Yeast. 2016 May 4. doi: 10.1002/yea.3171. [Epub ahead of print]

Yeast culture collections in the twenty-first century: New opportunities and challenges

Kyria L. Boundy-Mills[1], Ewald Glantschnig[2], Ian N. Roberts[3], Andrey Yurkov[4], Serge Casaregola[5], Heide-Marie Daniel[6], Marizeth Groenewald[7] and Benedetta Turchetti[8*]

All the authors contributed equally to the manuscript

[1]Phaff Yeast Culture Collection, Food Science and Technology, University of California Davis, One Shields Ave, Davis, CA, USA

[2]World Intellectual Property Organization, 34, chemin des Colombettes, CH-1211 Geneva 20

[3]National Collection of Yeast Cultures, Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, UK

[4]Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Inhoffenstrasse 7B, 38124 Braunschweig, Germany

Micalis Institute, INRA, AgroParisTech, CIRM-Levures, Université Paris-Saclay, 78350 Jouy-en-Josas, France

[6]Mycothèque de l'Université catholique de Louvain (BCCM/MUCL), Earth and Life Institute, Applied Microbiology, Laboratory of Mycology, Université catholique de Louvain, Louvain-la-Neuve, Belgium

[7]CBS Fungal Biodiversity Centre (CBS–KNAW), Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

[8]Department of Agricultural, Food and Environmental Science, Industrial Yeasts Collection DBVPG, University of Perugia, Borgo XX Giugno, 74, 06121 Perugia, Italy

*Corresponding author: Benedetta Turchetti (email: benedetta.turchetti@unipg.it, phone: +39 075 5856487, fax: +39 075 5856470)

Abstract: The twenty-first century has brought new opportunities and challenges to yeast culture collections, whether they are long-standing or recently established. Basic functions such as archiving, characterizing and distributing yeasts continue, but with expanded responsibilities and emerging opportunities.

In addition to a number of well-known, large public repositories, there are dozens of smaller public collections that differ in the range of species and strains preserved, field of emphasis and services offered.

Several collections have converted their catalogs to comprehensive databases and synchronize them continuously through public services, making it easier for users worldwide to locate a suitable source for specific yeast strains and the data associated with these yeasts.

In-house research such as yeast taxonomy continues to be important at culture collections. Because yeast culture collections preserve a broad diversity of species and strains within a species, they are able to make discoveries in many other areas as well, such as biotechnology, functional, comparative and evolution genomics, bioprocesses and novel products.

Due to the implementation of the Convention of Biological Diversity (CBD) and the Nagoya Protocol (NP), there are new requirements for both depositors and users to ensure that yeasts were collected following proper procedures and to guarantee that the country of origin will be considered if benefits arise from its utilisation. Intellectual Property Rights (IPRs) are extremely relevant to the current Access and Benefit Sharing (ABS) mechanisms; most research and development that involve genetic resources and associated traditional knowledge will be subject to this topic.

Keywords: Yeasts, culture collection, Convention of Biological Diversity, Nagoya Protocol, Intellectual Property Rights, yeast biotechnology

Running title: Yeast culture collections in the 21st century

Introduction

Collecting, preserving, characterizing, and exploiting Nature's diversity is a fundamental activity underpinning human civilization. Selection and breeding of diverse species over many millennia has resulted in crop plants and farm animals. This process is also evident in the yeast world considering the strains used today in brewing, baking and many other industrial applications Yeasts were among the first microorganisms to have been detected and purified in culture. Since then, the use of yeasts in science and technology has grown immensely (Steensels and Verstrepen, 2014).

This article presents a review of the yeast resources and associated information available to support the research community in addressing the scientific challenges of the twenty first century. Our primary aim is to highlight new opportunities arising from the expertly curated yeast culture collections that have been assembled worldwide over the past century. We also wish to encourage yeast scientists to continue to deposit important research strains into public domain collections to mutual benefit. Public availability has long been recognised as an essential step to safeguard the scientific reproducibility, ensure efficient distribution of important strains to the scientists of today, and preserve strains for discoveries in the future (Stackebrandt *et al.*, 2014).

History

The Carlsberg laboratory and its culture collection was one of the first founded (www.carlsberglab.dk/About/history/Pages/default.aspx). The oldest working service culture collections are recorded as being the Mycothèque de l'Université catholique de Louvain (MUCL) and the Centraalbureau voor Schimmelcultures (CBS) established in 1894 and 1906, respectively (Uruburu, 2003). The value of collecting, preserving and making publicly available the microbial diversity was recognised not long after E.C. Hansen isolated the first commercially valuable pure culture of *Saccharomyces cerevisiae* in 1883 (Barnett *et al.*, 2001). Considerable effort is being made to expand collections and services to meet contemporary needs. Global yeast culture collections continue to add new isolates, enhancing the microbial diversity to be used for future need. Indeed, some estimates suggest that more than 70% of existing fungal species diversity remains to be discovered (Mora *et al.*, 2011). In the post-genome era, new technologies such as next generation DNA sequencing and high performance computing are trying to understand yeast diversity to an ever greater extent. This, in turn, is leading to a wealth of new discoveries and new ideas about how to exploit our yeast heritage to maximum advantage. The following sections describe the resources available and some of the new opportunities arising in greater detail.

Core Capabilities

Biodiversity preservation, whether it be assemblages of novel species or genetic stocks of extensively researched species such as *S. cerevisiae* and *Schizosaccharomyces pombe*, is the key

function of culture collections. Depositing yeasts in culture collections implies that the biodiversity will be preserved in the long-term for future generations. This is important in the case of loss of biodiversity due to, for instance, the uniform use of starters in food fermentations, or the use of living modified organisms (LMO). The availability of preserved yeasts for a long period of time will allow the assessment of the effect of human activities on yeasts used in various biotechnological processes, or present in the environment (Purvis and Hector, 2000; Hoag, 2010).

Taxonomy is the second most important activity. Without a robust taxonomic framework, it would be impossible to recognise novel species or rare variants and the opportunities (or threats) arising from their discovery. Thus systematic biology and culture collections go hand in hand. The former is facilitated by the latter, and the latter would become mere archives of significantly reduced value without the former. Phylogenetic (and increasingly phylogenomic) methods are the currently preferred approaches to build a robust phylogenetic framework. The resultant family trees often require some re-interpretation of pre-existing frameworks based on physiologic characteristics and result in nomenclatural revisions.

In addition, freely available (or at least reasonably priced) authenticated cultures, safe deposit facilities for long-term preservation, patent repository, and identification services using the most up-to-date methods must be included in the core capabilities of a fully functioning culture collection. Finally, some collections offer screening services, in which genetic, physiological or biochemical properties of hundreds (or even thousands) of cultures are used for selecting industrially relevant strains.

Benefit arising

Many of the new scientific opportunities that public domain culture collections provide revolve around the availability of rapidly growing banks of novel isolates and the even faster development of associated genomics and metabolomics databases. Collections should no longer simply provide a pure, authenticated culture with reasonable provenance data. In an ideal world, whole genome sequences and comprehensive metabolomic profiles would be available for every strain in every collection. Each collection should aspire to this goal and aim at becoming the world's largest or most biodiverse or best characterised yeast resource. However, in a world with fluctuating public science spending, growing environmental problems, and novel yeasts that show strong potential for supporting production of renewable chemicals, e.g. biodiesel (Sitepu *et al.*, 2014a), it makes good sense for existing collections to collaborate globally, to build on each other's specialist skills, and to encourage capacity building in developing countries.

The availability of the Best Practice Guidelines published by the Organisation for Economic Co-operation and Development (OECD) is an important first step in harmonising collection activities

with regards to preservation and quality assurance data management. However, more connectivity is still needed. For example, culture collections need to cooperate to stay abreast of shipping, biosafety, and biosecurity regulations, to understand their obligations under the terms of the Convention of Biological Diversity (CBD) and the Nagoya Protocol (NP), and to share any market intelligence that might guide accession policies and aid decisions on commercial services offered by the collections. Furthermore, sharing of contractual best practices would be useful for both developed and emerging collections, for example, how to invest proceeds from culture collection screening and commercial development into collection maintenance, expansion and improvement. The following sections discuss these issues in more details.

Yeast collections of the world

Culture collections differ in size, form, and function. Different types of collections include public service collections with a primary taxonomic focus, private collections from industries or individuals, patented collections as well as diverse working or research collections. Some collections were originally established for specific research or application purposes; collections may focus on a specific topic and may be linked to a particular sector such as the environment, health care, education, or agriculture. There are also collections that are specialised in the safekeeping of parts of yeasts DNA such as genomes, plasmids and cDNAs. For example, addgene (http://www.addgene.org) preserves and distributes plasmids including a wide variety of yeast plasmids.

