

WESTERN SYDNEY UNIVERSITY



Does intraspecies variation in *Aspergillus fumigatus* affect infection outcomes? A phenotype/genotype study using an insect model.

Sam El-Kamand

Supervisor: Dr Charles Oliver Morton

Co-Supervisor: Dr Alexie Papanicolaou

School of Science and Health

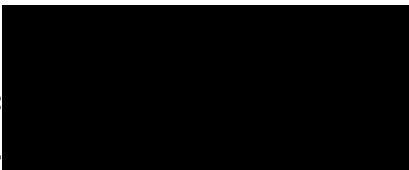
Western Sydney University

Sydney, Australia

December 2019

Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

— 

Abstract

Aspergillus fumigatus is a saprophytic soil-fungus and an opportunistic human pathogen. This haploid mould reproduces asexually using spores that can readily become airborne. In immunocompromised individuals, inhalation of *A. fumigatus* spores can lead to a pulmonary infection termed 'invasive aspergillosis' (IA). Despite extensive research on human immunity and treatment, the relative contribution of fungal genetic and phenotypic variation to the outcomes of infection is yet to be elucidated. In the present study, I sought to determine the pathogenic relevance of the intraspecies variation of *A. fumigatus*. Clinical isolates were characterised using phenotypic assays (UV resistance, amphotericin-B resistance, radial growth rate) and whole genome sequenced to determine genetic relatedness. These data were integrated with virulence data generated in an insect infection model, *Tenebrio molitor* larvae, to determine the relevance of fungal variation to clinical outcomes, identify potential virulence factors, and further our understanding of *A. fumigatus* pathogenesis in invasive aspergillosis. I observed a high level of intraspecies heterogeneity in all pathogenesis-associated phenotypic properties. The spectrum of core-genome single nucleotide polymorphisms (SNPs) present and virulence in *T. molitor* larvae also varied between isolates. Patterns of intraspecies variation aligned with clinical origin for two properties: growth rate on nutrient rich media and virulence in *T. molitor*. The correlation between clinical origin and both growth rate and virulence suggests a contribution of fungal biology towards clinical outcomes. The low level of virulence displayed by IA isolates relative to colonisers suggests the biology of IA isolates may be optimised for overcoming clinical challenges not modelled in *T. molitor* larvae. Finally, the absence of strong clustering of isolates based on their clinical origin

suggests more focused or non-SNP based assays of variation may be necessary to reveal any genomic markers of a strains ability to cause invasive disease.

Acknowledgements

The completion of this thesis was made possible with the support of many remarkable people. I am especially indebted to my supervisors. Thank you, Oliver and Alexie. You have been incredible mentors. I've really enjoyed working in your labs.

Thank you to all the fellow researchers who have helped me along the way, both at Campbelltown and the Hawkesbury Institute for the Environment. Your friendship made this journey all the more enjoyable and seeing you all work so hard was incredibly motivating. Thank you to Yağiz, Kay and Charlie, for taking the time to teach me so much about sequencing. Thank you to Paul and Benson, for taking the time to chat with me about biostats. I would also like to express my deepest gratitude to the technical staff at Campbelltown and Hawkesbury.

To my family, thank you. Zane, my brilliant brother, your humour and your company makes even the most terrible days brighter. Serene, thank you for your advice and support. I feel truly blessed to have such an incredible sister. To my father, Ghassan, thanks for reminding me to lean into whatever I do. Finally, to my mother, Sariah, I thank you with all my heart for teaching us the value of education, inspiring us with your strength, and making us laugh with your wit. You made this possible.

Table of Contents

Statement of Authentication	ii
Abstract	iii
Acknowledgements.....	v
Table of Contents	vi
Table of Figures	i
List of Tables	iii
Chapter 1 Introduction	1
1.1 Invasive Aspergillosis.....	1
1.2 <i>Aspergillus fumigatus</i> pathogenesis	1
1.3. Increasing relevance of <i>A. fumigatus</i> infection	2
1.4 Intraspecies variability of <i>A. fumigatus</i>	3
1.5 Aims of this work:.....	10
Chapter 2 Methods	12
2.1 Clinical <i>A. fumigatus</i> isolates	12
2.2 Non-standard statistical terminology	12
2.3 Phenotypic variation amongst clinical <i>A. fumigatus</i> isolates	13
2.4 Genomic variation amongst clinical <i>A. fumigatus</i>	14
2.5 Virulence Assays	18
Chapter 3 Phenotypic variation amongst clinical <i>A. fumigatus</i> isolates	23
3.1 Introduction	23
3.2 IA isolates grow more slowly than colonisers on nutrient rich media	24
3.3 IA isolates and colonisers show similar conidial UV resistance	26
3.4 AMB resistance varies amongst colonising isolates.	27
Chapter 4 Genotypic variation amongst clinical <i>A. fumigatus</i> isolates	31
4.1 Introduction	31

4.2 Quality control and trimming of raw reads	33
4.3 SPAdes assembly.....	35
4.4 Chromosomal level scaffolding	37
4.5 Contamination removal	41
4.6 Final frozen assemblies	42
4.7 Genome annotation	45
4.8 Variant analysis	46
Chapter 5 Virulence of clinical <i>A. fumigatus</i> isolates	48
5.1 Introduction	48
5.2 Infection model optimisation.....	49
5.3 Inter-isolate variation in virulence of <i>A. fumigatus</i> isolates	52
5.4 Correlation between inter-isolate variation and virulence in an invertebrate model.....	55
Chapter 6 Discussion.....	59
6.1 Clinical Relevance of phenotypic heterogeneity	59
6.2 Clinical relevance of genomic heterogeneity	65
6.3 Clinical relevance of fungal virulence	69
Conclusions	74
References	75
Appendix.....	84

Table of Figures

Figure 1.1. Overview of Invasive Aspergillosis	2
Figure 2.1. Position of 5 different injection sites tested for their effect on mealworm survival	19
Figure 3.1. Linearity of <i>Aspergillus fumigatus</i> growth on PDA at 37°C	25
Figure 3.2. Radial growth rates of clinical <i>A. fumigatus</i> isolates on potato dextrose agar	26
Figure 3.3. Resistance of clinical <i>A. fumigatus</i> isolates to UV irradiation.....	27
Figure 3.4. Resistance of colonising <i>A. fumigatus</i> isolates to Amphotericin B exposure	29
Figure 3.5. Heterogeneity in Amphotericin B resistance of colonising <i>A. fumigatus</i> isolates	30
Figure 4.1. Per base sequence quality (phred scores) of forward (R1) and reverse (R2) reads generated from paired-end Illumina sequencing of 10 clinical <i>A. fumigatus</i> genomes.....	34
Figure 4.2. Per base sequence content of forward (R1) and reverse (R2) reads generated from paired-end Illumina sequencing of 10 clinical <i>A. fumigatus</i> genomes	35
Figure 4.3. Evidence of misassemblies in SPAdes contigs	38
Figure 4.4. Detection of potential misassembly sites in the absence of multi-chromosome spanning SPAdes contigs.....	39
Figure 4.5. Phylogenetic analysis of 10 clinical <i>A. fumigatus</i> isolates.....	47
Figure 5.1. Optimisation of <i>Tenebrio molitor</i> larvae rearing and injection protocols.	50
Figure 5.2. Kaplan-Meier survival of <i>T. molitor</i> larvae infected with different doses of <i>A. fumigatus</i> spores	51
Figure 5.3. Kaplan-Meier survival of <i>T. molitor</i> larvae infected with clinical <i>A. fumigatus</i> isolates	53
Figure 5.4. Median survival time of <i>T. molitor</i> larvae infected with clinical <i>A. fumigatus</i> isolates	54
Figure 5.5. Virulence of clinical <i>A. fumigatus</i> isolates in a <i>T. molitor</i> infection model	55

Figure 5.6. Correlation between virulence of clinical <i>A. fumigatus</i> isolates in an invertebrate model of IA and isolate radial growth rate on nutrient rich media or conidial UV resistance	56
Figure 5.7. Virulence of <i>A. fumigatus</i> isolates by AMB resistance	56
Figure 5.8. Consensus phylogenetic tree constructed based on intronic variation within the core genome of <i>A. fumigatus</i> in the context of virulence in <i>Tenebrio molitor</i> larvae	58
Figure A1. Radial growth curves of clinical <i>A. fumigatus</i> isolates on potato dextrose agar at 37°C	84
Figure A2. Optimisation and validation of a conidial UV resistance assay. Mapping of biosafety cabinet UV exposure	85

List of Tables

Table 1. Raw read summary statistics from Illumina paired-end sequencing of clinical <i>A. fumigatus</i> isolate genomes.	33
Table 2. Contiguity statistics of SPAdes contigs and scaffolds generated assuming genome size of 28.8 Mb.....	36
Table 3. Complete list of contigs split based on evidence of SPAdes missassembly	40
Table 4. Chromosomal level scaffolds produced by Ragout from misassembly-corrected SPAdes contigs.....	41
Table 5. Quality of frozen <i>A. fumigatus</i> assemblies.....	43
Table 6. BUSCO completeness of clinical <i>A. fumigatus</i> genome assemblies	44
Table 7. Structural annotation of clinical <i>A. fumigatus</i> genomes.	45

Chapter 1 Introduction

1.1 Invasive Aspergillosis

Aspergillosis refers to an array of diseases caused by fungi of the *Aspergillus* genus¹. *Aspergillus*-associated diseases include allergic, chronic, and invasive aspergillosis². Of these, invasive aspergillosis (IA) is the most severe, with a mortality rate of 30-90% depending on underlying conditions and treatment regimen³. In IA, hyphae invade parenchymal tissues causing significant damage and necrosis⁴. IA is predominantly a pulmonary disease^{3,5} but can spread to other organs through processes known as angioinvasion and hematogenous dissemination⁶. Only a small subset of *Aspergillus spp.* are of significance in the context of human health and disease⁷. *Aspergillus fumigatus* is the most clinically relevant species, causing over 70% of IA cases⁵.

1.2 *Aspergillus fumigatus* pathogenesis

A. fumigatus is a ubiquitous soil saprophyte that asexually produces highly hydrophobic spores called conidia⁷. Being common fungal constituents of environmental⁸ and hospital air samples⁹⁻¹¹, hundreds of airborne *A. fumigatus* conidia may be inhaled daily¹². Despite enhanced evasion of mucociliary clearance due to their small size (2-3 μ m), pulmonary *A. fumigatus* conidia are asymptotically cleared by innate immune cells in immunocompetent individuals⁷. Alveolar macrophages phagocytose inhaled conidia and promote a proinflammatory response, recruiting neutrophils capable of destroying hyphae¹³. Host factors associated with reduced immunocompetence, including neutropenia, defects in NADPH oxidase and the use of corticosteroid immunosuppressants can leave individuals susceptible to invasive *A. fumigatus* infection (Figure 1.1)¹.

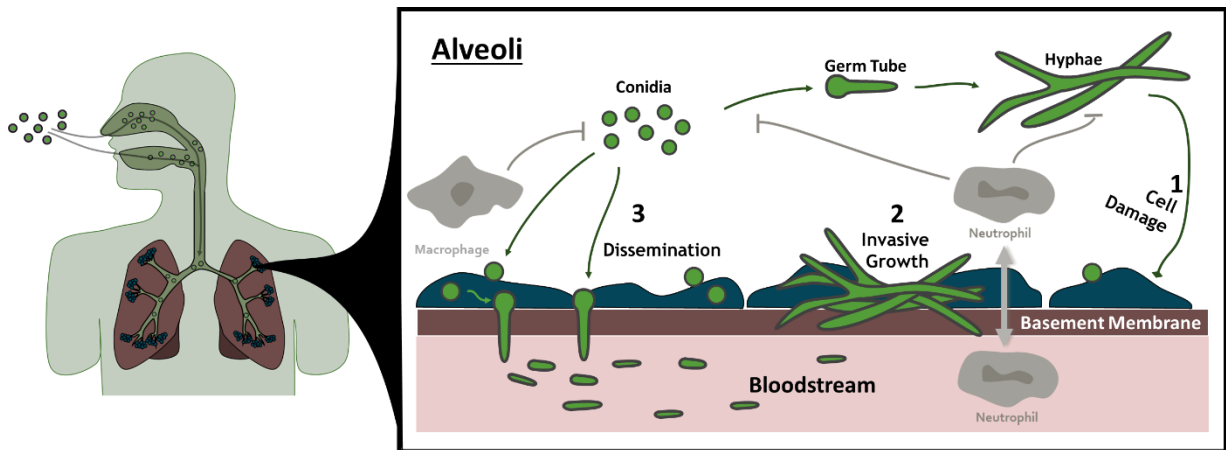


Figure 1.1. Overview of Invasive Aspergillosis. Hundreds of *Aspergillus fumigatus* conidia are inhaled daily. In healthy hosts, conidia that reach the alveoli are phagocytosed by resident macrophages. These macrophages also recruit neutrophils capable of clearing both conidia and hyphae. In an immunocompromised host, IA occurs when conidia are not effectively cleared. The fungus can secrete toxins which damage epithelia and immune cells, and hyphae grow invasively in lung tissues. Conidia can be internalised by epithelial cells and germinate across the basement membrane into the bloodstream, leading to dissemination to other organs. Adapted from Dagenais et al., 2009¹³.

1.3 Increasing relevance of *A. fumigatus* infection

IA affects ~10% of patients with acute leukemia, bone marrow or solid-organ transplants¹⁴. The increasing size of these at-risk populations highlight the growing relevance of the disease¹⁴. Further, resistance of *A. fumigatus* to azoles, the recommended first-line antifungal therapy^{15,16}, is becoming increasingly prevalent. This phenomenon is believed to be accelerated by the widespread agricultural use of related antifungals to combat plant pathogens^{2,17}. Exposure of environmental *A. fumigatus* to these agricultural fungicides induces cross-resistance to clinically important azole antifungals¹⁸. These azole-resistant fungi can then more easily overcome prophylaxis to cause breakthrough infection or withstand azole therapy.

Greater understanding of *A. fumigatus* in the context of IA can provide the foundation for development of novel preventative and therapeutic strategies. Characterising heterogeneity in the *A. fumigatus* population is a natural first step to building this

understanding. Fungal properties showing intraspecies heterogeneity can then be examined in relation to virulence: the ability of a microbe to overcome immune defences, cause infection and damage a host.

1.4 Intraspecies variability of *A. fumigatus*

Large scale studies of *A. fumigatus* isolates have shown that within the species, there is variation in both phenotypic and genotypic traits with putative links to pathogenesis^{19,20}.

1.4.1 Phenotypic variation and its clinical relevance

Recently, identification of phenotypic variation in *Cryptococcus*, another genus of fungi that causes human disease, revealed strong correlations between several morphological properties and clinical outcomes¹⁸. Such associations suggest heterogeneity of infective fungal species can contribute to disease onset and outcome. Like *Cryptococcus*, large scale studies of *Aspergillus fumigatus* isolates have shown there is intraspecies variation in phenotypic traits that may be important in pathogenesis¹⁹. For example, phenotypic properties such as growth rates^{19,21}, pigmentation²² and resistance to antifungal drugs such as Amphotericin B (AMB)²³, have all been shown to vary within the species. These properties are also putatively linked to virulence.

1.4.2 Genotypic variation and its clinical relevance

Due to decreasing costs, whole genome sequencing (WGS) is becoming an increasingly viable tool for assaying genomic variability at an intraspecies level. The first fungi to be sequenced were the yeasts. First *Saccharomyces cerevisiae*, then *Schizosaccharomyces pombe*²⁴. *Neurospora crassa* was the first mould sequenced²⁵. These fungi were all well-established model organisms used in genomic research for

over 70 years, with robust research communities and molecular toolkits. These early referencing studies yielded important discoveries, including that the *N. crassa* genome had almost double the number of genes of either yeast previously sequenced, with 41% of its genome lacking homologs to known proteins, suggesting yeasts are a poor model for the diverse range of fungal species²⁵. In these early days of sequencing, however, the time consuming and expensive nature of WGS rendered intra-specific surveys of genomic variation far less viable than they are today.

The progress of fungal WGS follows a similar pattern irrespective of genus. Early WGS projects attempt to “assemble”, i.e. construct, high-quality and high-level (chromosomal) reference genomes for widely available strains²⁶. In a process called resequencing, a reference genome of the same or closely related species can then be used to “guide” the genome assembly of target organisms. This provides a cost-effective approach to WGS studies, facilitating the pursuit of more specific biological questions. Initially these sequencing projects were time consuming and expensive but have become much more accessible since the introduction of the current generation of sequencing technologies, frequently referred to as next generation sequencing (NGS). Genomes of human pathogenic species from the genera *Cryptococcus*, *Aspergillus*, *Candida*, *Pneumocystis*, *Histoplasma*, *Coccidioides*, *Mucor*, *Blastomyces*, and *Scedosporium*, have been sequenced, assembled, and published in GenBank, the international repository for this kind of data²⁷.

Using WGS resequencing to resolve intra-specific variation has proven to be valuable in the study of fungal microevolution, antifungal resistance, and outbreak and virulence analysis.

Microevolution

Whole-genome sequencing of clinical *Candida albicans* isolates sub-cultured both *in vitro* and in a murine model has been used to characterize the mutations that arise. One study found microevolution to be driven primarily by *de novo* base substitutions and short-tract loss-of-heterozygosity (LOH) events that lead to recombination induced mutagenesis²⁸. WGS of *C. albicans* isolated from oral-samples taken from healthy human hosts also found short-tract LOH events to be important in generating within-host variation²⁹. Further, the high-resolution nature of WGS revealed intra-sample heterogeneity, highlighting the importance of considering intra-host variability when comparing serial isolates²⁹.

Recent *Cryptococcus neoformans* WGS projects have characterized the adaptation of this environmental fungus to the host environment. In these studies, isolates are serially sampled from patients over the course of infection then sequenced and compared to identify microevolution. WGS microevolution studies have demonstrated that: 1) isolates recovered after relapse in cryptococcal meningitis patients are usually clonally related to the original infection^{30,31}, 2) aneuploidy in chromosome 12^{30,31} and mutation of an AT-rich interaction domain protein may be important mechanisms of in-host adaptation³¹, and 3) nonsense mutations in DNA mismatch repair proteins can lead to a hypermutator state, accelerating microevolution³⁰. WGS data has also enabled identification of a genome amplification event that facilitates massive tandem gene amplification in response to environmental stimulus and drives microevolution³². Similar to *C. neoformans*, several recent *A. fumigatus* WGS projects have investigated in-host microevolution albeit with a greater focus on azole-resistance³³⁻³⁵.

Antifungal resistance

A. fumigatus microevolution studies suggest azole-resistance conferring *cyp51A* SNPs³³, a HapE SNP³⁴ and tandem repeats (TR120) in the *cyp51A* promoter are selected during infection³⁵. WGS of fluconazole-nonsusceptible isolates has implicated gain-of-function *Erg11* heterozygous and *Erg3* homozygous mutations and MDR1 promoter allele alterations in azole-resistance of *C. albicans*³⁶. WGS has also been used to identify mechanisms of resistance by genotyping strains in which antifungal resistance conferring genes have been deleted but resistance has been restored by experimental evolution. In the absence of *Rgd1*, an azole-resistance conferring gene, exposure to azoles induced amplification of several chromosomal regions as identified by WGS. Overexpression of a transporter gene, *NPR2* was found to confer resistance³⁷. A similar study exploring the effects of medium-chain fatty acids found susceptibility to be associated with trisomy of chromosome 7³⁸.

WGS of *C. albicans* and *S. cerevisiae* strains experimentally evolved for resistance to co-treatment with azoles and inhibitors of either *Hsp90* or calcineurin have revealed diverse resistance mechanisms including extensive aneuploidies and mutations in drug target genes and regulators of multidrug transporters, ergosterol biosynthesis, and sphingolipid biosynthesis³⁹. WGS has also been used to genotype experimentally evolved azole-resistant *A. fumigatus* isolates relative to their isogenic parental strains. Both medical and non-medical triazole fungicides have been used in experimental evolution studies. For example, Losada et al. found that variants contributing to medical triazole induced resistance include mutations in *erg11A* (*cyp51A*), *erg25*, multidrug transporters, and HMG-CoA reductase⁴⁰ while Zhang et al. showed that agricultural fungicides induced cross-resistance to medical triazoles and also observed mutations in *cyp51A* and HMG-CoA reductase⁴¹.

Outbreak analysis

Cryptococcus gattii is less common than *C. neoformans* but can infect immunocompetent individuals. While once considered endemic to tropical and subtropical environments, *C. gattii* outbreaks in the Pacific Northwest necessitated phylogenetic studies to identify outbreak origin, however the clonal nature of *C. gattii* sublineages impeded the ability of multilocus sequence typing (MLST) to resolve variation⁴². Thus, WGS of 118 genomes was employed to identify South America as the probable origin of Pacific Northwest lineages⁴³.

Virulence analysis

The high-resolution nature of WGS has already revealed clinically relevant heterogeneity in opportunistic fungal pathogens of humans. Since the first *C. albicans* genome (strain SC5314) was published in 2004, intraspecies variation has been assayed in several WGS studies. For example, in a 2017 study, WGS and construction of SNP based phylogenies showed that mucosal and bloodstream isolates are organized into separate clades⁴⁴. Similarly, WGS and phenotyping of two clinical *C. albicans* isolates of variable pathogenic potential found that major differentiating genetic variants are located in genes associated with biofilm production and first-line host barriers and vary in a manner that correlates to isolate-specific phenotypic differences⁴⁵.

Clinically informative resolution of intra-specific variation using WGS has not been restricted to *Candida*. Analysis of WGS data from 56 *C. neoformans* strains was recently used to identify intraspecies variation⁴⁶. Integration of this intra-specific heterogeneity with clinical data facilitated the identification of 40 genes putatively associated with human survival, immunologic response or clinical parameters. Using

gene deletion strains for these candidate genes, 6 were found to directly influence murine survival, four of which were novel⁴⁶. These studies highlight how intra-specific variation can be resolved using WGS, and how identifying this heterogeneity can act as the first step towards identifying clinically relevant fungal properties.

1.4.3 Virulence in animal models and its clinical relevance

Modelling the virulence of *A. fumigatus* isolates is an important tool for investigating pathobiology. When isolated from human hosts, clinical data can provide some indication of isolate virulence. For example, *A. fumigatus* strains isolated from patients without symptoms or histological evidence of invasive infection are likely less virulent than those isolated from IA patients. Infection by *A. fumigatus*, however, is infrequent, and the primary conditions of susceptible-hosts are highly variable. For example, the immunogenic profile and post-infection histology of corticosteroid-immunosuppressed transplant recipients is different to that of a chemotherapy patient¹. This means that high-sample size studies with standardisation of potentially confounding host-factors such as primary condition, age, gender and geographic region are difficult to accomplish. The reliability of virulence inferred from clinical origin is thus uncertain. By modelling *A. fumigatus* virulence in model organisms, we can reach higher sample sizes and standardise host-factors. The use of infection models also allows researchers to evaluate cause and effect, and not just correlation, by facilitating the use of knockout studies.

Rodent models have a long history of use in the evaluation of *A. fumigatus* virulence. Amongst the rodents, mice are the most commonly used IA model (85.8%) followed by rats (10.8%) and guinea pigs (3.8%)⁴⁷. Like human hosts, mice are endothermic with internal body temperatures of 37°C⁴⁸. They have both innate and adaptive immune defences. However, there are some differences between rodent and human

immunity that may be relevant to their use as models of invasive aspergillosis. For example, human blood is far richer in neutrophils compared to mice⁴⁹. Human neutrophils also produce antimicrobial peptides not produced by the neutrophils of mice⁴⁹. Furthermore, unlike humans, mice possess a significant amount of bronchus-associated lymphoid tissue⁴⁹. In saying this, mice show the greatest similarity to human biology of the commonly used IA models. As is the case in all models, deviation from a real-world scenario (such as human biology here) may limit the straightforward application of any results, however models are crucial in furthering our fundamental understanding of disease. This is especially true when any uniqueness present in the model is well known. Rodent models are advantageous in that well-established immunosuppression regimens are available as a means of more appropriately modelling host-conditions. These immunocompromised rodent models are frequently used. A recent review found 78% of murine IA models were immunocompromised, most commonly using steroids (44.3%), alkylating drugs (41.9%) or mutation/deletion in the rodents genetic background (18.4%)⁴⁷. Well-established histopathological techniques also exist, and are employed in over half of all IA studies that use mice models (53.2%)⁴⁷. The use of murine models is expensive, high maintenance and presents ethical challenges. It is often difficult to conduct studies with sufficient power to screen for biological properties potentially important in virulence, particularly due to the apparently multi-factorial and nuanced nature of *A. fumigatus* virulence⁵⁰. For these reasons, invertebrate models are often used as a proxy for mammals.

