Imperial College London
School of Public Health

The evolution of *Cryptococcus neoformans* var. *grubii* in the context of clinical disease using multilocus sequence typing

Sitali P. Simwami

September 2011

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in the School of Public Health
of Imperial College London and the Diploma of Imperial College London

Declaration

I herewith certify that all material in this dissertation which is not my own work has been properly acknowledged.

Sitali P. Simwami

Abstract

The global burden of HIV-associated cryptococcal meningitis (CM) is estimated at one million cases per year, causing up to a third of all AIDS-related deaths. *Cryptococcus neoformans* variety *grubii* (*Cng*) is the most ubiquitous cause of cryptococcal meningitis worldwide, however patterns of molecular diversity are understudied across some geographical regions experiencing significant burdens of disease. *Cryptococcus* species are notable in the degree that virulence differs amongst lineages, and highly virulent emerging lineages are changing patterns of human disease both temporally and spatially. Molecular epidemiology constitutes the main methodology for understanding the factors underpinning the emergence of this understudied, yet increasingly important, group of pathogenic fungi. A multilocus sequence typing (MLST) scheme was used to characterise a genetically depauperate *Cng* population in Thailand, and a contrastingly highly diverse *Cng* population in Cape Town, South Africa. Sequence types (STs) from these populations were integrated into a dataset comprising global STs of *Cng* and patterns of range expansion were traced. Evidence from haplotypic networks and coalescent analyses revealed an ancestral African population of molecular type VNB, from which emerged a VNI lineage. This VNI lineage expanded globally out of Africa and led to the introduction of a limited number of genotypes in novel regions, including Asia. Bayesian inference estimated this spread of VNI to have occurred between 1 600 and 70 500 years ago, putatively vectored by the anthropogenically mediated spread of domesticated pigeons, historically native African birds. Clinical data collected from patients presenting with AIDS-associated CM showed that infecting African *Cng* isolates were associated with poorer long-term survival compared to Asian isolates. As mortality rates reported for these patients in Sub-Saharan Africa are higher than those seen in Asia, supported by a *Galleria mellonella* virulence model, this apparently elevated virulence is postulated as being the result of recombinant progeny with diverse phenotypes being created through frequent meiotic recombination.

Acknowledgments

First, I would like to thank my principal supervisor, Dr. Matthew Fisher, for his vision for this project, guidance, and confidence in my abilities. I could not have asked for a better mentor. Thank you to the members of the Fisher Lab, past and present, and especially to Daniel who somehow knew the answers to all my questions. I also thank my second supervisor, Prof. Christl Donnelly, Dr. Jombart Thibaut and Dr. Miltiadis Mavrakakis for their valuable statistical advice, and the BBSRC for funding this work.

Thank you to members of the CBS (Ferry and Ning), St. George's and Hammersmith Hospitals (Tom, Tihana, Jo, Anne-Marie and Michael), and Duke University (Ana), whose valuable collaborations made this work possible.

Thank you to the the University of Cape Town's Institute of Infectious Disease and Molecular Medicine for hosting me during my fieldwork in South Africa (Widaad, Robert, Graeme). I also extend an especially warm thank you to those who made my stay so memorable (Tsungai, Vee, Slee, Joanna, Banda and Lee) — I will be back!

Thank you to my friends whose support never wavered, despite me regularly disappearing for long periods of time. Bee and Shmi, knowing our relationships remained unchanged made this process a lot easier. I also thank my little sister Moono; our relationship kept me grounded and means more to me than you know.

Thank you to my fiance, Milt, for being everything to me. As a wise Greek once said, 'A hungry bear does not dance'.

Finally, I would like to thank my parents for their constant love, support and wisdom. I owe you everything.

To my parents, Jonathan and Sharon, and to Milt.

Trust in the Lord with all your mind, in all your ways acknowledge him, and he shall direct your path.

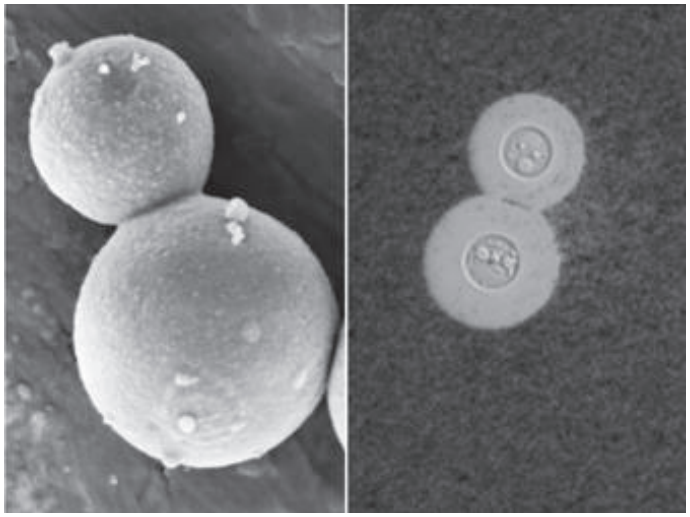
— Proverbs 3:5

Contents

List of Tables

List of Figures

1. Introduction



Kingdom: Fungi, Phylum: Basidiomycota, Genus: *Cryptococcus*
Photo credit: Xiaorong Lin,
`http://www.bio.tamu.edu/USERS/xlin/research1.html\#cn`

1.1. *Cryptococcus*

Cryptococcus neoformans is a basidiomycetous fungus. A pathogenic yeast of oval or spherical shape, 5 to 10 μm in diameter, it is capable of causing life-threatening disease in both immunocompetent and immunocompromised people (Kwon-Chung, 1992). The first environmental discovery of the fungus was in peach juice in 1894 by Sanfelice, the same year as the first clinical isolate was recovered from a human patient's bone lesion by the German pathologist Busse (Mitchell and Perfect, 1995). Since then, *Cryptococcus neoformans*, both a primary and opportunistic pathogen, has become a leading cause of mycotic morbidity and mortality (King and Dasgupta, 2005), causing major burden of disease through cyptococcal meningitis (CM).

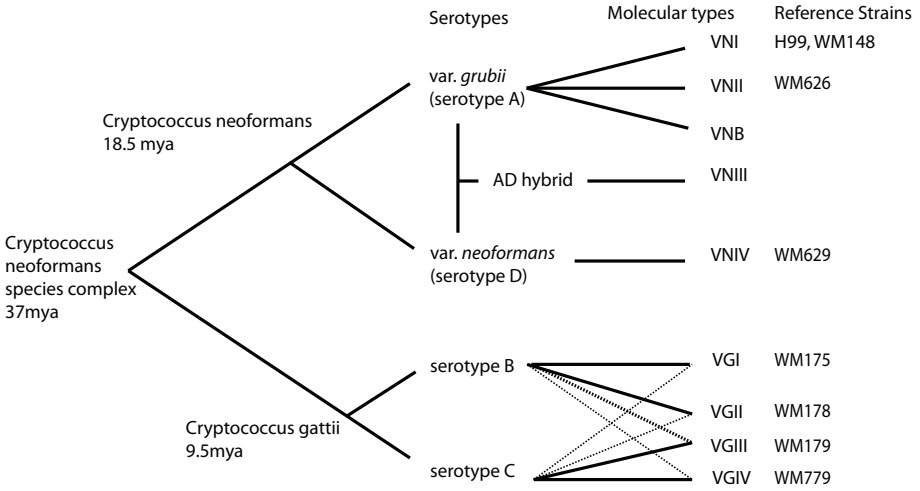
The HIV/AIDS epidemic has majorly contributed to CM's aetiology and known distribution by increasing the numbers of susceptible individuals, especially in Southeast Asia, Southern and Eastern Africa (Mitchell and Perfect, 1995; Bicanic and Harrison, 2004). The global burden of HIV-related CM is estimated at nearly one million cases per year (range: 371 700 – 1 540 000), with a 65% case fatality three months after the onset of symptoms (Park et al., 2009). Sub-Saharan Africa and Southeast Asia are the regions with the highest numbers of cases (720 000 and 120 000 cases per annum, respectively). Despite these rates of mortality being comparable to those caused by malaria (one million mortalities per annum), CM receives only a fraction of the attention, funding and control granted to more widely recognised diseases. The importance of this pathogen, its evolution and increasing threat to global burden cannot be ignored.

1.2. Species classification and distribution

The genus *Cryptococcus* includes 39 known species, and other than *Cryptococcus laurentii* and *albidus*, *Cryptococcus neoformans* is the only species complex found to be associated with clinical disease (Perfect, 2006). Originally believed to be a single species, investigations of the antigenic composition of glucuronoxylomannan within the fungus' polysaccharide capsule led to the detailing of four serotypes: serotypes A, B, C, and D (Evans, 1950; Evans and Kessel, 1951; Evans, 1949; Wilson et al., 1968). These serotypes correspond to two subspecies within the cryptococcal species complex which are distinguishable by biochemical tests (Salkin and Hurd, 1982; Kwon-Chung et al., 1982) as well as based on epidemiological, genetic, clinical, physiological and ecological differences: *Cryptococcus neoformans* (serotypes A, D and an AD hybrid; henceforth *Cn*) and *Cryptococcus gattii* (serotypes B and C; henceforth *C. gattii*; Kwon-Chung, 1975 & 1976; Kwon-Chung et al., 1982 & 2002; Lin and Heitman, 2006; Figure 1.1). Further genotypic and phenotypic evidence led to the separation of *Cn* into two varieties, which are increasingly being recognised as distinct and separate phylogenetic species of the organism: *Cn* var *grubii* (serotype A; henceforth *Cng*) and *Cn* var *neoformans* (serotype D), and their hybrids AD (Franzot et al., 1999).

Molecular typing has resulted in the *Cryptococcus neoformans* - *C. gattii* species complex being further divided into nine major molecular types: VNI, VNII and VNB (serotype A; *Cng*), VNIII (hybrid serotype AD; *Cn* var *neoformans*), VNIV (serotype D; *Cn* var *neoformans*), VGI, VGII, VGIII and VGIV (serotypes B and C; *C. gattii*; Boekhout et al., 2001; Meyer et al., 2003; Bovers et al., 2008a; Bovers et al., 2008c). Unlike with *Cng* and *Cn* var *neoformans*, no correlation has been found between serotype and molecular type for isolates of *C. gattii*.

Figure 1.1.: The evolution of the *Cryptococcus neoformans*' species complex. There are two subspecies, *Cryptococcus neoformans* and *Cryptococcus gattii*, diverged from a common ancestor 37 million years ago. These are further separated into five serotypes and nine molecular groups. Adapted from Lin and Heitman, 2006.



Although *Cryptococcus* is found worldwide, its species and their corresponding molecular types are not equally represented in geographical distribution and vary in their ecological niche and target human host (Jarvis and Harrison, 2007). Accounting for up to 80% of all cryptococcal isolates, *Cn* predominates both environmentally and clinically (Nishikawa et al., 2003). Distributed worldwide, it is commonly associated with soil and pigeon guano. Although unable to systemically colonize these birds, *Cng* can survive the elevated temperatures within their gastrointestinal tract (41-42°C), as well as remain alive for up to two years in the birds' excreta

(Lin and Heitman, 2006). This species tends to cause infections among the immunocompromised, although cases have been reported among the immunocompetent (Casadevall and Perfect, 1998; Dromer et al., 1996; Chen et al., 2008; Kwon-Chung and Bennett, 1984).

The less common species, *C. gattii*, is associated with infections among immunocompetent humans and behaves as a primary pathogen (Duncan et al., 2006; Wilbur and Heyborne, 2009; Qazzafi et al., 2007; Steenbergen and Casadevall, 2000; Bava et al., 1997). Having said this, there have been cases of disseminated cryptococcosis among immunocompromised patients caused by *C. gattii* (Chen et al., 2008; Chen et al., 2000; Litvintseva et al., 2005b). Originally found in the tropics and subtropics in countries such as Australia and Papua New Guinea, parts of Africa, India, Southeast Asia, Mexico, Brasil, Paraguay and Southern California (Ellis and Pfeiffer, 1990; Kwon-Chung and Bennett, 1984), this species is known to be associated with flowering decaying Eucalyptus and other trees (Randhawa et al., 2003; Granados and Castaneda, 2005; Ellis and Pfeiffer, 1992; Granados and Castaneda, 2006; Fortes et al., 2001; Lazera et al., 2000; Krockenberger et al., 2002) and infection can lead to major neurological morbidity (Day, 2004; Kwon-Chung, 1992; Speed and Dunt, 1995; Mitchell et al., 1995; Seaton et al., 1996; Rozenbaum and Goncalves, 1994; Casadevall and Perfect, 1998). Recent findings have been indicative of a wider distribution of this species, especially of VGI strains (Chen et al., 2008; Meyer et al., 2003; Kidd et al., 2004; Fraser et al., 2005; Bovers et al., 2006; Litvintseva et al., 2005a; Campbell and Carter, 2006; Bartlett et al., 2008), as well as a wider host range with an ongoing outbreak in British Columbia, Canada, showing evidence of infection not only among common domesticated animals such as dogs and cats, but also llamas and marine animals (Fraser et al., 2005; MacDougall et al., 2007; Stephen et al., 2002; Duncan et al., 2006).

This distribution of serotypes is reflected in the predominance of molecular types VNI and VGI throughout the world for *Cn* var. *grubii* and *C. gattii*, respectively, with most of the isolates recovered from AIDS patients belonging to VNI and VNIV (serotypes A and D; Meyer et al., 2003).

1.3. *Cryptococcus neoformans* var *grubii*

Within the *Cn* species, several studies report a global predominance of *Cng* (serotype A), the focus of this study (Kwon-Chung, 1992). Cosmopolitan, *Cng* is responsible for about 95% of cryptococcal infections worldwide (Casadevall and Perfect, 1998). The infection is mainly associated with patients of compromised immune status, be it due to external processes or diseases such as organ transplants, the administration of antirheumatic drugs or glucocorticoids, or cancers (particularly those of the bone marrow and blood cells), and accounts for 98% of infections among AIDS patients worldwide (Day, 2004). This being said, there has been evidence of *Cng* causing cryptococcosis among patients with no underlying disease (Jain et al., 2005; Chen et al., 2008; Chen et al., 2000). The same is true of *Cn* var *neoformans* (serotype D), although to a much lesser degree.

There is also bias within the *Cng* variety, with molecular type VNI predominating over VNII and VNB, and accounting for 78% of all *Cn* isolates typed (Meyer et al., 1999). Initially believed to be unique to Botswana, molecular group VNB, which is further subdivided into three groups (VNB-A, VNB-B and VNB-C), has been recently isolated from avian and clinical samples in Brazil, Portugal and Rwanda (Litvintseva et al., 2006; Bovers et al., 2008c).

Cng occurs in several environmental sources including avian excreta and other organic substrates (Randhawa et al., 2003; Nishikawa et al., 2003; Casadevall and Perfect, 1998; Viviani et al., 2001), such as soil contaminated with weathered bird excrement (Viviani et al., 2001), decaying trees (Nishikawa et al., 2003; Randhawa et al., 2003), and household dust (Swinne et al., 1989; Casadevall and Perfect, 1998; Nishikawa et al., 2003). *Cn* var *neoformans* is also found in many environmental sources, but only accounts for 0.4% of all clinical and environmental isolates (Nishikawa et al., 2003). Furthermore, the two *Cn* varieties differ in geographic distribution; *Cn* var *neoformans* is mainly found in Europe and other regions with similar temperate climates, with prevalence in patients reached 30% (Dromer et al., 1996), unlike *Cng* which has been found to occur worldwide. The restriction of *Cn* var *neoformans* to temperate climates is possibly due to its being more sensitive to high temperatures than *Cng*, a virulence phenotype

discussed in subsection 1.7 (Martinez et al., 2001).

Genetic evidence suggests that *Cn* var *neoformans* and *Cng* comprise conspecific biological species. Isolation and divergence, estimated at 18.5 million years ago, is hypothesised to have resulted in the loss of the possibility of "normal mating" which, when coupled with a lack of recent DNA exchange between these two varieties (*Cng* and *Cn* var *neoformans*), has led Bovers (2008) to suggest that the two varieties of *Cn* be elevated to separate species (Bovers et al., 2008c; Sun and Xu, 2007; Kavanaugh et al., 2006). Strong evidence against this is the fact that strains of the two varieties of *Cn* are capable of forming true haploid recombinants. This contrasts with the rarer *Cn*/*C. gattii* crosses, which are genetically similar enough to fuse and continue to form deuploid and aneuploid basidiospores, but fail to undergo meiosis to produce viable progeny (Katsu et al., 2004; Lin and Heitman, 2006; Bovers et al., 2006; Kwon-Chung and Varma, 2006; Bovers et al., 2008b). In some areas of Europe, incidence of AD hybrids in clinical samples is as high as 45-50% of all clinical isolates (Viviani et al., 2006). Further support for the maintenance of serotypes A and D as varieties of the same species is that the separation of isolates according to serotypes based on antigenic properties is flawed and appears not to be correlated with genetically defined clades. An example of this is *Cn* strain CBS132 which was separately determined to be both serotype A, D, and their hybrid AD (Lengeler et al., 2001; Boekhout et al., 2001).

1.4. Life cycle and sexual cycle of *Cn*

Cn's life cycle comprises two distinct fungal forms, sexual and asexual, and is controlled by a single locus, MAT, and a two-allele (idiomorph) bipolar mating type system (Kwon-Chung and Bennett, 1978; Hull and Heitman, 2002). This large single MAT locus (>100 kb) encodes over 20 genes and is responsible for pheromone production and sensing, the driving force of the sexual cycle (Lengeler et al., 2000; Karos et al., 2000; Lengeler et al., 2002; Moore and Edman, 1993). MAT also governs cell-type identity as seen in other sexually reproducing fungal organisms which possess idiomorphs (Glass and Nelson, 1994; Butler, 2007; Debuchy and Turgeon, 2006), and contains other essential genes, both involved and not involved in mating

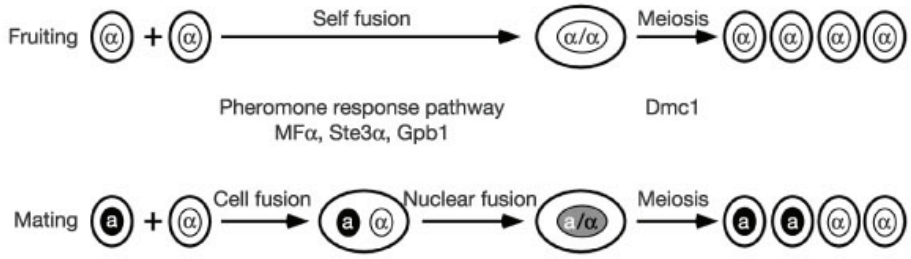
(Bovers et al., 2008a). The bipolar mating system is believed to be an ancestral organisation as it is encountered in species of ascomycete, basidiomycete, and zygomycete phyla (Metin et al., 2010).

As is the case in many fungi, reproduction is possible by either clonality or recombination between opposite mating type strains. The latter is also known as heterothallic mating, as opposed to homothallic mating found in other species of fungi, where the two mating types are packaged in the same genome. Although both have been observed under laboratory conditions, asexual clonal propagation by budding appears to be the most common mechanism (Wickes, 2002). First described in 1975 by Kwon-Chung, the sexual process involves the conjugation between two separate yeast cells of opposite mating type idiomorphs (MATα and MAT**a**) in the presence of limited nutrients. During this process, fusion of the two cells results in the formation of a dikaryon (binucleate) hyphae with a basidium at its tip. Within this basidium, the fusion of nuclei (karyogamy) leads to the **a**/α diploid. This is followed by meiosis, the production of 1 to 2 μm oval and haploid basidiospores, and sporulation (mitosis and basidiospore proliferation; Kwon-Chung, 1975; Kwon-Chung, 1976).

Globally, there is an overwhelming over-representation of the MATα idiomorph both in the environment, (40-fold more abundant) and clinically (30-fold more abundant) across serotypes (Kwon-Chung and Bennett, 1978; Lengeler et al., 2000; McClelland et al., 2004; Yan et al., 2002; Halliday et al., 1999; Madrenys et al., 1993; Barreto de Oliveira et al., 2004; Ohkusu et al., 2002). An exception are the less common AD hybrids of *Cn*, 68% of which possess the MAT**a** allele from serotype A, as well as the MATα allele from serotype D (Yan et al., 2002). Although this discrepancy in mating type prevalence is also observed in other pathogenic fungi, including *Histoplasma capsulatum* and several species of dermatophyte fungi (Kwon-Chung, 1974; Padhye and Carmichael, 1969; Padhye and Ajello, 1977; Kwon-Chung, 1975; Kwon-Chung et al., 1974), mating between isolates of opposite mating types should yield equal numbers of MATα and MAT**a** progeny (Mendelian segregation), as seen in laboratory crosses. This bias was initially postulated to be the result of MATα being more virulent, as observed in mice, as well as the result of a greater frequency of deleterious mutations in MAT**a** cells (Kwon-Chung et al., 1992; Sorrell and Ellis, 1997). A more recent

explanation is the evidence of haploid MATα cells undergoing monokaryotic fruiting (Figure 1.2) under conditions of nitrogen starvation and the absence of a mating partner; a process previously believed to be mitotic asexual mating (Lin et al., 2005).

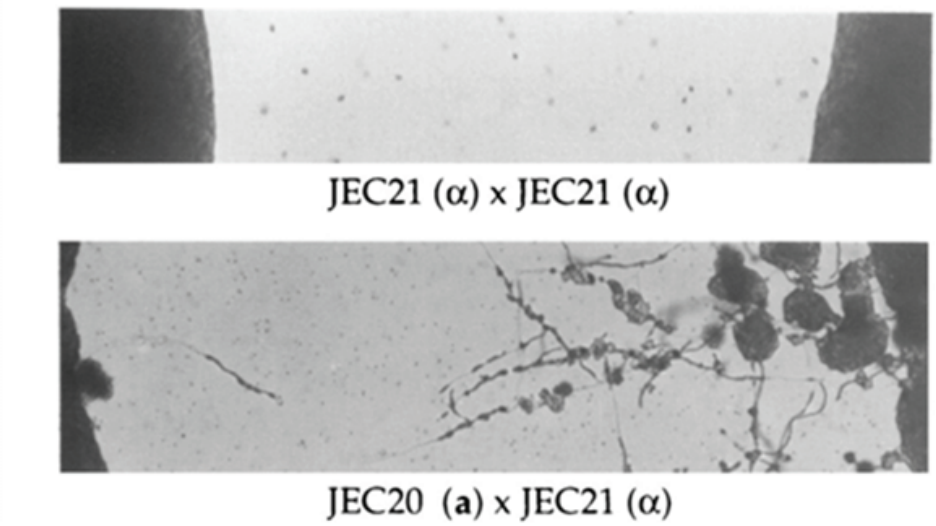
Figure 1.2.: Sexual recombination through monokaryotic fruiting, as well as heterothallic mating. During fruiting, MATα cells diploidize, filament, produce basidia, undergo meiosis, and sporulate, producing recombinant haploid progeny. During mating, nuclear fusion occurs in the basidium, where meiosis and sporulation also occur. Taken from Lin et al., 2005.



Fruiting is a modified form of sexual reproduction between strains of the same mating type, endoduplication, which maintains hallmark features of mating such as diploidization and meiosis and produces recombinant progeny with unique genotypes - the genetic constitution of an individual organism (Davidson et al., 2000; Campbell and Carter, 2006). This modified form of sexual reproduction produces basidiospores and is greatly upregulated in the presence of even a single MAT**a** cell, again through the release of MFa pheromone, and without the need for cell-cell or filament-filament contact (Wang et al., 2000; Wickes et al., 1996; Figure 1.3). This process can only be of evolutionary advantage when one of the two mating types is rare, MAT**a** in this case, increasing chances of long-term survival, similar to self-fertilization in plants. This novel cycle may well play an important role in the ecology and evolution of *Cn*, however its ultimate importance in nature has yet to be demonstrated.

Many pathogenic species have largely clonal population structures, and yet maintain the ability to undergo rare mating, *Cn* being no exception (Lin et al., 2005). The retention of the genetic mechanisms that allow

Figure 1.3.: Serotype D MAT**a** cells (JEC20 strain) regulate haploid fruiting in congenic cells of mating type MATα (JEC21 strain) without the need for contact (lower panel). Haploid fruiting did not occur when cells of mating type MATα were grown in confrontation with themselves (upper panel). Adapted from Wang et al., 2000.



sexual reproduction among species which appear to be primarily asexually recombining suggests the need to preserve the ability to generate novel genetic diversity, and thus the ability to adapt and evolve, in response to changes in the environmental or host (Goddard et et al., 2005). Studies of the model yeast *S. cerevisiae* revealed the benefits of this process: diploid strains/progeny experimentally created by sex were capable of undergoing meiotic recombination, unlike the asexually created ones which were only able to sporulate (Goddard et et al., 2005). The former were better able to cope under various challenging environmental conditions, an advantage conferred by sex also in *Chlamydomonas* (Goddard et al., 2005; Kaltz and Bell, 2002; Grimberg and Zeyl, 2005; Birdsell and Wills, 1996). Theoretically at least, in constant environments, asexual reproduction appears to outperform sex as it does not incur the 'twofold cost of sex'. Sexual reproduction transfers only half a sexual individual's genes to the next generation, while the pathenogenic individual multiplies twice as fast, assuming the number of offspring raised by parthenogenesis and offspring fitness are unaffected. Meiosis results in a breaking-up of co-adapted gene complexes (outbreeding; Smith, 1978; Xu, 2005). On the other hand, sexual reproduction contributes to novel gene combinations from successful alleles, as well as the removal of deleterious mutations, resulting in progeny more capable to adapt to different and new environments (Metin et al., 2010). When there is a lack of mating partners, monokaryotic fruiting may well generate spores that are able to survive under environmental stress (Yue et al., 1999).

1.5. Infection and disease caused by *Cng*

Cn does not require a host for replication but has proven to be both a primary and secondary pathogen in a wide range of mammalian hosts, including domestic animals (dogs, cats), llamas, and also marine animals, such as porpoises. The organism is capable of surviving in mechanical carriers such as pigeons, despite their elevated body temperature of 42°C (Casadevall and Perfect, 1998) and interactions with predatory soil micro-organisms such as *Acantamoeba* (Lin and Heitman, 2006). In humans, infection occurs through the inhalation of mitospores in the environment (Burtet et al., 2000), and there is no evidence of human-to-human or animal-to-human

spread, meaning that *Cn* is a non-transmissible infectious disease (Lin and Heitman, 2006; Aufderheide, 1998). Cells need to be of two microns or less in diameter in order to not only enter the lung but penetrate into the alveoli, followed by colonisation of the airways. This results in primary infection in the lungs (Day, 2004). It has been postulated that acapsular forms of the yeast are inhaled. Supporting this is the fact that the longer *Cn* cells are incubated in soil, the smaller the capsule becomes, whilst retaining viability and infectiousness (Fereshteh, 1970). Emmons reported that organisms isolated from pigeon excreta did not possess a visible capsule (Emmons, 1962). It is therefore believed that upon entry into the host, the acapsular yeast changes morphology, developing the capsule which is an important factor in its fending off attack by the immune system.

Although the separate varieties and species of *Cryptococcus* have shown differences such as in their target host and ability to induce migration of neutrophils (biochemical differences; Dong and Murphy, 1995), they result in similar disease progression (Perfect, 2006). Upon inhalation, asymptomatic infection is most common, leading us to believe that the human immune system is able to adequately control the infection, supported by serological studies (Yauch and Levitz, 2006; Casadevall and Perfect, 1998). A strong intracellular response is then elicited in order to contain infection, with complement-mediated phagocytosis being the main form of defence. If this fails to contain the cryptococcal invasion, there are two possible outcomes of infection. Firstly, haematogenous dissemination to other organs and organ systems may occur, as well as increased cryptococcal burden by budding, with there being a tropism for the central nervous system (Stevens et al., 1999; Tucker and Casadevall, 2002). Within the host, *Cn* is an intracellular parasite, surviving and replicating within phagocytic cells (mainly macrophages; Feldmesser et al., 2000). This feature is observed in the pathobiology of other fungi, including *Aspergillus fumigates* and *Histoplasma capsulum*, and is thought to have evolved as a mechanism to survive predation by *Acanthomoeba castellani* in the natural environment (Casadevall et al., 2003; Latge, 2001; Land, 2002; Levitz, 2001; Mansour and Levitz, 2002). This is one of two mechanisms by which *Cn* migrates to the brain (i.e. a Trojan horse scenario), the second being transcellular penetration of the blood-brain barrier (Chretien et al., 2002; Chang et al., 2004). Other

observed mechanisms involved in the movement of *Cn* within host tissue are the expulsion of live cells from the macrophage to the plasma, and lateral cell-to-cell transfer directly from one macrophage to another. In both processes, the macrophages are not ruptured and the pathogen therefore avoids triggering the host's localized inflammatory response (Ma et al., 2006; Ma et al., 2007). This first outcome is often the case among patients suffering from immunodeficiency. The yeast cells can assume latency, remaining dormant within a lymph node complex and within macrophages which conceal the pathogen from the immune system until reactivation at a later date or upon compromise of the immune system (Baker, 1976). At this point, *Cn* replicates and spreads as previously described (Newman, 2001; Perfect, 2006).

The result of exposure is dependent on various factors, such as the immune state and sex of the host, the variety of the infecting particles, as well as the dosage; in other words, the interaction between the host and pathogen. Infection by *Cn* results in a diverse spectrum of host responses and the outcome of infection is strongly influenced by the host's immune status (King and Dasgupta, 2005). The most common manifestation of cryptococcosis is cryptococcal meningitis (meningoencephalitis; henceforth CM; Banerjee et al., 2001; Stevens et al., 1999). Other recognised syndromes include dementia, pneumonia, Cushing's syndrome, prostatitis, myocarditis and optic nerve atrophy. If left untreated, cryptococcosis is uniformly fatal, with death occurring due to microbiological failure or complications such as raised intracranial pressure or cerebral infarction (Day, 2004).

1.6. Cryptococcosis and HIV/AIDS

A major risk factor of cryptococcosis is diminished cellular immunity, whether due to solid organ transplantation, systemic corticosteroid treatment, diabetes mellitus or hemopoietic disorders (Oursler et al., 1999). Previously mainly associated with cancers, CM's aetiology and distribution has been considerably altered by the HIV epidemic, and its prevalence is indicative of the levels of immune suppression within populations (Mitchell and Perfect, 1995; Schutte et al., 2000; Day, 2004). This fungal infection is now considered an AIDS-defining opportunistic infection, indicative of a CD4 cell

count below 100 cells/μL leading to an increased risk of disease onset (Day, 2004).

Between 5-20% of AIDS patients in the US present with cryptococcosis, whereas the collective percentage in western countries is 10-15% (Oursler et al., 1999; Imwidthaya and Poungvarin, 2000). In Southern African countries such as Zambia, this increases to a staggering 91%, with a mortality rate of 100% compared to the 22% among the US patients (Mwaba et al., 2001; Imwidthaya and Poungvarin, 2000; Oursler et al., 1999). This alarming disparity is, in part, driven by differences in access to services such as HIV testing, health education and the availability of HAART, this latter feature being the best prophylaxis for the prevention of HIV-associated cryptococcal disease (Bicanic and Harrison, 2004; Day, 2004). Differences in genotype, ecology and exposure may also account for the difference in prevalence and mortality of infection (Bicanic and Harrison, 2004; Simwami et al., 2011).

Readily treatable in immunocompetent patients, CM cannot be cured in AIDS patients. Patients with HIV-related CM must undergo maintenance antifungal therapy life-long or until immunoreconstitution is reached through antiretroviral therapy (Bicanic and Harrison, 2004). Conventional therapies have been investigated using models of immunosuppressed animals such as rats (Gross et al., 2006), as well as trial studies among human patients (Bicanic and Harrison, 2004; Bicanic et al., 2008). The recommended treatment is a three-stage regime of induction, consolidation and maintenance. During the first two weeks, a combination of amphotericin B at 0.7-1 mg/kg/day and flucytosine at 100 mg/day is administered. This combination prevents the developing of resistance to flucytosine. Although yielding the most rapid initial antifungal action, prolonged administration of this therapy results in nephrotoxicity. To counter this, and the high rate of relapse, the patient is then switched to fluconazole at 400 mg/day for eight weeks, followed by lifelong maintenance therapy at 200 mg/day thereafter (Bicanic and Harrison, 2004). Early diagnosis and drug susceptibility testing are very important for the outcome of the infection among these vulnerable patients.

1.6.1. AIDS in Thailand

The first reported case of AIDS in Thailand was in 1984 (Phanuohak et al., 1985), and 23 years later the number of people living with HIV/AIDS was estimated at 610 000 (UNAIDS/WHO, 2008). The epidemic's progression within this country is well documented due to Thailand's extensive HIV surveillance system, the first systems in place being the Ministry of Health's sentinel surveillance system (Ungchusak et al., 1989) and the Royal Thai Army's testing of military conscripts upon recruitment (Sirisopana et al., 1993). The earliest cases of HIV/AIDS incidence were predominantly among homosexual and bisexual men. This was followed by a rapid increase in infections among injecting drug users (IDUs) and female commercial sex workers (CSWs). A third subsequent wave occurred among CSW clients, as documented by the increased number of men attending government sexually transmitted disease clinics. The wives and girlfriends of these men made up the fourth wave of infections with numbers almost doubling annually and, finally, their offspring, as evidenced by increased paediatric AIDS cases (Brown et al., 1994; Weniger et al., 1993). Consistently higher rates of HIV prevalence were reported in the north of Thailand in all described population except for IDUs (Weniger et al., 1993; Brown et al., 1994; Chariyalertsak et al., 2001).

The HIV-control program in Thailand began in 1989 but initial response was weak, with the Thai government downplaying the threat, opting for denial and inaction (World Bank, 1998; UNDP, 2004). It was not until 1991 that AIDS prevention and control became a national priority of the highest level, and one of the most aggressive and extensive such programs implemented in Asia or the world (Owens, 1991). This offensive included the provision of basic public education targeting risk behaviors through a massive campaign, programs such as the '100 percent condom program' (Hanenberg et al., 1994), the insistence on confidentiality and anti-discrimination through the repealing of repressive policies (Porapakkham et al., 1995), and the increase in the annual allocated budget, from $684 000 in 1988 to $82 000 000 in 1997 (World Bank, 2000). Along with the mobilizing of community efforts, collaborations among ministries beyond those responsible for public health, as well as non-governmental organizations led to about 7 million new infections being averted (Brown et al., 1994; World Bank,

2000). In addition, the scaling up of the roll-out of antiretroviral drugs (ARVs) was made possible by the local manufacturing of generic versions and compulsory licensing of second-line drugs, including Merck & Co.'s Sustiva®(efavirenz; Alcorn, 2006). This halved the cost of the drug and resulted in the expansion of antiretroviral treatment (ART) coverage, which increased to 62% in 2009 (WHO et al., 2010). This was important, as there has been evidence of a decrease in cryptococcosis in the face of increased availability of ART in North America and Western Europe (Dromer et al., 2004; Kaplan et al., 2000; Mirza et al., 2003; van Elden et al., 2000).

1.6.2. AIDS in South Africa

There are an estimated 5.6 million people living with HIV/AIDS in South Africa. Despite being home to only 0.7% of the world's population, this country accounted for 17% of global HIV prevalence in 2007 (UNAIDS, 2008; UNAIDS, 2009; Abdool Karim et al., 2009). The impact of HIV is seen in various age groups, the most affected being the 25- to 49-year-olds. Forty-one percent of the 607 184 deaths in South Africa in 2006 was attributable to this demographic; up 12% from 1997 (Statistics South Africa, 2008).

Various strategies are being applied in the containment of this disease. Tackling mother-child transmission (MTCT) which contributed significantly to the 330 000 HIV+ children under the age of 15, the national Strategic Plan provided prophylaxis to 88% of pregnant seropositive women in 2010. This programme, PMTCT, had a 96.5% success rate in transmission prevention according to a study conducted by the Medical Research Council (WHO et al., 2010; Times Live, 2011). Voluntary male circumcision, proven to decrease transmission risk from women to men by 60%, is now part of the government's HIV counseling campaign (Auvert et al., 2005; Motsoaledi, 2010). Similarly to Thailand, the media were instrumental in raising awareness through campaigns targeting various age groups: the 'Soul City' campaign targeted adults, while the 'Soul Buddyz' campaign was aimed at children, and the 'LoveLife' campaign at teens (Shisana et al., 2009).

This regional HIV/AIDS hyper-epidemic has dramatically increased the immunocompromised population within a country which was already at

a disadvantage socially, economically and environmentally. The results of apartheid included inadequate health service, and squatter settlements occupied beyond capacity, conditions conducive to the spread of the virus mainly among black South Africans. This disadvantage was further compounded by the government's approach to the issue which included avoidance, denial and efforts to undermine scientific treatments. Despite positive results and changes in trends including increased condom use from 31% in 2002 to 64.8% in 2008, increased availability of ARV with 2205 health centers providing treatment in 2011 versus 490 in early 2010, and the considerable reduction of drug prices, further issues remain; these include human resource shortages, the delayed initiation of treatment, and the continually increasing number of cases (Motsoaledi, 2011; Department of Health, 2011; Avert, 2011).

1.7. Classical virulence phenotypes of Cn

Virulence factors increase the degree of pathogenicity of a microbe (Buchanan and Murphy, 1998). Classical virulence phenotypes of *Cn* are capsule production, melanin synthesis and growth at high temperatures.

The polysaccharide structure of unbranched chains of α1,3 linked mannose units, substituted with various xylosyl and β-glucuronyl groups, forms the distinctive and most studied fungal antiphagocytic feature of *Cn*, its capsule. This is very dynamic as it can vary in size from being barely visible on India ink preparations to surpassing the diameter of the yeast cell itself (Small, 1989). Capsular enlargement occurs in response to the environmental stimuli, including elevated ambient pCO_2 (Granger et al., 1985), low iron concentrations (Vartivarian et al., 1993), and the presence of mammalian serum (Zaragoza and Casadevall, 2004). Originally thought to be indicative of virulence, there is no defined correlation between the capsular size and virulence (Dykstraet et al., 1977; Torres et al., 2005). However, *Cn* mutants with fixed capsule sizes exhibited attenuated virulence, showing that the regulation of capsular size is important in response to changing environments (Granger et al., 1985). Having said this, two groups of *Cn*-infected patients showed similar clinical outcomes, despite one being infected with poorly encapsulated strains and the other with prominent capsular *Cn*

strains (Torres et al., 2005; Perfect, 2006).

In terms of assisting its proliferation within the host, *Cn*'s polysaccharide capsule is believed to possess antiphagocytic properties, with acapsular cells being phagocytosed at an average rate of 78%, versus only 24% for encapsulated yeasts (Bulmer, 1967). The capsule inhibits the phagocytosis of the fungus by macrophages (Granger et al., 1985; Bolanos and Mitchell, 1989; Levitz and Dibenedetto, 1989), dendritic cells (Vecchiarelli et al., 2003), neutrophils (Diamond et al., 1972; Kozel et al., 1984; Dong and Murphy, 1997), and endothelial cells (Ibrahim et al., 1995). Further actions attributable to the capsule are the induction of immune unresponsiveness in the form of the suppression of antibody synthesis (Kozel et al., 1977; Murphy and Cozad, 1972), inhibition of leukocyte migration, dysregulation of cytokine secretion, enhancement of HIV infection (Cherniak et al., 1995), induction of apoptosis, and complement depletion (Casadevall and Perfect, 1998).

Melanin synthesis occurs during infection and involves the conversion of diphenolic compounds by a laccase enzyme through a series of autoxidation steps, a process which is reported in several pathogenic fungi (Langfelder et al., 2003). This synthesis of melanin protects *Cn* as melanin scavenges host-produced antioxidants, protecting *Cn* from oxidative damage (Polacheck et al., 1990). Decreased virulence has been reported in mutants lacking melanin, and the disruption of the *LAC1* gene linked to melanin synthesis results in a hypovirulent phenotype (Salas et al., 1996; Kwon-Chung et al., 1982; Rhodes et al., 1982; Kwon-Chung and Rhodes, 1986). This ability to synthesise melanin is not restricted to *Cn* alone and has been observed among other human fungal pathogens such as *Histoplasma capsulatum* (Nosanchuk et al., 2002), *Aspergillus spp.* (Rosas et al., 2000), and *Wangiella dermatitidis* (Feng et al., 2001), as well among bacterial pathogens (Nosanchuk and Casadevall, 2003).

Elevated body temperature is a primary protective feature among mammalian hosts. Without being thermotolerant above the optimal temperatures of 25 to 35°C of most fungi, *Cn* would be unable to infect, and survive, in ectothermic hosts, as evidenced by the fact that only about 24 out of 1.5 million fungal species consistently cause disease among humans (Perfect, 2006). Other Cryptococcal species possess the ability to produce either a capsule, melanin or both, but this virulence phenotype is unique to

Cn and certain *C. laurentii* and *C. albidus* strains. Furthermore, *Cng* isolates are less sensitive to growth inhibition and exhibit higher virulence at higher temperatures compared to *Cn* var *neoformans*; this is hypothesised to confer *Cng* a survival advantage in warmer climates, thus leading to its broader global distribution (Martinez et al., 2001). The gene encoding calcineurin A, a serine-threonine specific phosphatase, is involved in the stress response of *Cn* — stress being defined as elevated temperature, 5% CO_2 and above, alkaline pH or high concentration of cations. This gene has been shown to be a basic requirement for the survival of *Cn* within the human host environment, with mutants being nonpathogenic in immunosupressed rabbits (Odom et al., 1997).

Several genes and genetic pathways have been identified, owing to their impact on virulence observed in animal models using site-specific gene disruptants (Table 1.1). For example, the *MAN1* gene encodes Phosphomannose Isomerise (PMI), an enzyme critical in mannose metabolism. Mannose makes up a large portion of the polysaccharides capsule's three components glucuronoxylomannan (GXM), galactoxylomannan (GalXM) and a mannoprotein (MP; Wills et al., 2001; Cherniak and Sundstrom, 1994). Without *MAN1*, the protective capsule is compromised, as is the pathogen's ability to survive and persist within the human host. Similarly, trehalose is a storage carbohydrate involved in regulating the stress response against heat in yeast (Devirgilio et al., 1994; Wiemken, 1990). Disruption of the *TPS1* gene, trehalose-6-phosphate synthase, results in a disruption of protective mechanisms crucial for *Cng*'s survival at elevated temperatures (Petzold et al., 2006). Disruption of the *ILV2* and *MAN1* genes results in the complete elimination of yeast from the host and full survival of mice. The relative impact of the disruption of genes involved in capsule formation (for instance *CAC1, GPA1, PKA1* and *SPE3/LYS9*; Table 1.1) is not as severe, resulting in a reduction in yeast counts compared to wild type, with the majority of animals surviving. The disruption of *AOX1, FHB1*, *PLB1*, *SOD1*, and *STE12D* and *20A* results in mild impact on virulence, with all animals succumbing to infection despite statistically reduced yeast counts and prolonged survival in the null mutants compared to wild type. This type of information is of use when it comes to the selection of antifungal targets and vaccine strategies (Perfect, 2006), and much recent research uses com-

parative genomics to identify virulence factors that are shared across fungal taxonomic groups (Lorenz, 2006; Stajich and Dietrich, 2006).

Table 1.1.: Genes linked to the major virulence phenotypes of *Cn* and their impact on virulence in animal models using site-specific gene disruptants[a]. Modified from Perfect, 2006.

Virulence Phenotype	Gene	Impact on virulence
Capsule formation	*ACA1, CAC1, CAP10, CAP59, CAP60, CAP64, CAS2, GPA1, MAN1, MET6, PKA1, SCH9, SPE3/LYS9, STE12D*[b]*, STE20A*[b]*, USX1, VPH1*	Attenuated virulence
	CAS1	Hypervirulent
Melanin production	*ACA1, CAC1, CLC1, GPA1, LAC1, MET3, PKA1, STE12D*[b]*, VPH1*	Attenuated virulence
	PKR1A[b]	Hypervirulent
Temp (high-temperature growth, $37 - 39°$C)	*CCN1, CNA1, CNB1, CPA1, ILV2, MPK1, RAS1, SOD2, SPE3/LYS9, STE20A*[b]*, TPS1, TSA1, UGD1, VPH1*	Attenuated virulence

[a] A Genes are given names and numbers consistent with *Saccharomyces* homologs in many cases of designated letters taken from the references which give specific descriptions of the genes; [b] A = Serotype A; D = Serotype D

Cn is a free-living organism and its pathogenesis in mammals, characterised by its replication and survival within macrophages, did not develop as the result of the obvious need for an animal host. Instead, mechanisms leading to survival in the wider environment, involving the polysaccharide capsule, phospholipase activity and melanin production observed in the *Cn*-amoeba complexes, must have been developed in response to environmental predators, namely amoebae, slime molds and nematodes (Mylonakis et al., 2002; Steenbergen et al. 2001; Steenbergen et al., 2003). Therefore, in order to better understand this fungus' pathogenicity, it is important to investigate it in the organism's environmental context, not just the clinical one.

1.8. Animal Host Models

Understanding the pathogenesis of *Cryptococcus neoformans* is important not only because of its propensity to cause disease in immunocompromised individuals — being a model for the switch from a benign environmental organism to a voracious pathogen of mammals and amoebae (Jones, 2007) — but also because it has been described as an excellent model fungal pathogen for the identification and study of virulence factors (Idnurm et al., 2004). The latter is due to its being a nonspecific pathogen, infecting mammals, birds, reptiles, amoebae, slime mold, flies, worms, and moths, giving range to a wide variety of potential experimentally tractable models.

One of the issues with studying the behavior of pathogenic organisms *in vitro* is that we cannot be sure of how representative our observations are of what actually happens *in vivo*. In an attempt to bridge the gap between *in vitro* and *in vivo*, model hosts are used in order to stimulate the processes which may be occurring in the host of interest.

The primary animal models used to investigate fungal pathogen virulence, host immunity and disease progression are mice (Najvar et al., 1999), rats and rabbits. In the past decade, there has been a shift from the use of vertebrate hosts to non-mammalian hosts including insects, worms, amoebae, fish, plants and slime molds. There are a number of rationales for using model hosts, including the potential for (i) the dissection of virulence mechanisms, (ii) comparative immunological studies, (iii) the investigation of the emergence and maintenance of virulence mechanisms originally acquired by certain microbes in response to environmental challenges , (iv) evolutionary studies, and (v) drug screening (Casadevall, 2005).

The discovery that the mechanisms and pathogenesis traits involved in microbial virulence remain unchanged in both mammals and non-vertebrate hosts makes these new host models relevant in the study of fungal mammalian pathogenesis (Salzet, 2001; Mylonakis et al., 2007). *Cn* is no exception to this. Genes and mechanisms associated with virulence such as signal transduction pathway genes *PKA1* and *RAS1*, laccase production and α-mating type in mammals also play key roles in the killing of various invertebrates such as nematode *Caenorhabditis elegans*, fruit fly *Drosophila melanogaster*, and insect *Galleria mellonella* (Mylonakis et al., 2002; My-

lonakis et al., 2005; Apidianakis et al., 2004). These new animal models for investigating host-fungal interactions are promising as they are relatively simple, genetically tractable, inexpensive, and ethically expedient. Some of these hosts can be manipulated to better mimic the actual environment of interest (e.g., immunocompromised *Dictyostelium discoideum*), and others, such as *C. elegans* and *Bombix mori*, have complete genome sequences, allowing for the study of host response).

Despite all these advantages, there are, of course, disadvantages. Practical limitations include the fact that *D. melanogaster* and *C. elegans* are not viable at 37°C, meaning that we cannot use them to investigate virulence expression at mammalian temperatures. Another technical challenge is the administration of inocula and/or antimicrobial substances is technically demanding due to size. Furthermore, there may be differences in the innate immune systems of metazoans that may reduce our ability to translate results between different model systems and clinical conditions (Mylonakis and Aballay, 2005).

1.8.1. *Cn* – *Galleria mellonella* pathosystem

With regards to *Cryptococcus*, recommended hosts include *Acanthamaebae catellanii* and the larvae of the wax moth *Galleria mellonella*. The latter is of particular interest as the wax moth larvae can tolerate mammalian temperatures and so experiments can be designed at 37°C. It is important to note, however, that *Cn* kills moths faster at 37°C than at 30°C, suggesting that either temperature plays a regulatory role in virulence or that there is a weakening in the moth's immune system at higher temperatures (Casadevall, 2005).

Of particular interest is the fact that *G. mellonella*'s hemocytes (plasmocytes and granulocytes) are capable of microbial killing by phagocytosis (Brennan et al., 2002; Bergin et al., 2005; Tojo et al., 2000). In response to sensing the pathogenicity of foreign bodies, *G. mellonella* is able to alter the diversity of its hemocyte, as seen when infected with *Stinerema* nematodes (Gagen and Ratcliffe, 1976; Tojo et al., 2000). The process of phagocytosis is similar to that of mammals in that opsonic ligands bind to the foreign particle's surface, followed by recognition by specific receptors and internal-

isation driven by an intracellular cascade. In the insect, the activation of the prophenoloxidase cascade and the production of reactive oxygen intermediates lead to the phagocytosis of *Cng* (Kavanagh and Reeves, 2004).

Further support for the use of *Galleria* in predicting the behaviour of the *Cn* pathogen in the human host is the striking correlation in the efficacy of antifungals in both the moth and humans, with the combination of amphotericin B and 5-flucytosine being most effective against *C. neoformans* in both (Casadevall, 2005; Mylonakis et al., 2006); this was found to be the best drug combination in human clinical trials (Brouwer et al., 2004). This is useful to rapidly screen and assess the efficacy of combinations of antifungal agents (Mylonakis et al., 2006).

1.9. Rationale and importance of study

Cn has been established not only as an important human pathogen, but as a model yeast for the study of fungal pathogens. As the population of immunosuppressed individuals increases, the potential for continued increase in disease burden of AIDS-related meningitis cannot be ignored, particularly in the developing world (Bicanic and Harrison, 2004; Day, 2004; Park et al., 2009). The HIV/AIDS epidemic has driven increased *Cryptococcus* infection rates via the rapid increase of immunosuppressed populations (Day, 2004; Mitchell and Perfect, 1995; Schutte et al., 2000).

Global typing is key to elucidating the origin, natural ecology, population structure, and epidemiology of environmental *Cn* var *grubii*. In addition, the genetic variation inter- and intra-species described by typing enables us to characterise and determine disease causation and progress. In collaboration with effective surveillance and monitoring, these can lead to the control of infectious diseases (Thompson et al., 1998).

Laboratory and analytical tools of population genetic approaches have contributed to our understanding of several infectious disease agents. Multilocus restriction fragment length polymorphism and microsatelite analysis revealed a clonal population structure in the protozoan *Toxoplasma gondii* consisting of two or three clonal lineages in animals and humans (Howe and Sibley, 1995; Ajzenberg et al., 2002). In *Batrachochytrium dendrobatidis*, genome sequencing lead to the description of the secretive fungus' life cycle,

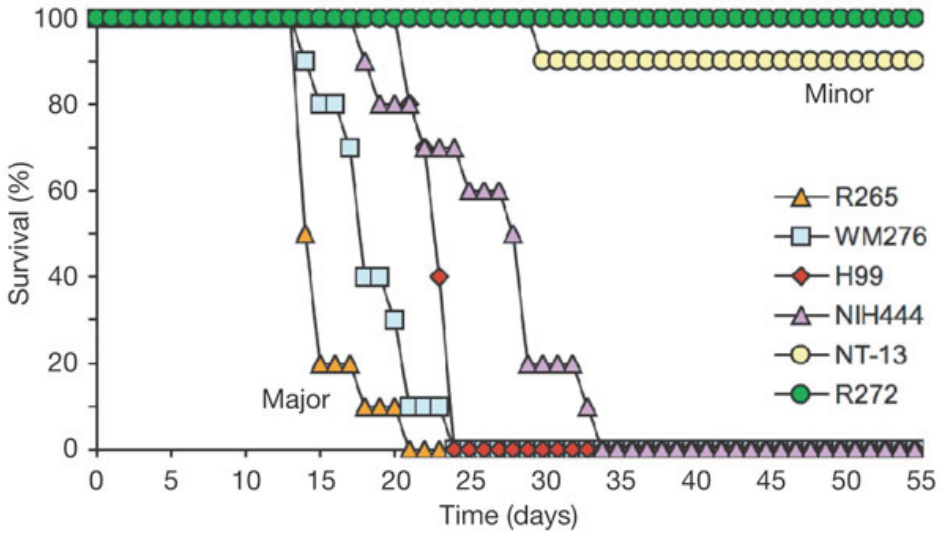
which in turn revealed the mechanisms of pathogenicity in the amphibian host (Rosenblum et al., 2008; Fisher, 2008). Similarly, genomic data have contributed to our understanding of the genetic basis of pathogenicity of human pathogens *Aspergillus fumigatus* and *Ehrlichiosis* agents (Nierman et al., 2005; Hotopp et al., 2006). These same population genetics tools have greatly contributed to our current understanding of the *Cryptococcus*-host interaction and the pathogen's life cycle (Lin and Heitman, 2006; Reedy et al., 2007).

In terms of epidemiology, the characterisation of globally distributed *Cng* isolates using MLST led to Litvintseva et al. (2005) uncovering a global distribution of clonally derived and genetically homogeneous VNI genotypes. Also elucidated was a highly diverse and recombining population of molecular type VNB in Southern Africa (Litvintseva et al, 2003; Litvintseva et al., 2006). These findings led to the proposal of the "Out of Africa" hypothesis which postulates that *Cng* has an evolutionary origin in Africa, followed by a global expansion. This spread was possibly vectored by the migration of avian species. The common pigeon (*Columba livia*), originating in Africa, is considered a mechanistic carrier and potential spreader of the fungus, its faeces being a common environmental source of *Cng* (Johnston, 1992; Lin and Heitman 2006; Swinne-Desgain, 1976). These birds were domesticated in Africa approximately 5 000 years ago and introduced to Europe, then subsequently distributed to many parts of the world during the European expansion in the last 500 years (Mooney, 2000; Grzimek, 2004); a range expansion that may have led to pigeon vectors allowing *Cng* to broaden its global ecological range. While wind transport has also been hypothesized as a potential method of the global dispersal of *Cng*, as demonstrated by the potential for dispersal of other infectious fungi (for instance *Coccidioides immitis* by wind-blown arthroconidia; Pappagianis and Einstein, 1978), Casadeval and Perfect (1998) state that this is unlikely, due to the *Cng* basidiospores being unsuitable for long-distance wind dispersal.

In terms of cryptococcal disease and its progression, an increased understanding of *Cng* virulence has been an important feature. Genetic variations between and within divergent strains have been linked to host-specific infection (Hull and Heitman, 2002): the two cryptococcal species target hosts with different immunity (competent or suppressed). Similarly, differences

in the progression of disease are linked to genetic variation: Dong and Murphy (1995) suggest that infections by *C. gattii* are more difficult to treat successfully than those by *Cn* var *neoformans*. Fraser et al. (2005) report two distinct genotypes of the *C. gattii* species involved in an outbreak of meningoencephalitis on Vancouver Island, Canada in 1999. Despite the two being closely related, they vary considerably in prevalence, likely due to differences in pathogenicity. An intranasal murine model revealed that the major outbreak genotype isolates were hypervirulent, whereas the minor ones were either avirulent or greatly attenuated (Figure 1.4; Fraser et al., 2005). Assessing intracellular proliferation within macrophages using *in vitro* and *in vivo* studies, live-cell imaging, and microarray analysis, Hansong et al. (2010) also found significant inter-strain variation between *C. gattii* isolates from Vancouver, with many of the hypervirulent isolates replicating better intracellularly than other global *C. gattii* strains (Levitz, 2008).

Figure 1.4.: Inter-strain virulence of strains of the *C. gattii* outbreak on Vancouver Island. Groups of ten A/Jcr mice were infected with one of six indicated *C. gattii* strains by intranasal inhalation. The major outbreak strains (R265, WM276, H99, NIH444) are hypervirulent while those of the minor genotype (NT-13 and R272) are attenuated and avirulent respectively. Taken from Fraser et al., 2005.



Molecular typing of *Cng* also determined that globally dominant *Cng* strains were being isolated from both environmental and clinical sources, indicative of the presence of "fully virulent" specimens in the environment (Casadevall and Perfect, 1998; Casadevall et al., 2003; Idnurm et al., 2005). Despite retaining the ability to cause disease in murine models, environmental isolates have been reported to be less virulent, either needing a larger dose to cause similar infection as clinical isolates, or with only a proportion of isolates being capable of causing disease at all (Fromtling et al., 1989; Hasenclever and Emmons, 1963; Silva et al., 2006). A more recent study by Litvintseva and Mitchell (2009) confirmed this diminished virulence with all but one of eleven environmental *Cng* strains being nonpathogenic and nonlethal over 60 days, in contrast to seven of ten lethal clinical isolates (median time to death = 19 to 40 days). Interestingly, environmental and clinical strains of identical genetic makeup, as determined by molecular typing, differed in virulence (Litvintseva and Mitchell, 2009).

As much as population genetics has unlocked the infectious disease epidemiology of *Cng*, many questions still remain. Less well established is our understanding of the fungus' life outside the host, its global spread, population structure, and how it evolved into such a deadly pathogen (Lin and Heitman, 2006). For instance, the discovery of same-sex mating forces us to reassess the contribution of various sexual cycles to the evolution of *Cng* and its current population structure (Figure 1.2; Lin et al., 2005). Furthermore, the potential to generate hypervirulent progeny through this newly described mating cycle (as witnessed in sister species *C. gattii* in the Vancouver Island outbreak) highlights how little we know about the evolutionary potential of the fungus and therefore increased threat of cryptococcal disease. Litvintseva et al.'s (2006) detailing of a geographically isolated VNB population whose charateristics are indicative of ancestry in the face of a predominantly clonal global VNI population challenges our understanding of the origin of of *Cng*. Finally, variations in virulence of strains identical in genotype but from different sources (environmental or clinical) leads us to question what we currently believe to be the environmental sources of human cryptococcosis. It also poses the question of whether or not all environmental isolates are equally capable of causing disease and, if not, what makes the subset of virulent strains so potent in the immunocompromised

human host.

1.10. Thesis overview

This study aims to tackle some of these important epidemiological questions using molecular typing tool multilocus sequence typing (MLST). Written in a conventional manner, this introductory chapter is followed by a description of methods used (Chapter 2); subsequently, the findings of this study are detailed and discussed in seven results chapters (Chapters 3 to 9). A final chapter is used to discuss and integrate these findings. Below is a brief overview of my contribution to various aspects of *Cng*'s epidemiology and the disease it causes, as covered in these chapters.

Chapter 3 begins by describing molecular methods used in the typing of fungal pathogens, and their advantages and disadvantages. Despite increased interest in the fungus, the multiple published studies are not informative, as a lack of reproducibility and consensus in methodology and data output renders cross-laboratory comparisons impossible. The Cryptococcal Working Group's standardisation of the typing method applied to *Cng*, MLST, and its nomenclature are presented (Meyer et al., 2008; Appendix A.1). This method is unambiguous and the data universally available for comparative analysis at `http://cneoformans.mlst.net/`.

In Chapters 4 and 5, MLST is applied to two previously undescribed populations of *Cng* from Thailand and Cape Town. From the results of analyses of molecular variance and phylogeny, it is clear that the two vary considerably in population structure, with the former being highly clonal while the Cape Town isolates are significantly more genetically diverse and feature more balanced molecular groups and mating types.

These two newly typed samples are integrated in to the current global *Cng* population in Chapter 6. The structure of this wider sample is characterised by globally distributed genotypes as well as geographically restricted genotypes within the African subpopulation. As my findings are in keeping with those of Litvintseva et al. (2006), the "Out of Africa" hypothesis is tested, and coalescent Bayesian analysis is used to estimate the pattern of *Cng*'s worldwide evolutionary spread.

In Chapter 7, I take advantage of data on the progression of cryptococcal disease among HIV+ CM patients available for the clinical Thai and Cape Town samples typed in the previous chapters (Chapters 4 and 5). Combined with MLST data, tests of association are applied in order to identify host and fungal factors predictive of outcome of infection. My findings are discussed in the context of treatments used for CM.

A field study was carried out in order to investigate the sources of cryptococcal infection among the CM patients of Cape Town origin whose infecting isolates were typed (Chapter 5). The aim of this was to isolate and type *Cng* from tree, soil and avian excreta niches in order to address questions of association between environmental and clinical *Cng* isolates and their variations in virulence, if any.

Finally, virulence assays were carried out using a *Galleria mellonella* animal model. Survival analyses of this simplified host-pathogen complex were carried out to assess associations between virulence and genetic makeup, geographical origin (Thai or South African) and source (clinical or environmental).

1.11. Aims of the project

1. To describe the genetic population structure of clinical and environmental *Cng* isolates from Thailand and Cape Town, South Africa, using MLST;
2. To compare these populations to the current global population of *Cng*;
3. To provide a universally available and up-to-date source of *Cng* population data;
4. To investigate whether or not there is an association between *Cng* genotype and clinical phenotype among immunocompromised AIDS patients; and
5. To test potential associations between genotype and virulence using an animal host model.

2. Methodology and data analysis

2.1. Molecular sequencing of *Cn*

2.1.1. Yeast cultivation and DNA extraction

All cryptococcal strains used in this project are listed in Appendix B.1. Isolates were cultured on pre-prepared malt extract agar (CM0059, Oxoid, Basingstoke, UK) and stored in an 18°C incubator. Five days later, DNA was extracted from these cultures using the DNeasy Blood and Tissue Kit (Qiagen, Crawley, UK) according to the following protocol:

180 μL of buffer ATL and 20 μL Proteinase K were added to beads into which culture was looped. The reagents and culture in each tube were homogenized by vortexing in a bead beater for one minute at full speed twice, centrifuged at 13 000 rpm for two minutes and then incubated at 56°C for between one and three hours; 200 μL of ethanol (96-100%) was then added to each tube and mixed by vortexing. The mixture (precipitate included) of each tube was pipetted into a DNeasy mini spin column and placed in a 2 mL collection tube. These were centrifuged at 8 000 rpm for one minute and the flow-through and collection tube discarded. The DNeasy mini spin columns were then placed in new 2 mL collection tubes and 500 μL of Buffer AW1 was added. They were centrifuged at 14 000 rpm for three minutes and the flow through and collection tubes discarded. This was repeated with 500 μL of Buffer AW2. Finally, the mini spin column was placed in a new 1.5 or 2 mL microcentrifuge tube and 100 μL Buffer AE was added to elute. The product was carefully labeled and stored at 4°C awaiting PCR.

Samples of the cultures were subsequently cryopreserved for storage. They were frozen down using the following protocol: YPD consisting of 2.5 g Bacto yeast, 5 g peptone, 5 g dextrose and 250 mL dH_2O was autoclaved

and 15% glycerol added. 1 mL was aliquoted in NUNC cryotubes and cryptococcal cultures looped in. These were then stored at −80°C.

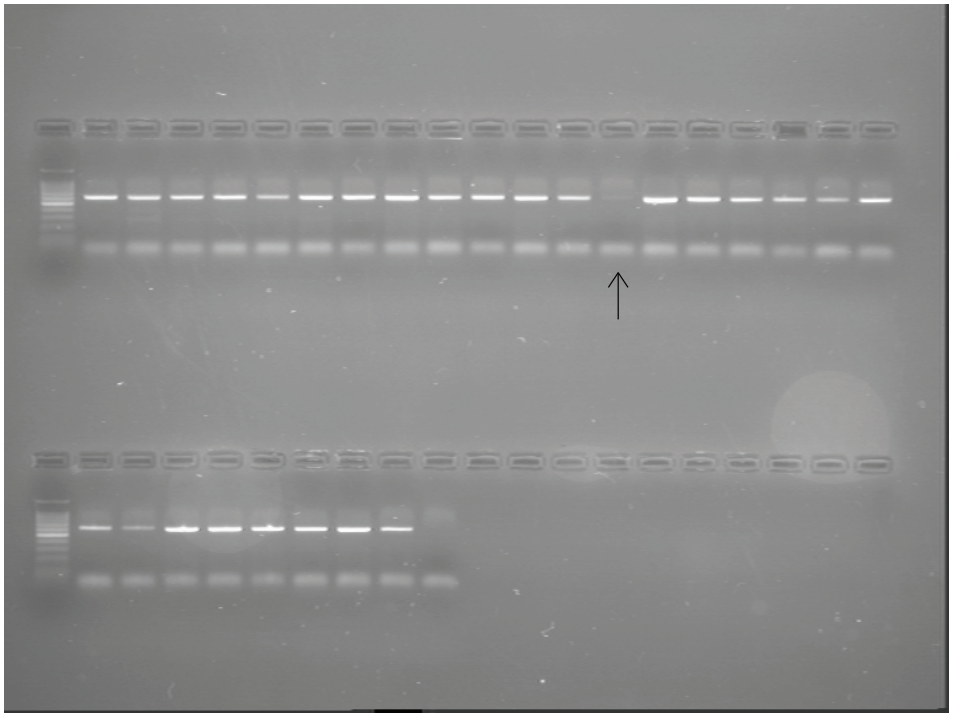
2.1.2. Polymerase Chain Reaction amplification (PCR)

In preparation for amplification by PCR, DNA extractions from clinical samples were diluted fivefold in molecular-grade sterile distilled water, and the environmental samples to 1 in 25. Each isolate was PCR-amplified in 50 μL reaction volumes for each of the seven MLST loci using the primers and protocols detailed by Meyer et al. (2009; Appendix A.1). The reagents were carefully handled and kept on ice during the preparation of the master mix for each locus consisting of 5 μL forward primer (10 μM dilution), 5 μL reverse primer (10 μM dilution), 5 μL Qiagen buffer (10X), 0.5 μL dNTPs (10 mM), 0.5 μL Qiagen Taq polymerase and 32 μL sterile distilled water. 48 μL of master mix and 2 μL of diluted DNA were added to each well on a 48-well microtitre PCR plate. A negative control free of DNA was always included. The plate was then sealed, vortexed and centrifuged at 1 000 rpm for one minute. It was then placed in the thermocycler DNA engine with the appropriate PCR condition for each locus (Meyer et al., 2009; Appendix A.1), and stored at −20°C ready for clean-up. Successful amplification at the appropriate loci was verified by electrophoresis on agarose gels, with a suitable molecular weight marker and positive controls (Figure 2.1).

2.1.3. PCR cleanup

60 μL of 20% PEG/2.5 M NaCl was added to each microtitre plate well. The plate was then resealed, vortexed, spun at 10 000 rpm for ten seconds then incubated at room temperature for 30 minutes or at 4°C overnight. The PCR products were then pelleted by centrifuging the plate at 3 200 rpm at 4°C for one hour. The seal was then removed and the supernatant discarded by inverting the plate onto folded paper towels, wrapping it in cling film and centrifuging it for 20 seconds at 500 rpm. The film was then removed and the plates inverted onto fresh paper towels and centrifuged once more for one minute at 500 rpm, removing any remaining PEG/NaCl.

Figure 2.1.: Verification of successful amplification of isolates at *CAP59*. Successfully amplified samples align with 560 base pairs on the ladder. The last well contained the control which is negative for contamination. Isolate CM30 failed to amplify (indicated by the arrow).



The DNA pellets were washed in 150 μL of 70% ethanol and centrifuged at 3 000–3 200 rpm for 15-20 minutes after the plate was resealed. The supernatant was again discarded by inverted centrifugation on paper towels for one minute at 500 rpm. This wash was repeated for a second time, and the supernatant discarded once more. A thermocycler was then used to dry the pellets at 37°C for two minutes with the plates being left unsealed and the thermocycler lid open to allow the remaining ethanol to dry. 80 to 100 μL of sterile H_20 was added to each well of clinical isolates, and 120 μL to wells of environmental isolates. The plate was then sealed, vortexed and spun at 1 000 rpm for seven seconds in order to resuspend the pellet. After ten minutes in the refrigerator, a second and final vortex and spin ensured that the clean PCR product was at the bottom of the well and prevented splashing and contamination at the next step.

2.1.4. Sequencing

Separate master mixes of forward and reverse primers and TaqFS (Big Dye) were set up for the sequencing reaction at each locus: 0.5 μL Tag-FS (Big Dye V1.1), 1.75 μL sequencing buffer, 1.75 μL water, 4 μL forward or reverse primer (1 pmol/μL). 2 μL of purified PCR product was added to 3 μL of the sequencing master mix in each well of a 96-well microtitre plate, each half consisting of one of the two primers. Again, this work was carried out on ice. The plate was sealed, vortexed and centrifuged at 1 000 rpm for ten seconds to mix the reagents and settle the liquid at the bottom of the well. The sealed plate was then placed in the thermocycler for the following sequencing program:

Control method: calculated

1. 96°C for 0:10
2. 50°C for 0:05
3. 60°C for 2:00
4. go to 1, 24 times
5. 0.10 /s to 4°C
6. 4°C for ever

2.1.5. Precipitation of sequencing cycling products

15 μL of 3M NaAc pH 5.2 (4:11) and 50 μL of 95% ethanol were added to each well. The plate was then sealed, vortexed, spun at 1 000 rpm for ten seconds and incubated at -20°C for an hour. The sequencing products were then pelleted by centrifuging the plate at 3 500 rpm for an hour at 4°C. Immediately after centrifuging was complete, the seal was removed and the plate inverted onto a paper towel to discard supernatant. Particular care was applied when handling the plates at this point as it was crucial not to disturb the pellets. The inverted plate was then transferred to a fresh paper towel and spun at 500 rpm for one minute in order to remove any residual ethanol from the wells. The pellets were then washed in 150 μL of 70% ethanol, and the plates sealed and spun at 3 500 rpm for 30 minutes. The supernatant was discarded, first by inversion onto a paper towel, then by inverted spinning at 500 rpm for one minute on a fresh paper towel.

The ethanol wash was repeated and the second spin conducted at 700–800 rpm in order to dry the pellets. The plates were finally sealed and stored at −20°C until ready for loading. An Applied Biosystems 3730xl DNA Analyzer (Warrington, UK) was used to determine the forward and reverse DNA sequences of all PCR products.

2.2. MLST determination

Sequences were manually edited using CodonCode Aligner (CodonCode Corporation, MA, USA), then aligned in MEGA 4.0 (Tamura et al., 2007). Alleles at each locus were aligned and assigned numbers (Allele Types; ATs) upon comparison with those identified in the global collection (Litvintseva et al., 2006). These were inserted into the database without gaps and with base pair length detailed. Alleles with insertions or deletions (indels) which were not of standard base pair length were tagged with an 'N', and a separate lookup table was created to allow for database queries to determine whether entered sequences were novel or previously described. Each isolate was therefore assigned a 7-digit allelic profile. Each unique allelic profile was concatenated and assigned a Sequence Type (ST) according to the MLST scheme (`http://cneoformans.mlst.net/`). Novel STs identified within the Thai population were assigned as additional STs within the global MLST database.

2.3. Mating-type and serotype analyses

The mating type of each of the isolates was determined by four different PCR amplification reactions. Primers specific to the MATα or MAT**a** allele of the *STE20* locus for either serotype A or D isolates were used: primers JOHE7270 and JOHE7272 (**a**A), JOHE7273/JOHE7275 (**a**D), JOHE7264/ JOHE7265 (αA) and JOHE7267/JOHE7268 (αD; Table 2.1; Barreto de Oliveira et al., 2004; Halliday et al., 1999; Halliday et al., 2003). PCR amplifications with a total volume of 25 μL contained 0.25 μL of 10 mM stock dNTPs, 0.25 μL Taq polymerase, 2.5 μL of buffer, 16.0 μL of sterilized distilled H_20, 1 μL of template DNA and 2.5 μL of each forward and reverse primer at a 10 μM final concentration.

Table 2.1.: Primer sequences used to determine the mating-type of cryptococcal isolates.

Mating type	Gene/Allele	Sequence ($5' - 3'$)
Cn var *grubii*[a]		
MATα	JOHE7264 (F)	CCAAAAGCTGATGCTGTGGA
	JOHE7265 (R)	AGGACATCTATAGCAGAT
MAT**a**	JOHE7270 (F)	ATCAGAGACAGAGGAGGAGCAAGAC
	JOHE7272 (R)	TCCACTGGCAACCCTGCGAG
Cn var *neoformans*[a]		
MATα	JOHE7267 (F)	ATAGGCTGGTGCTGTGAATTAAG
	JOHE7268 (R)	GTTCAAGTAATCTCACTACATGCG
MAT**a**	JOHE7273 (F)	GTTCATCAGATACAGAGGAGTGG
	JOHE7275 (R)	CTCCACTGTCAAACCTACGGC
Cryptococcus gattii		
MATα[b]	MFaU (F)	TTCACTGCCATCTTCACCACC
	MFaL (R)	TCTAGGCGATGACACAAAGGG
MAT**a**[c]	STE20aSF U (F)	TCCGATTGCTGCGATTTGCC
	STE20aSF L (R)	GCGCCTGCACCATAATTCACC

[a] Barreto de Oliveira et al., 2004; [b] Halliday and Carter, 1999; [c] Halliday and Carter, 2003 F = forward primer, R = reverse primer

2.4. Molecular data analyses

2.4.1. Diversity indices

Genetic variability

Genetic variability within individual populations was assessed through comparative sequence analyses performed in DnaSPv5 (Librado and Rozas, 2009). Four statistical measures were applied to the samples of *Cng*, both at individual loci and concatenated sequences:

1. The number of segregating sites (S; Watterson, 1975). This statistic is equal to the number of nucleotide positions at which polymorphism is found. When the polymorphisms observed are selectively neutral,

it provides us with the best estimate of nucleotide diversity as it has the lowest variance (Tajima, 1983).

2. Haplotypic diversity (*Hd*; Nei, 1987). This measures the uniqueness of a particular haplotype in a given population and is based on the presence of haplotypes (h). A haplotype is a set of nucleotide polymorphisms, also known as SNPs, located on a a single chromosome of a statistically associated chromosome pair. These values define "gene diversity", that is to say they represent the probability of picking two sequences at random from non-identical haplotypes in a sample set.

3. Nucleotide diversity (π; Nei, 1987). This estimates the average number of nucleotide differences per site between pairs of sequences within populations. An estimate of mean nucleotide heterozygosity, π gives information on the extent of DNA difference between two randomly chosen genes and — unlike S — takes into account the frequency of mutants, so that deleterious mutants low in frequency do not affect the estimate (Tajima, 1983; Nei and Kumar, 2000).

4. Population-scaled mutation rate estimated per site (θ; Watterson, 1975). This statistic is based on the expected number of segregating sites (S). Watterson's estimator assumes that each mutation occurs at a previously unmutated site and that recombination does not occur between sites. It is therefore an unbiased method-of-moments estimator regardless of the recombination rate (RoyChoudhury and Wakeley, 2010).

Tests of neutality

Tajima's D (1989) is a statistical test based on the neutral model of evolution and is used to assess whether the data are consistent with the population being at mutation-drift equilibrium. This estimate represents the correlation between the statistics S and π, and so in the case of neutral variation when a population is at a drift-mutation equilibrium, Tajima's D is zero. A positive estimate is due to an excess of rare polymorphisms and positive selection. This is the result of either a recent bottleneck in the population, or the presence of overdominant selection at the locus. A negative D value, on the other hand, is indicative of balancing selection resulting in

an excess of high-frequency variants. Processes which lead to this are purifying selection or recent population expansion, with both scenarios being characterised by an increase in mutations (S) while heterozygosity remains low (i.e. a low value of π; Oleksyk et al., 2010; Holsinger, 2010). As demographic changes affect all loci differently — especially when, as in this case, the loci have different physiological functions — this test is applied to the concatenated seven loci. Significance values were calculated by 10 000 coalescent simulations.

Ramos-Onsins & Rozas' R_2 (2006) is also a test for departure from the neutral model of molecular evolution. It is considered to be among the best statistical tests for detecting population growth, and is the recommended approach for small sample sizes and populations in which recombination is suspected to be occurring (Ramirez-Soriano et al., 2008; Ramos-Onsins and Rozas, 2002). Based on the distribution of mutation frequencies, this test can be used to detect recent severe population growth and, under this demographic scenario, values of R_2 are expected to be low. As with Tajima's D, significance was determined from 10 000 coalescent simulations.

2.4.2. Analysis of genetic structure based on allelic profiles

A hierarchical Analysis of Molecular Variance (AMOVA) was performed in GenAlEx 6.1 for Microsoft Excel (Peakall and Smouse, 2006) in order to examine the distribution of genetic variation, and to determine the extent of connectivity among populations and regions based on allelic profiles (Excoffier et al., 1992). AMOVA is a distance-based statistical technique that estimates the extent of genetic differentiation between individuals and populations directly from molecular data. The technique treats the raw molecular data as a pairwise matrix of genetic distances between all the possible combinations of *Cng* isolates, with sub-matrices corresponding to the different hierarchical data partitions (here, the genetic differences between *Cng* infecting different host individuals and geographical regions). These data are then analyzed within a nested analysis of variance (ANOVA) framework in which mean squares are computed for groupings at all levels of the hierarchy. Unlike in a simple ANOVA, this enables hypothesis tests of between- and within-group differences at several hierarchical levels (Excoffier et al., 1992). Under the null hypothesis of no genetic difference among the pop-

ulations, subpopulations can be considered part of a single large random mating genetic population. In subpopulation groups for which this is true, there is little difference between the arbitrary subpopulations (other than minor sampling effects), and randomization of the dataset samples will yield a good estimate of the expected value. In a randomly mating population, on the other hand, multiple shuffles of the samples in the dataset will result in values close to that expected by chance in a randomly mating population (Peakall and Smouse, 2006). An F-statistic analogue of the genetic differentiation among populations, ΦPT, and between pairs of groups (population pairwise ΦPT) is also reported (Excoffier et al., 1992), with significance estimated from 999 random permutations.

PCA

Spatial patterns of allelic variability among the MLST genotypes of the Thai isolates typed in this study were also explored by Principal Component Analysis (PCA). This multivariate method summarises genetic variability but does not make assumptions about an evolutionary model. This feature is particularly of use when little is known of the system being studied, as is often the case in landscape genetics (Jombart, 2008; Manel et al., 2003). PCA was performed using the Adegenet 1.1 package for statistical software R (version 2.6.1). This package is dedicated to the multivariate analysis of genetic markers, and is used to find and plot major patterns, thus illustrating population stratification within a set of genotypes (Jombart, 2008). This is done by orthogonal transformation of possibly correlated observations to uncorrelated ones along decreasingly variable axes. Diagrams obtained by PCA consist of dots, representing individual genotypes, clustered into groups. Isolates belonging to the same group are linked by matching colored lines, labeled and summarized by 95% ellipses. The major axes of variation are located within a multidimensional dataset and each successive axis explains proportionately less of the total variation. Therefore, when distinct groups are present, the first two or three axes reveal most of the separation among them (Peakall and Smouse, 2006). The contributions of the principal coordinates to the genetic structure of the population is represented by bar plots of eigenvalues. The multivariate tool PCA will enable us to unravel the genetic structuring among the seven MLST markers and the

spatial patterns of the genetic variability within the *Cng* population (Lalo et al., 2007).

2.4.3. Linkage disequilibrium and recombination

Evidence of linkage disequilibrium amongst the MLST loci was tested using two measures of index of association, I_A and $\bar{r}_d$ (Burt et al., 1996; Agapow and Burt, 2001; Brown et al., 1980; Smith et al., 1993). The significance of the pairwise statistics returned was determined by 1 000 randomizations. In the instance of significant clonality or population substructure, both values are expected to be greater than zero, while freely recombining populations would return a score of zero. These tests were also performed on clone-corrected samples, as recombination may sometimes be masked by an excess of clonally derived individuals (Litvintseva et al., 2003; Casadevall and Perfect, 1998; Smith et al., 1993; Xu and Mitchell, 2002). The proportion of phylogenetically compatible pairs of loci is also reported, with significance estimated with 1000 randomizations (Estabrook, 1975; Xu et al., 2009).

The minimum number of recombination events (R_m) was estimated both within an individual locus and between loci (Hudson and Kaplan, 1985). This summary statistic allows us to deduce that recombination has taken place between a pair of segregating sites when all four types of gametes are represented, and is also known as the 4-gamete test. R_m is a maximum-likelihood estimate, computed under the assumption that each new mutation takes place at a previously unmutated site (Wall, 2000). This statistic is of importance in the context of disease, as recombination breaks down genetic markers while genetic drift builds them up. These genetic markers and their association with disease phenotype are the basis for the mapping of disease-associated mutations (Pritchard and Przeworski, 2001).

2.4.4. Genetic differentiation between populations

The average pairwise number of nucleotide differences per site, D_{xy}, was used to estimate divergence among population groups (Nei, 1987). The inter-population measure is defined as the average number of differences between one sequence randomly chosen from one population and another sequence randomly chosen from another population (Takahata and Nei, 1979).

K^*_{ST} is a weighted measure of the ratio of the average pairwise differences within populations to the total average pairwise differences, and was used to determine population differentiation for all pairwise comparisons of the population subgroups (Hudson et al., 1992). This measure has been shown to have the highest power in detecting population differentiation under simple models of population structure (Hudson, 2000). Significance levels for this statistic were determined using 1 000 permutations (Hudson et al., 1992).

A third statistics used to assess differentiation between populations was S_{nn}. This is the proportion of nearest neighbours in sequence space found in the same population (Hudson and Kaplan, 1985; Hudson, 2000). All three statistics (D_{xy}, K^*_{ST} and S_{nn}) were calculated in DNASPv5, with significance levels assessed by 1 000 permutations.

2.5. Phylogenetic analyses and molecular type determination

Phylogenetic neighbor-joining trees were estimated for each aligned locus as well as the aligned concatenated sequences of the seven loci from all *Cng* samples. These imported alignments were unedited, meaning that all insertions and deletions were maintained. Evolutionary distances were computed using the maximum composite likelihood method in MEGA 4.0 (Tamura et al., 2007; Saitou and Nei, 1987). The percentage of replicate trees in which the associated taxa clustered together was estimated by the bootstrap test, inferred from 1 000 replicates (Felsenstein, 1985). Molecular VN groupings of the Thai isolates were inferred through phylogenetic and comparative analyses with reference strains of known major molecular types of the *C. neoformans*/*C. gattii* species complex: WM148 (serotype A, VNI), WM626 (serotype A, VNII), WM629 (serotype D, VNIV), WM179 (serotype B, VGI), WM178 (serotype B, VGII), WM175 (serotype B, VGIII), WM779 (serotype C, VGIV; Meyer et al., 2003) and the genome-project strain H99 (serotype A, VNI; Perfect et al., 1993).

2.6. Estimates of times of divergence and haplotype networks

A Bayesian Markov chain Monte Carlo (MCMC) method, implemented in the program BEAST version 1.5.3 (Drummond, 2007), was used to estimate the time of divergence between the geographically defined populations of the global sample of *Cng*, defined as the time to the most recent common ancestor (TMRCA). Sequence indels greater than a single nucleotide long were treated as single evolutionary events in the dataset, and a second partition reflecting these indels created in Beauti v1.5.3. The Hasegawa-Kishino-Yano (HKY) model of sequence evolution was assumed, and both a relaxed, uncorrelated lognormal, and a strict molecular clock model applied, depending on the presence of a fixed mutation rate. Simulations were run for 10^6 iterations with an initial burn-in of 10%. Parameters were logged every 1 000 steps over the course of the run. A neutral mutation rate of 2×10^{-9} per nucleotide per year was initially assumed, in accordance with the consensus mutation rate for protein coding genes (Li et al. 1987; Nei, 1987; Koufopanou et al. 1997). Following several runs and optimisations, this was finally changed to a rate of 4×10^{-9} substitutions per nucleotide per year. Credibility intervals were obtained using 95% highest posterior density (HPD) intervals, the shortest segment that includes 95% of the probability density of the parameter, and the effective sample sizes (ESS) for each parameter, depicted using Tracer v1.5.

Haplotype networks were also created for the STs of the global *Cng* population at each MLST locus. The inference of phylogenetic relationships among them using statistical parsimony was performed using the program TCS v1.21 (Clement et al., 2000).

2.7. MLST website eBURST tool

eBURST, a program available at `http://eburst.mlst.net/`, infers patterns of evolutionary descent among clusters of related genotypes from MLST data. eBURST utilises the MLST site's geographical mapping of MLST datasets to subdivide the STs into related groups of or clonal complexes, as well as to identify the founding genotype (ST) of each group (Feil

et al., 2004).

2.8. Clinical data and analysis

Clinical data indicative of the progression of cryptococcal infection were available for 165 Thai and Cape Town isolates typed in this study. We investigated potential associations between ST and baseline continuous variables using both ANOVA and multivariate ANOVA (MANOVA), with Fisher's exact test being applied to categorical variables. Logistic regression was used to determine factors associated with the clearance of *Cng* as well as death by 10 weeks. Survival analysis was determined using the Kaplan-Meier estimator (Kaplan and Meier, 1958) and the difference between groups detected by the logrank test. Survival analyses were also carried out using the Kaplan-Meier estimator and the logrank test applied to assess whether the differences in survival between groups, treatments, etc., were more than that expected by chance alone. All analyses were performed using the statistical software package R (version 2.6.1) and the code is available in Appendix C, as detailed within the text.

2.9. Environmental fieldwork

2.9.1. Sampling of *Cn*

From November 2009 to December 2009, a total of 295 samples of soil, bird droppings and feathers were collected from seven locations in Cape Town, South Africa. Swab samples were taken from tree hollows, preferably sheltered from the sun, while soil samples were collected in conical tubes. Avian excreta was sampled using both methods, as appropriate. The swabs used were sterile Sterlin Transport Swabs (VWR International) which contained media for the maintenance of the fungus during transport to the laboratory.

2.9.2. Preparation of niger seed agar

A solution of 140 g of pulverized niger seeds (*Guizotia abyssinica*) and 700 mL of distilled water was autoclaved for 15 minutes and filtered through

a double layer of cheesecloth. 200 mL of the niger seed extract was then added to 1 g of glucose, 20 g of agar and 800 mL of dilute water and autoclaved for 30 minutes, creating niger seed agar (NS). Following autoclaving, 200 mg of chloramphenicol (200 mg of 1 mL of stock solution: 200 mg/mL in 95% ethyl alcohol), 25 mg of gentamicin (1 mL of stock solution: 25 mg/mL in water) and 100mg of biphenyl (Diphenyl; 10 mL of stock solution: 10 mg/mL of 95% of ethyl alcohol) were added to the NS agar. The two agars, NS with and without antibiotics and biphenyl, were dispensed onto plates.

2.9.3. Isolation and identification

The tubes containing the faeces or soil were diluted in 10–25mL of sterile water, vortexed for two minutes, and the samples left to settle for five minutes at room temperature. A second dilution of 1:10 was made for each sample using dilute water. Next, 20–50 μL of the supernatant was inoculated onto NS plates supplemented with antibiotics and biphenyl, prepared as described above. Each sample was therefore spread onto two plates, one for each dilution, incubated at 30°C and examined daily for up to five days to identify smooth, beige to dark brown colonies suggestive of *Cryptococcus spp.* (Randhawa et al., 2005). Samples collected on swabs were also plated onto the surface of two NS + antibiotics and biphenyl petri dishes each, incubated at 30°C and examined for five days after inoculation for all yeasts resembling *Cn* (creamy-white to yellow-brown colonies). All such colonies were then sub-cultured onto NS media without antibiotics and biphenyl for 48 hours at 30°C by streaking for isolation, and identified by PCR and sequencing analysis at the *ITS1* gene. The primers used were *ITS1F* (Gardes and Bruns, 1993) and *ITS4* (White et al., 1990), both of which have specificity for ascomycetous, basidiomycetous, and zygomycetous fungi (Manter and Vivanco, 2007; Gardes and Bruns, 1993; Dickie et al., 2002; Klamer et al., 2002; Anderson et al., 2003a; Anderson et al., 2003b; Izzo et al., 2005). The resulting sequences were processed in Basic Local Alignment Search Tool (BLAST), a tool which can be used to find regions of similarity between biological sequences – achieved by comparing nucleotide or protein sequences to sequence databases, identifying members of gene families at `http://blast.ncbi.nlm.nih.gov/Blast.cgi` — and isolates of the *Cryp-*

tococcus species identified. These were then stored in distilled water at 4°C for genetic analysis according to MLST as described in earlier methods.

All species of *Cn* identified by NCBI BLAST were also assessed for relatedness among themselves and with reference strains of the *C. neoformans*/*C. gattii* species complex using phylogenetic neighbor-joining trees created in MEGA 4.0, as described in subsection 2.5.

2.10. *Galleria mellonella* wax moth model

Sixth-instar larvae of the wax moth *Galleria mellonella* were obtained from the Meal Worm Company (Sheffield, England). Prior to the experiment, larvae were stored in wood shavings in the dark at 15°C and used within five days of delivery. Caterpillars (330 ± 25 mg in body weight) were employed in the assay, with ten randomly chosen caterpillars of the required weight used per group.

Isolates pertaining to seven STs of *Cng* were used. One was unique to Thailand (ST4), three unique to Cape Town (STs 1, 23 and 53) and three STs shared between the two (STs 5, 6 and 32). Isolates of each were cultured on pre-prepared malt extract agar (CM0059, Oxoid, Basingstoke, UK) and incubated at 25°C. Three days later, the most viable were selected and cultured in liquid media. 10 g of Bacto yeast, 20 g of peptone, and 900 mL of H_20 are autoclaved separately from a sterilized solution of 20 g of glucose and 100 mL of H_20, the two solutions combined in a 9:1 ratio, stored at 28°C, and sub-cultured over three nights (with shaking). From these, ten dilutions of 10^4 cryptococcal cells suspended in PBS were made.

Using a 10 μL Hamilton syringe, 10 μL aliquots of the inoculum and antibacterial were injected into the hemocoel of each caterpillar via the last left proleg, which had been cleaned using an alcohol swab. In addition, 10 μL of 20 mg of ampicillin/kg of body weight was injected into the last right proleg in order to counter bacterial infection (Mylonakis et al., 2005). Control treatments were also included to ensure that neither the injection procedure nor the incubation period were responsible for any mortality observed. These controls involved *G. mellonella* larvae injected with 20 μL of PBS and a group to which nothing was done. *G. mellonella* larvae were placed in labeled Petri dishes and incubated in the dark at 30°C; a temperature which

was previously identified as being optimum for culturing larvae infected with fungal pathogens (Brennan et al., 2002; Cotter et al., 2000). Mortality rates were recorded for 21 days post-injection. Mortality was assessed based on lack of movement in response to stimulation and discolouration (melanisation) of the cuticle. Survival probabilities were estimated by the Kaplan-Meier method and differences between groups determined by the logrank test. All analyses were carried out in R and RStudio, and p-values of smaller than 0.05 were considered significant.

3. Multilocus Sequence Typing

Acknowledgements: The work presented in this chapter was made possible through the collaborative efforts of the Cryptococcal Working Group of the International Society for Human and Animal Mycology (ISHAM). Our work was published and is available in Appendix A.1.

Meyer, W., Aanensen, D. M., Boekhout, T., Cogliati, M., Diaz, M. R., Esposto, M. C., Fisher, M., Gilgado, F., Hagen, F., Kaocharoen, S., Litvintseva, A. P., Mitchell, T. G., Simwami, S. P., Trilles, L., Viviani, M. A. and Kwon-Chung, J. (2009) 'Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*', *Medical Mycology,* 47(6), 561-570.

3.1. Introduction

There has been an increase in the incidence of invasive fungal infections in recent years (Perfect and Casadevall, 2006). This is especially true among patients of immunosuppressed status, including patients on immunosuppressive therapy, receiving organ transplantation, and those with AIDS (Cuenca-Estrella et al., 2008). The high morbidity and mortality rates associated with these infections has fueled considerable growth in the development and application of rapid and accurate molecular methods in the analysis of the evolution of fungal species and their populations, *Cn* included (Xu, 2006).

Molecular epidemiology constitutes an important methodology for understanding the factors underpinning the emergence of pathogenic fungi. An accurate description of the temporal and spatial genetic composition of fungal pathogen populations is important from several standpoints: quantifying the amount and distribution of polymorphisms across space and time enables the identification of population-level processes that ultimately lead to an understanding of the process of infection, such as the reservoirs, transmissibility and longevity of populations and their component genotypes.

Increasingly, it is being recognised that specific genotypes act as markers of lineages that exhibit enhanced or reduced virulence (Illnait-Zaragozi et al., 2010; Kidd et al., 2004; Litvintseva et al., 2005; Byrnes et al., 2010), as evidenced by the invasion and expansion of two genotypes of *C. gattii* in the Pacific Northwest (Fraser et al., 2005). Therefore, an accurate understanding of the genetics of *Cn* will clarify its current and future evolutionary trajectories, and its potential to alter the burden of human disease.

