Genomics and Fungal Speciation: What Can We Expect in the Clinical Laboratory?

Jianping Xu

Abstract Fungal species concept has been evolving and different criteria have been used to define fungal species. Though molecular markers have been used to study human fungal pathogens for over 20 years, most clinical microbiology labs still rely on morphological and biochemical characters for fungal identifications. With increasing genetic and genomic information on fungal pathogens, such a practice is likely going to change. The rapid accumulation of genetic and genomic data, especially since the broad application of next generation DNA sequencing tools in the last few years, are providing unprecedented opportunities to identify signatures of fungal species, strains, and clinically important phenotypes such as antifungal drug susceptibilities. These advances will not only provide fundamental insights into fungal speciation, potentially unifying the different fungal species concepts, but also increase the sensitivity, specificity, and efficiency of clinical microbiology diagnosis.

J. Xu (✉)
Department of Biology, McMaster University, Hamilton, Ontario L8S 4K1, Canada
e-mail: jpxu@mcmaster.ca

Introduction

Fungi are a group of heterotrophic eukaryotic microorganisms enclosed in cell walls consisting of chitin and polysaccharides. About 100,000 fungal species have been reported and over 1000 new species are described each year [1]. Depending on the assumptions and methods of extrapolation, 1-10 million fungal species have been estimated to exist on Earth [2]. Fungi are found in almost all examined ecological niches on earth's biosphere, from the stratosphere to the bottoms of seas and oceans, from tropical rainforests to artic rocks and soils, from tranquil lakes on the Himalaya Plateau to rapidly flowing rivers, and from the skin and epidermis of animals and plants to the digestive and reproductive organs in humans [3, 4]. Morphologically, fungi are considered simple. Through most of their life cycles, fungi exist in either a unicellular (yeast) form or a multicellular (mycelial or hyphal) form. Common human yeast pathogens include *Candida albicans, Candida tropicalis, Candida parapsilosis, Candida glabrata,* and *Cryptococcus neoformans*. Common filamentous fungal pathogens include *Histoplasma capsulatum, Blastomyces dermatitidis, Coccodioides immitis*, and *Aspergillus fumigatus*.

The predominant mode of reproduction in fungi is asexual reproduction: budding or fission for yeasts and mycelial growth through apical extension and lateral branching for filamentous fungi [3–5]. However, under specific environmental conditions, many fungi can also undergo sexual reproduction through mating, nuclear fusion, meiosis, and sexual spore generation and dispersion. The fruiting bodies of many sexual fungi are highly visible to the naked eye, with some being picked and traded by humans for centuries as prized gourmet mushrooms [3]. Based on their sexual reproductive structures and genetic features, fungi are currently classified into the following seven groups/phyla: Microsporidia, Neocallimastigomycota, Blastocladiomycota,

Chitridiomycota, Glomeromycota, Zygomycota, Ascomycota, and Basidiomycota.

Among the 500 or so fungal species associated with human infections, with the majority in the Ascomycota phylum [1, 2]. While human fungal infections have been recognized for over a hundred years, their importance came into focus only over the last 30 years, primarily due to the increasing number in immunosuppressed/immunodeficient patients [2]. Indeed, most of the human fungal pathogens are opportunistic pathogens, inflicting primarily those with compromised immunity. Only about 50 of the 500 species are commonly associated with diseases in otherwise normal hosts [1]. The discoveries of a large number of opportunistic fungal pathogens have led to significant research efforts on these organisms. Coincidental with the rising importance of fungal pathogens is the rapid development of high-throughput biological research tools [6•]. Specifically, high-throughput DNA sequencing and the associated development of analytical tools are providing unprecedented insights on fungal genomes and opportunities for human fungal pathogen research. The results of such studies will have profound impact on clinical mycology.

Genomics

Genomics refers to the study of whole genomes—the complete complement of nucleic acids in a cell, an individual organism, or an environmental sample (commonly called a metagenome). Thus, genomics research investigates not only the structural and functional features of whole genomes but also the expression profiles of protein- and RNA-encoding genes. In addition, comparative genomics studies of multiple genomes from within the same species and among different species are becoming increasingly common [6•]. Specifically, such comparative studies are beginning to reveal the unique features that distinguish fungal pathogens from non-pathogens as well as between closely related pathogens [7, 8, 9•].

