FEMS Yeast Research, 2025, **25**, foaf037
DOI: 10.1093/femsyr/foaf037
Advance access publication date: 11 August 2025
Research Article – Population genomics and ecology

Taxogenomic analysis of *Pichia senei* sp. nov. and new insights into hybridization events in the *Pichia cactophila* species complex

Katharina O. Barros [1,2,3,†], Jassim Al-Oboudi[4,†], Larissa F. D. Freitas[1], Francisca M. P. Sousa[1], Thiago M. Batista [5], Ana Raquel O. Santos[1], Paula B. Morais[6], José Paulo Sampaio[7], Marc-André Lachance [8], Chris Todd Hittinger [2,3,4,*], Carlos A. Rosa [1,*]

[1]Departamento de Microbiologia, ICB, C.P. 486, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil
[2]Laboratory of Genetics, J. F. Crow Institute for the Study of Evolution, Wisconsin Energy Institute, Center for Genomic Science Innovation, University of Wisconsin-Madison, Madison, WI 53706-1580, United States
[3]DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53726-4084, United States
[4]Microbiology Doctoral Training Program, Laboratory of Genetics, Wisconsin Energy Institute, J. F. Crow Institute for the Study of Evolution, Center for Genomic Science Innovation, University of Wisconsin-Madison, Madison, WI 53706-1580, United States
[5]Instituto Nacional da Mata Atlântica, Santa Teresa, ES 29650-000, Brazil
[6]Laboratorio de Microbiologia Ambiental e Biotecnologia, Universidade Federal do Tocantins, Palmas, TO 77020-220, Brazil
[7]UCIBIO, i4HB, Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
[8]Department of Biology, University of Western Ontario, London, Ontario N6A 5B7, Canada
*Corresponding authors. Chris Todd Hittinger, Laboratory of Genetics, J. F. Crow Institute for the Study of Evolution, Wisconsin Energy Institute, Center for Genomic Dcience Innovation, University of Wisconsin-Madison, Madison, WI 53726-4084, United States. E-mail: cthittinger@wisc.edu; Carlos A. Rosa, Departamento de Microbiologia, ICB, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil. E-mail: carlrosa@icb.ufmg.br
†These authors contributed equally to this work
Editor: [Francisco Cubillos]

Abstract

Three strains of a novel yeast species were isolated from necrotic cactus tissues of *Cereus saddianus* and *Micranthocereus dolichospermaticus* and from phytotelmata of *Bromelia karatas*. DNA sequence analysis of the Internal Transcribed Spacer (ITS) region and D1/D2 domains of the large subunit ribosomal RNA, along with whole genome phylogenomic analysis, showed that this yeast is most closely related to *Pichia insulana*, *Pichia cactophila*, and *Pichia inconspicua*. The new species differs by 10–13 nucleotide substitutions from these species in D1/D2 sequences and exhibits <90% genome-wide average nucleotide identity to them. The name *Pichia senei* sp. nov. is proposed for the novel species, which is homothallic and produces asci with one to four hat-shaped ascospores. The holotype is CBS 16311 (MycoBank MB 858723). Taxogenomic analyses of the *P. cactophila* species complex, including *P. senei*, provide new insights about the hybridizations events that shaped this group. *Pichia insulana* and *P. inconspicua* are identified as the parental lineages that originated *P. cactophila*, and *P. senei* also appears closely related to one of the progenitors of *P. inconspicua*. We assess phylogeny, heterozygosity, and ploidy to explore the processes shaping diversity, showing how genomic data support yeast species delimitation and reveal complex hybridization.

Keywords: *Pichia senei* sp. nov.; cacti; bromeliads; *Pichia cactophila* complex; hybridizations

Introduction

The necrotic tissues of cacti are home to a diverse yeast community (Ganter et al. 2010, 2017, Freitas et al. 2020), whose members share some genotype–phenotype relationships (Gonçalves et al. 2024). Most yeasts found in decaying stems of cacti are restricted to cacti and do not overlap with other yeast habitats in the same or neighboring regions. In some cases, a yeast species is specific for a certain type of cactus or tissue in a limited geographic region (Ganter et al. 2017). Cactophilic species known by their global distribution include *Pichia cactophila*, *Candida sonorensis*, *Clavispora opuntiae*, and the species of the *Sporopachydermia cereana* complex (Ganter et al. 2017, Freitas et al. 2020). *Pichia cactophila* is one of the prevalent cactophilic yeast species in most cacti-associated yeast communities. However, this species was not found in cactus tissues collected on Curaçao, in the Caribbean region (Ganter et al. 2010). In that location, *Pichia insulana*, a close relative of *P. cactophila*, was the dominant species. *Pichia cactophila* is found throughout USA, Mexico, the Caribbean, Venezuela, southeastern Brazil, and Australia (Rosa et al. 1994, Ganter et al. 2010, 2017). According to Ganter et al. (2010), cryptic species related to *P. cactophila* could occur in some regions, as reported for *P. insulana*.

Pichia cactophila and *P. insulana* both belong to the *P. cactophila* species complex, which includes *Pichia norvegensis*, *Pichia pseudocactophila*, *Pichia inconspicua*, and the recently described *Pichia alaskaensis*. Members of this species complex have been isolated from environmental substrates, primarily cacti, as well as

Received 24 April 2025; **revised** 11 July 2025; **accepted** 8 August 2025

humans in clinical settings (Freitas et al. 2020, Mixão et al. 2024). Genetic evidence previously suggested that *P. inconspicua*, an opportunistic pathogen isolated from both human and environmental sources, represents a hybrid lineage of two other *Pichia* species (Mixão et al. 2019). Notably, this species exhibits losses of heterozygosity, a pattern often observed following hybridization events, as well as variations in ploidy (Mixão et al. 2019). Likewise, *P. alaskaensis*, which has only been isolated from a human clinical setting to date, has been suggested to result from a hybridization event (Mixão et al. 2024). Despite the characterization of these hybridization events, full parental lineages for these species are not yet identified in the literature, and it is unclear what factors shape the observed variation in heterozygosity and ploidy. It furthermore remains unclear what role hybridization and ploidy variation play, if any, in potential adaption to either environmental or clinical settings. In addition to these cryptic *Pichia* species that exhibit a strong cactophilic nature, diverse yeasts have also been isolated from bromeliads.

Bromeliads can occur in the same habitats of many species of cacti. Bromeliads, which are predominantly located in the Neotropical Region and characterized by a wide range of species, can host a rich diversity of living organisms within their phyllosphere, endosphere, and rhizosphere (Félix et al. 2022). These organisms play a positive role in the well-being of bromeliads by enhancing nutrient availability, providing protection against phytopathogens, and directly contributing to phytohormones (Dos Santos et al. 2023). Some members of this plant family have the capability of creating phytotelma environments (water tanks), wherein water and organic matter accumulate, serving as both substrate and nourishment for various organisms, including yeasts (Dos Santos et al. 2023, Rezende et al. 2023). In terms of diversity, total richness, and dominance, leaves and water tanks were reported to be significantly more diverse, richer, and with communities that are less dominant than those of flowers and fruits (Félix et al. 2022). *Papiliotrema laurentii*, *Papiliotrema nemorosus*, *Rhodotorula mucilaginosa*, *Sungouiella intermedia*, *Naganishia albida*, *Saturnispora silvae*, and *Pseudozyma hubeiensis* were reported to be the most frequent species associated with bromeliad phytotelmata (Gomes et al. 2015, Morais et al. 2020, Félix et al. 2022, Dos Santos et al. 2023). Nonetheless, some yeast species may exhibit specificity toward a particular plant type, influenced by geographical factors and the characteristics of these bromeliads (Morais et al. 2020, Félix et al. 2022). The yeast communities associated with bromeliads and cacti are sharply different, and only a few transient species are shared between them (Rosa et al. 1994, Morais et al. 2020, Freitas et al. 2020).

During a survey of yeast species associated with cacti and the phytotelmata of bromeliads in a tableland site (known in Brazil as "campos rupestres") in the state of Tocantins, Brazil, three isolates of a possible new species were found. Analysis of the Internal Transcribed Spacer (ITS) and D1/D2 sequences of the large subunit of the ribosomal RNA (rRNA) gene showed that this new species is phylogenetically related to *P. cactophila*, *P. insulana*, and *P. inconspicua*. Taxogenomic analysis also supports its separation from all species of *Pichia* with sequenced genomes. The new species was a minor component of the yeast communities associated with necrotic tissues of the cacti *Cereus saddianus* and *Micranthocereus dolichospermaticus*, as well as the phytotelmata of the bromeliad *Bromelia karatas*. The aim of this work is to describe this novel species, which taxogenomic analysis places within the *P. cactophila* species complex. In addition, we provide further analyses and insights into the intricate hybridization events that have shaped the evolution of this diverse group.

