated using the STRUCTURE software, using the GENALEX software, version 6.4.1 (Peakall & Smouse 2006). For comparison, we also calculated haploid diversity from two populations of *A. nidulans* using a set of seven microsatellite markers (Hosid *et al.* 2008), as well as from a population of *A. flavus* and a population of *A. parasiticus* using a different set of seven microsatellite markers (Tran-Dinh & Carter 2000). Note that the microsatellite markers used in these studies are different from the markers used in this study.

Linkage disequilibrium analysis

To calculate the degree of association between alleles in our set of 20 markers and to examine whether patterns of recombination are similar across the five populations delineated using the STRUCTURE software, we calculated linkage disequilibrium (LD) using the MULTILOCUS software, version 1.3b (Agapow & Burt 2001). Specifically, we evaluated the index of association (Ia) (Maynard Smith *et al.* 1993), which calculates the distance between all possible locus pairs and compares the variance of distances against results expected if there is no association between loci. Ia was calculated both globally for each population as well as between all locus pairs within each population. We assessed statistical significance by comparing the observed Ia value against the Ia values obtained from 1000 randomized multilocus genotype data matrices, using $P = 0.05$ as the significance value cut-off. The randomization step shuffles the alleles among isolates independently for each locus and assumes that the population is in linkage equilibrium. Because nonrecombining populations are expected to contain higher numbers of clonally related genotypes, we also calculated the number of identical genotypes in a population, or clonal richness, for each population as another indicator of reproductive mode.

Estimation of divergence times of Aspergillus fumigatus *populations*

We estimated pairwise divergence times between *A. fumigatus* populations by adapting the methods developed by Zhivotovsky (2001), which have been previously applied in *A. flavus* (Grubisha & Cotty 2010), for our set of microsatellite markers. We excluded markers 3A and 3C from this analysis, because they exhibited unusually high levels of variation that may deviate from the generalized stepwise mutation model. We calculated divergence time, in generation units, using the equation $T_D = (D_1/2w)–(V_0/w)$; D_1 is the average over loci of the average squared difference of repeat unit copy number between pairs of alleles sampled (one

each from the populations) over five replicates (Goldstein *et al.* 1995), w is the effective mutation rate (Thuillet *et al.* 2005) and V_0 is the average over all loci of the within-population variance in the repeat unit in the ancestral population. We estimated early and late boundaries of divergence by setting $V_0 = 0$ and V_0 = variance in the extant populations, respectively (Zhivotovsky 2001; Munkacsi *et al.* 2008); w was estimated by averaging the average mutation rate for each locus, where mutation rate (μ) = 0.00003 R−0.0001, where R is the repeat unit copy number (Thuillet *et al.* 2005).

Results

Aspergillus fumigatus *markers are evolving neutrally*

Six of our 20 markers reside in noncoding regions and are unlikely to be under selective pressure (Table S1, Supporting information). For the remaining 14 markers that reside within coding genes, we were able to reliably identify orthologues for 10 genes and estimate ω values for seven of them. The remaining three genes had d_S values larger than 2, making neutrality testing unreliable. The genes evaluated were associated with markers 2A, 2B, 3A, ANXC4, BGT1, CSP and RM7. None of the seven genes could reject the null M7 model in favour of the M8 model (2A: $\Delta L = 0$; 2B: $\Delta L = 0.098$; 3A: $\Delta L = 0$; ANXC4: $\Delta L = 0$; BGT1: $\Delta L = 0$; CSP: $\Delta L = 3.512$; and RM7: $\Delta L = 0$; P values for all tests are >0.01), suggesting that positive selection is unlikely to be acting on them. Consistent with these results, examination of Tajima's D statistic for the ANXC4 and BGT1 loci within *A. fumigatus* showed no evidence of positive or balancing selection (ANXC4: $D = -0.194$; and BGT1: $D = -0.522$; P values for all tests are >0.01).

Aspergillus fumigatus *genotypes and clonal correction*

Prior to any clonal correction, the entire collection of 255 isolates yielded 225 different genotypes. Microsatellite markers made the biggest contribution to the observed genotypic diversity. Specifically, 224/225 genotypes were recognized by the microsatellite markers alone, 106/225 by the RM markers, 20/225 by the CSP marker and 10/225 by the ANXC4 and BGT markers combined. In both ANXC4 and BGT1, we identified two new alleles. Since up to now in both genes only four different alleles were recognized, we have provisionally numbered the new alleles as the fifth and sixth allele of each marker.

To remove any additional clonally related genotypes, we generated series of clonally corrected data sets by eliminating the genotypes of all but one (randomly chosen) isolates with identical alleles for the n.3, ANXC4, BGT1, RM1–7 and CSP markers, as well as for 9–0 microsatellite markers (irrespectively of which microsatellite marker(s) were involved). This filter resulted in the elimination of 29–106 genotypes (Fig. S1, Supporting information), with the remaining number of nonclonal genotypes plateauing to ~150 genotypes when clonal correction of 5–0 identical (or 4–9 different) microsatellites was applied. Given these results, we decided to use the 'five identical microsatellites' clonal correction threshold, because it coincides with the reaching of the clonal correction plateau. Using this threshold, we removed 99 genotypes (94 with triazole susceptible alleles and five with the MTR allele), resulting in a final data set of 156 nonclonally related genotypes, including 20 with the MTR allele. Unless otherwise indicated, all subsequent analyses were performed on this data set.

