

Article

Not peer-reviewed version

Mitogenome Variations in a Global Population of *Aspergillus fumigatus*

Veronica Thorn and [Jianping Xu](#) *

Posted Date: 15 September 2023

doi: 10.20944/preprints202309.0977.v1

Keywords: human pathogen; aspergillus; mitochondria; single nucleotide polymorphism; clonal expansion; recombination; copy number variation



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Mitogenome Variations in a Global Population of *Aspergillus fumigatus*

Veronica Thorn¹ and Jianping Xu^{1,*}

¹ Department of Biology, McMaster University, Hamilton, ON, L8S 4K1, Canada; thornv2@mcmaster.ca (V.T.); jpxu@mcmaster.ca (J.X.)

* Correspondence: jpxu@mcmaster.ca

Abstract: *Aspergillus fumigatus* is a ubiquitous, critical priority human fungal pathogen. Despite its clinical importance, there is limited knowledge regarding the variations of the genome within mitochondria, the powerhouse organelle within eukaryotic cells. In this study, we leveraged publicly available raw, whole genome sequence data from 1939 isolates to investigate the variations in the mitochondrial genomes of *A. fumigatus*. These isolates were isolated from 22 countries on six continents, as well as from outer space and from within the International Space Station. In total, our analysis revealed 39 mitochondrial single nucleotide polymorphisms (mtSNPs) within this global sample and together, these 39 mtSNPs grouped the 1939 isolates into 79 mitochondrial multilocus genotypes (MLGs). Among the 79 MLGs, 39 were each distributed in at least two countries and 30 were each shared by at least two continents. The two most frequent MLGs were also broadly distributed: MLG11 represented 420 isolates from 11 countries and four continents and while MLG79 represented 418 isolates from 18 countries and five continents, consistent with long-distance dispersals of mitogenomes. Our population genetic analyses of the mtSNPs revealed limited differentiation among continental populations but highly variable genetic differences among national populations, largely due to localized clonal expansions of different MLGs. Phylogenetic analysis and Discriminant Analysis of Principal Components of mtSNPs suggested the presence of at least three mitogenome clusters. Linkage disequilibrium, Index of Association, and phylogenetic incompatibility analyses collectively suggested evidence for mitogenome recombination in natural populations of *A. fumigatus*. In addition, sequence read depth analyses revealed an average ratio of ~20 mitogenomes per nuclear genome in this global population but the ratios varied among strains within and between certain geographic populations. Together, our results suggest evidence for organelle dynamics, genetic differentiation, recombination, and both widespread and localized clonal expansion of the mitogenomes in the global *A. fumigatus* population.

Keywords: human pathogen; aspergillus; mitochondria; single nucleotide polymorphism; clonal expansion; recombination; copy number variation

1. Introduction

Aspergillus fumigatus is a ubiquitous airborne saprophytic fungus, and a leading cause of human fungal infections [1]. Playing an important role in the carbon and nitrogen cycles, *A. fumigatus*' ecological niche is primarily in organic debris and soil [2]. However, they can grow in a broad range of environmental conditions including indoor, outdoor, and aqueous and even extraterrestrial environments such as the International Space Station (ISS) [3,4]. Reproducing mainly asexually, *A. fumigatus* spreads via sporulation of minute conidia, roughly 2 – 3 µm in diameter, light enough for airborne dissemination and small enough to penetrate lung alveoli [5]. The combination of widespread environmental habitats and airborne spores leads to a high frequency of incidental inhalation every day for the majority of people throughout the world [6].

Despite the near constant inhalation of spores, a healthy host's innate immune system is usually able to eliminate the foreign cells, such that clinical symptoms are mild and rarely seen [6]. However, in immunocompromised patients, such as transplant recipients, leukemia patients, and those with acquired immunodeficiency syndrome (AIDS), inhaled *A. fumigatus* spores often develop into invasive aspergillosis (IA), with mortality rates reaching as high as 90% [5]. Consequently, the World

Health Organization (WHO) has designated *A. fumigatus* a Critical Priority fungal pathogen [7]. Antifungal treatment is essential for effectively clearing the fungal pathogen in acute aspergillosis cases [6]. However, shared fundamental metabolic pathways between fungal and human cells limit the availability of antifungal drug targets and treatment options. Furthermore, the emergence of drug-resistant strains in both environmental and clinical settings has contributed to a noticeable increase in mortality rates, surpassing non-resistant infections by more than 25% [8,9].

For effective growth and reproduction both in vitro and in vivo, like most other eukaryotes, *A. fumigatus* requires energy generated by mitochondria, through oxidative respiration. In addition, aside from energy generation, fungal mitochondria have also been implicated in various cellular processes, including temperature and oxidative stress responses, pathogenicity, antifungal resistance, iron homeostasis, and fungal dormancy [10–15]. In most fungal cells, there may be multiple mitochondria, with each containing one or more mitochondrial genome (mitogenome) that replicates independently of the nuclear genome and of the cell cycle [16,17]. Typically, each mitogenome is a single circular DNA molecule. The core gene content in fungal mitogenomes is relatively well conserved among species [18]. However, among species, the mitogenome sizes can vary significantly, from about 20kb to over 200kb [17,19,20]. Among strains within the same species, variations in mitogenome size and gene content have also been observed [17]. The inheritance of mitochondria in most eukaryotic organisms is uniparental, typically maternal, following a non-Mendelian pattern [16]. However, other modes of inheritance have been observed in fungi (for an in-depth review, see [17]). At present, the mode of mitochondrial inheritance in *A. fumigatus* and potential variations among strains in their mitogenomes of this fungal pathogen are unknown.

Since mitogenomes are small, relatively abundant, and well conserved yet often distinct among species, mitogenome analysis is a popular method for population genetics, physiological, and evolutionary studies [21–24]. Indeed, for most animals the mitochondrial cytochrome c oxidase 1 gene (COI) has been adopted as a DNA barcode, a standard for species identification in many animal taxa [25]. In fungi, however, COI is less commonly used due to the occurrence of introns within the gene and limited sequence divergence among species in certain fungal lineages [26]. Nevertheless, due to their small size and relatively high abundance within each cell, mitogenomic analysis could provide insights into fungal evolution, population structure, and reproduction at a fraction of the cost of conventional nuclear genomic sequencing [19].

In recent years, there has been a notable increase in the number of studies focusing on mitochondrial genomics. However, the application of mitochondria-based analyses to investigate fungal population structure, particularly within the *Aspergillus* species, is not well represented in the literature. Furthermore, investigations into the intraspecies structure or whole-genome level analysis of mitogenomic single nucleotide polymorphisms (mtSNPs) in *A. fumigatus* are currently nonexistent. Previous studies examining interspecies population dynamics and phylogenetics in the *Aspergillus* genus have predominantly relied on single gene such as cytochrome b or concatenated alignments of multiple mitochondrial genes, most of which were conducted more than a decade ago. Notably, Kozłowski and Stepień [27] pioneered the use of mitochondrial restriction enzyme analysis to establish interspecies phylogenetic relationships within the *Aspergillus* genus. Subsequent studies by Wang et al. [28,29] successfully investigated the phylogenetic relationships between *A. fumigatus* and other pathogenic *Aspergillus* species using cytochrome b gene sequences. More recently, Joardar et al. [19] published the first complete mitogenomes of nine *Aspergillus* and *Penicillium* species, including employing 14 concatenated mitochondrial proteins for inter-species and inter-genus phylogenetic analyses. Although relatively less common than single gene focused assays, genome-wide mtSNPs for phylogenetic and population analyses has been successfully demonstrated in studies on crops, cattle, humans, and selected fungi [30–33].

With the increasing abundance and availability of whole genome sequencing (WGS) reads in *A. fumigatus*, we conducted a large-scale intraspecific analysis of mtSNPs in a global *A. fumigatus* population. Specifically, raw sequencing reads of about 2000 isolates of *A. fumigatus* were available (as of May 2023) in the National Centre for Biology Information's (NCBI) Sequence Read Archive (SRA), each including both the nuclear and mitochondrial sequence reads. While the nuclear

genomes of these strains have been analysed and published in a diversity of papers and journals [34–38], the accompanying mitochondrial DNA (mtDNA) sequences have received limited attention. Furthermore, due to the higher abundance of mtDNA molecules than nuclear chromosomal DNA in the cell and the nature of gene sequencing, the sequence read depths of mtDNA are generally high, enabling accurate SNP calls [19]. Our study aimed to identify mtDNA sequence variation, examine the potential geographic patterns of the observed variation, and investigate evidence of recombination in shaping mitogenome diversity in this clinically significant human fungal pathogen.

2. Materials and Methods

2.1. Data Acquisition and Processing

NCBI's SRA, which includes raw data from published experiments, the European Bioinformatics Institute (EBI), and the DNA Database of Japan (DDBJ), served as the data source and repository for this study. A query was performed on the SRA to identify all accessions associated with "*Aspergillus fumigatus*" and Whole Genome Sequencing (WGS) and Whole Genome Amplicon (WGA) assay types in May 2023. The raw sequence data for each accession was obtained using the sra-toolkit (v3.0.0), and subsequently subjected to quality checking and trimming using Fastp (v0.23.1) [39].

To ensure data integrity, reads were filtered based on quality (requiring at least 60% of bases with a phred score of 30 or higher), length (minimum of 30 base pairs), and complexity. The filtered reads were aligned to a reference *A. fumigatus* mitochondrial genome (strain Af293, GenBank JQ346808.1) using Burrows-Wheeler Aligner (BWA-MEM v0.7.17-r1188) [40].

The alignments were further processed using samtools (v1.17) [41] to ensure quality, length, and alignment depth. Strains with alignments with an average sequence depth below 10x were considered of poor quality and excluded from further analysis. Duplicate reads were removed using MarkDuplicates in the Picard tool (v2.26.3) and genetic variants were called using GATK tools, including HaplotypeCaller (with the '-ERC GVCF' flag), GenomeDBImport, and GenotypeGVCFs (v4.2.5.0) [42]. The resultant multisample VCF file was subjected to additional filtering using bcftools (1v.16) [41] to remove insertions/deletions.

Country and continent of origin for each strain was initially determined from associated metadata collected from NCBI. In cases where metadata was incomplete, a manual literature search was conducted to confirm the origin information. Accessions with unconfirmed or artificial origins were labelled as having "Unknown" geographic origins. All isolates retrieved and analyzed, including their relevant metadata are shown in Supplementary Table S1

2.2. Data Analysis

2.2.1. Phylogeny

To investigate the relationships among strains and mitogenomes, we conduct phylogenetic analysis using concatenated SNPs. The SNPs were aligned using the program vcf2phylip (v2.0) [43] and the alignment was used to construct a neighbour joining phylogenetic tree using the R programming language (v4.2.3) of the fasttreeR package (v1.2.0) [44]. The tree was rooted using the ape package (v5.7) [45] with an *A. fischeri* isolate as the outgroup. The Interactive Tree Of Life (iTOL) web viewer (v5) [46] was used to visualize the resulting tree and incorporate sample metadata.