Materials and methods

Yeast isolation and identification

Collections were done in a tableland site ("campo rupestre") of the Cerrado ecosystem in Aurora of Tocantins, southeast region of the state of Tocantins (12°42′07″S and 46°25′04″W), northern Brazil. The campo rupestre vegetation is characterized by high levels of species richness and endemism (De Carvalho et al. 2012). *Cereus saddianus* and *M. dolichospermaticus* tissues were collected in November 2012, while collections of *B. karatas* phytotelmata were performed in February 2012. Thirty samples of bromeliad phytotelmata, 30 necrotic tissues of *C. saddianus*, and nine necrotic tissues of *M. dolichospermaticus* were collected. Water samples were collected aseptically with a sterile pipette and transferred to sterile flasks that were transported to the laboratory on ice for processing within 24 h. Aliquots of 0.2 ml of appropriate decimal dilutions were spread on Yeast extract-Malt (YM) medium (0.3% yeast extract, 0.3% malt extract, 0.5% peptone, 1% glucose, and 2% agar) supplemented with 0.02% chloramphenicol (Sousa et al. 2014). For cactus tissues, ∼1 g of decaying stems were collected aseptically in sterile plastic bags, transported on ice to the laboratory, and processed within a maximum of 5 h after collection. 0.1 g of each sample was suspended in 9 volumes of sterile distilled water and vortexed for 1 min. Aliquots of 0.1 ml of appropriate decimal dilutions were spread on YM agar. The plates were incubated at 25°C for 2–8 days (Freitas et al. 2020). Different yeast morphotypes were counted and purified for storage on liquid nitrogen for later identification. The yeasts were morphologically and physiologically characterized by standard methods (Kurtzman et al. 2011). Single ascospores were isolated from mature asci with a Zeiss Axio Scope.A1 microscope equipped with a micromanipulator. Prior to micromanipulation, the sporulated cultures were treated with a lyticase solution (1.5 mg ml^{-1}) for 15 min at 37°C.

Species identification was performed by sequencing of the ITS-5.8S region (primers ITS1 and ITS4) and the D1/D2 domains of the large subunit rRNA gene (primers NL1 and NL4) (White et al. 1990, Kurtzman and Robnett 1998, Lachance et al. 1999). The amplified DNA was concentrated, cleaned, and sequenced in an ABI 3130 Genetic Analyzer automated sequencing system (Life Technologies, California, USA) using BigDye v3.1 and POP7 polymer. The sequences were compared with those in the GenBank database using the Basic Local Alignment Search Tool (Altschul et al. 1990). They were aligned using MAFFT (Katoh et al. 2009) as implemented in the NGPhylogeny.fr online package (Lemoine et al. 2019). A phylogram based on the D1/D2 domains of the large subunit rRNA gene sequences was constructed by neighbor-joining analysis in the MEGA6 software package (Kumar et al. 2016) of 649 aligned positions using the number of substitutions as the distance metric. Bootstrap values were determined from 1000 pseudoreplicates.

Genome sequencing, assembly, and phylogenetic analysis

Genomic DNA of strain UFMG-CM-Y531 was isolated using a modified phenol:chloroform method (Shen et al. 2018). The sequencing was carried out in Illumina MiSeq (2 × 150) platform. We performed quality control and trimming of the whole-genome sequencing reads using Trimmomatic v0.30 (Bolger et al. 2014). Adapter removal and quality trimming were done with the following parameters: adapter sequences were clipped using the "trueseq.fasta" adapter file (ILLUMINACLIP:trueseq.fasta:2:30:10:2:keepBothReads), low-quality bases were removed from the start and end of reads (LEADING:3,

TRAILING:3), reads were trimmed using a sliding window approach (SLIDINGWINDOW:4:15), and reads shorter than 36 bases were discarded (MINLEN:36). Quality assessment was performed both before and after trimming using FastQC to ensure improvement in read quality. Paired-end reads were used to generate whole genome assembly using the tool SPADES v3.13.1 (Bankevich et al. 2012). The resulting genome assembly was evaluated for N50 and other summary statistics with QUAST v5.2.0 (Table S1) (Mikheenko et al. 2018). Genome quality was assessed by their completeness based on the expected gene content of the Benchmarking Universal Single-Copy Orthologs (BUSCO) (v5.2.2) (Manni et al. 2021).

Thirty-six *Pichia* genomes were selected for phylogenomic analysis due to their close evolutionary relationship to the novel species, and *Saturniaspora dispora* was used as an outgroup (Table S2; Opulente et al. 2024). Utilizing the BUSCO dataset for *Saccharomycetes*, 1847 single-copy BUSCO amino acid sequences were identified that were present in 90% of the 37 genomes. These sequences were aligned with MUSCLE v5.1 9 (Edgar 2004), trimmed with TrimAl v1.4 (Capella-Gutiérrez et al. 2009), and concatenated to generate a single sequence for each genome, comprising 916 048 columns that represent animo acids or ambiguous characters (in the absence of amino acid sequences). A maximum-likelihood tree was inferred from the concatenated alignment using IQ-Tree v2.2.2.3 (Minh et al. 2020) with 1000 ultrafast bootstrap replicates (-B 1000) and ModelFinder to select the appropriate substitution model. Among the 36 yeast species, we included *P. alaskaensis* in our analysis by downloading its raw reads from the NCBI Sequence Read Archive (SRA) under accession number SRR28437184 (Caldwell et al. 2024). In addition, we used the publicly available genome assembly ASM4093801v1 of *P. alaskaensis* (Mixão et al. 2024). The average nucleotide identity (ANI) was measured using orthoANI (Lee et al. 2016).

To identify putative mating types for *Pichia* species, protein sequences from *Pichia kudriavzevii* for *MAT*a*1* (XP_029320370.1), *MAT*a*2* (XP_029320371.1), *MAT*α*1* (AXK50428.1), and *MAT*α*2* (AXK50429.1) were obtained from NCBI and used in local TBLASTN 2.9.0+ (Camacho et al. 2009) searches of *Pichia* genomes. Matching protein sequences were aligned with MAFFT v7.520, and a phylogenetic tree was generated in IQ-TREE v2.3.6 with the substitution model (JTT + F + G4) selected by ModelFinder and 1000 UltraFast bootstraps (-bb 1000). Bootstrap values are displayed on branches and are only shown when below 100%.

Genome-wide sequence variants and phylogenetic network construction

To investigate the hybridization among *Pichia* species and whether the novel species candidate, which is closely related to *P. inconspicua*, is an interspecies hybrid, we undertook a multistep genomic analysis. The Genome Analysis Toolkit (GATK) v. 4.2.0.0 (O'Connor and Van der Auwera 2020) was used to identify genome-wide sequence variants among the *Pichia* species. Briefly, we used the Burroughs–Wheeler Aligner (Li 2013) to align short-read sequences from the novel species, *P. insulana*, *P. cactophila*, *P. galeolata*, *P. alaskaensis*, *P. norvegensis*, *P. pseudocactophila*, and *P. awuae* to the *P. inconspicua* CBS180 reference genome. Variant calling was performed using GATK HaplotypeCaller to generate individual Variant Call Format (VCF) files for each species. The –ploidy parameter in HaplotypeCaller was set according to the known ploidy of each species, ensuring biologically accurate variant detection. Genomic Variant Call Format (GVCF) files were then combined using GATK's GenomicsDBImport to create a unified database, enabling efficient data access and management across all species. Joint genotyping was subsequently performed using GenotypeGVCFs. During this step, we applied the parameters –max-alternate-alleles 60 and –max-genotype-count 10 240 to allow for high allelic diversity and accommodate potential complexity in variant sites. GATK's VariantFiltration was then used to produce a high-quality final variant dataset with the following parameters: QD < 2.0, SOR > 3.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5, ReadPosRankSum < −8.0. Then genotype calls were further refined using CalculateGenotypePosteriors by incorporating population-level information. Finally, heterozygosity and inbreeding coefficients were calculated for each species using VCFtools v0.1.14 (Danecek et al. 2011).

To better account for reticulation events, such as hybridization and recombination, genome-wide sequence variants were additionally used to construct a phylogenetic network with SplitsTree v6.3.34 (Huson and Bryant 2024) using the NeighborNet algorithm.

Hybridization and ploidy evaluation

To explore the hybrid status of the novel species candidate and related species, we utilized sppIDer (Langdon et al. 2018), a tool designed to detect hybrid genomes. The raw genome reads of *P. inconspicua* were mapped to the concatenated reference genomes of *C. awuae*, *P. galeolata*, *P. insulana*, *P. norvegensis*, *P. pseudocactophila*, and *Pichia senei*. Similarly, for *P. alaskaensis*, we mapped its raw reads to *C. awuae*, *P. galeolata*, *P. inconspicua*, *P. insulana*, *P. norvegensis*, *P. pseudocactophila*, and *P. senei*. In a second analysis using *P. alaskaensis* reads, we included *P. cactophila* as a reference genome in addition to those used in the first analysis. Lastly, the raw reads of *P. cactophila* were mapped to *C. awuae*, *P. galeolata*, *P. inconspicua*, *P. insulana*, *P. norvegensis*, *P. pseudocactophila*, and *P. senei*. Hybridization status and putative parentage were inferred by analyzing the read-mapping proportions and quality to the reference lineages. Ploidy state was evaluated from short read data using nQuire (Weib et al. 2018), applying both likelihood and linear regression approaches.

