Yeast evolutionary genomics

Bernard Dujon

Abstract | Over the past few years, genome sequences have become available from an increasing range of yeast species, which has led to notable advances in our understanding of evolutionary mechanisms in eukaryotes. Yeasts offer us a unique opportunity to examine how molecular and reproductive mechanisms combine to affect genome architectures and drive evolutionary changes over a broad range of species. This Review summarizes recent progress in understanding the molecular mechanisms — such as gene duplication, mutation and acquisition of novel genetic material — that underlie yeast evolutionary genomics. I also discuss how results from yeasts can be extended to other eukaryotes.

Clade
A group of taxa that forms a monophyletic unit. It is applicable to any level of the taxonomical hierarchy.

Institut Pasteur, Unité de Génétique Moléculaire des Levures, Centre National de la Recherche Scientifique/ University Pierre and Marie Curie-Paris, 25 rue du Docteur Roux, F75724 Paris-CEDEX 15, France.
e-mail: bernard.dujon@pasteur.fr
doi:10.1038/nrg2811

Yeasts offer unique advantages for evolutionary genomic studies among eukaryotic organisms. These unicellular fungi, which are characterized by their ability for unlimited clonal propagation by budding or fission and the absence of fruiting bodies[1], are easily amenable to microbial genetic techniques, and the limited size and compactness of their genomes facilitate the characterization of naturally or artificially evolved populations using sequencing. The comparison of yeast genomes, which is generally applied to infer evolutionary changes, can be complemented by experimental analyses to elucidate the underlying molecular mechanisms. More than a thousand different yeast species have now been identified, and many more are likely to exist. Contrary to the commonly held view, yeasts do not represent primitive unicellular eukaryotes but instead have repeatedly emerged from distinct phylogenetic lineages of 'modern' fungi[1,2,3]. The molecular evolutionary mechanisms identified in yeasts are therefore expected to have equivalents with those in other fungi and in all animals, given the common evolutionary origin of both kingdoms among Opisthokonta[4].

The complete genome sequence of the baker's yeast *Saccharomyces cerevisiae* was a milestone of early genomics in the 1990s that directed subsequent studies[5]. The abundance of novel genes stimulated the development of systematic functional genomic approaches. Today, *S. cerevisiae* offers an unparalleled reference source for studying basic molecular mechanisms of eukaryotic cells, as more than 80% of its ~5,780 protein-coding genes have been functionally characterized[6]. The second yeast genome to be completely sequenced was that of *Schizosaccharomyces pombe*[7], but this yeast is only distantly related to *S. cerevisiae* and its genetic architecture is very different, so comparing the two species did not offer many interpretable observations in terms of evolutionary genomics. Sequences of other yeasts at more evenly distributed evolutionary intervals were needed. About 40 different yeast species have been sequenced so far (FIG. 1), and genomic-level aspects of yeast evolution are gradually being unveiled. The most attention has been focused on the Saccharomycotina (or Hemiascomycetes, the large subphylum of Ascomycota to which *S. cerevisiae* belongs), and the evolution of their genomic architecture is now reasonably well understood (BOX 1). Genomic data on other phylogenetic lineages are presently too scarce to provide an equivalent degree of understanding, but these lineages nevertheless offer interesting outgroups for comparisons.

A major, and unexpected, lesson from yeast genomics is the extensive sequence divergence observed between different lineages. This divergence goes right down to the species level and reflects intense genomic changes that contrast with the conservation of biological properties of yeasts for very long evolutionary times. Rather than offering a continuous range of gradual evolutionary adaptations, as proposed by classical Darwinian theory, genomes from distinct yeast clades, or even from species of the same genus, differ from one another in an abrupt manner. This is consistent with yeast genomes being the remnants of repeated bottleneck events that occurred in essentially clonal populations. The stochastic drift resulting from such a mode of propagation is important as it offers the possibility for non-optimized variants to survive and eventually colonize novel niches to which they may be better adapted. The recurrent emergence of this unicellular mode of life — with its specific reproductive properties — during fungal evolution is one of the main

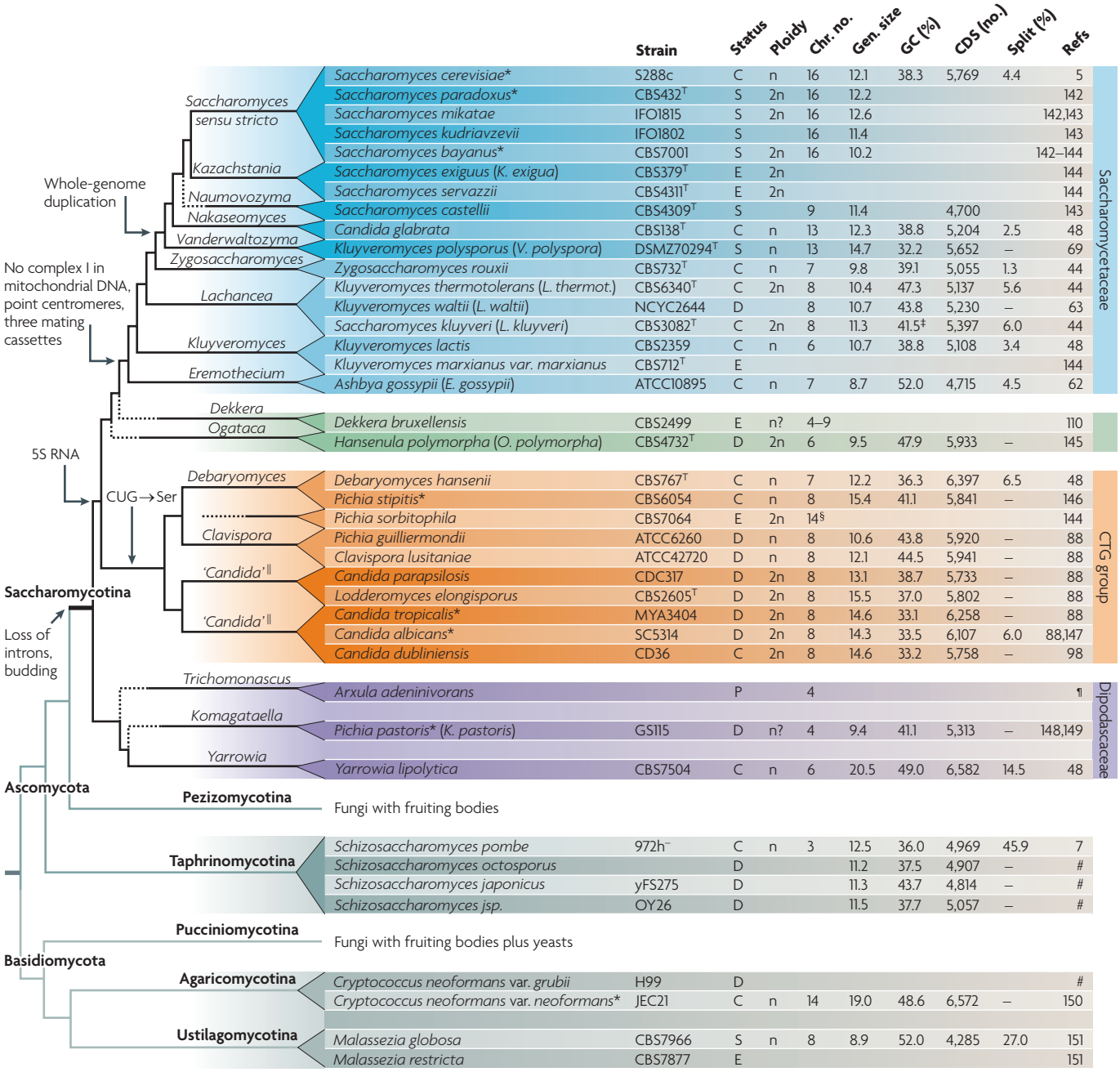

Figure 1 | **Overview of the sequenced yeast genomes.** The figure summarizes the list of sequenced yeast species with their original designation at the time of sequence publication (new taxonomies are between brackets). Coloured triangles represent clades or genera with their most recent designation shown on the left. The tree topology is adapted from REFS 137–139 and T. Boekhout and C.P. Kurtzman (personal communication). The branch lengths are arbitrary. Dotted lines illustrate uncertainty and/or incongruence between different published phylogenies. Deep-branch separations are very ancient[140,141]. Genomic architectures identify three major groups in Saccharomycotina (BOX 1), highlighted by coloured backgrounds (blue, Saccharomycetaceae family; orange, 'CTG group'; purple, Dipodascaceae and related families). The arrows point to major evolutionary events, as deduced from common genomic architecture in Saccharomycotina using a parsimony hypothesis. For each species, the sequenced the strain is indicated ([T], type strain) with the status of sequence completion (C, complete with finishing; D, draft assembly (limited number of supercontigs attributable to chromosomes); E, exploratory genome survey; P, work in progress; S, whole-genome shotgun (typically ~3–7× coverage and/or ~500–2,000 contigs)). The ploidy is also shown along with, for fully analysed sequences, the number of chromosomes (haploid set), the genome size (in megabases, excluding ribosomal DNA), the GC base content (GC), the total number of predicted protein-coding genes (the coding sequence (CDS)) and the percentage of the CDS that contains introns (the 'split'; '–' indicates that data are not available in computed or published form). *Species for which several strains have been sequenced. Only the first or the more complete sequence is indicated. [‡]Heterogeneous composition. [§]Hybrid, partially homozygotized (Génolevures Consortium, personal communication). [‖]*Candida* is commonly used to designate clades containing human pathogens, although this nomenclature only designates asexual yeasts, irrespective of their actual phylogeny (for example, *Candida glabrata*). [¶]Unpublished data from the Génolevures website. [#]Unpublished data from the Broad Institute Fungal Genome Initiative website.

Box 1 | **Summary of major architectural features of Saccharomycotina genomes**

Genomes of the budding yeasts range in size from ~9 to 20 megabases (for the haploid set) and contain a limited number of protein-coding genes (~4,700–6,500). They have few spliceosomal introns (~2–15% of split genes) and a variable number of tRNA genes (~160–510). Complete sets of genes for other non-coding RNAs occur within budding yeast genomes, in addition to limited numbers of mobile elements belonging to various families (mostly class I). Traces of RNAi machinery are generally absent, except in specific cases[123]. The presence of autonomous plasmids or viral elements is highly variable. All yeast genomes contain a large number of paralogous gene copies, which are highly diverged in their sequences and represent various types of ancestral duplications.