Yeast collections have three interdependent resources that are critical to yeast research. The most obvious is the yeast stocks. Just as crucial are the data associated with the yeasts, which are extensive and publicly available in some cases. A professional curator is the third essential resource, as they are able to assist users in locating the most appropriate strains for their application, based on their expertise, and knowledge of the collection and associated data. Even the largest collection would be of little value if this combination is incomplete.

Today, many of the public domain yeast collections are listed at the World Data Center for Microorganisms (WDCM; http://www.wdcm.org/), that is compiled and validated by the World Federation for Culture Collections (WFCC; http://www.wfcc.info). The amount and the nature of the data available for the conserved strains vary according to the culture collection considered. However, a number of culture collections (CBS, CABI, CLIB, now CIRM-Levures, DSMZ and other partners from 10 European countries), from the EU project MINE, defined a minimal dataset of at least 60 parameters, including carbon- and nitrogen-source assimilation, carbon-source fermentation and a number of diverse phenotypic characteristics. This trend of displaying strain-related information is now followed by others (PYCC and soon UCDFST). Aside from a few online

databases, it is usually not possible to conduct thorough searches using several criteria. It is desirable that all culture collections adopt similar types of information to make these databases more interoperable to facilitate searches in all yeast databases simultaneously. Global Catalogue of Microorganisms (GCM) (http://gcm.wfcc.info/) is an example of a single searchable catalog, in which the strain catalogs of over 100 collections have been merged into a single one. These WFCC resources have made it easier for collections with insufficient local support to post and sustain online catalogs. Additionally, species monographs, the traditional form to organise data, are assembled from culture collection strain catalogues and scientific publications to summarise the knowledge that has accumulated since the species description. In order to contrast the dispersion of data, initiatives to centralise data access are the objective of culture collection networks (e.g. Microbial Resource Research Infrastructure, MIRRI, http://www.mirri.org/user-service.html, Schüngel and Stackebrandt, 2015) and their host organizations (e.g. http://www.straininfo.net, Verslyppe *et al.*, 2014).

The long and growing list of yeast repositories and collections reflects the growing importance of yeasts in fundamental discoveries and in industrial applications. Table 1 lists global yeast culture collections that hold at least 500 strains in their public catalogs. The number of public yeast collections and the number of yeast strains held in these collections have grown considerably in the past few years (Boundy-Mills, 2012). Much of these data were taken from the WFCC (http://wfcc.info), which contains descriptive information on 715 culture collections.

Table 1 does not list all *ex situ* preserved yeast biodiversity. Numerous public microbial culture collections containing fewer than 500 yeast strains are not listed in the table, as well as industrial collections like Carlsberg Research Collection and Hefebank Weihenstephan. The research collection of Marc-André Lachance (UWOPS, University of Western Ontario, Ontario, Canada) is included in this list because of its large size (around 6000 strains), scope and impact. Additionally many academic yeast research collections that are not yet publicly available are difficult to enumerate.

Academic research collections, assembled with great care by dedicated researchers, have great potential: they are often specialized, or offer a wide range of diversity, or rare sets of mutants. Unfortunately, the fate of these research collections is often uncertain. Such collections risk becoming ‘orphans’ with the departure or retirement of the curator. Thanks to their incorporation in bigger collections (Biological Resource Centres - BRCs), these very valuable collections could be saved, ensuring that the strains will be available for future research and development purposes. For example, arrangements are currently being made for the cactus yeast collections of William T. Starmer (Syracuse University, USA) and Philip Ganter (Tennessee State University, USA) to be

transferred to the Phaff Yeast Culture Collection (UCDFST, University of California Davis, USA). The collection of soil yeasts of Lomonosov Moscow State University, Russia, founded by Inna Babjeva was partially safeguarded by the Russian National Collection of Industrial Microorganisms (VKPM) (Table 1). This was also done for the general collection as well as the deep sea yeast collection of University of Miami, USA, curated by Jack Fell and the valuable insect-associated yeast collection of Louisiana State University, USA, curated by Meredith Blackwell, that were transferred to the CBS. Dimorphic parasites of plants and fungi collected by Franz Oberwinkler and co-workers in the Eberhard Karls University in Tübingen (Germany) will be transferred to Leibniz Institute DSMZ in the future. The *S. pombe* collection, which was developed during the studies of Paul Nurse and others, and the use of which led to the Nobel Prize in Physiology or Medicine in 2001, and the collection of *S. cerevisiae* genetic mutants that had been assembled during work at the National Institute of Medical Research in London, are now preserved by NCYC National Collection of Yeast Cultures. The dairy yeast collection of the INA-PG school at Thiverval-Grignon was rescued by CIRM-Levures in the early 2000's, following the departure of Prof. Jean-Louis Bergère. The "Bergère collection" was constituted from French dairy facility environments and cheeses prior to the common use of adjunct starter in cheese making. Unfortunately, these cases are the exception: many potentially valuable research collections are discarded before rescue efforts are successful. Researchers are encouraged to make timely arrangements for the future of their collections, including stability of stocks, digitization of data, and formal transfer agreements.

In addition to the biodiversity collections that are the focus of this manuscript, there are collections of laboratory strains such as the Yeast Genetic Resource Center (YGRC) in Osaka, Japan (http://yeast.lab.nig.ac.jp/nig.v2.1/about_en), which contains over 35,000 laboratory strains of *S. cerevisiae* and *S. pombe*. The Yeast Genetic Stock Center (YGSC), formerly at the University of California Berkeley, is now located at the American Type Culture Collection (ATCC). In addition to 8,500 wild-type yeast strains, ATCC includes 23,500 *S. cerevisiae* deletion strains and other lab strains (Sung-Oui Suh, personal communication).

The growth in the number of prominent yeast culture collections, particularly in developing nations, may have been influenced by Article 9 of the CBD (https://www.cbd.int), which encourages parties to establish and maintain facilities to conserve biological diversity *ex situ*. Formal validation of the importance of preserving biodiversity provided by the CBD has led to increased governmental support of microbial collections in some countries. This *ex situ* preservation is particularly crucial in areas of high biodiversity where *in situ* preservation through conservation of natural habitats has been less effective than desired.

Table 2 lists the focus or specialty of some major yeast collections. Yeast collections have been established at different institutions for different purposes. For example, the UCD-V&E (University of California Davis, Viticulture & Enology Department, USA) collection contains exclusively wine yeasts, the NCYC historically focused on brewing yeasts, the ARS Culture Collection (NRRL) emphasizes microbes related to agricultural applications, and the UCDFST and UWOPS collections feature isolates from nature, primarily plants and insects. The holdings of the collections, and expertise of the curators, influence their uses and services.

Considering these assumptions, culture collections can be classified into two broad categories: the well-established generalist collections, and smaller collections that are more specialized and very often interested in specific characteristics and niches. The former can provide their users with a wide diversity of samples, whereas the latter can provide a wide variety of yeast strains within specific niche requirements or with specific features. When designing experiments, users should also explore the potential of smaller collections in addition to the well-known major repositories, because these specialised collections have different areas of emphasis and expertise: environmental isolates, food spoilage or fermentation yeasts including brewing or wine strains, human pathogens, genetic stocks of model organisms, etc (Table 2). Users are therefore encouraged to search for yeast strains in as many culture collections as possible and/or browse through the existing databases separately, or accessed jointly through the GCM (http://gcm.wfcc.info/), to take advantage of genetic resources and other associated data. In addition, users should not hesitate to call upon the expertise of curators in order to better explore the biodiversity maintained in yeast culture collections.