Aspergillus fumigatus *genotypes belong to five distinct populations*

We inferred genetic differentiation for the 156 nonclonally related genotypes using the STRUCTURE model-based approach (Pritchard *et al.* 2000) and the DAPC non-model-based approach (Jombart *et al.* 2010), independently for the microsatellite, the nonmicrosatellite and the full marker data sets. Although analyses with both approaches on the microsatellite and nonmicrosatellite data sets give different answers as to the optimal number of populations, DAPC analysis of the microsatellite data set and STRUCTURE analysis of the nonmicrosatellite data set both estimate that the optimal number of populations is five (Fig. S2, Supporting information). Analyses of the full marker data set with both the STRUCTURE and the DAPC approach support this inference. Specifically, both the Ln$P(D)$ (−4035.01) and ΔK (54.91) approaches in Structure analysis indicate that $K = 5$ (Fig. 2A), and so does the BIC in the DAPC analysis, which reaches its minimum value (393.37) for $K = 5$, as well as displays its smallest increase from $K = 5$ (393.37) to $K = 6$ (393.66) (Fig. 2C).

We also examined whether our genetic differentiation structure inferences differed as we imposed different clonal correction thresholds on our marker data set (Fig. S3, Supporting information). In the 225 genotype data set (generated by elimination of all but one genotypes with identical alleles in all nine microsatellite markers), the optimal number of populations predicted using the Ln$P(D)$ approach was 2. In contrast, the optimal number of populations predicted for all other clonally corrected data sets was either 4 (for the genotype data sets generated by elimination of all but one

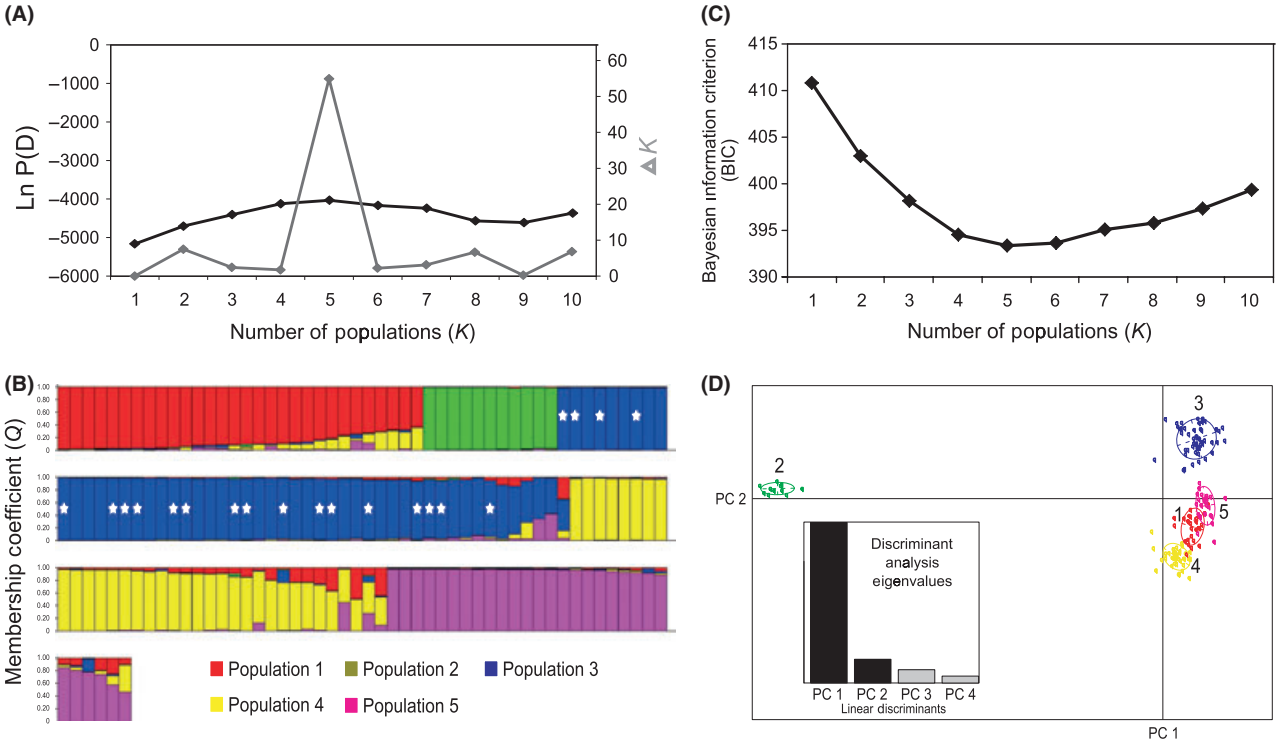

Fig. 2 Both STRUCTURE (panels A and B) and DAPC (panels C and D) analyses of 156 nonclonally related clinical and environmental genotypes identify the existence of five *Aspergillus fumigatus* populations in the Netherlands. (A) STRUCTURE analysis estimates that the optimal predicted number of populations *K* for our set of genotypes is five. This inference is supported by both the average log probability (Ln*P*(*D*)) of each *K* value (black line) and by the *ad hoc* statistic Δ*K* (grey line). (B) The STRUCTURE-based assignment of 156 genotypes into the five *A. fumigatus* populations. Each column on the *X*-axis corresponds to a different genotype. The *Y*-axis represents an individual's membership coefficient to each population. White stars indicate multi-triazole-resistant (MTR) individuals. Structure populations 1–5 are indicated by red, green, blue, yellow and pink colour, respectively. (C) DAPC analysis estimates that the optimal predicted number of populations *K* for our set of genotypes is five. The *Y*-axis corresponds to the Bayesian information criterion (BIC), a goodness-of-fit measurement calculated for each *K*. The lowest BIC value (*K* = 5) indicates the optimal number of populations. (D) DAPC clustering of the five populations using the first two principal components (*Y*-axis and *X*-axis, respectively). The first four eigenvalue components are show in the lower left panel. DAPC populations 1–5 are indicated by red, green, blue, yellow and pink colour, respectively, and are highly similar to STRUCTURE delineated populations.

