**REVIEW**

# The evolving species concepts used for yeasts: from phenotypes and genomes to speciation networks

Teun Boekhout[1,2] · M. Catherine Aime[3] · Dominik Begerow[4] · Toni Gabaldón[5,6,7] · Joseph Heitman[8] · Martin Kemler[4] · Kantarawee Khayhan[9] · Marc-André Lachance[10] · Edward J. Louis[11] · Sheng Sun[8] · Duong Vu[1] · Andrey Yurkov[12]

## Abstract

Here we review how evolving species concepts have been applied to understand yeast diversity. Initially, a phenotypic species concept was utilized taking into consideration morphological aspects of colonies and cells, and growth profiles. Later the biological species concept was added, which applied data from mating experiments. Biophysical measurements of DNA similarity between isolates were an early measure that became more broadly applied with the advent of sequencing technology, leading to a sequence-based species concept using comparisons of parts of the ribosomal DNA. At present phylogenetic species concepts that employ sequence data of rDNA and other genes are universally applied in fungal taxonomy, including yeasts, because various studies revealed a relatively good correlation between the biological species concept and sequence divergence. The application of genome information is becoming increasingly common, and we strongly recommend the use of complete, rather than draft genomes to improve our understanding of species and their genome and genetic dynamics. Complete genomes allow in-depth comparisons on the evolvability of genomes and, consequently, of the species to which they belong. Hybridization seems a relatively common phenomenon and has been observed in all major fungal lineages that contain yeasts. Note that hybrids may greatly differ in their post-hybridization development. Future in-depth studies, initially using some model species or complexes may shift the traditional species concept as isolated clusters of genetically compatible isolates to a cohesive speciation network in which such clusters are interconnected by genetic processes, such as hybridization.

Handling editor: Jian-Kui Liu

✉ Teun Boekhout
t.boekhout@wi.knaw.nl

1 Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

2 Institute of Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands

3 Dept Botany and Plant Pathology, College of Agriculture, Purdue University, West Lafayette, IN 47907, USA

4 Evolution of Plants and Fungi, Ruhr-University Bochum, 44801 Bochum, Germany

5 Barcelona Supercomputing Centre (BSC–CNS), Jordi Girona, 29, 08034 Barcelona, Spain

6 Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028 Barcelona, Spain

7 Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

8 Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

9 Department of Microbiology and Parasitology, Faculty of Medical Sciences, University of Phayao, Phayao 56000, Thailand

10 Department of Biology, University of Western Ontario, London, ON N6A 5B7, Canada

11 Department of Genetics and Genome Biology, Genetic Architecture of Complex Traits, University of Leicester, Leicester LE1 7RH, UK

12 German Collection of Microorganisms and Cell Cultures, Leibniz Institute DSMZ, Brunswick, Germany

# Introduction

The term 'species' is likely one of the most widely applied in biology, but it is also among those most debated. Names of yeast species are extensively used, e.g., to identify a subject analysed in scientific studies, to assess biodiversity and the intrinsic value of nature, to decide on treatment in the case of infections of plants, humans, and other animals, to protect intellectual property, to communicate information on species-specific properties, and to search related information in databases such as GenBank, PubMed or GBIF. Yet, the definition of a 'species' is still debated, especially in the areas of evolutionary biology and systematics. Many different species concepts have been applied (Table 1) that are also relevant for microbes, including yeasts (Lücking et al. 2020). Species concepts have evolved over time. In the early era of systematics, species described by John Ray, Carl Linnaeus, and Alphonse de Candolle were considered to be immutable and they used morphological and sexual features for grouping living organisms, especially plants, to place them into taxa, including genera and species (Ray 1686; Linnaeus 1758; De Candolle 1813). Since the publication of '*On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*' (Darwin 1859) various modern species concepts have emerged to address limitations of the traditional species definitions (Taylor et al. 2000; Mayden 2002; De Queiroz 2007; Aldhebiani 2018). The use of different species concepts is not a negative notion but is a consequence of an increase of scientific knowledge and the increasing types and number of data available for taxonomists.

The conceptual breakthrough in our understanding of species resulted from Charles Darwin's and Alfred Russel Wallace's thoughts on species as dynamic entities in nature that evolve over time (Darwin and Wallace 1858; Darwin 1859). This point of view was in stark contrast to the previously held belief in species as static and fixed units designed according to the taste of an almighty divine power. Next to this major philosophical paradigm shift, technological innovations helped to better understand the dynamic nature of species during their evolutionary past, but also in the present. Knowing DNA (and RNA) as the fundamental carrier of biological information, and how it changes over time, due to mutation and selection, is key to understanding mechanisms of speciation.

**Table 1** Some species concepts used in the systematics of yeasts (adapted from Taylor et al. 2000; De Queiroz 2007; Aldhebiani 2018)

| Species concept | Definition of species | Features used |
|---|---|---|
| **Phenotypic** | A set of organisms that look similar to each other and distinct from other such sets | Morphology, physiology, etc. |
| **Genetic** | Groups of isolates that share > 70% DNA similarity based on DNA reassociation experiments; 0-3 differences in the D1/D2 domains of the Large Subunit ribosomal DNA | DNA reassociation values; # nucleotide differences |
| **Biological** | Groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups | Crosses, fertility |
| Ecological | A species is a group of organisms that inhabit the same niche or habitat and that is dissimilar from other such species | Ecology, ecophysiology, fitness |
| Evolutionary | A single lineage of ancestor descendant populations of organisms which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate | All features |
| Cohesion/Genealogical | An evolutionary lineage that serves as the arena of action of basic microevolutionary forces, such as gene flow, genetic drift and natural selection | All features |
| **Phylogenetic** | A group of organisms that share unique traits, which are distinctive from other such groups, and that form a monophyletic cluster | Mainly nucleotide sequences, genomes |
| **Genealogical Concordance Phylogenetic Species Recognition** | A group of organisms that form cohesive clades for which gene phylogenies are concordant. | Nucleotide/amino acid sequences, genomes |
| Consolidated | Polyphasic approach combining aspects of phenotypic, ecological and phylogenetic species concepts | All features |

Species concepts widely used for recognizing yeast species are indicated in bold

In this review we will address: (1) the main species concepts that have been applied throughout time in yeast taxonomy; (2) the experimental approaches to test the biological species concept (BSC) and compare this with other species concepts; (3) the role and importance of hybrids; (4) the impact of comparative genomics on understanding yeast species; (5) The role and utility of DNA sequences in the practical recognition of yeast species, including barcode analysis on species identification; and (6) practicalities involved when describing yeast species, including nomenclatural aspects.

As unicellular fungi that are referred to as yeasts are a polyphyletic assemblage that occur in two fungal phyla, Ascomycota (viz. *Saccharomycotina* and *Taphrinomycotina*) and Basidiomycota (viz. *Ustilaginomycotina, Pucciniomycotina*, and *Agaricomycotina/Tremellomycetes*), we will present examples from these diverse lineages of fungi. Finally, we present a scheme that recognizes three levels of knowledge with increased confidence for species recognition and takes into account the genetic and evolutionary complexity of species (Fig. 1).
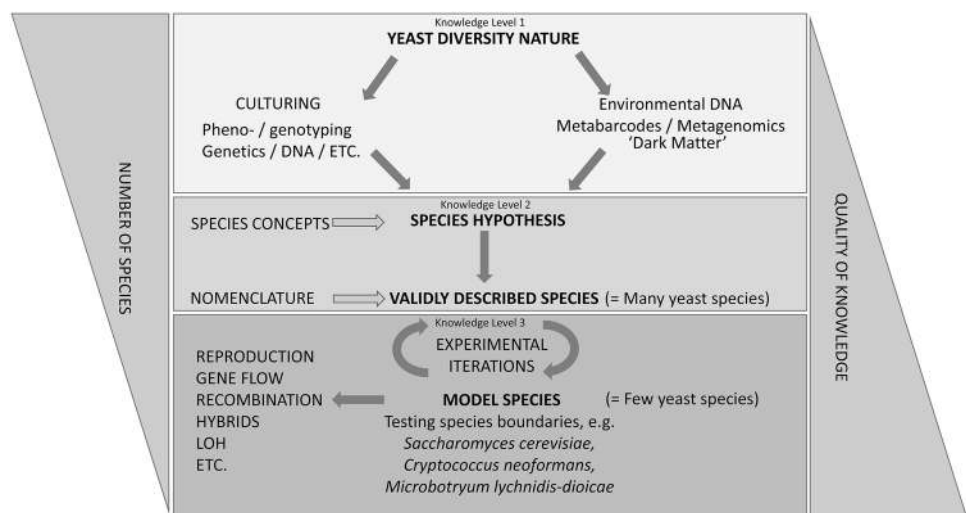
## Species concepts

Systematists have the mandate of discovering and delineating new species, documenting their properties, assigning them correctly to higher taxa, and providing means for accurate identification. These tasks are interrelated but distinct. In its simplest iteration, the description of a yeast species must entail the valid publication of a protologue that specifies a type specimen to which a binomial is permanently attached, as specified by the '*International Code of Nomenclature for Algae, Fungi, and Plants*' (ICNafp, Turland et al. 2018). Ideally, a good species description should go well beyond the mere fulfilment of formal requirements of the

ICNafp (Aime et al. 2021). Linnaeus (1758) famously stated that the species is the work of nature, while much of everything else in systematics is a human construct. The heart of the species description, therefore, is the demonstration that a strain or collection of strains represent a distinct species, a topic worthy of its own discussion.

Several species concepts have been applied to describe and interpret yeast diversity (Table 1), the most important being: 1. The Phenotypic Species Concept (PhenSC), with its molecular variant, the Genetic Species Concept (GSC). Ecological parameters have been used to recognize ecological species that represents a particular variety of the PhenSC under the so-called Ecological Species Concept (Kudryavtsev 1960; Barnett 2004). Growth properties on e.g. high sugar containing or high salt-contaning media, as well as nutritional growth profiles traditionally used to characterize yeast species can be seen as such ecological parameters, and the same holds for speciation due to host specificity; 2. The Biological Species Concept (BSC); and 3. The Phylogenetic Species Concept (PSC) with its operational version, the Genealogical Concordance Phylogenetic Species recognition (GCPSR). Recently, also the Consolidated Species Concept has been proposed that uses a polyphasic approach combining aspects of the PhenSC and PSC (Hawksworth and Lücking 2017; Quaedvlieg et al. 2014). This seems a pragmatic species concept that integrates various kinds of data into an operational species. It has also been emphasized that species boundaries are porous/semipermeable, at least under certain conditions (Hausdorf 2011; Steenkamp et al. 2017). This porosity may be due to cross-species gene flow between recently diverged species or closely related allopatric species and may confer a selective advantage. The authors listed hybridisation, introgressive hybridisation and horizontal gene flow as a consequence of this porosity/semipermeability (Steenkamp et al. 2017). Recently, the Differential Fitness Species Concept (DFSC) has been proposed as "groups

**Fig. 1** Scheme showing increase of knowledge of yeast biodiversity with three levels of knowledge on certainty of species recognition

of individuals that are reciprocally characterized by features that would have negative fitness effects in other groups and that cannot be regularly exchanged between groups upon contact" (Hausdorf 2011). This species concept considers all genetic changes that confer differences in fitness and differs from the BSC in that the exchange of species-specific features is not only restricted by reproductive isolation, but also by selection (Hausdorf 2011). As far as we know the DFSC has not yet been applied to fungi, but it is likely that some model species studied in speciation research, such as *Saccharomyces cerevisiae*, will be useful to test its relevance. Very recently, it has been proposed that variation in homoplasy of homologous DNA sequences when compared within and between species can be used as an additional feature to recognize species (Conti et al. 2021). Examples came from the genera *Candida, Debaryomyces, Kazachstania* and *Saccharomyces* (Conti et al. 2021).

For yeasts, the PhenSC was widely used before the introduction of molecular data (Barnett 2004; Giraud et al. 2008; Taylor et al. 2000). Yeast species have been recognized for decades based on phenotypic properties, including colony and cellular morphologies, growth requirements using arrays of different carbon and nitrogen sources, fermentation of various sugars, growth at different temperatures, starch formation, or gelatine liquefaction (Kurtzman et al. 2011b).

Features of DNA were introduced in the 1970s as quantitative criteria to distinguish and compare yeast species when Phaff and co-workers initiated the analysis of the guanine and cytosine content (mol% G+C) and DNA reassociation studies. From these studies it emerged that strains differing by 1–2 mol % G+C likely represented different species (Martini and Phaff 1973; Price et al. 1978; reviewed in Kurtzman et al. 2011c; Lachance 2018). Along similar lines, yeast strains differing by a ≤ 80% DNA reassociation value (later lowered to 70%) were considered to be conspecific, and values ranging between 40 and 70% were considered to indicate varieties or subspecies (Kurtzman 1987, 1998). In the early 1990s sequence analysis of parts of the ribosomal RNA (rRNA) gene, also referred to as ribosomal DNA (rDNA), was introduced (e.g., Peterson and Kurtzman 1991). These authors studied the D2 domain of the Large Subunit (LSU) rRNA and suggested that conspecific strains showed < 1% nucleotide divergence. Kurtzman and Robnett (1998) studied sequence variation in the D1/D2 domains of the LSU rDNA of ca. 500 species of ascomycetous yeasts and compared the application of the Biological Species Concept (BSC, see below) with the amount of sequence divergence present in the LSU rDNA D1/D2 domains. In the chapter '*Gene sequence analyses and other DNA-based methods for yeast species recognition*' published in the 5th edition of '*The Yeasts, a Taxonomic Study*' ('*The Yeasts*') the main findings were summarized as follows: conspecifics may differ by up to three nucleotide substitutions and species

differ with six or more substitutions of the approximate 600 nucleotides of the D1/D2 domains of the LSU rDNA (Kurtzman et al. 2011c). This of course leaves room for the interpretation of strains that differ with four or five nucleotides. Similar data were obtained by Fell et al. (2000) for basidiomycetous yeasts and these authors stated that conspecific strains had ≤ 3 nucleotides and different species ≥ 6 nucleotides in the D1/D2 domains. This interpretation leaves a gap for the range of 4–6 nucleotide differences in the D1/D2 domains of the LSU rDNA. In a follow-up study including both D1/D2 domains and the Internal Transcribed Spacer (ITS) 1 and 2, including the 5.8S rRNA gene, Scorzetti et al. (2002) came to different conclusions. They wrote "a boundary between species (a prerequisite number of nucleotide differences in either region to separate species) was not defined by our study. The answer undoubtedly lies not in the number of differences, but rather in the significance and function of the site mutations." In Table 10.1 of the 5th edition of *The Yeasts*, however, the following values were presented: % nDNA relatedness 70–100% correlated with 0–3 nucleotide differences in the D1/D2 domains of LSU rDNA implying conspecificity; 0–20% nDNA relatedness correlated with 6–250 nucleotide differences in the D1/D2 domains of LSU rDNA indicating different species (Kurtzman et al. 2011c). The above distance measures have been used by the late Cletus P. Kurtzman to introduce a Genetic Species Concept (GSC) that circumscribes yeast species based on a measure of genetic distance obtained indirectly from DNA reassociation kinetics or directly from the comparison of aligned sequences. Thus, the GSC holds a position intermediate to the Phenotypic-, Biological-, and Phylogenetic species concepts, and is best interpreted as an operational tool to detect and circumscribe potential new species. However, it is fair to state that values such as the number of nucleotide differences that predict species boundaries should only be used as a first approximation of what might represent actual yeast species. Lachance (2018) emphasized the lack of a thorough statistical analysis of such threshold values. In addition, some notable exceptions to this rule of thumb have been observed, e.g., the D2 domain of mating strains of *Clavispora lusitaniae* showed up to 32 nucleotide substitutions (Lachance et al. 2003).

In addition to the D1/D2 domains of LSU rDNA, other regions of the rRNA gene cluster have been used, e.g., the Small Subunit (SSU) rRNA gene, the ITS 1 + 2, and the Intergenic Spacer (IGS). Yeast taxonomists have also used partial sequences of protein-coding genes such as the Translation Elongation Factor 1α (*TEF1*) and RNA- and DNA-directed RNA polymerase II subunits (*RPB1* and *RBP*2). The phylogenetic interpretation of such sequence data has led to the use of the Phylogenetic Species Concept (PSC) in yeasts. In most cases, however, phylogenetic analyses have taken a secondary role, as sequence differences have

been interpreted as distance measures for an operational species recognition using the GSC. Taylor and collaborators presented a seminal paper on species concepts in use in mycology (Taylor et al. 2000). These authors proposed the Genealogical Concordance Phylogenetic Species Recognition (GCPSR), which compares genealogies resulting from several unrelated genes to infer species boundaries for fungi. Thus, the generation of multiple unlinked gene sequences is required for application of the GCPSR; see, e.g., Hagen et al. (2015). The analysis of multiple loci (a.k.a. Multi-Locus Sequence Analysis, MLSA) is likely a transition towards the use of fragmented or, preferably, entire genomes (see below) for species recognition.

## The biological species concept (BSC)

The BSC holds a central position in eukaryote biology and thus also in mycology, as it is based on the estimation of gene flow among and between populations that may, or may not, belong to different species. Full application of the BSC is in principle limited to organisms that depend on sexual reproduction for their reproductive success. *Prima facie,* yeasts would fare poorly in this context, as they propagate primarily by mitotic division. However, the prolificity of a species extends well beyond numbers, and the greatest benefit of sexual reproduction lies not in propagation, but in the fitness benefits accrued through genetic recombination (McDonald et al. 2016). A large proportion (45%) of ascomycetous yeast species are known exclusively from their asexual state (Kurtzman et al. 2011c), although many apparently asexual species may in fact be known from a single haploid mating type due to insufficient sampling. In many cases, this question can be answered by the characterization of mating genes in their genomes (Yurkov et al. 2015; Kijpornyongpan et al. 2018; Krassowski et al. 2019; Passer et al. 2019). Below we will explore the impact of the BSC in various fungal lineages to which yeasts belong.

### Lessons from *Saccharomyces, Saccharomycotina, Ascomycota*

The biological species concept has been used in *Saccharomyces* species for decades, well before genome sequences became available and before rRNA gene and ITS sequences were used routinely (Naumov 1987). Most wild isolates of *Saccharomyces* spp. [also referred to simply as 'yeast' or 'budding yeast' in many publications] are homothallic, due to the action of the HO endonuclease on the mating type locus, resulting in haploid mother cells switching mating type and forming diploid zygotes with daughter cells or undergoing mating between sister spores due to proximity (see for example Greig 2009; Ono et al. 2020). Despite

this, mating between different isolates can be accomplished using spore-to-spore mating (Naumov 1987) which can be enhanced with auxotrophic markers, allowing selection and confirmation of zygotes. Using tester strains from known species, crosses to strains of unknown lineage result in viable spores at reasonable frequencies (> 50%) in conspecifics (same species), or very low spore viability (~ 1% or less), between different species, due to chromosomal abnormalities or genetic problems leading to infertility. Several new species were identified solely on this basis (Naumov et al. 1995a, b), and their boundaries were subsequently supported by ITS sequence comparisons (Naumov et al. 2000). Whole genome sequences further refine species definitions by identifying hybrids, introgression, or cases of reproductive isolation due to Gross Chromosomal Rearrangements (GCRs) without notable sequence divergence, as is the case with *S. cariocanus*, now considered part of the American population of *S. paradoxus* but differing by four reciprocal translocations. Eight *Saccharomyces* species are now recognised, defined by a combination of the BSC and whole genome sequence comparisons (Dujon and Louis 2017; Ono et al. 2020).

Post-zygotic reproductive isolation can be caused by several mechanisms. GCRs can lead to unbalanced chromosomal segments and lethality as is the case with *S. cariocanus*, whose four reciprocal translocations theoretically reduce spore viability in heterozygotes to around 6%, close to the observed value (Liti et al. 2006; Dujon and Louis 2017). Other examples are found, such as the Malaysian population of *S. cerevisiae* (Liti et al. 2009) and other *S. cerevisiae* strains surveyed (Hou et al. 2014). Although GCRs do exist between many of the species, the genome of some distinct species are colinear, as for example *S. cerevisiae* and *S. paradoxus*, such that GCRs do not explain their reproductive isolation. A second mechanism involves Bateson–Dobshansky–Muller incompatibilities where genetic interactions cause sterility. There are certainly a few well-characterised examples of existing incompatibilities, particularly those observed between the nuclear and mitochondrial genomes (Lee et al. 2008), but these have been difficult to find, and it is not clear that they occurred prior to speciation (Louis 2009, 2011). The third mechanism is sequence divergence acted upon by the mismatch repair system during meiosis (Chambers et al. 1996; Hunter et al. 1996; Greig et al. 2003), where recombination intermediates are aborted resulting in lack of essential crossovers between homeologues required for proper chromosome segregation.