Invertebrate models have economic and bioethical advantages over mammalian models. Invertebrate models used to study fungal infections include nematodes (*Caenorhabditis elegans*), fruit flies (*Drosophila melanogaster*) and the larvae of moths (*Galleria mellonella*) and beetles (*T. molitor*; mealworms)^{51,52}. Unlike *Drosophila* and

nematode models, *G. mellonella* and *T. molitor* larvae can be reared at 37°C, the internal body temperature of humans⁵¹. They can also be inoculated via injection, allowing for precise control over infective load. The major limitation of invertebrate models of IA are the immunological differences between insect and human hosts. For example, while mealworms produce anti-microbial peptides and possess haemocytes capable of internalising and destroying foreign particles, they lack an adaptive immune system, possessing only innate defences⁵³. While major immunological differences between mammals and invertebrates exist, the major risk factor for IA is a compromised innate immune arm, particularly reduced levels of neutrophils capable of destroying both fungal spores and hyphae. Invertebrate models, such as mealworms, provide an environment in which their immune defences represent the major challenges a spore must overcome in the early stages of infection. This may explain why patterns of *A. fumigatus* virulence are often consistent between invertebrate and vertebrate models⁵⁴⁻⁵⁷, although this is not always the case⁵⁸. Due to their low cost and ease of use, invertebrate models often serve as a tool for large scale screening of fungal properties. The results of invertebrate studies can then inform more focused research in mammalian models.

Intra-specific heterogeneity of *A. fumigatus* virulence has been observed in both murine^{59,60} and invertebrate models^{61,62}.

1.5 Aims of this work:

Not all at-risk patients develop IA and the severity of host-damage can vary. For example, the disease occurs in only 7% of acute myeloid leukemia patients²⁰. While this is partly a result of treatment, prophylaxis and host factors, the contribution of intra-specific variation in *A. fumigatus* needs to be evaluated. In this study, I examine intraspecific genotypic and phenotypic variation in relation to the clinical origin of *A.*

fumigatus isolates. Clinical *A. fumigatus* isolates were taken from IA patients (“IA isolates”) and also from patients which had been colonised by the fungus but that did not develop IA (“colonising isolates”). These were characterised with respect to (1) phenotypic properties putatively linked to virulence (radial growth rate, UV resistance, and amphotericin B resistance), (2) genomic content, and (3) virulence in a *T. molitor* larvae model of IA.

Chapter 2 Methods

2.1 Clinical *A. fumigatus* isolates

Clinical *A. fumigatus* isolates were provided by Assoc. Prof. Sharon Chen from the Westmead Clinical School Centre for Infectious Diseases & Microbiology. Isolates were sampled from either IA patients or at-risk but asymptomatic patients colonised by *A. fumigatus*. To prepare single spore suspensions, sabouraud dextrose plates were spot inoculated with 10^6 conidia and incubated for 3-5 days at 37 °C. Each solid culture was flooded with 5 mL of 0.05 % v/v Tween 20-PBS (PBST) solution and conidia were dislodged from the mycelial mass using a sterile cotton swab. The resulting spore solutions were filtered (40 µm) and concentrations evaluated through Neubauer Chamber counts and dilution plating.

2.2 Non-standard statistical terminology

In several experiments below, technical replication occurs on the isolate-level. This leads to a hierarchical data structure when examining the effect of isolate origin on assayed properties. Where this hierarchical structure appears, the effect of clinical origin on the dependent variable of interest has been statistically queried using a 'nested t-test' in GraphPad Prism 8. This non-standard term refers to the fitting of a mixed-effect model where clinical origin is included as a fixed factor and the isolates as random factors. Note that several other statistically valid approaches exist for dealing with hierarchical data, such as reducing clusters to mean values prior to analysis. A mixed effect model approach was chosen as it is robust and powerful, explicitly accounting for clustering without losing information about individual observations⁶³.

2.3 Phenotypic variation amongst clinical *A. fumigatus* isolates

2.3.1 Radial growth rates at 37°C on nutrient rich media

For each *A. fumigatus* isolate, a potato dextrose agar plate was spot inoculated with 10^4 conidia. Cultures were incubated at 37°C over 3 days. Colony diameter was measured at regular intervals and radial growth rate calculated from 5 linearly distributed data points. Experiments were done in triplicate. Statistical significance of the effect of clinical origin on growth rate was identified by a nested t-test. Inter-isolate variation in radial growth rate was evaluated using single factor ANOVA with subsequent all-vs-all Tukey testing. Note that linearity of growth data was evaluated for each replicate and each isolate separately before being used to infer growth rate (Figure A1).

2.3.2 Conidial UV resistance

For each isolate, approximately 200 conidia were spread plated onto malt extract agar (MEA). Five plates inoculated with the same isolate were placed at different positions within a TopSafe PC2 Biosafety cabinet and UV irradiated (1.6 W/m^2) for 1 min (Figure A2). This was repeated for all 15 isolates. Following irradiation of each isolate, the biosafety cabinet was vented for 5 min to prevent ROS accumulation. The sequence in which isolates were irradiated was changed between replicates to achieve uniformity in average position in UV-order after 5 technical replicates. Colony forming units (CFU) on control plates were counted following 24 h of incubation at 37°C. Additional incubation for 24 h at 25°C preceded CFU counting of UV-irradiated plates. Percent survival for each isolate was calculated relative to a non-irradiated control and based on the average CFU counts across the five irradiated plates.

The effect of clinical origin on UV resistance was evaluated using a nested t-test. Isolate-level variation was further resolved statistically using a single-factor ANOVA and subsequent all vs all post-hoc Tukey tests.

2.3.3 Amphotericin B resistance

For each *A. fumigatus* isolate, acute AMB resistance was assayed by exposing 10^5 cfu/mL conidia in malt extract broth with 0.5, 1, or 1.5 $\mu\text{g/mL}$ AMB for 3 h at 37°C. Cultures were serially diluted, plated on sabouraud dextrose agar and incubated at 37°C for 24 h. Percentage spore survival was calculated based on CFU counts relative to drug-free controls. This experiment was repeated to obtain 5 technical replicates for each isolate.

For each AMB concentration, a single factor ANOVA was used to evaluate intra-specific variation in AMB resistance.

Paired t-tests were also used to identify significant differences between 0.5 and 1.5 $\mu\text{g/mL}$ AMB treatments. Isolates showing no significant difference between these low and high drug dosages were classed as resistant while significant variation was taken as an indicator of AMB susceptibility. Note that any large non-biological variation will bias this method towards an assumption of resistance.

2.4 Genomic variation amongst clinical *A. fumigatus*

2.4.1 DNA Isolation

For each of 10 clinical isolates, malt extract broth (20 mL) was inoculated with 10^4 conidia and incubated for 4-5 days at 37 °C with shaking. Fungal biomass was isolated through vacuum filtration and stored at -20°C until use. Genomic DNA was extracted from biomass using the Bioline ISOLATE II Genomic DNA Kit with pre-lysis steps supplemented by mechanical disruption. Lysis buffer (180 μL) and 100mg of biomass

was added to FastPrep Lysing Matrix G before bead milling in a FastPrep-24 (max speed; 30 seconds). RNase A (1 μ L of a 20 mg/mL) solution was then added and samples incubated at 37°C for 30 min. RNase was degraded by adding 25 μ L of Proteinase K solution and incubating at 56°C for 1 hr. Secondary lysis steps and column clean-up were conducted as per the Bioline ISOLATE II Genomic DNA Kit standard protocol.

2.4.2 Library preparation and sequencing

Whole-genome sequencing libraries were prepared from fungal gDNA using the TruSeq DNA PCR-Free Library Prep Kit. Libraries were sequenced using an Illumina NovaSeq 6000 at the Ramaciotti Sequencing Center at UNSW.

2.4.3 Read QC and trimming

FastQC⁶⁴ (v0.11.2) was used to evaluate read quality. The leading 7 bp, final base, Illumina adapters, and low-quality leading and trailing bases (phred < 3) were removed using trimmomatic⁶⁵ (v0.38). Reads were error corrected with LIGHTER⁶⁶ (v1.1.2).

2.4.4 De novo assembly

The SPAdes assembler⁶⁷ (v3.13) was used to produce *de novo* assemblies of the 10 *A. fumigatus* isolates. Quality of assemblies was assessed based on contiguity (scaffold/contig length statistics) and completeness (assessed by BUSCO⁶⁸, v2.0.1). Contiguity statistics were calculated assuming a true genome size of 28.831 Mb (median genome size of *A. fumigatus* genomes deposited in the NCBI genome database).

2.4.5 Chromosomal level scaffolding

Ragout⁶⁹ (v2.2) was used to arrange SPAdes contigs into a chromosomal level assembly based on synteny and phylogenomic relationships with the AF293 reference

genome (NCBI), the draft genome of A1163 (NCBI), and all the other contig-level isolate assemblies. Ragout was run with the solid scaffold flag to minimise false-synteny misassemblies. Ragout scaffolds were aligned to the AF293 genome with NUCmer⁷⁰ (v3.9.4α) and visualised using MUMmerplot⁷⁰ (v3.5). Scaffolds spanning multiple reference chromosomes or intra-scaffold changes in direction of alignment were identified as sites of potential misassembly. Coordinates of potential misassemblies were then identified using GMAJ⁷¹ (30.06.08) and sequence content probed in Geneious⁷² (v10.2.6). Sites of multi-chromosomal alignment or intra-scaffold changes in direction of alignment that occurred in a homopolymer region (>20 bp), or either side of a repeat region (as identified using RepeatMasker⁷³ (v4.0.6) softmasking & coverage levels) were taken as true misassemblies. Contigs were split to allow Ragout to re-arrange them to more appropriate positions. Ragout was re-run on misassemblies-corrected contigs.

2.4.6 Contamination removal

SPAdes contigs unplaced by Ragout were screened for potential contamination and those passing contamination QC were added to the final assembly. Augustus⁷⁴ (v3.2.2) was used to predict protein-coding sequences (species - *E. coli*; mode - intronless). BLASTp⁷⁵ (v2.6.0) of predicted proteins against the nr database was used to identify contigs with non-fungal genes. Hits were filtered and sorted by e-value (>10⁻⁴⁰) and contigs containing proteins with non-fungal top hits, or any non-fungal hits in the top 10 matches were identified.

In order to calculate depth of coverage, raw reads were then aligned to unplaced contigs. Contigs with over 5-fold higher median coverage than that of the placed contigs were suspected of being contaminants and run through a nucleotide BLAST⁷⁵

(v2.6.0) similarity search. Hits were filtered and sorted by e-value ($>10^{-40}$) so that contigs with non-fungal hits in the top 10 matches could be identified.

2.4.7 Genome annotation

Genomes were structurally annotated using the self-training gene-finder GeneMark-ET⁷⁶ (v4.48). Command-line options used are shown below:

```
gmes_petap.pl --soft_mask 1 --cores $CPUS --max_mask 10000 --ET introns.gff3  
--fungus --et_score 5 --sequence genome.fasta --max_intron 4000 \  
--min_gene_prediction 120
```

GeneMark predictions were informed by intron-exon coordinates. These were identified using messenger RNA sequences extracted from the official geneset of the AF293 reference genome. AF293 transcriptomic data was aligned to the genomes of each clinical isolate using GMAP⁷⁷ (v2019-05-12). This alignment data was also run through `augustus_RNAseq_hints.pl`, a script from the Just Annotate My genome⁷⁸ (JAMg) pipeline, to identify intron-exon junctions in the form of a general feature file. Additionally, a high-quality subset of gene data was identified from reference genome peptides using `prepare_golden_genes_for_predictors.pl`, another JAMg script. This 'golden' geneset and alignment information were combined to form the "introns.gff3" file fed to GeneMark.

Note that structural annotation of the sequenced isolate genomes was conducted as an important first step in facilitating more focused assays of genetic heterogeneity in the future. The structural annotations presented have not yet been used in any downstream analysis. Functional annotation and assaying of the heterogeneity of specific genes and pathways are beyond the scope of this thesis.

2.4.8 Variant analysis

The 10 assembled *A. fumigatus* genomes were aligned to the *AF293* reference using parSNP⁷⁹ (v1.2). Single nucleotide polymorphisms were also derived from the core genome alignment using parSNP. Harvest tools⁷⁹ (v1.2) was used to convert data to VCF format. Note that parSNP identifies SNPs based on core-genome alignments, which means SNPs are only identified if present in a genomic region that aligns across all 10 genomes and the *AF293* reference. Variants were filtered using VCFtools⁸⁰ (v0.1.15) to include only bi-allelic sites where the minor allele occurs in at least 2 genomes and the site itself is at least 10 bp away from the nearest SNP. SnpEff⁸¹ (v4.3t) was used to annotate variants based on the RefSeq annotation of the *AF293* reference genome. For each isolate, both intronic and non-synonymous SNPs were extracted, concatenated separately, and used to build phylogenies. Sites of intronic variation classed as splice_site_variants or splice_region_variants by SnpEff were excluded from downstream analysis as they may be affected by selection. Phylogenetic trees were constructed using IQ-Tree⁸² (v1.6.1) with a GTR model and ultrafast bootstrap approximation (10,000 replicates).

2.5 Virulence Assays

Virulence of clinical *A. fumigatus* isolates was established in a *T. molitor* larvae (mealworm) model of invasive aspergillosis. Mealworms were purchased from BioSupplies and reared in all experiments on a diet composed of wheat bran and LSA (linseeds, sunflower seeds and almonds) in a 5:1 ratio. Mealworms were size selected (100-150 mg) and checked for uniformity in colour and active response to physical stimulation before use in any of the experiments below. In all mealworm experiments, mortality was determined by response to physical stimulation. Where not explicitly specified, mealworms were incubated in Petri dishes (10 mealworms/Petri dish;

rearing density: 10 mealworms/58cm²) with 3 mL rearing diet and a slice of frozen carrot for moisture (0.4 cm³; 500 mg; changed daily).

2.5.1 Infection model optimisation

To identify optimal rearing conditions at 37°C, mealworms were incubated for 7 days with either different volumes of rearing diet volume (3 mL or 15 mL) or different amounts of carrot (0, 1, or 2pcs where 1pc = 500 mg). To assess validity of anesthetising mealworms by chilling on ice, mealworms were placed on ice for 5 min or unchilled and incubated at 37°C for 1 week. Each treatment group included 20 mealworms.

Based on the results of these optimisations, all experiments described below used 3 mL bran, a single piece of frozen carrot (~500 mg) and chilling of mealworms for 5 min prior to injection, to render the mealworms docile.

The site of mealworm injection was optimised to reduce mealworm mortality due to physical trauma. Mealworms were injected ventrally with 5 µL of PBST at the base of one of five sternites. Sternites 2-6 were tested (Figure 2). Survival was checked daily over 7 days of incubation at 37°C. Each treatment group included 20 mealworms. Based on the findings of this experiment, mealworms were injected at the base of sternite 5 in all virulence assays.

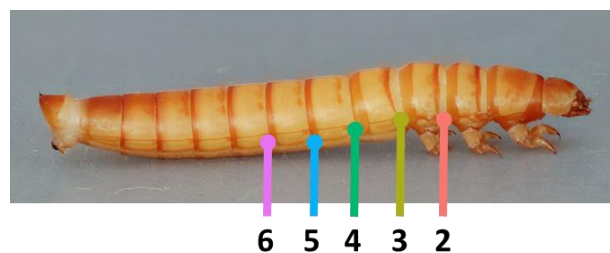


Figure 2.1. Position of 5 different injection sites tested for their effect on mealworm survival. Mealworms were injected at the base of sternites 2-6, where numbering of sternites starts at the prothorax and increases with posteriority⁸³.

To identify the fungal load with the greatest potential for resolving inter-isolate variation, I performed a dose response test of inoculations containing 0, 5, 5×10^1 , 5×10^2 , 5×10^3 , 5×10^4 , 5×10^5 , and 5×10^6 spores of isolate AF01. After inoculation, mealworms were kept at 37°C for 1 week with mortality checked daily. This treatment group also included 20 mealworms. The entire experiment was performed twice to assess level of variability.

In all experiments above Kaplan-Meier survival was calculated and log-rank testing used to evaluate differences amongst treatment groups. In experiments with more than two treatment groups, post-hoc testing was composed of Bonferroni corrected pairwise log-rank tests.

2.5.2 Inter-isolate variation in virulence of *A. fumigatus* isolates

The virulence of all 15 clinical *A. fumigatus* isolates was evaluated. For each isolate, 20 mealworms were inoculated with 5×10^4 conidia (in 5 μ L PBST) at sternite 5, and incubated for 1 week at 37°C. Each experimental replicate included three control groups: (1) Mealworms injected with sterile PBST, (2) Mealworms pierced at sternite 5 but with no solution injected, (3) Mealworms chilled on ice but otherwise untreated. PBST vehicle controls were the only controls included in downstream analysis. The others served only as QC metrics for each experimental run. A total of three replicate experiments were conducted.

Kaplan-Meier analysis was used to visualise survival curves and calculate median survival time. Associations between clinical origin and median survival time were evaluated using a nested t-test. Note that the virulence assays were conducted over only 7 days, limiting the statistical power of this approach. Nonetheless, Kaplan-Meier curves and median survival times of clinical *A. fumigatus* isolates in *T. molitor* larvae

have been reported in this study as they have not yet been published and may be beneficial as points of reference for future research in this developing model. Thus, when drawing conclusions about associations between clinical origin and virulence of isolates, greater emphasis was placed on more powerful statistical methods of resolving virulence, specifically Cox-regression modelling.

The Cox proportional hazards model is a semiparametric regression model that identifies associations between predictors (such as clinical origin) and time-to-event using hazard functions which represent the risk of dying at time t ⁸⁴. A Cox regression model was fitted to the survival data (excluding controls). Due to the hierarchical nature of the virulence data, a standard Cox regression model could not be used as isolate-level replication would artificially inflate power. The use of a mixed effects Cox regression model accounts for this nested data structure⁸⁵. Thus, clinical origin was added to the model as an independent variable, and replicate group added as a frailty term (i.e. a random effect). This model was used to evaluate differences in risk-of-mortality of colonising and IA isolates.

To evaluate variation in virulence at the isolate level, another Cox regression model was fit to the survival data. PBST vehicle controls were included as the reference group. Isolate ID was included as an independent variable, with replicate group added as a frailty term. This model was used to quantify the fold-increase in hazard (probability of mortality at any given timepoint) relative to the reference group, for each isolate. This metric of virulence represents the exponent form of the β coefficient of the Cox-regression equation and is the metric being referred to when describing “virulence in *T. molitor* larvae”.

All survival data analysis was conducted in R (version 3.6.0) using the survival and survminer packages.

2.5.3 Correlation between inter-isolate variation and virulence in invertebrate model

Pearson correlation between virulence in *T. molitor* larvae and both radial growth rate and UV resistance was evaluated in R using the Hmisc and stats packages. Associations between amphotericin resistance and virulence was evaluated using Welch's t-test.

Chapter 3 Phenotypic variation amongst clinical *A. fumigatus* isolates

3.1 Introduction

To evaluate the contribution of phenotypic and genotypic variation in clinical *A. fumigatus* isolates to disease severity, I first examined whether human colonising *A. fumigatus* isolates differed from IA isolates in phenotypic properties putatively associated with virulence. These properties include growth rate in a nutrient rich environment, conidial UV resistance and resistance to the antifungal drug, Amphotericin B.

The growth rates of *A. fumigatus* isolates have been found to be highly variable on both nutrient rich and minimal media¹⁹. Further, several studies have found *in vitro* growth rate of *A. fumigatus* strains to positively correlate with *in vivo* virulence, again using both nutrient rich and minimal media^{21,86}. Such a correlation is not always present⁵⁵.

Resistance of *A. fumigatus* conidia to solar UV radiation and UV-induced reactive oxygen species (ROS) is important for the survival of airborne conidia⁸⁷. Conidial defences against UV-induced oxidative damage, such as cell wall melanin, have been implicated in pathogenesis by promoting pre-germination concealment of immunogenic pathogen-associated molecular patterns (PAMPs)^{88,89}, evasion of internalization by phagocytes⁹⁰ and persistence within immune and alveolar epithelial cells⁹¹⁻⁹⁴. Several previous studies have also identified heterogeneity in pigmentation of clinical *A. fumigatus* isolates^{19,95}, likely representative of underlying variability in melanin biosynthesis pathways.

Amphotericin B (AMB) is a polyene class broad spectrum antifungal used in the treatment of IA⁹⁶. Thus, AMB resistance is directly related to virulence through its effect on treatment efficacy. AMB targets ergosterol in the fungal plasma membrane and can induce oxidative stress.⁹⁷ Proposed mechanisms of resistance include reduced membrane ergosterol levels and upregulation of anti-ROS enzymes.¹⁷ Ergosterol biosynthesis is linked to siderophore production via a shared precursor mevalonate⁹⁸. Thus, AMB resistance may contribute to virulence via effects on oxidative stress biology or iron sequestration.

This chapter describes variation observed amongst clinical IA and colonising *A. fumigatus* isolates in (1) radial growth rate in a nutrient rich environment and (2) conidial UV resistance. Evaluation of variation in AMB resistance amongst colonising *A. fumigatus* isolates is also presented.

3.2 IA isolates grow more slowly than colonisers on nutrient rich media

Radial growth rates on PDA at 37°C were calculated by measuring colony diameter over time and fitting a simple linear model. For all isolates, the 5 timepoints at which colony diameter was measured successfully captured linear regions of growth, with all R² values over 0.99 (Figure 3.1; Figure A1).

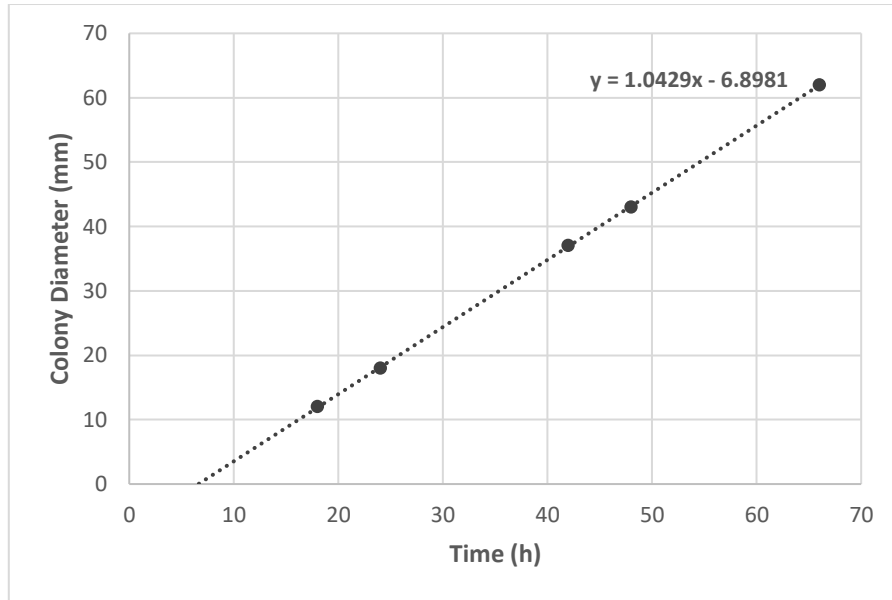


Figure 3.1. Linearity of *Aspergillus fumigatus* growth on PDA at 37°C. Data represents a single technical replicate for colonising isolate AF01. Radial growth rate was taken as the slope of the simple linear model (1.0429 mm/h). Timepoints chosen span a region of linear growth, as evidenced by high goodness-of-fit ($R^2 = 0.99995$). Growth rate data for all isolates and technical replicates are shown in Figure A1.

The radial growth rate of 10 colonising and 5 IA isolates was evaluated on PDA at 37°C (Figure 3.2A). Growth of IA isolates was significantly slower than colonising isolates (Figure 3.2B). On average, colonisers grew $124.1 \pm 46.69 \mu\text{m/h}$ faster than IA isolates. The trend was consistent, with 4 of the 5 slowest growers being IA isolates. The only major exception was AF12, an IA isolate with the second highest growth rate. Significant isolate level variation was also observed (Figure 3.2C).