Advances in molecular techniques have led to significant progress in our addressing fundamental questions about the evolution and population genetics of *Cn*; however, the variability of these techniques has resulted in a lack of uniformity in reported findings, making cross-referencing between research groups almost impossible. In order for us to assemble informative data on pathogenic fungi and their spread, collaborative efforts by the *Cn* research community to integrate genotyping approaches and results of spatial collections of isolates are crucial.

3.2. History of the identification of fungi

Traditionally, the recognition of fungal species relied on a combination of microscopy and culture-based techniques. Due to the understanding that differences in phenotype reflect genetic differences between strains, these techniques were based on morphological, biochemical and immunological criteria, as well as the presence and nature of reproductive structures (Gil-Lamaignere et al., 2003; Soll, 2000). Analyses of morphological and physiological traits, such as the standard laboratory method for the identification of *Cn* by Indian ink preparations of cerebrospinal fluid (CSF) or blood samples, were often laborious, time consuming and produced subjective results which were difficult to interpret (Wengenack and Binnicker, 2009; Mitchell et al., 1994). In addition to low discriminatory power and reproducibility, there was evidence of intrinsic limitations of these phenotyping procedures and their potential to hinder final diagnoses (Huston and Mody, 2009), such as phenotypic switching in fungi *Cryptococcus neoformans*, *Candida albicans*, *tropicalis* and *glabrata* (Goldman et al., 1998; Soll et al., 1988; Lachke et al., 2000). A second example is the discovery that antigen expression is not serotype-specific in *Candida albicans* (Poulain et al., 1985). There

was therefore an urgent need for rapid, accurate, simple and reproducible methods of fungal pathogen identification (Mitchell et al., 1994).

Nucleic acid hybridization techniques were integral in the development of molecular methods for several fungal pathogens (Aulakhet et al., 1981; Sambrook et al., 1989). DNA-fingerprinting techniques were developed around 1975 and enabled researchers to assess genetic relatedness of isolates within species at the genotypic level (Gil-Lamaignere et al., 2003). The most common of these is Restriction Fragment Length Polymorphism (RFLP), first without then with hybridization. The former was first successfully used to discriminate isolates of the *Candida spp.* (Magee et al., 1987), then expanded to other infectious fungi such as *Histoplasma capsulatum* (Vincent et al., 1986), *A. fumigates* (Burnie et al., 1992; Lin et al., 1995), and *Cryptococcus neoformans* (Currie et al., 1994). Relatively simple and fast, this procedure involves the visualisation of bands of digested, high molecular-weight DNA by electrophoresis in an agarose gel. These band patterns are created by repetitive DNA with greater variability resulting in greater discriminatory power. Limitations arise when DNA quality is low and when isolates are moderately related (Kwon-Chung and Wickes, 2006).

The development of DNA probes meant that visualization of restriction fragments could be limited to produce more highly resolved fingerprint patterns of, for example, a single gene. A major advantage of this technique in fungal epidemiology is the fact that the outputs are qualitative and one can easily determine the genetic relatedness of strains by visual inspection. Several effective probes used in fingerprinting by Southern blot hybridization include Ca3, 27A, and CARE-1 and CARE-2 for *C. albicans* (Lasker et al., 1991; Scherer and Stevens, 1988; Lockhart et al., 1995; Pujol et al., 1997; Mercure et al., 1993; Lasker et al., 1992; Muller et al., 1999), Afut, 3.11, 3.19, and M13 for *A. Fumigates* (Girardin et al., 1993; Anderson et al., 1996; Bart-Delabesse et al., 2001; Neuveglise et al., 1996), and Cntel and UT-4p for *Cn* (Spitzer and Spitzer, 1994; Spitzer and Spitzer, 1992).

A second method central to DNA-based methods of identification is the polymerase chain reaction (PCR). PCR arose in the late 1980s and is a technique by which data are obtained directly from specific regions of a sample's DNA through the amplification of a small, specific segment of the genome (Saiki et al., 1986; Sambrook et al., 1989). This fast, easily executable pro-

cess is not only cheaper than hybridisation using oligonucleotide probes, but can be applied to contaminated samples and/or mixed cultures (Sidrim et al., 2010), making it the main molecular technique used in the study of the etiological agents of cryptococcosis (Kang et al., 2009; Wengenack and Binnicker, 2009; Sidrim et al., 2010). Furthermore, its ability to be used in conjunction with other techniques makes it invaluable in advancing molecular and epidemiological studies of fungi. Based on PCR amplification, the use of random amplified polymorphic DNA (RAPD) fingerprinting is the easiest and most convenient tool used in the analysis of population genetic structure, evolution and epidemiology (Welsh and McClelland, 1991; Lehmann et al., 1992; Williams et al., 1990). This technique identifies polymorphisms detected as strain-specific DNA fragments generated by low stringency Taq (PCR) amplifications with single primers of arbitrary sequence. Advantages of RAPD are that (i) no previous sequence information is required for the organism in question, making all sorts of organisms accessible; (ii) the number of genetic markers used can be increased as required, increasing the number of detectable polymorphisms; and (iii) it is cost effective, allowing for the investigation of many individuals. Having said this, RAPD studies are largely unpublishable due to poor reproducibility between and within laboratories (Kwon-Chung and Wickes, 2006; Soll, 2000).

PCR techniques further diversified and have been implicated in various aspects of genomic variation at the individual level, as well as in the context of ecological and demographic events of populations (White, 1996). Nested PCR involves the use of two sets of primers in two successive PCR runs, resulting in heightened detection sensitivity and specificity (Bialek et al., 2002). Despite being prone to contamination, this process resulted in the successful diagnosis of cryptococcosis in patients suffering from neurocryptococcal and primary lymphonodular abdominal cryptococcosis (Rappelli et al., 1998; Zou et al., 2006), as well as cryptococcosis associated with inflammatory reconstitution syndrome (Putignani et al., 2008). Multiplex PCR is the process by which two or more loci are amplified in a single reaction, and has allowed for differentiation between both clinical and environmental isolates of the same *Cn* species, mating types, and serotypes in humans and animals (Esposto et al., 2004; Leal et al., 2008; Casali et al., 2003; Horta et al., 2002). Disadvantages of multiplex PCRs include the need for great

accuracy in the reagents and protocol in order to avoid erroneous results, as well as issues of inter-primer competition (Henegariu et al., 1997). A third variation of PCR is real-time PCR; this combines standard PCR with fluorescent probe detection, and has transformed the way in which infectious disease pathogens are now identified, *Cn* included (Bankowski and Anderson, 2004; Cockerill, 2003; Espy et al., 2006; Hsu et al., 2003; Levy et al., 2008; Veron et al., 2009). This is the fastest of PCR tests as it combines processing and detection in a single reaction, and is also able to provide quantitative results according to the quantity of nucleic acid present in a sample (Wengenack and Binnicker, 2009; Bialek et al., 2002; Vollmer et al., 2008). Despite being highly sensitive and specific, the modern materials and equipment needed render it expensive (Amjad et al., 2004; Espy et al., 2006).

Initially applied to *S. cervisiae*, the late 1980s saw the use of electrokaryotyping in the chromosomal analysis of *Cn* through the separation of large DNA fragments using alternate electric fields, referred to as pulsed field gel electrophoresis (PFGE; Polacheck and Lebens, 1989; Wickes et al., 1994; Zolan, 1995; Spitzer and Spitzer, 1997). Karyotyping, similar to RAPD, is sensitive to slight changes in the laboratory environment. In the 1990s, DNA-DNA hybridizations and electrokaryotyping were used together in the recognition of epidemiological markers as well as in the detection of genetic diversity of both *Cn* and *C. gattii* (Spitzer and Spitzer, 1992; Polacheck et al., 1992; Varma et al., 1995; Varma and Kwon-Chung, 1992). Another form of phylogenetic species recognition (PSR), involving the use of concordance of multiple gene genealogies in the recognition of species boundaries of *Cn*, was infrared spectroscopy (Dettman et al., 2003; Fischer et al., 2006; Schmalreck and Hotzel, 2000).

The protein-based multilocus enzyme elephoretic (MLEE) typing method outperforms several DNA-based typing methods and has been applied in epidemiological studies of bacterial and fungal diseases and at several levels of resolution (Selander et al., 1987; Lehmann et al., 1989; Soll, 2000), cryptococcosis included (Brandt et al., 1993). Determining relatedness between strains by evaluating isoenzyme or allozyme polymorphisms, MLEE allows us to compare multiple loci of multiple strains and detect microevolution (Soll, 2000). This typing method is highly discriminatory with low prob-

ability of homoplasy in clonal organisms, but is also laborious, subjective, and known to hide genetic variation (Taylor and Fisher, 2003; Enright and Spratt, 1999; Johnson et al., 2007). The concordance between MLEE and RAPD in terms of sensitivity in discriminatory power and reproducibility within related strains of a species was established in pathogens such as *Escherichia coli*, responsible for diarrheal disease worldwide (Wang et al., 1993), and parasitic protozoa (Tibayrenc et al., 1993). Furthermore, Brandt et al. (1995) found RAPD to display greater sensitivity in the subtyping of *Cn* var *neoformans* isolates, similar to findings by Wang et al. (1993).

3.3. Rationale for the standardisation of the typing of *Cn*

The ability to assess genetic relatedness allows us to (i) understand the dynamics of infectious organism in human populations; (ii) identify the origin of infections; (iii) monitor patterns of drug-resistant strains; and (iv) estimate potential threat (Soll, 2000). The evolution of multiple techniques for molecular determination in response to the increase in invasive fungal infections resulted in a broad range of epidemiological and population studies. Unfortunately, the variation in methods applied — both in technique and data output — made the accurate reproduction of the results and comparisons between typed populations impossible, and controversy in terms of differing taxonomical definition of the species ensued.

Currently, the main PCR-based typing methods being applied to discriminating between cryptococcal species and characterising their intraspecic genetic diversity are MLEE, PCR fingerprinting using microsatellite-specific wild-type phage M13 (Meyer et al., 1999; Meyer et al., 2003) or minisatellite-specific primers (e.g., (GACA)4 or (GTG)5; Viviani et al., 1997; Cogliati et al., 2000; Meyer and Mitchell, 1993), and Amplified Fragment Length Polymorphism PCR analysis (AFLP; Boekhout et al., 2001). In 1995, Brandt et al. typed a clinical population of *Cng* using MLEE and identified two distinct clusters which they proceeded to label ET1 and ET2 (Brandt et al., 1995; Brandt et al., 1996). Boekhout et al. (2001) subsequently described two distinct groups within global *Cng* isolates typed by AFLP and labeled them genotypes 1 and 1A. Most recently, PCR fingerprinting in local-scale

studies on patterns of genetic diversity identified two major and genetically isolated molecular types of *Cng*, VNI, VNII and VNB (Litvintseva et al., 2006; Meyer et al., 1999; Meyer et al., 2003). Finally, Litvintseva et al. used AFLP to detect two groups in a Botswanan *Cng* population and in a sample from the United States, described as I and II (Litvintseva et al., 2005). Although apparently identifying the same clades within the *Cng* variety, the absence of a cross-reference consensus means the corresponding major genotypes within *Cng* across various naming systems applied by different laboratories are not directly comparable between studies (Table 3.1). Clearly, there is a need for a rapid, affordable technique which is reproducible and whose results are universally available.

Table 3.1.: Relationships among different molecular groups in *Cn* var *grubii* according to typing method, as described by different authors (adapted from Litvintseva et al., 2006).

MLEE, PCR fingerprinting[a]	AFLP[b]	RFLP, PCR fingerprinting[c]	AFLP sequencing[d]	AFLP[e]	AFLP, MLST[f]
ET1 complex	1	VNI	II	II	VNI group
ET2 complex	1A	VNII	ND	I	VNII group
ND	ND	ND	I	ND	VNB group

[a] Brandt et al., 1995; [b] Boekhout et al., 2001; [c] Meyer et al., 2003; [d] Litvintseva et al., 2003; [e] Litvintseva et al., 2005; [f] Litvintseva et al., 2006; ND = not determined

3.4. ISHAM, our aims

In order to tackle the issues arising due to the need for a standard globally acceptable typing method, the Cryptococcal Working Group was formed under the International Society for Human and Animal Mycology (ISHAM). This group was established in 2006 in order to unite investigators involved in research on the epidemiology and population genetics of *Cryptococcus neoformans*. The working group "Genotyping of *Cryptococcus neoformans* and *C. gattii*" met in Torino, Italy, at the third Trends in Medical Mycology symposium (TIMM3, 2007), and addressed the need for concordance on a gold standard typing method, and nomenclature that can be applied universally to this pathogen. Our aims were

1. To standardise the naming system for the two species of *Cryptococcus neoformans* (*Cn*) and *C. gattii*;

2. To designate a set of standard strains for each of the eight molecular types and make them globally available;

3. To standardise the typing technique for cryptococcal strain typing; and

4. To create a platform from which the data can be accessed globally.

3.4.1. Consensus genotype nomenclature

The working group assessed the concordance of different molecular typing methods used for both *Cn* and *C. gattii* (Table 3.2). In order to determine the best naming system for the cryptococcal species complex, the two main typing systems at the time (PCR fingerprinting using primers specific for microsatellite or minisatellite specific primers, and AFLP) that had been applied to over 2 000 isolates (Meyer et al., 1999; Meyer et al., 2003; Boekhout et al., 2001) were grouped into eight major molecular types and correlated to serotype. For the *Cn* species, *Cng* consists of molecular types VNI = AFLP1, VNII = AFLP1A, while *Cn* var *neoformans*' serotype D consists of molecular type VNIV = AFLP2, and the AD hybrid of VNI = AFLP3 (Table 3.2). The working group therefore decided that the consensus genotype nomenclature would adhere to the VNI-VNIV scheme for *Cn* var *neoformans* and var *grubii*, and VGI-VGIV for *C. gattii* (Meyer et al., 2009), which represents the global population structure of the cryptococcal species (Appendix A.1).

3.4.2. Consensus standard strains

Reference strains were designated by the ISHAM working group for each of the eight molecular types in order to allow for global standardisation (*Cn* and *C. gattii*, Tables 3.3 and 3.4, respectively). These have been made publicly available from the CBS-Fungal Biodiversity Centre (CBS; `http://`

www.cbs.knaw.nl), the American Type Culture Collection (ATCC; http://www.atcc.org) and the Fungal Genetic Stock Centre (FGS; http://www.fgsc.net), and include strains and cultures used in major cryptococcal genome projects such as H99 and B-3501.

Table 3.2.: Concordance of different molecular typing methods used for *Cryptococcus neoformans* and *Cryptococcus gattii*. Adapted from Meyer et al., 2009.

		PCR-fingerprinting molecular type		AFLP genotype		*URA5* RFLP	*PLB1* RFLP	*IGS*	ITS
Subspecies	Serotype	Meyer[a],[b]	Viviani[c]	Boekhout[d]	Litvintseva[e]	type[b]	type[f]	genotype[g],[h]	genotype[i]
Cn var *grubii*	A	VNI	VN6 (VN5)	AFLP1	VNI	VNI	A1	1A, 1B	ITS1
		VNII		AFLP1A, AFLP1B	VNB	VNII		1A	ITS1
		VNII	VN7	AFLP1A, AFLP1B	VNII	VNII	A2	1C	ITS1
AD hybrid	AD	VNIII	VN3, VN4	AFLP3		VNIII	A3	2C	ITS1, ITS2
Cn var *neoformans*	D	VNIV	VN1 (VN2)	AFLP2		VNIV	A4	2A, 2B, 2C	ITS2
C. gattii	B/C	VGI		AFLP4A, AFLP4B		VGI	A5	4	ITS3, ITS7
		VGII		AFLP6		VGII	A6	3	ITS4
		VGIII		AFLP5A, AFLP5B, AFLP5C		VGIII	A7	5	ITS5
		VGIV		AFLP7		VGIV	A8	6	ITS6

[a]Meyer et al., 1999; [b] Meyer et al., 2003; [c] Viviani et al., 1997; [d] Boekhout et al., 2001; [e] Litvintseva et al., 2006; [f] Latouche et al., 2003; [g] Diaz et al., 2000; [h] Diaz et al., 2005;[i]Katsu et al., 2004.

Table 3.3.: Standard reference strains for *Cryptococcus neoformans* typing, determined by the ISHAM Cryptococcal Working Group.

CBS ID	Alternate ID	MAT & Serotype	Year	Source	References
Cryptococcus neoformans var. *grubii*					
VNI[a b]≡AFLP1[d]≡VN6 (VN5)[c]					
CBS 10085	WM 148	αA	1989	Australia, Sydney; clinical, CSF, HIV–	[a e]
CBS 8710	H99	αA	1978	USA, Durham; clinical CSF, Hodgkin's lymphoma; genome sequence strain	[f]
VNII[a b]≡AFLP1A[d]≡VN7[c]					
CBS 10084	WM 626	αA	1993	Australia, Sydney; clinical, CSF, HIV–	[a e]
AD hybrid					
VNIII[a b]≡AFLP3[d]≡VN33VN4[c]					
CBS 10080	WM 628	αA/aD	1988	Australia, Melbourne; clinical, CSF, HIV+	[a e]
CBS 132		αA/aD	1894	Italy; environmental, fermenting fruit juice	[g]
Cryptococcus neoformans var. *neoformans*					
VNIV[a b]≡AFLP2[d]≡VN1 (VN2)[c]					
CBS 10079	WM 629	αD	1987	Australia, Melbourne; clinical, blood, HIV+	[a]
CBS 6900	B-3501	αD	1975	USA, Bethesda; crossing of NIH 12 and 433	[h]

[a] Meyer et al., 1999; [b] Meyer et al., 2003; [c] Viviani et al., 1997; [d] Boekhout et al., 2001; [e] Chen et al., 1996; [f] Franzot et al., 1999; [g] Sanfelice, 1894; [h] Varma, 1989; ID = identifier.

Table 3.4.: Standard reference strains for *Cryptococcus gattii* typing, determined by the ISHAM Cryptococcal Working Group.

CBS ID	Alternate ID	MAT & Serotype	Year	Source	References
VGI[b]≡AFLP4[j]					
CBS 10078	WM 179	αB	1993	Australia, Sydney; clinical, CSF, HIV–	[b e]
CBS 6289		αB	1966	Congo, Kinshasa; clinical, CSF	[i]
CBS 10510	WM 276	αB	1993	Australia, Sydney; environmental,*Eycalyptus ereticornis* woody debris; genome sequence strain	[b e]
VGII[b]≡AFLP6[d]					
CBS 10082	WM 178	αB	1991	Australia, Sydney; clinical, CSF, HIV–	[b]
CBS 10514	WM 02.32	αB	2001	Canada, BC; clinical, bronchial wash; highly virulent Vancouver Island outbreak strain, VGIIa; genome sequence strain	[k]
VGIII[b]≡AFLP5[d]					
CBS 10081	WM 175	αB	1992	USA, California; environmental, *Eucalyptus spp.* woody debris	[b l]
CBS 6955		αC	Before 1970	USA, California; clinical, CSF	[m]
VGIV[b]≡AFLP7[d]					
CBS 10101	WM 779	αC	1994	South Africa, Johannesburg; veterinary (cheetah)	[b n]

[b]Meyer et al., 2003; [d]Boekhout et al., 2001; [e]Chen et al., 1996; [i]Gatti, 1970; [j]Boekhoutet al., 2003; [k]Kidd et al., 2004; [l]Sorrell et al., 1996; [m]Kwon-Chung et al., 1978; [n]Bolton et al., 1999; ID = identifier.

3.4.3. Consensus typing scheme MLST

The ISHAM group recognized the need for a cross-platform consensus-typing scheme for *Cn* in order to accurately discriminate between isolates of *Cn* and enable the rapid acquisition of global genotypic data. This typing scheme needed to be able to incorporate the findings from previous global-typing projects, while being universally applicable, publicly available, and able to integrate new data as it emerged. Multilocus sequence typing (MLST) was chosen as the as the "gold standard" technique in the progressive genotyping of the *Cn* species (Meyer et al., 2009), and seven genes were designated as loci for typing (Appendix A.1). The ideal MLST scheme should amplify and type the same genes from all eight molecular types using the same set of primers, as well as be discriminatory due to sufficient sequence diversity provided by the genes (Meyer, 2009). The genes should not be under strong or diversifying selection, as being subject to evolutionary pressure would bias genetic relatedness (Gil-Lamaignere et al., 2003). However, the final criteria had to be relaxed as one of the proposed genes was already widely used for typing (*CAP59*, see below).

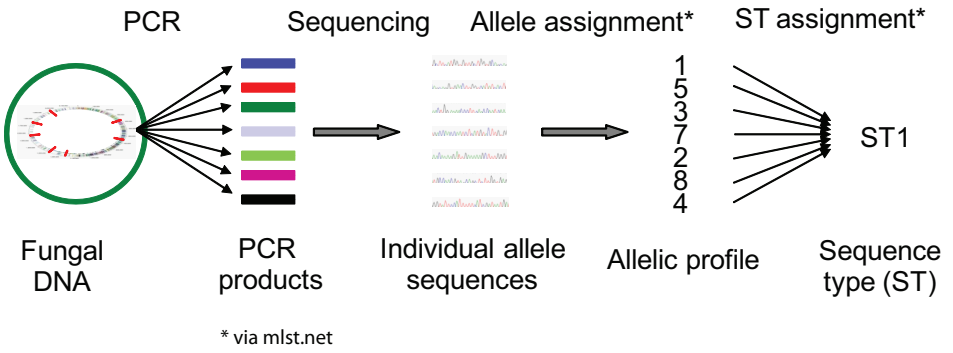
3.5. MLST

Introduced in 1998 by Maiden et al. and originally utilised in the typing of bacteria (Taylor and Fisher, 2003; Chan et al., 2001; Maiden, 2006; Spratt and Maiden, 1999; Urwin and Maiden, 2003), MLST has rapidly become recognised as one of the most widely applicable molecular typing approaches available today for medically important pathogens (Maiden et al., 1998; Revazishvili et al., 2004; Taylor and Fisher, 2003; Spratt et al., 2006), including *Candida albicans* (Bougnoux et al., 2004), *Candida glabrata* (Dodgson et al., 2003), *Candida tropicalis* (Tavanti et al., 2005), *Coccidioides spp.* (Koufopanou et al., 1997), *Histoplasma capsulatum* (Kasuga et al., 2003), and *Cryptococcus* (Fraser et al., 2005; Taylor and Fisher, 2003).

MLST is a conceptually simple method which is based on MLEE and comprises the amplification of DNA fragments of about 500 base pairs in length by PCR, followed by sequencing. The amplified allelic sequences are assigned distinct numbers (allele types, henceforth ATs) at each locus,

which collectively define an allelic profile. This, in turn, determines the sequence type (ST) of the isolate in question (Figure 3.1). Unlike MLEE, which relies on different electrophoretic mobilities of enzymes, MLST assigns sequence types to allelic profiles obtained directly from the nucleotide sequences of predetermined genes. This analysis therefore determines genotype, and not environmentally alterable enzymatic phenotype, and does so at a much higher resolution (Maiden et al., 1998). This unambiguious characterisation of strains is rapid and highly portable, because strains no longer need to be exchanged between laboratories. Through the analysis of the resulting sequence data using phylogenetic gene trees, we can elucidate the population structure of *Cn* (Boyd et al., 1996; Smith, 1992).

Figure 3.1.: Multilocus sequence typing involves the assigning of a sequence type (ST) to fungal isolates. The amplification and sequencing of fungal DNA yields an allelic type at individual loci, creating an allelic profile which dictates ST.



3.5.1. MLST loci

The genetic loci which constitute the MLST scheme would need to detect the largest number of different allele types, hence sufficient diversity between sequences. Ideally, the maximum level of discrimination would be achieved using the smallest possible number of loci. Furthermore, all loci included should be amplifiable at all five serotypes (eight molecular types).

In order to determine which loci would be included in the MLST scheme, we assessed two studies which used varying unlinked polymorphic loci to type 102 global isolates and 202 *C. gattii* strains pertaining to the Vancouver

Island outbreak. The former applied MLST at 12 loci: *MPD1*, *TOP1*, *MP88*, *CAP59*, *URE1*, *PLB1*, *CAP10*, *GPD1*, *TEF1*, *SOD1*, *LAC1* and *IGS1* (Litvintseva et al., 2006), and the latter at eight loci: *SXI***a**/α, *IGS1*, *TEF1*, *GPD1*, *LAC1*, *CAP10*, *PLB1* and *MPD1* (Fraser et al., 2005). Both studies applied highly polymorphic loci; as a result, the first separated the 102 globally sourced *Cng* isolates into three major clusters corresponding to VNI, VNII and VNB molecular groups, while the second identified four major molecular groups: VGI, VGII, VGIII and VGIV (Litvintseva et al., 2006; Fraser et al., 2005).

Computing Simpson's diversity index (Simpson, 1949) for the two studies, we determined that a minimal set of seven genetic loci resulted in optimal estimation of underlying relationships and evolutionary processes in each study (Meyer et al., 2009). The optimal combinations of loci were *MP88*, *CAP59*, *PLB1*, *GPD1*, *SOD1*, *LAC1* and *IGS1* for the global isolates (Simpson's diversity index 0.9632; Litvintseva et al., 2006), and *IGS1*, *TEF1*, *GPD1*, *LAC1*, *CAP10*, *PLB1* and *MPD1* (Simpson's diversity index 0.9319; Fraser et al., 2005) for the Vancouver isolates. Trial amplifications were conducted across all molecular types in six laboratories involved in developing the typing scheme in question, ours included (Fisher lab, Imperial College). On the basis of their findings, we selected the following seven loci to make up the standardised MLST scheme for the typing of the *Cryptococcus* species complex: *CAP59*, *GPD1*, *IGS1*, *LAC1*, *PLB1*, *SOD1*, and *URA5* (Table 3.5; Meyer et al., 2009).

All but one of the seven genes, *IGS1*, are housekeeping genes meaning they are involved in the obligatory sustenance and maintenance of the organism and are always expressed. More importantly, their expression remains constant across many conditions. Three of these six genes code for virulence factors. *CAP59* is involved in the formation of the polysaccharide capsule which exerts many immunoregulatory effects and is antiphagocytic (Petter et al., 2001; Garcia-Rivera et al., 2004). The secretion of cryptococcal phospholipase by the *PLB1* gene facilitates host cell evasion through cell lysis, enhancing the ability of the fungus to grow *in vivo* (Noverr et al., 2003; Chen et al., 2000). *LAC1* is involved in melanin synthesis, a gene product which scavenges host-produced antioxidants and protects *Cn* from oxidative damage (Jacobson and Emery, 1991; Polacheck et al., 1990). The intergenic

Table 3.5.: The seven designated loci for multilocus sequence typing, determined by ISHAM.

Gene	Gene product	Chromosome location	Length (bp)*
CAP59[a]	Capsular associated protein	1	561
GPD1[a]	Glyceraldehyde-3-phosphate dehydrogenase	7	547
IGS1[b]	Ribosomal RNA intergenic spacer	2	595
LAC1[a]	Laccase	8	475
PLB1[b]	Phospholipase	12	534
SOD1[c]	Cu, Zn superoxide dismutase	5	713
URA5[d]	Orotidine monophosphate pyrophosphorylase	8	606

[a] Fraser et al., 2005; [b] Litvintseva et al., 2006; [c] D'Souza et al., 2004; [d] Meyer et al., 2003; bp = base pairs

* may vary according to molecular type

spacer, *IGS1*, was included in the scheme as it proved to have high allelic diversity despite not being a housekeeping gene.

3.5.2. MLST website and *Cng* Database

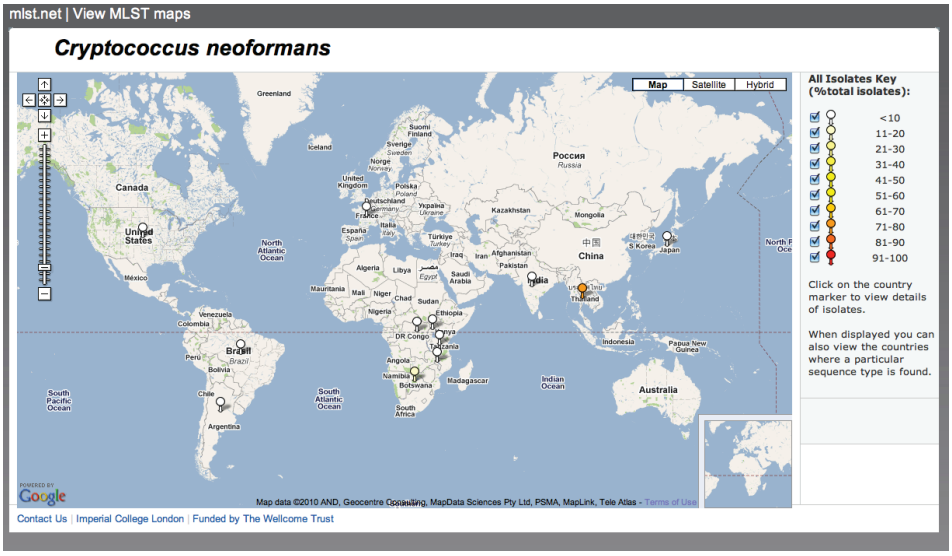
Sequence data produced by MLST is unambiguous and highly discriminatory, however, in order for its portability to be effective, there needs to be a single international resource and database where multiple users can deposit their data and query the data of other groups. The ISHAM working group elected to deposit the MLST data on the MLST website (`http://mlst.net/`), hosted at Imperial College, London. This website contains databases of reference allele sequence types and isolate epidemiological data for several human pathogens, including *E. coli*, *H. pylori* and *S. pneumonia*.

As a member of the Cryptococcal Working Group, I created a database consisting of the globally typed isolates, specifying the allelic profiles described so far according to the seven predetermined loci. This database forms the backbone of the MLST scheme for *Cn* var *grubii* and allows for the standardised assignment of the allele types (ATs) and sequence types (STs). As of 01/08/2010, the MLST scheme contained 50 STs from 230 clinical isolates, 21 environmental isolates and 11 of unknown source, from 14 countries worldwide (Litvintseva et al., 2006; Meyer et al., 2009; Simwami

et al., 2011). The data is available at `http://cneoformans.mlst.net/`, where interrogation and analysis software allow users to query their allele sequences and sequence types, enabling accurate inter-laboratory data comparison. As curator of this database, I was in charge of assessing submissions of novel ATs and STs made via the website, as well as uploading them to the database to ensure it is kept up to date. As the loci were selected due to their being applicable to the other *Cn* species, we intend to expand the database to include data on *Cn* var *neoformans* and *C. gattii*.

The automatic retrieval of previously described alleles and sequence types, as well as the assigning of new ones for submitted novel sequences, is implemented with the program NRDB (`http://linux.mlst.net/nrdb/nrdb.htm`). This software is utilised in allele assignment, comparing sets of sequences and finding those which are identical to each other, as well as those which are not yet in the database in question. It is important to note that this program is capable of processing DNA and amino acid sequence data, meaning it is not restricted to MLST data (Warren Gish, Washington University). MLST.net includes further bioinformatics tools such as eBurst, emaps and treemaps, which allow for further exploration of global biodiversity (`http://cneoformans.mlst.net/`). Emaps allows for the mapping of the global isolates according to location and thus the assessment of spatial relationships (Figure 3.2). Selection of a particular group allows for more in-depth analysis of the isolates within this group or in comparison to the global population. Eburst is a web-enabled clustering tool which can be used to identify mutually exclusive groups of related genotypes within populations, as well as the founding genotypes of each group (Feil et al., 2004). Other data analysis tools at the MLST website include Linkage Disequilibrium and Splits Tree. The former enables the detection of associations between genes at different loci and is appropriate when there is doubt about the extent of genetic recombination between members of a population (Smith et al., 1993). Splits Tree is used to analyse and visualise distance data using the split decomposition method (Dress et al., 1995; Hudson, 1998).

Figure 3.2.: MLST map of the current global *Cryptococcus neoformans* var *grubii* isolates. This screenshot of the current distribution of *Cng* isolates worldwide ($n = 261$) depicted by the MLST website represents the mapping tool utilised in comparative eBURST analysis of *Cng* populations.



3.6. Discussion

Great advances have been made in the detection and identification of fungi over the past decades. Initial phenotypic methods were to proven to be inadequate due to observed characteristics not only being variable within species, but also exhibiting major phenotypic plasticity in varying culture conditions (Kurtzman and Robnett, 1997; Franzot et al., 1998; Check, 1994). These mycological methods made way for molecular techniques which hinged on two main techniques: nucleic acid hybridisation and the polymerase chain reaction. The phylogenetic species recognition methods stemming from these resulted in great clarification of the taxonomy of *Cn* through increased speed and accuracy.

Despite these advances, there remained a need for uniformity in technique and the reporting of results in order to allow cross-referencing between studies. The Cryptococcal Working Group under the ISHAM working group on the 'Genotyping of *Cryptococcus neoformans* and *C. gattii*' proposed

the seven set of genetic loci as an international standard for multilocus sequence typing for *Cn* and *C. gattii*: *CAP59*, *GPD1*, *IGS1*, *LAC1*, *PLB1*, *SOD1* and *URA5*. This scheme allows for accurate and maximal discrimination between the eight various molecular groups of cryptococcal species, as well as between closely related strains using the minimal number of loci, based on 304 isolates from two studies (Litvintseva et al., 2006; Fraser et al., 2005). In conjunction with this scheme, the VN nomenclature is to be maintained: VNI–VNIV, VNB for *Cn* and VGI–VG for *C. gattii*.

MLST has proven to be unambiguous and highly discriminatory in the characterisation of individual *Cn* isolates. It is highly reproducible due to the advances made in nucleotide sequencing technology (Bougnoux et al., 2004) and the generation of standardised data. The resulting data is currently deposited on the MLST website `http://mlst.net/`, hosted at Imperial College, London, optimising its portability. This globally available dataset is up to date, allowing for data analysis and comparisons of the current global *Cng* database using tools such as eBURST. The standardisation of the typing method and nomenclature of this fungal pathogen will allow for better inter-laboratory and study collaboration in the mapping of the epidemiology and genetic makeup of this increasingly important fungus.

4. *Cng* in Thailand

Acknowledgements: The work presented in this chapter was made possible through a collaboration with the CBS fungal Biodiversity Centre. In particular, I'd like to thank Ferry Hagen and Kantarawee Khayhan whose contributions in the form of the sharing of isolates and typing data were invaluable.

Our work was published and is available in Appendix A.2:

Simwami, S. P., Khayhan, K., Henk, D. A., Aanensen, D. M., Boekhout, T., Hagen, F., Brouwer, A. E., Harrison, T. S., Donnelly, C. A. and Fisher, M. C. (2011) 'Low diversity *Cryptococcus neoformans* variety grubii multilocus sequence types from Thailand are consistent with an ancestral African origin', *PLoS pathogens,* 7(4), e1001343.

4.1. Introduction

There are an estimated 120 000 yearly cases of AIDS–associated CM in South and Southeast Asia. A 90-day case-fatality rate of 55% places this region second worldwide in terms of mortality rate from cryptococcal infection (Park et al., 2009). Mirroring the poor prognosis among adults with AIDS-associated CM within this region, cryptococcosis is a leading AIDS-defining systemic infection in Thailand, a country with 610 000 people living with HIV/AIDS (Pitisuttithum et al., 2001; UNAIDS/WHO, 2008). The high rates of mortality, readmission and relapse are attributed to a combination of factors; these include high poverty rates, resulting in few being able to afford timely antifungal treatment, the limitations of current antifungal drugs, the limited availability of highly active antiretroviral therapy (HAART), and the trend of late presentation due to religious and cultural influences (Wright and Inverarity, 2007). The HIV/AIDS epidemic has not

only resulted in a dramatic increase of immunocompromised populations, but also changed the aetiology of CM in Thailand, making this country of great interest in terms of the study of the fungus' epidemiology (Day, 2004; Mitchell and Perfect, 1995; Schutte et al., 2000; Sukroongreung et al., 1996).

4.1.1. *Cn* in Thailand

Cryptococcus was first isolated from natural substrates in Thailand in 1968, with the fungus being recovered from 8 of 43 samples from avian habitats in the northeast, center and southeast of Thailand (Taylor and Duangman, 1968). This study speculated that avian habitats were a potential source of infection, as the keeping of small birds including cuckoos and pigeons in cages throughout homes was customary in Thailand.

Between January 1988 and December 1993, 254 cases of systemic mycoses were recorded in Thailand, with cryptococcosis among the three most common. In 1992, thirty cases of CM were recorded, 27 of which were AIDS related. The following year, the 57 reported cases of systemic mycoses were also largely attributable to *Cryptococcus* ($n = 49$). A reporting system of AIDS– defining illnesses based on the CDC clinical case definition and requiring laboratory confirmation of cryptococcosis was then put in place by the Division of Epidemiology of the Ministry of Public Health (Chariyalertsak et al., 2001; Castro et al., 1993). Between 1994 and 1998, the annual incidence of AIDS–associated CM reported to Thailand's Ministry of Public Health more than doubled, from 2 148 to 4 348 patients. Cryptococcal infections made up around 18–19% of all 26 reported AIDS-defining illnesses, and 43% percent of these infections occurred in the north of Thailand, in keeping with the higher rates of HIV/AIDS in this region (Chariyalertsak et al., 2001). There was no evidence of these infections being seasonal, but regional variations were noted in the prevalence of extrapulmonary cryptococcosis, which appeared to be most common in the north (24.1%) and northeast, as opposed to the south of Thailand (only 6.9%; Chariyalertsak et al., 2001).

Prior to the AIDS epidemic, *C. gattii* was reported to have caused proportionally more disease than *Cng*, with a study reporting 66% of 187 isolates being of *C. gatti*. Of this collection of isolates, only 33% were *Cn* (28%

variety *grubii*, and 5% variety *neoformans*; Sukroongreung et al., 1996). This high distribution of *C. gattii* in Thailand reflected that found among a collection of 12 Southeast Asian isolates of the same era, all of which were of *C. gattii* (Kwon-Chung and Bennett, 1984). Sukroongreung et al. (1996) went on to compare the pre-AIDS cohort to a second collection of 169 cryptococcal samples isolated between January 1993 and March 1995. *C. gattii* prevalence had decreased to about 4% of the total, while *Cn* was responsible for about 97% of all infections (Sukroongreung et al., 1996). A more recent study of 139 *Cn* strains from Thai patients, 97 of which were HIV+, reflected this post-pandemic predominance of the *Cn* varieties, with *Cng* accounting for 96% of the isolates, while *C. gattii* was not found at all (Poonwan et al., 1997). The dramatic increase in environmental and clinical *Cng* prevalence can be explained by the large immunocompromised population that is susceptible to secondary infection by this variety of *Cn*. Despite the emerging importance of this pathogen and increased research effort (Bovers et al., 2008b; Litvintseva et al., 2006), aspects of the pathogen's global population genetic structure remain undetermined.

4.2. Aim

The aim of this study was to describe the population genetic structure of the previously untyped, but clinically important, population of *Cng* that infects HIV/AIDS patients in Thailand, Southeast Asia, with the intention of integrating these data into broader global patterns of genetic diversity within this species.

4.3. Materials and Methods

4.3.1. Isolates

Included in this study are 183 Thailand isolates of *Cng* acquired from three sources. Fifty-eight clinical isolates were collected during a randomized control trial at Sappasitprasong Hospital, Ubon Ratchathani, northeast Thailand. This study aimed to compare the efficacy of four randomly assigned antifungal treatment combinations in the initial treatment

of HIV-associated CM in an antiretroviral therapy (ART) naive population, enrolling 64 adults with a first episode of cryptococcal meningitis (CM; Brouwer et al., 2004). A further 108 clinical isolates were obtained from a collection of cryptococcal samples managed by the CBS-KNAW Fungal Biodiversity Center. These originated from patients at various hospitals in three Thai regions (76 in the north, 20 in the northeast and 9 from the south), with the remaining three isolates being of unknown provenance. Of the total 173 clinical isolates, 154 (89%) were from HIV/AIDS patients with culture-proven *Cn* isolated from cerebrospinal fluid ($n = 127$), blood ($n = 12$) and broncho-alveolar lavage ($n = 1$). Three were from blood samples of HIV-negative CM patients. A further eighteen cryptococcal isolates were provided by Dr. Pojana Sriburee, Chiang Mai University, ten of which were environmental and had been isolated from pigeon and dove guano (Sriburee et al., 2004). One of the eight remaining isolates recovered from cryptococcosis patients was of Japanese origin, and was not considered as part of the Thai dataset (isolate J1; Table B.1). In total, these three collections yielded 183 isolates from 11 provinces in three regions of Thailand: north ($n = 91$), northeast ($n = 79$) and south ($n = 9$), and a remaining four unknown, with 6% of the total isolates being environmental (Table 4.1).

Table 4.1.: Allelic profiles of the 183 Thai *Cng* isolates according to the seven MLST loci (base pairs in brackets). Haplotypes within each locus column are differentiated by colour.

Name	*CAP59* (501)	*GPD1* (489)	*IGS1* (709)	*LAC1* (471)	*PLB1* (533)	*SOD1* (527)	*URA* (637)	ST	Strain origin (if known)
CN5010	1	1	19	3	2	13	5	**4**	C. Rai, blood
CN4998	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4995	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4989	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4988	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4987	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4964	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4947	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4945	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4944	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN4943	1	1	19	3	2	13	5	**4**	C. Rai
CN4942	1	1	19	3	2	13	5	**4**	Lampang, CSF
CN4941	1	1	19	3	2	13	5	**4**	Thailand, CSF
CN4940	1	1	19	3	2	13	5	**4**	Thailand, CSF
CN4926	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4919	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4918	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4917	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4903	1	1	19	3	2	13	5	**4**	C. Rai, CSF
CN4901	1	1	19	3	2	13	5	**4**	C. Mai, CSF
CN49005	1	1	19	3	2	13	5	**4**	C. Mai
4-187	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
269	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
4-315	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
1-587	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
1219	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
4_83	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
1-588	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
4-202	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
1-846	1	1	19	3	2	13	5	**4**	K. Kaen, clinical
2551-07	1	1	19	3	2	13	5	**4**	Songkhla, CSF
2550 II-07	1	1	19	3	2	13	5	**4**	Songkhla, blood
2461-07	1	1	19	3	2	13	5	**4**	Songkhla, CSF
CM 1	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 6	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 7	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 8	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 12	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 13	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF

C. Rai = Chiang Rai; C. Mai = Chiang Mai; K. Kaen = Khon Kaen;
U. Ratchathani = Ubon Ratchathani; CSF = cerebrospinal fluid;
crypto patient = cryptococcosis patient; novel ATs are in bold.

(continued on next page)

Name	*CAP59* (501)	*GPD1* (489)	*IGS1* (709)	*LAC1* (471)	*PLB1* (533)	*SOD1* (527)	*URA* (637)	ST	Strain origin (if known)
CM 17	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 18	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 22	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 23	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 25	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 26	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 33	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 37	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 38	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 39	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 40	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 41	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 42	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 43	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 44	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 46	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 47	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 48	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 49	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 51	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 55	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 56	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 57	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 58	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 59	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 61	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
CM 63	1	1	19	3	2	13	5	**4**	U. Ratchathani, CSF
K 2	1	1	19	3	2	13	5	**4**	K. Kaen, crypto patient
Pg 1	1	1	19	3	2	13	5	**4**	C. Mai, pigeon dropping
D 6	1	1	19	3	2	13	5	**4**	C. Mai, dove dropping
D 1	1	1	19	3	2	13	5	**4**	C. Mai, dove dropping
CN5019	1	1	19	4	2	13	5	**6**	C. Rai, blood
CN5017	1	1	19	4	2	13	5	**6**	C. Rai, CSF
CN5014	1	1	19	4	2	13	5	**6**	C. Rai, blood
CN5013	1	1	19	4	2	13	5	**6**	C. Rai, CSF
CN5011	1	1	19	4	2	13	5	**6**	Thailand, clinical
CN5009	1	1	19	4	2	13	5	**6**	C. Rai, blood
CN5005	1	1	19	4	2	13	5	**6**	C. Rai, blood
CN5003	1	1	19	4	2	13	5	**6**	C. Rai, blood
CN5002	1	1	19	4	2	13	5	**6**	C. Rai, blood
CN5001	1	1	19	4	2	13	5	**6**	C. Rai, CSF
CN4970	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4968	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4957	1	1	19	4	2	13	5	**6**	C. Rai, CSF

C. Rai = Chiang Rai; C. Mai = Chiang Mai; K. Kaen = Khon Kaen;
U. Ratchathani = Ubon Ratchathani; CSF = cerebrospinal fluid;
crypto patient = cryptococcosis patient; novel ATs are in bold.

(continued on next page)

Name	*CAP59* (501)	*GPD1* (489)	*IGS1* (709)	*LAC1* (471)	*PLB1* (533)	*SOD1* (527)	*URA* (637)	ST	Strain origin (if known)
CN4956	1	1	19	4	2	13	5	**6**	C. Rai, CSF
CN4955	1	1	19	4	2	13	5	**6**	Thailand, BAL
CN4954	1	1	19	4	2	13	5	**6**	Lampang, CSF
CN4952	1	1	19	4	2	13	5	**6**	Tak, CSF
CN4950	1	1	19	4	2	13	5	**6**	Lampoon, CSF
CN4949	1	1	19	4	2	13	5	**6**	Lampoon, CSF
CN4938	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4937	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4936	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4934	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4933	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4932	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4931	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4927	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4915	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4914	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4909	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4907	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4905	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4904	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN4902	1	1	19	4	2	13	5	**6**	C. Mai, CSF
CN49008	1	1	19	4	2	13	5	**6**	C. Mai, CSF
4-319	1	1	19	4	2	13	5	**6**	K. Kaen, clinical
50NC2	1	1	19	4	2	13	5	**6**	Nan, clinical
50NC5	1	1	19	4	2	13	5	**6**	Nan, clinical
11112	1	1	19	4	2	13	5	**6**	K. Kaen, clinical
11109	1	1	19	4	2	13	5	**6**	K. Kaen, clinical
4-231	1	1	19	4	2	13	5	**6**	K. Kaen, clinical
P6	1	1	19	4	2	13	5	**6**	C. Mai, clinical
4-253	1	1	19	4	2	13	5	**6**	K. Kaen, clinical
4-381	1	1	19	4	2	13	5	**6**	K. Kaen, clinical
20662-07	1	1	19	4	2	13	5	**6**	Songkhla, blood
28170-07	1	1	19	4	2	13	5	**6**	Songkhla, CSF
1111I-08	1	1	19	4	2	13	5	**6**	Pattani, blood/HIV–
2895I-08	1	1	19	4	2	13	5	**6**	Pattani, blood/HIV–
4500-07	1	1	19	4	2	13	5	**6**	Pattani, blood
CM 2	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 3	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 4	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 5	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 10	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 14	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 11	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 15	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF

C. Rai = Chiang Rai; C. Mai = Chiang Mai; K. Kaen = Khon Kaen;
U. Ratchathani = Ubon Ratchathani; CSF = cerebrospinal fluid;
crypto patient = cryptococcosis patient; novel ATs are in bold.

(continued on next page)

Name	*CAP59* (501)	*GPD1* (489)	*IGS1* (709)	*LAC1* (471)	*PLB1* (533)	*SOD1* (527)	*URA* (637)	ST	Strain origin (if known)
CM 16	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 20	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 24	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 27	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 28	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 29	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 32	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 34	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 36	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 45	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 50	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 52	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 60	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
CM 64	1	1	19	4	2	13	5	**6**	U. Ratchathani, CSF
Pt 9	1	1	19	4	2	13	5	**6**	C. Mai, crypto patient
Pt 3	1	1	19	4	2	13	5	**6**	C. Mai, crypto patient
Pt 1	1	1	19	4	2	13	5	**6**	C. Mai, crypto patient
D 2	1	1	19	4	2	13	5	**6**	C. Mai, dove dropping
D 3	1	1	19	4	2	13	5	**6**	C. Mai, dove dropping
Pg 2	1	1	19	4	2	13	5	**6**	C. Mai, pigeon dropping
Pg 26	1	1	19	4	2	13	5	**6**	C. Mai, pigeon dropping
CN49004	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN48	1	3	19	5	2	13	1	**5**	K. Kaen, clinical
1-488	1	3	19	5	2	13	1	**5**	K. Kaen, clinical
1-489	1	3	19	5	2	13	1	**5**	K. Kaen, clinical
CM 30	1	3	19	5	2	13	1	**5**	U. Ratchathani, CSF
Pt 12	1	3	19	5	2	13	1	**5**	C. Mai, crypto patient
D 5	1	3	19	5	2	13	1	**5**	C. Mai, dove dropping
Pg 37	1	3	19	5	2	13	1	**5**	C. Mai, pigeon dropping
CN5015	1	3	19	5	2	13	1	**5**	C. Rai, CSF
CN5018	1	3	19	5	2	13	1	**5**	C. Rai, blood
CN5012	1	3	19	5	2	13	1	**5**	C. Rai, CSF
CN5008	1	3	19	5	2	13	1	**5**	C. Rai, CSF
CN4993	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4983	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4980	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4977	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4967	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4960	1	3	19	5	2	13	1	**5**	C. Rai, CSF
CN4948	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4946	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4924	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4921	1	3	19	5	2	13	1	**5**	Mae Hong Son, CSF
CN4920	1	3	19	5	2	13	1	**5**	C. Mai, CSF

C. Rai = Chiang Rai; C. Mai = Chiang Mai; K. Kaen = Khon Kaen;
U. Ratchathani = Ubon Ratchathani; CSF = cerebrospinal fluid;
crypto patient = cryptococcosis patient; novel ATs are in bold.

(continued on next page)

Name	*CAP59* (501)	*GPD1* (489)	*IGS1* (709)	*LAC1* (471)	*PLB1* (533)	*SOD1* (527)	*URA* (637)	ST	Strain origin (if known)
CN4916	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN4906	1	3	19	5	2	13	1	**5**	C. Mai, CSF
CN49006	1	3	19	5	2	13	1	**5**	C. Mai, CSF
50NC1	1	3	19	**10**	2	13	1	**50**	Nan, clinical
Pt 5	1	1	19	5	2	13	1	**48**	C. Mai, crypto patient
CN5007	1	1	20	3	4	13	1	**32**	C. Rai, CSF
1291-09	1	1	20	3	4	13	1	**32**	Pattani, blood/HIV–
CM 35	1	1	20	3	4	13	1	**32**	U. Ratchathani, CSF
K 45	1	1	19	3	4	13	5	**47**	K. Kaen, crypto patient
4_9	1	1	19	**9**	2	13	5	**49**	K. Kaen, clinical
D 9	1	1	19	4	2	13	**14**	**46**	C. Mai, dove dropping
CM 21	2	10	21	6	11	14	4	**45**	U. Ratchathani, CSF

C. Rai = Chiang Rai; C. Mai = Chiang Mai; K. Kaen = Khon Kaen;
U. Ratchathani = Ubon Ratchathani; CSF = cerebrospinal fluid;
crypto patient = cryptococcosis patient; novel ATs are in bold.

DNA was extracted from the 183 Thai isolates and utilized for molecular analyses of mating type, ST and molecular type, as outlined in section 2.1. Analytical tools used to investigate the genetic structure of this population include AMOVA, PCA, tests of recombination and neutral expectations, as well as phylogenetic methods, also described in Chapter 2.

4.4. Ethics Statement

Ethical approval was required for the randomized control trial at Sappasitprasong Hospital, Ubon Ratchathani, the source of some isolates typed in this study. This was approved by the ethical and scientific review subcommittee of the Thai Ministry of Public Health and by the research ethics committee of St George's Hospital, London, UK, with written informed consent obtained for all 64 adults enrolled in this study.

4.5. Results

4.5.1. Mating type, variety and MLST sequence type

All 183 Thai isolates typed in this study were *Cng* (serotype A) and of mating type MATα.

Ten sequence types identified in the Thai population

Sequence data were obtained for all 183 Thai isolates typed at the seven loci (Table 4.1). The aligned sequences of the concatenated loci were 3 959 base pairs in total, with 112 polymorphic sites (20 parsimony informative and 92 singleton sites).

The seven loci yielded 23 allele types (ATs), three of which were novel to the Thai population of *Cng*, having not been previously described (in bold; Table 4.1). Loci *LAC1* and *URA5* consisted of two and one novel ATs, respectively, while *CAP59*, *GPD1*, *IGS1*, *PLB1* and *SOD1* were composed of previously described ATs (Litvintseva et al., 2006). We identified a total of 10 multilocus sequence types (STs) within the Thai isolates. Four of these STs had previously been described within the global profile of *Cng* (STs 4, 5, 6 and 32) which consisted of 44 STs (described in Chapter 6; Table 6.1; Table B.2). The six STs novel to the Thai population of *Cng* were therefore allocated consecutive ST numbers 45 - 50 (Table 4.1), resulting in a dataset of 50 global STs for *Cng*. ST4 accounted for 38% of the Thai isolates (n = 70), ST5 for 14% (n = 26), and ST6 for 43% (n = 78 ; Table 4.1). STs 4 and 6 collectively contained 81% of all the isolates and differed only at the *LAC1* gene (nucleotide positions 36, 190, 232 and 338). STs 45 to 50 consisted of single isolates, all of which differed from at least one other ST at a single locus. Nine of the ten environmental isolates shared identical genotypes with clinical isolates (Table 4.1).

Analyses of genetic variation and phylogeny reveal a genetically depauperate Thai *Cng* population

The average nucleotide diversity within the Thai population was explored at all seven loci by calculating their haplotypic diversity (H_d), the num-

ber of nucleotide differences amongst sites (π) and Watterson's estimate of the population-scaled mutation rate (θ). The average estimates of these statistics for the concatenated sequences were low in comparison to those of African origin described in Chapter 5 ($H_d = 0.19$, $\pi = 0.001$ versus $H_d = 0.92$, $\pi = 0.006$, respectively; Tables 4.2 and 5.4), reflecting the low number of haplotypes, which ranged from two to six at the seven loci. Locus *LAC1*, 467 base pairs long, had the greatest number of segregating sites ($n = 61$), while *CAP59* had the lowest haplotypic diversity and population-scaled mutation rate (0.01 and 0.002, respectively).

Table 4.2.: Nucleotide diversity among the Thailand *Cng* isolates ($n = 183$).

Locus	pb^a	S^b	h^c	Hd^d	π^e	θ^f	D^g	R_2^h	${R_m}^i$
CAP59	501	5	2	0.01	0.0001	0.002	-1.819^{*}	0.0737	0
GPD1	489	6	3	0.27	0.0006	0.002	-1.431	0.0652	0
IGS1	721	12	3	0.04	0.0006	0.003	-1.960^{*}	0.0206	0
LAC1	467	61	6	0.64	0.0032	0.023	-2.635^{***}	0.0591	2
PLB1	533	7	3	0.05	0.0002	0.002	-1.938^{*}	0.0633	0
SOD1	529	11	2	0.01	0.0002	0.004	-2.259^{**}	0.0737	0
URA5	637	10	4	0.31	0.0006	0.003	-1.819^{*}	0.0612	0
		Average		0.19	0.0008	0.005			$5^{\#}$

[a]total number of sites in alignments, excluding indels and missing data; [b]number of segregating sites; [c]number of haplotypes; [d]haplotypic diversity; [e]average number of nucleotide differences per site; [f]Watterson's estimate of the population-scaled mutation rate, expressed per site (Watterson, 1975); [g]Tajima's D (Tajima, 1989); [h]Ramos-Onsins & Rozas' R_2 (Ramos-Onsins and Rozas, 2002); [i]minimum number of recombination events (Hudson and Kaplan, 1985); [#]average R_m between all seven loci; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

The spatial partitioning of genetic variability in the Thai *Cng* population typed in this study was examined using Analysis of Molecular Variance (AMOVA). This analysis demonstrated that only a small proportion (5%; $p < 0.013$) of the total estimated variance was attributable to the among-population variance component between the three Thai regions (north, northeast and south; Table 4.3).

A Principal Component Analysis (PCA) was then used to assess the hierarchical structuring of the genetic population of *Cng* in Thailand within 'R v.2.6.1' (R code C.1). The genetic structure captured by the first two principal components was depicted by the individual genotypes (represented by

Table 4.3.: Summary of AMOVA of Thai *Cng* isolates ($n = 179$), based on the seven polymorphic loci and according to geographical origin: north ($n = 92$), northeast ($n = 78$), south ($n = 9$).

	d.f.	Sum of squares	Variance components (%)	ΦPT	p-value[a]
Among populations	2	4	0.03 (5)	0.05	0.013
Within populations	176	114	0.65 (95)		
Total	178	118	0.68 (100)		

[a] p-value estimates are based on 999 permutations.

NB. Four isolates within the Thai population are of unknown origin and thus excluded from analysis.

dots) clustering into three groups, and summarized by 95% ellipses. Again, the typology of the individual allelic profiles revealed little differentiation between the 183 isolates from the three regions (Figure 4.1).

Phylogenetic analysis reveals two distinct molecular VN groups within the Thai *Cng* population

A maximum likelihood tree depicting the phylogenetic relationships within Thailand supported this genetic homogeneity, with all but the single isolate CM21 of ST45 clustering together with high bootstrap support (bootstrap 92%; Figure 4.2). CM21's allelic profile consisted of seven ATs which were not found in any other Thai isolate typed in this study. Also of interest is isolate 50NCI of ST50. Although identical to ST5 at six of the seven loci, 50NCI appears to be an outlier due to variations in its nucleotide sequence at *LAC1* (Table 4.1).

Figure 4.1.: Principal Component Analysis of the allelic profiles of the Thai *Cng* genotypes typed in this study. Individual genotypes (dots) are linked by colored lines to form clusters which are summarized by colored ellipses proportional in size to the number of isolates represented. The three groups depicted are numbered and defined according to Thai region: 1 = north (red; $n = 91$), 2 = northeast (blue; $n = 79$) and 3 = south (purple; $n = 9$). p-value showing inter-regional differentiation and eigenvalues are represented in the bar plot.

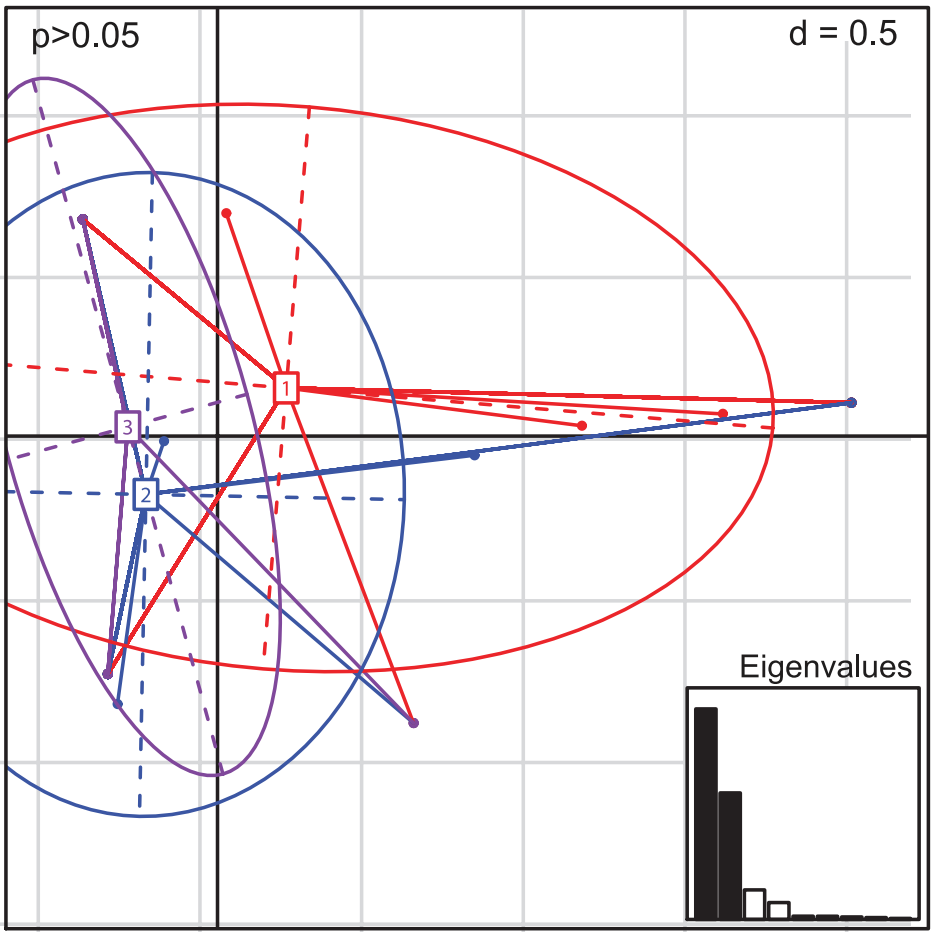
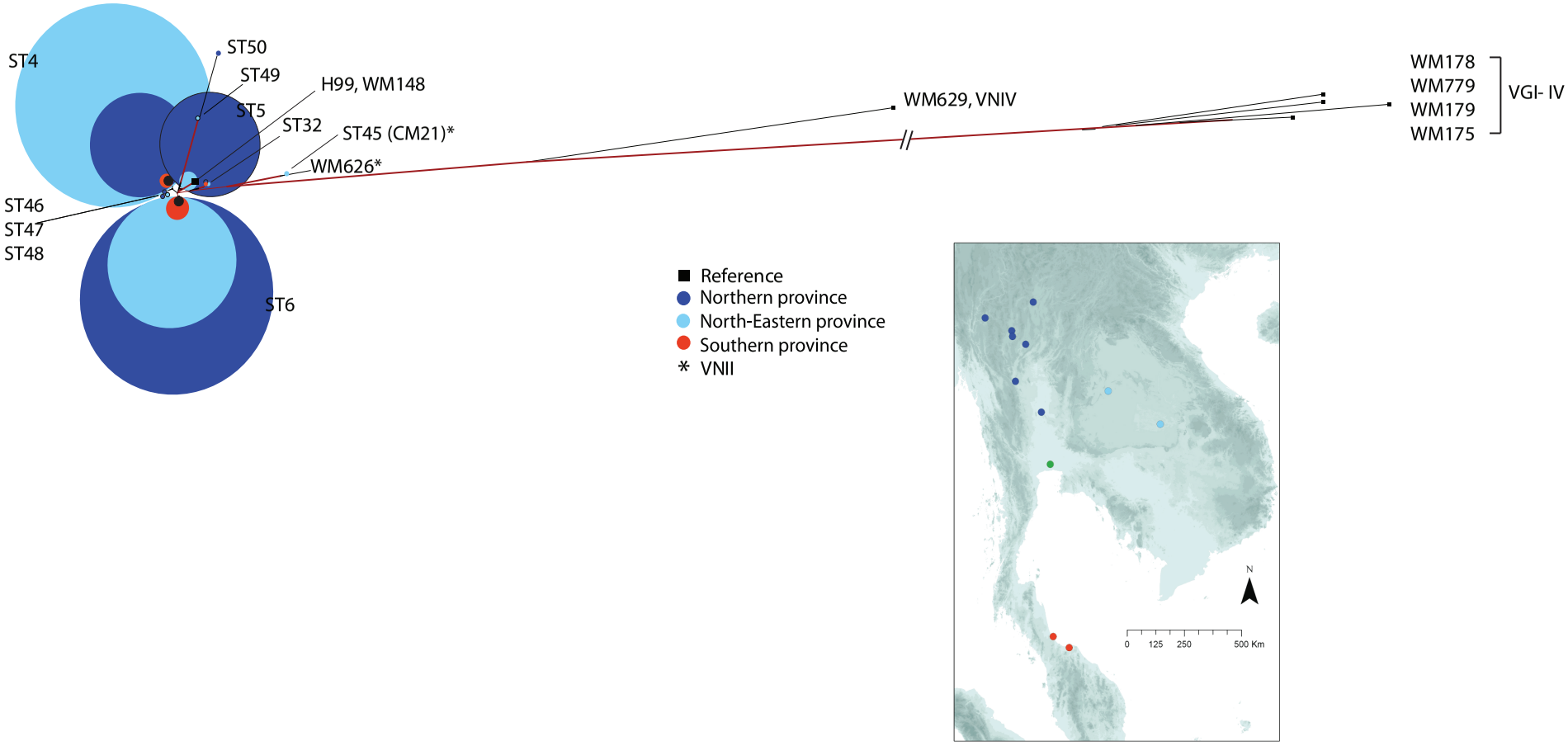


Figure 4.2.: Neighbour-joining tree illustrating the evolutionary relationships of the Thai isolates typed in this study ($n = 183$). Each circle represents a Sequence Type (ST) of the Thai isolates and is proportional in size to the number of isolates of this ST. The isolates are grouped according to three regions of Thailand, northern province in dark blue ($n = 91$), northeastern province in light blue ($n = 79$) and southern province in red ($n = 9$). The four Thai isolates of unknown origin are in black ($n = 4$). The percentage replicate trees in which the associated taxa clustered together in the bootstrap test (1 000 replicates) more than 70% of the time ($n \geq 70\%$) are indicated. The evolutionary distances were computed using the maximum composite likelihood method and are in the units of the number of base substitutions per site.



Molecular VN groupings of the Thai isolates were inferred through phylogenetic and comparative analyses with reference strains of known major molecular types of the *C. neoformans*/*C. gattii* species complex: WM148 (serotype A, VNI), WM626 (serotype A, VNII), WM629 (serotype D, VNIV), WM179 (serotype B, VGI), WM178 (serotype B, VGII), WM175 (serotype B, VGIII), WM779 (serotype C, VGIV; Meyer et al., 2003), and the genome-project strain H99 (serotype A, VNI; Perfect et al., 1993). The resulting phylogenetic dendrogram illustrating the relationships between the 183 Thai isolates delineated two major groups within the population: VNI and VNII (Figure 4.2). All but one of the Thai isolates typed in this study clustered with the VNI reference strains H99 and WM148 with significant bootstrap support ($n = 182$; 92%), with the single isolate CM21 of ST45 falling within molecular group VNII, along with the VNII reference strain WM626 (bootstrap support = 100%; Figure 4.2). Isolate CM21 being of a different VN group explains why its allelic profile was so different to the other nine STs of the Thai isolates (Table 4.1). In addition, although a seeming outlier of ST50, isolate 50NCI was found to correlate with the VNI group (WM148, H99), also supported by significant bootstrap value ($n = 76\%$; Figure 4.2).

Predominant clonality detected within the Thai *Cng* population

The index of association (I_A; Burt et al., 1996) and $\bar{r}_d$ (Agapow and Burt, 2001) were used to assess the overall association between alleles at the seven MLST loci, testing the null hypothesis of linkage equilibrium. A signature of clonal reproduction is the occurrence of non-random associations between loci, which generates linkage disequilibrium across the genome. Random association of alleles at the different loci rejected random outcrossing (panmixia) for the Thai *Cng* population ($\bar{r}_d = 0.28$, p-value < 0.001; Table 4.4). Clone-corrected data confirmed the predominance of clonal reproduction among these samples. The proportion of phylogenetically compatible pairs of loci was used to test for linkage disequilibrium in the dataset, with the null hypothesis of free recombination being rejected if there were fewer than two locus pairs with all four allele combinations, as expected under panmixia (Bennett et al., 2005). A significant percentage of phylogenetically compatible loci pairs was found (Table 4.4), and the hypothesis of random mating rejected.

Table 4.4.: Multilocus linkage disequilibrium analysis for samples of *Cn* var *grubii* from Thailand ($n = 183$).

Total sample			Clone-corrected sample[a]		
I_A [b]	$\bar{r}_d$ [c]	PcP [d]	I_A	$\bar{r}_d$	PcP
1.49***	0.34***	0.95***	1.98***	0.34***	0.95

[a]excluding replicate haplotypes; [b]index of association; [c]scaled index of association (I_A) by the number of loci ($m - 1$); [d]percentage of phylogenetically compatible pairs (PcP) of loci; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

The minimum number of recombination events (R_m; Hudson and Kaplan, 1985) was estimated both within an individual locus and between loci (R_m and average R_m respectively; Table 4.2). Despite the Thai population (n = 183) exhibiting extensive linkage disequilibria, some evidence for inter-locus recombination was detected (average $R_m = 5$; Table 4.2). The locus with the highest inferred intralocus R_m was *LAC1* ($n = 2$), which had the greatest number of segregating sites (Table 4.2). However as this locus has the highest number of segregating sites, this inferred recombination may in fact be due to homoplasies, owing to the high levels of diversity found. Therefore, the main feature of the Thai VNI *Cng* population is strong clonality, evidence of local clonal expansion within this geographical subset of the global VNI population.

Tajima's D tests the null hypothesis that populations are in mutation-drift equilibrium (Tajima, 1989). In the case of significant deviation from zero, the null hypothesis of neutral (random) evolution is rejected, a finding which can be attributed either to the occurrence of natural selection or to variable population dynamics. Significant departures from neutrality were detected at five of the seven loci of the Asian population (Table 4.2), all of which had negative values. Ramos-Onsins & Rozas' R_2 test (Ramos-Onsins and Rozas, 2006) did not detect any deviation from random evolution among any of the populations (Table 4.2).

4.6. Discussion

Although Thailand has achieved remarkable reductions in population-based transmission of HIV, there still remains a high number of people living with AIDS in this populous country (UNDP, 2004). This vulnerable population is further inflated by people suffering from diabetes mellitus, of which there will be an estimated 1.9 million over the age of 20 in 2025 in Thailand (King et al., 1998). Diabetes mellitus is among the most common underlying conditions associated with HIV-negative CM, accounting for an estimated 14% of infections following malignancies and the presence of systemic lupus erythematosus (16% each), and immunosuppressive drug treatment (41%; Kiertiburanakul et al., 2006). Furthermore, increased medical practices including transplantations of solid organs and bone marrow, immunosuppressive chemotherapy, intravenous drug use and the use of broad-spectrum antibiotics are also contributing to the amount of people with impaired immune systems within Thailand, a feature that is shared with countries across the rest of Asia (Xu et al., 2010). In addition, there is a real risk of resurgence in HIV/AIDS cases due to factors such as low awareness of HIV status, the high use of non-sterile injecting equipment (26–53% among IDUs in Chiang Mai, Songkla and Samut Praken) and the percentage of interspousal condom use being as low as 40% (USAID, 2010). All these predisposing conditions result in a large population which is highly susceptible to secondary fungal infections, cryptococcosis being among the most common with 44 061 cases/year, following *Mycobacterium tuberculosis* (81 955 cases/year) and *Pneumocystis jirovecii* (58 433 cases/year). The potential for the continued increase in the disease burden of AIDS-related meningitis cannot be ignored (Bicanic and Harrison, 2004).

We used MLST to describe the genetic structure of *Cn* in Thailand. All 183 isolates typed were of *Cng* (serotype A) and mating type MATα. This is consistent with previous reports that *Cng*, mating type MATα is the dominant cause of cryptococcosis among immunocompromised individuals in Thailand, after the onset of the AIDS epidemic (Sukroongreung et al., 1996; Poonwan et al., 1997). This is also similar to the global and environmental pattern of *Cn* distribution (Meyer et al., 1999; Meyer et al., 2003; McClelland et al., 2004; Mitchell and Perfect, 1995; Bicanic and Harrison, 2004; Kwon-Chung and Bennett, 1978; Yan et al., 2002). Similarly, all but

one Thai isolate, CM21, were of molecular type VNI (Figure 4.2), the most prevalent VN-type worldwide (Meyer et al., 1999; Casali et al., 2003; Meyer et al., 2003), as well as among Southeast Asian populations such as Thailand (Sukroongreung et al., 1996), China (Chen et al., 2008) and Malaysia (Tay et al., 2006).

MLST revealed ten sequence types (global STs 4, 5, 6 and 32 and novel STs 45–50), three of which accounted for 95% of the isolates typed. Two of these three Thai STs (STs 4 and 6) contained 81% of the 183 isolates (Table 4.1) and differed at only four nucleotide positions within the *LAC1* locus. The most diverse locus was *LAC1* which consisted of two new ATs (ATs 9 and 10). Furthermore, *LAC1* AT10 does not align with sequences of *Cng*, *Cn* var. *neoformans* or *C. gattii*. *LAC1* ATs for isolates 4_9 and 50NCI were ATs 9 and 10, respectively. These isolates were otherwise identical in allelic profile to STs 4, 6 (4_9 of ST49), and 5 (50NCI of ST50). This similarity is reflected in the fact that these two isolates correlate with the VNI group (WM148, H99; bootstrap support = 76%; Figure 4.2). A possible explanation for the marked difference of these ATs at *LAC1* may be that the sequence is in fact of the second putative laccase gene, *LAC2*. *LAC1* and *LAC2* are adjacent in position, share ≥75% nucleotide identity and can be amplified by similar primers (Pukkila-Worley et al., 2005). *LAC2* also shares significant homology with other fungal laccases including *Botryotinia fuckeliana* (Schouten et al., 2002) and *Agaricus bisporus* (Perry et al., 1993). This hypothesis could be tested by designing specific primers which would be used to PCR the two LAC genes from each isolate. The sequences of these would be analysed for levels of homology in order to ascertain whether there are gains or losses of the gene.

AMOVA revealed that only 5% of the observed genetic variation across Thailand could be attributed to differences among the three regions (Table 4.3), showing that *Cng* exhibits little spatial structure at this geographic scale. PCA (Figure 4.1) and phylogenetic analyses (Figure 4.2) support the conclusion that there is little geographical variation between the regional Thai *Cng* isolates typed in this study. This genetic pattern is consistent with the one found in environmental *Cng* isolates from five geographic locations within India (Hiremath et al., 2008). Typed by MLST at only five loci, these 78 Asian isolates were all of molecular type VNI and also presented

with low genetic variation across 16 sequence types.

Statistically significant tests of non-random association of alleles at the different loci (I_A, $\bar{r}_d$ and PcP; Table 4.4) demonstrated an overwhelmingly clonal Thai *Cng* population structure. These results are consistent with previous studies showing that non-meiotic reproduction is the predominate mode of descent in *Cng* worldwide (Litvintseva et al., 2006; Bovers et al., 2008a; Litvintseva et al., 2005; Taylor et al., 1999; Buchanan and Murphy, 1998). Hiremath et al. (2008) also reported an over-representation of certain genotypes among the Indian isolates mentioned above, evidence of clonality within this population. Having said this, this same Indian population also presented with low I_A and $\bar{r}_d$ values, evidence of recombination despite all 78 isolates being of mating type MATα. Recent investigation of the predominance of the α mating type in nature led to the finding that cryptococcal strains of the same mating type within *Cn* (serotypes A and D) are capable of sexual reproduction in the form of haploid and monokaryotic fruiting, a process previously believed to be mitotic and asexual (Lin et al., 2005). It is important to note that if meiosis was occuring between isolates of identical genotype, neither of these techniques would pick it up. That is to say, the occurrence of recombination between genetically highly similar isolates of the same mating type cannot be rejected. Statistical test R_m was therefore applied to the *Cng* population in order to detect the minimum number of recombination events necessary to explain the distribution of polymorphisms within and between loci. The test detected a degree of minimum recombination events in the Thai population ($R_m = 5$; Table 4.2), as well as low levels of intergenic recombination (1/7 loci). It is important to note that the detected signal may well be ancestral or due to homoplastic polymorphisms within genetically diverse loci. There have been previous reports of recombination within predominantly clonal populations of *Cng* (Bui et al., 2008; Xu and Mitchell, 2003; Litvintseva et al., 2003), including within an environmental sample consisting of only MATα alleles in India (Hiremath et al., 2008), although the recombination detected may well be ancestral or quite old, and that the signal may still be detected.

Estimates of haplotypic diversity (*Hd*), mutation rates (θ) and nucleotide differences (π) were consistently low (Table 4.2). Litvintesva et al. (2011) report values of *Hd* and π at the seven MLST loci for a global population of

Cng. The lowest value of haplotypic diversity recorded for this population was at *SOD1* and the highest at *LAC1* (global $Hd = 0.03$ and 0.73, respectively), therefore there is consistently less diversity in the Thai population than in the global population at every locus. Tajima's D is a statistical test for the identification of loci that are evolving under non-random processes, such as selection or demographic expansion or contraction, and I showed that 6/7 MLST loci in the Thai population were significantly negative for Tajima's D, hence non-neutral in their behaviour. One possible explanation for the negative D values is selective sweep. When purifying selection occurs, segregating sites increase in number (high S). However, they tend to accumulate at silent sites and not become common, meaning heterozygosity remains low (small π) and D is negative (Tajima, 1989). As the MLST loci used to type *Cng* are mostly in housekeeping genes (Meyer et al., 2009) and therefore unlikely to be under strong selection, these differences in Tajima's D, as well as consistently negative for all loci, are most likely due to demographic effects of population expansion following a population bottleneck.

The Asian population's comparatively low genetic diversity, high linkage disequilibrium, non-neutral evolution and lack of geographically defined structure are all consistent with a model of a rapid population expansion from a limited set of ancestors. Despite being characterised by similar limited genetic variation, *Cng* isolates from Northwest India were also found to be recombining in nature (Hiremath et al., 2008). This leads us to question how accurately this snapshot of Thailand represents *Cng*'s global structure. Further molecular typing on a wider scale is required to increase our understanding of the organism's gene flow over significant geographic distances.

5. *Cng* in Cape Town

5.1. Introduction

Cryptococcal meningitis (CM) is now among the leading causes of adult meningitis in South Africa, as well as being responsible for 63% of all microbiological diagnoses (Jarvis et al., 2010). This is in marked contrast to the pre-HIV era when *Neisseria meningitis* was most frequently isolated (Liebowitz et al., 1984). However, this is in keeping with reports from Central and Southern Africa, where cryptococcosis accounts for between 27% and 45% of all cases (Gordon et al., 2000; Hakim et al., 2000; Bisson et al., 2008; Scarborough et al., 2007; Heyderman et al., 1998). A retrospective study found CM to be responsible for 37% of deaths among South African miners in 1998 (Corbett et al., 2002). Two studies on hospital fatalities among HIV+ CM patients in Durban and Soweto, South Africa, reported mortality rates of 64% and 42%, respectively (McCarthy et al., 2006; Moosa and Coovadia, 1997; Bergemann and Karstaedt, 1996), and despite the greater availability of antifungal treatments, these mortality rates remain unacceptably high (Department of Health, 2008).

Table 5.1.: Summary of infectious disease prevalence among HIV-positive patients from various studies in South Africa.

Study	n	MTB	Pneumonia	Meningitis	Enteritis	*M. kansasii*
Corbett[a]	1374	135 (10%)				9 (1%)
Wood[b]	1206	342 (28%)				
Colvin[c]	274	94 (34%)	39 (14%)	15 (5%)	18 (7%)	
Corbett[d]	599	128 (21%)	111 (19%)	33 (6%)	27 (5%)	6 (1%)

MTB = Mycobacterial Tuberculosis; *M. kansasii* = *Mycobacterium kansasii*; [a]Corbett et al., 2000; [b]Wood et al., 2000; [c]Colvin et al., 2001; [d]Corbett et al., 2002.

Genetic homogeneity has been reported among *Cng* isolates worldwide

and a single clade, VNI, appears to predominate (Litvintseva et al., 2003; Brandt et al., 1996; Boekhout et al., 2001; Meyer et al., 2003; Litvintseva et al., 2005; Currie et al., 1994). In comparison, the pattern of genetic diversity among the *Cng* isolates in Southern Africa is markedly different to that found elsewhere. Described as highly genetically diverse by Litvintseva et al. (2006 and 2011), it boasts the presence of molecular types VNII and VNB as well as VNI, and there has been evidence that *Cng* strains within Sub-Saharan Africa may be sexually recombining, a process which generates new genotypes while destroying haplotypes — haplotypes being those combinations of polymorphisms which are inherited together (Litvintseva et al., 2003). The combination of high HIV/AIDS prevalence and the presence of great genetic diversity make South Africa an ideal country for investigating the interplay between potentially evolving environmental *Cng* populations and clinical infections.

In this chapter, I describe a clinical sample of cryptococcal isolates from Cape Town, South Africa, characterized by MLST. In keeping with other studies in Southern Africa, this population shows great variation in terms of its genetic diversity, mating-type, haplotypes and hybrids; this is in direct contrast to the genetically highly homogeneous Thai population described in Chapter 4.

5.2. Aims

I aimed to describe the genetic structure of *Cng* isolated from a population of AIDS-associated CM patients in Cape Town, South Africa, using MLST. This population was then compared against a Thai population also typed by MLST (Chapter 4).

5.3. Materials and Methods

5.3.1. Isolates

One hundred and eighteen clinical cryptococcal isolates were obtained from two studies performed in Jooste Hospital, Cape Town. The first was a

prospective observational study of 54 AIDS patients which aimed to compare early fungicidal activity of various regimens in the initial treatment of HIV-associated CM (Bicanic et al., 2007). Conducted between February and September 2005, patients over 21 years of age with cultures positive for *Cn* were enrolled and assigned to one of two treatment arms: (i) amphotericin B (AmB) at 1 mg/kg/day for a week, followed by oral fluconazole at 400 mg/day or (ii) fluconazole at 400 mg/day for ten weeks. Following these ten weeks, all were switched to 200 mg/day of fluconazole and those who were ART-naive initiated on ART (77%). The second was a randomized control study carried out between May 2005 and June 2006, and consisting of 64 HIV+, ART-naive patients experiencing first-episode CM (Bicanic et al., 2008). This study compared the administration of 0.7 mg/kg per day of AmB (standard protocol) to that of 1 mg/kg per day of AmB for two weeks, followed by oral fluconazole.

In both studies, the principal outcome measure of the effect of treatment was the rate of clearance of infection, determined by measurements of CSF cryptococcal culture (early fungicidal activity, EFA). The patients of the two studies were comparable in terms of average age, weight and percentage of known HIV infection (Table 5.2). The second study's patients had a higher HIV load (150 000 versus 55 000 copies/mL) as well as lower CD4 cell count (38 versus 49×10^6 cells/L) on average, although a greater percentage of patients presented with abnormal mental status in the observational study (24% versus 13%), a characteristic associated with increased mortality (Brouwer et al., 2004; Bicanic et al., 2009; Saag et al., 1992; Robinson et al., 1999).

5.3.2. Molecular methods

The cultivation and extraction of DNA from the 118 isolates, as well as their sequencing was conducted as per the molecular methods described in subsection 2.1.1. Molecular analyses were applied to the sequence data both in the form of allelic profiles and actual nucleotide sequences (subsection 2.4.1). These included tests of nucleotide diversity, random evolution (Tajima's D, Ramos-Onsins & Rozas' R_2 test), linkage disequilibrium (I_A and $\bar{r}_d$) and recombination (R_m). The phylogenetic structure of the population was assessed by maximum composite likelihood (subsection 2.5). In

Table 5.2.: Characteristics of patients with HIV-associated cryptococcal meningitis (CM) from two studies in Cape Town.

Characteristic	Study 1 ($n = 54$)[a]	Study ($n = 64$)[b]
Number of men (%)	14 (26)	24 (38)
Mean age, years (range)	34 (29 – 39)	33 (28 – 38)
Mean weight, kg (s.d.)	55 (12)	54 (15)
Number of patients with known HIV infection at presentation (%)	46 (85)	54 (84)
Number of patients with abnormal mental status (%)	13 (24)	8 (13)
Mean CD4 cell count, 10^6 cells/L (range)	49 (21 – 71)	38 (12 – 69)
Mean HIV load, copies/mL (range)	55 000 (1 060 – 250 000)	150 000 (48 000 – 540 000)
Number of patients presenting with relapse of CM (%)	9 (17)	0 (0)
Number of patients receiving HAART (%)	18 (33)	0 (0)
Death by week 2, number of patients (%)	n/a	4 (6)
Death by week 10, number of patients (%)	19/52 (37)	15 (24)
Death by 1 year, number of patients (%)	33/51 (65)	n/a

[a]Bicanic et al., 2007; [b]Bicanic et al., 2008; s.d. = standard deviation.

addition to reference strains of known major molecular types of the *Cn/C. gattii* species complex and the genome-project strain H99, Botswanan isolates of molecular group VNB were also included (bt1, bt24, bt27, bt31, bt33, bt35, bt46, bt60, bt63, bt65, bt76, bt84, bt85, bt88, bt89, bt109, bt125, bt131, bt204, bt206; Litvintseva et al., 2003). Finally, the African population of *Cng* isolates was compared to that from Thailand, using PCA performed in adegenet, and AMOVA performed using GenAlEx. Statistics K^*_{ST} and S_{nn} were also reported, analyzing the divergence between the two populations (Chapter 2).

5.4. Ethics Statement

Ethical approval was required for both studies mentioned above. The first was approved by the Research Ethics Committee of the University of Cape Town (Cape Town, South Africa). The randomized control trial was approved by research ethics committees of the University of Cape Town and St. George's Hospital (London, UK), as well as the Medicines Control Council of South Africa (Pretoria), and was conducted in accordance with the principles of the Helsinki Declaration of, 1975, as revised in 1983. The study was registered as ISRCTN68133435 (step 1; `http://www.controlled-trials.com`).

5.5. Results

5.5.1. Mating, molecular, and sequence types of isolates

Of the 118 isolates cultured, two failed to grow (CT8 and CT11). 107 of the remaining 116 were *Cng* (serotype A; 92%), six were diploids, two were *C. gattii* and one remained undetermined. All but one of the 107 *Cng* isolates were of mating type MATα, with isolate CT15 being of MAT**a**.

Thirty-five STs were allocated to the Cape Town *Cng* isolates

Complete sequence data was obtained for 107 Cape Town *Cng* isolates (Table 5.3). Concatenated at all seven loci and 3 892 base pairs long, the

aligned sequences revealed 153 variable sites, 134 of which are parsimony informative sites and 19 singleton sites.

Table 5.3.: Allelic profiles of the 107 Cape Town *Cng* isolates typed in this study and their corresponding Sequence Types. Haplotypes within each locus column are differentiated by colour.

Name	VN group	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*	ST
R16	VNI	7	1	1	1	1	1	1	1
R24	VNI	7	1	1	1	1	1	1	1
R49	VNI	7	1	1	1	1	1	1	1
R46	VNI	7	1	1	7	1	1	2	2
R57	VNI	1	1	1	3	4	1	1	3
CT1	VNI	1	3	1	5	2	1	1	5
CT3	VNI	1	3	1	5	2	1	1	5
CT7	VNI	1	3	1	5	2	1	1	5
CT13	VNI	1	3	1	5	2	1	1	5
CT18	VNI	1	3	1	5	2	1	1	5
CT30	VNI	1	3	1	5	2	1	1	5
CT33	VNI	1	3	1	5	2	1	1	5
CT34	VNI	1	3	1	5	2	1	1	5
CT45	VNI	1	3	1	5	2	1	1	5
CT48	VNI	1	3	1	5	2	1	1	5
R4	VNI	1	3	1	5	2	1	1	5
R5	VNI	1	3	1	5	2	1	1	5
R29	VNI	1	3	1	5	2	1	1	5
R64	VNI	1	3	1	5	2	1	1	5
R6	VNI	1	1	1	4	2	1	5	6
CT54	VNI	7	1	1	2	1	1	2	23
R10	VNI	7	1	1	2	1	1	2	23
R12	VNI	7	1	1	2	1	1	2	23
R23	VNI	7	1	1	2	1	1	2	23
R51	VNI	7	1	1	2	1	1	2	23
R58	VNI	7	1	1	2	1	1	2	23
CT6	VNI	1	1	10	3	4	1	1	32
CT9	VNI	1	1	10	3	4	1	1	32
CT16	VNI	1	1	10	3	4	1	1	32
CT21	VNI	1	1	10	3	4	1	1	32
CT22	VNI	1	1	10	3	4	1	1	32
CT23	VNI	1	1	10	3	4	1	1	32
CT24	VNI	1	1	10	3	4	1	1	32
CT36	VNI	1	1	10	3	4	1	1	32
CT37	VNI	1	1	10	3	4	1	1	32
CT43	VNI	1	1	10	3	4	1	1	32
R38	VNI	1	1	10	3	4	1	1	32
R43	VNI	1	1	10	3	4	1	1	32
R47	VNI	1	1	10	3	4	1	1	32
R52	VNI	1	1	10	3	4	1	1	32
CT2	VNI	1	1	1	1	1	1	2	37

ST = sequence type; CT = isolates from Bicanic et al., 2007;
R = isolates from Bicanic et al., 2008; novel ATs are in bold.

(continued on next page)

Name	VN group	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*	ST
R56	VNI	1	1	1	1	1	1	2	37
CT4	VNI	7	5	1	3	3	1	1	51
CT25	VNI	7	5	1	3	3	1	1	51
CT26	VNI	7	5	1	3	3	1	1	51
CT29	VNI	7	5	1	3	3	1	1	51
CT35	VNI	7	5	1	3	3	1	1	51
CT39	VNI	7	5	1	3	3	1	1	51
CT46	VNI	7	5	1	3	3	1	1	51
R2	VNI	7	5	1	3	3	1	1	51
R3	VNI	7	5	1	3	3	1	1	51
R7	VNI	7	5	1	3	3	1	1	51
R8	VNI	7	5	1	3	3	1	1	51
R25	VNI	7	5	1	3	3	1	1	51
R31	VNI	7	5	1	3	3	1	1	51
R32	VNI	7	5	1	3	3	1	1	51
R33	VNI	7	5	1	3	3	1	1	51
R35	VNI	7	5	1	3	3	1	1	51
R41	VNI	7	5	1	3	3	1	1	51
R45	VNI	7	5	1	3	3	1	1	51
R53	VNI	7	5	1	3	3	1	1	51
R55	VNI	7	5	1	3	3	1	1	51
R62	VNI	7	5	1	3	3	1	1	51
R13	VNI	7	5	1	3	3	1	2	53
R15	VNI	7	5	1	3	3	1	2	53
R27	VNI	7	5	1	3	3	1	2	53
R17	VNI	1	1	10	3	4	1	2	54
R44	VNI	1	1	10	3	4	1	2	54
R61	VNI	1	1	10	3	4	1	2	54
R18	VNI	1	3	1	5	2	1	2	55
R19	VNI	1	3	1	5	2	1	2	55
R63	VNI	1	3	1	5	2	1	2	55
CT10	VNI	1	1	10	1	2	1	1	57
CT12	VNI	7	1	1	3	1	1	1	58
CT19	VNI	1	1	1	1	4	**17**	2	61
CT27	VNI	1	1	1	7	1	1	1	62
CT40	VNI	1	7	1	1	4	3	9	63
CT41	VNI	1	1	10	3	4	12	1	64
CT47	VNI	7	1	1	7	4	1	2	65
R50	VNI	7	1	1	1	1	1	2	74
R54	VNI	7	1	10	1	1	1	1	75
R60	VNI	7	1	1	7	1	**18**	2	76
R59	VNII	2	9	14	8	11	1	4	30
CT17	VNII	2	9	14	8	11	12	4	52
CT20	VNII	2	9	14	8	11	12	4	52
CT31	VNII	2	9	14	8	11	12	4	52
CT32	VNII	2	9	14	8	11	12	4	52

ST = sequence type; CT = isolates from Bicanic et al., 2007;
R = isolates from Bicanic et al., 2008; novel ATs are in bold.

(continued on next page)

Name	VN group	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*	ST
CT38	VNII	2	9	14	8	11	12	4	52
CT42	VNII	2	9	14	8	11	12	4	52
CT52	VNII	2	9	14	8	11	12	4	52
CT53	VNII	2	9	14	8	11	12	4	52
R9	VNII	2	9	14	8	11	12	4	52
R11	VNII	2	9	14	8	11	12	4	52
R20	VNII	2	9	14	8	11	12	4	52
R39	VNII	2	9	14	8	11	12	4	52
CT49	VNII	9	9	14	6	11	**16**	**16**	66
CT51	VNII	8	10	**28**	8	2	3	11	67
R1	VNII	2	1	14	8	11	12	4	68
R30	VNII	8	10	1	8	12	3	2	70
R34	VNII	2	9	14	8	11	12	2	71
R40	VNII	10	9	14	8	11	12	4	72
R48	VNII	8	10	**28**	8	12	3	11	73
R21	VNB	4	11	**26**	6	6	1	3	56
R26	VNB	4	11	**26**	6	6	1	3	56
CT14	VNB	1	**14**	10	**11**	4	1	**15**	59
CT15	VNB	**13**	1	**25**	**11**	4	**19**	1	60
R22	VNB	9	11	**27**	6	6	**15**	2	69

ST = sequence type; CT = isolates from Bicanic et al., 2007;
R = isolates from Bicanic et al., 2008; novel ATs are in bold.