Against this common background, genomes of each of the three major 'families' that have been most extensively studied — the Saccharomycetaceae family, the 'CTG group' and the Dipodascaceae and related families (FIG. 1) — show distinctive features whose evolutionary origin is partially understood. Dipodascaceae and related yeasts have limited numbers of chromosomes (4–6) and dispersed 5S RNA genes. CTG yeasts and 'protoploid' Saccharomycetaceae have larger numbers of chromosomes (6–8) and 'duplicated' Saccharomycetaceae have twice as many (13–16). CTG yeasts are characterized by a common deviation of the genetic code (the CUG codon encodes serine) with corresponding changes of numerous codons[124], large centromeres and a unique mating-type locus (with two allelic forms). Saccharomycetaceae yeasts are characterized by point centromeres (which are highly conserved) and triplicate mating-type cassettes that ensure the simultaneous presence of both mating-type alleles in haploid cells (with some exceptions). A whole-genome duplication has occurred in this family, creating a subset of clades that have shorter chromosomes bearing the traces of the duplication followed by numerous gene deletions. Among them is the extensively studied baker's yeast *Saccharomyces cerevisiae*, which serves as the general reference organism for yeast genomics.

interests for evolutionary scientists studying yeasts, as is the conservation of this mode over long evolutionary periods. Now that population structures of yeasts can be readily studied at the genomic level, this opens a new perspective for evolutionary genomics. The flow of genetic material within and between populations or even species can now be precisely measured and its consequences carefully examined.

In this Review, I first discuss how the specificities of yeast reproduction influence genetic exchanges, propagation and population structures, and consequently the definition of species. I then examine the molecular mechanisms underlying evolutionary changes, covering aspects of gene duplications, mutations, chromosomal rearrangements, loss of heterozygosity (LOH), horizontal gene transfer and *de novo* gene formation, and try to place them in the broader perspective of other eukaryotic genomes. Other aspects of evolutionary significance, such as the conservation of gene families, functional repertoires, gene-expression variation, phenotypic robustness against mutational changes, the roles of episomal and mobile elements, expansions of satellite and other repeated sequences, and various features of RNA machineries, will not be comprehensively covered.

Yeast populations and species

Populations of yeasts are very different from those of multicellular organisms, in which sexual reproduction is obligatory. Under favourable conditions, all yeasts can propagate indefinitely by mitotic divisions (BOX 2), forming large clonal populations in either the haploid or diploid phase, or even in polyploids (although in this case various types of chromosomal rearrangements or loss can be observed). Note that, given their polyphyletic origin (FIG. 1), this ability for unlimited clonal growth in unicellular form constitutes one of the best criteria for differentiating yeasts from other fungi, next to the fact that their sexual states do not form within or on fruiting bodies[1]. The possibility of generating large subpopulations that can propagate without regular genetic exchange is obviously important for yeast evolution. Many species, however, conserve the ability to perform complete sexual cycles, but use it in various ways that affect (and often limit) the rates of genetic exchange in populations (BOX 3). Other yeasts are only known as asexual species, although some might undergo rare mating events. Polyploids and heterospecific hybrids are also encountered among yeast isolates, raising the difficult question of the definition of yeast species.

Population structures. The analysis of sequence polymorphisms at selected sets of loci suggested that, in nature, genetic exchanges and recombination are limited for both *Saccharomyces paradoxus*[8] and *S. cerevisiae*[9]. Genomic data precisely quantified the level of genetic exchanges in these populations. The distribution of segments of shared genealogy among three strains of *S. cerevisiae* revealed only 314 outcrossing events during ~16 million successive cell divisions, indicating a frequency of outcrossing as low as 2×10^{-5} per asexual generation[10]. A similar figure was reported for *S. paradoxus*, in which one sexual cycle was observed every 1,000 asexual generations and only 1% of these cycles corresponded to outcrossing[11]. With such low levels of outcrossing, the recombination of advantageous or deleterious alleles present in populations is limited, and subpopulations tend to form with independent accumulation of sequence variations, even in sexual species. Allopatric divergence accentuates the phenomenon[12,13]. From recent genome analysis, it was found that natural subpopulations of the wild yeast *S. paradoxus* remain well delineated within geographic boundaries[14]. Distinct subpopulations were also identified for *S. cerevisiae*, although with a higher incidence of mosaics between subpopulations, which was attributed to human domestication[14,15]. A similarly limited gene flow between yeast subpopulations was recently observed by sequencing all 18 known strains of *Saccharomyces kudriavzevii*. This led

Loss of heterozygosity
Formally, the loss of one active allele in a heterozygous pair. This loss can occur by any mechanism (mutation, deletion or gene conversion using the other allele as template). Loss of heterozygosity in yeast genomes corresponds to large-scale chromosomal regions encompassing multiple neighbouring genes.

Horizontal gene transfer
A process by which an organism incorporates genetic material from another organism that does not belong to its line of ancestry. This process is also called lateral gene transfer.

Allopatric
Refers to organisms, populations or species that inhabit distinct geographical regions.

Bateson–Dobzhansky–Muller effect
A negative effect of allelic reassortment by genetic recombination in hybrids between members of distinct populations. By extension, it is a lethal effect of reassortment by genetic recombination in crosses between parents that exhibit differential gene loss after genome duplication.

Differential gene loss
The loss of opposite copies in a pair of ohnologues between two cells that inherited the ohnologues from a common genome duplication.

Sympatric
Refers to organisms, populations or species inhabiting the same geographic area.

Ascospores
The four cellular products of a meiosis. The four ascospores are embedded in a sac called an ascus (observed in Ascomycota).

Protoploid
A general term created to designate all Saccharomycetaceae yeasts that do not originate from whole-genome duplication.

to the discovery of a 'balanced unlinked gene network polymorphism', in which in each strain, several genes involved in galactose utilization are all simultaneously either active (Portuguese strains) or pseudogenized (Japanese strains), despite their unlinked locations on multiple chromosomes[16].

Hybridization. Heterospecific hybridizations are not rare in yeasts. Successful interspecific hybridizations were experimentally obtained several years ago by mass mating between different *Saccharomyces* species[17]. Strains of the lager beer yeast *Saccharomyces pastorianus* are known from recent genomic studies to be distinct hybrids between *S. cerevisiae* and *Saccharomyces bayanus*[18,19]. Other beer strains are hybrids between *S. cerevisiae* and *S. kudriavzevii*[20,21]. These hybridization events may be accelerated by the stressful conditions imposed during alcoholic fermentations[22], but the formation of heterospecific hybrids is not limited to *Saccharomyces sensu stricto* yeasts. The asexual diploid pathogenic yeasts *Candida albicans* and *Candida dubliniensis* form tetraploid hybrids[23], and natural hybrids have also been reported between the Basidiomycota yeasts *Cryptococcus neoformans* and *Cryptococcus gattii*[24]. Similarly, two subgenomes were found in a wild isolate of *Zygosaccharomyces rouxii*[25] and in *Pichia sorbitophila* (Génolevures Consortium, personal communication). Despite the frequent occurrence of hybrids, the role of heterospecific hybridization in the evolution of yeasts remains unclear. In *Saccharomyces sensu stricto* hybrids, the two parental genomes often undergo non-reciprocal exchanges accompanied by loss of genes, chromosomal segments or even complete chromosomes, producing various chimaeras from which novel lineages might emerge[26,27]. With its megabase-long chromosomal fragment of distinct composition, *Lachancea kluyveri* might represent such a case[28], but additional examples are needed. The contradictory phylogenetic relationships of some yeasts[29] may be partly due to such hybridization phenomena.

Speciation. Given their above properties, the definition of yeast species becomes a complex question, as is often the case in organisms in which clonal propagation dominates sexual genetic exchanges. In the *Saccharomyces sensu stricto* clade (FIG. 1), in which this question has been most thoroughly addressed, the reduced meiotic fertility of heterospecific hybrids fulfils the most classical criterion used to define species[30]. However, several mechanisms combine to create this post-zygotic barrier, not all of which are applicable to other yeast clades. One mechanism is activation of the mismatch repair system by sequence divergence between two parental genomes[31]. Another mechanism is the consequence of chromosomal translocations[32–34]. A case of incompatibility was also reported between *S. cerevisiae* mitochondria and a nuclear gene of *S. bayanus*[35]. However, lineage-specific gene losses[36] are expected to have the largest effect in terms of lowering the meiotic fertility of hybrids between *Saccharomyces sensu stricto* species and other yeasts that have inherited the same genome duplication (see below). This corresponds to a special case of the Bateson–Dobzhansky–Muller effect[37], an interspecific genome incompatibility effect that is not expected to occur in other phylogenetic yeast lineages. One differential gene loss between two lineages results in half of the meiotic products of their hybrids inheriting an incomplete gene set owing to chromosomal reassortment. Combination of several differential losses between the two parental genomes further lowers the probability of forming meiotic products with a complete gene set. Comparisons with other phylogenetic branches of yeasts are, unfortunately, not presently possible as there is a lack of similar experimental data on the meiotic fertility of hybrids, a topic that needs further investigation. Prezygotic barriers also contribute to speciation in yeasts, allowing the sympatric existence of distinct species that are otherwise able to mate[38]. Germinating ascospores of *S. cerevisiae* and *S. paradoxus* show a preference for own-species mating over interspecific mating[39], whereas vegetative haploid cells of *S. paradoxus* do not[40]. The molecular basis for this phenomenon remains to be elucidated, but such a preference may have an important role in wild yeast populations, as mating in these species is suspected to occur primarily between newly germinated ascospores.

Gene duplication mechanisms

Building a minimal yeast genome, with one optimized gene per function, may sound like an attractive idea to synthetic biology engineers today, but nature has never done that. Instead, as postulated by Ohno 40 years ago[41], all genomes show numerous traces of gene duplications, and the evolutionary consequences of these duplications in creating novel gene functions have attracted considerable attention[42,43]. This is also true for yeast genomes that contain numerous series of paralogous genes from ancient or recent duplications. In *S. cerevisiae* and related species (FIG. 1), some paralogous gene pairs originate from complete genome duplication, but many dispersed paralogues and tandem gene arrays also exist. In the 'protoploid' Saccharomycetaceae species[44], the last two

Box 2 | Asymmetrical cellular division and evolutionary consequences

Although clonal yeast populations seem to increase exponentially in size under appropriate conditions, closer examination indicates a fundamentally asymmetrical mode of cellular division. This mode is immediately perceptible for the large group of Saccharomycotina (budding yeasts) but is also true for other groups[126,127]. In budding yeasts, mitoses produce a small cell, called a 'bud', that separates from a larger cell, called the 'mother'. Although both receive equivalent sets of genes, the bud is 'younger' than the mother because newly synthesized molecules tend to migrate to it, whereas ancient molecules tend to remain in the mother cell[128]. Clonal yeast populations are therefore composites of young cells, which can generate an unlimited succession of subsequent generations, and a series of gradually older cells that eventually stop dividing (FIG. 2). Asymmetrical cell division has two consequences for yeast genome evolution. First, equivalent mutational changes or genomic alterations will have different probabilities of propagation in populations depending on which cells they occur in. If this difference is small in *Saccharomyces cerevisiae*, in which mother cells can undergo a relatively large number of successive mitoses, it will increase if the proliferative ability of mother cells diminishes. Second, in haploid clones, young and old cells differ in their ability to undergo sexual cycles. In *S. cerevisiae*, at low mating-pheromone levels, the new bud is oriented towards the weak signal, increasing its probability of outcrossing[129,130]. By contrast, old cells undergo mating-type switching (BOX 3), which favours intraclonal mating.

