FEMS Yeast Research, 16, 2016, fov110

doi: 10.1093/femsyr/fov110
Advance Access Publication Date: 17 December 2015
Minireview

MINIREVIEW

The birth of a deadly yeast: tracing the evolutionary emergence of virulence traits in *Candida glabrata*

Toni Gabaldón[1,2,3,*] and Laia Carreté[1,2]

[1]Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr Aiguader 88, Barcelona 08003, Spain, [2]Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain and [3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain

[*]**Corresponding author:** Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr Aiguader 88, Barcelona 08003, Spain. Tel: +34-93316028; E-mail: tgabaldon@crg.es
One sentence summary: *Candida glabrata* is an opportunistic human pathogen; genomics analyses have revealed its evolutionary path to virulence.
Editor: Cécile Fairhead

ABSTRACT

The yeast *Candida glabrata* is an opportunistic human fungal pathogen whose incidence has increased in the last two decades. Despite its name, this yeast is only distantly related to the model fungal pathogen *C. albicans*, and more closely related to *Saccharomyces cerevisiae* and other yeasts that underwent an ancient whole-genome duplication. Understanding what specific traits make *C. glabrata* a successful opportunistic pathogen within a clade of mostly innocuous yeasts, and how these compare to virulence traits in distant pathogens such as *C. albicans* is a focus of intense research. From an evolutionary perspective, uncovering how the ability to infect humans has emerged multiple, independent times in different lineages may reveal new disease mechanisms and provide us with the capacity to predict which genomic features in a clade may confer a higher potential to develop virulence against humans.

Keywords: *Candida glabrata*; virulence; evolution; Nakaseomyces; *Candida*

INTRODUCTION

The incidence of fungal infections have increased over the last decades, partly as a result of medical progress. Indeed, the advance of life expectancy, the higher survival of pre-term neonates or patients with cancer or chronic diseases, the use of catheters and transplantation therapies, are all factors that have helped fungal pathogens come to the front line of clinical problems in developed countries (Brown *et al.* 2012; Diekema *et al.* 2012). In addition, the extensive use of broad-spectrum antibiotics eliminates bacterial competitors from the commensal microbiota and promotes fungal infections (Mason *et al.* 2012). Candidiasis is a general term used to designate a fungal infection caused by any type of *Candida* yeast. Such infections can affect several surface or mucosas in the body, or can become invasive and reach internal organs, or reach the bloodstream and spread throughout the body—in which case the infection is called candidaemia. Invasive candidiasis accounts for ∼75% of all systemic fungal infections and poses a serious threat to life, particularly in immunocompromised individuals, with mortality rates reaching 46%–75% (Wilson *et al.* 2002). Candidiasis can have several origins. In fact, more than 20 different *Candida* species have been identified that can cause candidiasis, although the incidence vary greatly from species to species. *Candida albicans*, *C. glabrata* and *C. parapsilosis* rank as the three most common species, generally in this order, with the incidence of the last two increasing over the last years (Diekema *et al.* 2012). From an evolutionary perspective, however, *C. glabrata* sticks out as an outlayer, as it is more closely related to the baker's yeast

Received: 26 August 2015; **Accepted:** 13 December 2015

Saccharomyces cerevisiae than to *C. albicans* and *C. parapsilosis*. Indeed, most typical *Candida* pathogens belong to the so-called CTG or Candida clade, a group of species that share a rare particularity in their genetic code so that the CUG codon encodes the amino acid serine instead of leucine (Santos *et al.* 2011). Although several yeast pathogens belong to the CTG-clade, this is a highly diverse clade whose origins have been estimated to date back 135 million years ago (Mya) and most species within this clade are non-pathogenic nor typical commensals of human (Kurtzman, Fell and Boekhout 2011; Mühlhausen and Kollmar 2014). *Candida glabrata*, in contrast, uses the standard genetic code and belongs to a clade called Saccharomycetaceae, which diverged from the CTG-clade more than 230 Mya (Mühlhausen and Kollmar 2014) and which also contains the model yeast *S. cerevisiae*. Many species within this clade, including *S. cerevisiae* and *C. glabrata* share another rare genomic particularity in that their genomes duplicated in a common ancestor that existed roughly 100–200 Mya (Wolfe and Shields 1997)—this ancestor is now proposed to have been a hybrid species (Marcet-Houben and Gabaldón 2015). As with the CTG-clade, most species in the Saccharomycetaceae are not associated with humans. Most described species within this clade have a cosmopolitan distribution and are often isolated from various environmental niches, such as tree barks, insects, fruits or soil (Kurtzman, Fell and Boekhout 2011). In addition, this clade contains the most intensively studied eukaryotic model organism: *S. cerevisiae*, and many species that are involved in industrial processes, such as *S. cerevisiae* itself, and hybrids thereof that are used in wine, beer and bread making.

Thus, *C. glabrata* and *C. albicans* are not only distantly related, but also both species are embedded within diverse clades formed by mostly non-pathogenic nor human commensal species. Hence, despite the common generic name of the species and the disease they cause, it stands out that *C. albicans* and *C. glabrata* must have followed independent evolutionary routes to pathogenesis. Uncovering such routes holds the promise of identifying virulence mechanisms and the genomic constituents that predispose a fungal clade to generate new pathogens. In addition, understanding how pathogenesis evolved can pave the way to understand differences and similarities in the mechanisms of infection in different clades. More generally, by understanding how pathogenesis can appear during evolution, we can have a deeper understanding of the many forms it can adopt, and the thin lines that may separate a pathogenic phenotype from the absence of it. The availability of the genome sequence for *C. glabrata* (Dujon *et al.* 2004) opened the possibility of comparing its genetic constituents with that of *C. albicans* (Jones *et al.* 2004) and *S. cerevisiae* (Goffeau *et al.* 1996). Initial comparisons relied on finding similarities between the two yeast pathogens and differences between these and the non-pathogenic model yeast (Roetzer, Gabaldón and Schüller 2011). However, recent advances in sequencing technologies and evolutionary approaches have enabled us to put these differences within a more resolved and detailed evolutionary framework (Gabaldón *et al.* 2013), facilitating the reconstruction of past genomic events that tightly correlate with the evolutionary origins of the lineages were pathogenesis towards humans is present. This mini-review will survey recent advances in understanding the evolutionary emergence of virulence in the *C. glabrata* clade and will highlight some of the questions that remain open. Of note, we do not try to provide a comprehensive overview on the extended research on *C. glabrata* virulence factors and we refer the interested reader to other recent reviews on the topic (Brunke and Hube 2013; Bolotin-Fukuhara and Fairhead 2014; Rodrigues, Silva and Henriques 2014).