Basic and applied research performed at yeast culture collections

Taxonomy creates the language to describe biodiversity (Agapow *et al.*, 2004). The dilemma of taxonomy is that only a tiny fraction of microbial biodiversity has been named so far despite the fact that microorganisms support most ecosystem processes and support human livelihood (Prakash *et al.*, 2013). Taxonomy is meant to serve effective communication in research, education and industry, including regulatory issues. Members of the same species are assumed to share similar characteristics and members of the higher classification units (genus, family, etc.) are assumed to do so in decreasing order. Hence, the available information on each taxonomic unit determines, for example, the decisions of technologists to react to a specific food contamination, of clinicians to treat a patient's infection, of R&D researchers to design industrial screening programs, and of policy makers to draw up lists of beneficial and dangerous microorganisms. However the short time frame and narrow, profit-focused aims of industry funding are not aligned with the needs of taxonomy research, which spans decades. This is why research on taxonomy is particularly suitable

for researchers of culture collections who know in details the characteristics of the biodiversity conserved and can count on a multitude of isolates for any kind of comparative tests. Unfortunately, because of the descriptive rather than innovative character of the endeavor, research on taxonomy suffers from insufficient economic support (Godfray, 2002). Challenges discussed by Godfray (2002) also include the dispersion of reference material and information, as well as poor integration of biological aspects in species descriptions. Reference material needs to be available for taxonomic research through culture collections without prohibitive fees, but in some cases the decreasing governmental support of culture collections and growing maintenance costs cause distribution fees to increase. Microbial species descriptions are often minimalistic without providing much insight to, for example, geographic distribution and ecology, because the necessary data only accumulate post-description. Taxonomy on the one hand must provide stable references to existing knowledge tagged by Linnaean binomial names, but also needs to link additional data to enlarge in the future the biological context of previous minimalistic species descriptions. Such additional data originate increasingly through emerging technologies, of which well-established examples are metabolic fingerprints (profiles of the metabolites produced by the strain) (Mapelli *et al.*, 2008), and next generation DNA sequencing technologies (Allendorf *et al.*, 2010). These technologies do not necessarily generate living cultures that lend themselves to the traditional assignment of Linnaean binomial names, and DNA sequence databases are increasingly populated by entries without taxonomical assignment (or with incomplete or wrong information). Reference material needs to be available for taxonomic research through culture collections without prohibitive fees, and culture collections should cooperate sharing specific and accurate information related to the conserved strains. Data like geographic distribution, ecology and physiology are determinant for the complete description of a taxa and, even if microbial species descriptions are often minimalistic without providing much insight, culture collection, and in particular their database, can be considered the place where all this information are physically accumulates at the moment of the description or post-description. A continuous update of databases should be one of the purpose of culture collection.

The community of yeast researchers has systematically accumulated a record of taxonomically relevant data. Initiated by Prof. A.J. Kluyver, then director of the Centraalbureau voor Schimmelcultures, the yeast cultures of this collection were studied and a series of monographs started by N.M. Stelling-Dekker (1931), cumulated in a multi-authored edition by C.P. Kurtzman *et al.* (2011). This series will be converted into an online encyclopedia linking a dynamic pool of data relevant to yeasts. On the other hand Eukaryote genomics was boosted by the publication of the *S. cerevisiae* genome (Goffeau et al., 1996) and the consequent follow up documented in the

306 Saccharomyces Genome Database (Engel and Cherry, 2013) and others. The addition of numerous
307 non-*Saccharomyces* genomes has enabled comparative genomics to address questions of
308 metabolomics, genome structure and eukaryote evolution (Sherman et al., 2009; Dujon, 2010;
309 Hittinger et al., 2015).
310 The cultivated microorganisms preserved in culture collections are widely assumed to represent
311 only a small fraction of microbial diversity, with an underrepresentation of ecologically specialised
312 and rare microorganisms. This is partly due to the common use of general and rich culture media
313 (Prakash *et al*., 2013). Next generation sequencing and other technologies reveal large numbers of
314 novel taxa. The challenge of culturing and conserving rare, recalcitrant and fastidious
315 microorganisms rises in importance. Taxonomy is expected to classify these taxa while culture
316 collections work on their safe preservation. Cultures are indispensable to study the organism's traits,
317 to test predicted metabolic pathways, and thereby to exploit the metagenomic data accumulating
318 from next generation sequence approaches. Metagenomic data may provide information on the
319 microbial potential, but realisation of the potential in eventual applications still depends on the
320 living, cultivable organism. Culture collections expertise in cultivating microorganisms can be
321 useful in converting information on microbial potential, via microbial culturomics (application of a
322 multitude of culture conditions) to the testable organisms (Prakash *et al*., 2013).
323 Next generation DNA sequencing technologies are expected to promote the understanding of
324 ecosystems involving microbes that have not yet been cultivated. However, such technologies bear
325 multiple challenges. Both amplicon-based fungal diversity assessments (metabarcoding) and
326 amplicon-free shotgun sequencing (metagenomics) require reference databases. A number of
327 databases contain fungal reference sequences such as NCBI RefSeq, Mycobank MycoID, ISHAM
328 ITS database, UNITE, and SILVA (Cuadros-Orellana *et al*., 2013; Quast *et al*., 2013; Schoch *et al*.,
329 2014; O'Leary *et al*., 2015; Irinyi *et al*., 2015; Nilsson *et al*., 2015). Shortcomings of these
330 databases include incomplete taxonomic representation, and uncertain taxonomic species
331 assignments (Santamaria *et al*., 2012; Scheuch *et al*., 2015; Tedersoo *et al.,* 2015). These
332 shortcomings can be addressed through utilizing sequences arising from preserved and
333 authenticated specimens, such as the ribosomal barcoding sequencing projects at CBS and the Phaff
334 Yeast Culture Collection.
335 Some major projects are currently addressing these shortcomings. One example is the 1000 Fungal
336 Genomes Project (http://1000.fungalgenomes.org), which aims to fill gaps remaining after the
337 Assembling the Fungal Tree of Life project. The goal is to generate reference fungal genome
338 sequence databases with at least two species per fungal family. The Génolevures project
339 (http://www.genolevures.org) provides annotated genome sequence data for eighteen ascomycetous

yeasts permitting large-scale comparative genome analysis. Another project that can be expected to enlarge the taxonomic breath of genome data of yeasts is the Y1000+ project (http://y1000plus.org/) with its focus on the subphylum Saccharomyotina. The 1002 Yeast Genomes project (http://1002genomes.u-strasbg.fr) intends to obtain comprehensive genomic data on a single species of yeast: *Saccharomyces cerevisiae*. Problems with the low-level taxonomic annotation of fungal data were addressed by Schoch *et al.* (2014). A large team, involving culture collection personnel and taxonomists, contributed to the curation of type specimen information and synonymies. This improved the efficiency of taxonomic updates of the NCBI taxonomy database, which serves as the standard nomenclature and classification repository for the International Nucleotide Sequence Databases Collaboration. This improved efficiency is hoped to have a positive effect on the use of recognised species names in the collaborating DNA sequence databases.

A common goal of metagenomics studies is to identify functions such as particular enzymatic activities. Metagenomics approaches have begun to discover novel biological functions in fungi or mixed ecosystems that comprise fungi (Damon *et al.*, 2011; Kuramae *et al*., 2013; Scully *et al*., 2013). Knowledge of such novel functions may help to guide the search in nature and culture collections to locate the most promising targets for discovery of the living microorganisms that possesses desired functions.

Culture collections are in a key position to influence developments in taxonomy by the accumulated biological material and knowledge. The progressive action of the yeast research community has resulted in an advanced state of yeast genome, proteome, and metabolome information, as well as yeast taxonomy, and should aim to remain at the forefront of the development and validation of new technologies.

Why culture collections have unique opportunities in applied research

Preserving yeast diversity is becoming particularly crucial as the use of other yeast species besides *S. cerevisiae* for industrial applications is increasing. The importance of *Saccharomyces* species is associated with human societies: *S. cerevisiae* is one of the oldest domesticated organisms on the planet (Steensels and Verstrepen, 2014). Its domestication and use in production of fermented beverages and foods has been proposed as an incentive for nomad human populations to become sedentary and develop agriculture (Legras *et al*., 2007). Nowadays *S. cerevisiae* and closely related ascomycetous yeasts species are the major producers of biotechnology products worldwide, out-performing other groups of microorganisms in productivity and economic impact (Johnson, 2013a/b). *Saccharomyces* is by far the most highly studied genus of yeast, widely studied for its basic and applied perspectives. Its principal role in industrial applications includes traditional food fermentations like wine, beer, bakery products, cider, sake, distilled spirits, sausages, cheese, and

local traditional foods and beverages, as well as industrial production of numerous products including bioethanol, single-cell protein (SCP), feeds, probiotics (such as *S. boulardii*), flavor ingredients, and industrial enzymes. Additionally, *S. cerevisiae* is used as an eukaryotic model organisms in medical research because of the high similarity in physiology with human cells. In fact, four Nobel prizes in physiology or medicine in the past 15 years resulted from discoveries made using yeast. Considering the number of strains developed for research, this species is the most abundant in the main yeasts culture collections (Daniel and Prasad, 2010); for example 29,328 strains (and genetic stocks) are deposited in ATCC**,** 1,443 in DBVPG, 1,147 strains in NCYC, 676 in UCDFST, 663 in CECT, 614 in NRRL, 382 in CBS, 348 in BCCM/MUCL, 320 in PYCC and 282 in CIRM-Levures (data obtained from online databases). Some strains are common to several collections, because the same strain was deposited in more than one collection and assigned different accession numbers, but the majority of the strains are unique. This vast diversity of natural and engineered strains, preserved in culture collections, (Table 1) virtually ensures the probability of finding the exceptional yeasts with desired properties.