genotypes with identical alleles in eight and seven microsatellite markers, respectively) or 5 (for the genotype data sets generated by elimination of all but one genotypes with identical alleles in six or fewer microsatellite markers). The very similar numbers of populations inferred by requiring elimination of all but one genotypes with identical alleles in eight or fewer microsatellite markers justify our use of the 'five identical microsatellites' clonal correction threshold. Unless otherwise indicated, all subsequent analyses use the 156 nonclonally related genotype data set generated by the 'five identical microsatellites' clonal correction threshold and its five inferred populations (as assigned by STRUCTURE).

Seventeen of the 20 markers used in this study contribute significantly to the observed genetic differentiation (with the CSP marker showing the strongest association with the different populations; Cramér's *V* statistic = 82.5%), whereas the remaining three markers (RM-1, RM-3 and RM-7) show a random distribution over the five populations (Figs S4–S6, Supporting information).

Our results also suggest that the ancestry of several genotypes in the five populations traces to more than one population; for example, the results from the STRUCTURE approach indicate that several genotypes from population 1 share contributions from population 4 and vice versa (Fig. 2B and Table S2, Supporting information), indicating either that these two populations share a more recent common ancestor and/or provide evidence of recent admixture between them. Furthermore, the results from the DAPC approach point to a clear separation of population 2, and to a lesser extent of population 3, from all other populations; in contrast, population 1 genotypes show some overlap with genotypes from population 4 and population 5 genotypes (Fig. 2D). Interestingly, we found that all 20 genotypes

containing the MTR allele nested within population 3 (Fig. 2B, Table S2, Supporting information), a result that was supported by both approaches.

Finally, we note that the assignment of individual genotypes into the five populations is highly concordant across STRUCTURE replicates, as well as between the DAPC and STRUCTURE analyses. Using CLUMPP (Jakobsson & Rosenberg 2007), we calculated average individual membership coefficients for the 100 STRUCTURE replicates and found that population assignments are identical across runs and membership coefficients nearly identical to the replicate used in this study (Fig. S7; Table S3, Supporting information). Between the DAPC and STRUCTURE analyses, only 8/156 genotypes, all of which have considerable STRUCTURE membership coefficients to more than one population, are assigned to different populations when analysed using the two different approaches (Table S2, Supporting information). For all subsequent analyses, we used the assignment of individual genotypes into the five populations from the STRUCTURE approach.

Aspergillus fumigatus *genotype geography is not associated with genetic differentiation*

Two different analyses failed to provide any evidence in support of the hypothesis that the geographical origin of genotypes was associated with genetic differentiation (as inferred by STRUCTURE). Specifically, we do not identify any global population differentiation when genotypes are grouped into populations by their city of origin (φ_{PT} = 0.004, P = 0.39). Similarly, we find only 4/66 cases of significant population differentiation when we calculate φ_{PT} values pairwise between cities (these were as follows: Alkmaar vs. Haarlem: φ_{PT} = 0.056, P = 0.036; Alkmaar vs. Rotterdam: φ_{PT} = 0.030, P = 0.038; Arnhem vs. Goes: φ_{PT} = 0.153, P = 0.020 and Arnhem vs. Veldhoven: φ_{PT} = 0.126, P = 0.026) (Fig. S3, Supporting information). Conversely, when genotypes are grouped by STRUCTURE, we find significant global (φ_{PT} = 0.223, P = 0.001) and pairwise differentiation for each population (P = 0.001 for all comparisons) (Fig. S8, Supporting information). Finally, the AMOVA analysis shows that almost none of the genetic variation among populations is explained when genotypes are grouped by city of origin; in contrast, 23% of the genetic variation among populations is explained when genotypes are grouped by the STRUCTURE-assigned populations (Fig. S8, Supporting information).

Using a second analysis, we also tested whether genotypes from particular cities are represented disproportionately in specific populations. For each population, our statistical analyses reject an association between geography and population assignment ($P_{population1}$ = 0.96, $P_{population2}$ = 0.83, $P_{population3}$ = 0.78, $P_{population4}$ = 0.45 and $P_{population5}$ = 0.90).

Haploid diversity in the five Aspergillus fumigatus *populations*

Levels of haploid diversity are variable across the five *A. fumigatus* populations. Populations 1, 3, 4 and 5 show relatively high levels of haploid diversity ($uh_{population1}$ = 0.599, $uh_{population3}$ = 0.540, $uh_{population4}$ = 0.586 and $uh_{population5}$ = 0.531), whereas population two exhibits relatively lower levels of diversity ($uh_{population2}$ = 0.388).

