Clonality Despite Sex: The Evolution of Host-Associated Sexual Neighborhoods in the Pathogenic Fungus *Penicillium marneffei*


1Department of Infectious Disease Epidemiology, Imperial College, Norfolk Place, London, United Kingdom, 2Departments of Microbiology, Regional Institute of Medical Sciences, Imphal, Manipur, India, 3Department of Genetics, University of Melbourne, Victoria, Australia, 4The Third People’s Hospital, Shenzhen City, Guangdong Province, China, 5The J. Craig Venter Institute, Rockville, Maryland, United States of America, 6Departments of Laboratory Medicine and Internal Medicine, National Taiwan University Hospital, National Taiwan University College of Medicine, Taipei, Taiwan, Republic of China, 7Department of Microbiology, The University of Hong Kong, Queen Mary Hospital, Hong Kong Island, Hong Kong, 8Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam, 9National Hospital for Tropical Diseases, Ha Noi, Viet Nam, 10Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Program Viet Nam National Hospital for Tropical Diseases, Ho Noi, Viet Nam, 11Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme Viet Nam Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam, 12Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand, 13Department of Microbiology, Imperial College London, London, United Kingdom

Abstract

Molecular genetic approaches typically detect recombination in microbes regardless of assumed asexuality. However, genetic data have shown the AIDS-associated pathogen *Penicillium marneffei* to have extensive spatial genetic structure at local and regional scales, and although there has been some genetic evidence that a sexual cycle is possible, this haploid fungus is thought to be genetically, as well as morphologically, asexual in nature because of its highly clonal population structure. Here we use comparative genomics, experimental mixed-genotype infections, and population genetic data to elucidate the role of recombination in natural populations of *P. marneffei*. Genome wide comparisons reveal that all the genes required for meiosis are present in *P. marneffei*, mating type genes are arranged in a similar manner to that found in other heterothallic fungi, and there is evidence of a putatively meiosis-specific mutational process. Experiments suggest that recombination between isolates of compatible mating types may occur during mammal infection. Population genetic data from 34 isolates from bamboo rats in India, Thailand and Vietnam, and 273 isolates from humans in China, India, Thailand, and Vietnam show that recombination is most likely to occur across spatially and genetically limited distances in natural populations resulting in highly clonal population structure yet sexually reproducing populations. Predicted distributions of three different spatial genetic clusters within *P. marneffei* overlap with three different bamboo rat host distributions suggesting that recombination within hosts may act to maintain population barriers within *P. marneffei*.

Introduction

Hypotheses of globally continuous populations and strict clonality in putatively asexual microbial pathogens are rarely supported [1,2,3,4,5]. Instead, genetic approaches detect recombination in microbes regardless of assumed asexuality, and pathogens are surprisingly promiscuous despite strong population genetic structure [6,7,8]. In some eukaryotic pathogens spatial structuring is readily attributable to dispersal limitations [9,10], but many fungal pathogens of humans display extensive spatial population genetic structure despite their ability to disperse via aerosolized spores [11,12,13,14,15]. Examples that cause extensive morbidity include *Cryptococcus neoformans*, *Coccidioides* sp., *Histoplasma capsulatum* and *Penicillium marneffei*. These fungi are maintained in natural environmental reservoirs that might contribute to structured populations via local adaptation, and they are thought to be largely clonal. However, fungi have many different mating systems that encompass asexual propagation through multiple forms of sexual and parasexual recombination, and clonal structure may be arrived at via very different mechanisms [16,17]. Evidence
Fungal Summary

Fungal pathogen populations show patterns ranging from globally recombining to endemic and clonal. Among the most genetically and spatially restricted fungi is the highly clonal pathogen Penicillium marneffei, an endemic AIDS-associated pathogen in Southeast Asia. Previous studies have shown that P. marneffei has a pattern of extreme clonality despite the ability to disperse across wide distances and the presence of mating type genes that are required for sexual recombination. In this study we used genetic markers, comparative genomics, experimental data, and spatial models to determine the influence of sex on P. marneffei populations, and we found that although there was substantial evidence of sexual recombination, most of the recombination in natural populations was limited to sexual neighborhoods, amongst genetically similar and spatially close individuals. Based on the results of experiments and spatial models we found support for sex occurring in bamboo rats that are known to harbor P. marneffei and the pathogens sexual neighborhoods. Our study suggests that the high levels of effective clonality and endemicity found in P. marneffei may have more to do with specific host interactions than with an innate inability to generate population genetic diversity through sexual recombination.

Materials and Methods

Genomic analyses

We used 84 sexual cycle genes (Table S1) to blast against the NCBI genome sequences NZ_AAHH00000000 (A. fumigatus), NZ_ABAB00000000 (P. marneffei), and NZ_ABAS00000000 (T. stipitatus). We screened transposon families for substitution bias by making BLAST based alignments to determine the dominant form of a functional integrase gene in each family. We counted the type of substitution based on differences of alleles as low as 70% identical to the dominant intact type. We compared gene sequences of the genomic region between sltB and aopB (the genes that flank the MAT idiomorph in related fungi) of strains FRR2161 and FRR3842.