2.2.2. Population Genetic Analysis

To investigate the potential influence of geographic isolation on mitogenomic population structure, population genetic analyses were performed using the R statistical software and the R package poppr (v.2.9.4) [47]. The first step involved calculating basic diversity statistics and examining the distribution of unique multi-locus genotypes (MLGs) based on mitogenome SNPs present in the entire population. From the initial multisample VCF, four distinct datasets were created to facilitate subsequent analyses. These datasets included: (1) the full dataset without the *A. fischeri* root, (2) a clone-corrected (CC) full pooled dataset, (3) a dataset excluding countries with fewer than nine

samples and samples with unknown or non-geographic origins, and (4) a country wise CC dataset where only one representative for each unique MLG type from each country was included.

Population structure was explored by performing a Discriminant Analysis of Principal Components (DAPC) on both the full and country wise clone corrected datasets. For both datasets, Kernel (K) values were chosen from where the Bayesian Information Criterion (BIC) began to decrease minimally, based on 10 repetitions using the *find.clusters* function from the *adeigenet* R package (v2.1.10) [48]. DAPC was then applied to the dataset clustered using the chosen K values using the *dapc* function from the *adeigenet* R package [48]. For the full dataset, a K of three was chosen and both three and four were chosen for the country wise clone corrected dataset for comparison.

The impact of geography on genetic distribution was further explored through an Analysis of MOlecular VAriance (AMOVA). The country filtered and country wise CC datasets were subjected to AMOVA at both the country and continental levels. Additionally, pairwise Nei's *F_{st}* values were calculated between pairs of countries to assess the genetic differentiation among them. To evaluate the correlation between geographic and genetic distances, a Mantel test was performed using the non-CC filtered dataset. This analysis involved comparing the geographic distances (as measured by latitude and longitude) with the corresponding genetic distances (measured as pairwise Nei's *F_{st}* in this implementation) between countries. Significance testing was conducted using 999 permutation bootstrap iterations, with the null hypothesis assuming no genetic differentiation among the geographic populations.

2.2.3. Recombination and Linkage Disequilibrium

Three tests were conducted to assess the presence of recombination in the mitogenome of *A. fumigatus*. The first test employed the common multilocus Index of Association (*I_A*) [49,50] to investigate allele associations across different loci. To determine the statistical significance of the observed *I_A*, we generated 999 artificially recombined datasets through random allele permutation while maintaining allelic proportions at each locus, and evaluated each for *I_A*. A significantly higher observed *I_A* compared to the randomly recombined permutations would indicate a lack of random recombination in the dataset, leading to the rejection of the null hypothesis of panmixia (i.e., no association of alleles between different loci). To facilitate comparisons between studies, we standardized *I_A* by dividing it by the number of loci analyzed, yielding the measure *rD*. *I_A* and *rD* calculations were performed on the clone-corrected dataset, where only one representative of each unique MLG was included. These calculations were conducted using the R package *poppr* [47].

To further investigate the potential recombination patterns in the mitogenome, the second test investigated allelic associations between all pairs of SNP loci across the mitogenome, using the clone-corrected pooled dataset. For each loci pair, both the coefficient (D) of linkage disequilibrium (LD) and the correlation coefficient (*r*²) were computed, with statistical significance assessed through a χ^2 test against the null hypothesis of panmixia (i.e., no association of alleles between different loci). Considering the extensive number of locus pairs, p-values from the pairwise χ^2 tests were adjusted using the Bonferroni method. Both LD calculations and χ^2 analyses were performed in R.

The third test complimented the LD analysis as it focused on determining the proportion of phylogenetically compatible pairs of loci using the Four Gamete Test between pairs of SNP loci. Phylogenetic compatibility of two loci is established when it is possible to account for all observed genotypes in the sample without inferring homoplasy (reversals, parallelisms, or convergences) or recombination [23]. For biallelic SNP loci observed in this study, a maximum of four possible haploid genotypes per pair of loci can exist. If no more than three of the four genotypes are observed in the sample, all combinations can be explained without homoplasy or recombination, indicating phylogenetic compatibility. However, when all four possible genotypes are present, the presence of all four cannot be explained without invoking homoplasy or recombination, signifying phylogenetic incompatibility [23]. The analysis was performed in R, and bootstrap procedures were manually implemented with the assistance of the *shufflepop* function in the *poppr* package [47]. Visualizations for both LD and the Four Gamete test were generated using the *corrplot* R package [51].

2.2.4. Mitochondria to Nuclear Genome Comparison

To assess the relative abundance of the mitogenome and the nuclear genome within a cell, we used two methods. In the first, we obtained the average read depth of a single-copy gene, *cyp51A* (GeneID: 3509526), to represent the nuclear genome and compared it with the average read depth of the mitogenome. The *cyp51A* read depth for each strain was obtained by aligning the adapter-trimmed and quality-filtered reads from each sample to the reference genome Af293 using the same methodology as described earlier. The entire mitogenome sequence was used to obtain the average read depth of the mitogenome for each strain. The read depth for each nucleotide position for both the *cyp51A* gene and the mitogenome in each strain was determined using samtools depth (v1.17) [41], incorporating the unmapped positions through the inclusion of the ‘-a’ flag. This approach ensured that even the unmapped positions were accounted for in the read depth calculation for the specific gene fragments. The resulting read depths were then averaged across all nucleotide positions, allowing for a more accurate estimation of the average depth by considering both mapped and unmapped sites. For each sample, we calculated the mitogenome to nuclear genome ratio by dividing the average mitogenome read depth by the *cyp51A* read depth.

In the second method, we estimated the nuclear genome read depth based on whole-genome sequencing coverage. To estimate coverage, the total count of bases in each sequence file that remained post quality filtering and trimming was divided by the aggregate genome size of the *A. fumigatus* reference strain, Af293 (GenBank: GCA_000002655.1), minus that of the mitogenome. The obtained coverage was then compared with the mitogenome coverage described in method #1 above. The obtained ratios were then analyzed for their potential phylogenetic and geographic patterns.

3. Results

In this study, we retrieved the raw DNA sequence reads of all the strains of *A. fumigatus* that had their whole-genome sequences deposited in the SRA by May 2023, and one *Aspergillus fischeri* sample used as an outgroup for later phylogenetic analysis. The retrieved 1939 strains came from 22 countries representing six continents, as well as from outer space (2 isolates) and from within the International Space Station (2 isolates) (Table 1). At the continental level, 53.43% of the isolates were from Europe, followed by 29.55% from North America, 9.33% from Asia, 1.24% from Oceania, 0.46% from Africa, and 0.1% from South America. About 5.9% of the isolates didn’t have any associated geographic information (Table 1). At the country level, USA had the highest number of isolates sequenced, followed by the Netherlands, Germany, United Kingdom, France, Japan, and Ireland, with the remaining 16 countries each contributed less than 30 isolates (Table 1).

Table 1. Summary geographic distribution of the *A. fumigatus* isolates analyzed in this study. Two geographic levels are shown: the country level and the continental level. For each region, total samples (N), percentage of total dataset (%), and number of both private and total mitogenome multi-locus genotypes (MLGs) are reported.

| Region | N | % | Private MLG | Total MLG | Region | N | % | Private MLG | Total MLG |
|---------------|-----|-------|-------------|-----------|---------------|------|--------|-------------|-----------|
| Africa | 9 | 0.46% | 0 | 2 | Europe | 1036 | 53.43% | 20 | 56 |
| Cote d'Ivoire | 9 | 0.46% | 0 | 2 | Austria | 2 | 0.10% | 0 | 1 |
| Asia | 181 | 9.33% | 4 | 24 | Belgium | 10 | 0.52% | 0 | 5 |
| China | 10 | 0.52% | 0 | 4 | France | 161 | 8.30% | 3 | 26 |
| India | 12 | 0.62% | 1 | 3 | Germany | 262 | 13.51% | 7 | 30 |
| Iran | 2 | 0.10% | 2 | 2 | Ireland | 72 | 3.71% | 3 | 17 |
| Japan | 155 | 7.99% | 1 | 18 | Netherlands | 282 | 14.54% | 1 | 17 |
| Singapore | 1 | 0.05% | 0 | 1 | Portugal | 8 | 0.41% | 0 | 5 |
| Thailand | 1 | 0.05% | 0 | 1 | Spain | 28 | 1.44% | 4 | 14 |

| | | | | | | | | | |
|----------------------|-----|--------|----|----|----------------------|-------------|----------------|-----------|-----------|
| Oceania | 24 | 1.24% | 0 | 5 | Sweden | 1 | 0.05% | 0 | 1 |
| New Zealand | 24 | 1.24% | 0 | 5 | United Kingdom | 210 | 10.83% | 2 | 25 |
| North America | 573 | 29.55% | 13 | 39 | South America | 2 | 0.10% | 0 | 2 |
| Canada | 10 | 0.52% | 1 | 6 | Brazil | 1 | 0.05% | 0 | 1 |
| USA | 563 | 29.33% | 12 | 38 | Peru | 1 | 0.05% | 0 | 1 |
| Space | 2 | 0.10% | 0 | 2 | Unknown | 112 | 5.78% | 3 | 18 |
| ISS | 2 | 0.10% | 0 | 2 | Grand Total | 1939 | 100.00% | 40 | 79 |

3.1. Distribution of Multilocus Genotypes Based on Mitogenome SNPs

Using the mitogenome sequence of Af293 as reference and the retrieved sequence reads from the 1939 isolates, we identified a total of 39 SNP sites in this global collection of isolates. These 39 SNPs enabled identification of 79 mitogenome multilocus genotypes. Similar to the sample size pattern, at the continental level, the largest number of MLGs were found in Europe, followed by North America, Asia, Oceania, Africa, and South America. At the country level, the highest number of MLGs were from the USA, followed by Germany, France, the UK, Japan, Ireland, the Netherlands, and Spain. The remaining 14 countries each had less than 10 mitogenome MLGs in the SRA.

Among the 79 MLGs, 39 were shared by at least two countries and 30 were shared by at least two continents. The remaining 40 were each found in only one country so far and 35 of these were only represented by one isolate each (Figure 1). The two most frequently shared MLGs were broadly distributed, representing 420 isolates from 11 countries and 418 isolates from 18 countries, respectively (Figure 1). The remaining MLGs were relatively less frequent. However, there were 19 other MLGs each shared among at least three countries by 10 or more isolates.