To make estimates of haploid diversity comparable to previously published microsatellite analyses from other *Aspergillus* species (Tran-Dinh & Carter 2000; Hosid *et al.* 2008), we also calculated uh using only our microsatellite markers (Fig. 3). Again, we found that populations 1, 3, 4 and 5 show higher relative levels of haploid diversity ($uh_{population1}$ = 0.788, $uh_{population3}$ = 0.702, $uh_{population4}$ = 0.787 and $uh_{population5}$ = 0.711), compared with population 2 ($uh_{population2}$ = 0.408). Interestingly, uh values obtained for populations 1, 3, 4 and 5 are comparable to those found in two populations of *A. nidulans* ($uh_{population1}$ = 0.675 and $uh_{population2}$ = 0.598) (Hosid *et al.* 2008), a population of *A. flavus* (uh = 0.733) (Tran-Dinh & Carter 2000) and, to a lesser extent, a population of *A. parasiticus* (uh =0.505) (Tran-Dinh & Carter 2000), even though different sets of

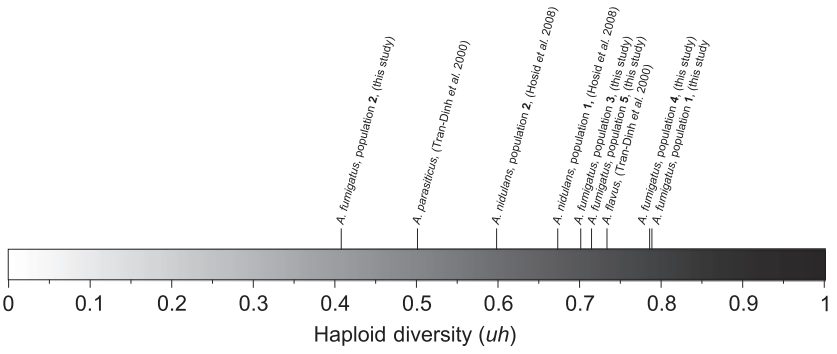

Fig. 3 Unbiased haploid diversity (uh) measures of the five *Aspergillus fumigatus* populations and other representative *Aspergillus* species. Microsatellite-based uh values from populations of other representative *Aspergillus* species are from the following studies: *A. flavus* (Tran-Dinh & Carter 2000), *A. parasiticus* (Tran-Dinh & Carter 2000) and two *A. nidulans* populations (Hosid *et al.* 2008).

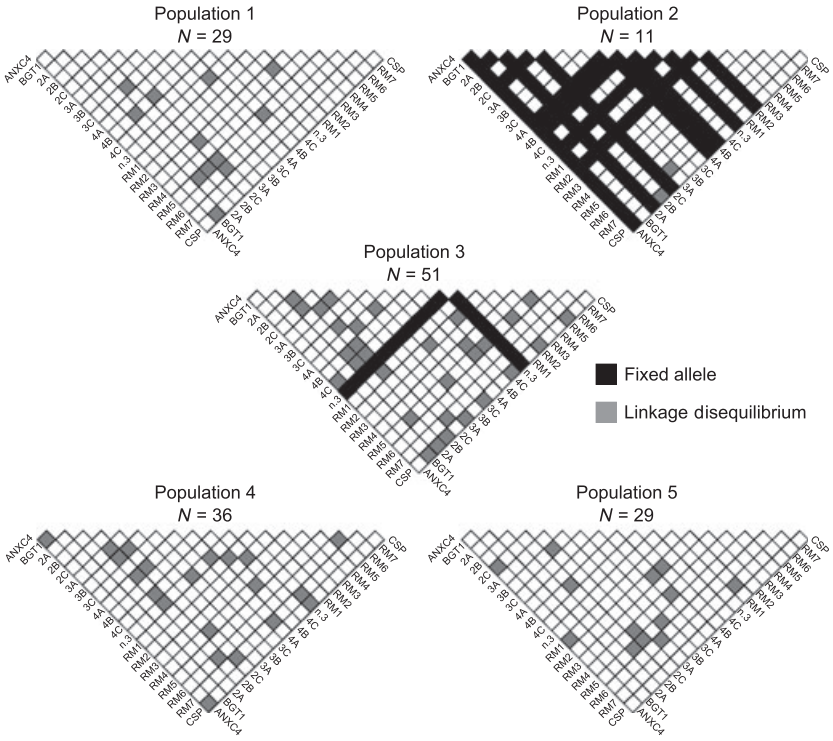

Fig. 4 Linkage disequilibrium (LD) patterns of the five *Aspergillus fumigatus* populations. The LD patterns of the five *A. fumigatus* populations. LD was determined by calculating the index of association (Ia) for all locus pairs independently for all populations. White, grey and black boxes represent loci in equilibrium, loci in significant LD and fixed loci, respectively.

markers were used in different studies. We observed few fixed loci in all populations (population 1 = 0, population 3 = 1, population 4 = 0 and population 5 = 0) with the exception of population 2, which contained seven fixed loci.

Recombination levels vary between the five Aspergillus fumigatus *populations*

Examination of LD suggests that *A. fumigatus* populations show varying levels of recombination (Fig. 4). Globally, populations 2–5 showed significant Ia values indicative of population-level LD (population 2: Ia = 0.588, $P = 0.003$; population 3: Ia = 0.429, $P < 0.001$; population 4: Ia = 0.374, $P < 0.001$; population 5: Ia = 0.471, $P < 0.001$), whereas values for population 1 (Ia = 0.106, $P = 0.114$) are indicative of a recombining population. Furthermore, of the 190 possible locus pairs tested, 12, 135, 47, 21 and 13 are fixed or in LD in populations 1–5, respectively (Fig. 3). Of the 76 observed locus pairs that were in LD, only 16 pairs involved markers located on the same chromosome, of which only nine involved neighbouring markers not separated by another marker.

Nonrecombining populations are expected to contain higher numbers of clonally related genotypes. Thus, clonal richness, i.e. the number of identical genotypes in a population, might be used as another indicator of reproductive mode. Levels of clonal richness across the five *A. fumigatus* populations were 50% (29 of 58 isolates belonging to population 1 were inferred to have identical genotypes) for population 1, 35% (six of 17) for population 2, 30% (22 of 73) for population 3, 45% (30 of 66) for population 4 and 29% (12 of 41) for population 5.