Speciation

The processes leading to the formation of two or more descendant species from a common ancestor is called speciation. Speciation research has been a central theme of biological research since Darwin's seminal publication over 150 years ago on the theory of evolution by natural selection [10]. Most speciation research works so far have focused on plants and animals [11] and relatively few studies have come from fungi, until very recently [12•]. Many factors have been identified that impact speciation and these factors can be grouped into three broad categories: geographic factors (allopatric speciation due to geographic barriers such as mountains, rivers, and oceans), ecological factors (e.g., host or niche specialization which may be sympatric), and genetic factors (e.g., mutation, hybridization, and horizontal gene transfer) [10, 11, 12•]. These factors may play different roles in the formation of different fungal species, with some involving factors in two or all three of the above categories.

The different processes that lead to new species formation can have significant influence on how fungal species are named and recognized. For example, if ecological factors such as substrate availability are the selective forces responsible for the divergence of the sister species, morphological and/or cross-fertility traits would likely be of limited use to discriminate such sibling species. Similarly, rapid evolution of reproductive barriers between sibling species would render mating tests very powerful methods to separate closely related biological species, while morphological and biochemical features might be of very limited utility in such cases. At present, though limited estimates based on experimental evolutionary investigations have been obtained for certain traits [13], the rates of evolution for morphological, biochemical, and reproductive traits are virtually unknown for fungi in nature.

Roles of Clinical Microbiology labs

Identifying infecting pathogens to the species level is one of the core functions of clinical microbiology labs. For some pathogens, there are systems available to identify clinical isolates to strain and genotype levels. In addition, clinical microbiology labs also test antimicrobial drug susceptibilities of the isolated pathogens. However, only those species that we can isolate and grow in pure culture are tested for drug susceptibilities. Furthermore, clinical microbiologists also monitor, analyze, and communicate the spatial and temporal trends of microbial pathogens and their drug susceptibilities to physicians and public health officials. Results from clinical microbiology labs thus provide the foundations from which hospitals and public health authorities formulate their decisions about prevention and treatment strategies for infectious diseases in individual jurisdictions.

At present, the data on fungal species and antifungal drug susceptibilities in clinical microbiology labs have been based on morphological, physiological, and growth information. However, DNA based molecular methods such as chromosomal typing through pulse field gel electrophoresis, Southern hybridization, PCR fingerprinting, and DNA sequencing have been increasingly used in human fungal pathogen research over the last two decades [3]. The molecular methods have shown several advantages over traditional methods, including high sensitivity (can detect very small amount of starting materials), excellent specificity (can discriminate closely related species and genotypes within species), and fast turnaround times. With the rapid accumulation of molecular

databases and search tools about human fungal pathogens, increasingly informative markers will become available for diagnosing fungal infections.

Fungal Species Concepts and Speciation

It is generally agreed that a species should represent a distinct evolutionary entity. However, how to define such entities remains controversial. For fungi, including human fungal pathogens, early studies relied on morphological features to define and identify species [1–4]. Some of the features include whether the organism exists in a unicellular yeast form or a multicellular mycelial form, its ability to form germ-tube and/ or pseudohyphae (for yeasts), the cell size and shape, its sexual and/or asexual reproductive structures, and colony morphology on specific medium etc. However, many species, especially yeasts, cannot be distinguished based on morphological features. As a result, polyphasic systems were introduced to include information on many different types of traits into classification [2, 4]. These include the utilization profiles of a panel of carbon and nitrogen sources, the metabolites that they produce (chemotyping), and DNA-DNA hybridization/ melting curve analyses, to complement the structural and morphological traits mentioned above. In addition, for fungal strains and species capable of mating and sexual reproduction, mating tests against standard testers of known mating types are also commonly used. Such mating tests for fungal species identification is similar to processes used for identifying species of plants and animals based on the biological species concept. However, unlike in plants and animals, many fungi cannot mate in the artificial laboratory conditions or it is difficult to observe their mating in nature. It is thus difficult to use biological species concept to define these fungi. Furthermore, due to convergent evolution and phenotypic plasticity, there is increasing evidence indicating that the specificity and sensitivity of morphological and physiological traits used to define fungal species and species complexes are relatively low. As a result, the application of evolutionary species concepts based on neutral genetic markers are attracting increasing attention.