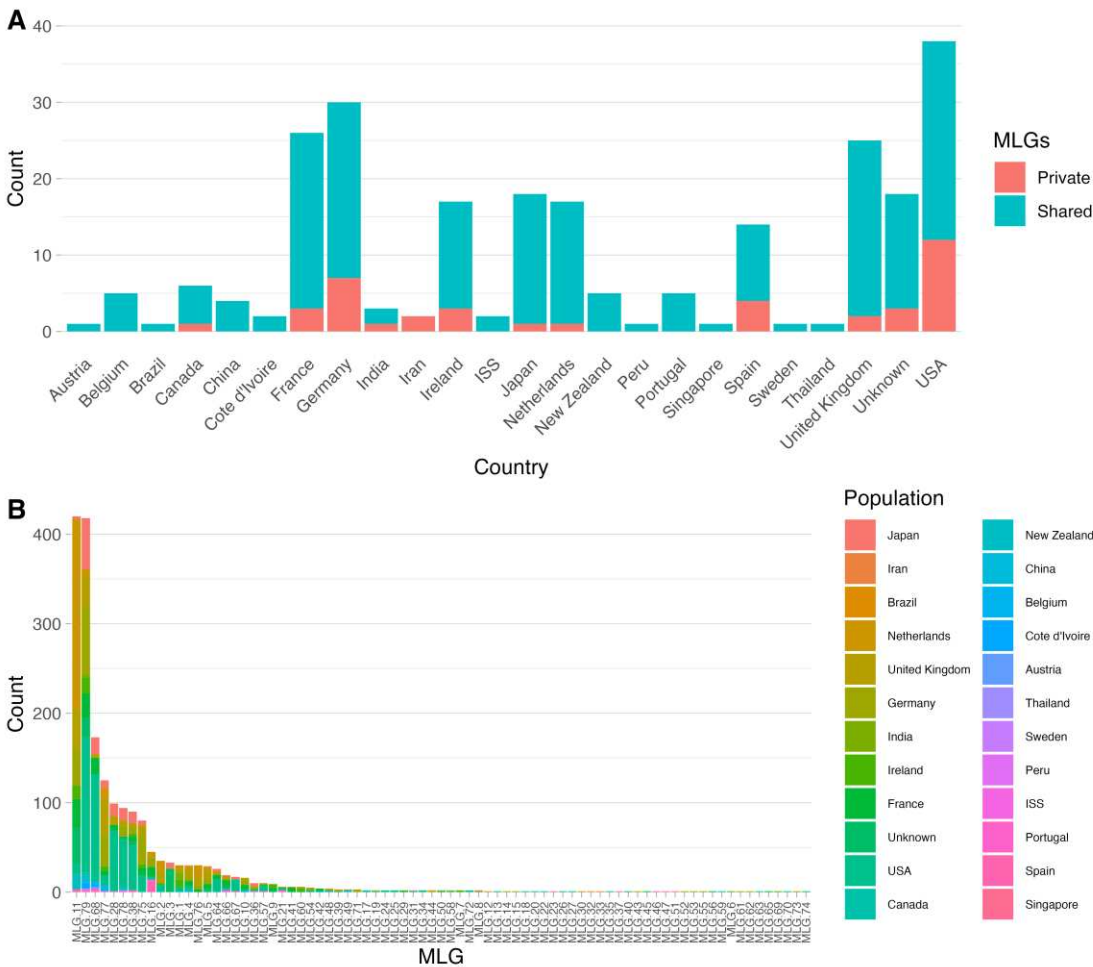


Figure 1. Geographic distributions of mitochondrial multilocus genotypes. (A) Stacked bar graph displaying the counts of private and shared MLGs per region at the “country” level for each region included in the total dataset. (B) Stacked bar graph displaying the geographic distribution of each of the 79 MLGs. Regions at the “country” level are denoted by colour, and the y-axis shows the raw strain count for each MLG in the total dataset.

3.2. *Phylogenetic Relationships among Strains and MLGs Based on Mitogenome SNPs*

The relationships among strains and MLGs based on mitogenome SNPs are shown in Figure 2. This neighbour-joining phylogenetic tree was constructed from the concatenated 39 SNPs. Overall, a few country- and/or continent-based small clusters of strains and MLGs were found (Figure 2). For example, small clusters containing two or more closely related MLGs were found within several countries such as the US, Germany, Japan, Spain, UK, etc. However, the overall phylogeny revealed that all major clusters of MLGs contained strains and MLGs from different countries and/or different continents. Taken together, the results are consistent with frequent gene flow of mitogenomes among geographic regions but that there are unique mitochondrial genotype(s) within many geographic regions.

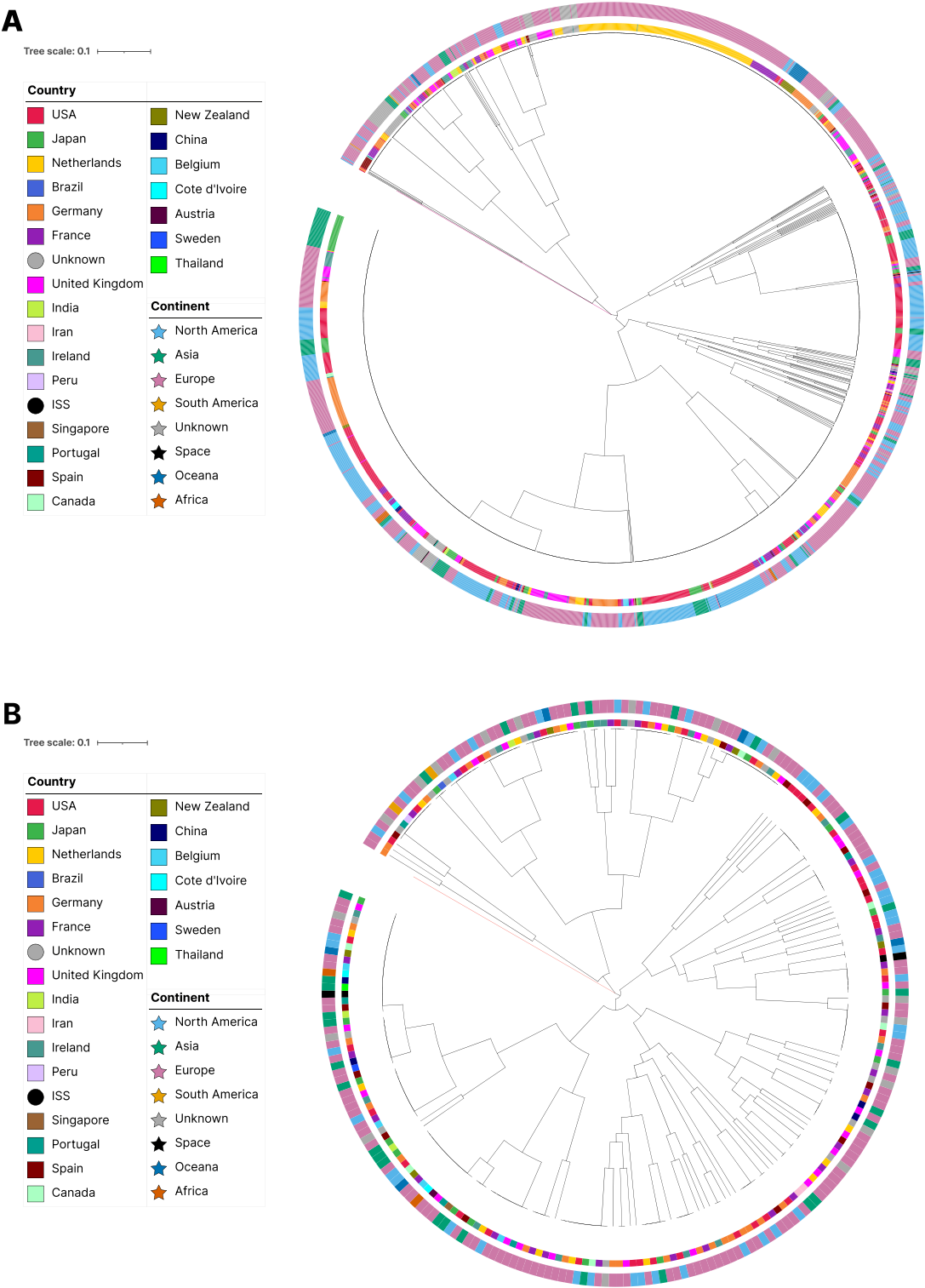


Figure 2. Neighbour joining phylogenetic trees showing relationships among *A. fumigatus* MLGs and strains. (A) all 1939 samples were included, with *A. fisherii* serving as outgroup, with its branch highlighted in red. (B) Only MLGs after clone correction at the country level are shown (n=244). In both A and B, the inner ring denotes the country of origin and the outer ring denotes continent of origin.

3.3. Genetic Clusters of the Global Mitogenome MLGs

We analyzed whether the 79 MLGs could be grouped into distinct genetic clusters. To accomplish this, we used the Discriminant Analysis of Principal Components (DAPC) of pairwise distances among MLGs. The DAPC analysis showed a clustering pattern very similar to that observed in the cladograms, with three groups appearing as the most parsimonious. However, there were four cases of incongruence between the DAPC grouping of three genetic clusters and the NJ tree, highlighted by different colors shown in Figure 3.

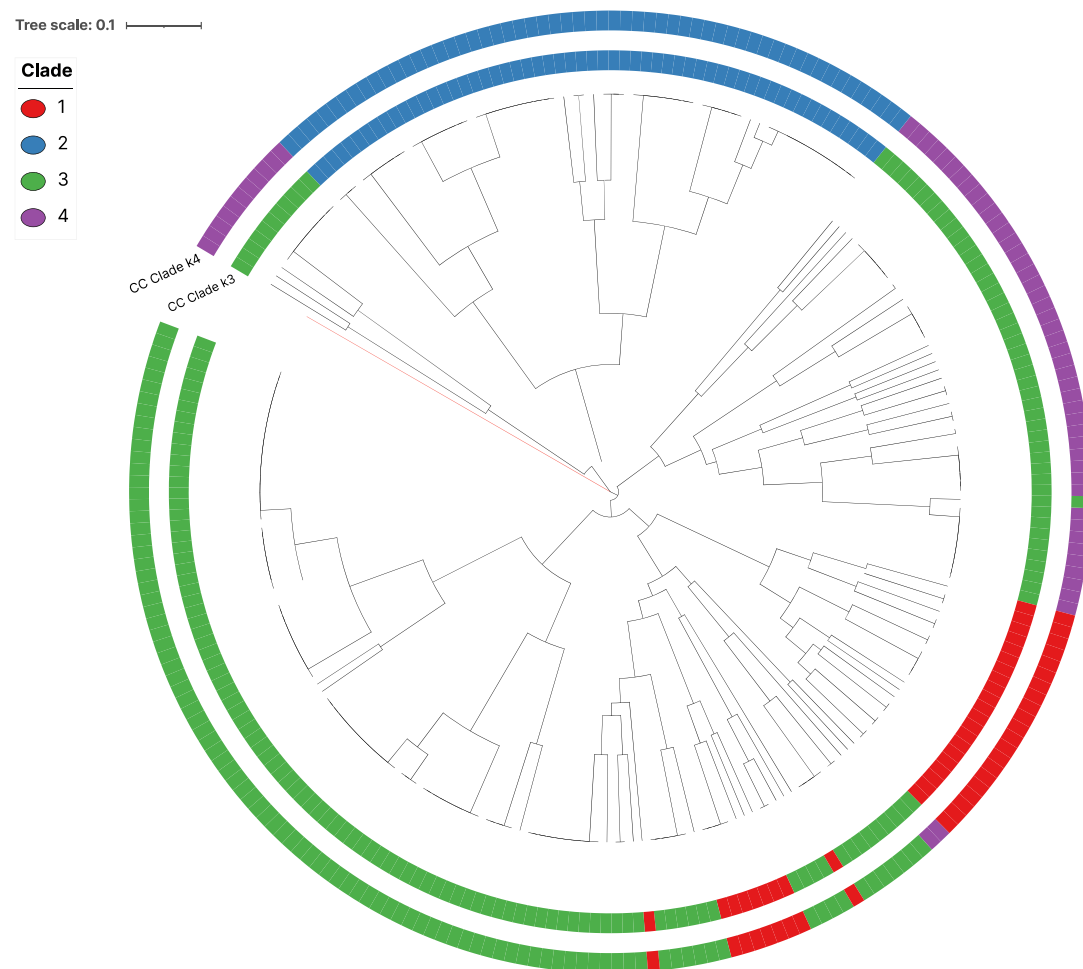


Figure 3. Country wise clone-corrected (CC) phylogenetic tree with DAPC analysis groupings. Different colors represent different clade assignment. The two rings were generated using two different DAPC parameters. The inner ring was generated using the kernel $k = 3$. The outer ring was created using the kernel $k = 4$.