Divergence times between Aspergillus fumigatus *populations*

We estimated the average mutation rate, w, at 2.97×10^{-4}, a value comparable to that obtained for *A. flavus* (Grubisha & Cotty 2010). The pairwise upper and lower divergence time estimates between population 3, which contains the MTR alleles, and all others were (in 1000 generation multiples) as follows: Population 3 vs. 1: 102–19; vs. 2: 80–23; vs. 4: 111–33 and vs. 5: 109–51 (Table 1). We also estimated upper and lower boundaries of divergence times (again in 1000 generation multiples) between the MTR allele containing genotypes within population 3 and populations 1 (113–20), 2 (49–6), 4 (94–16) and 5 (58–7) (Table 1).

Discussion

To investigate the role of reproductive mode and population structure in evolution of triazole drug resistance, we studied the *A. fumigatus* populations in the Netherlands, where the MTR TR/L98H allele was first reported (Verweij *et al.* 2007). Given the evidence that

Table 1 Estimated times of divergence between *Aspergillus fumigatus* populations with upper and lower divergence time estimates given in units of 1000 generations

Population	#1	#2	#3	#4	#5
#2	167–67				
#3	102–19	80–23			
#4	76–0	102–32	111–33		
#5	54–0	28–0	109–51	68–6	
Multiple-triazole-resistance-containing isolates within population #3	113–20	49–6	N/A	94–16	58–7

recombination, commonly associated with sexual reproduction, can be very advantageous in stressful environments (Goddard *et al.* 2005; Zeyl *et al.* 2005), such as inside a human host or upon exposure to a fungicide, we hypothesized that sexual reproduction was facilitating the emergence and/or spread of the TR/L98H allele. However, analysis of the data obtained in this study allowed us to reject our original hypothesis, suggesting instead that all TR/L98H alleles identified in our study nest within a single, predominantly asexual, population and have not spread across populations. These results emphasize the role of asexual reproduction and genetic differentiation in shaping the evolution of azole resistance in *A. fumigatus*.

High-resolution markers show genetic differentiation in Dutch Aspergillus fumigatus

To our knowledge, this is the first study that reports the existence of genetic differentiation in any part of the distribution of *A. fumigatus*, one of the most important opportunistic fungal pathogens of humans, and suggests that Dutch *A. fumigatus* genotypes group into five distinct populations. Consistent with our findings, another recent study reported the existence of two genetically differentiated lineages within *A. fumigatus* (Pringle *et al.* 2005), although in that case the authors argued, based on phylogenetic analysis, that these lineages were separate species. The detection of genetic differentiation in *A. fumigatus* argues against the 'everything is everywhere' hypothesis, which states that highly abundant microbial eukaryote species with cosmopolitan distributions lack genetic differentiation (Finlay 2002).

One potential explanation for this discrepancy between past studies and ours might be the difference in the density and scale of sampling between studies; for example, in the two most comprehensive multilocus studies to date, 63 and 70 isolates from five different continents were analysed, respectively (Pringle *et al.* 2005; Rydholm *et al.* 2006), whereas our study examined a much larger number of isolates from a relatively small geographical area.

Another potential explanation for the discordance of these results with those from earlier studies might be our use of a much larger and more highly informative panel of markers (Figs S4 and S5, Supporting information). All previous studies have employed either sequence-based or RFLP-based typing techniques (Debeaupuis *et al.* 1997; Pringle *et al.* 2005; Rydholm *et al.* 2006). Although these techniques are very reliable (Taylor *et al.* 1999b), they are typically less informative when compared to microsatellite markers (Bain *et al.* 2007). If the absence of differentiation in past studies of *A. fumigatus* population biology is to be explained by the use of less-informative markers, then we expect that future studies on isolates from other geographical regions of comparable size using similar or superior markers (e.g. Harris *et al.* 2010) to ours are highly likely to identify genetically differentiated *A. fumigatus* populations. Interestingly, a previous study of *A. fumigatus* isolates performed using the most informative nonmicrosatellite marker from our panel did not find any evidence of genetic differentiation in North America (Balajee *et al.* 2007), suggesting that the pattern of differentiation of this important human pathogen might vary across its range of distribution.

Our results also show that the five identified *A. fumigatus* populations do not correlate with geography. In many eukaryotes, the presence of genetically differentiated populations is often the result of geographical isolation or ecological niche preference. Surprisingly, the populations in our study do not show any correlation with geography or environment of origin (clinical vs. soil). A similar lack of correlation with geography was observed in *A. flavus* populations (Grubisha & Cotty 2010). In this study, the genetic diversity of 243 *A. flavus* isolates was analysed using 24 microsatellite loci and the mating-type locus. Notably, all *A. flavus* populations were clonal, with no evidence of gene flow between populations.

Varying levels of recombination among Dutch Aspergillus fumigatus *populations*

Four of the five Dutch populations show significant levels of LD, suggesting that recombination in these populations has been rare or absent. Significant LD can be because of several different reasons (Maynard Smith *et al.* 1993); for example, significant LD is expected in populations of organisms that do not possess any molecular mechanism for recombination, such that clonal propagation is their sole means of reproduction. This explanation is unlikely to hold true for *A. fumigatus*. The sexual reproduction machinery in the *A. fumigatus*

genome appears intact (Galagan *et al.* 2005; Rokas & Galagan 2008), the MAT loci are typically at near-equal frequencies in *A. fumigatus* populations (Paoletti *et al.* 2005)—this is also the case in our populations, and certain isolates can reproduce sexually in the laboratory (O'Gorman *et al.* 2009).