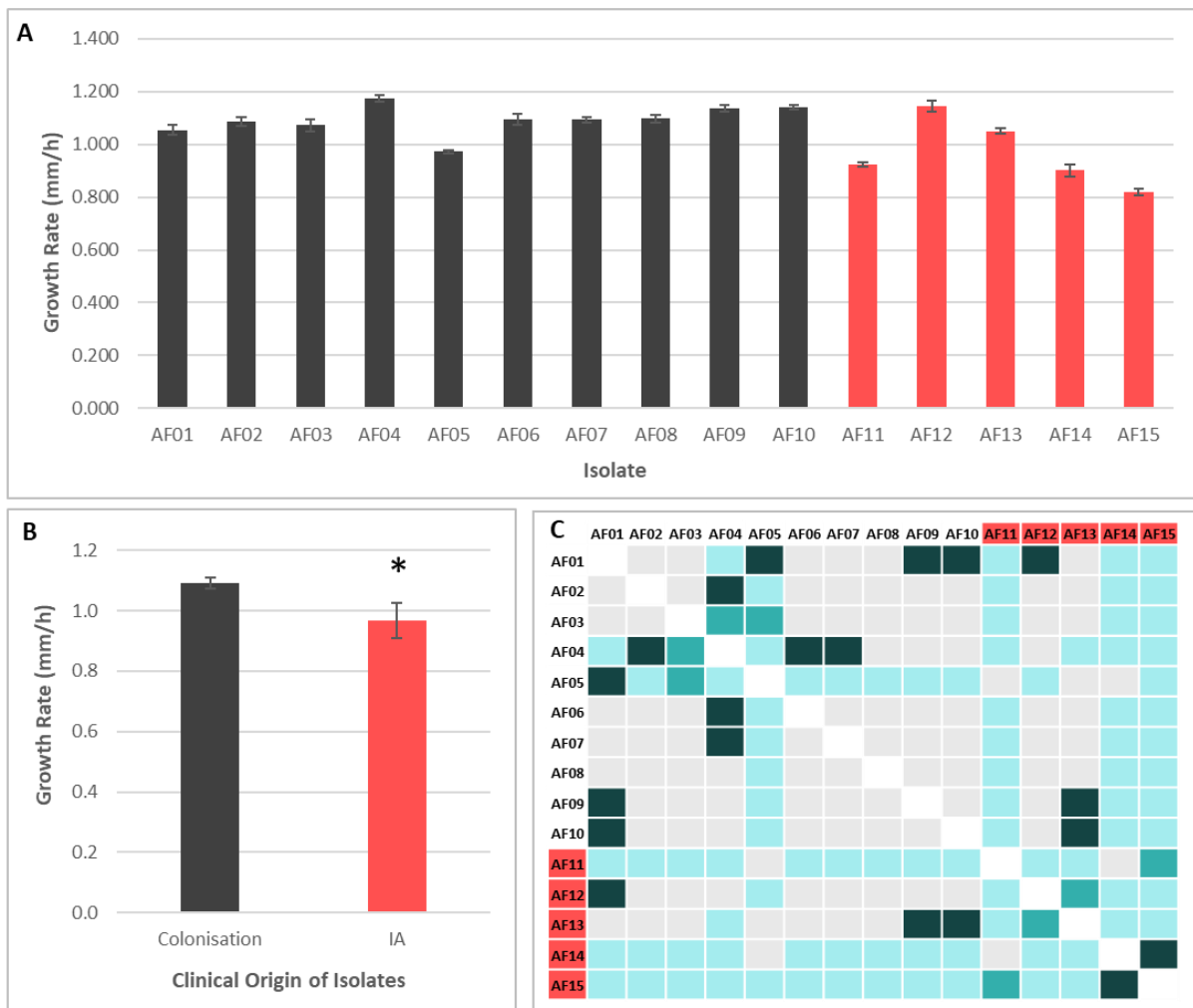


Figure 3.2. Radial growth rates of clinical *A. fumigatus* isolates on potato dextrose agar (37°C). (A) Growth rate of 10 colonising (grey) and 5 IA-associated (red) isolates (mean \pm SE; $n=3$). Significant isolate-level variation was observed (ANOVA: $p = 3.32 \times 10^{-16}$). (B) Mean growth rate of colonising and IA isolates (\pm SE; colonising: $n=10$, IA: $n=5$). Colonising isolates grew significantly faster than IA isolates (nested t-test: $p=0.0197$). (C) Isolate-level pairwise comparison of growth rates by Tukey test ($p < 0.001$: dark blue, $p < 0.01$: medium blue, $p < 0.05$: light blue, $p > 0.05$: grey).

3.3 IA isolates and colonisers show similar conidial UV resistance

The conidial UV resistance of 10 colonising and 5 IA isolates was evaluated with irradiation for 1 min at 1.6 W/m² (Figure 3.3A). *A. fumigatus* isolates of differing clinical origin did not differ significantly in conidial resistance to UV irradiation (Figure 3.3B). Significant variation was observed on the isolate level with colonising isolate AF01 possessing UV resistance significantly higher than the lowest 5 isolates (Figure 3.3C).

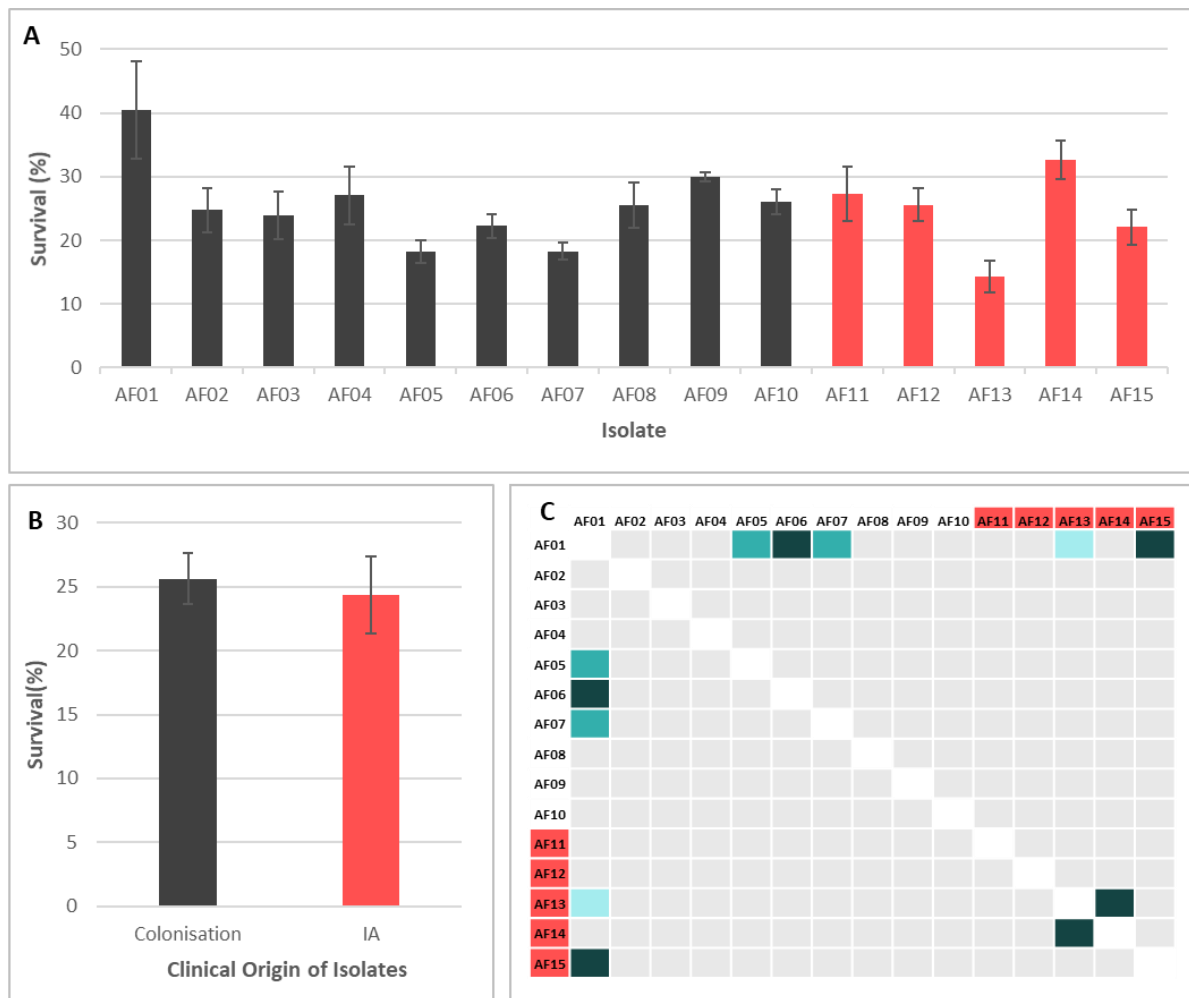


Figure 3.3. Resistance of clinical *A. fumigatus* isolates to UV irradiation (1 min at 1.6 W/m²). (A) UV resistance of 10 colonising (grey) and 5 IA-associated (red) *A. fumigatus* isolates (mean \pm SE; n=5). (B) Mean growth rate of colonising and IA isolates (\pm SE; colonising: n=10, IA: n=5). UV resistance of colonising and IA isolates did not differ significantly (nested t-test: p= 0. 728). (C) Significant variation in UV resistance was observed amongst the 15 isolates assayed (ANOVA: p = 5.95x10⁻⁶). Matrix represents results of pairwise Tukey tests (p < 0.001: dark blue, p < 0.01: medium blue, p < 0.05: light blue, p > 0.05: grey).

3.4 AMB resistance varies amongst colonising isolates.

The AMB resistance of 10 colonising *A. fumigatus* isolates was evaluated. Isolates did not differ significantly in their survival following exposure to 0.5, 1, or 1.5 μ g/mL AMB (Figure 3.4). Differences in isolate UV resistance were resolved by comparing survival following treatment at high and low levels of exposure. Isolates were classified as AMB

sensitive if survival at 0.5 µg/mL AMB and 1.5 µg/mL AMB significantly differed. Of 10 colonising isolates, 6 were classed as sensitive (Figure 3.5). No data is available for IA isolates, due to availability at the time of testing. Colonising isolate data is presented nonetheless, as it may be informative when integrated with virulence data described in chapter 5.

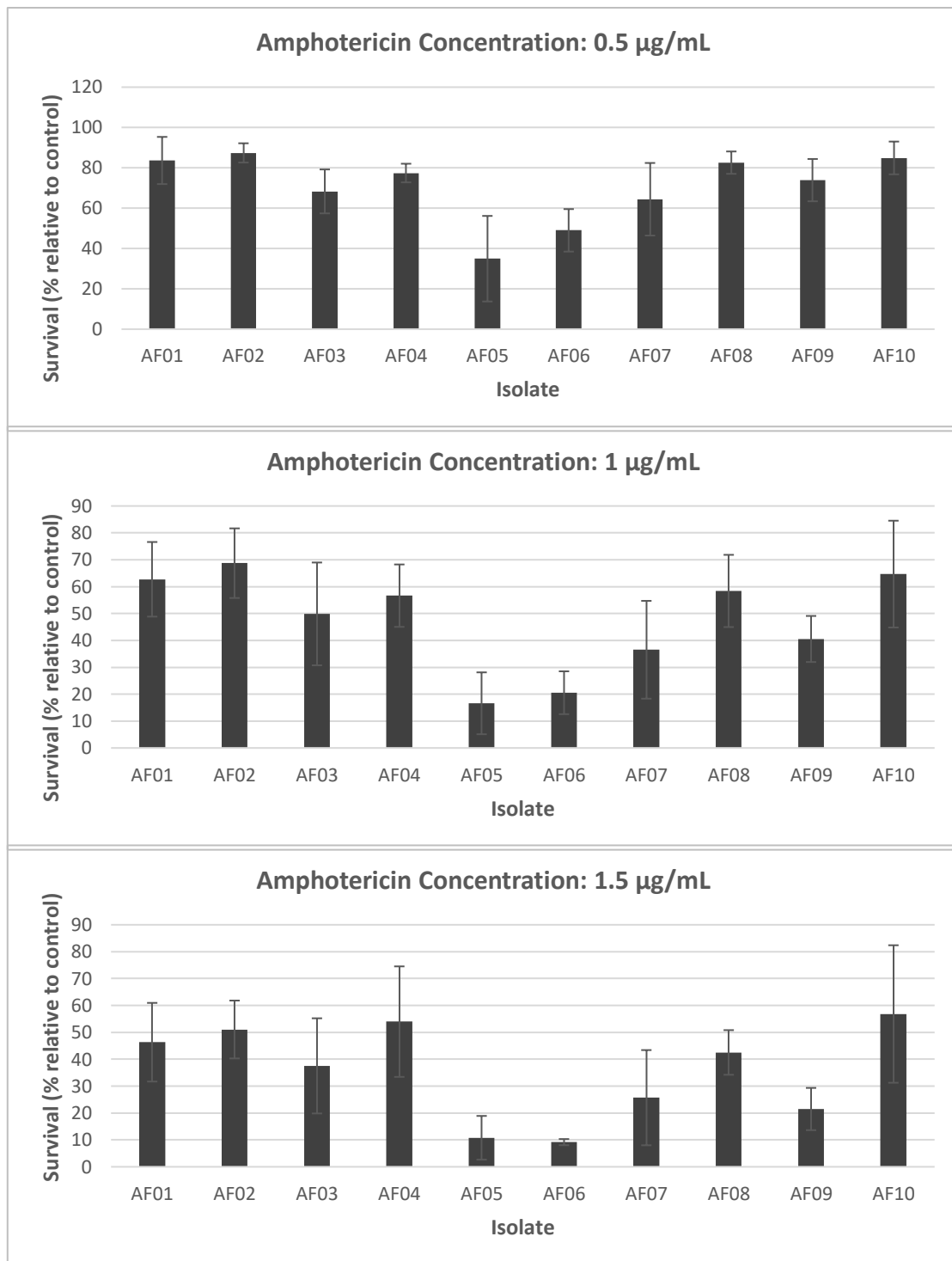


Figure 3.4. Resistance of colonising *A. fumigatus* isolates to Amphotericin B exposure (3 h; 37°C; AMB concentrations: 0.5, 1.0 and 1.5 µg/mL). Data is expressed as a percentage of CFUs surviving relative to non-exposed controls. Data represents mean and standard errors from three replicate experiments for each isolate. Data was analysed by one-way ANOVA (0.5: $p = 0.079$, 1.0: $p = 0.166$, 1.5: 0.254).

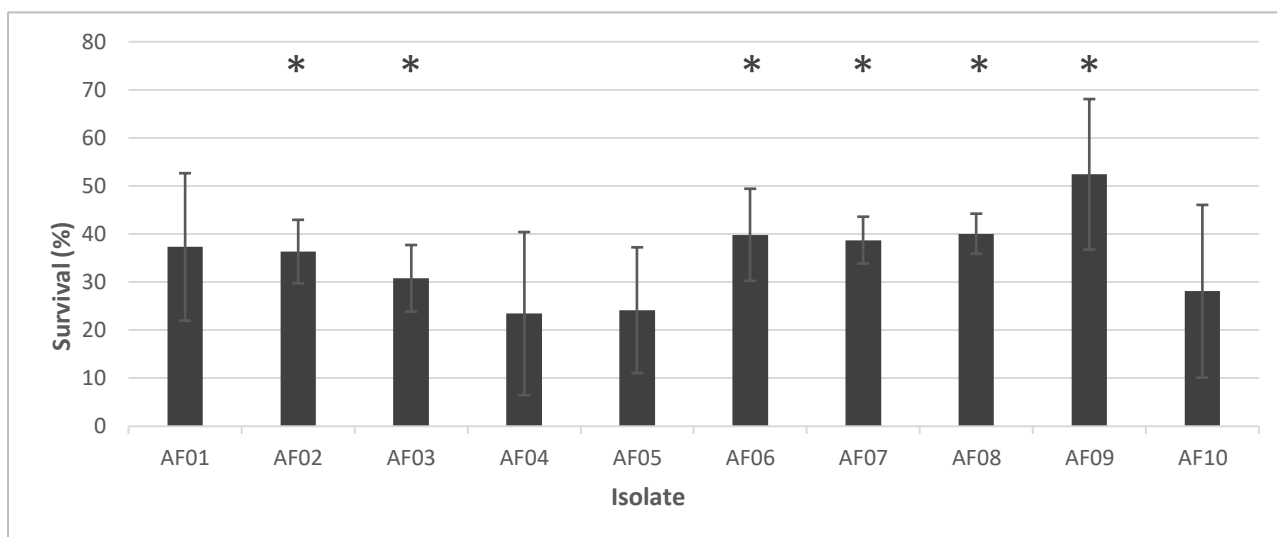


Figure 3.5. Heterogeneity in Amphotericin B resistance of colonising *A. fumigatus* isolates. Data represents difference in survival between spores exposed to 0.5 and 1.5 µg/mL AMB for 3 h at 37°C (mean ± standard error, n=3). Isolates showing significant decreases in survival when treated with the higher AMB concentrations were considered sensitive ($p < 0.05$: *). Data was analysed using a paired t-test (H_A : diff > 0).

Chapter 4 Genotypic variation amongst clinical *A. fumigatus* isolates

4.1 Introduction

In 2005, the first *Aspergillus fumigatus* genome sequence was published⁹⁹. The sequenced strain was AF293, a common laboratory strain derived from an invasive aspergillus isolate. Comparative genomics approaches rapidly revealed important characteristics of the genome. For example, the *A. fumigatus* genome was compared with the genomes of *Aspergillus nidulans* and *Aspergillus oryzae* revealing low intra-genus amino acid identity, as well as a genomic capacity for heterothallic sexual reproduction in *A. fumigatus*¹⁰⁰. As cost and throughput of WGS technologies improved, intraspecies studies of *A. fumigatus* have become more common, with WGS capable of resolving subtle genomic variation even amongst clonally related strains.

Assaying intra-specific genomic variation in *A. fumigatus* using WGS has proved useful as a tool for illuminating clinically relevant phenomena. For example, the study of *A. fumigatus* sequentially isolated from infected patients has revealed selection for several azole-resistance associated genomic variants including *cyp51A* SNPs³³, tandem repeats in the *cyp51* promoter³⁵ and a *hapE* SNP¹⁰¹. Azole-unrelated variants also selected for include Snf1 kinase and RNA polymerase II transcription factor SNPs³³. Despite the increasing frequency of *A. fumigatus* WGS studies, clear genetic distinctions between clinical subtypes of differing severity remain elusive.

In other fungal pathogens, WGS of genomic variation amongst isolates of differing pathogenic potential have identified clinically associated differences. For example, SNP based phylogenies have shown mucosal and bloodstream *C. albicans* isolates are organized into separate clades⁴⁴ and WGS and phenotyping of two clinical isolates

of variable pathogenic potential indicate major differentiating genetic variants are located in genes associated with biofilm production and first line host barriers and vary in a manner that correlates to isolate-specific phenotypic differences⁴⁵. Similarly, WGS of 56 *C. neoformans* strains has been used with GWAS and PCA analysis, deletion studies and animal models to identify virulence factors⁴⁶. In the present study, genomes of *A. fumigatus* clinical isolates that either caused IA or were colonising a human host were sequenced to further understanding of the clinical relevance of intra-specific genomic variation.

Methods used to examine intra-specific variation often require the identification of single nucleotide polymorphisms. The identification of SNPs typically starts with the alignment of reads or genome assemblies to a reference genome. The nucleotide composition of each genome at these sites of variation can then be retrieved, concatenated and used to evaluate phylogenetic relationships between different samples.

Phylogenetic frameworks are important tools to interpret SNP based genetic relatedness. Genetic distance between two genomes can result from differences in selection pressures, but also occurs over time as genomes pseudo-randomly mutate, independent of selection. If a certain set of SNPs are present in clinical subgroups of interest, it can be important to know whether this is a result of selective pressure or simply selection-independent evolutionary relatedness. Exonic SNPs are highly susceptible to selection pressures as they can directly inform protein structure but are also affected by random mutation events that occur over time. Intronic SNPs, however, particularly those outside of splice sites, can be used to construct phylogenetic trees that estimate selection-independent genetic relatedness. Note that this is only an estimate. Introns are functionally significant in the regulation of transcription¹⁰², and

thus not invisible to selection. However, the higher sequence variability and evolutionary neutrality of introns relative to exons make them useful in examining intra-specific relatedness¹⁰³. Once an evolutionary timeline is built, it can be used to interpret genetic relatedness quantified from variant sites more visible to selection.

This chapter describes the construction of clinical *A. fumigatus* genomes from WGS data, identification of inter-isolate genomic variation, and evaluation of clinical origin dependent clustering of this inter-isolate variation.

4.2 Quality control and trimming of raw reads

Paired-end genome sequencing of 10 clinical *A. fumigatus* isolates yielded a total of 46,486,871 151 bp sequences (Table 1). Reverse reads showed slightly lower quality than forward reads, however average per-base phred scores were above 32 at all sites (Figure 4.1).

Table 1. Raw read summary statistics from Illumina paired-end sequencing of 10 clinical *A. fumigatus* isolate genomes.

Property	Reads pairs
Total Sequences	46486871
Sequences flagged as poor quality	0
Average Sequence length (bp)	151
GC (%)	48

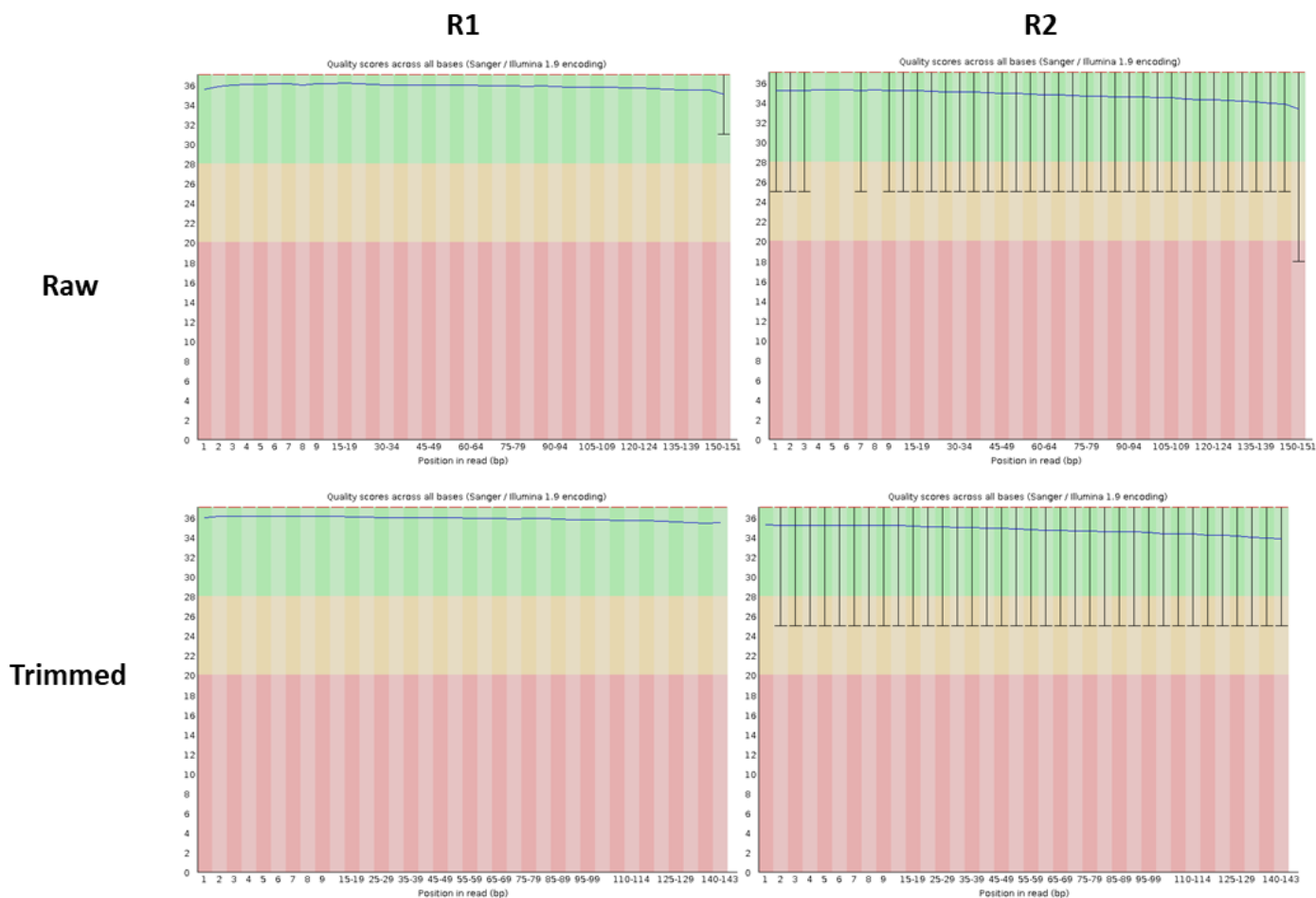


Figure 4.1. Per base sequence quality (phred scores) of forward (R1) and reverse (R2) reads generated from paired-end Illumina sequencing of 10 clinical *A. fumigatus* genomes.