Several research groups have used genetic modification (GM) or non GM techniques to overcome the limitation of *S.cerevisiae* such as limited spectrum of fermentations, production of off-flavours in food fermentations, or suboptimal fermentation performance. They have created superior variants of *S. cerevisiae* by, such selective breeding or adaptive evolution (Steensels *et al*., 2014). In recent decades, the potentialities of the Non-Conventional Yeasts has been addressed (Johnson, 2013a/b; Steensels and Verstrepen, 2014). There is no generally accepted definition for Non-Conventional Yeasts. Even though a number of scientists include *S. pombe* and *Kluyveromyces lactis* into the group of "conventional yeasts", many other consider Non-Conventional Yeasts as synonymous to "non-*Saccharomyces*" yeasts. Overall, Sibirny and Scheffers (2002) underlined that an increasing number of Non-Conventional Yeasts is gaining importance in fundamental and applied microbiological sciences. They are studied for many practical applications, such as: i) biocatalysts and multi-enzyme pathways for the synthesis of fine chemicals and small molecular weight molecules to be used as building blocks for medical and nutritional applications, ii) agents of biocontrol, bioremediation or indicators of environmental quality, iii) producers of specific industrial enzymes, iv) high-lipid producers including single-cell oils for biofuels (biodiesel), v) producers of carotenoids, surfactants, flavorants, organic acids, or vi) a host of heterologous proteins (Buzzini and Vaughan, 2005; Wolf, 2012; Johnson, 2013a/b).

Isolation and identification of yeast strains from natural environments is the starting point for the use of microorganisms in new biotechnologies. In this way, basic research supports applied research. The optimisation of specific biotechnology applications rarely starts by sampling and isolation of

new yeast strains. It is much more efficient to rely on other research groups experienced in basic microbiology and taxonomy: this is the role of service culture collections. Culture collections accumulate and conserve strains with distinctive characteristics, that were the result of countless explorations and sampling campaigns carried out over many decades in various environments. The culture collections are thus the place where geographic, temporal and genetic biodiversity is preserved, and provided for technological exploration.

In the last few years an increasing number of researchers have utilised service culture collections for large screening programs with different aims, but with the common intention to explore the potentialities of yeast natural biodiversity. Recent characterization of oleaginous yeasts in the Phaff Yeast Culture Collection has revealed 17 new oleaginous species (Sitepu *et al*., 2012; Sitepu *et al*., 2013), diversity in fatty acid profiles (Sitepu *et al*., 2013), and carbon source utilization and inhibitor tolerance (Sitepu *et al*., 2014b). A screen of 180 yeasts revealed some species, and strains within a species, that are highly tolerant of ionic liquids (Sitepu *et al*., 2014c). After screening more than 150 CBS strains, highly osmo- and thermotolerant strains were identified (Kurtzman *et al*., 2015). In these examples, yeasts isolated and preserved over many decades were selected for novel purposes. Diversity of characteristics within a species has consistently been observed, emphasising the importance of screening multiple strains of a given species, and thus the importance of preserving multiple, diverse strains within species in properly managed culture collections, to allow innovative uses by future generations of scientists.

Additionally, the market is changing: there is an increasing awareness of the role of microorganisms in industrial production, and acceptance of microbial products in food and non-food preparations. Culture collections benefit from the financial support that in turn allows continuous exploration of extreme, disappearing and unexplored environments, isolation of new strains and species and their long-term preservation. In this way, the support of culture collections by forward-thinking enterprises provides benefits for the future

Challenges and opportunities of the Nagoya Protocol

Microorganisms, animals, and plants are recognised as important genetic resources. The CBD was signed by 150 government leaders at the Earth Summit at Rio de Janeiro in 1992 and came into force in 1993. The Biodiversity Convention aims at (1) the preservation of biological diversity (both, *ex situ* and *in situ*), (2) the sustainable use of these resources and (3) the fair sharing of benefits arising from the use of such resources. Although the Convention links traditional conservation efforts to the economic goal of using biological resources but in a sustainable manner, it does entail more bureaucracy and investment of time. The Nagoya Protocol (NP) on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization is a

supplementary agreement to the CBD following its Article 15 (https://www.cbd.int/convention/). It provides a legal framework for the effective implementation of one of the three objectives of the CBD: the fair and equitable sharing of benefits arising out of the utilisation of genetic resources. The Protocol was adopted on 29 October 2010 in Nagoya (Japan), and entered into force on 12 October 2014. Implementation of CDB obliged each country to set up their own regulations and restrictions, which resulted in an unmanageable multitude of national regulations. By designing common ways of organizing access to resources, the NP aimed at easing access to countries' genetic resources and ensuring that potential benefits arising from the use of these resources would be shared with those countries. Importantly, the CBD and the NP are legally binding agreements, and countries that sign it are obliged to implement its provisions nationally. One year after the NP has entered into force, new national laws and regulations providing requirements for access to genetic resources, are changing traditional views on collecting, depositing and distributing of microorganisms, including yeasts.

Although the importance of understanding biological diversity was not questioned in microbiological research, microorganisms are not commonly considered covered by the existing CBD regulations in the same way as macroorganisms like plants and animals. Partly, this is due to outdated views on borderless abilities of microorganisms in general to propagate (e.g. Finlay, 2002; Wilkinson et al., 2012), where microbial cells can be distributed over longer distances with water and air currents (e.g. Smith et al., 2012) and grow in number where favorable conditions occurred. Indeed, yeasts are reported to be found among other microorganisms in air samples collected at high altitude, which suggests they may be travelling between continents (Fröhlich-Nowoisky et al, 2012; Smith et al., 2012). However, many yeasts species do not show global distribution (Starmer and Lachance, 2011): only a limited number of species were reported to be widespread, e.g. *Cryptococcus magnus, C. podzolicus, C. terricola, C. victoriae* (now reclassified as *Filobasidum magnum*, *Saitozyma podzolica*, *Solicoccozyma terricola*, *Vishniacozyma victoriae*, respectively) (Fonseca et al., 2011; Yurkov et al., 2012; Yurkov et al., 2015; Liu et al., 2016). Furthermore, biogeographic patterns have been shown for yeasts suggesting that endemism may be present in yeasts at the species (Lachance et al., 2003; Inácio et al., 2010) or population (Yurkov and Chernov, 2005; Lachance et al., 2011; David-Palma et al., 2014; Yurkov et al., 2015) levels. Therefore, the present knowledge suggests that the territorial origin of yeasts as a genetic resource in the sense of the NP should be acknowledged.

Any kind of research should conform to both CBD and NP, and basic research cannot be separated from applied studies. In the case of microorganisms, including yeasts, almost any activity is likely to be formally considered as research. For example, isolation and identification of a yeast strain

would provide information about its growth abilities and genetic properties. This problem is unique to microorganisms, as they cannot yet be studied without removing them from the environment. When a researcher intends to collect material in another country or deposit cultures in a collection outside the country of origin of the samples, permits to collect, transport, isolate, study and deposit genetic material (like yeast cultures) must be obtained. The contents of such permits may vary depending on the aim of a study. Because national regulations may vary between countries, a unified portal dedicated to ABS and CBD is being established to provide researchers an overview of the requirements and identify responsible authorities in the country (https://www.cbd.int/chm/). It is also planned that future permits will receive a common identifier, an internationally recognised certificate of compliance (IRCC-CBD number) and will be available at the ABS clearing house portal (https://absch.cbd.int/search/national-records/IRCC), so collections and any other users will be able to track the origin of the genetic material.

Unlike plants and animals, yeasts cannot be directly collected in the environment as single entity but they need to be isolated from the environmental samples. Therefore, proper documentation of sampling is essential to comply with both CBD and NP. This information will be required by culture collections for every deposited strain in the future. Although many culture collections were already requesting information regarding the origin of strains in their deposition forms, such data were not always mandatory. Researchers will still hold all responsibility for the material with which they work; culture collections have the additional responsibilities to collect the information related to each isolate. It is important to emphasize that collections are not official control agencies, but have to provide the information to users when a strain is ordered. Unless national governments and controlling agencies provide the minimal requirements for information, the following data fields are likely to be made mandatory in deposition forms in the future:

(1) Country and place of sampling, including geographical coordinates. Recording of approximate geographical coordinates is extremely important as it would provide users essential information to specify whether a culture has been collected from national or international waters. Sampling in international waters and territories regulated under other treaties (e.g. Antarctic Treaty, Svalbard Treaty) are not covered by both CBD and NP. Nevertheless, yeast cultures are being isolated from samples of ocean waters, sediments, Antarctic soil and glaciers collected during scientific expeditions. In these cases a collection may ask a researcher to prove the origin of samples by providing additionally the name of a scientific station or shipping vessel.