Another reason for significant LD values involves failure to account for the genetic differentiation of the species studied. In such cases, the presence of LD will probably reflect the lack of recombination between populations of the species, potentially masking recombination within populations; for example, analysis of the entire 156-genotype data set reveals significant levels of LD ($I_A = 0.718$, $P < 0.001$), masking the finding that population 1 is not in LD. Given that we first identified the pattern of genetic differentiation in *A. fumigatus* and then calculated LD separately for each genetically differentiated population, it is highly unlikely that our results are affected by this reason.

Furthermore, the observed varying levels of LD suggest that different *A. fumigatus* populations may exhibit different reproductive modes. Nearly two decades ago, Maynard Smith *et al.* (1993) distinguished microbial reproductive modes into three models: clonal (predominantly nonrecombining), panmictic (predominantly recombining) and epidemic (predominantly recombining, but shows significant associations between loci because of recent, explosive increases in particular genotypes). On a first level of analysis, it appears that the observed pattern of reproductive modes across *A. fumigatus* populations, where four of the five populations are predominantly nonrecombining, might be more similar to the epidemic model than to either the clonal or the panmictic one. One expectation of the epidemic model is that clonal populations are very recent and have undergone explosive growth, so they are expected to harbour little genetic diversity. However, neither the estimated times of divergence nor the levels of genetic diversity for most of the nonrecombining *A. fumigatus* populations support the very recent origins. Similarly, levels of genotypic diversity are comparable across most populations, irrespective of the presence or absence of recombination. Rather, it appears that different *A. fumigatus* populations fit into different reproductive mode models.

How this variation in reproductive mode across *A. fumigatus* populations is controlled and whether these distinct populations use either sexual or asexual reproductive modes to regulate gene flow regardless of environmental signals remain open questions. It is known that both asexual and sexual reproductive modes can confer significant advantages as well as disadvantages to fungal populations (Sun & Heitman 2011), so the pattern of reproductive modes across populations might be determined by their balance. It is also theoretically possible that the only population of *A. fumigatus* identified in this study as recombining may in fact be reproducing asexually. This is so because, in fungi, genetic recombination can occur both during meiosis as part of the sexual cycle or during mitosis, as part of the parasexual cycle (Clutterbuck 1996; Taylor *et al.* 1999a). The effect of the parasexual cycle on long-term genetic exchange within fungal populations is thought to be limited, because it typically takes place between genetically similar individuals (Clutterbuck 1996), but its true extent on recombination within *A. fumigatus* populations is unknown.

Population structure, recombination and their implications for the spread of the MTR allele

Our finding that all genotypes containing the TR/L98H allele were confined to a single, predominantly asexually reproducing, population and has not yet spread across populations rejects our hypothesis that sexual reproduction facilitated the emergence and/or spread of the TR/L98H allele in *A. fumigatus*. Notably, very recent global surveillance studies have found isolates with the TR/L98H allele in both China and India (Lockhart *et al.* 2011; Chowdhary *et al.* 2011), but it is not yet known whether these isolates stem from the same population.

This lack of recombination and lack of gene flow are puzzling; in environments lacking triazole drugs, retention of MTR-like alleles is likely to be costly (Cowen *et al.* 2001; Stergiopoulos *et al.* 2003; Cowen 2008), whereas in environments containing triazole drugs, MTR-like alleles are likely to be strongly advantageous. In both cases, recombination (and associated gene flow) should be favoured—to eliminate the costly resistant allele in the first case, or to fix the advantageous resistant allele in the second, and yet our evidence suggests that the population is predominantly asexual. One possible explanation is that the sexual reproduction mode in *A. fumigatus* populations cannot be 'switched on' instantaneously once a population has become asexual because of the accumulation of deleterious mutations in mating and meiosis genes (Sun & Heitman 2011), despite the high costs associated with clonal reproduction during episodic selection, e.g. upon sporadic exposure to a fungicide.

When did the TR/L98H allele originate? Recently, Verweij *et al.* (2009) raised the hypothesis that the evolution of MTR resistance in *A. fumigatus* might have been a by-product of the use of azole compounds in agriculture. Dating of the divergence of population 3 genotypes as well as of all TR/L98H allele containing genotypes within population 3 from the other populations

(Table 1) implies that both groups diverged from the other populations ~6000 generations ago. Thus, the genetic background of MTR allele containing genotypes is probably more ancient than the first use of azole drugs in agriculture or medicine. Nevertheless, it is entirely plausible that the origin of the TR/L98H allele in a population 3 genetic background occurred much more recently (Verweij *et al.* 2009), a hypothesis consistent with the pattern of resistance spreading from the Netherlands to the rest of Europe. Future studies that examine the global genetic structure of *A. fumigatus* using a similarly informative set of markers are likely to be highly instructive on how the genetic differentiation and reproductive mode of *A. fumigatus* populations shape the evolutionary dynamics of drug-resistance patterns of this deadly human pathogen.

From a practical standpoint, elucidating the role of genetic differentiation and reproductive mode in influencing the genetic structure of *A. fumigatus* can also facilitate the development of diagnostic tools to detect resistant infections and preventive measures aimed to curb the spread of azole resistance in fungal populations. As the frequency of azole-resistant *A. fumigatus* isolates continues to increase in several European countries and beyond, a better understanding of the origin and spread of MTR alleles across *A. fumigatus* is becoming critical.