Per base sequence content of raw reads showed biases in both the leading and trailing bases (Figure 4.2). Trimming the first 7 bases, the 151st bp, Illumina adaptor content, and any leading/trailing bases with phred-scores below 3 was sufficient to remove sequence content bias. Reads were then subsampled to 30x depth of coverage as assemblers operate best within that target range. Note that trimming of the ends does lead to a reduction of total read length, but an average of 143 bp is still substantial for short-read assembly.

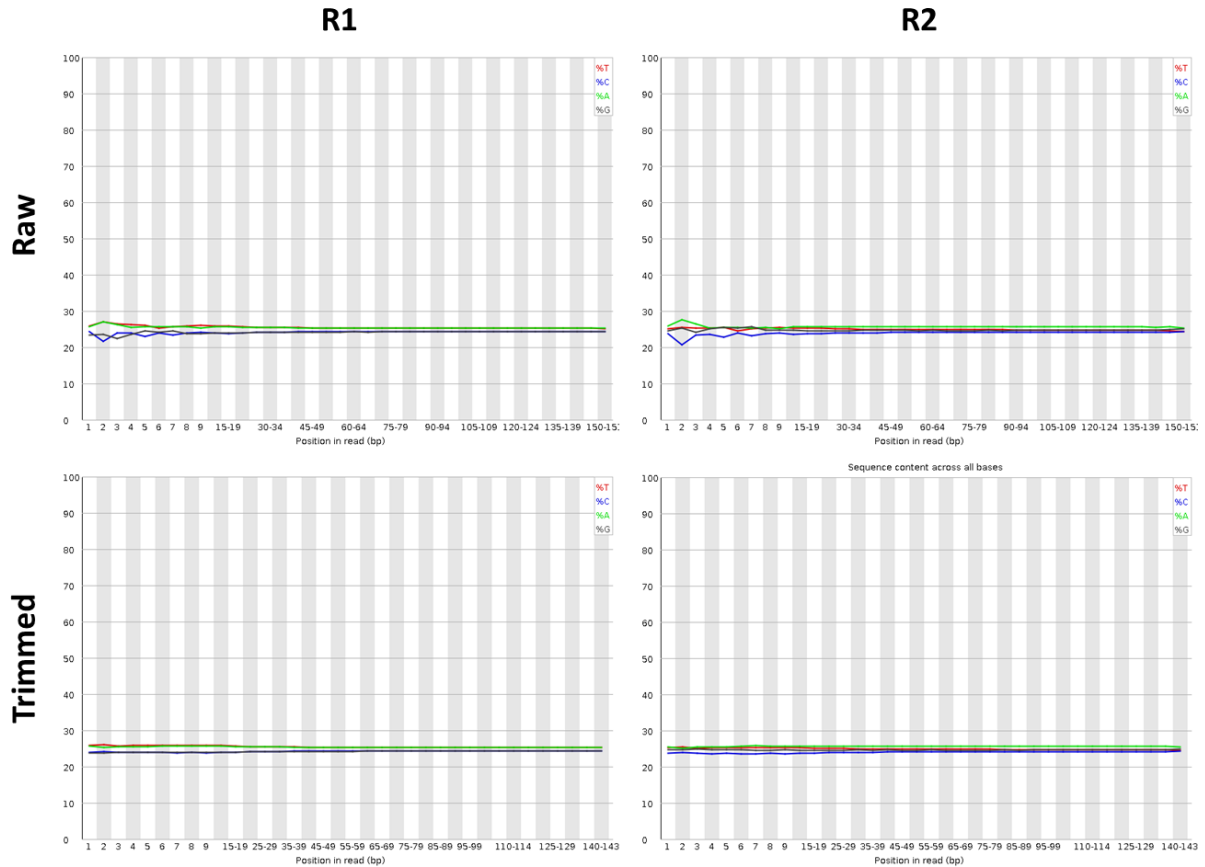


Figure 4.2. Per base sequence content of forward (R1) and reverse (R2) reads generated from paired-end Illumina sequencing of 10 clinical *A. fumigatus* genomes. Raw reads showed sequence content biases (top panel). Trimming the leading 7 bp, the 151st bp, adapter content and low quality leading/trailing bases (phred score < 3) removed regions showing positional biases (bottom panel).

4.3 SPAdes assembly

Trimmed, subsampled and error corrected reads were assembled using SPAdes. Both contig-level and scaffold-level assemblies were constructed (Table 2). At the contig-level, genomes constructed ranged from 28-29 Mb with NG50 averaging 40.9 (assuming true genome size of 28.831 Mb). Scaffolds showed increased contiguity, with average NG50 decreasing to 34.8.

Table 2. Contiguity statistics of SPAdes contigs and scaffolds generated assuming genome size of 28.8 Mb

	Property	AF01	AF02	AF03	AF04	AF06	AF10	AF11	AF12	AF13	AF14
Contigs	Total Size (bp)	28401846	28407426	28760795	28439013	28524037	28789999	29031520	28637354	28672234	28964372
	Contig Number	5676	5656	4854	4930	5389	5885	10679	5007	6067	5840
	Shortest (bp)	56	56	50	56	56	56	53	56	56	56
	Longest (bp)	612747	638238	612770	624750	624828	709162	699391	961310	722484	793871
	NG50 length (bp)	223909	259497	252318	215142	185738	236969	185433	266532	182557	266630
	NG50 count	37	37	41	42	48	38	49	32	50	35
	NG100 number	5676.0	5656.0	4854.0	4930.0	5389.0	5885.0	7260.0	5007.0	6067.0	3675.0
Scaffolds	Total size (bp)	28402990	28408277	28762536	28440462	28525202	28791359	29033859	28638234	28673262	28965797
	Scaffold Number	5649	5625	4819	4886	5349	5850	10626	4979	6025	5808
	Shortest (bp)	56	56	50	56	56	56	53	56	56	56
	Longest (bp)	828731	791597	662413	705843	624828	777272	719509	961310	837136	842094
	NG50 length (bp)	299511	323231	285207	249783	198603	274112	224074	299016	227641	322614
	NG50 count	31	32	35	37	40	35	40	29	39	30
	NG100 number	5649.0	5625.0	4819.0	4886.0	5349.0	5850.0	7171.0	4979.0	6025.0	3624.0

4.4 Chromosomal level scaffolding

Chromosomal level scaffolding was achieved using Ragout and the AF293 reference genome. Alignment of spades scaffolds to the reference genome showed several scaffolds spanned multiple reference chromosomes in gapped regions (where contigs were joined together), representing putative misassemblies. Thus, spades contigs were used in the reference-based Ragout assembly.

Several Ragout scaffolds constructed from SPAdes contigs still aligned to multiple reference chromosomes when Ragout was not permitted to fragment the contigs (Figure 4.3A). Coordinates of alignment breakpoints were identified using Gmaj (Figure 4.3B) and sequence content was investigated in Geneious. Several breakpoints occurred in homopolymer regions within contigs (Figure 4.3C). Similar misassembly-inducing homopolymers were identified in regions where chromosomal-level scaffolds aligned to the reverse complement of the AF293 reference (Figure 4.4A). These were taken as misassemblies and contigs were split in the centre of the homopolymer regions. In one case where a scaffold aligned to multiple reference chromosomes, no homopolymer region was identified. There was, however, a high-coverage segment adjacent to the breakpoint (Figure 4.4B). As this high-coverage segment likely represents a repeat region that would be difficult for assemblers to correctly place, this was also taken as a misassembly. For all 10 genomes assembled, a total of 7 contig-level misassemblies were identified, with contigs split either in the centre of homopolymer regions or at the edge of a repeat (Table 3). Ragout was then re-run with the manually corrected contigs to produce more accurate chromosome level assemblies (Table 4).

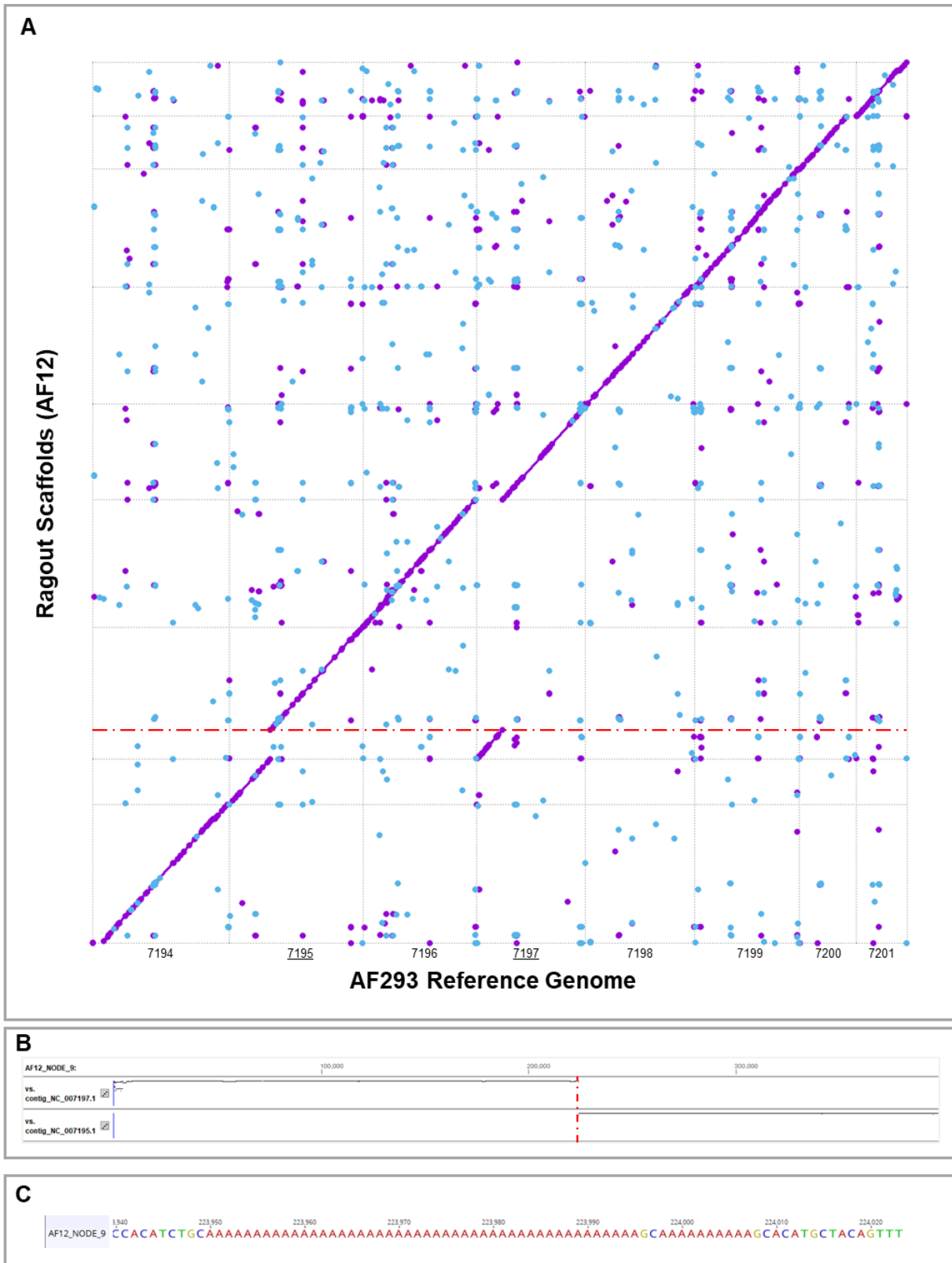


Figure 4.3. Evidence of misassemblies in SPAdes contigs. (A) Example of a Ragout scaffold aligning to multiple AF293 reference chromosomes, representing either chromosomal translocation or contig-level SPAdes misassembly. (B) Coordinates of alignment breakpoints could be identified using Gmaj. (C) In most cases, these alignment breakpoints occurred in homopolymeric regions.

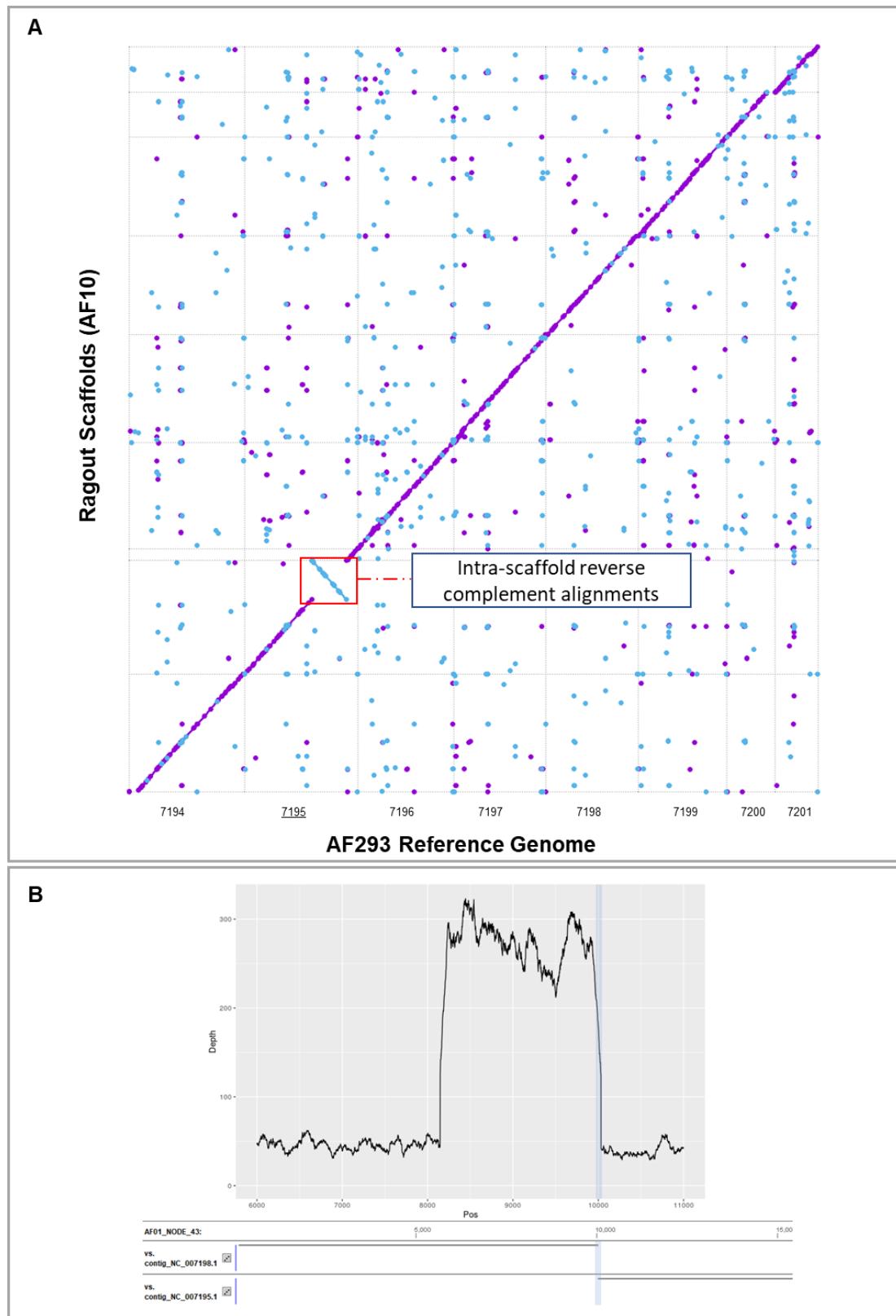


Figure 4.4. Detection of potential misassembly sites in the absence of multi-chromosome spanning SPAdes contigs. (A) Ragout scaffold aligning to reverse complement of AF293 reference, representing either an inversion or contig-level SPAdes misassembly. (B) Increased coverage in the region flanking the alignment breakpoint suggests a repeat-induced SPAdes misassembly is the most probable explanation.

Table 3. Complete list of contigs split based on evidence of SPAdes misassembly

Isolate	Contig of Interest	Reason for investigation	Chromosomes spanned	Gmaj Split	Evidence of misassembly	Split Point
AF01	AF01_NODE_13	Spanned multiple chromosomes	194;195;198	Y	Homopolymer: 45T	42806
	AF01_NODE_43	Spanned multiple chromosomes	194;195;198	Y	High coverage (5 fold increase) – Repeat	8146 & 10031
AF02	AF02_NODE_22	Spanned multiple chromosomes	196;197	Y	Homopolymer: 48T	43771
AF04	AF04_NODE_75	Spanned multiple chromosomes	200;201	Y	Homopolymer: 50A	63794
AF10	AF10_NODE_31	intra-scaffold inversion	195	Y	Homopolymer: 42T	7614
AF11	AF11_NODE_61	Spanned multiple chromosomes	197;200	Y	Homopolymer: 48T	67964
AF12	AF12_NODE_9	Spanned multiple chromosomes	195;197	Y	Homopolymer: 46A	223972

Ragout scaffolding of contigs split at sites of putative misassembly successfully produced chromosomal level scaffolds for all but one isolate (Table 4). In AF06, one of the *A. fumigatus* chromosomes is represented by two different scaffolds. Failure to collapse these two fragments may represent significant biological deviations from the reference at the site of fragmentation, an algorithmic quirk of Ragout, or an undetected misassembly.

Table 4. Chromosomal level scaffolds produced by Ragout from misassembly-corrected SPAdes contigs

Property	Total Size (Mbp)	Scaffold Number	Shortest (bp)	Longest (bp)
AF01	28.21	8	1,708,330	4,741,037
AF02	28.31	8	1,678,478	4,789,096
AF03	28.37	8	1,712,250	4,749,378
AF04	28.35	8	1,724,280	4,761,457
AF06	28.32	9	393,769	4,505,555
AF10	28.36	8	1,696,756	4,753,369
AF11	28.23	8	1,681,522	4,893,088
AF12	28.43	8	1,691,150	4,766,785
AF13	28.43	8	1,728,383	4,752,011
AF14	28.25	8	1,696,692	4,757,815
AF293*	29.38	8	1,833,124	4,918,979

* Reference genome

4.5 Contamination removal

Not all contigs from the SPAdes assemblies were used in the reference guided chromosome-level Ragout assembly. Such “unplaced” contigs may represent novel genomic content, which doesn’t align to the reference, or contamination. Augustus was used to predict genes in these unplaced contigs. BLASTp of predicted proteins against the NCBI non-redundant (nr; 08/01/19) database yielded no evidence of bacterial contamination. For all isolates, the top 10 hits (after an e-value cut-off of $< 10^{-40}$) were fungal. A nucleotide BLASTn against the NCBI nucleotide database also identified no bacterial DNA in top-10 hits.

After being screened for contamination, unplaced contigs were included in the final assemblies.

4.6 Final frozen assemblies

The contiguity of the final assemblies is much lower than the original Ragout scaffolds (Table 5). This is due to ‘dilution’ of the 8 chromosomal-length scaffolds produced by Ragout with thousands of “unplaced” contigs that may include novel genetic information. Note that many of these “unplaced” contigs are uninformative. For example, many are composed of near-homopolymeric sequences. Ideally, annotation data and sequence-content information can be used to filter out uninformative contigs, however as their inclusion does not negatively impair downstream analysis apart from increasing computation time, such work was considered beyond the scope of this thesis. All assemblies included at least 98% of BUSCOs (Benchmarking Universal Single-Copy Orthologs) typically present within the genome of ascomycetes (Table 6). This suggests sequencing depth was sufficient to capture an appropriate level of genetic information.

Table 5. Quality of frozen *A. fumigatus* assemblies.

Property	AF01	AF02	AF03	AF04	AF06	AF10	AF11	AF12	AF13	AF14	AF293*
Total Size (bp)	29.51	29.52	29.86	29.48	29.62	29.96	30.12	29.74	29.87	29.95	29.38
Scaffold Number	5,477	5,450	4,662	4,704	5,159	5,654	10,401	4,810	5,812	5,653	8
Shortest (bp)	56	56	50	56	56	56	53	56	56	56	1,833,124
Longest (Mbp)	4.74	4.79	4.75	4.76	4.51	4.75	4.89	4.77	4.75	4.76	4.92
NG50 length (bp)	3.98	4.03	4.05	4.00	3.98	4.03	3.93	4.09	4.02	3.89	3.95
NG50 count	4	4	4	4	4	4	4	4	4	4	4
NG100 number	130	136	20	64	52	15	175	19	28	36	8

* Reference genome

Table 6. BUSCO completeness of clinical *A. fumigatus* genome assemblies

Property	AF01	AF02	AF03	AF04	AF06	AF10	AF11	AF12	AF13	AF14	AF293*
Complete BUSCOs (C)	1296	1296	1296	1296	1297	1296	1296	1296	1297	1296	1295
Complete and single-copy BUSCOs (S)	1293	1293	1293	1293	1294	1293	1293	1293	1294	1293	1292
Complete and duplicated BUSCOs (D)	3	3	3	3	3	3	3	3	3	3	3
Fragmented BUSCOs (F)	7	7	7	7	6	7	7	7	6	7	7
Missing BUSCOs (M)	12	12	12	12	12	12	12	12	12	12	13
Total BUSCO groups searched	1315	1315	1315	1315	1315	1315	1315	1315	1315	1315	1315

* Reference genome

4.7 Genome annotation

Each of the 10 clinical *A. fumigatus* genomes contained between 8035-9365 genes (Table 7). The average number of introns and exons per gene were consistent across all isolates, but higher than the reference genome. The AF293 reference genome also has 260 more genes than the most gene-dense isolate. Determination of whether these trends are biological, or a result of annotation biases is beyond the scope of this thesis.

Table 7. Structural annotation of clinical *A. fumigatus* genomes.

Assembly	Genes	Introns	Exons	Introns per gene	Exons per gene
AF01	9175	20802	29977	2.27	3.27
AF02	9267	20944	30211	2.26	3.26
AF03	9184	20805	29989	2.27	3.27
AF04	9247	20919	30166	2.26	3.26
AF06	9282	20980	30262	2.26	3.26
F10	9365	21112	30477	2.25	3.25
AF11	9252	20906	30158	2.26	3.26
AF12	9318	21050	30368	2.26	3.26
AF13	8375	18915	27290	2.26	3.26
AF14	8035	18140	26175	2.26	3.26
AF293¹	9625	18625	28251	1.94	2.94

¹Statistics for AF293 were extracted from the RefSeq annotation. The public genome was not re-annotated.

4.8 Variant analysis

A total of 102,951 SNP sites were identified in the core-genome of *A. fumigatus* when SNPs were called against the *AF293* reference genome. At 160 sites, variation was not bi-allelic, and so filtered out. Of the remaining 102,791 variants, less than half (45,196) had a minor allele that occurred in more than one genome. Thinning of sites of variation such that no two occur within 10 bp of one another yielded a final core set of 41,432 SNPs. SNPeff annotation revealed that most of these SNP sites were intergenic (56%). A total of 8000 sites with putative non-synonymous mutations and 2,077 intronic variant sites were identified.

Consensus phylogenetic trees constructed from intronic variation showed no strong clustering based on clinical origin (Figure 4.5A). Three of the four IA isolates sequenced were most closely related to a colonising isolate. Low bootstrap support values surrounding AF14 suggest that different sections of the genome show similarity to different isolates. Removal of AF14 prior to tree construction produced a consensus tree with identical topology but higher branch support (Figure 4.5B). This consolidates AF14 as the source of noise.