The seven loci collectively yielded 54 allele types (ATs), 14 of which were novel to the South African population of *Cng* (Table 5.3; novel ATs are in bold). Nine of the 14 new ATs were within the isolates of molecular group VNB, three within VNII and two within VNI. *SOD1* and *IGS1* consisted of five and four new ATs, respectively, while no new ATs were found at *PLB1*, all having been previously described either by (Litvintseva et al., 2006), or in the Thai population (Table 4.1). Thirty-five multilocus STs were identified within the South African isolates. Twenty-seven percent of the Cape Town isolates shared STs with the Thai population: ST5 and 32 ($n = 14$ apiece) and ST6 ($n = 1$). Another 6 STs were shared with the global population (STs 1, 2, 3, 23, 30 and 37). The remaining 26 STs were allocated consecutive numbers 51 to 76 ($n = 64$; Table 5.3). ST51 contained 21 of the remaining 64 isolates (33%), ST52 12 (19%), STs 53, 54 and 55 three isolates each, and ST 56 two isolates. STs 57 to 76 all consisted of a single isolate.

5.5.2. Molecular analyses

Analyses of genetic variation and phylogeny reveal a genetically diverse Cape Town *Cng* population

Nucleotide diversity was estimated for both for the concatenated Cape Town isolate sequences, as well as at the seven individual loci (Table 5.4). The average estimate of haplotypic diversity (*Hd*) was 0.92. *IGS1* contained the greatest number of segregating sites ($S = 105$), average number of nucleotide differences per site ($\pi = 0.023$) and scaled mutation rate per site ($\theta = 0.099$), while *SOD1* had the lowest haplotypic diversity ($Hd = 0.39$). In six of the seven loci, Tajima's D values did not differ significantly from 0, as would be expected under the neutral theory. The one exception to this was at the *IGS1* locus which was significantly negative ($D = -2.53^{***}$; Table 5.4), with the remaining six loci yielding non-significant R_2 values.

Table 5.4.: Nucleotide diversity among the Cape Town *Cng* isolates ($n = 107$).

Locus	pb^a	S^b	h^c	Hd^d	π^e	θ^f	D^g	R_2^h	${R_m}^i$
CAP59	501	9	8	0.68	0.004	0.003	0.57	0.116	0
GPD1	489	11	7	0.71	0.006	0.004	0.80	0.125	0
IGS1	202	105	5	0.57	0.023	0.099	-2.53^{***}	0.022	0
LAC1	470	13	8	0.77	0.006	0.005	0.42	0.111	0
PLB1	533	12	7	0.82	0.006	0.004	1.13	0.136	0
SOD1	529	22	8	0.39	0.006	0.008	−0.65	0.077	3
URA5	636	15	9	0.69	0.005	0.005	0.03	0.096	0
		Average		0.92	0.006	0.005			$12^{\#}$

[a] total number of sites in alignments, excluding indels and missing data; [b] number of segregating sites; [c] number of haplotypes; [d] haplotypic diversity; [e] average number of nucleotide differences per site; [f] Watterson's estimate of the population-scaled mutation rate, expressed per site (Watterson, 1975); [g] Tajima's D (Tajima, 1989); [h] Ramos-Onsins & Rozas' R_2 (Ramos-Onsins and Rozas, 2002); [i] minimum number of recombination events (Hudson and Kaplan, 1985); [#] average R_m between all seven loci; $^{***}p < 0.001$.

The phylogenetic relationship between the Cape Town isolates is depicted in a neighbor-joining dendrogram (Figure 5.1). The resulting tree revealed the presence of all three molecular types of serotype A: VNI ($n = 82$; bootstrap = 51%), VNII ($n = 20$; bootstrap = 74%) and VNB isolates ($n = 5$).

Isolate CT15 of MAT**a** clustered with VNB isolates with 76% bootstrap support, while the four remaining VNB isolates CT14, R21, R26 and R22 did so with low bootstrap support (≤ 40; Figure 5.1). VNB and VNI clades also clustered together with high bootstrap support (93%) while the outliers of the *C. gattii* complex were grouped together (bootstrap = 100%).

Evidence of clonality and recombination detected within the Cape Town *Cng* population

The I_A and $\bar{r}_d$ statistics evaluating linkage disequilibrium were significantly greater than 0 ($p < 0.001$) for both the 107 African isolates (not shown) and the clone-corrected population which included only one representative of each ST ($n = 35$; Table 5.5). These results suggest rare or absent recombination (Burt et al., 2000; Smith et al., 1993) and lead to the rejection of the hypothesis of random mating, due to the existence of a significant percentage of phylogenetically compatible loci pairs (PcP = 0.19; p-value <0.001; Table 5.5). As the pooling of molecular groups VNI, VNII and VNB will always lead to significant I_A (because they will behave as separate species), independent calculations of I_A and $\bar{r}_d$ were also performed on the three clusters of isolates differentiated by phylogenetic analyses. These values were closer to zero for the VNI and VNB clone-corrected groups (p-value < 0.16; Table 5.5), meaning that the null hypothesis of recombination could not be rejected for the individual subgroups (Smith et al., 1993). Further analyses in DNASP revealed a minimum number of twelve recombination events across all seven loci within this population (Table 5.4); more than twice that found in Thailand ($R_m = 5$; Table 4.2).

Genetically differentiated Capetonian and Thai *Cng* populations

Thus, all analyses show that the typed African *Cng* population is considerably more genetically diverse than the Thai one described in Chapter 4 (Table 5.6). Differentiation between these two populations was further analysed using K^*_{ST} (the weighted measure of the ratio of the average pairwise differences within populations to the total average pairwise differences; Hudson et al., 1992) and S_{nn} (the nearest-neighbour statistic; Hudson and Kaplan, 1985; Hudson, 2000). Their values of 0.257 and 0.801, respectively, were

Figure 5.1.: Neighbour-joining tree illustrating the evolutionary relationships of the *Cng* isolates from Cape Town typed by MLST ($n = 107$). The evolutionary distances were computed using the maximum composite likelihood method in MEGA 4.0 and are in the units of the number of base substitutions per site. Reference strains of known major molecular types of the *C. neoformans*/*C. gattii* species complex are included: WM148 (serotype A, VNI), WM626 (serotype A, VNII), WM629 (serotype D, VNIV), WM179 (serotype B, VGI), WM178 (serotype B, VGII), WM175 (serotype B, VGIII), WM779 (serotype C, VGIV; Meyer et al., 2003) and the genome-project strain H99 (serotype A, VNI). In addition, isolates known to be of molecular group VNB are included. The Cape Town isolates cluster into three molecular groups of serotype A: VNI ($n = 82$), VNII ($n = 20$) and VNB isolates ($n = 5$). The percentage replicate trees in which the associated taxa clustered together in the bootstrap test (1 000 replicates) more than 75% of the time ($n \geq 75\%$) are indicated.

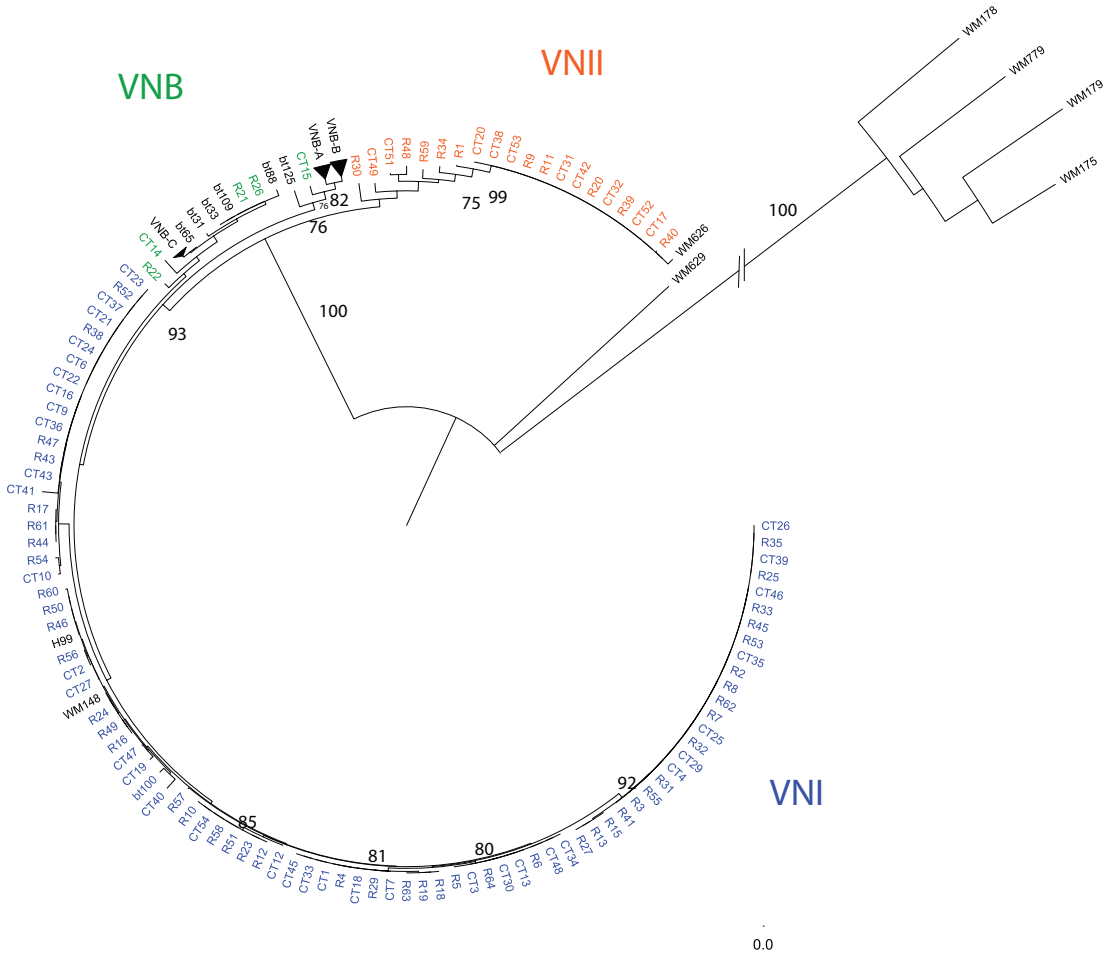


Table 5.5.: Multilocus linkage disequilibrium analyses for samples of *Cng* from Cape Town, South Africa ($n = 35$).

	Clone-corrected sample[a]		
	I_A[b]	$\bar{r}_d$ [c]	PcP[d]
Total Population	1.30***	0.22***	0.19***
VNI ($n = 22$)	0.13	0.02	0.48
VNII ($n = 9$)	2.01***	0.34***	1.00***
VNB ($n = 4$)	0.85	0.29	1.00

[a]excluding replicate haplotypes; [b]index of association; [c]scaled index of association (I_A) by the number of loci ($m - 1$); [d]percentage of phylogenetically compatible pairs (PcP) of loci; ***p <0.001.

found to be significant (p-value < 0.001), leading us to reject the null hypothesis of no differentiation between the populations of *Cng* in Southeast Asia and Africa.

Table 5.6.: Comparison of the average genetic diversity indices of the Thai and Cape Town *Cng* populations.

	Hd[a]	π[b]	θ[c]
Thailand	0.19	0.0008	0.005
Cape Town	0.92	0.0060	0.005

[a]haplotypic diversity; [b]average number of nucleotide differences per site; [c]Watterson's estimate of the population-scaled mutation rate, expressed per site.

PCA was applied to the allelic profiles of the two populations in order to investigate the underlying patterns of variability between them. The resulting topology illustrates that, despite population stratification between their genotypes at the first principal component, allelic profiles are shared between them (p-value = 0.01; Figure 5.2). This reflects the three STs shared between 27 Cape Town and 107 Thai isolates (STs 5, 6, and 32).

AMOVA reveals that 64% of the variation found in the collection of the 290 typed isolates was attributable to variation within the two geographically separate populations (Table 5.7), and principally reflects the high diversity within the Cape Town genotypes.

Figure 5.2.: Principal Component Analysis of the allelic profiles of the *Cng* genotypes from Cape Town and Thailand. Individual genotypes (dots) are linked by colored lines to form clusters which are summarized by colored ellipses proportional in size to the number of isolates represented. The isolates from Cape Town are in blue (group 1, $n = 107$) and the Thai isolates are in red (group 2, $n = 183$). p-value showing inter-population differentiation and eigenvalues are represented in the bar plot.

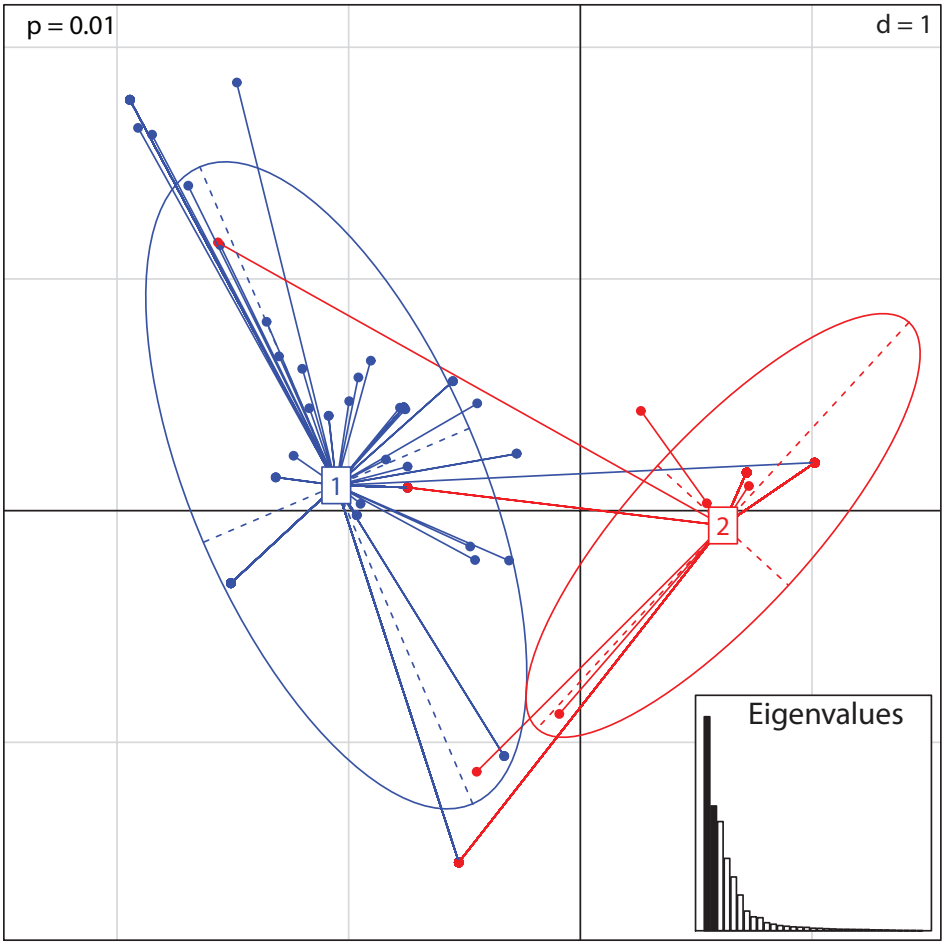


Table 5.7.: Summary of AMOVA of Cape Town and Thai *Cng* isolates (n = 290), based on the seven polymorphic loci and according to geographic origin.

Source	df	SS	Est. Var. (%)	ΦPT	p-value[a]
Among Population	1	98.9	0.73 (36)	0.365	0.001
Within Population	289	363.4	1.27 (64)		
Total	290	462.2	1.99 (100)		

[a] p-value estimates are based on 999 permutations.

5.6. Discussion

This chapter describes the genetic composition of clinical *Cng* isolates from Cape Town, South Africa. The data clearly show that that there is comparatively greater diversity in the Cape Town population than that found in the Thai *Cng* population described in Chapter 4.

MLST identified 35 STs among the 107 Cape Town isolates versus the 10 found in the Thai population (Table 5.3). The African population's STs were more evenly distributed, with no two STs accounting for more than 33% of the population. This was unlike the Thai population where two STs (ST4 and 6) accounted for 81% of the whole population — STs which differed by only four nucleotides (Table 4.1). Three frequently isolated STs found in the Thai population (STs 6, 5 and 32; n = 78, 27 and 3 respectively) were also identified within the Cape Town population, two of which predominated within the latter (ST5 and 32; 14 isolates each).

Diversity indices were consistently higher within the Cape Town sample than within the Thai population (Hd = 0.92 vs. 0.19, π = 0.006 vs. 0.001, respectively; Table 5.6), reflecting the higher number of haplotypes found in the former (range h = 7 to 9 versus 2 to 6, respectively; Tables 4.2 and 5.4). In both populations, *CAP59* contained the lowest number of segregating sites (S = 5 and 9) and the lowest average number of nucleotide differences per site (π = 0.001 and 0.004). The most diverse locus was *IGS1* in the African population and *LAC1* in the Thai population (Tables 4.2 and 5.4).

In terms of population growth within the Cape Town population, there was only one significant D value at *IGS1* (Table 5.4) as estimated by

Tajima's D. This meant that the null hypothesis of neutrality could not be rejected. This differs from the Thai population, in which only a single locus returned a non-significant value, and whose values were further from 0 and predominantly negative (*GPD1*; Table 4.2). These findings of high genetic variation and the suggestion of neutral evolution in this African *Cng* population is similar to that seen in the cyanobacterium *Microcystis aeruginosa*. This alga is commonly associated with warm, nutrient-enriched water, and its toxic blooms are a serious public health risk (Figueiredo et al., 2004). Contamination of drinking water by *M. aeruginosa* has resulted in liver damage, colorectal cancer, and death among humans, (Ueno et al., 1996; Falconer and Humpage, 1996; Jochimsen et al., 1998; Zhou et al., 2002). Tanabe et al. (2007) typed 164 *M. aeruginosa* strains and reported 79 STs and a high level of genetic diversity between them ($Hd = 0.95$), levels comparable to those found in this Cape Town *Cng* population ($Hd = 0.92$). It is important to note that Tanabe et al. (2007) also identified clone-corrected linkage equilibrium indicative of significant deviation from the null model of panmixia. The recombination within a clonal population structure found in *M. aeruginosa* is also identified in this African *Cng* population, and is further addressed later in this discussion.

In addition to more evenly distributed STs, there are more evenly distributed molecular types in the South African population. As expected, there was a predominance of serotype A, molecular type VNI (77%) among the clinical AIDS population from Cape Town, as is common worldwide (Meyer et al., 2003) and within our Thai samples. The Cape Town population also contained VNII ($n = 20$) and VNB isolates ($n = 5$), both less common molecular groups of serotype A, and only one of which was identified in the Thai population (CM21 of molecular type VNII; Table 4.1). This marked heterogeneity mirrors findings of molecular diversity within African *Cng* isolates, including a highly genetically diverse, area-specific and recombining population of VNB genotypes from neighbouring country Botswana (Litvintseva et al., 2006). Previous studies have also reported the global occurrence of VNB *Cn* Aα (also known as AFLP genotype 1A) both clinically and environmentally (Barreto de Oliveira et al., 2004; Boekhout et al., 2001; Bovers et al., 2008; Ngamskulrungroj et al., 2009). The origin of VNB is hypothesised to be the result of hybridisation between VNI (*Cng*; serotype A;

ALFP genotype 1) and VNIV (Cn var *neoformans*; serotype D; AFLP genotype 2; Boekhout et al., 2001; Litvintseva et al., 2006) — the two molecular groups which most isolates recovered from AIDS patients belong to (Meyer et al., 2003). This observation suggests a more recent divergence between VNB and VNI versus VNII, and is supported by the significant clustering of the VNI and VNB clades in the neighbor-joining tree of Cape Town *Cng* isolates (Figure 5.1; high bootstrap support = 93%), as well as by divergence estimates calculated using Bayesian inference reported and discussed in Chapter 6. Although no VNIV isolates were detected in this sample, six diploids were identified, evidenced by the presence of two bases at a single nucleotide position in sequence traces at at least one allele. These were isolates CT5, CT44, CT50 from observational study (Bicanic et al., 2007) and R14, R28 and R36 from the randomised control (Bicanic et al., 2008). As their allelic profiles were incomplete, the identity of the parental genotypes of these diploids remained undetermined. However, three of these six diploids consisted of (i) alleles which BLAST to VNIV and VNIII molecular types (serotypes D and AD hybrids respectively; QC and MI $\geq$98%), and (ii) at least one allele of molecular group VNI (isolates CT5, CT44 and R36).

Sex, although costly, is a major mechanism in generating progenies with diverse phenotypes, a mechanism which is important for the survival of species in adverse environments (Heitman, 2006; Wang and Lin, 2011). *Cn*, a largely unisexual population, undergoes genetic exchange through sexual monokaryotic fruiting as well as same-sex mating (Lin et al., 2005; Lin et al., 2007). The African population typed in this study supports hypotheses for both clonal expansion and recombination, similar to *M. aeruginosa* as described above (Tanabe et al., 2007). Clonality in the Cape Town *Cng* population is characterised by over-represented genotypes (Tibayrenc et al., 1991) as seen in the Thai population, where two almost identical STs represented over 80% of the 183 isolates. Similarly, significant clonality was identified in the Cape Town population as a whole, although clone-corrected I_A and values of the entire sample were closer to zero, indicating that this population was closer to being freely recombining than the Thai one (Tables 4.4 and 5.5). The high genetic diversity in the Cape Town sample — as shown by a greater number of STs, and the occurrence of all three serotype

A molecular types and hybrids — is in support of recombination. Statistical analysis of recombination and clone-corrected estimates of linkage disequilibrium according to molecular group further support the hypothesis of recombination within the Capetonian population (Table 5.5). A greater number of minimum recombination events was identified in this *Cng* population than that of the Thai isolates ($R_m = 12$ versus 5; Tables 5.4 and 4.2). Although there have previously been reports of sexual propagation within both clinical (Litvintseva et al., 2003) and environmental African isolates of *Cng* (Lin et al., 2009), several fungi exhibit both clonal and sexual reproduction, the former being preferred under uniform growth conditions while the latter is (theoretically at least) of benefit in changing environments (Burt et al., 2000; Taylor et al., 1999b; Xu and Mitchell, 2002; Litvintseva et al., 2003). In this South African population, same-sex mating may explain at least some of the recombination rates discovered, and, as there is more genetic diversity than that found in the Thai population, this is manifested as higher R_m estimate.

The preponderance of MATα is not unexpected (Kwon-Chung, Edman and Wickes, 1992), but the presence of a single MAT**a** isolate in the African population (CT15) sheds light on the apparent recombination that was observed. A similar bias was described by Litvintseva et al. (2011), where MAT**a** isolates made up only 4% of 273 global environmental isolates of *Cng*. Most interestingly, the ten isolates in question were all of Southern African origin. The presence of both mating types in the environment provides the opportunity for outcrossing, which would have the effect of increasing the observed R_m in my dataset. However, were substantial sexual reproduction taking place, one would expect more balanced mating types, especially since both mating types have been shown to be equally virulent in mice (Nielsen et al., 2003; Taylor et al., 1999a). A potential explanation for this bias is spatial isolation. All of the ten MAT**a** *Cng* strains were isolated from endemic Southern African trees in rural areas — none from avian excreta commonly found in urban areas (Litvintseva et al., 2001). The existence of a physical and ecological barrier could restrict the genetic exchange inter-strains between these two sources of *Cng*. Another possible explanation for this pattern is that the MAT**a** isolates are the remnants of a previously but no longer recombining population (Taylor et al., 2000), MATα strains being

proven to be more virulent than congenic MAT**a** strains in the murine tail vein injection model and thus thrive in human host from which the clinical samples are isolated (Kwon-Chung et al., 1992), or the influence of other factors such as haploid fruiting among the MATα strains (Wickes et al., 1996). MAT**a** cells are unable to undergo haploid fruiting as they lack STE12, a transcription factor required for filamentation and sporulation (Yue et al., 1999). Recent work investigating the role of a G-protein β subunit (*GPB1*) in *Cn* revealed that MAT**a** cells stimulate monokaryotic fruiting in MATα cells without the need for cell-cell or filament-filament contact (Figure 1.3; Wang et al., 2000). The former secrete a peptide mating pheromone which stimulates the differentiation pathway in the latter. This resulting rate of fruiting is considerably greater than that occurring in isolated *Cn* MATα cells, as well as MATα cells starved of nitrogen and in the absence of a mating partner (Wickes et al., 1997; Wang et al., 2000; Wickes et al., 1996). Wang et al. (2000) propose that haploid fruiting functions as a "prelude to mating" in this organism that manifests widespread unbalanced mating types. This means that even a single MAT**a** within a population can considerably upregulate haploid fruiting among MATα cells. Botts and Hull (2011) propose that the main agents of cryptococcosis are basidiospores, as is the case in various human fungal pathogens including *Coccidioides immitis* (Converse and Reed, 1966) and *Bacillus anthracis* (Casadevall and Pirofski, 2006). As these are produced by monokaryotic fruiting, the increased number of α spores in relation to **a** spores resulting in MATα bias is expected in the recombining Cape Town population. Mating pheromone signaling has also been shown to drive homothallic mating in *C. albicans*, as well as filamentous differentiation of haploid *S. cervisiae* cells, suggesting that pheromone induced filamentation and mating have remained conserved in divergent ascomycetous and basidiomycetous yeasts (Roberts et al., 2000; Erdman and Snyder, 2001; Wang et al., 2000).

The phylogenetic structure of the Cape Town population is further deepened by the presence of two *C. gattii* isolates: CT28 and R42, confirmed by sequences blasted in NBCI ($\geq$99% concordance). Although rare among immunocompromised patients, *C. gattii* causing disease within this population is not unheard of. There have been cases of *C. gattii* meningitis diagnoses in HIV+ patients in Mexico, Australasia, Thailand, Zaire and South Africa

(Castanon-Olivares et al., 2000; Poonwan et al., 1997; Kapenda et al., 1987; Karstaedt et al., 2002). These findings challenge our initial understanding of the infectious characteristics of the pathogen and further work is needed to understand the different host pathogen dynamics manifested by *Cng*.

The presence and frequency of recombination in this natural population has the potential to affect the incidence of clinical disease among humans via evolution of the pathogen (Klugman, 2003; Stefani and Agodi, 2000; Litvintseva et al., 2003; Fraser et al., 2005). The potential impact on public health of this adaptable pathogen is of particular importance in South Africa, where the AIDS epidemic has resulted in a high vulnerable population. The belief that increased coverage of ARV availability will lead to a decrease in opportunistic infections is also being challenged, with the opposite being reported in developing countries including South Africa, where HIV-associated CM mortality rates remain unacceptably high (McCarthy, 2007; Sloan et al., 2009a; Sloan et al., 2008; Sloan et al, 2009b; Park et al., 2009).

6. Global *Cng*

6.1. Introduction

Found worldwide, *Cng* accounts for the majority of all cryptococcal isolates found both in the clinical and environmental context (Kwon-Chung, 1992). Its increasing importance as an opportunistic fungal infection is evidenced by high mortality rates among immunocompromised patients (Park et al., 2009). In this chapter, I integrate typing data presented in the previous two chapters with pre-existing data comprising our understanding of the current global *Cng* population genetic structure. This is the first global analysis of the distribution of multilocus sequence typing (MLST) polymorphisms undertaken for this pathogen under the consensus scheme proposed by the ISHAM (Simwami et al., 2011). These findings contribute to our understanding of the patterns and processes underlying the generation of inter- and intra-varietal genetic variation and the evolution of global populations of *Cng*.

I will show that, other than the genetically diverse and highly recombining population in southern Africa for which 72% of all identified global STs were unique, genetic diversity in *Cng* does not significantly subdivide at the continental level. Non-African isolates of *Cng* on the whole consist of recently derived, and genetically bottlenecked, samples of 'African-like' alleles. These patterns of diversity, when combined with evidence from haplotype networks and coalescent analyses of global populations, are highly suggestive of an expansion of the *Cng* VNI clade out of Africa, leading to the recent founding of a limited number of genotypes found in Asia and North America. These findings show the power associated with the collection of global sequence databases in order to better understand the structure and evolution of major fungal pathogens.

6.2. Aim

The aim of this chapter is to describe the genetic structure of the current global collection of *Cng*, integrating the Thai and Cape Town *Cng* populations typed in this study.

My specific goals were:

1. To integrate the Thai and Cape Town *Cng* populations into the current global population;
2. To describe the genetic structure within and between geographically divided subpopulations of the global population of *Cng*;
3. To describe the divergence and expansion of the population of *Cng*;
4. To test the hypothesis of an African origin of *Cng*.

6.3. Materials and Methods

6.3.1. Global *Cng* isolates

The most current global MLST dataset as compiled by A. Litvintseva (Duke University, USA) consisted of 79 isolates from 13 countries whose genotypes and molecular groups had been previously determined by both amplified fragment length polymorphisms (AFLP) and MLST (Litvintseva et al., 2006). Fifty-seven of these genotypes were of clinical origin and eleven from pigeon excrement (11 unknown; Table 6.1). Ten percent of the clinical isolates were from HIV– patients, one of which had cancer. In terms of phylogeny, 51 isolates were of molecular group VNI, 20 of VNB and eight of VNII, with ten MAT**a** isolates, all of which were of African origin. The Thai and Cape Town isolates described in Chapters 4 and 5 were integrated to this population as follows:

Table 6.1.: Summary of the global isolates used in this study ($n = 79$).

Isolate	ST	Origin	Source	Year	Mating type	VN type
A3 38-20	1	N. Carolina	Pigeon excreta	2002[a]	α	VNI
A1	2	N. Carolina	Pigeon excreta	2002[a]	α	VNI
A1 35-8	2	N. Carolina	Pigeon excreta	2002[a]	α	VNI
arg1366	2	Argentina			α	VNI
arg1373	2	Argentina			α	VNI
c23	2	N. Carolina	BAL/HIV–	2001[a]	α	VNI
h99	2	N. Carolina		[b]	α	VNI
mal 120	2	Malawi	Blood/AIDS	1999[c]	α	VNI
Tn470	2	Tanzania	Blood/HIV+	1995[d]	α	VNI
CA 84-14	2	California	Pigeon excreta	2003[a]	α	VNI
JH125.91	3	Tanzania		[e]	**a**	VNI
th84	4	Thailand	Blood/HIV+	1997[f]	α	VNI
ug2463	4	Uganda	CSF/HIV+	2001[g]	α	VNI
A5 35-17	5	N. Carolina	Pigeon excreta	[a]	α	VNI
c48	5	N. Carolina	BAL/HIV–	2001[a]	α	VNI
c8	5	N. Carolina	CSF/HIV+	[a]	α	VNI
it743	5	Italy		[h]	α	VNI
jp1086	5	Japan	Human lung	1999[i]	α	VNI
jp1088	5	Japan	Human lung	1999[i]	α	VNI
th104	6	Thailand	Blood/HIV+	1997[f]	α	VNI
th206	6	Thailand	Blood/HIV+	1997[d]	α	VNI
Tn148	6	Tanzania	Blood/HIV+	1995[d]	α	VNI
bt130	13	Botswana	CSF/AIDS	2001[j]	**a**	VNI
bt68	13	Botswana	CSF/AIDS	2000[j]	α	VNI
A2 102-5	15	Texas	Pigeon excreta	2003[a]	α	VNI
A2 28-23	15	N. Carolina	Pigeon excreta	2003[a]	α	VNI
bt150	21	Botswana	CSF/AIDS	2001[j]	α	VNI
bt104	22	Botswana	CSF/AIDS	2001[j]	α	VNI
A3 1-1	23	N. Carolina	Pigeon excreta	2002[a]	α	VNI
it754	23	Italy		[h]	α	VNI
bt100	24	Botswana	CSF/AIDS	2001[j]	α	VNI
bt134	25	Botswana	CSF/AIDS	2001[j]	α	VNI
bt15	26	Botswana	CSF/AIDS	2000[j]	α	VNI
c26	31	N. Carolina	Blood/HIV+	2001[a]	α	VNI
br2362	32	Brazil		[k]	α	VNI
br794	32	Brazil		1998[k]	α	VNI
br795	32	Brazil		1998[k]	α	VNI
c27	32	N. Carolina	CSF/cancer	2001[a]	α	VNI
in2629	32	India	CSF/AIDS	2001[l]	α	VNI
in2632	32	India	CSF/AIDS	2001[l]	α	VNI

[a]Litvintseva et al., 2005; [b]Perfect et al., 1980; [c]Bell et al., 2001; [d]Archibald et al., 1998; [e]Lengeler et al., 2005; [f]Archibald et al., 1999; [g]S. A. Messer (University of Iowa); [h]Barchiesi et al., 1997; [i]Shigefumi Maesaki (Nagasaki University); [j]Litvintseva at al., 2003; [k]M. G. De Almeida (Universidad de Sao Paulo); [l]H. C. Gugnani (University of Delhi); [m]N. Myers (CDC); [n]Nielsen et al., 2003.

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
Tn10	32	Tanzania	Blood/HIV+	[d]	α	VNI
ug2458	32	Uganda	CSF/HIV+	2001[g]	α	VNI
za1345	32	DRC (Zaire)		[m]	α	VNI
za1346	32	DRC (Zaire)	CSF/HIV+	[m]	α	VNI
bt121	34	Botswana	CSF/AIDS	2001[j]	α	VNI
bt9	36	Botswana	CSF/AIDS	1999[j]	α	VNI
A4 1-12	37	N. Carolina	Pigeon excreta	2002[a]	α	VNI
in2637	38	India	CSF/HIV+	[l]	α	VNI
mal 9	38	Malawi	Blood/AIDS	[c]	α	VNI
ug2471	38	Uganda	CSF/HIV+	2001[g]	α	VNI
A4 34-6	39	N. Carolina	Pigeon excreta	2003[a]	α	VNI
ug2467	44	Uganda	CSF/HIV+	2001[g]	α	VNI
c12	30	N. Carolina	Lung/HIV–	[a]	α	VNII
c16	30	N. Carolina	Sputum/HIV–	2001[a]	α	VNII
JH8-1	30	N. Carolina		[n]	α	VNII
ug2472	40	Uganda	CSF/HIV+	2001[g]	α	VNII
c45	41	N. Carolina	Sputum/HIV–	2001[a]	α	VNII
c2	42	N. Carolina	BAL/HIV–	2002[a]	α	VNII
c44	42	N. Carolina	CSF/HIV–	2002[a]	α	VNII
A7	43	N. Carolina	Pigeon excreta	[a]	α	VNII
bt1	7	Botswana	CSF/AIDS	1999[j]	α	VNB
bt31	8	Botswana	CSF/AIDS	2000[j]	α	VNB
bt109	9	Botswana	CSF/AIDS	2001[j]	α	VNB
bt33	9	Botswana	CSF/AIDS	2000[j]	α	VNB
bt65	10	Botswana	CSF/AIDS	2000[j]	**a**	VNB
bt76	11	Botswana	CSF/AIDS	2000[j]	α	VNB
bt131	12	Botswana	CSF/AIDS	2001[j]	**a**	VNB
bt24	14	Botswana	CSF/AIDS	2000	**a**	VNB
bt206	16	Botswana	CSF/AIDS	2002[j]	**a**	VNB
bt89	16	Botswana	CSF/AIDS	2001[j]	α	VNB
bt84	17	Botswana	CSF/AIDS	2001[j]	α	VNB
bt35	18	Botswana	CSF/AIDS	2000[j]	α	VNB
bt46	19	Botswana	CSF/AIDS	2000[j]	α	VNB
bt63	20	Botswana	CSF/AIDS	2000[j]	**a**	VNB
bt85	27	Botswana	CSF/AIDS	2001[j]	**a**	VNB
bt88	28	Botswana	CSF/AIDS	2001[j]	**a**	VNB
bt204	29	Botswana	CSF/AIDS	2002[j]	**a**	VNB
bt60	33	Botswana	CSF/AIDS	2000[j]	α	VNB
bt27	35	Botswana	CSF/AIDS	2000[j]	α	VNB

[a]Litvintseva et al., 2005; [b]Perfect et al., 1980; [c]Bell et al., 2001; [d]Archibald et al., 1998; [e]Lengeler et al., 2005; [f]Archibald et al., 1999; [g]S. A. Messer (University of Iowa); [h]Barchiesi et al., 1997; [i]Shigefumi Maesaki (Nagasaki University); [j]Litvintseva at al., 2003; [k]M. G. De Almeida (Universidad de Sao Paulo); [l]H. C. Gugnani (University of Delhi); [m]N. Myers (CDC); [n]Nielsen et al., 2003.

Population structure of the wider Asian population of *Cng*

Three isolates from Litvintseva's previously typed *Cng* population originated from HIV-positive patients in Bangkok, Thailand (Archibald et al., 1999; Litvintseva et al., 2006), and were of ST6 (th104, th206) and ST4 (th84; Table 6.1). Seventy of the newly typed Thai isolates were also of ST4 and 78 of ST6 (Chapter 4). A further five isolates included in this study are of Asian origin from the global dataset: jp1086, jp1088 from Japan, and in2629, in2632 and in2637 from India (Table 6.1). Additionally, isolate J1 from Japan — typed alongside the Thai isolates but excluded from analyses due to origin — is included here. All three Japanese isolates were of ST5, while two Indian isolates, in2629 and in2632, were of ST32; also found within the typed Thai population. The third Indian isolate, in2637, was of ST38 and differs from ST32 by a single nucleotide at the *PLB1* locus. Results stemming from an analysis of molecular variance (AMOVA) of the allelic profiles of the 183 Thai isolates typed in this study and the nine isolates of wider Asian origin (global) supported this overlap in STs with only 19% of the variation being attributable to differences between the two (among-population variance = 0.161, $\Phi PT = 0.18$, p-value= 0.01; Table 6.2). These eight previously typed isolates of Asian origin were combined with the 183 Thai isolates typed in this study to form the Asian population ($n = 192$).

Population Structure of the wider African population of *Cng*

Forty-one isolates of African origin collected between 1991 and 2002 were used in the global analyses and had previously been described by Litvintseva et al. (2006). All of these are clinical isolates from HIV/AIDS patients, 68% of which originated from Botswana ($n = 28$), 12% from Uganda ($n = 5$) while four were from Tanzania, two from both Malawi and Zaire. MLST revealed 44 STs made up of 90 ATs (Table 6.1). Contributing 17 genotypes to this genetically variable African population is a distinct sample of molecular group VNB isolates from Botswana ($n = 19$; Litvintseva et al., 2006). Ten isolates are MAT**a**, one from Tanzania and the remaining nine from Botswana, and all but two being of molecular group VNB (the remaining of VNI). AMOVA established that only 17% of the variation of the African population is attributable to the variation between the allelic

Table 6.2.: Summary of AMOVA of the global *Cng* isolates, based on the seven polymorphic loci and according to geographical origin.

Asian population: Newly typed ($n = 184$), Global ($n = 8$)					
	d.f.	Sum of squares	Variance components (%)	ΦPT	p-value[a]
Among	1	3	0.16 (19)	0.18	0.001
Within	190	131	0.69 (81)		
Total	191	133	0.85 (100)		

African population: Newly typed ($n = 107$), Global ($n = 41$)					
	d.f.	Sum of squares	Variance components (%)	ΦPT	p-value[a]
Among	1	31	0.48 (17)	0.17	0.001
Within	146	341	2.33 (83)		
Total	147	372	2.84 (100)		

Global population[b]: Africa ($n = 148$), Asia ($n = 192$), North America ($n = 23$), South America ($n = 5$)					
	d.f.	Sum of squares	Variance components (%)	ΦPT	p-value[a]
Among	3	135	0.62 (29)	0.29	0.01
Within	364	563	1.55 (71)		
Total	367	698	2.18 (100)		

[a] p-value estimates are based on 999 permutations; [b] Europe was excluded due to small sample size ($n = 2$).

profiles of the Cape Town sample ($n = 107$) and the global African sample ($n = 41$; Table 6.2). The two were combined to create the African population ($n = 148$).

Global *Cng* population

All 370 *Cng* isolates were grouped according to geographic origin: Asia ($n = 192$), Africa ($n = 148$), North America ($n = 23$), South America ($n = 5$), Europe ($n = 2$). The MLST scheme yielded 76 STs from 338 clinical, 21 environmental isolates and 11 of unknown source, from 14 countries worldwide (Litvintseva et al., 2006; Meyer et al., 2009; Simwami et al., 2011; Table B.1).

6.3.2. Methods

The structure of the global *Cng* population was analysed using allelic profiles (PCA, AMOVA, linkage disequilibrium) and the raw nucleotide sequences of the isolates (phylogenetic analyses, diversity indices). eBURST and haplotype networks were used to infer founding and ancestral STs within the global population, while coalescent analysis estimated the divergence between the isolates, subdivided into geographically defined subpopulations. The analytical techniques and software used in this chapter are described in Chapter 2.

6.4. Results

6.4.1. Structure of the global population of *Cng* isolates based on allelic profiles

Seventy-one percent of the variation in the global population of *Cng* was attributed to differences within the five geographically defined subpopulations ($\Phi PT = 0.29$, $p = 0.01$; Table 6.2). When assessed among just the global VNI isolates, this percentage of variation attributed to within subpopulation difference is reduced to 66%, but remains significant (data not shown). Europe was excluded due to its small sample size ($n = 2$). Inter-class PCA is a technique which maximises the variance between predefined groups.

Performed on the global samples' allelic profiles, it distinguished the Asian population (first principal coordinate, pink ellipse, group 2) from the rest of the global population subsets (Africa, North and South America; $p < 0.01$, Figure 6.1; R code C.2). There was considerably more overlap between the three remaining subpopulations' ellipses, as depicted by the second principal coordinate. Africa (green ellipse, group 1) and North America (blue ellipse, group 3) shared considerably more individual genotypes (represented by coloured dots) than either shared with Asia. This reflects the ten genotypes which are shared between the continents: three between Africa and Asia, three between Africa and North America and another four shared between either three or four continents (Table 6.3). The two genotypes of South American origin (ST2 and ST23) were connected by a black line (group 4). The African population's 95% ellipse is the largest. This is in keeping with the fact that 54 of the 76 genotypes isolated in the global population were unique to Africa. PCA was was replicated for the global VNI isolates only, the topology of which remained unchanged relative to that of the global isolates of all three molecular groups (Figure B.1).

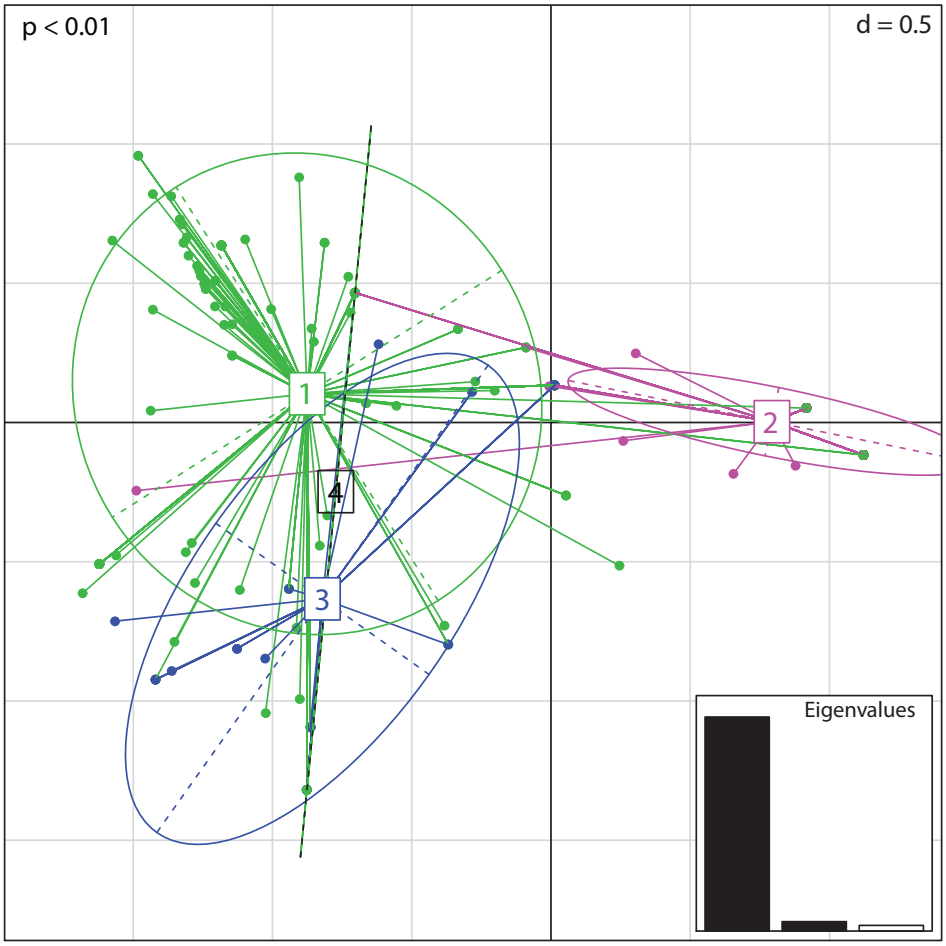
Table 6.3.: Sequence Types (STs) shared between two or more continents.

Continents	Shared STs
Africa, Asia	4, 6, 38
Africa, N. America	1, 30, 37
Africa, N. America, S. America	2
Africa, Asia, N. America, S. America	32
Africa, Europe, N. America	23
Africa, Asia, Europe, N. America	5

6.4.2. Phylogenetic analyses revealed three molecular groups within the global *Cng* population

The relationships between all 370 global *Cng* isolates were inferred by Maximum Composite Likelihood, with 1 000 replicate trees (subsection 2.5). Depicted in a dendrogram, the geographical origins of the isolates are represented by coloured rectangles: green = Africa ($n = 148$), red = Asia ($n = 192$), light blue = North America ($n = 23$), purple = South America ($n = 5$) and yellow = Europe ($n = 2$). Reference strains of known major

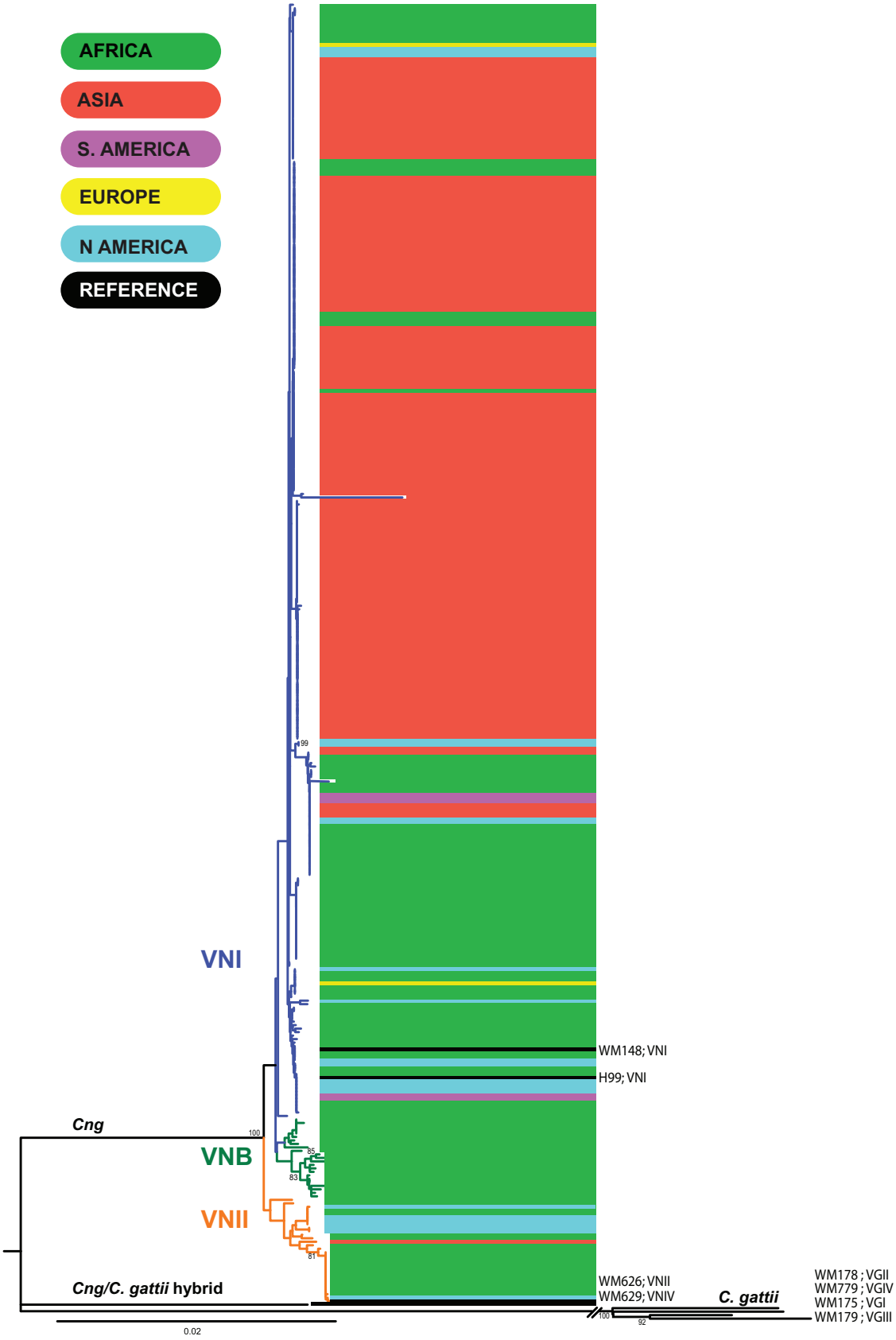
Figure 6.1.: Principal Component Analysis of the allelic profiles of the global *Cng* genotypes used in this study ($n = 368$), according to geographic origin. Individual genotypes (dots) are linked by coloured lines to form clusters which are summarised by coloured ellipses proportional in size to the number of isolates represented. The four groups depicted are numbered and defined according to continent: Africa = group 1 (green ellipse, $n = 148$), Asia = group 2 (pink ellipse, $n = 192$), North America = group 3 (blue ellipse, $n = 23$) and South America = group 4 (black line, $n = 5$). Europe was excluded due to small sample size. p-value showing inter-continental differentiation is shown and eigenvalues are represented in the bar plot.



molecular types of the *Cn* species complex were included in the analysis: VNI = WM148, H99; VNII = WM626; VNB isolates from Litvintseva et. al. (2003); and VNIV = WM629. Reference strains of the *C. gattii* complex were also included and served as an outgroup: VGI = WM175; VGII = WM178; VNGIII = WM179; VGIV = WM779. These reference strains are represented by black rectangles.

The phylogenetic analysis using these reference strains delineated three major groups within the global population: VNI (n = 317; type isolates WM148, H99), VNII (n = 29; type isolates WM626) and VNB (n = 24; Figure 6.2). VNII isolates clustered with bootstrap support of 69% while VNI and VNB groups grouped together with similar bootstrap support (70%). All seven South American and European isolates were of molecular group VNI, along with all but one Asian isolate (CM21) and 63% of North American isolates (n = 12). The remaining Asian and North American isolates (n = 7 and 1, respectively) were of molecular group VNII. The African isolates also reflected this trend of predominance of VNI (n = 105, 71% of isolates). The genotypes of these African VNI isolates exhibited no spatial differentiation between continental regions (AMOVA, $p-$value= 0.27; data not shown). In addition, molecular group VNB was only found in neighbouring African countries Botswana (n = 19) and South Africa (n = 4), and consisted of three subpopulations: VNB-A, VNB-B and VNB-C. Eight of the ten African isolates of the rare mating type MAT**a** were of this molecular group. Outliers WM178, 779, 175 and 179 of the *C. gattii* species clustered significantly with strong bootstrap support (100%), while WM629, comprising the variety *Cn* var *neoformans*, was isolated (Figure 6.2).

Figure 6.2.: Neighbour-joining tree inferring the evolutionary relationships of the global *Cng* isolates included in this study ($n = 370$). Significant bootstrap percentages ($n \geq 80\%$) are also displayed.



6.4.3. The African population exhibits the most genetic diversity of the global *Cng* subpopulations

The average nucleotide diversity within geographically defined subpopulations was calculated at each locus, and overall statistical tests included the number of segregating sites (S) and haplotypes (h), haplotypic diversity (Hd), the number of nucleotide differences per site (π) and Watterson's estimate of the population-scaled mutation rate (θ). Consistently higher average values of Hd, π and of θ indicated higher levels of within-population variation among the African isolates than were observed in the Asian and South American populations (Table 6.4). Similarly, the North American population's average values of Hd (0.93) and π (0.005) were lower than those of Africa (0.95 and 0.009, respectively; Table 6.4) although less so than the Asian population's, due to the greater proportion of VNII isolates (Table B.1).

Table 6.4.: Nucleotide diversity among the global *Cng* isolates according to geographic origin.

	Locus	pb^a	S^b	h^c	Hd^d	π^e	θ^f	D^g	${R_2}^h$	${R_m}^i$
Asia ($n = 192$)	*CAP59*	483	5	2	0.01	0.000	0.002	−1.81*	0.07	0
	GPD1	489	5	3	0.27	0.001	0.002	−1.24	0.06	0
	IGS1	721	12	3	0.07	0.001	0.002	−1.62	0.03	0
	LAC1	437	59	6	0.64	0.003	0.023	−2.60***	0.06	2
	PLB1	533	7	3	0.07	0.000	0.002	−1.90*	0.06	0
	SOD1	529	11	2	0.01	0.000	0.004	−2.25**	0.07	0
	URA5	637	10	4	0.33	0.001	0.003	−1.78*	0.06	0
			Average		0.67	0.001	0.005			$5^{\#}$
Africa ($n = 148$)	*CAP59*	483	14	13	0.75	0.005	0.005	−0.27	0.08	1
	GPD1	489	14	11	0.79	0.006	0.005	0.51	0.11	0
	IGS1	659	89	13	0.70	0.011	0.024	−1.83*	0.04	2
	LAC1	440	11	7	0.79	0.006	0.004	1.05	0.13	0
	PLB1	533	16	12	0.83	0.006	0.005	0.12	0.09	1
	SOD1	527	26	15	0.47	0.008	0.009	−0.34	0.08	4
	URA5	636	25	15	0.74	0.006	0.007	−0.52	0.08	1
			Average		0.95	0.007	0.009			$15^{\#}$
N. America (n=23)	*CAP59*	483	8	5	0.75	0.006	0.004	1.16	0.19	0
	GPD1	489	7	5	0.72	0.005	0.004	1.06	0.18	0
	IGS1	721	17	6	0.69	0.008	0.006	1.06	0.17	2
	LAC1	440	9	5	0.79	0.009	0.006	1.82	0.21	0
	PLB1	533	9	5	0.77	0.007	0.005	1.45	0.20	0
	SOD1	537	12	3	0.17	0.003	0.006	−1.85	0.10	0
	URA5	637	9	4	0.73	0.006	0.004	1.79	0.21	0
			Average		0.93	0.006	0.005			$5^{\#}$
S. America (n=5)	*CAP59*	483	1	2	0.60	0.001	0.001	1.22	0.30	0
	GPD1	489	0	1	0.00	0.000	0.000			
	IGS1	721	10	2	0.60	0.008	0.007	1.79	0.30	0
	LAC1	440	2	2	0.60	0.000	0.002	1.46	0.30	0
	PLB1	533	1	2	0.60	0.001	0.001	1.22	0.30	0
	SOD1	537	0	1	0.00	0.000	0.000			
	URA5	637	1	2	0.60	0.001	0.001	1.22	0.30	0
			Average		0.60	0.002	0.002			$0^{\#}$

[a]total number of sites in alignments, excluding indels and missing data; [b]number of segregating sites; [c]number of haplotypes; [d]haplotypic diversity; [e]average number of nucleotide differences per site; [f]Watterson's estimate of the population-scaled mutation rate, expressed per site (Watterson, 1975); [g]Tajima's D (Tajima, 1989); [h]Ramos-Onsins & Rozas' R_2 (Ramos-Onsins and Rozas, 2002); [i]minimum number of recombination events (Hudson and Kaplan, 1985); [#]average R_m between all seven loci; *p <0.05; **p <0.01; ***p <0.001.

Tajima's D tests the null hypothesis that populations are in mutation-drift equilibrium (Tajima 1989). In the case of significant deviation from zero, the null hypothesis of neutral (random) evolution is rejected, a finding which can be due to the occurrence of natural selection or variable population dynamics (Oleksyk et al., 2010; Holsinger, 2010). Significant departures from neutrality were detected at five of the seven loci of the Asian population (Table 6.4), all of which had negative values. The remaining three global populations (Africa, North and South America) only had one or no significant departures from zero (Table 6.4). Ramos-Onsins & Rozas' (2006) R_2 test, powerful at detecting population growth, did not detect any deviation from random evolution among any of the populations (Table 6.4).

The divergence among, and differentiation between, the four continental *Cng* populations were estimated using tests based on DNA diversity: the average nucleotide divergence between populations (D_{xy}; Nei 1987), a weighted measure of the ratio of the average pairwise differences within populations to the total average pairwise differences (K^*_{ST}; Hudson et al., 1992), and the nearest-neighbour statistic (S_{nn}; Hudson and Kaplan, 1985; Hudson, 2000). Low levels of nucleotide divergence were observed, with D_{xy} ranging from 0 to 1%, and no fixed differences were found between the various continental populations at the seven loci (Table 6.5). The total number of shared polymorphisms among populations ranged from ten for Asia vs. South America, to 68 for Africa vs. North America, with locus *IGS1* contributing the most in each case (Table 6.5). The null hypothesis of no differentiation among populations of *Cng* was rejected for all populations paired with Asia due to significant K^*_{ST} and S_{nn} values (Table 6.6). Africa and North America were also significantly differentiated, although considerably less so ($K^*_{ST} = 0.007$, $S_{nn} = 0.81$), reflecting the high number of shared polymorphisms (Table 6.5). Again, these findings match those found in both the first and second principal coordinates of the global PCA which showed the Asian population to be most differentiated from all other subpopulations, while Africa and North America were less so (Figure 6.1).

Table 6.5.: Divergence among groups of *Cng* isolates of African (n = 148), Asian (n = 192), North and South American origin (n = 23 and 5, respectively).

	Africa–Asia			Asia–N. America			Asia–S.America			Africa–N.America			Africa–S.America			N.America–S.America		
Locus	D_{xy}[a]	S_f[b]	S_s[c]	D_{xy}	S_f	S_s	D_{xy}	S_f	S_s	D_{xy}	S_f	S_s	D_{xy}	S_f	S_s	D_{xy}	S_f	S_s
CAP59	0.003	0	5	0.004	0	5	0.001	0	0	0.005	0	8	0.003	0	1	0.004	0	1
GPD1	0.004	0	5	0.004	0	5	0	0	0	0.001	0	6	0.004	0	0	0.003	0	0
IGS1	0.007	0	12	0.006	0	12	0.008	0	10	0.01	0	16	0.011	0	10	0.009	0	10
LAC1	0.006	0	4	0.008	0	2	0.004	0	0	0.008	0	9	0.005	0	2	0.007	0	2
PLB1	0.004	0	7	0.005	0	7	0.003	0	0	0.006	0	9	0.004	0	1	0.005	0	1
SOD1	0.005	0	10	0.002	0	11	0	0	0	0.006	0	11	0.005	0	0	0.001	0	0
URA5	0.005	0	9	0.005	0	8	0	0	0	0.006	0	9	0.004	0	1	0.004	0	1
Ave/total	0.005	0	52	0.005	0	50	0.003	0	10	0.007	0	68	0.005	0	15	0.005	0	15

[a]minimum estimate of the number of nucleotide differences per site between groups; [b]number of fixed differences between groups; [c]number of shared polymorphisms between groups.

Table 6.6.: Differentiation between groups of *Cng* isolates of African ($n = 148$), Asian ($n = 192$), North and South American origin ($n = 23$ and 5, respectively). K^*_{ST} values are displayed above the main diagonal and represent the weighted measure of the ratio of the average pairwise differences within groups to the total average pairwise differences. S_{nn} values are displayed below the main diagonal and represent the proportion of nearest neighbours in sequence space that are found in the same group.

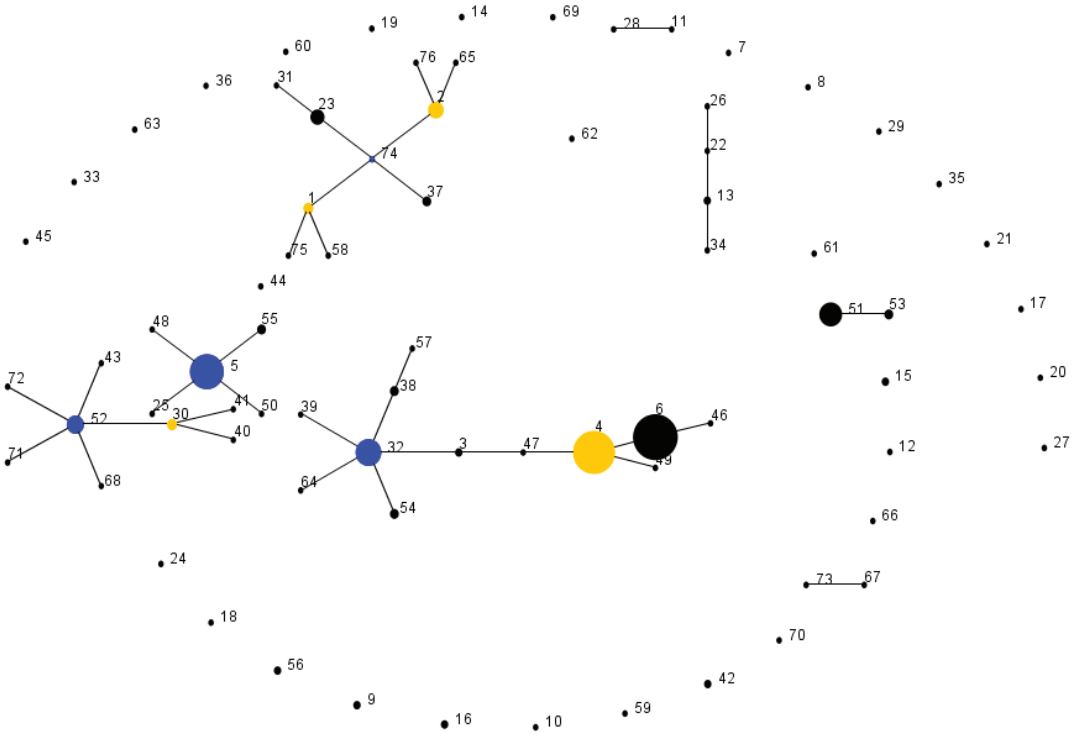
	Africa	Asia	N. America	S. America
Africa		0.17***	0.007*	0.00
Asia	0.88***		0.14***	0.07***
N. America	0.81*	0.95***		0.01
S. America	0.94	0.97***	0.75	

Significance levels for K^*_{ST} and S_{nn} were assessed using permutation tests, with 1000 permutations; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Europe has been excluded as it contains only two isolates.

6.4.4. Predominant clonality detected within the global *Cng* populations

eBURST, a web-enabled clustering tool (downloaded from `http://cneoformans.mlst.net/`), was used to infer patterns of evolutionary descent among clusters of related genotypes from the global MLST data. Capable of identifying mutually exclusive groups of related genotypes within populations, the technique revealed four clonal complexes within the global *Cng* population, each containing five to twelve STs (Figure 6.3). Two of these complexes' founding STs were linked to isolates from at least four continents (STs 5 and 32), while the other two consisted of exclusively African STs (STs 52 and 74). The results of the eBURST analysis mirror those of the PCA topology, demonstrating (i) the widespread relatedness within the Thai population as seen in the grouping of the all of the Asian STs into two eBURST groups linked by single-locus variants (SLVs; ST5, 48 and 50 in one, and ST4, 6, 32, 38, 46, 47 and 49 in the other), and (ii) the overlap between the African isolates and those of other continents in terms of shared STs, and groupings of unshared STs both in pairs as well as in clonal complexes (Figures 6.1 and 6.3).

Figure 6.3.: eBURST illustration comparing the global *Cng* isolates used in this study. No. of isolates = 370, no. STs = 76, no. re-samplings for bootstrapping = 1 000, no. loci per isolate = 7, no. identical loci for group def = 1, no. groups = 1. Founding genotypes are in blue, sub-founder STs in yellow and the size of the dots are representative of the number of isolates of that ST.



Assessing the overall association between alleles at the seven MLST loci, the index of association (I_A; Burt et al., 1996) and $\bar{r}_d$ (Agapow and Burt, 2001) were all significantly greater than 0 for subpopulations Africa, Asia and North America, as well as the global population as a whole (Table 6.7). Non-random association of alleles at the different loci was also assessed for clone-corrected datasets and these results also showed a predominance of clonal reproduction among the *Cng* samples, in keeping with the results of the eBURST analysis. A significant proportion of phylogenetically compatible loci pairs was found for all geographically defined subpopulations except for the global *Cng* population and the clone-corrected Asian subpopulation, for whom the hypothesis of random mating could not be rejected (Table 6.7). Clone-corrected VNI isolates were significantly greater than 0, but considerably less so in comparison to the other two molecular groups (Table 6.7). Taken together, these data show that linkage disequilibrium, likely resulting from extensive clonality, is exhibited by all molecular groups and all geographic populations of *Cng*.

I assessed the evidence for there being a component of sexual reproduction in the dataset. The minimum number of recombination events (R_m; Hudson and Kaplan 1985) was estimated both within an individual locus and between loci (R_m and average R_m respectively; Table 6.4) within described populations Africa, Asia and North America. Despite the main feature of the Asian population ($n = 192$) being strong clonality, some evidence for inter-locus recombination was detected (average $R_m = 5$; Table 6.4). However, this was low in comparison to the African population, where an average R_m of 15 was observed. Africa also exhibited more intralocus recombination, with 5/7 loci showing 1 or more inferred events, as opposed to 1/7 loci in Asia and North America. When analysed according to molecular group, recombination was detected within all three molecular groups, with VNB having the highest minimum number of recombination events ($R_m = 9$) in comparison to VNI and VNII ($R_m = 7$ and 6, respectively; data not shown). The global population is genetically recombining overall. However, there is also clonal expansion taking place within this population. This clonal expansion is what comprises populations such as the Thai one.

Table 6.7.: Multilocus linkage disequilibrium analyses for subpopulations of the global *Cng* population, according to geographic origin and molecular group.

Total sample				Clone-corrected sample[a]			
Population[#]	I_A[b]	$\bar{r}_d$[c]	PcP[d]	Population	I_A	$\bar{r}_d$	PcP
Africa ($n = 148$)	2.07***	0.35***	0.10***	Africa ($n = 64$)	1.28***	0.22***	0.10***
Asia ($n = 192$)	1.44***	0.32***	0.00***	Asia ($n = 11$)	1.70***	0.29***	0.9
N. America ($n = 23$)	2.38***	0.40***	1***	N. America ($n = 14$)	1.87***	0.31***	1***
VNI ($n = 317$)	1.12***	0.20***	0.14	VNI ($n = 41$)	0.26*	0.04*	0.14
VNII (n =29)	2.58***	0.44***	0.76***	VNII ($n = 14$)	1.46***	0.24***	0.76***
VNB ($n = 24$)	1.10***	0.19***	0.52***	VNB ($n = 21$)	0.86***	0.15***	0.52*
Global ($n = 370$)	0.46**	0.09**	0	Global ($n = 76$)	1.33***	0.22***	0

[a]excluding replicate haplotypes; [b]index of association; [c]scaled index of association (I_A) by the number of loci ($m-1$); [d]percentage of phylogenetically compatible pairs (PcP) of loci. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. [#] The South American and European populations were not individually analyzed due to their sample sizes being too small ($n = 5$ and 2, respectively), but were included in the global population ($n = 370$).

6.4.5. Divergence time estimates and haplotype networks support a hypothesis of African ancestry for global *Cng* isolates

The time of divergence between the global molecular group subpopulations was defined as the mean time to most recent common ancestor (tmrca). Estimated using Bayesian Markov chain Monte Carlo (MCMC) methods in BEAST, an initial 10^6 run assuming a relaxed log-normal was used to assess the divergence between the molecular groups. This run was calibrated at 4.5 million years ago (standard deviation = 0.5), the estimated time since divergence within *Cng* molecular groups as estimated by Xu et al. (2000). Describing the ancient evolutionary events splitting between the *Cng* varieties, estimates revealed that molecular group VNB was the oldest, with an estimated mean tmrca of 4.503 million years ago (95% highest posterior density, HPD = $(4.236 \times 10^6, 4.782 \times 10^6)$; ESS = 1 201). This was followed by molecular group VNII which was older than VNI by about 4 000 years.

The divergence between the geographically defined *Cng* populations was then assessed using an uncalibrated 10^6 run assuming a strict molecular clock model and an optimised mutation rate of 4×10^{-9} per nucleotide per year. The global population had an estimated mean tmrca of 4.57 million years ago (95% highest posterior density, HPD = $(3.23 \times 10^6, 5.86 \times 10^6)$; ESS= 245). The African population's tmrca was also estimated at 4.57 million years ago, representative of the speciation between the three molecular types it contains (Figure 6.4). Of the total 192 Asian isolates, 191 were of molecular group VNI. In order to assess the emergence of this population in relation to the African population, VNII and VNB isolates were removed from the sample. This was in order to avoid tmrca estimates describing the coalescence between VNI and older molecular groups. Estimated from four individual runs with a combined total of 1.2×10^7 states, divergence between the VNI isolates was estimated to have taken place 23 090 years ago (Figure 6.5). The 95% HPD of this event is (1 562, 70 464) years and spans the proposed time of spread suggested by Litvintseva et al. (2011) of 5 000 years.

To further explore the potential African ancestry of the *Cng* population, haplotype networks were constructed for each MLST locus (Figure 6.6).

Figure 6.4.: Trace file of the divergence estimates of *Cng* isolates according to geographic origin (Africa, Asia and North America). South America and Europe were excluded due to small sample sizes ($n = 5$ and 2, respectively), and therefore low ESS values.

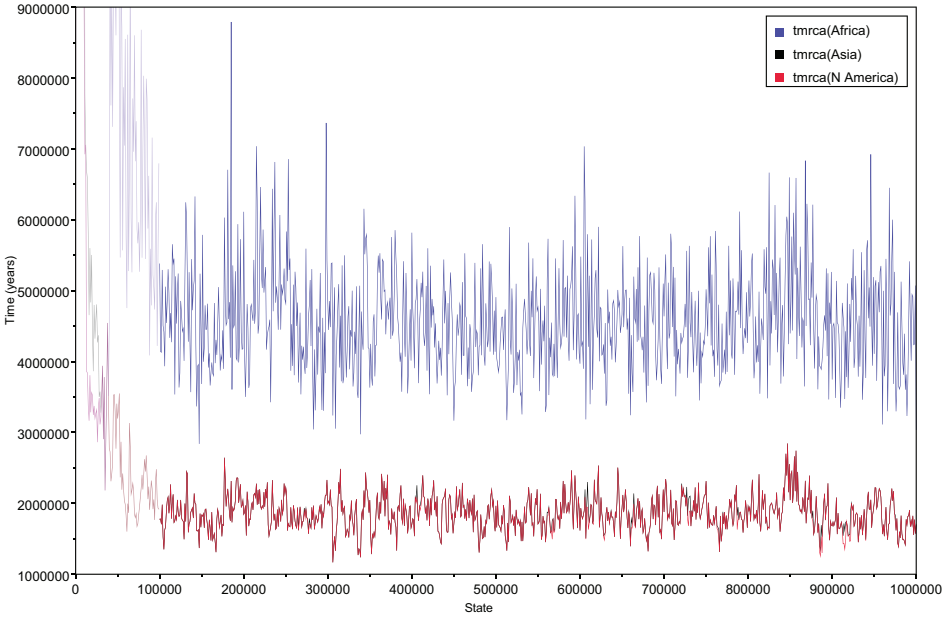
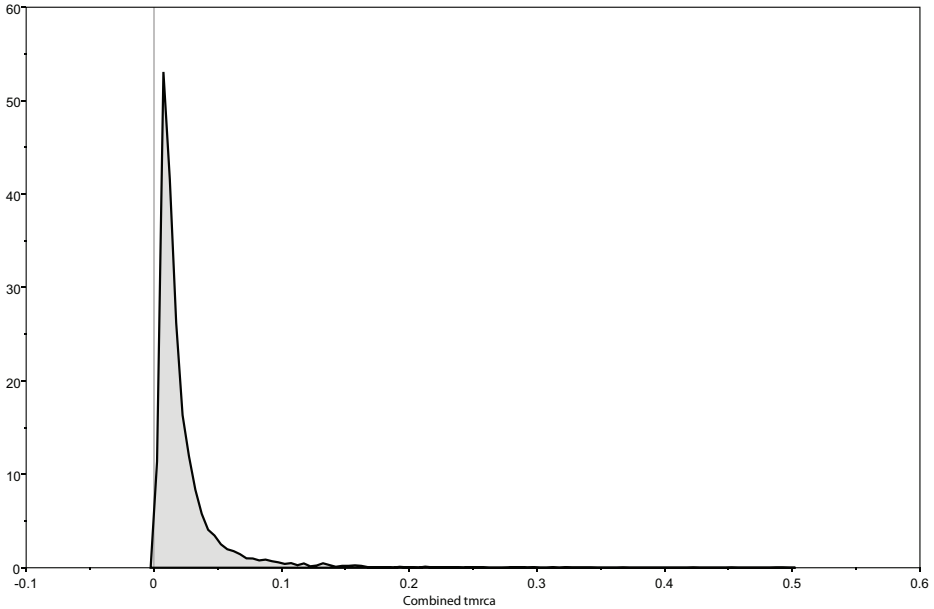
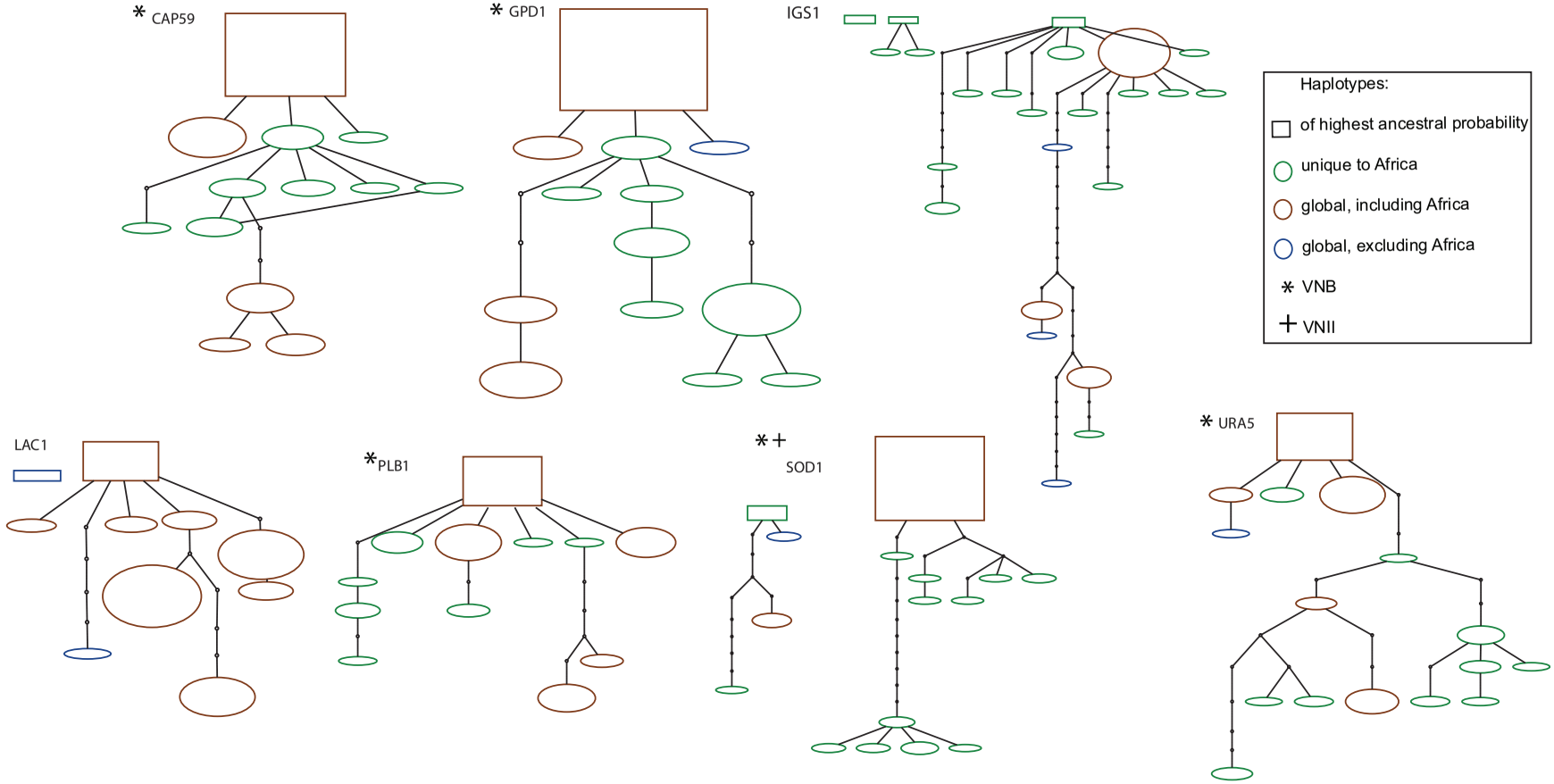


Figure 6.5.: Marginal density of the divergence estimates of *Cng* isolates of molecular group VNI ($n = 317$).



Sampled haplotypes are indicated by circles or rectangles colored according to the geographical region from which the sample was collected and proportional in size to observed haplotype frequency. Rectangles depict the haplotype with the highest ancestral probability and each branch indicates a single mutational difference. Internal nodes are representative of ancestral haplotypes, from which apical haplotypes evolved. The STs of non-African genotypes (shown in blue) were few and tended to be found at the apical (i.e. derived) positions of the networks, with the exception of locus *SOD1* where, although not apically positioned, the blue non-African STs are clearly derived from STs of African origin (Figure 6.6). The green circles, which represented STs of African origin only, were positioned throughout the networks but were only associated with clinical haplotypes. These seven networks pointed to an ancestral African population which had the highest variation in haplotype numbers and from which other global haplotypes were derived (Figure 6.6), as previously demonstrated by diversity indices and eBURST (Table 6.4; Figure 6.3). In terms of molecular group, the VNI and VNB subpopulations share ancestral haplotypes at five of the seven MLST loci (all except for *IGS1* and *LAC1*), indicating shared common origin. VNII, on the other hand, only has a single ancestral haplotype at *URA5* (Figure 6.6).

Figure 6.6.: Haplotype networks of the 76 STs of the global *Cng* population at each of the seven loci. Sampled haplotypes are indicated by circles or rectangles colored according to the geographical region from which the sample was collected. STs unique to the African population are shown in green and consist only of clinical isolates. Haplotypes found both in Africa and elsewhere are in brown, while those not found in Africa are represented in blue. Rectangles depict the haplotype with the highest ancestral probability. For each of these ancestral haplotypes, the inclusion of VNB and VNII isolates is indicated by an asterisk (*) or a plus sign (+), respectively. Each branch indicates a single mutational difference and black dots on the lines are representative of the number of mutational steps required to generate allelic polymorphisms. Circle size is proportional to observed haplotype frequency.



6.5. Discussion

In this chapter, I describe and compare the type and distribution of diversity between five different continental populations of *Cng*, including the Thai and Cape Town populations typed and described in Chapters 4 and 5. While sample sizes were low for two regions (Europe and South America), the power to detect differences between continents was satisfactory for the other regions (Africa, Asia and North America). The data and analyses clearly showed the following facets of *Cng*'s global population structure: (i) the fungus is widely clonally reproducing, (ii) recombination, where observed, is geographically proscribed, and (iii) the African subpopulation is the most genetically diverse and contains ancestral haplotypes, making it a candidate source of origin for *Cng*.

Multilocus sequence typing was initially applied to *Cng* populations Thailand ($n = 183$; Chapter 4) and Cape Town ($n = 107$; Chapter 5). These two populations were found to be genetically similar to the Asian and African populations of the global set by AMOVA (Table 6.2) and phylogenetic analyses (Figure 6.2) and were therefore combined to form two continentally separated subpopulations of *Cng* (Asia = 192, Africa = 148). Additional geographically defined subpopulations of *Cng* from the global population included in the analyses and were North America ($n = 23$), South America ($n = 5$) and Europe ($n = 2$; Table6.1).

MLST elucidated the population structure of global *Cng*, separated into three previously described molecular groups: VNI ($n = 317$), VNII ($n = 29$), and VNB ($n = 24$; Figure 6.2; Litvintseva et al., 2006). The neighbor-joining dendrogram revealed clustering between VNI and VNB with high bootstrap support ($n = 70\%$). This sharing of a common phylogenetic history between VNI and VNB was also identified in *Cng* by Litvintseva et al. (2011) who went on to state that the two are not cryptic species. In contrast, within the significantly clustered *Cng*, phylogenetic analysis suggests independent evolution of VNII. This mirrors the lack of ancestral haplotypes within the VNII population, as depicted by haplotype networks (Figure 6.6).

Molecular group VNB was restricted to Southern Africa, namely Botswana and Cape Town, while VNII isolates were found in North America as well as Africa with the two sharing a single genotype (ST30, Table B.1). Of the

76 global STs, 72% were unique to the African population, 39% of which were of molecular type VNB ($n = 21$). In addition, all ten rare MAT**a** isolates were found in the African population, eight VNB and two of VNI (Table B.1). PCA analysis revealed overlap in genotypes between isolates from Africa and the other three subpopulations of *Cng* (Figure 6.1). This was consistent with the distribution of African isolates across the phylogenetic tree (Figure 6.2) and showed a lack of geographical correlation within VNI genotypes, unlike that seen in VNB.

A predominantly clonal population structure both within the geographically defined subpopulations and the global population of *Cng* was demonstrated by statistically significant tests of non-random association (I_A and $\bar{r}_d$; Table 6.7). Despite this strong clonal component being consistent with previous studies showing that non-meiotic reproduction is the predominate mode of descent in *Cng* worldwide (Litvintseva et al., 2006; Bovers et al., 2008a; Litvintseva et al., 2005; Taylor et al., 1999; Buchanan and Murphy, 1998), there has also been evidence for recombination (Nielsen and Heitman 2007; Bui et al., 2008; Xu et al., 2000). The minimum number of recombination events necessary to explain the distribution of polymorphisms within and between loci demonstrated extensive spatial variation in the rates of recombination globally (Table 6.4). Importantly, the highest number of minimum recombination events was detected in the African subpopulation (Africa, $R_m = 15$; Asia, $R_m = 5$; North America, $R_m = 5$) and the majority of the MLST loci in Africa showed evidence of intergenic recombination, in comparison with much lower levels detected elsewhere (Africa, 5/7 loci; Asia, 1/7 loci; North America, 1/7 loci; Table 6.4). This mirrors the fact that linkage disequilibrium was lowest among the African haplotypes (clone-corrected= 0.22, $I_A = 1.28$; p-value< 0.001 each; Table 6.7), suggestive of a higher level of genetic recombination among these strains in comparison to the other subpopulations. These results are in keeping with studies reporting sexual propagation within both clinical (Litvintseva et al., 2003) and environmental African isolates of *Cng* (Lin et al., 2009; Litvintseva et al., 2011). Furthermore, subdivisions according to VN group reported a greater number of recombination evens in the African VNB population ($R_m = 9$) in comparison to the VNI group ($R_m = 7$; Table 6.4), likely due to the higher frequency of the **a**-mating type detected in the former (Litvintseva et al.,

2006). Recent findings on the ability of *Cn* strains of the same mating type to reproduce by haploid and monokaryotic fruiting, a process previously believed to be mitotic and asexual and upregulated in the presence of MAT**a** cells, may account for the genetically recombining global population within which there is clonal expansion (Lin et al., 2005; Wang et al., 2000).

Estimates of haplotypic diversity (Hd), mutation rates (θ) and nucleotide differences (π) were consistently greater for Africa, relative to populations in other continents (Table 6.4). Africa exhibited the greatest number of haplotypes (Africa = 86 > North America = 33 > Asia = 23), and the Asian population exhibited the least amount of haplotypic diversity (Africa = 0.95 > North America = 0.93 > Asia = 0.67). This reflects the large number of genotypes unique to Africa ($n = 54$ of 76), as well as its balanced composition of the different molecular groups. It is important to note that variation within North American isolates is considerably higher than that within Asia, to the point of being almost comparable to that found in Africa (Table 6.4). This is attributable to the presence of VNII isolates within North America ($n = 7$ of 23) which are phylogenetically diverged from VNI, and also older by approximately 4 000 years (Figure 6.2). Tajima's D is a statistical test to identify loci that are evolving under non-random processes, such as selection or demographic expansion or contraction, and showed that 5/7 MLST loci in Asia were significantly non-neutral, compared to 0/7 loci in North America and only 1/7 in Africa. The mean of D considers the difference between θ and π; the former gives more weight to rare alleles than the latter. A significant negative value at a single locus indicates a selective sweep at that locus (as seen at *IGS1* in Africa; Table 6.4), while many such values obtained at independent loci would suggest exponential population expansion or "genome-wide selective sweep" (as seen in the Thai population; Eswaran et al., 2005; Slatkin and Hudson 1991; Przeworski et al., 2000; Braverman et al., 1995). Positive D values imply balanced polymorphisms, hence common intermediate frequency alleles.

The model of rapid population expansion is associated with low nucleotide diversities and haplotypic pairwise sequence differences (Excoffier et al., 1992). The global population (excluding the African subpopulation) appears to fit this model (Global $Hd = 0.93$, $\pi = 0.004$ and $\theta = 0.008$; data not shown). This is consistent with other global analyses of genetic varia-

tion between global isolates and African isolates presented by Litvintseva et al. (2011). When subdivided into geographical location, this holds true for the Asian population, supported by diversity indices and Tajima's D results (Table 4.2).

Global analyses of pairwise population combinations led to the detection of significant genetic differentiation between all *Cng* populations (excepting the comparison between North and South America; Table 6.6), which shows that the different continental populations of *Cng* are experiencing divergent evolutionary trajectories. The global population's comparatively low genetic diversity, high linkage disequilibrium, non-neutral evolution, and lack of geographically defined structure are all consistent with a model of a rapid population expansion from a limited set of ancestors. These findings contrast with the African population of *Cng*, which is characterised by high genetic diversity, balanced mating and molecular types, and elevated recombination rates. Finding that the global isolates are genetically monomorphic in relation to African isolates led to the examination of the potential of an ancestral African origin of *Cng* using coalescent analyses in BEAST and haplotypic networks.

Xu et al. (2000) describe the major lineages of *Cn* as having diverged tens of millions of years prior, but that their dispersal is relatively recent. A substitution rate of 4×10^{-9} and a strict molecular model estimated the time to most recent common ancestor of the global population of *Cng* in this study to approximately 4.573 million years ago. This is identical to the time since divergence within the *Cng* species estimated by Xu et al. (2000), weighted for protein coding genes *LAC* and *URA5*. Apart from the single clinical isolates of Thai origin (CM21), molecular group VNII was isolated in Africa and North America only. In light of the results discussions above, VNII appears to be evolving independently from VNI and VNB and, as only one of the isolates are of environmental origin and numbers are low ($n = 29$), it is difficult to predict the pattern of dispersal and origin among these isolates.

Southern Africa's highly genetically diverse population is evidence of its potential as the origin of *Cng* (Goodwin et al., 1994; Xu et al., 1997; Avise, 2004). Haplotype networks for each MLST network show that haplotypes unique to the African population occupy both internal and apical positions

within the networks, whilst those unique to the global population are almost always in derived positions at the network-tips (Figure 6.6). In addition, founding STs of clonal groups identified by eBURST are all found in Africa (Figure 6.3). These data indicate a basal African population, which contains the ancestral molecular VNB group as determined by dating estimates. The VNI molecular group shares a common phylogenetic history with this ancestral VNB subpopulation, depicted by the phylogenetic tree (Figure 6.2) and shared ancestral haplotypes (Figure 6.6). This persuasive evidence for the derivation of molecular group VNI from VNB is supported by Litvintseva et al. (2011), who reported evidence of incomplete lineage sorting between these two molecular groups. This VNI lineage of VNB then became global approximately 23 000 years ago, hence the more recent invasion of Southeast Asia supported by low genetic diversity and negative Tajima's D values across all seven MLST loci (Table 6.4). It is important to note that the estimated date of the tmrca of the VNI population spans the proposed time of the domestication and spread of the common pigeon by humans through trade 5 000 years ago (Mooney, 2000; Grzimek, 2004; Litvintseva et al., 2011). The human vectoring of fungal pathogens of fungal pathogens is common, and has been reported in several pathogen pollutions including *Coccidiodes immitis* and *Toxoplasma gondii* (Daszak et al., 2000; Fisher et al., 2001; Su et al., 2003). eBURST analysis revealed two clonal clusters which contained the Asian STs (Figure 6.3). As this tool identifies likely founding STs, it would appear that there may have been not one, but two introductions of *Cng* VNI to Thailand: one by ST5 and the second by ST32. This hypothesis of recent expansion of VNI to Thailand can be applied to the global dispersal of VNI based on the fact that strains from different geographic regions which share molecular groups are more similar to each other phylogenetically than strains of different molecular groups from the same region — assuming sexual reproduction is taking place in nature (Xu et al., 2000). This rings true of the genetic structure seen in global population of *Cng* (excluding the ancestral African population) which shows no evidence of genetic correlation among the VNI STs.

Litvintseva et al. (2011) recently reported a significant association between environmental sources of *Cng* and geographic location in a Southern African population; further compelling evidence for the ancestral African

origin of *Cng*. They found that ancestral African *Cng* haplotypes were exclusively isolated from native trees *Colophospermum mopane* found in rural locations, while global strains restricted to urban areas were found in pigeon excreta. Additionally, these isolates found in two separate environmental niches were distinguishable by molecular group; the former (from endemic trees) being of VNB, while the latter were of VNI. The apparent physical separation of the two environmental Southern African *Cng* subpopulations was restricted genetic exchange. This separation of the two may provide an explanation for the different expansion trajectories seen among the isolates typed in this study. The isolation of VNI isolates from pigeon guano in Southern Africa and Thailand supports the proposed hypothesis of the common pigeon being key in the global dispersal of VNI *Cng* (Simwami et al., 2011; Litvintseva et al., 2011). Litvintseva et al. (2011) state that two African ancestral haplotypes account for the diversity within the entire global coprophilic *Cng* population. The exclusive presence of pigeons carrying VNI in urban areas of Southern Africa may serve to explain the high rate of infection by VNI seen in the the clinical African cohort of Cape Town (Chapter 5). Urban and rural areas vary in terms of human traffic, with most people being in regions where the pigeons carry VNI. Lower numbers of people being exposed to virulent VNB strains in restricted to rural areas may explain why only four patients within the typed Cape Town population were infected by *Cng* VNB: C14, C15, R22 and R26 (Table B.1). Similarly, as this study shows conclusive evidence of only the VNI lineage's global spread, worldwide exposure to this molecular group is high, hence the global trend of VNI predominance among clinical patients. In keeping with this, none of the Thai CM patients were infected with *Cng* of molecular group VNB (Table B.1).

Although confined to Botswana in this study, previous studies have reported the occurrence of VNB *Cng* (also known as AFLP genotype 1A) infecting AIDS patients in Rwanda, the USA and Belgium, isolated from the environment in Zaire and Australia, and from both clinical and environmental samples in Brazil (Barreto de Oliveira et al., 2004, Boekhout et al., 2001, Bovers et al., 2008b), South Africa and South America (Ngamskulrungroj et al., 2009; Bovers et al., 2008b). In addition, AD hybrid strains from three continents were found to possess the rare MAT**a** allele, and to

be related to VNB strains of Botswana origin (Litvintseva et al., 2007). This means that the dispersal "Out of Africa" may also be applicable to the VNB *Cng* population, but not through the avian carrier *Columba livia*, as no VNB strains were isolated from pigeon excreta (Litvintseva, 2011). As with VNII, more strains of this molecular type from numerous geographic areas must be analysed in order to provide evidence of the geographical origin and direction of dissemination.

The invasion of novel environments by fungal species is common and brought about either by anthropogenic or natural mechanisms (Schwartz et al., 2006; Parker and Gilbert 2004). For example, the emerging disease chytridiomycosis has recently been found to be causing decline in amphibian populations globally (Bosch et al., 2000; Daszak et al., 1999; Waldman et al., 2000). Comparative population genomics link the intercontinental introduction and spread of a hypervirulent lineage of *Batrachochytrium dendrobatidis*, which arose through anthropogenic mixing, through the international trade of amphibians (Walker et al., 2008; Farrer et al., unpublished; Daszak et al., 2001). The hypothesis of an African origin of *Cng* suggests human assisted dispersal of the fungus vectored by the avian species *Columba livia* (Litvintseva et al., 2011; Casadevall and Perfect 1998). This hypothesis is based on evidence found to be true within my study: (i) the high genetic diversity identified in Southern Africa compared to the global population of *Cng* (Table 6.4), (ii) the highly clonal global strains and lack of geographic structure (Table 6.7; Figure 6.2), (iii) the ancestral haplotypes being found in the African subpopulation (Figures 6.3 and 6.6), and (iv) the isolation of global STs from pigeon guano (Table 4.1). In conclusion, there may have been several expansions of *Cng* out of Africa — at least one per molecular type, and possibly two of VNI — and there may be several yet unsampled ecological niches of the yeast. Further sampling and typing of environmental sources may, in turn, reveal additional vectors and transmission routes which will contribute to our understanding of the structure and distribution of the fungus.