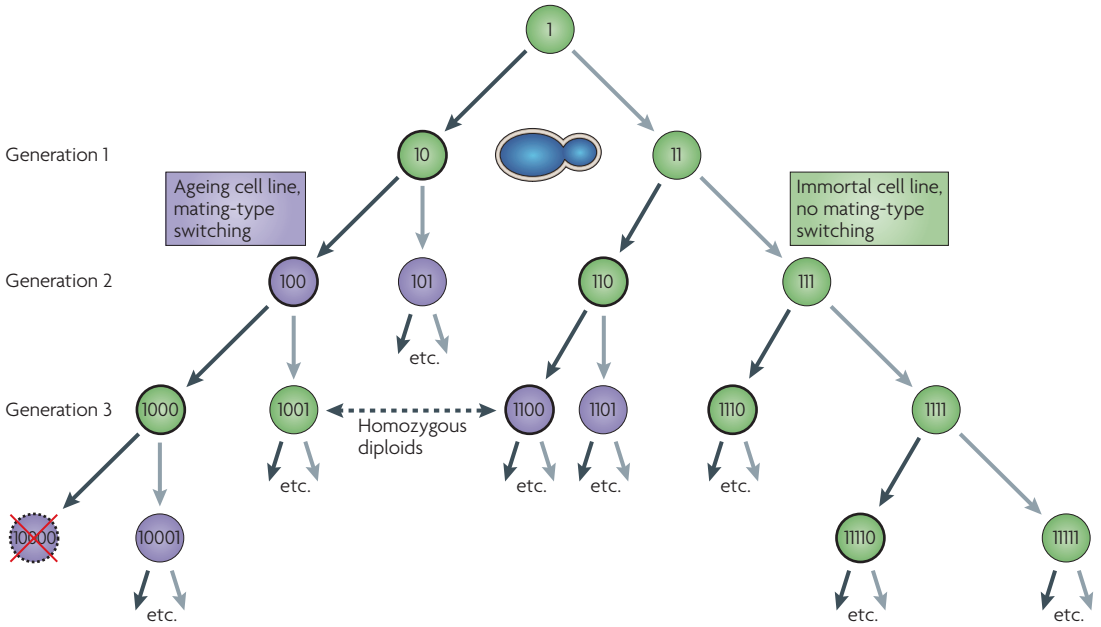

Figure 2 | **Asymmetrical cell divisions and consequences for population structures.** Mitotic divisions in Saccharomycotina are asymmetrical irrespective of the cellular ploidy. In both haploid and diploid cells, the parental cell produces a small bud, designated as the 'daughter cell' (circles with narrow outlines), and a larger sibling, designated as the 'mother' cell (circles with bold outlines), which keeps the original size and shape of the parental cell. After increasing in size, the buds reproduce the same phenomenon, producing a bud and becoming mother cells (for example, the cell marked 110). The bud lineages (right arrows) undergo an unlimited number of successive mitoses, whereas mother cells can either undergo mitosis (for example, cell 10) — producing both a new bud that will begin an immortal lineage (for example, cell 101) and a one-generation-older 'mother' cell (for example, cell 100) — or stop dividing (for example, cell 10000). In *Saccharomyces cerevisiae*, mother cells generally undergo 20–30 successive mitoses, producing the impression of a nearly exponential growth of the population in which, in any generation, one-half of the cells are newly formed buds (labelled 1*x*x1 in generation 3, in which *x* means 0 or 1), one-quarter are first-time 'mothers' (1x10), one-eighth are second-time 'mothers' (1100), and so on. This asymmetry points to similarities between yeasts and filamentous fungi in the sense that the immortal bud lineages are equivalent to the tips of hyphae, but the residual number of cellular divisions of mother lineages is specific to yeasts. Note that in Saccharomycetaceae species undergoing mating-type switching (symbolized by green and purple colours) by action of the HO endonuclease (BOX 3), immortal bud lineages never switch. Mating between members of the same descent (double-headed arrow) produces homozygous diploids, except for the mating locus itself.

phenomena occur so frequently that one-third of their protein-coding genes are members of paralogous families despite the absence of an ancestral whole-genome duplication. The numbers of paralogues are even higher in *Yarrowia lipolytica* and in species of the 'CTG group'. Four distinct mechanisms contribute to this universal genome redundancy in yeast genomes, and it would be interesting to better quantify their relative rates of action in future studies.

Expansions of tandem gene arrays. Clusters of identical or similar protein-coding genes exist in various eukaryotic genomes, but their role in evolution remains largely speculative. In yeasts, tandem gene arrays are generally not conserved[45] except in specific cases, such as B-type cyclin genes. Several examples indicate the role of tandem gene array expansions in rapid adaptive evolution. In *S. cerevisiae*, tandem expansion of the *CUP1* locus occurs when cells are grown under selective pressure for copper resistance[46]. Because this cluster is made of identical gene copies[47], its expansion (or shrinkage) can be accounted for by unequal homologous recombination events. However, in other cases tandem arrays are generally made of sequence-diverged gene copies, which is more consistent with functional diversification than with the simple need for copy-number increases, although the degree of possible concerted evolution in such tandems remains to be analysed. In the pathogenic yeast *Candida glabrata*, two large gene clusters are observed that are unique to this species[48]. One of them corresponds to the expansion of six additional YPS genes (compared with two YPS genes in *S. cerevisiae*), which encode extracellular glycosylphosphatidylinositol (GPI)-linked aspartyl proteases that are required for virulence[49]. Expansion of this cluster corresponds with the roles of these enzymes in processing a GPI-linked cell-wall adhesin that is necessary for the adherence of *C. glabrata* to mammalian cells. The other cluster is an expansion of the unique *S. cerevisiae MNT3* gene into eight copies in *C. glabrata*. This gene encodes an α-1,3-mannosyltransferase involved in cell wall biogenesis. Again, the sequence divergence between the tandem copies is consistent with functional diversification of the encoded proteins, and the cluster varies in size among clinical isolates of

CTG group
This term is used here to designate a monophyletic group of yeast species that share a common genomic architecture and a common deviation from the universal genetic code (the CUG codon specifies serine) but are taxonomically classified in diverse families, some of which contain yeasts that do not share these genomic properties.

Box 3 | **Variation of sexual cycles and evolutionary consequences**

Typically, mating in yeasts starts by the fusion of haploid cells that are similar in size and shape (isogamy), which produces zygotes[131]. Fusion involving diploid cells is also possible, and this forms polyploids[132]. Mating between haploid cells is usually followed by karyogamy, which produces a cell with a single diploid nucleus that either gives rise to a diploid clone or enters meiosis. Karyogamy is sometimes delayed (a process that can be enhanced by mutational alteration in *Saccharomyces cerevisiae*) such that dikaryotic cells may undergo mitosis before nuclear fusion. In such a case, the progeny may receive only one of the two nuclei that cohabited for some time in the same cell; this situation offers an opportunity for the two parents to exchange mitochondrial DNA (cytoduction) and various other autonomous genetic elements, such as plasmids, viruses or retrotransposons, without exchanging the bulk of their chromosomes. Depending on the species, mating can occur between genetically identical cells (homothallism, which is rare in yeasts) or requires two cells of distinct mating-types (heterothallism). In Ascomycota yeasts, mating-type is controlled by a unique chromosomal locus (designated *MAT* or *MTL*) that exists in two idiomorphic forms, usually designated *a* and *alpha*. This locus contains a limited set of genes that control the expression of numerous other genes by various molecular mechanisms that have been extensively elucidated in *S. cerevisiae* but that are beyond the scope of this Review. Alterations of the *MAT* (*MTL*) loci are frequently observed in yeast genomes, suggesting substantial rewiring of the cell identity and meiotic pathways during evolution[133]. For reasons that are not yet entirely clear, but that have important consequences for population structures and evolution, heterothallic yeast species have repeatedly recruited distinct molecular mechanisms during evolution[134–136] to allow mating-type switching of haploid cells and hence strongly reduce genetic exchange between different populations during sexual cycles.

C. glabrata[50]. Another dynamic large tandem gene array called *DUP240*, which is specific to *S. cerevisiae* and closely related species, encodes proteins that facilitate membrane trafficking[51]. Numerous other species-specific arrays are found in other yeasts, but their functional roles are not clearly understood.

Segmental duplications. Segmental duplications constitute a major signature of various eukaryotic genomes, although their abundance varies. In primates, for example, segmental duplications contribute to ~5% of the total genome size. They are distributed non-randomly in genomes with clustering near subtelomeric and pericentromeric regions[52]. Similar clustering in subtelomeric regions also exists in yeasts[44,53], but traces of segmental duplications remain scarce in other parts of yeast genomes. Yeasts, however, have been useful for elucidating the molecular mechanisms involved in the formation of segmental duplications and, given the structural similarities of segmental duplications between yeast and multicellular eukaryotes, it is likely that their varying abundance between genomes reflects different dynamic equilibriums rather than fundamental mechanistic differences[54]. In experimental cultures of *S. cerevisiae*, duplications of large chromosomal segments (tens to hundreds of kilobases) containing many genes occur spontaneously at high frequency (~10^{-7} per mitosis)[55–57]. Four types of chromosomal structures are formed (FIG. 3) with different degrees of stability during subsequent generations[58]. Intra- and interchromosomal duplications are also observed in mammalian genomes. In addition, extrachromosomal copies of a specific chromosome segment were reported[59] and similar structures have been observed in other experiments (A. Thierry, personal communication). This suggests that, after re-integration at ectopic chromosomal locations, extrachromosomal copies may contribute to genome reshuffling. It is now clear from yeast experiments that large segmental duplications result from untimely DNA synthesis events requiring polymerase-δ (Pol-δ)[60]. Dispersed repeated elements in genomes, such as remnants of Ty elements (class I retrotransposons), anchor the duplications by a *RAD52*-dependent break-induced replication (BIR) mechanism, often resulting in subtelomeric duplications or interchromosomal translocations. But duplications also occur in the absence of such elements, with no preference for genomic locations, as a result of a *RAD52*-independent microhomology/microsatellite-induced replication (MMIR) mechanism. Owing to the greater technical limitations involved, mechanisms are less precisely described for mammalian genomes, but the presence of sequence microhomology and topoisomerase binding sites at or near junctions also suggests a replication-based mechanism[52]. The influence of the segmental duplication mechanism in natural evolution of yeast genomes remains to be better quantified, but this mechanism is prone to leave behind duplicated gene copies and chromosomal rearrangements in yeast cultures that may contribute to the numerous dispersed paralogues observed.