Phylogenetic trees constructed from non-synonymous SNPs showed near-identical topology to the intronic consensus tree (Figure 4.5C). This suggests that selection pressures are not causing widespread single nucleotide variation in the core-genome of *A. fumigatus*.

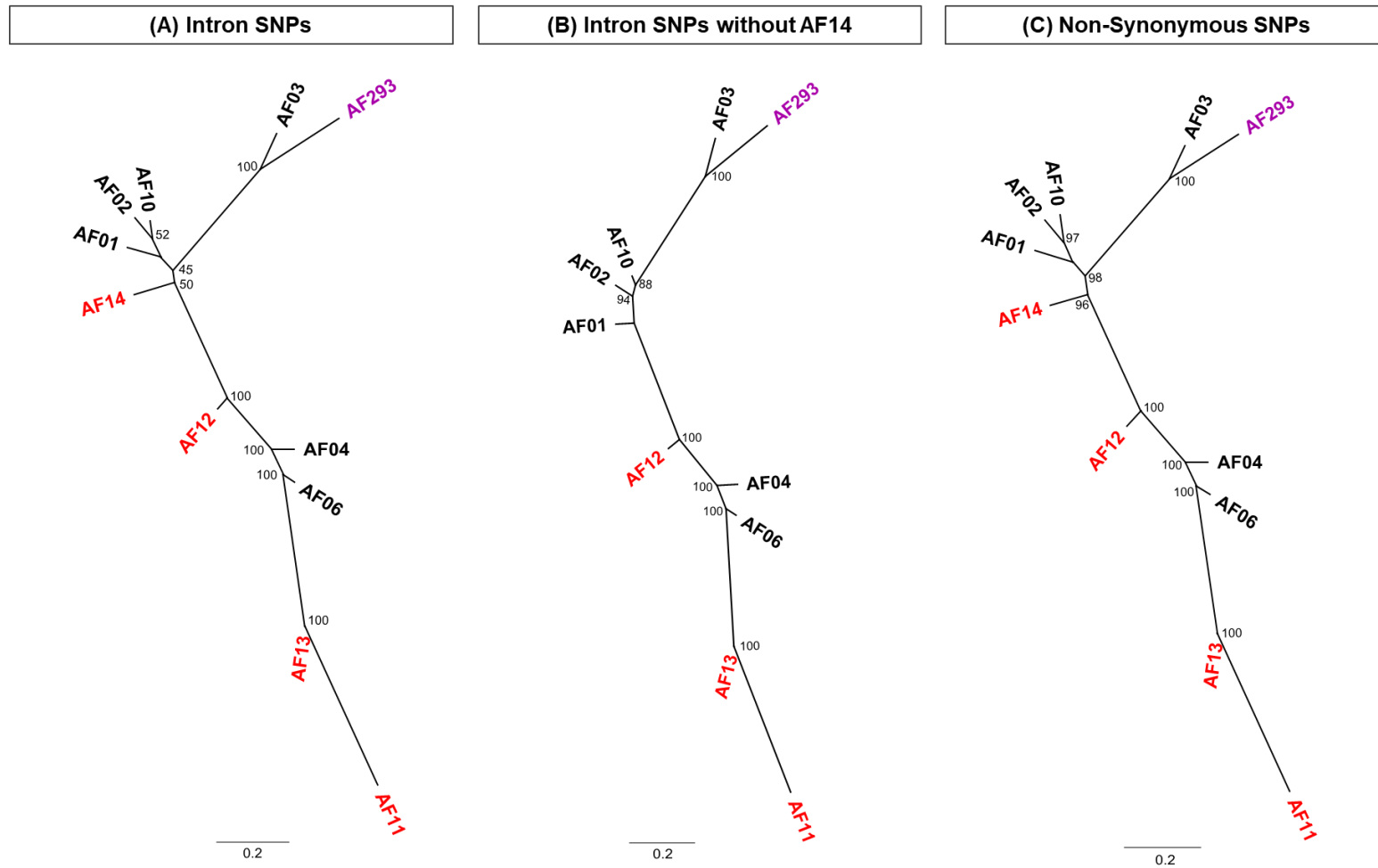


Figure 4.5. Phylogenetic analysis of 10 clinical *A. fumigatus* isolates. Trees were constructed from concatenated intronic SNPs (A), excluding AF14 (B), and from non-synonymous SNPs (C). Node values represent branch support from 10000 bootstraps. Intronic and non-synonymous variant-based phylogenies show similar topology.

Chapter 5 Virulence of clinical *A. fumigatus* isolates

5.1 Introduction

Virulence is the ability of a microbe to overcome host-defences, cause infection and damage a host. The clinical origin of an *A. fumigatus* isolate is not necessarily reflective of virulence. It is possible that the most clinically relevant difference between IA and colonising isolates is the host they happened to infect. Subtle differences in immune-profiles, treatment regimen and other host-factors may overwhelm the relevance of fungal factors. Quantifying virulence in an infection model allows for greater sample sizes and consistency in host-factors. If the contribution of *A. fumigatus* intraspecies variation to IA onset and severity is high, relative to host-factors, then virulence quantified in an infection model is likely to correlate with clinical origin. If *A. fumigatus* intra-specific variation informs clinical pathology to only a small degree, relative to host-factors, then patterns of virulence are expected to show independence from clinical origin. In the latter case, virulence data from an infection model can be used to detect the importance of phenotypic and genotypic variation in *A. fumigatus* to virulence, rather than relying on a heavily host-factor dependent clinical origin.

Tenebrio molitor is a species of darkling beetle. In recent years, *T. molitor* larvae (mealworms) have been used as an infection model in the study of bacteria such as *Staphylococcus aureus*¹⁰⁴ and fungi such as *C. albicans* and *C. neoformans*¹⁰⁵. Use of mealworms, like other insect models, presents ethical and economic advantages over mammalian models. Inoculation by injection also allows precise control of infective load, as opposed to ingestion-dependent inoculation in nematode models⁵². Further, *T. molitor* can be incubated at 37 °C, the internal body temperature of humans, unlike other potential model organisms such as *D. melanogaster* and *C.*

*elegans*¹⁰⁶. While lacking acquired immunity, antimicrobial defences of *T. molitor* include phagocytic haemocytes and antimicrobial peptides^{52,53}. Thus, the host presents challenges to the fungus similar to those faced by *A. fumigatus* in early-stage infection of the human lung.

This chapter describes (1) the optimisation and use of *T. molitor* larvae as an *A. fumigatus* infection model, (2) the relationship between the virulence in *T. molitor* larvae and clinical origin of *A. fumigatus* isolates and (3) correlation between inter-isolate variation in phenotype/genotype and virulence.

5.2 Infection model optimisation

The initial experimental design for virulence testing of clinical *A. fumigatus* isolates involved injection of spores into *T. molitor* larvae and week-long incubation at 37°C on a diet of LSA-supplemented oatmeal bran and a slice of carrot for moisture (1pc; ~500 mg). Uninfected mealworms showed high rates of survival in these experimental conditions at bran volumes of either 3 or 15 mL (Figure 5.1A). Absence of carrot in the diet significantly lowered mealworm survival, however increasing the amount available from 0.5 to 1.0 g did not affect survival (Figure 5.1B). Chilling mealworms on ice was evaluated as a means of reducing mealworm activity to increase feasibility of injection-site standardisation and minimisation of host-damage during injection. Incubating for 5 min on ice rendered mealworms docile without significantly affecting survival ($p = 0.32$; Figure 5.1C). Injection of mealworms with PBST negatively affects mealworm survival by up to 25%, but in a manner independent of injection site ($p = 0.96$, Figure 5.1D). Injecting mealworms at the base of sternite 5 (Figure 1D; inset) led to the lowest frequency of hemolymph leakage.

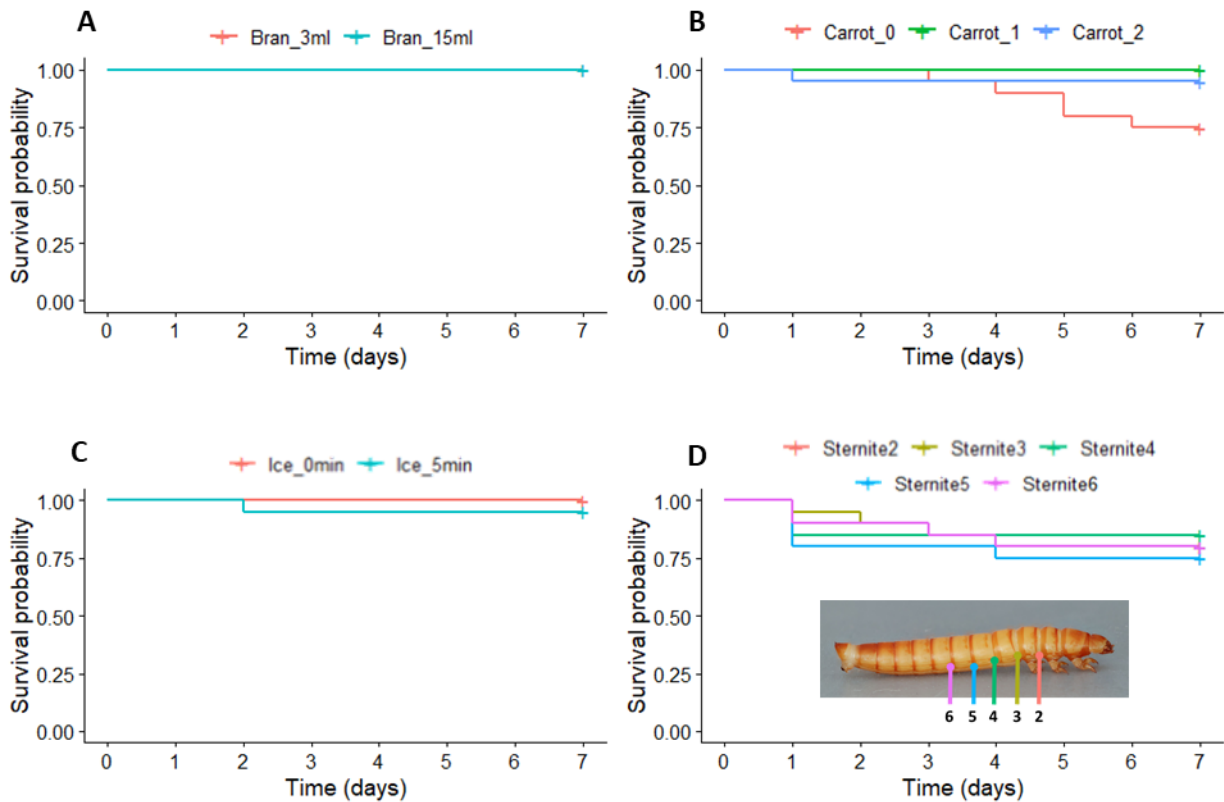


Figure 5.1. Optimisation of *T. molitor* larvae rearing and injection protocols. (A) Effect of bran volume on mealworm survival. Uninfected mealworms show high rates of survival when incubated at 37°C for 1 week. Increasing bran volume from 3 to 15 mL did not affect survival (logrank: $p = 1.00$; $n=20$). (B) Effect of carrot availability on mealworm survival. The amount of carrot supplied significantly affects survival (logrank: $p = 0.0228$; $n=20$). (C) Effect of ice-anaesthesia on mealworm survival. Chilling mealworms on ice for 5 min has no significant effect on the survival of uninfected mealworms (logrank: $p = 0.32$; $n=20$). (D) Effect of position-of-injection on mealworm survival. There is no significant difference in survival of mealworms injected with PBS-tween at the base of sternites 2-6 ($p = 0.963$; $n=20$).

Based on the above findings, in future experiments mealworms were chilled on ice for 5 min, injected at the base of sternite 5, and incubated with 3 mL LSA-supplemented bran and a single slice of carrot (~500 mg), replaced daily.

In the optimised experimental system, the effect of colonising isolate AF01 on mealworm survival was dose-dependent (Figure 5.2). All inoculum sizes tested (5×10^1 , 5×10^2 , 5×10^3 , 5×10^4 , 5×10^5 , or 5×10^6) significantly decreased survival rate relative to

vehicle controls. Inoculation with 5×10^4 spores consistently yielded mortality rates above 50%, which is required for the calculation of median-survival time, without killing at a rate so high that isolates of higher virulence would be difficult to resolve. All future experiments used an inoculum of 5×10^4 spores.

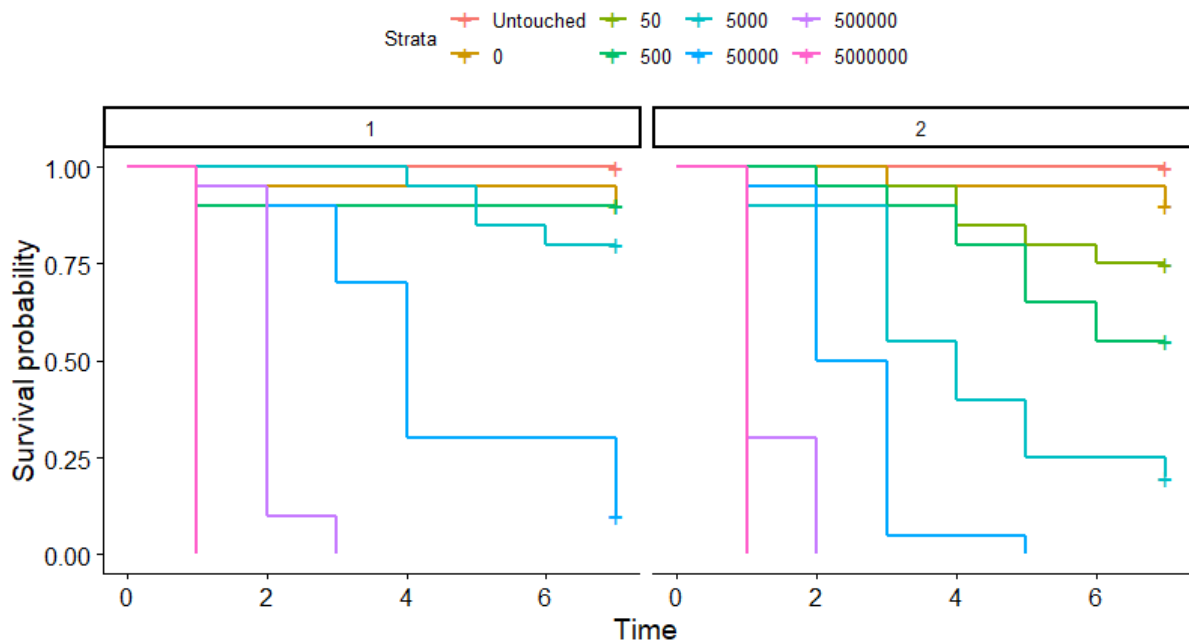


Figure 5.2. Kaplan-Meier survival of *T. molitor* larvae infected with different doses of *A. fumigatus* spores. Mealworms were inoculated with 0 (vehicle control), 5×10^1 , 5×10^2 , 5×10^3 , 5×10^4 , 5×10^5 , or 5×10^6 spores from colonising isolate AF01 and incubated at 37°C for 7 days ($n=20$). Data represents Kaplan-Meier survival probability. Panels represent results of two discrete experimental runs. Increasing inoculum size decreased rate of mealworm survival. Inoculation with 5×10^4 spores consistently resulted in mortality in over 50% of mealworms, which is required for the calculation of median-survival time. All subsequent experiments used an inoculum of 5×10^4 spores.

5.3 Inter-isolate variation in virulence of *A. fumigatus* isolates

5.3.1 Kaplan-Meier survival of *mealworms* infected with clinical *A. fumigatus* isolates

The virulence of 10 colonising and 5 IA-associated isolates was evaluated in a mealworm model. All isolates tested were virulent, decreasing Kaplan-Meier survival of infected mealworms relative to vehicle controls (0.05% v/v PBS-tween) (Figure 5.3). A high level of inter-replicate variability was observed. In all experiments, over 50% of isolate-infected mealworms were killed by day 7, allowing median survival time to be calculated. All future references to 'survival data' refer to this dataset.

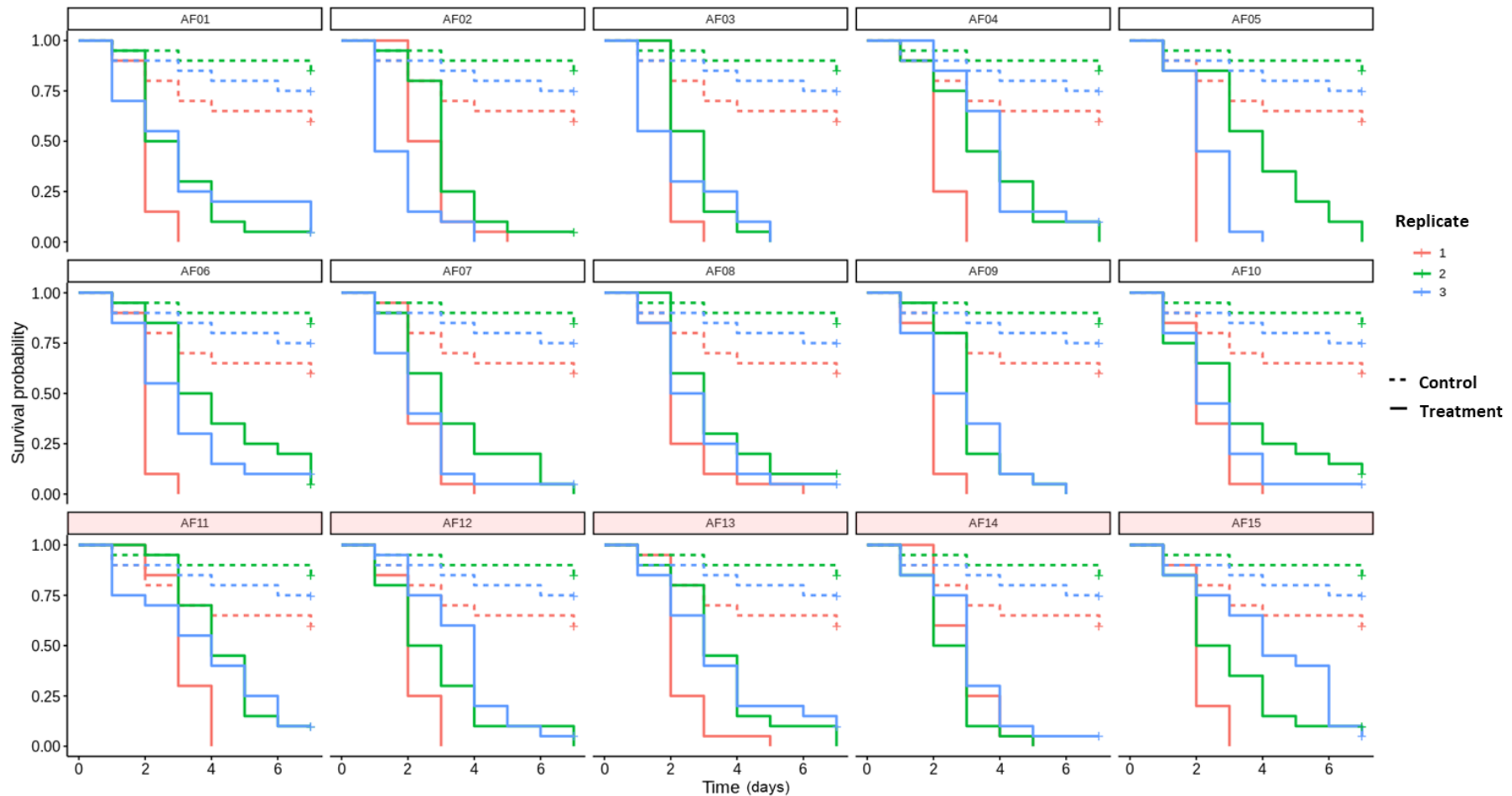


Figure 5.3. Kaplan-Meier survival of *T. molitor* larvae infected with clinical *A. fumigatus* isolates. For each isolate, 20 mealworms were injected with 5×10^4 spores (in 5 μ L of 0.05% PBS-tween) and incubated at 37°C for 7 days, with survival scored daily (solid lines). Control groups were injected with sterile PBS-tween (dashed lines). The experiment was repeated three times (orange, green & blue). Isolates AF01-AF10 were isolated from at-risk patients that never developed IA, while AF11-AF15 were from IA patients).

5.3.2 Median survival times

Median survival time of mealworms infected with 5×10^4 *A. fumigatus* spores ranged from 1-4 days. Infection with IA isolates tended to result in slightly longer median survival times. This trend, however, was not statistically significant at the 95% confidence threshold used ($p = 0.0684$, Figure 5.4).

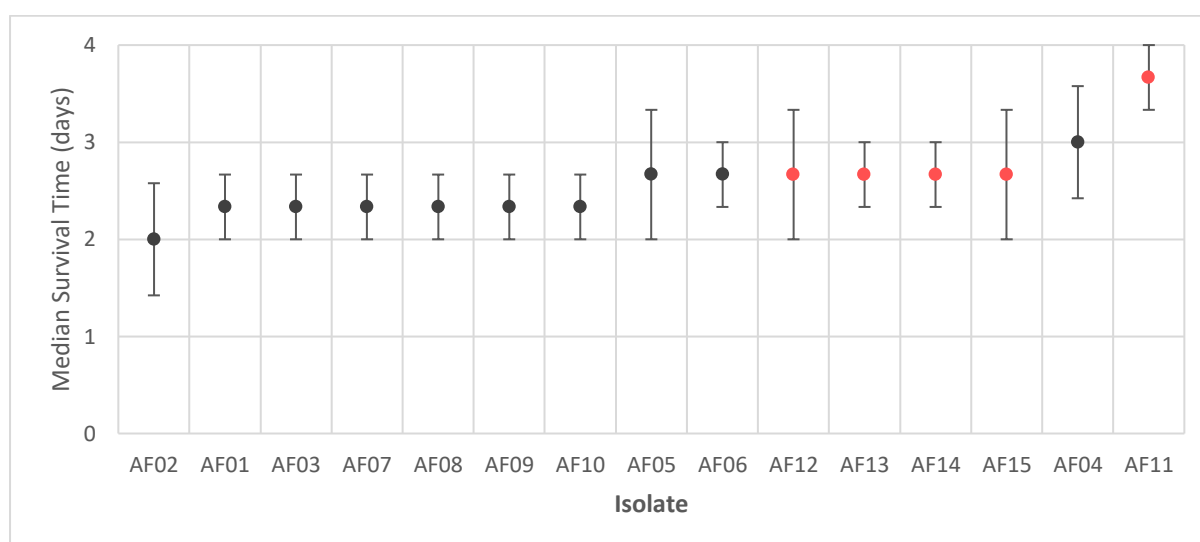


Figure 5.4. Median survival time of *T. molitor* larvae infected with clinical *A. fumigatus* isolates. Larvae were infected with 5×10^4 spores and incubated at 37°C for 7 days. Data represents mean and standard errors from three replicate experiments for each isolate (Colonising: grey, IA: red). Nested t-test indicated no significant effect of clinical origin on median survival time ($p = 0.0684$).

5.3.3 Cox-regression: inter-isolate variation and dependency on clinical origin

Cox regression models fit with frailty to account for inter-replicate variation indicate that the clinical origin of *A. fumigatus* isolates has a significant effect on their virulence within the invertebrate model ($p = 1.8 \times 10^{-5}$). Colonising isolates were 36.6% more likely to cause mortality in a host than IA isolates. All isolates tested were virulent, significantly increasing the risk of mortality in their host relative to PBS-tween controls (p values $< 10^{-11}$; Figure 5.5). IA isolate AF11 was the least virulent, with AF11-infected

hosts 2.8 times more likely to survive any given timepoint than those infected the most virulent isolate, AF03.