Isolates, MLMT barcoding and MAT discrimination

We acquired 307 isolates of P. marneffei from humans and vertebrate hosts (bamboo rats), covering the known global range of the fungus. Our study obtained 273 epidemiologically unlinked human isolates of P. marneffei from HIV-AIDS patients covering the time-period 1959 to 2005. Of these isolates, 258 were georeferenced to either the broad geographical region of collection or the patients home address. The remaining 15 isolates were recovered from patients whose infections were diagnosed in non-endemic regions, and no accurate geographical origin could be assigned. In addition to human isolates of P. marneffei, we obtained 34 isolates from the bamboo rats species Rhizomys pruinosus (n = 3), R. sumatrensis (n = 13), R. sinensis (n = 1) and Cannomys badius (n = 17). We also include the type isolate for P. marneffei Segretain et al. ATCC 18224, CBS 388:37, isolated from R. sinensis in 1959 [26]. All isolates were cultured on Sabouraud’s agar and DNA extracted as previously described [27]. Subsequently, isolates were genotyped at 21 microsatellite loci using the methods described in Fisher et al. [11,27].

The presence within each of isolate of the MAT1-1 α box and MAT1-2 high mobility group idiomorphs was determined using the PCR protocol detailed by Woo et al. [25].

Population genetic structure and spatial analyses

Genotypes were analysed using GenAlex 6.0 [28] to determine allelic diversity, genotypic diversity and spatial correlation across regions and the global distribution of P. marneffei. We used the package adegenet and its dependencies in R to conduct spatial PCA and DAPC analyses [29]. To compare our inferred results against a model of a single continuous population structured by a dingle dispersal kernel and mutation rate we used the coalescent based program IBDsim [30]. Additional distribution data for bamboo rats were collected from specimen databases AMNH, FMNH, NMMH, and the GBIF. We used the bioclim layers 1–21 at 30 sec from the world clim database in MAXENT to generate predicted distributions for bamboo rats and P. marneffei genetic clusters. We measured distributional overlap using Schoener’s D and a resampling approach [31]. We compared the relative overlap of genetic clusters to host distributions by generating null distributions of D based on resampling of R. sumatrensis and Cannomys (Text S1). Possible parental distances were compared to null distributions generated by choosing isolates randomly that met the genetic distance criteria from the population that met the parental criteria.
Co-infection analysis

Five co-housed outbred CD-1 male mice (16–18 g) were inoculated intranasally with 10^7 spores suspended in 40 μl of PBS. Conidia from two isolates, PM9, a MAT 1–2 isolate from Thailand, and the type strain ATCC18824 (FR2161) were mixed in a 1:1 ratio to form the inoculum (S9). Serial dilutions of homogenized saline samples were plated (no later than 6 hours after they were removed from the mice) on Sabouraud agar. Colonies were counted after 4 days in 27°C. Individual colonies used for DNA extraction and subsequent genotyping as before [11]. Isolate genotypes were compared to the initial genotypes of the inoculum and genotypes differing from inoculum were confirmed via DNA sequencing.

Ethics statement

All the clinical studies from which isolates are available were approved by the Wellcome Trust ethics committees at the study sites, in the UK and by the regulatory authorities of the countries involved. All patients or their next of kin gave written informed consent and all patient data are anonymised. This work strictly complied with the animal regulations and guidelines under UK law and was approved by Imperial College’s Ethical Review Process (ERP) Committee and the British Home Office. All murine work was carried out in a Biosafety level 3 secure animal facility under licensed approval from the British Home Office.