Whole-genome duplications. The original hypothesis that *S. cerevisiae* evolved from an ancient duplication of the entire genome[61] (a hypothesis that was subsequently confirmed by sequence comparisons with other yeasts of the same Saccharomycetaceae family[62,63]) had an important influence on our understanding of the evolutionary consequences of such duplication events (see below). However, much remains to be learned about the nature of the phenomenon and why successful duplications are so rare in evolution when autopolyploidization events (tetraploids) and allopolyploidization events (hybrids) seem to be so frequent in yeasts. Notwithstanding the possibility that other duplications may exist in the many lineages that have not yet been studied or that traces of ancient duplications have been erased beyond recognition, so far no other entire genome duplication event has been identified in yeasts, despite their long evolutionary history (although another example exists in primitive fungi[64]). This suggests that only a very small proportion of polyploidization events gives rise to successful lineages in yeast evolution for reasons that remain to be clarified. One possibility is that successive deletions of genes from the initial polyploid stage create phenotypically disadvantaged intermediates owing to gene dosage imbalance or the necessary rewiring of the protein-interaction network[65,66].

Meanwhile, the unique genome duplication known in yeast evolution provided information about post-duplication events, including massive gene deletions that are prone to completely reshuffle genome maps by forming two imperfect synteny blocks from the original one (FIG. 4). In *S. cerevisiae*, only ~550 duplicated

Autopolyploidization
The formation of cells or organisms with more than two pairs of homologous chromosomes as a result of self-fertilization or non-disjunctive segregation of chromosomes during mitosis or meiosis.

Allopolyploidization
The formation of cells or organisms with more than two pairs of homologous chromosomes as a result of hybridization between distinct species.

Synteny
The physical colocalization of genes, or genetic loci, along the same chromosome. It is more often used to designate subgroups of genes along a chromosomal segment.

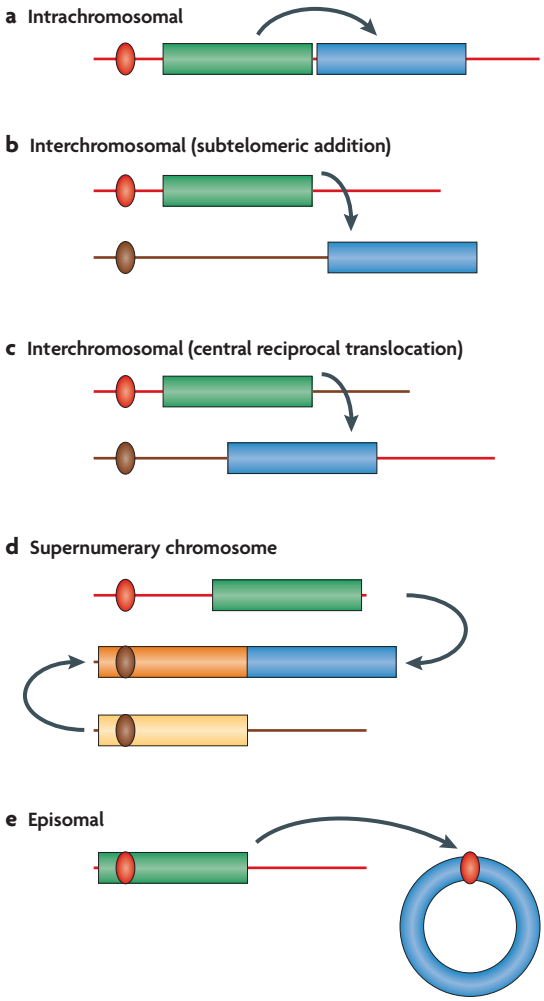

Figure 3 | **Types of spontaneous segmental duplications in *Saccharomyces cerevisiae* as observed in experimental evolution assays.** During clonal propagation, duplications of long chromosomal segments (tens to hundreds of kilobases long) containing numerous genes occur spontaneously at high frequency in both haploid and diploid *Saccharomyces cerevisiae* genomes. These duplications result in novel chromosomal structures of unequal stability during subsequent generations[58]. Blue sequences symbolize new copies of green sequences. Orange sequences symbolize new copies of yellow sequences. Tandem intrachromosomal duplications (**a**) are the most frequent structures observed and are unstable at meiosis. Interchromosomal duplications forming a novel chromosomal end (**b**) are also frequently observed. Interchromosomal duplications accompanied by reciprocal translocations of distal chromosome segments (**c**) are stable. Duplications encompassing centromeres (coloured ovals) can form supernumerary chromosomes (**d**) or circular episomes (**e**). Segmental duplications ignore gene boundaries and may create gene fusions at junctions. Some duplications originate at dispersed sequence repeats (often remnants of retrotransposons), whereas others involve sequence microhomology. Both correspond to replication accidents that depend on a novel mechanism involving the polymerase 32 (Pol32) subunit of Pol-δ[60].

Ohnologue
One of a pair of paralogues originating from a whole-genome duplication.

Neofunctionalization
The acquisition of a novel function by a gene after mutational changes. This usually applies to one of the two paralogues that are produced from a gene duplication.

Subfunctionalization
Functional specialization after mutational changes of the paralogues that are produced from a gene duplication.

pairs (ohnologues) still remain[67], and similar or lower figures are observed for other yeasts that emerged from the same duplication event, such as *C. glabrata*[48], *Naumovozyma castellii*[68] and *Vanderwaltozyma polymorpha*[69]. Although many of the deletions are common to several of these species, suggesting possible pre-speciation events, others are found in only one species, indicating post-speciation events[68,70]. The molecular mechanism by which so many individual genes disappear after duplication remains one of the problems to solve[71]. Although few relics of ancestral genes have been found among duplicated chromosomal segments[72,73], most events seem to be complete deletions of one or a few genes[74].

The remaining pairs of ohnologues shed light on post-duplication evolutionary mechanisms. Assuming that their sequences were identical to start with — that is, the duplication was an auto- rather than an allo-polyploidization event — the sequence divergence and differential expression among pairs of ohnologues suggest that newly arisen gene duplicates experienced altered selective regimes[75]. When compared with single-copy homologues from protoploid species of Saccharomycetaceae, most ohnologous pairs in duplicated species exhibit strong asymmetry in their evolution rate, which suggests neofunctionalization[76,77]. The serine/threonine protein kinases encoded by the *NPR1* and *PRR2* ohnologues of *S. cerevisiae* illustrate this: the slow-evolving NPR1p stabilizes plasma membrane amino-acid transporters, whereas the fast-evolving PRR2p is involved in the pheromone-induced signalling pathway. Interestingly, the fast-evolving copy is never an essential gene and its function is generally less well understood than that of the slow-evolving copy[77]. The expression of the two gene copies also frequently differs, consistent with the divergence of their *cis*-regulatory elements[78–80]. The *CYC1* gene of *S. cerevisiae,* which encodes isoform 1 of cytochrome *c*, is preferentially expressed in the presence of oxygen, whereas its ohnologue *CYC7*, which encodes isoform 2, is preferentially expressed under hypoxic conditions. Other cases of subfunctionalization involve subcellular localization, often separating mitochondrial from nuclear or cytoplasmic functions. About one-third of proteins encoded by pairs of ohnologues localize to distinct subcellular compartments[81]. Finally, functional redundancy persists for some ohnologous pairs[82].

Single-gene duplications. Despite the dispersion of most paralogues in genomes, no molecular mechanisms operating at the chromosomal level have been identified so far that would create single-gene duplicates at ectopic locations. The only known examples of such cases in yeasts and other eukaryotes proceed through RNA intermediates, the duplicated gene copy representing a retrogene reinserted into a new chromosomal location. This was directly demonstrated experimentally in *S. cerevisiae*, in which Ty retrotransposons can duplicate part of a gene and integrate the copy into a new chromosomal location[83]. The proposed mechanism involves template switching of the polymerase

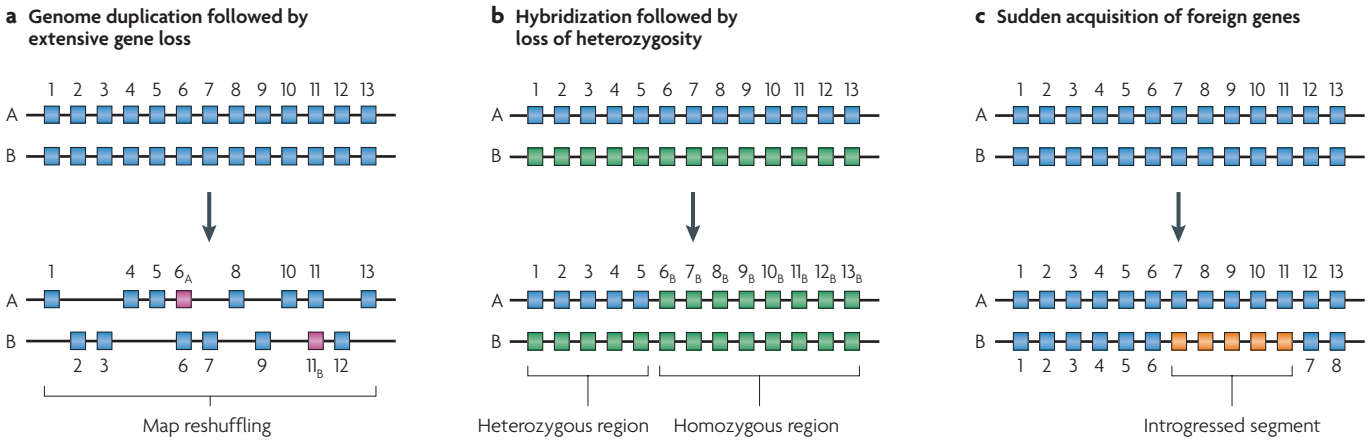

Figure 4 | **Main evolutionary mechanisms in yeasts and their functional and architectural consequences.** **a** | Genome duplication or hybridization between two isogenic yeasts creates cells with homozygous chromosomes. Subsequent gene loss results in extensive map reshuffling, and gene copy number gradually returns to the original pre-duplication status. The rare duplicated pairs that are maintained (ohnologues) often undergo asymmetrical evolution rates, with one copy acquiring a new or more specialized function. **b** | Crosses between two non-isogenic strains (distinct subpopulations) or hybridization between different species that have conserved a substantial degree of synteny creates diploids with heterozygous chromosomes. Subsequent loss of heterozygosity of large chromosomal segments often occurs. This phenomenon does not affect genomic maps (except locally at non-perfectly syntenic regions between parental yeasts (not shown)). **c** | Starting from a yeast with homozygous or heterozygous chromosomes, the acquisition of alien genes locally modifies the genomic map. The alien acquisition may concern a single gene (horizontal transfer, usually from bacteria) or a segment of chromosome from another yeast (introgression).

during the reverse transcription phase, such that a cDNA copy of a cellular mRNA is made and integrated into chromosomes. The general paucity of active retrotransposons in yeast genomes[84] suggests, however, that this mechanism plays only a limited part in their evolution except, possibly, during transient bursts of transposon activity. The scarcity of spliceosomal introns in many yeast species[85,86] also suggests that exon shuffling will be limited during yeast evolution.