### Lessons from other *Saccharomycotina, Ascomycota*

Among yeasts with a demonstrated sexual cycle, only 24% have been shown to have fixed mating types (heterothallic) that would ensure that a sexual cycle has the potential to

allow full participation in a common gene pool. In fact, a substantial proportion of heterothallic yeast species exist in nature as stable diploids whose meiotic division leads to the formation of short-lived ascospores that engage in sister-pair matings soon after gemination. In such cases, the potential for recombination may not differ much from that of strictly asexual species. Yeasts epitomize the paradox of sex. Speijer et al. (2015) argue that sexual reproduction lies at the very essence of what it is to be an eukaryote. Yet, yeasts have repeatedly evolved multiple, elegant mechanisms that enable continuous mating-type switching, which would tend to reduce genetic diversity in a population instead of fostering it. An analysis of 332 genomes (Krassowski et al. 2019) found that 42% of the sexual species studied at the genome level exhibit one of five forms of homothallism, which allows sexual reproduction within an essentially clonal population. In haplontic species, this can be accomplished by mating type switching within a population, which allows the conjugation of sister cells, as in *Zygosaccharomyces* species, or by fusion of sister nuclei, one of which returns to its mother cell from an aborted bud, as typified by *Debaryomyces*, *Torulaspora*, or *Nadsonia* species (Kurtzman et al. 2011c). The possibility exists that ascosporulation in such cases has been co-opted for dormancy or increased resistance to environmental stress (e.g., passage through insects guts). These mechanisms do not entirely preclude outcrossing but make it difficult to use the BSC as a criterion for species recognition. For example, the isolation of recombinants in hybrid crosses caused van der Walt and Johannsen (1984) to reduce several species of *Kluyveromyces* to varieties of *K. marxianus*, but later studies based primarily on genetic divergence (Lachance 2011) provided evidence that assignment to three distinct species, *K. marxianus*, *K. lactis*, and *K. dobzhanskii*, was a better reflection of biology. Naumov et al. (2016) found a good correspondence between hybrid fertility and phylogenetic relatedness in species of the homothallic genus *Komagataella*. Interestingly, the genus was used by Naumov (2015) as an example of a "genetic genus", where prezygotic isolation is used to delineate genera, a concept that he extended also to *Saccharomyces*, *Kluyveromyces*, *Arthroascus* (a.k.a. *Saccharomycopsis*), *Galactomyces*, *Metschnikowia*, *Phaffomyces*, *Williopsis*, and *Zygowilliopsis* (the last two are currently assigned to *Cyberlindnera* and *Barnettozyma*, respectively).

In a study of mating success among 34 haplontic heterothallic species of *Metschnikowia*, Lee et al. (2018) determined that the ability to form mating pairs depended primarily on recognition of the α pheromone by cells of mating type **a**. A similar conclusion had been reached with mostly homothallic species of *Saccharomyces* and closely related genera (Rogers et al. 2015). In haplontic *Metschnikowia* species, completion of meiosis and ascosporulation only occur between mating pairs that exhibited little DNA

sequence divergence (GSC). In this group of species, viable ascospores are formed in pairs, and asci arising from hybrid crosses give rise to large proportions of single-spored or empty asci. The single spores are generally sterile (Lachance and Bowles 2004; Marinoni and Lachance 2004). When recently diverged species are examined (Lee et al. 2020), a small amount of hybrid fertility may be observed, and the distinction becomes less clear. All other species examined in the *Metschnikowiaceae* share the same arrangement of mating genes and occur as haploid mating types or heterozygous diploids. Intriguingly, Muñoz et al. (2018) observed that isolates of *Candida auris* belong to distinct geographic clades, each of which features only one of the two known mating types. One has yet to report any attempts to mate the strains and observe ascospore formation in *C. auris*. Moving farther out into the larger CUG-ser-I clade, a few closely related species of *Kodamaea* with heterothallic haplontic life cycles do not exhibit any signs of hybridization (Lachance et al. 1999). Farther out yet, species of *Zygoascus* show a positive correlation between the formation of mature asci in crosses between the haploid strains and their DNA reassociation values, although many anomalies exist (Smith et al. 2005). In that case, interfertility served as an adjunct criterion for species circumscription.

One may rightly ask whether the BSC is in fact a genuine indicator that yeasts capable of fertile mating in the laboratory also belong to Mendelian populations in nature, which would contribute to maintaining their cohesion as evolutionary species. Five species of haplontic *Metschnikowia* have been examined for this very purpose. The endemic species *M. borealis*, *M. lochheadii*, *M. hawaiiensis*, and *M. hamakuensis*, which mate readily on agar media, exhibit independence among several genetic markers (Lachance et al. 2008, 2016; Wardlaw et al. 2009). In contrast, *M. ipomoeae*, characterized by a very low propensity for entering the sexual cycle, exhibits sharp linkage disequilibria and for all practical purposes, evolves clonally. A low but significant amount of disequilibrium was observed in *Kurtzmaniella cleridarum* (*Debaryomycetacae*, CUG-Ser-I clade), which can be accounted for by a less avid mating system and a larger separation between habitats, namely, beetles that live in cactus flowers in the Sonoran and Mojave deserts (Lachance et al. 2013).

## Lessons from *Cryptococcus*, *Tremellomycetes*, *Agaricomycotina*, *Basidiomycota*

In the *Basidiomycota*, sexual processes needed to fulfil the BSC are governed by mating types that are defined by two mating-type loci: the pheromone and pheromone receptor (*P/R*) loci that governs mating initiation, conjugation, and clamp cell fusion, and the *HD* locus that contains transcription factors, which regulates gene expression upon

successful zygote formation to initiate downstream sexual developments, including hyphal growth, basidia formation and maturation, as well as sporulation (Ni et al. 2011; Heitman et al. 2013; Coelho et al. 2017; Sun et al. 2019a). Thus, successful sexual reproduction occurs typically between cells that possess divergent and compatible alleles at both the *P/R* and *HD* loci. When the two *MAT* loci are unlinked and segregate independently during meiosis, the species is defined as having a tetrapolar mating system; when the two *MAT* loci are linked, either physically or genetically, the species is determined to have a bipolar mating system. Transitions between tetrapolar and bipolar mating systems have occurred during the evolution of basidiomycetous species, and they involved loss of mating specificity of one of the two *MAT* loci, or chromosomal translocation leading to physical linkage of the two *MAT* loci (James et al. 2006; Wu et al. 2015; Sun et al. 2017, 2019a, b; Zhu et al. 2017).

While sequence similarity can be employed as a good indicator of whether two isolates with a certain level of sequence divergence belong to different species, it alone does not necessarily prove whether or not two isolates with diverging genomes represent fully established, reproductively isolated species. As in all other eukaryotes, speciation in *Basidiomycota* requires reproductive isolation and restricted gene flow between lineages, the extent of which is however not well studied. Reproductive barriers maintaining such speciation can be prezygotic, where sexual reproduction is prevented by geographical isolation or incompatible pheromone/pheromone receptor interaction; these barriers can also be postzygotic where the production of viable mating progeny is prevented due to genetic variation existing between the different species, leading to either failed meiosis or meiotic products that have compromised genetic compositions that render them inviable. For example, in the human pathogenic *Cryptococcus* species complex, at least eight species are currently recognized using PSC/GCPSR that belong to three sub-complexes: *Cr. neoformans*, *Cr. deneoformans*, and *Cr. gattii* (Hagen et al. 2015; Farrer et al. 2019; Bahn et al. 2020). Genome analyses have shown that significant genetic variation has accumulated between these different species and that each species forms a well-supported clade with limited inter-specific gene flow, consistent with the existence of reproductive isolation between these species (Sun and Xu 2009; Desjardins et al. 2017; Rhodes et al. 2017). Thus, this is a nice example where the PSC-GCPSR/BSC and GSC largely coincide, while also phenotypic differences occur between these species, thus there is also support using the PhenSC. Interestingly, mating can still occur between certain *Cryptococcus* species and produce sexual structures that are morphologically indistinguishable from those arising from canonical intra-species mating, with fused clamp cells and abundant meiotic spores, suggesting that complete prezygotic reproductive isolation has yet to be established between these species. However, further analyses of the meiotic progeny revealed that they have very low germination rates and most of the viable spores are either aneuploid or diploid, suggesting that meiosis during interspecies mating has been significantly compromised, which is consistent with the fact that both high-level DNA sequence divergence and significant chromosomal rearrangements (e.g., inversions and translocations) are present between different *Cryptococcus* species. These divergent genomic properties collectively impede proper chromosomal pairing and recombination during meiosis and promote chromosomal mis-segregation (Sun and Xu 2007).

Gene flow can only be detected in population genetics/genomics data, which presents a problem when only a limited number of isolates of a species are available. In a recent study of four isolates that represent three non-pathogenic *Cryptococcus* species, *Cr. amylolentus*, *Cr. floricola*, and *Cr. wingfieldii*, which are closely related to the pathogenic *Cryptococcus* species complexes, the authors first obtained chromosome-level genome assemblies using PacBio and Nanopore long-read sequencing and de novo genome assembly. Subsequent genome comparisons revealed that while these non-pathogenic species share 93.5–94.4% sequence similarity, extensive chromosomal rearrangements, including inversions and translocations, are present, suggesting that they could represent distinct species. Additionally, these species have *MAT* loci configurations consistent with a tetrapolar mating system and possess divergent *MAT* alleles, indicating they should be mating compatible with each other, as confirmed by previous studies of the two *Cr. amylolentus* isolates (Findley et al. 2012). However, successful mating was only observed between *Cr. amylolentus* and *Cr. floricola*, and their meiotic progeny were either inviable or sterile, and showed reduced recombination during meiosis, suggesting postzygotic barriers had been established between the two species. Furthermore, physiological characterization of these three species revealed they have distinguishing phenotypic profiles that might reflect the differences in their natural habitats. Thus, multiple lines of evidence collectively support that *Cr. amylolentus*, *Cr. floricola*, and *Cr. wingfieldii* represent distinct species (Passer et al. 2019). A good correlation between BSC, PhenSC and PSC was observed in that study.

It should be noted that in some cases, isolates belonging to the same species with compatible mating types based on genome analyses fail to undergo sexual reproduction. It is possible that the laboratory conditions are not ideal to induce mating, or that the isolates might be sterile due to changes in the genes required for mating. Because sexual reproduction is not obligate in most basidiomycetous fungi, sterility is disadvantageous in the short term only if those genetic changes have pleiotropic effects that reduce the hosts' vegetative fitness, although the inability to undergo sexual reproduction is almost always detrimental during long-term evolution.

As mentioned earlier, both DNA sequence divergence and chromosomal rearrangements contribute to compromised meiosis and the establishment of postzygotic reproductive isolation between species. Because these two types of genetic variation almost always exist simultaneously when different species are compared, it has been difficult to dissect the effect of one in the absence of the other. In a recent study of *Cr. neoformans*, the authors induced double-strand breaks (DSBs) at the centromeres using CRISPR, and the subsequent repair of these DSBs by the cell resulted in a strain in which a large number of chromosomal arms were reshuffled, while all of the genetic information in the genome had been retained (Yadav et al. 2020). Genetic crosses between the original *Cr. neoformans* strain and the strain with a reshuffled genome revealed that while mating could still occur, the crosses produced almost no spores, suggesting chromosomal structural variation alone could be sufficient to establish a postzygotic reproductive barrier, although both sequence divergence and chromosomal rearrangements likely impose reproductive barriers collectively between closely related species in nature (Yadav et al. 2020).

There are also factors that enable species to overcome postzygotic reproductive barriers. For example, it has been shown in yeast species that the mismatch repair system is involved in repressing homeologous recombination between divergent sequences, and thus preventing chromosomal pairing and crossing over during hybridization, leading to compromised meiosis and chromosomal mis-segregation. Deletion of key genes in the mismatch repair pathway, such as *MSH2*, could significantly increase the spore viability during inter-species hybridization (Chambers et al. 1996; Hunter et al. 1996). It has been shown recently that in *Cryptococcus* deletion of *MSH2* also leads to increased spore viability during inter-species hybridization, even without significantly elevated recombination frequency during meiosis (Priest et al., 2021). Additionally, while sequence divergence and chromosomal rearrangements are detrimental during meiosis when the two isolates involved in the hybridization are haploid, diploidization of both isolates could significantly alleviate such deleterious effects because in this case all of the chromosomes exist in identical pairs and thus, are able to form homologous chromosomal pairs during meiosis I (Greig et al. 2002). Whether such factors also exist in basidiomycetous species has yet to be determined.

## Lessons from *Ustilaginomycotina, Basidiomycota* and *Microbotryales, Pucciniomycotina, Basidiomycota*

Classically, the 'smuts' comprise a group of fungi that are host-specific plant parasites that infect a large range of plant taxa. They exhibit a typical life cycle, whereby a basidium germinates from a diploid teliospore and produces haploid basidiospores, which can reproduce asexually by budding (Begerow et al. 2014). These asexually produced yeasts are also called sporidia. To initiate the plant parasitic phase, two basidiospores/sporidia with compatible mating types need to fuse on a suitable host and produce a dikaryotic infection hypha. After successful penetration of the plant epidermis through an appressorium the fungus proliferates within its host as a mycelium. At a given developmental stage of the host, the fungus starts forming spore-producing structures (i.e., sorus; pl. sori), which at maturity erupt from the plant tissue. In the majority of cases, sori are plant tissue specific for a given smut fungal species (Schäfer et al. 2010; Begerow et al. 2014). The combination of host plant and place of soral development together with other morphological characteristics (e.g., spore mass colour) are used for species identification in the field. The systematics and understanding of evolution of smut fungi have been mainly based on these characteristics, as well as on the additional morphological traits spore size and spore ornamentation (Vánky 2012), thus implying the applications of the phenotypic and ecological species concepts.

However, recent years have seen a revolution in the understanding of smut fungus evolution and systematics that has been driven by application of molecular data. The smut fungi themselves have been recognized as a polyphyletic group and fungi with the smut life history have evolved in both *Ustilaginomycotina* and the *Microbotryales* (*Pucciniomycotina*). In addition, it is now known that many species of *Ustilaginomycotina* appear to be yeasts without a known parasitic stage (e.g., *Malasseziomycetes* and *Moniliellomycetes*) (Begerow et al. 2000; Wang et al. 2014, 2015a) where, by default, the BSC cannot be applied. Because of the different life history strategies employed within *Ustilaginomycotina*, no single species concept has been applied within this group. While the PhenSC is still relevant for species delimitation and identification, it does not provide a conceptual idea of what a species actually represents, and which mechanisms maintain its cohesion. Species recognition of 'smuts' is mainly based on the PSC using molecular data, but also incorporating host taxa as an autapomorphy, thus leading to an integrative species concept. As with the phenotypic concept, a strict application of the PSC does not readily incorporate biological data. Therefore, investigating the BSC is the basis for studies focusing on mechanisms involved in the maintenance of species integrity in smut fungi. Several research lines related to the BSC thereby have received most attention: a. Function and evolution of smut fungal mating systems; b. Host adaptation as a mechanism important in creating and maintaining species; and c. The role of asexual lineages in *Ustilaginomycotina*. Aspects of hybridization of smut fungi are discussed in the appropriate text below.

## Function and evolution of smut fungal mating systems

Like most members of the *Basidiomycota*, smut fungi use a mating system consisting of two relevant factors. The *MAT* locus represents a pheromone/pheromone receptor (*P/R* locus) system, where one strain expresses the pheromone that interacts with the pheromone receptor of a second strain. The second locus consists of a homeodomain transcription factor formed by a heterodimer (*HD* locus), where both subunits are expressed by two different strains. While the *P/R* locus is essential for finding the mating partner of a compatible mating type, the HD locus confers the maintenance of the dikaryon and the proper development of hyphae (Bölker et al. 1992; Gillesen et al. 1992; Lee et al. 2010; Coelho et al. 2017). Both factors are involved in the development of the parasitic phase (Bölker et al. 1995; Kämper et al. 2020). All smut fungi are parasitic in the dikaryotic phase only and therefore mating genes were identified as virulence factors (Brachmann et al. 2001; Kahmann and Schirawski 2007). Current research indicates that both factors are part of all major lineages and are probably similar in function (Kijpornyongpan and Aime 2020). Although central to the maintenance of species, the mating system has undergone several modifications during the evolution of *Basidiomycota*. For instance, the number of alleles and the linkage of these factors varies significantly (Kellner et al. 2011). Many studies have focused on the presence and absence of bipolar and tetrapolar mating systems, which result from a recombination block between the two mating loci and transitions between the two mating systems seem quite common (Lee et al. 2010). Comparative genomics studies of smut fungi revealed a broad spectrum of evolutionary variations, including chromosome rearrangements and recombination suppression (Branco et al. 2017; Sun et al. 2017; Kijpornyongpan and Aime 2020; Yadav et al. 2020), and multi-allelic mating loci having tri-allelic pheromone loci (Kellner et al. 2011). Only a few species and lineages have been studied in this respect so far, but the implications for the BSC are already complex. Like in other *Basidiomycota* two haploid genotypes with compatible mating types fuse to form a pathogenic dikaryon. On the molecular level, however, smut fungi have evolved a plethora of mating systems to enable successful mating between individuals, all of which maintain the main function, i.e. the creation of new genotypes.

## Host adaptation as a mechanism important in creating and maintaining species

Reproductive isolation of well-adapted gene pools protects them from the invasion of non-adapted genetic material and is central to the BSC (Mayr 1969, as cited in Mayr 1996). Detecting and understanding mechanisms that generate and maintain this isolation are therefore a central point in research on biological species. Most *Ustilaginomycotina* and *Microbotryales* are parasitic and the host-parasite interaction has most likely resulted in a high host specificity (Begerow et al. 2014; Hartmann et al. 2019), and as such this fits the Ecological Species Concept. While co-evolution has been used to explain host specificity of smuts (Begerow et al. 2004), host jumps are also quite common, not only on closely related hosts, but also between different plant families (Kemler et al. 2020). In some cases, two or more species have adapted to a single host species and can co-occur in the same host population without a disintegration of parasite species delimitations (Begerow et al. 2002; Abbate et al. 2018). Studies focusing on virulence and pathogenicity factors are trying to identify factors conveying host specificity (e.g., Ökmen et al. 2018). However, as genes responsible for host adaptation are not necessarily involved in sexual reproduction, the way in which host adaptation plays a role in reproductive isolation and speciation of smut fungal populations on different hosts is still an active field of research. Phylogenetic studies support the hypothesis that host adaptation and speciation are linked, as the vast majority of smut fungi on different host species form evolutionary independent lineages (e.g., Begerow et al. 2002; Lutz et al. 2005; Stoll et al. 2005; Piątek et al. 2013; Denchev et al. 2019). On the other hand, work on smut fungal populations shows ongoing gene flow between closely related parasite species (Petit et al. 2017) and experimental work clearly indicates that even distantly related species are able to hybridize (e.g. Büker et al. 2020; Kellner et al. 2011). High selfing rates by fusion of meiotic products from the same teliospore, as is for instance common in *Microbotryum* anther smuts (Giraud et al. 2005), could restrict gene flow of sympatric populations on different host species enough for reproductive barriers between species to evolve on different hosts (Giraud et al. 2008). Thus, while host specificity is still important to identify species in the field, its role in generating and maintaining biological species remains unclear.

## The role of asexual lineages in the smut fungi

One obvious pitfall of the BSC is that it offers no assistant in delineating asexual species. Agamotaxon has been suggested by cladists as a term for uniparental lineages to indicate different mechanisms that maintain well-adapted genomes (Meier and Willmann 2000), but this convention is not widely used by mycologists. The *Ustilaginomycotina* contain numerous asexual yeasts that reproduce by budding, such as *Acaromyces* spp., *Meira* spp., *Tilletiopsis* spp. in the Exobasidiomycetes, *Pseudozyma* spp. and the *Violaceomycetales* in the *Ustilaginomycetes*, *Malassezia* in the *Malasseziomycetes*, and *Monilliella* in the *Monilliellomycetes* (Boekhout et al. 2003; Rush and Aime 2013; Begerow et al. 2014; Wang et al. 2014, 2015a; Albu et al. 2015). For most

of these lineages it is unknown whether the sexual stage has been found, whether these lineages have lost the sexual stage, or if sexuality takes on a different form compared to other *Ustilaginomycotina*. As with other crucial questions concerning basic fungal biology, DNA sequencing and especially whole genome sequencing (WGS) gave important insights into potentially asexual smut fungal lineages. Molecular phylogenetic analyses have shown that several of the asexual yeast taxa found ubiquitously in the environment represented the asexual stages of known smut fungi. The commonly isolated *Pseudozyma prolifica* for instance has been shown to be the anamorphic state of the well-known smut fungus *Mycosarcoma maydis* (formerly known as *Ustilago maydis*). Additionally, other yeast lineages are con-generic with well-known smut fungal genera in *Ustilaginomycotina* and *Microbotryales* (Begerow et al. 2000, 2014, 2017; Wang et al. 2015a, b), but whether they represent plant parasitic smut fungi is not known. WGS gives good indications that several of these 'asexual' species, potentially have the possibility for sexual reproduction. The genes controlling sexual reproduction in *Ustilaginomycotina* and *Microbotryales* have been identified (see mating type evolution, above) and analyses of whole genomes showed that potentially functional mating types are present in several asexual taxa (Coelho et al. 2017), including *Malassezia* species (Xu et al. 2007; Wu et al. 2015). Nevertheless, in some lineages (e.g., *Violaceomyces palustris*) essential components of the *P/R* locus, such as pheromone receptors and their pheromone precursors, are lacking (Kijpornyongpan and Aime 2020), indicating that not all smut fungal taxa undergo sexual reproduction.