Results/Discussion

Genomic evidence for sex in *P. marneffei*

Sexual reproduction leaves an imprint on fungal genomes by maintaining genes required for mating and by generating patterns of mutation and recombination restricted to meiotic processes [32,33,34,35,36]. Successful mating in fungi requires that a genome contains a functioning series of interconnected genetic pathways [37]. Using a comparative genomic approach we assessed the presence and functionality in *P. marneffei* of genes known to be involved in sexual development in fungi. First, comparing between strains FRR2161 and FRR3842 revealed that the region between genes *slab* and *apoB* resembled other fungal mating type idiomorphs. A region of complete dissimilarity was flanked by regions that were nearly identical between the strains (Fig. 1A and B). We found homologs for nearly all of the genes flanked by regions that were nearly identical between the strains (Fig. 1A and B). We found homologs for nearly all of the genes needed for a complete sexual cycle in yeast to be present and putatively functional (Table S1). Those genes not detected in *P. marneffei* were also not detected in *Talatymyces stipitatus*, a fungus with a complete sexual cycle, and most genes that were absent in those two fungi were also missing in the recently demonstrated heterothallic fungus *Aspergillus fumigatus*. Although these sex-related genes may be conserved to function in processes other than mating, their presence suggests that *P. marneffei* has preserved the ability to complete a sexual cycle. We detected another genomic signature of a functional sexual cycle, a type of mutation bias associated with meiosis. Repeat induced point mutation (RIP), a process by which some fungi silence genes involved in mobile genetic element function by preferentially mutating repeated sequences within their genomes, is associated with meiosis [38]. This process results in skewed base pair distributions due to the induced mutations. Using an approach similar to that of Clutterbuck [39], we found evidence of an excess of sliding windows with zero AG and CT dinucleotides and mutation bias in *P. marneffei* transposon family Ty-1 with a skew towards G to A and C to T transitions (Fig. 2). We also detected a putatively functional RID gene (Locus ID PMAA079888), the only conserved gene so far implicated in RIP [40,41,42]. Although the RID gene and the observed mutation bias can be explained by several factors including those acting during mitosis, they point towards a RIP or RIP-like process that is generally considered a feature of sexually reproducing fungi and an overall genomic pattern consistent with sex. Although these genomic signatures could represent relics from a sexual past rather than ongoing sexual recombination within *P. marneffei*, in the related human pathogenic fungus *Aspergillus fumigatus*, the discovery of mating type genes and evidence of RIP heralded the eventual description of a full sexual life cycle [35,43,44].

Spatially correlated genetic diversity

Microsatellite allelic diversity was high overall and within localities (Table 1). With the exception of Thai Central and Thai South, populations assigned a priori by locality were significantly differentiated from one another by Wright’s *F*<sub>ST</sub> [45]; this metric ascertains the proportion of genetic variance among geographical regions relative to the total variance. *F*<sub>ST</sub> values near zero mean that populations are not distinct and variation is shared equally within and between them, while higher *F*<sub>ST</sub> values mean that more genetic differences occur between populations compared to those within populations. The China and Taiwan populations were most different from the other a priori populations (Table 2). Phylogenetic analysis revealed associations between sampling area and the occurrence of phylogenetic clustering (Fig. 3). Using discriminant analysis of principal components (DAPC) to identify genetic clusters [46,47], we assigned individuals to 3 clusters based on the Bayesian information criterion. The clusters show some spatial association. Cluster 1 is composed mostly of isolates from central and southern Thailand, Cluster 2 of isolates from China, and Cluster 3 of isolates from northern Thailand (Fig. 3, Figure S1). As expected, we observed a strong pattern of spatial genetic correlation (*r^2 = 0.41, p < 0.01*). We also detected significant ‘global’ genetic structure (positive correlation between spatial and genetic distance) but no ‘local’ genetic structure (negative correlation) using spatial principal coordinate analysis [46,47]. To test for a homogenous neutral process of genetic differentiation we used a spatially explicit coalescent-based simulation of isolation by distance generated with IBDsim [30] to simulate a uniform dispersal/mutation process across our exact sampling scheme. This uniform genetic structure was then compare against our recovered spatial genetic pattern. By controlling for the spatial distribution of our sample sites we are able to determine if the apparent genetic clustering is simply an artificial product of clustered sampling and a single uniform process of genetic clustering. Because the hypothesised parameter space is nearly infinite, we concentrated on dispersal scenarios that most closely resembled the spatial genetic correlation present in our data, namely, the strength of spatial genetic correlation at the smallest spatial scale and the decay rate of the correlation. Although the simulated datasets largely overlapped with our recovered data we observed important departures between the two. The simulated datasets had a single peak in spatial genetic correlation at the smallest spatial scale and a decay in correlation dependent on the dispersal kernel, a feature of all single population isolation-by-distance models, but the observed pattern had additional peaks in certain distance classes that disrupted the uniform decay (Fig S2 and S3). One peak was composed of distances between individuals belonging to the outer edges of clusters 1 and 3. Another major peak comes at the spatial scale where the outer edges of Clusters 1 and 3 contact with Cluster 2. These results differ from previous results that observed different rates of decay for spatial genetic correlation [24], a feature that probably owes to the limited geographic scope and power of the earlier study. Our data now...
suggest that the observed genetic clusters are not the result of a process of uniform decay with geographic distance, and that other factors are also driving the heterogeneity observed in our dataset.

Recombination and linkage disequilibrium

Linkage disequilibrium was high throughout the sample with an overall $r^2$ of 0.113 (Table S2). We determined the relative frequency of mutation to recombination using the single locus variant approach applied by Fisher et al. [24] and found a mutation to recombination frequency of 0.083 suggesting that mutation is up to 12 times less frequent than recombination across the whole population. When restricted to only bamboo rat isolates, all single locus variants would be due to recombination, while for human-only isolates the ratios are unchanged in comparison to the entire dataset. Average fungal microsatellite mutation rates have been inferred from between $2.80 \times 10^{-6}$ and $2.50 \times 10^{-7}$ mutations.