Acknowledgements

We thank Zeev Frenkel, Abraham Korol and Nai Tran-Dinh for access to raw *Aspergillus* genotype data, Dr A. Chowdhary for access to unpublished data and Thibaut Jombart for assistance with the DAPC approach. We are also grateful to David McCauley and the reviewers for their helpful suggestions and advice. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. This project has been funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under contract numbers N01-AI30071 and/or HHSN272200900007C, the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH, NIAID: F31AI091343-01 to JGG), the Searle Scholars Program (AR) and the National Science Foundation (DEB-0844968 to AR). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIAID or the NIH.

References

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

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge.

Bain JM, Tavanti A, Davidson AD *et al.* (2007) Multilocus sequence typing of the pathogenic fungus *Aspergillus fumigatus*. *Journal of Clinical Microbiology*, **45**, 1469–1477.

Balajee SA, Tay ST, Lasker BA, Hurst SF, Rooney AP (2007) Characterization of a novel gene for strain typing reveals substructuring of *Aspergillus fumigatus* across North America. *Eukaryotic Cell*, **6**, 1392–1399.

Butler G (2010) Fungal sex and pathogenesis. *Clinical Microbiology Reviews*, **23**, 140–159.

Chowdhary A, Kathuria S, Randhawa HS *et al.* (2011) Isolation of multiple-triazole resistant *Aspergillus fumigatus* strains, carrying the TR/L98H mutations in the *cyp*51A gene in India. *Journal of Antimicrobial Chemotherapy*, DOI: 10.1093/JAC/dkr443.

Clutterbuck AJ (1996) Parasexual recombination in fungi. *Journal of Genetics*, **75**, 281–286.

Cowen LE (2008) The evolution of fungal drug resistance: modulating the trajectory from genotype to phenotype. *Nature Reviews Microbiology*, **6**, 187–198.

Cowen LE, Kohn LM, Anderson JB (2001) Divergence in fitness and evolution of drug resistance in experimental populations of *Candida albicans*. *Journal of Bacteriology*, **183**, 2971–2978.

Cramér H (1999) *Mathematical Methods of Statistics*. Princeton University Press, Uppsala.

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

Denning DW (1998) Invasive aspergillosis. *Clinical Infectious Diseases*, **26**, 781–803.

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.

Fedorova ND, Khaldi N, Joardar VS *et al.* (2008) Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. *PLoS Genetics*, **4**, e1000046.

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**(Suppl. 1), S34–S41.

Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science*, **296**, 1061–1063.

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, Timberlake WE, Arnold ML (1996) Loss of meiosis in *Aspergillus*. *Molecular Biology and Evolution*, **13**, 809–817.

Goddard MR, Godfray HCJ, Burt A (2005) Sex increases the efficacy of natural selection in experimental yeast populations. *Nature*, **434**, 636–640.

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

Grubisha LC, Cotty PJ (2010) Genetic isolation among sympatric vegetative compatibility groups of the aflatoxin-producing fungus *Aspergillus flavus*. *Molecular Ecology*, **19**, 269–280.

Harris SR, Feil EJ, Holden MTG *et al.* (2010) Evolution of MRSA during hospital transmission and intercontinental spread. *Science*, **327**, 469–474.

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

Herbrecht R, Denning DW, Patterson TF *et al.* (2002) Voriconazole versus amphotericin B for primary therapy of invasive aspergillosis. *New England Journal of Medicine*, **347**, 408–415.

Hittinger CT, Goncalves P, Sampaio JP *et al.* (2010) Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature*, **464**, 54–58.

Hosid E, Grishkan I, Frenkel Z *et al.* (2008) Diversity of microsatellites in natural populations of ascomycetous fungus, *Emericella nidulans*, in Israel on local and regional scales. *Mycological Progress*, **7**, 99–109.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jombart T (2008) ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Klaassen CH, de Valk HA, Balajee SA, Meis JF (2009) Utility of CSP typing to sub-type clinical *Aspergillus fumigatus* isolates and proposal for a new CSP type nomenclature. *Journal of Microbiological Methods*, **77**, 292–296.

Klaassen CH, de Valk HA, Curfs-Breuker IM, Meis JF (2010) Novel mixed-format real-time PCR assay to detect mutations conferring resistance to triazoles in *Aspergillus fumigatus* and prevalence of multi-triazole resistance among clinical isolates in the Netherlands. *Journal of Antimicrobial Chemotherapy*, **65**, 901–905.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Lockhart SR, Frade JP, Etienne KA *et al.* (2011) Azole resistance in *Aspergillus fumigatus* isolates from the ARTEMIS global surveillance is primarily due to the TR/L98H mutation in the *cyp51A* gene. *Antimicrobial Agents and Chemotherapy*, **55**, 4465–4468.

Lynch M, Sung W, Morris K *et al.* (2008) A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Sciences, USA*, **105**, 9272–9277.

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

Meis JF, Verweij PE (2001) Current management of fungal infections. *Drugs*, **61**(Suppl. 1), 13–25.

Mellado E, Garcia-Effron G, Alcazar-Fuoli L *et al.* (2007) A new *Aspergillus fumigatus* resistance mechanism conferring in vitro cross-resistance to azole antifungals involves a combination of cyp51A alterations. *Antimicrobial Agents and Chemotherapy*, **51**, 1897–1904.

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

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.

Nielsen K, Heitman J (2007) Sex and virulence of human pathogenic fungi. *Advances in Genetics*, **57**, 143–173.

Nierman WC, Pain A, Anderson MJ *et al.* (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature*, **438**, 1151–1156.

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

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

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

Pringle A, Baker DM, Platt JL *et al.* (2005) Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. *Evolution*, **59**, 1886–1899.