7. Clinical analysis

7.1. Introduction

Over the past 25 to 35 years, clinical studies have informed antifungal drug combinations key to the management of AIDS-associated cryptococcal meningitis (CM; Perfect, 2006a). Despite these advances, acute mortality rates attributable to CM remain unacceptably high — between 10 and 25% in developed countries, and higher yet in developing ones (Perfect and Casadevall, 2002). Continued efforts contributing to our understanding of the molecular pathogenesis of *Cng* are essential for the optimisation of prevention and treatment strategies against this opportunistic infection.

There is a strong body of evidence linking *Cng*'s genotype and disseminated cryptococcal infection. Serotype and cryptococcosis are associated, with *Cng* causing over 90% of infections among the immunocompromised. The invasion and expansion of two recombinant genotypes of *C. gattii* in the Pacific Northwest revealed that different genotypes of *Cn* can cause striking different clinical phenotypes, quantified by intramacrophagic and murine models (Byrnes et al., 2010; Ma and May, 2010). Previously believed to be host-dependent, varied expression of the pathogen's virulence factors were shown to influence this diversity in yeast-macrophage interactions and, ultimately, the clinical outcome of cryptococcal infection (Alanio, 2011). Further support of this is evidence of isolates of varied virulence eliciting different inflammatory reactions in the lungs of mice (Curtis et al., 1994). The potential to predict the outcome of cryptococcal meningitis in humans is therefore related both to the individual's health (measurable in immune status and susceptibility), as well as the genotype of the infecting fungal isolate.

This chapter takes advantage of clinical data available from three studies on HIV-associated CM patients and molecular data on the infecting isolates

obtained by MLST to identify prognostic factors associated with clinical outcome of disease. This approach is different from other studies as both host and fungal determinants are assessed as opposed to either one or the other.

7.2. Aims

This study includes a total of 116 patients from three trials, one in Thailand and two in Cape Town (Table 7.1). All three trials compared various antifungal treatments for AIDS-associated CM and clinical data were collected on each patient (Brouwer et al., 2004; Bicanic et al., 2007; Bicanic et al., 2008). The patient data were integrated with corresponding genotypic information on the infecting *Cng* isolate determined by MLST typing and described in previous chapters (Chapters 4 and 5; Table 7.1). Appropriate analyses were then undertaken in order to identify associations between clinical factors measured during the course of infection, and molecular data on the infecting pathogen.

My aims were

1. To describe the progression of cryptococcal disease in the combined clinical cohort of 116 HIV+ patients according to *Cng* ST;
2. To identify clinical characteristics of infection affected by ST; and
3. To identify prognostic factors of clinical outcome of disease, measured by death and rate of cryptococcal clearance.

7.3. Materials and methods

One hundred and eighty two patients from three trials, one in Thailand and two in Cape Town, were initially included in this study and are described in Chapters 3 and 4. Patients within these trials had been allocated to one of seven antifungal treatments for AIDS-associated CM (Table 7.1). In this study, they were further categorised into new groups according to the infecting *Cng* ST and these new patient groups analysed in this chapter.

As the patients were from two different countries, geographical region was assessed as a factor (continent). The clinical characteristics of disease

Table 7.1.: The seven treatment arms to which 182 patients from three clinical studies assessing AIDS-associated cryptococcal meningitis treatment therapies were assigned ($n = 182$).

	Antifungal treatment	Study 1[a] ($n = 64$)	Study 2[b] ($n = 54$)	Study 3[c] ($n = 64$)
1[#]	Amphotericin B only (AmB; 0.7 mg/kg/day)	16		
2[#]	AmB (0.7 mg/kg/day) + flucytosine (100 mg/kg/day)	16		30
3[#]	AmB (0.7 mg/kg/day) + fluconazole (400 mg/kg/day)	16		
4[#]	AmB (0.7 mg/kg/day) + flucytosine (100 mg/kg/day) + fluconazole (400 mg/ day)	16		
5[†]	AmB (1 mg/kg/day) for 7 days, followed by fluconazole (400 mg/day) for eight weeks		49	
6[†]	Fluconazole (400 mg/day) for ten weeks		5	
7[*]	AmB (1 mg/kg/day) + flucytosine (100 mg/kg/day)			34

[a] Brouwer et al., 2004 (trial period = May - December, 2002); [b] Bicanic et al., 2007 (trial period = February - September, 2005); [c] Bicanic et al., 2008 (trial period = May 2005 - June 2006)

[#] All four treatment arms lasted for two weeks and were followed by 8 weeks of 400mg/day of oral fluconazole.

[†] All patients on treatments (5) and (6) were switched to fluconazole (200mg/day) after ten weeks.

[*] This last treatment lasted for two weeks and was followed by 8 weeks of 400mg/day of oral fluconazole.

N.B. ART was generally unavailable in Thailand during the time period over which the study was conducted, but patient counseling and ART initiation occurred at a median interval of 47 days (Cape Town) following antifungal therapy (Bicanic et al., 2009b).

recorded for the patients in the three studies include baseline measurements of cerebrospinal fluid (CSF) opening pressure (cm H_2O), quantitative cryptococcal colony forming units (CFU) in CSF culture (CFU/mL CSF) and logarithmic interferon gamma (IFNγ; Table 7.2). Cerebrospinal dysfunction upon presentation was also reported.

The outcome variables of interest were death, measured at weeks 2 and 10, and early fungicidal activity (EFA). Eighteen percent of all the patients had died by week 2, with the cumulative figure rising to 46% by week 10 ($n = 31$ of 67). Fungicidal activity, EFA, is a summary statistic which measures the decrease in CSF cryptococcal colony-forming units (CFU) from quantitative CSF cultures. It is defined as the average daily decrease in log CFU per mL CSF (Brouwer et al., 2004). For each patient, EFA was estimated by fitting a simple linear regression model with CFU as the response and time (in days) as the explanatory variable. Thus, the patient's EFA is the estimated slope coefficient of this model. Patients with fewer than two observations were not included in this calculation.

Descriptive statistics of the patients' baseline characteristics were computed. Continuous variables were summarised by their overall median and interquartile range, as well as the mean grouped by ST. Categorical variables were described by group proportions. Statistical comparisons of the continuous and categorical baseline variables in relation to ST were conducted using ANOVA and Fisher's exact test, respectively. Significant associations were further analysed by allele types at the seven MLST loci. Univariable and multivariable linear models were used to determine strong prognostic factors of early death, while associations of clinical variables with EFA were assessed using linear regression. Survival analyses were also carried out using the Kaplan-Meier estimator (Kaplan and Meier, 1958) and the log rank test applied to a test for differences in the survival distribution. Continuous variables were grouped into quartiles or deciles. All calculations were performed in R2.12.1 using the RStudio interface, and p-values of less than 0.05 were considered significant.

7.4. Results

7.4.1. Study population according to ST

One hundred and nineteen patients were categorised into six groups according to ST: ST4 (n = 33, 28%), ST6 (n = 23, 19%), ST5 (n = 15, 13%) and ST32 (n = 15, 13%), ST51 (n = 21, 18%) and ST52 (n = 12, 10%). STs with fewer than twelve patients were excluded from analyses. ST4 was unique to Thai patients, while STs 51 and 52 were exclusive to African patients. ST52 was the only infecting ST of molecular group VNII (Table 7.2). The six ST-defined patient groups were comparable in age and weight (with means of 34 years and 50.7 kg, respectively). Continuous variables measured prior to the commencement of therapy were opening base pressure (Opbase, with a mean of 25.8 cm H_2O) and white blood cell count (CSFWBC, with a mean of 106). CSFWBC was highly variable, ranging from 0 to 2,800 cells/mL, with 32% of patients presenting with 0 cells/mL ($n = 38$). These variables did not differ significantly according to ST ($p-$value > 0.05; Table 7.3).

Table 7.2.: Patient characteristics by MLST sequence type (ST).

		All patients	ST4 ($n = 33$)	ST6 ($n = 23$)	ST5 ($n = 15$)	ST32 ($n = 15$)	ST51 ($n = 21$)	ST52 ($n = 12$)
Continuous variable	N	Median (IQR)	Mean (unless otherwise indicated)					
Weight (kg)	110	49.3, (44.154)	47.4	47.3	55.3	51.8	53.4	57.3
Age (years)	116	33, (2938)	34	32	33	37	34	36
Cryptococcal Antigen Titer (CRAG)	118	1 024, (5123 584)	1 024	1 024	1 024	512	2 048	1 536
Opening pressure cm H2O (Opbase)	102	24.0, (14.6 – 33)	28.3	23.3	29.7	21.8	25.4	23.1
Quantitative Cryptococcal Colony Forming Unit Count, CSF CFU/mL (QCCC)	116	297 167 (21 979 – 1 276 439)	820 352†	485 289†	174 181†	19 498†	332 660†	53 456†
CD4 cell count	96	21, (8 – 36)	18	20	43	62	35.6	53.5
White Blood Cells count /mL CSF (CSFWBC)	118	8.5, (0 – 70.8)	0†	3†	13†	25†	1†	43†
log Interferon γ (IFNγ)	116	1.37, (0.8 – 1.9)	0.98	1.1	1.58	1.62	1.44	1.75
Categorical variable	N		Count (percentage in brackets)					
Death (week 2)	118		8 (24)	1 (4)	1 (7)	1 (7)	2 (10)	1 (8)
Death (week 10)#	118		9 (27)	3 (13)	1 (7)	6 (40)	6 (29)	6 (50)
Altered mental status(altment)	119		10 (30)	2 (9)	3 (20)	2 (13)	3 (14)	4 (33)
Antiretroviral therapy (ART)	119		0	0	3 (20)	4 (27)	1 (5)	1 (8)
Male	119		22 (67)	15 (65)	5 (33)	2 (13)	7 (33)	5 (42)

Cumulative death count (includes death by 2 weeks); † Median

7.4.2. ST and baseline variables

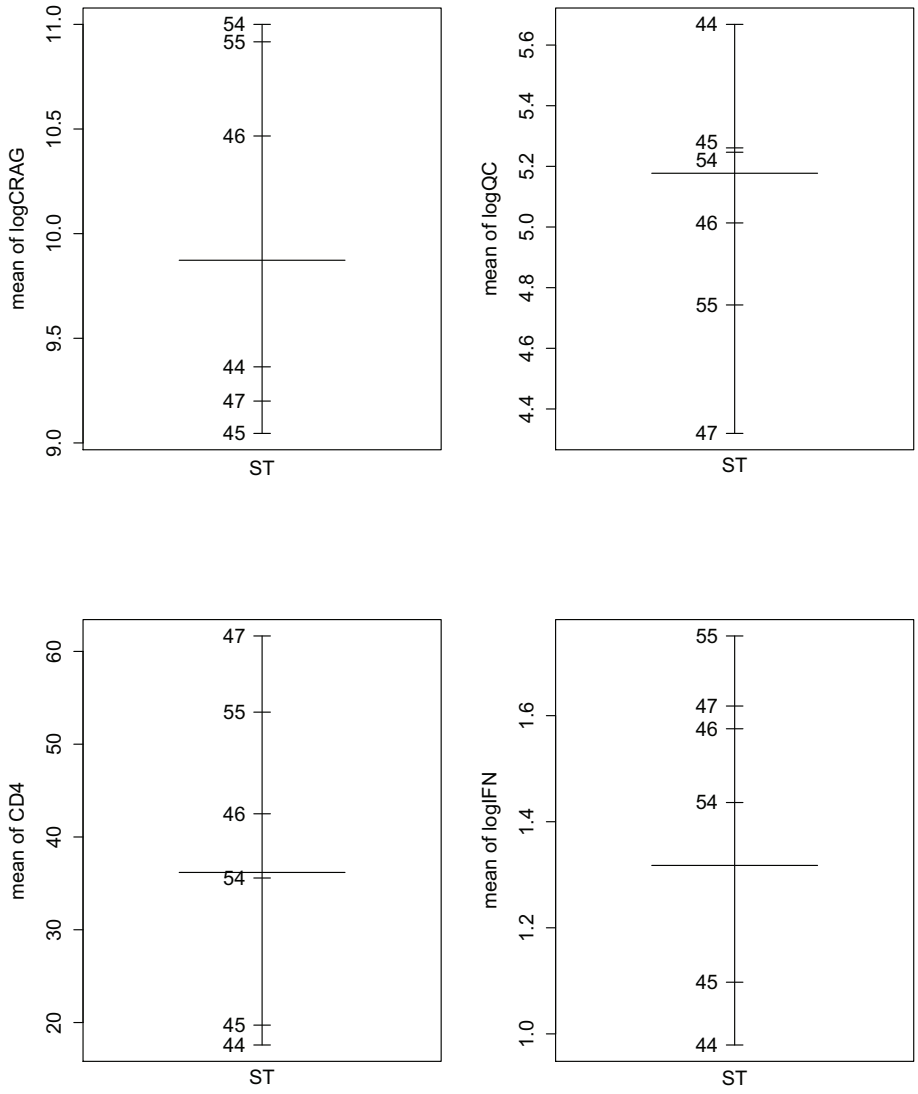
The percentage of men in each patient group varied significantly between STs (Fisher's exact test p-value < 0.001; Table 7.3). The proportion of males infected by STs 4 and 6 were 67% and 65%, respectively. All patients infected by ST4, and all but one patient infected by ST6 (isolate R6), were Thai (Table 7.2). STs 5, 32, 51 and 52 collectively contained one or no patients from the Thai study; the percentage of males for these STs ranged from 13% to 42%, with an overall percentage of 30% for these four STs.

ST significantly affected logCRAG, weight, logQC, CD4 and logIFN (ANOVA p-value $\leq 0 - 0.01$; Table 7.3). The differences between the group means for these ANOVAs can be seen in the main effects plots (Figure 7.1). Patients infected with ST51 had the highest mean baseline CRAG titre (mean = 6 334), while those infected by *Cng* ST32 presented with low mean CRAG and QC values. ST32 was also associated with elevated mean logIFN and CD4 cell counts in comparison to the average. Patients infected with STs 4 and 6 consistently had the lowest means for logIFN, CD4 cell count and weight, in contrast to ST52 infected patients, who were the heaviest on average, and had the highest logIFN and CD4 cell counts.

As ST frequency varies between the African and Asian patient populations, the baseline clinical variables were also assessed according to geographic origin. As expected, the two are strongly confounded. Continent was significantly associated with weight, sex, logIFN, and CD4 cell count (p-value < 0.001), as well as logCRAG (p-value $= 0.003$) and logQC(p-value $=$ 0.05). This pattern is identical to that found for ST displayed in Table 7.3.

Substituting ST with allele type at each locus revealed further significant associations with logQC and logCRAG, both measures of fungal burden within the host. The mean QC count was 213 039 cells/mL greater in patients infected with *IGS1* AT19 than AT20, similar to the difference of 232 085 cells/mL between AT2 and AT4 of *PLB1* (p-value ≤ 0.01; data not shown). In relation to logCRAG, patients infected with AT1 of both *CAP59* and *GPD1* had lower mean CRAG measures in comparison to those infected by ATs 2 and 7 (*CAP59*) and 5 and 9 (*GPD1*).

Figure 7.1.: Main effects plot of the significant association between ST and the mean of baseline clinical factors identified by ANOVA: log cryptococcal antigen titre, log quantitative cryptococcal culture, CD4 cell count and log interferon-gamma. The large horizontal line is the global mean for the variable in question. The small horizontal tick marks represent the means of each individual group - in this instance, patients within each ST.



7.4.3. Death by weeks 2 and 10

An association between death and altered mental status at presentation was detected by use of Fisher's exact test. This characteristic, defined by the presence of a decrease in Glasgow Coma scale or seizures, resulted in 11.8-fold increased odds of death within 10 weeks (95% CI = $(3.8, 40.4)$; p-value ≤ 0.001), similar to that estimated for death within 2 weeks (OR = 11.2; p-value ≤ 0.001). In a multivariable logistic regression, this association remained unchanged (Table 7.4). In addition to altered mental status, an increase of 1 kg in weight decreased the log odds of death within 2 weeks by 0.11 (p-value = 0.05). However, weight was not significant in the model fit for death by week 10 (p-value = 0.202). Elevated baseline opening pressure (Opbase) was significant, with a single unit change resulting in a 0.96 decrease in the log odds of death (p-value = 0.04). There was no significant interaction between death and ST.

7.4.4. Early fungicidal activity and treatment

There was a fall in CSF fungal load over time in all but two African patients infected with CT52 and CT7 (Table B.1). As clinically expected, EFA was strongly confounded with treatment effect (ANOVA p-value ≤ 0.001; data not shown). Patients on treatment 7 (1 mg AmB plus 100 mg of flucytosine per kg/day, $n = 14$) benefited from the greatest initial decrease in CSF CFU (mean EFA = -0.54 log CFU). This was followed by treatments 2 and 5, which differed in mean decrease in EFA from treatment 7 by 0.01 and 0.02 log CFU per day, respectively (Figure 7.2). The treatment which had the smallest effect on cryptococcal load reduction was treatment 6; consisting of 400 mg of daily fluconazole alone administered over ten weeks. The four patients within this group had mean EFAs ranging from -0.09 to 0.03 log CFU, with two patients experiencing an increase in cryptococcal load rather than a decrease.

In order to identify the factors contributing to the decrease in cryptococcal load in the patients, a multivariable linear regression model was fitted, with EFA as the response. Each patient's EFA estimate was weighted in proportion to the number of initial observations it was calculated from (num_obs; code C.3), because there were different numbers of observations for different

Table 7.3.: Summary of ANOVA and Fisher's exact test between infecting *Cng* ST continuous and categorical baseline variables.

Continuous variable	Df	Sum Sq	Mean Sq	F-stat	p-value
logCRAG	5	75.4	15.1	2.69	0.025
Age	5	255.9	51.2	1.00	0.422
weight	5	1 469.8	294	3.07	0.013
Opbase	5	788.7	157.7	0.63	0.676
logQC	5	21.1	4.2	3.49	0.006
CSFWBC	5	569 439	113 888	1.14	0.344
CD4	5	25 042	5 008.3	5.06	0.000
logIFN	5	9.6	1.9	3.76	0.004

Categorical variable	p-value
Sex	0.00
Altered mental status	0.31
Death by 2 weeks	0.33
Death by 10 weeks[#]	0.08

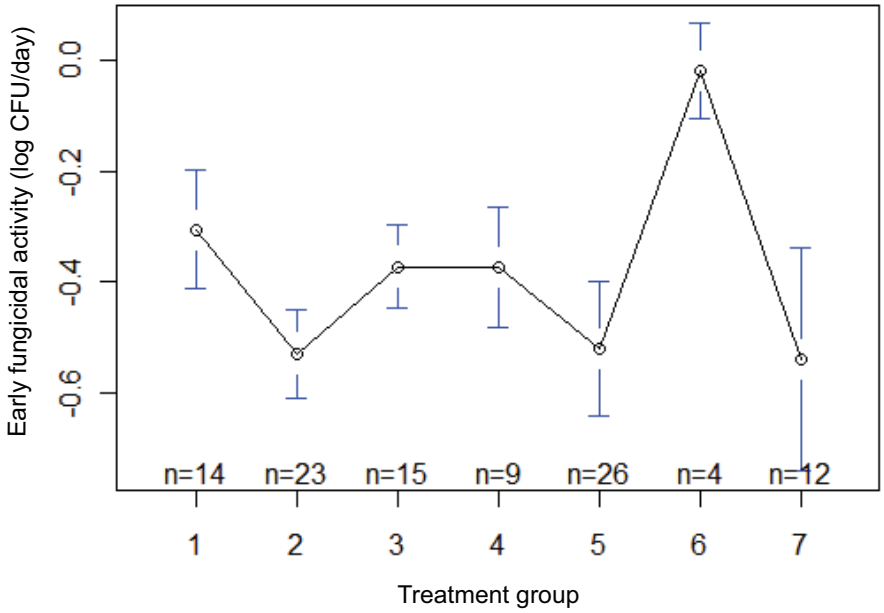
[#] cumulative death count (includes death by 2 weeks)

Table 7.4.: Coefficients of the best model fit for death at two and ten weeks following the initiation of treatment.

Death (2 weeks)	Estimate	Std. Error	T-stat	p-value
(Intercept)	2.271	2.6158	0.868	0.385
weight	−0.1139	0.0583	−1.953	0.051
altment	2.4203	0.7543	3.209	0.001

Death (10 weeks)	Estimate	Std. Error	T-stat	p-value
(Intercept)	−0.8130	0.5013	−1.622	0.105
Opbase	−0.0415	0.0198	−2.098	0.036
altment	3.0936	0.6565	4.712	0.000

Figure 7.2.: Mean early fungicidal activity of HIV+ patients according to treatment group, measured in change in log cryptococcal colony-forming units count over two weeks. 95% Confidence intervals are also displayed.



patients (range = 2 to 7). The final model contained as explanatory variables the weight, treatment, logCRAG, altered mental status, CSFWBC and logIFN (Table 7.5). Note that — since the response variable, EFA, represents the daily decrease in log CSF CFU — positive slope coefficients in the final model indicate variables where an increase leads to a slower rate of fungal clearance.

Inclusion of these additional variables did not substantially alter the differences between the drug regimens. Treatments 2 and 5 led to significantly faster initial clearance in comparison to treatment 1 (difference of 0.24, p-value ≤ 0.001, respectively), as did treatment 7 (difference of 0.2, p-value $= 0.02$). Increases in weight and cryptococcal antigen titre led to slower clearance in cryptococcal load, while increased white blood cell count and IFNγ levels increase the rate of clearance, as did a lack of altered mental status.

Table 7.5.: Final linear regression model for early fungicidal activity.

	Estimate	Std. Error	T-stat	p-value
(Intercept)	−0.5160	0.1260	−4.108	0.000
weight	0.0041	0.0021	2.021	0.047
treat2	−0.2360	0.0646	−3.653	0.001
treat3	−0.1290	0.0670	−1.919	0.059
treat4	−0.0827	0.0789	−1.047	0.298
treat5	−0.2370	0.0689	−3.443	0.001
treat6	0.1560	0.1290	1.209	0.230
treat7	−0.1980	0.0795	−2.485	0.015
logCRAG	0.0202	0.0094	2.142	0.035
altment	0.1430	0.0694	2.068	0.042
CSFWBC	−0.0001	0.0001	−2.232	0.028
logIFN	−0.1250	0.0282	−4.437	0.000

7.4.5. Genotypic effect on EFA at the allelic profile level but not ST level

Although ST was not found to have a significant effect on EFA, a number of allele types for loci *IGS1* and *PLB1* were significant in the final regression model. Infection with *IGS1* AT20 as opposed to AT19 decreased the log

clearance rate by 0.15 (p-value = 0.02) while infection with *PLB1* AT4 resulted in a 0.14 increase in comparison to AT2 (p-value = 0.04; Tables 7.6 and 7.7). In both these models, the variables, intercepts and overall regressions were all significant (p-value < 0.001; Tables 7.6 and 7.7). The model containing *IGS1* had higher adjusted R-squared (0.443 versus 0.439) and slightly lower residual standard error (0.3205 versus 0.3216) than the one with *PLB1*. ANOVA was used to compare the two regression lines and found them to not be significantly different (p-value = 0.51), meaning that the ATs of one locus are as useful as the ATs of the other in predicting early fungicidal activity in the patients analysed. Furthermore, combining the two in a single model resulted in a loss of significance of both due to insufficient data.

Table 7.6.: Final linear regression model for *IGS1*.

	Estimate	Std. Error	T-stat	p-value
(Intercept)	−0.5560	0.1240	−4.484	0.000
weight	0.0046	0.0020	2.282	0.025
treat2	−0.2549	0.0637	−4.001	0.000
treat3	−0.1337	0.0656	−2.038	0.045
treat4	−0.1019	0.0777	−1.312	0.193
treat5	−0.3012	0.0727	−4.144	0.000
treat6	0.1266	0.1271	0.996	0.322
treat7	−0.2394	0.0803	−2.983	0.004
logCRAG	0.0231	0.0094	2.465	0.016
altment	0.1458	0.0682	2.138	0.036
IGS120	0.1541	0.0654	2.358	0.021
IGS121	0.0757	0.0842	0.899	0.371
CSFWBC	−0.0001	0.0001	−2.218	0.029
logIFN	−0.1320	0.0281	−4.701	0.000

7.4.6. Survival analysis

The survival functions for each of the three studies were estimated by the Kaplan-Meier estimator, implemented with the 'survival' package in R. Using the log-rank test (equivalent to the Mantel-Haenszel test), significant differences in survival time were found between patients grouped by the risk factors ST, treatment group and altered mental status (p-value = 0.004,

Table 7.7.: Final linear regression model for *PLB*.

	Estimate	Std. Error	T-stat	p-value
(Intercept)	−0.5712	0.1265	−4.516	0.000
weight	0.0048	0.0020	2.341	0.022
treat2	−0.2418	0.0669	−3.616	0.001
treat3	−0.1361	0.0660	−2.064	0.042
treat4	−0.1014	0.0780	−1.300	0.197
treat5	−0.2934	0.0739	−3.972	0.000
treat6	0.1414	0.1294	1.092	0.278
treat7	−0.2176	0.0869	−2.504	0.014
logCRAG	0.0240	0.0095	2.529	0.013
altment	0.1425	0.0686	2.077	0.041
PLB13	−0.0417	0.0624	−0.669	0.505
PLB14	0.1405	0.0687	2.044	0.044
PLB111	0.0636	0.0864	0.736	0.464
CSFWBC	−0.0001	0.0001	−2.272	0.026
logIFN	−0.1323	0.0282	−4.693	0.000

0.014 and ≤ 0.001, respectively; Figures 7.3, 7.6 and 7.7). With only 20% survival by day 150, infection with *Cng* ST32 (Figure 7.3; green line) yielded the lowest probability of survival compared to STs 4, 5 and 6 (Figure 7.3; blue, black and red lines, respectively). The proportional survival for patients within these three ST groups did not fall below 70%.

As the survival probabilities associated with the different ST's may be confounded by the geographic structure, further survival curves were plotted. STs 4, 5, and 6 infecting the Thai population remained associated with high survival probability ($>$ 70% by day 150; Figure 7.4). African STs 5 and 6 resulted in similarly high survival probabilities by day 150 ($>$ 75%; Figure 7.5). STs restricted to the Cape Town patients resulted in 40% mortality by day 150. Most interesting was the effect on continental origin on infecting ST32. Overall, this was associated with the lowest probability of survival (Figure 7.3; orange line), an effect which remained constant in the African population. Among the Asian patients, however, ST32 was associated with no deaths at all (Figure 7.5).

Survival analysis echoed the previous findings on variables linked to poor outcome of infection. Estimated percentage of survival was at 0 before 100 days for those treated with fluconazole only. Seventy-five percent survival

duration was highest for treatment 2, at 300 days (Figure 7.6). Survival beyond 400 days was estimated at 60% for patients presenting without altered mental status, compared to 30% for those with evidence of cerebral dysfunction (Figure 7.7). Again, ATs of loci *IGS1* and *PLB1* were of significance (logrank, p-value = 0.002 and 0.001, respectively). By day, 200, the probability of survival was estimated at 80% for *PLB1* AT2, but 60% or less for ATs 3, 4 and 11 (data not shown). *IGS1* AT19 was associated with significantly better survival probability than ATs 20 and 21.

Figure 7.3.: Survival of HIV-positive patients suffering from cryptococcal meningitis according to the six infecting Sequence Types (ST). ST4 contained only Thai patients while ST6 contained all but one Thai patients. ST5 and ST32 each consisted of all but one African patients. ST51 and ST52 contained only African patients.

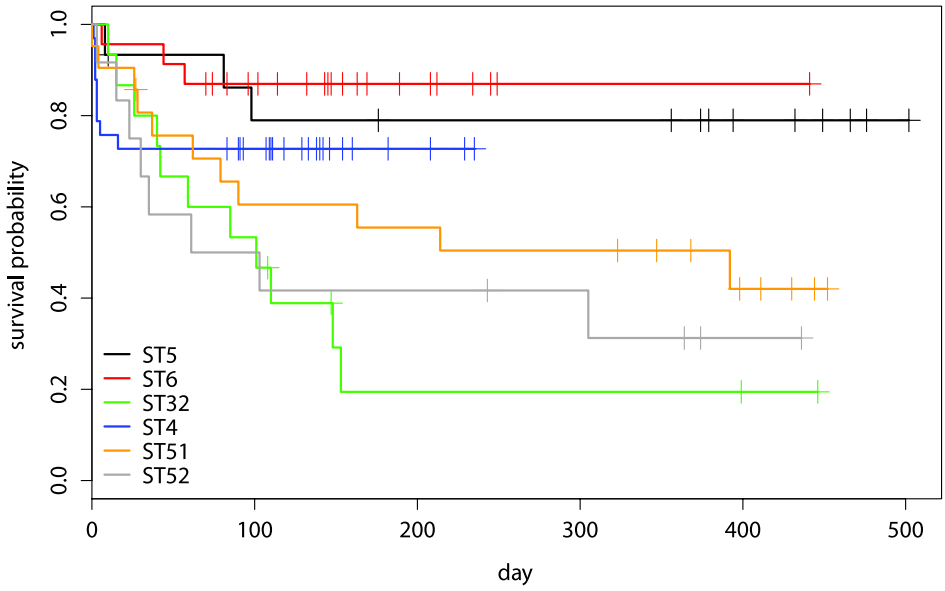


Figure 7.4.: Survival of HIV-positive patients from Thailand suffering from cryptococcal meningitis according to infecting ST.

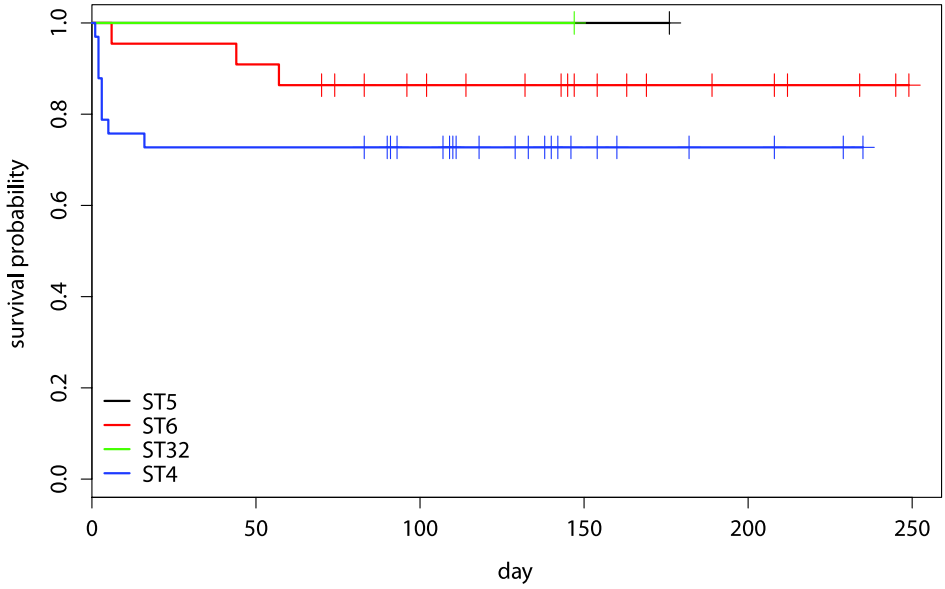


Figure 7.5.: Survival of HIV-positive patients from South Africa suffering from cryptococcal meningitis according to infecting ST.

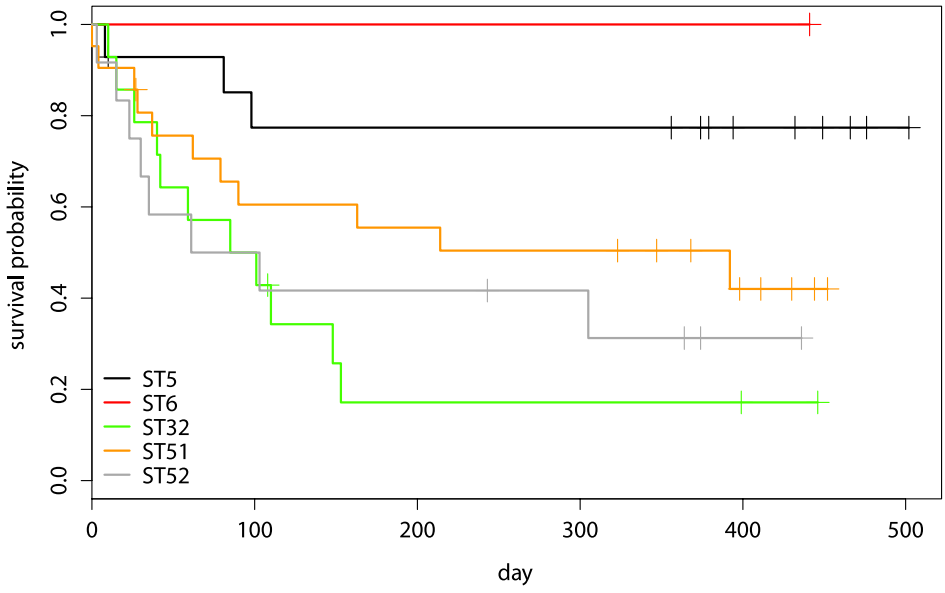


Figure 7.6.: Survival of HIV-positive patients suffering from cryptococcal meningitis according to treatment group. Treat 1 = AmB only; Treat 2 = AmB (0.7 mg/kg/day) + flucytosine; Treat 3 = AmB (0.7 mg/kg/day) + fluconazole; Treat 4 = AmB + flucytosine + fluconazole; Treat 5 = AmB (1 mg/kg/day) + fluconazole ; Treat 6 = Fluconazole only; Treat 7 = AmB (1 mg/kg/day) + flucytosine.

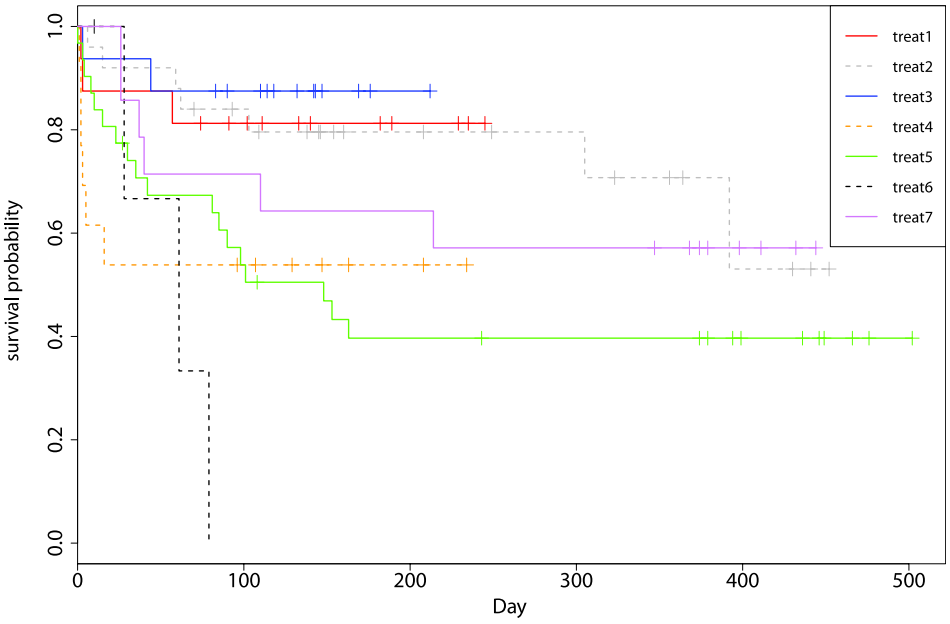
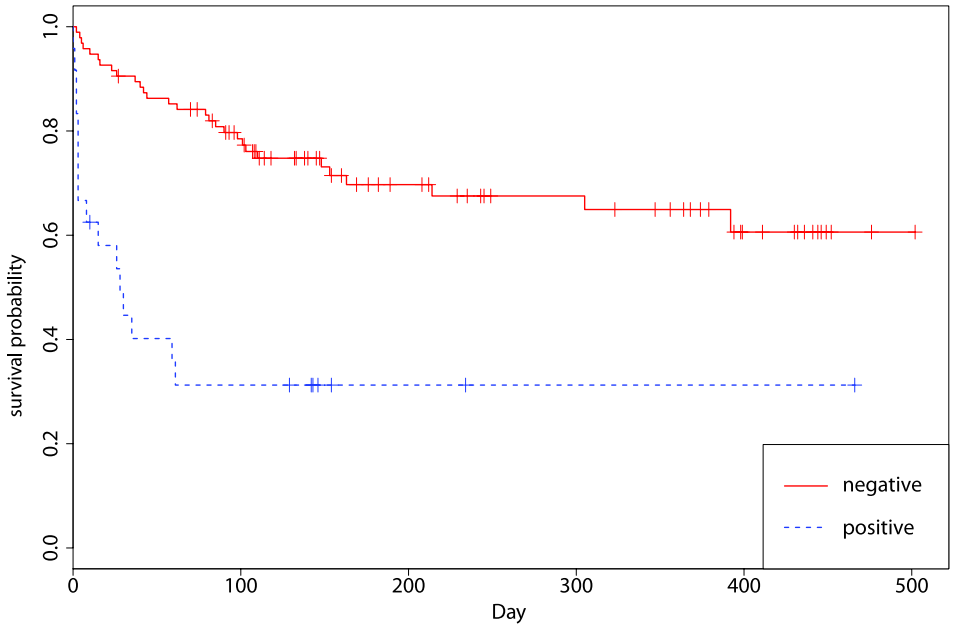


Figure 7.7.: Survival of HIV-positive patients suffering from cryptococcal meningitis according to diagnosis of altered mental status at presentation. The negative (red) line represents those who did not present with an altered mental state, while the purple line represents patients who did.



7.5. Discussion

The aim of this chapter was to identify factors which can predict the outcome of HIV-associated cryptococcal meningitis. Infection and disease are dependent on both the host and on the pathogen (Finlay and Falkow, 1989), and so both clinical and molecular data were analyzed. The patients of three clinical trials were stratified by genotype of infecting *Cng*, determined by MLST earlier in this study (Chapters 3 and 4). Prognostic factors of death and early fungicidal activity (EFA), the outcome measures, included altered mental status at presentation, low interferon gamma (IFNγ) levels in the cerebrospinal fluid, and high cryptococcal antigen titres (CRAG; Tables 7.4 and 7.5). Although sequence type (ST) was not associated with either of the outcome measures including death by weeks two or ten, it was significantly associated with probability of survival longer term. In addition, associations between clinical phenotypes and allele types of individual loci were identified.

ST was significantly associated with a number of baseline variables (Table 7.2). Ninety-seven percent of patients (61 out of 63) infected by STs 5, 32, 51 and 52 were from Cape Town, and over 67% were female (Fisher's exact p-value ≤ 0.001; Tables 7.1 and 7.2). A retrospective study of CM infections in Durban, South Africa, between January, 1991 and December, 1994 reported a similar female predominance, with a male to female ratio of 0.63 (Moosa and Coovadia, 1997). These findings reflect the pattern of the AIDS epidemic in South Africa, which is driven by heterosexual infections, resulting in the majority of cases being among women (Department of Health, 1995). In Asia, on the other hand, an estimated 75 million men regularly buy sex from female sex workers and are the single most powerful driving force of the epidemic (Hein, 2008). This contrast is reflected by elevated percentages of males in the two ST groups, which collectively contain all but one of the Thai patients, STs 4 and 6 (Table 7.2).

Comparisons of the means of baseline variables significantly associated with ST revealed a pattern: patients with low immune responses had high viral loads of *Cng* and vice versa. For example, patients infected with *Cng* ST4 presented with low IFNγ levels and CD4 cell counts, while their quantitative cryptococcal culture counts (QCC) were elevated (Figure 7.1). The

opposite is seen in patients of ST32, whose elevated immune parameters correspond to comparatively lower markers of fungal burden (CRAG and QCC). A comparatively high level of cryptococcal burden suggests a high proliferation rate of the invading pathogen, attributable to either the host mounting a low immune response, or the to the successful evasion or neutralisation mechanisms implemented by the pathogen (Bulmer et al., 1967a).

Continental differences in genotypes were identified in the frequency of STs. Survival analysis revealed that STs 4 and 6 of Thai origin resulted in the highest probability of survival, while the predominantly African STs (32, 51 and 52) resulted in the lowest (Figure 7.3). The log rank test showed that the mean survival rates of the STs within Thailand were not significantly different from each other (log rank p-value = 0.16). Similarly, the African STs were not significantly different from each other in mean survival probability (log rank p-value = 0.38). This suggests that infecting isolates of Cape Town origin are more virulent than those responsible for cryptococcal disease in the Thai population. Of the STs which presented in both the Asian and African patient groups, STs 5 and 6 correlated in high survival probabilities despite geographic origin (Figures 7.4 and 7.5). Interestingly, survival plots of geographically defined STs revealed that ST32 which resulted in 50% mortality by day 100 in the African patients was associated with 0% mortality in the Thai population. Although there was only one Thai patient infected by this ST, this finding strengthens the possibility of geographical influence on the isolates. This hypothesis is investigated using a *Galleria mellonella* animal model, the results of which are detailed in Chapter 9.

Microbial pathogenicity is governed by factors which can be classified into two groups, both of which can be targets for antifungal drugs. The first group consists of virulence factors unique to the fungus which are involved in the direct harming or circumventing of the host cells, e.g. the production of toxins and hydrolytic enzymes, and tissue adhesions (Finlay and Falkow, 1989). The factors of *Cng* which fall within this group are well documented and include melanin synthesis and the ability to proliferate at high temperatures (Polacheck et al., 1990; Perfect, 2006b; Kwonchung et al., 1982). The second group contains factors which maintain metabolic functions essential to the survival and proliferation of the pathogen (Perfect

et al., 1993). Both groups of factors contribute to the successful persistence of cryptococcal infection and are governed by specific genes, as evidenced by knock-out animal models (Perfect, 2006a; Buchanan and Murphy, 1998). Because of this, associations between *Cng*'s genetic makeup and virulence were assessed at the seven individual loci typed by MLST, and not just as a complete allelic profile represented by ST. Two clinical measures of fungal burden were associated with allele types of five *Cng* genes: quantitative cryptococcal culture count (QCC) with *IGS1*, *PLB1* and *URA5*; and cryptococcal antigen titre (CRAG) with *CAP59*, *GPD1* and *PLB1* (data not shown). Also, *IGS1*, *PLB1* are independently significant in the final regression model for EFA (Tables 7.6 and 7.6). *IGS1* and *URA5* are involved in "housekeeping" functions, allowing for the proliferation and pathogenicity of the cells within the host, while *PLB1* and *CAP59* are directly involved in the pathogenesis of *Cng*. Owing to high levels of linkage disequilibrium in my dataset, these alleles are likely markers for other, unknown, virulence factors and not necessarily alleles at these genes (which were chosen as are putatively neutral).

Two genes are found within tandem repeats of fungal ribosomal RNA and between three slowly evolving and relatively conserved coding regions (18S, 5.8S and 28S; White et al., 1990). The first is the internal transcribed spacer (ITS), which has been applied in the characterization of several medically important microorganisms, including *Candida*, *Malassezia* and the *Trichosporon* species (Turenne et al., 2000; Sugita et al., 1999; Makimura et al., 2000; Lott et al., 1998; Gupta et al., 2000; Chen et al., 2000). The second is the intergenic spacer region (*IGS1*), which has been also been used in the comparison of the genetic diversity within species and subspecies. The latter is the most variable region in the rRNA gene, and is therefore more powerful than ITS in the differentiation of phylogenetically closely related species — the varieties of *Cn* included (Xu et al., 2000; Diaz et al., 2000; Sugita et al., 2001; Sugita et al., 2002). Complete *IGS1* regions have been detailed for *Neurospora crassa* (Dutta and Verma, 1990), *Saccharomyces cerevisiae* (Molina et al., 1993), *Neotyphodium lolii* (Ganley and Scott, 1998), the *Trichosporon* species (Sugita et al., 2002) and *Cn* (Fan et al., 1995). Analyses described in Chapter 6 of in this study identify *IGS1* as having the highest number of segregating sites and haplotypes within the

African, North American and South American *Cng* populations (Table 6.4), in keeping with it being the most variable single locus investigated so far (Sutar et al., 2004). Despite not being directly implicated in the virulence of strains, the association identified with QCC supports the occurrence of recombination being linked to virulence (Litvintseva et al., 2011) — recombination potential which has previously resulted in the formation of new and highly virulent genotypes in the outbreak in British Columbia (Fraser et al., 2005).

Also associated with CSF QCC was *URA5* (ANOVA p-value < 0.05). This is the protein encoding pyrimidine biosynthesis, a "housekeeping" function referred to as orotidine-5'-monophosphate pyrophosphorylase (OMPase) by Yu et al. (1999). A murine model showed significantly decreased virulence in mutant *ura5 Cn* cells in comparison to wild-type parent cells (Varma et al., 1992). This pathway is important in other fungi including *Candida albicans* (Fasoli et al., 1990; Kirsch and Whitney, 1991), *Coccidiodes immitis* (Yu et al., 1999), and *Histoplasma capsulatum* (Retallack et al., 2000). Importantly, the antimetabolite drug flucytosine acts on this pathway. By competitively inhibiting pyrimidine uptake, fungal DNA and RNA synthesis is also inhibited (Bennett, 1977). As an estimated 2% of *Cn* are resistant to this drug from the onset of treatment, flucytosine is administered with AmB as opposed to as monotherapy. This also results in synergism between the two (Richardson and Warnock, 2008).

The secretion of phospholipase is a virulence mechanism which several fungi such as *Candida albicans*, *Candida glabrata* and the *Aspergillus* species (Ghannoum, 2000) use to cause damage to host cells. This genetic strategy is also observed in parasitic protozoa such as *Toxoplasma gondii* and *Entamoeba histolytica* (Saffer et al., 1989; Saffer and Schwartzman, 1991; Longkrug et al., 1985), and bacteria (e.g. *Clostridium perfringens*, *Listeria monocytogenes*, *Pseudomonas aeruginosa*, *Bacillus cereus*, *Rickettisia* and *Corynebacterium pseudotuberculosis*; Titball and Rood, 2000; Songer, 1997; Titball, 1993; Smith et al., 1995; Rahmet-Alla et al., 1990; Silverman et al., 1992; McNamara et al., 1994). First described in *Cryptococcus* in 1997 (Vidotto et al., 1997), the secretion of phospholipase was correlated with virulence through the quantification of *Cryptococcus* in the lungs and brains of mice (Chen et al., 1997). The phospholipase gene, *PLB1*, is in-

volved in intracellular growth and the production of eicosanoids. The latter are potent regulators of host immune responses which, when released by the pathogen, are protective and downregulate macrophage function at the site of infection (Noverr et al., 2003a; Noverr et al., 2003b).

CAP59 is essential for capsule formation in *Cn* (Chang and Kwonchung, 1994), with acapsular mutants being characterised by reduced virulence (Bulmer et al., 1967b; Fromtling et al., 1982; Zaragoza et al., 2003; Kozel and Gotschlich, 1982). The polysaccharide capsule is composed mainly of glucuronoxylomannan (∼88%; GXM), galactoxylomannan and mannoprotein (MP; Cherniak and Sundstrom, 1994). It is antiphagocytic, protective against environmental stress (Casadevall and Perfect, 1998; Zaragoza et al., 2008; Kozel and Mastroianni, 1976), and interferes with the development of effective immune responses by the host, both as part of the cell and when shed (Kozel et al., 1977; Murphy and Cozad, 1972; Buchanan and Murphy, 1998, Vecchiarelli, 2000; Vecchiarelli et al., 2003). The variation in virulence across serotypes has been correlated to the chemical structure of GXM. Evidence of this is the selective infection of immunocompromised patients by *Cng* (serotype A) strains, as well as the finding that the serotypes of acapsular *Cn* cells are untypeable (Jacobson et al., 1982; Kozel and Cazin, 1971). Similarly, the polysaccharide's ability to induce immune unresponsiveness through the inhibition of neutrophil migration and antibody recognition has also been experimentally correlated to the chemical structure of the major antigen of the polysaccharide capsule (Ellerbroek et al., 2004). The association of *CAP59* with variations in clinical patterns observed in the HIV+ patients of this study is therefore not surprising, as cryptococcal antigen titre is a measure of virulence mechanisms involving the capsule and governed by the gene in question. In terms of research on the effects of antifungals on the pathogenesis of *Cn*, a murine model of infection revealed that AmB reduces the size of the capsule and decreases serum polysaccharide (Zaragoza et al., 2005).

Cerebral dysfunction at baseline was found to be a strong prognostic factor of early and longer term death under univariate analysis, in keeping with previous findings (Brouwer et al., 2004; Saag et al., 1992; Heyderman et al., 1998). Fitting a logistic regression model to the probability of death by week 10, altered mental status was found to have a significant effect, as

did opening base pressure (Table 7.4). Seventy-five percent of CM patients suffered from elevated intracranial pressure. This is caused by the migration of high levels of inflammatory cells attempting to counter cryptococcal invasion to the CSF, impeding CSF reabsorption at the arachnoid granulations (Loyse et al., 2010). Satishchandra et al. (2007) and Graybill et al. (2000) report significant differences in survival probabilities are reported between patients with baseline CSF pressure < 190 and those with ≥ 350 and even ≥ 250 mm H_2O (p-value $= 0.008$ and 0.05, respectively; Satishchandra et al., 2007; Graybill et al., 2000). Graybil et al. also report an association between elevated baseline opening pressure and decreased long-term survival, similar to what is seen in this study, with Opbase being significant at week 10, but not at week 2 (Table 7.4). Seizures, decreased mental status and intracranial pressure are indicative of considerably advanced infection and successful haematogenous dissemination to the CNS. This tends to occur in the later stages of disease upon the failure of the host's immune system to contain the infection at the primary site (the lungs), leading to increased cryptococcal burden at the *Cng*'s preferred site, the CNS (Stevens et al., 1999). It is therefore not surprising that altered mental status and high opening pressure are significantly associated with death, a finding mirrored in the survival analysis (Figure 7.7).

The most effective antifungal regimens in the management of CM among the HIV+ patient sample were combination therapies of Amphoterecin B and flucytosine (Figure 7.2, Table 7.5), as found in an independent dataset (van der Horst et al., 1997). Bicanic et al (2008) detected faster clearance rates in patients placed on a higher dose of AmB, with 1 mg/kg per day plus flucytosine being more rapidly fungicidal than the standard dose of 0.7 mg/kg per day (difference of 0.11 log CFU per day; p-value $= 0.05$). Several studies report poor outcomes of fluconazole therapy alone, with a reported median survival time of 19 days (Mwaba et al., 2001; Schaars et al., 2006; Larsen et al., 1990). As with AmB, better prognosis has been associated with a higher dosage of daily fluconazole monotherapy, with 1 200 mg resulting in a mean EFA of -0.18 ± 0.11 log CFU/mL. This is still comparatively poorer than clearance rates reported for AmB-based combination therapies, both in this study and previous ones (clearance rates of $0.31-0.56$ log CFU/mL per day; Brouwer et al., 2004; Bicanic et al., 2007; Bicanic et

al., 2008). It is important to note that the patients included in this study were placed on this treatment arm as they presented with the most severe disease symptoms (Glasgow Coma Score of < 10; Bicanic et al., 2007).

In a final multivariable regression, logCRAG, altered mental status, weight, CSFWBC and logIFN were found to be strong prognostic factors of clearance of cryptococcal culture. Cryptococcal antigen titre is a simple, rapid and sensitive screening method used in the diagnosis of CM (positive diagnosis >1:8; Chuck and Sande, 1989). Although elevated CSF CRAG levels are indicative of poor prognosis in patients with AIDS (Asawavichienjinda et al., 1999; Antinori et al., 2001; Larsen et al., 1990) and CRAG screening is believed to be a highly cost-effective strategy in the prevention of CM-related mortality (Meya et al., 2010), the use of CRAG in the monitoring of the efficacy of treatment is challenged. Studies report no correlation between changes in CSF or serum cryptococcal antigen and the outcome of CM, which suggests that CRAG is a poor index of cure (Powderly et al., 1994; Antinori et al., 2005; Lu et al., 2005). Despite not being a suitable host marker of response to treatment and progress in CM therapy, CRAG is the only fungal determinant found to be associated with early fungicidal activity of therapy in this study, along with the associated *PLB1* gene.

In HIV+ patients, CM is associated with profound immunosupression characterised by poor CSF inflammatory response, quantified by low CD4 and white blood cell counts (<100 and <20 per μL, respectively; Bicanic and Harrison, 2004; Chuck and Sande, 1989). The levels of IFNγ — a cytokine critical for innate and adaptive immunity — at the site of infection determine the rate of clearance of *Cn* in CM, with significantly higher rates reported among survivors (Siddiqui et al., 2005; Chen et al., 2005). Murine models show that the administration of therapeutic IFNγ potentiates AmB therapy in the treatment of CM, a finding which is of particular interest among the immunocompromised who suffer from diminished immune responses (Joly et al., 1994; Lutz et al., 2000). Elevated white bloods cell counts in the CSF, an indication of dissemination of *Cng* to the CNS, was found to be significantly associated with improved clearance of *Cng* (Table 7.5). CD4 cell count, also a measure of host immune response, was not found to be associated with either death or EFA. The combined cohort used in this study was made up of patients with already very low CD4 cell count

(median = 21, IQR = $8 - 36.2$), most likely the result of delayed diagnosis, delayed imitation of therapy, and advanced HIV+ infection. In the Mycoses Study Group study, 60% of patients had opening CSF pressures of at least 25 cm H_2O (Graybill et al., 2000). Unlike the MSC study, however, I did not find an elevated pressure to be significantly associated with cryptococcal clearance ($\geq$20 cm H_2O; $n = 63$ of 102). This was likely the result of the CSF pressure of patients included in these studies being monitored and controlled through additional lumbar punctures, administered to patients with symptoms attributable to elevated pressure (Saag et al., 2000; Bicanic et al., 2007; Bicanic et al., 2008).

The findings presented in this chapter are in keeping with reports of combination antifungal therapy (AmB and flucytosine) being the current optimal treatment for HIV-associated CM. There is evidence of improved outcome with changes in dose, and so further work needs to be carried out to optimise treatment of this opportunistic infection within the growing vulnerable population. The identification of suitable markers of response to treatment is key to progress in CM therapy, and has been well established. This study identified some factors affected by the genetic makeup of the pathogen. The associations at the AT level show the value of using genes as targets for the action of antifungal agents, genes which have been rightly included in the MLST loci profile. Limitations of analyses were mainly due to the retrospective nature of the study and therefore the small sample sizes. Even so, this study makes a strong case for the prediction of and progress monitoring of CM infection using molecular genetics and clinical data.

8. Environmental Cape Town isolates

Acknowledgements: The work presented in this chapter was made possible through a collaboration with the University of Cape Town and the members of the Institute of Infectious Disease and Molecular Medicine.

8.1. Introduction

Several molecular typing studies have identified genetically identical clinical and environmental strains of *Cn* (Litvintseva et al., 2005; Currie, Freundlich and Casadevall, 1994; Sorrell et al., 1996; Yamamoto et al., 1995; Litvintseva et al., 2006). This is evidence of the presence of virulent strains in the environment and indicative of potential sources of cryptococcal infection within the human patient population. In Chapter 5, MLST typing was applied to *Cn* samples isolated from 118 patients enrolled in two clinical studies in Cape Town, South Africa (Bicanic et al. 2007; Bicanic et al. 2008). Here, I describe a field study that I undertook during the month of November 2009 in Cape Town, South Africa. During this time, I attempted to isolate *Cng* from the environment in which the patients from the Cape Town clinical trials were exposed to the fungus. These isolates would be typed and the structure of environmental *Cng* described.

The aims of the field study were

1. To identify niches of environmental sources of *Cng* in Cape Town;
2. To investigate the potential link between environmental *Cng* and the disease described in the Cape Town HIV+ patient population using MLST; and

3. To establish whether the all environmental isolates are equally virulent, or if there exists a genetically isolated sub-set within these responsible for the disease observed in the Capetonian patient population.

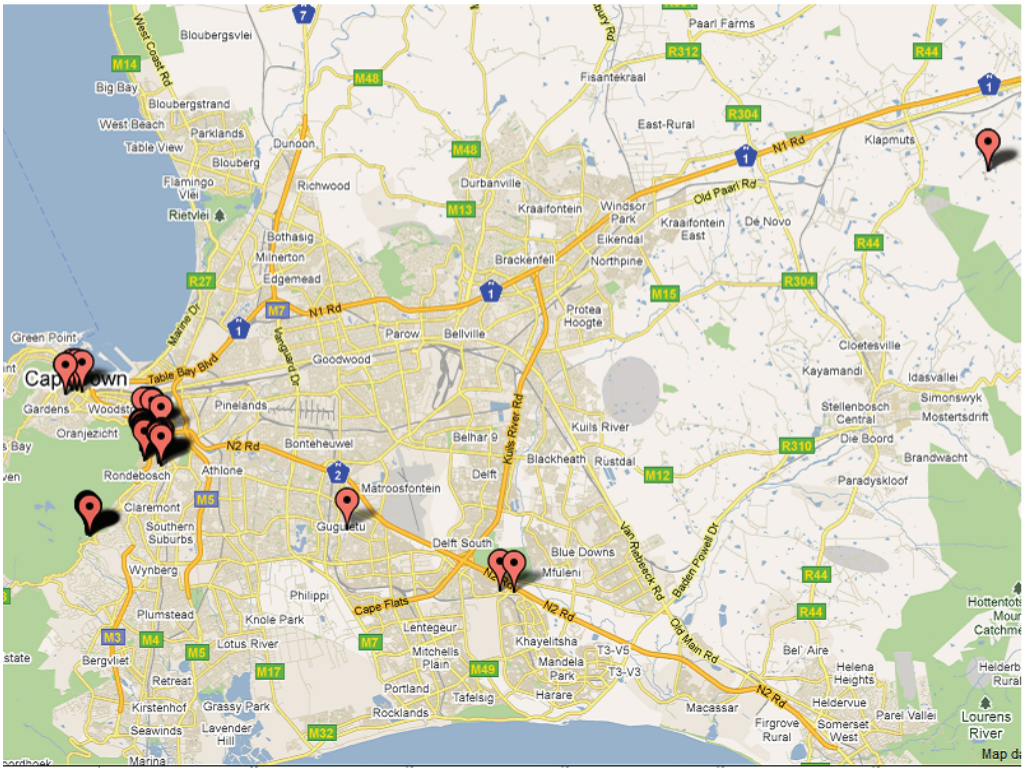
8.2. Material and Methods

The study was conducted in Cape Town, South Africa, between the 1st and 30th of November and December, 2009. Soil, pigeon excreta and feather samples were collected from multiple environmental locations across the city (Figure 8.1) and according to the protocol described in Section 2.9. The samples were then cultured and stored within two days of collection. A month later, DNA was extracted from sub-cultured plates of the environmental samples and sequenced at *ITS1F* (Gardes and Bruns, 1993) and *ITS4* (White et al., 1990) — primers which allow for the determination of the genus of fungi. Sequences were aligned and species queried using the Basic Local Alignment Search Tool (BLAST) available at the National Center for Biotechnology Information (NCBI, `http://blast.ncbi.nlm.nih.gov/Blast.cgi`). The samples identified as being of the cryptococcal species were then sequenced at the seven MLST loci in order to confirm the isolation of *Cryptococcus neoformans* var *grubii* and, if so, to determine their genotype. The sequence types (STs) of any *Cng* isolates were to be compared against the clinical Cape Town *Cng* population (described in Chapter 5), and integrated into the global *Cng* population (Table B.1).

8.3. Results

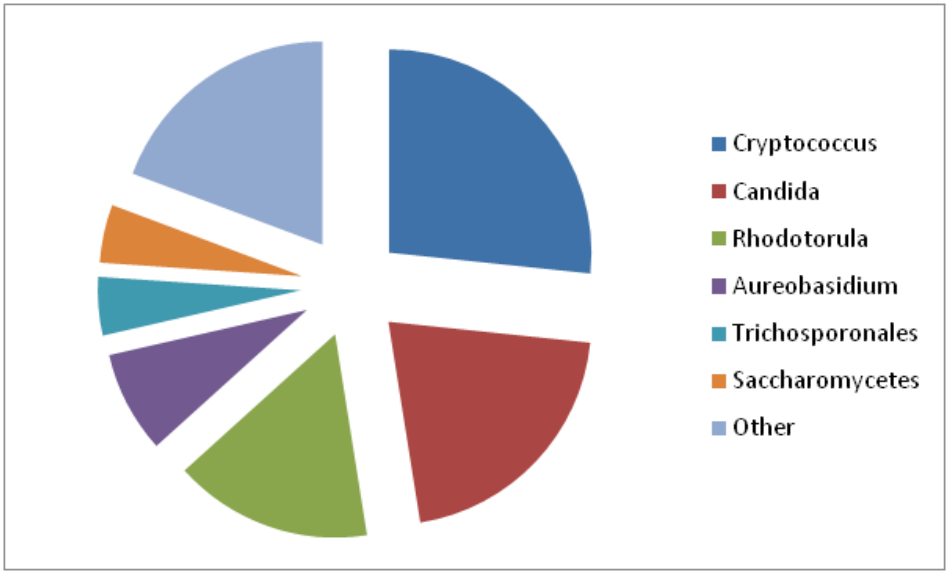
A total of 282 samples were collected, 259 of which were successfully sequenced at the *ITS1* gene. Three of the four phyla of the fungal kingdom were represented among these isolates: *Ascomycota* ($n = 122$), *Basidiomycota* ($n = 126$) and *Zygomycota* ($n = 3$). Sixty nine of the 259 samples were of the cryptococcal species (27%). Other frequently isolated fungal species were *Candida spp.* ($n = 54$), *Rhodotorula spp.* ($n = 41$) and *Aureobasidium spp.* ($n = 21$; Figure 8.2).

Figure 8.1.: Map of the areas of environmental sampling of *Cryptococcus* in Cape Town, South Africa.



Of the 69 samples of the *Cryptococcus spp.*, BLAST identified 24 *C. laurentii* isolates, 12 *C. flavescens* and four *C. gattii* isolates. The remaining 29 isolates were not differentiated further than by species in BLAST (*Cryptococcus spp.*). No isolates aligned with *Cng*. Thirty-three of the cryptococcal samples were isolated from trees (48%), twenty-two from soil and the remaining 14 from avian excreta (32% and 20%, respectively). PCR amplifications and sequencing at all seven loci confirmed the presence of four *C. gattii* isolates (E64, E156, E157 and E159; data not shown). All but E64 were taken from tree swabs, the latter being from soil. A fifth isolate, E194, was of allele type AT5 at the *GPD1* locus, but failed to amplify at the remaining loci. Amplification with serotype-specific mating primers (Table 2.1, page 49) also failed to determine the molecular group of this isolate. All five samples were isolated from the Lower Campus of the University of Cape Town, (long: -33.9547, lat: 18.469658).

Figure 8.2.: The distribution of fungal species isolated from the environment in Cape Town, South Africa.



8.4. Discussion

In order to determine the environmental occurrence of *Cng* in the city of Cape Town, South Africa, 282 samples were taken from trees, soil and bird droppings at typically crowded places. These sites included two townships, Gugalethu and Khayelitsha, a central bus station and municipal park in the city center, and university campuses. A total of 65 sites yielded 69 isolates of the *Cryptococcus* species, but none were confirmed to be *Cng* by MLST. *C. gattii* was recovered from one of the university of Cape Town campuses, along with a sample which was of undetermined serotype (E194). Isolate E194 consisted of AT5 at the *GPD1* locus and its being untypeable at the remaining six could be indicative of an AD hybrid, contamination or impure DNA.

Unfortunately, no comparisons could be made between the genotypes of the clinical and environmental populations of *Cng* in Cape Town as intended. The low success in isolating *Cng* from the environment contradicts findings of environmental strains being commonly associated with pigeon excrement, soil and decaying wood, as well as the evidence of human infections occurring as a result of exposure to environmental sources. Most recently, environmental sampling of 440 locations in Botswana and South Africa yielded 273 *Cng* isolates (Litvintseva et al., 2011). Of the 22 sites which were positive for the fungus, 16 were tree or soil sites and six of avian guano. Molecular typing revealed both the presence of molecular types VNI and VNB within this sample, with the latter being associated with the endemic and widely used Mopane tree, while the former was mainly found in avian excreta and in urban areas. The findings of this study stresses the importance of global typing in identifying new ecological niches of the fungus. Almost thirty percent of the sampled Mopane trees, *Colophospermum mopane*, hosted potentially virulent *Cng* strains. These trees are a source of firewood, hut-making, local medicinal remedies, as well as the sole home of the culinary delicacy mopane worm (*Gonimbrasia belina*). Similar to a study of patients in Burundi which detailed the potential of reinfection of cured CM – AIDS patients upon returning to their homes, the original environment of infection (Swinne et al., 1989), findings such as these have important Public Health implications.

There are several possible reasons for *Cng* not being successfully isolated from Cape Town. As mentioned above, the main tree associated with environmental *Cng* in Southern Africa is the mapone tree (*Colophospermum mopane*; Litvintseva et al., 2011). This is not native to Cape Town and so tree sampling was performed at random. A more systematic study noting tree species sampled would be more informative of potential and novel environmental cryptococcal niches.

Studies from Asia report seasonal variations in the isolation of *Cng* from environmental niches. Kuroki et al. (2004) reported a decrease from 24% in the dry season to 4% during the rainy season. Similarly, a seven-year retrospective study by Randhawa et al. (2011) described the prevalence of *Cng* as lowest during the rainy season. October and November mark the end of the rainy season and the beginning of the summer in South Africa and could be characterised by low aerial dispersal of the cryptococcal basidiospores, hence no isolated *Cng*. As *Cng* is thermotolerant (Casadevall and Perfect, 2006; Kwon-Chung and Bennet, 1992), its prevalence in Cape Town may be increased between the summer months of January to March.

The isolation of four *C. gattii* samples suggests that this species is more abundant in the environment than initially believed (Karstaedt et al., 2002; Litvintseva, 2005). Three of the isolates were from trees, and the fourth from a soil sample in keeping with strong evidence of the association between *C. gattii* and trees, but not with avian excreta. Granados and Castaneda (2004) postulate that this may be due to the higher pH levels in ecrement compared to that in vegetal samples. It is interesting to note that there were two *C. gattii* infections among the Cape Town patients (isolates C28 and R42; data not shown). This suggests that, although uncommon, there exist virulent *C. gattii* strains in the environment which are causing disease within the HIV+ population. This association of *C. gattii* with AIDS patients has been reported in Botswana (14% of patients), Malawi (13% of patients) and Soweto, South Africa (Litvintseva et al. 2005; Karstaedt et al., 2002). A retrospective study of the trends in AIDS-associated CM in Kigali, Rwanda, reported 2% incidence (Bogaerts et al., 1999), closer to the percentage of *C. gattii* isolated in the clinical Cape Town population typed in this study. In light of the greater severity of infection by this species of *Cryptococcus*, Litvintseva et al. (2005) suggest that the incidence of *C.*

gattii in this highly vulnerable population may well contribute to the substantially poor prognosis of cryptococcal infeciton in the important setting of the sub-Saharan African region.

Although no new information has been contributed to the ecology and epidemiology of *Cng* in Cape Town, this work reveals that the methodology applied is robust for the isolation of cryptococcal strains. In addition, it suggests that perhaps the environmental sources of *Cng* are not exhausted. Lastly, *C. gattii* is readily present in the environment of Cape Town, South Africa, unlike previously thought. It is clear that further sampling and typing is required from a wider range of environmental substrates to elucidate the true population structure of this pathogen, identify principal environmental sources of infection and ascertain the distribution of STs recovered from both patients and the environment. In addition, molecular methods may reveal the genetic factors which are allowing *C. gattii* to expand its target host into this region.

9. *Galleria mellonella* - *Cng* virulence assay

9.1. Introduction

The characterisation of *Cng* in Thailand and Cape Town using MLST revealed genetic variation between isolates (Chapters 4 and 5). The integration of these data with clinical indicators of cryptococcal infection and dissemination among HIV+ CM patients hinted at a correlation between geographically defined genotypes and survival outcome of disease (Chapter 7). Disease is one of the several possible outcomes of infection, and, more importantly, is not purely a host nor parasite phenomenon (Casadevall and Pirofski, 2001). Instead, it is an outcome of the host-pathogen interaction which, in the case of the Thai and Cape Town patients, is complex and further confounded by factors such as variations in treatment regime. In order to better understand these variations in virulence, a simpler host-pathogen model is required; hence the use of a *Galleria mellonella* - *Cng* model. This model was selected due to the host-pathogen interaction being similar to that seen in cryptococcosis among humans. Discussed in Chapter 1, these include the potential to experiment at high temperatures, *Cng*'s ability to proliferate and kill *G. mellonella*, and the fact that the host's immune response to *Cng* involves phagocytosis.

9.2. Aims

The aim of the work presented in this chapter was to ascertain whether clinical *Cng* isolates of different STs vary in virulence, using *G. mellonella* as an animal model, and, if so, to establish whether there is a correlation between virulence and geographic origin.

9.3. Materials and Method

9.3.1. Isolates

Seven STs of molecular group VNI were represented in this assay: STs 1, 4, 5, 6, 23, 32 and 53. These were selected according to frequency, as the genotypes responsible for the majority of the disease observed in the clinical population are likely to be the most virulent and fit. Three of these STs were shared between Thailand and South Africa (STs 5, 6 and 32), and a single isolate was taken from each continent. ST4 was unique to the Thai population, while STs 1, 23 and 53 were unique to the Cape Town *Cng* population. The isolates used were as follows:

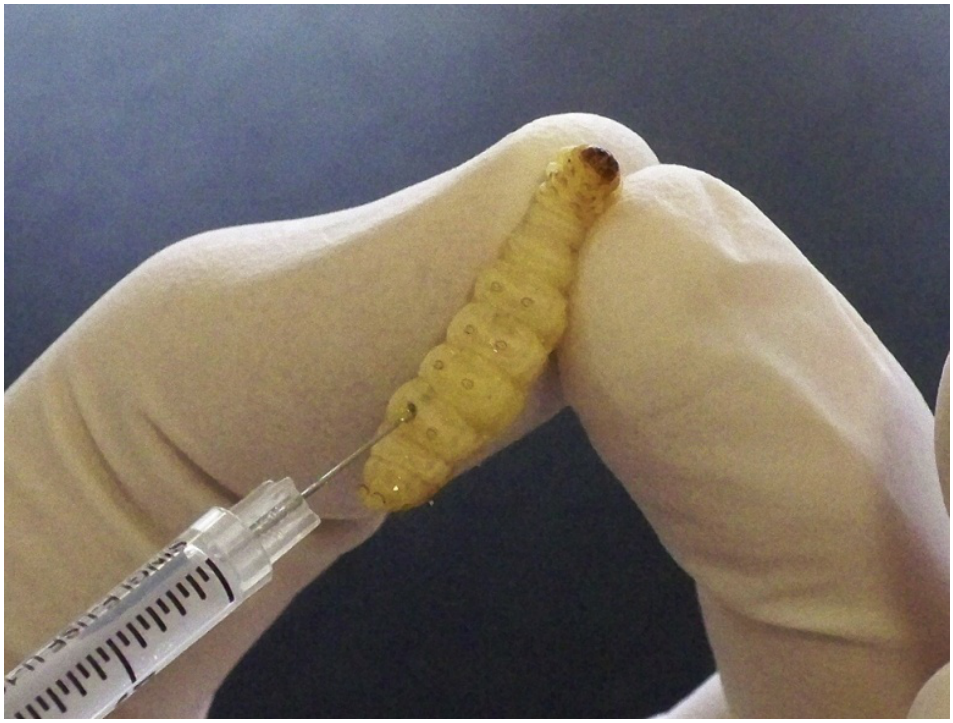
1. ST4 : CM39
2. ST5 : CM30, CT13
3. ST6 : CM34, R6
4. ST32 : CM35, CT37
5. ST1 : R16
6. ST23 : C54
7. ST53 : R27

Ten dilutions of 10^4 *Cng* cells were made using PBS (section 2.10), one per isolate. 10 μL of each inoculum was injected into a group of 10 larvae, alongside antibacterial ampicillin (Figure 9.1). Two control groups of 10 larvae each were also included, the first injected with PBS and the second with nothing. The number of dead worms was then recorded daily over a period of 21 days. The larvae were considered dead if they did not respond to stimuli or had turned black. Survival analyses were performed in R and the code used is detailed in the appendix (R code C.4).

9.4. Result

The virulence of genetically variable strains from two geographical locations was assessed through the survival analysis of the *Cng* - *G. mellonella* assay. Survival curves for the first 21 days post-infection are shown for all 10 STs

Figure 9.1.: The method by which *G. mellonella* larvae are inoculated with *Cng*. Intra-haemocoel infection is achieved by the insertion of a Hamilton syringe into the invertebrate's proleg. Taken from Kavanagh and Fallon, 2010.



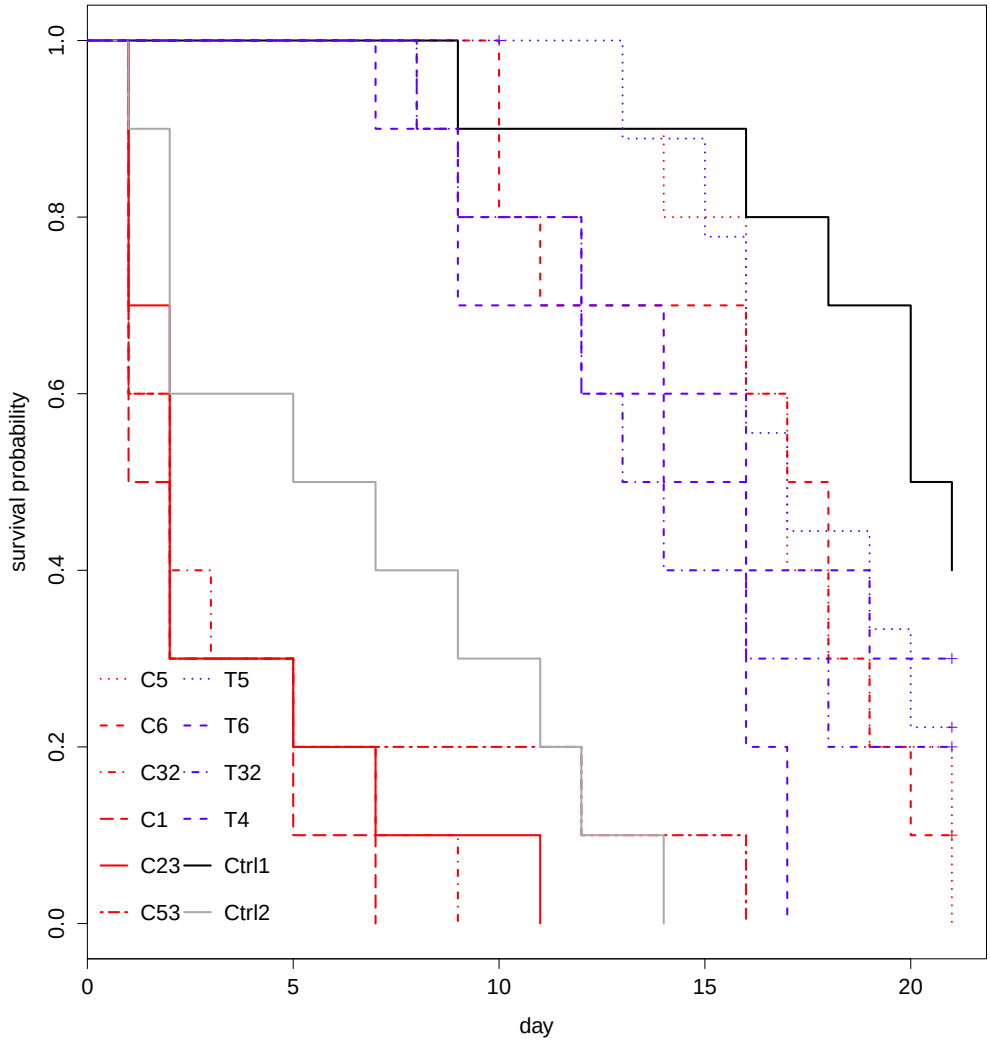
collectively, with the red lines representing Cape Town strains, and the blue lines Thai strains (Figure 9.2). The two control groups are in grey and black. In addition, strains of identical ST but different geographic origin are depicted with identical line types.

Isolates of STs unique to Cape Town (R16, C54 and R27 of STs 1, 23 and 53, respectively) were highly virulent for *G. mellonella*, all three resulting in 50% mortality by day 2. In contrast, single isolate CM39 of ST4 unique to Thailand was considerably less virulent, with the same 50% mortality rate in wax moth larvae taking 13 days to achieve (0% mortality probability by day 2). The three remaining Thai isolates of STs 5, 6 and 32 also showed attenuated virulence within *G. mellonella* (Figure 9.2). Survival probability for the STs unique to Cape Town (STs 1, 23 and 53) was estimated at $\leq 20\%$ by day 5 compared to 100% for all isolates of Thai origin. The difference between these two groups (pooled Cape Town STs 1, 2, 3 and 53 versus pooled Thai STs 4, 5, 6 and 32) was significant (logrank p-value < 0.001).

Having assessed overall virulence patterns according to country of origin, variations within the isolates of the same country were also investigated. The Thai strains were not significantly differentiated from each other in mean survival rate (logrank p-value $= 0.35$), that is to say that they all exhibited attenuated virulence in the wax moth larvae. The six Cape Town infecting strains, on the other hand, were significantly differentiated from one another ($p < 0.001$). This was to be expected as STs 5 and 6 were attenuated in virulence, while STs 1, 23 and 53 were hypervirulent (Figure 9.2).

Three STs used in this assay were shared between the Thai and Cape Town isolates: 5, 6 and 32. Survival estimates of *G. mellonella* infection were not significantly different between the Thai and Cape Town isolates of ST5 (logrank p-value $= 0.45$). This was not the case between CM34 and R6, both of ST6. Although similar in effect up until day 16, a sharp increase in death among the wax moth larvae infected by the Thai strain followed (logrank p-value $= 0.04$). ST32, recovered from Thai and Cape Town, showed the most marked and significant difference according to origin (logrank p-value < 0.001), with a 70% probability of death by day four among larvae infected by the Cape Town strain (isolate CT37), at which point no larvae infected by Thai strain CM35 of the same ST had died (0%

Figure 9.2.: Survival plot of *G. mellonella* larvae infected with *Cng* according to ST and geographic origin. Red lines indicate Cape Town origin, blue lines Thai origin and black lines the two control groups, PBS control to the left and the group to which nothing was done on the right. Line types shared between isolates of different geographic origin are identical. Ctrl1 = PBS Control, Ctrl2 = Nothing.



probability of death; Figure 9.3).

Figure 9.3.: *Galleria mellonella* - *Cng* virulence assay on day 8 of the experiment. Both plates contain wax moth larvae infected by *Cng* sequence type 32. Those in the left plate were infected with a Cape Town strain (CT37), while those on the right were inoculated with a Thai strain (CM35). Seven of the larvae in the first plate are dead, while none are in the second plate.

In order to integrate this virulence study with findings from the clinical analyses reported in Chapter 7, the allelic profiles of the STs used in this assay were reviewed (Table 9.1). The hypervirulent strains of STs 1, 23 and 53 were unique in AT at two loci: *CAP59* and *PLB1*. AT7 of *CAP59* and ATs 1 and 3 of *PLB1* may contribute to the increased virulence seen in the Cape Town-specific STs. AT1 (*CAP59*) and ATs 2 and 4 of *PLB1* are associated with low virulence.

Table 9.1.: Allelic profiles of the seven STs used in the *Cng* - *G. mellonella* virulence assay.

Origin	ST	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*
Thailand	4	1	1	1	3	2	1	5
Shared	5	1	3	1	5	2	1	1
Shared	6	1	1	1	4	2	1	5
Shared	32	1	1	10	3	4	1	1
Cape Town	1	7	1	1	1	1	1	1
Cape Town	23	7	1	1	2	1	1	2
Cape Town	53	7	5	1	3	3	1	2

9.5. Discussion

Several mechanisms by which *Cryptococcus* is able to cause infection remain homologous between insects and vertebrates. This is due to a high degree of similarities in innate immune response between the two, despite their having diverged an estimated 500 million years ago (Hoffman, 1995; Fallon and Sun, 2001; Boman and Hultmark, 1987). These similarities allow alternative host models to provide indications of vertebrate response to infection by microbial pathogens such as *Cn* (Kavanagh, 2004; Mylonakis, Calderwood and Ausubel 2006). Larvae of *Galleria mellonella*, the Greater wax moth, have successfully been used in investigations of virulence of *Pseudomonas aeruginosa* and *Candida albicans* mutants - the results of which were on par with those obtained from mice assays (Dunphy et al., 1986; Jander et al., 2000; Brennan et al., 2002). This invertebrate host's first line of defence against infection is its cuticle, similar to the mammalian skin. This physical barrier also contains anti-microbial agents (Brennan et al., 2006). The haemocel of the larvae (the body cavity) contains haemolymph. Likened to blood, this transports nutrients, waste and immune cells about the insect's body (Kavanagh and Fallon, 2010). Known as haemocytes, there are six described immune cells involved in the killing of microbes both by phagocytosis (plasmatocytes and granulocytes; Mylonakis et al., 2006), and by the secretion of antimicrobial peptides which act on the invading organism's cell wall (Ratcliffe, 1985). *G. mellonella*'s immune response also has a humoral component involved in melanisation, the clotting of the haemolymph and

healing of wounds (Kavanagh and Reeves, 2004). Inexpensive and easy to culture and inoculate, the Greater wax moth larva is ideal for the investigation of variations in virulence hinted at in Chapter 7. This *Galleria mellonella* - *Cn* model simplifies the complex host-pathogen interactions of the patient population and aims to shed light on associations between ST, geographic origin and *Cng* virulence.

This experiment demonstrates that the larvae of *G. mellonella* are susceptible to infection by *Cng* and that the strains used vary in virulence. The results of this assay clearly show that the cryptococcal isolates are widely divergent in their pathogenic potential according to ST and geographic origin. STs 1, 23 and 53 of African origin were hypervirulent while STs 4, 5 and 6 of Asian origin were attenuated in the *G. mellonella*. The difference in mean survival rates between the two were significant (logrank p-value $<$ 0.001).

Examining the variation in survival according to allelic profiles, additional parallels were found between the wax moth larvae assay and the clinical findings. Grouped according to virulence, the attenuated isolates of Thai and shared origin (4, 5 and 6) form a group, while the hypervirulent Cape Town isolates represented by STs 1, 23 and 53 form another. ST32, represented by Thai isolate CM35 and Cape Town isolate CT37, was not included in either group as there was no correlation in virulence between the two origins (logrank p-value < 0.001), further addressed below. Loci *CAP59* and *PLB1* may be markers of a virulent phenotype in the *G. mellonella* assay as they contain ATs which are not shared between the two pathogenically different groups. *CAP59* is involved in the production, maintenance and shedding of the polysaccharide capsule, while *PLB1* is involved in the intracellular growth of *Cn*, particularly during mobilisation from the lungs into the blood (Santangelo et al, 2004). Both phenotypes are key elements of *Cng*'s pathogenesis and have been demonstrated to be important in the survival of *Cn* in amoebae. *A. castellanii* kills non-encapsulated *Cn* and, similarly, phospholipase-deficient *Cn* mutants are also susceptible to the amoebae (Steenbergen, 2001; Casadevall, 2006). As previously discussed, virulence mechanisms of *Cn* are believed to have evolved in response to predation by amoebae, supported by the homology between the insect and mammalian interactions with the pathogen (Levitz 2001, Steenbergen and Casadevall

2003). *CAP59* was found to be statistically associated with cryptococcal antigen titre (CRAG; Chapter 7), with AT1 being associated with lower levels of CRAG, and STs 2 and 7 with higher levels. In this experiment, *CAP59* AT1 was present in genotypes unique to Cape Town, while AT7 was found in the attenuated STs 4, 5 and 6. It is important to note that these results are not evidence of the loci causing the phenotype, but rather of their link to the phenotype, possibly via linkage disequilibrium. *PLB1* was also associated with cryptococcal burden, but of greater relevance was its association with survival among the HIV+ patients. Among these patients, infection with *PLB1* AT2 resulted in 80% survival, while infection with AT3 resulted in $\leq$60% survival probability. In the *G.mellonella* assay, *PLB1* AT3 of the Cape Town genotypes was also associated with lower survival rates in comparison to those of attenuated Thai isolates containing AT2 (Figure 9.2). These findings further support the homology in the interaction of *Cng* with both hosts (humans and *G. mellonella*) and validates the use of this simpler model system in the investigation of its pathogenesis.

As mentioned above, striking variations in virulence was observed between isolates of shared ST32 (Figure 9.2). Within two days of the commencement of the experiment, the Cape Town isolate CT37 had attained 50% mortality, while that of Thai origin (CM35) only achieved this at day 13 (Figure 9.2). Although it appears strange, this result is identical to that seen in the survival analysis of the CM patients analysed in Chapter 7. Within this clinical population, ST32 of Asian origin was associated with 100% survival at 150 days of follow up. Survival probability among patients infected with ST32 of African origin, on the other hand, was below 20% by day 150. The fact that this pattern is repeated in the invertebrate animal model rules out the possibility of this variation being attributable to host differences. This geographically defined difference in virulence between genetically identical isolates could be due to the fact that *Cng*'s status is ever-changing. An example of this is evidenced by serotype switching of *Cn* within the human host, similar to that seen in *C. albicans* (Cherniak et al., 1995; Garcia-Hermoso et al., 2004; Soll, 1989). Fries et al. (2001) linked this switching to changes in the polysaccharide capsule, which in turn inhibited phagocytosis by alveolar macrophages in a murine model and affected the outcome of infection. The *CAP59* gene involved in the maintenance of *Cn*'s

polysaccharide capsule structure was identified as one of two genes potentially involved in determining the virulence of the isolates used in this assay (Table 9.1). The loss or acquisition of *Cn*'s virulence phenotype during storage, transportation and use has been shown to be dynamic but is, as yet, not fully understood (Perfect, 2006). CT37 mirrored the pathogenesis of the remaining Cape Town isolates in this assay — as well as that seen among clinical patients infected by ST32 who had the lowest probability of survival (Figure 7.3) — while CM35 mirrored that of the attenuated Thai ones; this fact hints at external environmental factors resulting in microevolution and influencing virulence, despite the two isolates being genetically identical according to MLST. This demonstrates the importance of controlling practical aspects and careful use of standard storage methods in order to minimise further effects on the pathogen's already complex virulence composite.

There were several limitations to this study. Deaths of larvae were recorded for two control groups. The first contained ten larvae to which nothing was done and mortality was low, as expected. The 50% mortality rate by day 20 is an indicator of the natural course of the larvae's life cycle, meaning results past day 15 may not be solely attributable to the infectious agent. This leads us to reconsider the apparent difference in survival probability seen between isolates of shared ST6 which were identical up until day 16 (Thai isolate CM34 and Capetonian isolate R6; logrank p-value = 0.04). The second group aimed to control for trauma using mock inoculations. In this control group, 50% of the larvae were dead within five days. The replication of the experiment with greater numbers of larvae per group would likely improve the confidence in the results. There are several additional changes to the study which would decrease uncertainty. For example, instead of two separate injections — one of inocula and the second of ampicillin — the two can be combined into a single 10 μL injection, reducing handling and physical trauma. Mowlds et al. (2008) showed that mild shaking of *G. mellonella* prior to inoculation resulted in decreased susceptibility to infection with *C. albicans*. Inoculated 24 hours after shaking, an increased number of circulating haemocytes was reported. The current experiment was not blinded and, as this was a novel technique, the inexperience of handling the worms and successfully injecting them may have translated in greater physical stress to the first group of wax worm larvae. Injected with Thai

isolates, these larvae's probability of survival was greater than the second group of larvae injected with Cape Town isolates who, due to practice, may have been less physically traumatized. Blinding of the trial may dispel such bias. There was also an element of fatigue to consider as all inoculations were carried out on the same day. All these may have contributed to the variation seen in virulence, and, again, repetition of the assay and the use of greater numbers of larvae per group would increase the confidence in the results.

Further work which could be carried out using this model includes the quantification of fungal burden in the *G. mellonella* larvae over time. This would allow for the monitoring of phagocytosis and clearance, and give us further insight to the interaction between the host and the pathogen. As not all larvae deaths are due to cryptococcal burden, this increases confidence in death attributed to the proliferation of the fungus. The quantified rates of fungal growth will enable comparative analyses of the virulence impact of the STs and identified target genes.

The use of *G. mellonella* in the study of fungal pathogenesis of *Cng* is appropriate, reducing time and cost, and overcoming legal and ethical barriers. These wax moth larvae are readily available and can produce results within 48 hours (Cotter et al., 2000). Benefits in comparison to other invertebrate hosts include the fact that space is not of concern, unlike with *D. melanogaster* which need to fly, and no additional equipment is needed, such as incubators and microscopes required for the study of *C. elegans*. With a host response parallel to that of mammalians, assays can be performed at high temperatures, and significant correlations have been shown with clinical trials of antifungal treatments within humans (Mylonakis et al., 2005). This model will be key in bringing new insights to the study of *Cn*'s pathogenesis. The only drawback is the lack of a sequenced genome for *G. mellonella*, a resource which would improve our scope for rapid comparative analyses of this host's responses (Mylonakis, 2007).

10. Final Discussion

The work presented in this PhD thesis demonstrates the value of global typing in elucidating the evolution and pathogenicity of the human fungal pathogen *Cryptococcus neoformans* variety *grubii* (*Cng*). These findings contribute to the current understanding of the pathogen's evolutionary epidemiology, genetic diversity and likely patterns of global spread. These molecular data also provided genetic markers that may be associated with variation in the virulence of *Cng* which, in turn, may have implications for controlling clinical disease burden and mortality, particularly in developing countries.

At the root of this study is the collaborative work of the International Society of Human and Animal Mycoses working group, whose overall aim was to resolve anomalies in *Crytpococcus*' taxonomy and integrate the rapidly increasing genotypic data. As described in our communication, Meyer et al. (2009), we (i) proposed a standardised typing method for the *Cryptococcus* species complex: multilocus sequence typing (MLST), (ii) proposed a consensus genotype nomenclature for both *Cryptococcus neoformans* (*Cn*) and *Cryptococcus gattii* (*C. gattii*), and (iii) designated a set of reference strains for each molecular group. Most important was the creation of a platform from which the data pertaining to this pathogen could be stored, universally accessed and updated, as the possibility of easy and accurate cross-referencing is crucial for the advancement of global typing. *Cng* now joins other active fungal MLST schemes available at `http://www.mlst.net/`, including the human pathogenic fungi *Candida albicans*, *Candida glabrata*, *Candida tropicalis*, and *Candida krusei* (Bain et al., 2007).

I successfully used MLST to characterise *Cng* isolates from two populations, Thailand ($n = 183$) and Cape Town ($n = 107$). This genotyping uncovered genetic diversity that was used to infer their phylogenetic and population genetic structure. The predominantly clinical cohort of the

Thai population was found to be highly homogenous, consisting of only 10 sequence types (STs). Of these, all but one isolate was of molecular group VNI. This domination of by VNI echoes that reported in other Southeast Asian countries including China (Chen et al., 2008), Malaysia (Tay et al., 2006), and another Thai study (Sukroongreung et al., 1996). I showed using several population genetic analyses that this was a highly clonal population. I also demonstrated a lack of geographically defined structure consistent with non-neutral evolution, and rapid population expansion from a limited set of ancestors. My evidence of recent dispersal to the Asian region was also observed in the neighbouring country of India, where genetic variation of environmental *Cng* from five regions was also shown to be limited (Hiremath et al., 2008). The South African population, on the other hand, showed much higher levels of genetic diversity. Similarly high genetic diversity had previously been reported in the African population of Botswana, described by Litvintseva et al. (2006), which was dominated by an unique African population of VNB isolates. In this study, the typed Cape Town population also contained VNB isolates, evidence of wider dispersal of VNB across southern Africa than was previously believed.

The value of cross-referencing between studies was then shown by the integration of these data on the Thai and Cape Town *Cng* populations into a global dataset of isolates compiled from several other studies (Litvintseva et al., 2006). A complete dataset of 370 isolates from 14 countries across five continents yielded 76 STs. The previously described structures of Thailand and Cape Town remained unchanged following their integration with isolates from wider Asian and African subpopulations, evidence of their being representative of general patterns across these regions. Interestingly, analyses of genetic divergence and differentiation between the geographically defined subpopulations of *Cng* revealed a recombining global population within which there exists clonal propagation, as evidenced by the widespread over-representation of certain genotypes. Many microbial eukaryote populations are characterized by infrequent recombination alongside clonal dispersal and expansion of a few genotypes including human pathogenic fungi *Coccidioides immitis* (Burt et al., 1996; Koufopanou et al., 1997), *Candida albicans* (Pujol et al., 1993; Graser et al., 1996), and *Aspergillus nidulans* (Geiser et al., 1994).