Figure 1. Mating type idiomorph of *P. marneffei*. A) A dot plot showing similarity between strains FRR2161 and FRR3482 similarity across the 7 kb covering MAT loci. B) Gene cartoon showing arrangement of genes in the MAT idiomorphs. Grey regions are areas of high homology, and black lines are small segments of rearrangement. Coding regions of MAT alpha and HMG genes are shown in black, flanking genes with high homology are shown in grey, and hypothetical genes with low homology are shown in white. Arrows indicate the direction of transcription.

doi:10.1371/journal.ppat.1002851.g001

Figure 2. Biased mutation in *P. marneffei*. Nucleotide bias in *P. marneffei*. A) AG dinucleotide frequency across the entire genome of *P. marneffei* shows a non-normal distribution consistent with patterns of RIP-like process. B) Comparison between the types of mutations from the inferred ancestral state shows a bias in transitions and towards G to A and C to T mutations in particular.

doi:10.1371/journal.ppat.1002851.g002
per generation [48], making the inferred recombination rate in *P. marneffei* between $3 \times 10^{-5}$ and $2.50 \times 10^{-4}$, a rate about half that observed in wild yeast [49]. This approach only detects single locus recombination events, which may be a minority in eukaryotic populations, while it should detect virtually all mutations that have not otherwise been masked by recombination. However, the method could be strongly biased towards inferring recombination due to convergent mutations in microsatellite length. The measure of minimal recombination ($R_{MI}$), which represents the minimum number of recombination events necessary to explain alleles failing the four gamete test [50] given the arrangement of the alleles in a contig, showed that recombination did occur within contigs (Figure S4). Complete clonality and complete panmixia are rejected for *P. marneffei*, but similar to previous results the inferred levels of clonality remain among the highest observed for fungi [24]. To explain the high level of clonal structure either recombination must be rare or it must occur largely between closely related individuals.

**Distribution of MAT genes**

The entire sample population of *P. marneffei* showed a distribution of mating types that was significantly skewed ($p = 0.02$ or $p = 0.04$ when clone corrected) from a ratio of 1:1 in favour of an overabundance of MAT1-1 alleles, but some local populations were skewed towards MAT1-2 alleles (Table 1). Two of the genetic clusters inferred by DAPC were skewed towards more MAT1-1 alleles, but MAT genes within the central cluster did not differ from a 1:1 ratio (Table 3). As predicted in work prior to the discovery of MAT loci in *P. marneffei*, highly skewed MAT ratios would be expected in a predominately asexual population [24]. On one hand, in the absence of sex and selection, MAT genes at an initial frequency of 0.5 are expected to be fixed in a population on average by $\ln (2/N)$ generations. Alternatively, in a completely sexual population without selection associated with a mating type, MAT alleles would be maintained at frequencies near 0.5 with very limited variance because all individuals in each generation will possess MAT alleles according to a binomial distribution, and there is no opportunity for drift beyond a single generation. MAT allele counts can be used to represent the reduction in effective population size caused by drift in MAT ratios [51,52], but this assumes a fully sexual population. When sex is limited, the average allele frequencies for populations that do not lose sex and become fixed remain 0.5, but the variance in MAT allele frequency depends on population size and the frequency of sex. Based only on the differences in MAT allele frequencies between clusters and an intrinsic restriction on sex, *P. marneffei* would have an intrinsic