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

Rokas A (2009) The effect of domestication on the fungal proteome. *Trends in Genetics*, **25**, 60–63.

Rokas A, Galagan JE (2008) The *Aspergillus nidulans* genome and a comparative analysis of genome evolution in *Aspergillus*. In: *The Aspergilli: Genomics, Medical Applications, Biotechnology, and Research Methods* (eds Goldman GH and Osmani SA), pp. 43–55. CRC Press, Boca Raton, FL.

Rokas A, Payne G, Fedorova ND *et al.* (2007) What can comparative genomics tell us about species concepts in the genus *Aspergillus*? *Studies in Mycology*, **59**, 11–17.

Rydholm C, Szakacs G, Lutzoni F (2006) Low genetic variation and no detectable population structure in *Aspergillus fumigatus* compared to closely related *Neosartorya* species. *Eukaryotic Cell*, **5**, 650–657.

Scannell DR, Zill OA, Rokas A *et al.* (2011) The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3: Genes, Genomes, Genetics*, **1**, 11–25.

Snelders E, van der Lee HA, Kuijpers J *et al.* (2008) Emergence of azole resistance in *Aspergillus fumigatus* and spread of a single resistance mechanism. *PLoS Medicine*, **5**, e219.

Stergiopoulos I, van Nistelrooy JGM, Kema GHJ, De Waard MA (2003) Multiple mechanisms account for variation in base-line sensitivity to azole fungicides in field isolates of *Mycosphaerella graminicola*. *Pest Management Science*, **59**, 1333–1343.

Sun S, Heitman J (2011) Is sex necessary? *BMC Biology*, **9**, 56.

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

Taylor J, Jacobson D, Fisher M (1999a) The evolution of asexual fungi: reproduction, speciation and classification. *Annual Review of Phytopathology*, **37**, 197–246.

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

Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D (2006) Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom

Fungi. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, **361**, 1947–1963.
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 microsatellite data. *Genetics*, **169**, 1589–1599.
Tran-Dinh N, Carter D (2000) Characterization of microsatellite loci in the aflatoxigenic fungi *Aspergillus flavus* and *Aspergillus parasiticus*. *Molecular Ecology*, **9**, 2170–2172.
de Valk HA, Meis JF, Curfs IM *et al.* (2005) Use of a novel panel of nine short tandem repeats for exact and high-resolution fingerprinting of *Aspergillus fumigatus* isolates. *Journal of Clinical Microbiology*, **43**, 4112–4120.
Varga J, Toth B (2003) Genetic variability and reproductive mode of *Aspergillus fumigatus*. *Infection, Genetics and Evolution*, **3**, 3–17.
Verweij PE, Mellado E, Melchers WJ (2007) Multiple-triazole-resistant aspergillosis. *New England Journal of Medicine*, **356**, 1481–1483.
Verweij PE, Snelders E, Kema GH, Mellado E, Melchers WJ (2009) Azole resistance in *Aspergillus fumigatus*: a side-effect of environmental fungicide use? *Lancet Infectious Diseases*, **9**, 789–795.
Yang Z (2006) *Computational Molecular Evolution*. Oxford University Press, Oxford.
Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.
Zeyl C, Curtin C, Karnap K, Beauchamp E (2005) Antagonism between sexual and natural selection in experimental populations of *Saccharomyces cerevisiae*. *Evolution*, **59**, 2109–2115.
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.

C.H.W.K. is a molecular microbiologist with a major interest in the development and application of molecular diagnostic tools for detection- identification- and characterisation of pathogenic micro-organsism, especially fungi. J.G.G. and A.R. have a long-standing interest in using comparative functional genomic techniques to understand the molecular basis and evolution of the fungal lifestyle. N.D.F research is focused on integration of high-throughput data to facilitate the development of therapeutic and diagnostic approaches against fungal infectious disease. J.F.M. is a consultant microbiologist focussing on diagnosis and therapy of opportunistic fungal infections.

Data accessibility

Genotype data and sample information: DRYAD entry doi:10.5061/dryad.7m0797t0.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 The effect of different clonal correction thresholds on the number of nonclonal genotypes identified from 255 Dutch *A. fumigatus* isolates.

Fig. S2 STRUCTURE and DAPC analysis of the microsatellite marker (panels A and B), nonmicrosatellite marker (panels C and D) and full marker (panels E and F) data sets.

Fig. S3 The optimal number of populations *K* predicted by STRUCTURE analysis on data sets with varying levels of clonal correction.

Fig. S4 The distribution of alleles for each of the nine microsatellite markers and the one indel marker used in this study across the five *A. fumigatus* populations, as delineated by the STRUCTURE analysis.

Fig. S5 The distribution of alleles for each of the 10 sequence/PCR-typing markers used in this study across the five *A. fumigatus* populations, as delineated by the STRUCTURE analysis.

Fig. S6 Strength of association between markers and populations according to Cramér's *V* statistic (Cramér 1999).

Fig. S7 STRUCTURE analysis is highly consistent across replicates.

Fig. S8 Geographical origin is not associated with genetic differentiation.

Table S1 The nomenclature, description, marker type, genomic location and origin of the 20 markers used in this study.

Table S2 The population assignment of the 156 nonclonally related clinical and environmental *A. fumigatus* genotypes from the Netherlands into five populations using the STRUCTURE and DAPC approaches.

Table S3 The individual membership coefficients of the 156 nonclonally related clinical and environmental Dutch *A. fumigatus* genotypes from the STRUCTURE replicate used in this study against the average individual membership coefficients from the 100 STRUCTURE replicates calculated using the CLUMPP software.

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.