Received: 7 September 2023 | Revised: 13 December 2023 | Accepted: 19 December 2023

DOI: 10.1002/yea.3922

RESEARCH ARTICLE

Fantastic Yeast

Yeast WILEY

Reinstatement of the fission yeast species *Schizosaccharomyces versatilis* Wickerham et Duprat, a sibling species of *Schizosaccharomyces japonicus*

Michael Brysch-Herzberg[1] | Guo-Song Jia[2] | Matthias Sipiczki[3] | Martin Seidel[1] | Wen-Cai Zhang[2] | Li-Lin Du[2,4]

[1]Laboratory for Wine Microbiology, Department International Business, Heilbronn University, Heilbronn, Germany

[2]National Institute of Biological Sciences, Beijing, China

[3]Department of Genetics and Applied Microbiology, University of Debrecen, Debrecen, Hungary

[4]Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing, China

Correspondence

Michael Brysch-Herzberg, Heilbronn University of Applied Sciences, Max-Planck-Str. 39, Heilbronn 74081, Germany.
Email: michael.brysch-herzberg@hs-heilbronn.de

Matthias Sipiczki, Department of Genetics and Applied Microbiology, University of Debrecen, Debrecen 4032, Hungary.
Email: gecela@post.sk

Funding information

Ministry of Science and Technology of the People's Republic of China; Deutsche Bundesstiftung Umwelt DBU (German Federal Environmental Foundation),
Grant/Award Number: 34053/01-32

Abstract

Schizosaccharomyces japonicus Yukawa et Maki (1931) and *Schizosaccharomyces versatilis* Wickerham et Duprat (1945) have been treated as varieties of *S. japonicus* or as conspecific, based on various approaches including mating trials and nDNA/nDNA optical reassociation studies. However, the type strains of *S. japonicus* and *S. versatilis* differ by five substitutions (99.15% identity) and one 1-bp indel in the sequences of the D1/D2 domain of the 26S rRNA gene, and 23 substitutions (96.3% identity) and 31-bp indels in the sequences of internal transcribed spacer (ITS) of rRNA, suggesting that they may not be conspecific. To reassess their taxonomic status, we conducted mating trials and whole-genome analyses. Mating trials using the type strains showed a strong but incomplete prezygotic sterility barrier, yielding interspecies mating products at two orders of magnitude lower efficiency than intraspecies matings. These mating products, which were exclusively allodiploid hybrids, were unable to undergo the haplontic life cycle of the parents. We generated chromosome-level gap-less genome assemblies for both type strains. Whole genome sequences yielded an average nucleotide identity (ANI) of 86.4%, indicating clear separation of *S. japonicus* and *S. versatilis*. Based on these findings, we propose the reinstatement of *S. versatilis* as a distinct species (holotype strain: CBS 103^T and ex-types: NRRL Y-1026, NBRC 1607, ATCC 9987, PYCC 7100; Mycobank no.: 847838).

KEYWORDS

genomic rearrangements, reproductive incompatibility, *Schizosaccharomyces versatilis*, species delimitation

Michael Brysch-Herzberg, Matthias Sipiczki, and Guo-Song Jia are equal first author of this study.

1 | INTRODUCTION

Because of their phylogenetic position within the *Taphrinomycotina* (formerly *Archiascomycota*), fission yeasts occupy a unique phylogenetic position among yeasts (Ebersberger et al., 2012; Liu et al., 2009; Sipiczki, 1995, 2000). Currently, six fission yeast species in the genus *Schizosaccharomyces* are recognized: *Schizosaccharomyces pombe* (Lindner, 1893), *Schizosaccharomyces octosporus* (Beijerinck, 1894), *Schizosaccharomyces japonicus* (Yukawa & Maki, 1931), *Schizosaccharomyces cryophilus* (Helston et al., 2010), *Schizosaccharomyces osmophilus* (Brysch-Herzberg et al., 2019) and *Schizosaccharomyces lindneri* (Brysch-Herzberg et al., 2023). An overview of the long lasting discourse concerning species delimitation and phylogeny in the genus is given by Brysch-Herzberg et al. (2019). Proposal were made to divide the genus into three genera (Kudrjawzew, 1960; Yamada & Banno, 1987). However, as these proposals were not accepted by later authors, we will name the species according to Vaughan-Martini and Martini (2011) as given in the book 'The Yeast, a taxonomic study'.