At present, the dominant fungal evolutionary species concept is the phylogenetic species concept (PSC). PSC states that a species should represent a distinct entity containing its own unique phylogenetic signal(s) not shared with other closely related entities. In practical terms to identify such phylogenetic species, DNA sequences from multiple genetic loci are typically analyzed. Groups of strains that are always clustered together at all loci but show differences from other groups of strains would constitute a distinct phylogenetic entity (Fig. 1). In other words, these gene loci consistently separate the strains into the same groups and there was no phylogenetic incongruence at the group level. The smallest

such entity in these analyses within which recombination is frequent would be called a phylogenetic species. Thus, within a phylogenetic species, the relationships among strains as reflected by DNA sequences can be and often are different among the analyzed loci (Fig. 1). This is especially true in sexual species where mating and meiosis reshuffle allelic combinations among loci to create different genotypes.

In addition, within most phylogenetic species, the amounts of sequence divergence among alleles are typically smaller than that between species. However, there is no clear consensus as to the level of DNA sequence divergence that can be used to separate two closely related groups (e.g., varieties, serotypes, or genetically differentiated geographic or ecological populations) of organisms into different species. Population genetic analyses of strains from closely related taxa often show a continuum of sequence divergence within and between sister species. As a result, it is difficult to arrive at a consensus on sequence-based species nomenclature that applies equally to all fungal species.

Roles of Population Genetics and Genomics in Fungal Systematics and Taxonomy

One major reason for the lack of consensus on the relationship between the amount of sequence divergence and species delimitation is that fungal species have been defined differently over the last century, and with different species concepts applied to different groups of organisms. In addition, species defined based on variations in carbohydrate utilization
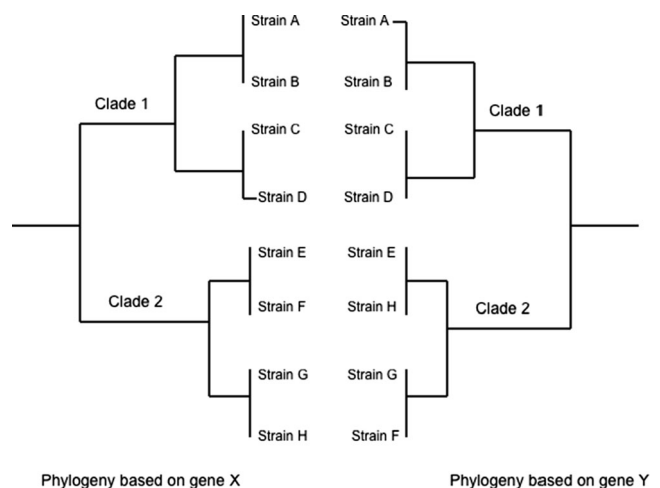

Phylogeny based on gene X   Phylogeny based on gene Y

Fig. 1 Using multiple gene genealogies to identify reproductively isolated entities. DNA sequences are obtained for two loci from eight isolates of a species complex and two gene genealogies are constructed. Isolates were separated into two clades, "clade 1" and "clade 2". Within "clade 1" the two gene trees had identical shape, consistent with clonality and continued asexual speciation. In contrast, within "clade 2", the two gene trees were not consistent, consistent with recombination and the four strains belonging to the same "biological species"

patterns will not necessary reflect the relationships based on morphological features, metabolites, or DNA sequence variations at loci unrelated to either morphological features or physiological traits. Even for sequence-based markers, different genes may show different degrees of sequence variation for both within species (e.g., due to mating and recombination) and between species (due to incomplete lineage sorting of ancestral polymorphisms and/or hybridization) [3, 14].

One approach to resolve these conflicts is through comprehensive large-scale population genetic and/or genomic studies of geographically and ecologically representative strains from closely related species. Sequencing a large number of homologous single-copy genes for strains in such populations would allow the identification of distinct phylogenetic lineages (i.e., reproductively isolated groups in nature), revealing both population genetic structure within species as well as the patterns of historical divergence among closely related species. Indeed, such studies have helped resolve the relationships among many closely related species or species complexes [3, 14]. In addition, many of the traditionally established species have also been found to contain two or more phylogenetically distinct entities (cryptic species) [15]. An additional potential benefit of these population genetic studies to clinical microbiology is the discovery of species-specific signature sequences that could be developed into markers to allow rapid discrimination of closely related species.

Comparisons between Bacterial and Fungal Genomes

Most clinical microbiology labs deal with more bacterial pathogens than fungal pathogens. For both groups of organisms, studies so far have identified that within individual (cryptic) species of microbial pathogens, evidence for both clonality and recombination are often found in natural and clinical populations [14, 16, 17]. While certain features may be shared between them, these two groups of organisms are fundamentally different from each other. For example, genomic studies have revealed the relatively constant genome size and gene content among strains within individual fungal species. This is unlike in bacteria where genomes from multiple strains of the same species often differ significantly from each other in genome size and gene content, mainly due to frequent horizontal gene transfers [18]. Furthermore, what has been found for intra-specific genome size and gene content consistencies in fungi also hold true for most other microbial eukaryotes as well as plants and animals, except polyploid hybrids.