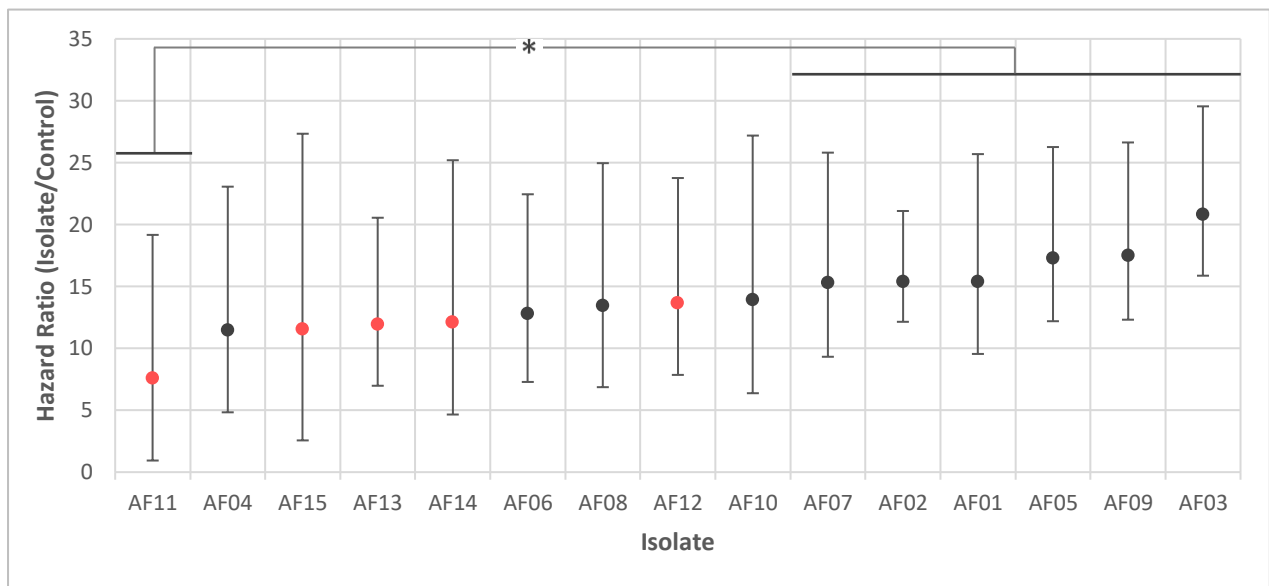


Figure 5.5. Virulence of clinical *A. fumigatus* isolates in a *T. molitor* infection model. A Cox-regression model was fit to survival data, with frailty to account for inter-replicate variability. Data represents hazard ratios of each isolate to PBS-tween control groups \pm 95% confidence intervals. Note hazard refers to the probability of the event (mealworm death) occurring at any given timepoint. These hazard ratios represent the exponent form of the β coefficient of the Cox model fit. All isolates significantly increased hazard relative to vehicle controls ($p < 10^{-11}$). Significant inter-isolate variation is shown (bonferroni-corrected $p < 0.05$: *).

5.4 Correlation between inter-isolate variation and virulence in an invertebrate model

5.4.1 Phenotypic variation

Correlation between phenotypic properties and virulence in the mealworm model of IA was evaluated. No significant correlation was observed between isolate-virulence (hazard-ratio of infected mealworms relative to PBST controls) and either radial growth rate or conidial UV resistance (Figure 5.6). Virulence of isolates as modelled in mealworms was also independent of their AMB sensitivity (Figure 5.6).

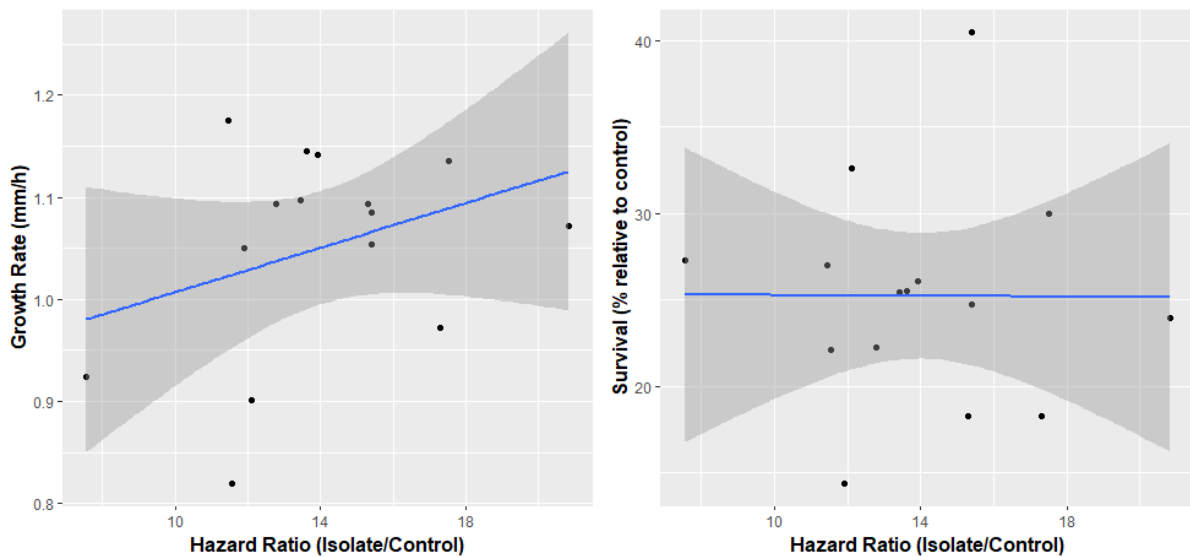


Figure 5.6. Correlation between virulence of clinical *A. fumigatus* isolates in an invertebrate model of IA and isolate radial growth rate on nutrient rich media (left) or conidial UV resistance (right). Hazard ratio of isolate-infected mealworms relative to PBST controls were used as the metric of virulence, where hazard is the probability of mortality occurring at any given timepoint. Radial growth rate was measured on PDA at 37°C. UV resistance was measured as percent survival of conidia following irradiation for 1 min at 1.6 W/m². Simple linear models fit to data are shown in blue (\pm 95% confidence intervals; grey). No significant Pearson correlation between virulence in the mealworm model and either radial growth rate or UV resistance was observed (growth rate: $p = 0.218$, UV resistance: 0.9848).

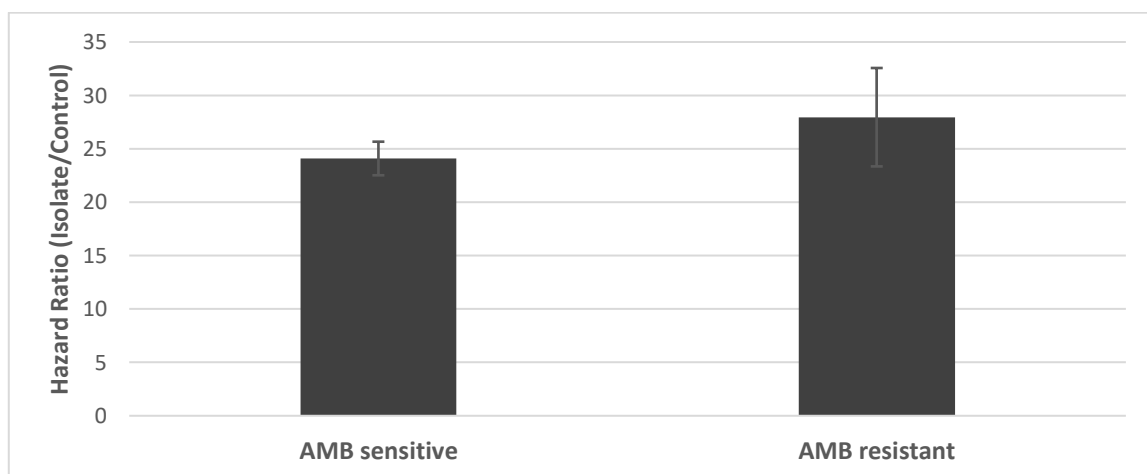


Figure 5.7. Virulence of *A. fumigatus* isolates by AMB resistance. Data represents hazard ratios of 10 colonising *A. fumigatus* isolates relative to PBST control groups. The mean hazard ratios of six AMB sensitive isolates and 4 AMB resistant isolates are shown. No significant association between AMB sensitivity and isolate virulence was found. Data was analysed by Welch t-test ($p = 0.475$).

5.4.2 Genotypic variation

Phylogenetic analysis revealed no strong clustering of *A. fumigatus* isolates based on clinical origin (Figure 4.5). For example, colonising isolates AF06 and AF04 were far more closely related to all the IA isolates than the remaining colonisers. These two colonising isolates represent the least virulent colonising isolates, of those sequenced (Figure 5.8).

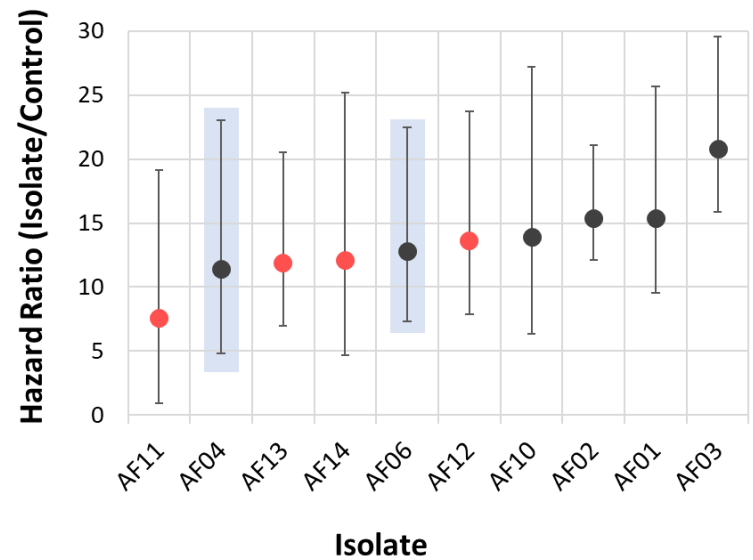
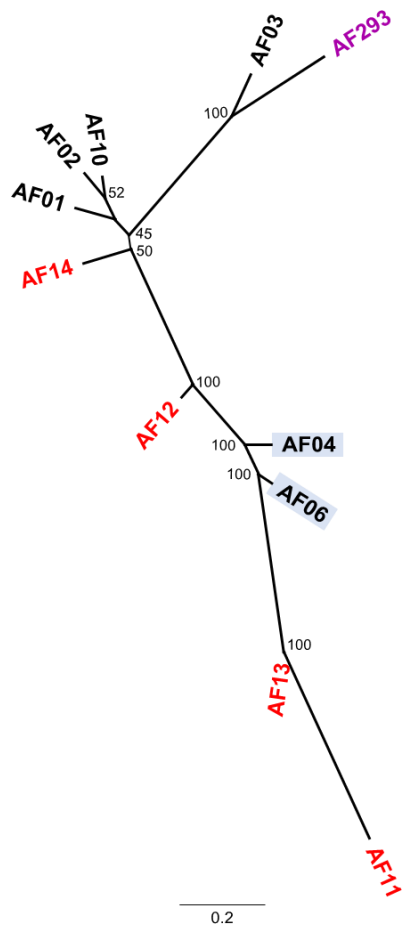


Figure 5.8. Consensus phylogenetic tree constructed based on intronic variation within the core genome of *A. fumigatus* in the context of virulence in *Tenebrio molitor* larvae. Hazard ratios of each isolate to vehicle controls, calculated by fitting a Cox-regression model with frailty to survival data described in Figure 5.3, were used as an estimate of virulence. The two colonisers most closely related to the IA isolates (AF06 and AF04) represent the least virulent colonising isolates, of those sequenced. Error bars represent 95% confidence intervals. Scale bar units are substitutions per site.

Chapter 6 Discussion

6.1 Clinical Relevance of phenotypic heterogeneity

High levels of phenotypic heterogeneity of *A. fumigatus* has been observed on an intraspecies level and with regards to many properties putatively associated with virulence. In this study, I have characterized 15 clinical *A. fumigatus* isolates with respect to phenotypic properties with theoretical links to pathogenesis. By examining how patterns of intra-specific phenotypic variation relate to the clinical origin and virulence of isolates in *T. molitor* larvae, the contribution of these fungal properties to clinical pathology can be estimated.

6.1.1 Growth rate

Growth rate is one of many *A. fumigatus* properties potentially important in pathogenesis. It is theorised that higher rates of *A. fumigatus* growth lead to increased fungal biomass, which makes it more difficult for the immune system to eradicate the infection. In 1995, Rinyu et al. identified a high level of heterogeneity in the growth rates of 61 *A. fumigatus* strains at 37°C, both on minimal and nutrient rich solid growth media¹⁹. Intra-specific variability of *A. fumigatus* growth rates at 37°C has also been noted in several smaller scale studies which assayed growth in liquid culture using nutrient rich media²¹ and RPMI¹⁰⁷, as well as those which assayed it in solid culture on minimal media⁸⁶. The 15 clinical *A. fumigatus* isolates examined in the present study also showed significant variability in their growth rate, as assayed on PDA, a solid, nutrient rich growth medium (Figure 3.2A,C).

The natural variability of *A. fumigatus* growth rate has previously been examined in the context of virulence. In 1996, a double knockout of two *A. fumigatus* chitin synthases yielded a strain with reduced growth rate as assayed at 37°C on solid,

nutrient reach media¹⁰⁸. The slow growing mutant also showed reduced virulence in a murine model of IA, relative to the wild-type parent. Several subsequent studies in which *A. fumigatus* deletion mutants were generated also found decreased radial growth rate was often accompanied with decreased virulence in murine models¹⁰⁹⁻¹¹¹. Further, an association between growth rate and virulence has been found in studies examining natural variation in growth rate. In 2005, Paisley et al. evaluated the growth rates of 9 *A. fumigatus* isolates in liquid culture using spectrophotometric methods and nutrient rich sabouraud dextrose broth²¹. The authors found rank-order growth rate was positively correlated with rank order LD90 in a neutropenic murine model. A similar study in 2013 spectrophotometrically assayed the growth curves of thirty *A. fumigatus* isolates in RPMI and examined the virulence of half of these isolates in a murine model of disseminated aspergillosis¹⁰⁷. Higher growth rates once again correlated with virulence in a murine model. Most recently, A 2014 study found the virulence of two *A. fumigatus* strains in neutropenic mice to positively correlate with growth rate assayed using aspergillus minimal media⁸⁶.

Growth rate does not always positively correlate with virulence. For example, in 2010, a knockout study of the gene *RacA* yielded an *A. fumigatus* strain with reduced growth rate on both nutrient rich and minimal media at 37°C, but comparable virulence to the wild type strain in both an insect model and two different murine models of IA⁵⁵. Similarly, when *A. fumigatus* growth was spectrophotometrically assayed in liquid yeast nitrogen base media, no correlation with virulence in *Toll*-deficient *Drosophila* was observed⁶². In saying this, the former study involved only two strains, an experimental mutant and a wild type and although the latter examined 20 *A. fumigatus* isolates, both growth rate and virulence were assayed at 29°C due to limitations of the *Drosophila* model, rather than at 37°C as used in all other studies described. Thus,

taken together, the literature suggests a clinical relevance of *A. fumigatus* growth rate predicated on the assumption that virulence metrics generated in murine models are, themselves, clinically relevant.

In our study, we examined 15 clinical *A. fumigatus* isolates and found those from cases of IA to be slower growing than colonising isolates (Figure 3.2B,C) on PDA at 37°C. The observed correlation between growth rate and clinical origin suggests intra-specific variation in growth rate is relevant to clinical pathology. However, this study indicates that selection for slower growth rates may occur in human hosts. This seems to be at odds with the large body of evidence that positively correlates growth rate and virulence in animal models. There are several possible explanations for these conflicting results. These include (1) inter-study variation in the experimental system within which growth rate is assayed and (2) a disconnect between virulence in animal models and clinical pathology.

The growth rate of *A. fumigatus* isolates is dependent on growth conditions. Thus, differences in incubation temperature, composition of growth medium, or whether liquid or solid cultures were examined may explain the variation. We assayed growth rate on at 37°C on PDA, a nutrient rich solid growth medium. The temperature represents internal human body temperature and is used in all studies which positively correlated growth rate and virulence. Conversely, both the medium used in these studies and whether growth rate is assayed in liquid or solid culture is highly variable. In saying this, growth rate positively correlated with virulence has been observed in a range of experimental conditions that encompasses those used in the present study. Studies have been conducted using nutrient rich solid growth media such as SDA¹⁰⁸, minimal solid growth media such as aspergillus minimal media⁸⁶, and in liquid culture using both nutrient rich sabouraud broth²¹ and cell-culture medium RPMI¹⁰⁷. Thus, it

is unlikely that the nutrient-composition of the media or mode of growth led to the apparent difference in results.

As differences in growth conditions are unlikely to explain our observation of a negative correlation between growth rate and clinical severity in the context of numerous studies that identify positive correlation between growth rate and virulence, other possibilities should be explored. Most studies evaluating the contribution of growth rate to pathogenesis rely on virulence measured in animal models. Thus, it is possible that a disconnect between virulence and clinical origin may be present. Indeed, my findings are at odds with previously published data only under the assumption that if we were to assay virulence in an animal model, that the slow-growing IA isolates would be more virulent than colonisers. In this study, IA isolates were not only slow growing relative to colonising isolates, but also were less virulent in *T. molitor* larvae. Thus, clinical severity of isolates does not necessarily equate to virulence in IA models. It is possible that faster growth rates may be associated with a property selected for in animal models but not in human hosts. Perhaps some fitness trade-off is occurring in human hosts in response to human specific immune factors or prophylactic therapy not appropriately modelled in the experimental systems used. The relationship between clinical data and infection models is discussed in more detail in chapter 6.3.

It is worth noting that despite both growth rate and virulence in *T. molitor* larvae being negatively correlated with clinical occurrence of IA, no statistically significant correlation between growth rate and virulence was observed (Figure 5.6A). This suggests that while subtle differences in growth rate between IA and non-IA isolates exist, an isolate's growth rate is unlikely to inform virulence in a strong, quantitative and consistent manner. This may be a result of the complex, multi-factorial nature of

growth rate which means it can be modulated by changes in many different facets of *A. fumigatus* biology^{55,108-110}.

6.1.2 UV Resistance

Resistance of *A. fumigatus* conidia to solar UV radiation and UV-induced reactive oxygen species (ROS) is important for the survival of airborne conidia⁸⁷. Conidial defences against UV-induced oxidative damage, such as cell wall melanin, have been implicated in pathogenesis by promoting pre-germination concealment of immunogenic PAMPs^{88,89}, evasion of internalization by phagocytes⁹⁰, and persistence within immune and alveolar epithelial cells⁹¹⁻⁹⁴. Several previous studies have also identified heterogeneity in pigmentation of clinical *A. fumigatus* isolates^{19,95}, likely representative of underlying variability in melanin biosynthesis pathways. In this study we also observed significant variation amongst 15 clinical *A. fumigatus* isolates, possibly reflective of variability in UV resistance conferring properties (Figure 3.3A,C).

Despite theoretical and experimental links between virulence and UV-resistance conferring properties, we found no association between conidial UV resistance of *A. fumigatus* isolates and their clinical origin (Figure 3.3B,C) or virulence in *T. molitor* larvae (Figure 5.6B). There are several reasons this might be. Firstly, much of the research tying conidial melanin to phenomena potentially important to IA onset is conducted by comparing *A. fumigatus* strains that represent extreme examples of melanin variation. For example, studies often compare pigment free mutants with a wild type parental strain^{93,112}. While these studies are useful when it comes to suggesting putative mechanisms by which melanisation affects virulence, the frequency with which such variation occurs in nature is dubious, and so the clinical relevance of natural melanin variability, such as that present in the clinical isolates tested, remains ambiguous. Secondly, UV-resistance is a multifactorial property.

Conidial defences against UV-induced damage includes not only cell wall pigments such as melanin, but also ROS detoxifying enzymes such as catalases, cellular metabolites such as polyols and DNA repair systems such as nucleotide excision repair and photoreactivation⁸⁷. Thus, conidial UV resistance represents a metric that integrates many facets of *A. fumigatus* biology, some of which may not affect virulence at all or affect it in opposing directions.

6.1.3 AMB Resistance

Amphotericin B is a polyene class broad spectrum antifungal used in the treatment of IA⁹⁶. Thus, AMB resistance is directly related to virulence through its effect on treatment efficacy. AMB targets ergosterol in the fungal plasma membrane and can induce oxidative stress⁹⁷. Proposed mechanisms of resistance include reduced membrane ergosterol levels and upregulation of anti-ROS enzymes.¹⁷ Ergosterol biosynthesis is linked to siderophore production via a shared precursor mevalonate.⁹⁸ Thus, AMB resistance may also contribute to virulence via effects on oxidative stress biology or iron sequestration. In this study, the resistance of acute exposure of 10 colonising *A. fumigatus* isolates to 0.5, 1.0 or 1.5 µg/mL of AMB did not yield any significant inter-isolate variation (Figure 3.4). Inter-isolate variability in AMB resistance could be resolved, however, when AMB sensitive isolates were taken as those whose post-exposure survival significantly decreased when treated with 1.5 µg/mL AMB relative to 0.5 µg/mL (Figure 3.5).

AMB resistance in *C. albicans* has been previously associated with fitness trade-offs and reduced virulence in murine models¹¹³. Similarly, a study comparing the virulence of 2 AMB resistant and 3 AMB sensitive strains of *A. terreus* in *G. mellonella* larvae also found AMB resistant isolates to be less virulent¹¹⁴. Here, we found the virulence of AMB resistant isolates in *T. molitor* larvae to be comparable to AMB sensitive

isolates (Figure 5.7). This does not necessarily suggest a strong species-dependent variation in the fitness-cost of AMB resistance. We assayed acute AMB resistance (3 h of exposure) in malt extract broth using dilution plating, whereas the aforementioned studies assayed chronic AMB exposure (24-48 h) in RPMI using visual signs of growth. Thus, all we can conclude is that in our *A. fumigatus* isolates, variability in acute AMB resistance appears unimportant to virulence. It is possible that fitness and virulence trade-offs do not occur below a certain threshold of AMB resistance which our isolates did not surpass. Future experiments using standard EUCAST antifungal resistance testing to examine the effect of chronic AMB exposure would allow a more direct comparison of resistance profiles across studies.

6.2 Clinical relevance of genomic heterogeneity

There have been several studies of intra-specific variation of *A. fumigatus* isolates^{33,115-117}, however associations between genomic variation and clinical severity have proved elusive, particularly when using broad-scale phylogenetic comparison. For example, a WGS study of 17 isolates from Japanese patients compared the genomic profiles of two clinical forms of aspergillosis: pulmonary aspergilloma (PA), a fungal tumour, and chronic necrotizing pulmonary aspergillosis (CNPA), a locally invasive form of aspergillosis¹¹⁵. The authors identified SNPs by aligning reads to the AF293 reference genome and identified 'consensus' sites where SNPs occurred in all 17 strains. Phylogenomic analysis was then based on concatenated sequences at all these consensus sites. The resulting phylogenetic tree did not cluster based on pathological condition or medication history profile (whether the isolate came from a patient treated with itraconazole, voriconazole, micafungin or no antifungals). Further, no SNPs unique to either PA or CNP were observed. A very recent WGS study examined genomes of 9 *A. fumigatus* isolates, including 7 clinical and 2 environmental

isolates. Screening for presence of 244 virulence associated genes showed all genes were present in every isolate, with the exception of one clinical and one environmental isolate which both lacked an putative ABC transporter¹¹⁷. This suggests presence/absence of virulence genes isn't a defining factor of *A. fumigatus* virulence. In the same study, SNP calling and comparative genomics analysis was completed for one colonising isolate, two IA isolates and one clinical experiment strain, no obvious relationship between clinical origin and genomic properties could be found, even when examining only SNPs within putative virulence genes, although this is unsurprising given the small sample size.

In agreement with previous work in this area, the SNP-based phylogenies produced in this study showed no strong clustering based on clinical origin (Figure 4.5), with IA isolates most closely related to colonisers in all but one case. It is possible that the WGS SNP-based approaches may be too broad in their assessment of *A. fumigatus* intra-specific variation. Indeed, the similar topology of our intronic and non-synonymous variant based trees suggest any differential selection pressures acting on the fungus are not generating large scale changes in the core *A. fumigatus* genome. Thus, while the larger scale genomic variation identified in our isolates does not seem to possess a strong clinical relevance, it remains possible that more focused analysis may be able to resolve isolates of differing clinical origins. For example, one could focus on variation between genes of known function, or in a subset of repetitive regions of the genome. In any case, the regions of the genome compared in attempts to resolve clinical origin will likely need to differ from those targeted in common genotyping methods, with both Microsatellite (STRAf) and rep-PCR based genotyping also failing to resolve colonising and IA isolates^{62,118}.