## Hybridization and species concepts in yeasts

Fungi are broadly used as experimental systems to detect and study hybridization. Both the species concept itself and species delimitation are challenged by the existence of hybrids. The term hybridization was coined in biology to define individuals resulting from the cross of genetically distinct organisms, without necessarily referring to organisms of different species, e.g., Gregor Mendel used the word to refer to the progeny resulting from crossing different breeds of a single pea species, *Pisum sativum* (Mendel 1866). Nowadays, the definition of hybrids is commonly related to that of species, distinguishing between intra- (within) or inter- (across) species hybrids (Abbott et al. 2013). Here, we shall discuss several outstanding issues that complicate the use of a species concept in hybrid species/lineages.

Inter-species hybrids are common among yeasts. They are not only present as recently formed organisms in industrial, clinical, or environmental niches (Morales and Dujon 2012; Mixão and Gabaldón 2018), but also can be at the root of

established lineages (Gabaldón 2020a). For instance, species such as the important opportunistic pathogen *Candida albicans*, or the so-called post whole genome duplication clade of *Saccharomycotina* species have likely originated from hybridization events (Wolfe and Shields 1997; Dietrich et al. 2004; Marcet-Houben and Gabaldón 2015; Mixão and Gabaldón 2020).

Due to technical improvements in genome analysis and sequencing technologies, bio-informatics, as well as awareness, more yeast hybrids are recognized to occur in various yeast lineages, e.g. in the *Cryptococcus neoformans/ Cr. gattii* complex (Boekhout et al. 2001; Bovers et al. 2006, 2008; Samarasinghe and Xu 2018) and *Trichosporonales* (*Basidiomycota, Agaricomycotina, Tremellomycetes*) (Sriswasdi et al. 2019), *Malassezia* (*Basidiomycota, Ustilaginomycotina, Malasseziomycetes*) (Wu et al. 2015), *Candida* (*Ascomycota, Saccharomycotina*) (Mixão and Gabaldón 2018, 2020; Mixão et al. 2019; Pryszcz et al. 2014), *Pichia* (*Ascomycota, Saccharomycotina*) (Leh-Louis et al. 2012), and *Schizosaccharomyces* (*Ascomycota, Taphrinomycotina*) (Zanders et al. 2014). These hybrids may be important in applied areas as they may have beneficial properties e.g., in fermentations, but similarly this may also result in novel virulent lineages (Gabaldón 2020a; Morales and Dujon 2012). Hybridization may be followed by extensive postzygotic genome modification, such as loss of heterozygosity (LOH), structural mutations and point mutations, and copy number changes resulting in fitness differences (Piotrowski et al. 2012, Smukowski Heil et al. 2017). See also below for more on this topic.

### Hybrids in the *Saccharomycotina*

Non-vertical evolutionary processes, such as introgression and hybridization, may complicate the proper recognition of species (Gabaldón 2020a; see below). This, of course, may cause inconvenience for the investigator, but eventually will result in a better understanding of the nature of species and how to utilize such information. An illustrative example is the brewer's yeast *Saccharomyces pastorianus*, including its synonym *S. carlsbergensis* that is widely used for the production of lager beers. Initially believed to be a unique species, *S. pastorianus* was later found to be the result of a hybridization event between *S. cerevisiae* and *S. eubayanus* (Libkind et al. 2011). Following the discovery of the hybrid nature of this species subsequent research focused on how to utilize this knowledge in the beer making process (Langdon et al. 2019). Another example is *Saccharomyces bayanus*, a hybrid involving contributions from three distinct species as parents: *S. cerevisiae, S. eubayanus*, and *S. uvarum* (Libkind et al. 2011; Pérez-Través et al. 2014). The proper recognition of this hybrid resulted in the circumscription of *S. uvarum*

as a genetically well-defined species (Nguyen and Gaillardin 2005; Nguyen and Boekhout 2017).

Contrary to the situation in plants or animals, where hybrids are traditionally recognized by the presence of intermediate phenotypic traits between well-defined parental species, hybrids in yeasts are usually first recognized upon analysis of several genomic markers or complete genomic sequences (e.g., Boekhout et al. 2001; Nguyen and Gaillardin 2005; Nguyen and Boekhout 2017; Gabaldón 2020b, Naranjo-Ortiz and Gabaldón 2020). This means that the parental species that originated the hybrid are not necessarily known. If the parental species are described and sequenced, the description and naming of the hybrid is straightforward. However, in many cases only one of the parental species, or none is known, and their existence can only be inferred from the divergence of the two hybrid sub-genomes. To illustrate this issue, we will refer to two hybrid lineages within the *Candida parapsilosis* species complex (Pryszcz et al. 2014, 2015) as well as three hybrids of *Trichosporonales*, *Basidiomycota* (see below Takashima et al. 2018, Sriswadi et al. 2016, 2019). Although strains within the *C. parapsilosis* clade were initially described as a single species, *Candida parapsilosis*, this clade was re-organized into three distinct species, *C. parapsilosis* (sensu stricto), *C. orthopsilosis*, and *C. metapsilosis* (Tavanti et al. 2005). Further sequencing of their genomes revealed that both *C. orthopsilosis* (Pryszcz et al. 2014; Schröder et al. 2016) and *C. metapsilosis* (Pryszcz et al. 2015) comprised, respectively, a majority or exclusively hybrid strains. This raises the issue of whether the hybrid strains of *C. orthopsilosis*, which are the result of the crossing of the lineage represented by the type strains and an unknown parental, should be considered a different species and given a different name. The situation of *C. metapsilosis* is somewhat different as the type strain, which served as the basis for describing the species, is a hybrid itself, and none of the sequenced strains seem to represent a pure, homozygous lineage. Hence two different parental lineages can be inferred, but these are (thus far) unknown, a situation that precludes their formal description, at least using standard approaches. Another issue relates to the parallel emergence of hybrids. Species that can form successful hybrids by crossing can form the same hybrids multiple times. This seems to be the case in various hybrids such as *C. orthopsilosis*, where at least four independent events of hybrid formation can be inferred (Schröder et al. 2016). Shall one consider independently formed hybrids between the same parental lineages as the same species or as different ones? If the former assumption is made, this will mean that the same species can originate several times and in different locations independently. This concept of convergent species formation is at odds with our standard understanding of the origin of species.

Finally, the definition of a hybrid lineage as a new species is clearly rooted in the very definition of species, as hybrids between different species may be considered as candidates for new species, whether intra-species hybrids may be better accommodated in the concept of variety.

Hybrids have been extensively found in the genus *Saccharomyces*, including *S. pastorianus/carlsbergensis* mentioned above, but also *S. cerevisiae* x *S. kudriavzevii, S. uvarum* × *S. eubayanus* and others indicating that interspecies *Saccharomyces* hybrids can form readily. Although there is some evidence of prezygotic reproductive isolation (Greig 2009; Ono et al. 2020) most of the reproductive isolation between species is post-zygotic that is the basis of the BSC (see above). Population genomics surveys reveal that introgression between species occurs regularly (Dujon and Louis 2017) with one of the first described being between *S. cerevisiae* and *S. paradoxus* (Liti et al. 2006) where the European population of *S. paradoxus* contains a segment of *S. cerevisiae* sequence as a homologous replacement of the *S. paradoxus* sequences found in other populations. This indicates that reproductive isolation is not complete and that some gene flow between species can occur, generally thought to be due to a rare viable spore being successively backcrossed to one of the parents. The spore viability of hybrids between *Saccharomyces* species is generally around 0.5% which is the expected frequency of obtaining a viable set of chromosomes when 16 pairs are randomly segregating. The chromosomes in these viable spore progenies have few if any crossovers and tend to be aneuploid. In *S. cerevisiae* × *S. paradoxus* hybrids, where there are no GCRs, no nuclear Bateson–Dobshansky–Muller incompatibilities have been found (Greig et al. 2002; Greig 2007; Kao et al. 2010). Instead, fertility is increased when the mismatch repair system is defective (Chambers et al. 1996; Hunter et al. 1996; Greig et al. 2003), which results in an increase in meiotic crossovers and a decrease in mis-segregation of chromosomes. As the mismatch repair system has crucial functions throughout the lifecycle of yeast, complete fertility restoration was not possible. Recently, using a reduction of mismatch repair specifically in meiosis along with meiotic loss of the *SGS1* helicase which unwinds most meiotic recombination intermediates as well as has crucial mitotic functions, Greig and colleagues were able to restore most of the fertility of hybrid spores (from 0.5% to over 30%) successfully breaking down the post-zygotic species barrier (Ozan Bozdag et al. 2021). Given that the major reproductive isolation mechanism leading to species in *Saccharomyces* appears to be the lack of recombination and the resulting missegregation of homeologous chromosomes, how do we explain introgressions that are found between species throughout the genus? In a hybrid between *S. cerevisiae* and *S. uvarum* under selective conditions, a recurrent rearrangement was found that included the introgression of a segment of the *S.*

*uvarum* genome into the *S. cerevisiae* genome (Dunn et al. 2013). If a viable spore from this modified hybrid contained mostly *S. cerevisiae* chromosomes or the *S. uvarum* genome was lost as can happen in unstable hybrids, a rapid introgression into the *S. cerevisiae* population can occur without numerous backcrosses. Very recently a number of isolates from the Alpechin population of *S. cerevisiae* were analysed by whole genome sequencing (Peter et al. 2018) and these contain a large number of introgressions from *S. paradoxus* (up to 8% of the genome in some isolates). In this population an *S. cerevisiae* × *S. paradoxus* hybrid was found that upon whole genome sequence analysis and proper haplotyping was found to be the living ancestor of the strains with *S. paradoxus* introgressions (D'Angiolo et al. 2020). This hybrid has suffered massive genome instability with mitotic recombination between the homeologues and many regions of loss of heterozygosity such that every chromosome pair now has homologous regions interspersed between the homoelogous regions. This living ancestor has one *S. cerevisiae* genome with numerous small introgressions from the *S. paradoxus* genome along with one *S. paradoxus* genome with a few introgressions from the *S. cerevisiae* genome. The introgressions found in the extant Alpechin *S. cerevisiae* population are found in the living ancestor's *S. cerevisiae* genome. The now homologous segments are sufficient to allow meiotic recombination between the homeologues and proper segregation resulting in some viable spores. This fortuitous finding of the progenitor hybrid in an intermediate state with the introgressions indicates how extant populations across the *Saccharomyces* clade may have arisen and reconciles the paradox of introgressions in species whose reproductive isolation barrier is lack of homologous recombination.

## Hybrids in the *Microbotryales, Pucciniomycotina, Ustilaginomycotina* and the *Tremellomycetes, Agaricomycotina*

Population studies of *Microbotryales* smut fungi revealed that several members can hybridize and gene flow between species occurs (Devier et al. 2010; Gibson et al. 2014; Petit et al. 2017; Hartmann et al. 2020). A study of these pathogens even indicated the relevance of hybrids in the development of new virulence patterns (Büker et al. 2013). Kellner et al. (2011) showed that non-specificity in the pheromone/pheromone receptor system of *Ustilaginales* might provide the molecular basis for such hybrids. However, pre- and postzygotic reproduction barriers in agreement with the BSC are the norm in these fungi, as even artificial mating experiments in the laboratory resulted in few viable hybrids. Still, the hybrids that survived showed rare recombination events (Büker et al. 2020), which could be important for accessing new hosts and the evolution of new smut fungal species.

Hybrids between different pathogenic *Cryptococcus* species, in particular between *Cr. neoformans* and *Cr. deneoformans*, have been isolated from the clinic and the environment, suggesting that inter-species hybridization may still occur in nature (Xu et al. 2000, 2002; Boekhout et al. 2001; Lengeler et al. 2001; Bovers et al. 2006, 2008; Lin et al. 2007; Litvintseva et al. 2007). *Cryptococcus* hybrids can still reproduce asexually as yeast and in some cases have higher fitness compared to haploid *Cryptococcus* species due to their increased ploidy level (aneuploid/diploid) and hybrid vigor arising from heterozygosity (Lin et al. 2007, 2008).

Recently, whole genome analysis has shown that three species of *Trichosporonales*, *Trichosporon coremiiforme*, *T. ovoides* and *Cutaneotrichosporon mucoides* are hybrids (Takashima et al. 2018, 2019). The genomes of *Cu. mucoides*, *T. coremiiforme* and *T. ovoides* were estimated to be 40.8, 42.35 and 40.3 Mb in size, respectively, which is significantly larger than those of putative haploid genomes of *Trichosporonales*, which range from 17.2 to 36.6 Mb and the number of genes of the three hybrids ranged from 12,979 to 14,292 compared to 5,647–9,805 in the putative haploid species (Takashima et al. 2018, 2019, Aliyu et al. 2020). A bioinformatics analysis showed that the two subgenomes of *Cu. mucoides* were most closely related to that of *Cu. dermatis*, with subgenome A being closest with 85% similarity. The subgenomes of *T. ovoides* were most closely related to that of *T. inkin* with subgenome A being the closest. The two subgenomes of *T. coremiiforme* yielded a surprise as subgenome A was found to be closely related to the genome of *T. asahii*, whereas subgenome B was closer to the genome of *T. faecale*. It was estimated that the *T. ovoides* and *T. coremiiforme* hybrids originated 14–22 Mya, which is much older than the recently formed hybrids of *Saccharomyces* species. The authors made a distinction between recently formed hybrids such as those occurring in *Saccharomyces*, the so-called Class I hybrids, and those that originated in a distant evolutionary past, the so-called Class II hybrids. The authors considered that the latter could be interpreted as species (see also below), and in several cases they were already described as separate species—see the species of *Trichosporonales* and those of the *C. parapsilosis* complex. The close phylogenetic relatedness between haploid and hybrid species in, e.g., the *Trichosporonales* hybrids, may be the reason for the difficulty of identifying them by MALDI-TOF MS that targets the proteome (Kolecka et al. 2013).

An in-depth sequence comparison and a transcriptome analysis of two of these *Trichosporonales* hybrids and their putative parental ancestors yielded some insights in the posthybridization processes of the three hybrids (Sriswasdi et al. 2016, 2019). The genome comparisons revealed that redundant genes underwent a deceleration, and not an acceleration of their evolutionary rates. It was suggested that gene conversion might have contributed to the genome stability

of the initial unbalanced hybrid genomes by restricting the functional differences between homologous gene sites from both subgenomes present in the hybrid. Moreover, large-scale gene loss, in particular from the transcriptional and translational machineries, formed a global compensatory mechanism against increased gene dosages resulting from the hybridization events (Sriswasdi et al. 2019). In the other study it was observed that the genome and transcriptome evolved separately after the hybrids were formed. Moreover, the various hybrids showed differences in the involvement of both subgenomes in the massive gene loss, where either only one or both subgenomes were involved (Sriswasdi et al. 2019). It was concluded that the stabilization of the genomes and the transcriptomes are two different evolutionary processes in young allopolyploids, and that closely related hybrids may follow different evolutionary trajectories (Sriswasdi et al. 2019).

Gabaldón (2020b) has proposed three Hybrid Genetic Zones (HGZ) defined by the interaction between gene flow and genetic divergence to differentiate between various hybrids. Sexually reproducing species show high gene flow and limited genetic divergence and crosses yield highly viable progeny. HGZ 1 hybrids show a relatively high gene flow with relatively low genetic divergence. These hybrids occur relatively often, and their progeny has moderate to low viability. Recently formed hybrids among *Saccharomyces* species and the *Cryptococcus neoformans* x *Cr. deneoformans* hybrids, which also can be generated easily in the laboratory, may belong here. HGZ 2 hybrids are characterized by limited gene flow and limited genetic divergence. It is probable that *Cu. mucoides* belongs in this category as its genes were still 85% identical with those of its ancestor *Cu. dermatis* (Takashima et al. 2018). Several of the hybrids described within *Candida* species with genetic divergences ranging from 3 to 5%, such as *C. orthopsilosis* (Pryszcz et al. 2014), *C. metapsilosis* (Pryszcz et al. 2015) or *C. inconspicua* (Mixao et al. 2019). Finally, *C. albicans* may represent a highly evolved HGZ 2 hybrid, in which continued LOH has extensively reduced the amount of heterozygous regions (Mixão and Gabaldón 2020). Finally, HGZ 3 hybrids are characterized by extensive genetic divergence and moderate gene flow. These hybrids might occur more rarely, and may undergone rapid ploidy changes and losses of one of the subgenomes to minimize genomic incompatibilities. *Millerozyma (Pichia) farinosa,* resulting from hybridization of two parental species diverging > 10% at the nucleotide level may represent a hybrid of this class (Leh-Louis et al. 2012). Note that hybrids categorized in each of the HGZ classes from Gabaldón may correspond to one of the two temporal classes proposed by Takashima and collaborators, as both categorizations focus on properties that are 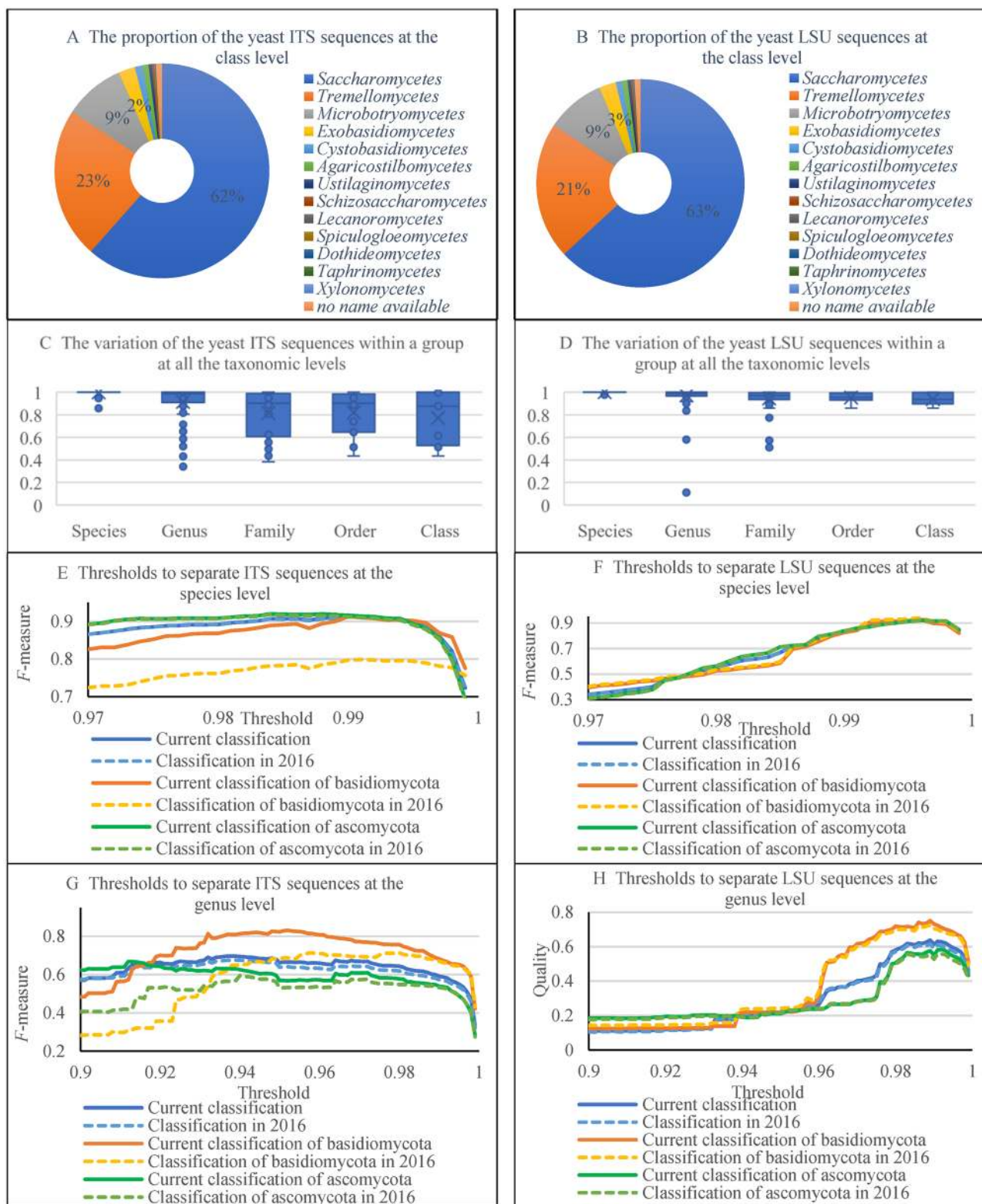orthologonal to each other (divergence of the hybridizing species in the former, or how long ago the hybrids were formed in the latter).