Other mechanisms of genome evolution

Mutational rates and loss of synteny: two independent clocks. The extensive sequence divergence between orthologous genes in distinct yeast species is a characteristic feature of yeast evolution that is often overlooked in comparative genome annotations. Even species that belong to the same genus and share highly conserved gene synteny can exhibit large sequence divergence that indicates that they are not closely related. For example, *S. bayanus* proteins differ as much, on average, from their *S. cerevisiae* orthologues as human and mouse proteins do[87]. Similar observations were made in the *Lachancea*[44], *Clavispora*[88] and *Nakaseomyces* genera (Génolevures Consortium, personal communication). Between genera, sequence divergence rapidly jumps to even higher figures (*S. cerevisiae* and *C. glabrata* orthologous proteins differ as much as those of humans and fishes do).

In the absence of clear-cut fossils for yeasts, it is interesting to examine what such figures represent in terms of successive clonal generations made by each yeast lineage, based on the mutational rates experimentally determined in *S. cerevisiae*. Depending on the genes studied and their chromosomal environments, slightly variable figures were obtained, but the average is close to 3×10^{-10} mutations per nucleotide per mitotic generation[89]. Similar estimates were obtained by whole-genome resequencing of *S. cerevisiae* cultures[90]. In these experiments, indels were roughly ten times less frequent than nucleotide substitutions. A similar ratio is observed in comparisons between natural isolates of *S. cerevisiae* or *S. paradoxus*[14]. Assuming that independent and neutral mutations occur at the above rate over successive generations, the proportion of non-mutated nucleotides follows an exponential decrease, such that 50% of the nucleotides will have been mutated at least once after ~2 billion successive clonal generations which, for yeasts, translates into about 2 million years (or less with generation rates as stated by REF. 91). Although this figure is based on improbable assumptions of neutrality and unlimited population size, it gives a minimal time limit to the observed sequence divergences within and between yeast clades. With this view, the degrees of intraspecific sequence polymorphism observed between isolates of *S. cerevisiae* or *S. paradoxus* (~0.1–0.3%)[14,15,92] correspond to ~5 million successive generations (or a few thousand years). This suggests that these yeast species represent very recent clonal expansions from small samples of large previous populations that had undergone similar successive bottlenecks. Note that, at this mutational rate, traces of a neutral genetic element will disappear from yeast genomes in very short time periods, consistent with the limited number of pseudogenes observed[93].

Indels
Mutations due to the insertion or deletion of DNA sequences. In practice the term is often used to designate insertions or deletions that affect only one or a few nucleotides.

Pseudogenes
Genomic DNA sequences that are similar to normal genes but are rendered non-functional after mutations.

Compared with the rate of sequence divergence, yeast genetic maps seem to be much more robust. The distinct clades of protoploid Saccharomycetaceae share synteny blocks of ~20 genes on average in pairwise comparisons, indicating that their genomes have only experienced about 300 rearrangements from their common ancestor[44], whereas their orthologous proteins conserve only 55–60% sequence identity. The more limited synteny for most other pairwise comparisons between yeast genomes[44,48,88] only reminds us of the large evolutionary distances separating them. Within clades, synteny conservation can be very high. For example, the species of the *Saccharomyces sensu stricto* clade have nearly identical chromosomal maps interrupted by only a few chromosomal translocations[72,87]. The loss of synteny between distinct lineages that originate from a common ancestor results from a combination of several mechanisms. Gross chromosomal rearrangements, such as translocations or interchromosomal segmental duplications, break gene adjacency over long distances, but considerable accumulation of such events is needed before effects on local gene adjacency levels become measurable. In addition, preferred breakage points seem to exist, such as at dispersed tRNA gene copies, which are often encountered at extremities of conserved synteny blocks. By contrast, in large conserved synteny blocks, local gene adjacency disruptions are frequent, mostly as a result of small inversions[94,95] or single-gene deletions or insertions[96]. Finally, selective pressures help to conserve genome maps, resulting in some degree of clustering of essential genes, a phenomenon that is linked to recombination hot spots[97].

Loss of heterozygosity. Interestingly, in addition to mating-type switching (BOX 3), yeasts have evolved another mechanism that reduces heterozygosity in diploid cells or hybrids (FIG. 4). The asexual diploid yeast *C. albicans* is an interesting model in which homologous chromosome pairs show a mosaic of heterozygous and homozygous regions[88]. In this case, sequence divergence between alleles is erased by a general chromosomal mechanism, LOH. By comparing commensal isolates, most LOH events were found to encompass entire chromosomes or large regions extending to telomeres, as predicted by a BIR mechanism[98]. The related yeast *C. dubliniensis* also has three homozygous chromosomes, and distinct chromosomal bands for other chromosomes correspond to the rest of the *C. dubliniensis* genome[99]. To what extent LOH applies to other yeast evolutionary lineages is presently unclear, although genes involved in LOH have been identified in *S. cerevisiae*[100]. In *C. albicans*, long-range LOH is increased (relative to *in vitro* cultivation) by passage though a mammalian host[101].

Introgression. The presence in yeast genome sequences of large chromosomal segments that are identical or nearly identical in sequence to other species is puzzling. The genomes of wine strains of *S. cerevisiae* contain DNA fragments from *S. paradoxus*, *S. kudriavzevii*, *Saccharomyces uvarum* and even the distantly related *Zygosaccharomyces bailii*[33,102–105], suggesting that very recent introgressions have occurred (as the sequences are nearly identical to those in the donor genomes). Introgression was also reported between varieties of the Basidiomycota yeast *C. neoformans*[106], indicating that this is a general phenomenon in yeast genomes, although its importance for evolution has not yet been determined. The mechanism of introgression remains unknown. Classically in plants, introgression is obtained by hybridization followed by successive backcrossing. It is possible, although not likely, that the same process applies in yeasts. Alternatively, delayed karyogamy in heterospecific hybrids (BOX 3) may allow the transfer of chromosomal fragments from one nucleus to the other. In crosses involving *S. cerevisiae* mutants delayed in karyogamy, haploid segregants can inherit intranuclear autonomous elements of the other parent. The formation of circular chromosomal fragments by segmental duplication mechanisms (FIG. 3) may provide the source of DNA material transferred to the other nucleus. The ecological proximity and selective pressures to adapt to high-sugar, low-nitrogen and high-ethanol conditions during fermentations may facilitate the phenomenon, explaining the frequent introgressions observed in industrial *S. cerevisiae* strains.

Horizontal gene transfer from bacterial origin. The acquisition of genes from bacteria has long been regarded as a marginal phenomenon in yeasts. However, various cases have been reported, and some of these have contributed to important functional innovations or reacquisitions of genes. A notable example is provided by the acquisition of a bacterial gene encoding a dihydroorotate dehydrogenase, possibly originating from a *Lactococcus*, in the ancestor of the Saccharomycetaceae family followed by its vertical transmission to many extant members of this family to form the *URA1* gene[107,108]. This gene either cohabits with the ancestral eukaryotic enzyme encoded by the *URA9* gene (in protoploids) or has replaced it (in *Saccharomyces sensu stricto* species). Interestingly, the cytosolic bacterial enzyme allows the synthesis of uracil in the absence of oxygen, whereas the mitochondrial ancestral eukaryotic one does not, a property that could have allowed the emergence of facultative anaerobe yeasts. Several other examples of acquisition or reacquisition of metabolic functions in various yeast lineages have been reported[109–112]. Furthermore, a recent systematic analysis of inserted genes in conserved synteny blocks[96] and a systematic computer analysis using phylogenomic criteria[113] suggest that the phenomenon is not rare in yeasts. Although the mechanism of transfer has not been elucidated, it is probably facilitated by the fact that most yeast species are terrestrial saprobes living in close association with other organisms. Interestingly, transferred genes have a tendency to duplicate in their new yeast host[96].

De novo *gene formation and acquisition of other alien sequences.* Every sequenced yeast genome has revealed a small but significant number of specific genes without clear-cut homologues in other species, including the

Commensal
Refers to an organism living at the expense of another one without causing substantial damage (a case of non-deleterious parasitism).

Introgression
The incorporation of genes of one species into the genetic pool of another. It is classically viewed as the result of hybridization followed by backcrossing, but could result from other mechanisms in yeasts.

Karyogamy
The fusion of distinct nuclei present in the same cell.

most closely related ones. These genes are generally of unknown function and unclear origin. If these genes have not been acquired by horizontal transfer, they may be examples of *de novo* gene formation. In complex eukaryotes, such genetic innovations often result from exonization of mobile sequences or exon shuffling. However, the dearth of spliceosomal introns in Saccharomycotina — despite their intron-rich ancestors[86] — and the inconstant presence of active class I transposons[84] predict that exonization and exon shuffling should have a more limited role in yeast evolution. Nevertheless, these two mechanisms may partly explain the fusions and fissions of protein domains that have been observed in yeast genomes[114]. The *de novo* formation of protein-coding genes by mutational sequence changes is also possible in yeasts. In *S. cerevisiae*, a short protein-coding gene, *BSC4*, that is involved in DNA repair during stationary phase seems to have emerged from a transcribed but non-coding sequence that is present in related species[115]. Another example is provided by the *S. cerevisiae*-specific gene *MDF1*, which seems to have emerged from translation of the antisense transcript from the ancestral *ADF1* gene, which is conserved across yeast species[116]. The translation product of *MDF1* suppresses mating efficiency of *S. cerevisiae* by binding to the MATα2 protein, hence promoting vegetative growth. The integration of fragments of mitochondrial DNA (NUMTs) into nuclear chromosomes is also a strong mutagenic force that could, theoretically, create novel gene fusions; for example, 5′ extension of a protein-coding gene was recently reported in *Debaryomyces hansenii*[117]. The process of transfer of mitochondrial DNA to the nucleus is not entirely understood, except that it must involve the fragmentation of mitochondrial DNA before integration at double-strand breaks of chromosomes, as judged from the frequent occurrence of mosaics in both natural and experimental cases[117,118]. Similarly, the presence of nuclear sequences of plasmid and viral origin (NUPAVs) was recognized in about 40% of the sequenced Saccharomycotina genomes[119]. Although most correspond to pseudogenes, some active protein-coding genes were shown to originate from the non-retroviral RNA viruses that are occasionally found in various yeast and fungal species[119,120].