Results and discussion

Isolation and phenotypic characterization of the novel species

Strain UFMG-CM-Y411 was isolated from necrotic tissues of *C. saddianus* in counts of 5×10^3 cfu ml^{-1}, while strain UFMG-CM-Y6546 was isolated from necrotic tissues *M. dolichospermaticus* in counts of 2.4×10^6 cfu ml^{-1}. Strain UFMG-CM-Y531 was isolated from the phytotemata of *B. karatas* in counts of 1×10^2 cfu ml^{-1}. This new species can be regarded as a minor component of the yeast communities of these plants. The prevalent species in the yeast communities associated with rots of *C. saddianus* were previously found to be *Kluyveromyces starmeri*, *Magnusiomyces capitatus*, and an undescribed *Candida* species (Freitas et al. 2020). In rots of *M. dolichospermaticus*, the yeasts *C. sonorensis*, *Aureobasidium pullulans*, *Kwoniella heveanensis*, and *Papiliotrema terrestris* were prevalent (Freitas et al. 2020). The species isolated from the same sample as strain UFMG-CM-Y411 were *C. sonorensis*, a member of the *S. cerevisiae* species complex sometimes referred to as "*Sporopachydermia obscura*," *Pseudozyma jejuensis*, and *A. pullulans*. The sample with strain UFMG-CM-Y6546 also contained *Cereus sonorensis*. Strain UFMG-CM-Y531 was isolated along with *Meyerozyma caribbica*, *Ps. hubeiensis*, *Hagleromyces aurorensis*, *P. laurentii*, and *R. mucilaginosa*. At this time, we cannot confirm whether the new species is a true cactophilic yeast, as it is important to sample other cacti and

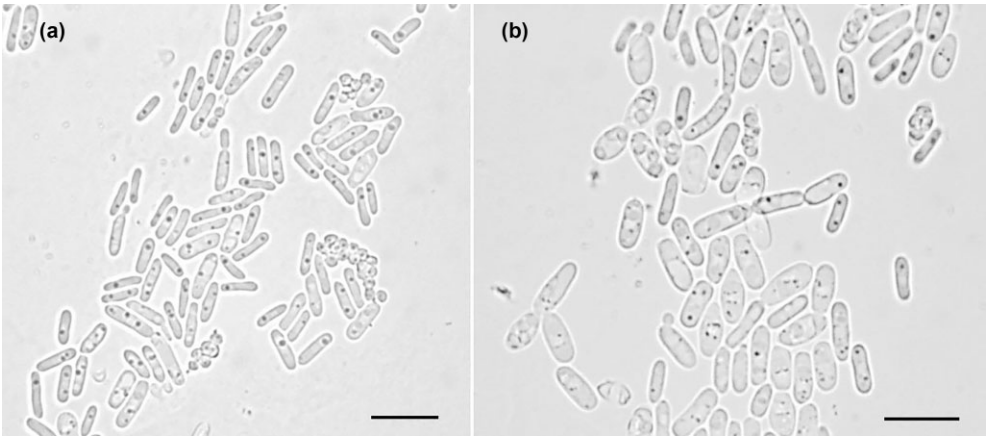

Figure 1. Budding cells of *P. senei* sp. nov. UFMG-CM-Y531 on 5% malt medium after 3 days of incubation at 25°C. Bars, 10 µm. (A) Budding cells and released and agglutinated ascospores. (B) Budding cells, asci with four hat-shaped ascospores, and rudimentary pseudohypha.

different substrates of this region and other regions of northern Brazil to determine its ecological niche. Our results suggest, based on collections in other Brazilian regions, that the species could be restricted to the cactus rots of northern Brazil, representing a local population very similar physiologically and morphologically to *P. cactophila* (Freitas et al. 2020), as suggested by Ganter et al. (2010) for *P. insulana* in Curaçao. The presence of the novel species in other habitats, such as the phytotelmata of *B. karatas*, could be due to insect vectors that visit both cacti and bromeliads in this locality.

The three strains of the novel species produced one to four hat-shaped ascospores (Fig. 1). To test whether these yeasts represent self-sporulating diploid strains, asci of strain UFMG-CM-Y531 were dissected with a micromanipulator, and the germination of 11 ascospores was tested. All cultures formed by these haploid clones were able to produce ascospores on acetate McClary agar, indicating that the species is homothallic. The new species has a restricted growth profile, assimilating glucose, ethanol, glycerol, DL-lactate, and D-glucosamine, which is similar to its closest relatives (Opulente et al. 2023). There is currently no phenotypic trait that allows one to distinguish the novel species from *P. cactophila*, *P. insulana*, and *P. inconspicua*, and DNA sequencing is recommended for the correct identification of these species.

Phylogenetic and phylogenomic analyses

A phylogram (Fig. 2) based on the D1/D2 domains of the LSU rRNA gene identified *P. insulana* as sister to the novel species. The D1/D2 sequences of the three isolates of the novel species were identical, as were those of their ITS regions (accession number KT373981). These joined a subclade containing *P. cactophila*, *P. insulana*, and *P. inconspicua*. The D1/D2 sequence of the new species (KT373981) differs by 10 nucleotide substitutions from that of *P. insulana* (NG_055166) and 13 substitutions each from *P. cactophila* (NG_055115) and *P. inconspicua* (NG_055114). The ITS sequence of the novel species (KT373981) differed by 15, 20, and 21 nucleotide substitutions, respectively, from *P. insulana* (KM252834), *P. cactophila* (NR_138243), and *P. inconspicua* (NR_111116).

The Whole Genome Shotgun project of the strain UFMG-CM-Y531 has been deposited at DDBJ/ENA/GenBank under accession JAZAQD000000000. The genome assembly statistics of this strain are shown in Table S1. Phylogenomic analysis confirmed the novelty of the strain UFMG-CM-Y531 as distinct from *P. cactophila*, *P. insulana*, and *P. inconspicua* with high bootstrap support (Fig. 3). The analysis placed the novel yeast species in an outgroup to the subclade identified in the tree presented in Fig. 3. In the phylogenomic analysis, we included *P. alaskaensis*, a species proposed for the causative agent of a fatal case of fungemia following cardiothoracic surgery (Caldwell et al. 2024), which was recently identified as an interspecies hybrid in this species complex (Mixão et al. 2024). The genome assembly of this species was highly fragmented and exhibited low BUSCO completeness [C:47.6% (S:43.8%, D:3.8%), F:27.2%, M:25.2%, n:2137]. This issue prompted us to address the missing data by including BUSCO proteins that were complete and single-copy in 90% of the input samples. Only 169 shared single-copy protein-coding genes were found in 100% of the genomes. Despite these shortcomings, the resulting tree is likely to be a better representation of the true phylogeny than that obtained from barcode sequences (James et al. 2020). Our phylogenomic analysis indicated that *P. alaskaensis* grouped with *P. inconspicua* and that the novel species was in an outgroup to the subclade containing *P. alaskaensis*, *P. inconcpicua*, *P. insulana*, and *P. cactophila* (Fig. 3).

The ANI of strain UFMG-CM-Y531 compared to the members of the *P. cactophila* species complex ranged from 74.1% (*C. awuae*) to 86.9% (*P. inconspicua* (Fig. 4). Similar genetic identities for this strain were observed for the other closely related lineages [*P. insulana* (86.4%), *P. alaskaensis* (86.2%), and *P. cactophila* (86.8%)]. In yeasts, values of 95% ANI or more can be considered indicative of strains belonging to the same species (Lachance et al. 2020, Barros et al. 2024). Based on these data, as well as the D1/D2 divergence and the topology resolved by phylogenomic analysis, we conclude that the three isolates represent a distinct species and propose the name *P. senei* sp. nov. to accommodate them.