Based on these observations, one can state that the genetic background of various yeast hybrids may differ widely. Hybrids belonging to class II/HGZ2/3 may represent (small) speciation networks whose ancestral species may or may not be known. Moreover, the demonstration of a hybrid origin of the whole genome duplication lineage of *Saccharomyces* (Marcet-Houben and Gabaldón (2015), as well as the hybrid origin of *C. albicans* (Mixão and Gabaldón 2020) and *C. stellatoidea* (Mixão et al. 2021) clearly demonstrate that hybridization may also drive speciation.

## The role of DNA sequences in the practical recognition of yeast species

From the above it is clear that the advent of DNA sequencing has had an unprecedented impact on the recognition of yeast species. Nucleotide sequence databases such as NCBI Taxonomy (https://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/), UNITE (https://unite.ut.ee/), and theyeastst.org (https://theyeasts.org/) can be used to report validly described species, as well as unassigned strains and sequences from clone libraries and metabarcoding. It is important to realize that these databases are prone to contamination with inaccuracies in the sequences, the taxonomy, or both. We are greatly indebted to C.P. Kurtzman for his vision of giving the highest priority to populating a complete database of a single orthologous genomic region using type and reference strains, i.e., the D1/D2 domains of the Large Subunit (LSU) rRNA gene (Kurtzman and Robnett 1998). The sequence database may also serve as a forum that can bring together various authors to share their respective knowledge of a novel species. Databases can be exploited for automated and semi-automated data mining on species distributions across geography and habitats, and the evaluation of sequence polymorphisms. The counterpart is that sequences are so easily accessible that the effort needed to document morphology and growth responses seems enormous in comparison. Yet, if the description is to achieve its full purpose, *it must describe*. The cycler cannot replace the microscope or the fermentation tube, and a sequence is not a living organism.

Sequencing technologies have stimulated the design of sophisticated algorithms and associated software implementations. Because the various tasks of systematics are both similar and distinct, the systematist must seek to use appropriate methods. At the very start of the process, online searches for a ranked list of the best matching sequences are subject to the vagaries of artefacts specific to the laboratories that generated both the query and the target sequences (Stavrou et al. 2018; Pentinsaari et al. 2020). The default ranking

A The proportion of the yeast ITS sequences at the class level

- Saccharomyces
- Tremellomycetes
- Microbotryomycetes
- Exobasidiomycetes
- Cystobasidiomycetes
- Agaricostilbomycetes
- Ustilaginomycetes
- Schizosaccharomycetes
- Lecanoromycetes
- Spiculogloeomycetes
- Dothideomycetes
- Taphrinomycetes
- Xylonomycetes
- no name available

B The proportion of the yeast LSU sequences at the class level

- Saccharomyces
- Tremellomycetes
- Microbotryomycetes
- Exobasidiomycetes
- Cystobasidiomycetes
- Agaricostilbomycetes
- Ustilaginomycetes
- Schizosaccharomycetes
- Lecanoromycetes
- Spiculogloeomycetes
- Dothideomycetes
- Taphrinomycetes
- Xylonomycetes
- no name available

C The variation of the yeast ITS sequences within a group at all the taxonomic levels

D The variation of the yeast LSU sequences within a group at all the taxonomic levels

E Thresholds to separate ITS sequences at the species level

F Thresholds to separate LSU sequences at the species level

- Current classification
- Classification in 2016
- Current classification of basidiomycota
- Classification of basidiomycota in 2016
- Current classification of ascomycota
- Classification of ascomycota in 2016

G Thresholds to separate ITS sequences at the genus level

H Thresholds to separate LSU sequences at the genus level

- Current classification
- Classification in 2016
- Current classification of basidiomycota
- Classification of basidiomycota in 2016
- Current classification of ascomycota
- Classification of ascomycota in 2016

of a BLAST search (Altschul et al. 1990) is a function of both identity and query coverage, such that the rank order may not truly reflect relatedness. Depending on the length of the compared sequences, different BLAST algorithms (megablast, discontinuous megablast and blastn) can yield different orders of best matches (discussed in Lücking et al. 2020). Once a suitable set of published sequences is selected for further analysis, a competent alignment must be

◄**Fig. 2** Analysis of ITS1+2 and D1/D2 LSU ribosomal RNA gene sequences of yeasts. **A**, **B** Proportion of yeast ITS and LSU sequences at the class level. **C**, **D** Variation of the median ITS and LSU sequence similarity scores of yeasts at various taxonomic levels. **E–H** Optimal thresholds and associated highest *F*-measures predicted at the species and genus levels from a previous analysis (Vu et al. 2016) and a current dataset updated based on recent taxonomic revisions (date of analysis December 2020). The sequences were compared with each other using BLAST (Altschul et al. 1997). For each of the resulting local alignments of two sequences, a BLAST similarity score was calculated as the percentage of matches *s* if the alignment length *l* was greater than 300 bp (the minimum length of ITS sequences, Vu et al. 2019). Otherwise, the score was recomputed as $sl/300$. The names of the taxa associated with the sequences were downloaded from MycoBank (Robert et al. 2013)

generated, the details of which are beyond the scope of this article (but see Lücking et al. 2020, 2021). The construction of a tree then follows with the dual purpose of (1) identifying the phylogenetic position of the novel species in the hierarchy of classification and (2) adding evidence that a new species has indeed been discovered. The first task must be regarded as preliminary, as a robust phylogeny simply cannot be obtained from insufficient data (in terms of both sequence data and taxon sampling), no matter what tree inference algorithm is used (Aime et al. 2021). The second objective must consider what species concept is to be applied. If the BSC is to be given precedence, then the kind of phylogenetic tree is irrelevant. But if, as is usually the case with yeasts, the argument is founded primarily on Kurtzman's GSC, the selected tree must retain distance information in the form of terminal branch lengths (the maximum likelihood method often truncates terminal branches).
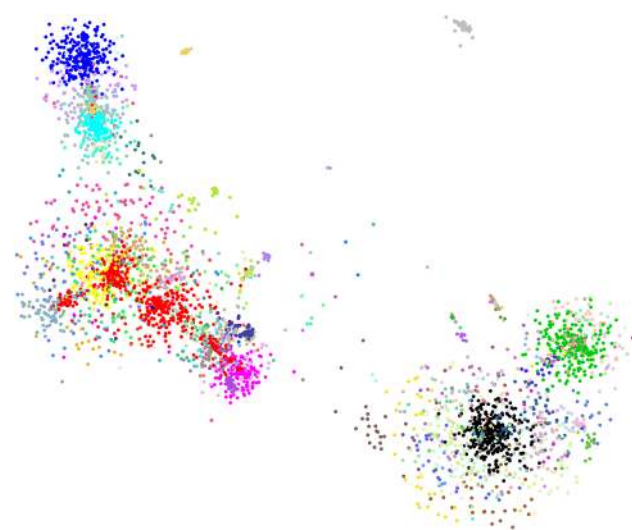
Since the introduction of molecular phylogenetic datasets and methods it has become clear that many previously recognized phenotypic species represent species complexes/cryptic species sharing similar morphological and growth characters. Examples are the *Cr. neoformans*/*Cr. gattii* species complex (Hagen et al. 2015, 2017; Kwon-Chung et al. 2017), *Papiliotrema flavescens*/*P. terrestris* species complex (Yurkov et al. 2015), *C. albicans*/*C. dubliniensis* (Sullivan et al. 1995), and lipid dependent *Malassezia* species (e.g., *M. arunalokei*, *M. restricta*, *M. globosa*) (Guého et al. 1996; Honnavar et al. 2016). In the case of such cryptic species, the BSC and/or the PSC/GCPSR are helpful to understand species delineation. (Giraud et al. 2008; Taylor et al. 2000). GCPSR can recognize cryptic species that lack distinguishing phenotypic features and that do not mate due to prezygotic (i.e., pre-mating) isolation, and in which inferences on the BSC can only be made indirectly, e.g., by showing recombination using molecular tests (Giraud et al. 2008; Hagen et al. 2015; Taylor et al. 2000).

## The value of barcodes for species identification

As indicated above, sequence analyses of the D1/D2 domains of the LSU rRNA gene and the ITS regions of rDNA have been widely used for yeast identification as well as species delineation. Here we address the reliability and accuracy of these two barcode regions for the identification of yeasts using the barcode datasets generated at the Westerdijk Fungal Biodiversity Institute, Utrecht, the Netherlands, which consists of 4,356 ITS and 4,213 LSU sequences representing 80% of the yeast species present in the CBS collection (Fig. 2A–H). Vu et al. (2016) conducted such an analysis previously, allowing us to address the effect of some recent taxonomic revisions (e.g., Liu et al. 2015, Wang et al. 2015a, b). Based on the classification obtained at the time of the release of the barcode datasets in 2016, there were 1227 (1249) species, 184 (180) genera, 52 (52) families, 25 (21) orders, and 8 (11) classes present in the yeast ITS and LSU (in parentheses) datasets. Since 2016 340 (316) sequences of 17 (17) genera, including 135 (151) *Candida* and 56 (49) *Cryptococcus* sequences have been reassigned to other genera. To evaluate if these name changes resulted in an improvement for species and genus identification, we computed two metrics, a reasonable sequence similarity cut-off (optimal threshold) and its associated *F*-measure (Paccanaro et al. 2006, Vu et al. 2014) (Figs. 2E–H). The *F*-measure (having a value between 0 and 1) evaluates the confidence of a threshold used to assign a sequence to a species or genus which is computed by comparing the groupings obtained by clustering the sequences with the given threshold and with the groupings equivalent to the associated taxon names. This measure has been widely used for evaluating the results of clustering approaches and its formula can be found in Paccanaro et al. (2006) and Vu et al. (2014, 2016). The higher the value of the *F*-measure, the closer the grouping of sequences by sequence similarity to the grouping of the sequences based on taxon names. The optimal threshold has the highest *F*-measure for sequence identification. For the estimation of the optimal threshold for species identification, indistinguishable species that shared the same ITS and/or LSU sequences with the others, i.e., they were in the same group when clustering the datasets with a 100% identity score, were removed. The analysis shows the accuracy of identification of a random yeast sequence at different sequence similarity levels (threshold) and across the taxonomic ranks. The threshold value reflects the observed average similarity of sequences of members of the given taxonomic rank but does not consider a barcoding gap. The analysis is a good proxy for best-score sequence similarity comparisons, like those researchers routinely perform against GenBank and MycoBank sequence databases. At the species level, the best *F*-measures obtained in most of the datasets were high (~0.92), showing that except for

the indistinguishable species (~5.3 and ~7.9% for ITS and LSU sequences, respectively), both ITS and D1/D2 LSU sequences had a strong identification power to identify yeasts. Within the *Basidiomycota*, a significant improvement for species identification was seen after the major changes in classification, as evidenced by the highest *F*-measure, which increased from 0.8 to 0.914 using ITS barcodes. The thresholds to predict species boundaries using ITS barcodes were 0.984 and 0.99 for *Ascomycota* and *Basidiomycota*, respectively, while they were 0.996 for both groups using the LSU barcode. Note that if the indistinguishable species were not excluded from the prediction, the obtained *F*-measures were < 0.82 and 0.81, although the thresholds for species boundaries were > 0.993 and 0.998 using ITS and LSU barcodes, respectively. At the genus level the obtained highest *F*-measures were low and varied from 0.56 to 0.83. Based on the *F*-measures obtained, ITS had a better identification power than LSU for yeast genera. In the *Basidiomycota*, an improvement was observed for genus identification in the updated taxonomy, as the best *F*-measure increased from 0.71 to 0.8 for ITS, and from 0.72 to 0.75 for LSU. For *Ascomycota,* the best *F*-measure increased from 0.6 to 0.68 for ITS, and from 0.56 to 0.58 for LSU. These low values of the *F*-measures obtained in the *Ascomycota* at the genus level, may suggest a need for further taxonomic improvement. This

is likely due to the current circumscription of the polyphyletic genus *Candida*, which contains many species that are poised to be reassigned to other genera as more data become available (Fig. 3). Based on the whole datasets, the optimal thresholds predicted for genus identification using ITS (LSU) barcodes were 0.937 (0.989) (currently) and 0.939 (0.989) (2016 analysis). For the *Basidiomycota*, thresholds of 0.952 (0.989) (currently) and 0.958 (0.989) (2016 analysis) were obtained, and for the *Ascomycota*, 0.912 (0.992) (currently) and 0.940 (0.991) (2016 analysis).

Although the study comprised the largest collection of sequence data derived from authenticated collection cultures, the analysis does not suggest that sequence similarity cut-offs should be used as a sole criterion for the identification of yeast species and genera. Different clades may require different sequence similarity cut-offs, a reflection of different rates of substitution in various species or genera (Fig. 2A–D). Some genera show limited barcode sequence variation among species (e.g., *Debaryomyces, Cryptococcus*), while others are highly heterogeneous (e.g., *Clavispora, Metschnikowia*). A combination of the sequence similarity cut-off and its confidence is useful to evaluate the consequences of changes in classification upon the efficacy of barcode sequences in identifying yeast isolates. Overall, the recent revisions in the taxonomy of yeasts, especially those proposed for the *Basidiomycota*, resulted in an improvement in the ability of both barcode sequences to identify yeasts at the genus level.

Barcode sequences, especially ITS, are widely used to estimate fungal species richness in so-called metabarcoding studies of such diverse environments as soils, water, plants, and the human body (Findley et al. 2013; Tedersoo et al. 2014; Rämä et al. 2016; Botschuijver et al. 2017; Canini et al. 2019; Nilsson et al. 2019). ITS metabarcoding has also resulted in the recognition of many putative, yet undescribed new species collectively referred to as 'fungal dark matter' (Hibbett 2016; Lücking et al. 2020, 2021). A major debate is going on among mycologists if and how to publish formal descriptions of 'dark matter' species that are only documented from sequence data obtained from environmental DNA (Lücking et al. 2020, 2021), and whether environmental DNA can serve as type material (Hawksworth et al. 2016; Hongsanan et al. 2018; Zamora et al. 2018). Regardless of the outcome of this debate, DNA barcoding will remain a major tool, both in metabarcoding studies and in the identification of yeast isolates.



**Fig. 3** Taxonomic clustering of 4,356 fungal ITS barcodes of the current taxonomic classification. The sequences are coloured based on the genus name. *Candida* (with 593 sequences) is indicated in red, followed by *Cryptococcus* (329) in black, *Saccharomyces* (288) in blue, *Rhodotorula* (218) in green, *Kluyveromyces* (146) in cyan, *Pichia* (135) in pink, *Debaryomyces* (113) in yellow, *Malassezia* (106) in grey. The remaining 188 genera with less than 85 ITS sequences are coloured using the distinctColorPalette function of the randomcoloR package. The coordinates of the sequences were generated using fMLC (Vu et al. 2018). The sequences are visualised using the rgl R package (https://r-forge.r-project.org/projects/rgl/)

# Species, genomes and genomics

Recent innovations in genome sequencing techniques, coupled with advances in tools and computational powers for genome assembly and analyses, have led to a better understanding of the diversity, phylogeny, speciation, and genome evolution in the fungal kingdom. Fungal genomes undergo extensive genome reduction/expansion and rearrangements even between closely related species. Fungal genome sizes range from 2 Mb in *Microsporidia* to > 2 Gb in *Pucciniales* (Aime et al. 2017), and the total number of predicted protein-coding genes vary from a couple of thousands in *Microsporidia* to more than 35,000 in *Sphaerobolus* in the *Agaricomycotina* (Stajich 2017).

Whole-genome sequencing is rapidly becoming commonplace and can contribute greatly to a sound species recognition (James et al. 2020; Libkind et al. 2020). In addition to enlightening the evolution and genetics of the phenotype of the novel species, whole-genome data can provide a robust phylogeny, which can be of great utility not only to situate the species hierarchically, but also if the PSC is to be applied. Overall genome similarity metrics are useful in the application of the genetic species concept (GSC, Libkind et al. 2020). This will, however, require empirical validation. Having examined 36 well-delineated haplontic, heterothallic *Metschnikowia* species, Lachance et al. (2020) determined that an average nucleotide identity (ANI) value of 95% provides a good guideline to delineate species following the GSC, although the sharp boundary observed for prokaryotic microorganisms was not found in these yeast species. Whole genome analyses can resolve the problem of incongruent phylogenies obtained from different marker genes (Capella-Gutierrez et al. 2014). Improvements in high-throughput whole-genome sequencing, and the efforts of individual researchers and several large-scale initiatives such as the 1000 Fungal Genomes Project led by the United States Department of Energy Joint Genome Institute (https://mycocosm.jgi.doe.gov/mycocosm/home/1000-fungal-genomes) and the Broad Institute's Fungal Genome Initiative (https://www.broadinstitute.org/fungal-genome-initiative/update-website-resources; Spatafora et al. 2017; Shen et al. 2018), have led to a major increase in the number of available fungal genomes. Fungal genome repositories such as FungiDB (https://fungidb.org/fungidb/app) have facilitated comparative genomics studies across the fungal kingdom (Stajich et al. 2011; Basenko et al. 2018) and phylogenomic analyses have allowed resolution of previously unresolved lineages, such as the *Ustilaginomycotina* (Kijpornyongpan et al. 2018) and the *Agaricomycetes* (Dentinger et al. 2016), that both include yeasts. Genome data can also allow inference of mating and life history strategies in cryptic fungi (Xu et al. 2007; Toome et al. 2014; Wu et al. 2015). Phylogenomic

analyses supply ample markers for testing the GCPSR and provide a robust framework for the analysis of phenotypic evolution of fungi and the underlying genetic changes, including horizontal gene transfers, underlying such transitions. For example, it has been shown that the acquisition of genes required for anaerobic growth through horizontal gene transfer, followed by gene duplication and clustering, allowed *Saccharomyces* to become an ideal brewing yeast (Gojković et al. 2004; Piškur et al. 2006).

The increased availability of novel sequencing technologies, including single-cell sequencing and broad sequencing of environmental samples (e.g., metagenomics), promises to reveal even more genotypic and ecological diversity and complexity in the fungal kingdom (Gabaldón 2020c; Rosling et al. 2011; James et al. 2013; Ahrendt et al. 2018). In recent years, advancements in long-read sequencing techniques, such as PacBio and Oxford Nanopore MinION, allow the assembly of complete fungal genomes at the chromosome level, as they are able to generate sequencing reads that are long enough to span genomic regions that are highly repetitive (e.g., rDNA) or enriched with transposon repeats (e.g. centromeres), which are difficult to assemble from short reads generated by sequencing platforms such as Illumina. Improvements in genome assemblies from long-read sequencing has also facilitated better allele phasing (assignment of alleles to their respective chromosome) in diploid species or diploid hybrids of haploid species, and thus, more robust analyses of mitotic recombination and their effects on fitness and evolution of the hosts.

Complete genome data, including full-length chromosomes sequences and the best possible gene annotation is necessary to understand many evolutionary processes, such as gene duplication, mutations and acquisition of novel genetic material that underlie yeast evolutionary genomics (Dujon 2010; Dujon and Louis 2017). Complete genome assemblies enable more robust analyses of genome rearrangements such as inversions and translocations, which have shown that: (1) such chromosomal changes occur frequently between closely related species, even among isolates of the same species; (2) many of these chromosomal rearrangements were mediated by transposable elements, and in some cases by chromosomal regions (such as centromeres) that are enriched with transposons and their remnants; and (3) chromosomal rearrangements alone are sufficient to establish reproductive barriers, and could have played significant roles in fungal speciation (Janbon et al. 2014; Sun et al. 2017; Yadav et al. 2020). Together with genetic and molecular biology studies, analyses of complete fungal genomes also revealed that fungal centromeres are highly diverse, ranging in size from 125 bp in *S. cerevisiae*, to a few kilobases in certain *Candida* species, to tens or even hundreds of kilobases in some *Cryptococcus* species that harbour large regional centromeres (Guin et al. 2020).

Fungal centromeres also differ significantly in their structures, varying from well-defined point centromeres to short regional centromeres with terminal repeats, to large regional centromeres enriched with transposons and their remnants. While regional centromeres are likely defined epigenetically, centromere-specific sequence motifs may still play critical roles in centromere characterization and function (Yadav et al. 2018; Navarro-Mendoza et al. 2019; Guin et al. 2020; Sankaranarayanan et al. 2020).