While the exact reasons for such consistencies remain largely unknown, the genome cohesiveness among strains of the same species in eukaryotes may be related to the subcellular compartmentalization and the meiotic process during sexual reproductive cycle. Subcellular compartmentalization creates additional barriers (e.g., the nuclear and mitochondrial membranes) that could have limited the invasion of foreign genetic elements into nucleus (and mitochondria). In addition, meiosis and sexual reproduction help to eliminate rare genome structural variants that could have caused chromosomal nondisjunction, progeny inviability, and sterility. However, the effects of these processes and structures on within-species genome size and gene content stability remain to be quantified. It should also be noted that evidence for invasions of eukaryotic genomes by both endogenous and exogenous mobile genetic elements have been observed in fungi [19]. Indeed, in several plant fungal pathogens, mobile genetic elements have been proposed as responsible for host specialization and species-specific differences [20, 21].

DNA Barcoding

Because of the genome cohesiveness within individual eukaryotic species, sequence information from a single locus often can reflect the relationships among species and that the specific sequence information can be used to identify individual species. Indeed, this observation laid the foundation for the international barcode of life (iBOL) project. In animals, the core barcode locus is subunit I of the mitochondrial cytochrome C oxidase gene located in the mitochondria [22]. In terrestrial plants, a two-gene combination rbcL+matK in chloroplasts is the recommended core barcode [23]. In fungi, the core barcode is the internal transcribed spacer (ITS) regions within the nuclear ribosomal RNA gene cluster located in the nucleus [24•]. These barcodes were selected based on several factors [22, 23, 24•], including the ease from which to design broadly applicable PCR primers, the efficiency of PCR amplification, the quality of DNA sequences, the amount of sequence variation within and between closely- related species, the quality and quantity of available databases, and the specificity and selectivity of barcode sequences for identifying individual species. For many closely related species in plants, animals, and fungi, the core barcodes may be insufficient or applicable to distinguish them. As a result, sequences from other loci are often needed in order to obtain accurate species identifications [22, 23, 24•]. For human fungal pathogens, the ITS sequences are sufficient for identifying most specimens to species or species complex level. However, in a few cases, additional information from other loci is needed.

Below I illustrate how genomic analyses of the *Cryptococcus neoformans* species complex (CNSC) are impacting our understanding of fungal speciation and clinical identification.

Genomics and Speciation in CNSC

Introduction to CNSC

The CNSC has become a model for fungal pathogen research, not only for understanding the molecular mechanisms of fungal pathogenesis and host-pathogen interactions but also for revealing the fundamental ecology and evolution of fungal pathogens [25]. Interestingly, the recent genetic and genomic analyses of CNSC have also brought challenging issues for clinical microbiology labs. Some of the issues include the potential number of "species" within CNSC; the criteria that such putative "species" should be defined; and the specific markers that such "species" could be recognized by clinicians [26, 27]. As our data and knowledge about fungi expand, similar issues will likely be encountered in other fungal pathogen groups.

The taxonomy of CNSC has had a complicated history. Early studies treated CNSC as a single species *Cryptococcus neoformans*, with two varieties, var. *neoformans* and var. *gattii*. Strains in var. *neoformans* were further grouped into three serotypes A, D, and AD while those in var. *gattii* were grouped into serotypes B and C [26]. There were also strains not typable by the commercial serotyping kits (which are no longer available at present). In 2004, the original species *C. neoformans* were formally split into two species, *C. neoformans* and *C. gattii* [28].

Strains of CNSC are geographically broadly distributed and can be found in many types of ecological niches, including soil, trees, bird droppings, and other organic matter [29, 30]. Human infections by CNSC are common in certain parts of the world such as sub-Sahara Africa where a large number of immuno-compromised patients are located due to the ongoing AIDS epidemic [31]. The most common manifestation of CNSC infection is meningitis. However, other body sites can also be infected by strains of CNSC, including the skin, lung, and bloodstream [32]. While *C. neoformans* infects primarily immuno-compromised individuals and is distributed globally, *C. gattii* infections are often associated with immuno-competent hosts and are predominantly found in tropical, subtropical, and low-latitude temperate regions [29, 30]. At present, standard laboratory tests based on colony morphology and carbohydrate utilization profiles cannot distinguish the two sister species or the serotypes within each species.