GENERAL CONSIDERATIONS ON VIRULENCE AS AN EMERGENT PHENOTYPE

Virulence is defined as the capacity of an organism to cause disease in a given host. Hence, microbial virulence is necessarily the outcome of a specific host–microbe interaction, and changes in either the host or the microbe can affect the degree of virulence, including the presence or absence of it. In addition, virulence can be regarded as an emergent property (Casadevall, Fang and Pirofski 2011). As such it cannot be entirely explained by the sum of its parts and is unpredictable. However, once it has appeared, the dissection of the steps that led to the emergence of virulence is possible. These considerations highlight the difficulties of studying virulence under an evolutionary perspective. For instance, virulence is only expressed under certain conditions. This is particularly true for opportunistic pathogens, and in contrast to specialized pathogens in which the pathogenic interaction with the host is part of their life cycle. Hence, opportunistic pathogens may only express their virulence potential when they encounter a debilitated host or when they find themselves in the wrong host or in the wrong tissue. Moreover, these encounters may happen only too rarely for adaptive traits to be subject to selection in the context of the whole population and evolutionary history of a given species. In this regard, it is difficult to consider virulence in opportunistic pathogens as an adaptive phenotype *per se*. Rather, we may better regard virulence as a secondary effect, one that was not developed for that purpose but is one of the possible outcomes of adapting to some other selective pressure such as a particularly natural environment or interaction with a healthy host. Indeed, in species that regularly interact with the host such as commensals, it is expected that selection acts to attenuate virulence under standard conditions, because the survival of the host is linked to the survival of the commensal. In summary, virulence in opportunistic pathogens should be regarded as an evolutionary accident, rather than as an evolutionary goal in itself. Moreover, in opportunistic pathogens, we may regard virulence as a mostly latent phenotype, one that only expresses itself under particular conditions.

As mentioned above, the increase of incidence of *Candida* opportunistic pathogens is associated with factors linked to modern technological progress. Moreover, most healthy persons carry several yeast opportunistic pathogens as normal members of their microbiota (Cui, Morris and Ghedin 2013). Thus, it seems reasonable to assume that *Candida* virulence generally results from an alteration in the host–microbe interactions, mostly initiated in the host side. Nevertheless, it is also true that, given the same alterations on the host side, some species such as *C. albicans* or *C. glabrata* have a higher likelihood of causing disease as compared to other co-existing yeasts. Thus, it seems also clear that the ability to infect humans is developed to different degrees in different yeast species. Considering that this ability is a 'latent' phenotype that was likely present long before the recent increase in incidence of *Candida* infections, one can ask the question of when and how it appeared. In other words: What past genomic re-arrangements underlie the emergence of the ability to infect humans in pathogens such as *C. glabrata*? A possible approach to address this question, which has been used in *C. glabrata* is to dissect its past genomic evolution and study its genetic specificities that differentiate it from close

Table 1. Relevant virulence traits in *C. glabrata*. Some of the traits discussed with relation to *C. glabrata* virulence are listed. The columns indicate, respectively, the trait, its relevance for the virulence phenotype, and its possible evolutionary origin. The traits and their relevance for virulence are discussed throughout the text, and inferences of their evolutionary origins are taken from Gabaldón *et al.* (2013).

Trait of *C. glabrata*	Relevance for virulence	Evolutionary origin
Growth at 37°C	Standard body temperature	Very variable trait, widespread in all *Nakaseomyces*
High stress resistance	Response to stresses present in human tissues and related to immune system	Common in *Nakaseomyces*, but lack of experimental data
Resistance to starvation	Allows survival within macrophages	Genes present across *Nakaseomyces*, but lack of experimental data
Drug resistance	Resistant to commonly used azole drugs	Not explored
High adherence	Adhesion to human tissues, clinical material, formation of biofilms	Result of a recent expansion of the adhesin repertoire. Predicted to be high in other pathogenic *Nakaseomyces* and low in non-pathogenic ones, but lack of experimental data
Presence of several auxotrophies	May be related to a higher dependence on the host	Ancient trait within *Nakaseomyces*.

non-pathogenic relatives. This quest started by simple pairwise comparisons with the model pathogen *C. albicans* and the (mostly) non-pathogenic yeast *S. cerevisiae*.

Candida glabrata differs in many ways from *C. albicans*

Candida glabrata was first named as *Cryptococcus glabratus*, and described in 1917 as a component of the human gut microbiota (Anderson 1917; reviewed in Bolotin-Fukuhara and Fairhead 2014). Years later, it started to be identified as the source of several infections either named as *Cryptococcus glabratus* (Plaut 1950), or as *Torulopsis glabrata* (Grimley, Wright and Jennings 1965). In the late 1980s, now renamed as *Candida glabrata*, it began to be commonly identified as the etiological agent of candidaemia in immunocompromised patients, and later on was recognized as an emergent pathogen (Hazen 1995). The first molecular analyses of ribosomal RNA already revealed that *C. glabrata* was only distantly related to *C. albicans* and more closely related to *S. cerevisiae* (Kurtzman and Robnett 1998). The large evolutionary distance between *C. albicans* and *C. glabrata* is reflected in many phenotypical differences between the species that likely result in different virulence mechanisms. Importantly, these differences may influence disease outcome and should be taken into account to manage disease treatment (Table 1). For instance, one of the first differences noted by clinicians is an intrinsically lower susceptibility to commonly used azole antifungals in *C. glabrata* (Diekema *et al.* 2012). Other differences include the growth as haploid cells in *C. glabrata* which cannot form true hyphae, as compared to diploid cells in *C. albicans* that have the ability to grow hyphae. Indeed switching to hyphal growth is a known virulence mechanism in *C. albicans*, which enables it to be more invasive and escape macrophage engulfment (Mayer, Wilson and Hube 2013). In contrast, *C. glabrata* copes with macrophage engulfment in a totally different way. Notably, it has been demonstrated that *C. glabrata* lets itself to be taken up by macrophages, where it can persist for long periods of time and even divide (Roetzer *et al.* 2010; Rai *et al.* 2012). Although both *C. albicans* and *C. glabrata*, as many other human fungal pathogens, have been described as asexual species, they also differ in this respect. A so-called parasexual cycle has been described in *C. albicans* by which cells of different or the same mating type fuse to form a transient tetraploid which then reverts to a diploid form by concerted loss of chromosomes (Bennett, Forche and Berman 2014). Such cycle has not been described in *C. glabrata*, where genetic data suggest that true sexual mating may occur although at very residual frequencies (Dodgson *et al.* 2005). Finally, according to the MetaPhOrs orthology database (Pryszcz, Huerta-Cepas and Gabaldón 2011), 1557 (29.5%) *C. glabrata* protein-coding genes lack an ortholog in *C. albicans*, and the reverse is true for 2257 (36.3%) *C. albicans* genes. Thus large physiological differences are expected simply due to the intrinsically different genetic composition of both species. In addition, even if the corresponding orthologs are present, the function may have diverged to an extent that it may affect its role in virulence. This may be especially true for versatile proteins such as transcription factors and other regulators than can easily rewire its network of targets. An exemplary such case is provided by the finding that null mutants in the transcription factor ACE2 had slightly attenuated virulence in *C. albicans*, whereas they are hypervirulent in *C. glabrata* (MacCallum *et al.* 2006). All these considerations underscore the fact that *C. albicans* is not a good model for *C. glabrata*'s virulence, as they seem to use quite different mechanisms for virulence (Brunke and Hube 2013).