As sequencing prices drop, comparative genomics studies can address genomic diversity within single species, or even populations. Population genomics studies have revealed that fungal species are genetically heterogeneous. This diversity extends well beyond polymorphisms of shared loci, also eliciting differences in the presence and absence of genes among genomes from strain to strains. Although only a limited number of model fungal species have been explored in this regard, it seems that the concept of pan-genome, developed for bacteria, also applies to fungi (Naranjo-Ortiz and Gabaldón 2020). This large genetic diversity within fungal species as well as the realization that non-vertical modes of evolution occur in fungi, challenges the identification and separation of species based on the basis of specific marker loci (Lücking et al. 2020). The selection of barcoding genes heavily relies on historical choices and not on the empirical testing of the ability of each locus to properly differentiate strains from different species. As different genes may show different levels of conservation in different clades, the very concept of whether there is a universal fungal barcoding gene remains a matter of debate. In this context, the continuous developments in sequencing and computing technologies coupled with more affordable and user-friendly platforms may soon lead to a situation in which low coverage sequencing and comparison with extensive databases become the cost-effective solution for the accurate delineation of a fungal species (Gabaldón 2020b). However, for this to be a reality several issues remain to be resolved, namely establishing standardized pipelines, curating errors from public databases, and superimposing the boundaries of our current morphology-based species definition to the expanding knowledge of genomic divergence (Stavrou et al. 2018; Xu 2020).

## The species description, best practices

As stated by Mayden (2002), "species descriptions represent testable hypotheses, just the same as what is presented in other areas of the natural sciences." The mere isolation of a contaminant whose DNA-barcode sequence differs by four substitutions from an existing species hardly justifies the formal description of a novel species. Whether or not the observed differences are sufficient to erect a new species must be viewed in the context of a species concept. An overview of species concepts used in mycology is provided by Lücking et al. (2020) and those frequently used for yeasts are discussed in more detail here (see Table 1). The choice of a species concept determines the characters and tools used for its delineation and identification. For example, the demonstration of reciprocal monophyly required by the PSC makes no sense if the extent of sequence polymorphism is not known because a species is represented by only a single isolate. Application of the criterion of genealogical concordance, which combines the PSC with the BSC requires examination of several individuals. Once the existence of a novel species is demonstrated from its genealogy (PSC/GCPSR), reproduction (BSC), and phenotype (PhenSC), efforts should be made to identify the habitat, that is, where the species occurs naturally, and its niche. This entails repeating the isolation from similar substrates, using a sound sampling strategy that will help identify the actual habitat, or search for its presence in metabarcoding libraries.

Some researchers advance that naming of yeast species should be delayed until all taxa within a phylogenetic grouping have been identified (see e.g., Kwon-Chung et al. 2017). The reality is that there will be a near continual recognition of new taxa as science advances. Once a species is described formally it can be thoroughly tested as new specimens are identified (Fig. 1). A striking example is *Candida auris*, which was described from only a single isolate obtained from the ear of an inpatient in Japan (Satoh et al. 2009). Within one decade after its discovery *C. auris* has become one of the major global fungal pathogens (Lockhart et al. 2017a, b; de Jong and Hagen 2019).

A scrupulous scrutiny of yeast species descriptions published in the past might well reveal that a substantial proportion would not meet modern best practices (Yurkov et al. 2021). Perhaps one of the most important aspects of describing new taxa should be a sensitivity to the user community. Descriptions should be informative and founded on the highest quality biology and provide enough detail for unambiguous identification (Aime et al. 2021). In an editorial published in the journal '*Yeast*', Lachance (2020) listed a number of essential requirements, namely high-quality work, reference to the relevant literature including the latest edition of '*The Yeasts, a Taxonomic Study*' (Kurtzman et al. 2011a) or its electronic successor https://theyeasts.org, expert etymology, a standard list of growth properties, high-quality photographs of all stages of the life cycle, evidence of due diligence in observing sexual reproduction, deposition of barcode sequences (e.g. ITS and D1/D2 domains of the LSU rRNA gene), a lucid statement of applicable species concepts, a sequence-based tree appropriate to the species concept invoked, and clear information on the geography and habitats of isolates. In addition to these essential features, the

editorial recommended that the proposal of a new species be documented from distinct, multiple isolates, in the absence of which a strong justification should be given to warrant publication of a description. Additionally, deposition and annotation of a (draft) genome were considered desirable, and authors were encouraged to couch the description in an ecologically meaningful framework. Karyotype information was also recommended. These and other recommendations have been taken up by the *International Commission on the Taxonomy of Fungi*, and formalized in a publication that includes model manuscripts, a checklist, and templates for descriptions (Aime et al. 2021). Taken together, these requirements ensure that a new species is thoroughly characterised and distinct from other species, that it can be identified by other researchers, and that its reference material is available for further research.

Yeast morphology has yielded the way to nutritional tests and later sequence information for the purpose of delineation and identification of species. It remains that morphology contributes a large proportion of what we regard as representative of a yeast species. Microscopes are unlikely to improve much in the foreseeable future and so the onus is on individual systematists to choose the most appropriate approaches. The ability to generate good images is based almost entirely on adequate training. Characteristics such as cell wall composition, nucleotide ratios, principal ubiquinone isoprenoid length have had their historical importance, but it is fair to question their current relevance. Growth tests formulated primarily by L.J. Wickerham provide much insight on the adaptation of yeast species to their environment but are laborious and fraught with much response variation between laboratories. Most of the commercially available kits used for identification of yeast isolates can also be used to perform growth tests, but they do not follow the traditional standard method, i.e., Wickerham's tube method or auxanograms, and lack validation against time-tested methods. As the kits focus on the identification of a limited number of clinical and food-borne yeasts, their results are often less informative for species that are not normally associated with those substrates. For example, the multi-drug resistant emerging pathogen, *Candida auris* (see also above), has been repeatedly misidentified by commercial kits (Lockhart et al. 2017b; Mizusawa et al. 2017). Although specialized identification with some of these test kits has improved (Arastehfar et al. 2019), their use in a clinical setting is rapidly being displaced by MALDI-TOF and PCR or sequence-based tests. Effort is needed to find better ways to capture growth responses. It is difficult to imagine how one can describe a yeast species without such data, as growth responses are a window on the fundamental niche of a yeast species and thus represent valuable information relevant to ecology. Examples are yeasts found in sugar-rich substrates (Péter et al. 2017), soil yeasts able to assimilate the by-products of cellulose and lignin degradation (Botha 2011), or the need for yeasts that are pathogenic to mammals, including humans to grow above 35 °C (Robert et al. 2015).

## Naming of species and the rules of nomenclature

There are three general ways in which fungal species have been named. The first one is in honour or homage to a distinguished scientist, such the basidiomycete genus *Kwoniella*, named in honour of Dr. June Kwon-Chung. Second, species are named in recognition of a location associated with their discovery, such as *Kwoniella botswanensis*, discovered from Mopane Trees in Botswana. Third, species are named after an attribute, such as the substrate on which they were found as in the case of *Cryptococcus floricola*, isolated from the nectar of flowers on the Canary Islands. Given the vast number of fungi that remain to be discovered and named creativity in the generation of new names will be an asset. A consensus among fungal biologists is lacking on what approach to naming is most desirable. For example, yeasts named after a locality may later be found to be more widespread, and the same holds for the source of isolation.

Two aspects are important when formally describing any new species including yeasts. The first is taxonomy, which includes the characters used to differentiate species. Although no formal requirements exist for good taxonomy, community standards and best practices, such as those outlined in the sections above, exist to guide authors and journals in the publication of new taxa that meet rigorous scientific standards as well as the needs of the user community. The other realm is that of nomenclature, which concerns the legal requirements that dictate how to publish and name a new taxon. In other words, while taxonomy is concerned with what constitutes a species and how to describe it, nomenclature concerns how to name a species, regardless of how it has been defined.

Nomenclatural rules applicable to yeasts are found in the '*International Code of Nomenclature for algae, fungi, and plants*' (ICNafp, or the *Code*) (Turland et al. 2018; https://www.iapt-taxon.org/nomen/main.php). The ICNafp is updated every six years, and the requirements in each new code supersede those of prior codes. For publication of a new name to be *Code*-compliant it must fulfil the rules for: (1) effective publication, in other words that it be published in a sanctioned journal, explained in detail in Arts. 29–31; (2) valid publication, which primarily deals with correct designation of holotypes or basionyms, which are explained in detail in Arts. 32–45; and (3) legitimate publication, which broadly concerns the usage of names based on previously published taxa as well as publication of superfluous names,

as detailed in Art. 6. Any names that do not meet the requirements for all three are unavailable for use. Illegitimate names may or may not also be invalid. General guidelines for how to avoid publishing invalid or illegitimate names are provided in Aime et al. (2021); more detailed guidelines for yeast taxa follow.

The most common errors in erecting new names for yeasts species are: (1) improper designation of the holotype and holotype institution (Arts. 8.4 and 40.7). Most commonly, this is the failure to designate a *single* institution in which the holotype resides; (2) not designating a metabolically inactive specimen as the holotype. To fulfil the latter requirement of the *Code*, either a dried specimen or a metabolically inactive culture should be prepared, designated and deposited in a repository. This can be achieved by drying a culture plate and preserving it in a fungarium/herbarium or to deposit the culture in a culture collection that preserves holotypes as lyophilized samples or cryopreserved at ultra-low temperatures. As of January 1, 2019, the words "metabolically inactive" need to be included along with the institution designation in the holotype designation to show this clearly (Arts. 40.7 and 40.8).

In the ICNafp, a holotype is a single collection, specimen, or isolate that serves as the reference for its name for all contemporary and future researchers. A holotype should not only be selected as the best representative of a name but should also be deposited in an internationally recognized and easily accessible biorepository (fungarium, herbarium, or culture collection), and wherever possible, accompanied by an isotype (metabolically inactive duplicates of the holotype). When a researcher deposits subcultures of a strain that was selected to represent a nomenclatural type of a name in two (or more) culture collections, and if all of these deposits are preserved in a metabolically inactive state, only one can be designated as a metabolically inactive holotype and other deposits should be indicated as isotypes. The ICNafp recommends the deposition of living cultures prepared from the holotype in at least two culture collections (Rec. 8B1). These cultures are designated as ex-types, i.e., derived from the metabolically inactive holotype.

While there are no formal requirements for depositing viable cultures, it goes without saying that ex-type cultures should be deposited in publicly accessible international collections and be available to the scientific community, a list of which are available through the *Culture Collections Information Worldwide* of the World Federation for Culture Collections and World Data Center for Microorganisms (http://www.wfcc.info). It is preferable that duplicates of the type material be deposited in more than one country.

Fortunately, invalid names can be later validated by correcting the errors in a subsequent publication. For example, a new name that was published invalidly because several cultures were listed without indicating which was the holotype, can be validated in a later publication by selection and designation of one of the cultures to serve as holotype (as long as it is preserved in a metabolically inactive state). In these cases, the original author names are followed by "ex" and the names of the authors who validated the name. Illegitimate names can, in rare cases, be nominated for conservation, but this is rarely done, and illegitimately published names of distinct taxa are usually corrected by publication of entirely new names for the replaced name. In these cases, the new name is followed by the authors of the new name and the designation "nom. nov.", and the basionym is cited as the original name and publication. In cases where names were both invalidly and illegitimately published, a correct holotype designation must also be included in the nom. nov. description. Other formal requirements for valid and legitimate publication of names are explained in May et al. (2019) and Turland (2019).

The ICNafp provides indications on how to name hybrids between defined species (Turland et al. 2018). However, these indications allow enough room for alternative choices. For instance, a hybrid taxon can be given a standard species name, without any indication of the hybrid nature in the name. This is the case of many naturally occurring hybrids that were described before their hybrid nature was discovered. For instance, *S. pastorianus* (Vaughan Martini and Martini 1987). Hybrids between two or more described taxa can also be named using the multiplication sign "×", which is placed before the name of an intergeneric hybrid or before the epithet in the name of an inter-specific hybrid. This is usually the choice for artificially created hybrids in which both parental species are known, e.g., *S. cerevisiae* × *S. uvarum*. Nguyen and Boekhout (2017) provided some guidelines for naming yeast hybrids that can also be used for other hybrid yeasts. The species that comprise the hybrid are given in alphabetical order, e.g., strain CBS 8615 (CID1) is a *S. cerevisiae* × *S. kudriavzevii* × *S. uvarum* hybrid, and if the ploidy of the parent species is known this is given after each species name. For example, *S. bayanus* CBS 380$^T$ is *S. cerevisiae* ˂1% × *S. eubayanus* 37% × *S. uvarum* 63%. Note that different hybrids may differ in their evolutionary trajectories and a distinction is to be made between recently formed hybrids and those that have evolved over a long history that even may represent species (see discussion above).

Additionally, the use of the prefix "notho-" (optionally abbreviated "n-") added to the term denoting the rank of the taxon can be used to emphasize the hybrid nature of a taxon, although this convention is not commonly used in yeasts. Of note, the prefix notho- and the multiplication sign are not part of the actual species name and should be disregarded for nomenclatural purposes, such as synonymy or homonymy. Furthermore, the designation as nothotaxon of any taxon that is believed to be of hybrid origins is optional, and a nothotaxon cannot be designated unless at least one

parental taxon is known or can be postulated, but this leaves ample room to consider different possible ways of postulating a parental taxon.

## Conclusions and perspectives

From the above it is clear that many species concepts have been used over time to describe yeast diversity. Initially a PhenSC was used, mainly taking into consideration cellular and colony morphologies and growth responses. Mating experiments with or without a genetic analysis of the offspring were used to assess the presence of reproductive barriers important to our understanding of species boundaries under the BSC. However, the increasing number of yeast hybrids observed indicate that the application of the BSC is not without risk and must be based on a thorough genetic analysis. The use of (multiple) molecular markers, such as sequences of the D1/D2 domains of the LSU rRNA gene, the ITS rDNA region, and other (protein coding) sequences, as proxies for species delimitation, and comparison with species boundaries set under the BSC has yielded insights into thresholds of nucleotide divergence values that may be considered to set species hypotheses under the GSC. The phylogenetic grouping of yeast isolates using partial or entire genome DNA sequences under the PSC/GCPSR has recently gained in popularity. Finally, unknown species identified by ITS metabarcoding from environmental DNA, usually referred to as 'fungal dark matter' (Lücking et al. 2020, 2021), need attention as well. We believe that the species hypotheses originating from the above approaches will need thorough experimental testing, using genetics, ecology, and comparative and functional genomics approaches that will take (several or many) experimental iterations (Fig. 1). Note that testing of species hypotheses usually starts once these have been postulated, e.g., in a formal species description (Nguyen and Gaillardin 2005; Louis 2009; Büker et al. 2013; Pérez-Través et al. 2014; Pryzcz et al. 2014, 2015; Hagen et al., 2015, 2017; Marcet-Houben and Gabaldón 2015; Kwon-Chung et al. 2017; Nguyen and Boekhout 2017; Passer et al. 2019: Priest et al. 2021). This process of testing species hypotheses has started with a few model species, such as *S. cerevisiae*, but also the *Cr. neoformans* species complex and *M. lychnidis-dioicae* (Fig. 1), and this will be expanded to other species complexes in the future. To gain better insight in the genomic dynamics of evolving species we strongly recommend analysing completed rather than draft genomes. Only detailed knowledge of yeast genome organization, for isolates and their possible offspring—including structural genetic elements, such as chromosomes, centromeres, genes/open reading frames, exons/introns, repeated sequences, non-coding RNAs including tRNAs, rRNAs, snoRNAs, microRNAs, and transposons—will give us insight into the evolution of the genomes within and betweens species, how this affects selective advantage, and the data for implementation of DFSC. Such analyses may include aspects of structural biology, such as the three-dimensional structure of proteins of, e.g., pheromone peptides, pheromone receptors, and transcriptomics studies of various mating systems and physiology, hybridization, chromosome evolution, and the meiosis-related cell machinery, to better understand mechanistic aspects of pre- and postzygotic isolation that determine gene flow between populations. Also (in)compatibility between nuclear and mitochondrial genomes in such processes can be considered. Eventually, this will result in a more accurate, 'holistic' species concept that also will take into consideration intrinsic complexities, e.g., due to hybridization and introgression, forces that illustrate, reflect, and reinforce species boundaries but continue to confuse many investigators into considering (and even inaccurately asserting and espousing) the misconception that species boundaries are artificial edifices. Such a species concept likely will take the form of (small) speciation networks that relate to other such networks, e.g., via introgression and hybridization processes, and that address the genetic cohesiveness of species complexes more fully and accurately than the current tree-like models. Given the rapid developments of omics and bioinformatics technologies, the future will hold many surprises with respect to our understanding of the (fuzzy) boundaries between yeast species.

**Author contributions** All authors have equally contributed to drafting and reviewing the manuscript. All authors agree with the submission of the manuscript to Fungal Diversity. The manuscript does not contain supplementary materials.

## Declarations

**Conflict of interest** There is no conflict of interest.

# References

Abbate J, Gladieux P, Hood ME, de Vienne DM, Antonovics JA, Snirc A, Giraud T (2018) Co-occurrence among three divergent plant-castrating fungi in the same *Silene* host species. Mol Ecol 27:3357–3370. https://doi.org/10.1111/mec.14805

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Möst M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, Rice AM, Ritchie MG, Seifert B, Smadja CM, Stelkens R, Szymura JM, Väinölä R, Wolf JBW, Zinner D (2013) Hybridization and speciation. J Evol Biol 26:229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Ahrendt SR, Quandt CA, Ciobanu D, Clum A, Salamov A, Andreopoulos B, Cheng JF, Woyke T, Pelin A, Henrissat B, Reynolds NK, Benny GL, Smith ME, James TY, Grigoriev IV (2018) Leveraging single-cell genomics to expand the fungal tree of life. Nat Microbiol 3:1417–1428. https://doi.org/10.1038/s41564-018-0261-0

Aime MC, McTaggart AR, Mondo SJ, Duplessis S (2017) Phylogenetics and phylogenomics of rust fungi. Adv Genet 100:267–307. https://doi.org/10.1016/bs.adgen.2017.09.011

Aime MC, Miller A, Aoki T, Bensch K, Cai L, Crous PW, Hawksworth DL, Hyde K, Kirk P, Lücking R, May T, Malosso E, Redhead SA, Rossman A, Stadler M, Thines M, Yurkov Y, Zhang N, Schoch L (2021) How to publish a new fungal species, or name, version 3.0. IMA Fungus 12:11. https://doi.org/10.1186/s43008-021-00063-1

Albu S, Toome M, Aime MC (2015) *Violaceomyces palustris* gen. et sp. nov. and a new monotypic lineage, *Violaceomycetales* ord. nov. in *Ustilaginomycetes*. Mycologia 107:1193–1204

Aldhebiani AY (2018) Species concept and speciation. Saudi J Biol Sci 25:437–440. https://doi.org/10.1016/j.sjbs.2017.04.013

Aliyu A, Olga Gorte O, de Maayer P, Neumann A, Ochsenreither K (2020) Genomic insights into the lifestyles, functional capacities and oleagenicity of members of the fungal family Trichosporonaceae. Sci Rep 10:2780. https://doi.org/10.1038/s41598-020-59672-2

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389

Arastehfar A, Daneshnia F, Kord M, Roudbary M, Zarrinfar H, Fang W, Hashemi SJ, Najafzadeh MJ, Khodavaisy S, Pan W, Liao W (2019) Comparison of 21-Plex PCR and API 20C AUX, MALDI-TOF MS, and rDNA sequencing for a wide range of clinically isolated yeast species: Improved identification by combining 21-Plex PCR and API 20C AUX as an alternative strategy for developing countries. Front Cell Infect Microbiol 9:21

Bahn Y-S, Sun S, Heitman J, Lin X (2020) Microbe profile: *Cryptococcus neoformans* species complex. Microbiology 166:797–799. https://doi.org/10.1099/mic.0.000973

Barnett JA (2004) A history of research on yeasts 8: taxonomy. Yeast 21:1141–1193

Basenko EY, Pulman JA, Shanmugasundram A, Harb OS, Crouch K, Starns D, Warrenfeltz S, Aurrecoechea C, Stoeckert CJ, Kissinger JC, Roos DS, Hertz-Fowler C (2018) FungiDB: an integrated bioinformatic resource for fungi and oomycetes. J Fungi 4:39

Begerow D, Bauer R, Boekhout T (2000) Phylogenetic placements of ustilaginomycetous anamorphs as deduced from nuclear LSU rDNA sequences. Mycol Res 104:53–60

Begerow D, Bauer R, Oberwinkler F (2002) The Exobasidiales: an evolutionary hypothesis. Mycol Progr 1:187–199. https://doi.org/10.1007/s11557-006-0018-7

Begerow D, Göker M, Lutz M, Stoll M (2004) On the evolution of smut fungi on their hosts. Frontiers in basidiomycete mycology. IHW-Verlag Eching, pp 81–98

Begerow D, Schäfer AM, Kellner R, Yurkov A, Kemler M, Oberwinkler F, Bauer R (2014) Ustilaginomycotina. In: McLaughlin DJ, Spatafora JW (eds) The mycota: systematics and evolution, Part A. Springer, Heidelberg, pp 295–329