**Table 1. Local population diversity.**

<table>
<thead>
<tr>
<th>Population</th>
<th>n</th>
<th>MT</th>
<th>MAT1-1(Mi)</th>
<th>MAT1-2(Mi)</th>
<th>Na</th>
<th>Ne</th>
<th>Hs</th>
<th>Ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHINA</td>
<td>66</td>
<td>57</td>
<td>42(70)</td>
<td>20(30)</td>
<td>5.524(0.872)</td>
<td>2.716(0.519)</td>
<td>0.519(0.050)</td>
<td>1.238(0.425)</td>
</tr>
<tr>
<td>INDIA</td>
<td>11</td>
<td>6</td>
<td>7(27)</td>
<td>2(22)</td>
<td>1.857(0.186)</td>
<td>1.582(0.282)</td>
<td>0.282(0.055)</td>
<td>0.095(0.006)</td>
</tr>
<tr>
<td>THAI-C</td>
<td>22</td>
<td>19</td>
<td>16(76)</td>
<td>5(24)</td>
<td>3.952(0.417)</td>
<td>2.281(0.470)</td>
<td>0.470(0.044)</td>
<td>0.046(0.048)</td>
</tr>
<tr>
<td>THAI-E</td>
<td>25</td>
<td>23</td>
<td>4(16)</td>
<td>5(20)</td>
<td>3.619(0.519)</td>
<td>2.344(0.434)</td>
<td>0.434(0.056)</td>
<td>0.190(0.088)</td>
</tr>
<tr>
<td>THAI-N</td>
<td>120</td>
<td>62</td>
<td>72(70)</td>
<td>31(30)</td>
<td>3.619(0.434)</td>
<td>1.834(0.331)</td>
<td>0.331(0.057)</td>
<td>0.333(0.159)</td>
</tr>
<tr>
<td>THAI-S</td>
<td>19</td>
<td>18</td>
<td>9(47)</td>
<td>10(53)</td>
<td>4.095(0.539)</td>
<td>2.726(0.515)</td>
<td>0.515(0.047)</td>
<td>0.429(0.235)</td>
</tr>
<tr>
<td>TAIWAN</td>
<td>8</td>
<td>3</td>
<td>8(100)</td>
<td>0(0)</td>
<td>1.095(0.066)</td>
<td>1.042(0.028)</td>
<td>0.028(0.020)</td>
<td>0.000(0.000)</td>
</tr>
<tr>
<td>VIETNAM-N</td>
<td>10</td>
<td>8</td>
<td>4(40)</td>
<td>4(60)</td>
<td>3.333(0.319)</td>
<td>2.561(0.512)</td>
<td>0.512(0.052)</td>
<td>0.238(0.136)</td>
</tr>
<tr>
<td>VIETNAM-S</td>
<td>11</td>
<td>11</td>
<td>1(9)</td>
<td>10(91)</td>
<td>2.667(0.374)</td>
<td>2.073(0.363)</td>
<td>0.363(0.061)</td>
<td>0.046(0.048)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>292</td>
<td>197</td>
<td>163(65)</td>
<td>87(35)</td>
<td>3.307(0.177)</td>
<td>2.129(0.384)</td>
<td>0.384(0.020)</td>
<td>0.000(0.000)</td>
</tr>
</tbody>
</table>

n, number of samples; MT, number of microsatellite types (= haplotypes); Na, mean number of alleles; Ne, mean effective number of alleles; Hs, mean haploid genetic diversity; Ps, mean number of private alleles. Standard errors are shown in parentheses for Na, Ne, Hs and Ps.

doi:10.1371/journal.ppat.1002851.t001

**Table 2. Local population differentiation.**

<table>
<thead>
<tr>
<th>Population</th>
<th>CHINA</th>
<th>INDIA</th>
<th>THAI-C</th>
<th>THAI-E</th>
<th>THAI-N</th>
<th>THAI-S</th>
<th>TAIWAN</th>
<th>VIETNAM-N</th>
<th>VIETNAM-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHINA</td>
<td>0.815</td>
<td>0.716</td>
<td>0.706</td>
<td>0.927</td>
<td>0.752</td>
<td>0.277</td>
<td>0.412</td>
<td>0.808</td>
<td></td>
</tr>
<tr>
<td>INDIA</td>
<td>0.386</td>
<td>0.268</td>
<td>0.373</td>
<td>0.322</td>
<td>0.257</td>
<td>1.211</td>
<td>0.416</td>
<td>0.462</td>
<td></td>
</tr>
<tr>
<td>THAI-C</td>
<td>0.319</td>
<td>0.220</td>
<td>0.214</td>
<td>0.048</td>
<td>0.070</td>
<td>1.118</td>
<td>0.414</td>
<td>0.471</td>
<td></td>
</tr>
<tr>
<td>THAI-E</td>
<td>0.332</td>
<td>0.301</td>
<td>0.154</td>
<td>0.216</td>
<td>0.189</td>
<td>1.240</td>
<td>0.532</td>
<td>0.434</td>
<td></td>
</tr>
<tr>
<td>THAI-N</td>
<td>0.462</td>
<td>0.345</td>
<td>0.059</td>
<td>0.242</td>
<td>0.077</td>
<td>1.304</td>
<td>0.576</td>
<td>0.487</td>
<td></td>
</tr>
<tr>
<td>THAI-S</td>
<td>0.309</td>
<td>0.196</td>
<td>0.019</td>
<td>0.123</td>
<td>0.097</td>
<td>1.023</td>
<td>0.417</td>
<td>0.350</td>
<td></td>
</tr>
<tr>
<td>TAIWAN</td>
<td>0.267</td>
<td>0.746</td>
<td>0.562</td>
<td>0.595</td>
<td>0.647</td>
<td>0.527</td>
<td>0.740</td>
<td>1.214</td>
<td></td>
</tr>
<tr>
<td>VIETNAM-N</td>
<td>0.200</td>
<td>0.283</td>
<td>0.208</td>
<td>0.276</td>
<td>0.405</td>
<td>0.183</td>
<td>0.521</td>
<td>0.641</td>
<td></td>
</tr>
<tr>
<td>VIETNAM-S</td>
<td>0.359</td>
<td>0.382</td>
<td>0.290</td>
<td>0.297</td>
<td>0.410</td>
<td>0.214</td>
<td>0.687</td>
<td>0.317</td>
<td></td>
</tr>
</tbody>
</table>