Pichia cactophila is a hybrid of *P. inconspicua* and *P. insulana*

We investigated the genetic diversity and potential hybridization among *Pichia* species, particularly those in the *P. cactophila* species complex, which are thought to have been the subject of multiple complex hybridizations (Caldwell et al. 2024, Mixão et al. 2024), and of which *P. senei* is the most recently discovered member. *P. cactophila* exhibited high heterozygosity (25.7% of variant sites) when mapped to a *P. inconspicua* reference genome and a low inbreeding coefficient ($F = 0.17$), suggesting that it represents a hybrid lineage (Table 1). It was recently proposed that the parents of *P. cactophila* are *P. inconspicua* and a yet-unidentified *Pichia* species

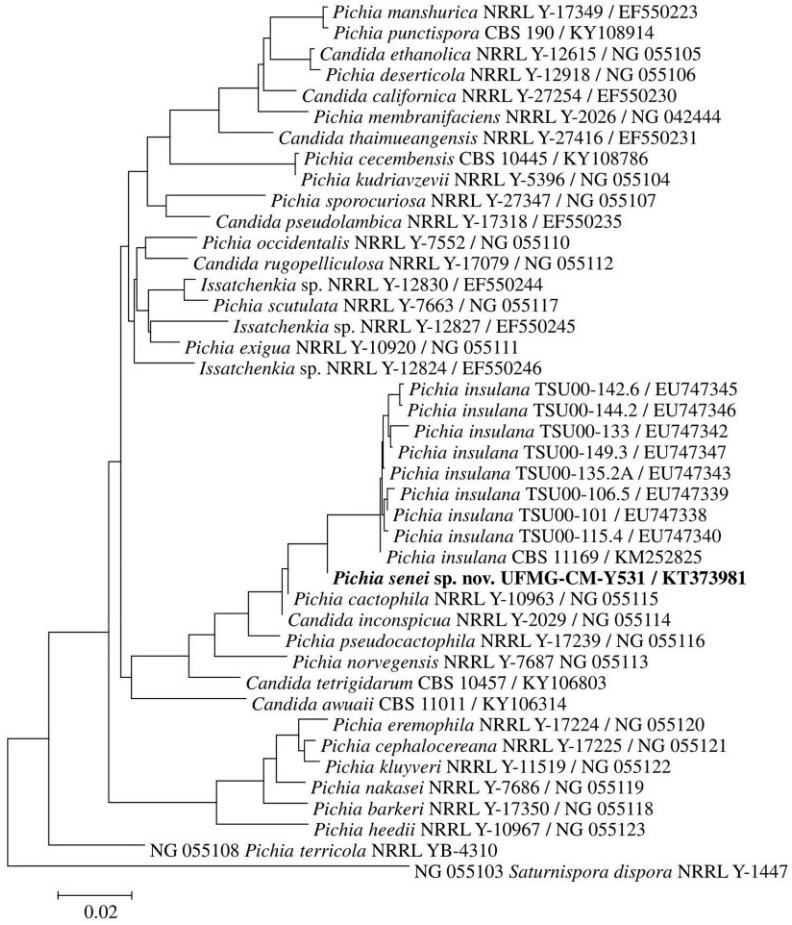

Figure 2. Phylogenetic tree of novel species and closely related species. The neighbor-joining tree was reconstructed from an alignment of 649 nucleotide positions of the large subunit rRNA gene D1/D2 region using the number of substitutions as the distance metric.

(Mixão et al. 2024). A reference-based read mapping approach performed with sppIDer confirmed the parental contribution of *P. inconspicua* and further revealed *P. insulana* as the other parent to *P. cactophila* (Fig. 5a; Fig. S1). This hybridization scenario was further supported by a phylogenetic network analysis of genome-wide sequence variants where *P. cactophila* was located on two hybrid edges of similar lengths between the branches of *P. inconspicua* and *P. insulana* (Fig. 6). The lack of additional branching at the *P. cactophila* node further suggests that the hybridization that produced it was a recent occurrence.

Mapping quality plots (Fig. 5b), which represent the confidence level of short-read alignment to a specific reference genome, further revealed that *P. cactophila* reads map with similar accuracy to *P. insulana* and *P. inconspicua*, which is consistent with the proposed parentage scenario. The lower overall amount of reads that map to *P. insulana* could be consistent with a scenario in which one of the *P. cactophila* parents was closely related to the extant *P. inconspicua* lineage but where the other parent had a more distant relationship to *P. insulana*. However, this model is inconsistent with the phylogenomic topology of these species (Fig. 3), where *P. cactophila* is more closely related to *P. insulana* than to *P. inconspicua*. If one of the parental lines is not directly represented by extant lineages, this discrepancy may also explain the observed low-percentage read mapping to *P. senei* (Fig. 5a). Given that *P. insulana* is diploid (Table 1), the results suggest a scenario involving ploidy reduction either after or before hybridization, as was suggested for the parental contribution of *P. inconspicua* (Mixão et al. 2024).

Recent genetic evidence suggested that the unknown parent of *P. cactophila* had the *MATα* cell type since *P. cactophila* harbors an apparent recombination error in the *MAT* locus with intact *MAT*a and *MATα* ideomorphs and an extra *MATα2* that matches to the *MATα2* of *P. inconspicua* (Mixão et al. 2024). For this reason, we sought to evaluate the mating type of *P. insulana*, which we propose represents the hitherto unknown parental lineage of *P. cactophila*. Phylogenetic analysis revealed that *P. insulana* does indeed harbor the *MATα* mating type, with *MATα1* and *MATα2* both present in the genome (Fig. S2), which is consistent with a scenario in which *P. insulana* or a very closely related lineage hybridized with *P. inconspicua*, donating its *MATα* genes to the hybrid progeny.

Additional hybridizations in the *P. cactophila* species complex

Pichia alaskaensis also exhibited relatively high heterozygosity (13.8% of variant sites) when mapped to *P. inconspicua* (Table 1). However, the inbreeding coefficient (F = 0.55) was not as low as *P. cactophila* (F = 0.17), and our read mapping approach suggests a strong relationship between *P. alaskaensis* and *P. inconspicua* (~82% of mapped reads) (Fig. 7a; Fig. S3), aligning with the phylogenomic relationship of these species (Fig. 3). Additionally, *P. alaskaensis* exhibited higher mapping quality to *P. inconspicua* (Fig. 7b). Regardless, ploidy evaluation indicates that *P. alaskaensis* is indeed triploid (Table 1) with the high mapping percentage of *P. alaskaen-*

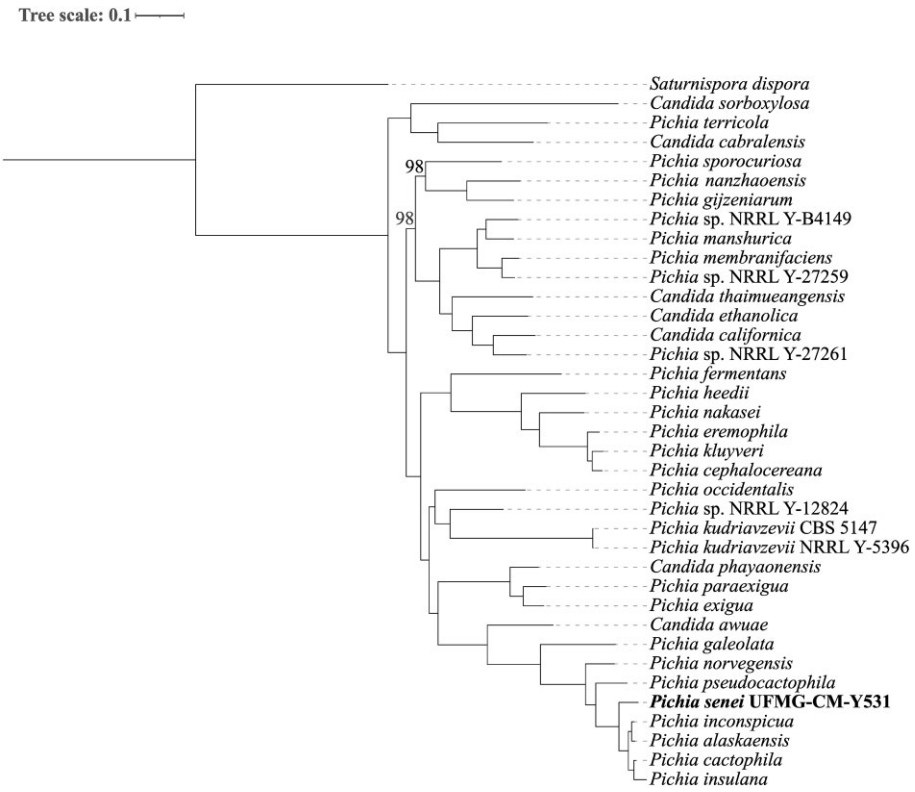

Figure 3. Phylogenomic tree of select species in the *Pichia* genus. Phylogenetic tree reconstructed from the alignment of 1847 orthologous proteins of *P. senei* compared with 35 other species plus *S. dispora* as outgroup. Bootstrap values (1000 pseudoreplicates) are displayed only for cases where the values were <100%.