Genome Variation within *C. neoformans*

Within *C. neoformans*, there are currently two varieties, var. *grubii* and var. *neoformans*, that correspond to serotypes A and D respectively as determined based on differences in cell-surface antigenic properties. Genome sequence analyses revealed that these two varieties differed by about 5-10 % at the nucleotide level [7]. Genome structure comparisons between two sequenced strains H99 (representing var. grubii) and JEC21 (representing var. neoformans) revealed a total of 32 unambiguous chromosome rearrangements, including five translocations, nine simple inversions, and 18 complex rearrangements [7]. Most of the complex rearrangement regions were located in putative centromeric regions. Interestingly, almost all the translocations and inversions were fixed within the two varieties, suggesting that these rearrangements likely predated the divergence of the current populations of these two varieties [7]. Recent analyses suggest that var. *neoformans* and var. *grubii* most likely diverged allopatrically about 20 million years ago, with var. *grubii* originated in Africa and var. *neoformans* in temperate regions in the northern hemisphere [27, 29]. Their current global distributions were due to recent anthropogenic activities that have helped spread strains of both varieties, increased their chances of contacts, and contributed to the emergence of hybrid serotype AD strains [27]. However, a genomic region of unusually low sequence divergence between the two sequenced strains (H99 and JEC21) has been found, suggesting infrequent genetic exchange has occurred between the two varieties [33].

Genome Structure Variation, Hybridization, and Speciation

Comparisons of the genome structure maps between var. *neoformans* and var. *grubii* with genetic linkage maps revealed that overall, recombination frequencies around rearranged regions were about half of those for syntenic regions [7]. However, the relative contributions of the structural rearrangements to the low spore viability in the hybrids remain to be critically examined. A high percentage (80-90 %) of basidiospores from hybrid crosses between strains of these two varieties have shown to be inviable [34•]. Genes contributing to the low spore viability would be excellent candidate speciation genes between var. *neoformans* and var. *grubii* in *C. neoformans*.

As expected, research on speciation genetics has been focused on species or species complexes defined following the biological species concept. Here, speciation genes refer to those that confer reproductive isolation between sister taxa. These genes may act pre-zygotically (e.g., preventing mating) or post-zygotically (e.g., producing non-viable or sterile progeny). Unlike in animals, in the majority of ascomycete and basidiomycete fungi, pre-zygotic reproductive barrier seemed relatively weak and mating between closely related species or varieties can occur [35]. However, the mated products (termed heterokaryons) may not be stable and/or meiosis may not proceed smoothly to generate viable offspring. Indeed, in CNSC, evidence for successful mating between strains from the two species (i.e., *C. neoformans* and *C. gattii*) and several serotype combinations have been reported in both laboratory crosses and natural environments (as natural hybrids) [29, 36, 37], consistent with weak pre-zygotic reproductive barriers in

CNSC. In contrast, the low viabilities of meiotic spores from such crosses suggest strong post-zygotic reproductive barriers in CNSC. These results also point to the quantitative nature of reproductive isolation. Indeed, absolute reproductive isolation (for both pre-zygotic and post-zygotic) is very rare between fungal sister species. At present, the genetic determinants and molecular mechanisms governing post-zygotic isolation in CNSC are unknown.

In obligate sexual organisms such as most animals, the failure to produce viable and fertile progeny would result in the extinction of the hybrids and accelerate speciation. However, in facultative sexual organisms such as most fungi, sterile hybrids may persist through asexual reproduction. For example, in C. neoformans, the serotype AD hybrids have been found in both patients and the natural environments [29, 32]. Indeed, some of these hybrids have shown superiority in several key phenotypic traits over the progenitor serotypes A and D strains, such as high temperature growth and resistance to antifungal drugs [38]. In addition, the novel gene-gene interactions and the plasticity within the hybrid genomes could accelerate the evolution and adaptation of these pathogens in both environmental and human populations.