Begerow D, Kemler M, Feige A, Yurkov A (2017) Parasitism in yeasts. In: Buzzini P, Lachance MA, Yurkov AM (eds) Yeasts in natural ecosystems: ecology. Springer, Cham, pp 179–210

Boekhout T, Theelen B, Diaz M, Fell JW, Hop WC, Abeln EC, Dromer F, Meyer W (2001) Hybrid genotypes in the pathogenic yeast *Cryptococcus neoformans*. Microbiology 147:891–907

Boekhout T, Theelen B, Houbraken J, Robert V, Scorzetti G, Gafni A, Gerson U, Sztejnberg A (2003) Novel anamorphic mite-associated fungi belonging to the Ustilaginomycetes: *Meira geulakonigii* gen. nov., sp. nov., *Meira argovae* sp. nov. and *Acaromyces ingoldii* gen. nov., sp. nov. Int J Syst Evol Microbiol 53:1655–1664. https://doi.org/10.1099/ijs.0.02434-0

Bölker M, Urban M, Kahmann R (1992) The a mating type locus of *U. maydis* specifies cell signaling components. Cell 68:441–450. https://doi.org/10.1016/0092-8674(92)90182-c

Bölker M, Genin S, Lehmler C, Kahmann R (1995) Genetic regulation of mating, and dimorphism in *Ustilago maydis*. Can J Bot 73:320–325. https://doi.org/10.1007/978-94-011-0177-6_35

Botha A (2011) The importance and ecology of yeasts in soil. Soil Biol Biochem 43:1–8

Botschuijver S, Roeselers G, Levin E, Jonkers DM, Welting O, Heinsbroek SEM, de Weerd HH, Boekhout T, Fornai M, Masclee AA, Schuren FHJ, de Jonge WJ, Seppen J, van den Wijngaard RM (2017) Intestinal fungal dysbiosis is associated with visceral hypersensitivity in patients with irritable bowel syndrome and rats. Gastroenterology 153:1026–1039. https://doi.org/10.1053/j.gastro.2017.06.004

Bovers M, Hagen F, Kuramae EE, Diaz MR, Spanjaard L, Dromer F, Hoogveld HL, Boekhout T (2006) Unique hybrids between the fungal pathogens *Cryptococcus neoformans* and *Cryptococcus gattii*. FEMS Yeast Res 6:599–607. https://doi.org/10.1111/j.1567-1364.2006.00082.x

Bovers M, Hagen F, Kuramae EE, Hoogveld HL, Dromer F, St-Germain G, Boekhout T (2008) AIDS patient death caused by novel *Cryptococcus neoformans* x *C. gattii* hybrid. Emerg Infect Dis 14:1105–1108. https://doi.org/10.3201/eid1407.080122

Brachmann A, Weinzierl G, Kämper J, Kahmann R (2001) Identification of genes in the bW/bE regulatory cascade in *Ustilago maydis*. Mol Microb 42:1047–1063. https://doi.org/10.1046/j.1365-2958.2001.02699.x

Branco S, Badouin H, Rodríguez de la Vega RC, Gouzy J, Carpentier F, Aguileta G, Siguenza S, Brandenburg JT, Coelho MA, Hood ME, Giraud T (2017) Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. Proc Nat Acad Sci USA 114:7067–7072. https://doi.org/10.1073/pnas.1701658114

Büker B, Petit E, Begerow D, Hood ME (2013) Experimental hybridization and backcrossing reveal forces of reproductive isolation in

*Microbotryum*. BMC Evol Biol 13:224. https://doi.org/10.1186/1471-2148-13-224

Büker B, Guerreiro MA, Hood ME, Brachmann A, Rahmann S, Begerow D (2020) Meiotic recombination in the offspring of *Microbotryum* hybrids and its impact on pathogenicity. BMC Evol Biol 20:123. https://doi.org/10.1186/s12862-020-01689-2

Canini F, Zucconi L, Pacelli C, Selbmann L, Onofri S, Geml J (2019) Vegetation, pH and water content as main factors for shaping fungal richness, community composition and functional guilds distribution in soils of Western Greenland. Front Microbiol 10:2348. https://doi.org/10.3389/fmicb.2019.02348

Capella-Gutierrez S, Kauff F, Gabaldón T (2014) A phylogenomics approach for selecting robust sets of phylogenetic markers. Nucleic Acids Res 42:e54. https://doi.org/10.1093/nar/gku071

Chambers SR, Hunter N, Louis EJ, Borts RH (1996) The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. Mol Cell Biol 16:6110–6120. https://doi.org/10.1128/mcb.16.11.6110

Coelho MA, Bakkeren G, Sun S, Hood ME, Giraud T (2017) Fungal sex: the Basidiomycota. In: Heitman J, Howlett BJ, Crous PW, Stukenbrock EH, James TY, Gow NAR (eds) The fungal kingdom. ASM Press, Washington DC. https://doi.org/10.1128/9781555819583.ch7

Conti A, Casagrande Pietoni D, Robert V, Cardinali G, Corte L (2021) Homoplasy as an auxiliary criterion for species delimitation. Microroganisms 9:273. https://doi.org/10.3390/microorganisms9020273

D'Angiolo M, De Chiara M, Yue JX, Irizar A, Stenberg S, Persson K, Llored A, Barré B, Schacherer J, Marangoni R, Gilson E, Warringer J, Liti G (2020) A yeast living ancestor reveals the origin of genomic introgressions. Nature 587:420–425. https://doi.org/10.1038/s41586-020-2889-1

Darwin C (1859) On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life [Internet]. John Murray, London. https://search.library.wisc.edu/catalog/9934839413602122

Darwin C, Wallace A (1858) On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. J Proc Linn Soc Zool 3:45–62. https://doi.org/10.1111/j.1096-3642.1858.tb02500.x

De Candolle AP (1813) Théorie élémentaire de la botanique; ou, Exposition des principes de la classification naturelle et de l'art de décrire et d'étudier les végétaux. Déterville, Paris. https://doi.org/10.5962/bhl.title.39705

De Jong AW, Hagen F (2019) Attack, defend and persist: how the fungal pathogen *Candida auris* was able to emerge globally in healthcare environments. Mycopathologia 184:353–365. https://doi.org/10.1007/s11046-019-00351-w

De Queiroz K (2007) Species concepts and species delimitation. Syst Biol 56:879–886. https://doi.org/10.1080/10635150701701083

Denchev TT, Kemler M, Begerow D, Denchev CM (2019) Molecular and morphological evidence reveals a new smut fungus, *Microbotryum arcticum* (Microbotryaceae), on *Silene uralensis* (Caryophyllaceae) from Greenland and Canada. Willdenowia 49:241–255. https://doi.org/10.3372/wi.49.49212

Dentinger BTM, Gaya E, O'Brien H, Suz LM, Lachlan R, Diaz-Valderrama JR, Koch RA, Aime MC (2016) Tales from the crypt: genome mining from fungarium specimens improves resolution of the mushroom tree of life. Biol J Linn Soc 117:11–32. https://doi.org/10.1111/bij.12553

Desjardins CA, Giamberardino C, Sykes SM, Yu C-H, Tenor JL, Chen Y, Yang T, Jones AM, Sun S, Haverkamp MR, Heitman J, Litvintseva AP, Perfect JR, Cuomo CA (2017) Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. Genome Res 27:1207–1219. https://doi.org/10.1101/gr.218727.116

Devier B, Aguileta G, Hood ME, Giraud T (2010) Using phylogenies of pheromone receptor genes in the *Microbotryum violaceum* species complex to investigate possible speciation by hybridization. Mycologia 102:689–696. https://doi.org/10.3852/09-192

Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pöhlmann R, Luedi P, Choi S, Wing RA, Flavier A, Gaffney TD, Philippsen P (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. Science 304:304–307. https://doi.org/10.1126/science.1095781

Dujon B (2010) Yeast evolutionary genomics. Nat Rev Genet 11:512–524. https://doi.org/10.1038/nrg2811

Dujon BA, Louis EJ (2017) Genome diversity and evolution in the budding yeasts (Saccharomycotina). Genetics 206:717–750

Dunn B, Paulish T, Stanbery A, Piotrowski J, Koniges G, Krol E, Louis EJ, Liti G, Sherlock G, Rosenzweig F (2013) Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgressions. PLoS Genet 9:e1003366. https://doi.org/10.1371/journal.pgen.1003366

Farrer RA, Chang M, Davis MJ, van Dorp L, Yang D-H, Shea T, Sewell TR, Meyer W, Balloux F, Edwards HM, Chanda D, Kwenda G, Vanhove M, Chang YC, Cuomo CA, Fisher MC, Kwon-Chung KJ (2019) A new lineage of *Cryptococcus gattii* (VGV) discovered in the central Zambezian Miombo woodlands. mBIO 10:e2306–e2319. https://doi.org/10.1128/mbio.02306-19

Fell JW, Boekhout T, Fonseca A, Scorzetti G, Statzell-Tallman A (2000) Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. Int J Syst Evol Microbiol 50:1351–1371. https://doi.org/10.1099/00207713-50-3-1351

Findley K, Sun S, Fraser JA, Hsueh Y-P, Averette AF, Li W, Dietrich FS, Heitman J (2012) Discovery of a modified tetrapolar sexual cycle in *Cryptococcus amylolentus* and the evolution of *MAT* in the *Cryptococcus* species complex. PLoS Genet 8:e1002528. https://doi.org/10.1371/journal.pgen.1002528

Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Intramural sequencing center comparative sequencing program NIH, Kong HH, Segre JA (2013) Topographic diversity of fungal and bacterial communities in human skin. Nature 498:367–370. https://doi.org/10.1038/nature12171

Gabaldón T (2020a) Patterns and impacts of nonvertical evolution in eukaryotes: a paradigm shift. Ann NY Acad Sci 1476:78–92. https://doi.org/10.1111/nyas.14471

Gabaldón T (2020b) Hybridization and the origin of new yeast lineages. FEMS Yeast Res 20:foaa040. https://doi.org/10.1093/femsyr/foaa040

Gabaldón T (2020c) Grand challenges in fungal genomics and evolution. Front Fungal Biol 1:594855. https://doi.org/10.3389/ffunb.2020.594855

Gibson AK, Refrégier G, Hood ME, Giraud T (2014) Performance of a hybrid fungal pathogen on pure species and hybrid host plants. Int J Plant Sc 175:724–730. https://doi.org/10.1086/676621

Gillissen B, Bergemann J, Sandmann C, Schroeer B, Bölker M, Kahmann R (1992) A two-component regulatory system for self/nonself recognition in *Ustilago maydis*. Cell 68:647–657. https://doi.org/10.1016/0092-8674(92)90141-x

Giraud T, Jonot O, Shykoff JA (2005) Selfing propensity under choice conditions in a parasitic fungus, *Microbotryum violaceum*, and parameters influencing infection success in artificial inoculations. Int J Plant Sci 166:649–657. https://doi.org/10.1086/430098

Giraud T, Refrégier G, Le Gac M, de Vienne DM, Hood ME (2008) Speciation in fungi. Fungal Genet Biol 45:791–802. https://doi.org/10.1016/j.fgb.2008.02.001

Gojković Z, Knecht W, Zameitat E, Warneboldt J, Coutelis JB, Pynyaha Y, Neuveglise C, Møller K, Löffler M, Piškur J (2004) Horizontal gene transfer promoted evolution of the ability to propagate under

anaerobic conditions in yeasts. Mol Genet Genom 271:387–393. https://doi.org/10.1007/s00438-004-0995-7

Greig D (2007) A screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast. PLoS Genet 3:e21

Greig D (2009) Reproductive isolation in *Saccharomyces*. Heredity 102:39–44

Greig D, Borts RH, Louis EJ, Travisano M (2002) Epistasis and hybrid sterility in *Saccharomyces*. Proc Biol Sci 269:1167–1171. https://doi.org/10.1098/rspb.2002.1989

Greig D, Travisano M, Louis EJ, Borts RH (2003) A role for the mismatch repair system during incipient speciation in *Saccharomyces*. J Evol Biol 16:429–437

Guého E, Midgley G, Guillot J (1996) The genus *Malassezia* with description of four new species. Antonie Van Leeuwenhoek 69:337–355. https://doi.org/10.1007/BF00399623

Guin K, Sreekumar L, Sanyal K (2020) Implications of the evolutionary trajectory of centromeres in the Fungal Kingdom. Annu Rev Microbiol 74:835–853. https://doi.org/10.1146/annurev-micro-011720-122512

Hagen F, Khayhan K, Theelen B, Kolecka A, Polacheck I, Sionov E, Falk R, Parnmen S, Lumbsch HT, Boekhout T (2015) Recognition of seven species in the *Cryptococcus gattii/Cryptococcus neoformans* species complex. Fungal Genet Biol 78:16–48. https://doi.org/10.1016/j.fgb.2015.02.009

Hagen F, Lumbsch HT, Arsic Arsenijevic V, Badali H, Bertout S, Billmyre RB, Bragulat MR, Cabañes FJ, Carbia M, Chakrabarti A, Chaturvedi S, Chaturvedi V, Chen M, Chowdhary A, Colom MF, Cornely OA, Crous PW, Cuétara MS, Diaz MR, Espinel-Ingroff A, Fakhim H, Falk R, Fang W, Herkert PF, Ferrer Rodríguez C, Fraser JA, Gené J, Guarro J, Idnurm A, Illnait-Zaragozi MT, Khan Z, Khayhan K, Kolecka A, Kurtzman CP, Lagrou K, Liao W, Linares C, Meis JF, Nielsen K, Nyazika TK, Pan W, Pekmezovic M, Polacheck I, Posteraro B, de Queiroz Telles F Filho, Romeo O, Sánchez M, Sampaio A, Sanguinetti M, Sriburee P, Sugita T, Taj-Aldeen SJ, Takashima M, Taylor JW, Theelen B, Tomazin R, Verweij PE, Wahyuningsih R, Wang P, Boekhout T (2017) Importance of resolving fungal nomenclature: the case of multiple pathogenic species in the *Cryptococcus* genus. mSphere 2:e00238-17. https://doi.org/10.1128/msphere.00238-17

Hartmann FE, Rodríguez de la Vega RC, Carpentier F, Gladieux P, Cornille A, Hood ME, Giraud T (2019) Understanding adaptation, coevolution, host specialization, and mating system in castrating anther-smut fungi by combining population and comparative genomics. Annu Rev Phytopathol 57:431–457. https://doi.org/10.1146/annurev-phyto-082718-095947

Hartmann FE, Rodríguez de la Vega RC, Gladieux P, Wen-Juan M, Hood ME, Giraud T (2020) Higher gene flow in sex-related chromosomes than in autosomes during fungal divergence. Mol Biol Evol 37:668–682. https://doi.org/10.1093/molbev/msz252

Hausdorf B (2011) Progress toward a general species concept. Evolution 65:923–931

Hawksworth D, Lücking R (2017) Fungal diversity revisited: 2.2 to 3.8 million species. Microbiol Spectr 5:FUNK-0052-2016. https://doi.org/10.1128/microbiolspec.funk-0052-2016

Hawksworth DL, Hibbett DS, Kirk PM, Lücking R (2016) Proposals to permit DNA sequence data to serve as types of names of fungi. Taxon 65:899–900. https://doi.org/10.12705/654.31

Heitman J, Sun S, James TY (2013) Evolution of fungal sexual reproduction. Mycologia 105:1–27. https://doi.org/10.3852/12-253

Hibbett D (2016) The invisible dimension of fungal diversity. Science 351:1150–1151

Hongsanan S, Jeewon R, Purahong W et al (2018) Can we use environmental DNA as holotypes? Fung Div 92:1–30. https://doi.org/10.1007/s13225-018-0404-x

Honnavar P, Prasad GS, Ghosh A, Dogra S, Handa S, Rudramurthy SM (2016) *Malassezia arunalokei* sp. nov., a novel yeast species isolated from seborrheic dermatitis patients and healthy individuals from India. J Clin Microbiol 54:1826–1834. https://doi.org/10.1128/JCM.00683-16

Hou J, Friedrich A, de Montigny J, Schacherer J (2014) Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in *Saccharomyces cerevisiae*. Curr Biol 24:1153–1159

Hunter N, Chambers SR, Louis EJ, Borts RH (1996) The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. EMBO J 15:1726–1733

James TY, Srivilai P, Kües U, Vilgalys R (2006) Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. Genetics 172:1877–1891. https://doi.org/10.1534/genetics.105.051128

James TY, Pelin A, Bonen L, Ahrendt S, Sain D, Corradi N, Stajich JE (2013) Shared signatures of parasitism and phylogenomics unite Cryptomycota and Microsporidia. Curr Biol 23:1548–1553. https://doi.org/10.1016/j.cub.2013.06.057

James TY, Stajich JE, Hittinger CT, Rokas A (2020) Toward a fully resolved fungal tree of life. Ann Rev Microbiol 74:291–313

Janbon G, Ormerod KL, Paulet D, Byrnes EJ, III, Yadav V, Chatterjee G, Mullapudi N, Hon C-C, Billmyre RB, Brunel F, Bahn Y-S, Chen W, Chen Y, Chow EWL, Coppée J-Y, Floyd-Averette A, Gaillardin C, Gerik KJ, Goldberg J, Gonzalez-Hilarion S, Gujja S, Hamlin JL, Hsueh Y-P, Ianiri G, Jones S, Kodira CD, Kozubowski L, Lam W, Marra M, Mesner LD, Mieczkowski PA, Moyrand F, Nielsen K, Proux C, Rossignol T, Schein JE, Sun S, Wollschlaeger C, Wood IA, Zeng Q, Neuvéglise C, Newlon CS, Perfect JR, Lodge JK, Idnurm A, Stajich JE, Kronstad JW, Sanyal K, Heitman J, Fraser JA, Cuomo CA, Dietrich FS (2014) Analysis of the genome and transcriptome of *Cryptococcus neoformans* var. *grubii* reveals complex RNA expression and microevolution leading to virulence attenuation. PLoS Genet 10:e1004261. https://doi.org/10.1371/journal.pgen.1004261

Kahmann R, Schirawski J (2007) Mating in the smut fungi: from a to b to the downstream cascades. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA (eds) Sex in fungi. ASM Press, Washington DC, pp. 377–387. https://doi.org/10.1128/9781555815837.ch22

Kämper J, Friedrich JM, Kahmann R (2020) Creating novel specificities in a fungal nonself recognition system by single step homologous recombination events. New Phytol 228:1001–1010. https://doi.org/10.1111/nph.16755

Kao KC, Schwartz K, Sherlock G (2010) A genome-wide analysis reveals no nuclear dobzhansky-muller pairs of determinants of speciation between *S. cerevisiae* and *S. paradoxus*, but suggests more complex incompatibilities. PLoS Genet 6:e1001038

Kellner R, Vollmeister E, Feldbrügge M, Begerow D (2011) Interspecific sex in grass smuts and the genetic diversity of their pheromone-receptor system. PLoS Genet 7:e1002436. https://doi.org/10.1371/journal.pgen.1002436

Kemler M, Denchev TT, Denchev CM, Begerow D, Piątek M, Lutz M (2020) Host preference and sorus location correlate with parasite phylogeny in the smut fungal genus *Microbotryum* (Basidiomycota, Microbotryales). Mycol Prog 19:481–493. https://doi.org/10.1007/s11557-020-01571-x

Kijpornyongpan T, Aime MC (2020) Investigating the smuts: common cues, signaling pathways, and the role of *MAT* in dimorphic switching and pathogenisis. J Fungi 6:368. https://doi.org/10.3390/jof6040368

Kijpornyongpan T, Urbina H, Suh S-O, Luangsa-ard J, Aime MC, Blackwell M (2018) The Suhomyces clade: from single isolate to multiple species to disintegrating sex loci. FEMS Yeast Res 19:foy125. https://doi.org/10.1093/femsyr/foy125

Kolecka A, Khayhan K, Groenewald M, Theelen B, Arabatzis M, Velegraki A, Kostrzewa M, Mares M, Taj-Aldeen SJ, Boekhout T (2013) Identification of medically relevant species of arthroconidial yeasts by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 51:2491–2500. https://doi.org/10.1128/JCM.00470-13

Krassowski T, Kominek J, Shen XX, Opulente DA, Zhou X, Rokas A, Hittinger CT, Wolfe KH (2019) Multiple reinventions of mating-type switching during budding yeast evolution. Curr Biol 29:2555–2562

Kudrjavfsev WI (1960) Die Systematik der Hefen. Akademie-Verlag, Berlin

Kurtzman CP (1987) Prediction of biological relatedness among yeasts from comparisons of nuclear DNA complementarity. In: de Hoog GS, Smith MTh, Weijman ACM (eds) The expanding realm of yeast-like fungi. Elsevier, Amsterdam, pp 459–468

Kurtzman CP (1998) Yeast systematics: from phenotype to genotype. Food Technol Biotechnol 36:261–266

Kurtzman CP, Robnett CJ (1998) Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie Van Leeuwenhoek 73:331–371. https://doi.org/10.1023/a:1001761008817

Kurtzman CP, Fell JW, Boekhout T (2011a) The yeasts, a taxonomic study. Elsevier, Amsterdam, 2080 pp.

