Going wild: ecology and genomics are crucial to understand yeast evolution
Duccio Cavalieri\textsuperscript{1,*}, Beatrice Valentini\textsuperscript{2} and Irene Stefanini\textsuperscript{2}

Improved and more accessible genome-sequencing approaches have allowed the analysis of large sets of natural yeast isolates. As a consequence, this unprecedented level of description of yeast-genome characteristics and variations in natural environments has provided crucial insights on yeast ecology and evolution. Here, we review some of the most relevant and intriguing aspects of yeast evolution pointed out, thanks to the combination of yeast ecology and genomics, and critically examine the resulting improvement of our knowledge on this field. Only integrated approaches, taking into consideration not only the characteristics of the microbe but also those of the hosting environment, will significantly move forward the exploration of yeast diversity, ecology, and evolution.

Addresses
\textsuperscript{1} Department of Biology, University of Florence, Florence, Italy
\textsuperscript{2} Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy

Corresponding author: Irene Stefanini (irene.stefanini@unito.it)
*Twitter account: @yeastCellCell

Current Opinion in Genetics & Development 2022, 75(1):101922
Edited by Christian Landry and Gianni Liti

This review comes from a themed issue on Evolutionary Genetics
For complete overview of the section, please refer to the article collection, “Evolutionary Genetics”
Available online 9th June 2022
https://doi.org/10.1016/j.gde.2022.101922
0959-437X/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction
Yeasts are almost ubiquitous and have been found in several biomes of the world and are associated with other organisms in various environmental systems [1–3]. Hence, the study of yeast ecology is relevant not only to disclose relevant information on these yeasts but also on the associated organisms and hosts. The increasing accessibility of whole-genome sequencing provides an unprecedented opportunity to assess and compare the genome characteristics of a previously inconceivably large number of yeast species. Over the last five years, the genomes of 427 Saccharomycotina taxa (more than 1700 genomes, corresponding to more than one-third of the species belonging to this subphylum [4]), have been deposited in Sequence Read Archive (accession on February 2022). Despite 826 of these new genome sequences belonging to Saccharomyces isolates (688 to S. cerevisiae) [5], the genomes available for other yeast genera are rapidly increasing (Figure 1). This imbalanced characterization can be ascribed not only to the fact that S. cerevisiae is among the most-studied microorganisms used for studying fundamental aspects of cell biology [6] or to the recently renewed interest in its ecology [7], but also to technical issues associated with the sequencing and analysis of other species’ genomes [8]. Ploidy variations, high heterozygosity, and lack of protocols for manipulating the genomes severely complicate the genomic analysis of most yeast species. However, the steadily improving set of bioinformatics tools available for genomic data processing is considerably helping to tackle these issues [8]. As a consequence, the last five years have been fruitful for the sequencing of ecologically and biotechnologically interesting nonconventional yeast genomes: Yarrowia, Torulaspora, Kazachstania, Hanseniaspora, Ogataea, Pichia, Brettanomyces, Clavispora, and Metschnikowia [9,10] (Figure 1).

Exploring new niches
A wealth of information on physiology, molecular mechanisms, and evolution of yeasts has been obtained through in vitro analyses (e.g. experimental evolution [11]) based on the assumption that a given condition is perceived as stressful by every strain belonging to the same species. Whereas it is likely true for most organisms, the wide intraspecific phenotypic and genetic variability of microorganisms suggests a different scenario [12,13]. A representative example is provided by the well-known variability in one of the traits representing one of yeasts’ most important acquisitions, sex. The production of gametes and transition to the haploid phase in S. cerevisiae requires the formation of spores and occurs when yeast cells are subjected to nitrogen starvation. The efficiency of this process varies significantly among strains [14], with some strains sporulating in the absence of this stress [15], a feature that has been associated with the possibility of improving purifying selection (further discussed in the “Genome renewal hypothesis updated” box). The recent assessment revealed a correlation between sporulation efficiency and the infection by endosymbiotic dsRNA...
viruses [16], previously found to have a role in the development of yeast social behaviors [17]. Furthermore, a recent study has provided fundamental information linking \textit{S. cerevisiae} domestication and sporulation: whereas wild \textit{S. cerevisiae} strains set off the meiosis and sporulate, domesticated yeasts enter the G0 quiescence stage, hindering sporulation [18]. Hence, domesticated yeasts have a reduced chance to cleanse their genomes from deleterious mutations through meiotic recombination and generate the broad diversity of genotypes observed in wild isolates. Overall, natural yeasts have highlighted that the importance of these dsRNA viruses in yeast ecology and evolution has yet to be completely appreciated.