The Bayesian Information Criterion (BIC) plot exhibited a sharp drop from a k value of one to two, followed by a less steep decline from two to three, and three to four (Figure 4.A). Subsequent increases in k values showed minimal differences and less pronounced drops (Figure 4.A). Plotting the first and second loadings for each grouping value between two and five revealed distinct clusters for a k value of three (Figure 4.B). Increasing the k value from three to four resulted in cluster #3 being split into two clusters, #3 and #4 (Figure 4.C-D). However, according to the NJ tree, with $k=4$, one MLG, MLG 74, belonging to cluster #3 showed a closer phylogenetic relationship with cluster #4 MLGs than with the remaining cluster #3 MLGs. Similarly, two MLGs, MLGs 32 and 33, were categorized as cluster #4 despite a closer phylogenetic relationship with clusters #3 and #1.

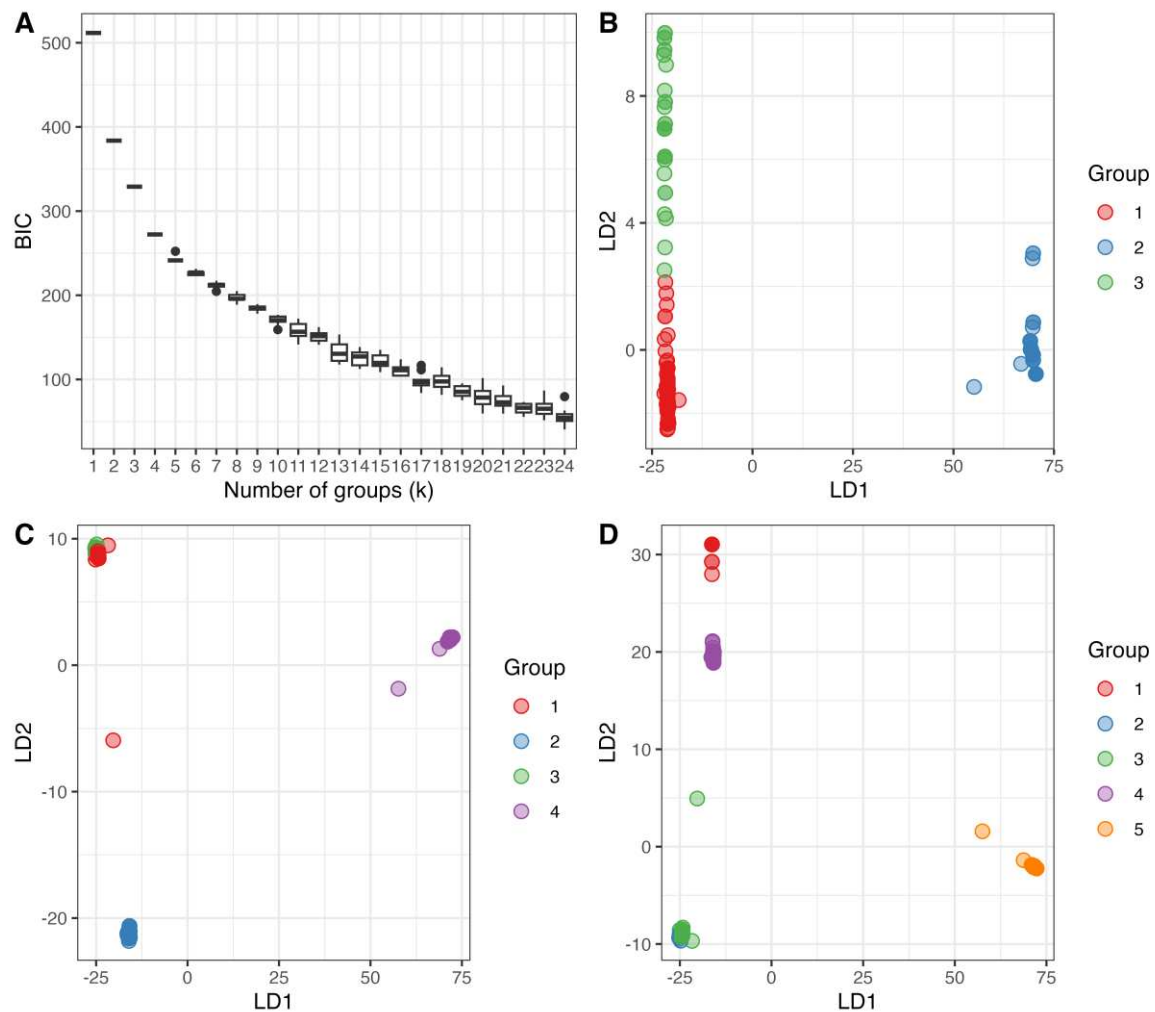


Figure 4. Bayesian Information Criterion (BIC) boxplot and Discriminant Analysis of Principal Components (DAPC) scatterplots based on mitogenome SNPs. (A) BIC boxplot for each putative grouping kernel (k) value for the country wise clone corrected dataset. BIC values were obtained from 10 repetitions of the *find.clusters* function in *r*, which applies successive K-means on Principal Component Analysis (PCA) transformed data then measures for goodness of fit (BIC) of the model. (B) Scatter plot of the first two Discriminant Analysis of Principal Components (DAPC) loadings of the country wise clone corrected dataset when k=3. Despite groups #1 and #3 clustering close to each other, there is little overlap between them. (C) Scatter plot of the first two DAPC loadings of the country wise clone corrected dataset when k=4. Noticeable overlap between groups #1 and #3 can be seen in the top left corner of the plot. (D) Scatter plot of the first two DAPC loadings of the country wise clone corrected dataset when k=5. Noticeable overlap between groups #2 and #3 can be seen in the bottom left corner of the plot.

3.4. Geographic Structure

To investigate the geographic structure of mitogenome SNPs and estimate the potential quantitative relationships between geographic distance and genetic distance among our samples, we conducted AMOVA at the continental and country levels and the Mantel analyses at the country level.

AMOVA was conducted using three different datasets: (1) the non-clone-corrected dataset including all countries with sample sizes greater than eight, (2) the same dataset after clone correction on a country wise basis, and (3) the country wise clone corrected dataset but included only countries with more than nine MLGs. The AMOVA test of the first dataset revealed a statistically significant genetic differentiation among countries, but the differentiation was statistically insignificant among continents. However, after clone correction, no statistically significant genetic differentiations were

observed at either the country or the continental levels for both the second and third dataset. Taken together, the results suggested that localized clonal expansion of mitogenome MLGs were responsible for the observed statistically significant genetic differentiations in the total samples among countries.

To further identify which pairs of national *A. fumigatus* samples were genetically differentiated, we obtained F_{st} values between pairs of national samples using both the non-clone-corrected data and the clone-corrected data. Similar to the analyses above, only countries with an initial sample size exceeding eight for the original dataset (i.e., dataset #1 above) and exceeding eight MLGs for clone-corrected data (i.e., dataset #3 above) were analyzed. The preliminary pairwise comparison revealed a considerable and diverse distribution of F_{st} values between countries (Figure 5). Among the 91 paired national samples in the non-clonal-corrected dataset, 51 pairs showed significant ($p < 0.01$, 999 permutations) differentiation. The two pairings that resulted in the highest F_{st} values both included Cote d'Ivoire. The two highest F_{st} values, 0.7538 and 0.7308, were observed between Cote d'Ivoire and the Netherlands and New Zealand, respectively. However, after clone correction, there was a substantial reduction in variability among country pairs. In addition, six of the initial 14 countries assessed failed to meet our sample size criteria for comparison, and thus only eight countries were compared (Figure 5). As shown in Figure 5, clone-correction led to the reduction in pairwise F_{st} values across the board and none of the observed differences between countries were statistically significant in the clone-corrected samples. Consistent with the observations of frequent gene flows, Mantel test revealed slightly positive but no statistically significant correlation between geographic distance and population genetic differentiation between pairs of national populations (the Mantel r statistic was 0.109 and a p -value of 0.271 based on 999 bootstrap iterations).