Kurtzman CP, Fell JW, Boekhout T, Robert V (2011b) Methods for isolation, phenotypic characterization and maintenance of yeasts. In: Kurtzman CP, Fell JW, Boekhout T (eds) The yeasts, a taxonomic study, chapter 7. Elsevier, Amsterdam, pp 87–110

Kurtzman CP, Fell JW, Boekhout T (2011c) Gene sequence analyses and other DNA-based methods for yeasts species recognition. In: Kurtzman CP, Fell JW, Boekhout T (eds) The yeasts, a taxonomic study, chapter 10. Elsevier, Amsterdam, pp 137–144

Kwon-Chung KJ, Bennett JE, Wickes BL, Meyer W, Cuomo CA, Wollenburg KR, Bicanic TA, Castañeda E, Chang YC, Chen J, Cogliati M, Dromer F, Ellis D, Filler SG, Fisher MC, Harrison TS, Holland SM, Kohno S, Kronstad JW, Lazera M, Levitz SM, Lionakis MS, May RC, Ngamskulrongroj P, Pappas PG, Perfect JR, Rickerts V, Sorrell TC, Walsh TJ, Williamson PR, Xu J, Zelazny AM, Casadevall A (2017) The case for adopting the "species complex" nomenclature for the etiologic agents of cryptococcosis. mSphere 2:e00357-16. https://doi.org/10.1128/msphere.00357-16

Lachance MA (2011) Kluyveromyces van der Walt, (1971). In: Kurtzman CP, Fell JW, Boekhout T (eds) The yeasts, a taxonomic study, chapter 35. Elsevier, Amsterdam, pp 471–482

Lachance MA (2018) C. P. Kurtzman's evolving concepts of species, genus and higher categories. FEMS Yeast Res 18:foy103. https://doi.org/10.1093/femsyr/foy103

Lachance MA (2020) Guidelines for the publication of novel yeast species descriptions in yeast. Yeast 37:251–252

Lachance MA, Bowles JM (2004) Metschnikowia similis sp. nov. and Metschnikowia colocasiae sp.nov., two ascomycetous yeasts isolated from Conotelus spp. (Coleoptera: Nitidulidae) in Costa Rica. Stud Mycol 50:69–76

Lachance MA, Bowles JM, Starmer WT, Barker JSF (1999) Kodamaea kakaduensis and Candida tolerans, two new yeast species from Australian Hibiscus flowers. Can J Microbiol 45:172–177

Lachance MA, Daniel HM, Meyer W, Prasad GS, Gautam SP, Boundy-Mills K (2003) The D1/D2 domain of the large-subunit rDNA of the yeast species Clavispora lusitaniae is unusually polymorphic. FEMS Yeast Res 4:253–258. https://doi.org/10.1016/S1567-1356(03)00113-

Lachance MA, Lawrie D, Dobson J, Piggott J (2008) Biogeography and population structure of the Neotropical endemic yeast species Metschnikowia lochheadii. Antonie Van Leeuwenhoek 94:403–414

Lachance MA, Wijayanayaka TM, Bundus JD, Wijayanayaka DN (2011) Ribosomal DNA sequence polymorphism and the delineation of two ascosporic yeast species, Metschnikowia agaves and Starmerella bombicola. FEMS Yeast Res 11:324–333

Lachance MA, Perri A, Farahbakhsh A, Starmer WT (2013) Genetic structure of Kurtzmaniella cleridarum, a cactus flower beetle yeast of the Sonoran and Mojave Deserts: speciation by distance? FEMS Yeast Res 13:674–681

Lachance MA, Collens JD, Peng XF, Wardlaw AM, Bishop L, Hou LY, Starmer WT (2016) Spatial scale, genetic structure, and speciation of Hawaiian endemic yeasts. Pac Sci 70:389–408

Lachance MA, Lee DK, Hsiang T (2020) Delineating yeast species with genome average nucleotide identity: a calibration of ANI with haplontic, heterothallic Metschnikowia species. Antonie Van Leeuwenhoek 113:2097–2106. https://doi.org/10.1007/s10482-020-01480-9

Langdon QK, Peris D, Baker EP, Opulente DA, Nguyen HV, Bond U, Gonçalves P, Sampaio JP, Libkind D, Hittinger CT (2019) Fermentation innovation through complex hybridization of wild and domesticated yeasts. Nat Ecol Evol 3:1576–1586. https://doi.org/10.1038/s41559-019-0998-8

Lee HY, Chou JY, Cheong L, Chang NH, Yang SY, Leu JY (2008) Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell 135:1065–1073

Lee SC, Ni M, Li W, Shertz C, Heitman J (2010) The evolution of sex: a perspective from the fungal kingdom. Microbiol Mol Biol Rev 74:298–340. https://doi.org/10.1128/MMBR.00005-10

Lee DK, Hsiang T, Lachance MA (2018) Metschnikowia mating genomics. Antonie Van Leeuwenhoek 111:1935–1953

Lee DK, Santos ARO, Hsiang T, Rosa CA, Lachance MA (2020) Catching speciation in the act-act 2: Metschnikowia lacustris sp. nov., a sister species to Metschnikowia dekortorum. Antonie Van Leeuwenhoek 113:753–762. https://doi.org/10.1007/s10482-020-01395-5

Leh-Louis V, Despons L, Friedrich A, Martin T, Durrens P, Casarégola S, Neuvéglise C, Fairhead C, Marck C, Cruz JA, Straub M-L, Kugler V, Sacerdot C, Uzunov Z, Thierry A, Weiss S, Bleykasten C, De Montigny J, Jacques N, Jung P, Lemaire M, Mallet S, Morel G, Richard G-F, Sarkar A, Savel G, Schacherer J, Seret M-L, Talla E, Samson G, Jubin C, Poulain J, Vacherie B, Barbe V, Pelletier E, Sherman DJ, Westhof E, Weissenbach J, Baret PV, Wincker P, Gaillardin C, Dujon B, Souciet J-L (2012) Pichia sorbitophila, an interspecies yeast hybrid, reveals early steps of genome resolution after polyploidization. G3 (Bethesda) 2:299–311

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

Libkind D, Hittinger CT, Valério E, Gonçalves C, Dover J, Johnston M, Gonçalves P, Sampaio JP (2011) Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proc Natl Acad Sci USA 108:14539–14544. https://doi.org/10.1073/pnas.1105430108

Libkind D, Cadez N, Opulente D, Langdon Q, Rosa C, Sampaio J, Goncalves P, Hittinger C, Lachance MA (2020) Towards yeast taxogenomics: lessons from novel species descriptions based on complete genome sequences. FEMS Yeast Res 20:foaa042

Lin X, Litvintseva AP, Nielsen K, Patel S, Floyd A, Mitchell TG, Heitman J (2007) αADα hybrids of Cryptococcus neoformans: evidence of same-sex mating in nature and hybrid fitness. PLoS Genet 3:1975–1990

Lin X, Nielsen K, Patel S, Heitman J (2008) Impact of mating type, serotype, and ploidy on the virulence of Cryptococcus neoformans. Infect Immun 76:2923–2938

Linnaeus C (1758) Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis,

synonymis, locis. Tomus I. Editio decima, reformata. pp. [1–4], 1–824. Holmiæ. (Salvius)

Liti G, Barton DBH, Louis EJ (2006) Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. Genetics 174:839–850

Liti, G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V, Tsai IJ, Bergman CM, Bensasson D, O'kelly MJT, Van Oudenaarden A, Barton DBH, Bailes E, Jones M, Quail MA, Goodhead I, Sims S, Smith F, Blomberg A, Durbin R, Louis EJ (2009) Population genomics of domestic and wild yeasts. Nature 458:337–341

Litvintseva AP, Lin X, Templeton I, Heitman J, Mitchell TG (2007) Many globally isolated AD hybrid strains of *Cryptococcus neoformans* originated in Africa. PLoS Path 3:e114

Liu XZ, Wang QM, Göker M, Groenewald M, Kachalkin AV, Lumbsch HT, Millanes AM, Wedin M, Yurkov AM, Boekhout T, Bai FY (2015) Towards an integrated phylogenetic classification of the Tremellomycetes. Stud Mycol 81:85–147. https://doi.org/10.1016/j.simyco.2015.12.001

Lockhart SR, Berkow EL, Chow N, Welsh RM (2017a) *Candida auris* for the clinical microbiology laboratory: not your grandfather's *Candida* species. Clin Microbiol Newsl 39:99–103

Lockhart SR, Etienne KA, Vallabhaneni S, Farooqi J, Chowdhary A, Govender NP, Colombo AL, Calvo B, Cuomo CA, Desjardins CA, Berkow EL, Castanheira M, Magobo RE, Jabeen K, Asghar RJ, Meis JF, Jackson B, Chiller T, Litvintseva AP (2017b) Simultaneous emergence of multidrug-resistant *Candida auris* on 3 continents confirmed by whole-genome sequencing and epidemiological analyses. Clin Infect Dis 64:134–140. https://doi.org/10.1093/cid/ciw691

Louis EJ (2009) Evolutionary genetics: origins of reproductive isolation. Nature 457:549–550

Louis EJ (2011) Population genomics and speciation in yeasts. Fung Biol Rev 25:136–142

Lücking R, Aime MC, Robbertse B, Miller AN, Ariyawansa HA, Aoki T, Cardinali G, Crous PW, Druzhinina IS, Geiser DM, Hawksworth DL, Hyde KD, Irinyi L, Jeewon R, Johnston PR, Kirk PM, Malosso E, May TW, Meyer W, Öpik M, Robert V, Stadler M, Thines M, Vu D, Yurkov AM, Zhang N, Schoch CL (2020) Unambiguous identification of fungi: where do we stand and how accurate and precise is fungal DNA barcoding? IMA Fungus 11:14. https://doi.org/10.1186/s43008-020-00033-z

Lücking R, Aime MC, Robbertse B, Miller AN, Ariyawansa HA, Aoki T, Cardinali G, Crous PW, Druzhinina IS, Geiser DM, Hawksworth DL, Hyde KD, Irinyi L, Jeewon R, Johnston PR, Kirk PM, Malosso E, May TW, Meyer W, Öpik M, Robert V, Stadler M, Thines M, Vu D, Yurkov AM, Zhang N, Schoch CL (2021) Fungal taxonomy and sequence-based nomenclature. Nat Microbiol 6:540–548

Lutz M, Göker M, Piatek M, Kemler M, Begerow D, Oberwinkler F (2005) Anther smuts of *Caryophyllaceae*: molecular characters indicate host-dependent species delimitation. Mycol Progr 4:225–238. https://doi.org/10.1007/s11557-006-0126-4

Marcet-Houben M, Gabaldón T (2015) Beyond the whole-genome duplication: phylogenetic evidence for an ancient interspecies hybridization in the baker's yeast lineage. PLoS Biol 13:e1002220. https://doi.org/10.1371/journal.pbio.1002220

Marinoni G, Lachance MA (2004) Speciation in the large-spored Metschnikowia clade and establishment of a new species, *Metschnikowia borealis* comb. nov. FEMS Yeast Res 4:587–596

Martini A, Phaff HJ (1973) The optical determination of DNA–DNA homologies in yeasts. Ann Microbiol 23:59–68

May TW, Redhead SA, Bensch K, Hawksworth DL, Lendemer J, Lombard L, Turland NJ (2019) Chapter F of the International Code of Nomenclature for algae, fungi, and plants as approved by the 11th International Mycological Congress, San Juan, Puerto Rico, July 2018. IMA Fungus 10:1–14

Mayden RL (2002) On biological species, species concepts and individuation in the natural world. Fish Fisheries 3:171–196

Mayr E (1969) Principles of systematic zoology. McGraw Hill, New York/London

Mayr E (1996) What is a species, and what is not? Philos Sci 63:262–277. https://doi.org/10.1086/289912

McDonald M, Rice D, Desai M (2016) Sex speeds adaptation by altering the dynamics of molecular evolution. Nature 531:233–236

Meier R, Willmann R (2000) The Hennigian Species concept. In: Wheeler QD, Meier R (eds) Species concepts and phylogenetic theory. Columbia University Press, New York, pp 30–43

Mendel G (1866) Versuche über Pflanzen-Hybriden. Verhand naturforsch Ver Brunn 4:3–47

Mixão V, Gabaldón T (2018) Hybridization and emergence of virulence in opportunistic human yeast pathogens. Yeast 35:5–20. https://doi.org/10.1002/yea.3242

Mixão V, Gabaldón T (2020) Genomic evidence for a hybrid origin of the yeast opportunistic pathogen *Candida albicans*. BMC Biol 18:48. https://doi.org/10.1186/s12915-020-00776-6

Mixão V, Hansen AP, Saus E, Boekhout T, Lass-Florl C, Gabaldón T (2019) Whole-genome sequencing of the opportunistic yeast pathogen *Candida inconspicua* uncovers its hybrid origin. Front Genet 10:383. https://doi.org/10.3389/fgene.2019.00383

Mixão V, Saus E, Boekhout T, Gabaldón T (2021) Extreme diversification driven by parallel events of massive loss of heterozygosity in the hybrid lineage of *Candida albicans*. Genetics 217:iyaa004. https://doi.org/10.1093/genetics/iyaa004

Mizusawa M, Miller H, Green R, Lee R, Durante M, Perkins R, Hewitt C, Simner PJ, Carroll KC, Hayden RT, Zhang SX (2017) Can multidrug-resistant *Candida auris* be reliably identified in clinical microbiology laboratories? J Clin Microbiol 55:638–640

Morales L, Dujon B (2012) Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. Microbiol Mol Biol Rev 76:721–739. https://doi.org/10.1128/MMBR.00022-12

Muñoz JF, Gade L, Chow NA, Loparev VN, Juieng P, Berkow EL, Farrer RA, Litvintseva AP, Cuomo CA (2018) Genomic insights into multidrug-resistance, mating and virulence in *Candida auris* and related emerging species. Nat Commun 9:5346. https://doi.org/10.1038/s41467-018-07779-6

Naranjo-Ortiz MA, Gabaldón T (2020) Fungal evolution: cellular, genomic and metabolic complexity. Biol Rev Camb Philos Soc 95:1198–1232. https://doi.org/10.1111/brv.12605

Naumov GI (1987) Genetic basis for classification and identification of the ascomycetes yeasts. Stud Mycol 30:469–475

Naumov GI (2015) The yeast *Komagataella*: a genetic genus in accordance with interspecies hybridization. Microbiology (Moscow) 84:538–543

Naumov GI, Naumova ES, Hagler AN, Mendonca-Hagler LC, Louis EJ (1995a) A new genetically isolated population of the *Saccharomyces* sensu stricto complex from Brazil. Antonie Van Leeuwenhoek 67:351–355

Naumov GI, Naumova ES, Louis EJ (1995b) Two new genetically isolated populations of the *Saccharomyces* sensu stricto complex from Japan. J Gen Appl Microbiol 41:499–505

Naumov GI, James SA, Naumova ES, Louis EJ, Roberts IN (2000) Three new species in the *Saccharomyces* sensu stricto complex: *Saccharomyces cariocanus*, *Saccharomyces kudriavzevii* and *Saccharomyces mikatae*. Int J Syst Evol Microbiol 50:1931–1942

Naumov GI, Kondratieva VI, Meshcheryakova EV, Naumova ES (2016) Taxonomic genetics of methylotrophic yeast genus *Komagataella*: new biological species *K. kurtzmanii*. Russ J Genet 52:378–382

Navarro-Mendoza MI, Pérez-Arques C, Panchal S, Nicolás FE, Mondo SJ, Ganguly P, Pangilinan J, Grigoriev IV, Heitman J, Sanyal

K, Garre V (2019) Early diverging fungus *Mucor circinelloides* lacks centromeric histone CENP-A and displays a mosaic of point and regional centromeres. Curr Biol 29:3791–3802.e3796. https://doi.org/10.1016/j.cub.2019.09.024

Nguyen H-V, Boekhout T (2017) Characterization of *Saccharomyces uvarum* (Beijerinck, 1898) and related hybrids: assessment of molecular markers that predict the parent and hybrid genomes and a proposal to name yeast hybrids. FEMS Yeast Res 17:fox014. https://doi.org/10.1093/femsyr/fox014

Nguyen HV, Gaillardin C (2005) Evolutionary relationships between the former species *Saccharomyces uvarum* and the hybrids *Saccharomyces bayanus* and *Saccharomyces pastorianus*; reinstatement of *Saccharomyces uvarum* (Beijerinck) as a distinct species. FEMS Yeast Res 5:471–483. https://doi.org/10.1016/j.femsyr.2004.12.004

Ni M, Feretzaki M, Sun S, Wang X, Heitman J (2011) Sex in fungi. Annu Rev Genet 45:405–430. https://doi.org/10.1146/annurev-genet-110410-132536

Nilsson HR, Anslan R, Bahram M, Wurzbacher C, Baldrian P, Tedersoo L (2019) Mycobiome diversity: high-throughput sequencing and identification of fungi. Nat Rev Microb 17:95–109

Ökmen B, Mathow D, Hof A, Lahrmann U, Aßmann D, Doehlemann G (2018) Mining the effector repertoire of the biotrophic fungal pathogen *Ustilago hordei* during host and non-host infection. Mol Plant Pathol 19:2603–2622. https://doi.org/10.1111/mpp.12732

Ono J, Greig D, Boynton PJ (2020) Defining and disrupting species boundaries in *Saccharomyces*. Annu Rev Microbiol 74:477–495. https://doi.org/10.1146/annurev-micro-021320-014036

Ozan Bozdag G, Ono J, Denton JA, Karakoc E, Hunter N, Leu J-Y, Greig D (2021) Breaking a species barrier by enabling hybrid recombination. Curr Biol 31:R180–R181. https://doi.org/10.1016/j.cub.2020.12.038

Paccanaro P, Casbon JA, Saqi MA (2006) Spectral clustering of proteins sequences. Nucl Acids Res 34:1571

Passer AR, Coelho MA, Billmyre RB, Nowrousian M, Mittelbach M, Yurkov AM, Averette AF, Cuomo CA, Sun S, Heitman J (2019) Genetic and genomic analyses reveal boundaries between species closely related to *Cryptococcus* pathogens. mBio 10:e00764-00719. https://doi.org/10.1128/mbio.00764-19

Pentinsaari M, Ratnasingham S, Miller SE, Hebert PDN (2020) BOLD and GenBank revisited: do identification errors arise in the lab or in the sequence libraries? PLoS ONE 15:e0231814. https://doi.org/10.1371/journal.pone.0231814

Pérez-Través L, Lopes CA, Querol A, Barrio E (2014) On the complexity of the *Saccharomyces bayanus* taxon: hybridization and potential hybrid speciation. PLoS ONE 9:e93729. https://doi.org/10.1371/journal.pone.0093729

Péter G, Takashima M, Čadež N (2017) Yeast habitats: different but global. In: Buzzini P, Lachance MA, Yurkov A (eds) Yeasts in natural ecosystems: ecology. Springer, Cham, pp 39–71

Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, Sigwalt A, Barre B, Freel K, Llored A, Cruaud C, Labadie K, Aury J-M, Istace B, Lebrigand K, Barbry P, Engelen S, Lemainque A, Wincker P, Liti G, Schacherer J (2018) Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature 556:339–344

Peterson SW, Kurtzman CP (1991) Ribosomal RNA sequence divergence among sibling species of yeasts. Syst Appl Microbiol 14:124–129

Petit E, Silver C, Cornille A, Gladieux P, Rosenthal L, Bruns E, Yee S, Antonovics J, Giraud T, Hood ME (2017) Co-occurrence and hybridization of anther-smut pathogens specialized on *Dianthus* hosts. Mol Ecol 26:1877–1890. https://doi.org/10.1111/mec.14073

Piątek M, Lutz M, Kemler M (2013) *Microbotryum silenes-saxifragae* sp. nov. sporulating in the anthers of *Silene saxifraga* in southern European mountains. IMA Fungus 4:29–40. https://doi.org/10.5598/imafungus.2013.04.01.04

Piotrowski JS, Nagarajan S, Kroll E, Stanbery A, Chiotti KE, Kruckeberg AL, Dunn B, Sherlock G, Rosenzweig F (2012) Different selective pressures lead to different genomic outcomes as newly-formed hybrid yeasts evolve. BMC Evol Biol 12:46. https://doi.org/10.1186/1471-2148-12-46

Piškur J, Rozpędowska E, Polakova S, Merico A, Compagno C (2006) How did *Saccharomyces* evolve to become a good brewer? Trends Genet 22:183–186. https://doi.org/10.1016/j.tig.2006.02.002