Another bias related to experimental evolution is that these experiments are carried out in human-influenced environments, in which rapid adaptation is frequently observed [19]. Such settings poorly mimic the environmental fluctuations to which yeasts are exposed in nature. The rapid adaptation previously observed in \textit{Saccharomyces paradoxus} has been recently confuted by the analysis of strains subjected to monthly-scale selective biotic and abiotic environmental fluctuations in their native habitats [20]. In this perspective, the exploration of genetic and phenotypic variations among yeasts isolated from natural environments holds the promise to disclose previously neglected genomic features and evolutionary paths. Natural environments are sources of new diversity of already-known and previously unknown yeast species. Over the last two years, 76 new yeast species have been identified, only seven of which were isolated from human-related environments (Figure 2a). These 76 new species belong to eight yeast families (Debaryomycetaceae, Saccharomycetales incertae sedis, Metschnikowiaceae, Pichiaceae, Saccharomycetaceae, Phaffiomycetaceae, Trichomonascaceae, and Lipomycetaceae). Some new species belonging to a few yeast families are associated with specific isolation sources, suggesting a potential selection: the Lipomycetaceae family (14 species, of which 12 \textit{Lipomyces} spp. and 2 \textit{Babjevia} spp. [21]) in soil, \textit{Limtongella siamensis} in industry specimens (grease waste) [22], and \textit{Saccharomyces oxydans} in dairy products [23] (Chi-square test p. value < 0.05). Conversely, trees [24–29], plants and flowers [30–36], fruit, water [37], and insects [38–45] provided new species belonging to multiple families (Figure 2a). The situation slightly changes when
considering the newly sequenced genomes, which are largely represented by clinical or fermentation-related strains (Figure 2b), a bias possibly resulting from biotechnological interests. And yet, some natural sources have provided broad yeast biodiversity: soil, plants/flowers, and insects (Figure 2). Soil is a well-known source of yeast biodiversity, with the composition of soil yeast populations being more resilient than in other natural sources [46], but differing among geographical locations and impacting other organisms (protists, bacteria, insects, plants, and nematodes) by providing nutrients [47]. The yeast–plant association, despite being less acknowledged than the one with the soil, provides useful insights into yeast ecology and evolution [48]. Recent analyses have highlighted species-specific associations between oaks and Saccharomyces species: whereas S. paradoxus, S. jurei, and S. kudriavzevii are selectively associated with Quercus spp., S. cerevisiae, S. uvarum, and S. eubayanus are broadly found on Fagales trees [49]. The investigation of yeast-nectar associations has provided controversial insights on the impact of these associations on pollinators [50,51]. The known role of yeast species-specific metabolites in promoting the onset of cactophilic Drosophila niche partitioning [52] can help resolve this dispute: the key is considering the specificity of the association. Moreover, yeasts–insect associations have also provided insightful information on yeast spread and evolution [30,53]. Several yeast species are hosted by various insect species, suggesting the potential of insects as reservoirs and vectors for a
wide range of yeast diversity [54,55] of biotechnological interest such as the xylose-metabolizing yeasts isolated from wood-inhabiting insects [44,56–58]. Furthermore, the association between \textit{S. cerevisiae} and social wasps (\textit{Vespa crabro}, \textit{Polistes} spp., and \textit{Vespula} spp.) has proven to be crucial for both the maintenance of \textit{S. cerevisiae} in wild environments across seasons and years [59] and the generation of new yeast genetic setups, as the insect intestine promotes \textit{Saccharomyces} spp. intraspecific and interspecific mating [60].