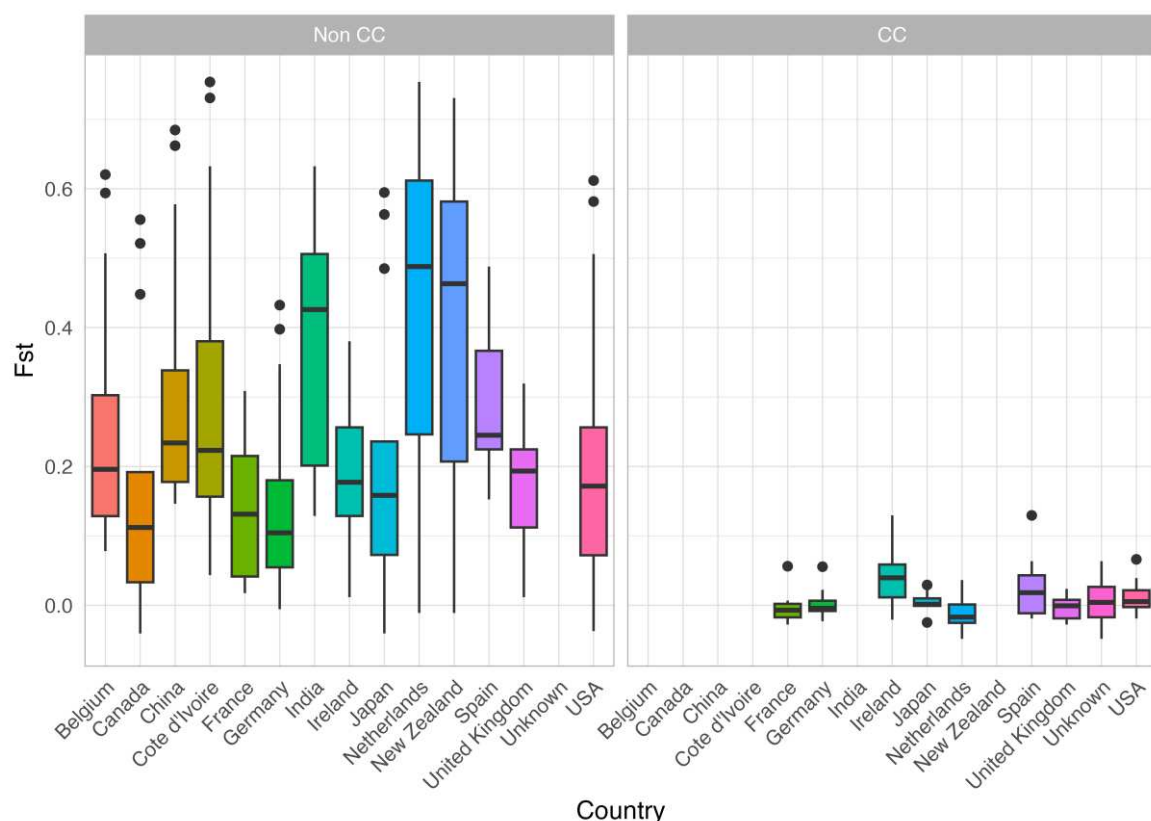


Figure 5. Pairwise Nei's F_{st} values for countries with sample sizes exceeding 8 (left) or total MLGs exceeding 8 (right). The left panel was generated using the original sample data original sample data (non-CC), while the right consisted of country wise clone corrected (CC) data (filtered for sample size post clone correction). The value for each pair of countries is represented twice in each figure, once towards each country in the pair.

3.5. Recombination and Phylogenetic Incompatibility

Three analyses were performed to investigate the possibility of recombination in the mitogenome in *A. fumigatus*. Here, only the global sample of 79 clone-corrected MLGs were included in the analyses. In the first analysis of multilocus index of association (I_A) among the 39 SNP loci, our sample showed significant deviations from random mating and recombination, with the overall I_A and standardized rD values of 2.50 and 0.07, respectively. Both values were significantly higher than what would be expected in randomly recombining populations ($p < 0.01$). In the second test, phylogenetic compatibility analysis of the 39 SNP loci also demonstrated a significant departure from random mating. Approximately 80% of the loci pairs displayed phylogenetic compatibility, which is significantly higher than what would be observed under the hypothesis of random recombination ($p < 0.01$). In addition, out of the 741 SNP loci pairs, 73 were found to exhibit significant allelic associations, indicative of linkage disequilibrium ($p < 0.01$, Bonferroni correction). Taken together, these results indicated no evidence of random recombination in the mitogenome of the global *A. fumigatus* population.

However, signatures of recombination were found in the mitogenomes of this global sample. Specifically, 145 of the 741 pairs of SNP loci showed phylogenetic incompatibility (Figure 6), consistent with recombination between these SNP sites. In addition, upon comparison of the LD and four gamete test results per loci pair, the majority (86%) of the phylogenetically incompatible pairs failed to reject the null hypothesis of recombination (Figure 6). Taken together, our results suggest that while non-random mating and linkage disequilibrium dominate, there is unambiguous evidence for recombination in the mitogenomes of the global population of *A. fumigatus*.

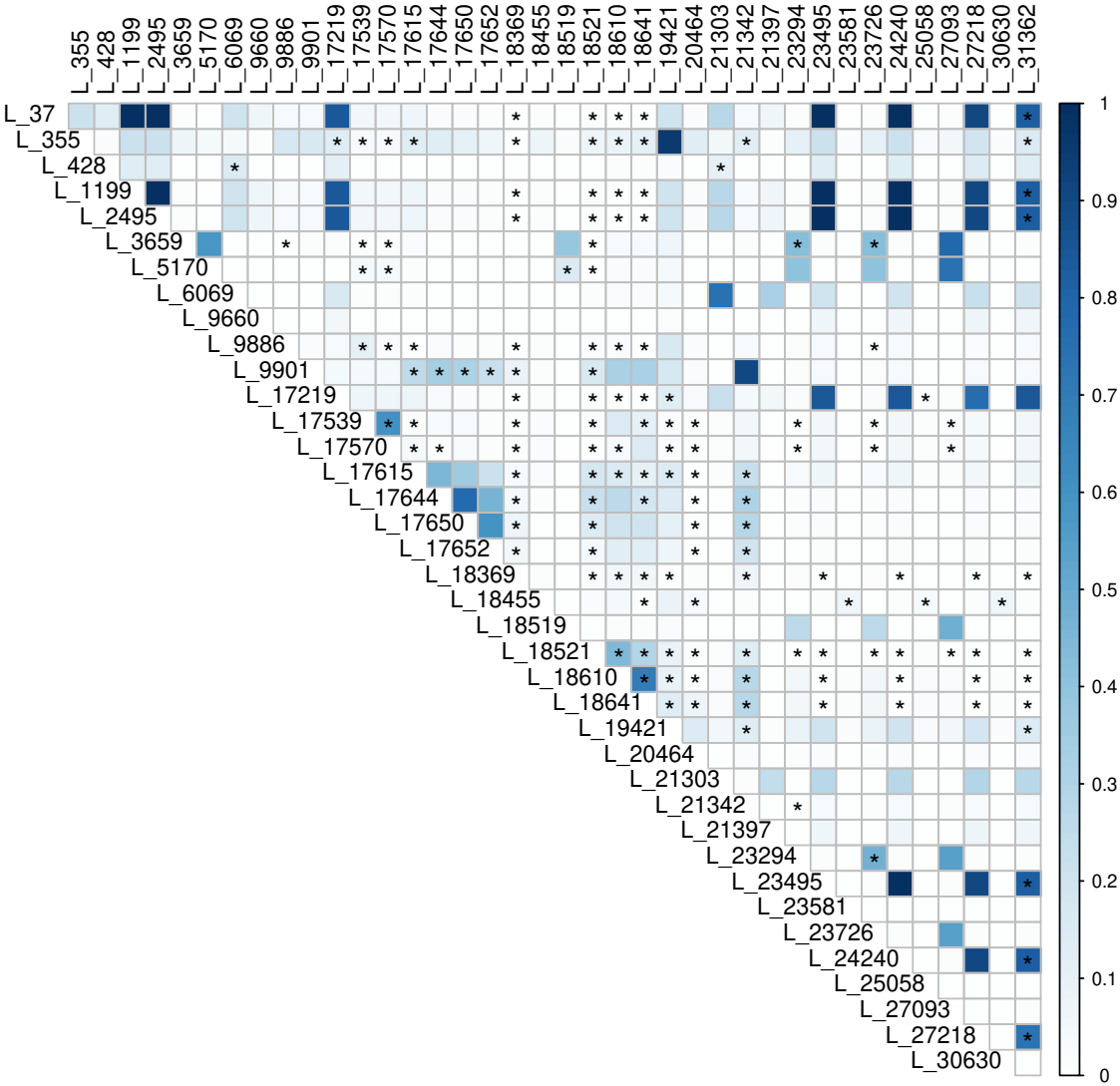


Figure 6. Allelic correlation/linkage disequilibrium and Four Gamete Test results. Colour intensity represents r^2 value, with higher colour intensity indicating higher correlation. Pairs of loci marked with ‘*’ contained all four possible pairs of alleles, indicative of phylogenetic incompatibility and consistent with recombination.

3.6. Relative Mitogenome to Nuclear Genome Copy Number Ratio

The average sequencing depth per nucleotide across the entire *cyp51A* gene varied widely among the 1939 isolates, from a few reads to over 400 reads (Figure 7.A). As expected, the sequencing depths estimated based on *cyp51A* and the whole-genome coverage are highly correlated with each other ($r^2 = 0.95$, $p < 0.001$; Figure 7.B). Similarly, the average sequencing depth per nucleotide across the mitogenome varied widely among the 1939 isolates, from a few reads to over 15000 reads (Figure 7.C). Both the *cyp51A* read depth and the mitogenome read depth data showed a heavy right skew, indicated by the high values of γ (shape parameter), 1.79 and 2.83, respectively (Figure 7.A and C). Interestingly, the ratio between mitogenome sequencing depth and *cyp51A* sequencing depth also showed a very broad distribution, from less than 1 to over 100, with one ratio close to 600.

We checked the metadata and citations of strains with either very low or very high ratios. Our analyses identified that ten of the strains with either very high (>200) or very low (<1) ratios were either transcriptome data or synthetic metagenome data but were deposited as whole-genome sequence data. These ten strains were removed from subsequent analyses of mitogenome to nuclear

genome ratios. Interestingly, though the raw ratios had a significantly skewed distribution ($\gamma = 1.10$), the square-rooted ratios had a normal distribution ($\gamma = 0.18$), with a mean ratio of 20.70 among the 1929 isolates (95% CI = (18.06, 23.34); Figure 7.D).