Candida glabrata also differs in many ways from *S. cerevisiae*

Given the large differences between *C. albicans* and *C. glabrata*, it was soon realized that the closer relative model yeast *S. cerevisiae* would be a much better source of information. Indeed the number of *C. glabrata* genes without ortholog in *S. cerevisiae* is more limited (446 genes, representing 8.6% of the gene repertoire) and thus the physiology of these two species is expected to be more conserved. This higher genetic similarity encouraged a focus on the differences between these two species, as this may provide mechanistic explanations of why *C. glabrata* is an important pathogen while *S. cerevisiae* is not (Dujon *et al.* 2004; Roetzer, Gabaldón and Schüller 2011). Grossly, the main differences between *C. glabrata* and *S. cerevisiae* that were interpreted as adaptation to the human host in the former can be enumerated as follows: (i) the optimal growth temperature of *C. glabrata* is close to 37°C, (ii) *C. glabrata* has a higher stress resistance and an enhanced ability to sustain prolonged starvation as compared to

S. cerevisiae, (iii) *C. glabrata* genome has remodeled the cell-wall components, resulting in a higher adherence, (iv) *C. glabrata* has lost more genes than *S. cerevisiae* from its common ancestor, meaning that it may have a higher dependence of the host. The advantage for a human pathogen or commensal species of being able to grow well at 37°C, the normal internal temperature of the human body, is obvious. However, although the first sequenced lab strain of *S. cerevisiae* grows poorly at this temperature, optimal temperature for growth is an adaptive trait that is highly variable across *S. cerevisiae* strains (Salvadó *et al*. 2011). Furthermore, some *S. cerevisiae* strains can be part of the healthy human microbiota (Ghannoum *et al*. 2010). Thus, adaptation to grow at 37°C may be indeed useful for pathogenesis, but does not seem to be the key evolutionary event that led to the specific emergence of virulence in the *C. glabrata* lineage. In the following paragraphs, we further elaborate the mentioned observations on stress resistance, host dependence and higher adherence.

A high level of resistance to various types of stresses can be key to become a successful pathogen. Mucosal areas in mammalian tissues cause shortage of nutrients as well as osmotic and other stresses due to the presence of other microbial organisms and the protective mechanisms of the host. In particular, the adaptation to use alternative nutrients and survive long periods of starvation is key in the abovementioned ability of *C. glabrata* to survive macrophage engulfment. The ability to adhere to the host epithelial tissue is thought to play a key role in the virulence of *Candida* species. Adhesion is mediated by cell-wall-associated proteins termed adhesins that belong to diverse protein families (de Groot *et al*. 2013). Adherence is also necessary to form biofilms, which are complex surface-associated cell communities embedded in an extra-cellular matrix (Fanning and Mitchell 2012). Biofilms can increase resistance to antifungals and the host immune system, thereby facilitating increased persistence in the host. Moreover, biofilms formed in other surfaces such as catheters or other materials can be the source of hospitalary infections. A phenotype of increased adherence of a particular species is often correlating with a higher virulence and with an underlying expansion of genes mediating adhesion (Butler *et al*. 2009; Gabaldón *et al*. 2013). Thus, a higher adherence in *C. glabrata* as compared to *S. cerevisiae* and the presence of a high number of a particular type of adhesins encoded by the EPA family of genes was always considered a key addition to a pathogenic toolkit in *C. glabrata* (Cormack, Ghori and Falkow 1999; Roetzer, Gabaldón and Schüller 2011).

Finally, genome reduction is a general consequence of commensal and pathogenic lifestyles (Andersson and Kurland 1998; Ehrlich, Hiller and Hu 2008). As compared to most other environments, a host is much more stable, with temperature and other parameters being regulated through homeostasis and with some nutrients or factors readily provided by the host metabolism. In fungi, microsporidian pathogens represent an extreme case of genome reduction due to a specialized pathogenic lifestyle, with some of their genomes coding for less than 2000 proteins (Corradi and Slamovits 2011). Genome reduction in *C. glabrata*, which is not a specialized pathogen, may seem comparatively meager. Its 5202 protein-coding gene repertoire is 'only' 688 genes less than that of *S. cerevisiae*, which represents a relative size reduction of 11.7%. Nevertheless, the nature of some of the losses was pointing to possible a higher dependence on the host to overcome auxotrophies such as those of pyridoxine, thiamine and nicotinic acid (Kaur *et al*. 2005). Particularly, the loss of the whole pathway to synthesize nicotinic acid was considered as highly relevant. Nicotinic acid is a precursor of NAD^+, and without the ability to be synthesized *de novo*, *C. glabrata* should rely on an external source (e.g. the host). The finding that the de-repression of some EPA adhesins, and thereby adhesion, was regulated by the depletion of NAD^+ precursors in some human tissues (e.g. the urinary tract), pointed to a sophisticated adaptation to the human host (Domergue *et al*. 2005).

Recently, a comparative genomics approach to search for *C. glabrata* particularities was among those driving the selection of several genes to enter the first large-scale gene knock out collection in *C. glabrata* (Schwarzmüller *et al*. 2014). These included, for instance, an aspartate racemase gene inferred to have been recently acquired in *C. glabrata* through horizontal gene transfer from bacteria (Marcet-Houben and Gabaldón 2010). Further functional screening of the knock out mutants revealed that not having an ortholog in *S. cerevisiae* was not a factor increasing the likelihood of the deletion resulting in an attenuated virulence. These results indicate that most differences in gene content are not related to virulence. Of note, at the level of sequence divergence among orthologous proteins, *C. glabrata* and *S. cerevisiae* are as divergent as human and zebra fish (Gabaldón *et al*. 2013; Howe *et al*. 2013). Thus, comparing these two species to answer the question of what genomic differences may underlay the higher virulence in *C. glabrata* as compared to *S. cerevisiae* can be considered extremely naïve. Indeed it is conceptually equivalent as trying to understand how fishes can breathe under water by looking at differences between human and zebra fish genomes. Clearly, two divergent species differ in many more phenotypes—known or unknown—than the one of interest. One way of minimizing such effect is by using a higher resolution of genomic comparisons by increasing the number of compared species and their evolutionary relatedness.