(2) Collection date, as well as the name of the person who performed the sampling. The date when the sample was collected is necessary to understand whether ABS rules apply to a strain, since cultures isolated before ratification of the NP by a country are not covered by the regulation. Some

regions within countries that did not yet ratify the NP may require collection permits as well, such as national nature reserves. It is important to document that the NP is not applied retroactively. Persons who collected the samples and isolated cultures, as well as the depositor, are expected to keep the original permits and data regarding the origin of strains. Although strains used in research should be of known origin and sampling date, this information is not always available for old strains. However, the competent authorities would still need a formal proof that the material was collected before the CBD was ratified. Since collections keep records of their accessions, it might be a practical solution that they could formally use the accession date (when a strain was deposited) to stand in place for the date of sampling to confirm that a yeast culture has been isolated before the CBD. It is however unclear how individual researchers would overcome this problem.

(3) Information regarding sampling permits, so-called Prior Informed Consent (PIC) and Mutually Agreed Terms of Access (MAT). Collection of samples, and subsequent isolation of strains deposited in a collection, is expected to comply with both CBD and NP and respective national laws. Consequently, researchers should provide information regarding sampling permits when they deposit strains. If a country does not restrict access to their resources, this information should be also stated. Mutual or formal agreements with a landowner or a field research station authority can be provided together with the respective contact details. This, however, applies only if the national law has provided them with this power. It is important to mention that national laws may differ from general CBD and NP guidelines. This information should be obtained from the local ABS office. A sampling permit should ideally (among others) allow the transport of samples outside a country, the isolation of strains thereof, the necessary analyses and the deposit of cultures in a public culture collection from where the strains can be further distributed.

(4) Restrictions regarding deposition and distribution of strains to culture collections and their users. Since national laws may restrict access to the genetic material originating from their countries, a depositor should obtain this information from the person that did the sampling. This information is in all probability included in the MAT. A combination of sampling permits and MAT establishes the basis for a Material Transfer Agreement (MTA), which is used to distribute a culture to users. It is important to note that each culture collection would need to decide whether or not they can take a strain with certain restrictions in the open collection. Collections may need to take the decision not to accept deposits with too strong restrictions, when this would cause additional unacceptable distribution costs or would result in an excessive workload.

Collections are considered service collections if they collect, store, identify, accept, and distribute material on behalf of the scientific community. Activities in direct connection with deposition and release of cultures may need to be distinguished from other services rendered by a collection to

comply with the CBD. Storage of cultures in the open collection implies that the material is characterised according to the internal quality standards and has all relevant information regarding the origin, required permits, PIC and MAT. Also, strains from restricted collections, which are not immediately available to the public, such as unpublished, safe and patent deposits, are covered by these regulations. Identification services are not generally covered by the CBD since the material (samples or strains) will be destroyed after analyses and would not enter the collection.

Regulations of the European Union

In the European Union (EU), access to genetic resources is regulated by the Regulation No 511/2014 of the European Parliament and of the Council of 16 April 2014 on compliance measures for users from the NP on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization in the Union. Both companies and academic researchers in the EU should comply with both CBD and NP and ensure that they use the genetic material in accordance with the applicable access and benefit-sharing legislation. In particular, these regulations regarding access and benefit sharing apply when (1) the material has been collected after the ratification of the NP by a member state; (2) the country where material has been sampled signed the CBD and ratified the NP; (3) national regulations regarding ABS have been approved by the government and have been made available to the date of sampling. Importantly, also recipients of publicly funded research will need to comply with the new regulations. Funding agencies may request required permits among other documents to be submitted with a research proposal in the future. As regards the temporal scope, the EU regulations apply to the material collected in a Party to the NP (see above) after the entry into force of the NP for the EU, specifically after 12th of October 2014. It is not unlikely that some Parties may adopt a wider temporal scope, which then may also include "new use" of materials collected earlier in the country of origin, and for example deposited in *ex situ* collections. Some of these concerns are discussed below under *other restrictions*.

The regulations recognise collections as major suppliers of genetic resources and traditional knowledge in the EU. In order to help users to comply with their obligations and to supply the genetic material, a system of registered collections within the EU will be put in place. This system will ensure that collections, which are included in this register, uniformly apply their management practices to comply with both CBD and NP regulations by providing evidence of legal access to their resources, and ensure the establishment of mutually agreed terms, where required. The European Commission will maintain this register of collections and offer culture collections within the EU the possibility to voluntarily apply for admittance to this register. Each EU country should verify if a collection meets the requirements for recognition as a collection for inclusion in the

Register of collections. Unless EU Member States provide further details to the requirements, it is hard to estimate whether additional administrative demands would be affordable for most of collections. Also, there is presently no common opinion whether implementation of the CBD/NP regulations in the EU would result in a reduced number of strains deposited, or supplied by public collections.

Other restrictions

Both CBD and NP regulations are expected to provide a common transparent legal framework for access and benefit sharing among Parties. However, it is important that researchers should always consider national regulations of the country of origin of strains they would like to work with. A country which did not sign the CBD, or which did not sign the NP, may have national laws that govern access to their resources. Although in most cases regulations on the CBD do not account for material collected before the CBD ratification date, some countries extended their laws to cover genetic material collected prior to CBD. In this case, it is useful to contact the local ABS authorities for more details.

Apart from the requirements specified in permits such as the PIC and MAT, any restrictions regarding distribution and availability of strains should be provided in the form of a MTA, which will be signed by the depositor (a researcher or a collection) and the culture collection. If further transfer of a strain in not prohibited, this MTA will accompany the strain to a user ordering this strain. Based on our experience, examples of restrictions on the distribution or usage of strains may include: deposition of cultures outside the country of origin is completely prohibited, or only allowed when the strain is also deposited in parallel in a national collection; deposition of cultures outside of country of origin is permitted in a few selected and specified collections only; the deposited strain is available for research use only; DNA of deposited strain cannot be partially sequenced or genome-sequenced; specific features/characteristics of a deposited strain may not be studied in future research or used in industrial applications.

As it has been mentioned above, each culture collection would need to decide whether or not they can take a strain with certain restrictions in the open collection. Restrictions may affect collections in different ways by increasing administrative, maintenance and shipment costs but also to interfere with their quality control practices, e.g. by prohibiting DNA sequencing. Accordingly, depositors may be charged for depositing material accompanied with a restrictive MTA. Although depositors may find an MTA important to secure genetic resources, this may be redundant in some cases because distribution policies of culture collections often limit distribution of strains to the use for scientific purposes only.

The Importance of the Budapest Treaty and International Depositary Authority (IDAs) for the Protection of Biotechnological Inventions Regarding Microorganisms Including Yeasts

The World Intellectual Property Organization (WIPO) administered Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure plays an important role in the field of biotechnological inventions.

Disclosure of the invention is a requirement for the grant of patents. Normally, an invention is disclosed by means of a written description. Where an invention involves a microorganism or the use of a microorganism, disclosure is not possible in writing but can only be effected by the deposit, with a specialized institution, of a sample of the microorganism. In practice, the term "microorganism" is interpreted in a broad sense, covering biological material the deposit of which is necessary for the purposes of disclosure, in particular regarding inventions in the field of biotechnological inventions such as food, agriculture, pharmaceutics, health, environment and biofuels.

It is in order to eliminate the need to deposit in each country in which protection is sought, that the Treaty (Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. World Intellectual Property Organization, Geneva - www.wipo.int/budapest) provides that the deposit of a microorganism with any "international depositary authority" suffices for the purposes of patent procedure before the national patent offices of all of the Contracting States and before any regional patent office (if such a regional office declares that it recognizes the effects of the Treaty).

What the Treaty calls an "international depositary authority" is a scientific institution - typically a "culture collection" - which is capable of storing microorganisms. Such an institution acquires the status of "international depositary authority" through the furnishing by the Contracting State in the territory of which it is located of assurances to the Director General of WIPO to the effect that the said institution complies and will continue to comply with certain requirements of the Treaty.