$F_{ST}$, the proportion of variance among geographical regions relative to the total variance. Pairwise population values of $F_{ST}$ are below the diagonal with significant values ($p < 0.001$ from 999 permutations) in bold. Pairwise Nei’s genetic distances are shown above the diagonal.

doi:10.1371/journal.ppat.1002851.t002
Figure 3. Correspondence between genetic diversity and spatial location. A) Unrooted neighborjoining tree of genetic distances between isolates labelled with the colour each isolation locality. B) Genetic clusters inferred using DAPC and isolates coloured according to cluster. C) Spatial distribution of DAPC clusters.

doi:10.1371/journal.ppat.1002851.g003
upper bound of sexual recombination frequency at less than 4.5% given a modest population size of 1000. This small level of sex could explain the highly skewed ratio of MAT alleles in Cluster 3 and still accommodate the 1:1 ratio in Cluster 1 while avoiding any fixation of MAT alleles. However, if the intrinsic sexual recombination rate explained the distribution of MAT alleles it would predict equal frequencies of clone detection across populations. Instead, percent clonality tracks the MAT allele skew, suggesting that sexual recombination in Cluster 3 is reduced relative to Cluster 1 (Table 3). We do note, however, that MAT allele frequencies would not be informative about where sex occurs if unisexual mating occurs in *P. marneffei* as is known in *C. neoformans* [22].

**Sexual neighbourhoods**

Given that recombination occurs in *P. marneffei*, we wanted to determine the geographic and genetic scope of sex. Out of 43 clonal groups inferred with EBURST [53], six contained both MAT alleles, and four otherwise genetically identical multi-locus microsatellite types contained both MAT alleles (Fig. 4). Otherwise genetically identical isolates that differ only at mating type have also been detected in *Cryptococcus gattii* populations [54,55,56]. Clones with both MAT alleles represent the smallest possible genetic scale of sex, and are unequivocal evidence for recombination. To detect the more divergent recombination events we defined putative recombinants as any genotype that had no unique alleles, yet differed from the most similar genotype for at least 1.5%. We describe here the most divergent recombination events as putative recombinants.

**Table 3.** Mating type counts within clusters and the percent of isolates with MLMT identical to at least one other isolate in the sample.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Mat1-1-1</th>
<th>Mat1-2-1</th>
<th>Ne(MAT)*</th>
<th>%MTClonesb</th>
<th>%MAT1-1-1 Clonesc</th>
<th>%MAT1-2-1 Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33</td>
<td>30</td>
<td>1.00</td>
<td>10</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
<td>27</td>
<td>0.93</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>74</td>
<td>36</td>
<td>0.88</td>
<td>40</td>
<td>46</td>
<td>28</td>
</tr>
</tbody>
</table>

*Ne(MAT) is the estimated effective population size based on mating type frequency within the cluster.

%MTClones are those isolates that are identical to other isolates including the MAT.

%The percent of isolates within each mating type that are identical to another isolate in the cluster.

**Figure 4.** EBURST diagram showing clones with mating types. Mating types are shown for each clone composed of more than a single isolate. MAT1-1-1 isolates are labelled with a red 1, and MAT1-2-1 isolates are labelled with a blue 2. The four genotypes that have both mating types are labelled by Microsatellite Type (MT) number.

doi:10.1371/journal.ppat.1002851.g004
three loci. Putative parents or ancestral parents were defined as all isolates that together could complete the multi-locus microsatellite type (MLMT) of the recombinant genotype (Figure S5). This allows us to compare between observed distances of maximal observed recombination against a null hypothesis that any two isolates could recombine. We identified 11 potential recombinants with 54 possible parent genotype combinations. Geographic distance between putative parents was shorter, 382 km, and genetic similarity higher, 60.06% identical, than random potential parents drawn from the entire population, 675 km (p = 0.005) and 49.29% identical (p = 0.025) respectively.