Pathogenic close relatives of *C. glabrata* come into the picture

For a long time, *C. glabrata* remained the only clear pathogenic species among the Saccharomycetaceae. As such its pathogenicity appeared as a drastic discontinuity in evolutionary terms. Early molecular evolutionary analyses placed *C. glabrata* within a specific genus of the Saccharomycetaceae named *Nakaseomyces* (Kurtzman 2003). By that time, this genus contained only three known species besides *C. glabrata* and these were all environmental: *C. castellii*, *Nakaseomyces* (*Kluyveromyces*) *delphensis* and *N.* (*Kluyveromyces*) *bacillisporus*. The precise habitat and ecology of these species is unknown, and the only existing information is that collected from their isolations. More specifically, *C. castellii* appeared in a sample from acidic soil in a pine tree bog in Finland (Capriotti 1961), *N. delphensis* was isolated from a sugary deposit of dried figs in South Africa and *N. bacillisporus* was isolated from exudates from Emory oak in Arizona (USA) in an environment that also had cotton wood and picky pear cacti, and collectors noted that these trees were visited by yeast-carrying flies (Kurtzman, Fell and Boekhout 2011). Thus, no close relative of *C. glabrata* seemed to be tightly associated with humans. Instead, the natural history of the clade seemed to suggest an environmental lifestyle, probably in association with plants, and perhaps feeding on fruit sources and being dispersed by insects. Soon after the *Nakaseomyces* clade was defined, however, two new species within this group were described from human clinical samples: *C. nivariensis* (Alcoba-Flórez *et al*. 2005) and *C. bracarensis* (Correia *et al*. 2006). Their description led to re-evaluation of isolate collections putatively assigned to *C. glabrata* showing that 0.2%–2% of the clinical isolates of *C. glabrata* may actually belong to these newly

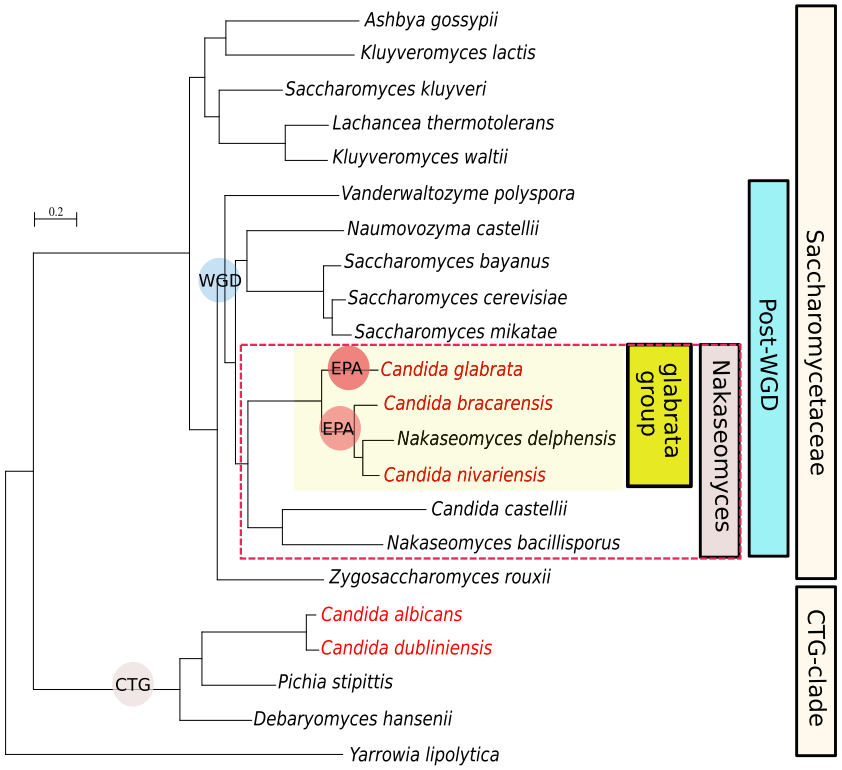

Figure 1. Phylogenetic relationships among *C. glabrata*, the other sequenced *Nakaseomyces* species and other Saccharomycotina species. Pathogenic species are indicated in red. Relevant taxonomic groups and significant evolutionary events discussed in the text are indicated with boxes at the right side of the tree and colored circles, respectively. CTG indicates a transition of the genetic code; WGD indicates the ancestral whole-genome duplication /hybridization; and EPA indicates the lineages where the two independent expansions of the EPA genes occurred. (Adapted from Gabaldón *et al.* 2013).

described species (Bishop *et al.* 2008; Lockhart *et al.* 2009). Thus, their incidence seems to be much lower than that of *C. glabrata*. We still lack much phenotypic information about these newly described emerging pathogens but early analyses seem to indicate that these species also present some of the virulence attributes of *C. glabrata*, although to a lesser degree (Moreira *et al.* 2015). Interestingly, besides clinical samples, *C. nivariensis* have also been found in isolates from floralicious insects and hibiscus flowers in Australia, suggesting it may be only sporadically associated with humans (Kurtzman, Fell and Boekhout 2011). This newer picture narrowed down the context of *C. glabrata*'s pathogenicity, pointing to an origin from an environmental clade which also contained milder pathogens. Notably, the presence of closely related pathogens with somewhat intermediate virulence towards humans brought about the idea of a stepwise increase of virulence within the clade. In any case, the availability of isolates from all these species paved the way for using genomic tools to understand what genomic re-arrangements underlay the emergence of pathogenesis.

Comparative genomics of the *Nakaseomyces*

In 2013, the sequencing of the five described relatives of *C. glabrata*, all the remaining species in the *Nakaseomyces* clade, was undertaken (Gabaldón *et al.* 2013). The genomes were sequenced with a combination of illumina and 454 technologies and resulted in high-quality assemblies with a close correspondence between scaffolds and chromosome number (Gabaldón *et al.* 2013). These genomes and their corresponding annotations can be downloaded or browsed in the Genome Resources for Yeast Chromosomes (GRYC) database, a genome repository and browser focused on high-quality genome sequences of yeasts (http://gryc.inra.fr). The availability of these genomes and its comparison with other 17 sequenced saccharomycotina species was performed using sophisticated phylogenomics techniques. In particular, the evolutionary history of every single protein-coding gene in the clade was reconstructed using maximum likelihood phylogenetic methods. This so-called phylome is publicly available at phylomeDB (http://phylomedb.org) were all reconstructed gene phylogenies can be interactively explored (Huerta-Cepas *et al.* 2014). Computational examination of these gene trees enabled accurate establishment of orthology and paralogy relationships across saccharomycotina genomes, and the inference and dating of past gene loss and gene duplication events (Gabaldón 2008). In addition, to detect more recent selective pressures, gene sequences were compared in a phylogenetic context to detect deviations in rates of synonymous and non-synonymous mutations, and the existence of arrays of tandemly duplicated genes. The recent evolution of *C. glabrata* within the *Nakaseomyces* was therefore decoded, which revealed some surprising results and served to contrast previously raised hypotheses.