Conclusions and perspective

Many lessons have been learnt from the recent progress of yeast evolutionary genomics, even if many lineages still remain to be explored given the broad evolutionary diversity and multiple phylogenetic origins of yeasts among fungi. New sequencing technologies mean that investigations are no longer limited to model systems or the important biotechnological or infectious agents on which most of yeast evolutionary genomics presently relies. Sequencing projects specifically directed at biodiversity studies are in sight. Natural as well as artificially evolved populations are becoming amenable to complete genome analyses by rapid resequencing to discover mutational changes, copy-number variations and other genome alterations[57,90,121]. Integration of these results with comparative genomic data should rapidly increase our understanding of eukaryotic genome evolution processes, using yeasts as models. But more data will probably be needed to understand the complexity of the genetic determinism of the various functional adaptations of yeasts. The sequence divergence between extant yeast lineages is generally so extensive that it suggests that a considerable drift has been superimposed onto gradual Darwinian adaptations during yeast evolution. This view is consistent with the rapid propagation of essentially clonal unicellular populations undergoing repeated bottleneck events that facilitate the emergence of not-necessarily-optimized variants[122]. The recurrent loss of sex and the variety of mechanisms used to limit actual genetic exchanges suggest that this mode of propagation offers advantages to ensure the long-term evolutionary success of yeasts.

With this view of yeast evolution, two major questions deserve further studies. The first concerns the multiple emergences of various yeast forms from more complex fungal lineages (FIG. 1). More data are needed to address this question. At present, it looks likely that these events correspond to non-reversible regressive changes to fulfil conditions under which rapid clonal expansions become more important than biological complexity. It will be interesting to examine the nature and the diversity of such changes. The second relates to the very long evolutionary conservation of major yeast lineages, such as the Saccharomycotina, over several hundreds of billions of successive generations. By comparison, the diversification of animal lineages is several orders of magnitude shorter in terms of generations (even at the cellular level). The currently known rates of mutations, gene duplications and losses and horizontal acquisitions suggest that Saccharomycotina genomes have experienced a much larger range of mutational changes since their origin than have Metazoa genomes, as if they were trapped in a sort of evolutionary deadlock in which changes constantly occur in the genomes without significant innovation. One possibility is that the loss of spliceosomal introns in these yeasts[86] reached a threshold beyond which the role of RNA-mediated mechanisms for novel gene formation (which seem to have important roles in multicellular eukaryote evolution) becomes insufficient to compensate for the losses. The lack of traces of RNAi machinery in genomes of Saccharomycotina, except in specific cases[123], also suggests that RNA-mediated mechanisms have a limited role in yeasts. In addition, the frequent loss of active transposons in yeast lineages, which probably correlates with the loss or reduction of their sexual exchanges, is expected to accentuate the phenomenon. The alteration of the genetic code in an entire yeast lineage[124], the successive changes of tRNA sparing rules[125], and other unexpected features of non-coding RNAs suggest that additional investigations of the exact role of RNA-mediated mechanisms in yeast evolution may be promising. But whatever lies ahead for yeast evolutionary genomics, it is already clear that studying yeast genomes is a rewarding route towards elucidating novel general principles of evolution.

1. Kurtzman, C. P., Fell, J. W. & Boekhout, T. *The Yeasts: A Taxonomic Study* 5th edn (Elsevier, Amsterdam) (in the press).
2. Hughes, A. L. & Friedman, R. Parallel evolution by gene duplication in the genomes of two unicellular fungi. *Genome Res.* **13**, 794–799 (2003).
3. Liu, Y. J. & Hall, B. D. Body plan evolution of ascomycetes, as inferred from an RNA polymerase II phylogeny. *Proc. Natl Acad. Sci. USA* **101**, 4507–4512 (2004).
4. Keeling, P. J. *et al.*, The tree of eukaryotes. *Trends Ecol. Evol.* **20**, 670–676 (2005).
5. Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 563–567 (1996).
The first sequencing of a eukaryotic genome, that of *Saccharomyces cerevisiae*.
6. Pena-Castillo, L. & Hughes, T. R. Why are there still over 1000 uncharacterized yeast genes? *Genetics* **176**, 7–14 (2007).
7. Wood, V. *et al.* The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**, 871–880 (2002).
8. Johnson, L. J. *et al.* Population genetics of the wild yeast *Saccharomyces paradoxus*. *Genetics* **166**, 43–52 (2004).
9. Aa, E., Townsend, J. P., Adams, R. I., Nielsen, K. M. & Taylor, J. W. Population structure and gene evolution in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **6**, 702–715 (2006).
10. Ruderfer, D. M., Pratt, S. C., Seidel, H. S. & Kruglyak, L. Population genomic analysis of outcrossing and recombination in yeast. *Nature Genet.* **38**, 1077–1081 (2006).
11. Tsai, I. J., Bensasson, D., Burt, A. & Koufopanou, V. Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle. *Proc. Natl Acad. Sci. USA* **105**, 4957–4962 (2008).
12. Koufopanou, V., Hughes, J., Bell, G. & Burt, A. The spatial scale of genetic differentiation in a model organism: the wild yeast *Saccharomyces paradoxus*. *Philos. Trans. R. Soc. Lond. B* **361**, 1941–1946 (2006).
13. Kuehne, H. A., Murphy, H. A., Francis, C. A. & Sniegowski, P. D. Allopatric divergence, secondary contact, and genetic isolation in wild yeast populations. *Curr. Biol.* **17**, 407–411 (2007).
14. Liti, G. *et al.* Population genomics of domestic and wild yeasts. *Nature* **458**, 337–341 (2009).
The most comprehensive recent analysis of the genomic diversity of numerous isolates of *S. cerevisiae* and *S. paradoxus*, highlighting population structures of these two yeasts.
15. Schacherer, J., Shapiro, J. A., Ruderfer, D. M. & Kruglyak, L. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* **458**, 342–346 (2009).
16. Hittinger, C. T. *et al.* Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* **464**, 54–60 (2010).
The discovery of a stably maintained multilocus polymorphism among populations of *S. kudriavzevii*.
17. Marinoni, G. *et al.* Horizontal transfer of genetic material among *Saccharomyces* yeasts. *J. Bacteriol.* **181**, 6488–6496 (1999).
18. Dunn, B. & Sherlock, G. Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Gene Res.* **18**, 1610–1623 (2008).
19. Nakao, Y. *et al.* Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res.* **16**, 115–129 (2009).
20. Gonzalez, S., Barrio, E. & Querol, A. Molecular characterization of new natural hybrids of *Saccharomyces cerevisiae* and *S. kudriavzevii* in brewing. *Appl. Environ. Microbiol.* **74**, 2314–2320 (2008).
21. Belloch, C. *et al.* Chimeric genomes of natural hybrids between *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii*. *Appl. Environ. Microbiol.* **75**, 2534–2544 (2009).
22. Replansky, T., Koufopanou, V., Greig, D. & Bell, G. *Saccharomyces sensu stricto* as a model system for evolution and ecology. *Trends Ecol. Evol.* **23**, 494–501 (2008).
23. Pujol, C. *et al.* The closely related species *Candida albicans* and *Candida dubliniensis* can mate. *Eukaryot. Cell* **3**, 1015–1027 (2004).
24. Bovers, M. *et al.* Unique hybrids between the fungal pathogens *Cryptococcus neoformans* and *Cryptococcus gattii*. *FEMS Yeast Res.* **6**, 599–607 (2006).
25. Gordon, J. L. & Wolfe, K. H. Recent allopolyploid origin of *Zygosaccharomyces rouxii* strain ATCC 42981. *Yeast* **25**, 449–456 (2008).
26. Greig, D., Louis, E. J., Borts, R. H. & Travisano, M. Hybrid speciation in experimental populations of yeast. *Science* **298**, 1773–1775 (2002).
27. Usher, J. & Bond, U. Recombination between homoeologous chromosomes of lager yeasts leads to loss of function of the hybrid *GPH1* gene. *Appl. Environ. Microbiol.* **75**, 4573–4579 (2009).
28. Payen, C. *et al.* Unusual composition of a yeast chromosome arm is associated with its delayed replication. *Genome Res.* **19**, 1710–1721 (2009).
29. Wu, Q., James, S. A., Roberts, I. N., Moulton, V. & Huber, K. T. Exploring contradictory phylogenetic relationships in yeasts. *FEMS Yeast Res.* **8**, 641–650 (2008).
30. Naunov, G. I., James, S. A., Naumova, E. S., Louis, E. J. & Roberts, I. N. Three new species in the *Saccharomyces sensu stricto* complex: *Saccharomyces cariocanus*, *Saccharomyces kudriavzevii* and *Saccharomyces mikatae*. *Int. J. Syst. Evol. Micobiol.* **50**, 1931–1942 (2000).
31. Chambers, S. R., Hunter, N., Louis, E. J. & Borts, R. H. The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. *Mol. Cell. Biol.* **16**, 6110–6120 (1996).
32. Delneri, D. *et al.* Engineering evolution to study speciation in yeasts. *Nature* **422**, 68–72 (2003).
33. Liti, G., Barton, D. B. H. & Louis, E. J. Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. *Genetics* **174**, 839–850 (2006).
34. Fischer, G., James, S. A., Roberts, I. N., Oliver, S. G. & Louis, E. J. Chromosomal evolution in *Saccharomyces*. *Nature* **405**, 415–454 (2000).
35. Lee, H.-Y. *et al.* Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. *Cell* **135**, 1065–1073 (2008).
36. Scannell, D. R. *et al.* Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* **440**, 341–345 (2006).
A comprehensive analysis of post-duplication consequences in terms of speciation.
37. Lynch, M. & Force, A. G. The origin of interspecific genomic incompatibility via gene duplication. *Am. Nat.* **156**, 590–605 (2000).
38. Sampaio, J. P. & Gonçalves, P. Natural populations of *Saccharomyces kudriavzevii* in Portugal are associated with oak bark and are sympatric with *S. cerevisiae* and *S. paradoxus*. *Appl. Environ. Microbiol.* **74**, 2144–2152 (2008).
39. Maclean, C. J. & Greig, D. Prezygotic reproductive isolation between *Saccharomyces cerevisiae* and *Saccharomyces paradoxus*. *BMC Evol. Biol.* **8**,1 (2008).
40. Murphy, H., Kuehne, H., Francis, C. & Sniegowski, P. Mate choice assays and mating propensity differences in natural yeast populations. *Biol. Lett.* **2**, 553–556 (2006).
41. Ohno, S. *Evolution by Gene Duplication* (Springer, New York, 1970).
42. Conant, G. C. & Wolfe, K. H. Turning a hobby into a job: how duplicated genes find new functions. *Nature Rev. Genet.* **9**, 938–950 (2008).
43. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nature Rev. Genet.* **11**, 97–108 (2010).
44. Souciet, J.-L. *et al.* Comparative genomics of protoploid Saccharomycetaceae. *Genome Res.* **19**, 1696–1709 (2009).
The basic protein repertoire, distribution of paralogues and conservation of synteny in Saccharomycetaceae yeasts.
45. Despons, L. *et al.* Genome-wide computational prediction of tandem gene arrays: application in yeasts. *BMC Genomics* **11**, 56 (2010).
46. Fogel, S. & Welch, J. W. Tandem gene amplification mediates copper resistance in yeast. *Proc. Natl Acad. Sci. USA* **79**, 5342–5346 (1982).
47. Johnston, M. *et al.* Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. *Science* **265**, 2077–2082 (1994).
48. Dujon, B. *et al.* Genome evolution in yeasts. *Nature* **430**, 35–44 (2004).
The first multispecies genomic comparison across the entire Saccharomycotina subphylum.
49. Kaur, R., Ma, B. & Cormack, B. P. A family of glycosylphosphatidylinositol-linked aspartyl proteases is required for virulence of *Candida glabrata*. *Proc. Natl Acad. Sci. USA* **104**, 7628–7633 (2007).
50. Müller, H. *et al.* Genomic polymorphism in the population of *Candida glabrata:* gene copy number variation and chromosomal translocations. *Fungal Genet. Biol.* **46**, 264–276 (2009).
51. Despons, L., Wirth, B., Louis, V. L., Potier, S. & Souciet, J.-L. An evolutionary scenario for one of the largest yeast gene families. *Trends Genet.* **22**, 10–15 (2006).
52. Marques-Bonet, T., Girirajan, S. & Eichler, E. E. The origins and impact of primate segmental duplications. *Trends Genet.* **25**, 443–454 (2009).
53. Fairhead, C. & Dujon, B. Structure of *Kluyveromyces lactis* subtelomeres: duplication and gene content. *FEMS Yeast Res.* **6**, 428–441 (2006).
54. Koszul, R. & Fischer, G. A prominent role for segmental duplications in modeling eukaryotic genomes. *C. R. Biol.* **332**, 254–266 (2009).
55. Koszul R., Caburet S., Dujon B. & Fischer G. Eukaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J.* **23**, 234–243 (2004).
56. Schacherer, J., Tourrette, Y., Potier, S., Souciet, J.-L. & de Montigny J. Spontaneous duplications in diploid *Saccharomyces cerevisiae* cells. *DNA Repair (Amst.)* **6**, 1441–1452 (2007).
57. Gresham, D. *et al.* The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* **4**, e1000303 (2008).
A genome-wide analysis of evolutionary adaptations in experimental cultures of *S. cerevisiae*, based on genome resequencing.
58. Koszul, R., Dujon, B. & Fischer, G. Stability of large segmental duplications in the yeast genome. *Genetics* **172**, 2211–2222 (2006).
59. Libuda, D. E. & Winston, F. Amplification of histone genes by circular chromosome formation in *Saccharomyces cerevisiae*. *Nature* **443**, 1003–1007 (2006).
60. Payen, C., Koszul, R., Dujon, B. & Fischer, G. Segmental duplications arise from pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet.* **4**, e1000175 (2008).
The discovery of the molecular mechanisms responsible for spontaneous segmental duplications.
61. Wolfe, K. H. & Shields, D. C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (1997).
The first hypothesis of whole-genome duplication in *S. cerevisiae*.
62. Dietrich, F. S. *et al.* The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science* **304**, 304–307 (2004).
63. Kellis, M., Birren, B. W. & Lander, E. S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617–624 (2004).
64. Ma, L.-J. *et al.* Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. *PLoS Genet.* **5**, e1000549 (2009).
65. Presser, A., Elowitz, M. B., Kellis, M. & Kishony, R. The evolutionary dynamics of the *Saccharomyces cerevisiae* protein interaction network after duplication. *Proc. Natl Acad. Sci. USA* **105**, 950–954 (2008).
66. Vinogradov, A. E. & Anatskaya, O. V. Loss of protein interactions and regulatory divergence in yeast whole-genome duplicates. *Genomics* **93**, 534–542 (2009).
67. Byrne, K. P. & Wolfe, K. H. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* **15**, 1456–1461 (2005).
68. Cliften, P. F., Fulton, R. S., Wilson, R. K. & Johnston, M. After the duplication: gene loss and adaptation in *Saccharomyces* genomes. *Genetics* **172**, 863–872 (2006).
69. Scannell, D. R. *et al.* Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. *Proc. Natl Acad. Sci. USA* **104**, 8397–8402 (2007).
70. Gordon, J. L., Byrne, K. P. & Wolfe, K. H. Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. *PLoS Genet.* **5**, e1000485 (2009).