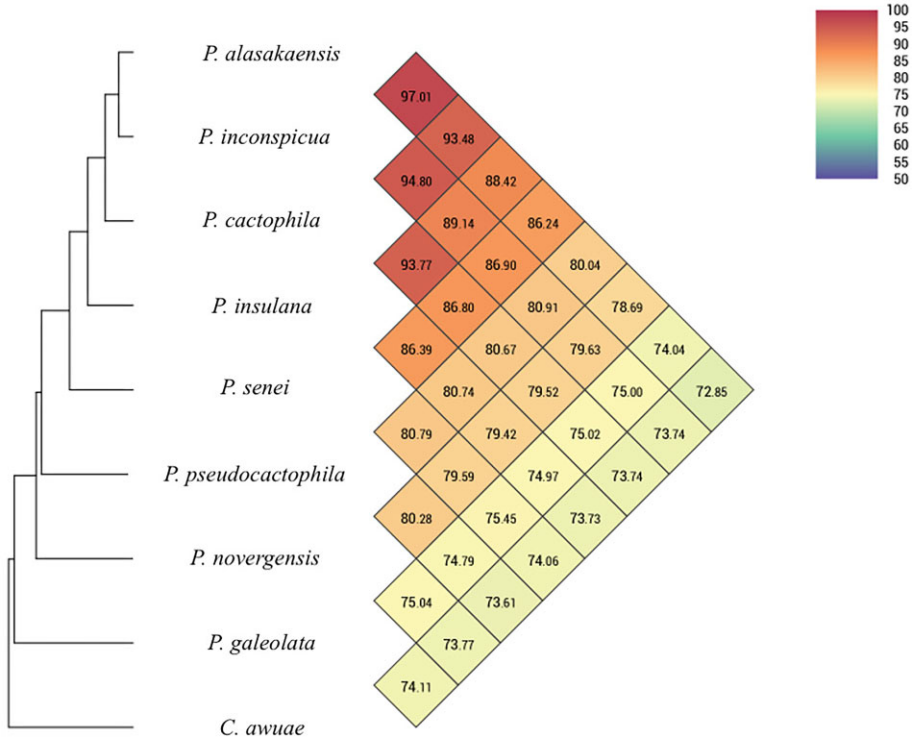

Figure 4. Heatmap generated with OrthoANI values calculated from the OAT software. Each cell in the heatmap represents the percentage of ANI similarity between two organisms. The values are shown numerically within each cell and are also color-coded according to the scale on the right. Higher similarity values (e.g., >95%) appear in dark red tones at the top of the scale, while lower values (e.g., <55%) appear in blue tones at the bottom.

Table 1. Inbreeding coefficients (F), levels of heterozygosity, and ploidy state for members of the *P. cactophila* species complex.

Species	Inbreeding coefficient (F)	% Heterozygous	nQuire ploidy
P. alaskaensis	0.81	6.56	Triploid
P. awuae	1	0.00	Haploid
P. cactophila	0.20	27.4	Diploid
P. galeolata	1	0.00	Haploid
P. inconspicua	0.81	6.67	Diploid
P. insulana	0.97	0.98	Diploid
P. norvegensis	0.97	0.96	Diploid
P. pseudocactophila	1	0.00	Haploid
P. senei	1	0.00	Haploid

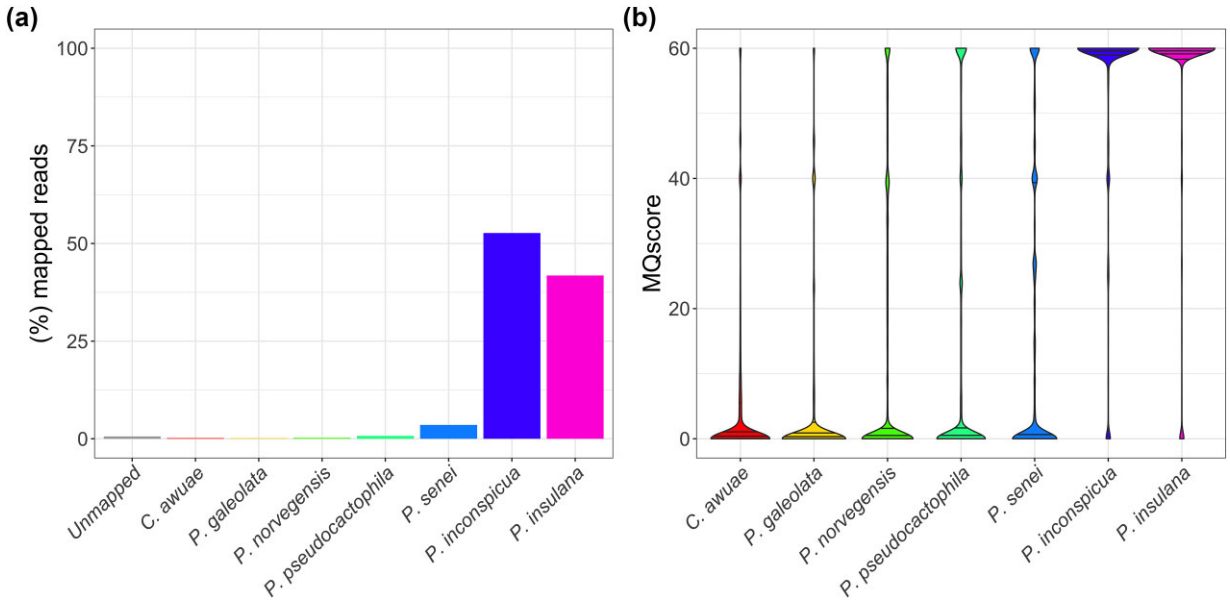

Figure 5. Hybridization analysis of *P. cactophila*. Short reads for *P. cactophila* were mapped to multiple reference genomes of closely related *Pichia* species. Mapping percentages (a) indicate the proportion of short reads that map to a particular reference genome with mapping quality (b) indicating the quality of short-read mapping to a particular reference genome.

sis to *P. inconspicua* being consistent with the previously proposed scenario (Mixão et al. 2024), wherein *P. alaskaensis* and *P. inconspicua* share a parental lineage, rather than *P. inconspicua* serving as a direct parent to *P. alaskaensis*. It could be the case that the extant *P. inconspicua* lineage is the closest representative of one of the parents of *P. alaskaensis* with the other parental lineage being as-yet-unidentified but also closely related to *P. inconspicua*. Such a scenario could explain the polyploidy and apparent hybrid nature of *P. alaskaensis*, as well as its high percentage mapping to *P. inconspicua*. The additional mapping to *P. senei* and *P. insulana* was in accordance with the phylogenomic relationships observed between these lineages (Fig. 3).

In a second analysis, with the interspecies hybrid *P. cactophila* also included as a reference genome, *P. alaskaensis* reads predominantly mapped to *P. inconspicua* (~68%) and *P. cactophila* (~25%) (Fig. 7c), with mapping quality remaining higher for *P. inconspicua* (Fig. 7d; Fig. S4). Since *P. cactophila* is a hybrid with *P. inconspicua* as one of its parents, *P. alaskaensis* likely shares genetic material with both species. The substantial read mapping to *P. cactophila* indicates that *P. alaskaensis* may have inherited genomic material from an ancestor that had either hybridized with or was closely related to *P. cactophila*. This pattern raises the possibility that *P. alaskaensis* and *P. cactophila* share a partially overlapping hybrid ancestry or that *P. alaskaensis* underwent introgression with lineages related to *P. cactophila*. This finding further suggests the polyploid nature and complex evolutionary history of *P. alaskaensis*, which is in line with the proposed scenario that *P. alaskaensis* and *P. inconspicua* share a parental lineage.

Pichia inconspicua is known for its clinical relevance and ability to adapt to various environments (Mixão et al. 2019). Hybridization in *P. inconspicua* can lead to increased genetic diversity and may enhance its adaptability and pathogenic potential (Guitard et al. 2015, Mixão et al. 2019). *Pichia inconspicua* itself has also been suggested to be a hybrid (Mixão et al. 2024). Interestingly, when its own reference genome was excluded from the sppIDer concatenated reference genome, *P. inconspicua* reads mapped roughly 75% to the recently discovered *P. insulana* and 25% to the novel species proposed here, *P. senei* (Fig. 8a; Fig. S5), but mapping quality was uneven (Fig. 8b). While the phylogenetic network analysis does not suggest a direct parental contribution from *P. senei*, it nonetheless suggests that *P. senei* represents a close relative to at least one of the parental lineages of *P. inconspicua* and, additionally, that *P. insulana* also bears some relationship to the parental lineages of *P. inconspicua*. In a recently proposed scenario (Mixão et al. 2024), the parental lineage of *P. inconspicua* that is not shared with *P. alaskaensis* is suggested to be MATa, whereas *P. insulana* appears