Integration of virulence data generated in *T. molitor* larvae revealed that genetic relatedness of the clinical *A. fumigatus* isolates may be related to their virulence in animal models. Of the 10 sequenced isolates, the two least virulent colonisers were more closely related to the 4 IA isolates than any of the other colonisers (Figure 5.8). These findings are similar to those of a study in *Drosophila* which found an association between *A. fumigatus* clades identified using rep-PCR genotyping and virulence in the *Toll*-deficient fruit flies⁶². The differences in genotyping methods between the two studies, combined with the small sample size in this study means these findings remain preliminary. Higher power assays of virulence and sequencing of more isolates is required to determine whether these results represent a real, biological trend. It must also be considered that *A. fumigatus* is now a species complex, comprised of several sub-species. Thus, it is very possible the clustering observed in our intron-based phylogenies is based on the sub-species present and that these sub-species may have differing virulence. Future work can include mining of ribosomal sequences from WGS data and subspecies identification.

The phylogenetic framework produced in this study can serve as a tool for guiding future analysis. For example, selection-independent phylogenies reveal IA isolate AF12 is most closely related to colonising isolate AF04 (Figure 5.8). To identify what genetic variants are important for IA isolates, comparison between these closely related pairs will produce the least noise, as they are less likely to differ in clinically irrelevant sites due to random mutation. This approach can also mitigate confounding that can occur due to differences between *A. fumigatus* subspecies. It must be noted, however, that *A. fumigatus* colonising immunocompromised hosts can progress to invasive infection. Thus, it is possible that these colonising *A. fumigatus* isolates which resemble IA isolates in both virulence within invertebrate models and patterns of

genetic variation, simply represent potential IA isolates that never caused invasive disease due to differences in host-factors. The probability of such a phenomenon is difficult to estimate, requiring either accurate animal models of IA or larger sample sizes.

One problem that we are faced with in comparative genomics research is that *A. fumigatus* SNPs are called against one of two reference genomes, AF293 or A1163. Both AF293 and A1163 are clinical *A. fumigatus* strains, with 99.5% genomic identity, but with 828 genes not common in both genomes¹¹⁹. A 2018 study examining 28 *A. fumigatus* genomes sequenced in-house and 73 public genomes found that genomes fell into 4 well defined clusters, with the two references falling into two different clusters.¹¹⁶ The number of SNPs identified in each genome is dependent on the reference used. In the future, it may be worth characterising genomic variation of our isolates with respect to both potential reference genomes.

The intra-specific genomic variation characterised in this study is only a small fraction of that which can be inferred from WGS data. For example, genomes can vary not only in the SNPs present but also in which genes are present and how many copies of each gene are present. Gene copy number is highly variable in *A. fumigatus* isolates, particularly in genes related to transposable element and secondary metabolism functions²⁰, and limited data on their clinical relevance is available. Future work can include evaluation of these different types of genomic variation. More focused analysis of variation within genes potentially relevant to pathogenesis may reduce noise from selection independent genetic relatedness. Analysis of the differences between closely related IA and colonising isolates may also facilitate identification of clinically important variants, although must be interpreted with care as clinical relevance relies on the assumption that these low virulence colonising isolates

do not simply represent isolates with IA potential that was never reached due to differences in the host. Investigating the SNPs associated with different clades and further assessment of the evolutionary timeline of the isolates by including a non-*A. fumigatus* outgroup in phylogenetic analyses may also improve our understanding of the clinical relevance of genomic variation within *A. fumigatus*.

6.3 Clinical relevance of fungal virulence

Investigation of *A. fumigatus* virulence in model organisms is invaluable in the study of IA. Hosts susceptible to invasive *A. fumigatus* infections are almost always immunodeficient, however their immunological profiles can vary wildly⁴. IA is also both uncommon and often misdiagnosed, making it difficult to obtain a large set of clinical isolates standardised with respect to potentially noise-creating host-factors such as primary condition, therapeutic history or geographical region. The use of animal models allows the virulence of *A. fumigatus* isolates from different hosts to be compared in an experimental system where host-factors are standardised. The clinical relevance of data coming out of these studies depends on how accurately virulence in animal models relates to clinical data.

Intra-specific heterogeneity of *A. fumigatus* virulence has been observed in both murine^{59,60} and invertebrate models^{61,62}. In this study, we also observed intra-specific variation in *A. fumigatus* virulence as modelled in *T. molitor* larvae (Figure 5.5). Thus, we have strong evidence that intra-specific variation in *A. fumigatus* virulence occurs, and this variation can be resolved in animal models. The clinical relevance of this variation is less clear.

Several studies comparing clinical and environmental *A. fumigatus* isolates have been conducted. In immunosuppressed mice, environmental isolates were found to be less

virulent than clinical isolates using mortality-based metrics⁵⁹. A similar trend was observed in mixed infection murine models, where mice were co-infected with a clinical and corresponding environmental isolate and relative virulence inferred from the ratio of recovery after the mice shows signs of pulmonary distress¹²⁰. This trend in virulence is not mammal specific. Clinical isolates are also more virulent than environmental isolates in *G. mellonella* larvae⁶¹. Taken together, these studies suggest a clinical relevance of virulence data produced in animal models. They suggest that either (1) some environmental *A. fumigatus* isolates possess phenotypic profiles more conducive to causing infection than others, (2) within a human-host, virulence-enhancing micro-evolution occurs, or (3) some combination of both. In any case, the increased virulence of clinical isolates in animal models suggests that the fungal factors selected for in human hosts are also selected for in the animal models, at least to some degree.

In the present study, rather than comparing environmental and clinical isolates we compared two clinical subgroups: colonising and IA isolates. In *T. molitor* larvae, IA isolates were less virulent than colonisers (Figure 5.5). This suggests that clinically important fungal properties selected for in human hosts are not being selected for in our invertebrate model. One possibility is that virulence factors important in overcoming a clinical barrier to infection have fitness-costs that become visible when the selective pressure is lifted due to differences between the clinical environment and the experimental system used to assess virulence. This phenomenon may also explain the decreased growth rate of IA isolates relative to colonisers (Figure 3.2). The ability of an isolate to survive prophylactic and response therapy is important clinically, however was not modelled in *T. molitor* larvae. In *Candida*, AMB resistance is associated with extreme fitness costs¹¹³. Azole resistance has also been associated

with fitness trade-offs in *C. albicans*¹²¹, although some of the common azole-resistance conferring mutations do not appear detrimental to *A. fumigatus* fitness in either immunosuppressed¹²² or immunocompetent mice¹²³.

In addition to virulence factors with fitness-costs, the reduced virulence of IA isolates in *T. molitor* larvae may be a result of differences in the composition or strength of the immunological response. Mealworms lack the adaptive immune defences present in mammals⁵². While a compromised innate immune arm is the primary risk factor for IA, it is very possible the dual action of adaptive and innate immune mechanisms in human hosts selects for phenotypic profiles that differ from those selected for in mealworms. Further, while innate immune-mechanisms in mealworms and humans overlap, there are important differences between the two. The immune system of *T. molitor* larvae is composed primarily of phagocytic cells and antimicrobial peptides⁵², two core components of innate immunity in humans. Thus, while phagocytic responses are conserved between mammals and invertebrates, mechanisms of dealing with hyphae too large to phagocytose differ. In humans, neutrophils can autolyse, releasing hyphae-impeding nucleic acids coated in antimicrobial peptides in a process termed NETosis¹²⁴. Neutrophils can also bind to hyphae and degranulate, releasing antimicrobial compounds¹²⁵. In invertebrates, haemocytes aggregate around large foreign particles to form a complex that often becomes melanised, segregating the pathogen which is presumably killed by oxidative damage or starvation⁵³. These differing forms of hyphal-killing may impose different selective profiles on the infective species.

It is also important to note that the mealworm model used was not immunosuppressed. In mammalian models, some fungal factors, such as gliotoxin production, affect virulence in immunocompetent hosts but not in those that are immunocompromised¹³.

Thus, it is very possible that the immune-profile of immunocompetent *T. molitor* larvae select for a different phenotypic profile to immunocompromised human-hosts.

The data presented in this thesis is not the first to suggest inverse fitness of clinical *A. fumigatus* isolates in human and invertebrate hosts. A 2018 study examined the virulence of clinical *A. fumigatus* strains serially isolated from an IA patient over two years using *G. mellonella* larvae³³. Despite WGS suggesting they were isogenic, of the four isolates tested, only two showed increased virulence relative to the precursor, one exhibited comparable virulence and one had both a growth-defect and attenuated virulence. The study suggests isolates under selection in a human host can attain growth-defects and attenuated virulence in invertebrate models. However, this evidence does not suggest a trend ubiquitous enough to, in isolation, explain what we observed in our study. Further, a study comparing the virulence of colonising and IA isolates in *Toll*-deficient *Drosophila* found no significant differences between the two clinical sub-groups in either *A. fumigatus* or *A. terreus*⁶². It is possible that differences in these findings is due to differences in the *Drosophila* and mealworm models. For example, *Drosophila* were incubated at 29°C whereas mealworms can be incubated at 37°C. Differences in in-host selection factors due to inter-continental variation in prophylactic or therapeutic approaches may also be important. It's possible some antifungal resistance adaption associated with fitness costs is being selected for in one country but not the other due to use of different antifungals. Alternatively, it remains possible that most of our IA causing isolates just happened to be sampled from a genomic cluster with low-virulence in invertebrate models and that with increased sample size we would also observe no significant trends. It is worth noting that in both studies genotyping identified a potential relationship between virulence

and genetic relatedness and found no strong evidence of clustering based on clinical origin⁶².

Future research should be conducted on a much larger scale to determine the breadth of any clinical subtype dependent variation. Expanding the mealworm model to include immunosuppression, prophylaxis, and breakthrough therapy may also help narrow down which facets of *A. fumigatus* biology are being selected for in human hosts.

Conclusions

This thesis examined the phenotypic and genotypic differences between colonising and IA-associated *A. fumigatus* isolates. Relative to colonisers, invasive aspergillosis isolates were slower growing *in vitro* and less virulent in *T. molitor* larvae. The identification of phenotypic variation consistent with clinical origin suggests that intraspecific variation contributes to the clinical occurrence of IA. The inverse correlation between clinical severity and virulence suggests invertebrate models and human hosts select for different, but related phenotypic profiles and thus care must be taken when investigating putative virulence factors in invertebrate models. Broad-scale SNP based phylogenetic comparison of isolates is unable to resolve colonising and IA isolates, and so more focused or non-SNP based assays of variation may be necessary to reveal any genomic markers of a strains ability to cause invasive disease. Future work may include: (1) continued mining of genomic variation from WGS data to guide hypothesis generation; (2) assaying of greater numbers of isolates with respect to phenotypic profiles putatively associated with clinical origin; and (3) investigation of the validity of the mealworm infection model by comparing virulence data to that generated by mammalian models, use of more clinical isolates, and expansion of the model to include immunosuppression, prophylaxis and antifungal therapy.

References

- 1 Segal, B. H. Aspergillosis. *N. Engl. J. Med.* **360**, 1870-1884, doi:10.1056/NEJMra0808853 (2009).
- 2 Berger, S., El Chazli, Y., Babu, A. F. & Coste, A. T. Azole Resistance in *Aspergillus fumigatus*: A Consequence of Antifungal Use in Agriculture? *Front. Microbiol.* **8**, 6, doi:10.3389/fmicb.2017.01024 (2017).
- 3 Lin, S. J., Schranz, J. & Teutsch, S. M. Aspergillosis case-fatality rate: systematic review of the literature. *Clin. Infect. Dis.* **32**, 358-366, doi:10.1086/318483 (2001).
- 4 Hohl, T. M. Immune responses to invasive aspergillosis: new understanding and therapeutic opportunities. *Curr. Opin. Infect. Dis.* **30**, 364-371, doi:10.1097/qco.0000000000000381 (2017).
- 5 Steinbach, W. J. *et al.* Clinical epidemiology of 960 patients with invasive aspergillosis from the PATH Alliance registry. *J. Infect.* **65**, 453-464, doi:10.1016/j.jinf.2012.08.003 (2012).
- 6 Kosmidis, C. & Denning, D. W. The clinical spectrum of pulmonary aspergillosis. *Thorax* **70**, 270-277, doi:10.1136/thoraxjnl-2014-206291 (2015).
- 7 Paulussen, C. *et al.* Ecology of aspergillosis: insights into the pathogenic potency of *Aspergillus fumigatus* and some other *Aspergillus* species. *Microb. Biotechnol.* **10**, 296-322, doi:10.1111/1751-7915.12367 (2017).
- 8 O’Gorman, C. M. Airborne *Aspergillus fumigatus* conidia: a risk factor for aspergillosis. *Fungal Biol Rev* **25**, 151-157, doi:10.1016/j.fbr.2011.07.002 (2011).
- 9 Caggiano, G. *et al.* Mold contamination in a controlled hospital environment: a 3-year surveillance in southern Italy. *BMC Infect. Dis.* **14**, 595, doi:10.1186/s12879-014-0595-z (2014).
- 10 Cavallo, M., Andreoni, S., Martinotti, M. G., Rinaldi, M. & Fracchia, L. Monitoring environmental *Aspergillus* spp. contamination and meteorological factors in a haematological unit. *Mycopathologia* **176**, 387-394, doi:10.1007/s11046-013-9712-6 (2013).
- 11 Tong, X. *et al.* High diversity of airborne fungi in the hospital environment as revealed by meta-sequencing-based microbiome analysis. *Sci. Rep.* **7**, 39606, doi:10.1038/srep39606 (2017).
- 12 Kwon-Chung, K. J. & Sugui, J. A. *Aspergillus fumigatus*—What Makes the Species a Ubiquitous Human Fungal Pathogen? *PLoS Pathog.* **9**, e1003743, doi:10.1371/journal.ppat.1003743 (2013).
- 13 Dagenais, T. R. T. & Keller, N. P. Pathogenesis of *Aspergillus fumigatus* in Invasive Aspergillosis. *Clin. Microbiol. Rev.* **22**, 447-465, doi:10.1128/cmr.00055-08 (2009).
- 14 Brown, G. D. *et al.* Hidden killers: human fungal infections. *Science translational medicine* **4**, 165rv113, doi:10.1126/scitranslmed.3004404 (2012).
- 15 Patterson, T. F. *et al.* Practice Guidelines for the Diagnosis and Management of Aspergillosis: 2016 Update by the Infectious Diseases Society of America. *Clin. Infect. Dis.* **63**, e1-e60, doi:10.1093/cid/ciw326 (2016).
- 16 Ullmann, A. J. *et al.* Diagnosis and management of *Aspergillus* diseases: executive summary of the 2017 ESCMID-ECMM-ERS guideline. *Clinical microbiology and infection : the official publication of the European Society of*

- Clinical Microbiology and Infectious Diseases* **24 Suppl 1**, e1-e38, doi:10.1016/j.cmi.2018.01.002 (2018).
- 17 Gonçalves, S. S., Souza, A. C. R., Chowdhary, A., Meis, J. F. & Colombo, A. L. Epidemiology and molecular mechanisms of antifungal resistance in *Candida* and *Aspergillus*. *Mycoses* **59**, 198-219, doi:10.1111/myc.12469 (2016).
- 18 Ren, J. *et al.* Fungicides induced triazole-resistance in *Aspergillus fumigatus* associated with mutations of TR46/Y121F/T289A and its appearance in agricultural fields. *J Hazard Mater* **326**, 54-60, doi:10.1016/j.jhazmat.2016.12.013 (2017).
- 19 Rinyu, E., Varga, J. & Ferenczy, L. Phenotypic and genotypic analysis of variability in *Aspergillus fumigatus*. *Journal of clinical microbiology* **33**, 2567-2575 (1995).
- 20 Zhao, S. & Gibbons, J. G. A population genomic characterization of copy number variation in the opportunistic fungal pathogen *Aspergillus fumigatus*. *PLoS One* **13**, e0201611, doi:10.1371/journal.pone.0201611 (2018).
- 21 Paisley, D., Robson, G. D. & Denning, D. W. Correlation between in vitro growth rate and in vivo virulence in *Aspergillus fumigatus*. *Medical mycology* **43**, 397-401, doi:10.1080/13693780400005866 (2005).
- 22 Fuller, K. K., Cramer, R. A., Zegans, M. E., Dunlap, J. C. & Loros, J. J. *Aspergillus fumigatus* Photobiology Illuminates the Marked Heterogeneity between Isolates. *mBio* **7**, doi:10.1128/mBio.01517-16 (2016).
- 23 Lima, F. R. *et al.* Surveillance for azoles resistance in *Aspergillus* spp. highlights a high number of amphotericin B resistant isolates. *Mycoses*, doi:10.1111/myc.12759 (2018).
- 24 Wood, V. *et al.* The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**, 871-880, doi:10.1038/nature724 (2002).
- 25 Galagan, J. E. *et al.* The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**, 859-868, doi:10.1038/nature01554 (2003).
- 26 El-Kamand, S., Papanicolaou, A. & Morton, C. O. The Use of Whole Genome and Next-Generation Sequencing in the Diagnosis of Invasive Fungal Disease. *Current Fungal Infection Reports* **13**, 284-291, doi:10.1007/s12281-019-00363-5 (2019).
- 27 Prakash, P. Y. *et al.* Online Databases for Taxonomy and Identification of Pathogenic Fungi and Proposal for a Cloud-Based Dynamic Data Network Platform. *J Clin Microbiol* **55**, 1011-1024, doi:10.1128/JCM.02084-16 (2017).
- 28 Ene, I. V. *et al.* Global analysis of mutations driving microevolution of a heterozygous diploid fungal pathogen. *Proc Natl Acad Sci U S A* **115**, E8688-e8697, doi:10.1073/pnas.1806002115 (2018).
- 29 Sitterle, E. *et al.* Within-Host Genomic Diversity of *Candida albicans* in Healthy Carriers. *Sci Rep* **9**, 2563, doi:10.1038/s41598-019-38768-4 (2019).
- 30 Rhodes, J. *et al.* A Population Genomics Approach to Assessing the Genetic Basis of Within-Host Microevolution Underlying Recurrent Cryptococcal Meningitis Infection. *G3 (Bethesda, Md.)* **7**, 1165-1176, doi:10.1534/g3.116.037499 (2017).
- 31 Ormerod, K. L. *et al.* Comparative Genomics of Serial Isolates of *Cryptococcus neoformans* Reveals Gene Associated With Carbon Utilization and Virulence. *G3 (Bethesda, Md.)* **3**, 675-686, doi:10.1534/g3.113.005660 (2013).
- 32 Chow, E. W., Morrow, C. A., Djordjevic, J. T., Wood, I. A. & Fraser, J. A. Microevolution of *Cryptococcus neoformans* driven by massive tandem gene

- amplification. *Mol Biol Evol* **29**, 1987-2000, doi:10.1093/molbev/mss066 (2012).
- 33 Ballard, E. *et al.* In-host microevolution of *Aspergillus fumigatus*: A phenotypic and genotypic analysis. *Fungal genetics and biology : FG & B* **113**, 1-13, doi:10.1016/j.fgb.2018.02.003 (2018).
- 34 Camps, S. M. *et al.* Discovery of a HapE mutation that causes azole resistance in *Aspergillus fumigatus* through whole genome sequencing and sexual crossing. *PLoS One* **7**, e50034, doi:10.1371/journal.pone.0050034 (2012).
- 35 Hare, R. K. *et al.* In Vivo Selection of a Unique Tandem Repeat Mediated Azole Resistance Mechanism (TR120) in *Aspergillus fumigatus* cyp51A, Denmark. *Emerg Infect Dis* **25**, 577-580, doi:10.3201/eid2503.180297 (2019).
- 36 Castanheira, M., Deshpande, L. M., Davis, A. P., Rhomberg, P. R. & Pfaller, M. A. Monitoring Antifungal Resistance in a Global Collection of Invasive Yeasts and Molds: Application of CLSI Epidemiological Cutoff Values and Whole-Genome Sequencing Analysis for Detection of Azole Resistance in *Candida albicans*. *Antimicrobial agents and chemotherapy* **61**, doi:10.1128/aac.00906-17 (2017).
- 37 Mount, H. O. *et al.* Global analysis of genetic circuitry and adaptive mechanisms enabling resistance to the azole antifungal drugs. *PLoS Genet* **14**, e1007319, doi:10.1371/journal.pgen.1007319 (2018).
- 38 Ma, Q., Ola, M., Iracane, E. & Butler, G. Susceptibility to Medium-Chain Fatty Acids Is Associated with Trisomy of Chromosome 7 in *Candida albicans*. *mSphere* **4**, doi:10.1128/mSphere.00402-19 (2019).
- 39 Hill, J. A., Ammar, R., Torti, D., Nislow, C. & Cowen, L. E. Genetic and genomic architecture of the evolution of resistance to antifungal drug combinations. *PLoS Genet* **9**, e1003390, doi:10.1371/journal.pgen.1003390 (2013).
- 40 Losada, L. *et al.* Genetic Analysis Using an Isogenic Mating Pair of *Aspergillus fumigatus* Identifies Azole Resistance Genes and Lack of MAT Locus's Role in Virulence. *PLoS Pathog* **11**, e1004834, doi:10.1371/journal.ppat.1004834 (2015).
- 41 Zhang, J. *et al.* Evolution of cross-resistance to medical triazoles in *Aspergillus fumigatus* through selection pressure of environmental fungicides. *Proc Biol Sci* **284**, doi:10.1098/rspb.2017.0635 (2017).
- 42 Gillece, J. D. *et al.* Whole genome sequence analysis of *Cryptococcus gattii* from the Pacific Northwest reveals unexpected diversity. *PLoS One* **6**, e28550, doi:10.1371/journal.pone.0028550 (2011).
- 43 Engelthaler, D. M. *et al.* *Cryptococcus gattii* in North American Pacific Northwest: whole-population genome analysis provides insights into species evolution and dispersal. *mBio* **5**, e01464-01414, doi:10.1128/mBio.01464-14 (2014).
- 44 Bradford, L. L., Chibucos, M. C., Ma, B., Bruno, V. & Ravel, J. Vaginal *Candida* spp. genomes from women with vulvovaginal candidiasis. *Pathog Dis* **75**, doi:10.1093/femspd/ftx061 (2017).
- 45 Cavalieri, D. *et al.* Genomic and Phenotypic Variation in Morphogenetic Networks of Two *Candida albicans* Isolates Subtends Their Different Pathogenic Potential. *Frontiers in immunology* **8**, 1997-1997, doi:10.3389/fimmu.2017.01997 (2018).
- 46 Gerstein, A. C. *et al.* Identification of Pathogen Genomic Differences That Impact Human Immune Response and Disease during <span class="named-content genus-species" id="named-content-

- 1"Cryptococcus neoformans"; Infection. *mBio* **10**, e01440-01419, doi:10.1128/mBio.01440-19 (2019).
- 47 Desoubeaux, G. & Cray, C. Rodent Models of Invasive Aspergillosis due to *Aspergillus fumigatus*: Still a Long Path toward Standardization. *Front Microbiol* **8**, 841, doi:10.3389/fmicb.2017.00841 (2017).
- 48 Reitman, M. L. Of mice and men – environmental temperature, body temperature, and treatment of obesity. *FEBS Letters* **592**, 2098-2107, doi:10.1002/1873-3468.13070 (2018).
- 49 Mestas, J. & Hughes, C. C. W. Of Mice and Not Men: Differences between Mouse and Human Immunology. *The Journal of Immunology* **172**, 2731, doi:10.4049/jimmunol.172.5.2731 (2004).
- 50 Askew, D. S. *Aspergillus fumigatus*: virulence genes in a street-smart mold. *Current opinion in microbiology* **11**, 331-337, doi:10.1016/j.mib.2008.05.009 (2008).
- 51 Arvanitis, M., Glavis-Bloom, J. & Mylonakis, E. Invertebrate models of fungal infection. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1832**, 1378-1383, doi:10.1016/j.bbadis.2013.03.008 (2013).
- 52 Canteri de Souza, P., Custodio Caloni, C., Wilson, D. & Sergio Almeida, R. An Invertebrate Host to Study Fungal Infections, Mycotoxins and Antifungal Drugs: *Tenebrio molitor*. *Journal of fungi (Basel, Switzerland)* **4**, doi:10.3390/jof4040125 (2018).
- 53 Hillyer, J. F. Insect immunology and hematopoiesis. *Developmental and comparative immunology* **58**, 102-118, doi:10.1016/j.dci.2015.12.006 (2016).
- 54 Amich, J., Schafferer, L., Haas, H. & Krappmann, S. Regulation of sulphur assimilation is essential for virulence and affects iron homeostasis of the human-pathogenic mould *Aspergillus fumigatus*. *PLoS pathogens* **9**, e1003573-e1003573, doi:10.1371/journal.ppat.1003573 (2013).
- 55 Li, H. *et al.* The small GTPase RacA mediates intracellular reactive oxygen species production, polarized growth, and virulence in the human fungal pathogen *Aspergillus fumigatus*. *Eukaryotic cell* **10**, 174-186, doi:10.1128/ec.00288-10 (2011).
- 56 O'Hanlon, K. A. *et al.* Targeted disruption of nonribosomal peptide synthetase *pes3* augments the virulence of *Aspergillus fumigatus*. *Infection and immunity* **79**, 3978-3992, doi:10.1128/IAI.00192-11 (2011).
- 57 Slater, J. L., Gregson, L., Denning, D. W. & Warn, P. A. Pathogenicity of *Aspergillus fumigatus* mutants assessed in *Galleria mellonella* matches that in mice. *Medical mycology* **49**, S107-S113, doi:10.3109/13693786.2010.523852 (2011).
- 58 Jackson, J. C., Higgins, L. A. & Lin, X. Conidiation color mutants of *Aspergillus fumigatus* are highly pathogenic to the heterologous insect host *Galleria mellonella*. *PLoS One* **4**, e4224-e4224, doi:10.1371/journal.pone.0004224 (2009).
- 59 Mondon, P., De Champs, C., Donadille, A., Ambroise-Thomas, P. & Grillot, R. Variation in virulence of *Aspergillus fumigatus* strains in a murine model of invasive pulmonary aspergillosis. *J. Med. Microbiol.* **45**, 186-191, doi:10.1099/00222615-45-3-186 (1996).
- 60 Kowalski, C. H. *et al.* Heterogeneity among Isolates Reveals that Fitness in Low Oxygen Correlates with *Aspergillus fumigatus* Virulence. *mBio* **7**, doi:10.1128/mBio.01515-16 (2016).