On November 1, 2015, there were 45 such authorities: seven in the United Kingdom, four in the Republic of Korea, three in Italy, Russia and the United States of America, two each in Australia, China, India, Japan, Poland and Spain, and one each in Belgium, Bulgaria, Canada, Chile, the Czech Republic, Finland, France, Germany, Hungary, Latvia, Mexico, the Netherlands and Slovakia. The institution which most recently became an IDA was the Colección de Microorganismos del Centro National de Recursos Genéticos (CM-CNRG) in Mexico, which obtained the status of IDA in August 2015. The most widely accepted kinds of microorganisms are yeasts and bacteria, both kinds accepted by more than 30 IDAs.

The Treaty makes the patent system of the contracting State more attractive because it is primarily advantageous to the depositor if he is an applicant for patents in several contracting States; the deposit of a microorganism under the procedures provided for in the Treaty will reduce his costs and increase his security. It will reduce his costs because, instead of depositing the microorganism in each and every Contracting State in which he files a patent application referring to that microorganism, he will deposit it only once, with one depositary authority. The Treaty increases the security of the depositor because it establishes a uniform system of deposit, recognition and furnishing of samples of microorganisms.

The Budapest Treaty was concluded in 1977 and entered into force on August 19, 1980. On November 1, 2015, there were a total of 79 Contracting States of the Budapest Treaty.

The most recent statistics relating to deposits and samples furnished under the Budapest Treaty are available at: http://www.wipo.int/ipstats/en/statistics/micros/. The total number of deposits made between the year in which the Budapest Treaty became operational (1981) and the end of 2014 amounts to 92,017, the top four IDAs being as follows:

1. American Type Culture Collection (ATCC) (US) 30,461
2. China General Microbiological Culture Collection Center (CGMCC) (CN) 10,332
3. International Patent Organism Depositary (IPOD), National Institute of Technology and Evaluation (NITE) (JP) 10,182
4. Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (DSMZ) (DE) 7,768

Conclusions

Culture collections have increased their skills and their quality standards and thanks to ISO 9001, ISO 17025 and NF S 96-900 certifications, some European culture collections such as BCCM, CBS,CECT, CIRM and DSMZ, have acquired the status of Biological Resource Centre (BRC) in the past 10 years or so. For example, most of the partners of the EU project EMbaRC, http://www.embarc.eu/, were ISO 9001 or NF S 96-900 certified (Table 1). Certified or some accredited culture collections also take into account biosafety and biosecurity issues by following management guidelines, such as those of the OECD (http://www.oecd.org/sti/biotech/38778261.pdf), and the Code of Conduct for Biosecurity (http://www.embarc.eu/EMbaRC_CoC_Biosecurity_final.pdf) generated during the course of the EU projects GBRCN (http://www.gbrcn.org/) and EMbaRC (http://www.embarc.eu). All yeast collections should work toward such professional management in order to ensure that their holdings will be properly preserved as viable and pure cultures, and that the method for the identification of yeasts to the species level will be validated. The use of well-established protocols and at least two

different modes of preservation contribute to long-term preservation with the aims to minimize genetic drift and reduce strain-identity errors that often occur when yeasts are poorly stored and/or transferred continuously from one laboratory to another. One of the main purposes of culture collections is the proper preservation of biodiversity. Unfortunately, even if the International Code of Nomenclature for algae, fungi, and plants (Melbourne Code) is regulating nomenclature of fungi, including yeasts, it does not govern the mandatory deposition of the living type material in yeast culture collections to ensure the safekeeping availability of at least the type strains (Table 1, Table 2) for the scientific community. According to commonly accepted practice most journals publishing new descriptions of yeasts ask researchers to provide a confirmation that the type material was deposited in a properly managed culture collections, sometimes even in a specified collection. Unfortunately, the present practice of the deposition of the material in a collection does not necessarily result in unrestricted availability to researchers and other users. In addition to the restrictions arising from the implementation of both CBD and NP regulations (e.g. legally obtained permits), national laws of the country of origin and restrictions provided by depositors in the form of MTA may further limit the distribution of yeast cultures, including type strains. Unlike the filamentous fungi, almost all type strains of described yeast species are available in one or more internationally recognized BRCs. However, it is unlikely that published strains other than type strains will also be deposited in culture collections. Therefore, a proposition to systematically deposit microbial strains in public culture collections, as part of the publication process, was made by some curators; it is hoped that this proposition will be supported by other collections, and that journal policies will also support this proposal (Stackebrandt *et al.*, 2014). This practice would help encourage innovation by reducing obstacles to access of genetic material.

Acknowledgements: INR thanks the BBSRC for a National Capability Grant. Dr. Dagmar Fritze (DSMZ) is acknowledged for her valuable comments on CBD/NP regulations. This material is based upon work supported by an award to KLB from the US National Science Foundation under award number 1349395. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the NSF.

712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745

746

747 **References:**

748 Agapow PM, Bininda-Emonds OR, Crandall KA, *et al.* 2004. The impact of species concept on
749 biodiversity studies. *Q Rev Biol* **79(2)**: 161-179.

750 Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics.
751 *Nature Rev Genet* **11**: 697-709.

752 Barnett JA, Lichtenthaler FW. 2001. A history of research on yeasts 3: Emil Fischer, Eduard
753 Buchner and their contemporaries, 1880-1900. *Yeast* **18(4)**: 363-88.

754 Blackwell M. 2011. The Fungi: 1, 2, 3… 5.1 million species? *Am J Bot* **98(3)**: 426-438.

755 Boundy-Mills K. 2012. Yeast culture collections of the world: meeting the needs of industrial
756 researchers. *J Ind Microbiol Biotechnol* **39**: 673-680.

757 Buzzini P, Vaughan-Martini AE. 2005. Yeast biodiversity and Biotechnology. In *Biodiversity and*
758 *Ecophysiology of Yeasts*, Rosa CA, Peter G. (eds). Springer-Verlag: Godalming; 533-559.

759 Cuadros-Orellana S, Leite LR, Smith A, *et al*. 2013. Assessment of Fungal Diversity in the
760 Environment using Metagenomics: a Decade in Review. *Fungal Genom Biol* **3**: 110.

761 Damon C, Vallon L, Zimmermann S, *et al*. 2011. A novel fungal family of oligopeptide transporters
762 identified by functional metatranscriptomics of soil eukaryotes. *ISME J* **5**: 1871-1880.

763 Daniel HM, Prasad GS. 2010. The role of culture collections as an interface between providers and
764 users: the example of yeasts. *Res Microbiol* **161**: 488-496.

765 David Palma M, Libkind D, Sampaio JP. 2014. Global distribution, diversity hot spots and niche
766 transitions of an astaxanthin-producing eukaryotic microbe. *Mol Ecol* **23(4)**: 921-932.

767 Dujon B. 2010. Yeast evolutionary genomics. *Nat Rev Genet* **11**: 512–524.

768 Engel SR, Cherry JM. 2013. The new modern era of yeast genomics: community sequencing and
769 the resulting annotation of multiple *Saccharomyces cerevisiae* strains at the *Saccharomyces*
770 Genome Database. *Database* Vol. 2013: article ID bat012; doi: 10.1093/database/bat012.

771 Finlay BJ. 2002. Global dispersal of free-living microbial eukaryote species. *Science*, **296(5570)**:
772 1061-1063.

773 Fonseca A, Boekhout T, Fell JW. 2011. *Cryptococcus* Vuillemin. In: *The yeasts, a taxonomic study,*
774 *5th edn,* Kurtzman CP, Fell JW, Boekhout T. (eds). Elsevier, Amsterdam; 1661–1737.

775 Freedman LP, Cockburn IM, Simcoe TS. 2015. The Economics of Reproducibility in Preclinical
776 Research. PLoS Biol **13**(6) e1002165. doi:10.1371/journal.pbio.1002165.

777 Fröhlich-Nowoisky J, Burrows SM, Xie Z, et al. 2012. Biogeography in the air: fungal diversity
778 over land and oceans. *Biogeosciences*. **9(3)**: 1125-1136.

779 Godfray HCJ.. 2002 Challenges for taxonomy. *Nature* **417**: 17-19.

780 Goffeau A, Barrell BG, Bussey H, *et al.* 1996. Life with 6000 genes. *Science* **274(5287)**: 546, 563–
781 567.

782 Hittinger CT, Rokas A, Bai F-Y, *et al.* 2015. Genomics and the making of yeast biodiversity.
783 *Current opinion in genetics & development* **35**: 100-109.