That recombination appears to be more frequent within the African subpopulation, may be the due to a higher frequency of the rare but sexually viable *Cng* **a** mating type (Litvintseva et al., 2003; Litvintseva et al., 2011). When combined with consistently high genetic diversity and the presence of ancestral haplotypes, this is compelling evidence of African ancestry for *Cng* — especially in comparison to the remaining global subpopulations which show more biased distributions of mating types and rarer recombination. These findings led to the investigation of the potential origin of *Cng* which, according to Litvintseva et al. (2011)'s "Out of Africa" hypothesis, is southern Africa. Reporting two ecological niches between which genetic exchange was limited, Litvintseva et al (2011)'s study described an association between (i) global *Cng* strains and of rock dove faeces, and (ii) VNB genotypes and endemic African strains. Coalescent analysis of the isolates included in my study revealed that the African VNB population was the eldest (time to most recent common ancestor $\approx$ 4.5 million years ago) and tests of phylogeny showed evidence of a common history between molecular groups VNI and VNB. My findings demonstrate strong evidence of the VNI lineage having arisen from a basal African VNB population. VNI then became globally distributed between 1 600 and 70 500 years ago. The guano of the common pigeon (*Columba livia*) is the predominant ecological niche of *Cng* VNI, and this time frame encompasses that of this bird's domestication in Africa and its global spread through human trade (Mooney, 2000; Grzimek, 2004). In addition, it appears that there may have been several expansion events out of Africa and into Asia owing to the identification of two founding VNI genotypes within the Asian population.

The emergence of microbial pathogens in novel geographical locales is common, and appears to be widely vectored by humans (Daszak et al., 2000). The proposed model for the spread of *Cng* is similar to that of the etiological agent of potato blight, *Phytophthora infestans*; the cause of several pandemics, including the Irish Potato Famine in 1845. DNA fingerprinting described the dominance of a single clonal lineage of the oomycete in a collection of 200 isolates (Goodwin et al. 1994a). This pan-global genotype was believed to have expanded out of a geographically restricted and recombining population in Mexico, the center of diversity for the blight pathogen (Goodwin et al., 1994b; Fry and Goodwin, 1997). Another exam-

ple of disease epidemics resulting from human trade is that of the fungal pathogen *Ustilago scitaminea* which causes yield loss in sugarcane through smut disease. Raboin et al. (2007) used microsatellite typing to reveal high genetic diversity of *U. scitaminea* in the Philippines, which contrasted with extremely low diversity elsewhere. Identifying a single worldwide lineage which shared allelic combinations with Asian strains, they determined an Asian origin of *U. scitaminea*. The global spread of the pathogen is believed to have been by plant associated migration of infected sugarcane material (Rabaoin et al., 2007). One such proposed event was the production of the first sugarcane hybrids in Indonesia and India, which were distributed worldwide in the 1920s and '30s. Additional evidence of human behavior broadly affecting microbial evolution is zoonotic parasite *Toxoplasma gondii*, a fatal opportunistic pathogen in AIDS patients (Sibley and Boothroyd, 1992). Its unusual clonal population structure consists of only three genotypes which are found worldwide (Dard et al., 1992; Howe and Sibley, 1995). Su et al. (2003) attribute the successful clonal expansion of these three lineages to the parasite's direct oral infectivity. This virulence trait and subsequent spread of *Toxoplasma gondii* is believed to have evolved at the time of human agricultural expansion and the domestication of the cat, a host species in which the fungus is able to circumvent sexual reproduction (Hole, 1994; Serpell, 2000; Su et al., 2003). *Cryptococcus gattii*, a sister species of *Cng*, was previously believed to be geographically restricted to tropical and subtropical regions (Australia, Brazil, Kenya, Zaire, Southeastern Asia, and Southern California; Ellis, 1987). Recently, however, a hypervirulent *C. gattii* strain believed to be the progeny of asexual mating was implicated in the Vancouver Island outbreak, which began in 1999 (Fraser et al., 2005; Kidd et al., 2004; Byrnes et al., 2009). Not only is this causing substantial ongoing mortality among humans and animals with no immunological impairment (CDC, 2008), but it also appears to be spreading with cases being reported on British Colombia (mainland), the San Juan Islands and in Oregon and Washington states (CDC, 2008; Kidd et al., 2004; MacDougall et al., 2007; Kidd et al., 2007; Upton et al., 2007; Bartlett et al., 2008). *C. gattii*'s establishment in a novel environmental niche is associated with *E. camaldulensis*, a species with which molecular type VGI has a strong environmental association (Fraser et la., 2005; Pfeiffer and Ellis, 1991). Saul et al. (2008) showed that both same-sex and opposite-sex mating is likely

occurring in the hollows of this tree. This eucalyptus species is native to Australia and has been exported and planted in areas of California, Mexico, and other parts of the world (Pfeiffer and Ellis, 1991), and may have originally vectored *C. gattii* into this region from its native heartland in Australia.

Clonality is expected in most microbial species as they all undergo mitotic reproduction to some degree, and therefore all have the latent, or realised, capacity to reproduce asexually (Xu, 2004; Brut et al., 1996). Asexual reproduction confers a two-fold advantage over sexual reproduction and eliminates the need to find a mate (Maynard-Smith, 1978). Even so, many species elect to reproduce sexually when they could do so asexually instead (Williams, 1975). *Cn* also retains the ability to reproduce sexually, confirmed by laboratory crosses (Kwon-Chung, 1975). This persistence of sexual reproduction therefore suggests wider evolutionary benefits. Although not yet fully understood, the theories on the advantages of sex have been classified as mutational or ecological (Kondrashov, 1993). Largely asexually reproducing lineages run the risk of becoming extinct due to mutation accumulation. Although individual mutations are not detrimental on their own, additional mutations increasingly affect the fitness of the organism (synergistic epistasis; Muller 1964; Lynch and Gabriel 1990; Charlesworth et al. 1993). The deterministic mutation model proposed by Kondrashov (1988) argues that sexual reproduction recombines genotypes and groups deleterious genes in individuals which are cleared by selection, thus purging mutations (Muller, 1964). An ecological advantage of reproduction is the creation of novel variants which are favored by selection in rapidly changing environments (Jaenike 1978; Hamilton 1980). Genomic rearrangement resulting in novel virulence profiles in fungal pathogens is witnessed in the emergence of several plant and animal fungal diseases including *T. gondii* and *Batrachochytrium dendrobatidis*, as previously described (Fisher et al., 2009; Su et al. 2003). Griggs et al. (2001) showed that random genetic reassortment of two avirulent strains of *T. gondii* can give rise to highly virulent progeny, underlining the ability of recombination to potentiate virulence in otherwise attenuated lineages. Therefore, while we do not know exactly how or why *Cng* is recombining, we can assert that the manifest benefits of sex are preserved to ensure the evolutionary long-term 'health'

of individuals of the species.

Of greater relevance is the anthropogenically mediated inter-lineage mixing of *C. gattii* strains at the root of the Vancouver Island epidemic. Here, same-sex mating was responsible for the hypervirulent progeny (Fraser et al., 2005). This closely related fungal population shows how sexual recombination can impact the current global distribution of *Cng* (Bovers, 2008; Xu et al., 2000). In my study, *Cng* was clonal but there was evidence of recombination, especially within the African isolates. In this African population, the presence of MAT**a** may be upregulating haploid fruiting of α cells. This same-sex mating mechanism is advantageous as there is a lack of opposite mating types. This process also generates robust MATα spores, the infectious propagules of *Cng* (Velagapudi et al., 2009; Botts and Hull, 2010). Infectious in murine and invertebrate models, these spores are more resistant to stress than cryptococcal yeasts (oxidative, heat and chemical insult; Botts et al, 2009). In addition, Botts and Hull (2010) demonstrated that spores resulted in higher disease burden owing to differences in host-pathogen interactions: spores were readily phagocytosed by macrophages in a murine model, unlike yeasts, and were therefore able to replicate and disseminate to other organs and organ systems, resulting in greater fungal burdens (Feldmesser et al., 2000). Same-sex mating can, therefore, be producing novel, highly fit basidiospores reflected in the dominant MATα cells observed in the typed African population. This potential for the evolution of new virulent strains through same-sex recombination may well be the cause of the high mortality rates reported among AIDS-associated CM patients in Africa. This same process could well be the driving force underpinning the natural evolution of virulence in other successful pathogens such as the predominantly clonal *Pneumocystis jirovecii*, for which only one mating type has been identified so far (Thomas and Limper, 2004; Fraser et al., 2005).

Having established the value of global typing in the epidemiological context, these molecular data were applied to the clinical context of HIV/AIDS disease in which *Cng* is responsible for 98% of clinical infections worldwide (Day, 2004). Recent estimates of the global burden of *Cng* by Park et al. (2009) reported 957 000 cases of AIDS-associated cryptococcal meningitis (CM) in 2006, 600 000 of which were fatal (Park et al., 2009). This is of

particular concern in developing countries of Sub-Saharan Africa and South and Southeast Asia where the HIV/AIDS epidemic is thriving. Case fatality rates are highest in these regions, accounting for 92% of the total global deaths attributable to AIDS-associated CM (81% and 11%, respectively; Park et al. 2009). This is the first study which combines molecular typing data and detailed clinical indicators of disease progress within human cryptococcal patients. These clinical data were available for both Thai and Capetonian patients ($n = 119$) infected by cryptococcal isolates of six STs. Cerebral dysfunction at baseline was found to be a strong prognostic factor for death both by two and ten weeks. This was expected as the invasion of the central nervous system is indicative of successful proliferation and dissemination of the pathogen within the host, as reported in previous studies (Brouwer et al., 2004; Saag et al., 1992; Heyderman et al., 1998). Using early fungicidal activity as an outcome variable, regression analysis revealed independently significant associations between loci *IGS1* and *PLB1*, the former being a "housekeeping" gene while the latter is directly involved in *Cng* pathogenesis. Both these loci remained significant in determination of survival probabilities of the HIV+ CM patients and may be genetic markers of *Cng* pathogenesis.

Survival analyses hinted at a correlation between the geographically defined genotypes and survival outcome of disease: isolates of African origin resulted in poorer long-term survival probabilities than those of Thai origin. In order to assess this geographically determined variation in virulence, a simple animal host model was used. *Galleria mellonella* was the choice invertebrate host as phagocytosis by its hemocytes mimics microbial killing by humans immune cells (Brennan et al., 2002; Bergin et al., 2005; Tojo et al., 2000), and assays can be carried out at elevated temperatures that are, to some extent, representative of the human host environment (Casadevall, 2005). Although this invertebrate host assay had its limitations, the results of the experiment support the clinical findings. Hypervirulence was uniform among genotypes specific to Cape Town, while those found in both Cape Town and Thailand were attenuated in virulence. STs 5 and 6 predominate in the global VNI population, and are also present in the African cohort. All isolates of these STs were found to be uniformly attenuated in virulence, irrespective of origin. This genetic and pathogenic correlation confirms recent

gene flow between the two geographically separated populations, especially as ST5 was determined to be a founding genotype introduced into Thailand from Africa during a recent expansion event. Importantly, the African population also contained unique and novel variants which were shown to be hypervirulent. These are, potentially, progeny of same-sex mating as suggested above. As 60% of the global CM-attributable deaths per year are from HIV/AIDS patients in Africa, we turn back to the global data as we consider whether this high mortality is owing to the exposure to more virulent African strains. These data reveal that two of the three hypervirulent African STs used in the assay are not restricted to Cape Town (Table B.1), but have been reported elsewhere in low frequencies. Isolate A3-38-20 from pigeon guano is of North American origin and of ST1. Similarly, isolates A3-1-1, also isolated from pigeon excreta in North America, and it754 of Italian origin are of ST23. These regions (North America and Western Europe) have comparatively low mortality estimates for AIDS-related CM (700 and 45 deaths per year, respectively; Park et al., 2009). The North American isolates are environmental isolates from pigeon excreta (the potential vector of *Cng*) and so we are unable to determine pathogenicity in humans. This reinforces the multi-faceted applications of global typing data, but highlights the need for continued assimilation of data. It is likely that a combination of the existence of hypervirulent strains (created by same-sex recombination), and their presence in high numbers (due to selection favouring these genotypes of increased fitness) is resulting in the high disease burden witnessed in Sub-Saharan Africa.

A second finding from the clinical analyses mirrored in the *G. mellonella* assay was that of variations in virulence between isolates of identical genotype according to geographical origin: African ST32 was hypervirulent while Asian ST32 was attenuated in virulence. This suggests two possibilities, either MLST was not robust enough to discriminate between two very close genotypes, or that there are geographically determined environmental factors affecting the virulence of the isolates. Litvintseva and Mitchell (2009) also reported differences in virulence of genetically identical according to source, with environmental strains being less pathogenic than clincial ones. The characterisation of the adaptations leading to variations in virulence which appear to transcend genotype is a potential avenue for further work.

This study was the first successful implementation of the new MSLT scheme in the description of *Cng* (Simwami et al., 2011), and has shown the gains that are associated with the collection of global MLST datasets, setting the stage for integrating future MLST datasets, as well as utilising new deep-sequencing approaches to genotype whole *Cng* genomes in parallel. Also novel was analysis of the integration of molecular and clinical data in the analysis of prognostic factors of outcome of cryptococcal disease, with implications to public health and antifungal drug therapy. Importantly, my work highlights the lack of sufficient data on *Cng* strains in regions such as Europe and South America. Continued global typing, especially from these regions, is the key to elucidating the population structure of *Cng* in order to understand the contribution of the pathogen's genotype to the epidemiology of this infection. Further sampling may reveal previously unknown natural ecological niches of the organism, shedding more light on potential modes of expansion of *Cn*'s molecular types. Additional work on the role of recombination in the adaptation of this sexually competent fungus's pathogenicity will lend to more accurate quantification of its potential for disease in different regions and among different populations. This can, in turn, inform public health resources and antifungal treatment and vaccine development in our fight against CM. This is especially exciting as we are in the era of full genome sequencing, which will undoubtedly result in unprecedented progress in combating this life-threatening pathogen.

A. Published work

Published in final edited form as:
Med Mycol. 2009 ; 47(6): 561–570. doi:10.1080/13693780902953886.

Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*

Wieland Meyer[*], **David M. Aanensen**[†], **Teun Boekhout**[‡], **Massimo Cogliati**[§], **Mara R. Diaz**[+], **Maria Carmela Esposto**[§], **Matthew Fisher**[†], **Felix Gilgado**[*], **Ferry Hagen**[‡], **Sirada Kaocharoen**[*,#], **Anastasia P. Litvintseva**[$], **Thomas G. Mitchell**[$], **Sitali P. Simwami**[†], **Luciana Trilles**[*,^], **Maria Anna Viviani**[§], and **June Kwon-Chung**[~]

[*]Molecular Mycology Research Laboratory, Centre for Infectious Diseases and Microbiology, Westmead Millennium Institute, University of Sydney Western Clinical School at Westmead Hospital, Westmead, Australia [†]Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College London, London, UK [‡]CBS Fungal Biodiversity Centre, Utrecht, The Netherlands [§]Laboratory of Medical Mycology, Dip. Sanitá Pubblica-Microbiologia-Virologia, Università degli Studi di Milano, Milano, Italy [+]Rosentiel School of Marine and Atmospheric Science, University of Miami, Miami, Fl, USA [#]Inter-Department of Medical Microbiology, Graduate School, Chulalongkorn University, Bangkok, Thailand [$]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham NC, USA [^]Laboratório de Micologia, Instituto de Pesquisa Clínica Evandro Chargas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil [~]Molecular Microbiology Section, Laboratory of Clinical Infectious Diseases, NIAID, NIH, Bethesda, MD, USA

Abstract

This communication describes the consensus multi-locus typing scheme established by the Cryptococcal Working Group I (Genotyping of *Cryptococcus neoformans* and *C. gattii*) of the International Society for Human and Animal Mycology (ISHAM) using seven unlinked genetic loci for global strain genotyping. These genetic loci include the housekeeping genes *CAP59, GPD1, LAC1, PLB1, SOD1, URA5* and the IGS1 region. Allele and sequence type information are accessible at http://www.mlst.net/.

Keywords

Cryptococcus neoformans; *Cryptococcus gattii*; Genotyping; Multi Locus Sequence Typing

Introduction

Cryptococcus neoformans, the agent of cryptococcosis, had been considered a homogeneous species until 1949 when the existence of four serotypes was revealed based on the antigenic properties of its polysaccharide capsule [1]. Such heterogeneity of the species, however, remained obscure until the two morphologically distinct teleomorphs of *C. neoformans* were discovered during the mid 1970s [2,3]. The teleomorph *Filobasidiella neoformans* was found

Correspondence: J. Kwon-Chung, Molecular Microbiology Section, Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases/NIH, Bethesda, MD 20892, USA. Tel: +1 301 496 1602; fax: +1 301 480 3240; june_kwonchung@nih.go.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

to be produced by strains of serotype A and D [2] while *F. bacillispora* was found to be produced by strains of serotype B and C [3]. Ensuing studies revealed numerous differences between the anamorphs of the two *Filobasidiella* species with regards to their ecology, epidemiology, pathobiology, biochemistry and genetics.

Presently, the etiologic agent of cryptococcosis is classified into two species [4], *C. neoformans*, with two varieties: *C. neoformans* var. *grubii* (serotype A) [5] and *C. neoformans* var. *neoformans* (serotype D) [6], as well as an AD hybrid, and *C. gattii* (serotypes B and C) [7]. Intra-species genetic diversity has also been revealed as more genotyping methods have been applied for each serotype. In addition inter-species hybrid strains of AB and BD serotypes have been described [8,9]. As a result, the number of scientifically valid species within *C. neoformans* has become a controversial issue because of the differing opinions among taxonomists as to the appropriate definition of a species. There are several research groups focusing on the molecular determination of the number of genetically diverse sub-groups within each serotype. The molecular methods employed by each group to define these sub-groups vary from DNA fingerprinting [10,11] and PCR fingerprinting based on microsatellite-(M13) or minisatellite-specific primers (e.g., $(GACA)_4$ or $(GTG)_5$) [12–16], over random amplification of polymorphic DNA (RAPD) analysis [17–20], amplified fragment length polymorphism (AFLP) analysis [21–23], restriction fragment length polymorphism (RFLP) analysis of the *URA5* [16,24] and *PLB1* genes [25], the use of IGS sequences [26], multigene sequence analysis [27, Meyer *et al.* unpublished data], to multi-locus sequence typing (MLST) [23,28] and multi-locus microsatellite typing (MLMT) [29,30]. This research has revealed associations between geographic origin and particular genotypes, implying an epidemiologic significance of certain genotypes. Different methods have resulted in various numbers of sub-groups or different nomenclature of those sub-groups. However, due to the lack of a cross-reference consensus between the results obtained by different genotyping method, there is currently no concordance on a universally acceptable genotyping method for this important human pathogen.

Recognizing the urgent need for a standardized globally acceptable typing method, a Cryptococcus working group I, 'Genotyping of *Cryptococcus neoformans* and *C. gattii*', was formed under the umbrella of the International Society of Human and Animal Mycoses (ISHAM) in the beginning of 2007 which united all the major research groups that were involved in molecular strain typing of *C. neoformans* complex. The members of this ISHAM working group met at the 3rd Trends in Medical Mycology (TIMM3) Meeting in Torino, Italy in October 2007, and reviewed all the typing techniques in use. The group selected multi-locus sequence typing (MLST) as the method of choice for future strain typing in light of its high discriminatory power as well as reproducibility between different laboratories. The working group also chose standard reference strains representing the eight known major molecular types of the agent of cryptococcosis as well as the nomenclature of each genotype.

Consensus genotype nomenclature

As a result of the Torino meeting, the working group recognized that the different genotyping methods used by the different research groups lead to corresponding major genotypes for the agents of cryptococcosis (Table 1). Principally, the two main typing systems being used are: PCR fingerprinting using primers specific for microsatellite (M13) [14,16] or minisatellite $(GACA)_4$ DNA [13,15] and AFLP analysis [21]. In both typing schemes, over 2000 isolates were grouped into eight major molecular types. With some exceptions [26,31], the molecular types of *C. neoformans* are correlated with the serotypes: *C. neoformans* var. *grubii*, serotype A, consists of molecular types VNI=AFLP1 and VNII=AFLP1A; the hybrid serotype AD comprises VNIII=AFLP3; and *C. neoformans* var. *neoformans*, serotype D, corresponds to VNIV=AFLP2. *C. gattii* consists of VGI=AFLP4, VGII = AFLP6, VGIII=AFLP5, and

VGIV=AFLP7, which all correspond to both serotypes B or C [16,21, unpublished data]. Based on these findings, it was agreed by all cryptococcal working group members present in Torino to use the VNI–VNIV and VGI–VGIV nomenclature [16] since it correlated with the current concept of two species and represents the global population structure based on more than 2000 *C. neoformans* and *C. gattii* isolates among which *C. neoformans* var. *grubii* (serotype A=VNI) being the most prevalent molecular type world-wide.

Consensus standard strains

To enable global standardization, the working group also agreed to use a set of standard strains representing each of the eight major molecular types. This included the molecular type strains used in PCR fingerprinting or *URA5*-RFLP analysis [16] plus additional strains representing type cultures or strains, which are used in major cryptococcal genome projects (Table 2). All standard strains are publicly available from the CBS-Fungal Biodiversity Centre (CBS) (http://www.cbs.knaw.nl), the American Type Culture Collection (ATCC) (http://www.atcc.org) or the Fungal Genetic Stock Center (FGS) (http://www.fgsc.net). The corresponding collection numbers are listed in Table 2.

Consensus multi-locus sequence typing loci

To overcome problems arising from inter-laboratory reproducibility associated with the two commonly used typing techniques, such as PCR fingerprinting or AFLP analysis, the working group decided to use multi-locus sequence typing (MLST) as the method of choice for future cryptococcal strain typing. MLST has become the number one typing approach for epidemiological investigations of microorganisms [32]. MLST, originally developed for bacteria [32], indexes the sequence variation in approximately 400–500 bp of five to ten genes composed primarily of housekeeping genes. This technique has proven to be highly discriminatory for a number of human pathogenic fungi: *C. albicans* [33], *C. glabrata* [34], *C. tropicalis* [35], *Coccidioides* spp. [36] and *Histoplasma capsulatum* [37]. Most of the published MLST schemes are developed as tools for the wider scientific community, by being made publicly available as online databases at http://www.mlst.net/ and http://pubmlst.org/. In the case of the *Cryptococcus* species complex, two different MLST typing schemes have been introduced to type isolates of *C. neoformans* [23], and *C. gattii* [28], using twelve and eight unlinked loci respectively.

In the first study, 12 unlinked polymorphic loci: *MPD1, TOP1, MP88, CAP59, URE1, PLB1, CAP10, GPD1, TEF1, SOD1, LAC1* and the IGS1 ribosomal RNA intergenic spacer region, which are dispersed on nine different chromosomes, were used to type 102 globally obtained serotype A strains [23]. MLST differentiated three major groups among the studied isolates, corresponding to VNI, VNII and VNB, a Botswana specific genotype closely related to VNI. In connection with this study a central web based database was created at www.mlst.net (http://cneoformans.mlst.net/) allowing for an online determination of the alleles and sequence types of *C. neoformans* serotype A strains.

The second study used eight unlinked polymorphic loci: *SXIa* or *SXIα*, IGS1, *TEF1, GPD1, LAC1, CAP10, PLB1*, and *MPD1*, of which two are mating type locus specific and can not be amplified for all strains, to type 202 *C. gattii* strains. These loci were supplemented for a more detailed analysis of 9 closely related strains by 22 additional gene loci: *HOG1, BWC1, CNB1, TOR1, CAC1, CRG1, URE1, FHB1, BWC2, CNA1, CBP1, TSA1, STE7, FTR1, PAK1, CAP59, ICL1, GPA1, GPB1, RAS1, CCP1*, and *TRR1* to investigate the origin of the Vancouver Island outbreak isolates [28]. MLST differentiated all four major molecular types of *C. gattii* (VGI, VGII, VGIII and VGIV) and highlighted two possible origins (Australia or South America) for the outbreak strains.

Statistical analysis using the Simpsons's index of diversity [38] revealed that for both previously studied MLST data sets, a minimum of seven loci are required to differentiate between the sequence types of all strains (Fig. 1). For the Litvintseva *et al.* [23] MLST data set, the following loci resulted in the highest discrimination of the investigated strains: *CAP59*, IGS1, *GPD1, LAC1, PLB1, MP88* and *SOD1*, with a Simpson's index of diversity of 0.9632. For the Fraser *et al.* [28] MLST data set, the most discriminatory loci were: *GPD1*, IGS1, *TEF1, LAC1, MPD1, CAP10* and *PLB1*, which resulted in a Simpson's index of diversity of 0.9319.

Both MLST schemes utilized highly polymorphic loci, which resulted in stable and reproducible typing systems that were able to distinguish between closely related strains. While using as many genetic loci as possible would enhance the discriminatory power of the MLST scheme, it would be pragmatic to achieve the maximal level of differentiation with a minimal set of genetic loci. The ideal MLST scheme for the *Cryptococcus* species complex should fulfill two criteria: (i) it should amplify and type the same genes from all five serotypes/eight molecular types using the same set of primers, and (ii) the selected genes should contain sufficient sequence diversity to produce a discriminatory typing scheme. Taking these facts into account, the working group has selected a set of seven gene loci for a cryptococcal consensus MLST scheme based on the results obtained in the previously published studies by Litvintseva *et al.* [23], Fraser *et al.* [28], and additional unpublished data obtained by Meyer *et al.* and Fisher *et al.* Special emphasis was placed on using loci that exhibited the largest number of different allele types, as well as the potential to use the same primers with all eight major molecular types identified previously for *C. neoformans* and *C. gattii.* These gene loci included six housekeeping genes *CAP59, GPD1, LAC1, PLB1, SOD1, URA5*, from which three genes code for cryptococcal virulence factors: the polysaccharide capsule (*CAP59*), melanin synthesis (*LAC1*) and cell invasion (*PLB1*), and the intergenic spacer, IGS1, which was selected based on its high allelic diversity.

All the herein proposed MLST loci, except for the *CAP59* locus, are similar to the ones used previously enabling the incorporation of, and comparisons with all previously obtained data. The region of the *CAP59* locus proposed for the consensus MLST scheme represents a different fragment of the *CAP59* gene used by Litvintseva *et al.* [23] (Fig. 2). This new locus was chosen based on the fact that it can be amplified from all eight molecular types using the same primers.

An additional locus, *TEF1*, which also showed high discriminatory power when used for *C. neoformans* var. *grubii* and for *C. gattii* molecular type VGII, was excluded from the consensus typing scheme. This was based on the fact that sequence data are only available for *C. neoformans* var. *grubii* and technical problems had been encountered when amplifying this locus. However, this locus may offer additional discrimination in some of the eight major molecular types.

To enable amplification of all seven loci from the eight major molecular types of *C. neoformans* and *C. gattii*, the previously published primers were tested on all eight major molecular types in three of the six laboratories (Teun Boekhout's laboratory at the CBS, June Kwon Chung's laboratory at the NIH, Matthew Fisher's laboratory at the Imperial College, Wieland Meyer's laboratory at the University of Sydney, Tom Mitchell's laboratory at Duke University, and Maria Anna Viviani's laboratory at the Università degli Studi di Milano) that collaborated in the development of the herein presented consensus MLST scheme. Satisfactory amplifications were obtained for all loci except for the *SOD1* locus, where two different sets of primers were finally used to amplify either VNI–VNIV for *C. neoformans* or VGI–VGIV for *C. gattii* (Table 3). The specific primers and the suggested amplification conditions to amplify the seven gene loci are given in Table 3. Primer directions are listed according to the orientation in the genome sequence of the strain H99 at the Broad Institute

(http://www.broad.mit.edu). Variations in the quality of the amplification products, resulting from either the Taq DNA polymerase enzyme or the PCR machine and PCR conditions used, were observed between participating laboratories. For that reason, the amplification conditions given in Table 3 should only serve as a guideline that may be optimized by individual laboratories.

Automatic allele type and sequence type retrieval

Allele types for *C. neoformans* were assigned according to Litvintseva *et al.* [23] and for *C. gattii* according to by Fraser *et al.* [28], if applicable. The exact start- and endpoints for the sequence of each analyzed locus are given in Table 3 based on the H99 genome sequence at the Broad Institute (http://www.broad.mit.edu/), these may change over time if more strains are studied. The latest sequence cut points are listed at the webpage for each locus. To standardize the assignment of allele types (AT) and sequence types (ST), a centralized globally accessible MLST database will be established at www.mlst.net/. The online software NRDB (http://linux.mlst.net/nrdb/nrdb.htm) allows for an automatic retrieval of allele and sequence types and will assign a new allele and sequence type for any submitted unknown sequence. These are then uploaded to the database *via* a database curator. The designated curators are contactable *via* the website.

Conclusion

In conclusion the ISHAM working group on 'Genotyping of *Cryptococcus neoformans* and *C. gattii'* proposes the following set of genetic loci as an international standard for multi-locus sequence typing for *C. neoformans* and *C. gattii: CAP59, GPD1, LAC1, PLB1, SOD1, URA5* and IGS1.

Acknowledgments

The authors would like to thank Matthew O'Sullivan for allowing us to use the software page developed as part of his PhD to determine the number of gene loci to be essential for an MLST scheme based on the Simpson's index of diversity. This work was supported by an NH&MRC project grant #352303 to Wieland Meyer. June Kwon-Chung was supported by funds from the intramural program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health, USA. Matthew Fisher and David Aanensen were supported by the Wellcome Trust. Sitali Simwami was supported by the BBSRC, UK. Ferry Hagen was supported by funds from the Odo van Vloten Foundation. Anastasia P. Litvintseva and Thomas G. Mitchell were supported by a US Public Health Service NIH grant AI 25783. Luciana Trilles was supported by CAPES scholarship from the Ministério da Educação, Brazil. Sirada Kaocharoen was supported by the Chulalongkorn University Graduate Scholarship to commemorate the 72th anniversary of his majesty King Bhumibol Adulyadej, Thailand.

References

1. Evans EE. The antigenic composition of *Cryptococcus neoformans.* I. A serologic classification by means of the capsular and agglutination reactions. J Immunol 1950;64:423–430. [PubMed: 15415610]
2. Kwon-Chung KJ. Morphogenesis of *Filobasidiella neoformans*, the sexual state of *Cryptococcus neoformans*. Mycologia 1976;68:821–833. [PubMed: 790172]
3. Kwon-Chung KJ. A new species of *Filobasidiella*, the sexual state of *Cryptococcus neoformans* B and C serotypes. Mycologia 1976;68:943–946. [PubMed: 790173]
4. Kwon-Chung KJ, Varma A. Do major species concepts support one, two or more species within *Cryptococcus neoformans*? FEMS Yeast Res 2006;6:574–587. [PubMed: 16696653]
5. Franzot SP, Salkin IF, Casadevall A. *Cryptococcus neoformans* var. *grubii:* separate varietal status for *Cryptococcus neoformans* serotype A isolates. J Clin Microbiol 1999;37:838–840. [PubMed: 9986871]
6. Kwon-Chung KJ. A new genus, *Filobasidiella*, the perfect state of *Cryptococcus neoformans*. Mycologia 1975;67:1197–1200. [PubMed: 765816]

7. Kwon-Chung KJ, Boekhout T, Fell JW, Diaz M. (1557) Proposal to conserve the name *Cryptococcus gattii* against *C. hondurianus* and *C. basillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae). Taxon 2002;51:804–806.
8. Bovers M, Hagen F, Kuramae EE, et al. AIDS patient death caused by novel *Cryptococcus neoformans* x *C. gattii* hybrid. Emerg Infect Dis 2008;14:1105–1108. [PubMed: 18598632]
9. Bovers M, Hagen F, Kuranae EE, et al. Unique hybrids between the fungal pathogens *Cryptococcus neoformans* and *Cryptococcus gattii*. FEMS Yeast Res 2006;6:599–607. [PubMed: 16696655]
10. Spitzer ED, Spitzer SG. Use of a dispersed repetitive DNA element to distinguish clinical isolates of *Cryptococcus neoformans*. J Clin Microbiol 1992;30:1094–1097. [PubMed: 1349898]
11. Varma A, Kwon-Chung KJ. DNA probe for strain typing of *Cryptococcus neoformans*. J Clin Microbiol 1992;30:2960–2967. [PubMed: 1452666]
12. Meyer, W.; Mitchell, TG. PCR fingerprinting to distinguish species and strains of yeast. In: Maresca, B.; Kobayashi, GS., editors. Molecular biology of pathogenic fungi: A laboratory manual. New York: Telos Press; 1993. p. 293-302.
13. Viviani MA, Wen H, Roverselli A, et al. Identification by polymerase chain reaction fingerprinting of *Cryptococcus neoformans* serotype AD. J Med Vet Mycol 1997;35:355–360. [PubMed: 9402529]
14. Meyer W, Marszewska K, Amirmostofian M, et al. Molecular typing of global isolates of *Cryptococcus neoformans* var *neoformans* by polymerase chain reaction fingerprinting and randomly amplified polymorphic DNA – a pilot study to standardize techniques on which to base a detailed epidemiological survey. Electrophoresis 1999;20:1790–1799. [PubMed: 10435451]
15. Cogliati M, Allaria M, Tortorano AM, Viviani MA. Genotyping *Cryptococcus neoformans* var. *neoformans* with specific primers designed from PCR-fingerprinting bands sequenced using a modified PCR-based strategy. Med Mycol 2000;38:97–103. [PubMed: 10817225]
16. Meyer W, Castaneda A, Jackson S, Huynh M, Castaneda E. Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates. Emerg Infect Dis 2003;9:189–195. [PubMed: 12603989]
17. Ruma P, Chen SC, Sorrell TC, Brownlee AG. Characterization of *Cryptococcus neoformans* by random DNA amplification. Lett Appl Microbiol 1996;23:312–316. [PubMed: 8987712]
18. Chen SCA, Brownlee A, Sorrell T, et al. Identification by random amplification of polymorphic DNA (RAPD) of a common molecular type of *C. neoformans* var *neoformans* in patients with AIDS. J Infect Dis 1996;173:754–758. [PubMed: 8627047]
19. Sorrell TC, Chen SC, Ruma P, et al. Concordance of clinical and environmental isolates of *Cryptococcus neoformans* var *gattii* by random amplification of polymorphic DNA analysis and PCR fingerprinting. J Clin Microbiol 1996;34:1253–1260. [PubMed: 8727912]
20. Boekhout T, van Belkum A, Leenders AC, et al. Molecular typing of *Cryptococcus neoformans:* taxonomic and epidemiological aspects. Int J Syst Bacteriol 1997;47:432–442. [PubMed: 9103633]
21. Boekhout T, Theelen B, Diaz M, et al. Hybrid genotypes in the pathogenic yeast *Cryptococcus neoformans*. Microbiology 2001;147:891–907. [PubMed: 11283285]
22. Halliday CL, Bui T, Krockenberger M, Malik R, Ellis DH, Carter DA. Presence of alpha and a mating types in environmental and clinical collections of *Cryptococcus neoformans* var *gattii* strains from Australia. J Clin Microbiol 1999;37:2920–2926. [PubMed: 10449476]
23. Litvintseva AP, Thakur R, Vilgalys R, Mitchell TG. Multilocus sequence typing reveals three genetic subpopulations of *Cryptococcus neoformans* var *grubii* (serotype A), including a unique population in Botswana. Genetics 2006;172:2223–2238. [PubMed: 16322524]
24. Velegraki A, Kiosses VG, Kansouzidou A, et al. Prospective use of RFLP analysis on amplified *Cryptococcus neoformans URA5* gene sequences for rapid identification of varieties and serotypes in clinical samples. Med Mycol 2001;39:409–417. [PubMed: 12054051]
25. Latouche GN, Huynh M, Sorrell TC, Meyer W. PCR-restriction fragment length polymorphism analysis of the phospholipase B (*PLB1*) gene for subtyping of *Cryptococcus neoformans* isolates. Appl Environ Microbiol 2003;69:2080–2086. [PubMed: 12676686]
26. Diaz MR, Boekhout T, Theelen B, Fell JW. Molecular sequence analyses of the intergenicspacer (IGS) associated with rDNA of the two varieties of the pathogenic yeast, *Cryptococcus neoformans*. Systemat Appl Microbiol 2000;23:535–545.

27. Bovers M, Hagen F, Kuramae, Boekhout T. Six monophyletic lineages identified within *Cryptococcus neoformans* and *Cryptococcus gattii* by multi-locus sequence typing. Fungal Gen Biol 2008;45:400–421.
28. Fraser JA, Giles SS, Wenink EC, et al. Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak. Nature 2005;437:1360–1364. [PubMed: 16222245]
29. Hanafy A, Kaocharoen S, Jover-Botella A, et al. Multi-locus microsatellite typing for *Cryptococcus neoformans* var *grubii*. Med Mycol 2008;46:685–696. [PubMed: 18608927]
30. Karaoglu H, Man Ying Lee C, Carter D, Meyer W. Development of polymorphic microsatellite markers for *Cryptococcus neoformans*. Mol Ecol Res 2008;8:1136–1138.
31. Diaz MR, Boekhout T, Kiesling T, et al. Comparative analysis of the intergenic spacer regions and population structure of the species complex of the pathogenic fungus: *Cryptococcus neoformans*. FEMS Yeast Res 2005;5:1129–1140. [PubMed: 16061425]
32. Taylor JW, Fisher MC. Fungal multi-locus sequence typing - it's not just for bacteria. Curr Opin Micro 2003;6:1–6.
33. Bougnoux ME, Aanensen DM, Morand S, et al. Multi-locus sequence typing of *Candida albicans:* strategies, data exchange and applications. Infect Genet Evol 2004;4:243–252. [PubMed: 15450203]
34. Dodgson, Ar; Pujol, C.; Denning, DW.; Soil, DR.; Fox, AJ. Multi-locus sequence typing of *Candida glabrata* reveals geographically enriched clades. J Clin Micro 2003;41:5709–5717.
35. Tavanti A, Davidson AD, Johnson EM, et al. Multilocus sequence typing for differentiation of strains of *Candida tropicalis*. J Clin Micro 2005;43:5593–5600.
36. Koufopanou V, Burt A, Taylor J. Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. Proc Natl Acad Sci USA 1994;94:5478–5482. [PubMed: 9144263]
37. Kasuga T, White TJ, Koenig G, et al. Phylogeography of the fungal pathogen *Histoplasma capsulatum*. Mol Ecol 2003;12:73–84.
38. Simpson EH. Measurement of diversity. Nature 1949;163:688.
39. Feng X, Yao Z, Ren D, Liao W. Simultaneous identification of molecular and mating types within the *Cryptococcus* species complex by PCR-RFLP analysis. J Med Microbiol 2008;57:1481–1490. [PubMed: 19018017]
40. Katsu M, Kidd S, Ando A, et al. The internal transcribed spacers and 5.8S rRNA gene show extensive diversity among isolates of the *Cryptococcus neoformans* species complex. FEMS Yeast Res 2004;4:377–388. [PubMed: 14734018]
41. Sanfelice F. Contributo alla morfologia e biologia dei blastomiceti che si sviluppano nei succhi di alcuni frutti. Ann Igien 1894:463–495.
42. Varma A, Kwon-Chung KJ. Restriction fragment polymorphism in mitochondrial DNA of *Cryptococcus neoformans*. J Gen Microbiol 1989;135:3353–3362. [PubMed: 2576873]
43. Gatti F, Eeckels H. An atypical strain of *Cryptococcus neoformans* (Sanfelice) Vuillemin. Part I: Description of the diseases and of the strain. Ann Soc Belge Méd Trop 1970;50:689–694.
44. Kidd S, Hagen F, Tscharke R, et al. A rare genotype of *Cryptococcus gattii* caused the Cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). PNAS 2004;107:17258–17263. [PubMed: 15572442]
45. Kwon-Chung KJ, Bennett JE, Theodore TS. *Cryptococcus bacillisporus* sp. nov. serotype B-C of *Cryptococcus neoformans*. Int J Syst Bacteriol 1978;28:616–620.
46. Bolton LA, Lobetti RG, Evezard DN, et al. Cryptococcosis in captive cheetah (*Acinonyx jubatus*): two cases. J S Afr Vet Assoc 1999;70:35–39. [PubMed: 10855821]
47. D'Souza CA, Hagen F, Boekhout T, Cox GM, Heitman J. Investigation of the basis of virulence in serotype A strains of *Cryptococcus neoformans* from apparently immunocompetent individuals. Curr Genet 2004;46:92–102. [PubMed: 15150669]

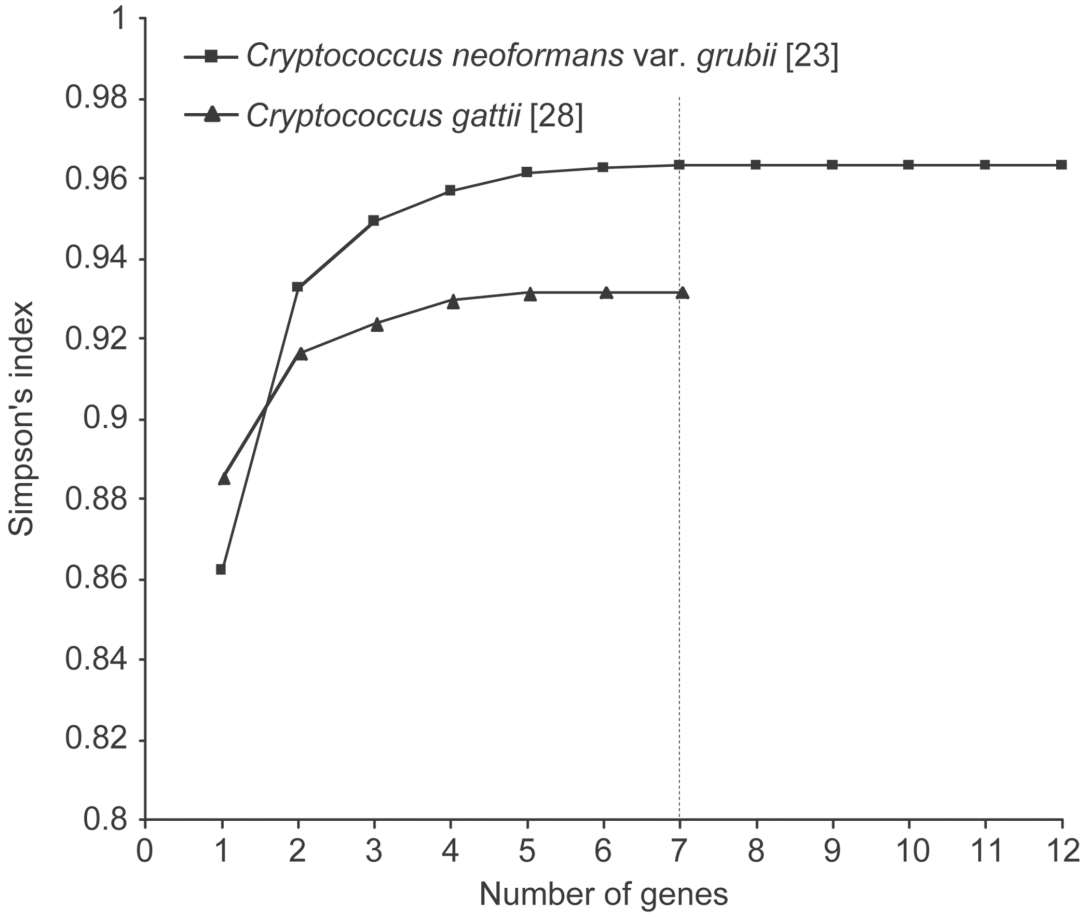


Fig. 1.
Number of genes necessary to differentiate all sequence types obtained in the two previously published cryptococcal MLST schemes based on Simpson's index of diversity [38]. For the Fraser *et al.* [28] data set the mating type locus specific genes, *SXI***a** or *SXIα*, have been excluded for this analysis since they can't be amplified from all isolates.

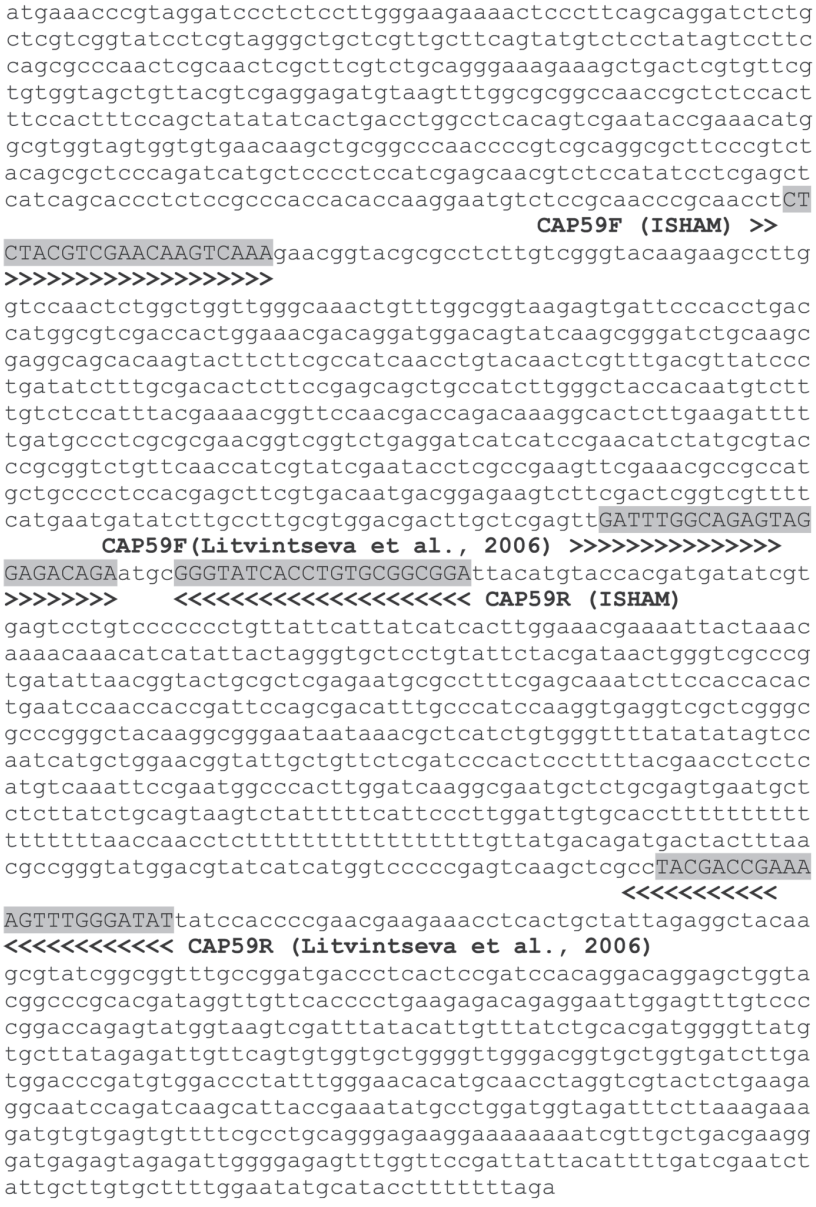


Fig. 2.

CAP59 genomic sequence of strain H99 (*Cryptococcus neoformans* var. *grubii*, VNI, http://www.broad.mit.edu) used as master sequence, indicating the location of the Cryptococcal Working Group (ISHAM) proposed consensus primers used in Fraser *et al.* [28] and the location of the primers used in Litvintseva *et al.* [23].

Table 1

Concordance of different molecular typing methods used for *Cryptococcus neoformans* and *Cryptococcus gattii*

Species/Variety/ Hybrid	Serotype	PCR-fingerprinting molecular type Meyer *et al.* [14,16,39]	PCR-fingerprinting molecular type Viviani *et al.* [13]	AFLP genotype Boekhout *et al.* [21]	AFLP genotype Litvintseva *et al.* [23]	*URA5* RFLP type Meyer *et al.* [16]	*PLB1* RFLP type Latouche *et al.* [25]	IGS genotype Diaz *et al.* [26,31]	ITS genotype Katsu *et al.* [40]
C. neoformans var. grubii	A	VNI	VN6 (VN5)	AFLP1	VNI	VNI	A1	1A/1B	ITS1
	A	VNII		AFLP1A/AFLP1B	VNB	VNII		1A	ITS1
	A	VNII	VN7	AFLP1A/AFLP1B	VNII	VNII	A2	1C	ITS1
AD Hybrid	AD	VNIII	VN3/VN4	AFLP3		VNIII	A3	2C	ITS1/ITS2
C. neoformans var. neoformans	D	VNIV	VN1 (VN2)	AFLP2		VNIV	A4	2A/2B /2C	ITS2
C. gattii	B/C	VGI		AFLP4A/AFLP4B		VGI	A5	4	ITS3/ITS7
	B/C	VGII		AFLP6		VGII	A6	3	ITS4
	B/C	VGIII		AFLP5A/AFLP5B/ AFLP5C		VGIII	A7	5	ITS5
	B/C	VGIV		AFLP7		VGIV	A8	6	ITS6

Table 2

Standard/reference strains for *Cryptococcus neoformans* and *Cryptococcus gattii* strain typing

CBS #	ATCC#	FGS#	Other numbers	MAT & Serotype	Comments	References
Cryptococcus neoformans						
Cryptococcus neoformans var. grubii						
VNI (Meyer *et al*. [14,16])=AFLP1 (Boekhout *et al*. [21])=VN6 (VN5) (Viviani *et al*. [13])						
CBS 10085	ATCC MYA-4564	10415	WM 148; W10; Brown	αA	1989, Australia, NSW, Sydney, clinical, CSF, HIV–, isolated by Sharon Chen	[14,18]
CBS 8710	ATCC 48922	9487	DUMC 135.97; H99; NYSD 1649; CBS 10515; WM 04.15	αA	1978, USA, NC, Durham, clinical, CSF, patient with Hodgkin's lymphoma, isolated by John Perfect/Wiley Schell, type culture of *C. neoformans* var. *grubii* , genome sequence strain	[5]
VNII (Meyer *et al*. [14/16])=AFLP1A (Boekhout *et al*. [21])=VN7 (Viviani *et al*. [13])						
CBS 10084	ATCC MYA-4565	10416	WM 626, W20; Cetin	αA	1993, Australia, NSW, Sydney, clinical, CSF, HIV–, isolated by Sharon Chen	[14,18]
AD hybrid						
VNIII (Meyer *et al*. [14,16])=AFLP3 (Boekhout *et al*. [21])=VN33VN4 (Viviani *et al*. [13])						
CBS 10080	ATCC MYA-4566	10417	WM 628; 88B5400; Zapf	αA/*a*D	1988, Australia, VIC, Melbourne, clinical, CSF, HIV+, isolated by Bryan Speed	[14,18]
CBS 132	ATCC 32045	-	CCRC 20528; DBVPG 6010; IFO 0608; IGC 3957; NRRL Y-2534	αA/*a*D	1894, Italy, environmental, fermenting fruit juice, isolated by F. Sanfelice, type culture for *C. neoformans*	[41]
Cryptococcus neoformans var. neoformans						
VNIV (Meyer *et al*. [14,16])=AFLP2 (Boekhout *et al*. [21])=VN1 (VN2) (Viviani *et al*. [13])						
CBS 10079	ATCC MYA-4567	10418	WM 629; B 87455, Borg, F 14	αD	1987, Australia, VIC, Melbourne, clinical, blood, HIV+, isolated by Bryan Speed	[14]
CBS 6900	ATCC 34873	10423	B-3501; DBVPG 6228; CBS 7697	αD	1975, USA, MD, Bethesda, NIH, crossing of NIH 12 x NIH 433, isolated by June Kwon-Chung	[42]
Cryptococcus gattii						
VGI (Meyer *et al*. [16])=AFLP4 (Boekhout *et al*. [21])						
CBS 10078	ATCC MYA-4560	10419	WM 179; Bryon; H33.1; MH56	αB	1993, Australia, NSW, Sydney, clinical, CSF, HIV −, isolated by Sharon Chen	[16,18]
CBS 6289	ATCC 32269	-	MUCL 30449, RV 20186; CBS 8273	αB	1966, Congo, Kinshasa, clinical, CSF, isolated by E. Gatti/R. Eeckels, type strain of *C. neoformans* var. *gattii*,	[43]
CBS 10510	-	-	WM 276; TCS -SC1	αB	1993, Australia, NSW, Mt Annan National Park, environmental, *Eycalyptus tereticornis* woody debris, isolated by Tania Sorrell/Sharon Chen, genome sequence strain	[16]

CBS #	ATCC#	FGS#	Other numbers	MAT & Serotype	Comments	References
VGII (Meyer *et al.* [16])=AFLP6 (Boekhout *et al.* [21])						
CBS 10082	ATCC MYA-4561	10420	WM 178; 49435; Colter; IFM 50894	αB	1991, Australia, NSW, Sydney, clinical, CSF, HIV –, isolated by Sharon Chen	[16]
CBS 10514	-	-	CDC R265; WM 02.32	αB	2001, Canada, BC, Duncan, Vancouver Island, clinical, bronchial wash, isolated by British Columbia CDC, high virulent Vancouver Island outbreak strain, VGIIa, genome sequence strain	[44]
VGIII (Meyer *et al.* [16])=AFLP5 (Boekhout *et al.* [21])						
CBS 10081	ATCC MYA-4562	10421	WM 175; WM 161; E698; 689; TP 0689; D1.13H	αB	1992, USA, California, San Diego, Blind Recreation Center/Park Boulevard UPAS street, environmental, Eucalyptus spp. woody debris, isolated by Tania Pfeifer/David Ellis	[16,19]
CBS 6955	ATCC 32608	10424	DBVPG 6225; MUCL 30454; NIH 191; CBS 6916	αC	Before 1970, USA, San Fernando, California, clinical, CSF.	[45]
VGIV (Meyer *et al.* [16])=AFLP7 (Boekhout *et al.* [21])						
CBS 10101	ATCC MYA-4563	10422	WM 779; King Cheetah; IFM 50896	αC	1994, South Africa, Johannesburg, veterinary, Cheetah, isolated by Valarie Davis	[16,46]

Table 3

MLST loci information

Gene locus	Gene product	Chromo-some location[a]	Primer name and sequence (If not specified differently primers listed will work for *C.n.* and *C.g.*)	Amplification conditions	No. of bases analysed (bp)[a]	Analysed sequence fragment, start (5′) and end (3′) points[a]	Ref.
CAP59	Capsular associated protein	1	**CAP59F** 5′ CTCTACGTCGA GCAAGTCAAG 3′ **CAP59R** 5′ TCCGCTGCA CAAGTGATACCC 3′	94°C 3min; 35 cycles: 94°C 30s, 568C 30s, 72°C 1min **Alternative conditions:** 30 cycles: 94°C 30s, 64°C 30s, 72°C 1min **or:** 30 cycles: 95°C 3min, 95°C 30s, 54°C 30s, 72°C 1min	559	5′- ACG**G**TA**C**GCGCCG **G**AG**A**CAGAATG-3′	[28]
			Alternative primers: **CAP59LF** 5′ GTGAACAA GCTGCGGC 3′ **CAP59LR** 5′ GGATTCAG TGTGGTGGAAGA 3′	35 cycles: 94°C 30s, 60°C 30s, 72°C 1min			[current study]
GPD1	Glyceraldehyde-3-phosphate dehydrogenase	7	**GPD1F** 5′ CCACCGAACCC TTCTAGGATA 3′ **GPD1R** 5′ CTTCTTGGCA CCTCCCTTGAG 3′	94°C 3min; 35 cycles: 94°C 45s, 63°C 1min, 72°C 2min **Alternative conditions:** 12 cycles; 62 − 56°C step-down 2°C every 2 cycles 95°C 3 min; 95°C 30 sec, 62 −56°C 30 s, 72°C 1 min; followed by 25 cycles: 95°C 30 s, 56°C 30 s, 72°C 1 min	543	5′- GGTTTCGGTACGG GACCC**T**GCCAA-3′	[28]
LAC1	Laccase	8	**LAC1F** 5′ AACATGTTCCCT GGGCCTGTG 3′ **LAC1R** 5′ ATGAGAATTG AATCGCCTTGT 3′	94°C 3min; 30 cycles: 94°C 30s, 58°C 30s, 72°C 1min **Alternative conditions:** 30 cycles: 95°C 30s, 50°C 30s, 72°C 1min	469	5′- GTAAGTA**T**CAGCT C**AA**GCTA**A**ACA-3′	[28]
PLB1	Phospholipase	12	**PLB1F** 5′ CTTCAGGCGGA GAGAGGTTT 3′ **PLB1R** 5′ GATTTGGCGT TGGTTTCAGT 3′	94°C 3min; 30 cycles: 94°C 45s, 61°C 45s, 72°C 1min **Alternative conditions:** 12 cycles; 62 −56°C step-down 2°C every 2 cycles 95°C 3 min; 95°C 30 s, 62 −56°C 30 s, 72°C 1 min; followed by 25 cycles: 95°C 30 s, 56°C 30 s, 72°C 1 min	532	5′- TGT**T**AC**T**TGGATT **C**TGGAACATCG-3′	[23]
SOD1	Cu, Zn superoxide dismutase	5	**Primers for *C.n.*** **SOD1CNF** 5′AAGCCTCT CATCCATATCTT 3′ **SOD1CNR** 5′TTCAACCAC GAATATGTA 3′ **Primers for *C.g.*** **SOD1CGF** 5′ GATCCTCAC GCCATTACG 3′ **SOD1CGR** 5′ GAATGATG CGCTTAGTTGGA 3′	94°C 3min; 35 cycles: 94°C 30s, 52°C 30s, 72°C 1.5min	700	5′ - **C**CACGTGCTC**GC**A CC**TGTCAA**TGC-3′	[46]

Gene locus	Gene product	Chromo-some location[a]	Primer name and sequence (If not specified differently primers listed will work for *C.n.* and *C.g.*)	Amplification conditions	No. of bases analysed (bp)[a]	Analysed sequence fragment, start (5′) and end (3′) points[a]	Ref.
			Alternative primers for *C.n.*: **SOD1-f** 5′ TCTAATCGAAA TGGTCAAGG 3′ **SOD1-r** 5′ CGCAGCTGTT CGTCTGGATA 3′	12 cycles; 62 −56°C step-down 2°C every 2 cycles 95°C 3 min; 95°C 30 sec, 62 −56°C 30 sec, 72°C 1 min; followed by 25 cycles: 95°C 30 sec, 56°C 30 sec, 72°C 1 min	535	5′ -ATCGCTCACCGCT GCCCATTGTCA-3′	[23]
URA5	Orotidine monophosphate pyrophosphorylase	8	**URA5F** 5′ ATGTCCTCCCA AGCCCTCGAC 3′ **URA5R** 5′ TTAAGACCTCT GAACACCGTACTC 3′	94°C 3min; 35 cycles: 94°C 45s, 63°C 1min, 72°C 2min **Alternative conditions:** 30 cycles: 94°C 45s, 63°C 1min, 72°C 2min (*C.n.*) 26 cycles: 94°C 30 s, 68°C 30s, 72°C 30s (*C.g.*) **or:** 30 cycles: 95°C 3 min; 95°C 30 sec, 63°C 30 sec, 72°C 1min	601	5′ - TTTTCGGCAACTCT TGGAA**A**GCTC-3′	[16]
IGS1	Ribosomal RNA intergenic spacer	2	**IGSF** 5′ ATCCTTTGCAGA CGACTTGA 3′ **IGSR** 5′ GTGATCAGTGC ATTGCATGA 3′	94°C 3min; 35 cycles: 94°C 30s, 60°C 30s, 72°C 1min **Alternative conditions:** 30 cycles: 94°C 30s, 56°C 30s, 72°C 1min	723	5′ - TAAGCCCTTGTTA**A** AGAT**TTA**T**TG**-3′	[23]

Note:

[a]The sequences of the genome of strain H99 (*C. neoformans* var. *grubii*, VNI) at the Broad Institute (http://www.broad.mit.edu) were used as the master sequences. Nucleotide bases shown in bold typeface denote nucleotide bases that could vary between the different molecular types.

Med Mycol. Author manuscript; available in PMC 2010 June 12.

OPEN ACCESS Freely available online PLoS PATHOGENS

Low Diversity *Cryptococcus neoformans* Variety *grubii* Multilocus Sequence Types from Thailand Are Consistent with an Ancestral African Origin

Sitali P. Simwami[1], Kantarawee Khayhan[2,3], Daniel A. Henk[1], David M. Aanensen[1], Teun Boekhout[2], Ferry Hagen[2], Annemarie E. Brouwer[4,5], Thomas S. Harrison[6], Christl A. Donnelly[1], Matthew C. Fisher[1]*

1 Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College London, London, United Kingdom, **2** CBS Fungal Biodiversity Centre, Utrecht, The Netherlands, **3** Department of Microbiology and Parasitology, School of Medical Science, Naresuan University Phayao, Phayao, Thailand, **4** Department of General Internal Medicine and Nijmegen University Center for Infectious Diseases, Radboud University Medical Centre, Nijmegen, The Netherlands, **5** St. Elisabeth Hospital, Tilburg, The Netherlands, **6** Department of Infectious Diseases, St George's Hospital Medical School, London, United Kingdom

Abstract

The global burden of HIV-associated cryptococcal meningitis is estimated at nearly one million cases per year, causing up to a third of all AIDS-related deaths. Molecular epidemiology constitutes the main methodology for understanding the factors underpinning the emergence of this understudied, yet increasingly important, group of pathogenic fungi. *Cryptococcus* species are notable in the degree that virulence differs amongst lineages, and highly-virulent emerging lineages are changing patterns of human disease both temporally and spatially. *Cryptococcus neoformans* variety *grubii* (*Cng*, serotype A) constitutes the most ubiquitous cause of cryptococcal meningitis worldwide, however patterns of molecular diversity are understudied across some regions experiencing significant burdens of disease. We compared 183 clinical and environmental isolates of *Cng* from one such region, Thailand, Southeast Asia, against a global MLST database of 77 *Cng* isolates. Population genetic analyses showed that Thailand isolates from 11 provinces were highly homogenous, consisting of the same genetic background (globally known as VNI) and exhibiting only ten nearly identical sequence types (STs), with three (STs 44, 45 and 46) dominating our sample. This population contains significantly less diversity when compared against the global population of *Cng*, specifically Africa. Genetic diversity in *Cng* was significantly subdivided at the continental level with nearly half (47%) of the global STs unique to a genetically diverse and recombining population in Botswana. These patterns of diversity, when combined with evidence from haplotypic networks and coalescent analyses of global populations, are highly suggestive of an expansion of the *Cng* VNI clade out of Africa, leading to a limited number of genotypes founding the Asian populations. Divergence time testing estimates the time to the most common ancestor between the African and Asian populations to be 6,920 years ago (95% HPD 122.96 - 27,177.76). Further high-density sampling of global *Cng* STs is now necessary to resolve the temporal sequence underlying the global emergence of this human pathogen.

Citation: Simwami SP, Khayhan K, Henk DA, Aanensen DM, Boekhout T, et al. (2011) Low Diversity *Cryptococcus neoformans* Variety *grubii* Multilocus Sequence Types from Thailand Are Consistent with an Ancestral African Origin. PLoS Pathog 7(4): e1001343. doi:10.1371/journal.ppat.1001343

Editor: Joseph Heitman, Duke University Medical Center, United States of America

Received July 14, 2010; **Accepted** April 15, 2011; **Published** April 28, 2011

Funding: This work was funded by grants from the Wellcome Trust to MC Fisher, (http://www.wellcome.ac.uk/), the Biotechnology and Biological Sciences Research Council, grant number BB/D52637X/1 (www.bbsrc.ac.uk) as well as the Naresuan University Phayao Staff Development Project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mathew.fisher@imperial.ac.uk

Introduction

Cryptococcus neoformans (*Cn*) is an encapsulated basidiomycetous yeast, and the etiological agent of the invasive fungal infection cryptococcosis. The first clinical discovery of *Cn* was in 1894, and this pathogen has since become one of the leading causes of mycotic morbidity and mortality worldwide [1,2,3]. Capable of causing disease among both immunocompetent and immunocompromised individuals, the most common manifestation of cryptococcosis is cryptococcal meningitis (CM) [4,5]. The HIV/AIDS epidemic has driven increased *Cryptococcus* infection rates via the rapid increase of immunosuppressed populations [1,6,7]. Patients with HIV-related CM must undergo maintenance anti-fungal therapy life-long or until immunoreconstitution is reached by antiretroviral therapy [8], and mortality rates remain unacceptably high [3].

Originally believed to be a single species, two distinct varieties of *Cn* have been described, corresponding to three serotypes: *Cn* var *grubii* (serotype A; henceforth *Cng*), *Cn* var *neoformans* (serotype D) and AD hybrids [9]. *C. gattii*, a second species of the genus *Cryptococcus*, consists of serotypes B and C [10], and is also capable of forming hybrids with *Cn* [11,12,13]. Molecular typing has resulted in these two species being further subdivided into eight major molecular types: VNI and VNII (serotype A; var *grubii*), VNIII (hybrid serotype AD; var *neoformans*), VNIV (serotype D; var *neoformans*), VGI, VGII, VGIII and VGIV (serotypes B and C; var *gattii*) [12,13,14,15]. Within *Cng*, VNI predominates worldwide, including in Southeast Asian countries such as Thailand [16] and Malaysia [17]. *Cn* has two mating types, *MAT*α and *MAT***a**, controlled by a single locus, two allele mating system [18]. Globally, there is a predominance of mating-type *MAT*α among

Author Summary

Cryptococcus neoformans is a species complex of often highly pathogenic fungi that cause significant disease in humans. *Cryptococcus* is notable in the degree that virulence differs amongst genotypes, and highly-virulent emerging lineages are changing patterns of disease in time and space. *Cryptococcus neoformans* variety *grubii* (*Cng*) causes meningitis among HIV/AIDS patients, up to 1 million cases/year resulting in over 600,000 mortalities. Despite these rates of mortality being comparable to those caused by malaria (one million mortalities per annum), cryptococcal meningitis receives only a fraction of the attention, funding and control granted to more widely recognised diseases. This study uses multilocus sequence typing to compare the genetic diversity of *Cng* in a largely unstudied country with an emerging HIV epidemic, Thailand, against the diversity seen elsewhere. We found that *Cng* in Thailand exhibits significantly less genetic diversity in comparison to other areas of the world, especially Africa. Analyses dating the pathogen's origin in Thailand support the introduction of a limited number of genotypes into Southeast Asia from an ancestral African population within the last 7,000 years. These findings show the power associated with the collection of global sequence databases in order to better understand the evolution of major fungal pathogens.

both environmental and clinical samples across serotypes [19,20,21,22,23]. An exception is the less common AD hybrid, 68% of which possess the *MAT*a allele from serotype A as well as the *MAT*α allele from serotype D [19]. This discrepancy in mating type prevalence is also observed in other pathogenic fungi including *Histoplasma capsulatum* and several species of dermatophyte fungi [24,25,26,27,28].

Cng (serotype A) is widely associated with avian excreta and other organic substrates [29,30,31,32], and is known to infect mainly immunocompromised hosts [1,33], although there has been evidence of cryptococcosis due to *Cng* among patients with no underlying disease [34,35,36]. Distributed nearly worldwide and commonly isolated from the environment, this variety is responsible for about 95% of cryptococcal infections worldwide [31] and 98% of infections among AIDS patients [6]. However, despite the emerging importance of this pathogen and increased research effort [13,37], aspects of the pathogen's global population genetic structure remain undetermined. This is especially true for Southeast Asia where cryptococcosis affects nearly 20% of HIV infected patients [38] in this highly populous region.

An accurate description of the genetic composition of fungal pathogen populations is important from several standpoints: quantifying the amount and distribution of polymorphisms across space and time enables the identification of population-level processes that ultimately lead to an understanding of the process of infection, such as the reservoirs, transmissibility and longevity of populations and their component genotypes. Increasingly, it is being recognised that specific genotypes act as markers of lineages that exhibit enhanced or reduced virulence [39,40,41,42]. Therefore, an accurate understanding of the genetics of these pathogens clarifies their current and future evolutionary trajectories, and their potential to alter the burden of human disease.

To accurately discriminate between isolates of *Cng* and to enable the rapid acquisition of global genotypic data, the International Society of Human and Animal Mycoses (ISHAM) special working group on *Cryptococcus* and cryptococcosis recognized the need for a cross-platform consensus-typing scheme for *Cn*. This typing scheme needed to be able to incorporate the findings from previous global-typing projects, while being universally applicable, publicly available and able to integrate new data as they emerged. Previously, PCR fingerprinting with the minisatellite-specific core sequence of the wild-type phage M13 or microsatellites was utilized in local-scale studies on patterns of genetic diversity, identifying three major molecular types of *Cng*, VNI, VNII and VNB [37,43]. The ISHAM group has selected multi-locus sequence typing (MLST) using seven loci as the method of choice for global molecular epidemiological typing of *Cryptococcus* species *Cng* [44]. The molecular type (VN system) [15] has been maintained as the standardized naming system for specific related clades of sequence types (STs). Using MLST-approaches, Litvintseva *et al.* (2006) have demonstrated marked heterogeneity in the global distribution of VN-types with a highly genetically diverse, area-specific and recombining population of VNB genotypes in Africa (Botswana) [37].

Increasingly, it is recognised that many human infectious diseases have emerged within the last 11,000 years, following the rise of agriculture and domestication of animals [45]. The consequential globalisation of microbes that have been carried along with this human expansion has left its mark in the population genetic structure of both transmissible [46] and non-transmissible environmental pathogens [47]. One such pathogen is the sister species of *Cng*, *C. gattii*, which has seen a rapid rise in human infections in the non-tropical Pacific Northwest areas of Canada and the United States. Here the introduction of *C. gattii* is believed to have occurred more recently, perhaps vectored by the international trade in Eucalyptus trees from Australia where the species is most commonly found [40,42,48]. The discovery of a population displaying ancestral characteristics in southern Africa, and a global distribution of clonally-derived and genetically homogenous VNI genotypes [37], has led Litvintseva *et al*, 2006 to hypothesise that *Cng* has an evolutionary origin in Africa followed by a global expansion, possibly vectored by the migration of avian species (conference abstract, Fungal Genetics Reports: 56S). The common pigeon (*Columba livia*), originating in Africa, is considered a mechanistic carrier and potential spreader of the fungus, its faeces being a common environmental source of *Cng* [49,50,51]. Although unable to systemically colonize these birds, *Cng* can survive the elevated temperatures within their gastrointestinal tract (41 - 42°C), as well as remain alive for up to two years in the birds' excreta [50]. These birds were domesticated in Africa approximately 5,000 years ago and introduced to Europe, then subsequently distributed to many parts of the world during the European expansion in the last 500 years [52,53]; a range expansion that may have led to pigeon vectors allowing *Cng* to broaden its global ecological range. While wind transport has also been hypothesized as a potential method of the global dispersal of *Cng*, as demonstrated by the potential for dispersal of *Coccidioides immitis* by wind-blown arthroconidia [54], Casadeval and Perfect state that this is unlikely, due to the *Cng* basidiospores being unsuitable for long-distance wind dispersal [31].

The aim of this study was to describe the population genetic structure of the previously untyped, but clinically important, population of *Cng* that infects HIV/AIDS patients in Thailand, Southeast Asia, with the intention of integrating these data into broader global patterns. Our specific goals were (i) to describe the genetic structure of this population of *Cng* using MLST, (ii) to compare the population genetic structure of these isolates against the global collection of *Cng* STs and (iii) to investigate potential associations between infecting genotypes of *Cng* and disease progression among HIV-AIDS patients.

Results

Mating-type and serotypes of *Cng* isolates

All 183 Thai isolates typed in this study were *Cng (*serotype A) and of mating type MATα. Ten were from environmental sources in Chiang Mai, Northern Thailand, while 83 of the 173 clinical isolates (48%) originated from the North, 78 from the Northeast (45%) and 9 (5%) from the South of Thailand (three were of unknown origin; table 1). All 77 of the global isolates were also *Cng*. Thirteen percent of these ($n = 10$) were of mating type MAT**a**, nine originating from Botswana, and one from Tanzania (table S1) [41]. Previously typed by both Amplified Fragment Length Polymorphism (AFLP) and MLST, three molecular groups within serotype A were present in the global isolates: VNI = 48 (62%), VNII = 9 (12%) and VNB = 20 (26%) [37].

MLST determination

Sequence data were obtained for all 183 Thai isolates typed at the seven loci (table 1). The aligned sequences of the concatenated loci were 3,959 base pairs in total, with 112 polymorphic sites (20 parsimony informative and 92 singleton sites). The seven loci yielded 23 allele types (ATs), eight of which were novel to the Thai population of *Cng* (table 1). Loci *IGS1* and *SOD1* consisted entirely of novel ATs, while *CAP59*, *GPD1* and *PLB1* were made up of previously described ATs [37]. We identified 10 multilocus sequence types (STs) within the Thai isolates.

The collection of 77 global isolates of *Cng* yielded 86 ATs and 43 STs. The concatenated sequences were 3,970 base pairs in length, with 190 variable sites. The ten new STs described in Thailand were allocated consecutive numbers ST 44-53 (table 1), resulting in a complete dataset of 53 global STs for *Cng* (table S1). ST44 accounted for 38% of the Thai isolates ($n = 70$), ST45 for 43% ($n = 78$) and ST46 for 14% ($n = 26$) (table 1). STs 44 and 45 collectively contained 81% of all the isolates and differed only at the *LAC1* gene (nucleotide positions 36, 190, 232 and 338). STs 48 to 53 consisted of single isolates, all of which differed from at least one other ST at a single locus. Nine of the ten environmental isolates shared identical genotypes with clinical isolates.

Analyses of genetic variation and phylogeny reveal a genetically depauperate Thai *Cng* population

Initial analyses using eBURST, a web-enabled clustering tool at http://cneoformans.mlst.net/, revealed spatial differentiation between the Thai *Cng* population when compared to the current global population (figure 1). This tool infers patterns of evolutionary descent among clusters of related genotypes from MLST data and identifies mutually exclusive groups of related genotypes within populations. Widespread relatedness was demonstrated within Thailand, shown by the grouping of the majority of Thai STs into a single eBURST group linked by single-locus variants (SLVs; ST44, 45, 49, 50 and 52). STs identified by eBURST as present both in Thailand and elsewhere in the global dataset were highlighted (pink text; ST4, 6, 46; figure 1) and those only found in Thailand shown in green (ST44, 45, 47, 48, 49, 50, 51, 52, 53).

The average nucleotide diversity within the Thai population was explored at all seven loci using haplotypic diversity (H_d), the number of nucleotide differences per site (π) and Watterson's estimate of the population scaled mutation rate (θ). The average estimates of these statistics for the concatenated sequences were low ($Hd = 0.19$, $\pi = 0.001$ and $\theta = 0.005$ respectively; table S2), reflecting the low number of haplotypes which ranged from two to six at the seven loci. Locus *LAC1,* 467 base pairs long, had the greatest number of segregating sites ($n = 61$), while *CAP59* had the lowest haplotypic diversity and population scaled mutation rate (0.01 and 0.002, respectively).

The spatial partitioning of genetic variability in the Thai *Cng* population typed in this study ($n = 183$) was examined using Analysis of Molecular Variance (AMOVA). This analysis demonstrated that only a small proportion, 5% ($p<0.013$), of the total estimated variance was attributable to the among-population variance component between the three Thai regions (table 2).

A Principal Component Analysis (PCA) was used to assess the hierarchical structuring of the genetic population of *Cng* in Thailand. The genetic structure captured by the first two principal components was depicted by the individual genotypes (represented by dots) clustering into three groups and summarised by 95% ellipses. The typology of the individual allelic profiles revealed little differentiation between the 183 isolates from the three regions (figure 2). A maximum likelihood tree depicting the phylogenetic relationships within Thailand supported this genetic homogeneity, with all but the single isolates of STs 48 and 53 (CM21 and 50NC1 respectively; table 1) clustering together with high bootstrap support (bootstrap 100%; figure 3). Although identical to ST46 at six of the seven loci, 50NCI of ST53 was an outlier due to variations in its nucleotide sequence at *LAC1* (table 1). CM21's allelic profile, on the other hand, consisted of seven ATs which were not found in any other Thai isolate typed in this study.

Population structure of the wider Asian population of *Cng*

Three isolates from the previously typed *Cng* population originated from HIV positive patients in Bangkok, Thailand [37,55], and were of ST4 (th84, th206) and ST6 (th104; table S1). The STs of the newly typed Thai isolates consisted of a 12 nucleotide insertion at the *IGS1* locus, as well as a six and a three-nucleotide insertion at *SOD1*; these mutations were not found within the ATs of the previously typed Thai isolates (table S3). A further five isolates included in this study are of Asian origin: jp1086, jp1088 and J1 from Japan, and in2629 and in2632 from India (table S1). 25% of the variation between the Thai isolates typed in this study and the eight isolates of wider Asian origin was due to among population differences (data not shown). These eight previously typed isolates of Asian origin were combined with the 183 Thai isolates typed in this study to form the Asian population ($n = 191$) which was then compared to the remaining global isolates, also grouped according to geographic location: Africa ($n = 44$), North America ($n = 19$) and South America ($n = 5$).

Genetic structure of the global population subdivided into geographically defined subpopulations

AMOVA attributed 52% of the variation in the global population of *Cng* to differences between the four geographically defined sub-populations ($\Phi PT = 0.52$, $p = 0.001$; table 2). We excluded Europe due to a small sample size ($n = 2$). The first principal coordinate in the inter-class PCA for the global samples' allelic profiles distinguished the Asian population (pink ellipse, group 1) from the rest of the global population subsets (Africa, North and South America), $p<0.001$ (figure 4). A dendrogram inferring the relationships between all isolates delineated three major groups within the global population: VNI ($n = 230$; type isolates WM148, H99), VNII ($n = 10$; type isolates WM626) and

Table 1. The allelic profiles of the 183 *Cng* isolates from Thailand typed by MLST in this study.

Name	CAP59 allele (501 bp)	GPD1 allele (489 bp)	IGS1 allele (709 bp)	LAC1 allele (471 bp)	PLB1 allele (533 bp)	SOD1 allele (527 bp)	URA5 allele (637 bp)	ST	Strain origin (if known)
CN5010	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, blood
CN4998	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4995	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4989	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4988	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4987	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4964	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4947	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4945	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4944	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN4943	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand
CN4942	1	1	**19**	3	2	**13**	5	44	Lampang, Thailand, CSF
CN4941	1	1	**19**	3	2	**13**	5	44	Thailand, CSF
CN4940	1	1	**19**	3	2	**13**	5	44	Thailand, CSF
CN4926	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4919	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4918	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4917	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4903	1	1	**19**	3	2	**13**	5	44	Chiang Rai, Thailand, CSF
CN4901	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, CSF
CN49005	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand
4-187	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
269	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
4-315	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
1-587	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
1219	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
4_83	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
1-588	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
4-202	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
1-846	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, clinical
2551-07	1	1	**19**	3	2	**13**	5	44	Songkhla, Thailand, CSF
2550 II-07	1	1	**19**	3	2	**13**	5	44	Songkhla, Thailand, blood
2461-07	1	1	**19**	3	2	**13**	5	44	Songkhla, Thailand, CSF
CM 1	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 6	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 7	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 8	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 12	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 13	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 17	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 18	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 22	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 23	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 25	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 26	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 33	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 37	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 38	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF

Table 1. Cont.

Name	CAP59 allele (501 bp)	GPD1 allele (489 bp)	IGS1 allele (709 bp)	LAC1 allele (471 bp)	PLB1 allele (533 bp)	SOD1 allele (527 bp)	URA5 allele (637 bp)	ST	Strain origin (if known)
CM 39	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 40	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 41	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM42	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 43	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 44	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 46	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 47	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 48	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 49	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 51	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 55	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 56	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 57	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 58	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 59	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 61	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
CM 63	1	1	**19**	3	2	**13**	5	44	Ubon Ratchathani, Thailand, CSF
K 2	1	1	**19**	3	2	**13**	5	44	Khon Kaen, Thailand, crypto patient
Pg 1	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, pigeon dropping
D 6	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, dove dropping
D 1	1	1	**19**	3	2	**13**	5	44	Chiang Mai, Thailand, dove dropping
CN5019	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, blood
CN5017	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, CSF
CN5014	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, blood
CN5013	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, CSF
CN5011	1	1	**19**	4	2	**13**	5	45	Thailand, clinical
CN5009	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, blood
CN5005	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, blood
CN5003	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, blood
CN5002	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, blood
CN5001	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, CSF
CN4970	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4968	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4957	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, CSF
CN4956	1	1	**19**	4	2	**13**	5	45	Chiang Rai, Thailand, CSF
CN4955	1	1	**19**	4	2	**13**	5	45	Thailand, BAL
CN4954	1	1	**19**	4	2	**13**	5	45	Lampang, Thailand, CSF
CN4952	1	1	**19**	4	2	**13**	5	45	Tak, Thailand, CSF
CN4950	1	1	**19**	4	2	**13**	5	45	Lampoon, Thailand, CSF
CN4949	1	1	**19**	4	2	**13**	5	45	Lampoon, Thailand, CSF
CN4938	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4937	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4936	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4934	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4933	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF

Table 1. Cont.

Name	CAP59 allele (501 bp)	GPD1 allele (489 bp)	IGS1 allele (709 bp)	LAC1 allele (471 bp)	PLB1 allele (533 bp)	SOD1 allele (527 bp)	URA5 allele (637 bp)	ST	Strain origin (if known)
CN4932	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4931	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4927	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4915	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4914	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4909	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4907	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4905	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4904	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN4902	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
CN49008	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, CSF
4-319	1	1	**19**	4	2	**13**	5	45	Khon Kaen, Thailand, clinical
50NC2	1	1	**19**	4	2	**13**	5	45	Nan, Thailand, clinical
50NC5	1	1	**19**	4	2	**13**	5	45	Nan, Thailand, clinical
11112	1	1	**19**	4	2	**13**	5	45	Khon Kaen, Thailand, clinical
11109	1	1	**19**	4	2	**13**	5	45	Khon Kaen, Thailand, clinical
4-231	1	1	**19**	4	2	**13**	5	45	Khon Kaen, Thailand, clinical
P6	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, clinical
4-253	1	1	**19**	4	2	**13**	5	45	Khon Kaen, Thailand, clinical
4-381	1	1	**19**	4	2	**13**	5	45	Khon Kaen, Thailand, clinical
20662-07	1	1	**19**	4	2	**13**	5	45	Songkhla, Thailand, blood
28170-07	1	1	**19**	4	2	**13**	5	45	Songkhla, Thailand, CSF
1111I-08	1	1	**19**	4	2	**13**	5	45	Pattani, Thailand, blood/HIV-
2895I-08	1	1	**19**	4	2	**13**	5	45	Pattani, Thailand, blood/HIV-
4500-07	1	1	**19**	4	2	**13**	5	45	Pattani, Thailand, blood
CM 2	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 3	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 4	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 5	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 10	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 14	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 11	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 15	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM16	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 20	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 24	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 27	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 28	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 29	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 32	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 34	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 36	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 45	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 50	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 52	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 60	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
CM 64	1	1	**19**	4	2	**13**	5	45	Ubon Ratchathani, Thailand, CSF
Pt 9	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, crypto patient

Table 1. Cont.

Name	CAP59 allele (501 bp)	GPD1 allele (489 bp)	IGS1 allele (709 bp)	LAC1 allele (471 bp)	PLB1 allele (533 bp)	SOD1 allele (527 bp)	URA5 allele (637 bp)	ST	Strain origin (if known)
Pt 3	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, crypto patient
Pt 1	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, crypto patient
D 2	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, dove dropping
D 3	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, dove dropping
Pg 2	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, pigeon dropping
Pg 26	1	1	**19**	4	2	**13**	5	45	Chiang Mai, Thailand, pigeon dropping
CN49004	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN48	1	3	**19**	5	2	**13**	1	46	Khon Kaen, Thailand, clinical
1-488	1	3	**19**	5	2	**13**	1	46	Khon Kaen, Thailand, clinical
1-489	1	3	**19**	5	2	**13**	1	46	Khon Kaen, Thailand, clinical
CM 30	1	3	**19**	5	2	**13**	1	46	Ubon Ratchathani, Thailand, CSF
Pt 12	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, crypto patient
D 5	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, dove dropping
Pg 37	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, pigeon dropping
CN5015	1	3	**19**	5	2	**13**	1	46	Chiang Rai, Thailand, CSF
CN5018	1	3	**19**	5	2	**13**	1	46	Chiang Rai, Thailand, blood
CN5012	1	3	**19**	5	2	**13**	1	46	Chiang Rai, Thailand, CSF
CN5008	1	3	**19**	5	2	**13**	1	46	Chiang Rai, Thailand, CSF
CN4993	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4983	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4980	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4977	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4967	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4960	1	3	**19**	5	2	**13**	1	46	Chiang Rai, Thailand, CSF
CN4948	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4946	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4924	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4921	1	3	**19**	5	2	**13**	1	46	Mae Hong Son, Thailand, CSF
CN4920	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4916	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN4906	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
CN49006	1	3	**19**	5	2	**13**	1	46	Chiang Mai, Thailand, CSF
50NC1	1	3	**19**	**10**	2	**13**	1	53	Nan, Thailand, clinical
Pt 5	1	1	**19**	5	2	**13**	1	51	Chiang Mai, Thailand, crypto patient
CN5007	1	1	**20**	3	4	**13**	1	47	Chiang Rai, Thailand, CSF
1291-09	1	1	**20**	3	4	**13**	1	47	Pattani, Thailand, blood/HIV-
CM 35	1	1	**20**	3	4	**13**	1	47	Ubon Ratchathani, Thailand, CSF
K 45	1	1	**19**	3	4	**13**	5	50	Khon Kaen, Thailand, crypto patient
4_9	1	1	**19**	**9**	2	**13**	5	52	Khon Kaen, Thailand, clinical
D 9	1	1	**19**	4	2	**13**	**14**	49	Chiang Mai, Thailand, dove dropping

Table 1. Cont.

Name	CAP59 allele (501 bp)	GPD1 allele (489 bp)	IGS1 allele (709 bp)	LAC1 allele (471 bp)	PLB1 allele (533 bp)	SOD1 allele (527 bp)	URA5 allele (637 bp)	ST	Strain origin (if known)
CM 21	2	10	**21**	6	11	**14**	4	48	Ubon Ratchathani, Thailand, CSF

bp = base pairs; crypto patient = cryptococcosis patient; novel ATs are in bold.
doi:10.1371/journal.ppat.1001343.t001

VNB ($n = 21$; figure 5). Molecular group VNB was mostly found in Botswana, and consisted of three previously described sub-populations which were geographically and genetically isolated from lineages of *Cng* found elsewhere: VNB-A, VNB-B [56] and VNB-C [41]. Although confined to Botswana in this study, previous studies have reported the occurrence of VNB *Cn* Aα (also known as AFLP genotype 1A) infecting AIDS patients in Rwanda, the USA and Belgium, from the environment in Zaire and Australia and from both clinical and environmental samples in Brazil [13,14,22], South Africa and Columbia [57]. The origin of VNB has previously been hypothesised to be the result of hybridisation between VNI (serotype A, ALFP genotype 1) and VNIV (serotype D, AFLP genotype 2) [14,37]. Eight of the ten African isolates of the rare mating type MAT**a** were from this group. All but one of the Thai isolates typed in this study clustered with the global VNI isolates, with the single isolate, CM21 of ST48 (table 1), falling within molecular group VNII along with reference strain WM626 (bootstrap value 100%; figure 5). Isolate CM21 being of a different VN group explains why it was an outlier in the maximum likelihood tree analysis of the phylogenetic relationships within the Thai STs (figure 3). In addition, isolate 50NCI, the second outlier of ST53, was found to correlate with the VNI group (WM148, H99), also supported by significant bootstrap value ($n = 90\%$; figure 5). In accordance with our PCA, the global phylogenetic analysis showed the previously typed Thai isolates (th84, th206 and th104) grouped with the newly typed Thai isolates (bootstrap support = 70%), while the remaining Asian isolates (J1, jp1086, jp1088, in2629 and in2632) clustered with the Thai isolates within the VNI group (figure 5).

Predominant clonality detected within the Asian *Cng* populations

The Index of Association (I_A) [58] and $\bar{r}_d$ [59] were used to assess the overall association between alleles at the seven MLST loci, testing the null hypothesis of linkage equilibrium. A signature of clonal reproduction is the generation of non-random associations between loci, the amount of which can be estimated using linkage disequilibrium. Random association of alleles at the different loci was rejected for the sub-populations of isolates divided by geographic origin, with Africa having the lowest $\bar{r}_d$ value (0.28, $p<0.001$; table 3). Clone-corrected data confirmed the predominance of clonal reproduction among the *Cng* samples. The proportion of phylogenetically compatible pairs of loci was used to test for linkage disequilibrium in the dataset, with the null hypothesis of free recombination being rejected if there were fewer than two locus pairs with all four allele combinations than expected under panmixis [60]. A significant percentage of phylogenetically compatible loci pairs was found for all geographically defined sub-populations (table 3), and the hypothesis of random mating rejected. The minimum number of recombination events (R_m) [61] was estimated both within an individual locus and between loci (R_m and average R_m respectively; table 4) within described populations Africa, Asia and North America. Despite the main feature of the Asian population ($n = 191$) being strong clonality, some evidence for inter-locus recombination was detected (average $R_m = 5$; table 4). This was low in comparison with the African population, where an average R_m of 12 was observed. Africa also exhibited more intralocus recombination with 5/7 loci showing 1 or more inferred events, as opposed to 1/7 loci in Asia and North America. The locus with the highest inferrred intralocus R_m was *IGS1* for African, Asian and North American populations (table 4); a feature that is perhaps related to the multicopy nature of this locus. When analysed according to molecular group, recombination was detected within the VNI ($n = 230$) and VNB ($n = 10$) populations of the global isolates ($R_m = 6$ and 7, respectively; data not shown) and less so within the VNII population ($n = 21$, $R_m = 1$). The main feature of the Thai VNI *Cng* population is strong clonality, evidence of local clonal expansion within this geographical subset of the recombining global VNI population.

Subpopulations of the global *Cng* population are genetically divergent and differentiated

The average nucleotide diversity within geographically defined subpopulations was calculated at each locus and overall statistical tests included the number of segregating sites (S) and haplotypes (h), haplotypic diversity (H_d), the number of nucleotide differences per site (π) and Watterson's estimate of the population scaled mutation rate (θ). Consistently higher average values of H_d, π and of θ indicated higher levels of within-population variation among the African isolates than were observed in the Asian and South American populations. Similarly, the North American population's average values of H_d (0.75) and θ (0.005) were lower than those of Africa (0.79 and 0.007, respectively; table 4).

Tajima's D tests the null hypothesis that populations are in mutation-drift equilibrium [62]. In the case of significant deviation from zero, the null hypothesis of neutral (random) evolution is rejected, a finding which can be due to the occurrence of natural selection or variable population dynamics. Significant departures from neutrality were detected at five of the seven loci of the Asian population (table 4), all of which had negative values. The remaining three global populations (Africa, North and South America) only had one or no significant departure from zero (table 4). Ramos-Onsins & Rozas' R_2 test which is more powerful at detecting population growth [63] did not detect any deviation from random evolution among any of the populations (table 4).