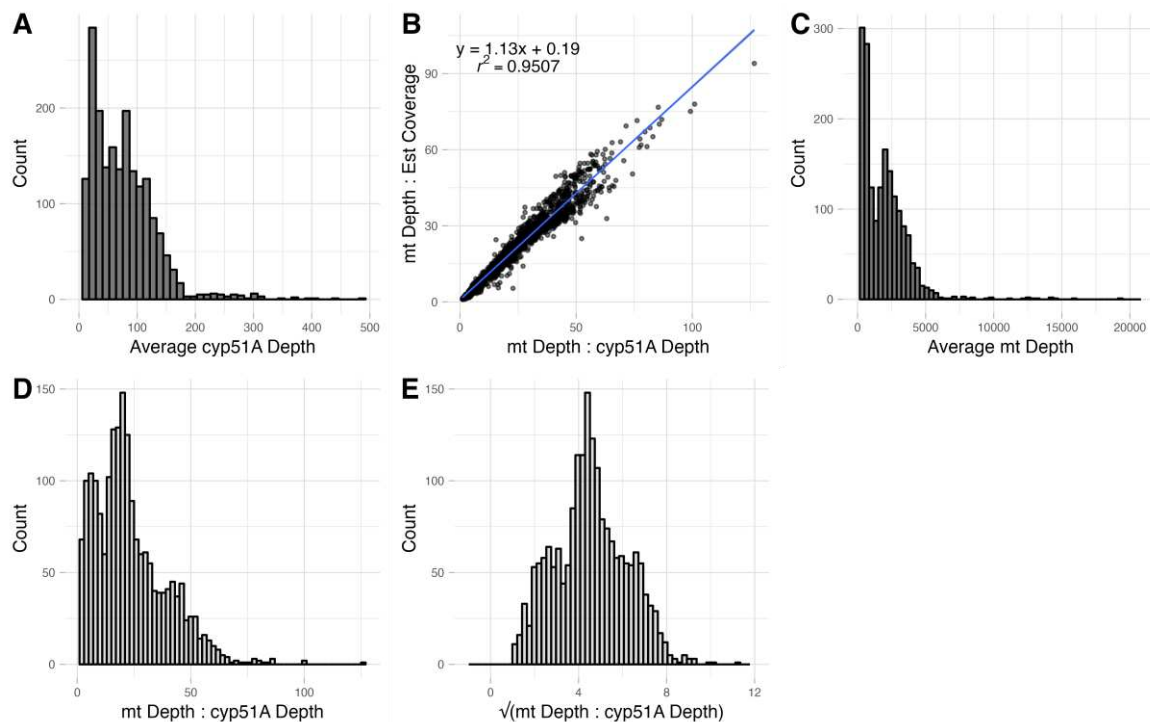


Figure 7. Distributions of sequence read depths and mitogenome to nuclear genome read depth ratios among 1939 isolates of *A. fumigatus*. (A) Right skewed histogram of the average *cyp51A* read depth per isolate, $n = 1939$. (B) Relationship between mitogenome to *cyp51A* read depths ratio (x-axis) and the mitogenome read depth to estimated genome coverage ratio (y-axis). The values are strongly correlated with an $r^2 = 0.9507$ ($p < 0.01$), $n = 1929$. (C) Right skewed histogram of the average mt genome read depth per isolate, $n = 1939$. (D) Histogram of mt genome to *cyp51A* average read depths ratio displaying heavy right skew, $n = 1929$. (E) Histogram of the square root transformed values from D displaying a more normal distribution, $n = 1929$.

To investigate the potential patterns of mitogenome to nuclear genome ratios among strains and genetic and geographic populations, we compared the ratio patterns across both the DAPC grouping and among countries with sample size (>8). Kruskal Wallis testing revealed statistically significant heterogeneity in ratio distribution within both the DAPC groups and the countries examined ($p < 0.05$, $p < 0.01$). Subsequent post hoc pairwise Wilcoxon rank sum testing uncovered significant ($p < 0.05$) differences between Group #3 and both Group #1 and Group #2. Specifically, the Group #3 strains had an overall higher mitogenome to nuclear genome ratio than those in Groups #1 and #2, 29 compared to 21 and 18, respectively. In contrast, no significant differences were observed between Groups #1 and #2.

Among the 14 compared countries (each country with sample size greater than 8 isolates), 31 pairs (out of 91 total pairs) showed statistically significant ($p < 0.01$) differences in their mitogenome to nuclear genome ratios. Canada had the highest average ratio of 46.2 mitochondrial to nuclear genome depth ratios, more than double the overall average of 20.7. Conversely, Spain had the lowest average ratio, 11.4, just over half of the overall average. Canada had a significantly higher ratio than 11 of the 13 compared countries. In contrast, Spain had a significantly lower ratio than 6 of the 13 compared countries. Interestingly, China stood out as the only country that did not display any significant differences when compared to other 13 countries, primarily due to the large variations among strains within China. Furthermore, China and Canada were not significantly different, but Canada and New Zealand were (at a 0.05 significance level) despite New Zealand having a higher

average ratio than China (22.0 and 20.0, respectively). Interestingly, there is evidence of a statistically significant negative correlation between the mitochondrial to nuclear genome ratio and climate. Comparing the samples from the 14 countries to the average annual temperature for 2021 resulted in a Pearson correlation coefficient of -0.61 ($p < 0.05$).

4. Discussion

In this study, we conducted a broad analysis of mitogenome SNPs in a global collection of 1939 isolates of *A. fumigatus*, a critical priority human fungal pathogen. Our analyses identified 39 SNPs and 79 multi-locus genotypes among the analyzed isolates. In addition, we found a range of mitogenome to nuclear genome copy number ratios, from about 11 to over 100, with a mean of ~20. Our examination of concatenated SNP sequences separated the strains into at least three genetic groups. However, though multiple genotypes are often overrepresented in certain geographic regions, there was limited geography-based clustering at either the country or the continent level for the 79 mitochondrial genotypes. Interestingly, evidence for mitogenome recombination was found between several pairs of SNP loci. Below, we discuss the relevance and potential implications of our results.

4.1. Geographic Structuring

The analyzed global population included 1939 *A. fumigatus* isolates representing 22 countries in 6 continents, including two strains from the International Space Station. Among the 79 MLGs from these 1939 isolates, 19 were each distributed in at least three countries, indicative of continental and global clonal expansion. Indeed, all the sequenced samples from Africa, Oceania, South America, and the International Space Station shared MLGs with samples from other geographic regions and none had private mitochondrial MLGs. Interestingly, the two samples from the International Space Station were genetically divergent from each other, similar to the extensive differences found in their nuclear genomes [3]. On the other hand, 40 of the 79 MLGs were found only in one country each, and AMOVA analysis of the total samples revealed statistically significant genetic differentiations among national samples. However, clone-corrected samples revealed limited genetic differentiations among national samples. Together, the results suggested that localized clonal expansion of certain MLGs played a major role in the observed genetic differentiations among countries in the total sample. For instance, MLGs 11 and 79 were the most abundant, representing 420 isolates from 18 countries and 418 isolates from 11 countries, respectively (Figure 1). However, over half and one third of MLG 11 and MLG 79 isolates were found in the Netherlands and the USA, respectively. Localized clonal expansion can significantly bias allele frequencies, contributing to the observed genetic difference among geographic populations [52].

Our phylogenetic and DAPC analyses showed our samples can be optimally clustered into three major clades. This finding differed from previous studies conducted by Pringle et al. [53], Klaassen et al. [54], and Ashu et al. [55] who found two, five, and eight clades, respectively, in their global samples. These three studies used different markers as ours: they all relied on nuclear genomic data, based on either multilocus sequence typing or microsatellite markers. Interestingly, while both Pringle et al. [53] and Laassen et al. [54] conclude no significant correlation between clades and geography, Ashu et al. [55] found limited but statistically significant geographic contributions to genetic variation with a statistically significant geographic contributions in the non-clone corrected AMOVA, but a non-significant Mantel test. The difference in population structure among the studies likely reflect the different samples used and the difference in inheritance patterns between nuclear and mitochondrial genomes, especially where mitochondria largely inherit in a non-Mendelian pattern during sexual mating in filamentous ascomycetes [56,57].

We note that despite the large sample size (1939 isolates) and broad geographic representation (22 countries from across six continents), the analyzed samples were still limited in their geographic distributions. Specifically, the 22 countries represented only about 12% of the total countries in the world. A previous study by Ashu et al. [55] revealed the Cameroonian population of *A. fumigatus* to be genetically highly distinct from other national populations. It's likely that the mitogenomes from

the Cameroonian populations will be genetically distinct as well. However, there is yet no genome sequence information from Cameroonian isolates of *A. fumigatus*. In addition, among the 22 sampled countries, the distributions of isolates were also highly skewed. Specifically, over half of the isolates (1036) originated from western Europe, and an additional 30% (563) were from the United States. Furthermore, we were unable to find the geographic origins for 114 of the isolates. Together, we believe that better data submission and including more strains from under-sampled regions and strains from not-yet sampled regions will reveal additional novel mitogenome diversity in this species.

4.2. Recombination

We observed signatures of recombination in the mitogenome of this global population of *A. fumigatus*. In *A. fumigatus*, the mitochondrial inheritance has not been critically examined in sexual crosses in the lab. However, sexual mating through hyphal fusion between strains of different mating types should result in significant cytoplasmic mixing, including the potential for mitochondrial fusion and mitogenome recombination. The observation of cleistothecia development along junction zones between mating partners is consistent with hyphal fusion being a dominant mode of mating in *A. fumigatus* [58]. Our observed mitogenome recombination in natural populations of *A. fumigatus* contrasts with those observed in the majority of eukaryotic organisms where mitogenomes are inherited uniparentally in sexual crosses without recombination [57]. However, a growing number of fungal species have shown evidence of mitogenome recombination in their natural populations. These included several mushroom species such as *Agaricus bisporus* [24], *Thelephora ganbajun* [59], *Cantharellus cibarius* species complex [60], and several human pathogenic fungi, such as *Cryptococcus neoformans* [22], *Cryptococcus gattii* [23], *Candida albicans* [61], and *Candida auris* [62]. Interestingly, studies on other members of the *Aspergillus* genus have provided potential evidence of intra- and inter-species mitochondrial recombination. For instance, investigations on *Aspergillus nidulans* using genetic crosses, gel electrophoresis, Southern hybridization, and electron microscopy suggested genetic exchange between mitogenomes [63,64]. Similarly, novel mitogenomes were observed in the progeny of a sexual cross in *Aspergillus niger*, a related opportunistic human pathogen, caused by the homing of mobile mitochondrial group I introns [65]. In both cases, restriction fragment length differences were observed between parental and progeny mitogenomes, consistent with horizontal transfer of mobile introns. However, in our *A. fumigatus* natural populations, evidence for recombination was observed between SNPs located across the mitogenome.

4.3. Relative Copy Number Ratios of Mitogenome to Nuclear Genome

Aside from mitogenome SNPs, we also investigated the relative copy number differences between mitochondrial and nuclear genomes among the 1939 strains of *A. fumigatus*. Wide variations of the ratio, from about 11 to over 100, were found among strains and among geographic populations. Based on the redundancy in their assembled sequence, a previous study estimated a clinical isolate of *A. fumigatus* Af293 having a mitochondrial to nuclear DNA ratio of ~12:1 [66]. In a more recent study based on real-time PCR of single copy genes in the nuclear and mitochondrial genomes, Neubauer et al. [67] estimated that the mitochondrial to nuclear genome ratio in a wild-type strain AfS35 was about 60(±15) :1. At present, the underlying mechanisms for such variations are unknown. Based on the literature, all the 1929 genome-sequenced strains were grown in favorable laboratory conditions before their mycelia were harvested and DNA extracted for sequencing. While some variations may exist in growth conditions among studies in preparing their strains for DNA extraction and genome sequencing, the observed differences among geographic populations suggest that the identified range of ratios might reflect their adaptations to the natural environments from which the strains originally resided. Specifically, we hypothesize that the relatively high mitochondrial to nuclear genome ratio for Canadian isolates may reflect the low temperature these isolates were adapted to and the need to have a high number of mitochondrial genes to generate more ATP to fulfill their metabolic needs in such environments. Controlled experimental evolution under different temperatures are needed to test this hypothesis. Furthermore, due to potential

differences in their capacity to generate ATP among strains with different mitochondrial to nuclear genome ratios, we speculate that such strains may exhibit different properties such as antifungal drug resistance and pathogenesis. Targeted experimental studies are needed to determine the biological significance of the observed variations in mitochondrial to nuclear genome ratios among strains.