784 Hoag H. 2010. Confronting the biodiversity crisis. *Nat Rep Climate Change*, **4**: 51–54.

Inácio J, Ludwig W, Spencer-Martins I, Fonseca Á. 2010. Assessment of phylloplane yeasts on selected Mediterranean plants by FISH with group-and species-specific oligonucleotide probes. *FEMS Microbiol Ecol* **71(1)**: 61-72.

International Commission on Zoological Nomenclature. 2012. Amendment of articles 8, 9, 10, 21 and 78 of the International Code of Zoological Nomenclature to expand and refine methods of publication. *Zootaxa* **3450**: 1-7.

Irinyi L, Serena C, Garcia-Hermoso D, *et al*. 2015. International Society of Human and Animal Mycology (ISHAM)-ITS reference DNA barcoding database—the quality controlled standard tool for routine identification of human and animal pathogenic fungi. *Med Mycol* myv008.

Johnson EA. 2013a. Biotechnology of non-*Saccharomyces* yeasts—the ascomycetes. *Appl Microbiol Biotechnol* **97**: 503–517.

Johnson EA. 2013b. Biotechnology of non-*Saccharomyces* yeasts—the basidiomycetes. *Appl Microbiol Biotechnol* **97**: 7563–7577.

Kuramae EE, Hillekens RHE, de Hollander M, *et al*. 2013. Structural Screening of a soil metatransrriptomic library by functional complementation of *Saccharomyces cerevisiae* mutants and functional variation in soil fungal communities associated with litter bags containing maize leaf. *FEMS Microbiol Ecol* **84**: 519-531.

Kurtzman CP, Fell JW, Boekout T. (eds). 2011. *The Yeasts: a taxonomic study*. Elsevier, London, UK.

Kurtzman CP, Quintilla Mateo R, Kolecka A, *et al*. 2015. Advances in yeast systematics and phylogeny and their use as predictors of biotechnologically important metabolic pathways. *FEMS Yeast Res* **15(6)**: fov050.

Lachance MA, Bowles JM, Starmer WT. 2003. Geography and niche occupancy as determinants of yeast biodiversity: the yeast–insect–morning glory ecosystem of Kípuka Puaulu, Hawaii. *FEMS Yeast Res* **4(1)**: 105-111.

Lachance MA, Wijayanayaka TM., Bundus JD, Wijayanayaka DN. 2011. Ribosomal DNA sequence polymorphism and the delineation of two ascosporic yeast species: *Metschnikowia* agaves and *Starmerella bombicola*. *FEMS Yeast Res*, **11(4)**: 324-333.

Legras JL, Merdinoglu D, Cornuet JM, Karst F. 2007. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* **16**: 2091-102.

Liu X-Z, Wang Q-M, Göker M, *et al.* 2015. Towards an integrated phylogenetic classification of the Tremellomycetes. *Stud Micol* **81**: 85–147.

Mapelli V, Olsson L, Nielsen J. 2009. Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology. *Trends Microbiol*, **26(9)**: 490-497.

Mora C, Tittensor DP, Adl S, *et al*. 2011. How many species are there on Earth and in the ocean? PLoS Biol 9(8): e1001127. doi: 10.1371/journal.pbio.1001127.

Nilsson RH, Tedersoo L, Ryberg M, *et al*. 2015. A Comprehensive, Automatically Updated Fungal ITS Sequence Dataset for Reference-Based Chimera Control in Environmental Sequencing Efforts. *Microbes Environ* **30(2)**: 145-150.

O'Leary NA, Wright MW, Brister JR, *et al*. 2015. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **Nov 8**: gkv1189.

Prakasch O, Shouche Y, Jangid K, Kostka JE. 2013. Microbial cultivation and the role of microbial resource centers in the omics era. *Appl Microbiol Biotechnol* **97**: 51-62.

Purvis A, Hector A. 2000. Getting the measure of biodiversity. *Nature* **405**: 212-219.

Quast C, Pruesse E, Yilmaz P, *et al*. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*, **41**(Database issue): D590-6.

Santamaria M, Fosso B, Consiglio A, *et al*. 2012. Reference databases for taxonomic assignment in metagenomics. *Brief Bioinform* **13(6)**: 682-95.

Scheuch M, Höper D, Beer M. 2015. RIEMS: A software pipeline for sensitive and comprehensive taxonomic classification of reads from metagenomics datasets. *BMC Bioinformatics* **16(1)**: 69.

Schoch CL, Robbertse B, Robert V, *et al*. 2014. Finding needles in haystacks: linking scientific names, reference specimens and molecular data for Fungi. Database, ID bau061.

Schüngel M, Stackebrandt E. 2015. Microbial Resource Research Infrastructure (MIRRI): Infrastructure to foster academic research and biotechnological innovation. *Biotechnol J* **10**: 17-19.

Scully ED, Geib SM, Hoover K, *et al*. 2013. Metagernomics profiling reveals lignocellulose degrading systems in a microbial community associated with a wood-feeding beetle. *Plos One* **8**: e73827.

Sherman DJ, Martin T, Nikolski M, *et al*. 2009. Génolevures: protein families and synteny among complete hemiascomycetous yeast proteomes and genomes. *Nucleic Acids Res* **37**(Database issue): D550-4.

Sibirny AA, Scheffers L. 2002. Thematic section "Biochemistry, Genetics, Biotechnology and Ecology of Non-conventional Yeasts". *FEMS Yeast Res*. **2**: 293.

Sitepu I, Ignatia L, Franz A, *et al.* 2012. An improved high-throughput Nile red fluorescence assay for estimating intracellular lipids in a variety of yeast species. *Journal of Microbiological Methods* **91**: 321-328.

Sitepu IR, Sestric R, Ignatia L, *et al*. 2013. Manipulation of culture conditions alters lipid content and fatty acid profiles of a wide variety of known and new oleaginous yeasts species. *Bioresour Technol* **144**: 360-369.

Sitepu I, Garay L, Sestric R, *et al*. 2014a. Oleaginous yeasts for biodiesel: Current and future trends in biology and production. *J Biotechnol Adv* **32**: 1336-1360.

Sitepu I, Selby T, Zhu S, *et al*. 2014b. Carbon source utilization and inhibitor tolerance of 45 oleaginous yeast species. *J Ind Microbiol Biotechnol* **41**: 1061-1070.

Sitepu I, Shi S, Simmons BA, *et al*. 2014c. Yeast tolerance to the ionic liquid 1-ethyl-3-methylimidazolium acetate. *FEMS Yeast Res* **14**: 1286-1294.

Smith DJ, Jaffe DA, Birmele MN, *et al*. 2012. Free tropospheric transport of microorganisms from Asia to North America. *Microb Ecol* **64(4)**: 973-985.

Smith D, McCluskey, K, Stackebrandt E. 2014. Investment into the future of microbial resources: Culture collection funding models and BRC business plans for Biological Resource Centres. *SpringerPlus* **3**: 1-12.

Stackebrandt E, Smith D, Casaregola S, *et al*. 2014. Deposit of microbial strains in public service collections as part of the publication process to underpin good practice in science. *SpringerPlus* **3**: 208.

Stackebrandt E. 2011. Diversification and focusing: strategies of microbial culture collections. *Trends Microbiol*, **18(7)**: 283-287.

Starmer WT, Lachance MA. 2011. Yeast ecology. In *The Yeasts: A Taxonomic Study, 5th ed.* Kurtzman CP, Fell JW, Boekhout T. (eds). Elsevier, Amsterdam; 65-83.

Steensels J, Verstrepen KJ. 2014. Taming Wild Yeast: Potential of onventional and nonconventional yeasts in industrial fermentations. *J Annu Rev Microbiol* **68**: 61-80.

Steensels J, Snoek T, Meersma E, et al. 2014. Improving industrial yeast strains: exploiting natural and artificial diversity. FEMS Microbiol Rev **38(5)**: 947-95.

Stelling-Dekker NM. 1931. Die sporogenen Hefen. *Verh K Ned Akad Wet Afd Natuurk* **28**: 1-547.

Tedersoo L, Anslan S, Bahram M, *et al*. 2015. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys* **10**: 1-43.

Uruburu F. 2003. History and services of culture collections. *Int Microbiol* **6(2):** 101-103.

Verslyppe B, De Smet W, De Vos P, De Baets B, Dawyndt P. 2010. Semantic integration of isolation habitat and location in StrainInfo. *BMC Bioinformatics* **11**(Suppl 5): P3.

Verslyppe B, De Smet W, De Baets B, *et al*. 2014. StrainInfo introduces electronic passports for microorganisms. *Syst Appl Microbiol* **37**: 42-50.