The divergence among, and differentiation between, the four continental *Cng* populations were estimated using tests based on DNA sequences: the average nucleotide divergence between populations (D_{xy}) [64], a weighted measure of the ratio of the average pair-wise differences within populations to the total average pairwise differences (K^*_{ST})[65] and the

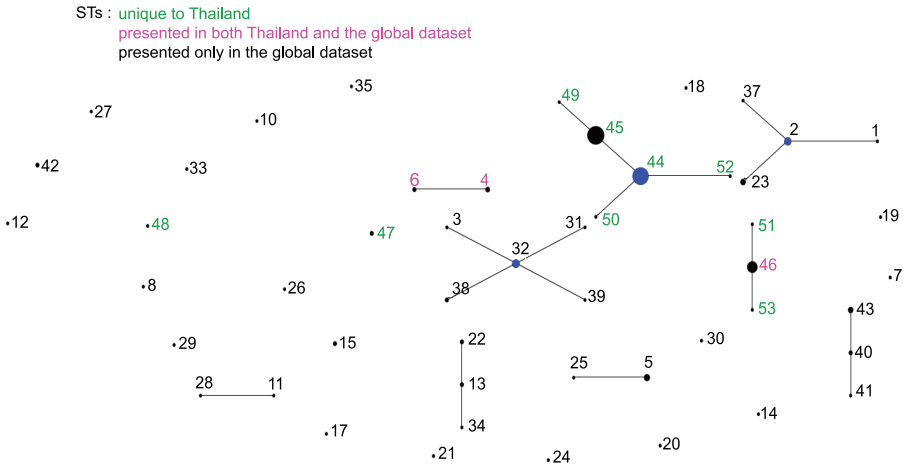


Figure 1. eBURST illustration comparing the isolates from Thailand with the global population of *Cng* used in this study. No. isolates = 176, no. STs = 53, no. re-samplings for bootstrapping = 1000, no. loci per isolate = 7, no. identical loci for group def = 1, no. groups = 1. STs identified by eBURST as present in Thailand and elsewhere in the global dataset are highlighted pink text, those only found in Thailand highlighted green and those only in the global population and not in Thailand are black. Founding genotypes are in blue, and the size of the dots are representative of the number of isolates of that ST.
doi:10.1371/journal.ppat.1001343.g001

nearest-neighbour statistic (S_{nn}) [61,66]. Low levels of nucleotide divergence were observed, with D_{xy} ranging from 0.3 and 0.7%, and no fixed differences found between the various continental populations at the seven loci (table 5A). The total number of shared polymorphisms among populations ranged from ten for Asia vs. South America, to 62 for Africa vs. North America, with locus *IGS1* contributing the most in each case (table 5A). The null hypothesis of no differentiation among populations of *Cng* was rejected for all populations paired with Asia due to significant K^*_{ST} and S_{nn} values (table 5B). Africa and North America were also significantly differentiated, although considerably less so ($K^*_{ST} = 0.03$, $S_{nn} = 0.83$), reflecting the high number of shared polymorphisms (table 5).

Divergence time estimates and haplotype networks support a hypothesis of African ancestry for Asian *Cng* isolates

The time of divergence between the global subpopulations is defined as the mean time to most common recent ancestor (TMRCA) and was estimated using Bayesian markov-chain monte carlo (MCMC) methods in BEAST. Estimates obtained from runs of 10^7 generations, according to three fixed substitution rates

Table 2. Summary of AMOVA of *Cng* isolates, based on the seven polymorphic loci and according to geographical origin.

	d.f.	Sum of squares	Variance components (%)	ΦPT	*P* - value[a]
(i) Thai population: North (*n* = 92), Northeast (*n* = 78), South (*n* = 9)					
Among populations	2	4	0.03 (5)	0.05	0.013
Within populations	176	114	0.65 (95)		
Total	178	118	0.68 (100)		
(ii) Asian and Global populations: Asia (*n* = 191), Global (*n* = 70)					
Among populations	1	12	1.22 (49)	0.49	0.010
Within populations	259	333	1.28 (51)		
Total	260	459	2.51 (100)		
(iii) Global population[b]: Africa (*n* = 44), Asia (*n* = 191), North America (*n* = 19), South America (*n* = 5)					
Among populations	3	145	1.29 (52)	0.52	0.001
Within populations	255	308	1.21 (48)		
Total	258	452	2.5 (100)		

[a]*P* - value estimates are based on 999 permutations.
[b]Europe was excluded due to small sample size (*n* = 2).
doi:10.1371/journal.ppat.1001343.t002

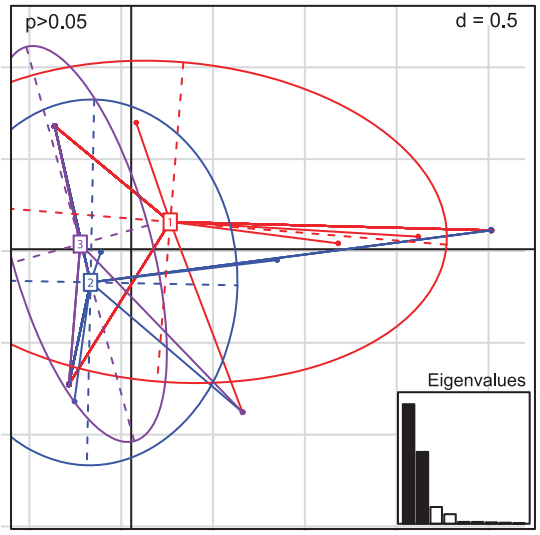


Figure 2. Principle Components Analysis of the allelic profiles of the Thai *Cng* genotypes typed in this study. Individual genotypes (dots) are linked by coloured lines to form clusters which are summarised by coloured ellipses proportional in size to the number of isolates represented. The three groups depicted are numbered and defined according to Thai region: 1 = North (red; $n = 91$), 2 = Northeast (blue; $n = 79$) and 3 = South (purple; $n = 9$). *P* - value is shown and eigenvalues represented in the bar plot.
doi:10.1371/journal.ppat.1001343.g002

estimated for *Eurotiomycetes* [67] and assuming a relaxed log-normal clock, are shown in table 6. Two of the three mutation rates (0.9×10^{-9}, 8.8×10^{-9}) resulted in a TMRCA estimate whose upper and lower bounds span 5,000 years before present (y.b.p.). These values encompass the time of divergence proposed by the "Out of Africa" hypothesis for the global radiation of *Cng*. The highest effective sample size (ESS) was for an estimated rate of 0.9×10^{-9} substitutions per generation. We therefore estimated the mean TMRCA of the African and Asian population to be ≈ 6,921 y.b.p. (95% highest posterior density, HPD = 122–27,178) according to the best representative sample of the model used (XML file, dataset S1). Estimates of mean time to divergence for the two remaining populations were 5,090±1,419 y.b.p. (ESS = 42.09) for North America ($n = 19$) and 4,528±1,287 y.b.p. (ESS = 41.60) for South America ($n = 5$; data not shown).

To further explore the potential African ancestry of the *Cng* population, haplotype networks were constructed for each MLST locus (figure 6), as well as for the concatenated loci (figure S1). Sampled haplotypes are indicated by circles or rectangles colored according to the geographical region from which the sample was collected and proportional in size to observed haplotype frequency. Rectangles depict the haplotype with the highest ancestral probability and each branch indicates a single mutational difference. Internal nodes are representative of ancestral haplotypes, from which apical haplotypes evolved. The STs of non-African genotypes (shown in blue) were few and tended to be found at the apical (*ie.* derived) positions of the networks. The green circles, which represented STs of African origin only, were positioned throughout the networks but were only associated with clinical haplotypes. The combination of the seven networks pointed to an ancestral African population which had the highest variation in haplotype numbers and from which other global haplotypes were derived (figure S1).

Associations between clinical variables and ST

There were no significant associations between the infecting ST and any of the reported baseline clinical variables indicative of disease progression. This lack of association is not surprising, as the genetically highly-related nature of these Thai genotypes is unlikely to lead to detectable variability in their clinical phenotype. The statistical power in this experiment was however sufficient to detect associations between clinical variables and disease progression as we found elevated baseline quantitative cryptococcal culture (range = 30 to 9,200,000) to be significantly associated with early death, with a 500,000 increment in CFU/ml/CSF resulting in a 30.6% increase in odds of death within ten weeks ($p = 0.02$). Similarly, altered mental status at presentation, defined by the presence of a decrease in Glasgow Coma scale or seizures, resulted in a 5.4 fold increased likelihood of death within 10 weeks (95% CI = 1.097 to 27.5; $p = 0.02$). These findings were consistent with previous observations made by Brouwer *et al.*, 2004 [68]. The regression model best describing the prognostic factors of early death also included logarithmic interferon gamma (range = 0.32 to 2.23), which, when decreased by 0.1 in CSF, results in a 29% increase in odds of death within ten weeks ($p = 0.02$; table S4).

Discussion

Affecting nearly 20% of HIV-AIDS patients nationwide, cryptococcosis is a leading AIDS-defining systemic infection in Thailand [38]. The high rates of mortality, re-admissions and relapses are attributed to a combination of factors that include high poverty rates resulting in few being able to afford timely

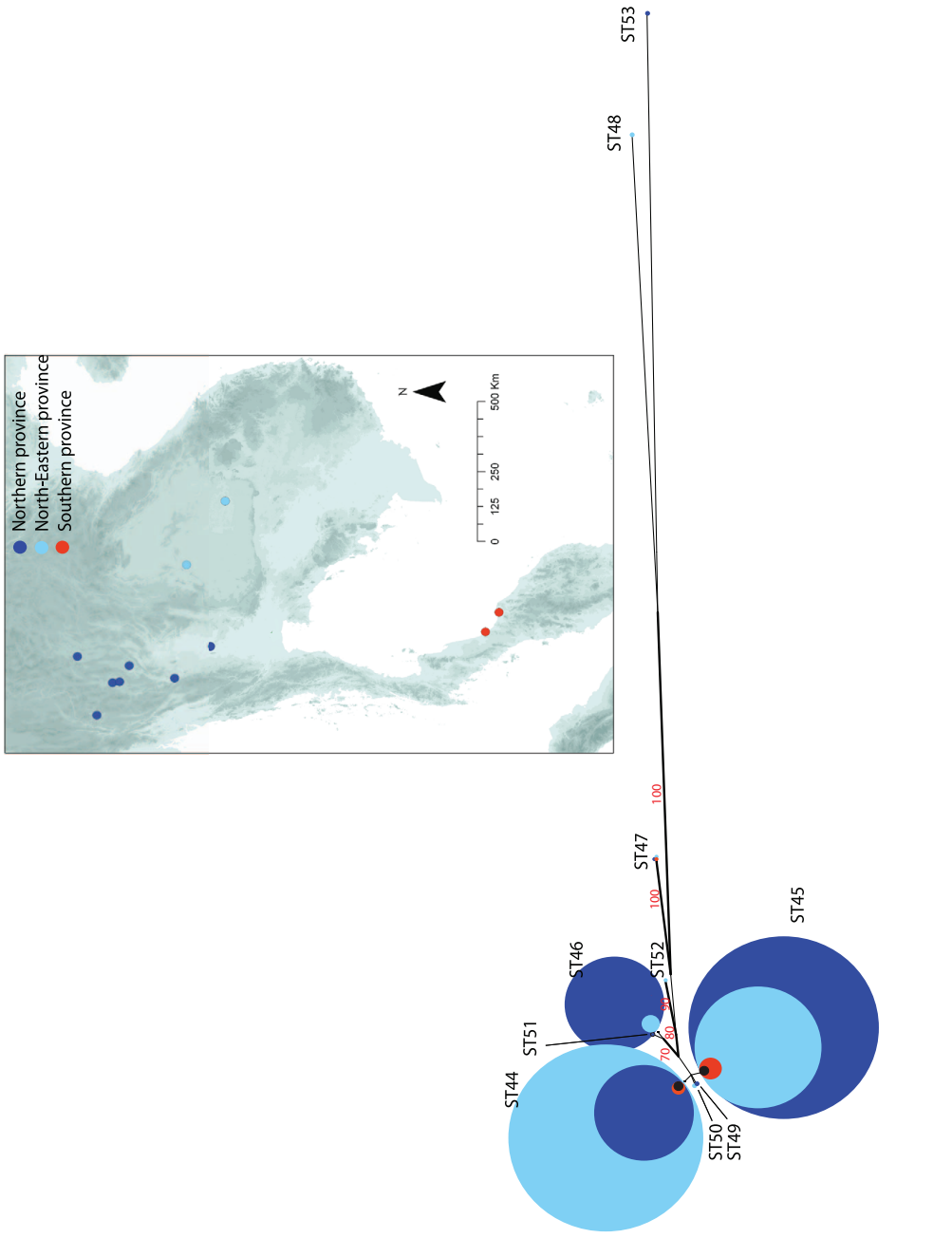

Northern province
North-Eastern province
Southern province
N
0 125 250 500 Km
ST53
ST48
ST47
ST45
ST46
ST52
ST51
ST44
ST50
ST49
100
100
80
70

Figure 3. Neighbour-joining tree inferring the evolutionary relationships of the Thai isolates typed in this study (*n* = 183). Each circle represents a Sequence Type (ST) of the Thai isolates and is proportional in size to the number of isolates of this ST. The isolates are grouped according to three regions of Thailand, Northern province in dark blue (*n* = 91), Northeastern province in light blue (*n* = 79) and Southern province in red (*n* = 9). The four Thai isolates of unknown origin are in black (*n* = 4). The percentage replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) more than 70% of the time (*n*≥70%) are indicated. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site.
doi:10.1371/journal.ppat.1001343.g003

antifungal treatment, the limitations of current antifungal drugs, the limited availability of highly active anti-retroviral therapy (HAART) and the trend of late presentation due to religious and cultural influences [69]. As the population of immunosuppressed individuals increases, the potential for the continued increase in the disease burden of AIDS–related meningitis cannot be ignored, particularly in the developing countries of Southeast Asia [8]. Continued global typing is the key to elucidating the population structure of *Cng* in order to understand the contribution of the pathogen's genotype to the epidemiology of this infection. Therefore, standardisation by ISHAM of the typing methodologies and nomenclature in the study of *Cng* has the potential to greatly facilitate global health efforts to increase our knowledge and surveillance of this pathogenic fungus [44].

We initially used MLST to describe the genetic structure of *Cng* in Thailand. All 183 isolates typed were of *Cng* (serotype A) and mating type α, consistent with previous reports that serotype A, mating type α, is the dominant cause of cryptococcosis among immunocompromised individuals, as well as predominating in the environment [1,8,15,19,43,70,71]. Similarly, all but one Thai isolate, CM21, were of molecular type VNI (figure 5), which is the most prevalent VN-type worldwide [15,43,72], as well as among Southeast Asian populations such as Thailand [16] and Malaysia [17]. MLST revealed ten sequence types (ST44 to 53), three of which accounted for 95% of the isolates typed. Two of these three STs (44 and 45) contained 81% of the 183 isolates (table 1) and differed at only four nucleotide positions within the *LAC1* locus. AMOVA showed that only 5% of the observed genetic variation across Thailand could be attributed to differences among the three regions (table 2), showing that *Cng* exhibits little spatial structure at this geographic scale. PCA (figure 2) and phylogenetic analyses (figure 3) support the conclusion that there is little geographical variation between the regional Thai *Cng* isolates that were typed in this study. This genetic pattern is consistent with that found in Cng isolates from five geographic locations within another Asian country, India [73].

Eight isolates within the previously typed *Cng* population [38] were of Asian origin (table S1). AMOVA revealed 25% of the molecular variance to be due to diversity between this wider Asian population (*n* = 8) and that of the Thai isolates typed in this study (*n* = 183). All the previously typed isolates clustered within groups of the Thai isolates with high bootstrap support, showing that they are highly related; for this reason they were subsequently combined to form the Asian population of *Cng* which was subsequently tested against the global sample of *Cng*.

Our analyses then focused on comparing the type and distribution of diversity between the different continental populations of *Cng*, and is the first time that a global analysis of the distribution of MLST polymorphisms has been undertaken for this

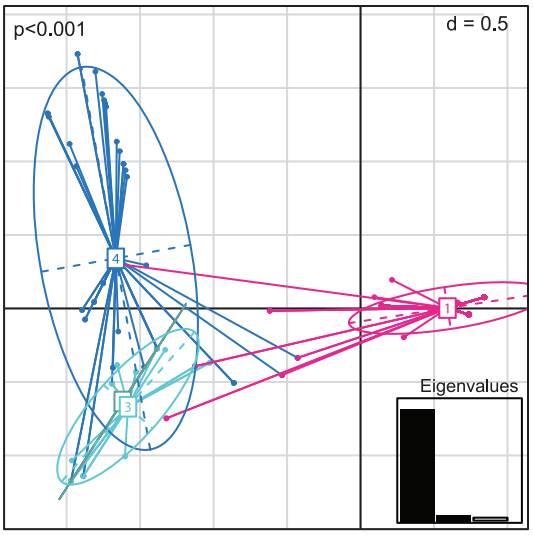


Figure 4. Principle Components Analysis of the allelic profiles of the global *Cng* genotypes analysed in this study. Individual genotypes (dots) are linked by coloured lines to form clusters which are summarised by coloured ellipses proportional in size to the number of isolates represented. The four groups are numbered and defined according continent: 1 = Asia (pink; *n* = 191), 2 = South America (grey; *n* = 5), 3 = North America (light blue; *n* = 19), 4 = Africa (dark blue; *n* = 44). *P* - value is shown and eigenvalues represented in the bar plot.
doi:10.1371/journal.ppat.1001343.g004

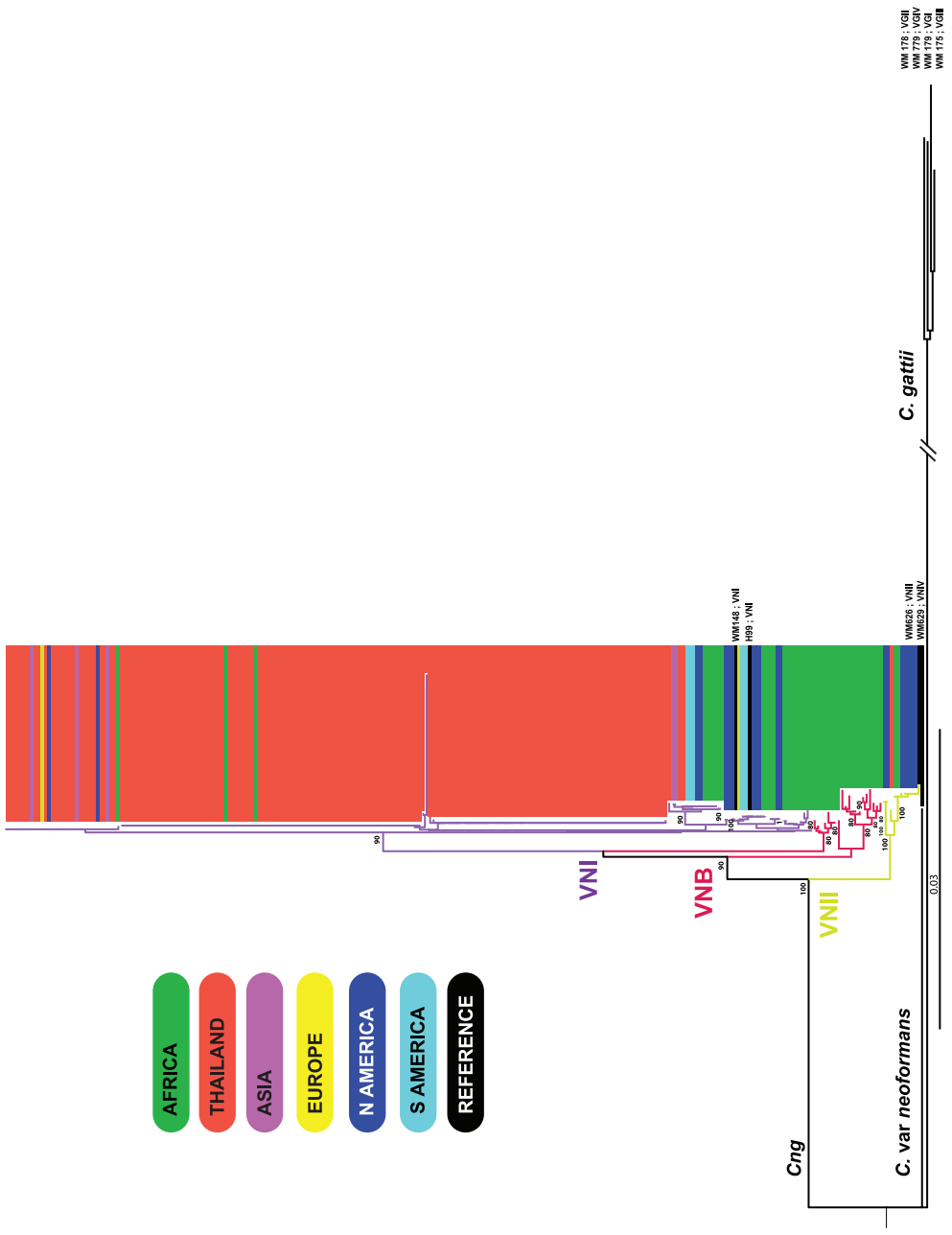

WM 178 : VGIII
WM 179 : VGIV
WM 179 : VGI
WM 175 : VGII
C. gattii
WM148 : VNI
H99 : VNI
WM626 : VNII
WM629 : VNIV
VNI
VNB
VNII
0.03
Cng
C. var neoformans
AFRICA
THAILAND
ASIA
EUROPE
N AMERICA
S AMERICA
REFERENCE

Figure 5. Neighbour-joining tree inferring the evolutionary relationships of the global *Cng* isolates included in this study (*n* = 261). The geographical origins of the isolates are represented by coloured rectangles: green = Africa (*n* = 44), red = Thailand (isolates typed in this study; *n* = 186), purple = remaining Asian isolates (*n* = 5), dark blue = North America (*n* = 19), light blue = South America (*n* = 5) and yellow = Europe (*n* = 2). Black rectangles represent reference strains of known VN molecular types that are detailed on the figure for VNI (WM148, H99; *n* = 232), VNII (WM626; *n* = 11) and VNB (*n* = 21). Reference strains of the *C. gattii* complex (molecular groups VGI – IV) are labelled and serve as an outgroup: WM179, WM178, WM175 and WM779. The percentage replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are indicated if supported by significant bootstrap values (*n*≥80%). The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site.
doi:10.1371/journal.ppat.1001343.g005

pathogen. While sample sizes were low for two regions (Europe and South America), our power to detect differences between continents was satisfactory for the other sampled regions (North America, Africa and Asia). Our data and analyses clearly showed the following facets of *Cng*'s global population structure: 1. the fungus is widely clonally reproducing, 2. recombination, where observed, is geographically proscribed and 3. continental populations are differentiated and vary in their levels of diversity. Below, we discuss and integrate these findings.

Statistically significant tests of non-random association of alleles at the different loci (I_A, $\bar{r}_d$ and PcP; table 3) demonstrated an overwhelmingly clonal population structure within the Asian population of *Cng*. Elsewhere, a similar pattern of clonality was seen for populations of *Cng* sampled from Africa and North America (clone corrected $\bar{r}_d = 0.21$ and 0.36 respectively, $p<0.001$). These results are consistent with previous studies showing that non-meiotic reproduction is the predominate mode of descent in *Cng* worldwide [12,37,41,74,75]. Having said this, recent investigation of the predominance of the α mating type in nature led to the finding that cryptococcal strains of the same mating type within serotypes A and D are capable of sexual reproduction in the form of haploid and monokaryotic fruiting, a process previously believed to be mitotic and asexual [76]. As there have been previous reports of recombination within predominantly clonal populations of *Cng* [56,77,78], including an environmental sample consisting of only MAT-α alleles in the Asian country of India [73], R_m was applied to the different subpopulations of *Cng* despite the strong clonal component detected. This technique detects the minimum number of recombination events that are necessary to explain the distribution of polymorphisms within and between loci. The test demonstrated a high degree of spatial variation in the rates of recombination globally (table 4). Importantly, the highest number of minimum recombination events was detected in the African population (Africa $R_m = 12$; Asia $R_m = 5$; North America $R_m = 4$) and the majority of the MLST loci in Africa showed evidence of intergenic recombination, in comparison with much lower levels detected elsewhere (Africa 5/7 loci; Asia 1/7 loci; North America 1/7 loci). These results are in keeping with studies reporting sexual propagation within both clinical [56] and environmental African isolates of *Cng* [79]. Furthermore, sub-divisions according to VN group showed the African VNB population (n = 21) to be highly recombining (Rm = 7) in comparison to the African VNI group (n = 21, Rm = 3; data not shown), likely due to the high frequency of the **a**-mating type detected in the former (table S1) [37].

Estimates of haplotypic diversity (H_d), mutation rates (θ) and nucleotide differences (π) were consistently greater for Africa relative to populations in other continents (table 4). Africa exhibited the greatest number of haplotypes (Africa = 74> North America = 34> Asia = 24), and the Asian population exhibited the least amount of haplotypic diversity (Africa = 0.79> North America = 0.75> Asia = 0.20). Tajima's D is a statistical test that identifies loci that are evolving under non-random processes, such as selection or demographic expansion or contraction, and showed that 5/7 MLST loci in Asia were significantly non-neutral, compared to only 1/7 loci in North America and 0/7 in Africa. As the MLST loci used to type *Cng* are mostly in housekeeping genes [44], and therefore unlikely to be under strong selection, these differences in Tajima's D are most likely due to demographic effects such as population expansion following a population bottleneck. The possibility of neutrality could not be rejected within any of geographically defined population groups, according to the more powerful R_2 statistical test (table 4), however the results qualitatively mirror those found for Tajima's D (table 4).

Global analyses of pairwise population combinations detected significant genetic differentiation between all *Cng* populations

Table 3. Multilocus linkage disequilibrium analyses for samples of *Cn* var *grubii*.

Population[#]	Total sample			Population	Clone-corrected sample[a]		
	I_A[b]	$\bar{r}_d$[c]	PcP[d]		I_A	$\bar{r}_d$	PcP
Africa (*n* = 44)	1.67***	0.28***	0.43***	Africa (*n* = 33)	1.25***	0.21***	0.43***
Asia (*n* = 191)	1.54***	0.30***	0.67***	Asia (*n* = 14)	1.11***	0.19***	0.67***
North America (*n* = 19)	3.45***	0.58***	1***	North America (*n* = 10)	2.13***	0.36***	1***
Global (*n* = 261)	3.18***	0.53***	0.19***	Global (*n* = 53)	1.53***	0.53***	0.38***

[a]excluding replicate haplotypes;
[b]index of association;
[c]scaled index of association (I_A) by the number of loci (*m* – 1);
[d]percentage of phylogenetically compatible pairs (PcP) of loci.
***P<0.001.
[#]The South American and European populations were not individually analyzed due to their sample sizes being too small (*n* = 5 and 2, respectively), but were included in the global population (*n* = 261).
doi:10.1371/journal.ppat.1001343.t003

Table 4. Polymorphism summary and tests neutral evolution for groups of isolates of *Cn* var *grubii* according to geographic origin.

	Locus	pb[a]	S[b]	*h*[c]	*Hd*[d]	π[e]	θ[f]	D[g]	R_2[h]	Rm[i] #
Africa ($n = 44$)	*CAP59*	501	11	10	0.82	0.004	0.005	-0.79_{ns}	0.08_{ns}	1
	GPD1	489	16	11	0.82	0.006	0.008	-0.55_{ns}	0.09_{ns}	0
	IGS1	704	22	12	0.83	0.006	0.007	-0.50_{ns}	0.10_{ns}	2
	LAC1	470	12	8	0.75	0.006	0.006	0.03_{ns}	0.11_{ns}	0
	PLB1	533	15	11	0.8	0.004	0.006	-1.09_{ns}	0.07_{ns}	1
	SOD1	524	24	10	0.64	0.011	0.011	0.30_{ns}	0.12_{ns}	1
	URA5	636	24	12	0.86	0.008	0.009	-0.43_{ns}	0.10_{ns}	1
			Average		0.79	0.007	0.007			12
Asia ($n = 191$)	*CAP59*	501	5	2	0.01	0.0001	0.002	-1.81*	0.07_{ns}	0
	GPD1	489	6	3	0.28	0.0007	0.002	-1.40_{ns}	0.06_{ns}	0
	IGS1	707	11	3	0.06	0.0008	0.003	-1.71_{ns}	0.03_{ns}	0
	LAC1	474	61	6	0.64	0.0031	0.022	-2.62***	0.06_{ns}	2
	PLB1	533	8	4	0.07	0.0003	0.003	-1.97*	0.05_{ns}	0
	SOD1	526	11	2	0.01	0.0002	0.004	-2.25**	0.07_{ns}	0
	URA5	637	10	4	0.33	0.0007	0.003	-1.78*	0.06_{ns}	0
			Average		0.2	0.0001	0.005			5
North America ($n = 19$)	*CAP59*	501	8	5	0.78	0.006	0.005	1.39_{ns}	0.20_{ns}	0
	GPD1	489	7	5	0.76	0.006	0.004	1.28_{ns}	0.20_{ns}	0
	IGS1	708	16	6	0.77	0.008	0.006	1.09_{ns}	0.18_{ns}	2
	LAC1	471	9	5	0.8	0.008	0.005	1.77_{ns}	0.22_{ns}	0
	PLB1	533	9	5	0.81	0.007	0.005	1.65_{ns}	0.21_{ns}	0
	SOD1	526	12	4	0.57	0.01	0.007	1.80_{ns}	0.21_{ns}	0
	URA5	637	9	4	0.75	0.006	0.004	2.06*	0.23_{ns}	0
			Average		0.75	0.007	0.005			4
South America ($n = 5$)	*CAP59*	501	1	2	0.6	0.001	0.001	1.22_{ns}	0.3_{ns}	0
	GPD1	489	0	1	0	0	0	ND	ND	ND
	IGS1	709	43	2	0.6	0.037	0.03	1.88*	0.3_{ns}	0
	LAC1	470	2	2	0.6	0.003	0.002	1.46_{ns}	0.3_{ns}	0
	PLB1	533	1	2	0.6	0.001	0.001	1.22_{ns}	0.3_{ns}	0
	SOD1	527	0	1	0	0	0	ND	ND	ND
	URA5	637	1	2	0.6	0.001	0.001	1.22_{ns}	0.3_{ns}	0
			Average		0.4	0.006	0.005			0

[a]total number of sites in alignments, excluding indels and missing data;
[b]number of segregating sites;
[c]number of haplotypes;
[d]haplotypic diversity;
[e]average number of nucleotide differences per site;
[f]Watterson's estimate of the population scaled mutation rate, expressed per site [95];
[g]Tajima's D [62];
[h]Ramos-Onsins & Rozas' R_2 [99];
[i]minimum number of recombination events [61];
[#]average Rm = Rm between all seven loci; ND not determined because of no polymorphism. [ns] non-significant ($P>0.05$),
*$P<0.05$, **$P<0.01$, ***$P<0.001$.
doi:10.1371/journal.ppat.1001343.t004

excepting the comparison between North and South America (table 5B), showing that the different continental populations of *Cng* are experiencing divergent evolutionary trajectories. The Asian population's comparatively low genetic diversity, high linkage disequilibrium, non-neutral evolution and lack of geographically defined structure are all consistent with a model of a rapid population expansion from a limited set of ancestors. This is supported by evidence of limited genetic variation within isolates from Northwest India, suggestive of recent origin and/or dispersal of Asian *Cng* isolates [73]. These findings contrast with the African population of *Cng*, which is characterised by high genetic diversity, balanced mating types and elevated recombination rates. This finding that the Asian isolates are genetically monomorphic in relation to African isolates led to our examining the potential of an ancestral African origin of *Cng* using coalescent analyses in BEAST. A substitution rate of 0.9×10^{-9} and a relaxed log-normal

Table 5. (A) Divergence among the sub-populations of the global *Cng* isolates. (B) Differentiation between sub-populations of the global *Cng* isolates.

A.	Africa - Asia			Asia - North America			Asia - South America			Africa - North America			Africa - South America			North America - South America		
Locus	D_{xy} [a]	S_f [b]	Ss [c]	D_{xy}	S_f	Ss	D_{xy}	S_f	Ss	D_{xy}	S_f	Ss	D_{xy}	S_f	Ss	D_{xy}	S_f	Ss
CAP59	0.003	0	3	0.005	0	1	0.001	0	0	0.006	0	6	0.003	0	1	0.005	0	1
GPD1	0.007	0	5	0.005	0	5	0.003	0	0	0.007	0	6	0.005	0	0	0.004	0	0
IGS1	0.004	0	13	0.008	0	13	0.008	0	9	0.009	0	13	0.009	0	9	0.009	0	9
LAC1	0.006	0	4	0.008	0	2	0.004	0	0	0.008	0	9	0.005	0	2	0.007	0	2
PLB1	0.004	0	8	0.006	0	8	0.003	0	1	0.006	0	8	0.003	0	1	0.005	0	1
SOD1	0.008	0	11	0.008	0	12	0.000	0	0	0.013	0	11	0.008	0	0	0.008	0	0
URA5	0.007	0	9	0.006	0	8	0.002	0	0	0.001	0	9	0.006	0	1	0.005	0	1
Average/total	0.005	0	53	0.006	0	49	0.003	0	10	0.007	0	62	0.006	0	14	0.006	0	14
B.[d, e]			Africa		Asia		N. America		S. America									
	Africa				0.11***		0.03**		0.01ns									
	Asia		**0.95*** **				0.08***		0.04***									
	N. America		**0.83*** **		**0.96*** **				0.02ns									
	S. America		**0.86ns**		**0.99*** **		**0.74ns**											

The isolates are subdivided by continent: Africa ($n = 44$), Asia ($n = 191$), North and South America ($n = 19$ and 5, respectively).
[a]minimum estimate of the number of nucleotide differences per site between groups;
[b]number of fixed differences between groups;
[c]number of shared polymorphisms between groups.
[d]**K_{ST}*** values are displayed above the diagonal and represent the weighted measure of the ratio of the average pair-wise differences within groups to the total average pair-wise differences.
[e]**S_{nn}** values are displayed below the diagonal and in bold and represent the proportion of nearest neighbours in sequence space that are found in the same group.
Significance levels for K_{ST} and S_{nn} were assessed using permutation tests, with 1000 permutations:
ns = non-significant, $^{**}P<0.01$, $^{***}P<0.001$.
Europe has been excluded as it contains only two isolates.
doi:10.1371/journal.ppat.1001343.t005

model estimated the time to ancestry of Africa/Asia to be at 6,920 y.b.p. with the 95% HDP levels of 123 – 27,178 (table 6). Ancestral estimations report a mean TMRCA of 5,090±1,419 y.b.p. for North America and 4,528±1,287 y.b.p for South America. However, these last two populations are considerably smaller ($n = 19$ and 5, respectively) leading to wide uncertainty. If a hypothesis of human trade-associated pigeon migration vectoring *Cng* is correct, one would expect Europe to follow Africa, but the current lack of data on *Cng* MLST genotypes in Europe means this cannot currently be tested. However, despite uncertainty in the exact order of the phylogenetic relationships, the 95% HPD estimates for ancestry between the Africa/Asia populations encompass the time frame of the domestication of the birds in Africa approximately 5,000 years ago prior to their introduction to Europe and subsequent distribution worldwide at two of the three substitution rates that we examined. Importantly, haplotype networks for each MLST network show that haplotypes unique to the African population occupy both internal and apical positions within the networks, whilst those unique to the global population are almost always at the derived positions at the network-tips. These data are persuasive evidence for the derivation of these lineages from an ancestral African population (figure 6, figure S1).

Table 6. Bayesian estimates of time (in years) to the most recent common ancestor of *Cng* populations, according to geographic location, calculated under the assumption of three mutation rates and adopting the relaxed uncorrelated lognormal molecular clock model as implemented in BEAST v.1.4.1.

TMRCA	Mutation rates per site per year		
	0.9×10^{-9}	8.8×10^{-9}	16.7×10^{-9}
Africa/Asia	6,921	60,572	1.05×10^{6}
95% HPDI	(123 - 27,178)	(28 - 2.8×10^{5})	(3.8×10^{5}-2.0×10^{6})
ESS	58.9	22.9	44.1
Global	7,103	60,739	1.05×10^{6}
95% HPDI	(123 - 27,178)	(28-2.8×10^{5})	(3.8×10^{5}-2.0×10^{6})
ESS	57.0	22.8	44.0

ESS = Effective sample size.
95% HPDI = 95% highest posterior densities intervals.
doi:10.1371/journal.ppat.1001343.t006

The invasion and expansion of two recombinant genotypes of *C. gattii* in the Pacific Northwest, and their differential virulence, has shown that genotypes of *Cryptococcus* can encode striking different clinical phenotypes [42]. We hypothesised that the bottlenecked diversity that we observe in our Thailand populations of *Cng* would translate into negligible difference in the progression of clinical disease between these highly-related ST's. The fact that one cohort of isolates collected from Sappasitprasong Hospital, Ubon Ratchathani, were highly characterised with respect to the progression of clinical disease following infection led us to test for a relationship between ST and the various clinical variables indicative of the progression of cryptococcosis in AIDS patients. While these sample sizes were sufficient to detect associations between clinical variables and disease progression, as has been previously described by Brouwer *et al* [68], we found no association

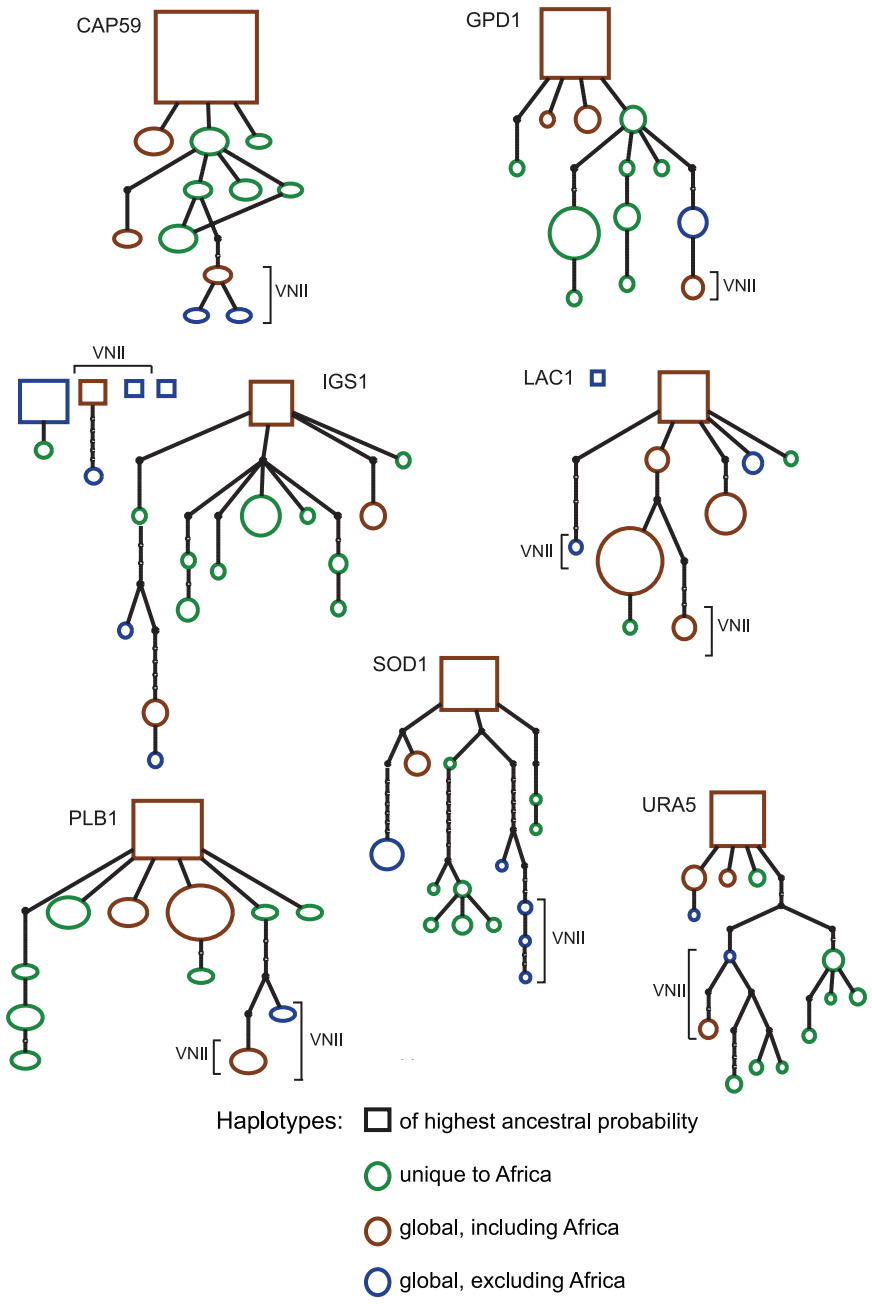

CAP59
VNII
GPD1
VNII
VNII
IGS1
LAC1
VNII
VNII
SOD1
VNII
PLB1
VNII
VNII
URA5
VNII
Haplotypes:
of highest ancestral probability
unique to Africa
global, including Africa
global, excluding Africa

Figure 6. Haplotype networks of the 53 STs of the global *Cng* population at each of the seven loci. Sampled haplotypes are indicated by circles or rectangles colored according to the geographical region from which the sample was collected. STs unique to the African population are shown in green and consist only of clinical isolates. Haplotypes found both in Africa and elsewhere are in brown, while those not found in Africa are represented in blue. Rectangles depict the haplotype with the highest ancestral probability. Each branch indicates a single mutational difference and black dots on the lines are representative of the number of mutational steps required to generate allelic polymorphisms. Circle size is proportional to observed haplotype frequency.
doi:10.1371/journal.ppat.1001343.g006

between ST and disease progression. This is likely due to the fact that 95% of theses isolates were either of ST 44 or ST 45, which differ at only a single locus. As low genetic diversity appears to be the general condition in Asia *Cng*, the variation in clinical phenotype seen in this clinical sample appears overwhelmingly due to host effects as opposed to *Cng* genotype, whereas were we to look at an African cohort, effects owing to *Cng* genotype might be more apparent. A robust comparative analysis between African and Asian *Cng* using either experimental models or further clinical cohorts will be necessary to definitively answer this question.

Our study has shown that a genetically depauperate population of *Cng* infecting Thai HIV-AIDs patients shows many signatures of having been derived from a recombining African population across a timeframe that broadly encapsulates the anthropogenically driven globalisation of many major human infectious diseases. Further, our study has shown the gains that are associated with the collection of global MLST datasets, and sets the stage for integrating future MLST datasets, as well as utilising new deep-sequencing approaches to genotype whole *Cng* genomes in parallel. Further collaborative efforts by the *Cng* research community to integrate such genotyping approaches with spatial collections of isolates and clinical studies will lead to a better understanding of the evolution of this increasingly important, and understudied, emerging human pathogen.

Materials and Methods

Ethics statement

Ethical approval was required for the randomised control trial at Sappasitprasong Hospital, Ubon Ratchathani, the source of some isolates typed in this study. This was approved by the ethical and scientific review subcommittee of the Thai Ministry of Public Health and by the research ethics committee of St George's Hospital, London, UK, with written informed consent obtained for all 64 adults enrolled in this study.

Isolates

The 183 Thailand isolates of *Cng* were acquired from three sources. Fifty-eight clinical isolates were collected during a randomised control trial at Sappasitprasong Hospital, Ubon Ratchathani, Northeast Thailand. This study aimed to compare the efficacy of four randomly assigned anti-fungal treatment combinations in the initial treatment of HIV-associated CM in an antiretroviral therapy (ART) naïve population, enrolling 64 adults with a first episode of cryptococcal meningitis [68]. A further 108 clinical isolates were obtained from a collection of cryptococcal samples managed by the CBS-KNAW Fungal Biodiversity Centre and originated from patients at various hospitals in three Thai regions: 76 in the North, 20 in the Northeast and 9 from the South. Three of these isolates were of unknown provenance. Of the total 173 clinical isolates, 154 (89%) were from HIV/AIDS patients with culture-proven *Cn* isolated from cerebrospinal fluid ($n = 127$), blood ($n = 12$) and broncho-alveolar lavage ($n = 1$). Three were from blood samples of HIV- negative CM patients. Eighteen cryptococcal isolates were provided by Dr. Pojana Sriburee, Chiang Mai University, ten of which were environmental and had been isolated from pigeon and dove guano [80]. One of the eight remaining isolates recovered from cryptococcosis patients was of Japanese origin, and was not considered as part of the Thai dataset (isolate J1; table S1). In total, these three collections yielded 183 isolates from 11 provinces in three regions of Thailand: North ($n = 91$), Northeast ($n = 79$) and South ($n = 9$), four unknown, 6% of which are environmental (table 1, figure 3).

These isolates were then compared to the global MLST dataset as compiled by A. Litvintseva [37], which consisted of 77 isolates whose genotypes and molecular groups had been previously determined by both amplified fragment length polymorphisms (AFLP) and MLST. All 261 *Cng* isolates, including the Japanese isolate J1, were grouped according to geographic origin: Asia ($n = 191$), Africa ($n = 44$), North America ($n = 19$), South America ($n = 5$), Europe ($n = 2$; table S1). As of the 2nd of November, 2009, the MLST scheme contained 53 STs from 232 clinical, 20 environmental isolates and nine unknown of source, from 19 countries worldwide [37,44] (table S1).

Cultivation and DNA extraction

Isolates were cultured on pre-prepared malt extract agar (CM0059, Oxoid, Basingstoke, UK) and DNA extracted using the DNEASY Blood and Tissue Kit (Qiagen, Crawley, UK), then stored at 4°C prior to PCR-amplification. Samples of all cultures were subsequently cryopreserved in YPD (2.5 g Bacto yeast, 5 g Peptone, 5 g Dextrose and 250 ml dH_2O) and 15% glycerol at -80°C.

Mating-type and serotype analyses

The mating type of each of the isolates was determined by four different PCR amplification reactions. Primers specific to the MATα or MAT**a** allele of the STE20 locus for either serotype A or D isolates were used: primers JOHE7270 and JOHE7272 (aA), JOHE7273/JOHE7275 (aD), JOHE7264/JOHE7265 (αA) and JOHE7267/JOHE7268 (αD) [19,22,81]. PCR amplifications with a total volume of 25 µL contained 0.25 µL of 10 mM stock dNTPs, 0.25 µL Taq polymerase, 2.5 µL of buffer, 16.0 µl of sterilised distilled H_20, 1 µl of template DNA and 2.5 µL of each forward and reverse primer at a 10 µM final concentration.

MLST determination

Each isolate was PCR-amplified in 50 µl reaction volumes for each of the seven MLST loci using the primers and protocols detailed in Meyer *et al.*, 2009 [45]. Each locus was subsequently sequenced using TaqFS (Big Dye V1.1) and an Applied Biosystems 3730XL sequencer (Warrington, UK) to determine the forward and reverse DNA sequences of all PCR products.

Sequences were manually edited using CodonCode Aligner (CodonCode Corporation, MA, USA), then aligned in MEGA 4.0 [82]. Alleles at each locus were assigned numbers (Allele Types; ATs) upon comparison with those identified in the global collection [37], resulting in a 7-digit allelic profile for each isolate. Each unique allelic profile was concatenated and assigned a Sequence Type (ST) according to the MLST scheme (http://cneoformans.mlst.net/). Novel STs identified within the Thai population were assigned as additional STs within the global

MLST database. Data analyses were performed on both the Thai population of *Cn* typed in this study ($n = 183$), and on the complete global collection of strains ($n = 261$).

Analysis of genetic structure based on allelic profiles

A hierarchical Analysis of Molecular Variance (AMOVA) was performed in GenAlEx 6.1 for Excel [83] in order to examine the distribution of genetic variation, and to determine the extent of connectivity among populations based on allelic profiles [84]. AMOVA is a statistical technique that estimates the extent of genetic differentiation between individuals and populations directly from molecular data. The technique treats the raw molecular data as a pairwise matrix of genetic distances between all the possible combinations of *Cng* isolates, with sub-matrices corresponding to the different hierarchical data-partitions (here, the genetic differences between *Cng* infecting different host individuals and geographical regions). The data is then analysed within a nested analysis of variance (ANOVA) framework. An F-statistic analogue of the genetic diversity among populations, ΦPT, and between pairs of groups (population pair wise ΦPT) is also reported [84], with significance estimated from 999 random permutations.

Patterns of allelic variability among the MLST genotypes of the Thai isolates typed in this study were investigated by Principle Component Analysis (PCA) using the Adegenet 1.1 package for statistical software R (version 2.6.1). This package is dedicated to the multivariate analysis of genetic markers, illustrating population stratification within a set of genotypes [85]. Diagrams obtained by PCA consist of dots, representing individual genotypes, clustered into groups. Isolates belonging to the same group are linked by matching coloured lines, labelled and summarised by 95% ellipses. Bar plots represent eigenvalues which describe the contributions of the principal coordinates to the genetic structure of the population depicted. Inter-class PCA was performed on the global population of *Cng*, also using Adegenet v1.1. This technique maximizes the variance between pre-defined groups as opposed to the total variance [86]. In order to assess the significance of this hierarchical data-structure, a Monte-Carlo procedure was applied.

Phylogenetic analyses and molecular type determination

Phylogenetic neighbour-joining trees were inferred for each locus as well as concatenated sequences for both the Thai and the global populations, with evolutionary distances computed using the Maximum Composite Likelihood method in MEGA 4.0 [82,87]. The percentage of replicate trees in which the associated taxa clustered together was estimated by the bootstrap test, inferred from 1000 replicates [88]. Molecular VN groupings of the Thai isolates were inferred through phylogenetic and comparative analyses with the global isolates ($n = 77$; table S1). The VN groupings of global isolates were previously determined using phylogenetic methods and non-hierarchical ordination analyses of both AFLP and MLST data [37]. We also included reference strains of known major molecular types of the *C. neoformans*/*C. gattii* species complex: WM148 (serotype A, VNI), WM626 (serotype A, VNII), WM629 (serotype D, VNIV), WM179 (serotype B, VGI), WM178 (serotype B, VGII), WM175 (serotype B, VGIII), WM779 (serotype C, VGIV) [15] and the genome-project strain H99 (serotype A, VNI) [89].

Linkage disequilibrium and recombination

Evidence of linkage disequilibrium was tested for using two measures of index of association, I_A [58] and $\bar{r}_d$ [59,90,91]. The significance of the pairwise statistics returned was determined by 1000 randomizations. In the instance of significant clonality or population substructure, both values are expected to be greater than zero, while freely recombining populations would return a score of zero. These tests were also performed on clone corrected samples as recombination may sometimes be masked by clonal reproduction. The proportion of phylogenetically compatible pairs of loci is also reported, with significance estimated with 1000 randomizations [92,93].

The minimum number of recombination events (R_m) was estimated based on the four-gametic test [61], both within individual locus and between loci within described subpopulations.

Genetic variability and testing neutral expectations within individual populations

Comparative sequence analyses were performed in DnaSPv5 [94]. For each locus and each taxon, the number of segregating sites (*S*), haplotypes (*h*) and haplotypic diversity (*Hd*) [64] were calculated. The average number of nucleotide differences between pairs of sequences (π) [64] and the population scaled mutation rate estimated per site (θ) [95] are also reported. Tajima's D [62] and Ramos-Onsins and Rozas' R2 [63] were used to test for departures from the neutral model of molecular evolution, based on the site frequency spectrum. For both tests, significance was obtained from 10000 coalescent simulations.

Genetic differentiation between populations

The average pair-wise number of nucleotide differences per site, D_{xy}, was used to estimate divergence among population groups[64], while K^*_{ST} (a weighted measure of the ratio of the average pair-wise differences within populations to the total average pairwise differences) [65] and S_{nn} (the proportion of nearest neighbours in sequence space found in the same population)[61,66], were used to assess differentiation between the populations. These statistics were also calculated in DNASPv5, with significance levels assessed by 1000 permutations.

Estimates of times of divergence and haplotype networks

A Bayesian Markov Chain Monte Carlo (MCMC) method, implemented in the program BEAST version 1.5.3 [96], was used to estimate the time of divergence between the geographically-defined populations of the global sample of *Cng*, defined as the time to the most recent common ancestor (TMRCA). Sequence indels greater than a single nucleotide long were treated as single evolutionary events in the dataset, and a second partition reflecting these indels created in Beauti v1.5.3 (XML file, dataset S1). The Hasegawa-Kishino-Yano (HKY) model of sequence evolution was assumed, and a relaxed, uncorrelated lognormal molecular clock model applied due to initial runs revealing standard deviation estimates of branch rates to be greater than the mean rate ($\sigma > 1$), indicative of substantial rate heterogeneity among data lineages [96]. Simulations were run for 10^7 with an initial burn-in of 10%. Parameters were logged every 1000 steps over the course of the run. We applied fixed substitution rates, allowing us to convert parameter estimates to calendar years. The rates used were 0.9×10^{-9}, 8.8×10^{-9} and 16.7×10^{-9} mutations per site per year. These are the lower, mean, and upper bounds of a range of substitution rates estimated for *Eurotiomycetes*, based on a calibration date of 400 Myr [67]. Credibility intervals were obtained using 95% highest posterior density (HPD) intervals, the shortest segment that includes 95% of the probability density of the parameter, and the effective sample sizes (ESS) for each parameter, depicted using Tracer v1.5.

Haplotype networks were also created for the STs of the global *Cng* population at each MLST locus. The inference of phylogenetic

relationships among them using statistical parsimony was performed using the program TCS v1.21 [97].

Clinical data and analysis

Clinical data indicative of the progression of cryptococcal infection was available for 58 of the 174 Thai clinical isolates typed in this study. These data were collected previously during a randomized control trial at Sappasitprasong Hospital, Ubon Ratchathani, Thailand. The study aimed to compare the efficacy of four randomly assigned anti-fungal treatment combinations in the initial treatment of HIV-associated CM [68]. Data available included baseline measurements of cerebrospinal fluid (CSF) opening pressure (cm), quantitative cryptococcal CSF culture (CFU/ml CSF), and logarithmic interferon gamma levels. Fungicidal activity was defined by the reduction in CSF cryptococcal colony-forming units (CFU) from quantitative CSF cultures measured at three intervals over the two weeks of treatment. Cerebral dysfunction upon presentation and time to death were also reported [68].

We investigated potential associations between ST and baseline continuous variables using both ANOVA and multivariate ANOVA (MANOVA), with Fisher's exact test being applied to categorical variables. Logistic regression was used to determine factors associated with death by 10 weeks. All analyses were performed using statistical software package R (version 2.6.1).

MLST website eBURST tool

eBURST, a program available at http://eburst.mlst.net/, infers patterns of evolutionary descent among clusters of related genotypes from MLST data. eBURST utilises the MLST site's geographical mapping of MLST data sets (figure S2) to subdivide the STs into related groups of or clonal complexes, as well as to identify the founding genotype (ST) of each group [98].

Accession numbers

All genotypes mentioned within this manuscript are publically available on the MLST database at http://cneoformans.mlst.net/, numbered according to ST as detailed in table S1.

Supporting Information

Table S1 The allelic profiles of the 261 global *Cng* isolates typed at the seven loci as determined by the ISHAM MLST included in this study.
Found at: doi:10.1371/journal.ppat.1001343.s001 (0.55 MB DOC)

Table S2 Diversity indices of the Thai *Cng* population.
Found at: doi:10.1371/journal.ppat.1001343.s002 (0.03 MB DOC)

Table S3 Distribution of nucleotide polymorphisms and insertions within MLST genes IGS1 and SOD1 *Cng* allele types according to the respective position at which it was observed.
Found at: doi:10.1371/journal.ppat.1001343.s003 (0.19 MB DOC)

Table S4 Logistic regression model best describing the prognostic factors of early death (by 10 weeks) among the Thai HIV/AIDS patients.
Found at: doi:10.1371/journal.ppat.1001343.s004 (0.03 MB DOC)

Figure S1 Haplotype networks of the 53 concatenated STs of the global *Cng* population. Sampled haplotypes are indicated by circles or rectangles colored according to the geographical region from which the sample was collected. STs unique to the African population are shown in green and consist only of clinical isolates. Haplotypes found both in Africa and elsewhere are in brown, while those not found in Africa are represented in blue. Rectangles depict the haplotype with the highest ancestral probability. Each branch indicates a single mutational difference and black dots on the lines are representative of the number of mutational steps required to generate allelic polymorphisms. Circle size is proportional to observed haplotype frequency.
Found at: doi:10.1371/journal.ppat.1001343.s005 (0.17 MB PDF)

Figure S2 MLST map of the current global *Cng* isolates. This screenshot of the current distribution of Cng isolates worldwide ($n = 261$) depicted by the MLST website represents the mapping tool utilised in comparative eBURST analysis of Cng populations.
Found at: doi:10.1371/journal.ppat.1001343.s006 (0.33 MB PNG)

Dataset S1 XML file of the current global population of *Cng* assuming a relaxed log-normal clock and a fixed substitution rate of 0.9 x 10-9 per generation.
Found at: doi:10.1371/journal.ppat.1001343.s007 (1.09 MB XML)

Acknowledgments

Isolates were kindly donated by the following people: P. Sriburee, V. Vuthakul and K. Chaicumpar and some sequences made available by A. Litvintseva and the CBS-KNAW. We thank A. Litvintseva for valuable information and input, and T. Jombart for assistance with analyses.

Author Contributions

Conceived and designed the experiments: SPS MCF. Performed the experiments: SPS KK FH AEB. Analyzed the data: SPS DAH CAD. Contributed reagents/materials/analysis tools: SPS DMA TB TSH. Wrote the paper: SPS MCF. MLST website creation: DMA.

References

1. Mitchell TG, Perfect JR (1995) Cryptococcosis in the era of AIDS—100 years after the discovery of *Cryptococcus neoformans*. Clin Microbiol Rev 8: 515–548.
2. King J, Dasgupta A (2005) Cryptococcosis. Updated 30th October, 2009. Available: http://emedicine.medscape.com/article/215354-overview. Accessed 24 April 2010.
3. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, et al. (2009) Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. AIDS 23: 525–530.
4. Banerjee U, Datta K, Majumdar T, Gupta K (2001) Cryptococcosis in India: the awakening of a giant? Med Mycol 39: 51–67.
5. Stevens DA, Denning DW, Shatsky S, Armstrong RW, Adler JD, et al. (1999) Cryptococcal meningitis in the immunocompromised host: intracranial hypertension and other complications. Mycopathologia 146: 1–8.
6. Day J (2004) Cryptococcal meningitis. Pract Neurol 4: 274–285.
7. Schutte CM, Van der Meyden CH, Magazi DS (2000) The impact of HIV on meningitis as seen at a South African Academic Hospital (1994 to 1998). Infection 28: 3–7.
8. Bicanic T, Harrison TS (2004) Cryptococcal meningitis. Br Med Bull 72: 99–118.
9. Franzot SP, Salkin IF, Casadevall A (1999) *Cryptococcus neoformans* var. *grubii*: Separate varietal status for *Cryptococcus neoformans* serotype A isolates. J Clin Microbiol 37: 838–840.
10. Kwon-Chung KJ, Boekhout T, Fell JW, Diaz M (2002) (1557) Proposal to conserve the name *Cryptococcus gattii* against *C. hondurianus* and *C. bacillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae). Taxon 51: 804–806.
11. Bovers M, Hagen F, Kuramae E, Diaz M, Spanjaard L, et al. (2006) Unique hybrids between the fungal pathogens *Cryptococcus neoformans* and *Cryptococcus gattii*. FEMS Yeast Res 6: 599–607.
12. Bovers M, Hagen F, Boekhout T (2008) Diversity of the *Cryptococcus neoformans* (*Cryptococcus gattii*) species. Rev Iberoam Micol 25: S4–12.

13. Bovers M, Hagen F, Kuramae EE, Boekhout T (2008) Six monophyletic lineages identified within *Cryptococcus neoformans* and *Cryptococcus gattii* by multi-locus sequence typing. Fungal Genet Biol 45: 400–421.
14. Boekhout T, Theelen B, Diaz M, Fell JW, Hop WCJ, et al. (2001) Hybrid genotypes in the pathogenic yeast *Cryptococcus neoformans*. Microbiology 147: 891–907.
15. Meyer W, Castaneda A, Jackson S, Huynh M, Castaneda E (2003) Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates. Emerg Infect Dis 9: 189–195.
16. Sukroongreung S, Nilakul C, Ruangsomboon O, Chuakul W, Eampokalap B (1996) Serotypes of *Cryptococcus neoformans* isolated from patients prior to and during the AIDS era in Thailand. Mycopathologia 135: 75–78.
17. Tay ST, Lim HC, Tajuddin TH, Rohani MY, Hamimah H, et al. (2006) Determination of molecular types and genetic heterogeneity of *Cryptococcus neoformans* and *C. gattii* in Malaysia. Med Mycol 44: 617–622.
18. Kwon-Chung KJ, Bennett JE (1978) Distribution of alpha and **a** mating types of *Cryptococcus neoformans* among natural and clinical Isolates. Am J Epidemiol 108: 337–340.
19. Yan Z, Li XG, Xu JP (2002) Geographic distribution of mating type alleles of *Cryptococcusneoformans* in four areas of the United States. J Clin Microbiol 40: 965–972.
20. Halliday CL, Bui T, Krockenberger M, Malik R, Ellis DH, et al. (1999) Presence of alpha and **a** mating types in environmental and clinical collections of *Cryptococcus neoformans* var. *gattii* strains from Australia. J Clin Microbiol 37: 2920–2926.
21. Madrenys N, Devroey C, Raeswuytack C, Torresrodriguez JM (1993) Identification of the perfect state of *Cryptococcus neoformans* from 195 clinical isolates including 84 from AIDS patients.. Mycopathologia 123: 65–68.
22. Barreto de Oliveira MT, Boekhout T, Theelen B, Hagen F, Baroni FA, et al. (2004) *Cryptococcus neoformans* shows a remarkable genotypic diversity in Brazil. J Clin Microbiol 42: 1356–1359.
23. Ohkusu M, Tangonan N, Takeo K, Kishida E, Ohkubo M, et al. (2002) Serotype, mating type and ploidy of *Cryptococcus neoformans* strains isolated from patients in Brazil. Rev Inst Med Trop S Paulo 44: 299–302.
24. Kwon-Chung KJ (1974) Genetics of fungi pathogenic for man. CRC Cr Rev Microbiol 3: 115–133.
25. Padhye AA, Carmichael JW (1969) Mating behavior of *Trichophyton mentagrphytes* varieties paried with *Arthroderma benhamiae* mating types. Sabouraudia 7: 178–181.
26. Padhye AA, Ajello L (1977) Taxonomic status of hedgehog fungus *Trichophyton erinacei*. Sabouraudia 15: 103–114.
27. Kwon-Chung KJ (1975) Perfect state (*Emmonsiella capsulata*) of fungus causing large form African histoplasmosis. Mycologia 67: 980–990.
28. Kwon-Chung KJ, Weeks RJ, Larsh HW (1974) Studies on *Emmonsiella capsulata* (*Histoplasma capsulatum*): II. Distribution of two mating types in 13 endemic states of the United States. Am J Epidemiol 99: 44–49.
29. Randhawa HS, Kowshik T, Khan ZU (2003) Decayed wood of *Syzygium cumini* and *Ficus religiosa* living trees in Delhi/New Delhi metropolitan area as natural habitat of *Cryptococcus neoformans*. Med Mycol 41: 199–209.
30. Nishikawa MM, Lazera MS, Barbosa GG, Trilles L, Balassiano BR, et al. (2003) Serotyping of 467 *Cryptococcus neoformans* isolates from clinical and environmental sources in Brazil: analysis of host and regional patterns. J Clin Microbiol 41: 73–77.
31. Casadevall A, Perfect JR (1998) *Cryptococcus neoformans*. Washington, D.C.: ASM Press.
32. Viviani MA, Esposto MC, Cogliati M, Montagna MT, Wickes BL (2001) Isolation of a *Cryptococcus neoformans* serotype A MAT**a** strain from the Italian environment. Med Mycol 39: 383–386.
33. Kwon-Chung KJ BJ (1992) Mucormycosis. In: Medical Mycology. Philadelphia: Lea & Febiger.
34. Jain N, Wickes BL, Keller SA, Fu J, Casadevall A, et al. (2005) Molecular epidemiology of clinical *Cryptococcus neoformans* strains from India. J Clin Microbiol 43: 5733–5742.
35. Chen J, Varma A, Diaz M, Litvintseva A, Wollenberg K, et al. (2008) *Cryptococcus neoformans* strains and infection in apparently immunocompetent patients, China. Emerg Infect Dis 14: 755–762.
36. Chen S, Sorrell T, Nimmo G, Speed B, Currie B, et al. (2000) Epidemiology and host- and variety-dependent characteristics of infection due to *Cryptococcus neoformans* in Australia and New Zealand. Clin Infect Dis 31: 499–508.
37. Litvintseva AP, Thakur R, Vilgalys R, Mitchell TG (2006) Multilocus sequence typing reveals three genetic subpopulations of *Cryptococcus neoformans* var. *grubii* (Serotype A), including a unique population in Botswana. Genetics 172: 2223–2238.
38. Pitisuttithum P, Tansuphasawadikul S, Simpson AJH, Howe PA, White NJ (2001) A prospective study of AIDS-associated cryptococcal meningitis in Thailand treated with high-dose amphotericin B. J Infection 43: 226–233.
39. Illnait-Zaragozi MT, Martinez-Machin GF, Fernandez-Andreu CM, Boekhout T, Meis JF, et al. (2010) Microsatellite typing of clinical and environmental *Cryptococcus neoformans* var. *grubii* isolates from Cuba shows multiple genetic lineages. Plos One 5(2): e9124. doi:10.1371/journal.pone.0009124.t004.
40. Kidd SE, Hagen F, Tscharke RL, Huynh M, Bartlett KH, et al. (2004) A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). Proc Natl Acad Sci USA 101: 17258–17263.
41. Litvintseva AP, Kestenbaum L, Vilgalys R, Mitchell TG (2005) Comparative analysis of environmental and clinical populations of *Cryptococcus neoformans*. J Clin Microbiol 43: 556–564.
42. Byrnes EJ, Li W, Lewit Y, Ma H, Voelz K, et al. (2010) Emergence and pathogenicity of highly virulent *Cryptococcus gattii* genotypes in the northwest United States. PLoS Pathog 6(4): e1000850. doi:10.1371/journal.ppat.1000850.
43. Meyer W, Marszewska K, Amirmostofian M, Igreja RP, Hardtke C, et al. (1999) Molecular typing of global isolates of *Cryptococcus neoformans* var. *neoformans* by polymerase chain reaction fingerprinting and randomly amplified polymorphic DNA - a pilot study to standardize techniques on which to base a detailed epidemiological survey. Electrophoresis 20: 1790–1799.
44. Meyer W, Aanensen DM, Boekhout T, Cogliati M, Diaz MR, et al. (2009) Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*. Med Mycol 47: 561–570.
45. Wolfe N, Dunavan C, Diamond J (2007) Origins of major human infectious diseases. Nature 447: 279–283.
46. Falush D, Wirth T, Linz B, Pritchard J, Stephens M, et al. (2003) Traces of human migrations in *Helicobacter pylori* populations. Science 299: 1582–1585.
47. Fisher M, Koenig G, White T, San-Blas G, Negroni R, et al. (2001) Biogeographic range expansion into South America by *Coccidioides immitis* mirrors New World patterns of human migration. Proc Natl Acad Sci 98: 4558–4562.
48. Fraser JA, Giles SS, Wenink EC, Geunes-Boyer SG, Wright JR, et al. (2005) Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak. Nature 437: 1360–1364.
49. Johnston R (1992) Birds of North America. Philadelphia: American Ornithologist's Union and Academy of Natural Sciences Of Philadelphia.
50. Lin X, Heitman J (2006) The biology of the *Cryptococcus neoformans* species complex. Annu Rev Microbiol 60: 69–105.
51. Swinne-Desgain D (1976) *Cryptococcus neoformans* in Crops of Pigeons Following Its Experimental Administration. Sabouraudia 14: 313–317.
52. Mooney HA, HRJ, ed (2000) Invasive species in a changing world. Washington DC: Island Press Washington DC.
53. Grzimek BSN, Olendorf D (2004) Grzimek's animal life encyclopedia. Farmington Hills, Michigan: Gale.
54. Pappagianis D, Einstein H (1978) Tempest from Tehachapi takes toll or *Coccidioides* conveyed aloft and afar. West J Med 129: 527–530.
55. Archibald LK, McDonald LC, Rheanpumikankit S, Tansuphaswadikul S, Chaovanich A, et al. (1999) Fever and Human Immunodeficiency Virus infection as sentinels for emerging mycobacterial and fungal bloodstream infections in hospitalized patients >/ = 15 years old, Bangkok. J Infect Dis 180: 87–92.
56. Litvintseva AP, Marra RE, Nielsen K, Heitman J, Vilgalys R, et al. (2003) Evidence of sexual recombination among *Cryptococcus neoformans* serotype A isolates in sub-Saharan Africa. Eukaryot Cell 2: 1162–1168.
57. Ngamskulrungroj P, Gilgado F, Faganello J, Litvintseva AP, Leal AL, et al. (2009) Genetic diversity of the *Cryptococcus* species complex suggests that *Cryptococcus gattii* deserves to have varieties. PLoS ONE 4(6): e5862. doi:10.1371/journal.pone.0005862.
58. Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. Proc Natl Acad Sci 93: 770–773.
59. Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. Mol Ecol Notes 1: 101–102.
60. Bennett RS, Milgroom MG, Bergstrom GC (2005) Population structure of seedborne *Phaeosphaeria nodorum* on New York wheat. Phytopathology 95: 300–305.
61. Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111: 147–164.
62. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.
63. Ramos-Onsins SE, Rozas J (2006) Statistical properties of new neutrality tests against population growth (vol 19, pg 2092, 2002). Mol Biol Evol 23: 1642–1642.
64. Nei M (1987) Molecular Evolutionary Genetics. New York: Columbia University Press.
65. Hudson RR, Boos DD, Kaplan NL (1992) A Statistical test for detecting geographic subdivision. Mol Biol Evol 9: 138–151.
66. Hudson RR (2000) A new statistic for detecting genetic differentiation. Genetics 155: 2011–2014.
67. Kasuga T, White TJ, Taylor JW (2002) Estimation of nucleotide substitution rates in eurotiomycete fungi. Mol Biol Evol 19: 2318–2324.
68. Brouwer AE, Rajanuwong A, Chierakul W, Griffin GE, Larsen RA, et al. (2004) Combination antifungal therapies for HIV-associated cryptococcal meningitis: a randomised trial. Lancet 363: 1764–1767.
69. Wright P, Inverarity D (2007) Human immunodeficiency virus (HIV) related cryptococcal meningitis in rural central Thailand - treatment difficulties and prevention strategies. Southeast Asian J Trop Med Public Health 38: 58–61.
70. McClelland CM, Chang YC, Varma A, Kwon-Chung KJ (2004) Uniqueness of the mating system in *Cryptococcus neoformans*. Trends Microbiol 12: 208–212.
71. Kwon-Chung KJ, Bennett JE (1978) Distribution of alpha and alpha mating types of *Cryptococcus neoformans* among natural and clinical isolates. Am J Epidemiol 108: 337–340.

72. Casali AK, Goulart L, Silva LKR, Silva KRE, Ribeiro AM, et al. (2003) Molecular typing of clinical and environmental *Cryptococcus neoformans* isolates in the Brazilian state Rio Grande do Sul. FEMS Yeast Res 3: 405–415.
73. Hiremath SS, Chowdhary A, Kowshik T, Randhawa HS, Sun S, et al. (2008) Long-distance dispersal and recombination in environmental populations of *Cryptococcus neoformans* var. *grubii* from India. Microbiology 154: 1513–1524.
74. Taylor JW, Geiser DM, Burt A, Koufopanou V (1999) The evolutionary biology and population genetics underlying fungal strain typing. Clin Microbiol Rev 12: 126–146.
75. Buchanan KL, Murphy JW (1998) What makes *Cryptococcus neoformans* a pathogen? Emerg Infect Dis 4: 71–83.
76. Lin XR, Hull CM, Heitman J (2005) Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*. Nature 434: 1017–1021.
77. Bui T, Lin X, Malik R, Heitman J, Carter D (2008) Isolates of *Cryptococcus neoformans* from infected animals reveal genetic exchange in unisexual, alpha mating type populations. Eukaryot Cell 7: 1771–1780.
78. Xu JP, Mitchell TG (2003) Comparative gene genealogical analyses of strains of serotype AD identify recombination in populations of serotypes A and D in the human pathogenic yeast *Cryptococcus neoformans*. Microbiology 149: 2147–2154.
79. Lin XR, Patel S, Litvintseva AP, Floyd A, Mitchell TG, et al. (2009) Diploids in the *Cryptococcus neoformans* serotype A population homozygous for the alpha mating type originate via unisexual mating. Plos Pathogens 5(1): e1000283. doi:10.1371/journal.ppat.1000283.
80. Sriburee P, Khayhan S, Khamwan C, Panjaisee S, Tharavichitkul P (2004) Serotype and PCR-fingerprints of clinical and environmental isolates of *Cryptococcus neoformans* in Chiang Mai, Thailand. Mycopathologia 158: 25–31.
81. Lengeler KB, Cox GM, Heitman J (2001) Serotype AD strains of *Cryptococcus neoformans* are diploid or aneuploid and are heterozygous at the mating-type locus. Infect Immun 69: 115–122.
82. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.
83. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288–295.
84. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. Genetics 131: 479–491.
85. Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403–1405.
86. Doledec S, Chessel D (1987) Seasonal successions and spatial variables in fresh-water environments. 1. Description of a complete 2-way layout by projection of variables. Acta Oecol-Oec Gen 8: 403–426.
87. Saitou N, Nei M (1987) The Neighbor-joining method - a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.
88. Felsenstein J (1985) Confidence-limits on phylogenies - an approach using the bootstrap. Evolution 39: 783–791.
89. Perfect JR, Ketabchi N, Cox GM, Ingram CW, Beiser CL (1993) Karyotyping of *Cryptococcus neoformans* as an epidemiological tool. J Clin Microbiol 31: 3305–3309.
90. Brown AHD, Feldman MW, Nevo E (1980) Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics 96: 523–536.
91. Smith JM, Smith NH, Orourke M, Spratt BG (1993) How clonal are bacteria. Proc Natl Acad Sci 90: 4384–4388.
92. Estabrook GF, Landrum L (1975) A simple test for the possible simultaneous evolutionary divergence of two amino acid positions. Taxon 24: 609–613.
93. Xu JP, Yan Z, Guo H (2009) Divergence, hybridization, and recombination in the mitochondrial genome of the human pathogenic yeast *Cryptococcus gattii*. Mol Ecol 18: 2628–2642.
94. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.
95. Watterson GA (1975) Number of segregating sites in genetic models without recombination. Theor Popul Biol 7: 256–276.
96. Drummond AJ HS, Rawlence N, Rambaut A (2007) A rough guide to BEAST 1.4. Available: http://beast.bio.ed.ac.uk/Main_Page. Accessed 11 November 2009.
97. Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9: 1657–1659.
98. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG (2004) eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol 186: 1518–1530.
99. Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. Mol Biol Evol 19: 2092–2100.

B. Additional tables and figures

Table B.1.: Summary of all isolates strains used in this study ($n = 370$).

Isolate	ST	Origin	Source	Year	Mating type	VN type
A3-38-20	1	N. Carolina	Pigeon excreta	2002[a]	α	VNI
R16	1	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R24	1	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R49	1	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
A1-38-2	2	N. Carolina	Pigeon excreta	2002[a]	α	VNI
A1-35-8	2	N. Carolina	Pigeon excreta	2002[a]	α	VNI
arg1373	2	Argentina			α	VNI
arg1366	2	Argentina			α	VNI
c23	2	N. Carolina	BAL/HIV−	2001[a]	α	VNI
CA84-14	2	California	Pigeon excreta	2003[a]	α	VNI
H99	2	N. Carolina		[c]	α	VNI
mal-120	2	Malawi	Blood/AIDS	1999[d]	α	VNI
Tn470	2	Tanzania	Blood/HIV+	1995[e]	α	VNI
R46	2	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
125.91	3	Tanzania		[f]	**a**	VNI
R57	3	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
th84	4	Thailand	Blood/HIV+	1997[g]	α	VNI
ug2463	4	Uganda	CSF/HIV+	2001[h]	α	VNI
CN5010	4	Thailand	Blood	[i]	α	VNI
CN4998	4	Thailand	CSF	[i]	α	VNI
CN4995	4	Thailand	CSF	[i]	α	VNI
CN4989	4	Thailand	CSF	[i]	α	VNI
CN4988	4	Thailand	CSF	[i]	α	VNI
CN4987	4	Thailand	CSF	[i]	α	VNI
CN4964	4	Thailand	CSF	[i]	α	VNI
CN4947	4	Thailand	CSF	[i]	α	VNI
CN4945	4	Thailand	CSF	[i]	α	VNI
CN4944	4	Thailand	CSF	[i]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001; [e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999; [h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004; [k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University); [n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003; [r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi); [t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CN4943	4	Thailand	CSF	[i]	α	VNI
CN4942	4	Thailand	CSF	[i]	α	VNI
CN4941	4	Thailand	CSF	[i]	α	VNI
CN4940	4	Thailand	CSF	[i]	α	VNI
CN4926	4	Thailand	CSF	[i]	α	VNI
CN4919	4	Thailand	CSF	[i]	α	VNI
CN4918	4	Thailand	CSF	[i]	α	VNI
CN4917	4	Thailand	CSF	[i]	α	VNI
CN4903	4	Thailand	CSF	[i]	α	VNI
CN4901	4	Thailand	CSF	[i]	α	VNI
CN49005	4	Thailand	CSF	[i]	α	VNI
4-187	4	Thailand	Clinical	[i]	α	VNI
269	4	Thailand	Clinical	[i]	α	VNI
4-315	4	Thailand	Clinical	[i]	α	VNI
1-587	4	Thailand	Clinical	[i]	α	VNI
1219	4	Thailand	Clinical	[i]	α	VNI
4_83	4	Thailand	Clinical	[i]	α	VNI
1-588	4	Thailand	Clinical	[i]	α	VNI
4-202	4	Thailand	Clinical	[i]	α	VNI
1-846	4	Thailand	Clinical	[i]	α	VNI
2551-07	4	Thailand	CSF	[i]	α	VNI
2550 II-07	4	Thailand	Blood	[i]	α	VNI
2461-07	4	Thailand	CSF	[i]	α	VNI
CM 1	4	Thailand	CSF	2002[j]	α	VNI
CM 6	4	Thailand	CSF	2002[j]	α	VNI
CM 7	4	Thailand	CSF	2002[j]	α	VNI
CM 8	4	Thailand	CSF	2002[j]	α	VNI
CM 12	4	Thailand	CSF	2002[j]	α	VNI
CM 13	4	Thailand	CSF	2002[j]	α	VNI
CM 17	4	Thailand	CSF	2002[j]	α	VNI
CM 18	4	Thailand	CSF	2002[j]	α	VNI
CM 22	4	Thailand	CSF	2002[j]	α	VNI
CM 23	4	Thailand	CSF	2002[j]	α	VNI
CM 25	4	Thailand	CSF	2002[j]	α	VNI
CM 26	4	Thailand	CSF	2002[j]	α	VNI
CM 33	4	Thailand	CSF	2002[j]	α	VNI
CM 37	4	Thailand	CSF	2002[j]	α	VNI
CM 38	4	Thailand	CSF	2002[j]	α	VNI
CM 39	4	Thailand	CSF	2002[j]	α	VNI
CM 40	4	Thailand	CSF	2002[j]	α	VNI
CM 41	4	Thailand	CSF	2002[j]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001; [e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999; [h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004; [k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University); [n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003; [r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi); [t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CM42	4	Thailand	CSF	2002[j]	α	VNI
CM 43	4	Thailand	CSF	2002[j]	α	VNI
CM 44	4	Thailand	CSF	2002[j]	α	VNI
CM 46	4	Thailand	CSF	2002[j]	α	VNI
CM 47	4	Thailand	CSF	2002[j]	α	VNI
CM 48	4	Thailand	CSF	2002[j]	α	VNI
CM 49	4	Thailand	CSF	2002[j]	α	VNI
CM 51	4	Thailand	CSF	2002[j]	α	VNI
CM 55	4	Thailand	CSF	2002[j]	α	VNI
CM 56	4	Thailand	CSF	2002[j]	α	VNI
CM 57	4	Thailand	CSF	2002[j]	α	VNI
CM 58	4	Thailand	CSF	2002[j]	α	VNI
CM 59	4	Thailand	CSF	2002[j]	α	VNI
CM 61	4	Thailand	CSF	2002[j]	α	VNI
CM 63	4	Thailand	CSF	2002[j]	α	VNI
K 2	4	Thailand	Crypto patient	2002[k]	α	VNI
Pg 1	4	Thailand	Pigeon dropping	2002[k]	α	VNI
D 6	4	Thailand	Dove dropping	2002[k]	α	VNI
D 1	4	Thailand	Dove dropping	2002[k]	α	VNI
A5-35-17	5	N. Carolina	Pigeon excreta	[a]	α	VNI
c48	5	N. Carolina	BAL/HIV−	2001[a]	α	VNI
c8	5	N. Carolina	CSF/HIV+	[a]	α	VNI
it743	5	Italy		[l]	α	VNI
jp1086	5	Japan	Human lung	1999[m]	α	VNI
jp1088	5	Japan	Human lung	1999[m]	α	VNI
CN49004	5	Thailand	CSF	[i]	α	VNI
CN48	5	Thailand	Clinical	[i]	α	VNI
1-488	5	Thailand	Clinical	[i]	α	VNI
1-489	5	Thailand	Clinical	[i]	α	VNI
CM 30	5	Thailand	CSF	2002[j]	α	VNI
J 1	5	Japan	Crypto patient	2002[k]	α	VNI
Pt 12	5	Thailand	Crypto patient	2002[k]	α	VNI
D 5	5	Thailand	Dove dropping	2002[k]	α	VNI
Pg 37	5	Thailand	Pigeon dropping	2002[k]	α	VNI
CN5015	5	Thailand	CSF	[i]	α	VNI
CN5018	5	Thailand	Blood	[i]	α	VNI
CN5012	5	Thailand	CSF	[i]	α	VNI
CN5008	5	Thailand	CSF	[i]	α	VNI
CN4993	5	Thailand	CSF	[i]	α	VNI
CN4983	5	Thailand	CSF	[i]	α	VNI
CN4980	5	Thailand	CSF	[i]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001; [e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999; [h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004; [k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University); [n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003; [r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi); [t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CN4977	5	Thailand	CSF	[i]	α	VNI
CN4967	5	Thailand	CSF	[i]	α	VNI
CN4960	5	Thailand	CSF	[i]	α	VNI
CN4948	5	Thailand	CSF	[i]	α	VNI
CN4946	5	Thailand	CSF	[i]	α	VNI
CN4924	5	Thailand	CSF	[i]	α	VNI
CN4921	5	Thailand	CSF	[i]	α	VNI
CN4920	5	Thailand	CSF	[i]	α	VNI
CN4916	5	Thailand	CSF	[i]	α	VNI
CN4906	5	Thailand	CSF	[i]	α	VNI
CN49006	5	Thailand	CSF	[i]	α	VNI
CT1	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT13	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT18	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT3	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT30	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT33	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT34	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT44	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT47	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT7	5	S. Africa	CSF/HIV+	2005[n]	α	VNI
R29	5	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R4	5	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R5	5	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R64	5	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
th104	6	Thailand	Blood/HIV+	1997[g]	α	VNI
th206	6	Thailand	Blood/HIV+	1997[e]	α	VNI
Tn148	6	Tanzania	Blood/HIV+	1995[e]	α	VNI
CN5019	6	Thailand	Blood	[i]	α	VNI
CN5017	6	Thailand	CSF	[i]	α	VNI
CN5014	6	Thailand	Blood	[i]	α	VNI
CN5013	6	Thailand	CSF	[i]	α	VNI
CN5011	6	Thailand	Clinical	[i]	α	VNI
CN5009	6	Thailand	Blood	[i]	α	VNI
CN5005	6	Thailand	Blood	[i]	α	VNI
CN5003	6	Thailand	Blood	[i]	α	VNI
CN5002	6	Thailand	Blood	[i]	α	VNI
CN5001	6	Thailand	CSF	[i]	α	VNI
CN4970	6	Thailand	CSF	[i]	α	VNI
CN4968	6	Thailand	CSF	[i]	α	VNI
CN4957	6	Thailand	CSF	[i]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001;
[e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999;
[h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004;
[k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University);
[n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003;
[r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi);
[t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CN4956	6	Thailand	CSF	[i]	α	VNI
CN4955	6	Thailand	BAL	[i]	α	VNI
CN4954	6	Thailand	CSF	[i]	α	VNI
CN4952	6	Thailand	CSF	[i]	α	VNI
CN4950	6	Thailand	CSF	[i]	α	VNI
CN4949	6	Thailand	CSF	[i]	α	VNI
CN4938	6	Thailand	CSF	[i]	α	VNI
CN4937	6	Thailand	CSF	[i]	α	VNI
CN4936	6	Thailand	CSF	[i]	α	VNI
CN4934	6	Thailand	CSF	[i]	α	VNI
CN4933	6	Thailand	CSF	[i]	α	VNI
CN4932	6	Thailand	CSF	[i]	α	VNI
CN4931	6	Thailand	CSF	[i]	α	VNI
CN4927	6	Thailand	CSF	[i]	α	VNI
CN4915	6	Thailand	CSF	[i]	α	VNI
CN4914	6	Thailand	CSF	[i]	α	VNI
CN4909	6	Thailand	CSF	[i]	α	VNI
CN4907	6	Thailand	CSF	[i]	α	VNI
CN4905	6	Thailand	CSF	[i]	α	VNI
CN4904	6	Thailand	CSF	[i]	α	VNI
CN4902	6	Thailand	CSF	[i]	α	VNI
CN49008	6	Thailand	CSF	[i]	α	VNI
4-319	6	Thailand	Clinical	[i]	α	VNI
50NC2	6	Thailand	Clinical	[i]	α	VNI
50NC5	6	Thailand	Clinical	[i]	α	VNI
11112	6	Thailand	Clinical	[i]	α	VNI
11109	6	Thailand	Clinical	[i]	α	VNI
4-231	6	Thailand	Clinical	[i]	α	VNI
P6	6	Thailand	Clinical	[i]	α	VNI
4-253	6	Thailand	Clinical	[i]	α	VNI
4-381	6	Thailand	Clinical	[i]	α	VNI
20662-07	6	Thailand	Blood	[i]	α	VNI
28170-07	6	Thailand	CSF	[i]	α	VNI
1111I-08	6	Thailand	Blood/HIV−	[i]	α	VNI
2895I-08	6	Thailand	Blood/HIV−	[i]	α	VNI
4500-07	6	Thailand	Blood	[i]	α	VNI
CM 2	6	Thailand	CSF	2002[j]	α	VNI
CM 3	6	Thailand	CSF	2002[j]	α	VNI
CM 4	6	Thailand	CSF	2002[j]	α	VNI
CM 5	6	Thailand	CSF	2002[j]	α	VNI
CM 10	6	Thailand	CSF	2002[j]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001; [e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999; [h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004; [k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University); [n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003; [r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi); [t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CM 14	6	Thailand	CSF	2002[j]	α	VNI
CM 11	6	Thailand	CSF	2002[j]	α	VNI
CM 15	6	Thailand	CSF	2002[j]	α	VNI
CM16	6	Thailand	CSF	2002[j]	α	VNI
CM 20	6	Thailand	CSF	2002[j]	α	VNI
CM 24	6	Thailand	CSF	2002[j]	α	VNI
CM 27	6	Thailand	CSF	2002[j]	α	VNI
CM 28	6	Thailand	CSF	2002[j]	α	VNI
CM 29	6	Thailand	CSF	2002[j]	α	VNI
CM 32	6	Thailand	CSF	2002[j]	α	VNI
CM 34	6	Thailand	CSF	2002[j]	α	VNI
CM 36	6	Thailand	CSF	2002[j]	α	VNI
CM 45	6	Thailand	CSF	2002[j]	α	VNI
CM 50	6	Thailand	CSF	2002[j]	α	VNI
CM 52	6	Thailand	CSF	2002[j]	α	VNI
CM 60	6	Thailand	CSF	2002[j]	α	VNI
CM 64	6	Thailand	CSF	2002[j]	α	VNI
Pt 9	6	Thailand	Crypto patient	2002[k]	α	VNI
Pt 3	6	Thailand	Crypto patient	2002[k]	α	VNI
Pt 1	6	Thailand	Crypto patient	2002[k]	α	VNI
D 2	6	Thailand	Dove dropping	2002[k]	α	VNI
D 3	6	Thailand	Dove dropping	2002[k]	α	VNI
Pg 2	6	Thailand	Pigeon dropping	2002[k]	α	VNI
Pg 26	6	Thailand	Pigeon dropping	2002[k]	α	VNI
R6	6	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
bt1	7	Botswana	CSF/AIDS	1999[o]	α	VNB
bt31	8	Botswana	CSF/AIDS	2000[o]	α	VNB
bt33	9	Botswana	CSF/AIDS	2000[o]	α	VNB
bt109	9	Botswana	CSF/AIDS	2001[o]	α	VNB
bt65	10	Botswana	CSF/AIDS	2000[o]	**a**	VNB
bt76	11	Botswana	CSF/AIDS	2000[o]	α	VNB
bt131	12	Botswana	CSF/AIDS	2001[o]	**a**	VNB
bt130	13	Botswana	CSF/AIDS	2001[o]	**a**	VNI
bt68	13	Botswana	CSF/AIDS	2000[o]	α	VNI
bt24	14	Botswana	CSF/AIDS	2000[o]	**a**	VNB
A2-102-5	15	Texas	Pigeon excreta	2003[a]	α	VNI
A2-38-23	15	N. Carolina	Pigeon excreta	2003[a]	α	VNI
bt206	16	Botswana	CSF/AIDS	2002[o]	**a**	VNB
bt89	16	Botswana	CSF/AIDS	2001[o]	α	VNB
bt84	17	Botswana	CSF/AIDS	2001[o]	α	VNB
bt35	18	Botswana	CSF/AIDS	2000[o]	α	VNB

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001; [e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999; [h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004; [k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University); [n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003; [r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi); [t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
bt46	19	Botswana	CSF/AIDS	2000[o]	α	VNB
bt63	20	Botswana	CSF/AIDS	2000[o]	**a**	VNB
bt150	21	Botswana	CSF/AIDS	2001[o]	α	VNI
bt104	22	Botswana	CSF/AIDS	2001[o]	α	VNI
A3-1-1	23	N. Carolina	Pigeon excreta	2002[a]	α	VNI
it754	23	Italy		[p]	α	VNI
CT54	23	S. Africa	CSF/HIV+	2005[n]	α	VNI
R10	23	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R12	23	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R23	23	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R51	23	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R58	23	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
bt100	24	Botswana	CSF/AIDS	2001[o]	α	VNI
bt134	25	Botswana	CSF/AIDS	2001[o]	α	VNI
bt15	26	Botswana	CSF/AIDS	2000[o]	α	VNI
bt85	27	Botswana	CSF/AIDS	2001[o]	**a**	VNB
bt88	28	Botswana	CSF/AIDS	2001[o]	**a**	VNB
bt204	29	Botswana	CSF/AIDS	2002[o]	**a**	VNB
c12	30	N. Carolina	Lung/HIV−	[a]	α	VNII
c16	30	N. Carolina	Sputum/HIV−	2001[a]	α	VNII
8-1	30	N. Carolina		[q]	α	VNII
R59	30	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
c26	31	N. Carolina	Blood/HIV+	2001[a]	α	VNI
br2362	32	Brazil		[r]	α	VNI
br794	32	Brazil		1998[r]	α	VNI
br795	32	Brazil		1998[r]	α	VNI
c27	32	N. Carolina	CSF/cancer	2001[a]	α	VNI
in2629	32	India	CSF/AIDS	2001[s]	α	VNI
in2632	32	India	CSF/AIDS	2001[s]	α	VNI
Tn10	32	Tanzania	Blood/HIV+	[e]	α	VNI
ug2458	32	Uganda	CSF/HIV+	2001[h]	α	VNI
za1345	32	DRC (Zaire)		[t]	α	VNI
za1346	32	DRC (Zaire)	CSF/HIV+	[t]	α	VNI
CN5007	32	Thailand	CSF	[i]	α	VNI
1291-09	32	Thailand	Blood/HIV−	[i]	α	VNI
CM 35	32	Thailand	CSF	2002[j]	α	VNI
CT16	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT21	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT22	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT23	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT24	32	S. Africa	CSF/HIV+	2005[n]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001;
[e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999;
[h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004;
[k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University);
[n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003;
[r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi);
[t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CT36	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT37	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT43	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT6	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT8	32	S. Africa	CSF/HIV+	2005[n]	α	VNI
R38	32	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R43	32	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R47	32	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R52	32	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
bt60	33	Botswana	CSF/AIDS	2000[o]	α	VNB
bt121	34	Botswana	CSF/AIDS	2001[o]	α	VNI
bt27	35	Botswana	CSF/AIDS	2000[o]	α	VNB
bt9	36	Botswana	CSF/AIDS	1999[o]	α	VNI
A4-1-12	37	N. Carolina	Pigeon excreta	2002[a]	α	VNI
CT2	37	S. Africa	CSF/HIV+	2005[n]	α	VNI
R56	37	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
in2637	38	India	CSF/HIV+	[s]	α	VNI
mal-9	38	Malawi	Blood/AIDS	[d]	α	VNI
ug2471	38	Uganda	CSF/HIV+	2001[h]	α	VNI
A4-34-6	39	N. Carolina	Pigeon excreta	2003[a]	α	VNI
ug2472	40	Uganda	CSF/HIV+	2001[h]	α	VNII
c45	41	N. Carolina	Sputum/HIV−	2001[a]	α	VNII
c2	42	N. Carolina	BAL/HIV−	2002[a]	α	VNII
c44	42	N. Carolina	CSF/HIV−	2002[a]	α	VNII
A7-35-23	43	N. Carolina	Pigeon excreta	[a]	α	VNII
ug2467	44	Uganda	CSF/HIV+	2001[h]	α	VNI
CM 21	45	Thailand	CSF	2002[j]	α	VNII
D 9	46	Thailand	Dove dropping	2002[k]	α	VNI
K 45	47	Thailand	Crypto patient	2002[k]	α	VNI
Pt 5	48	Thailand	Crypto patient	2002[k]	α	VNI
4_9	49	Thailand	Clinical	[i]	α	VNI
50NC1	50	Thailand	Clinical	[i]	α	VNI
CT25	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT26	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT29	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT35	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT39	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT4	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT45	51	S. Africa	CSF/HIV+	2005[n]	α	VNI
R2	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R25	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001;
[e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999;
[h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004;
[k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University);
[n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003;
[r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi);
[t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
R3	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R31	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R32	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R33	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R35	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R41	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R45	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R53	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R55	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R62	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R7	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R8	51	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
CT17	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT20	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT31	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT32	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT38	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT42	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT52	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT53	52	S. Africa	CSF/HIV+	2005[n]	α	VNII
R11	52	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R20	52	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R39	52	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R9	52	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R13	53	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R15	53	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R27	53	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R17	54	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R44	54	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R61	54	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R18	55	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R19	55	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R63	55	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R21	56	S. Africa	CSF/HIV+	2005/06[b]	α	VNB
R26	56	S. Africa	CSF/HIV+	2005/06[b]	α	VNB
CT10	57	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT12	58	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT14	59	S. Africa	CSF/HIV+	2005[n]	α	VNB
CT15	60	S. Africa	CSF/HIV+	2005[n]	**a**	VNB
CT19	61	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT27	62	S. Africa	CSF/HIV+	2005[n]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001;
[e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999;
[h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004;
[k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University);
[n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003;
[r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi);
[t]N. Myers (CDC).