First of all, since molecular phylogenetic analyses had been performed before the description of *C. bracarensis* and *C. nivariensis* (Kurtzman 2003), it remained an open question whether all pathogenic species within the clade would form a monophyletic clade, which in turn would indicate a single origin for the virulence trait and a subsequent stepwise increase in the virulence potential. The results, however, strongly indicated otherwise (Fig. 1). Two of the environmental species *N. bacillisporus* and *C. castellii* were indeed very divergent from the pathogen species forming an early branching and highly diverged group

within the clade. *Nakaseomyces delphensis*, however, was branching within the two newly recognized pathogens in a compact clade which also contained *C. glabrata* as the first diverging lineage. This clade of four species was named the glabrata group. Orthologous protein sequence identity within the glabrata group species ranged 77%–88% and high levels of conserved synteny were observed. Thus, the first shocking result was the polyphyletic origin of the three distinct human pathogens, suggesting the virulence trait may have appeared multiple (up to three), independent times. The alternative scenario involves a single origin of virulence potential at the base of the glabrata group with a subsequent secondary loss in the lineage leading to *N. delphensis*. We stress that we refer to the trait as a potential for virulence, underscoring the above discussed concept that virulence can be dealt with as a latent phenotype. Considering the divergence between the species in the glabrata group and the recent origin of our species, it seems unplausible that such an ancestor would have been a human commensal.

A second sobering result was the finding that many of the peculiarities of *C. glabrata* in comparison to *S. cerevisiae* that were thought to be important for the emergence of the virulence phenotype in the former, were shared by most, if not all, Nakaseomyces species. In particular, the loss of the synthesis of nicotinic acid was a shared trait in all Nakaseomyces, indicating this was an ancestral trait in the clade and obviously not the result of adaptation to human commensalism. As many of the Nakaseomyces species are environmental, this and other shared auxotrophies should be complemented in their respective environments or, alternatively, a salvage pathway must exist. Of note, loss of the biosynthesis pathway for nicotinic acid has been reported in more distant species such as *Kluyveromyces lactis* or *Schizosaccharomyces pombe* and appears to be a non-essential trait in many yeasts, thanks indeed to the presence of a salvage pathway. Such salvage pathway, which uses nicotinate phosphoribosyltransferase is highly conserved in eukaryotes and serves to recycle nicotinamide produced from various NAD-consuming reactions (Li and Bao 2007). Remarkably, all gene losses of *C. glabrata* relative to *S. cerevisiae* were shared by all species in the glabrata group, leaving no gene loss specific to all pathogens in the clade. In addition, all Nakaseomyces species analyzed are able to grow well at 37°C.

Final, and perhaps the most important result of that study, was the observation that variations in the number of genes encoding members of the EPA family of adhesins were highly correlating with the ability to infect humans (Gabaldón *et al*. 2013). In fact the genome of the type strain of *C. glabrata*, the most pathogenic species, encoded 18 distinct EPA genes, followed by 12 in *C. bracarensis* and 9 in *C. nivarensis*, the two newly described pathogens. In contrast, the non-pathogenic *N. delphensis* harbors a single copy. Considering the already discussed important role of adhesins in virulence, this finding was remarkable, providing the first evolutionary link between adhesion and virulence properties in the *Nakaseomyces*. To investigate whether these expansions shared by all pathogenic species were independent or the result of a common ancestral gene expansion—implying a secondary gene contraction in *N. delphensis*—the researchers performed an in-depth phylogenetic study of this family. The results (Fig. 2) indicated that EPA expansions in *C. glabrata* had occurred independently from those in *C. bracarensis* and *C. nivariensis*, whereas most of the expansions in these two latter species could be traced back to ancestral duplication events in their ancestor. That expansions in *C. glabrata* and the other Nakaseomyces pathogens have an independent origin is also supported by the fact that expanded EPA genes appear in different relative locations of their respective genomes (Fig. 2b). This observation has two important implications. First, increased adherence—and perhaps, by extension, the virulence potential—appeared at least twice in the *Nakaseomyces* clade. Second, the lack of multiple EPA copies in *N. delphensis*—and perhaps its inability to colonize or infect humans—seems to result from a secondary contraction of its adhesion repertoire.

The fact that two rapid expansions of such scale occurred in the same family of adhesins and in closely related lineages may seem highly unlikely. There are, however, some factors that may predispose the EPA adhesins for a fast amplification. First, most EPA genes are localized in subtelomeric regions. In *S. cerevisiae*, these regions have been observed to form clusters near the nuclear periphery (Maillet *et al*. 1996), which may facilitate non-homologous recombination. Indeed subtelomeric regions are prone to re-arrangements and duplications of genes (Lafontaine *et al*. 2004). Second, EPA genes are often surrounded by long stretches of DNA with high identity, extending sometimes several hundred bases. Finally, the EPA genes contain internal tandemly repeated regions which encode serine/threonine repeats that act as a spacer between the cell-wall anchoring site and the environment. These repeats are found in varying number and the length of the encoded spacer is thought to affect adhesin specificity to different substrates (Diderrich *et al*. 2015). The propensity to duplicate and change functional specificity via recombination mediated by the presence of repeats, and the particularly genomic location, may have been key contingency factors that have facilitated a rapid adaptation of the adherence properties in this lineage, provided that a selection for higher or more versatile adherence was in place.

Altogether these results provide an accurate, yet still rough, reconstruction of the genomic changes that appeared close to the appearance of the virulence potential. Altogether the obtained results suggested that many traits are ancestral in the *Nakaseomyces* clade, clearly resulting from adaptation to environments other than the human gut. This finding does not diminish the relevance of some ancient traits for present-day virulence, but it highlights the fact that these were not the key recent evolutionary steps that can be linked to the emergence of virulence. In evolutionary terms, this process is referred to as exaptation—a more appropriate term than the more commonly used pre-adaptation—which indicates that a trait that was selected in the past for another purpose may now serve to a completely different one. The finding of other pathogens that are evolutionary close to *C. glabrata* and the observation that they formed a tight clade in which three of the four known species can infect humans, although they are not monophyletic, suggests that the glabrata group shares some genomic background that predisposes the clade to the appearance of virulence traits. From the many traits that are relevant for *C. glabrata*'s pathogenesis it seems that adherence stands out as one of the ultimate steps in the path to become a species with a pathogenic potential. The fact that expansions of adhesins had appeared twice independently and that these correspond to lineages with pathogenic species reinforces the idea that shift in adherence properties, given an appropriate genomic background, may predispose species in this clade to become opportunistic pathogens.