Price CW, Fuson GB, Phaff HJ (1978) Genome comparison in yeast systematics: delimitation of species within the genera *Schwanniomyces, Saccharomyces, Debaryomyces,* and *Pichia*. Microbiol Rev 42:161–193

Priest SJ, Coelho MA, Mixão V, Clancey SA, Xu Y, Sun S, Gabaldón T, Heitman J (2021) Factors enforcing the species boundary between the human pathogens *Cryptococcus neoformans* and *Cryptococcus deneoformans*. PLoS Genet 17:e1008871. https://doi.org/10.1371/journal.pgen.1008871

Pryszcz LP, Nemeth T, Gacser A, Gabaldón T (2014) Genome comparison of *Candida orthopsilosis* clinical strains reveals the existence of hybrids between two distinct subspecies. Genome Biol Evol 6:1069–1078. https://doi.org/10.1093/gbe/evu082

Pryszcz LP, Németh T, Saus E, Ksiezopolska E, Hegedűsová E, Nosek J, Wolfe KH, Gacser A, Gabaldón T (2015) The genomic aftermath of hybridization in the opportunistic pathogen *Candida metapsilosis*. PLoS Genet 11:e1005626. https://doi.org/10.1371/journal.pgen.1005626

Quaedvlieg W, Binder M, Groenewald JZ, Summerell BA, Carnegie AJ, Burgess TI, Crous PW (2014) Introducing the consolidated species concept to resolve species in the *Teratosphaeriaceae*. Persoonia 33:1–40. https://doi.org/10.3767/003158514X681981

Rämä T, Davey ML, Nordén J, Halvorsen R, Blaalid R, Mathiassen GH, Alsos IG, Kauserud H (2016) Fungi sailing the Arctic Ocean: speciose communities in North Atlantic driftwood as revealed by high-throughput amplicon sequencing. Microb Ecol 72:295–304. https://doi.org/10.1007/s00248-016-0778-9

Ray J (1686) Historia plantarum. Clark, London

Rhodes J, Desjardins CA, Sykes SM, Beale MA, Vanhove M, Sakthikumar S, Chen Y, Gujja S, Saif S, Chowdhary A, Lawson DJ, Ponzio V, Colombo AL, Meyer W, Engelthaler DM, Hagen F, Illnait-Zaragozi MT, Alanio A, Vreulink J-M, Heitman J, Perfect JR, Litvintseva AP, Bicanic T, Harrison TS, Fisher MC, Cuomo CA (2017) Tracing genetic exchange and biogeography of *Cryptococcus neoformans* at the global population level. Genetics 207:327–346. https://doi.org/10.1534/genetics.117.203836

Robert V, Vu D, Amor AB, van de Wiele N, Brouwer C, Jabas B, Szoke S, Dridi A, Triki M, Daoud SB, Chouchen O, Vaas L, de Cock A, Stalpers JA, Stalpers D, Borges dos Santos F, Li W, Wu L, Zhang R, Ma J, Zhou M, Eurwilaichitr L, Ingsriswang S, Hansen K, Schoch C, Robbertse B, Irinyi L, Meyer W, Hawksworth DL, Taylor JW, Crous PW (2013) MycoBank gearing up for new horizons. IMA Fungus 4:371–379

Robert V, Cardinali G, Casadevall A (2015) Distribution and impact of yeast thermal tolerance permissive for mammalian infection. BMC Biol 13:18

Rogers DW, Denton JA, McConnell E, Greig D (2015) Experimental evolution of species recognition. Curr Biol 25:1753–1758

Rosling A, Cox F, Cruz-Martinez K, Ihrmark K, Grelet GA, Lindahl BD, Menkis A, James TY (2011) Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. Science 333:876–879. https://doi.org/10.1126/science.1206958

Rush TA, Aime MC (2013) The genus *Meira*: phylogenetic placement and description of a new species. Antonie Van Leeuwenhoek 103:1097–1106. https://doi.org/10.1007/s10482-013-9889-1

Samarasinghe H, Xu J (2018) Hybrids and hybridization in the *Cryptococcus neoformans* and *Cryptococcus gattii* species complexes. Infect Genet Evol 66:245–255. https://doi.org/10.1016/j.meegid.2018.10.011

Sankaranarayanan SR, Ianiri G, Coelho MA, Reza MH, Thimmappa BC, Ganguly P, Vadnala RN, Sun S, Siddharthan R, Tellgren-Roth C, Dawson TLJ, Heitman J, Sanyal K (2020) Loss of centromere function drives karyotype evolution in closely related *Malassezia* species. eLife 9:e53944. https://doi.org/10.7554/elife.53944

Satoh K, Makimura K, Hasumi Y, Nishiyama Y, Uchida K, Yamaguchi H (2009) *Candida auris* sp. nov., a novel ascomycetous yeast isolated from the external ear canal of an inpatient in a Japanese hospital. Microbiol Immunol 53:41–44

Schäfer AM, Kemler M, Bauer R, Begerow D (2010) The illustrated life cycle of *Microbotryum* on the host plant *Silene latifolia*. Can J Bot 88:875–885. https://doi.org/10.1139/b10-061

Schröder MS, Martinez de San Vicente K, Prandini THR, Hammel S, Higgins DG, Bagagli E, Wolfe KH, Butler G (2016) Multiple origins of the pathogenic yeast *Candida orthopsilosis* by separate hybridizations between two parental species. PLoS Genet 12:e10064 04. https://doi.org/10.1371/journal.pgen.1006404

Scorzetti G, Fell JW, Fonseca A, Statzell-Tallman A (2002) Systematics of basidiomycetous yeasts: a comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. FEMS Yeast Res 2:495–517. https://doi.org/10.1111/j.1567-1364.2002.tb00117.x

Shen X-X, Opulente DA, Kominek J, Zhou X, Steenwyk JL, Buh KV, Haase MAB, Wisecaver JH, Wang M, Doering DT, Boudouris JT, Schneider RM, Langdon QK, Ohkuma M, Endoh R, Takashima M, Manabe R-i, Čadež N, Libkind D, Rosa CA, DeVirgilio J, Hulfachor AB, Groenewald M, Kurtzman CP, Hittinger CT, Rokas A (2018) Tempo and mode of genome evolution in the budding yeast subphylum. Cell 175:1533–1545.e1520. https://doi.org/10.1016/j.cell.2018.10.023

Smith MTh, Robert V, Poot GA, Epping W, de Cock AWAM (2005) Taxonomy and phylogeny of the ascomycetous yeast genus *Zygoascus*, with proposal of *Zygoascus meyerae* sp. nov. and related anamorphic varieties. Int J Syst Evol Microbiol 55:1353–1363

Smukowski Heil CS, DeSevo CG, Pai DA, Tucker CM, Hoang ML, Dunham MJ (2017) Loss of heterozygosity drives adaptation in hybrid yeast. Mol Biol Evol 34:1596–1612. https://doi.org/10.1093/molbev/msx098

Spatafora JW, Aime MC, Grigoriev IV, Martin F, Stajich JE, Blackwell M (2017) The Fungal Tree of Life: From molecular systematics to genome scale phylogenies. Microb Spect 5:FUNK-0053-2016. https://doi.org/10.1128/microbiolspec.funk-0053-2016

Speijer D, Lukeš J, Eliáš M (2015) Sex is an old, inherent attribute of eukaryotes. Proc Natl Acad Sci 112:8827–8834

Sriswasdi S, Takashima M, Manabe R, Ohkuma M, Sugita T, Iwasaki W (2016) Global deceleration of gene evolution following recent genome hybridizations in fungi. Genome Res 26:1081–1090. https://doi.org/10.1101/gr.205948.116

Sriswasdi S, Takashima M, Manabe RI, Ohkuma M, Iwasaki W (2019) Genome and transcriptome evolve separately in recently hybridized *Trichosporon* fungi. Commun Biol 2:263. https://doi.org/10.1038/s42003-019-0515-2

Stajich JE (2017) Fungal genomes and insights into the evolution of the kingdom. Microbiol Spectr 5:FUNK-0055-2016. https://doi.org/10.1128/microbiolspec.FUNK-0055-2016

Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, Harb OS, Kissinger JC, Li W, Nayak V, Pinney DF, Stoeckert CJ Jr, Roos DS (2011) FungiDB: an integrated functional genomics database for fungi. Nucleic Acids Res 40(D1):D675–D681. https://doi.org/10.1093/nar/gkr918

Stavrou AA, Mixão V, Boekhout T, Gabaldón T (2018) Misidentification of genome assemblies in public databases: the case of *Naumovozyma dairenensis* and proposal of a protocol to correct misidentifications. Yeast 35:425–429. https://doi.org/10.1002/yea.3303

Steenkamp ET, Wingfield MJ, McTaggart AR, Wingfield BD (2017) Fungal species and their boundaries matter: definitions, mechanisms and practical implications. Fung Biol Rev 32:104–116

Stoll M, Begerow D, Oberwinkler F (2005) Molecular phylogeny of *Ustilago, Sporisorium,* and related taxa based on combined analyses of rDNA sequences. Mycol Res 109:342–356. https://doi.org/10.1017/s0953756204002229

Sullivan DJ, Westerneng TJ, Haynes KA, Bennett DE, Coleman DC (1995) *Candida dubliniensis* sp. nov.: phenotypic and molecular characterization of a novel species associated with oral candidosis in HIV-infected individuals. Microbiology 141:1507–1521. https://doi.org/10.1099/13500872-141-7-1507

Sun S, Xu J (2007) Genetic analyses of a hybrid cross between serotypes A and D strains of the human pathogenic fungus *Cryptococcus neoformans*. Genetics 177:1475–1486. https://doi.org/10.1534/genetics.107.078923

Sun S, Xu J (2009) Chromosomal rearrangements between serotype A and D strains in *Cryptococcus neoformans*. PLoS ONE 4:e5524. https://doi.org/10.1371/journal.pone.0005524

Sun S, Yadav V, Billmyre RB, Cuomo CA, Nowrousian M, Wang L, Souciet J-L, Boekhout T, Porcel B, Wincker P, Granek JA, Sanyal K, Heitman J (2017) Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. PLoS Biol 15:e2002527. https://doi.org/10.1371/journal.pbio.2002527

Sun S, Coelho MA, David-Palma M, Priest SJ, Heitman J (2019a) The evolution of sexual reproduction and the mating-type locus: links to pathogenesis of *Cryptococcus* human pathogenic fungi. Annu Rev Genet 53:417–444. https://doi.org/10.1146/annurev-genet-120116-024755

Sun S, Coelho MA, Heitman J, Nowrousian M (2019b) Convergent evolution of linked mating-type loci in basidiomycete fungi. PLoS Genet 15:e1008365. https://doi.org/10.1371/journal.pgen.1008365

Takashima M, Sriswasdi S, Manabe RI, Ohkuma M, Sugita T, Iwasaki WA (2018) *Trichosporonales* genome tree based on 27 haploid and three evolutionarily conserved 'natural' hybrid genomes. Yeast 35:99–111. https://doi.org/10.1002/yea.3284

Takashima M, Manabe RI, Nishimura Y, Endoh R, Ohkuma M, Sriswasdi S, Sugita T, Iwasaki W (2019) Recognition and delineation of yeast genera based on genomic data: lessons from *Trichosporonales*. Fungal Genet Biol 130:31–42. https://doi.org/10.1016/j.fgb.2019.04.013

Tavanti A, Davidson AD, Gow NAR et al (2005) *Candida orthopsilosis* and *Candida metapsilosis* spp. nov. to replace *Candida parapsilosis* groups II and III. J Clin Microbiol 43:284–292. https://doi.org/10.1128/JCM.43.1.284-292.2005

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, Fisher MC (2000) Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol 31:21–32. https://doi.org/10.1006/fgbi.2000.1228

Tedersoo L, Bahram M, Põlme S, Kõljalg U, Yorou NS, Wijesundera R, Villarreal Ruiz L, Vasco-Palacios AM, Thu PQ, Suija A, Smith ME, Sharp C, Saluveer E, Saitta A, Rosas M, Riit T, Ratkowsky D, Pritsch K, Põldmaa K, Piepenbring M, Phosri C, Peterson M, Parts K, Pärtel K, Otsing E, Nouhra E, Njouonkou AL, Nilsson RH, Morgado LN, Mayor J, May TW, Majuakim L, Lodge DJ, Lee SS, Larsson KH, Kohout P, Hosaka K, Hiiesalu I, Henkel TW, Harend H, Guo LD, Greslebin A, Grelet

G, Geml J, Gates G, Dunstan W, Dunk C, Drenkhan R, Dearnaley J, De Kesel A, Dang T, Chen X, Buegger F, Brearley FQ, Bonito G, Anslan S, Abell S, Abarenkov K (2014) Fungal biogeography. Global diversity and geography of soil fungi. Science 346:1256688. https://doi.org/10.1126/science.1256688

Toome M, Ohm RA, Riley RW, James TY, Lazarus KL, Henrissant B, Albu S, Boyd A, Chow J, Clum A, Heller G, Lipzen A, Nolan M, Sandor L, Zvenigorodsky N, Grigoriev IV, Spatafora JW, Aime MC (2014) Genome sequencing provides insight into the reproductive biology, nutritional mode, and ploidy of the fern pathogen *Mixia osmundae*. New Phytol 202:554–564. https://doi.org/10.1111/nph.12653

Turland N (2019) Coded decoded: a user's guide to the international code of nomenclature for algae, fungi, and plants, 2nd edn. Pensoft Publishers, Sofia

Turland NJ, Wiersema JH, Barrie FR, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Kusber W-H, Li D-Z, Marhold K, May TW, McNeill J, Monro AM, Prado J, Price MJ, Smith GF (2018) International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Regnum Vegetabile 159. Koeltz Botanical Books, Glashütten. https://doi.org/10.12705/Code.2018

van der Walt JP, Johannsen E (1984) Genus 13. *Kluyveromyces* van der Walt emend. van der Walt. In: NJW Kreger-van Rij (ed) The yeasts, a taxonomic study, 3rd ed. Elsevier Science Publishers BV, Amsterdam, pp 224–251

Vánky K (2012) Smut fungi of the world. APS, St. Paul

Vaughan Martini A, Martini A (1987) Three newly delimited species of *Saccharomyces* sensu stricto. Antonie Van Leeuwenhoek 53:77–84

Vu D, Szöke S, Wiwie C, Baumbach J, Cardinali G, Röttger R, Robert V (2014) Massive fungal biodiversity data re-annotation with multilevel clustering. Sci Rep 4:6837

Vu D, Groenewald M, Szöke S, Cardinali G, Eberhardt U, Stielow B, de Vries M, Verklej GJM, Crous PW, Boekhout T, Robert V (2016) DNA barcoding analysis of more than 9000 yeast isolates contributes to quantitative thresholds for yeast species and genera delimitation. Stud Mycol 85:91–105

Vu D, Georgievska S, Szöke S, Kuzniar A, Robert V (2018) fMLC: fast multi-level clustering and visualization of large molecular datasets. Bioinformatics 34:1577–1579

Vu D, Groenewald M, de Vries M, Gehrmann T, Stielow B, Eberhardt U, Al-Hatmi A, Groenewald JZ, Cardinali G, Houbraken J, Boekhout T, Crous PW, Robert V, Verkley GLM (2019) Large-scale analysis of filamentous fungal DNA barcodes reveals thresholds for species and higher taxon delimitation. Stud Mycol 92:135–154

Wang QM, Theelen B, Groenewald M, Bai FY, Boekhout T (2014) Moniliellomycetes and Malasseziomycetes, two new classes in Ustilaginomycotina. Persoonia 33:41–47. https://doi.org/10.3767/003158514X682313

Wang QM, Begerow D, Groenewald M, Liu XZ, Theelen B, Bai FY, Boekhout T (2015a) Multigene phylogeny and taxonomic revision of yeasts and related fungi in the Ustilaginomycotina. Stud Mycol 81:55–83. https://doi.org/10.1016/j.simyco.2015.10.004

Wang QM, Yurkov AM, Göker M, Lumbsch HT, Leavitt SD, Groenewald M, Theelen B, Liu XZ, Boekhout T, Bai FY (2015b) Phylogenetic classification of yeasts and related taxa within Pucciniomycotina. Stud Mycol 81:149–189. https://doi.org/10.1016/j.simyco.2015.12.002

Wardlaw AM, Berkers TE, Man KC, Lachance MA (2009) Population structure of two beetle-associated yeasts: comparison of a New World asexual and an endemic Nearctic sexual species in the Metschnikowia clade. Antonie Van Leeuwenhoek 96:1–15

Wolfe KH, Shields DC (1997) Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387:708–713. https://doi.org/10.1038/42711

Wu G, Zhao H, Li C, Rajapakse MP, Wong WC, Xu J, Saunders CW, Reeder NL, Reilman RA, Scheynius A, Sun S, Billmyre BR, Li W, Averette AF, Mieczkowski P, Heitman J, Theelen B, Schr^der MS, De Sessions PF, Butler G, Maurer-Stroh S, Boekhout T, Nagarajan N, Dawson TL Jr (2015) Genus-wide comparative genomics of *Malassezia* delineates its phylogeny, physiology, and niche adaptation on human skin. PLoS Genet 11:e1005614. https://doi.org/10.1371/journal.pgen.1005614

Xu J (2020) Fungal species concepts in the genomics era. Genome 63:459–468. https://doi.org/10.1139/gen-2020-0022

Xu J, Vilgalys R, Mitchell TG (2000) Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*. Mol Ecol 9:1471–1481

Xu J, Luo G, Vilgalys RJ, Brandt ME, Mitchell TG (2002) Multiple origins of hybrid strains of *Cryptococcus neoformans* with serotype AD. Microbiology 148:203–212

Xu J, Saunders CW, Hu P, Grant RA, Boekhout T, Kuramae EE, Kronstad JW, Deangelis YM, Reeder NL, Johnstone KR, Leland M, Fieno AM, Begley WM, Sun Y, Lacey MP, Chaudhary T, Keough T, Chu L, Sears R, Yuan B, Dawson TL Jr (2007) Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. Proc Natl Acad Sci USA 104:18730–18735. https://doi.org/10.1073/pnas.0706756104

Yadav V, Sun S, Billmyre RB, Thimmappa BC, Shea T, Lintner R, Bakkeren G, Cuomo CA, Heitman J, Sanyal K (2018) RNAi is a critical determinant of centromere evolution in closely related fungi. Proc Natl Acad Sci USA 115:3108–3113. https://doi.org/10.1073/pnas.1713725115

Yadav V, Sun S, Coelho MA, Heitman J (2020) Centromere scission drives chromosome shuffling and reproductive isolation. Proc Natl Acad Sci USA 117:7917–7928. https://doi.org/10.1073/pnas.1918659117

Yurkov A, Guerreiro MA, Sharma L, Carvalho C, Fonseca A (2015) Multigene assessment of the species boundaries and sexual status of the basidiomycetous yeasts *Cryptococcus flavescens* and *C. terrestris* (*Tremellales*). PLoS ONE 10:e0120400

Yurkov A, Alves A, Bai F-Y, Boundy-Mills K, Buzzini P, Čadež N, Cardinali G, Casaregola S, Vishnu V, Collin V, Fell JW, Girard V, Groenewald M, Hagen F, Hittinger CT, Kachalkin AV, Kostrzewa M, Kouvelis V, Libkind D, Liu X-Z, Maier T, Meyer W, Péter G, Piątek M, Robert R, Rosa CA, Sampaio JP, Sipiczki M, Marc Stadler M, Sugita T, Sugiyama J, Takagi H, Takashima M, Turchetti B, Wang Q-M, Boekhout T (2021) Nomenclatural issues of concern to those working with yeasts and other fungi in culture: why to avoid unneeded name changes. IMA Fungus. https://doi.org/10.1186/s43008-021-00067-x

Zamora JC, Svensson M, Kirschner R et al (2018) Considerations and consequences of allowing DNA sequence data as types of fungal taxa. IMA Fungus 9:167–175. https://doi.org/10.5598/imafungus.2018.09.01.10

Zanders SE, Eickbush MT, Yu JS, Kang JW, Fowler KR, Smith GR, Malik HS (2014) Genome rearrangements and pervasive meiotic drive cause hybrid infertility in fission yeast. eLife 24:e02630. https://doi.org/10.7554/elife.02630

Zhu Y, Engström PG, Tellgren-Roth C, Baudo CD, Kennell JC, Sun S, Billmyre RB, Schröder MS, Andersson A, Holm T, Sigurgeirsson B, Wu G, Sankaranarayanan SR, Siddharthan R, Sanyal K, Lundeberg J, Nystedt B, Boekhout T, Dawson JTL, Heitman J, Scheynius A, Lehtiö J (2017) Proteogenomics produces comprehensive and highly accurate protein-coding gene annotation in a complete genome assembly of *Malassezia sympodialis*. Nucleic Acids Res 45:2629–2643. https://doi.org/10.1093/nar/gkx006