71. Martin, N., Ruedi, E. A., Leduc, R., Sun, F. J. & Caetano-Anolles, G. Gene-interleaving patterns of synteny in the *Saccharomyces cerevisiae* genome: are they proof of an ancient genome duplication event? *Biol. Direct* **2**, 23 (2007).
72. Fischer, G., Neuvéglise, C., Durrens, P., Gaillardin, C. & Dujon, B. Evolution of gene order in the genomes of two related yeast species. *Genome Res.* **11**, 2009–2019 (2001).
73. Lafontaine, I., Fischer, G., Talla, E. & Dujon, B. Gene relics in the genome of the yeast *Saccharomyces cerevisiae*. *Gene* **335**, 1–17 (2004).
74. Byrnes, J. K., Morris, G. P. & Li, W. H. Reorganization of adjacent gene relationships in yeast genomes by whole-genome duplication and gene deletion. *Mol. Biol. Evol.* **23**, 1136–1143 (2006).
75. Scannell, D. R. & Wolfe, K. H. A burst of protein sequence evolution and a prolonged period of asymmetric evolution follow gene duplication in yeast. *Genome Res.* **18**, 137–147 (2008).
76. Kim, S.-H. & Yi, S. V. Correlated asymmetry of sequence and functional divergence between duplicate proteins of *Saccharomyces cerevisiae*. *Mol. Biol. Evol.* **23**, 1068–1075 (2006).
77. Byrne, K. P. & Wolfe, K. H. Consistent patterns of rate asymmetry and gene loss indicate widespread neofunctionalization of yeast genes after whole-genome duplication. *Genetics* **175**, 1341–1350 (2007).
An analysis of the post-duplication divergence of genes and the functional consequences of duplication.
78. Papp, B., Pal, C. & Hurst, L. D. Evolution of *cis*-regulatory elements in duplicated genes of yeast. *Trends Genet.* **19**, 417–422 (2003).
79. Gu, X., Zhang, Z. & Huang, W. Rapid evolution of expression and regulatory divergences after yeast gene duplication. *Proc. Natl Acad. Sci. USA* **102**, 707–712 (2005).
80. Wapinski, I., Pfeffer, A., Friedman, N. & Regev, A. Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**, 54–64 (2007).
81. Marques, A. C., Vinckenbosch, N., Brawand, D. & Kaessmann, H. Functional diversification of duplicate genes through subcellular adaptation of encoded proteins. *Genome Biol.* **9**, R54 (2008).
82. Dean, E. J., Davis, J. C., Davis, R. W. & Petrov, D. A. Pervasive and persistent redundancy among duplicated genes in yeast. *PLoS Genet.* **4**, e1000113 (2008).
83. Schacherer, J., Tourette, Y., Souciet, J.-L., Potier, S. & De Montigny, J. Recovery of a function involving gene duplication by retroposition in *Saccharomyces cerevisiae*. *Genome Res.* **14**, 1291–1297 (2004).
A demonstration of the RNA-mediated mechanism of single-gene duplication at ectopic locations in yeast.
84. Neuvéglise, C., Feldman, H., Bon, E., Gaillardin, C. & Casarégola, S. Genomic evolution of the long terminal repeat retrotransposons in hemiascomycetous yeasts. *Genome Res.* **12**, 930–943 (2002).
85. Bon, E. *et al.* Molecular evolution of eukaryotic genomes: hemiascomycetous yeast spliceosomal introns. *Nucleic Acids Res.* **31**, 1121–1135 (2003).
86. Stajich, J. E., Dietrich, F. S. & Roy, S. W. Comparative genomic analysis of fungal genomes reveals intron-rich ancestors. *Genome Biol.* **8**, R223 (2007).
87. Dujon, B. Yeasts illustrate the molecular mechanisms of eukaryotic genome evolution. *Trends Genet.* **22**, 375–387 (2006).
88. Butler G. *et al.* Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature* **459**, 657–662 (2009).
The most comprehensive recent multispecies comparative analysis of yeasts of the CTG group.
89. Lang, G. L. & Murray, A. W. Estimating the per-base mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* **178**, 67–82 (2008).
90. Lynch, M. *et al.* A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl Acad. Sci. USA* **105**, 9272–9277 (2008).
The first genome-wide estimation of the mutational spectrum in experimental cultures of *S. cerevisiae* that have not been submitted to limiting nutritional conditions. The estimation was based on resequencing.
91. Fay, J. C. & Benavides, J. A. Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet.* **1**, 66–71 (2005).
92. Gresham, D. *et al.* Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. *Science* **311**, 1932–1936 (2006).
93. Lafontaine, I. & Dujon, B. Origin and fate of pseudogenes in hemiascomycetes: a comparative analysis. *BMC Genomics* **11**, 260 (2010).
94. Seoighe, C. *et al.* Prevalence of small inversions in yeast gene order evolution. *Proc. Natl Acad. Sci. USA* **97**, 14433–14437 (2000).
95. Fischer, G., Rocha, E. P., Brunet, F., Vergassola, M. & Dujon, B. Highly variable rates of genome rearrangements between hemiascomycetous yeast lineages. *PLoS Genet.* **2**, e32 (2006).
96. Rolland, T., Neuvéglise, C., Sacerdot, C. & Dujon, B. Insertion of horizontally transferred genes within conserved syntenic regions of yeast genomes. *PLoS ONE* **4**, e6515 (2009).
97. Keller, P. J. & Knop, M. Evolution of mutational robustness in the yeast genome: a link to essential genes and meiotic recombination hotspots. *PLoS Genet.* **5**, e1000533 (2009).
98. Diogo, D., Bouchier, C., d'Enfert, C. & Bougnoux, M.-E. Loss of heterozygosity in commensal isolates of the asexual diploid yeast *Candida albicans*. *Fungal Genet. Biol.* **46**, 159–168 (2009).
99. Jackson, A. P. *et al.* Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Res.* **19**, 2231–2244 (2009).
100. Andersen, M. P., Nelson, Z. W., Hetrick, E. D. & Gottschling, D. E. A genetic screen for increased loss of heterozygosity in *Saccharomyces cerevisiae*. *Genetics* **179**, 1179–1195 (2008).
101. Forche, A., Magee, P. T., Selmecki, A., Berman, J. & May, G. Evolution in *Candida albicans* populations during a single passage through a mouse host. *Genetics* **182**, 799–811 (2009).
102. Naumova, E. S., Naumov, G. I., Masneuf-Pomarède, I. & Aigle, M. Molecular genetic study of introgression between *Saccharomyces bayanus* and *S. cerevisiae*. *Yeast* **22**, 1099–1115 (2005).
103. Doniger, S. W. *et al.* A catalog of neutral and deleterious polymorphisms in yeast. *PLoS Genet.* **4**, e1000183 (2008).
104. Muller, L. A. H. & McCusker, J. H. A multispecies-based taxonomic microarray reveals interspecies hybridization and introgression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **9**, 143–152 (2009).
105. Novo, M. *et al.* Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proc. Natl Acad. Sci. USA* **106**, 16333–16338 (2009).
The most notable recent example of introgression in yeasts.
106. Kavanaugh, L. A., Fraser, J. A. & Dietrich, F. S. Recent evolution of the human pathogen *Cryptococcus neoformans* by intervarietal transfer of a 14-gene fragment. *Mol. Biol. Evol.* **23**, 1879–1890 (2006).
107. Gojdovic, Z. *et al.* Horizontal gene transfer promoted evolution of the ability to propagate under anaerobic conditions in yeasts. *Mol. Genet. Genomics* **271**, 387–393 (2004).
108. Hall, C., Brachat, S. & Dietrich, F. S. Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. *Eukaryotic Cell* **4**, 1102–1115 (2005).
109. Hall, C. & Dietrich, F. S. The reacquisition of biotin prototrophy in *Saccharomyces cerevisiae* involved horizontal gene transfer, gene duplication and gene clustering. *Genetics* **177**, 2293–2307 (2007).
110. Wei, W. *et al.* Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *Proc. Natl Acad. Sci. USA* **104**, 12825–12830 (2007).
111. Woolfit, M., Rozpedowska, E, Piskur, J. & Wolfe, K. H. Genome survey sequencing of the wine spoilage yeast *Dekkera (Brettanomyces) bruxellensis*. *Eukaryot. Cell* **6**, 721–733 (2007).
112. Fitzpatrick, D. A., Logue, M. E. & Butler, G. Evidence of recent interkingdom horizontal gene transfer between bacteria and *Candida parapsilosis*. *BMC Evol. Biol.* **8**, 181 (2008).
113. Marcet-Houben, M. & Gabaldón, T. Acquisition of prokaryotic genes by fungal genomes. *Trends Genet.* **26**, 5–8 (2010).
114. Durrens, P., Nikolski, M. & Sherman, D. Fusion and fission of genes define a metric between fungal genomes. *PLoS Comput. Biol.* **4**, e1000200 (2008).
115. Cai, J., Zhao, R., Jiang, H. & Wang, W. *De novo* origination of a new protein-coding gene in *Saccharomyces cerevisiae*. *Genetics* **179**, 487–496 (2008).
116. Li, D. *et al.* A *de novo* originated gene depresses budding yeast mating pathway and is repressed by the protein encoded by its antisense strand. *Cell Res.* **20**, 408–420 (2010).
117. Sacerdot, C. *et al.* Promiscuous DNA in the nuclear genomes of hemiascomycetous yeasts. *FEMS Yeast Res.* **8**, 846–857 (2008).
118. Ricchetti, M., Fairhead, C. & Dujon, B. Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* **402**, 96–100 (1999).
119. Frank, A. C. & Wolfe K. H. Evolutionary capture of viral and plasmid DNA by yeast nuclear chromosomes. *Eukaryot. Cell* **8**, 1521–1531 (2009).
120. Taylor, D. J. & Bruenn, J. The evolution of novel fungal genes from non-retroviral RNA viruses. *BMC Biol.* **7**, 88 (2009).
121. Araya, C. L., Payen, C., Dunham, M. J. & Fields, S. Whole-genome sequencing of a laboratory-evolved yeast strain. *BMC Genomics* **11**, 88 (2010).
122. Lynch, M. *The Origins of Genome Architecture* (Sinauer Associates, Sunderland, Massachusetts, 2007).
123. Drinnenberg, I. A. *et al.* RNAi in budding yeast. *Science* **326**, 544–550 (2009).
124. Massey, S. E. *et al.* Comparative evolutionary genomics unveils the molecular mechanism of reassignment of the CTG codon in *Candida* spp. *Genome Res.* **13**, 544–557 (2003).
125. Marck, C. *et al.* The RNA polymerase III-dependent family of genes in hemiascomycetes: comparative RNomics, decoding strategies, transcription and evolutionary implications. *Nucleic Acids Res.* **34**, 1816–1835 (2006).
126. Slaughter, B. D. Smith, S. E. & Li, R. Symmetry breaking in the life cycle of the budding yeast. *Cold Spring Harb. Perspect. Biol.* **1**, a003384 (2009).
127. Barral, Y. & Liakopoulos, D. Role of spindle asymmetry in cellular dynamics. *Int. Rev. Cell Mol. Biol.* **278**, 149–213 (2009).
128. Shcheprova, Z., Baldi, S., Frei, S. B., Gonnet, G. & Barral, Y. A mechanism for asymmetric segregation of age during yeast budding. *Nature* **454**, 728–734 (2008).
129. Paliwal, S. *et al.* MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature* **446**, 46–51 (2007).
130. Moore, T. I., Chou, C.-S., Nie, Q., Jeon, N. L. & Yi, T.-M. Robust spatial sensing of mating pheromone gradients by yeast cells. *PloS ONE* **3**, e3865 (2008).
131. Knop, M. Evolution of the hemiascomycete yeasts: on the life styles and the importance of inbreeding. *Bioessays* **28**, 696–708 (2006).
132. Albertin, W. *et al.* Evidence for autotetraploidy associated with reproductive isolation in *Saccharomyces cerevisiae*: towards a new domesticated species. *J. Evol. Biol.* **22**, 2157–2170 (2009).
133. Reedy, J. L., Floyd, A. M. & Heitman, J. Mechanistic plasticity of sexual reproduction and meiosis in the *Candida* pathogenic species complex. *Curr. Biol.* **19**, 891–899 (2009).
134. Fabre, E. *et al.* Comparative genomics in hemiascomycete yeasts: evolution of sex, silencing and subtelomeres. *Mol. Biol. Evol.* **22**, 856–873 (2005).
135. Barsoum, E., Martinez, P. & Aström, S. U. *Alpha3*, a transposable element that promotes host sexual reproduction. *Genes Dev.* **24**, 33–44 (2010).
136. Egel, R. Fission yeast mating-type switching: programmed damage and repair. *DNA Repair (Amst.)* **4**, 525–536 (2005).
137. Kurtzman, C. P. & Robnett, C. J. Phylogenetic relationships among yeasts of the '*Saccharomyces* complex' determined from multigene sequence analyses. *FEMS Yeast Res.* **3**, 417–432 (2003).
138. Fitzpatrick, D. A., Logue, M. E., Stajich, J. E. & Butler, G. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evol. Biol.* **6**, 99 (2006).
139. Tsui, C. K. M. Daniel, H.-M., Robert, V. & Meyer, W. Re-examining the phylogeny of clinically relevant *Candida* species and allied genera based on multigene analysis. *FEMS Yeast Res.* **8**, 651–659 (2008).
140. Hedges, S. B., Blair, J. E., Venturi, M. L. & Shoe, J. L. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol. Biol.* **4**, 2 (2004).