The genus *Schizosaccharomyces* includes *S. pombe*, one of the best studied eukaryotic organisms (Fantes & Hoffman, 2016; Hoffman et al., 2015). Another fission yeast species, *S. japonicus*, is gaining more attention in studies about cell biology, genetics, and evolution (Rutherford et al., 2022). *S. japonicus* was described by Yukawa and Maki (1931) on the basis of a strain isolated from strawberry wine in Japan. Wickerham and Duprat (1945) described *S. versatilis* based on a strain isolated from home-canned grape juice, in Michigan, USA. In the 1952 edition of 'The yeast, a taxonomic study' (Lodder & Kreger-van Rij, 1952), *S. japonicus* is listed as synonym of *S. versatilis* although *S. japonicus* was validly described earlier. The authors explicitely stated that they felt unsure about treating both as conspecific.

The initial standard descriptions of both species only differed in a positive starch test with iodine in the ascospore cell wall for *S. versatilis* and the fermentation of galactose by *S. japonicus*. Both test results could not be reproduced by Lodder and Kreger-van Rij (1952). Thus, the type strains were phenotypically indistinguishable. Mating trials with auxotrophic mutants of *S. japonicus* and *S. versatilis* led to the formation of prototrophic colonies (Sipiczki et al., 1982) but the nature of these colonies was not investigated at the time. Because prototrophic colonies were produced at a much lower efficiency in interspecies mating compared to intraspecies mating, the authors concluded that *S. versatilis* should be treated as a variety of *S. japonicus*. This view was adopted by Yarrow (1984) in the 1984 version of 'The yeast, a taxonomic study'. Later, Martini (1991) rejected this assesment and concluded from the results of nDNA/nDNA optical reassociation studies that *S. versatilis* is conspecific with *S. japonicus*. This view is generally accepted to this day (Vaughan-Martini & Martini, 2011).

We here present new data from mating trials as well as from whole-genome sequence comparisons which lead us to the conclusion that *S. versatilis* and *S. japonicus* are separate species. We propose the validation and reinstatement of *S. versatilis*.

Take-away

- It is shown that an almost complete reproductive barrier exists between *S. japonicus* and *S. versatilis*.
- Gapless genomes of the type strains of *S. japonicus* and *S. versatilis* are reported here.
- Genome-based analysis estimates that *S. japonicus* and *S. versatilis* diverged more than 10 million years ago.
- *S. versatilis* is reinstated and a valid species description is provided.

2 | MATERIALS AND METHODS

2.1 | Analysis of rRNA gene repeat sequences

Sequences of the rRNA gene repeat were retrieved from Genbank or the culture institutes that host the strains (Table 1). For four strains the ITS-region was sequenced as described before (Brysch-Herzberg et al., 2019). Sequence alignments were generated in MAFFT (Katoh et al., 2019) online (URL: https://mafft.cbrc.jp/alignment/server/). For aligning the D1/D2 region sequences the 'strategy' E-INS-i and for the ITS region sequences the L-INS-i was chosen. Except of the 'strategy' the presettings were left unaltered. The resulting alignments were trimmed in Jalview (Waterhouse et al., 2020). Mega X was used for phylogenetic analysis (Kumar et al., 2018). Before analysis all sites with gaps were deleted from the alignment ('complete deletion' option). The Tamura-Nei model (Tamura & Nei, 1993) was applied as substitution model. Phylogenetic trees were inferred for the D1/D2 domain and the ITS region. For each locus a tree was inferred by employing the Neighbor joining algorithm (Felsenstein, 1981) and alternatively the maximum likelihood method (Kashyap & Subas, 1974; Neyman, 1971), separately. Statistical support was estimated with the bootstrap approach (Felsenstein, 1985) from 1000 iterations. The Tree was edited in TreeGraph 2 (Stöver & Müller, 2010). In TreeGraph 2 the tree was rooted with the outgroup species.