From fruits to guts?

Have we answered the question of how pathogenesis emerged in the *C. glabrata* lineage? We are probably far away from the

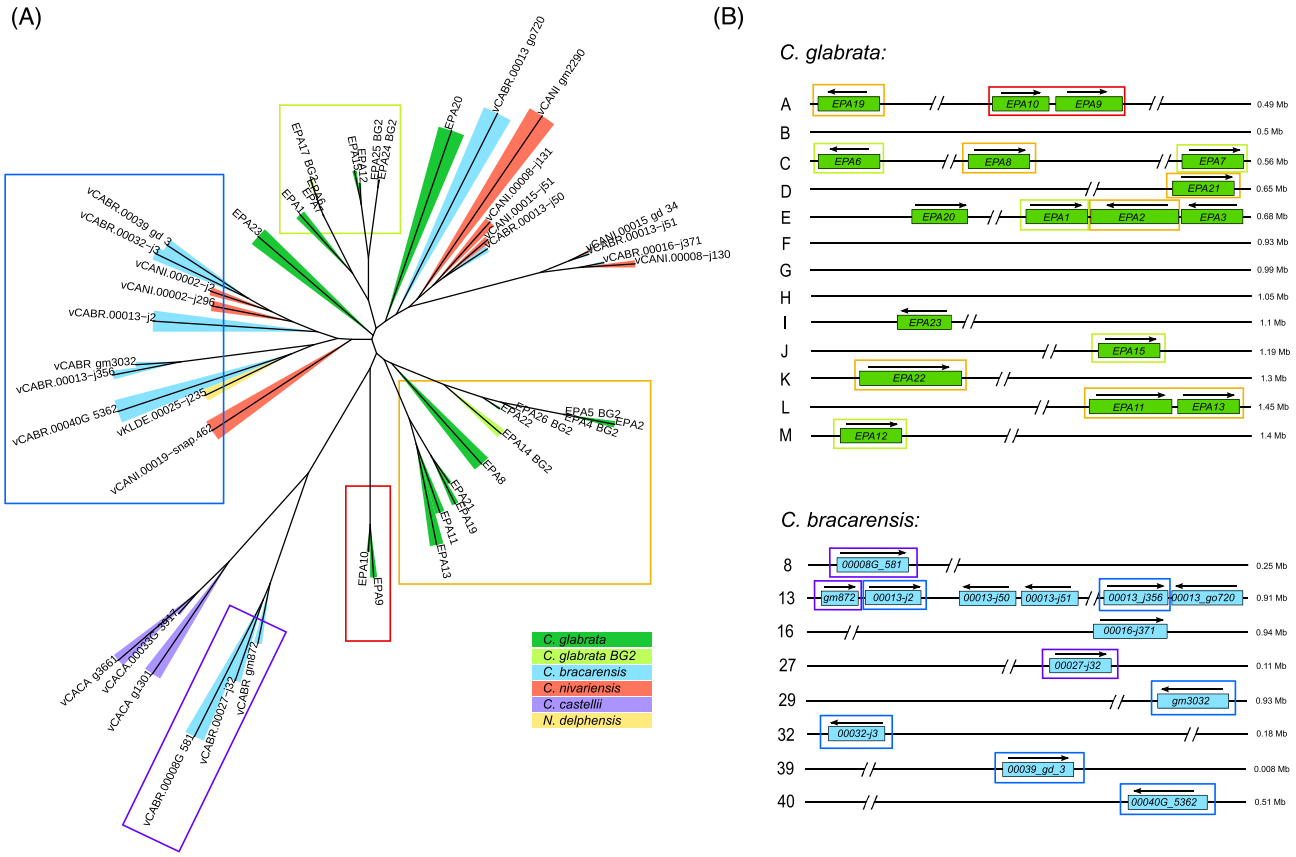

Figure 2. Multiple expansions of EPA adhesins. (**A**) Maximum likelihood phylogenetic tree representing the evolutionary relationships of the members of the EPA family of adhesins present in two strains of *C. glabrata* and other *Nakaseomyces* species. Different species and strains are indicated with colors. Main gene expansions are indicated with colored frames. (Adapted from Gabaldón *et al.* 2013). (**B**) Relative genomic location in *C. glabrata* chromosomes and *C. bracarensis* contigs of some of the recently duplicated EPA genes. Colored frames correspond to expansions highlighted in A.

real answer but the growing availability of data makes it possible to hypothesize possible scenarios. What follows is mostly wide speculation based on limited available data and we apologize for that. However, we think that by proposing plausible scenarios we will stimulate the search for additional data, perhaps supporting alternative narratives. One recurrent question is how tight *C. glabrata* is associated with humans. Although the species was first described as a commensal of the human gut, and it recurrently appears in surveys of human oral and gut microbiomes (Plaut 1950; Ghannoum *et al.* 2010; Bolotin-Fukuhara and Fairhead 2014), the fact is that we lack information of whether it is also present in some environmental niches. In addition, its prevalence in the human mouth or gut is generally much lower than *C. albicans* and even *S. cerevisiae,* and may be associated with older ages. In fact, a recent study found that relatively low number of individuals carried *C. glabrata* (15%) and that all of them were older than 40 years, with the likelihood of carrying *C. glabrata* increasing drastically over the age of 60 years (Malani *et al.* 2011). A possibility is thus that *C. glabrata* is also an opportunistic commensal of humans, us being only a secondary niche for this yeast. Some features of *C. glabrata* genome also point to this possibility. For instance, like *S. cerevisiae*, *C. glabrata* has conserved six duplicated paralogs encoding glycolytic enzymes and is an efficient alcohol producing yeast even under anaerobic conditions (Hagman *et al.* 2013). It is unclear why such trait would be conserved after long adaptation to the human gut or mouth, if other alternative niches would not be available. The idea of a wider distribution outside our guts and mouths is consistent with the observation that association to environmental niches—possibly plants and their fruits—-is an ancient adaptation of the clade, and probably still is in many present-day species. One could envision that, as other related yeasts, some *C. glabrata* ancestor was present in the environment, associated with plants and possibly feeding from a variety of sources, perhaps preferring fruits and other sugary parts of plants, where the abilities for rapid fermentation and production of alcohol may represent an advantage shared with many other yeasts of the Saccharomycetaceae. In addition, adaptation to varying alternative carbon sources and prolonged periods of starvation were present (Roetzer *et al.* 2010). Although dispersion was probably already ensured by insects, acquiring the ability to survive in the mouth and guts of the mammals that were eating the fruits may be also considered as an advantage, higher if some auxotrophies were already present. Increasing adherence, by increasing the strength of adherence and the different surfaces a cell can adhere to, seems a good way of providing a higher persistence in the mammalian digestive tract. Alternatively, adherence could have been selected earlier to facilitate transport by insects or formation of biofilms on food sources and only the more sticky species were able to persist in the digestive tract when eaten. In this framework, *C. glabrata* may have been often long colonizing us and other mammals from the environment, and this may have resulted in some further adaptations. In this context, pathogenesis would be an inevitable consequence of being a highly persistent and sticky yeast in the face of a debilitated immune system.