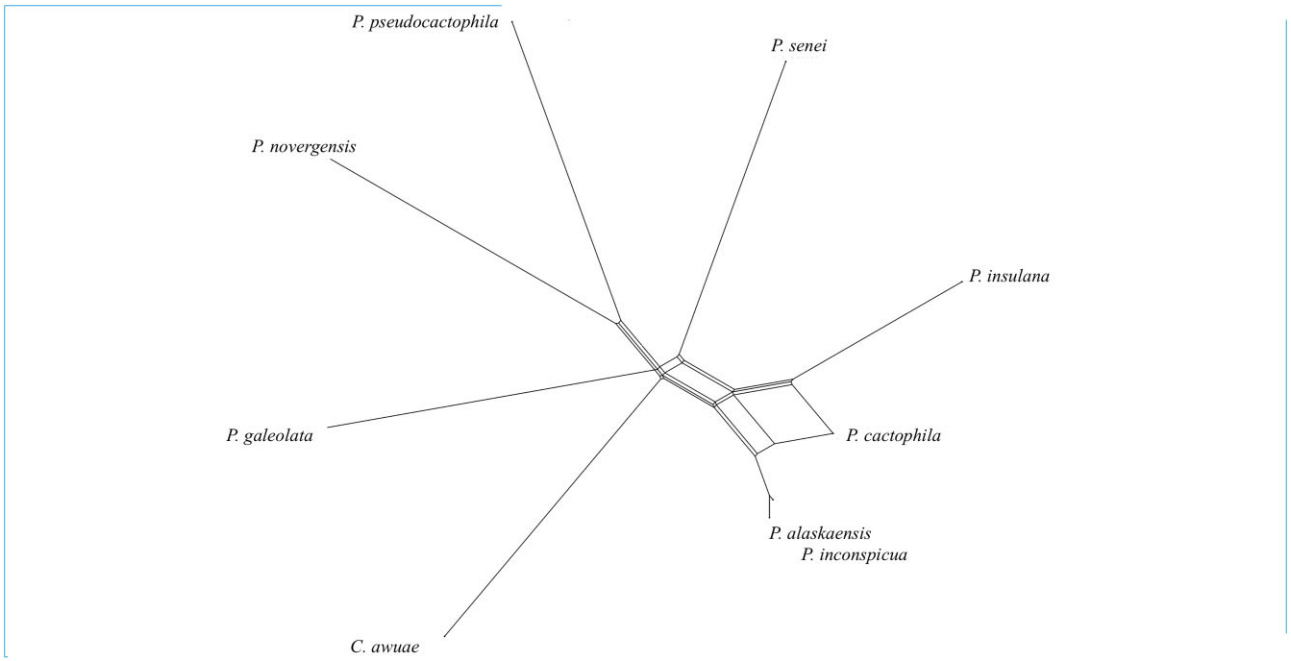

Figure 6. Phylogenetic network of the *P. cactophila* species complex. Genome-wide sequence variants were used to construct a phylogenetic network. The resulting network is displayed as a splits graph showing the phylogenetic relationship between the members of the *P. cactophila* species complex, as well as edges, which represent genomic partitions between the taxa.

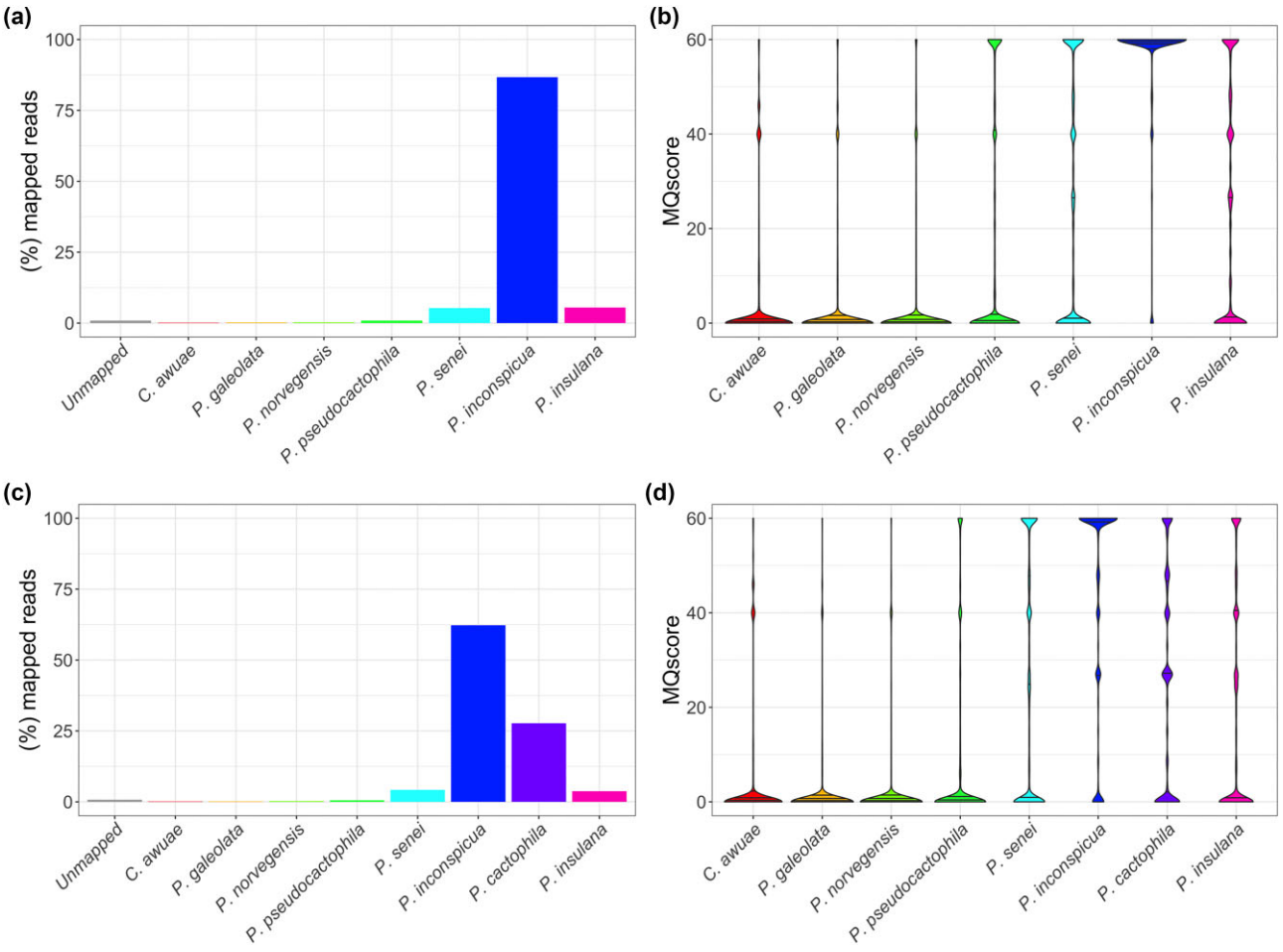

Figure 7. Hybridization analysis of *P. alaskaensis*. Short reads for *P. alaskaensis* (a) were mapped to multiple reference genomes of closely related *Pichia* species with (b) higher mapping quality to *P. inconspicua*. In a subsequent analysis, (c) the reference genome of the interspecies hybrid *P. cactophila* was included (d) with mapping quality remaining higher for *P. inconspicua*.

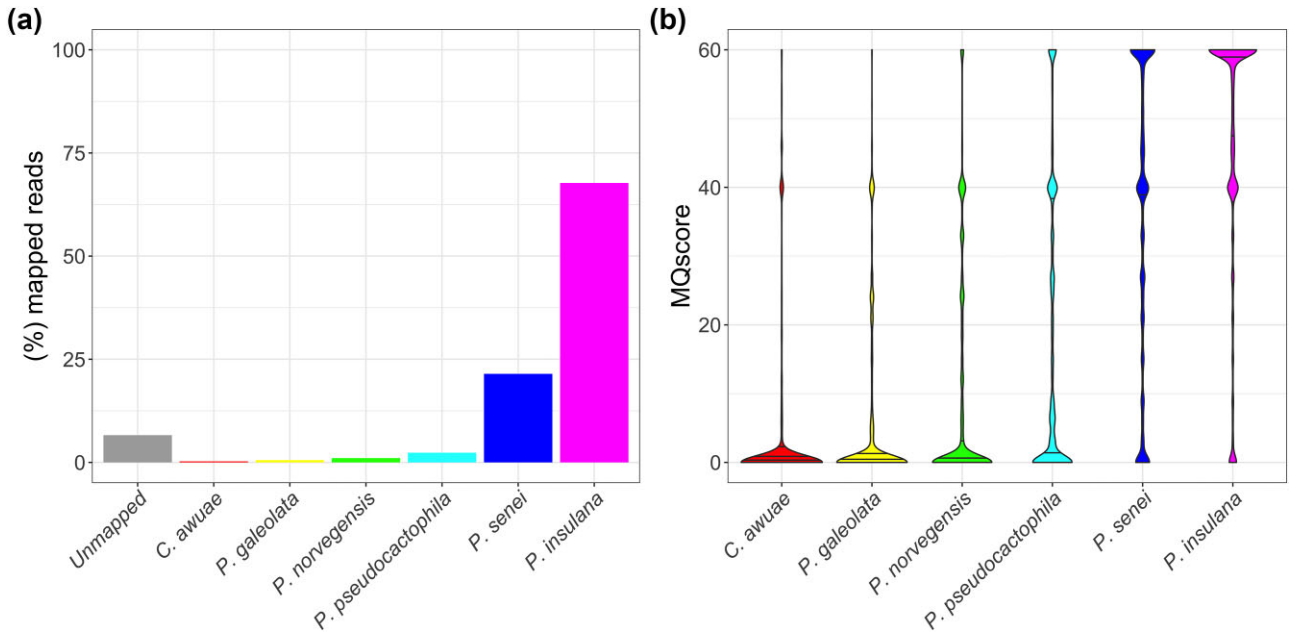

Figure 8. Hybridization analysis of *P. inconspicua*. Short reads from *P. inconspicua* (a) were aligned to several reference genomes of closely related *Pichia* species, demonstrating (b) higher mapping quality with *P. insulana*.