2.2 | Sequencing the genomes of the type strains of *S. versatilis* and *S. japonicus*

The ex-type strain NRRL Y-1026 of *S. versatilis* (=CBS 103 = NBRC 1607 = ATCC 9987 = PYCC 7100; obtained from NRRL in 2017) and the ex-type strain NBRP FY16936 of *S. japonicus* (=CBS 354 = NBRC 1609 = ATCC10660 = NRRL Y-1361 = yFS275; obtained from NBRP/YGRC in 2008) were grown on solid YES plates at 30°C. Cells were collected and washed before being frozen using liquid nitrogen. To sequence the genomes of the *S. versatilis* and *S. japonicus* type strains using Oxford Nanopore ultra-long DNA sequencing technology, we dispatched the frozen cells to GrandOmics Biosciences for high molecular weight genomic DNA extraction, sequencing library

TABLE 1 List of yeast strains, auxotrophic mutants and DNA sequence Genbank accession no.

Species	Strain	Origin	D1/D2 domain sequence accession no.	ITS region sequence accession no.	References (origin of the strain)
Schizosaccharomyces japonicus	CBS 354^T wildtype	Japan	U94943.1	NR_121199.1	Yukawa and Maki (1931)
S. japonicus	CBS 354^T 7-83 ade$^-$	This study	-	-	-
S. japonicus	CBS 354^T 7-106 lys$^-$	This study	-	-	-
S. japonicus	NBRC 1646	Japan	NBRC	OR807356	-
S. japonicus	NBRC 1712	Japan	NBRC	OR807355	Hayashibe et al. (1974)
S. japonicus	PYCC 6842	Japan	KT767196.1	KT767197.1	-
Schizosaccharomyces versatilis	CBS 103^T wildtype	USA, Michigan	CBS	CBS	Wickerham and Duprat (1945)
S. versatilis	CBS 103^T 8-5 ura$^-$	This study	-	-	-
S. versatilis	CBS 103^T 8-8 ade$^-$	This study	-	-	-
S. versatilis	CBS 103^T 8-22 ura$^-$	This study	-	-	-
S. versatilis	CBS 103^T 8-28 ura$^-$	This study	-	-	-
S. versatilis	CBS 5679	USA, California	CBS	CBS	Phaff et al. (1964)
S. versatilis	CBS 2629	Portugal	CBS	CBS	-
S. versatilis	CBS 7116	North Italy	CBS	CBS	Delfini et al. (1983)
S. versatilis	CBS 8543	Netherlands	CBS	CBS	-
S. versatilis	CBS 15803	Germany, Ruwer valley	MK690476	OR807353	Brysch-Herzberg et al. (2019)
S. versatilis	CBS 15809	France, Provence	MK690482	OR807354	Brysch-Herzberg et al. (2019)

Abbreviations: CBS, Westerdijk Fungal Biodiversity Institute, Utrecht, the Netherlands; NBRC, National Institute of Technology and Evaluation (NITE) Biological Resource Centre, Chiba, Japan; PYCC, Portuguese Yeast Culture Collection, Lisbon, Portugal.

construction and single-molecule long-read sequencing on the PromethION platform under the ultra-long sequencing mode. The resulting fast5 signals were decoded into fastq format. Low-quality reads were discarded, and adapter sequences were trimmed. A total of 94.2 Gb (for *S. versatilis*) and 79.8 Gb (for *S. japonicus*) of reads were retained. These reads have been deposited at the NCBI SRA database under the accession numbers SRR25208122 (for *S. versatilis*) and SRR25243140 (for *S. japonicus*) (Supporting Information S7: Table 1).