141. Taylor, J. W. & Berbee, M. L. Dating divergence in the fungal tree of life: review and new analyses. *Mycologia* **98**, 838–849 (2006).
142. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241–254 (2003).
143. Cliften, P. *et al.* Finding functional features in *Saccharomyces cerevisiae* by phylogenetic footprinting. *Science* **301**, 71–76 (2003).
144. Souciet, J.-L. *et al.* Genomic exploration of the hemiascomycetous yeasts *FEBS Lett.* **487**, 3–147 (2000).
145. Ramezani-Rad, M. *et al.* The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. *FEMS Yeast Res.* **4**, 207–215 (2003).
146. Jeffries, T. W. *et al.* Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nature Biotech.* **25**, 319–326 (2007).
147. Jones, T. *et al.* The diploid genome sequence of *Candida albicans*. *Proc. Natl Acad. Sci. USA* **101**, 7329–7334 (2004).
148. De Schutter K. *et al.* Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nature Biotech.* **27**, 561–566 (2009).
149. Mattanovich, D. *et al.* Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb. Cell Fact.* **8,** 29 (2009).
150. Loftus, B. J. *et al.* The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* **307**, 1321–1324 (2005).
151. Xu, J. *et al.* 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 (2007).

Acknowledgements
I thank the members of the Génolevures Consortium (GDR2354 Centre National de la Recherche Scientifique) for sharing unpublished data and for numerous stimulating discussions. B.D. is a member of Institut Universitaire de France.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION

Author's homepage: http://www.pasteur.fr/recherche/unites/Gmlev/en/accueil-en.html
Broad Institute Fungal Genome Initiative: http://www.broadinstitute.org/annotation
Génolevures — Genomic Exploration of the Hemiascomycete Yeasts: http://www.genolevures.org

ALL LINKS ARE ACTIVE IN THE ONLINE PDF