to only carry *MATα* ideomorphs (Fig. S1). Furthermore, if *P. insulana* is related to one of the parents of *P. inconspicua*, it cannot be the one that is shared with *P. alaskaensis* under the previously proposed scenario (Mixão et al. 2024) since *P. alaskaensis* mapped poorly to *P. insulana* (Fig. 7a). This result suggests some relationship of both *P. senei* and *P. insulana* to the parental lineages of *P. inconspicua*, but the precise scenario that gave rise to these relationships remains unclear, especially given how low the ANI values are for *P. senei* in all cases.

Heterozygosity and inbreeding estimates hint at genetic isolation for some lineages but are not directly matched with ploidy

In contrast to the interspecies hybrid lineages, where high heterozygosity and low inbreeding coefficients reflect their hybrid origins, *P. pseudocactophila*, *P. galeolata*, *P. insulana*, *P. norvegensis*, and *P. senei* showed low heterozygosity and high inbreeding coefficients (Table 1). Notably, these patterns were not necessarily reflected in their inferred ploidy level. Although the hybrid lineages were either diploid (*P. cactophila* and *P. inconspicua*) or triploid (*P. alaskaensis*), and many of the nonhybrids appeared to be haploid, the nonhybrid lineages of *P. insulana* and *P. norvegensis* appeared to be diploid (nQuire histotest $r^2 = 0.99$ and $r^2 = 0.93$, respectively) and yet had low levels of heterozygosity and high inbreeding coefficients like haploids. These data suggest that species, such as *P. senei*, are mainly haplontic, are relatively genetically isolated, and may exist in ecological scenarios that involve smaller populations. The diploid *P. insulana* and *P. norvegensis* lineages also appeared to be similarly genetically isolated and, while ploidy has been suggested to be clinically significant in closely related pathogenic *Pichia* species (Mixão et al. 2019), it remains unclear why these two species have maintained an apparent diploid state. However, it is notable that *P. insulana* was isolated from necrotic cactus tissue (Ganter et al. 2010), while *P. norvegensis* has been isolated from both human and environmental sources (Mixão et al. 2024), indicating that they may experience some of the same ecological pressures that select for diploidy in the pathogenic lineages.

Conclusions and future prospects for the *P. cactophila* species complex

Overall, *P. senei* is the most recently discovered member of the *P. cactophila* species complex and appears to represent another nonhybrid lineage that may be closely related to a yet-unknown parental lineage to *P. inconspicua*. Nonetheless, given the sequence divergence, *P. senei* is not likely to be a direct ancestor of *P. inconspicua*. More microbial sampling efforts, particularly in bromeliad phytotelmata and necrotic cactus tissues, should identify additional representatives of the members of this species complex and will aid population genomic analyses to understand the ecological history of these lineages beyond what is reflected in the genomes of a handful of isolates. Furthermore, additional sampling of diverse human-associated and environmental sources is likely to identify additional members of this species complex and further elucidate the complicated patterns of hybridization observed among them.

Description of *P. senei* K. O. Barros, J. Al-Oboudi, F.M.P. Sousa, L. F. D. Freitas, T. M Batista, A. R. O. Santos, P. B. Morais, Sampaio, Lachance, Hittinger, and *Rosa* sp. nov.

Pichia senei (se.ne´i. N.L. gen. masc. n. *senei*, in honor of Fábio de Melo Sene, in recognition of his contributions to the studies of genetics, ecology, and taxonomy of *Drosophila* in Brazil).

In 2% glucose—0.5% yeast extract broth after 3 days at 25°C, cells are spherical to ellipsoid and ~2–3 × 4–7 μm (Fig. 3). True hyphae or pseudohyphae are not formed. Colonies are white, convex, smooth, butyrous, and with a glistening surface. Ascospores were produced one to four hat-shaped formed on malt 5% and acetate McClary agars after 3 days at 25°C. Species is homothallic. Glucose is not fermented. Glucose, ethanol, glycerol, DL-lactate, succinate, and D-glucosamine are assimilated. Galactose, inulin,

sucrose, raffinose, melibiose, galactose, lactose, trehalose, maltose, melezitose, soluble starch, cellobiose, salicin, L-sorbose, L-rhamnose, D-xylose, L-arabinose, D-arabinose, D-ribose, methanol, erythritol, ribitol, galactitol, D-mannitol, D-glucitol, *myo*-inositol, citrate, D-gluconate, *N*-acetyl-D-glucosamine, hexadecane, xylitol, 2-propanol, and ethyl acetate are not assimilated. Lysine is utilized as sole nitrogen source but not nitrate or nitrite. Growth on amino acid-free medium is positive. Growth at 37°C and 40°C is positive. Growth in the presence of 0.01% cycloheximide is positive. Growth on YM agar with 10% sodium chloride and 5% glucose is negative. Growth in the presence of 50% glucose and 1% acetic acid is negative. Starch-like compounds are not produced. Production of acetic acid is positive. Urease activity is negative. Diazonium Blue B reaction is negative. The habitat is cactus rots and phytotelmata of bromeliads in Cerrado ecosystems of northern Brazil. The holotype is CBS 16311, and it is preserved in a metabolically inactive state in the CBS Yeast Collection of the Westerdijk Fungal Diversity Institute, Utrecht, the Netherlands. It was isolated from phytotelmata of the bromeliad *B. karatas* in the Cerrado ecosystem of Aurora do Tocantins, Tocantins, Brazil. An isotype of *P. senei* sp. nov. is deposited in the Collection of Microorganisms and Cells of Federal University of Minas Gerais (Coleção de Microrganismos e Células da Universidade Federal de Minas Gerais, UFMG), Belo Horizonte, Minas Gerais, Brazil, as strain UFMG-CM-Y531. The Mycobank number is MB 858723. The GenBank accession number for the sequences of the ITS-5.8S region and D1/D2 domains of the rRNA gene determined in this study is KT373981. The Whole Genome Shotgun project of strain *P. senei* UFMG-CM-Y531 has been deposited at DDBJ/ENA/GenBank under accession JAZAQD000000000. The version described in this paper is version JAZAQD000000000.

Supplementary data

Supplementary data is available at *FEMSYR Journal* online.

Conflict of interest: The authors declare that there is no conflict of interest.

Funding

This work is part of the project "INCT Yeasts: Biodiversity, preservation and biotechnological innovation", funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, grant number 406564/2022-1. This work was also funded by CNPq process numbers 0457499/2014-1, 141586/2017-6, 313088/2020-9, and 408733/2021, Fundação do Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG, process numbers APQ-01525-14, APQ-02552-15, and APQ-03071–17), the Natural Science and Engineering Research Council of Canada (MAL), and by FCT (Portugal) through grants UIDB/04378/2020 and UIDP/04378/2020 (UCIBIO), and LA/P/0140/2020 (i4HB). Research in the Hittinger Lab is funded by the National Science Foundation (DEB-2110403), USDA National Institute of Food and Agriculture (Hatch Project 7005101), in part by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE–SC0018409), and an H. I. Romnes Faculty Fellowship (Office of the Vice Chancellor for Research with funding from the Wisconsin Alumni Research Foundation).

Data availability

The raw data supporting the findings have been deposited to NCBI's SRA under the BioProject accession PRJNA1061143.

References

Altschul SF, Gish W, Miller W *et al.* Basic Local Alignment Search Tool. *J Mol Biol* 1990;**215**:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.

Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77. https://doi.org/10.1089/cmb.2012.0021.

Barros KO, Batista TM, Soares RC *et al. Spathaspora marinasilvae* sp. nov., a xylose-fermenting yeast isolated from galleries of passalid beetles and rotting wood in the Amazonian rainforest biome. *Yeast* 2024;**41**:437–47. https://doi.org/10.1002/yea.3966.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Caldwell AT, Gabaldón T, Mixão V *et al.* An inconspicuous identification: isolation and identification of a novel *Pichia* species presenting as fungemia following cardiac surgery. *Int J Infect Dis* 2024;**143**:107040. https://doi.org/10.1016/j.ijid.2024.107040.

Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: architecture and applications. *BMC Bioinf* 2009;**10**:421. https://doi.org/10.1186/1471-2105-10-421.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;**25**:1972–3. https://doi.org/10.1093/bioinformatics/btp348.

Danecek P, Auton A, Abecasis G *et al.* The variant call format and VCFtools. *Bioinformatics* 2011;**27**:2156–8. https://doi.org/10.1093/bioinformatics/btr330.

De Carvalho F, De Souza FA, Carrenho R *et al.* The mosaic of habitats in the high-altitude Brazilian rupestrian fields is a hotspot for arbuscular mycorrhizal fungi. *Appl Soil Ecol* 2012;**52**:9–19. https://doi.org/10.1016/j.apsoil.2011.10.001.

Dos Santos VL Silva UC Santos EH, Resende AA *et al.* Exploring the mycobiota of bromeliads phytotelmata in Brazilian Campos Rupestres. *Braz J Microbiol* 2023;**54**:1885–97.

Edgar RC. MUSCLE RE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004;**5**:113. https://doi. org/10.1186/1471-2105-5-113.

Félix CR, da Silva Nascimento BE, Valente P *et al.* Different plant compartments, different yeasts: the example of the bromeliad phyllosphere. *Yeast* 2022;**39**:363–400.

Freitas LF, Batista TM, Santos AR *et al.* Yeast communities associated with cacti in Brazil and the description of *Kluyveromyces starmeri* sp. nov. based on phylogenomic analyses. *Yeast* 2020;**37**:625–37. https://doi.org/10.1002/yea.3528.

Ganter PF, Cardinali G, Boundy-Mills K. *Pichia insulana* sp. nov., a novel cactophilic yeast from the Caribbean. *Int J Syst Evol Microbiol* 2010;**60**:1001–7. https://doi.org/10.1099/ijs.0.014258-0.

Ganter PF, Morais PB, Rosa CA. Yeasts in cacti and tropical fruit. In: *Yeasts in Natural Ecosystems: Diversity*. Cham: Springer, 2017, 225–64. https://doi.org/10.1007/978-3-319-62683-3_8.

Gomes FCO, Gomes FC, Safar SV *et al.* The diversity and extracellular enzymatic activities of yeasts isolated from water tanks of *Vriesea minarum*, an endangered bromeliad species in Brazil, and the description of *Occultifur brasiliensis* fa, sp. nov. *Antonie Van Leeuwenhoek* 2015;**107**:597–611. https://doi.org/10.1007/s10482-014-0356-4.

Gonçalves C, Harrison MC, Steenwyk JL *et al.* Diverse signatures of convergent evolution in cactus-associated yeasts. *PLoS Biol* 2024;**22**:e3002832. https://doi.org/10.1371/journal.pbio.3002832.

Guitard J, Atanasova R, Brossas JY *et al. Candida inconspicua* and *Candida norvegensis*: new insights into identification in relation to

sexual reproduction and genome organization. *J Clin Microbiol* 2015;**53**:1655–61. https://doi.org/10.1128/JCM.02913-14.

Huson DH, Bryant D. The SplitsTree App: interactive analysis and visualization using phylogenetic trees and networks. *Nat Methods* 2024;**21**:1773–4. https://doi.org/10.1038/s41592-024-02406-3.

James TY, Stajich JE, Hittinger CT *et al.* Toward a fully resolved fungal tree of life. *Annu Rev Microbiol* 2020;**74**:291–313. https://doi.org/10.1146/annurev-micro-022020-051835.

Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with MAFFT. In: *Bioinformatics for DNA Sequence Analysis*. Vol. **537**. Totowa, NJ: Humana Press, 2009, 39–64. https://doi.org/10.1007/978-1-59745-251-9_3.

Kumar S, Stecher G, Mega KT. Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;**33**:1870–4. https://doi.org/10.1093/molbev/msw054.

Kurtzman CP, Fell JW, Boekhout T *et al.* Methods for isolation, phenotypic characterization and maintenance of yeasts. In: *The Yeasts*. Amsterdam: Elsevier, 2011, 87–110.

Kurtzman CP, Robnett CJ. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenhoek* 1998;**73**:331–71. https://doi.org/10.1023/A:1001761008817.

Lachance MA, Bowles JM, Starmer WT *et al. Kodamaea kakaduensis* and *Candida tolerans*, two new ascomycetous yeast species from *Australian Hibiscus* flowers. *Can J Microbiol* 1999;**45**:172–7. https://doi.org/10.1139/w98-225.

Lachance MA, Lee DK, Hsiang T. Delineating yeast species with genome average nucleotide identity: a calibration of ANI with haplontic, heterothallic *Metschnikowia* species. *Antonie Van Leeuwenhoek* 2020;**113**:2097–106. https://doi.org/10.1007/s10482-020-01480-9.

Langdon QK, Peris D, Kyle B *et al.* Sppider: a species identification tool to investigate hybrid genomes with high-throughput sequencing. *Mol Biol Evol* 2018;**35**:2835–49.

Lee I, Ouk Kim Y, Park SC *et al.* OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 2016;**66**:1100–3. https://doi.org/10.1099/ijsem.0.000760.

Lemoine F, Correia D, Lefort V *et al.* NGPhylogeny. Fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res* 2019;**47**:W260–5. https://doi.org/10.1093/nar/gkz303.

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, 2013. https://doi.org/10.48550/arXiv.1303.3997.

Manni M, Berkeley MR, Seppey M *et al.* BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 2021;**38**:4647–54. https://doi.org/10.1093/molbev/msab199.

Mikheenko A, Prjibelski A, Saveliev V *et al.* Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 2018;**34**:i142–50. https://doi.org/10.1093/bioinformatics/bty266.

Minh BQ, Schmidt HA, Chernomor O *et al.* IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;**37**:1530–4. https://doi.org/10.1093/molbev/msaa015.

Mixão V, Hansen AP, Saus E *et al.* Whole-genome sequencing of the opportunistic yeast pathogen *Candida inconspicua* uncovers its hybrid origin. *Front Genet* 2019;**10**:383. https://doi.org/10.3389/fgene.2019.00383.

Mixão V, Saus E, Hagen F *et al.* A yeast love triangle: multiple hybridizations shape genome evolution in the *Pichia cactophila* species complex. *Mycosphere* 2024;**15**:1083–109. https://doi.org/10.5943/mycosphere/15/1/9.

Morais PB, Morais PB, de Sousa FM *et al.* Yeast in plant phytotelmata: is there a "core" community in different localities of rupestrian savannas of Brazil?. *Braz J Microbiol* 2020;**51**:1209–18. https://doi.org/10.1007/s42770-020-00286-1.

O'Connor BD, Van der Auwera GA. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. Sebastopol, CA: O'Reilly Media, 2020.

Opulente DA, LaBella AL, Harrison MC *et al.* Genomic factors shape carbon and nitrogen metabolic niche breadth across *Saccharomycotina* yeasts. *Science* 2024;**384**:eadj4503. https://doi.org/10.1126/science.adj4503.

Opulente DA, Langdon QK, Jarzyna M *et al.* Taxogenomic analysis of a novel yeast species isolated from soil, *Pichia galeolata* sp. nov. *Yeast* 2023;**40**:608–15. https://doi.org/10.1002/yea.3905.

Rezende D, Lopes Filho DR, Mantovano T *et al.*. Bromeliad phytotelmata: the first scientometric study. *Limnetica* 2023;**42**:143–54.

Rosa CA, Morais PB, Hagler AN *et al.* Yeast communities of the cactus *Pilosocereus arrabidae* and associated insects in the sandy coastal plains of southeastern Brazil. *Antonie Van Leeuwenhoek* 1994;**65**:55–62. https://doi.org/10.1007/BF00878279.

Shen XX, Opulente DA, Kominek J *et al.* Tempo and mode of genome evolution in the budding yeast subphylum. *Cell* 2018;**175**:1533–1545.e20. https://doi.org/10.1016/j.cell.2018.10.023.

Sousa FM, Morais PB, Lachance MA *et al. Hagleromyces* gen. nov., a yeast genus in the Saccharomycetaceae, and description of *Hagleromyces aurorensis* sp. nov., isolated from water tanks of bromeliads. *Int J Syst Evol Microbiol* 2014;**64**:2915–8. https://doi.org/10.1099/ijs.0.063883-0.

Weib CL, Pais M, Cano LM *et al.* nQuire: a statistical framework for ploidy estimation using next generation sequencing. *BMC Bioinf* 2018;**19**:122.

White TJ, Bruns T, Lee SJ *et al.* Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Prot Guide Methods Appl* 1990;**18**:315–22.

Received 24 April 2025; **revised** 11 July 2025; **accepted** 8 August 2025