Genomics and Species Signatures

Though most current researches on speciation genes have focused on those related to reproductive isolation, the availability of genomic sequences may also help reveal genes responsible for species-specific morphological, biochemical, and/or ecological features. However, identifying the correlation between genomic variants and phenotypic differences between species can be difficult. Instead, two types of species signatures have often been sought in genome comparisons: (i) genes that are uniquely found/lost in only one species; and (ii) genes that show unusually high levels of sequence divergence compared to the rest of the genome [12•]. Genes showing extreme sequence divergence are often under positive diversifying selection and may be related to ecological niche specialization and/or host-pathogen interactions. In plant fungal pathogens, genes coding for effectors often show significantly higher sequence divergence than the rest of the genomes and these genes commonly define the host-specific pathogen species [12•, 20, 35]. Such signature sequences could be developed into excellent molecular markers from which to identify species and strains.

Genomics and Antifungal Drug Resistance

Aside from helping identifying species and identifying the genes responsible for reproductive and ecological isolation, genomics studies can also help reveal the genes and mutations underlying antifungal drug resistance. There have been several platforms for conducting such genome-wide comparative studies between drug-resistant and drug susceptible strains, including whole genome microarrays for monitoring transcription profiles, whole genome sequencing, and whole cell transcriptome sequencing (RNA-seq). Microarrays were used frequently in model organisms until the last few years when cost effective next generation DNA sequencing platforms were introduced. In combination with traditional molecular biology approaches, these genomic studies have identified many common variants associated with resistance to the common antifungal drugs [39, 40, 41•]. Some of those variants can now be directly assayed using gene-specific markers, thus contributing to speedy diagnosis and to designing effective treatment strategies.

Metagenomics of the Human Mycobiome

A typical human body contains more microbial cells than human cells. These microbes are distributed throughout the skin and mucosal surfaces and along the digestive, respiratory and reproductive tracts. Metagenomic surveys of the human microbiome suggest that over 1000 microbial species are found in a typical healthy host [42]. Recent analyses suggest that human microbiomes play important roles not only in acute infections but also in chronic diseases such as diabetes and cancer. While most of the human microbiome consists of prokaryotic bacteria, in certain body sites such as the vaginal tract, fungal communities (the mycobiome) can be significant. This is especially true in immuno-compromised hosts [16, 17]. For example, yeast communities can dominate the oral microbial flora in AIDS patients, causing oral thrash.

Most common human fungal pathogens can be easily grown in laboratory conditions. As a result, they can be usually identified based on morphological and physiological traits of pure cultures. However, in some instances, direct metagenomic approach might be necessary for pathogen identification if the infections are caused by (i) a small number of pathogen cells that traditional methods are not sensitive enough to detect; (ii) a pathogen that no longer have live cells but still contains DNA from dead cells; and (iii) a pathogen suspected to be a biosafety level 3 organism that most clinical microbiology labs do not have the containment facility to grow. In these instances, direct amplification using conserved PCR primers targeting the 16S rRNA gene (for suspected bacterial infection) or the ITS regions (for suspected fungal infection) of the sample followed by sequencing could provide reliable diagnosis. Furthermore, such a metagenomic approach could also diagnose the pathogen(s) for other types of specimens using the next generation sequencing platforms.

Conclusions

Medical microbiology, in particular medical mycology, is undergoing a renaissance. The impetus for significant changes have come from several fronts: (i) the rising importance of fungal infections in humans; (ii) the expanding number of fungal species identified capable of causing such infections; (iii) the increasing affordability of high throughput DNA sequencing for analyzing both the genome sequences and transcriptome profiles; and (iv) the continued development of databases and bioinformatics software. These developments will allow us to develop accurate and effective molecular markers associated with species-specific and strain-specific features as well as drug-resistant profiles.

Accurately and efficiently identifying clinically important features of fungal pathogens in patients is a top priority for medical mycologists and infectious disease physicians. With a concerted effort, the goal is achievable in the near future. As discussed above, a major requirement for achieving the goal is the generation of population genetic and genomic datasets on representative populations of individual human fungal pathogens. Such datasets should also help us address several fundamental issues in fungal biology. For example, how much concordance is there between fungal species defined based on morphological, substrate utilization, chemical, and reproductive criteria? Are the correlations clade-specific or universal in fungi? How could we better define fungal species that will not only reflect their biological/evolutionary history but also be clinically informative? Should species be defined based on overall genome sequence similarities or specific signature marker(s)? If based on overall sequence similarities, what should be the cutoff point(s) and how should the cutoff points be established? For barcoding, a significant barcode gap are expected to be present between the within-species sequence variation and the among species sequence divergence. However, because different fungal species/species complexes have different evolutionary and biogeographic histories, criteria for selecting strains for analyses should be examined carefully for different groups of organisms in order to obtain accurate information on intra- and inter-species genetic variations. At present, only a few representative isolates (typically five or fewer) are used to establish the ITS barcode gaps [24•], similar to those in plants and animals [22, 23]. On the other hand, if species are identified based on signature sequences, how many and how unique should the signature sequences be? The time is right for large-scale population genetic and genomic studies that will help us address these and other issues, leading to more robust, efficient, sensitive, and specific clinical mycology diagnostic criteria.