**The ecology fingerprints on genomes**

As previously mentioned, laboratory analyses aimed at the dissection of evolutionary and ecological paths are based on the modeling of natural environmental settings [61]. The modeling is needed for technical reasons, but the simplification of environmental characteristics hinders the observation of phenomena occurring in natural environments and rarely takes into account an important feature of natural systems: they are unstable. Experimental evolution assesses the effects of changes in one or a few environmental parameters, whereas multiple parameters simultaneously and repeatedly change in natural settings. Neglecting these fluctuations may result in severely biased hypotheses on ecology and evolution. Once again, the model \textit{S. cerevisiae} provides useful insights. Hybridization generates genomics and phenotypic settings providing an ecological advantage compared with the parental lineages [62] at the expense of offspring survival, which is impaired by lethal combinations of the parental genomes. Return-To-Growth (RTG) is the recently described process promoting the increase of hybrids’ fertility through extensive loss-of-heterozygosity [63,64] mediated by recombination among not highly divergent chromosomal regions. In the lab, RTG is induced by subjecting meiotically induced diploid cells to an environment rich in nutrients before their commitment to complete meiosis. Different compartments of animal intestines are characterized by different chemical–physical features and nutrient availability, and yeast cells passing through them may be subjected to the environmental fluctuations promoting RTG, which may lay at the basis of the increased mating rate observed within the insect intestines [60]. However, since \textit{S. cerevisiae} strains isolated from human intestines, showing low heterozygosity, are unlikely the result of hybridization or mating [65], the environmental variations imposed by the animal intestine are not sufficient to promote yeast hybridization. A better characterization of these natural environments will be fundamental to understanding how yeasts can bypass the reproductive barriers leading to speciation [66].

Other natural environments have provided crucial insights on yeast evolution, such as those obtained on the \textit{Wickerhamiella} and \textit{Starmerella} genera. Over their evolution, species belonging to these genera have faced loss and reinstatement of alcoholic fermentation mediated by horizontal gene transfer and are characterized by fructophily, the unusual preference for fructose over glucose as the carbon source. In-depth genomics analyses on natural strains have led to the identification of a new pathway clarifying the mechanisms underlying fructophily [67]. From the ecological viewpoint, \textit{Wickerhamiella} and \textit{Starmerella} yeasts may take advantage of this pathway to outgrow other microorganisms competing for the same substrate in natural environments, as most \textit{S. cerevisiae} strains do through the Crabtree effect [68]. Conversely, other yeast species such as \textit{Yarrowia lipolytica} have maintained variegate metabolic functions allowing them, rather than improving their competitiveness in a limited set of substrates, to thrive on a broader range of environments [69]. For instance, genomics features of \textit{Yarrowia} spp. symbionts associated with the beetle \textit{Nicrophorus vespilloides} suggest that these yeasts can utilize microbial metabolites associated with putrefaction and grow in environments with extreme pH and salt concentrations, both characteristics present in the host cells [70].

**Conclusions**

Yeast ecology has enormously benefited from the progress in genome sequencing. The vast set of genome sequences available for strains and species isolated from natural environments has highlighted evolutionary and ecological aspects that could not be predicted in the laboratory. However, our knowledge on most yeast species is still limited by the scarce characterization of natural environments and of potential reciprocal influences between yeasts and the hosting environment. Further investigations should be aimed at addressing these issues to fully understanding the ecological systems in which yeasts can live and evolve. Aiming at this, interdisciplinary teams, combining expertise in yeast physiology and genetics but also in the environmental system under investigation, will have higher chances of fully depicting the interlaced relationship between yeasts and their hosting environment.