Wilkinson DM, Koumoutsaris S, Mitchell EA, Bey I. 2012. Modelling the effect of size on the aerial dispersal of microorganisms. *J Biogeogr* **39(1**): 89-97.

Wolf K. (ed). 2012. *Nonconventional Yeasts in Biotechnology: A Handbook.* Springer-Verlag: Godalming; 619.

Yurkov AM, Chernov IY. 2005. Geographical races of certain species of ascomycetous yeasts in the Moscow and Novosibirsk regions. *Microbiol* **74(5)**: 597-601.

Yurkov A, Inácio J, Chernov IY, Fonseca A. 2015. Yeast Biogeography and the Effects of Species Recognition Approaches: The Case Study of Widespread Basidiomycetous Species from Birch Forests in Russia. *Curr Microbiol*, **70(4)**: 587-601.

Yurkov A, Wehde T, Kahl T, Begerow D. 2012. Aboveground deadwood deposition supports development of soil yeasts. *Diversity* **4(4)**: 453-474.

Table 1: Culture collections with public catalogs containing at least 500 wild-type strains of yeast, in descending order of the number of publicly available strains. Information was collected from the World Data Centre for Microorganisms (WDCM, http://wdcm.nig.ac.jp/) and from curators as indicated by the footnotes

Acronym	Name	Country	Number of yeast strains in public catalog	Website URL
ATCC	American Type Culture Collection	USA	8,500 wild type plus 23,500 genetic stocks[2]	http://www.atcc.org
NRRL	National Center for Agricultural Utilization Research, USDA	USA	18,060[2]	http://nrrl.ncaur.usda.gov/cgi-bin/usda/
CBS	Centraalbureau voor Schimmelcultures	Netherlands	10,800[1]	http://www.cbs.knaw.nl
UCDFST	Phaff Yeast Culture Collection, University of California Davis	USA	7,490[2]	http://phaffcollection.ucdavis.edu
BCCM/IHEM & BCCM/MUCL	Belgian Coordinated Collections of Microorganisms: IHEM Biomedical Fungi and Yeasts Collection Scientific Institute of Public Health; MUCL Environmental and Applied Mycology, Université catholique de Louvain	Belgium	7,485[2]	http://bccm.belspo.be

DBVPG	Industrial Yeast Collection, Department of Agricultural, Food and Environmental Science, University of Perugia	Italy	6,730[2]	http://www.dbvpg.unipg.it
UWOPS	Department of Plant Sciences, University of Western Ontario	Canada	5,000[1]	n/a
BCRC	Bioresource Collection and Research Center	Taiwan	4,804[1]	www.bcrc.firdi.org.tw
CGMCC	China General Microbiological Culture Collection Center, Chinese Academy of Sciences	China	4,700[1]	http://www.cgmcc.net
NCYC	National Collection of Yeast Cultures	UK	4,074[2]	http://www.ncyc.co.uk
CICIM	The Culture and Information Centre of Industrial Microorganisms of China Universities, Southern Yangtze University	China	3,600[1]	http://CICIM-CU.sytu.edu.cn
CCY	Culture Collection of Yeasts, Institute of Chemistry, Slovak Academy of Sciences	Slovakia	3,500[1]	http://www.chem.sk/activities/yeast/ccy/
JCM	Japan Collection of Microorganisms, RIKEN BioResource Center	Japan	3,391[1]	www.jcm.riken.jp
CICC	China Center for Industrial Culture Collection, China National Research Institute of Food and Fermentation Industries	China	3,318[1]	http://www.china-cicc.org
YM	Yunnan Institute of Microbiology, Yunnan University	China	3,154[1]	n/a
PYCC	Portuguese Yeast Culture	Portugal	3,100[1]	http://pycc.bio-

	Collection, Universidade Nova de Lisboa			aware.com
NBRC (formerly IFO)	NITE Biological Resource Center	Japan	3,081[1]	www.nbrc.nite.go.jp
VKM	All-Russian Collection of Microorganisms, Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences	Russia	2,990[1]	http://www.vkm.ru
VKPM	Russian National Collection of Industrial Microorganisms, State Research Institute of Genetics and Selection of Industrial Microorganisms	Russia	2942	http://eng.genetika.ru/service-offer/vkpm/
CIRM-Levures	Centre International de Ressources Microbiennes-Levures	France	2900[1]	http://www.inra.fr/cirm/Levures
UFS	Yeast Culture Collection, Department of Microbial, Biochemistry and Food Biotechnology, University of The Free State	South Africa	2843[2]	n/a
AWRI MCC	AWRI Microorganisms Culture Collection	Australia	2700[1]	http://www.awri.com.au/
CECT	Coleccion Espanola de Cultivos Tipo	Spain	2,495[2]	http://www.uv.es/cect
KCTC	Korean Collection for Type Cultures	Republic of Korea	2,472[1]	http://kctc.kribb.re.kr/English/index.aspx
ZIM	ZIM Collection of Industrial Microorganisms, University of Ljubljana	Slovenia	2,312[1]	www.bf.uni-lj.si/zt/biotech/chair/index.html
URM	Universidade Federal de	Brazil	2,025[1]	https://www.ufpe.br/mico

	Pernambuco, Micoteca do Departmento de Micologia			teca/
UCD VEN	Wine Yeast and Bacteria Collection, University of California Davis	USA	500 wild type plus 1,500 genetic stocks[2]	http://wineserver.ucdavis.edu/industry/enology/culture/index.html
LYCC	Lallemand Yeast Culture Collection, Lallemand Inc.	Canada	1,800[1]	www.lallemand.com/
NCAIM	National Collection of Agricultural and Industrial Microorganisms	Hungary	1,660[1]	http://ncaim.uni-corvinus.hu/
VTTCC	VTT Technical Center	Finland	1,429[1]	http://culturecollection.vtt.fi/
USRCB	Ukranian Scientific-Research Cell Bank	Ukraine	1,339[1]	n/a
AHU	AHU Culture Collection, Hokkaido University	Japan	835[1]	http://www.agr.hokudai.ac.jp/oukin/index.html
DSMZ	Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures	Germany	700[2]	http://www.dsmz.de
NBIMCC	National Bank for Industrial Microorganisms and Cell Cultures, University of Chemical Technology and Metallurgy, Sofia	Bulgaria	693[1]	http://www.nbimcc.org
UOA/HCPF	University of Athens/Hellenic Collection of Pathogenic Fungi	Greece	606[1]	http://www.med.uoa.gr/~aveleg/index_files/Page596.htm
MTCC	Microbial Type Culture Collection, Institute of Microbial Technology	India	575[1]	http://mtcc.imtech.res.in/aboutmtcc.php
InaCC	Lembaga Ilmu Pengetahuan Indonesia (Indonesian Institute of Sciences, LIPI)	Indonesia	550[1]	n/a

IAFB	Collection of Industrial Microorganisms, Institute of Agricultural and Food Biotechnology, Warsaw	Poland	520[1]	http://cim.ibprs.pl
FRR	Food Science Australia, Ryde, CSIRO	Australia	500[1]	http://www.foodscience.csiro.au/fcc/services.htm
UAMH	University of Alberta Microfungus Collection and Herbarium (currently moving to University of Toronto)	Canada	500[1]	http://www.uamh.devonian.ualberta.ca

[1]Data obtained from the World Data Centre for Microorganisms online database, http://wdcm.nig.ac.jp/CCINFO. n/a, information not available.

[2]Data from curators.

Table 2. Emphasis of selected major yeast collections.

Acronym	**Specialty/focus**
ATCC	Resources for life science research. *S. cerevisiae* genetic stocks.
DBVPG	Yeasts isolated from industrial and natural habitats. *S. cerevisiae* from wineries. Basidiomycetes from cold habitats.
CBS	Resources for research on taxonomy, biodiversity and biotechnology; standards for long-term preservation and data storage.
CIRM-Levures	Yeasts from traditional fermentations (wine, cider, cheese) and environment. Genomics, taxonomy, biotechnology.
DSMZ	Dimorphic yeasts such as plant- and mycoparasites.
NCYC	Biodiversity, taxonomy, phylogenetics/genomics, brewing, biorefining.
NRRL	Agricultural production, food safety, public health, and economic development.
PYCC	Yeasts from Mediterranean ecosystems and related regional fermented foods and beverages.
UCDFST	Food and environmental isolates, contract screening.
UCD VEN	Wine yeasts and bacteria; wine fermentations and spoilage.
UWOPS	Yeasts recovered from nature, with special emphasis on necrotic cactus, drosophila, ephemeral flowers, and nitidulid beetles. Some species are represented by hundreds of independent isolates. Mostly ascomycetous yeasts.