5. Conclusions and Perspectives

Our study provides novel insights into mitogenome variations in a large global collection isolates of *A. fumigatus*, a critical human fungal pathogen. Our analyses revealed evidence for both global and local clonal expansions of mitogenomes. The over-representation of relatively few mitochondrial genotypes in the global sample suggested potential selective advantages of these mitochondrial MLGs and/or their associated nuclear genomes in both natural environments and patients. In addition, evidence of mitogenome recombination was found among several mtSNP loci, indicating a departure from typical uniparental mitochondrial inheritance pattern in eukaryotes. Similarly, the investigation of relative copy number differences between mitochondria and nuclear genomes revealed a wide range of ratios. At present, the potential biological significance of the observed variations in mitogenome to nuclear genome ratios is not known. However, it's tempting to speculate that the variable ratios likely reflect the dynamic nature of mitochondrial genome replication and maintenance which are regulated by various environmental and metabolic cues. Our study underscores the need for further investigation into the functional implications of the observed mitogenome variations in *A. fumigatus*. Collectively, these findings help lay the foundation for future studies into the genetic mechanisms underlying the diversity and evolution of this critical fungal pathogen.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Table S1: Information about the 1939 isolates analyzed in this study.

Author Contributions: Conceptualization, J.X.; methodology, V.T.; software, V.T.; validation, V.T. and J.X.; formal analysis, V.T.; investigation, V.T.; resources, V.T. and J.X.; data curation, V.T.; writing—original draft preparation, V.T.; writing—review and editing, J.X.; visualization, V.T.; supervision, J.X.; project administration, J.X.; funding acquisition, J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by an internal grant from the David Braley Centre for Antibiotic Discovery at McMaster University, grant number 2022-5.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data described in this study are presented in the manuscript and public available.

Acknowledgments: We thank Yue Wang, George Long, Megan Hitchcock, and Greg Korfanty for technical assistance.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Xu, J. Assessing Global Fungal Threats to Humans. *mLife* **2022**, *1*, 223–240, doi:10.1002/mlf2.12036.
2. Latgé, J.-P.; Chamilos, G. *Aspergillus fumigatus* and Aspergillosis in 2019. *Clin Microbiol Rev* **2019**, *33*, e00140-18, doi:10.1128/CMR.00140-18.
3. Knox, B.P.; Blachowicz, A.; Palmer, J.M.; Romsdahl, J.; Huttenlocher, A.; Wang, C.C.C.; Keller, N.P.; Venkateswaran, K. Characterization of *Aspergillus fumigatus* Isolates from Air and Surfaces of the International Space Station. *mSphere* **2016**, *1*, e00227-16, doi:10.1128/mSphere.00227-16.
4. Paulussen, C.; Hallsworth, J.E.; Álvarez-Pérez, S.; Nierman, W.C.; Hamill, P.G.; Blain, D.; Rediers, H.; Lievens, B. Ecology of Aspergillosis: Insights into the Pathogenic Potency of *Aspergillus fumigatus* and Some Other *Aspergillus* Species. *Microb. Biotechnol.* **2017**, *10*, 296–322, doi:10.1111/1751-7915.12367.

5. Nywening, A.V.; Rybak, J.M.; Rogers, P.D.; Fortwendel, J.R. Mechanisms of Triazole Resistance in *Aspergillus fumigatus*. *Environmental Microbiology* **2020**, *22*, 4934–4952, doi:10.1111/1462-2920.15274.
6. Dagenais, T.R.T.; Keller, N.P. Pathogenesis of *Aspergillus fumigatus* in Invasive Aspergillosis. *Clin Microbiol Rev* **2009**, *22*, 447–465, doi:10.1128/CMR.00055-08.
7. World Health Organization WHO Fungal Priority Pathogens List to Guide Research, Development and Public Health Action 2022.
8. Barber, A.E.; Sae-Ong, T.; Kang, K.; Seelbinder, B.; Li, J.; Walther, G.; Panagiotou, G.; Kurzai, O. *Aspergillus fumigatus* Pan-Genome Analysis Identifies Genetic Variants Associated with Human Infection. *Nat Microbiol* **2021**, *6*, 1526–1536, doi:10.1038/s41564-021-00993-x.
9. Szalewski, D.A.; Hinrichs, V.S.; Zinniel, D.K.; Barletta, R.G. The Pathogenicity of *Aspergillus fumigatus*, Drug Resistance, and Nanoparticle Delivery. *Can. J. Microbiol.* **2018**, *64*, 439–453, doi:10.1139/cjm-2017-0749.
10. Black, B.; Lee, C.; Horianopoulos, L.C.; Jung, W.H.; Kronstad, J.W. Respiring to Infect: Emerging Links between Mitochondria, the Electron Transport Chain, and Fungal Pathogenesis. *PLOS Pathogens* **2021**, *17*, e1009661, doi:10.1371/journal.ppat.1009661.
11. Grahl, N.; Dinamarco, T.M.; Willger, S.D.; Goldman, G.H.; Cramer, R.A. *Aspergillus fumigatus* Mitochondrial Electron Transport Chain Mediates Oxidative Stress Homeostasis, Hypoxia Responses and Fungal Pathogenesis. *Molecular Microbiology* **2012**, *84*, 383–399, doi:10.1111/j.1365-2958.2012.08034.x.
12. Ma, H.; Hagen, F.; Stekel, D.J.; Johnston, S.A.; Sionov, E.; Falk, R.; Polacheck, I.; Boekhout, T.; May, R.C. The Fatal Fungal Outbreak on Vancouver Island Is Characterized by Enhanced Intracellular Parasitism Driven by Mitochondrial Regulation. *Proceedings of the National Academy of Sciences* **2009**, *106*, 12980–12985, doi:10.1073/pnas.0902963106.
13. Magnani, T.; Soriani, F.M.; Martins, V. de P.; Policarpo, A.C. de F.; Sorgi, C.A.; Faccioli, L.H.; Curti, C.; Uyemura, S.A. Silencing of Mitochondrial Alternative Oxidase Gene of *Aspergillus fumigatus* Enhances Reactive Oxygen Species Production and Killing of the Fungus by Macrophages. *J Bioenerg Biomembr* **2008**, *40*, 631–636, doi:10.1007/s10863-008-9191-5.
14. Misslinger, M.; Lechner, B.E.; Bacher, K.; Haas, H. Iron-Sensing Is Governed by Mitochondrial, Not by Cytosolic Iron–Sulfur Cluster Biogenesis in *Aspergillus fumigatus*. *Metallomics* **2018**, *10*, 1687–1700, doi:10.1039/c8mt00263k.
15. Song, J.; Zhou, J.; Zhang, L.; Li, R. Mitochondria-Mediated Azole Drug Resistance and Fungal Pathogenicity: Opportunities for Therapeutic Development. *Microorganisms* **2020**, *8*, 1574, doi:10.3390/microorganisms8101574.
16. Basse, C.W. Mitochondrial Inheritance in Fungi. *Current Opinion in Microbiology* **2010**, *13*, 712–719, doi:10.1016/j.mib.2010.09.003.
17. Sandor, S.; Zhang, Y.; Xu, J. Fungal Mitochondrial Genomes and Genetic Polymorphisms. *Appl Microbiol Biotechnol* **2018**, *102*, 9433–9448, doi:10.1007/s00253-018-9350-5.
18. Yan, Z.; Xu, J. Fungal mitochondrial inheritance and evolution. In *Evolutionary Genetics of Fungi*; Horizon bioscience: Wymondham, 2005; pp. 221–252 ISBN 978-1-904933-15-1.
19. Joardar, V.; Abrams, N.F.; Hostetler, J.; Paukstelis, P.J.; Pakala, S.; Pakala, S.B.; Zafar, N.; Abolude, O.O.; Payne, G.; Andrianopoulos, A.; et al. Sequencing of Mitochondrial Genomes of Nine *Aspergillus* and *Penicillium* Species Identifies Mobile Introns and Accessory Genes as Main Sources of Genome Size Variability. *BMC Genomics* **2012**, *13*, 698, doi:10.1186/1471-2164-13-698.
20. Song, S.-N.; Tang, P.; Wei, S.-J.; Chen, X.-X. Comparative and Phylogenetic Analysis of the Mitochondrial Genomes in Basal Hymenopterans. *Sci Rep* **2016**, *6*, 20972, doi:10.1038/srep20972.
21. Lewin, R. The Unmasking of Mitochondrial Eve. *Science* **1987**, *238*, 24–26, doi:10.1126/science.3116666.
22. Wang, Y.; Xu, J. Mitochondrial Genome Polymorphisms in the Human Pathogenic Fungus *Cryptococcus neoformans*. *Frontiers in Microbiology* **2020**, *11*.
23. Xu, J.; Yan, Z.; Guo, H. Divergence, Hybridization, and Recombination in the Mitochondrial Genome of the Human Pathogenic Yeast *Cryptococcus gattii*. *Molecular Ecology* **2009**, *18*, 2628–2642, doi:10.1111/j.1365-294X.2009.04227.x.
24. Xu, J.; Zhang, Y.; Pun, N. Mitochondrial Recombination in Natural Populations of the Button Mushroom *Agaricus bisporus*. *Fungal Genetics and Biology* **2013**, *55*, 92–97, doi:10.1016/j.fgb.2012.09.004.
25. Hebert, P.D.N.; Cywinska, A.; Ball, S.L.; deWaard, J.R. Biological Identifications through DNA Barcodes. *Proc Biol Sci* **2003**, *270*, 313–321, doi:10.1098/rspb.2002.2218.
26. Xu, J. Fungal DNA Barcoding. *The 6th International Barcode of Life Conference* **2017**, *01*, 913–932, doi:10.1139/gen-2016-0046@gen-iblf.issue01.
27. Kozłowski, M.; Stepień, P.P. Restriction Enzyme Analysis of Mitochondrial DNA of Members of the Genus *Aspergillus* as an Aid in Taxonomy. *Microbiology* **1982**, *128*, 471–476, doi:10.1099/00221287-128-3-471.