Compliance with Ethics Guidelines

Conflict of Interest Jianping Xu has declared no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance

1. Brandt ME, Warnock DW. Taxonomy and classification of fungi. Manual of clinical microbiology. 10th ed. Washington DC: ASM Press; 2011. p. 1747–55.
2. Guarro J, Gene J, Stchigel AM. Developments in fungal taxonomy. Clin Microbiol Rev. 1999;12:454–500.
3. Xu J. Evolutionary genetics of fungi. UK: Horizon Bioscience Press; 2005.
4. Alexopoulos CJ, Mims CW, Blackwell M. Introductory mycology. 4th ed. New York: Wiley; 1996.
5. Xu J. Reproduction and sex in microorganisms. Encyclopedia of life supporting systems. UN-UNESCO. 2009.
6.• Xu J. Next generation sequencing: current technologies and applications. UK: Horizon Bioscience Press; 2014. *A forthcoming book highlighting next generation sequencing tools and its application in a diversity of biomedical research.*
7. Sun S, Xu J. Chromosomal rearrangements between serotype A and D strains in *Cryptococcus neoformans*. PLoS ONE. 2009;4(5): e5524.
8. Wei W, McCusker JH, Hyman RW, et al. Genome sequencing and comparative analysis of Saccharomyces cerevisiae strain YJM789. Proc Natl Acad Sci U S A. 2007;104(31):12825–30.
9.• D'Souza CA, Kronstad JW, Taylor G. Genome variation in cryptococcus gattii, an emerging pathogen of immunocompetent hosts. mBio. 2011;2(1):e00342-10. *An article describing the patterns of genomic variation in the Cryptococcus neoformans species complex.*
10. Darwin C. On the origin of species by natural selection. London: John Murray; 1859.
11. Coyne JA, Orr HA. Speciation. MA, USA: Sinaur Associates; 2004.
12.• Stukenbrock EH. Evolution, selection and isolation: a genomic view of speciation in fungal plant pathogens. New Phytol. 2013;199:895–907. *A review article summarizing recent advances of fungal speciation research.*
13. Xu J. Rapid evolution in experimental populations of all major life forms. In: Singh RS, Xu J, Kulathinal R, editors. Rapidly evolving genes and genetic systems. Oxford: Oxford University Press; 2012. p. 40–52.
14. Xu J. Microbial population genetics. UK: Horizon Bioscience Press; 2010.
15. Matute DR, McEwen JG, Montes BA, San-Blas G, Bagagli E, Rauscher JT, et al. Cryptic speciation and recombination in the fungus *Paracoccidioides brasiliensis* as revealed by gene genealogies. Mol Biol Evol. 2006;23:65–73.
16. Mitchell TG. Population genetics of pathogenic fungi in humans and other animals. In: Xu J, editor. Microbial population genetics. UK: Horizon Bioscience Press; 2010. p. 139–58.

17. Pujol C, Dodgson A, Soll DR. Population genetics of ascomycetes pathogenic to humans and animals. In: Evolutionary genetics of fungi. UK: Horizon Bioscience Press; 2005. p. 149–88.

18. Hao W, Golding GB. Patterns of horizontal gene transfer in bacteria. UK: Horizon Bioscience Press; 2010. p. 49–60.

19. Murat C, Payen T, Petitpierre D, Labbé J. Repeated elements in filamentous fungi with a focus on wood-decay fungi. In: Martin F, editor. The ecological genomics of fungi. UK: Wiley-Blackwell; 2013. p. 21–40.

20. de Jonge R, Bolton MD, Thomma BPHJ. How filamentous pathogens coopt plants: the ins and outs of fungal effectors. Curr Opin Plant Biol. 2011;14:400–6.

21. Raffaele S, Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Microbiol. 2012;10:417–30.

22. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identification through DNA barcodes. Proc R Soc Lond B. 2003;270: 313–21.