- 61 Alshareef, F. & Robson, G. D. Genetic and virulence variation in an environmental population of the opportunistic pathogen *Aspergillus fumigatus*. *Microbiology (Reading, England)* **160**, 742-751, doi:10.1099/mic.0.072520-0 (2014).
- 62 Ben-Ami, R., Lamarinis, G. A., Lewis, R. E. & Kontoyiannis, D. P. Interstrain variability in the virulence of *Aspergillus fumigatus* and *Aspergillus terreus* in a Toll-deficient *Drosophila* fly model of invasive aspergillosis. *Medical mycology* **48**, 310-317, doi:10.3109/13693780903148346 (2010).
- 63 Galbraith, S., Daniel, J. A. & Vissel, B. A study of clustered data and approaches to its analysis. *J Neurosci* **30**, 10601-10608, doi:10.1523/JNEUROSCI.0362-10.2010 (2010).
- 64 Andrews, S. *et al.* *FastQC: A quality control tool for high throughput sequence data.*, <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>> (2014).
- 65 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
- 66 Song, L., Florea, L. & Langmead, B. Lighter: fast and memory-efficient sequencing error correction without counting. *Genome Biology* **15**, 509, doi:10.1186/s13059-014-0509-9 (2014).
- 67 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).
- 68 Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212, doi:10.1093/bioinformatics/btv351 (2015).
- 69 Kolmogorov, M. *et al.* Chromosome assembly of large and complex genomes using multiple references. *Genome research* **28**, 1720-1732, doi:10.1101/gr.236273.118 (2018).
- 70 Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome biology* **5**, R12-R12, doi:10.1186/gb-2004-5-2-r12 (2004).
- 71 Blanchette, M. *et al.* Aligning multiple genomic sequences with the threaded blockset aligner. *Genome research* **14**, 708-715, doi:10.1101/gr.1933104 (2004).
- 72 *Geneious 10.2.6*, <<https://www.geneious.com/>> (2018).
- 73 Smit, A., Hubley, R. & Green, P. *RepeatMasker Open-4.0.6*, <<http://www.repeatmasker.org>> (2013-2015).
- 74 Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637-644, doi:10.1093/bioinformatics/btn013 (2008).
- 75 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421, doi:10.1186/1471-2105-10-421 (2009).
- 76 Lomsadze, A., Burns, P. D. & Borodovsky, M. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res* **42**, e119, doi:10.1093/nar/gku557 (2014).
- 77 Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859-1875, doi:10.1093/bioinformatics/bti310 (2005).
- 78 Papanicolaou, A. & Brian, H. Just Annotate My genome. (2016).

- 79 Treangen, T. J., Ondov, B. D., Koren, S. & Phillippy, A. M. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* **15**, 524, doi:10.1186/s13059-014-0524-x (2014).
- 80 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158, doi:10.1093/bioinformatics/btr330 (2011).
- 81 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92, doi:10.4161/fly.19695 (2012).
- 82 Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**, 518-522, doi:10.1093/molbev/msx281 (2017).
- 83 Smid, H. *Tenebrio molitor* Linnaeus, 1758, <https://e-insects.wageningenacademic.com/tenebrio_molitor> (2017).
- 84 George, B., Seals, S. & Aban, I. Survival analysis and regression models. *J Nucl Cardiol* **21**, 686-694, doi:10.1007/s12350-014-9908-2 (2014).
- 85 Austin, P. C. A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications. *Int Stat Rev* **85**, 185-203, doi:10.1111/insr.12214 (2017).
- 86 Amarsaikhan, N. *et al.* Isolate-dependent growth, virulence, and cell wall composition in the human pathogen *Aspergillus fumigatus*. *PLoS One* **9**, e100430, doi:10.1371/journal.pone.0100430 (2014).
- 87 Braga, G. U. L., Rangel, D. E. N., Fernandes, É. K. K., Flint, S. D. & Roberts, D. W. Molecular and physiological effects of environmental UV radiation on fungal conidia. *Curr. Genet.* **61**, 405-425, doi:10.1007/s00294-015-0483-0 (2015).
- 88 Bayry, J. *et al.* Surface structure characterization of *Aspergillus fumigatus* conidia mutated in the melanin synthesis pathway and their human cellular immune response. *Infect. Immun.* **82**, 3141-3153, doi:10.1128/iai.01726-14 (2014).
- 89 Chai, L. Y. *et al.* *Aspergillus fumigatus* conidial melanin modulates host cytokine response. *Immunobiology* **215**, 915-920, doi:10.1016/j.imbio.2009.10.002 (2010).
- 90 Mech, F., Thywissen, A., Guthke, R., Brakhage, A. A. & Figge, M. T. Automated Image Analysis of the Host-Pathogen Interaction between Phagocytes and *Aspergillus fumigatus*. *PLoS One* **6**, 10, doi:10.1371/journal.pone.0019591 (2011).
- 91 Amin, S., Thywissen, A., Heinekamp, T., Saluz, H. P. & Brakhage, A. A. Melanin dependent survival of *Aspergillus fumigatus* conidia in lung epithelial cells. *Int. J. Med. Microbiol.* **304**, 626-636, doi:10.1016/j.ijmm.2014.04.009 (2014).
- 92 Akoumianaki, T. *et al.* *Aspergillus* Cell Wall Melanin Blocks LC3-Associated Phagocytosis to Promote Pathogenicity. *Cell Host Microbe.* **19**, 79-90, doi:10.1016/j.chom.2015.12.002 (2016).
- 93 Jahn, B. *et al.* Isolation and characterization of a pigmentless-conidium mutant of *Aspergillus fumigatus* with altered conidial surface and reduced virulence. *Infect. Immun.* **65**, 5110-5117 (1997).
- 94 Thywissen, A. *et al.* Conidial Dihydroxynaphthalene Melanin of the Human Pathogenic Fungus *Aspergillus fumigatus* Interferes with the Host Endocytosis Pathway. *Front Microbiol* **2**, 96, doi:10.3389/fmicb.2011.00096 (2011).

- 95 Hagiwara, D. *et al.* Temperature during conidiation affects stress tolerance, pigmentation, and tryptacin accumulation in the conidia of the airborne pathogen *Aspergillus fumigatus*. *PLoS One* **12**, e0177050, doi:10.1371/journal.pone.0177050 (2017).
- 96 Chamilos, G. & Kontoyiannis, D. P. Update on antifungal drug resistance mechanisms of *Aspergillus fumigatus*. *Drug Resist Updat* **8**, 344-358, doi:10.1016/j.drug.2006.01.001 (2005).
- 97 Anderson, T. M. *et al.* Amphotericin forms an extramembranous and fungicidal sterol sponge. *Nature chemical biology* **10**, 400-406, doi:10.1038/nchembio.1496 (2014).
- 98 Dhingra, S. & Cramer, R. A. Regulation of Sterol Biosynthesis in the Human Fungal Pathogen *Aspergillus fumigatus*: Opportunities for Therapeutic Development. *Front. Microbiol.* **8**, 92, doi:10.3389/fmicb.2017.00092 (2017).
- 99 Nierman, W. C. *et al.* Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* **438**, 1151-1156, doi:10.1038/nature04332 (2005).
- 100 Galagan, J. E. *et al.* Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **438**, 1105-1115, doi:10.1038/nature04341 (2005).
- 101 Camps, S. M. T. *et al.* Discovery of a HapE mutation that causes azole resistance in *Aspergillus fumigatus* through whole genome sequencing and sexual crossing. *PLoS One* **7**, e50034-e50034, doi:10.1371/journal.pone.0050034 (2012).
- 102 Carmel, L. & Chorev, M. The Function of Introns. *Frontiers in Genetics* **3**, doi:10.3389/fgene.2012.00055 (2012).
- 103 Creer, S. Choosing and using introns in molecular phylogenetics. *Evol Bioinform Online* **3**, 99-108 (2007).
- 104 Dobson, A. J., Purves, J. & Rolff, J. Increased survival of experimentally evolved antimicrobial peptide-resistant *Staphylococcus aureus* in an animal host. *Evol Appl* **7**, 905-912, doi:10.1111/eva.12184 (2014).
- 105 de Souza, P. C. *et al.* *Tenebrio molitor* (Coleoptera: Tenebrionidae) as an alternative host to study fungal infections. *Journal of microbiological methods* **118**, 182-186, doi:10.1016/j.mimet.2015.10.004 (2015).
- 106 Desalermos, A., Fuchs, B. B. & Mylonakis, E. Selecting an Invertebrate Model Host for the Study of Fungal Pathogenesis. *PLOS Pathogens* **8**, e1002451, doi:10.1371/journal.ppat.1002451 (2012).
- 107 Mavridou, E. *et al.* Composite survival index to compare virulence changes in azole-resistant *Aspergillus fumigatus* clinical isolates. *PLoS One* **8**, e72280, doi:10.1371/journal.pone.0072280 (2013).
- 108 Mellado, E., Aufauvre-Brown, A., Gow, N. A. & Holden, D. W. The *Aspergillus fumigatus* *chsC* and *chsG* genes encode class III chitin synthases with different functions. *Mol Microbiol* **20**, 667-679, doi:10.1046/j.1365-2958.1996.5571084.x (1996).
- 109 Zhao, W. *et al.* Deletion of the Regulatory Subunit of Protein Kinase A in *Aspergillus fumigatus* Alters Morphology, Sensitivity to Oxidative Damage, and Virulence. *Infection and immunity* **74**, 4865, doi:10.1128/IAI.00565-06 (2006).
- 110 Fortwendel, J. R. *et al.* A Fungus-Specific Ras Homolog Contributes to the Hyphal Growth and Virulence of *Aspergillus fumigatus*. *Eukaryotic cell* **4**, 1982, doi:10.1128/EC.4.12.1982-1989.2005 (2005).

- 111 Cramer, R. A. *et al.* Calcineurin Target CrzA Regulates Conidial Germination, Hyphal Growth, and Pathogenesis of *Aspergillus fumigatus*. *Eukaryotic cell* **7**, 1085, doi:10.1128/EC.00086-08 (2008).
- 112 Jahn, B., Langfelder, K., Schneider, U., Schindel, C. & Brakhage, A. A. PKSP-dependent reduction of phagolysosome fusion and intracellular kill of *Aspergillus fumigatus* conidia by human monocyte-derived macrophages. *Cell Microbiol* **4**, 793-803 (2002).
- 113 Vincent, B. M., Lancaster, A. K., Scherz-Shouval, R., Whitesell, L. & Lindquist, S. Fitness Trade-offs Restrict the Evolution of Resistance to Amphotericin B. *PLOS Biology* **11**, e1001692, doi:10.1371/journal.pbio.1001692 (2013).
- 114 Maurer, E. *et al.* *Galleria mellonella* as a host model to study *Aspergillus terreus* virulence and amphotericin B resistance. *Virulence* **6**, 591-598, doi:10.1080/21505594.2015.1045183 (2015).
- 115 Takahashi-Nakaguchi, A. *et al.* Genome sequence comparison of *Aspergillus fumigatus* strains isolated from patients with pulmonary aspergilloma and chronic necrotizing pulmonary aspergillosis. *Medical mycology* **53**, 353-360, doi:10.1093/mmy/myv003 (2015).
- 116 Garcia-Rubio, R., Monzon, S., Alcazar-Fuoli, L., Cuesta, I. & Mellado, E. Genome-Wide Comparative Analysis of *Aspergillus fumigatus* Strains: The Reference Genome as a Matter of Concern. *Genes* **9**, doi:10.3390/genes9070363 (2018).
- 117 Puertolas-Balint, F. *et al.* Revealing the Virulence Potential of Clinical and Environmental *Aspergillus fumigatus* Isolates Using Whole-Genome Sequencing. *Front Microbiol* **10**, 1970, doi:10.3389/fmicb.2019.01970 (2019).
- 118 Escribano, P., Pelaez, T., Bouza, E. & Guinea, J. Microsatellite (STRAf) genotyping cannot differentiate between invasive and colonizing *Aspergillus fumigatus* isolates. *Journal of clinical microbiology* **53**, 667-670, doi:10.1128/jcm.02636-14 (2015).
- 119 Rokas, A. *et al.* What can comparative genomics tell us about species concepts in the genus *Aspergillus*? *Stud Mycol* **59**, 11-17, doi:10.3114/sim.2007.59.02 (2007).
- 120 Aufauvre-Brown, A., Brown, J. S. & Holden, D. W. Comparison of Virulence between Clinical and Environmental Isolates of *Aspergillus fumigatus*. *European Journal of Clinical Microbiology and Infectious Diseases* **17**, 778-780, doi:10.1007/s100960050184 (1998).
- 121 Sasse, C. *et al.* The stepwise acquisition of fluconazole resistance mutations causes a gradual loss of fitness in *Candida albicans*. *Mol Microbiol* **86**, 539-556, doi:10.1111/j.1365-2958.2012.08210.x (2012).
- 122 Valsecchi, I., Mellado, E., Beau, R., Raj, S. & Latgé, J.-P. Fitness Studies of Azole-Resistant Strains of *Aspergillus fumigatus*. *Antimicrobial agents and chemotherapy* **59**, 7866-7869, doi:10.1128/AAC.01594-15 (2015).
- 123 Lackner, M. *et al.* Azole-resistant and -susceptible *Aspergillus fumigatus* isolates show comparable fitness and azole treatment outcome in immunocompetent mice. *Medical mycology* **56**, 703-710, doi:10.1093/mmy/myx109 (2018).
- 124 Veerdonk, F. L., Gresnigt, M. S., Romani, L., Netea, M. G. & Latgé, J.-P. *Aspergillus fumigatus* morphology and dynamic host interactions. *Nat Rev Micro* **15**, 661-674, doi:10.1038/nrmicro.2017.90 (2017).

- 125 Diamond, R. D., Krzesicki, R., Epstein, B. & Jao, W. Damage to hyphal forms of fungi by human leukocytes in vitro. A possible host defense mechanism in aspergillosis and mucormycosis. *Am J Pathol* **91**, 313-328 (1978).

Appendix

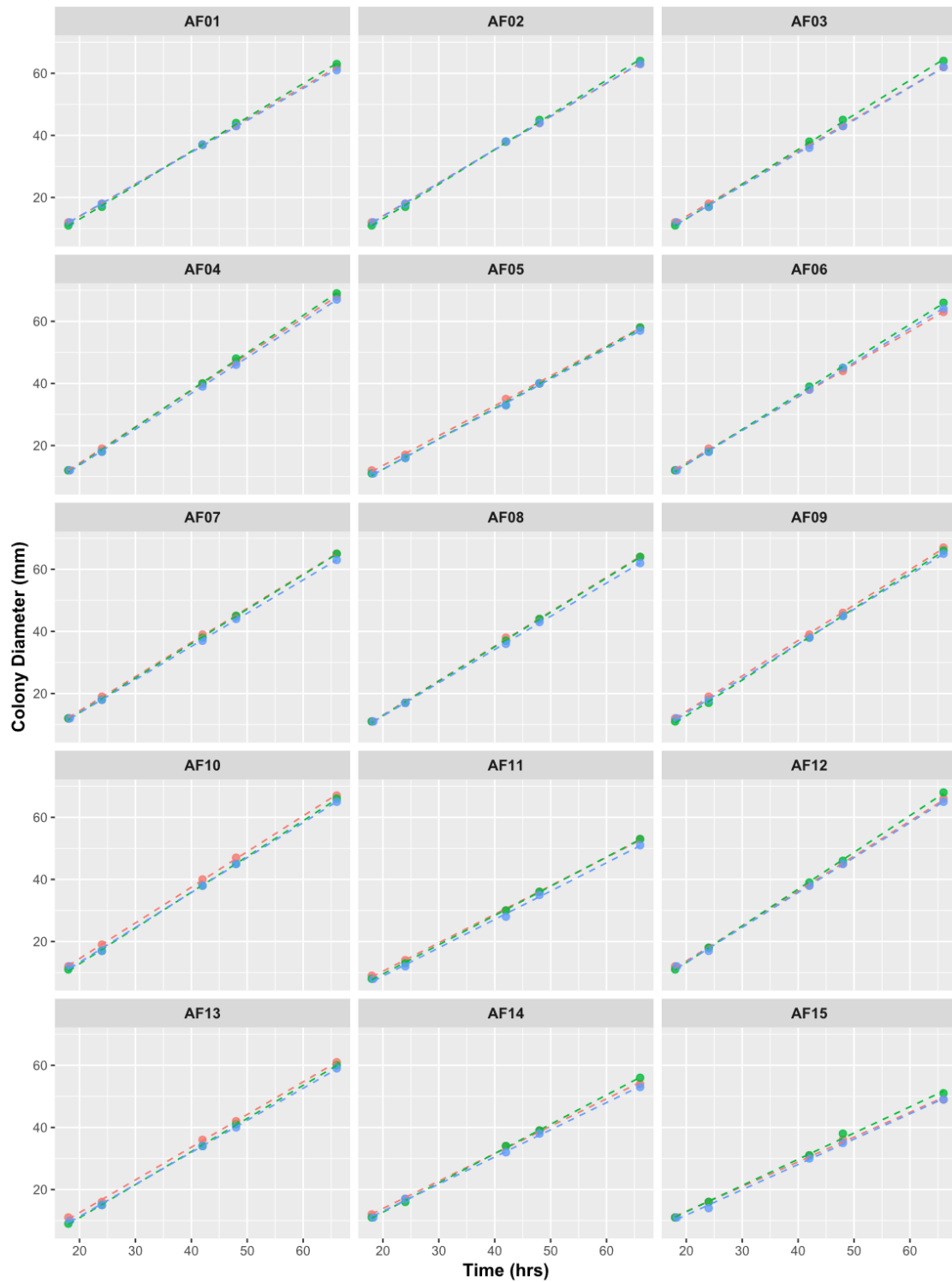


Figure A1. Radial growth curves of clinical *A. fumigatus* isolates on potato dextrose agar at 37°C. For all isolates and all replicates (orange, green, blue), the 5 timepoints at which colony diameter was measured successfully captured linear regions of growth.

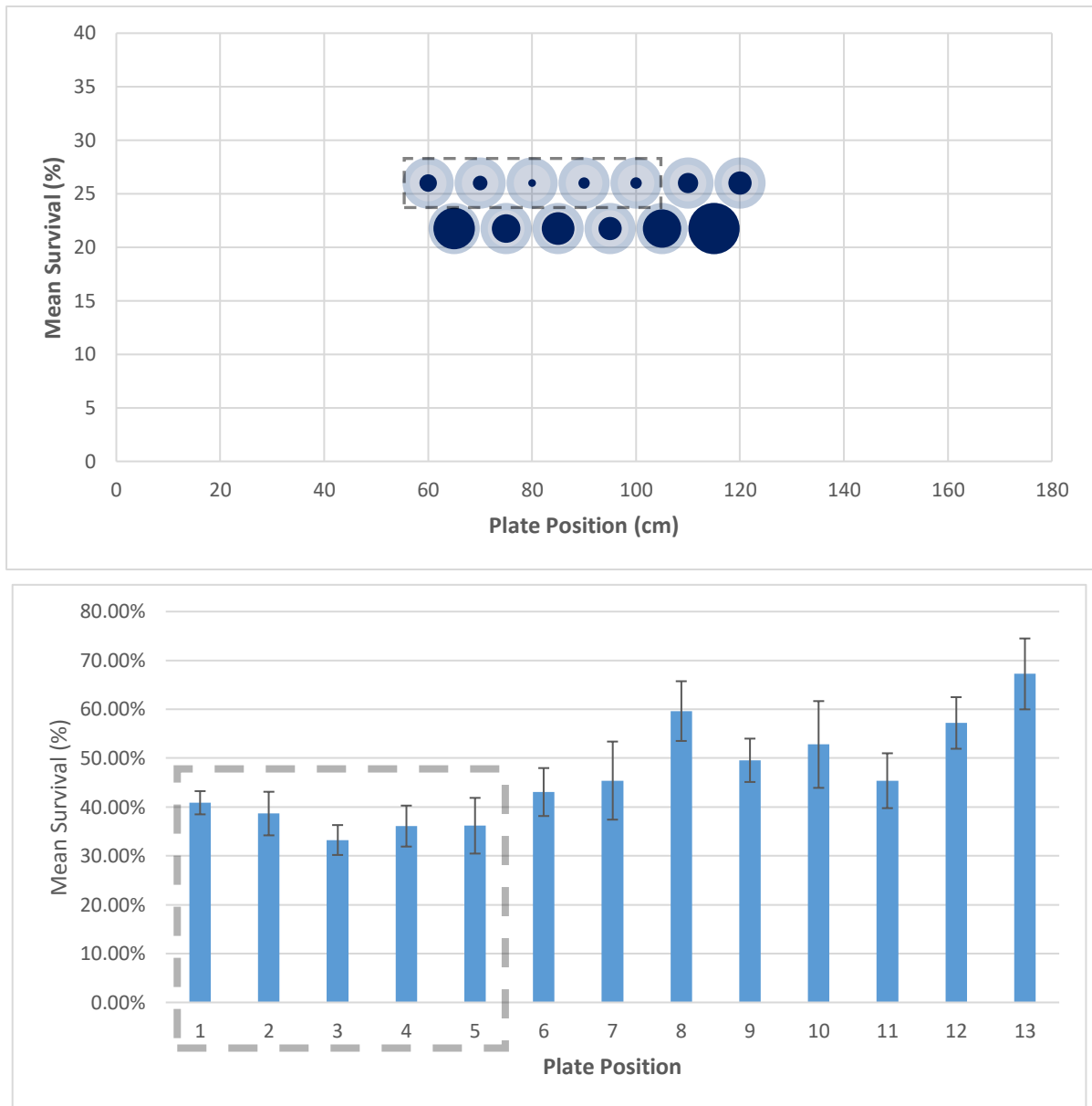


Figure A2. Optimisation and validation of a conidial ultraviolet radiation (UV) resistance assay. Mapping of biosafety cabinet UV exposure. (Top panel) Malt extract agar plates inoculated with a single clinical *Aspergillus fumigatus* isolate were positioned throughout a biosafety cabinet (grey circles). Mean spore survival (%) at each position following 1 min of UV irradiation is shown as a percentage of grey circle width (blue; n=3; biological replicates). Note that for visual clarity, spore survival was re-scaled to 0.15–1, and thus represents relative differences in survival rather than absolute survival at each position. The least variable region was identified (dashed box). (Bottom panel) Unscaled Spore Survival (%) shown for each of 13 plate positions with each experimental replicate (mean +/- SEM shown (right) (n=3)).