CONCLUSION AND OUTLOOK

Recent advances in genomics and evolutionary approaches have enabled digging into the past of human yeast pathogens, such as *C. glabrata*. The emerging picture is a complex one, with both recent traits specific to pathogens and ancient traits shared with non-pathogenic relatives playing a role in virulence. Recent results suggest that the ability to infect humans has emerged multiple times within the *Nakaseomyces* and that the appearance of this trait is linked with an expansion of the genomic repertoire of adhesins. It is also apparent that the glabrata group within Nakaseomyces already presents a genomic background that may facilitate the appearance of virulence towards humans. In this respect, we may expect that new potential emerging pathogenic yeasts may be described in the future that belong to this clade. Much progress has been performed in detailing the virulence mechanisms in *C. glabrata,* thanks to the development of novel genetic tools for this organism (Ho and Haynes 2015). As we better understand the key traits that make *C. glabrata* an efficient pathogen, we will have a better perspective for interpreting virulence within an evolutionary framework. In this respect, characterization of virulence traits in *C. bracarensis* and *C. nivariensis* is key, as they represent yet a different evolutionary acquisition of the pathogenesis towards humans. Finally, the next frontier is definitely increasing the evolutionary resolution by exploiting population genomics tool. This should include sequencing not only clinical strains from different world locations, but also commensal strains and potentially those that may be found in the environment. Clearly, understanding possible paths to pathogenesis require a deeper knowledge on the natural ecology of *C. glabrata* and its close relatives. Population genomics holds the promise of revealing most recent adaptations in the pathogenic *Nakaseomyces* as well as revealing how tight has been their co-evolution with human populations.

FUNDING

TG group acknowledges support of the Spanish Ministry of Economy and Competitiveness grants, 'Centro de Excelencia Severo Ochoa 2013–2017' SEV-2012-0208, and BIO2012-37161 cofounded by European Regional Development Fund (ERDF); from the European Union and ERC Seventh Framework Programme (FP7/2007-2013) under grant agreements FP7-PEOPLE-2013-ITN-606786 and ERC-2012-StG-310325, and grant from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No H2020-MSCA-ITN-2014-642095.

Conflict of interest. None declared.

REFERENCES

Alcoba-Flórez J, Méndez-Alvarez S, Cano J *et al.* Phenotypic and molecular characterization of *Candida nivariensis* sp. nov., a possible new opportunistic fungus. *J Clin Microbiol* 2005;**43**:4107–11.

Anderson HW. Yeast-like fungi of the human intestinal tract. *J Infect Dis* 1917;**21**:341–86.

Andersson SG, Kurland CG. Reductive evolution of resident genomes. *Trends Microbiol* 1998;**6**:263–8.

Bennett RJ, Forche A, Berman J. Rapid mechanisms for generating genome diversity: whole ploidy shifts, aneuploidy, and loss of heterozygosity. *Cold Spring Harb Perspect Med* 2014;**4**:a019604.

Bishop JA, Chase N, Magill SS *et al.* *Candida bracarensis* detected among isolates of *Candida glabrata* by peptide nucleic acid fluorescence *in situ* hybridization: susceptibility data and documentation of presumed infection. *J Clin Microbiol* 2008;**46**:443–6.

Bolotin-Fukuhara M, Fairhead C. *Candida glabrata*: a deadly companion? *Yeast Chichester Engl* 2014;**31**:279–88.

Brown GD, Denning DW, Gow NAR *et al.* Hidden killers: human fungal infections. *Sci Transl Med* 2012;**4**:165rv13.

Brunke S, Hube B. Two unlike cousins: *Candida albicans* and *C. glabrata* infection strategies. *Cell Microbiol* 2013;**15**:701–8.

Butler G, Rasmussen MD, Lin MF *et al.* Evolution of pathogenicity and sexual reproduction in eight Candida genomes. *Nature* 2009;**459**:657–62.

Capriotti A. *Torulopsis castellii* sp. nov. a yeast isolated from a Finnish soil. *J Gen Microbiol* 1961;**26**:41–3.

Casadevall A, Fang FC, Pirofski L-A. Microbial virulence as an emergent property: consequences and opportunities. *PLoS Pathog* 2011;**7**:e1002136.

Cormack BP, Ghori N, Falkow S. An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells. *Science* 1999;**285**:578–82.

Corradi N, Slamovits CH. The intriguing nature of microsporidian genomes. *Brief Funct Genomics* 2011;**10**:115–24.

Correia A, Sampaio P, James S *et al.* *Candida bracarensis* sp. nov., a novel anamorphic yeast species phenotypically similar to *Candida glabrata*. *Int J Syst Evol Micr* 2006;**56**:313–7.

Cui L, Morris A, Ghedin E. The human mycobiome in health and disease. *Genome Med* 2013;**5**:63.

Diderrich R, Kock M, Maestre-Reyna M *et al.* Structural hot spots determine functional diversity of the *Candida glabrata* epithelial adhesin family. *J Biol Chem* 2015;**290**:19597–613.

Diekema D, Arbefeville S, Boyken L *et al.* The changing epidemiology of healthcare-associated candidemia over three decades. *Diagn Micr Infec Dis* 2012;**73**:45–8.

Dodgson AR, Pujol C, Pfaller MA *et al.* Evidence for recombination in *Candida glabrata*. *Fungal Genet Biol* 2005;**42**:233–43.

Domergue R, Castaño I, De Las Peñas A *et al.* Nicotinic acid limitation regulates silencing of Candida adhesins during UTI. *Science* 2005;**308**:866–70.

Dujon B, Sherman D, Fischer G *et al.* Genome evolution in yeasts. *Nature* 2004;**430**:35–44.

Ehrlich GD, Hiller NL, Hu FZ. What makes pathogens pathogenic. *Genome Biol* 2008;**9**:225.

Fanning S, Mitchell AP. Fungal biofilms. *PLoS Pathog* 2012;**8**:e1002585.

Gabaldón T. Large-scale assignment of orthology: back to phylogenetics? *Genome Biol* 2008;**9**:235.

Gabaldón T, Martin T, Marcet-Houben M *et al.* Comparative genomics of emerging pathogens in the *Candida glabrata* clade. *BMC Genomics* 2013;**14**:623.