The scale of recombination determines the efficacy of adaptation and the adaptive potential of populations. Although recombination across large distances allows generation of greater genetic diversity and more rapid spread of advantageous alleles, it disrupts locally advantageous combinations reducing local ecological genetic correlation. When sex is limited to small geographic distances it can reinforce local adaptation, and when limited to smaller genetic distances can reinforce genomic coadaptation. Together these effects can promote ecological speciation [57,58,59]. When ecological adaptation acts to reinforce genetic differentiation, strong correlations between key ecological factors and population distributions will exist [60,61,62,63,64]. To assess the possibility that ecological adaptation drives population differentiation in P. marneffei we used MAXENT [64] to predict overlap between the ecological niches of the genetic clusters (Fig. 5). Cluster 1, the cluster with the ratio of MAT alleles nearest 1:1, had the widest predicted range and overlapped with the entire predicted range of Cluster 3, including the predicted range that was not sampled in Myanmar. Cluster 1, Cluster 2, and Cluster 3 isolates are all found in bamboo rats, but 16 of 17 samples from Cannomys badius and 13 of 14 unique genotypes were from cluster 1 and distributed among India, Thai Central, Thai North, and Thai South sampling localities. None of the Cluster 1 isolates were among the 13 recovered from Rhizomys sumatrensis, which were all in Cluster 3. Cannomys badius is relatively more abundant in the western portion of the range of P. marneffei. The predicted distributions of bamboo rats were similar to the IUCN species ranges and had overlap with P. marneffei distribution. Although our spatial sample of Cluster 2 was geographically restricted it was entirely within the distribution of Rhizomys sumatrensis, a species that has been shown to consistently harbour clinically relevant P. marneffei [65]. The distributions for R. sumatrensis overlapped with Cannomys and Clusters 1 and 3 (Fig. 5). However, using ENMTools [66] to account for sampling error we found that Cluster 1 predicted distributions overlapped more with the Cannomys distribution than R. sumatrensis distribution, and Cluster 3 similarly overlapped more with R. sumatrensis than Cannomys distribution (S8). This observed range overlap supports a host specific effect on P. marneffei population structure.

Hosts may structure populations of pathogenic fungi in many ways, including by providing an environment in which recombination can occur and by acting as a selective filter on population genetic diversity [67,68]. We used a murine inhalation model of co-infection with genetically distinct strains to investigate the effect of host infection on P. marneffei (Text S2). Isolates of different mating types were used for experimental co-infection of 5 mice. Subsequent culture after 15 days from the livers showed a strong bias towards recovery of the MAT1-1 genotype for each of the mice. However, in two mice, genotypes of 4 isolates recovered from co-infections also revealed infrequent transfer of alleles between isolates of different mating type and genetic cluster (Table 4), suggesting that recombination may be possible across genetic barriers if multiple strains are within a host. In a smaller but similar in vitro experiment we did not observe significant bias towards MAT1-1, and from our scan of partial genotypes we did not recover any recombinants (S9, Table 4). We do not rule out regular recombination outside of hosts, but in the context of our spatial genetic evidence, the result of experimental infections indicate that hosts may play an important role in the development of sexual neighborhoods in populations of P. marneffei. However, the evolution of that role may involve restricted mating with or without host adaptation and remains to be explored.

Causes and consequences of sexually driven clonal structure

Asexual spores are common in vitro and likely a feature of natural P. marneffei populations, but sexual recombination may be an unexpectedly common occurrence in natural populations. The evidence supports the occurrence of recombination and perhaps even frequent sex, yet the natural populations remain strongly clonal and spatially structured. Although many mycologists might perceive this as a paradox because clonality is usually used as a proxy for asexual, many fungi, including key pathogens, also employ same clone mating or sibling mating [7,16,69]. Three key hypotheses could explain the perceived clonality in P. marneffei: 1) Spatially restricted dispersal keeps individuals in contact with only closely related individuals; 2) Genetic incompatibility between dissimilar individuals restricts sex to genetically similar individuals; 3) Local adaptation restricts the ability of dissimilar genotypes to penetrate habitats ensuring mating between genetically similar individuals. All three are likely to be partially correct. Although the genetic evidence shows spatial limitations to effective dispersal, the physical dispersal of airborne conidia is not likely to be a limiting factor, and four genetically identical clones are dispersed across distances over 800 km. We have little information about the effect of genetic similarity on mating success in P. marneffei, but genetic restrictions on successful recombination are present in some plant pathogens [63,70,71] and should not be completely discounted. Local adaptation is not fully supported by ecological niche models that show overlap between distinct genetic clusters, but there is limited evidence of host specialisation. A key question unanswered in all of these hypotheses is why have sex at all?

Previous work has focused on the consequences of selectively neutral loss of sex in P. marneffei [72], but the persistence of a sexual cycle in P. marneffei despite abundant asexual reproduction in the lab suggests that there is a selective advantage for sex not associated with the advantages of greater adaptive potential provided by outcrossing. One major consequence of sexual clonality is release from Muller’s Ratchet compared to asexual [51]. Large numbers of haploid offspring and wide dispersal maximize environmental exposure of genes and increase the efficiency of purging deleterious alleles [73,74]. However, the sexual process itself may also reduce the accumulation of deleterious mutations independent of recombinational effects [75]. Among the close relatives of P. marneffei in the subgenus Biverticillium, outcrossing has not been shown to occur, but self-fertility is common [76,77], and the distribution of mating systems in the subgenus suggests that inbreeding may not reduce the evolutionary longevity of this group [78].