**Box: genome-renewal hypothesis updated**

The study of phenotypic, genetic, and genomic variation in \textit{S. cerevisiae} builds on the profound knowledge acquired from a single strain, S288c, that served as a model for yeast genetics. When working in Polsinelli’s laboratory, RK Mortimer, Beadle medal for genetics and father of S288c, realized that, in contrast with the lab strain S288c, many yeast strains isolated from wine and grapes can form gametes on rich media [15]. The vast majority (90%) of grape, wine, and bark strains are homothallic, thus capable of undergoing mating-type switching followed by mother–daughter mating. This process makes the diploid completely homozygous, preventing the
propagation of deleterious or lethal alleles [15]. The proposed evolutionary scenario leading to this situation, called ‘Genome Renewal’ by Mortimer and supported by mathematical models [71], involves long periods of clonal reproduction, in which diploid strains accumulate recessive heterozygous alleles, followed by autodiploidization that exposes these alleles to selection.

Nowadays, the genome-renewal hypothesis appears in contradiction with \textit{S. cerevisiae} environmental isolates harboring abundant polymorphisms [72,73] and frequent hybridization observed over genome sequencing of thousands of yeast isolates [74]. These discrepancies could be explained by divergent evolutionary paths in different ecological environments and from the extreme rarity of the haploid–meiotic cycle, that so far has been demonstrated to occur in a limited number of environments, mainly the gut of insects [65]. Disentangling the contradiction between the proven genome-renewal hypothesis and the high heterozygosity observed in natural isolates will represent an exciting opportunity for future research aimed at the understanding of \textit{S. cerevisiae} ecology and evolution.

**Conflict of interest statement**

The authors declare no conflict of interest.

**Acknowledgements**

This work was supported by the Human Frontier Science Program HFSP Research Grant RGP0060/2021.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


With this mini-review, the authors provide useful insights and hints on the tools and protocols suitable for yeast genomics analyses and also report a quite broad overview of the major features characterizing \textit{Saccharomyces} spp. and hybrids genomes.


By tracking the sexual and asexual reproduction capability of 1011 \textit{Saccharomyces cerevisiae} strains, the authors observed a neat differentiation between domesticated and wild strains: domesticated strains show impaired or absent sexual life cycle, together with an enhanced fermentation capability and a reduced respiratory asexual growth. Thanks to the availability of the genomic sequences of the studied strains and to the application of genome-wide association analysis, the authors have disclosed the quantitative effects of the significant differences between domesticated and wild strains.


To investigate the potential for monthly-scale adaptation of \textit{Saccharomyces paradoxus}, the authors measured the population structure and killer-related phenotypes (production and susceptibility) of \textit{S. paradoxus} strains isolated from forest specimens collected over four seasons. This study provides the first evidence rejecting the previously hypothesized rapid adaptation of \textit{S. paradoxus} when considering native environmental settings.

6 Evolutionary Genetics


The authors provide here an accurate depiction of the current knowledge of biodiversity, ecology, and population structure of the *Saccharomyces* species, with a particular focus on the ecology and biogeography of some of the most frequently observed inter-specific hybrids. Russell AL, Ashman T-L: Associative learning of flowers by generalist bumble bees can be mediated by microbes on the petals. *Behav Ecol* 2019, 30:746-755.


Through a meta-analysis, the authors summarize here the most known and well-documented yeast-insects associations. The results highlight the extent of this association, both in terms of geographical distribution and species involved. Also, the authors have collected and cured YeastFinder, a cured online database providing a collection of evidence of the isolation of yeast species from insects.


64. By observing the segregation of selectable genetic markers introduced to both arms of all S. cerevisiae chromosomes, the authors evaluated whether loss-of-heterozygosity (LoS) occurs randomly across the genome or forms discerning patterns. The results clearly show that LoS preferentially occurs through recombination among highly similar sites, in short rather than long chromosomal arms and regions farther from centromeres and closer or adjacent to telomeres.


Using a combination of classical microbiology and Next Generation Sequencing NGS techniques on a large set of Saccharomyces strains, the authors pinpoint the mechanism of RTG clearly showing that it occurs in regions with low local heterozygosity and near meiotic recombination hotspots. This is the first experimental hint on a path allowing some natural sterile yeast hybrids to overcome the post-zygotic reproductive barriers underlying speciation.