Ghannoum MA, Jurevic RJ, Mukherjee PK *et al.* Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. *PLoS Pathog* 2010;**6**:e1000713.

Goffeau A, Barrell BG, Bussey H *et al.* Life with 6000 genes. *Science* 1996;**274**:546, 563–7.

Grimley PM, Wright LD, Jennings AE. *Torulopsis glabrata* infection in man. *Am J Clin Pathol* 1965;**43**:216–23.

de Groot PWJ, Bader O, de Boer AD *et al.* Adhesins in human fungal pathogens: glue with plenty of stick. *Eukaryot Cell* 2013;**12**:470–81.

Hagman A, Säll T, Compagno C *et al.* Yeast "make-accumulate-consume" life strategy evolved as a multi-step process

that predates the whole genome duplication. *PloS One* 2013;**8**:e68734.
Hazen KC. New and emerging yeast pathogens. *Clin Microbiol Rev* 1995;**8**:462–78.
Ho H-L, Haynes K. *Candida glabrata*: new tools and technologies-expanding the toolkit. *FEMS Yeast Res* 2015;**15**, DOI: 10.1093/femsyr/fov066.
Howe K, Clark MD, Torroja CF *et al*. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 2013;**496**:498–503.
Huerta-Cepas J, Capella-Gutiérrez S, Pryszcz LP *et al*. PhylomeDB v4: zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Res* 2014;**42**:D897–902.
Jones T, Federspiel NA, Chibana H *et al*. The diploid genome sequence of *Candida albicans*. *P Natl Acad Sci USA* 2004;**101**:7329–34.
Kaur R, Domergue R, Zupancic ML *et al*. A yeast by any other name: *Candida glabrata* and its interaction with the host. *Curr Opin Microbiol* 2005;**8**:378–84.
Kurtzman C, Fell JW, Boekhout T. *The Yeasts: A Taxonomic Study*. Amsterdam: Elsevier, 2011.
Kurtzman CP. Phylogenetic circumscription of Saccharomyces, Kluyveromyces and other members of the Saccharomycetaceae, and the proposal of the new genera Lachancea, Nakaseomyces, Naumovia, Vanderwaltozyma and Zygotorulaspora. *FEMS Yeast Res* 2003;**4**:233–45.
Kurtzman CP, Robnett CJ. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenhoek* 1998;**73**:331–71.
Lafontaine I, Fischer G, Talla E *et al*. Gene relics in the genome of the yeast *Saccharomyces cerevisiae*. *Gene* 2004;**335**:1–17.
Li Y-F, Bao W-G. Why do some yeast species require niacin for growth? Different modes of NAD synthesis. *FEMS Yeast Res* 2007;**7**:657–64.
Lockhart SR, Messer SA, Gherna M *et al*. Identification of *Candida nivariensis* and *Candida bracarensis* in a large global collection of *Candida glabrata* isolates: comparison to the literature. *J Clin Microbiol* 2009;**47**:1216–7.
MacCallum DM, Findon H, Kenny CC *et al*. Different consequences of ACE2 and SWI5 gene disruptions for virulence of pathogenic and nonpathogenic yeasts. *Infect Immun* 2006;**74**:5244–8.
Maillet L, Boscheron C, Gotta M *et al*. Evidence for silencing compartments within the yeast nucleus: a role for telomere proximity and Sir protein concentration in silencer-mediated repression. *Genes Dev* 1996;**10**:1796–811.
Malani AN, Psarros G, Malani PN *et al*. Is age a risk factor for *Candida glabrata* colonisation? *Mycoses* 2011;**54**:531–7.
Marcet-Houben M, Gabaldón T. Acquisition of prokaryotic genes by fungal genomes. *Trends Genet TIG* 2010;**26**:5–8.
Marcet-Houben M, Gabaldón T. Beyond the whole-genome duplication: phylogenetic evidence for an ancient interspecies hybridization in the Baker's yeast lineage. *PLoS Biol* 2015;**13**:e1002220.
Mason KL, Erb Downward JR, Mason KD *et al*. *Candida albicans* and bacterial microbiota interactions in the cecum during recolonization following broad-spectrum antibiotic therapy. *Infect Immun* 2012;**80**:3371–80.
Mayer FL, Wilson D, Hube B. *Candida albicans* pathogenicity mechanisms. *Virulence* 2013;**4**:119–28.
Moreira A, Silva S, Botelho C *et al*. *Candida bracarensis*: evaluation of virulence factors and its tolerance to amphotericin B and fluconazole. *Mycopathologia* 2015;**5–6**:305–15.
Mühlhausen S, Kollmar M. Molecular phylogeny of sequenced saccharomycetes reveals polyphyly of the alternative yeast codon usage. *Genome Biol Evol* 2014;**6**:3222–37.
Plaut A. Human infection with *Cryptococcus glabratus*; report of case involving uterus and fallopian tube. *Am J Clin Pathol* 1950;**20**:377–80.
Pryszcz LP, Huerta-Cepas J, Gabaldón T. MetaPhOrs: orthology and paralogy predictions from multiple phylogenetic evidence using a consistency-based confidence score. *Nucleic Acids Res* 2011;**39**:e32.
Rai MN, Balusu S, Gorityala N *et al*. Functional genomic analysis of *Candida glabrata*-macrophage interaction: role of chromatin remodeling in virulence. *PLoS Pathog* 2012;**8**:e1002863.
Rodrigues CF, Silva S, Henriques M. *Candida glabrata*: a review of its features and resistance. *Eur J Clin Microbiol Infect Dis* 2014;**33**:673–88.
Roetzer A, Gabaldón T, Schüller C. From *Saccharomyces cerevisiae* to *Candida glabrata* in a few easy steps: important adaptations for an opportunistic pathogen. *FEMS Microbiol Lett* 2011;**314**:1–9.
Roetzer A, Gratz N, Kovarik P *et al*. Autophagy supports *Candida glabrata* survival during phagocytosis. *Cell Microbiol* 2010;**12**:199–216.
Salvadó Z, Arroyo-López FN, Guillamón JM *et al*. Temperature adaptation markedly determines evolution within the genus saccharomyces. *Appl Environ Microb* 2011;**77**:2292–302.
Santos MAS, Gomes AC, Santos MC *et al*. The genetic code of the fungal CTG clade. *C R Biol* 2011;**334**:607–11.
Schwarzmüller T, Ma B, Hiller E *et al*. Systematic phenotyping of a large-scale *Candida glabrata* deletion collection reveals novel antifungal tolerance genes. *PLoS Pathog* 2014;**10**:e1004211.
Wilson LS, Reyes CM, Stolpman M *et al*. The direct cost and incidence of systemic fungal infections. *Value Health* 2002;**5**:26–34.
Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 1997;**387**:708–13.