(continued on next page)

Isolate	ST	Origin	Source	Year	Mating type	VN type
CT40	63	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT41	64	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT46	65	S. Africa	CSF/HIV+	2005[n]	α	VNI
CT48	66	S. Africa	CSF/HIV+	2005[n]	α	VNII
CT50	67	S. Africa	CSF/HIV+	2005[n]	α	VNII
R1	68	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R22	69	S. Africa	CSF/HIV+	2005/06[b]	α	VNB
R30	70	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R34	71	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R40	72	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R48	73	S. Africa	CSF/HIV+	2005/06[b]	α	VNII
R50	74	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R54	75	S. Africa	CSF/HIV+	2005/06[b]	α	VNI
R60	76	S. Africa	CSF/HIV+	2005/06[b]	α	VNI

[a]Litvintseva et al., 2005; [b]Bicanic et al., 2008; [c]Perfectet al., 1980; [d]Bellet al.,2001;
[e]Archibaldet al., 1998; [f]Lengeleret al., 2000; [g]Archibaldet al., 1999;
[h]S. A. Messer (University of Iowa); [i]Simwami et al., 2011; [j]Brouwer et al., 2004;
[k]Sriburee et al., 2004; [l]Barchiesiet al.,1997; [m]Shigefumi Maesaki (Nagasaki University);
[n]Bicanic et al., 2007; [o]Litvintseva et al., 2003; [p]Barchiesiet al., 1997; [q]Nielsenet al., 2003;
[r]M. G. De Almeida (Universidad de Sao Paulo); [s]H. C. Gugnani (University of Delhi);
[t]N. Myers (CDC).

Table B.2.: Allelic profiles of the 107 Cape Town *Cng* isolates typed in this study and their corresponding Sequence Types.

ST	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*
1	7	1	1	1	1	1	1
2	7	1	1	7	1	1	2
3	1	1	1	3	4	1	1
4	1	1	1	3	2	1	5
5	1	3	1	5	2	1	1
6	1	1	1	4	2	1	5
7	9	11	8	6	10	7	6
8	4	11	9	6	7	1	12
9	4	11	7	6	4	1	3
10	9	11	2	6	4	1	3
11	5	11	7	6	6	7	10
12	3	12	8	6	6	6	8
13	1	5	1	1	4	1	1
14	6	8	21	6	9	8	8
15	1	2	6	3	2	1	1
16	4	11	2	6	4	7	3
17	5	11	16	6	6	5	3
18	4	11	8	6	6	6	3

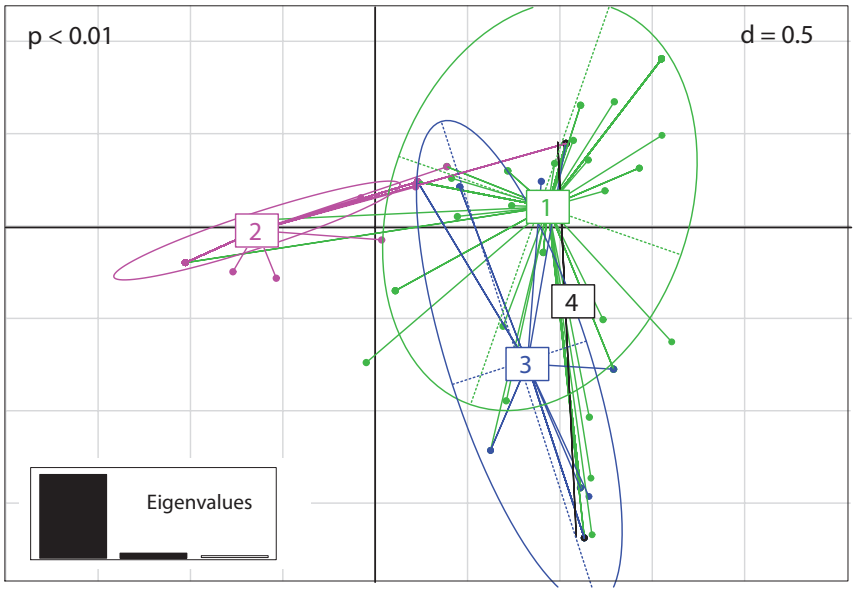
(continued on next page)

ST	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*
19	11	11	2	6	6	7	6
20	6	8	13	6	8	8	7
21	7	5	22	1	3	4	9
22	1	6	1	1	4	1	1
23	7	1	1	2	1	1	2
24	7	5	1	1	5	2	9
25	1	3	23	5	2	1	1
26	1	6	19	1	4	1	1
27	12	8	17	6	9	9	8
28	5	11	7	6	6	1	10
29	6	13	12	6	9	8	13
30	2	9	14	8	11	1	4
31	7	1	1	2	1	3	2
32	1	1	10	3	4	1	1
33	4	8	24	6	9	9	7
34	1	5	18	1	4	1	1
35	5	11	8	6	4	14	6
36	1	7	3	1	4	1	9
37	1	1	1	1	1	1	2
38	1	1	10	3	2	1	1
39	1	1	11	3	4	1	1
40	2	9	20	8	11	1	4
41	10	9	14	8	11	1	4
42	8	10	15	8	12	1	11
43	2	9	14	8	11	11	4
44	1	1	1	2	2	1	2
45	2	10	14	6	11	11	4
46	1	1	1	4	2	1	14
47	1	1	1	3	4	1	5
48	1	1	1	5	2	1	1
49	1	1	1	9	2	1	5
50	1	3	1	10	2	1	1
51	7	5	1	3	3	1	1
52	2	9	14	8	11	12	4
53	7	5	1	3	3	1	2
54	1	1	10	3	4	1	2
55	1	3	1	5	2	1	2
56	4	11	26	6	6	1	3
57	1	1	10	1	2	1	1
58	7	1	1	3	1	1	1
59	1	14	10	11	4	1	15
60	13	1	25	11	4	19	1
61	1	1	1	1	4	17	2
62	1	1	1	7	1	1	1
63	1	7	1	1	4	3	9
64	1	1	10	3	4	12	1
65	7	1	1	7	4	1	2
66	9	9	14	6	11	16	16

(continued on next page)

ST	*CAP59*	*GPD1*	*IGS1*	*LAC1*	*PLB1*	*SOD1*	*URA5*
67	8	10	28	8	2	3	11
68	2	1	14	8	11	12	4
69	9	11	27	6	6	15	2
70	8	10	1	8	12	3	2
71	2	9	14	8	11	12	2
72	10	9	14	8	11	12	4
73	8	10	28	8	12	3	11
74	7	1	1	1	1	1	2
75	7	1	10	1	1	1	1
76	7	1	1	7	1	18	2

Figure B.1.: Principle Components Analysis of the allelic profiles of the global VNI *Cng* genotypes used in this study ($n = 315$), according to geographic origin. Individual genotypes (dots) are linked by coloured lines to form clusters which are summarised by coloured ellipses proportional in size to the number of isolates represented. The four groups depicted are numbered and defined according to continent: Africa = group 1 (green ellipse, $n = 102$), Asia = group 2 (pink ellipse, $n = 191$), North America = group 3 (blue ellipse, $n = 17$) and South America = group 4 (black line, $n = 5$). Europe was excluded due to small sample size. p-value showing inter-continental differentiation is shown and eigenvalues are represented in the bar plot.



C. R Code

C.1. PCA commands

```
install.packages("adegenet", dep=TRUE)
library(adegenet)
library(ade4)
SEA<-read.delim("clipboard", head=TRUE)
attach(SEA)
obj<-df2genind(X=SEA[,3:9], pop=Pop, ploidy=1)
# PCA
pca1<-dudi.pca(obj$tab, cent=TRUE, scale=FALSE, scannf=FALSE, nf=3)
# Pot
s.class(pca1$li, obj$pop, lab=obj$pop.names)
# Add eigenvalues
add.scatter.eig(pca1$eig, nf=2, xax=1, yax=2, pos="bottomright", csub=1.2)
# Monte-Carlo test
d1 <- randtest(between(pca1, fac=obj$pop, scan = FALSE), 99)
 rand1
```

C.2. Inter-class PCA commands

```
global<-read.delim("clipboard", head=TRUE)
attach(global)
obj<-df2genind(X=global[,3:9], ploidy=1, pop=Pop)
# Interclass PCA maximises the variance between populations
# instead of the total variance
bet1<-between(pca1, fac=obj$pop, scannf=FALSE, nf=3)
# Plot
```

```
 s.class(bet1$ls, fac=obj$pop, clab=1.2, lab=obj$pop.names)
# Add eigenvalues
 add.scatter.eig(bet1$eig, nf=2, xax=1, yax=2, pos="bottomright")
rand1 <- randtest(between(pca1, fac=obj$pop, scan = FALSE), 99)
rand1
```

C.3. Clinical analysis

```
# Variable manipulation
ST<- as.factor(ST)

##Summary statistics of patient characteristics.
summary(fulldata$logCRAG)
summary(fulldata$age)
summary(fulldata$weight)
etc.

# Summary statistics of continuous variables stratified according to ST.
summary((fulldata$logCRAG[fulldata$ST=="4"]), na.rm=TRUE)
summary((fulldata$logCRAG[fulldata$ST=="5"]), na.rm=TRUE)
summary((fulldata$logCRAG[fulldata$ST=="6"]), na.rm=TRUE)
etc.

# Summary statistics of categorical variables stratified according to ST.
table(fulldata$died2w,fulldata$ST)
table(fulldata$sex,fulldata$ST)
table(fulldata$altment,fulldata$ST)
etc.

##Univariable analysis
# ANOVA - associations between ST and continuous variables.
# Repeated for explanatory variables "continent", "died2wk", "died10wk" and "EFA".
summary(aov(logCRAG~ST,data=fulldata))
summary(aov(age~ST,data=fulldata))
```

```
summary(aov(weight~ST,data=fulldata))
etc.

# Main effects plot of significant ST and continuous variables
par(mfrow=c(2,2))
plot.design(fulldata[c("logCRAG","ST")])
plot.design(fulldata[c("logQC","ST")])
plot.design(fulldata[c("CD4","ST")])
plot.design(fulldata[c("logIFN","ST")])

# ANOVA - associations between loci and continuous variables
# (CAP59, GPD1, IGS1, LAC1, PLB1, SOD1, URA5).
summary(aov(logCRAG~CAP59,data=fulldata))
summary(aov(age~CAP59,data=fulldata))
summary(aov(weight~CAP59,data=fulldata))
etc.

# Fisher's exact test - associations between ST and categorical variables.
# Repeated for explanatory variables "continent", "died2wk", "died10wk".
fisher.test(fulldata$ST,fulldata$sex)
fisher.test(fulldata$ST,fulldata$altment)
fisher.test(fulldata$ST,fulldata$ART)

# ANOVA - associations between EFA and categorical variables.
summary(aov(EFA~treat,data=fulldata))
summary(aov(EFA~sex,data=fulldata))
summary(aov(EFA~altment,data=fulldata))
summary(aov(EFA~ART,data=fulldata))

## Multivariable analysis
 # Final model for death by week 2 (death2w)
summary(glm(formula= died2w ~weight + altment, data=fulldata, family=binomial))

# Final model for death by week 10 (died10w)
summary(glm(formula= died10w ~ Opbase +  altment,  data=fulldata, family=binomial))
```

```
# Final model for EFA
summary(lm(slope ~  weight + treat + logCRAG + altment +  CSFWBC + logIFN,
weights= num_obs, data=fulldata))

# Final model for EFA with PLB1
summary(lm(slope ~  weight + treat + logCRAG + altment + PLB1  +  CSFWBC + logIFN,
weights= num_obs, data=fulldata))

# Final model for EFA with IGS1
summary(lm(slope ~  weight + treat + logCRAG + altment + IGS1  +  CSFWBC + logIFN,
weights= num_obs, data=fulldata))

# Survival analysis
install.packages("survival")
install.packages("survival")
install.packages("KMsurv") # for Kaplan-Meier estimator
library(foreign)
library(survival)

# Create survival object.
surv.object <- Surv(fulldata$FU, fulldata$FO)
# Survival analysis
 fit1 <- survfit(Surv(fulldata$FU, fulldata$FO) ~ 1, data = fulldata)
# Plot.
 plot(fit1)

# Survival analysis according to categorical variables, plot and egend.
fit.altment <- survfit(Surv(fulldata$FU, fulldata$FO) ~ altment, data = fulldata)
plot(fit.altment, lw=2, cex=2, col= c("blue" , "red", "blue", "orange", "green",
"black", "yellow"), lty= 1:2,xlab="Day", ylab="Survival probability")
legend("bottomleft",lw=2, 1.0, c("positive", "negative"), col= c("blue" , "red"),
lty= 1:2)


# Continuous variables divided into quartiles, followed by survival analysis and plot.
quartiles <- gtools::quantcut (fulldata$logIFN)
```

```
table(quartiles)
fit.logIFN <- survfit(Surv(fulldata$FU, fulldata$FO) ~ quartiles, data = fulldata)
plot(fit.logIFN, col= c("red" , "grey", "blue", "orange"), lty= 1:2)

# Log rank test of differences between groups
survdiff(Surv(fulldata$FU, fulldata$FO) ~ ST, data = fulldata)

# Survival analysis according to ST
fit.all <- survfit(Surv(FU, FO) ~ ST_new, data = fulldata)
plot(fit.all, col=c(1:4, "orange", "darkgrey"), lwd=2, cex=2, xlab="day",
ylab="survival probability")
legend("bottomleft", legend = c("ST5", "ST6", "ST32", "ST4", "ST51", "ST52"),
col= c(1:4, "orange", "darkgrey"), lwd=2, bty = "n")

# Survival analysis according to ST and continents Africa and Asia
fit.africa <- survfit(Surv(FU, FO) ~ ST_new,
data = fulldata[fulldata$continent == "Africa",])
plot(fit.africa, col=c(1:3, "orange", "darkgrey" ), lwd=2, cex=2, xlab="day",
ylab="survival probability")
legend("bottomleft", legend = c("ST5", "ST6", "ST32", "ST51", "ST52"),
col= c(1:3, "orange", "darkgrey"), lwd=2, bty = "n")

fit.asia <- survfit(Surv(FU, FO) ~ ST_new,
data = fulldata[fulldata$continent == "Asia",])
plot(fit.asia, col=c(1:4), lwd=2, cex=2, xlab="day", ylab="survival probability")
legend("bottomleft", legend = c("ST5", "ST6", "ST32", "ST4"),col= c(1:4),
lwd=2, bty = "n")
```

C.4. *Galleria mellonella* assay anaylsis

```
# Survival analysis
install.packages("survival")
install.packages("survival")
install.packages("KMsurv") # for Kaplan-Meier estimator
library(foreign)
```

```
library(survival)

GM<-read.delim("clipboard", head=TRUE)
attach(GM)

# Create survival object.
surv.object <- Surv(GM$FU, GM$FO)
# Survival analysis
fit1 <- survfit(Surv(GM$FU, GM$FO) ~ 1, data = fulldata)
# Plot.
plot(fit.ST,col=c(rep(2,6),rep(1,2),rep(4,4)),lwd=2,lty=1,
mark=c(1,2,3,0,0,0,0,0,0,1,2,3))

#Survival analysis according to continent and ST.
plot(fit.ST,col=c(rep(2,6),1,"darkgrey",rep(4,4)),lwd=2,
lty=c(2,3,4,7,6,5,1,1,8,2,3,4), xlab="day", ylab="survival probability")
legend("bottomleft", legend = c("C5", "C6", "C32", "C1","C23", "C53", "T5",
"T6", "T32", "T4", "Ctrl1", "Ctrl2"), col=c(rep(2,6),rep(4,4),1,"darkgrey"),
lwd=2,lty=c(3,2,4,5,7,6,3,2,4,8,1,1), ncol=2, bty = "n")

# Log rank test of differences between groups of shared STs
survdiff(Surv(FU, FO) ~ ST, data = GM[GM$ST %in% levels(GM$ST)[c(1,10)],])
```

D. Glossary

AIDS Acquired immunodeficiency syndrome

allele One of the alternative forms of a gene.

allopatric speciation Speciation that occurs when two or more populations of a species are geographically isolated from one another sufficiently that they do not interbreed.

antibacterial Having the ability to kill bacteria.

asexual reproduction A type of reproduction involving only one parent that usually produces genetically identical offspring.

avian Of, relating to, or characteristic of birds (members of the class *Aves*).

bacteria Tiny, single-celled, prokaryotic organisms that can survive in a wide variety of environments.

base The DNA molecule is a chain of nucleotide units; each unit consists of a backbone made of a sugar and a phosphate group, with a nitrogenous base attached. The base in a unit is one of adenine (A), guanine (G), cytosine (C), or thymine (T).

cell The basic structural and functional unit of most living organisms.

CSF Cerebrospinal fluid. A watery fluid, continuously produced and absorbed, which flows in the ventricles (cavities) within the brain and around the surface of the brain and spinal cord.

chromosome A structure in the cell nucleus that carries DNA.

clade A set of species descended from a common ancestral species - synonymous with monophyletic group.

cladogram A branching diagram that illustrates hypotheses about the evolutionary relationships among groups of organisms.

class A category of taxonomic classification between order and phylum, a class comprises members of similar orders.

common ancestor The most recent ancestral form or species from which two different species evolved.

convergence The process by which a similar character evolves independently in two species.

distance In taxonomy, referring to the quantitatively measured difference between the phenetic appearance of two groups of individuals, such as populations or species (phenetic distance), or the difference in their gene frequencies (genetic distance).

DNA Deoxyribonucleic acid, the molecule that controls inheritance.

DNA base sequence A chain of repeating units of deoxyribonucleotides (adenine, guanine, cytosice, thymine) arranged in a particular pattern.

electrophoresis The method of distinguishing entities according to their motility in an electric field.

epistasis An interaction between the genes at two or more loci, such that the phenotype differs from at would be expected if the loci were expressed independently.

eukaryote Any organism made up of eukaryotic cells. Almost all multicellular organisms are eukaryotes.

evolution Darwin defined this term as "descent with modification." It is the change in a lineage of populations between generations. In general terms, biological evolution is the process of change by which new species develop from preexisting species over time; in genetic terms, evolution can be defined as any change in the frequency of alleles in populations of organisms from generation to generation.

family The category of taxonomic classification between order and genus.

fitness The success of an individual (or allele or genotype in a population) in surviving and reproducing, measured by that individual's (or allele's or genotype's) genetic contribution to the next generation and subsequent generations.

fungi A group of organisms comprising the kingdom Fungi, which includes molds and mushrooms.

gene A sequence of nucleotides coding for a protein (or, in some cases, part of a protein); a unit of heredity.

genetic Related to genes. A gene is a sequence of nucleotides coding for a protein (or, in some cases, part of a protein); a unit of heredity.

genome The full set of DNA in a cell or organism.

genotype The set of two genes possessed by an individual at a given locus. More generally, the genetic profile of an individual.

geographic isolation See reproductive isolation.

geographic speciation See allopatric speciation.

reproductive isolation Two populations or individuals of opposite sex are considered reproductively isolated from one another if they cannot together produce fertile offspring.

haploid The condition of having only one set of genes or chromosomes. In normally diploid organisms such as humans, only the gametes are haploid.

haplotype A set of genes at more than one locus inherited by an individual from one of its parents. It is the multi-locus analog of an allele.

heredity The process by which characteristics are passed from one generation to the next.

HIV Acronym for the Human Immunodeficiency Virus, the cause of AIDS (acquired immunodeficiency syndrome)

homology A character shared by a set of species and present in their common ancestor.

homozygote An individual having two copies of the same allele at a genetic locus. Also sometimes applied to larger genetic entities, such as a whole chromosome; a homozygote is then an individual having two copies of the same chromosome.

homozygous Having identical alleles for a particular trait. See also homozygote.

hybrid The offspring of a cross between two species.

hypothesis An explanation of one or more phenomena in nature that can be tested by observations, experiments, or both. In order to be considered scientific, a hypothesis must be falsifiable, which means that it can be proven to be incorrect.

immigration The movement of organisms into an area.

immunocompetent Able to develop an immune response or able to recognize antigens and respond to them.

immunocompromised Having an immune system that has been impaired by disease or treatment.

immunodeficient Lacking immunity and so susceptible to infection

immunosupressed Having a suppressed immune response, as by drugs or radiation.

induction The process of deriving general principles from particular facts.

inference A conclusion drawn from evidence.

inheritance of acquired characters Historically influential but factually erroneous theory that an individual inherits characters that its parents acquired during their lifetimes.

isolation Synonym for reproductive isolation.

reproductive isolation Two populations or individuals of opposite sex are considered reproductively isolated from one another if they cannot together produce fertile offspring.

kingdom The second highest level of taxonomic classification of organisms (below domains).

larva The prereproductive stage of many animals. The term is particularly apt when the immature stage has a different form from the adult. For example, a caterpillar is the larval stage of a butterfly or moth.

lineage An ancestor-descendant sequence of populations, cells or genes.

linkage disequilibrium A condition in which the haplotype frequencies in a population deviate from the values they would have if the genes at each locus were combined at random. (When no deviation exists, the population is said to be in linkage equilibrium.)

linked Of genes, present on the same chromosome.

locus The location in the DNA occupied by a particular gene.

mammals The group (specifically, a class) of animals, descended from a common ancestor, that share the derived characters of hair or fur, mammary glands, and several distinctive features of skeletal anatomy, including a particular type of middle ear.

meiosis A special kind of cell division that occurs during the reproduction of diploid organisms to produce the gametes.

Mendelian inheritance The mode of inheritance of all diploid species, and therefore of nearly all multicellular organisms. Inheritance is controlled by genes, which are passed on to the offspring in the same form as they were inherited from the previous generation. At each locus an individual has two genes – one inherited from its father and the other from its mother. The two genes are represented in equal proportions in its gametes.

microbe A nonscientific and very general term, with no taxonomic significance, sometimes used to refer to microscopic (not visible to the unaided eye) organisms. The term often refers to bacteria or viruses that cause disease or infection.

microevolution Evolutionary changes on the small scale, such as changes in gene frequencies within a population.

mimicry A case in which one species looks more or less similar to another species.

mitosis Cell division. All cell division in multicellular organisms occurs by mitosis except for the special division called meiosis that generates the gametes.

molecular clock The theory that molecules evolve at an approximately constant rate. The difference between the form of a molecule in two species is then proportional to the time since the species diverged from a common ancestor, and molecules become of great value in the inference of phylogeny.

monophyletic group A set of species containing a common ancestor and all of its descendants, and not containing any organisms that are not the descendants of that common ancestor.

morphology The study of the form, shape, and structure of organisms.

mutation A change in genetic material that results from an error in replication of DNA. Mutations can be beneficial, harmful, or neutral.

natural selection The differential survival and reproduction of classes of organisms that differ from one another in one or more usually heritable characteristics. Through this process, the forms of organisms in a population that are best adapted to their local environment increase in frequency relative to less well-adapted forms over a number of generations. This difference in survival and reproduction is not due to chance.

nervous system An organ system, composed of a network of cells called neurons, that allows an animal to monitor its internal and external environment, and to move voluntarily or in response to stimulation.

niche The ecological role of a species; the set of resources it consumes and habitats it occupies.

nucleotide A unit building block of DNA and RNA. A nucleotide consists of a sugar and phosphate backbone with a base attached.

organisms Living things.

panmixia Random mating throughout a population.

parasite An organism that lives on or in a plant or animal of a different species, taking nutrients without providing any benefit to the host.

parsimony The principle of phylogenetic reconstruction in which the phylogeny of a group of species is inferred to be the branching pattern requiring the smallest number of evolutionary changes.

pathogen A microorganism that causes disease.

pathological Related to or caused by disease.

phenotype The physical or functional characteristics of an organism, produced by the interaction of genotype and environment during growth and development.

phenotypic characters Individual traits that can be observed in an organism (including appearance and behavior) and that result from the interaction between the organism's genetic makeup and its environment.

phylogeny The study of ancestral relations among species, often illustrated with a "tree of life" branching diagram, which is also known as a phylogenetic tree.

phylum One of the highest levels of taxonomic classification.

polymorphism A condition in which a population possesses more than one allele at a locus. Sometimes it is defined as the condition of having more than one allele with a frequency of more than five percent in the population.

population A group of organisms, usually a group of sexual organisms that interbreed and share a gene pool.

population genetics The study of processes influencing gene frequencies.

postulate A basic principle.

prokaryote A cell without a distinct nucleus. Bacteria and some other simple organisms are prokaryotic.

protein A molecule made up of a sequence of amino acids.

recombination An event, occurring by the crossing-over of chromosomes during meiosis, in which DNA is exchanged between a pair of chromosomes of a pair. Thus, two genes that were previously unlinked, being on different chromosomes, can become linked because of recombination, and linked genes may become unlinked.

selection Synonym of natural selection.

spacer region A sequence of nucleotides in the DNA between coding genes.

species A set of interbreeding organisms.

theory A well-substantiated explanation of some aspect of the natural world that typically incorporates many confirmed observations, laws, and successfully verified hypotheses.

trait A characteristic or condition.

variance A measure of how variable a set of numbers are.

vertebrates The group (specifically, a subphylum) of animals, descended from a common ancestor, that share the derived character of an internal skeleton made of bone or cartilage.

virulence The disease-producing ability of a microorganism.

virus A kind of intracellular parasite that can replicate only inside a living cell.

Bibliography

[1] Agapow, P. M. and Burt, A. (2001) 'Indices of multilocus linkage disequilibrium', *Molecular Ecology Notes,* 1(1-2), 101-102.

[2] Ajzenberg, D., Banuls, A. L., Tibayrenc, M. and Darde, M. L. (2002) 'Microsatellite analysis of *Toxoplasma gondii* shows considerable polymorphism structured into two main clonal groups', *International Journal for Parasitology,* 32(1), 27-38.

[3] Alanio, A., Desnos-Ollivier, M. and Dromer, F. (2011) 'Dynamics of *Cryptococcus neoformans* – macrophage interactions reveal that fungal background influences outcome during cryptococcal meningoencephalitis in humans', *mBio,* 2(4).

[4] Alcorn, K. (2006) 'Thailand to issue compulsory license for efavirenz', *Aidsmap*, available: `http://www.aidsmap.com/Thailand-to-issue-compulsory-license-for-efavirenz/page/1425713/` [accessed 12 February, 2011].

[5] Amjad, M., Kfoury, N., Cha, R. and Mobarak, R. (2004) 'Quantification and assessment of viability of *Cryptococcus neoformans* by LightCycler amplification of capsule gene mRNA', *Journal of Medical Microbiology,* 53(12), 1201-1206.

[6] Anderson, I. C., Campbell, C. D. and Prosser, J. I. (2003a) 'Diversity of fungi in organic soils under a moorland – Scots pine (*Pinus sylvestris L.*) gradient', *Environmental Microbiology,* 5(11), 1121-1132.

[7] Anderson, I. C., Campbell, C. D. and Prosser, J. I. (2003b) 'Potential bias of fungal 18S rDNA and internal transcribed spacer polymerase chain reaction primers for estimating fungal biodiversity in soil', *Environmental Microbiology,* 5(1), 36-47.

[8] Anderson, M. J., Gull K., and Denning D. W. (1996) 'Molecular typing by random amplification of polymorphic DNA and m13 southern hybridization of related paired isolates of *Aspergillus fumigatus*', *Journal of Clinical Microbiology,* 34(1), 87-93.

[9] Antinori, S., Galimberti, L., Magni, C., Casella, A., Vago, L., Mainini, F., Piazza, M., Nebuloni, M., Fasan, M., Bonaccorso, C., Vigevani, G. M., Cargnel, A., Moroni, M. and Ridolfo, A. (2001) '*Cryptococcus neoformans* infection in a cohort of Italian AIDS patients: Natural history, early prognostic parameters, and autopsy findings', *European Journal of Clinical Microbiology & Infectious Diseases,* 20(10), 711-717.

[10] Antinori, S., Radice, A., Galimberti, L., Magni, C., Fasan, M. and Parravicini, C. (2005) 'The role of cryptococcal antigen assay in diagnosis and monitoring of cryptococcal meningitis', *Journal of Clinical Microbiology,* 43(11), 5828-5829.

[11] Apidianakis, Y., Rahme, L. G., Heitman, J., Ausubel, F. M., Calderwood, S. B. and Mylonakis, E. (2004) 'Challenge of *Drosophila melanogaster* with *Cryptococcus neoformans* and role of the innate immune response', *Eukaryotic Cell,* 3(2), 413-419.

[12] Archibald, L. K., McDonald, L. C., Rheanpumikankit, S., Tansuphaswadikul, S., Chaovanich, A., Eampokalap, B., Banerjee, S. N., Reller, L. B. and Jarvis, W. R. (1999) 'Fever and Human Immunodeficiency Virus Infection as Sentinels for Emerging Mycobacterial and Fungal Bloodstream Infections in Hospitalized Patients $>/=15$ Years Old, Bangkok', *The Journal of Infectious Diseases,* 180(1), 87-92.

[13] Asawavichienjinda, T., Sitthi-Amorn, C. and Tanyanont, V. (1999) 'Serum cryptococcal antigen : Diagnostic value in the diagnosis of AIDS-related cryptococcal meningitis', *Journal of the Medical Association of Thailand,* 82(1), 65-71.

[14] Aufderheide Arthur C. and Rodriguez-Martin C. (1998) *The Cambridge Encyclopedia of Human Paleopathology,* New York, NY: Cambridge University Press.

[15] Aulakh, H. S., Straus, S. E. and Kwon-Chung, K. J. (1981) 'Genetic relatedness of Filobasidiella neoformans (*Cryptococcus neoformans*) and Filobasidiella bacillispora (*Cryptococcus* bacillisporus) as determined by deoxyribonucleic acid base composition and sequence homology studies', *International Journal of Systematic Bacteriology,* 31(1), 97-103.

[16] Auvert, B., Taljaard, D., Lagarde, E., Sobngwi-Tambekou, J., Sitta, M. and Puren, A. (2005) 'Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 trial', *Plos Medicine,* 2(11), 1112-1122.

[17] Avert (2011) 'HIV and AIDS in South Africa', available: `http://www.avert.org/aidssouthafrica.htm` [accessed 09/02/2011].

[18] Avise, J. (2004) *Molecular Markers, Natural History, and Evolution.,* 2nd ed.*,* New York: Chapman & Hall.

[19] Bain J. M., Tavanti A., Davidson A. D., Jacobsen M. D.,Shaw D., Gow N. A. R., Odds F. C. (2007). Multilocus Sequence Typing of the pathogenic f ungus *Aspergillus fumigates*, J Clin Microbiol., 45(5): 1469-1477.

[20] Baker, R. D. (1976) 'The primary pulmonary lymph node complex of crytptococcosis', *American Journal of Clinical Pathology,* 65(1), 83-92.

[21] Banerjee, U., Datta, K., Majumdar, T. and Gupta, K. (2001) 'Cryptococcosis in India: the awakening of a giant?', *Med Mycol,* 39(1), 51-67.

[22] Bankowski, M. and Anderson, S. (2004) 'Real-time nucleic acid amplification in clinical microbiology', *Clinical Microbiology Newsletter,* 26(2), 9 - 15.

[23] Barreto de Oliveira, M. T., Boekhout, T., Theelen, B., Hagen, F., Baroni, F. A., Lazera, M. S., Lengeler, K. B., Heitman, J., Rivera, I. N. G. and Paula, C. R. (2004) '*Cryptococcus neoformans* shows a remarkable genotypic diversity in Brazil', *J Clin Microbiol,* 42(3), 1356-9.

[24] Bart-Delabesse, E., Sarfati, J., Debeaupuis, J. P., van Leeuwen, W., van Belkum, A., Bretagne, S. and Latge, J. P. (2001) 'Comparison of restriction fragment length polymorphism, microsatellite length polymorphism, and random amplification of polymorphic DNA analyses for fingerprinting *Aspergillus fumigatus* isolates', *Journal of Clinical Microbiology,* 39(7), 2683-2686.

[25] Bartlett, K. H., Kidd, S. E. and Kronstad, J. W. (2008) 'The emergence of *Cryptococcus gattii* in British Columbia and the Pacific Northwest', *Current infectious disease reports,* 10(1), 58-65.

[26] Bava, A. J., Negroni, R., Arechavala, A., Robles, A. M. and Bianchi, M. (1997) 'Cryptococcosis associated with AIDS in the Muniz Hospital of Buenos Aires', *Mycopathologia,* 140(1), 13-7.

[27] Bell M., Archibald L. K., Nwanyanwu O., Dobbie H., Tokars J., Kazembe P. N., Reller L. B. and Jarvis W. R. (2001) 'Seasonal variation in the etiology of bloodstream infections in a febrile inpatient population in a developing country'. *Int J Infect Dis* 5: 63-69.

[28] Bennett, J. E. (1977) 'Flucytosine', *Annals of Internal Medicine,* 86(3), 319-322.

[29] Bennett, R. S., Milgroom, M. G. and Bergstrom, G. C. (2005) 'Population structure of seedborne Phaeosphaeria nodorum on New York wheat', *Phytopathology,* 95(3), 300-305.

[30] Bergemann, A. and Karstaedt, A. S. (1996) 'The spectrum of meningitis in a population with high prevalence of HIV disease', *Qjm-Monthly Journal of the Association of Physicians,* 89(7), 499-504.

[31] Bergin, D., Reeves, E. P., Renwick, J., Wientjes, F. B. and Kavanagh, K. (2005) 'Superoxide production in Galleria mellonella hemocytes: Identification of proteins homologous to the NADPH oxidase complex of human neutrophils', *Infection and Immunity,* 73(7), 4161-4170.

[32] Besnard, G., Henry, P., Wille, L., Cooke, D. and Chapuis, E. (2007) 'On the origin of the invasive olives (Olea europaea L., Oleaceae)', *Heredity,* 99(6), 608-619.

[33] Bialek, R., Weiss, M., Bekure-Nemariam, K., Najvar, L. K., Alberdi, M. B., Graybill, J. R. and Reischl, U. (2002) 'Detection of *Cryptococcus neoformans* DNA in tissue samples by nested and real-time PCR assays', *Clinical and Diagnostic Laboratory Immunology,* 9(2), 461-469.

[34] Bicanic, T., Brouwer, A. E., Meintjes, G., Rebe, K., Limmathurotsakul, D., Chierakul, W., Teparrakkul, P., Loyse, A., White, N. J., Wood, R., Jaffar, S. and Harrison, T. (2009a) 'Relationship of cerebrospinal fluid pressure, fungal burden and outcome in patients with cryptococcal meningitis undergoing serial lumbar punctures', *Aids,* 23(6), 701-706.

[35] Bicanic, T. and Harrison, T. S. (2004) 'Cryptococcal meningitis', *Br Med Bull,* 72, 99-118.

[36] Bicanic, T., Meintjes, G., Wood, R., Hayes, M., Rebe, K., Bekker, L. G. and Harrison, T. (2007) 'Fungal burden, early fungicidal activity, and outcome in cryptococcal meningitis in antiretroviral-naive or antiretroviral-experienced patients treated with amphotericin B or fluconazole (vol 45, pg 76, 2007)', *Clinical Infectious Diseases,* 45(4), 526-526.

[37] Bicanic, T., Muzoora, C., Brouwer, A. E., Meintjes, G., Longley, N., Taseera, K., Rebe, K., Loyse, A., Jarvis, J., Bekker, L. G., Wood, R., Limmathurotsakul, D., Chierakul, W., Stepniewska, K., White, N. J., Jaffar, S. and Harrison, T. S. (2009b) 'Independent Association between Rate of Clearance of Infection and Clinical Outcome of HIV-Associated Cryptococcal Meningitis: Analysis of a Combined Cohort of 262 Patients', *Clinical Infectious Diseases,* 49(5), 702-709.

[38] Bicanic, T., Wood, R., Meintjes, G., Rebe, K., Brouwer, A., Loyse, A., Bekker, L. G., Jaffar, S. and Harrison, T. (2008) 'High-dose amphotericin B with flucytosine for the treatment of cryptococcal meningitis in HIV-infected patients: A randomized trial', *Clinical Infectious Diseases,* 47(1), 123-130.

[39] Birdsell, J. and Wills, C. (1996) 'Significant competitive advantage conferred by meiosis and syngamy in the yeast Saccharomyces cerevisiae', *Proceedings of the National Academy of Sciences of the United States of America,* 93(2), 908-912.

[40] Bisson, G. P., Lukes, J., Thakur, R., Mtoni, I. and MacGregor, R. R. (2008) '*Cryptococcus* and lymphocytic meningitis in Botswana', *Samj South African Medical Journal,* 98(9), 724-725.

[41] Boekhout, T., Theelen, B., Diaz, M., Fell, J. W., Hop, W. C. J., Abeln, E. C. A., Dromer, F. and Meyer, W. (2001) 'Hybrid genotypes in the pathogenic yeast *Cryptococcus* neoformans', *Microbiology-Uk,* 147, 891-907.

[42] Bogaerts, J., Rouvroy, D., Taelman, H., Kagame, A., Aziz, M. A., Swinne, D. and Verhaegen, J. (1999) 'AIDS-associated Cryptococcal meningitis in Rwanda (1983-1992): Epidemiologic and diagnostic features', *Journal of Infection,* 39(1), 32-37.

[43] Bolanos, B. and Mitchell, T. G. (1989) 'Phagocytosis of *Cryptococcus*-Neoformans by Rat Alveolar Macrophages', *Journal of Medical and Veterinary Mycology,* 27(4), 203-217.

[44] Boman H.G., Hultmark D. (1987) 'Cell-free immunity in insects.' *Ann. Rev. Microbiol.*, 41, 103-126.

[45] Bosch J., Martinez-Solano I. and Garcia-Paris M. (2000) 'Evidence of a chytrid fungus infection involved in the decline of the common midwife toad in protected areas of Central Spain', *Conference Proceedings: Getting the Jump on Amphibian Disease*, Cairns, Australia.

[46] Botts M.R., Giles S.S., Gates M.A., Kozel T.R., Hull C.M. (2009) 'Isolation and characterization of *Cryptococcus neoformans* spores reveal a critical role for capsule biosynthesis genes inspore biogenesis. *Eukaryot Cell*, 8 , pp. 595-605.

[47] Bougnoux, M.-E., Aanensen, D. M., Morand, S., Theraud, M., Spratt, B. G. and d'Enfert, C. (2004) 'Multilocus sequence typing of Candida albicans: strategies, data exchange and applications', *Infection Genetics and Evolution,* 4(3), 243-252.

[48] Bovers, M., Hagen, F. and Boekhout, T. (2008a) 'Diversity of the *Cryptococcus neoformans* (*Cryptococcus gattii*) species', *Revista Iberoamericana De Micologia,* 25(1), S4-S12.

[49] Bovers, M., Hagen, F., Kuramae, E., Diaz, M., Spanjaard, L., Dromer, F., Hoogveld, H. and Boekhout, T. (2006) 'Unique hybrids between the fungal pathogens *Cryptococcus* neoformans and *Cryptococcus gattii*', *Fems Yeast Research*, 599-607.

[50] Bovers, M., Hagen, F., Kuramae, E., Hoogveld, H., Dromer, F., St-Germain, G. and Boekhout, T. (2008) 'AIDS patient death caused by novel *Cryptococcus neoformans* x C.*gattii* hybrid', *Emerging Infectious Diseases*, 1105-1108.

[51] Bovers, M., Hagen, F., Kuramae, E. E. and Boekhout, T. (2008b) 'Six monophyletic lineages identified within *Cryptococcus neoformans* and *Cryptococcus gattii* by multilocus sequence typing', *Fungal Genetics and Biology,* 45(4), 400-421.

[52] Boyd, E. F., Wang, F. S., Whittam, T. S. and Selander, R. K. (1996) 'Molecular genetic relationships of the salmonellae', *Applied and Environmental Microbiology,* 62(3), 804-808.

[53] Brandt, M. E., Bragg, S. L. and Pinner, R. W. (1993) 'Multilocus enzyme typing of *Cryptococcus neoformans*', *Journal of Clinical Microbiology,* 31(10), 2819-2823.

[54] Brandt, M. E., Hutwagner, L. C., Klug, L. A., Baughman, W. S., Rimland, D., Graviss, E. A., Hamill, R. J., Thomas, C., Pappas, P. G., Reingold, A. L., Pinner, R. W., Stephens, D., Farley, M., Bardsley, M., Siegel, B., Jackson, G., Lao, C., Otte, J., Harvey, C., Gillespie, R., Rothrock, G., Pattni, B. and Daily, P. (1996) 'Molecular subtype distribution of *Cryptococcus neoformans* in four areas of the United States', *Journal of Clinical Microbiology,* 34(4), 912-917.

[55] Brandt, M. E., Hutwagner, L. C., Kuykendall, R. J., Pinner, R. W., Stephens, D., Farley, M., Rimland, D., Baughman, W., Lao, C., Otte, J., Harvey, C., Hamill, R., Graviss, E., Pappas, P., Thomas, C., Gillespie, R., Reingold, A. L., Rothrock, G., Pattni, B., Daily, P., Klug, L., Brockwell, C. and Shihata, N. (1995) 'Comparison of multilocus enzyme electrophoresis and random amplified polymorphic DNA analysis for molecular subtyping of *Cryptococcus-neoformans*', *Journal of Clinical Microbiology,* 33(7), 1890-1895.

[56] Braverman, J. M., Hudson, R. R., Kaplan, N. L., Langley, C. H. and Stephan, W. (1995) 'The hitchhiking effect on the site frequency spectrum of DNA polymorphisms', *Genetics,* 140(2), 783-796.

[57] Brennan, M., Thomas, D. Y., Whiteway, M. and Kavanagh, K. (2002) 'Correlation between virulence of *Candida albicans* mutants in mice and *Galleria mellonella* larvae', *Fems Immunology and Medical Microbiology,* 34(2), PII S0928-8244(02)00374-7.

[58] Brouwer, A. E., Rajanuwong, A., Chierakul, W., Griffin, G. E., Larsen, R. A., White, N. J. and Harrison, T. S. (2004) 'Combination antifungal therapies for HIV-associated cryptococcal meningitis: a randomised trial', *Lancet,* 363(9423), 1764-1767.

[59] Brown, A. H. D., Feldman, M. W. and Nevo, E. (1980) 'Multilocus Structure of Natural-Populations of Hordeum-Spontaneum', *Genetics,* 96(2), 523-536.

[60] Brown, T., Sittitrai, W., Vanichseni, S. and Thisyakorn, U. (1994) 'The recent epidemiology of HIV and AIDS in Thailand', *Aids,* 8, S131-S141.

[61] Buchanan, K. L. and Murphy, J. W. (1998) 'What makes *Cryptococcus neoformans* a pathogen?', *Emerg Infect Dis,* 4(1), 71-83.

[62] Bui, T., Lin, X., Malik, R., Heitman, J. and Carter, D. (2008) 'Isolates of *Cryptococcus* neoformans from Infected Animals Reveal Genetic Exchange in Unisexual, alpha Mating Type Populations', *Eukaryotic Cell,* 7(10), 1771-1780.

[63] Bulmer, G., MD Sans (1967) '*Cryptococcus neoformans.* III. Inhibition of phagocytosis', *J. Bacteriol.,* (95), 5-8.

[64] Bulmer, G. S., Sans, M. D. and Gunn, C. M. (1967) '*Cryptococcus neoformans.* I. Nonencapsulated mutants', *Journal of bacteriology,* 94(5), 1475-9.

[65] Burnie, J. P., Coke, A. and Matthews, R. C. (1992) 'Restriction endonuclease analysis of *Aspergillus fumigatus* DNA', *Journal of Clinical Pathology,* 45(4), 324-327.

[66] Burt, A., Carter, D. A., Koenig, G. L., White, T. J. and Taylor, J. W. (1996) 'Molecular markers reveal cryptic sex in the human pathogen Coccidioides immitis', *Proceedings of the National Academy of Sciences of the United States of America,* 93(2), 770-773.

[67] Burt, A., Koufopanou, V. and Taylor, J. W. (2000) 'Population genetics of human-pathogenic fungi', *Molecular Epidemiology of Infectious Diseases*, Arnold, London, 229-244.

[68] Butler, G. (2007) 'The evolution of MAT: The ascomycetes', *Sex in Fungi: Molecular Determination and Evolutionary Implications-Print*, 3.

[69] Byrnes, E. J., Li, W., Lewit, Y., Ma, H., Voelz, K., Ren, P., Carter, D. A., Chaturvedi, V., Bildfell, R. J., May, R. C. and Heitman, J. (2010) 'Emergence and pathogenicity of highly virulent *Cryptococcus gattii* genotypes in the northwest United States', *PLoS Pathog,* 6(4), e1000850.

[70] Campbell, L. T. and Carter, D. A. (2006) 'Looking for sex in the fungal pathogens *Cryptococcus neoformans* and *Cryptococcus gattii*', *Fems Yeast Research,* 6(4), 588-598.

[71] Campbell, L. T., Currie, B. J., Krockenberger, M., Malik, R., Meyer, W., Heitman, J. and Carter, D. (2005) 'Clonality and recombination in genetically differentiated subgroups of *Cryptococcus gattii*', *Eukaryotic Cell,* 4(8), 1403-1409.

[72] Casadevall, A. (2005) 'Host as the variable: Model hosts approach the immunological asymptote', *Infection and Immunity,* 73(7), 3829-3832.

[73] Casadevall, A. and Perfect, J. R. (1998) *Cryptococcus neoformans,* Washington, D.C.: ASM Press.

[74] Casadevall A., Pirofski L. A. (2006)'The weapon potential of human pathogenic fungi', *Med Mycol*, 44:689-696.

[75] Casadevall A., Pirofski L. (2001) 'Host-pathogen interactions: The attributes of virulence', *Journal of Infectious Diseases*, Vol. 184: 3, pp. 337-344.

[76] Casadevall, A., Steenbergen, J. N. and Nosanchuk, J. D. (2003) ' 'Ready made' virulence and 'dual use' virulence factors in pathogenic environmental fungi - the *Cryptococcus* neoformans paradigm', *Current Opinion in Microbiology,* 6(4), 332-337.

[77] Casali, A. K., Goulart, L., Silva, L. K. R., Silva, K. R. E., Ribeiro, A. M., Amaral, A. A., Alves, S. H., Schrank, A., Meyer, W. and Vainstein, M. H. (2003) 'Molecular typing of clinical and environmental *Cryptococcus neoformans* isolates in the Brazilian state Rio Grande do Sul', *Fems Yeast Research,* 3(4), 405-415.

[78] Castanon-Olivares, L. R., Arreguin-Espinosa, R., Ruiz-Palacios y Santos, G. and Lopez-Martinez, R. (2000) 'Frequency of *Cryptococcus* species and varieties in Mexico and their comparison with some Latin American countries', *Rev Latinoam Microbiol,* 42(1), 35-40.

[79] Castro, K. G., Ward, J. W., Slutsker, L., Buehler, J. W., Jaffe, H. W., Berkelman, R. L. and Curran, J. W. (1993) '1993 Revised Classification System for HIV Infection and Expanded Surveillance Case Definition for AIDS Among Adolescents and Adults (reprinted from MMWR, Vol 41, PG RR 17, 1992)', *Clinical Infectious Diseases,* 17(4), 802-810.

[80] Centre for Disease Control, British Columbia (2008) 'BC *Cryptococcus gattii* surveillance summary, 1999-2006', available from `http://www.bccdc.org/topic.php?item=109`. Accessed November 2010.

[81] Chan, M. S., Maiden, M. C. J. and Spratt, B. G. (2001) 'Database-driven multi locus sequence typing (MLST) of bacterial pathogens', *Bioinformatics,* 17(11), 1077-1083.

[82] Chang, Y. C. and Kwon-Chung, K. J. (1994) 'Complementation of a capsule-deficient mutation of *Cryptococcus neoformans* restores its virulence', *Molecular and Cellular Biology,* 14(7), 4912-4919.

[83] Chang, Y. C., Stins, M. F., McCaffery, M. J., Miller, G. F., Pare, D. R., Dam, T., Paul-Satyaseela, M., Mm, K. S. and Kwon-Chung, K. J. (2004) 'Cryptococcal yeast cells invade the central nervous system via transcellular penetration of the blood-brain barrier (vol 72, pg 4985, 2004)', *Infection and Immunity,* 72(11), 6753-6753.

[84] Chariyalertsak, S., Sirisanthana, T., Saengwonloey, O. and Nelson, K. (2001) 'Clinical presentation and risk behaviors of patients with acquired immunodeficiency syndrome in Thailand, 1994-1998: Regional variation and temporal trends', *Clinical Infectious Diseases*, 955-962.

[85] Charlesworth D., Morgan M. T., Charlesworth B. (1993) 'Mutation accumulation in finite outbreeding and inbreeding populations', *Genet Res*, 61:39-56.

[86] Check, W. A. (1994) '*Molecular techniques shed light on fungal genetics.*', *ASM News.,* 60, 593-596.

[87] Chen, G. H., McDonald, R. A., Wells, J. C., Huffnagle, G. B., Lukacs, N. W. and Toews, G. B. (2005) 'The gamma interferon receptor is required for the protective pulmonary inflammatory response to *Cryptococcus neoformans*', *Infection and Immunity,* 73(3), 1788-1796.

[88] Chen, J., Varma, A., Diaz, M., Litvintseva, A., Wollenberg, K. and Kwon-Chung, K. (2008) '*Cryptococcus neoformans* strains and infection in apparently immunocompetent patients, China', *Emerging Infectious Diseases*, 755-762.

[89] Chen, S., Sorrell, T., Nimmo, G., Speed, B., Currie, B., Ellis, D., Marriott, D., Pfeiffer, T., Parr, D., Byth, K. and Australasian Cryptococcal Study, G. (2000) 'Epidemiology and host- and variety-dependent characteristics of infection due to *Cryptococcus neoformans* in Australia and New Zealand', *Clinical Infectious Diseases,* 31(2), 499-508.

[90] Chen, S. C. A., Muller, M., Zhou, J. Z., Wright, L. C. and Sorrell, T. C. (1997) 'Phospholipase activity in *Cryptococcus neoformans*: A new virulence factor?', *Journal of Infectious Diseases,* 175(2), 414-420.

[91] Chen, S. C. A., Wright, L. C., Golding, J. C. and Sorrell, T. C. (2000) 'Purification and characterization of secretory phospholipase B, lysophospholipase and lysophospholipase/transacylase from a virulent strain of the pathogenic fungus *Cryptococcus neoformans*', *Biochemical Journal,* 347, 431-439.

[92] Chen, Y. C., Eisner, J. D., Kattar, M. M., Rassoulian-Barrett, S. L., LaFe, K., Yarfitz, S. L., Limaye, A. P. and Cookson, B. T. (2000) 'Identification of medically important yeasts using PCR-based detection of DNA sequence polymorphisms in the internal transcribed spacer 2 region of the rRNA genes', *Journal of Clinical Microbiology,* 38(6), 2302-2310.

[93] Cherniak, R., Morris, L. C., Belay, T., Spitzer, E. D. and Casadevall, A. (1995) 'Variation in the Structure of Glucuronoxylomannan in Isolates from Patients with Recurrent Cryptococcal Meningitis', *Infection and Immunity,* 63(5), 1899-1905.

[94] Cherniak, R. and Sundstrom, J. B. (1994) 'Polysaccharide antigens of the capsule of *Cryptococcus neoformans*', *Infection and Immunity,* 62(5), 1507-1512.

[95] Chretien, F., Lortholary, O., Kansau, I., Neuville, S., Gray, F. and Dromer, F. (2002) 'Pathogenesis of cerebral *Cryptococcus neoformans* infection after fungemia', *Journal of Infectious Diseases,* 186(4), 522-530.

[96] Chuck, S. L. and Sande, M. A. (1989) 'Infections with *Cryptococcus neoformans* in the acquired immunodeficiency syndrome', *New England Journal of Medicine,* 321(12), 794-799.

[97] Ciosi, M., Miller, N. J., Kim, K. S., Giordano, R., Estoup, A. and Guillemaud, T. (2008) 'Invasion of Europe by the western corn rootworm, Diabrotica virgifera virgifera: multiple transatlantic introductions with various reductions of genetic diversity', *Molecular Ecology,* 17(16), 3614-3627.

[98] Clement, M., Posada, D. and Crandall, K. A. (2000) 'TCS: a computer program to estimate gene genealogies', *Molecular Ecology,* 9(10), 1657-1659.

[99] Cockerill, F. R. (2003) 'Application of rapid-cycle real-time polymerase chain reaction for diagnostic testing in the clinical microbiology laboratory', *Archives of Pathology & Laboratory Medicine,* 127(9), 1112-1120.

[100] Cogliati, M., Allaria, M., Tortorano, A. M. and Viviani, M. A. (2000) 'Genotyping *Cryptococcus neoformans* var. neoformans with specific primers designed from PCR-fingerprinting bands sequenced using a modified PCR-based strategy', *Medical Mycology,* 38(2), 97-103.

[101] Colvin, M., Dawood, S., Kleinschmidt, I., Mullick, S. and Lallo, U. (2001) 'Prevalence of HIV and HIV-related diseases in the adult medical wards of a tertiary hospital in Durban, South Africa', *International Journal of Std & Aids,* 12(6), 386-389.

[102] Converse J. L., Reed R.E. (1996) Experimental epidemiology of *coccidioidomycosis. Bacteriol Rev*, 30:678-695.

[103] Corbett, E. L., Churchyard, G. J., Charalambos, S., Samb, B., Moloi, V., Clayton, T. C., Grant, A. D., Murray, J., Hayes, R. J. and De Cock, K. M. (2002) 'Morbidity and mortality in South African gold miners: Impact of untreated disease due to human immunodeficiency virus', *Clinical Infectious Diseases,* 34(9), 1251-1258.

[104] Cotter, G., Doyle, S. and Kavanagh, K. (2000) 'Development of an insect model for the in vivo pathogenicity testing of yeasts', *Fems Immunology and Medical Microbiology,* 27(2), 163-169.

[105] Cuenca-Estrella, M., Bernal-Martinez, L., Buitrago, M. J., Castelli, M. V., Gomez-Lopez, A., Zaragoza, O. and Rodriguez-Tudela, J. L. (2008) 'Update on the epidemiology and diagnosis of invasive fungal infection', *International Journal of Antimicrobial Agents,* 32, S143-S147.

[106] Currie, B. P., Freundlich, L. F. and Casadevall, A. (1994) 'Restriction fragment length polymorphism analysis of *Cryptococcus neoformans* isolates from environmental (pigeon excreta) and clinical sources in New York City', *Journal of Clinical Microbiology,* 32(5), 1188-1192.

[107] Curtis, J. L., Huffnagle, G. B., Chen, G. H., Warnock, M. L., Gyetko, M. R., McDonald, R. A., Scott, P. J. and Toews, G. B. (1994) 'Experimental murine pulmonary cryptococcosis. Differences in pulmonary inflammation and lymphocyte recruitment induced by two encapsulated strains of *Cryptococcus neoformans.*', *Laboratory Investigation,* 71(1), 113-126.

[108] Dard M. L., Bouteille B., Pestre-Alexandre M. (1992) 'Isoenzyme Analysis of 35 *Toxoplasma gondii* isolates and the biological and epidemiological implications.' *The Journal of Parasitology*, Vol. 78, No. 5, pp. 786-794.

[109] Daszak, P., Berger, L., Cunningham, A.A., Hyatt, A.D., Green, D.E. and Speare, R. (1999) 'Emerging infectious diseases and amphibian population declines', *Emerg. Infect. Dis.,* 5, 735748.

[110] Daszak P., Cunningham A. A. and Hyatt A. D. (2000). 'Emerging Infectious diseases of wildlife – threats to biodiversity and human health.' *Science*, Vol. 287 no. 5452 pp. 443-449. DOI: 10.1126/science.287.5452.443.

[111] Daszak, P., Cunningham, A. A. and Hyatt, A. D. (2001) 'Anthropogenic environmental change and the emergence of infectious diseases in wildlife', *Acta Tropica,* 78(2), 103-116.

[112] Datta, K., Bartlett, K. H., Baer, R., Byrnes, E., Galanis, E., Heitman, J., Hoang, L., Leslie, M. J., MacDougall, L., Magill, S. S., Morshed, M. G., Marr, K. A. and *Cryptococcus* Gattii Working, G. (2009) 'Spread of *Cryptococcus gattii* into Pacific Northwest Region of the United States', *Emerging Infectious Diseases,* 15(8), 1185-1191.

[113] Datta K., Bartlett K. H., Marr K. A. (2009). '*Cryptococcus gattii*: emergence in western North America: Exploitation of a novel ecological niche', Volume 2009, Article ID 176532, 8 pages doi:10.1155/2009/176532

[114] Davidson, R. C., Moore, T. D. E., Odom, A. R. and Heitman, J. (2000) 'Characterization of the MF alpha pheromone of the human fungal pathogen *Cryptococcus neoformans*', *Molecular Microbiology,* 38(5), 1017-1026.

[115] Day, J. (2004) 'Cryptococcal meningitis', *Practical Neurology,* (4), 274-285.

[116] Debuchy, R. and Turgeon, B., eds. (2006) *Mating-Type structure, evolution, and function in euascomycetes,* Berlin Heidelberg: Springer-Verlag.

[117] De Figueiredo D. R., Azeiteiro U. M., Esteves S. M., Goncalves F. J. M., Pereira M. J. (2004) 'Microcystin-producing blooms - A serious global public health issue Ecotoxicology and Environmental Safety', 59 (2), pp. 151-163.

[118] Department of Health (1995) *AIDS in South Africa, 1995*, 2.

[119] Department of Health (2008) *The burden of Cryptococcosis in South Africa.*

[120] Department of Health (2011) *Massive reduction in ARV prices.*,

[121] Dettman, J. R., Jacobson, D. J. and Taylor, J. W. (2003) 'A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote Neurospora', *Evolution,* 57(12), 2703-2720.

[122] Devirgilio, C., Hottiger, T., Dominguez, J., Boller, T. and Wiemken, A. (1994) 'The role of trehalose synthesis for the acquisition of thermotolerance in yeast. I. Genetic evidence that trehalose is a thermoprotectant.', *European Journal of Biochemistry,* 219(1-2), 179-186.

[123] Diamond, R. D., Root, R. K. and Bennett, J. E. (1972) 'Factors Influencing Killing of *Cryptococcus*-Neoformans by Human Leukocytes in-Vitro', *Journal of Infectious Diseases,* 125(4), 367-&.

[124] Diaz, M. R., Boekhout, T., Theelen, B. and Fell, J. W. (2000) 'Molecular sequence analyses of the intergenic spacer (IGS) associated with rDNA of the two varieties of the pathogenic yeast, *Cryptococcus neoformans*', *Systematic and Applied Microbiology,* 23(4), 535-545.

[125] Dickie, I. A., Xu, B. and Koide, R. T. (2002) 'Vertical niche differentiation of ectomycorrhizal hyphae in soil as shown by T-RFLP analysis', *New Phytologist,* 156(3), 527-535.

[126] Dodgson, A. R., Pujol, C., Denning, D. W., Soll, D. R. and Fox, A. J. (2003) 'Multilocus sequence typing of Candida glabrata reveals geographically enriched clades', *Journal of Clinical Microbiology,* 41(12), 5709-5717.

[127] Dong, Z. and Murphy, J. (1997) 'Cryptococcal polysaccharides bind to CD18 on human neutrophils', *Infection and Immunity,* 65(2), 557-563.

[128] Dong, Z. M. and Murphy, J. W. (1995) 'Effects of the 2 Varieties of *Cryptococcus*-Neoformans Cells and Culture Filtrate Antigens on Neutrophil Locomotion', *Infection and Immunity,* 63(7), 2632-2644.

[129] Dromer, F., Mathoulin, S., Dupont, B., Laporte, A., Chardon, H., Greze, H., Hayette, M. P., Chabasse, D., Martin, M., Pulik, M., Hautefort, B., LePennec, M. P., Larfouilloux, J., Faller, J. P., Barale, T., BougesMichel, C., Delzanig, G., Couprie, B., Bougnoux, M. E., Masure, O., Lefort, C., Duhamel, C., Bidault, C., Zaegel, M., Allard, C., Laurens, E., Dormont, J., Beytout, J., Mollo, J. L., Boussougant, Y., Cordier, F., Bretagne, S., Gaillard, J. P., Portier, H., Grise, G., Leblanc, A., Bouchard, I., Nauciel, C., Lebeau, B., Bouree, P., Lagarde, A., Vincent, J., DesaillyChanson, M. A., Eloy, O., Grillot, M. L., Gosset, X., Capbern, P., Darde, M. L., Lemercier, Y., Piens, M. A., Janin, G., Tremolieres, F., Quilici, M., Eme, A., Piquet, M., Reynes, J., Gettler, V., Trevoux, A., Kures, L., Morin, O., Kernbaum, S., GariToussaint, M., Delage, A., Barthez, J. P., Giudicelli, Y., Jacquemin, J. L., Bouvry, M., Pateyron, F., Toubas, D., Guigen, C., Brasseur, C., Lafaye, J. M. P., Sirondelle, G., Gauthron, M., Janvier, M., Godineau, N., Dorche, G., Fegueux, S., Roue, R., Simoneau, M., Malherbe, P., Koenig, H., Cahen, P., Muzellec, Y., Linas, M. D., Baixench, M. T., Declosets, F., Verger, J., Saliba, F., Patey, O., Datry, A., Lavarde, V., deRecondo, J., Ponsot, G., Lamer, C., Hennequin, C., Traore, F., Arborio, M., Dematons, C., Roux, P., et al. (1996) 'Epidemiology of cryptococcosis in France: A 9-year survey (1985-1993)', *Clinical Infectious Diseases,* 23(1), 82-90.

[130] Dromer, F., Mathoulin-Pelissier, S., Fontanet, A., Ronin, O., Dupont, B., Lortholary, O. and French Cryptococcosis Study, G. (2004) 'Epidemiology of HIV-associated cryptococcosis in France (1985-2001): comparison of the pre- and post-HAART eras', *Aids,* 18(3), 555-562.

[131] Drummond AJ, H. S., Rawlence N, Rambaut A (2007) 'A rough guide to BEAST 1.4', available: `http://beast.bio.ed.ac.uk/Main\_Page\#A\_Rough\_Guide\_to\_BEAST\_1.4` [accessed 10th December 2009].

[132] Duncan, C., Schwantje, H., Stephen, C., Campbell, J. and Bartlett, K. (2006) '*Cryptococcus gattii* in wildlife of Vancouver Island, British Columbia, Canada', *Journal of Wildlife Diseases,* 42(1), 175-178.

Dunphy G.B., Thurston G.S. (1990) *Insect Immunity.* CRC Press, Boca Raton, Florida.

[133] Dutta, S. K. and Verma, M. (1990) 'Primary structure of the non-transcribed spacer region and flanking sequences of the ribosomal DNA of Neurospora crassa and comparison with other organisms', *Biochemical and Biophysical Research Communications,* 170(1), 187-193.

[134] Dykstra, M. A., Friedman, L. and Murphy, J. W. (1977) 'Capsule size of *Cryptococcus* neoformans: control and relationship to virulence', *Infect Immun,* 16(1), 129-35.

[135] Ellerbroek, P. M., Lefeber, D. J., van Vegbel, R., Scharringa, J., Brouwer, E., Gerwig, G. J., Janbon, G., Hoepelman, A. I. M. and Coenjaerts, F. E. J. (2004) 'O-acetylation of cryptococcal capsular glucuronoxylomannan is essential for interference with neutrophil migration', *Journal of Immunology,* 173(12), 7513-7520.

[136] Ellis, D. and Pfeiffer, T. (1992) 'The ecology of *Cryptococcus neoformans*', *European Journal of Epidemiology,* 8(3), 321-325.

[137] Ellis, D. H. and Pfeiffer, T. J. (1990) 'Natural habitat of *Cryptococcus neoformans* var. *gattii*', *Journal of Clinical Microbiology,* 28(7), 1642-1644.

[138] Emmons, C. (1962) 'Natural occurrence of opportunistic fungi', *Lab. Invest.,* 11, 1026-1032.

[139] Enright, M. C. and Spratt, B. G. (1999) 'Multilocus sequence typing', *Trends Microbiol,* 7(12), 482-7.

[140] Erdman, S. and Snyder, M. (2001) 'A filamentous growth response mediated by the yeast mating pathway', *Genetics,* 159(3), 919-928.

[141] Esposto, M. C., Cogliati, M., Tortorano, A. M. and Viviani, M. A. (2004) 'Determination of *Cryptococcus neoformans* var. neoformans mating type by multiplex PCR', *Clinical Microbiology and Infection,* 10(12), 1092-1094.

[142] Espy, M. J., Uhl, J. R., Sloan, L. A., Buckwalter, S. P., Jones, M. F., Vetter, E. A., Yao, J. D. C., Wengenack, N. L., Rosenblatt, J. E., Cockerill, F. R. and Smith, T. F. (2006) 'Real-time PCR in clinical microbiology: Applications for routine laboratory testing (vol 19, pg 165, 2006)', *Clinical Microbiology Reviews,* 19(3), 595-595.

[143] Estabrook, G. F., and L. Landrum. (1975) 'A simple test for the possible simultaneous evolutionary divergence of two amino acid positions', *Taxon,* 24, 609-613.

[144] Eswaran, V., Harpending, H. and Rogers, A. (2005) 'Genomics refutes an exclusively African origin of humans', *Journal of Human Evolution*, 1-18.

[145] Evans, E. (1949) 'An immunologic comparison of twelve strains of *Cryptococcus neoformans* (*Torula histolytica*). ', *Proc. Soc. Exptl. Biol.Med.,* 71, 644-66.

[146] Evans, E. E. (1950) 'The antigenic composition of *Cryptococcus neoformans*. I. A serologic classification by means of the capsular and agglutination reactions', *Journal of immunology (Baltimore, Md. : 1950),* 64(5), 423-30.

[147] Evans, E. E. and Kessel, J. F. (1951) 'The antigenic composition of *Cryptococcus* neoformans. II. Serologic studies with the capsular polysaccharide', *Journal of immunology (Baltimore, Md. : 1950),* 67(2), 109-14.

[148] Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992) 'Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes - Application to Human Mitochondrial-DNA Restriction Data', *Genetics,* 131(2), 479-491.

[149] Falconer I. R., Beresford A. M., Runnegar M. T. C. (1983). 'Evidence of liver damage by toxin from a bloom of the blue-green alga *Microcystis aeruginosa. Med. J. Aust.*, 1:511-514

[150] Fan, M., Chen, L. C., Ragan, M. A., Gutell, R. R., Warner, J. R., Currie, B. P. and Casadevall, A. (1995) 'The 5S rRNA and the rRNA intergenic spacer of the two varieties of *Cryptococcus neoformans*', *Journal of Medical and Veterinary Mycology,* 33(4), 215-221.

[151] Farrer, R. A., Weinert, L. A., Bielby, J., Garner, T. W. J., Balloux, F., Clare, F., Bosch, J., Cunningham, A. A., Weldon, C., du Preez, L. H., Anderson, L., Kosakovsky Pond, S. L., Shahar-Golan, R., Henk, D. A. and Fisher, M. C. 'Multiple emergences of genetically diverse amphibian-infecting chytrids include a globalised hypervirulent recombinant lineage', *PNAS*, in press.

[152] Fasoli, M. O. F., Kerridge, D. and Ryley, J. F. (1990) 'Pathogenicity of 5-fluorocytosine resistant strains of Candida albicans', *Journal of Medical and Veterinary Mycology,* 28(1), 27-34.

[153] Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P. and Spratt, B. G. (2004) 'eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data', *Journal of Bacteriology,* 186(5), 1518-1530.

[154] Feldmesser, M., Kress, Y., Novikoff, P. and Casadevall, A. (2000) '*Cryptococcus* neoformans is a facultative intracellular pathogen in murine pulmonary infection', *Infection and Immunity,* 68(7), 4225-4237.

[155] Felsenstein, J. (1985) 'Confidence-Limits on Phylogenies - an Approach Using the Bootstrap', *Evolution,* 39(4), 783-791.

[156] Feng, B., Wang, X., Hauser, M., Kaufmann, S., Jentsch, S., Haase, G., Becker, J. M. and Szaniszlo, P. J. (2001) 'Molecular cloning and characterization of WdPKS1, a gene involved in dihydroxynaphthalene melanin biosynthesis and virulence in Wangiella (Exophiala) dermatitidis', *Infection and Immunity,* 69(3), 1781-1794.

[157] Fereshteh, F., GS Bulmer, JR Tacker (1970) '*Cryptococcus neoformans* IV. The not-so-encapsulated yeast', *Infection and Immunity,* 1(6), 526-531.

[158] Finlay, B. B. and Falkow, S. (1989) 'Common themes in microbial pathogenicity', *Microbiological Reviews,* 53(2), 210-230.

[159] Fischer, G., Braun, S., Thissen, R. and Dott, W. (2006) 'FT-IR spectroscopy as a tool for rapid identification and intra-species characterization of airborne filamentous fungi', *Journal of Microbiological Methods,* 64(1), 63-77.

[160] Fisher, M. C. (2008) 'Molecular toolkit unlocks life cycle of the panzootic amphibian pathogen Batrachochytrium dendrobatidis', *Proceedings of the National Academy of Sciences of the United States of America,* 105(45), 17209-17210.

[161] Fisher M. C., Garner T. W. J., Walker S. F. (2009) 'Global emergence of *Batrochochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host.' *Annu. Rev Microbiol.*, 63, 291-310.

[162] Fortes, S. T., Lazera, M. S., Nishikawa, M. M., Macedo, R. C. L. and Wanke, B. (2001) 'First isolation of *Cryptococcus neoformans* var. *gattii* from a native jungle tree in the Brazilian Amazon rainforest', *Mycoses,* 44(5), 137-140.

[163] Franzot, S. P., Mukherjee, J., Cherniak, R., Chen, L. C., Hamdan, J. S. and Casadevall, A. (1998) 'Microevolution of a standard strain of *Cryptococcus neoformans* resulting in differences in virulence and other phenotypes', *Infection and Immunity,* 66(1), 89-97.

[164] Franzot, S. P., Salkin, I. F. and Casadevall, A. (1999) '*Cryptococcus neoformans* var. grubii: Separate varietal status for *Cryptococcus neoformans* serotype A isolates', *Journal of Clinical Microbiology,* 37(3), 838-840.

[165] Fraser, J. A., Giles, S. S., Wenink, E. C., Geunes-Boyer, S. G., Wright, J. R., Diezmann, S., Allen, A., Stajich, J. E., Dietrich, F. S., Perfect, J. R. and Heitman, J. (2005) 'Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak', *Nature,* 437(7063), 1360-4.

[166] Fries B. C., Taborda C. P., Serfass E., Casadevall A. (2001). 'Phenotypic switching of *Cryptococcus neoformans* occurs in vivo and influences the outcome of infection, *J Clin Invest*, 108(11):1639-48.

[167] Fromtling, R. A., Abruzzo, G. K. and Ruiz, A. (1989) 'Virulence and antifungal susceptibility of environmental and clinical isolates of *Cryptococcus neoformans* from Puerto Rico', *Mycopathologia,* 106(3), 163-166.

[168] Fromtling, R. A., Shadomy, H. J. and Jacobson, E. S. (1982) 'Decreased virulence in stable, acapsular mutants of *Cryptococcus neoformans*', *Mycopathologia,* 79(1), 23-29.

[169] Fry W., Goodwin S. B. (1997) 'Resurgence of the Irish potato famine fungus', *Bioscience*, 47, 363 - 371

[170] Gagen, S. J. and Ratcliffe, N. A. (1976) 'Studies on the in vivo cellular reactions and fate of injected bacteria in Galleria mellonella and Pieris brassicae larvae', *Journal of Invertebrate Pathology,* 28(1), 17-24.

[171] Ganendren R., Widmer F., Singhal V., Wilson C., Sorrell T., Wright L. (2004) 'In vitro antifungal activities of inhibitors of phospholipases from the fungal pathogen *Cryptococcus neoformans*. *Antimicrob Agents Chemother.*, 48(5): 1561-1569. doi: 10.1128/AAC.48.5.1561-1569.2004.

[172] Ganley, A. R. D. and Scott, B. (1998) 'Extraordinary ribosomal spacer length heterogeneity in a neotyphodium endophyte hybrid: Implications for concerted evolution', *Genetics,* 150(4), 1625-1637.

[173] Garcia-Rivera, J., Chang, Y. C., Kwon-Chung, K. J. and Casadevall, A. (2004) '*Cryptococcus* neoformans CAP59 (or cap59p) is involved in the extracellular trafficking of capsular glucuronoxylomannan', *Eukaryotic Cell,* 3(2), 385-392.

[174] Gardes, M. and Bruns, T. D. (1993) 'ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts', *Molecular Ecology,* 2(2), 113-118.

[175] Geiser D. M. Arnold M. L., Timberlake W. E. (1994) 'Sexual origins of British *Aspergillus nidulans* isolates.', *Proc Natl Acad Sci U S A*, 91(6):2349-52.

[176] Ghannoum, M. A. (2000) 'Potential role of phospholipases in virulence and fungal pathogenesis', *Clinical Microbiology Reviews,* 13(1), 122-+.

[177] Gil-Lamaignere, C., Roilides, E., Hacker, J. and Muller, F. M. C. (2003) 'Molecular typing for fungi - a critical review of the possibilities and limitations of currently and future methods', *Clinical Microbiology and Infection,* 9(3), 172-185.

[178] Girardin, H., Latge, J. P., Srikantha, T., Morrow, B. and Soll, D. R. (1993) 'Development of DNA probes for fingerprinting *Aspergillus fumigatus*', *Journal of Clinical Microbiology,* 31(6), 1547-1554.

[179] Glass, N. L. and Nelson, M. A. (1994) 'Mating-type genes in mycelial ascomycetes', *The Mycota, I; Growth, differentiation and sexuality*, 295-306.

[180] Goddard, M. R., Godfray, H. C. J. and Burt, A. (2005) 'Sex increases the efficacy of natural selection in experimental yeast populations', *Nature,* 434(7033), 636-640.

[181] Goldman, D. L., Fries, B. C., Franzot, S. P., Montella, L. and Casadevall, A. (1998) 'Phenotypic switching in the human pathogenic fungus *Cryptococcus neoformans* is associated with changes in virulence and pulmonary inflammatory response in rodents', *Proceedings of the National Academy of Sciences of the United States of America,* 95(25), 14967-14972.

[182] Goodwin, S. B., Cohen, B. A. and Fry, W. E. (1994) 'Panglobal distribution of a single clonal lineage of the Irish potato famine fungus', *Proceedings of the National Academy of Sciences of the United States of America,* 91(24), 11591-11595.

[183] Goodwin S. B., Cohen B. A., Deahl K. L., Fry, W. E. (1994b) 1Migration from Northern Mexico as the probable cause of recent genetic changes in populations of *Phytophthora infestans* in the United States and Canada.' *Phytopathology*, 84, 553558

[184] Gordon, S. B., Walsh, A. L., Chaponda, M., Gordon, M. A., Soko, D., Mbwvinji, M., Molyneux, M. E. and Read, R. C. (2000) 'Bacterial meningitis in Malawian adults: Pneumococcal disease is common, severe, and seasonal', *Clinical Infectious Diseases,* 31(1), 53-57.

[185] Granados, D. P. and Castaneda, E. (2005) 'Isolation and characterization of *Cryptococcus* neoformans varieties recovered from natural sources in Bogota, Colombia, and study of ecological conditions in the area', *Microbial Ecology,* 49(2), 282-290.

[186] Granados, D. P. and Castaneda, E. (2006) 'Influence of climatic conditions on the isolation of members of the *Cryptococcus neoformans* species complex from trees in Colombia from 1992-2004', *Fems Yeast Research,* 6(4), 636-644.

[187] Granger, D. L., Perfect, J. R. and Durack, D. T. (1985) 'Virulence of *Cryptococcus*-Neoformans - Regulation of Capsule Synthesis by Carbon-Dioxide', *Journal of Clinical Investigation,* 76(2), 508-516.

[188] Graser Y., Volovsek M., Arrington J., Schonian G., Presber W., Mitchell T. G., Vilgalys R. (1996) 'Molecular markers reveal that population structure of the human pathogen *Candida albicans* exhibits both clonality and recombination. *Proc Natl Acad Sci U S A*., 93(22): 12473-12477.

[189] Graybill, J. R., Sobel, J., Saag, M., van der Horst, C., Powderly, W., Cloud, G., Riser, L., Hamill, R., Dismukes, W., Grp, N. M. S. and Grp, A. C. T. (2000) 'Diagnosis and management of increased intracranial pressure in patients with AIDS and cryptococcal meningitis', *Clinical Infectious Diseases,* 30(1), 47-54.

[190] Grigg M. E., Bonnefoy S., Hehl A. B., Suzuki Y., Boothroyd J. C.(2001) 'Success and virulence in *Toxoplasmaas* the result of sexual recombination between two distinct ancestries'. *Science* 294: 161-165.

[191] Grimberg, B. and Zeyl, C. (2005) 'The effects of sex and mutation rate on adaptation in test tubes and to mouse hosts by Saccharomyces cerevisiae', *Evolution,* 59(2), 431-438.

[192] Gross, N. T., Guerrero, O. M., Chinchilla, M. and Jarstrand-Hall, C. (2006) 'Trypanosoma lewisi-induced immuno suppression: The effects on alveolar macrophage activities against *Cryptococcus neoformans*', *Experimental Parasitology,* 113(4), 262-266.

[193] *Grzimek's animal life encyclopedia* (2004).

[194] Gupta, A. K., Kohli, Y. and Summerbell, R. C. (2000) 'Molecular differentiation of seven Malassezia species', *Journal of Clinical Microbiology,* 38(5), 1869-1875.

[195] Hakim, J. G., Gangaidzo, I. T., Heyderman, R. S., Mielke, J., Mushangi, E., Taziwa, A., Robertson, V. J., Musvaire, P. and Mason, P. R. (2000) 'Impact of HIV infection on meningitis in Harare, Zimbabwe: a prospective study of 406 predominantly adult patients', *Aids,* 14(10), 1401-1407.

[196] Halliday, C. L., Bui, T., Krockenberger, M., Malik, R., Ellis, D. H. and Carter, D. A. (1999) 'Presence of alpha and a mating types in environmental and clinical collections of *Cryptococcus neoformans* var. *gattii* strains from Australia', *Journal of Clinical Microbiology,* 37(9), 2920-2926.

[197] Hanenberg, R. S., Rojanapithayakorn, W., Kunasol, P. and Sokal, D. C. (1994) 'Impact of Thailand's HIV-control programme as indicated by the decline of sexually transmitted diseases', *Lancet,* 344(8917), 243-245.

[198] Hasenclever, H. F. and Emmons, C. W. (1963) 'The prevalence and mouse virulence of *Cryptococcus neoformans* strains isolated from urban areas', *American journal of hygiene,* 78, 227-31.

[199] Hein, M. (2008) *Redefining AIDS in Asia: Crafting an Effective Response* India.

[200] Heitman, J. (2006) 'Sexual reproduction and the evolution of microbial pathogens', *Current Biology,* 16(17), R711-R725.

[201] Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H. and Vogt, P. H. (1997) 'Multiplex PCR: Critical parameters and step-by-step protocol', *Biotechniques,* 23(3), 504-511.

[202] Heyderman, R. S., Gangaidzo, I. T., Hakim, J. G., Mielke, J., Taziwa, A., Musvaire, P., Robertson, V. J. and Mason, P. R. (1998) 'Cryptococcal meningitis in human immunodeficiency virus-infected patients in Harare, Zimbabwe', *Clinical Infectious Diseases,* 26(2), 284-289.

Hiremath, S. S., Chowdhary, A., Kowshik, T., Randhawa, H. S., Sun, S. and Xu, J. P. (2008) 'Long-distance dispersal and recombination in environmental populations of *Cryptococcus neoformans* var. grubii from India', *Microbiology-Sgm,* 154, 1513-1524.

[203] Hole F. (1994). In Bunney S, ed. 1994. *Origins of Agriculture*. Cambridge: Cambridge University Press, pp. 373-379.

[204] Holsinger, K. E. (2010) 'Patterns of selection on nucleotide polymorphisms'. available:http://darwin.eeb.uconn.edu/eeb348/lecturenotes/molevol-polymorphism.pdf [accessed 11 February, 2011].',

[205] Horta, J. A., Staats, C. C., Casali, A. K., Ribeiro, A. M., Schrank, I. S., Schrank, A. and Vainstein, M. H. (2002) 'Epidemiological aspects of clinical and environmental *Cryptococcus neoformans* isolates in the Brazilian state Rio Grande do Sul', *Medical Mycology,* 40(6), 565-571.

[206] Hotopp, J. C. D., Lin, M., Madupu, R., Crabtree, J., Angiuoli, S. V., Eisen, J. A., Seshadri, R., Ren, Q., Wu, M., Utterback, T. R., Smith, S., Lewis, M., Zhang, K. C., Niu, H., Lin, Q., Ohashi, N., Zhi, N., Nelson, W., Brinkac, L. M., Dodson, R. J., Rosovitz, M. J., Sundaram, J., Daugherty, S. C., Davidsen, T., Durkin, A. S., Gwinn, M., Haft, D. H., Selengut, J. D., Sullivan, S. A., Zafar, N., Zhou, L., Benahmed, F., Forberger, H., Halpin, R., Mulligan, S., Robinson, J., White, O., Rikihisa, Y. and Tettelin, H. (2006) 'Comparative genomics of emerging human ehrlichiosis agents (vol 2, pg 2, 2006)', *Plos Genetics,* 2(12), 2176-2176.

[207] Howe D. K., Sibley L. D. (1995) '*Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease', *Journal of Infectious Diseases,* 172(6), 1561-1566.

[208] Hsu, M. C., Chen, K. W., Lo, H. J., Chen, Y. C., Liao, M. H., Lin, Y. H. and Li, S. Y. (2003) 'Species identification of medically important fungi by use of real-time LightCycler PCR', *Journal of Medical Microbiology,* 52(12), 1071-1076.

[209] Hudson, R. R. (2000) 'A new statistic for detecting genetic differentiation', *Genetics,* 155(4), 2011-2014.

[210] Hudson, R. R., Boos, D. D. and Kaplan, N. L. (1992) 'A statistical test for detecting geographic subdivision', *Molecular Biology and Evolution,* 9(1), 138-151.

[211] Hudson, R. R. and Kaplan, N. L. (1985) 'Statistical properties of the number of recombination events in the history of a sample of DNA sequences', *Genetics,* 111(1), 147-164.

[212] Hull, C. M. and Heitman, J. (2002) 'Genetics of *Cryptococcus neoformans*', *Annual Review of Genetics,* 36, 557-615.

[213] Huston, S. M. and Mody, C. H. (2009) 'Cryptococcosis: An Emerging Respiratory Mycosis', *Clinics in Chest Medicine,* 30(2), 253-+.

[214] Ibrahim, A. S., Filler, S. G., Alcouloumre, M. S., Kozel, T. R., Edwards, J. E. and Ghannoum, M. A. (1995) 'Adherence to and Damage of Endothelial-Cells by *Cryptococcus*-Neoformans in-Vitro - Role of the Capsule', *Infection and Immunity,* 63(11), 4368-4374.

[215] Idnurm, A., Bahn, Y. S., Nielsen, K., Lin, X. R., Fraser, J. A. and Heitman, J. (2005) 'Deciphering the model pathogenic fungus *Cryptococcus neoformans*', *Nature Reviews Microbiology,* 3(10), 753-764.

[216] Idnurm, A., Reedy, J. L., Nussbaum, J. C. and Heitman, J. (2004) '*Cryptococcus* neoformans virulence gene discovery through insertional mutagenesis', *Eukaryotic Cell,* 3(2), 420-429.

[217] Illnait-Zaragozi, M. T., Martinez-Machin, G. F., Fernandez-Andreu, C. M., Boekhout, T., Meis, J. F. and Klaassen, C. H. W. (2010) 'Microsatellite Typing of Clinical and Environmental *Cryptococcus neoformans* var. grubii Isolates from Cuba Shows Multiple Genetic Lineages', *Plos One,* 5(2).

[218] Imwidthaya, P. and Poungvarin, N. (2000) 'Cryptococcosis in AIDS', *Postgrad Med J,* 76(892), 85-8.

[219] Izzo, A., Agbowo, J. and Bruns, T. D. (2005) 'Detection of plot-level changes in ectomycorrhizal communities across years in an old-growth mixed-conifer forest', *New Phytologist,* 166(2), 619-630.

[220] Jochimsen E. M., Carmichael W. W., An J. S., Cardo D. M. , Cookson S. T., Holmes C. E. M., Antunes M. B. D. , DeMelo F. D. A. , Lyra T. M.,Bar reto V. S. T, Azevedo S. M. F. O., and Jarvis W. R. (1998) 'Liver failure and death after exposure to microcystins at a hemodialysis center in Brazil'. *N. Engl. J. Med.*, 338:873-878.

[221] Jacobson, E. S., Ayers, D. J., Harrell, A. C. and Nicholas, C. C. (1982) 'Genetic and phenotypic characterization of capsule mutants of *Cryptococcus neoformans*', *Journal of Bacteriology,* 150(3), 1292-1296.

[222] Jacobson, E. S. and Emery, H. S. (1991) 'Catecholamine uptake, melanization, and oxygen toxicity in *Cryptococcus neoformans*', *Journal of Bacteriology,* 173(1), 401-403.

[223] Jain, N., Wickes, B. L., Keller, S. A., Fu, J., Casadevall, A., Jain, P., Ragan, M. A., Banerjee, U. and Fries, B. C. (2005) 'Molecular epidemiology of clinical *Cryptococcus neoformans* strains from India', *Journal of Clinical Microbiology,* 43(11), 5733-5742.

[224] Jarvis, J. N. and Harrison, T. S. (2007) 'HIV-associated cryptococcal meningitis', *Aids,* 21(16), 2119-29.

[225] Jarvis, J. N., Meintjes, G., Williams, A., Brown, Y., Crede, T. and Harrison, T. S. (2010) 'Adult meningitis in a setting of high HIV and TB prevalence: findings from 4961 suspected cases', *Bmc Infectious Diseases,* 10.

[226] Johnson, J. K., Arduino, S. M., Stine, O. C., Johnson, J. A. and Harris, A. D. (2007) 'Multilocus sequence typing compared to pulsed-field gel electrophoresis for molecular typing of Pseudomonas aeruginosa', *J Clin Microbiol,* 45(11), 3707-12.

[227] Johnston, R. (1992) *Birds of North America,* Philadelphia: American Ornithologist's Union and Academy of Natural Sciences Of Philadelphia.

[228] Joly, V., Saintjulien, L., Carbon, C. and Yeni, P. (1994) 'In vivo activity of interferon-gamma in combination with amphotericin B in the treatment of experimental cryptococcosis', *Journal of Infectious Diseases,* 170(5), 1331-1334.

[229] Jombart, T. (2008) 'adegenet: a R package for the multivariate analysis of genetic markers', *Bioinformatics,* 24(11), 1403-1405.

[230] Jones, S. (2007) 'New way out for *Cryptococcus*', *Nat Rev Micro,* 5(1), 3-3.

[231] Kaltz, O. and Bell, G. (2002) 'The ecology and genetics of fitness in Chlamydomonas. XII. Repeated sexual episodes increase rates of adaptation to novel environments', *Evolution,* 56(9), 1743-1753.

[232] Kang, Y. Q., Tanaka, H., Moretti, M. L. and Mikami, Y. (2009) 'New ITS genotype of *Cryptococcus gattii* isolated from an AIDS patient in Brazil', *Microbiology and Immunology,* 53(2), 112-116.

[233] Kapenda, K., Komichelo, K., Swinne, D. and Vandepitte, J. (1987) 'Meningitis due to *Cryptococcus neoformans* biovar *gattii* in a Zairean AIDS patient', *European Journal of Clinical Microbiology & Infectious Diseases,* 6(3), 320-321.

[234] Kaplan, J. E., Hanson, D., Dworkin, M. S., Frederick, T., Bertolli, J., Lindegren, M. L., Holmberg, S. and Jones, J. L. (2000) 'Epidemiology of human immunodeficiency virus-associated opportunistic infect-ions in the United States in the era of highly active antiretroviral therapy', *Clinical Infectious Diseases,* 30, S5-S14.

[235] Kaplan E.L. and Meier P. (1958) 'Nonparametric estimation from incomplete observations', *Journal of the American Statistical Association*, 53, 457-481 (1958).

[236] Karim, A. S. S., Churchyard, G. J., Abdool Karim, Q. and Lawn, S. D. (2009) 'HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response', *Lancet,* 374(9693), 921-33.

[237] Karos, M., Chang, Y. C., McClelland, C. M., Clarke, D. L., Fu, J., Wickes, B. L. and Kwon-Chung, K. J. (2000) 'Mapping of the *Cryptococcus neoformans MAT* locus: presence of mating type-specific mitogen-activated protein kinase cascade homologs', *J Bacteriol,* 182, 6222-6227.

[238] Karstaedt, A. S., Crewe-Brown, H. H. and Dromer, F. (2002) 'Cryptococcal meningitis caused by *Cryptococcus neoformans* var. *gattii*, serotype C, in AIDS patients in Soweto, South Africa', *Medical Mycology,* 40(1), 7-11.

[239] Kasuga, T., White, T. J., Koenig, G., McEwen, J., Restrepo, A., Castaneda, E., Lacaz, C. D., Heins-Vaccari, E. M., De Freitas, R. S., Zancope-Oliveira, R. M., Qin, Z. Y., Negroni, R., Carter, D. A., Mikami, Y., Tamura, M., Taylor, M. L., Miller, G. F., Poonwan, N. and Taylor, J. W. (2003) 'Phylogeography of the fungal pathogen Histoplasma capsulatum', *Molecular Ecology,* 12(12), 3383-3401.

[240] Kasuga, T., White, T. J. and Taylor, J. W. (2002) 'Estimation of nucleotide substitution rates in eurotiomycete fungi', *Molecular Biology and Evolution,* 19(12), 2318-2324.

[241] Katsu, M., Kidd, S., Ando, A., Moretti-Branchini, M. L., Mikami, Y., Nishimura, K. and Meyer, W. (2004) 'The internal transcribed spacers and 5.8S rRNA gene show extensive diversity among isolates of the *Cryptococcus neoformans* species complex', *Fems Yeast Research,* 4(4-5), 377-388.

[242] Kavanagh K., Fallon J. P. (2010) '*Galleria mellonella* larvae as models for studying fungal virulence. Vol 24: 1-2, pp. 79-83.

[243] Kavanagh, K. and Reeves, E. P. (2004) 'Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens', *Fems Microbiology Reviews,* 28(1), 101-112.

[244] Kavanaugh, L. A., Fraser, J. A. and Dietrich, F. S. (2006) 'Recent evolution of the human pathogen *Cryptococcus neoformans* by intervarietal transfer of a 14-gene fragment', *Molecular Biology and Evolution,* 23(10), 1879-1890.