Another compelling scenario favouring sex recognises the opportunity presented by mating itself for dramatic shifts in clonal structure. Local adaptation is not fully supported by ecological niche models that show overlap between distinct genetic clusters, but there is limited evidence of host specialisation. A key question unanswered in all of these hypotheses is why have sex at all?
withstand stress by mating when it would otherwise not survive. If that were the case, recombination in *P. marneffei* might be clustered in space or time where or when stress occurs. Unfortunately, little is known about the natural ecology of *P. marneffei*, and any conditions that might allow mating to occur are unknown. Isolates are commonly recovered from bamboo rats, yet the epizology of the fungus is poorly known including unknown routes of infection and unknown course and outcomes of the zoonosis. In other dimorphic fungi including *Histoplasma* and *Blastomyces*, mating in natural populations is also poorly known, but in these species it mating has long been studied outside of the host and at lower temperatures in vitro [83,84]. Nevertheless, the association with small mammals may be the best starting point for a search for the natural sexual niche of *P. marneffei*.

Cryptic mating and inbreeding in *P. marneffei* has some parallels with other fungal and non-fungal eukaryotic pathogens [7]. There is growing support for high inbreeding in addition to asexual reproduction in *Leishmania brasiliensis* [85]. Experimental data support a role for within-vector recombination and it is thought that this restricted recombination in *Leishmania* results in sexual reproduction. If that were the case, recombination in *P. marneffei* might be clustered in space or time where or when stress occurs. Unfortunately, little is known about the natural ecology of *P. marneffei*, and any conditions that might allow mating to occur are unknown. Isolates are commonly recovered from bamboo rats, yet the epizology of the fungus is poorly known including unknown routes of infection and unknown course and outcomes of the zoonosis. In other dimorphic fungi including *Histoplasma* and *Blastomyces*, mating in natural populations is also poorly known, but in these species it mating has long been studied outside of the host and at lower temperatures in vitro [83,84]. Nevertheless, the association with small mammals may be the best starting point for a search for the natural sexual niche of *P. marneffei*.

Cryptic mating and inbreeding in *P. marneffei* has some parallels with other fungal and non-fungal eukaryotic pathogens [7]. There is growing support for high inbreeding in addition to asexual reproduction in *Leishmania brasiliensis* [85]. Experimental data support a role for within-vector recombination and it is thought that this restricted recombination in *Leishmania* results in sexual reproduction.
neighbourhoods of pathogen genotypes with high differentiation at multiple spatial scales [86]. In the malaria parasite *Plasmodium falciparum*, high inbreeding has been linked to faster emergence of drug resistance in some low infection intensity regions, but inbreeding has also emerged as a general property of *P. falciparum* populations regardless of infection intensity [87,88]. In *Toxoplasma gondii* and *Sarcocystis neurona*, ‘clonal’ emergence is enabled by high rates of selfing and results in spatially structured populations [89,90]. The most common fungal infection of humans, *Candida albicans*, also undergoes same-sex mating that can facilitate inbreeding and the ability to accelerate evolution via sex within clonal populations [91]. In *C. neoformans*, multiple ecological niches where recombination between strains with opposite mating type occurs have been found [92,93]. However, most mating in *C. neoformans* globally is likely to occur between strains of the same mating type, and although there is as yet no indication of what factors alter the probability of this kind of inbreeding across natural populations, it may have facilitated the global emergence of a single mating type and highly clonal populations [15,22,94,95]. Although some of the mechanisms underlying these inbreeding eukaryotic pathogens remain mysterious and likely differ between organisms, there is an emerging consistent pattern of clonality resulting from inbreeding rather than strictly asexual propagation even in the absence of a recognized sexual stage.

**Supporting Information**

**Figure S1** DAPC cluster analysis. (PDF)

**Figure S2** Dispersal kernel effects on spatial genetic correlation. (PDF)

**References**


Author/s:
Henk, DA; Shahar-Golan, R; Devi, KR; Boyce, KJ; Zhan, N; Fedorova, ND; Nierman, WC; Hsueh, P-R; Yuen, K-Y; Sieu, TPM; Nguyen, VK; Wertheim, H; Baker, SG; Day, JN; Vanittanakom, N; Bignell, EM; Andrianopoulos, A; Fisher, MC

Title:
Clonality Despite Sex: The Evolution of Host-Associated Sexual Neighborhoods in the Pathogenic Fungus Penicillium marneffei

Date:
2012-10-01

Citation:

Persistent Link:
http://hdl.handle.net/11343/264368

File Description:
Published version

License:
CC BY