23. CBOL Plant Working Group. A DNA barcode for land plants. Proc Natl Acad Sci U S A. 2009;106(31):12794–7.

24.• Schoch CL, the Fungal Barcode Consortium. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci U S A. 2012;109:6241–6. *The article that formally established ITS as the barcode for fungi.*

25. Heitman J, Kozel TR, Kwon-Chung KJ, Perfect JR, Casadevall A. Cryptococcus: from human pathogen to model yeast. Washington DC: ASM Press; 2010.

26. Kwon-Chung KJ, Varma A. Do major species concepts support one, two or more species within *Cryptococcus neoformans*? FEMS Yeast Res. 2006;6:574–87.

27. Xu J, Vilgalys RJ, Mitchell TG. Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*. Mol Ecol. 2000;9:1471–81.

28. Kwon-Chung KJ, Boekhout T, Fell JW, Diaz M. Proposal to conserve the name *Cryptococcus gattii* against *C. hondurianus* and *C. bacillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae). Taxon. 2004;51:804–6.

29. Litvintseva AP, Xu J, Mitchell TG. Population structure and ecology of cryptococcus. In: Heitman J, Kozel TR, Kwon-Chung KJ, Perfect JR, Casadevall A, editors. Cryptococcus: from human pathogen to model yeast. Washington DC: ASM Press; 2011. p. 97–111.

30. Xu J, Manosuthi M, Banerjee U, Zhu LP, Chen JH, Kohno S, et al. Cryptococcosis in Asia. In: Heitman J, Kozel TR, Kwon-Chung KJ, Perfect JR, Casadevall A, editors. Cryptococcus: from human pathogen to model yeast. Washington DC: ASM Press; 2011. p. 287–97.

31. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, Chiller TM. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. AIDS. 2009;23(4):525–30.

32. Casadevall A, Perfect JR. Cryptococcus neoformans. Washington, DC: American Society for Microbiology Press; 1998.

33. Kavanaugh LA, Fraser JA, Dietrich FS. Recent evolution of the human pathogen *Cryptococcus neoformans* by intervarietal transfer of a 14-gene fragment. Mol Biol Evol. 2006;23:1879–90. doi:10.1093/molbev/msl070.

34.• Vogan AA, Khankhet J, Xu J. Evidence for mitotic recombination within the basidia of a hybrid cross of *Cryptococcus neoformans*. PLoS ONE. 2013;8(5):e62790. doi:10.1371/journal.pone.0062790. *This paper described unusual recombination processes in fungi that could expand genetic diversity in hybrids.*

35. Giraud T, Refrégier G, Le Gac M, de Vienne DM, Hood ME. Speciation in fungi. Fungal Genet Biol. 2008;45(6):791–802.

36. Xu J, Yan Z, Guo H. Divergence, hybridization, and recombination in the mitochondrial genome of the human pathogenic yeast *Cryptococcus gattii*. Mol Ecol. 2009;18:2628–42.

37. Xu J, Luo G, Vilgalys R, Brandt ME, Mitchell TG. Multiple origins of hybrid strains of *Cryptococcus neoformans* with serotype AD. Microbiology. 2002;148:203–12.

38. Shahid M, Han S, Yoell H, Xu J. Fitness distribution and transgressive segregation across 40 environmental conditions in a hybrid progeny population of the human pathogenic yeast Cryptococcus neoformans. Genome. 2008;51:272–81.

39. Oliver BG, Silver PM, White TC. Evolution of drug resistance in pathogenic fungi. In: Xu J, editor. Evolutionary genetics of fungi. UK: Horizon Bioscience Press; 2005. p. 253–88.

40. Chowdhary A, Kathuria S, Xu J, Sharma C, Sundar G, et al. Clonal expansion and emergence of environmental multiple-triazole-resistant *Aspergillus fumigatus* strains carrying the TR34/L98H mutations in the cyp51A gene in India. PLoS ONE. 2012;7:e52871.

41.• Chowdhary A, Kathuria S, Xu J, Meis JF. Emergence of azole-resistant *Aspergillus fumigatus* strains due to agricultural azole use creates an increasing threat to human health. PLoS Pathog. 2013;9(10):e1003633. *The short review summarizes the origin and rapid spread of multi-azole resistant Aspergillus fumigatus in nature and in clinics.*

42. Nelson KE, Madupu R, Szpakowski S, Goll JB, Krampis K, Methé BA. Next-generation sequencing, metagenomes and the human Microbiome. In: Xu J, editor. Next generation sequencing: current technologies and applications. UK: Horizon Bioscience Press; 2014. p. 141–52.