[245] Kidd S. E., Chow Y., Mak S., Bach P.J., Chen H., Hingston A.O., Kronstad J. W., Bartlett K. H. (2007) 'Characterization of environmental sources of the human and animal pathogen *Cryptococcus gattii* in British Columbia, Canada, and the Pacifc Northwest of the United States', *Applied and Environmental Microbiology*, vol. 73: 5, pp. 1433-1443.

[246] Kidd, S. E., Hagen, F., Tscharke, R. L., Huynh, M., Bartlett, K. H., Fyfe, M., MacDougall, L., Boekhout, T., Kwon-Chung, K. J. and Meyer, W. (2004) 'A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada)', *Proceedings of the National Academy of Sciences of the United States of America,* 101(49), 17258-17263.

[247] Kiertiburanakul, S., Wirojtananugoon, S., Pracharktam, R. and Sungkanuparph, S. (2006) 'Cryptococcosis in human immunodeficiency virus-negative patients', *International Journal of Infectious Diseases,* 10(1), 72-78.

[248] King, H., Aubert, R. and Herman, W. (1998) 'Global burden of diabetes, 1995-2025 - Prevalence, numerical estimates, and projections', *Diabetes Care*, 1414-1431.

[249] King, J. and Dasgupta, A. (2005) 'Cryptococcosis', *Living Rev. Relativity*, 8available: `http://www.emedicine.com/med/topic482.htm` [accessed 23/12/07].

[250] Kirsch, D. R. and Whitney, R. R. (1991) 'Pathogenicity of Candida albicans auxotrophic mutants in experimental infections', *Infection and Immunity,* 59(9), 3297-3300.

[251] Klamer, M., Roberts, M. S., Levine, L. H., Drake, B. G. and Garland, J. L. (2002) 'Influence of elevated CO2 on the fungal community in a coastal scrub oak forest soil investigated with terminal-restriction fragment length polymorphism analysis', *Applied and Environmental Microbiology,* 68(9), 4370-4376.

[252] Klugman, K. P. (2003) 'The role of clonality in the global spread of fluoroquinolone-resistant bacteria', *Clinical Infectious Diseases,* 36(6), 783-785.

[253] Kolar, C. S. and Lodge, D. M. (2001) 'Progress in invasion biology: predicting invaders', *Trends in Ecology & Evolution,* 16(4), 199-204.

[254] Kondrashov A. S. (1988) 'Deleterious mutations and the evolution of sexual reproduction'. *Nature*, 336, 435 - 440; doi:10.1038/336435a0

[255] Kondrashov A. S. (1993) 'Classification of hypotheses on the advantage of amphimixis'. *J Hered*, 1993;84:372-387.

[256] Koufopanou, V., Burt, A. and Taylor, J. W. (1997) 'Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus Coccidioides immitis', *Proceedings of the National Academy of Sciences of the United States of America,* 94(10), 5478-5482.

[257] Kozel, T. R. and Cazin, J. (1971) 'Nonencapsulated variant of *Cryptococcus neoformans*. I. Virulence studies and characterization of soluble polysaccharide.', *Infection and Immunity,* 3(2), 287-&.

[258] Kozel, T. R. and Gotschlich, E. C. (1982) 'The capsule of *Cryptococcus neoformans* passively inhibits phagocytosis of the yeast by macrophages', *Journal of Immunology,* 129(4), 1675-1680.

[259] Kozel, T. R., Gulley, W. F. and Cazin, J. (1977) 'Immune-Response to *Cryptococcus*-Neoformans Soluble Polysaccharide - Immunological-Unresponsiveness', *Infection and Immunity,* 18(3), 701-707.

[260] Kozel, T. R., Highison, B. and Stratton, C. J. (1984) 'Localization on Encapsulated *Cryptococcus*-Neoformans of Serum Components Opsonic for Phagocytosis by Macrophages and Neutrophils', *Infection and Immunity,* 43(2), 574-579.

[261] Kozel, T. R. and Mastroianni, R. P. (1976) 'Inhibition of phagocytosis by cryptococcal polysaccharide: dissociation of the attachment and ingestion phases of phagocytosis', *Infection and Immunity,* 14(1), 62-67.

[262] Krockenberger, M. B., Canfield, P. J. and Malik, R. (2002) '*Cryptococcus neoformans* in the koala (Phascolarctos cinereus): colonization by C-n. var. *gattii* and investigation of environmental sources', *Medical Mycology,* 40(3), 263-272.

Koufopanou V., Burt A., Taylor J. W. (1997). 'Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*.' *Proc Natl Acad Sci U S A*, 13;94(10):5478-82.

[263] Kuroki M., Phichaichumpon C., Yasuoka A., Chiranairadul P., Chosa T., Sirinirund P., Miyazaki T., Kakeya H., Higashiyama Y., Miyazaki Y., Ishida Y., Kohno S. (2004) 'Environmental isolation of *Cryptococcus neoformans* from an endemic region of HIV-associated cryptococcosis in Thailand'. *Yeast*, 2004;21:809-12.

[264] Kurtzman, C. P. and Robnett, C. J. (1997) 'Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5' end of the large-subunit (26S) ribosomal DNA gene', *Journal of Clinical Microbiology,* 35(5), 1216-1223.

[265] Kwon-Chung, K. J. (1974) 'Genetics of fungi pathogenic for man', *CRC Critical Reviews in Microbiology,* 3(2), 115-133.

[266] Kwon-Chung, K. J. (1975) 'A new genus, filobasidiella, the perfect state of *Cryptococcus* neoformans', *Mycologia,* 67(6), 1197-200.

[267] Kwon-Chung, K. J. (1976) 'A new species of Filobasidiella, the sexual state of *Cryptococcus neoformans* B and C serotypes', *Mycologia,* 68(4), 943-6.

[268] Kwon-Chung, K. J. and Bennett, J. E. (1978) 'Distribution of alpha and alpha mating types of *Cryptococcus neoformans* among natural and clinical isolates', *Am J Epidemiol,* 108(4), 337-40.

[269] Kwon-Chung KJ, B. J. (1992) *Medical Mycology,* Philadelphia: Lea & Febiger.

[270] Kwon-Chung, K. J., Boekhout, T., Fell, J. W. and Diaz, M. (2002) '(1557) Proposal to conserve the name *Cryptococcus gattii* against C. hondurianus and C. bacillisporus (Basidiomycota, Hymenomycetes, Tremellomycetidae)', *Taxon,* 51(4), 804-806.

[271] Kwon-Chung, K. J., Edman, J. C. and Wickes, B. L. (1992) 'Genetic association of mating types and virulence in *Cryptococcus neoformans*', *Infect Immun,* 60(2), 602-5.

[272] Kwon-Chung, K. J. and Varma, A. (2006) 'Do major species concepts support one, two or more species within *Cryptococcus neoformans*?', *FEMS Yeast Res,* 6(4), 574-87.

[273] Kwon-Chung, K. J. and Wickes, B. L. (2006) 'The conversion from classical studies in fungal pathogenesis to the molecular era', *Molecular Principles of Fungal Pathogenesis-Print*, 49-+.

[274] Kwon-Chung.Kj, Weeks, R. J. and Larsh, H. W. (1974) 'Studies on Emmonsiella capsulata. (Histoplasma capsulatum). II. Distribution of the two mating types in 13 endemic states of the United States.', *American Journal of Epidemiology,* 99(1), 44-49.

[275] Kwon-Chung, K. J. (1975) 'Perfect state (Emmonsiella capsulata) of the fungus causing large-form African histoplasmosis', *Mycologia,* 67(5), 980-990.

[276] Kwon-Chung, K. J. and Bennett, J. E. (1978) 'Distribution of Alpha and a Mating Types of *Cryptococcus*-Neoformans among Natural and Clinical Isolates', *American Journal of Epidemiology,* 108(4), 337-340.

[277] Kwon-Chung, K. J. and Bennett, J. E. (1984) 'High prevalence of *Cryptococcus neoformans* var. *gattii* in tropical and subtropical regions', *Zentralblatt Fur Bakteri-*

ologie Mikrobiologie Und Hygiene Series a-Medical Microbiology Infectious Diseases Virology Parasitology, 257(2), 213-218.

[278] Kwon-Chung, K. J., Bennett, J. E. and Rhodes, J. C. (1982a) 'Taxonomic studies on Filobasidiella species and their anamorphs', *Antonie Van Leeuwenhoek Journal of Microbiology,* 48(1), 25-38.

[279] Kwon-Chung, K. J., Polacheck, I. and Bennett, J. E. (1982b) 'Improved Diagnostic Medium for Separation of *Cryptococcus*-Neoformans Var Neoformans (Serotype-a and Serotype-D) and *Cryptococcus*-Neoformans Var Gattii (Serotype-B and Serotype-C)', *Journal of Clinical Microbiology,* 15(3), 535-537.

[280] Kwon-Chung, K. J., Polacheck, I. and Popkin, T. J. (1982c) 'Melanin-Lacking Mutants of *Cryptococcus*-Neoformans and Their Virulence for Mice', *Journal of Bacteriology,* 150(3), 1414-1421.

[281] Kwon-Chung, K. J. and Rhodes, J. C. (1986) 'Encapsulation and melanin formation as indicators of virulence in *Cryptococcus neoformans*', *Infection and Immunity,* 51(1), 218-223.

[282] Lachke, S. A., Srikantha, T., Tsai, L. K., Daniels, K. and Soll, D. R. (2000) 'Phenotypic switching in Candida glabrata involves phase-specific regulation of the metallothionein gene MT-II and the newly discovered hemolysin gene HLP', *Infection and Immunity,* 68(2), 884-895.

[283] Laloë, D., Jombart, T., Dufour, A. B. and Moazami-Goudarzi, K. (2007) 'Consensus genetic structuring and typological value of markers using multiple co-inertia analysis', *Genet Sel Evol.*, Sep-Oct, 39(5), 545-567.

[284] Land, K. (2002) 'Soil amoeba help to unravel fungal pathogenesis', *Trends in Microbiology,* 10(3), 116-116.

[285] Langfelder, K., Streibel, M., Jahn, B., Haase, G. and Brakhage, A. A. (2003) 'Biosynthesis of fungal melanins and their importance for human pathogenic fungi', *Fungal Genetics and Biology,* 38(2), 143-158.

[286] Larsen, R. A., Leal, M. A. E. and Chan, L. S. (1990) 'Fluconazole compared with amphotericin B plus flucytosine for cryptococcal meningitis in AIDS. A randomized trial', *Annals of Internal Medicine,* 113(3), 183-187.

[287] Lasker, B. A., Page, L. S., Lott, T. J. and Kobayashi, G. S. (1992) 'Isolation, characterization, and sequencing of Candida albicans repetitive element 2', *Gene,* 116(1), 51-57.

[288] Lasker, B. A., Page, L. S., Lott, T. J., Kobayashi, G. S. and Medoff, G. (1991) 'Characterization of CARE-1: Candida albicans repetitive element-1', *Gene,* 102(1), 45-50.

[289] Latge, J. P. (2001) 'The pathobiology of *Aspergillus fumigatus*', *Trends in Microbiology,* 9(8), 382-389.

[290] Lazera, M. S., Cavalcanti, M. A. S., Londero, A. T., Trilles, L., Nishikawa, M. M. and Wanke, B. (2000) 'Possible primary ecological niche of *Cryptococcus neoformans*', *Medical Mycology,* 38(5), 379-383.

[291] Leal, A. L., Faganello, J., Bassanesi, M. C. and Vainstein, M. H. (2008) '*Cryptococcus* species identification by multiplex PCR', *Medical Mycology,* 46(4), 377-383.

[292] Lehmann, P. F., Kemker, B. J., Hsiao, C. B. and Dev, S. (1989) 'Isoenzyme biotypes of Candida species', *Journal of Clinical Microbiology,* 27(11), 2514-2521.

[293] Lehmann, P. F., Lin, D. M. and Lasker, B. A. (1992) 'Genotypic identification and characterization of species and strains within the genus Candida by using random amplified polymorphic DNA', *Journal of Clinical Microbiology,* 30(12), 3249-3254.

[294] Lengeler, K. B., Cox, G. M. and Heitman, J. (2001) 'Serotype AD strains of *Cryptococcus* neoformans are diploid or aneuploid and are heterozygous at the mating-type locus', *Infection and Immunity,* 69(1), 115-122.

[295] Lengeler, K. B., Fox, D. S., Fraser, J. A., Allen, A., Forrester, K., Dietrich, F. S. and Heitman, J. (2002) 'Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes', *Eukaryotic Cell,* 1(5), 704-718.

[296] Lengeler, K. B., Wang, P., Cox, G. M., Perfect, J. R. and Heitman, J. (2000) 'Identification of the MATa mating-type locus of *Cryptococcus neoformans* reveals a serotype A MATa strain thought to have been extinct', *Proceedings of the National Academy of Sciences of the United States of America,* 97(26), 14455-14460.

[297] Levitz, S. (2008) 'The Immune Response', in *7th International Conference on Cryptococcus and Cryptococcosis*, Nagasaki, Japan, 1/2008, European Confederation of Medical Mycology,

[298] Levitz, S. M. (2001) 'Commentary - Does amoeboid reasoning explain the evolution and maintenance of virulence factors in *Cryptococcus neoformans*?', *Proceedings of the National Academy of Sciences of the United States of America,* 98(26), 14760-14762.

[299] Levitz, S. M. and Dibenedetto, D. J. (1989) 'Paradoxical Role of Capsule in Murine Bronchoalveolar Macrophage-Mediated Killing of *Cryptococcus*-Neoformans', *Journal of Immunology,* 142(2), 659-665.

[300] Levy, P. Y., Raoult, D., Rolain, J. M., Habib, G. and Reynaud-Gaubert, M. (2008) 'Pericardial effusion due to *Cryptococcus neoformans* in a patient with cystic fibrosis following lung transplantation', *International Journal of Infectious Diseases,* 12(4), 452-452.

[301] Li, W. H., Tanimura, M. and Sharp, P. M. (1987) 'An evaluation of the molecular clock hypothesis using mammalian DNA sequences', *Journal of Molecular Evolution,* 25(4), 330-342.

[302] Librado, P. and Rozas, J. (2009) 'DnaSP v5: a software for comprehensive analysis of DNA polymorphism data', *Bioinformatics,* 25(11), 1451-1452.

[303] Liebowitz, L. D., Koornhof, H. J., Barrett, M., Bracken, C., Davis, A., Fraser, N., Lezzi, M., Moraes, M. F., Palexas, G. N., Potgieter, D., Stratling, A. and Trumpelman, M. (1984) 'Bacterial meningitis in Johannesburg 1980-1982', *South African Medical Journal,* 66(18), 677-679.

[304] Lin, D. M., Lehmann, P. F., Hamory, B. H., Padhye, A. A., Durry, E., Pinner, R. W. and Lasker, B. A. (1995) 'Comparison of three typing methods for clinical and environmental isolates of *Aspergillus fumigatus*', *Journal of Clinical Microbiology,* 33(6), 1596-1601.

[305] Lin, X. and Heitman, J. (2006) 'The biology of the *Cryptococcus neoformans* species complex', *Annu Rev Microbiol,* 60, 69-105.

[306] Lin, X. R., Hull, C. M. and Heitman, J. (2005) 'Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*', *Nature,* 434(7036), 1017-1021.

[307] Lin, X. R., Litvintseva, A. P., Nielsen, K., Patel, S., Floyd, A., Mitchell, T. G. and Heitman, J. (2007) 'alphaADalpha hybrids of *Cryptococcus neoformans*: evidence of same-sex mating in nature and hybrid fitness', *PLoS Genetics,* 3(10), e186.

[308] Lin, X. R., Patel, S., Litvintseva, A. P., Floyd, A., Mitchell, T. G. and Heitman, J. (2009) 'Diploids in the *Cryptococcus neoformans* Serotype A Population Homozygous for the alpha Mating Type Originate via Unisexual Mating', *Plos Pathogens,* 5(1).

[309] Litvintseva, A. P., Carbone, I., Rossouw, J., Thakur, R., Govender, N. P. and Mitchell, T. G. (2011) 'Evidence that the Human Pathogenic Fungus *Cryptococcus neoformans* var. grubii May Have Evolved in Africa', *Plos One,* 6(5).

[310] Litvintseva, A. P., Kestenbaum, L., Vilgalys, R. and Mitchell, T. G. (2005a) 'Comparative analysis of environmental and clinical populations of *Cryptococcus neoformans*', *Journal of Clinical Microbiology,* 43(2), 556-564.

[311] Litvintseva, A. P., Lin, X., Templeton, I., Heitman, J. and Mitchell, T. G. (2007) 'Many globally isolated AD hybrid strains of *Cryptococcus neoformans* originated in Africa', *Plos Pathogens,* 3, 1109-1117.

[312] Litvintseva, A. P., Marra, R. E., Nielsen, K., Heitman, J., Vilgalys, R. and Mitchell, T. G. (2003) 'Evidence of sexual recombination among *Cryptococcus neoformans* serotype A isolates in sub-Saharan Africa', *Eukaryot Cell,* 2(6), 1162-8.

[313] Litvintseva, A. P. and Mitchell, T. G. (2009) 'Most Environmental Isolates of *Cryptococcus* neoformans var. grubii (Serotype A) Are Not Lethal for Mice', *Infection and Immunity,* 77(8), 3188-3195.

[314] Litvintseva, A. P., Thakur, R., Reller, L. B. and Mitchell, T. G. (2005b) 'Prevalence of clinical isolates of *Cryptococcus gattii* serotype C among patients with AIDS in sub-Saharan Africa', *Journal of Infectious Diseases,* 192(5), 888-892.

[315] Litvintseva, A. P., Thakur, R., Vilgalys, R. and Mitchell, T. G. (2006) 'Multilocus sequence typing reveals three genetic subpopulations of *Cryptococcus neoformans* var. grubii (Serotype A), including a unique population in Botswana', *Genetics,* 172(4), 2223-2238.

[316] Lockhart, S. R., Fritch, J. J., Meier, A. S., Schroppel, K., Srikantha, T., Galask, R. and Soll, D. R. (1995) 'Colonizing populations of Candida albicans are clonal in origin but undergo microevolution through C1 fragment reorganization as demonstrated by DNA fingerprinting and C1 sequencing', *Journal of Clinical Microbiology,* 33(6), 1501-1509.

[317] Longkrug, S. A., Fischer, K. J., Hysmith, R. M. and Ravdin, J. I. (1985) 'Phospholipase A enzymes of Entamoeba histolytica: description and subcellular localization', *Journal of Infectious Diseases,* 152(3), 536-541.

[318] Lonsdale, W. M. (1999) 'Global patterns of plant invasions and the concept of invasibility', *Ecology,* 80(5), 1522-1536.

[319] Lorenz, M. C. (2006) 'Studying fungal virulence by using genomics', *Molecular Principles of Fungal Pathogenesis*, 591-609.

[320] Lott, T. J., Burns, B. M., Zancope-Oliveira, R., Elie, C. M. and Reiss, E. (1998) 'Sequence analysis of the internal transcribed spacer 2 (ITS2) from yeast species within the genus Candida', *Current Microbiology,* 36(2), 63-69.

[321] Loyse, A., Wainwright, H., Jarvis, J. N., Bicanic, T., Rebe, K., Meintjes, G. and Harrison, T. S. (2010) 'Histopathology of the arachnoid granulations and brain in HIV-associated cryptococcal meningitis: correlation with cerebrospinal fluid pressure', *Aids,* 24(3), 405-410.

[322] Lu, H. Z., Zhou, Y. J., Yin, Y. K., Pan, X. Z. and Weng, X. H. (2005) 'Cryptococcal antigen test revisited: Significance for cryptococcal meningitis therapy monitoring in a tertiary Chinese hospital', *Journal of Clinical Microbiology,* 43(6), 2989-2990.

[323] Lutz, J. E., Clemons, K. V. and Stevens, D. A. (2000) 'Enhancement of antifungal chemotherapy by interferon-gamma in experimental systemic cryptococcosis', *Journal of Antimicrobial Chemotherapy,* 46(3), 437-442.

[324] Lynch M., Gabriel W. (1990) 'Mutation load and the survival of small populations'. *Evolution*, 44:1725-1737.

[325] Ma, H., Croudace, J., Lammas, D. and May, R. (2006) 'Expulsion of live pathogenic yeast by macrophages', *Current Biology,* (16), 2156-2160.

[326] Ma, H. and May, R. C. (2010) 'The fatal fungal outbreak on Vancouver Island is characterized by enhanced intracellular parasitism driven by mitochondrial regulation (vol 106, pg 12980, 2009)', *Virulence,* 1(3), 197-201.

[327] Ma, H. S., Croudace, J. E., Lammas, D. A. and May, R. C. (2007) 'Direct cell-to-cell spread of a pathogenic yeast', *Bmc Immunology,* 8.

[328] MacDougall, L., Kidd, S. E., Galanis, E., Mak, S., Leslie, M. J., Cieslak, P. R., Kronstad, J. W., Morshed, M. G. and Bartlett, K. H. (2007) 'Spread of *Cryptococcus gattii* in British Columbia, Canada, and detection in the Pacific Northwest, USA', *Emerging Infectious Diseases,* 13(1), 42-50.

[329] Madrenys, N., Devroey, C., Raeswuytack, C. and Torresrodriguez, J. M. (1993) 'Identification of the perfect state of *Cryptococcus neoformans* from 195 clinical isolates including 84 from AIDS patients', *Mycopathologia,* 123(2), 65-68.

[330] Magee, B. B., Dsouza, T. M. and Magee, P. T. (1987) 'Identification of Candida species by PCR and restriction fragment length polymorphism analysis of intergenic spacer regions of ribosomal DNA', *Journal of Bacteriology,* 169(4), 1639-1643.

[331] Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M. and Spratt, B. G. (1998) 'Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms', *Proc Natl Acad Sci U S A,* 95(6), 3140-5.

[332] Maiden, M. C. J. (2006) 'Multilocus sequence typing of bacteria', *Annual Review of Microbiology,* 60, 561-588.

[333] Makimura, K., Tamura, Y., Kudo, M., Uchida, K., Saito, H. and Yamaguchi, H. (2000) 'Species identification and strain typing of Malassezia species stock strains and clinical isolates based on the DNA sequences of nuclear ribosomal internal transcribed spacer 1 regions', *Journal of Medical Microbiology,* 49(1), 29-35.

[334] Manel, S., Schwartz, M. K., Luikart, G. and Taberlet, P. (2003) 'Landscape genetics: combining landscape ecology and population genetics', *Trends in Ecology and Evolution*, 18, 189-197.

[335] Mansour, M. K. and Levitz, S. M. (2002) 'Interactions of fungi with phagocytes', *Current Opinion in Microbiology,* 5(4), 359-365.

[336] Martinez, L. R., Garcia-Rivera, J. and Casadevall, A. (2001) '*Cryptococcus neoformans* var. neoformans (serotype D) strains are more susceptible to heat than C. neoformans var. grubii (serotype A) strains', *Journal of Clinical Microbiology,* 39(9), 3365-3367.

[337] McCarthy, K. (2007) *Cryptococcal Surveillance, South Africa, 2006*, 5.

[338] McCarthy, K. M., Morgan, J., Wannernuehler, K. A., Mirza, S. A., Gould, S. M., Mhlongo, N., Moeng, P., Maloba, B. R., Crewe-Brown, H. H., Brandt, M. E., Hajjeh, R. A. and Gauteng Cryptococcal, S. (2006) 'Population-based surveillance for cryptococcosis in an antiretroviral-naive South African province with a high HIV seroprevalence', *Aids,* 20(17), 2199-2206.

[339] McClelland, C. M., Chang, Y. C., Varma, A. and Kwon-Chung, K. J. (2004) 'Uniqueness of the mating system in *Cryptococcus neoformans*', *Trends Microbiol,* 12(5), 208-12.

[340] McNamara, P. J., Bradley, G. A. and Songer, J. G. (1994) 'Targeted mutagenesis of the phospholipase D gene results in decreased virulence of Corynebacterium pseudotuberculosis', *Molecular Microbiology,* 12(6), 921-930.

[341] McVean, G., Awadalla, P. and Fearnhead, P. (2002) 'A coalescent-based method for detecting and estimating recombination from gene sequences', *Genetics,* 160(3), 1231-1241.

[342] Mercure, S., Poirier, S., Lemay, G., Auger, P., Montplaisir, S. and Derepentigny, L. (1993) 'Application of biotyping and DNA typing of Candida albicans to the epidemiology of recurrent vulvovaginal candidiasis', *Journal of Infectious Diseases,* 168(2), 502-507.

[343] Metin, B., Findley, K. and Heitman, J. (2010) 'The Mating Type Locus (MAT) and Sexual Reproduction of *Cryptococcus* heveanensis: Insights into the Evolution of Sex and Sex-Determining Chromosomal Regions in Fungi', *Plos Genetics,* 6(5), 961-961.

[344] Meya, D. B., Manabe, Y. C., Castelnuovo, B., Cook, B. A., Elbireer, A. M., Kambugu, A., Kamya, M. R., Bohjanen, P. R. and Boulware, D. R. (2010) 'Cost-Effectiveness of Serum Cryptococcal Antigen Screening to Prevent Deaths among HIV-Infected Persons with a CD4(+) Cell Count <= 100 Cells/mL Who Start HIV Therapy in Resource-Limited Settings', *Clinical Infectious Diseases,* 51(4), 448-455.

[345] Meyer, W., Aanensen, D. M., Boekhout, T., Cogliati, M., Diaz, M. R., Esposto, M. C., Fisher, M., Gilgado, F., Hagen, F., Kaocharoen, S., Litvintseva, A. P., Mitchell, T. G., Simwami, S. P., Trilles, L., Viviani, M. A. and Kwon-Chung, J. (2009) 'Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*', *Medical Mycology,* 47(6), 561-570.

[346] Meyer, W., Castaneda, A., Jackson, S., Huynh, M. and Castaneda, E. (2003) 'Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates', *Emerging Infectious Diseases,* 9(2), 189-195.

[347] Meyer, W., Marszewska, K., Amirmostofian, M., Igreja, R. P., Hardtke, C., Methling, K., Viviani, M. A., Chindamporn, A., Sukroongreung, S., John, M. A., Ellis, D. H. and Sorrell, T. C. (1999) 'Molecular typing of global isolates of *Cryptococcus* neoformans var. neoformans by polymerase chain reaction fingerprinting and randomly amplified polymorphic DNA - a pilot study to standardize techniques on which to base a detailed epidemiological survey', *Electrophoresis,* 20(8), 1790-1799.

[348] Meyer, W. and Mitchell T. G. (1993) 'PCR fingerprinting to distinguish species and strains of yeast', *Molecular biology of pathogenic fungi: A laboratory manual*, New York: Telos Press, 293-302.

[349] Mirza, S. A., Phelan, M., Rimland, D., Graviss, E., Hamill, R., Brandt, M. E., Gardner, T., Sattah, M., de Leon, G. P., Baughman, W. and Hajjeh, R. A. (2003) 'The changing epidemiology of cryptococcosis: An update from population-based active surveillance in 2 large metropolitan areas, 1992-2000', *Clinical Infectious Diseases,* 36(6), 789-794.

[350] Mitchell, D. H., Sorrell, T. C., Allworth, A. M., Heath, C. H., McGregor, A. R., Papanaoum, K., Richards, M. J. and Gottlieb, T. (1995) 'Cryptococcal disease of the CNS in immunocompetent hosts: influence of cryptococcal variety on clinical manifestations and outcome', *Clinical Infectious Diseases,* 20(3), 611-616.

[351] Mitchell, T. G. and Perfect, J. R. (1995) 'Cryptococcosis in the era of AIDS–100 years after the discovery of *Cryptococcus neoformans*', *Clin Microbiol Rev,* 8(4), 515-48.

[352] Mitchell, T. G., Sandin, R. L., Bowman, B. H., Meyer, W. and Merz, W. G. (1994) 'Molecular mycology: DNA probes and applications of PCR technology', *Journal of Medical and Veterinary Mycology,* 32, 351-366.

[353] Molina, F. I., Jong, S. C. and Huffman, J. L. (1993) ' PCR amplification of the 3' external transcribed and intergenic spacers of the ribosomal DNA repeat unit in three species of Saccharomyces', *Fems Microbiology Letters,* 108(3), 259-263.

[354] Mooney H.A., H. R. J., ed. (2000) *Invasive species in a changing world,* Washington DC: island Press Washington DC.

[355] Mooney, H. A. and Cleland, E. E. (2001) 'The evolutionary impact of invasive species', *Proceedings of the National Academy of Sciences of the United States of America,* 98(10), 5446-5451.

[356] Moore, T. D. E. and Edman, J. C. (1993) 'The alpha-mating type locus of *Cryptococcus neoformans* contains a peptide pheromone gene', *Molecular and Cellular Biology,* 13(3), 1962-1970.

[357] Moosa, M. Y. S. and Coovadia, Y. M. (1997) 'Cryptococcal meningitis in durban, South Africa: A comparison of clinical features, laboratory findings, and outcome for human immunodeficiency virus (HIV)-positive and HIV-negative patients', *Clinical Infectious Diseases,* 24(2), 131-134.

[358] Motsoaledi, A. (2010) *Outline of the national HIV Counselling and Testing (HCT) campaign.*, South Africa National AIDS Council (`http://www.sanac.org.za`)

[359] Motsoaledi, A. (2011) *Health Budget Vote Policy Speech, presented at the National Assembly*, South African Government Information.

[360] Muller H. J. (1964) 'The relation of recombination to mutational advance', *Mutat Res* 1:2-9.

[361] Muller F. M. C., Lischewski A., Harmsen D., Hacker J. (1999) 'Standardized molecular typing', *Mycoses,* 42, 69-72.

[362] Murphy, J. W. and Cozad, G. C. (1972) 'Immunological Unresponsiveness Induced by Cryptococcal Capsular Polysaccharide Assayed by Hemolytic Plaque Technique', *Infection and Immunity,* 5(6), 896-&.

[363] Musser, J. M. (1996) 'Molecular population genetic analysis of emerged bacterial pathogens: Selected insights', *Emerging Infectious Diseases,* 2(1), 1-17.

[364] Mwaba, P., Mwansa, J., Chintu, C., Pobee, J., Scarborough, M., Portsmouth, S. and Zumla, A. (2001) 'Clinical presentation, natural history, and cumulative death rates of 230 adults with primary cryptococcal meningitis in Zambian AIDS patients treated under local conditions', *Postgrad Med J,* 77(914), 769-73.

[365] Mylonakis, E. and Aballay, A. (2005) 'Worms and flies as genetically tractable animal models to study host-pathogen interactions', *Infect Immun,* 73(7), 3833-41.

[366] Mylonakis, E., Ausubel, F. M., Perfect, J. R., Heitman, J. and Calderwood, S. B. (2002) 'Killing of Caenorhabditis elegans by *Cryptococcus neoformans* as a model of

yeast pathogenesis', *Proceedings of the National Academy of Sciences of the United States of America,* 99(24), 15675-15680.

[367] Mylonakis, E., Calderwood, S. B. and Ausubel, F. A. (2006) 'Heterologous hosts and the evolution and study of fungal pathogenesis', *Molecular Principles of Fungal Pathogenesis*, 215-225.

[368] Mylonakis, E., Casadevall, A. and Ausubel, F. M. (2007) 'Exploiting amoeboid and non-vertebrate animal model systems to study the virulence of human pathogenic fungi', *Plos Pathogens,* 3(7), 859-865.

[369] Mylonakis, E., Moreno, R., El Khoury, J. B., Idnurm, A., Heitman, J., Calderwood, S. B., Ausubel, F. M. and Diener, A. (2005) 'Galleria mellonella as a model system to study *Cryptococcus neoformans* pathogenesis', *Infect Immun,* 73(7), 3842-50.

[370] Najvar, L. K., Bocanegra, R. and Graybill, J. R. (1999) 'An alternative animal model for comparison of treatments for cryptococcal meningitis', *Antimicrobial Agents and Chemotherapy,* 43(2), 413-414.

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

[372] Neuveglise, C., Sarfati, J., Latge, J. P. and Paris, S. (1996) 'Afut1, a retrotransposon-like element from *Aspergillus fumigatus*', *Nucleic Acids Research,* 24(8), 1428-1434.

[373] Newman, S. L. (2001) 'Cell-mediated immunity to Histoplasma capsulatum', *Semin Respir Infect,* 16(2), 102-8.

[374] Ngamskulrungroj, P., Gilgado, F., Faganello, J., Litvintseva, A. P., Leal, A. L., Tsui, K., Mitchell, T. G., Vainstein, M. H. and Meyer, W. (2009) 'Genetic diversity of the *Cryptococcus* species complex suggests that *Cryptococcus gattii* deserves to have varieties', *PLoS ONE,* (June), e5862.

[375] Nielsen, K., Cox, G. M., Wang, P., Toffaletti, D. L., Perfect, J. R. and Heitman, J. (2003) 'Sexual cycle of *Cryptococcus neoformans* var. grubii and virulence of congenic a and alpha isolates', *Infection and Immunity,* 71(9), 4831-4841.

[376] Nielsen, K. and Heitman, J. (2007) 'Sex and Virulence of Human Pathogenic Fungi', *Fungal Genomics,* 57, 143-173.

[377] Nierman, W. C., Pain, A., Anderson, M. J., Wortman, J. R., Kim, H. S., Arroyo, J., Berriman, M., Abe, K., Archer, D. B., Bermejo, C., Bennett, J., Bowyer, P., Chen, D., Collins, M., Coulsen, R., Davies, R., Dyer, P. S., Farman, M., Fedorova, N., Feldblyum, T. V., Fischer, R., Fosker, N., Fraser, A., Garcia, J. L., Garcia, M. J., Goble, A., Goldman, G. H., Gomi, K., Griffith-Jones, S., Gwilliam, R., Haas, B., Haas, H., Harris, D., Horiuchi, H., Huang, J., Humphray, S., Jimenez, J., Keller, N., Khouri, H., Kitamoto, K., Kobayashi, T., Konzack, S., Kulkarni, R., Kumagai, T., Lafton, A., Latge, J. P., Li, W. X., Lord, A., Majoros, W. H., May, G. S., Miller, B. L., Mohamoud, Y., Molina, M., Monod, M., Mouyna, I., Mulligan, S., Murphy, L., O'Neil, S., Paulsen, I., Penalva, M. A., Pertea, M., Price, C., Pritchard, B. L.,

Quail, M. A., Rabbinowitsch, E., Rawlins, N., Rajandream, M. A., Reichard, U., Renauld, H., Robson, G. D., de Cordoba, S. R., Rodriguez-Pena, J. M., Ronning, C. M., Rutter, S., Salzberg, S. L., Sanchez, M., Sanchez-Ferrero, J. C., Saunders, D., Seeger, K., Squares, R., Squares, S., Takeuchi, M., Tekaia, F., Turner, G., de Aldana, C. R. V., Weidman, J., White, O., Woodward, J., Yu, J. H., Fraser, C., Galagan, J. E., Asai, K., Machida, M., Hall, N., Barrell, B. and Denning, D. W. (2005) 'Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*', *Nature,* 438(7071), 1151-1156.

[378] Nishikawa, M. M., Lazera, M. S., Barbosa, G. G., Trilles, L., Balassiano, B. R., Macedo, R. C., Bezerra, C. C., Perez, M. A., Cardarelli, P. and Wanke, B. (2003) 'Serotyping of 467 *Cryptococcus neoformans* isolates from clinical and environmental sources in Brazil: analysis of host and regional patterns', *J Clin Microbiol,* 41(1), 73-7.

[379] Nosanchuk, J. D. and Casadevall, A. (2003) 'The contribution of melanin to microbial pathogenesis', *Cellular Microbiology,* 5(4), 203-223.

[380] Nosanchuk, J. D., Gomez, B. L., Youngchim, S., Diez, S., Aisen, P., Zancope-Oliveira, R. M., Restrepo, A., Casadevall, A. and Hamilton, A. J. (2002) 'Histoplasma capsulatum synthesizes melanin-like pigments in vitro and during mammalian infection', *Infection and Immunity,* 70(9), 5124-5131.

[381] Noverr, M. C., Cox, G. M., Perfect, J. R. and Huffnagle, G. B. (2003a) 'Role of PLB1 in pulmonary inflammation and Cryptococcal eicosanoid production', *Infection and Immunity,* 71(3), 1538-1547.

[382] Noverr, M. C., Erb-Downward, J. R. and Huffnagle, G. B. (2003b) 'Production of eicosanoids and other oxylipins by pathogenic eukaryotic microbes', *Clinical Microbiology Reviews,* 16(3), 517-+.

[383] Odom, A., Muir, S., Lim, E., Toffaletti, D. L., Perfect, J. and Heitman, J. (1997) 'Calcineurin is required for virulence of *Cryptococcus neoformans*', *Embo Journal,* 16(10), 2576-2589.

[384] Ohkusu, M., Tangonan, N., Takeo, K., Kishida, E., Ohkubo, M., Aoki, S., Nakamura, K., Fujii, T., de Siqueira, I. C., Maciel, E. A. P., Sakabe, S., Almeida, G. M. D., Heins-Vaccari, E. M. and Lacaz, C. d. S. (2002) 'Serotype, mating type and ploidy of *Cryptococcus neoformans* strains isolated from patients in Brazil', *Revista do Instituto de Medicina Tropical de Sao Paulo,* 44(6), 299-302.

[385] Oleksyk, T. K., Smith, M. W. and O'Brien, S. J. (2010) 'Genome-wide scans for footprints of natural selection', *Philosophical Transactions of the Royal Society B-Biological Sciences,* 365(1537), 185-205.

[386] Oursler, K. A., Moore, R. D. and Chaisson, R. E. (1999) 'Risk factors for cryptococcal meningitis in HIV-infected patients', *AIDS Res Hum Retroviruses,* 15(7), 625-31.

[387] Owens, C. (1991) 'Alarming spread of AIDS virus in Thailand may threaten country's recent economic gains', *Asian Wall Street Journal,* 13(51), 4.

[388] Padhye, A. A. and Ajello, L. (1977) 'The taxonomic status of the hedgehog fungus Trichophyton erinacei', *Sabouraudia-Journal of Medical and Veterinary Mycology,* 15(JUL), 103-114.

[389] Padhye, A. A. and Carmichael, J. W. (1969) 'Mating behaviour of Trichophyton mentagrophytes varieties paired with Arthroderma benhamiae mating types', *Sabouraudia,* 7(3), 178-181.

[390] Pappagianis, D. and Einstein, H. (1978) 'Tempest from Tehachapi takes toll or coccidioides conveyed alor and afar', *Western Journal of Medicine*, 527-530.

[391] Park, B. J., Wannemuehler, K. A., Marston, B. J., Govender, N., Pappas, P. G. and Chiller, T. M. (2009) 'Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS', *Aids,* 23(4), 525-530.

[392] Parker, I. M. and Gilbert, G. S. (2004) 'The evolutionary ecology of novel plant-pathogen interactions', *Annual Review of Ecology Evolution and Systematics,* 35, 675-700.

[393] Peakall, R. and Smouse, P. E. (2006) 'GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research', *Molecular Ecology Notes,* 6(1), 288-295.

[394] Perfect, J. R. (2006) '*Cryptococcus neoformans*: a sugar-coated killer', *Molecular Principles of Fungal Pathogenesis*, 281-303.

[395] Perfect, J. R. (2006) '*Cryptococcus neoformans*: the yeast that likes it hot', *Fems Yeast Research,* 6(4), 463-468.

[396] Perfect, J. R. and Casadevall, A. (2002) 'Cryptococcosis', *Infectious Disease Clinics of North America,* 16(4), 837-+.

[397] Perfect, J. R. and Casadevall, A. (2006) 'Fungal molecular pathogenesis: What can it do and why do we need it?', *Molecular Principles of Fungal Pathogenesis-Print*, 3-11.

[398] Perfect, J. R., Ketabchi, N., Cox, G. M., Ingram, C. W. and Beiser, C. L. (1993a) 'Karyotyping of *Cryptococcus neoformans* as an epidemiological tool', *Journal of Clinical Microbiology,* 31(12), 3305-3309.

[399] Perfect, J. R., Toffaletti, D. L. and Rude, T. H. (1993b) 'The gene encoding phosphoribosylaminoimidazole carboxylase (ADE2) is essential for growth of *Cryptococcus neoformans* in cerebrospinal fluid', *Infection and Immunity,* 61(10), 4446-4451.

[400] Perry, C. R., Matcham, S. E., Wood, D. A. and Thurston, C. F. (1993) 'The structure of laccase protein and its synthesis by the commercial mushroom Agaricus bisporus', *Journal of General Microbiology,* 139, 171-178.

[401] Petter, R., Kang, B. S., Boekhout, T., Davis, B. J. and Kwon-Chung, K. J. (2001) 'A survey of heterobasidiomycetous yeasts for the presence of the genes homologous to virulence factors of Filobasidiella neoformans, CNLAC1 and CAP59', *Microbiology-Sgm,* 147, 2029-2036.

[402] Petzold, E. W., Himmelreich, U., Mylonakis, E., Rude, T., Toffaletti, D., Cox, G. M., Miller, J. L. and Perfect, J. R. (2006) 'Characterization and regulation of the trehalose synthesis pathway and its importance in the pathogenicity of *Cryptococcus* neoformans', *Infection and Immunity,* 74(10), 5877-5887.

[403] Pfeiffer T., Ellis D. (1991) 'Environmental isolation of *Cryptococcus gattii* from California'. *J Infect Dis* 163: 929-930.

[404] Phanuohak, P., Locharernkul, C., Panmuong, W. and Wilde, H. (1985) 'A report of three cases of AIDS in Thailand', 3:195-199.

[405] Pitisuttithum, P., Tansuphasawadikul, S., Simpson, A. J. H., Howe, P. A. and White, N. J. (2001) 'A prospective study of AIDS-associated cryptococcal meningitis in Thailand treated with high-dose amphotericin B', *Journal of Infection,* 43(4), 226-233.

[406] Polacheck, I., Lebens, G. and Hicks, J. B. (1992) 'Development of DNA probes for early diagnosis and epidemiological study of cryptococcosis in AIDS patients', *Journal of Clinical Microbiology,* 30(4), 925-930.

[407] Polacheck, I. and Lebens, G. A. (1989) 'Electrophoretic Karyotype of the Pathogenic Yeast *Cryptococcus*-Neoformans', *Journal of General Microbiology,* 135, 65-71.

[408] Polacheck, I., Platt, Y. and Aronovitch, J. (1990) 'Catecholamines and Virulence of *Cryptococcus*-Neoformans', *Infection and Immunity,* 58(9), 2919-2922.

[409] Poonwan, N., Mikami, Y., Poosuwan, S., BoonLong, J., Mekha, N., Kusum, M., Yazawa, K., Tanaka, R., Nishimura, K. and Konyama, K. (1997) 'Serotyping of *Cryptococcus* neoformans strains isolated from clinical specimens in Thailand and their susceptibility to various antifungal agents', *European Journal of Epidemiology,* 13(3), 335-340.

[410] Porapakkham, Y., Pramarnpol, S. and Athibhoddhi, S. (1995) 'The Evolution of HIV/AIDS Policy in Thailand: 1984-1994', available: pdf.usaid.gov/pdf_docs/PNACG546.pdf [accessed 16th November, 2010].

[411] Poulain, D., Hopwood, V. and Vernes, A. (1985) 'Antigenic variability of Candida albicans', *Crc Critical Reviews in Microbiology,* 12(3), 223-270.

[412] Powderly, W. G., Cloud, G. A., Dismukes, W. E. and Saag, M. S. (1994) 'Measurement of cryptococcal antigen in serum and cerebrospinal fluid: value in the management of AIDS-associated cryptococcal meningitis', *Clinical Infectious Diseases,* 18(5), 789-792.

[413] Pritchard, J. K. and Przeworski, M. (2001) 'Linkage disequilibrium in humans: Models and data', *American Journal of Human Genetics,* 69(1), 1-14.

[414] Przeworski, M., Hudson, R. R. and Di rienzo, A. (2000) 'Adjusting the focus on human variation', *Trends in Genetics,* 16(7), 296-302.

[415] Pujol, C., Joly, S., Lockhart, S. R., Noel, S., Tibayrenc, M. and Soll, D. R. (1997) 'Parity among the randomly amplified polymorphic DNA method, multilocus enzyme electrophoresis, and southern blot hybridization with the moderately repetitive

DNA probe Ca3 for fingerprinting Candida albicans', *Journal of Clinical Microbiology,* 35(9), 2348-2358.

[416] Pujol C., Reynes J., Renaud F., Raymond M., Tibayrenc M ., Ayala F. J., Janbon F., Malli M., Bastide J. M. (1993) 'The yeast *Candida albicans* has a clonal mode of reproduction in a population of infected human immunodeficiency virus-positive patients.' *Proc Natl Acad Sci U S A.*, 90(20): 9456-9459.

[417] Pukkila-Worley, R., Gerrald, Q. D., Kraus, P. R., Boily, M. J., Davis, M. J., Giles, S. S., Cox, G. M., Heitman, J. and Alspaugh, J. A. (2005) 'Transcriptional network of multiple capsule and melanin genes governed by the *Cryptococcus neoformans* cyclic AMP cascade', *Eukaryotic Cell,* 4(1), 190-201.

[418] Putignani, L., Antonucci, G., Paglia, M. G., Vincenzi, L., Festa, A., De Mori, P., Loiacono, L. and Visca, P. (2008) 'Cryptococcal lymphadenitis as a manifestation of immune reconstitution inflammatory syndrome in an HIV-positive patient: a case report and review of the literature', *International Journal of Immunopathology and Pharmacology,* 21(3), 751-756.

[419] Qazzafi, Z., Thiruchunapalli, D., Birkenhead, D., Bell, D. and Sandoe, J. A. T. (2007) 'Invasive *Cryptococcus neoformans* infection in an asplenic patient', *Journal of Infection,* 55(6), 566-568.

[420] Rahmet-Alla, M., Rowley, A. F. and Alla, M. R. (1990) 'Studies on the cellular defense reactions of the Madeira cockroach, Leucophaea maderae: in vitro phagocytosis of different strains of Bacillus cereus and their effect on hemocyte viability', *Journal of Invertebrate Pathology,* 55(3), 350-356.

[421] Ramirez-Soriano, A., Ramos-Onsins, S. E., Rozas, J., Calafell, F. and Navarro, A. (2008) 'Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination', *Genetics,* 179(1), 555-567.

[422] Ramos-Onsins, S. E. and Rozas, J. (2006) 'Statistical properties of new neutrality tests against population growth (vol 19, pg 2092, 2002)', *Molecular Biology and Evolution,* 23(8), 1642-1642.

[423] Randhawa, H. S., Kowshik, T. and Khan, Z. U. (2003) 'Decayed wood of Syzygium cumini and Ficus religiosa living trees in Delhi/New Delhi metropolitan area as natural habitat of *Cryptococcus neoformans*', *Medical Mycology,* 41(3), 199-209.

[424] Randhawa, H. S., Kowshik, T. and Khan, Z. U. (2005) 'Efficacy of swabbing versus a conventional technique for isolation of *Cryptococcus neoformans* from decayed wood in tree trunk hollows', *Medical Mycology,* 43(1), 67-71.

[425] Rappelli, P., Are, R., Casu, G., Fiori, P. L., Cappuccinelli, P. and Aceti, A. (1998) 'Development of a nested PCR for detection of *Cryptococcus* neoformans in cerebrospinal fluid', *Journal of Clinical Microbiology,* 36(11), 3438-3440.

[426] Reedy, J. L., Bastidas, R. J. and Heitman, J. (2007) 'The virulence of human pathogenic fungi: notes from the South of France', *Cell Host Microbe,* 2(2), 77-83.

[427] Retallack, D. M., Heinecke, E. L., Gibbons, R., Deepe, G. S. and Woods, J. P. (2000) 'The URA5 gene is necessary for Histoplasma capsulatum growth during infection of mouse and human cells (vol 67, pg 624, 1999)', *Infection and Immunity,* 68(1), 426-426.

[428] Revazishvili, T., Kotetishvili, M., Stine, O. C., Kreger, A. S., Morris, J. G., Jr. and Sulakvelidze, A. (2004) 'Comparative analysis of multilocus sequence typing and pulsed-field gel electrophoresis for characterizing Listeria monocytogenes strains isolated from environmental and clinical sources', *J Clin Microbiol,* 42(1), 276-85.

[429] Rhodes, J. C., Polacheck, I. and Kwon-Chung, K. J. (1982) 'Phenoloxidase Activity and Virulence in Isogenic Strains of *Cryptococcus*-Neoformans', *Infection and Immunity,* 36(3), 1175-1184.

[430] Richardson, D. M. and Pysek, P. (2007) 'Classics in physical geography revisited: Elton, C.S. 1958: The ecology of invasions by animals and plants', *Progress in Physical Geography,* 31(6), 659-666.

[431] Richardson, M. D. and Warnock, D. W. (2008) 'Antifungal drugs', *Fungal Infection: Diagnosis and Management*,Third edition, Oxford, UK: Blackwell Publishing Ltd.

[432] Roberts, C. J., Nelson, B., Marton, M. J., Stoughton, R., Meyer, M. R., Bennett, H. A., He, Y. D. D., Dai, H. Y., Walker, W. L., Hughes, T. R., Tyers, M., Boone, C. and Friend, S. H. (2000) 'Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles', *Science,* 287(5454), 873-880.

[433] Robinson, P. A., Bauer, M., Leal, M. E., Evans, S. G., Holtom, P. D., Diamond, D. M., Leedom, J. M. and Larsen, R. A. (1999) 'Early mycological treatment failure in AIDS-associated cryptococcal meningitis', *Clinical Infectious Diseases,* 28(1), 82-92.

[434] Rosas, A. L., Nosanchuk, J. D., Gomez, B. L., Edens, W. A., Henson, J. M. and Casadevall, A. (2000) 'Isolation and serological analyses of fungal melanins', *Journal of Immunological Methods,* 244(1-2), 69-80.

[435] Rosenblum, E. B., Stajich, J. E., Maddox, N. and Eisen, M. B. (2008) 'Global gene expression profiles for life stages of the deadly amphibian pathogen Batrachochytrium dendrobatidis', *Proceedings of the National Academy of Sciences of the United States of America,* 105(44), 17034-17039.

[436] RoyChoudhury, A. and Wakeley, J. (2010) 'Sufficiency of the number of segregating sites in the limit under finite-sites mutation', *Theoretical Population Biology,* 78(2), 118-122.

[437] Rozenbaum, R. and Goncalves, A. J. R. (1994) 'Clinical epidemiological study of 171 cases of cryptococcosis', *Clinical Infectious Diseases,* 18(3), 369-380.

[438] Saag, M. S., Graybill, R. J., Larsen, R. A., Pappas, P. G., Perfect, J. R., Powderly, W. G., Sobel, J. D., Dismukes, W. E. and Mycoses Study Grp Cryptococcal, S. (2000) 'Practice guidelines for the management of cryptococcal disease', *Clinical Infectious Diseases,* 30(4), 710-718.

[439] Saag, M. S., Powderly, W. G., Cloud, G. A., Robinson, P., Grieco, M. H., Sharkey, P. K., Thompson, S. E., Sugar, A. M., Tuazon, C. U., Fisher, J. F., Hyslop, N., Jacobson, J. M., Hafner, R. and Dismukes, W. E. (1992) 'Comparison of amphotericin B with fluconazole in the treatment of acute AIDS-associated cryptococcal meningitis. The NIAID Mycoses Study Group and the AIDS Clinical Trials Group.', *New England Journal of Medicine,* 326(2), 83-89.

[440] Saffer, L. D., Krug, S. A. L. and Schwartzman, J. D. (1989) 'The role of phospholipase in host cell penetration by *Toxoplasma gondii*', *American Journal of Tropical Medicine and Hygiene,* 40(2), 145-149.

[441] Saffer, L. D. and Schwartzman, J. D. (1991) 'A soluble phospholipase of *Toxoplasma gondii* associated with host cell penetration', *Journal of Protozoology,* 38(5), 454-460.

[442] Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1986) 'Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes', *Nature,* 324(6093), 163-166.

[443] Saitou, N. and Nei, M. (1987) 'The Neighbor-Joining Method - a New Method for Reconstructing Phylogenetic Trees', *Molecular Biology and Evolution,* 4(4), 406-425.

[444] Salas, S. D., Bennett, J. E., KwonChung, K. J., Perfect, J. R. and Williamson, P. R. (1996) 'Effect of the laccase gene, CNLAC1, on virulence of *Cryptococcus neoformans*', *Journal of Experimental Medicine,* 184(2), 377-386.

[445] Salkin, I. F. and Hurd, N. J. (1982) 'New medium for differentiation of *Cryptococcus neoformans* serotype pairs', *Journal of Clinical Microbiology,* 15(1), 169-171.

[446] Salzet, M. (2001) 'Vertebrate innate immunity resembles a mosaic of invertebrate immune responses', *Trends in Immunology,* 22(6), 285-288.

[447] Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) 'Molecular cloning: a laboratory manual', *Molecular cloning: a laboratory manual.,* (Ed. 2), xxxviii + 1546 pp.

[448] Santangelo, R., Zoellner, H., Sorrell, T., Wilson, C., Donald, C., Djordjevic, J., Shounan, Y. and Wright, L. (2004) 'Role of extracellular phospholipases and mononuclear phagocytes in dissemination of cryptococcosis in a murine model', *Infection and Immunity,*72, 2229-2239.

[449] Satishchandra, P., Mathew, T., Gadre, G., Nagarathna, S., Chandramukhi, A., Mahadevan, A. and Shankar, S. K. (2007) 'Cryptococcal meningitis: Clinical, diagnostic and therapeutic overviews', *Neurology India,* 55(3), 226-232.

[450] Saul N., Krockenberger M., Carter D. (2008) 'Evidence of recombination in mixed-mating-type and alpha-only populations of *Cryptococcus gattii* sourced from single eucalyptus tree hollows',*Eukaryot Cellm,* 7: 727-734.

[451] Scarborough, M., Gordon, S. B., Whitty, C. J. M., French, N., Njalale, Y., Chitani, A., Peto, T. E. A., Lalloo, D. G. and Zijlstra, E. E. (2007) 'Corticosteroids for bacterial meningitis in adults in sub-Saharan Africa', *New England Journal of Medicine,* 357, 2441-2450.

[452] Schaars, C. F., Meintjes, G. A., Morroni, C., Post, F. A. and Maartens, G. (2006) 'Outcome of AIDS-associated cryptococcal meningitis initially treated with 200 mg/day or 400 mg/day of fluconazole', *Bmc Infectious Diseases,* 6.

[453] Scherer, S. and Stevens, D. A. (1988) 'A Candida albicans dispersed, repeated gene family and its epidemiologic applications', *Proceedings of the National Academy of Sciences of the United States of America,* 85(5), 1452-1456.

[454] Schmalreck, A. F. and Hotzel, H. (2000) 'Fourier-Transform Infrared Spectroscopy, molecular biological methods and antifungal susceptibility patterns for identification and differentiation of *Cryptococcus* species', *Mycoses,* 43, 61-68.

[455] Schouten, A., Wagemakers, L., Stefanato, F. L., van der Kaaij, R. M. and van Kan, J. A. L. (2002) 'Resveratrol acts as a natural profungicide and induces self-intoxication by a specific laccase', *Molecular Microbiology,* 43(4), 883-894.

[456] Schutte, C. M., Van der Meyden, C. H. and Magazi, D. S. (2000) 'The impact of HIV on meningitis as seen at a South African Academic Hospital (1994 to 1998)', *Infection,* 28(1), 3-7.

[457] Schwartz, M. W., Hoeksema, J. D., Gehring, C. A., Johnson, N. C., Klironomos, J. N., Abbott, L. K. and Pringle, A. (2006) 'The promise and the potential consequences of the global transport of mycorrhizal fungal inoculum', *Ecology Letters,* 9(5), 501-515.

[458] Seaton, R., Nafaqi, S., Wembri, J. and Warrell, D. (1996) 'Correct predictors of outcome in *Cryptococcus neoformans* var. *gattii* meningitis.', *Quarterly Journal of Medicine,* 89, 423-28.

[459] Selander, R. K., Musser, J. M., Caugant, D. A., Gilmour, M. N. and Whittam, T. S. (1987) 'Population genetics of pathogenic bacteria', *Microbial Pathogenesis,* 3(1), 1-7.

[460] Serpell J. (2000). In D. Turner, P. P. G. Bateson, Eds. *The Domestic Cat.* Cambridge: Cambridge Univ. Press.

[461] Shisana, O., Rehle, T., Simbayi, L., Zuma, K., Jooste, S., Pillay-van-Wyk, V., Mbelle, N., Van Zyl, J., Parker, W., NP, Z., Pezi, S. and Team, t. S. I. I. (2009) *South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2008: A Turning Tide Among Teenagers.*, Cape Town: Human Sciences Research Council.

[462] Sibley L. D., Boothroyd J. C. (1992). 'Virulent strains of Toxoplasma gondii comprise a single clonal lineage.' *Nature*, 359, 82 - 85 (03 September 1992); doi:10.1038/359082a0.

[463] Siddiqui, A. A., Brouwer, A. E., Wuthiekanun, V., Jaffar, S., Shattock, R., Irving, D., Sheldon, J., Chierakul, W., Peacock, S., Day, N., White, N. J. and Harrison, T. S. (2005) 'IFN-gamma at the site of infection determines rate of clearance of infection in cryptococcal meningitis', *Journal of Immunology,* 174(3), 1746-1750.

[464] Sidrim, J. J., Costa, A. K., Cordeiro, R. A., Brilhante, R. S., Moura, F. E., Castelo-Branco, D. S., Neto, M. P. d. A. and Fabio, M. (2010) 'Molecular methods for the diagnosis and characterization of *Cryptococcus*: a review ', 56(6), 445 - 58.

[465] Silber, E., Sonnenberg, P., Ho, K. C., Koornhof, H. J., Eintracht, S., Morris, L. and Saffer, D. (1999) 'Meningitis in a community with a high prevalence of tuberculosis and HIV infection', *Journal of the Neurological Sciences,* 162(1), 20-26.

[466] Silva, E. G. d., Baroni, F. d. A., Viani, F. C., Ruiz, L. d. S., Gandra, R. F., Auler, M. E., Dias, A. L. T., Gambale, W. and Paula, C. R. (2006) 'Virulence profile of strains of *Cryptococcus neoformans* var. grubii evaluated by experimental infection in BALB/c mice and correlation with exoenzyme activity', *Journal of medical microbiology,* 55(Pt 2), 139-42.

[467] Silverman, D. J., Santucci, L. A., Meyers, N. and Sekeyova, Z. (1992) 'Penetration of host cells by Rickettsia rickettsii appears to be mediated by a phospholipase of rickettsial origin', *Infection and Immunity,* 60(7), 2733-2740.

[468] Simpson, E. (1949) 'Measurement of diversity', *Nature,* 163(4148).

[469] Simwami, S. P., Khayhan, K., Henk, D. A., Aanensen, D. M., Boekhout, T., Hagen, F., Brouwer, A. E., Harrison, T. S., Donnelly, C. A. and Fisher, M. C. (2011) 'Low diversity *Cryptococcus neoformans* variety grubii multilocus sequence types from Thailand are consistent with an ancestral African origin', *PLoS pathogens,* 7(4), e1001343.

[470] Sirisopana, N., Torugsa, K., Jugsudee, A., Chuenchitra, C., Supapongse, T. and Armed Forces Res. Inst. Med. Sicences-Royal Thai Army, B., Thailand (1993) *Prevalence of HIV-1 infection in young men entering the royal Thai army,* translated by Berlin, Germany: IXth International Conference on AIDS {a}, Berlin, Germany, 680.

[471] Slatkin, M. and Hudson, R. R. (1991) 'Pairwise Comparisons of Mitochondrial DNA Sequences in Stable and Exponentially Growing Populations', *Genetics,* 129(2), 555-562.

[472] Sloan, D., Dlamini, S. and Dedicoat, M. (2009) 'Management of cryptoccocal meningitis in resource-limited settings: A systematic review', *Samj South African Medical Journal,* 99(5), 310-312.

[473] Sloan, D., Dlamini, S., Paul, N. and Dedicoat, M. (2008) 'Treatment of acute cryptococcal meningitis in HIV infected adults, with an emphasis on resource-limited settings', *Cochrane Database Syst Rev,* (4), CD005647.

[474] Sloan, D. J., Dedicoat, M. J. and Lalloo, D. G. (2009) 'Treatment of cryptococcal meningitis in resource limited settings', *Current Opinion in Infectious Diseases,* 22(5), 455-463.

[475] Small, J. M., Mitchell Thomas G (1989) 'Strain variation in antiphagocytic activity of capsular polysaccharides from *Cryptococcus neoformans* serotype A', *Infection and Immunity,* 57(12), 3751-3756.

[476] Smith, G. A., Marquis, H., Jones, S., Johnston, N. C., Portnoy, D. A. and Goldfine, H. (1995) 'The two distinct phospholipases C of Listeria monocytogenes have overlap-

ping roles in escape from a vacuole and cell-to-cell spread', *Infection and Immunity,* 63(11), 4231-4237.

[477] Smith, J. M. (1978) 'The evolution of sex', *The Evolution of Sex*, 222.

[478] Smith, J. M. (1992) 'Analyzing the mosaic structure of genes', *Journal of Molecular Evolution,* 34(2), 126-129.

[479] Smith, J. M., Smith, N. H., Orourke, M. and Spratt, B. G. (1993) 'How Clonal Are Bacteria', *Proceedings of the National Academy of Sciences of the United States of America,* 90(10), 4384-4388.

[480] Soll, D. R. (2000) 'The ins and outs of DNA fingerprinting the infectious fungi', *Clinical Microbiology Reviews,* 13(2), 332-+.

[481] Soll, D. R., Staebell, M., Langtimm, C., Pfaller, M., Hicks, J. and Rao, T. V. G. (1988) 'Multiple Candida strains in the course of a single systemic infection', *Journal of Clinical Microbiology,* 26(8), 1448-1459.

[482] Sommerhalder, R. J., McDonald, B. A. and Zhan, J. (2007) 'Concordant evolution of mitochondrial and nuclear genomes in the wheat pathogen Phaeosphaeria nodorum', *Fungal Genetics and Biology,* 44(8), 764-772.

[483] Songer, J. G. (1997) 'Bacterial phospholipases and their role in virulence', *Trends in Microbiology,* 5(4), 156-161.

[484] Sorrell, T. C., Chen, S. C. A., Ruma, P., Meyer, W., Pfeiffer, T. J., Ellis, D. H. and Brownlee, A. G. (1996) 'Concordance of clinical and environmental isolates of *Cryptococcus neoformans* var *gattii* by random amplification of polymorphic DNA analysis and PCR fingerprinting', *Journal of Clinical Microbiology,* 34(5), 1253-1260.

[485] Sorrell, T. C. and Ellis, D. H. (1997) 'Ecology of *Cryptococcus* neoformans', *Revista Iberoamericana de Micologia,* 14(2), 42-43.

[486] Sorvillo F., Beall G., Turnaer P. A., Beer V. L., Kovacs A. A., Kerndt P. R. (1997) 'Incidence and factors associated with extrapulmonary cryptococcosis among persons with HIV infection in Los Angeles Country'. *AIDS*, 11: 673-679.

[487] Speed, B. and Dunt, D. (1995) 'Clinical and host differences between infections with the two varieties of *Cryptococcus neoformans*', *Clinical Infectious Diseases,* 21(1), 28-34.

[488] Spitzer, E. D. and Spitzer, S. G. (1992) 'Use of a dispersed repetitive DNA element to distinguish clinical isolates of *Cryptococcus neoformans*', *Journal of Clinical Microbiology,* 30(5), 1094-1097.

[489] Spitzer, S. G. and Spitzer, E. D. (1994) 'Characterization of the CNRE-1 family of repetitive DNA elements in *Cryptococcus neoformans*', *Gene,* 144(1), 103-106.

[490] Spitzer, S. G. and Spitzer, E. D. (1997) 'Isolation of *Cryptococcus* neoformans chromosome-specific probes using expressed sequence tags', *Journal of Medical and Veterinary Mycology,* 35(4), 257-261.

[491] Spratt, B. G. and Maiden, M. C. J. (1999) 'Bacterial population genetics, evolution and epidemiology', *Philosophical Transactions of the Royal Society B-Biological Sciences,* 354(1384), 701-710.

[492] Spratt, B. G., Staley, J. T. and Fisher, M. C. (2006) 'Introduction: species and speciation in micro-organisms', *Philosophical Transactions of the Royal Society B-Biological Sciences,* 361(1475), 1897-1898.

[493] Sriburee, P., Khayhan, S., Khamwan, C., Panjaisee, S. and Tharavichitkul, P. (2004) 'Serotype and PCR-fingerprints of clinical and environmental isolates of *Cryptococcus neoformans* in Chiang Mai, Thailand', *Mycopathologia,* 158(1), 25-31.

[494] Stajich, J. E. and Dietrich, F. S. (2006) 'Genomic perspectives on the fungal kingdom', *Molecular Principles of Fungal Pathogenesis*, 657-666.

[495] Statistics South Africa (2008) *Mortality and causes of death in South Africa: Findings from death notification, 2006*, Pretoria.

[496] Steenbergen, J. N. and Casadevall, A. (2000) 'Prevalence of *Cryptococcus* neoformans var. neoformans (Serotype D) and *Cryptococcus neoformans* var. grubii (Serotype A) isolates in New York City', *J Clin Microbiol,* 38(5), 1974-6.

[497] Steenbergen, J. N., Nosanchuk, J. D., Malliaris, S. D. and Casadevall, A. (2003) '*Cryptococcus neoformans* virulence is enhanced after growth in the genetically malleable host Dictyostelium discoideum', *Infection and Immunity,* 71(9), 4862-4872.

[498] Steenbergen, J. N., Shuman, H. A. and Casadevall, A. (2001) '*Cryptococcus* neoformans interactions with amoebae suggest an explanation for its virulence and intracellular pathogenic strategy in macrophages', *Proceedings of the National Academy of Sciences of the United States of America,* 98(26), 15245-15250.

[499] Stefani, S. and Agodi, A. (2000) 'Molecular epidemiology of antibiotic resistance', *International Journal of Antimicrobial Agents,* 13(3), 143-153.

[500] Su C., Evans D., Cole R. H., Kissinger J. C., Ajioka J. W., Sibley L. D. (2003) Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science*, 299: 414-416.

[501] Stephen, C., Lester, S., Black, W. and Fyfe, M. (2002) 'Multispecies outbreak of cryptococcosis on southern Vancouver Island, British Columbia', *Canadian Veterinary Journal-Revue Veterinaire Canadienne,* 43(10), 792-794.

[502] Stevens, D. A., Denning, D. W., Shatsky, S., Armstrong, R. W., Adler, J. D. and Lewis, B. H. (1999) 'Cryptococcal meningitis in the immunocompromised host: intracranial hypertension and other complications', *Mycopathologia,* 146(1), 1-8.

[503] Stukenbrock E. H., McDonald B. A. (2008) 'The origins of plant pathogens in agro-ecosystems' *Ann Rev Phytopathol*, 46, 75-100.

[504] Sugita, T., Ikeda, R. and Shinoda, T. (2001) 'Diversity among strains of *Cryptococcus neoformans* var. *gattii* as revealed by a sequence analysis of multiple genes and a chemotype analysis of capsular polysaccharide', *Microbiology and Immunology,* 45(11), 757-768.

[505] Sugita, T., Nakajima, M., Ikeda, R., Matsushima, T. and Shinoda, T. (2002) 'Sequence analysis of the ribosomal DNA intergenic spacer 1 regions of Trichosporon species', *Journal of Clinical Microbiology,* 40(5), 1826-1830.

[506] Sugita, T., Nishikawa, A., Ikeda, R. and Shinoda, T. (1999) 'Identification of medically relevant Trichosporon species based on sequences of internal transcribed spacer regions and construction of a database for Trichosporon identification', *Journal of Clinical Microbiology,* 37(6), 1985-1993.

[507] Sukroongreung, S., Nilakul, C., Ruangsomboon, O., Chuakul, W. and Eampokalap, B. (1996) 'Serotypes of *Cryptococcus neoformans* isolated from patients prior to and during the AIDS era in Thailand', *Mycopathologia,* 135(2), 75-78.

[508] Sun, S. and Xu, J. (2007) 'Genetic analyses of a hybrid cross between serotypes A and D strains of the human pathogenic fungus *Cryptococcus* neoformam', *Genetics,* 177(3), 1475-1486.

[509] Sutar, R., David, J. K., Ganesan, K., Ghosh, A. K., Singhi, S., Chakrabarti, A. and Bachhawat, A. K. (2004) 'Comparison of ITS and IGS1 regions for strain typing of clinical and non-clinical isolates of Pichia anomala', *Journal of Medical Microbiology,* 53(2), 119-123.

[510] Swinne, D., Deppner, M., Laroche, R., Floch, J. J. and Kadende, P. (1989) 'Short Communication Isolation of *Cryptococcus-Neoformans* from Houses of Aids-Associated Cryptococcosis Patients in Bujumbura (Burundi)', *Aids,* 3(6), 389-390.

[511] Swinne-Desgain, D. (1976) '*Cryptococcus*-Neoformans in Crops of Pigeons Following Its Experimental Administration', *Sabouraudia-Journal of Medical and Veterinary Mycology,* 14(Nov), 313-317.

[512] Tajima, F. (1983) 'Evolutionary Relationship of DNA Sequences in Finite Populations', *Genetics,* 105(2), 437-460.

[513] Tajima, F. (1989) 'Statistical method for testing the neutral mutation hypothesis by DNA polymorphism', *Genetics,* 123(3), 585-595.

[514] Takahata, N. and Nei, M. (1985) 'Gene genealogy and variance of interpopulational nucleotide differences', *Genetics,* 110(2), 325-344.

[515] Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) 'MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0', *Molecular Biology and Evolution,* 24(8), 1596-1599.

[516] Tanabe Y., Kasai F., Watanabe M. M. (2007) 'Multilocus sequence typing (MLST) reveals high genetic diversity and clonal population structure of the toxic cyanobacterium *Microcystis aeruginosa*', *Microbiology*, 153(Pt 11):3695-703.

[517] Tavanti, A., Davidson, A. D., Johnson, E. M., Maiden, M. C. J., Shaw, D. J., Gow, N. A. R. and Odds, F. C. (2005) 'Multilocus sequence typing for differentiation of strains of Candida tropicalis', *Journal of Clinical Microbiology,* 43(11), 5593-5600.

[518] Tay, S. T., Lim, H. C., Tajuddin, T. H., Rohani, M. Y., Hamimah, H. and Thong, K. L. (2006) 'Determination of molecular types and genetic heterogeneity of *Cryptococcus neoformans* and C-*gattii* in Malaysia', *Medical Mycology,* 44(7), 617-622.

[519] Taylor, J. W. and Fisher, M. C. (2003) 'Fungal multilocus sequence typing – it's not just for bacteria', *Curr Opin Microbiol,* 6(4), 351-6.

[520] Taylor, J. W., Geiser, D. M., Burt, A. and Koufopanou, V. (1999a) 'The evolutionary biology and population genetics underlying fungal strain typing', *Clin Microbiol Rev,* 12(1), 126-46.

[521] Taylor, J. W., Jacobson, D. J. and Fisher, M. C. (1999b) 'The evolution of asexual fungi: Reproduction, speciation and classification', *Annual Review of Phytopathology,* 37, 197-246.

[522] Taylor, J. W., Jacobson, D. J., Kroken, S., Kasuga, T., Geiser, D. M., Hibbett, D. S. and Fisher, M. C. (2000) 'Phylogenetic species recognition and species concepts in fungi', *Fungal Genetics and Biology,* 31(1), 21-32.

[523] Taylor, R. L. and Duangman.C (1968) 'Occurrence of *Cryptococcus neoformans* in Thailand', *American Journal of Epidemiology,* 87(2), 318-&.

Thomas C. F. Jr., Limper A. H. (2004) '*Pneumocystis pneumonia.*' *N. Engl. J. Med.*, 350, 2487-2498

[524] Thompson, R. C. A., Constantine, C. C. and Morgan, U. M. (1998) 'Overview and significance of molecular methods: what role for molecular epidemiology?', *Parasitology,* 117, S161-S175.

[525] Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Breniere, S. F., Darde, M. L. and Ayala, F. J. (1991) 'Are eukaryotic microorganisms clonal or sexual? A population genetics vantage.', *Proceedings of the National Academy of Sciences of the United States of America,* 88(12), 5129-5133.

[526] Tibayrenc, M., Neubauer, K., Barnabe, C., Guerrini, F., Skarecky, D. and Ayala, F. J. (1993) 'Genetic characterization of 6 parasitic protozoa – parity between random-primer DNA typing and multilocus enzyme electrophoresis', *Proceedings of the National Academy of Sciences of the United States of America,* 90(4), 1335-1339.

[527] Times Live (2011) 'SA nearly wipes out infant AIDS infections.', available: `http://www.timeslive.co.za/lifestyle/health/2011/06/09/sa-nearly-wipes-out-infant-aids-infections` [accessed 10/08/2011].

[528] Titball, R. W. (1993) 'Bacterial phospholipases C', *Microbiological Reviews,* 57(2), 347-366.

[529] Titball, R. W. and Rood, J. I. (2000) 'Bacterial phospholipases', *Handbook of Experimental Pharmacology; Bacterial protein toxins,* 145, 529-556.

[530] Tojo, S., Naganuma, F., Arakawa, K. and Yokoo, S. (2000) 'Involvement of both granular cells and plasmatocytes in phagocytic reactions in the greater wax moth, Galleria mellonella', *Journal of Insect Physiology,* 46(7), 1129-1135.

[531] Torres, H. A., Prieto, V. G., Raad, I. and Kontoyiannis, D. P. (2005) 'Proven pulmonary cryptococcosis due to capsule-deficient *Cryptococcus neoformans* does not differ clinically from proven pulmonary cryptococcosis due to capsule-intact Cr. neoformans', *Mycoses,* 48(1), 21-24.

[532] Tucker, S. C. and Casadevall, A. (2002) 'Replication of *Cryptococcus* neoformans in macrophages is accompanied by phagosomal permeabilization and accumulation of vesicles containing polysaccharide in the cytoplasm', *Proceedings of the National Academy of Sciences of the United States of America,* 99(5), 3165-3170.

[533] Turenne, C. Y., Sanche, S. E., Hoban, D. J., Karlowsky, J. A. and Kabani, A. M. (2000) 'Rapid identification of fungi by using the ITS2 genetic region and an automated fluorescent capillary electrophoresis system (vol 37, pg 1846, 1999)', *Journal of Clinical Microbiology,* 38(2), 944-944.

[534] Ueno Y., Nagata S., Tsutsumi T., Hasegawa A., Watanabe M.F., Park H.D., Chen G., Yu S.Z. (1996) 'Detection of microcystin, a blue-green algal hepatotoxin, in drinking water sampled in Haimen and Fusui, endemic areas of primary liver cancer in China, by highly sensitive immunoassay' *Carcinogenesis*, 17 (6), 1317-1321.

[535] UNAIDS (2008) *Report on the global AIDS epidemic. Geneva, UNAIDS.*, Geneva.

[536] UNAIDS (2009) *AIDS epidemic update: December 2009*, Geneva.

[537] UNAIDS/WHO, (2008) 'UNAIDS/WHO Epidemiological fact sheet on HIV and AIDS: Core data on epidemiology, Thailand, 2008 Update', available: `http://apps.who.int/globalatlas/predefinedReports/EFS2008/full/EFS2008\_TH.pdf` [accessed

[538] UNDP (2004) *United Nations Development Programme: Thailand's response to HIV/AIDS: Progress and Challenges*, Bangkok.

[539] Ungchusak, K., Sriprapandh, S., Pinichapongsa, S., Kunasol, P. and Thanprasertsuk, S. (1989) 'First national sentinel seroprevalence survey of HIV-1 infection in Thailand, June, 1989', *Thai AIDS Journal,* 1(2), 57-74.

[540] Upton A., Fraser J. A., Kidd S. E., Bretz C., Bartlett K. H., Heitman J., Marr K. A. (2007) 'First contemporary case of human infection with *Cryptococcus gattii* in Puget sound: evidence for spread of the Vancouver Island outbreak', *Journal of Clinical Microbiology*, vol. 45:9, pp. 3086-3088.

[541] Urwin, R. and Maiden, M. C. J. (2003) 'Multi-locus sequence typing: a tool for global epidemiology', *Trends in Microbiology,* 11(10), 479-487.

[542] USAID (2010) *Thailand: HIV/AIDS health profile*.

[543] van Elden, L. J. R., Walenkamp, A. M. E., Lipovsky, M. M., Reiss, P., Meis, J., de Marie, S., Dankert, J. and Hoepelman, A. I. M. (2000) 'Declining number of patients with cryptococcosis in the Netherlands in the era of highly active antiretroviral therapy', *Aids,* 14(17), 2787-2788.

[544] van der Horst, C. M., Saag, M. S., Cloud, G. A., Hamill, R. J., Graybill, J. R., Sobel, J. D., Johnson, P. C., Tuazon, C. U., Kerkering, T., Moskovitz, B. L., Powderly, W.

G., Dismukes, W. E., Riser, L., Thomas, C., Lacke, C., White, A. C., Patterson, S., Davis, D., Flanigan, C., Harden, T., Phillips, D., Ullom, I., Vasquez, J., McMullen, S., Flowers, D., Bosha, M., Britton, M., Fisher, J. F., Newman, C., Willis, B., Henderson, H., Johnson, E., Dungo, L., Pierce, M., Morgan, M., Stern, J., Petrosky, N., Bamberger, D., Farnan, R., Feinberg, J., Apuzzo, L., Royal, W., Thompson, S., Barrett, K., Kauffman, C., Gutsch, H., McKinsey, D., Lee, B., Hyslop, N., Greenspan, D., Strata, R., Fessel, W. J., VanRaalte, G., Pankey, G., Schneider, D., Briggs, N. D., Pierce, D., Lee, S., Vangness, I., Longmire, B., Stansell, J., Jacobson, M., Gray, D., Coleman, R., Voorhees, J., Klebert, M., Royal, M., Mildvan, D., Friedland, G., Cooney, E., Walesky, M., Leedom, J. M., Larsen, R., Diamond, D., Wheat, L. J., Nixon, H., Craft, J., Neidig, J. L., Fass, R. J., Russell, J., MacGregor, R. R., Kappes, R., Tselis, A., Skahan, K. J., Jackson, B., Daniel, P., PaulsonWhite, C., Hewitt, R., Cohn, S., Lewis, M., Greisberger, C., Heller, H., Sugar, A., Karchmer, A. W., Fitch, H., Flam, R., Joseph, M., ElSadr, W., Soeiro, R., Stein, D., et al. (1997) 'Treatment of cryptococcal meningitis associated with the acquired immunodeficiency syndrome', *New England Journal of Medicine,* 337(1), 15-21.

[545] Varma, A., Edman, J. C. and Kwon-Chung, K. J. (1992) 'Molecular and genetic analysis of URA5 transformants of *Cryptococcus neoformans*', *Infection and Immunity,* 60(3), 1101-1108.

[546] Varma, A. and Kwon-Chung, K. J. (1992) 'DNA probe for strain typing of *Cryptococcus neoformans*', *Journal of Clinical Microbiology,* 30(11), 2960-2967.

[547] Varma, A., Swinne, D., Staib, F., Bennett, J. E. and Kwonchung, K. J. (1995) 'Diversity of DNA fingerprints in *Cryptococcus neoformans*', *Journal of Clinical Microbiology,* 33(7), 1807-1814.

[548] Vartivarian, S. E., Anaissie, E. J., Cowart, R. E., Sprigg, H. A., Tingler, M. J. and Jacobson, E. S. (1993) 'Regulation of Cryptococcal Capsular Polysaccharide by Iron', *Journal of Infectious Diseases,* 167(1), 186-190.

[549] Vecchiarelli, A. (2000) 'Immunoregulation by capsular components of *Cryptococcus neoformans*', *Medical Mycology,* 38(6), 407-417.

[550] Vecchiarelli, A., Pietrella, D., Lupo, P., Bistoni, F., McFadden, D. C. and Casadevall, A. (2003) 'The polysaccharide capsule of *Cryptococcus neoformans* interferes with human dendritic cell maturation and activation', *Journal of Leukocyte Biology,* 74(3), 370-378.

[551] Velagapudi R., Hsueh Y.P., Geunes-Boyer S., Wright J.R., Heitman J. (2009) 'Spores as infectious propagules of *Cryptococcus neoformans.*' *Infect Immun*, 77, pp. 4345-4355.

[552] Veron, V., Simon, S., Blanchet, D. and Aznar, C. (2009) 'Real-time polymerase chain reaction detection of *Cryptococcus neoformans* and *Cryptococcus gattii* in human samples', *Diagnostic Microbiology and Infectious Disease,* 65(1), 69-72.

[553] Vidotto, V., Sinicco, A., Di Fraia, D., Cardaropoli, S., Aoki, S. and Ito-Kuwa, S.

(1997) 'Phospholipase activity in *Cryptococcus neoformans*', *Mycopathologia,* 136(3), 119-123.

[554] Vincent, R. D., Goewert, R., Goldman, W. E., Kobayashi, G. S., Lambowitz, A. M. and Medoff, G. (1986) 'Classification of Histoplasma capsulatum isolates by restriction fragment polymorphisms', *Journal of Bacteriology,* 165(3), 813-818.

[555] Viviani, M. A., Cogliati, M., Esposto, M. C., Lemmer, K., Tintelnot, K., Valiente, M. F. C., Swinne, D., Velegraki, A., Velho, R. and Ecmm (2006) 'Molecular analysis of 311 *Cryptococcus neoformans* isolates from a 30-month ECMM survey of cryptococcosis in Europe', *Fems Yeast Research,* 6(4), 614-619.

[556] Viviani, M. A., Esposto, M. C., Cogliati, M., Montagna, M. T. and Wickes, B. L. (2001) 'Isolation of a *Cryptococcus neoformans* serotype A MATa strain from the Italian environment', *Medical Mycology,* 39(5), 383-386.

[557] Viviani, M. A., Wen, H., Roverselli, A., CaldarelliStefano, R., Cogliati, M., Ferrante, P. and Tortorano, A. M. (1997) 'Identification by polymerase chain reaction fingerprinting of *Cryptococcus neoformans* serotype AD', *Journal of Medical and Veterinary Mycology,* 35(5), 355-360.

[558] Vollmer, T., Stormer, M., Kleesiek, K. and Dreier, J. (2008) 'Evaluation of novel broad-range real-time PCR assay for rapid detection of human pathogenic fungi in various clinical specimens', *Journal of Clinical Microbiology,* 46(6), 1919-1926.

[559] Waldman B., Andjic V., Bishop P., Klena J. and Norman R. (2000) 'Discovery of chytridiomycosis in New Zealand', *Conf. Proceedings: Getting the Jump on Amphibian Disease*, Cairns, Australia.

[560] Wall, J. D. (2000) 'A comparison of estimators of the population recombination rate', *Molecular Biology and Evolution,* 17(1), 156-163.

[561] Wang, G., Whittam, T. S., Berg, C. M. and Berg, D. E. (1993) ' RAPD (arbitrary primer) PCR is more sensitive than multilocus enzyme electrophoresis for distinguishing related bacterial strains', *Nucleic Acids Research,* 21(25), 5930-5933.

[562] Wang, L. and Lin, X. (2011) 'Mechanisms of unisexual mating in *Cryptococcus neoformans*', *Fungal Genetics and Biology,* in press.

[563] Wang, P., Perfect, J. R. and Heitman, J. (2000) 'The G-protein beta subunit GPB1 is required for mating and haploid fruiting in *Cryptococcus neoformans*', *Molecular and Cellular Biology,* 20(1), 352-362.

[564] Watterson, G. A. (1975) 'On the number of segregating sites in genetical models without recombination', *Theoretical Population Biology,* 7(2), 256-276.

[565] Welsh, J. and McClelland, M. (1991) 'Genomic fingerprinting using arbitrarily primed PCR and a matrix of pairwise combinations of primers', *Nucleic Acids Research,* 19(19), 5275-5279.

[566] Wengenack, N. L. and Binnicker, M. J. (2009) 'Fungal Molecular Diagnostics', *Clinics in Chest Medicine,* 30(2), 391-+.

[567] Weniger, B. G., Limpakarnjanarat, K., Ungchusak, K., Thanprasertsuk, S., Choopanya, K., Vanichseni, S., Uneklabh, T., Thongcharoen, P. and Wasi, C. (1993) 'The epidemiology of HIV infection and AIDS in Thailand (AIDS. 1991;5 Suppl 2:S71-85.)', *Aids,* 7(1), 149-149.

[568] White, T. (1996) 'The future of PRC technology: diversification of technologies and applications.', *Trends Biotechnol.,* 14(12), 478-83.

[569] White, T. J., Bruns, T., Lee, S. and Taylor, J. (1990) 'Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics', *Innis, M. a. et al. (Ed.). Pcr Protocols: a Guide to Methods and Applications. Xviii* Academic Press, Inc.: San Diego, California, USA; London, England, Uk. Illus, 315-322.

[570] WHO/UNAIDS/UNICEF (2010) *Towards universal access: Scaling up priority HIV/AIDS interventions in the health sector, Progress Report 2010*.

[571] Wickes, B. L. (2002) 'The role of mating type and morphology in *Cryptococcus neoformans* pathogenesis', *Int J Med Microbiol,* 292(5-6), 313-29.

[572] Wickes, B. L., Edman, U. and Edman, J. C. (1997) 'The *Cryptococcus* neoformans STE12 alpha gene: a putative Saccharomyces cerevisiae STE12 homologue that is mating type specific', *Molecular Microbiology,* 26(5), 951-960.

[573] Wickes, B. L., Mayorga, M. E., Edman, U. and Edman, J. C. (1996) 'Dimorphism and haploid fruiting in *Cryptococcus neoformans*: Association with the alpha-mating type', *Proceedings of the National Academy of Sciences of the United States of America,* 93(14), 7327-7331.

[574] Wickes, B. L., Moore, T. D. E. and Kwon-Chung, K. J. (1994) 'Comparison of the electrophoretic karyotypes and chromosomal location of 10 genes in the 2 varieties of *Cryptococcus neoformans*', *Microbiology-Uk,* 140, 543-550.

[575] Wiemken, A. (1990) 'Trehalose in yeast, stress protectant rather than reserve carbohydrate', *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology,* 58(3), 209-217.

[576] Wilbur, L. and Heyborne, R. (2009) 'Transient loss of consciousness caused by cryptococcal meningitis in an immunocompetent patient: a case report', *Cases journal,* 2(1), 60.

[577] Williams, G. C., 1975. *Sex and Evolution.* Princeton University Press, ISBN 0-691-08152-2

[578] Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990) 'DNA polymorphisms amplified by arbitrary primers are useful as genetic markers', *Nucleic Acids Research,* 18(22), 6531-6535.

[579] Wills, E. A., Roberts, I. S., Del Poeta, M., Rivera, J., Casadevall, A., Cox, G. M. and Perfect, J. R. (2001) 'Identification and characterization of the *Cryptococcus neoformans* phosphomannose isomerase-encoding gene, MAN1, and its impact on pathogenicity', *Molecular Microbiology,* 40(3), 610-620.

[580] Wilson, D. E., Bennett, J. E. and Bailey, J. W. (1968) 'Serologic grouping of *Cryptococcus neoformans*', *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.),* 127(3), 820-3.

[581] World Bank (1998) 'Confronting AIDS: public priorities in a global epidemic', *Confronting AIDS: public priorities in a global epidemic.*, 275-276.

[582] World Bank (2000) *Thailand's response to AIDS: Building on Success, Confronting the Future*, 5.

[583] Wright, P. and Inverarity, D. (2007) 'Human immunodeficiency virus (HIV) related cryptococcal meningitis in rural central Thailand - Treatment difficulties and prevention strategies', *Southeast Asian Journal of Tropical Medicine and Public Health,* 38(1), 58-61.

[584] Xu, J. (2006) 'Fundamentals of fungal molecular population genetic analyses', *Current Issues in Molecular Biology,* 8, 75-89.

[585] Xu, J., Kerrigan, R. W., Callac, P., Horgen, P. A. and Anderson, J. B. (1997) 'Genetic structure of natural populations of Agaricus bisporus, the commercial button mushroom', *Journal of Heredity,* 88(6), 482-488.

[586] Xu, J., Manosuthi, W., Banerjee, U., Zhu, L., Chen, J., Kohno, S., Izumikawa, K., Chen, Y., Sungkanuparph, S., Harrison, T. and Fisher, M. (2010) 'Cryptococcosis in Asia' in Heitman, J., Kozel, T., Kwon-Chung, J., Perfect, J. and Casadevall, A., eds., *Cryptococcus*: From Human Pathogen to Model Yeast,Washington, DC: ASM Press, 287 - 297.

[587] Xu, J. and Mitchell, T. G. (2002) 'Strain variation and clonality in *Candida spp.* and *Cryptococcus neoformans*', *Fungal pathogenesis: principles and clinical applications*, 739-749.

[588] Xu, J. P. (2005) 'Cost of interacting with sexual partners in a facultative sexual microbe', *Genetics,* 171(4), 1597-1604.

[589] Xu, J. P. and Mitchell, T. G. (2003) 'Comparative gene genealogical analyses of strains of serotype AD identify recombination in populations of serotypes A and D in the

[590] Xu, J. P., Vilgalys, R. and Mitchell, T. G. (2000) 'Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*', *Molecular Ecology,* 9(10), 1471-1481.

[591] Xu, J. P., Yan, Z. and Guo, H. (2009) 'Divergence, hybridization, and recombination in the mitochondrial genome of the human pathogenic yeast *Cryptococcus gattii*', *Molecular Ecology,* 18(12), 2628-2642.

[592] Yamamoto, Y., Kohno, S., Koga, H., Kakeya, H., Tomono, K., Kaku, M., Yamazaki, T., Arisawa, M. and Hara, K. (1995) 'Random amplified polymorphic DNA analysis of clinically and environmentally isolated *Cryptococcus neoformans* in Nagasaki', *Journal of Clinical Microbiology,* 33(12), 3328-3332.

[593] Yan, Z., Li, X. G. and Xu, J. P. (2002) 'Geographic distribution of mating type alleles of *Cryptococcus neoformans* in four areas of the United States', *Journal of Clinical Microbiology,* 40(3), 965-972.

[594] Yauch, L. E. and Levitz, S. M. (2006) 'Interactions of immune cells with the facultative intracellular pathogen *Cryptococcus neoformans*', *Molecular Principles of Fungal Pathogenesis*,Washington: Amer Soc Microbiology, 537-554.

[595] Yu, J. J., Zheng, L., Thomas, P. W., Szaniszlo, P. J. and Cole, G. T. (1999) 'Isolation and confirmation of function of the Coccidioides immitis URA5 (orotate phosphoribosyl transferase) gene', *Gene,* 226(2), 233-242.

[596] Yue, C. L., Cavallo, L. M., Alspaugh, J. A., Wang, P., Cox, G. M., Perfect, J. R. and Heitman, J. (1999) 'The STE12 alpha homolog is required for haploid filamentation but largely dispensable for mating and virulence in *Cryptococcus* neoformans', *Genetics,* 153(4), 1601-1615.

[597] Zaragoza, O. and Casadevall, A. (2004) 'Experimental modulation of capsule size in *Cryptococcus neoformans*', *Biological Procedures Online,* 6(1), 10-15.

[598] Zaragoza, O., Chrisman, C. J., Castelli, M. V., Frases, S., Cuenca-Estrella, M., Rodriguez-Tudela, J. L. and Casadevall, A. (2008) 'Capsule enlargement in *Cryptococcus neoformans* confers resistance to oxidative stress suggesting a mechanism for intracellular survival', *Cellular Microbiology,* 10(10), 2043-2057.

[599] Zaragoza, O., Mihu, C., Casadevall, A. and Nosanchuk, J. D. (2005) 'Effect of amphotericin B on capsule and cell size in *Cryptococcus neoformans* during murine infection', *Antimicrobial Agents and Chemotherapy,* 49(10), 4358-4361.

[600] Zaragoza, O., Taborda, C. P. and Casadevall, A. (2003) 'The efficacy of complement-mediated phagocytosis of *Cryptococcus neoformans* is dependent on the location of C3 in the polysaccharide capsule and involves both direct and indirect C3-mediated interactions', *European Journal of Immunology,* 33(7), 1957-1967.

[601] Zolan, M. E. (1995) 'Chromosome-length polymorphism in fungi', *Microbiological Reviews,* 59(4), 686.

[602] Zou, C. C., Yu, Z. S., Tang, L. F., Liang, L. and Zhao, Z. Y. (2006) 'Primary abdominal lymphonodular cryptococcosis in children: 2 case reports and a literature review', *Journal of Pediatric Surgery,* 41(3).

[603] Zhou L., Yu H., Chen K. (2002) 'Relationship between microcystin in drinking water and colorectal cancer.' *Biomed. Environ. Sci.*, 15 (2), 166-171.