28. Wang, L.; Yokoyama, K.; Miyaji, M.; Nishimura, K. The Identification and Phylogenetic Relationship of Pathogenic Species of *Aspergillus* Based on the Mitochondrial Cytochrome b Gene. *Med Mycol* **1998**, *36*, 153–164.
29. Wang, L.; Yokoyama, K.; Miyaji, M.; Nishimura, K. Mitochondrial Cytochrome b Gene Analysis of *Aspergillus fumigatus* and Related Species. *Journal of Clinical Microbiology* **2000**, *38*, 1352–1358, doi:10.1128/JCM.38.4.1352-1358.2000.
30. Álvarez-Iglesias, V.; Mosquera-Miguel, A.; Cerezo, M.; Quintáns, B.; Zarrabeitia, M.T.; Cuscó, I.; Lareu, M.V.; García, Ó.; Pérez-Jurado, L.; Carracedo, Á.; et al. New Population and Phylogenetic Features of the Internal Variation within Mitochondrial DNA Macro-Haplogroup R0. *PLOS ONE* **2009**, *4*, e5112, doi:10.1371/journal.pone.0005112.
31. Chung, H. Phylogenetic Analysis and Characterization of Mitochondrial DNA for Korean Native Cattle. **2013**, 2013, doi:10.4236/ojgen.2013.31003.
32. Sabir, J.S.M.; Arasappan, D.; Bahieldin, A.; Abo-Aba, S.; Bafeel, S.; Zari, T.A.; Edris, S.; Shokry, A.M.; Gadalla, N.O.; Ramadan, A.M.; et al. Whole Mitochondrial and Plastid Genome SNP Analysis of Nine Date Palm Cultivars Reveals Plastid Heteroplasmy and Close Phylogenetic Relationships among Cultivars. *PLOS ONE* **2014**, *9*, e94158, doi:10.1371/journal.pone.0094158.
33. Zhu, Q.; Gao, P.; Liu, S.; Amanullah, S.; Luan, F. Comparative Analysis of Single Nucleotide Polymorphisms in the Nuclear, Chloroplast, and Mitochondrial Genomes in Identification of Phylogenetic Association among Seven Melon (*Cucumis Melo* L.) Cultivars. *Breeding Science* **2016**, *66*, 711–719, doi:10.1270/jsbbs.16066.
34. Barber, A.E.; Riedel, J.; Sae-Ong, T.; Kang, K.; Brabetz, W.; Panagiotou, G.; Deising, H.B.; Kurzai, O. Effects of Agricultural Fungicide Use on *Aspergillus fumigatus* Abundance, Antifungal Susceptibility, and Population Structure. *mBio* **2020**, *11*, 10.1128/mbio.02213-20, doi:10.1128/mbio.02213-20.
35. Hagiwara, D.; Takahashi, H.; Watanabe, A.; Takahashi-Nakaguchi, A.; Kawamoto, S.; Kamei, K.; Gonoi, T. Whole-Genome Comparison of *Aspergillus fumigatus* Strains Serially Isolated from Patients with Aspergillosis. *Journal of Clinical Microbiology* **2014**, *52*, 4202–4209, doi:10.1128/JCM.01105-14.
36. Lofgren, L.A.; Ross, B.S.; Cramer, R.A.; Stajich, J.E. The Pan-Genome of *Aspergillus Fumigatus* Provides a High-Resolution View of Its Population Structure Revealing High Levels of Lineage-Specific Diversity Driven by Recombination. *PLOS Biology* **2022**, *20*, e3001890, doi:10.1371/journal.pbio.3001890.
37. Rhodes, J.; Abdolrasouli, A.; Dunne, K.; Sewell, T.R.; Zhang, Y.; Ballard, E.; Brackin, A.P.; van Rhijn, N.; Chown, H.; Tsitsopoulou, A.; et al. Population Genomics Confirms Acquisition of Drug-Resistant *Aspergillus fumigatus* Infection by Humans from the Environment. *Nat Microbiol* **2022**, *7*, 663–674, doi:10.1038/s41564-022-01091-2.
38. Winter, D.J.; Weir, B.S.; Glare, T.; Rhodes, J.; Perrott, J.; Fisher, M.C.; Stajich, J.E.; Digby, A.; Dearden, P.K.; Cox, M.P. A Single Fungal Strain Was the Unexpected Cause of a Mass Aspergillosis Outbreak in the World's Largest and Only Flightless Parrot. *iScience* **2022**, *25*, 105470, doi:10.1016/j.isci.2022.105470.
39. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* **2018**, *34*, i884–i890, doi:10.1093/bioinformatics/bty560.
40. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. **2013**, doi:10.48550/ARXIV.1303.3997.
41. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve Years of SAMtools and BCFtools. *GigaScience* **2021**, *10*, giab008, doi:10.1093/gigascience/giab008.
42. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data. *Genome Res.* **2010**, *20*, 1297–1303, doi:10.1101/gr.107524.110.
43. Ortiz, E.M. Vcf2phylip v2.0: Convert a VCF Matrix into Several Matrix Formats for Phylogenetic Analysis. **2019**.
44. Gkanogiannis, A. fasttreeR: Phylogenetic, Distance and Other Calculations on VCF and Fasta Files **2023**.
45. Paradis, E.; Schliep, K. Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R. *Bioinformatics* **2019**, *35*, 526–528, doi:10.1093/bioinformatics/bty633.
46. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Research* **2021**, *49*, W293–W296, doi:10.1093/nar/gkab301.
47. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr : An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction. *PeerJ* **2014**, *2*, e281, doi:10.7717/peerj.281.
48. Jombart, T.; Devillard, S.; Balloux, F. Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations. *BMC Genet* **2010**, *11*, 94, doi:10.1186/1471-2156-11-94.
49. Brown, A.H.D.; Feldman, M.W.; Nevo, E. Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics* **1980**, *96*, 523–536, doi:10.1093/genetics/96.2.523.

50. Smith, J.M.; Smith, N.H.; O'Rourke, M.; Spratt, B.G. How Clonal Are Bacteria? *Proceedings of the National Academy of Sciences* **1993**, *90*, 4384–4388, doi:10.1073/pnas.90.10.4384.
51. Wei, T.; Simko, V. R Package “Corrplot”: Visualization of a Correlation Matrix 2021.
52. Xu, J. Fundamentals of Fungal Molecular Population Genetic Analyses. *Current Issues in Molecular Biology* **2006**, *8*, 75–90, doi:10.21775/cimb.008.075.
53. Pringle, A.; Baker, D.M.; Platt, J.L.; Wares, J.P.; Latgé, J.P.; Taylor, J.W. Cryptic Speciation in the Cosmopolitan and Clonal Human Pathogenic Fungus *Aspergillus fumigatus*. *Evolution* **2005**, *59*, 1886–1899, doi:10.1111/j.0014-3820.2005.tb01059.x.
54. Klaassen, C.H.W.; Gibbons, J.G.; Fedorova, N.D.; Meis, J.F.; Rokas, A. Evidence for Genetic Differentiation and Variable Recombination Rates among Dutch Populations of the Opportunistic Human Pathogen *Aspergillus fumigatus*. *Molecular Ecology* **2012**, *21*, 57–70, doi:10.1111/j.1365-294X.2011.05364.x.
55. Ashu, E.E.; Hagen, F.; Chowdhary, A.; Meis, J.F.; Xu, J. Global Population Genetic Analysis of *Aspergillus fumigatus*. *mSphere* **2017**, *2*, 10.1128/msphere.00019-17, doi:10.1128/msphere.00019-17.
56. Verma, S.; Shakya, V.P.S.; Idnurm, A. Exploring and Exploiting the Connection between Mitochondria and the Virulence of Human Pathogenic Fungi. *Virulence* **2018**, *9*, 426–446, doi:10.1080/21505594.2017.1414133.
57. Xu, J.; Li, H. Current Perspectives on Mitochondrial Inheritance in Fungi. *Cell Health and Cytoskeleton* **2015**, *7*, 143–154, doi:10.2147/CHC.S59508.
58. Korfanty, G.; Stanley, K.; Lammers, K.; Fan, Y.Y.; Xu, J. Variations in Sexual Fitness among Natural Strains of the Opportunistic Human Fungal Pathogen *Aspergillus fumigatus*. *Infection, Genetics and Evolution* **2021**, *87*, 104640. <https://doi.org/10.1016/j.meegid.2020.104640>
59. Li, H.; Xu, J.; Wang, S.; Wang, P.; Rao, W.; Hou, B.; Zhang, Y. Genetic Differentiation and Widespread Mitochondrial Heteroplasmy among Geographic Populations of the Gourmet Mushroom *Thelephora Ganbajun* from Yunnan, China. *Genes* **2022**, *13*, 854, doi:10.3390/genes13050854.
60. Zhang, Y.; Wang, S.; Li, H.; Liu, C.; Mi, F.; Wang, R.; Mo, M.; Xu, J. Evidence for Persistent Heteroplasmy and Ancient Recombination in the Mitochondrial Genomes of the Edible Yellow Chanterelles from Southwestern China and Europe. *Front. Microbiol.* **2021**, *12*, 699598, doi:10.3389/fmicb.2021.699598.
61. Jacobsen, M.D.; Rattray, A.M.J.; Gow, N.A.R.; Odds, F.C.; Shaw, D.J. Mitochondrial Haplotypes and Recombination in *Candida albicans*. *Medical Mycology* **2008**, *46*, 647–654, doi:10.1080/13693780801986631.
62. Wang, Y.; Xu, J. Population Genomic Analyses Reveal Evidence for Limited Recombination in the Superbug *Candida auris* in Nature. *Computational and Structural Biotechnology Journal* **2022**, *20*, 3030–3040, doi:10.1016/j.csbj.2022.06.030.
63. Earl, A.J.; Turner, G.; Croft, J.H.; Dales, R.B.G.; Lazarus, C.M.; Lünsdorf, H.; Küntzel, H. High Frequency Transfer of Species-Specific Mitochondrial DNA Sequences between Members of the Aspergillaceae. *Curr Genet* **1981**, *3*, 221–228, doi:10.1007/BF00429824.
64. Rowlands, R.T.; Turner, G. Three-Marker Extranuclear Mitochondrial Crosses in *Aspergillus nidulans*. *Molec. Gen. Genet.* **1975**, *141*, 69–79, doi:10.1007/BF00332379.
65. Hamari, Z.; Tóth, B.; Beer, Z.; Gácsér, A.; Kucsera, J.; Pfeiffer, I.; Juhász, Á.; Kevei, F. Interpretation of Intraspecific Variability in mtDNAs of *Aspergillus niger* Strains and Rearrangement of Their mtDNAs Following Mitochondrial Transmissions. *FEMS Microbiology Letters* **2003**, *221*, 63–71, doi:10.1016/S0378-1097(03)00165-4.
66. Nierman, W.C.; Pain, A.; Anderson, M.J.; Wortman, J.R.; Kim, H.S.; Arroyo, J.; Berriman, M.; Abe, K.; Archer, D.B.; Bermejo, C.; et al. Genomic Sequence of the Pathogenic and Allergenic Filamentous Fungus *Aspergillus fumigatus*. *Nature* **2005**, *438*, 1151–1156, doi:10.1038/nature04332.
67. Neubauer, M.; Zhu, Z.; Penka, M.; Helmschrott, C.; Wagener, N.; Wagener, J. Mitochondrial Dynamics in the Pathogenic Mold *Aspergillus fumigatus*: Therapeutic and Evolutionary Implications. *Molecular Microbiology* **2015**, *98*, 930–945, doi:10.1111/mmi.13167.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.