Additionally, the genomes of *S. versatilis* and *S. japonicus* type strains were sequenced using the PacBio highly accurate long-reads (HiFi reads) sequencing technology. We sent the frozen cells to Frasergen Bioinformatics for high molecular weight genomic DNA extraction, sequencing library preparation, and single-molecule long-read sequencing on the Sequel II platform in the circular consensus sequencing (CCS) mode. DeepConsensus v1.1, a deep learning method for correcting errors in PacBio CCS data, was run to generate a total of 3397.7 Mb (for *S. versatilis*) and 887.43 Mb (for *S. japonicus*) of HiFi reads (Delaluna et al., 2023). These HiFi reads have been deposited at the NCBI SRA database under the accession numbers SRR25243141 (*S. japonicus*) and SRR25208123 (*S. versatilis*) (Supporting Information S7: Table 1).

We also sequenced the genomes of the type strains of *S. versatilis* and *S. japonicus* using Illumina short-read sequencing technology. Illumina sequencing library construction using home-made Tn5 transposase was performed as described previously (Tao et al., 2019). The resulting libraries were sent to Novogene for paired-end sequencing on the Illumina NovaSeq. 6000 system, generating 2 × 150 bp read pairs. Approximately 2.5 Gb (for *S. versatilis*) and 3.00 Gb (for *S. japonicus*) raw Illumina read data were obtained. These reads have been deposited at the NCBI SRA database under the accession number SRR24463226 (for *S. versatilis*) and SRR25243143 (for *S. japonicus*) (Supporting Information S7: Table 1). We performed preprocessing and quality assessment of the Illumina sequencing data using fastp v0.2.0 (https://github.com/OpenGene/fastp) with the parameter '--length_required 70' (Chen et al., 2018). About 2.3 Gb (for *S. versatilis*) and 2.9 Gb (for *S. japonicus*) Illumina read data remained after fastp processing. These reads have been deposited at the NCBI SRA database under the accession numbers SRR24463225 (for *S. versatilis*) and SRR25243142 (for *S. japonicus*) (Supporting Information S7: Table 1).

2.3 | De novo assembly of the *S. japonicus* genome

We employed nextdenovo v2.5.0 (https://github.com/Nextomics/NextDenovo) to perform de novo assembly using Oxford Nanopore reads longer than 200 kb (about 211× coverage based on a genome size of 12 Mb) (Hu, Wang, Sun, et al., 2023) (Figure S1a). A total of four contigs were produced with a combined length of 15,372,310 bp. We performed three rounds of polishing using nextpolish v1.4.1 (https://github.com/Nextomics/NextPolish) and Oxford Nanopore reads longer than 100 kb (about 1569× coverage) (Hu et al., 2020). The assembly was then further polished for three rounds using nextpolish v1.4.1 and PacBio HiFi reads (about 74× coverage). Finally, both PacBio HiFi reads and fastp-processed Illumina reads were used to polish the assembly by running nextpolish2 (https://github.com/Nextomics/NextPolish2) for three rounds (Hu, Wang, Liang, et al., 2023). The resulting genome assembly contains four contigs (ctg00, ctg10, ctg20 and ctg30) and has a total length of 15,401,221 bp. Additionally, we assembled the PacBio HiFi reads using hifiasm v0.16.5 (https://github.com/chhylp123/hifiasm) to generate a total of 117 contigs with a total length of 17,348,380 bp (Cheng et al., 2021).

To facilitate visual inspection of the polished nextdenovo-based genome assembly, we retrieved from JaponicusDB (https://www.japonicusdb.org/) the annotations of protein-coding genes in the reference genome of *S. japonicus* and used liftoff v1.6.3 (https://github.com/agshumate/Liftoff) to transfer the annotations to our assembly (Rutherford et al., 2022; Shumate & Salzberg, 2021). We also annotated tRNA and rRNA genes using tRNAscan-SE v2.0.12, (https://github.com/UCSC-LoweLab/tRNAscan-SE) (Lowe & Chan, 2016) and barrnap v0.9 (Seemann, 2018), respectively. Subsequently, we imported the genome assembly with annotation information into Geneious Prime (Dotmatics) and manually inspected the assembly. Genome alignment between the newly assembled *S. japonicus* contigs and the reference genome of *S. japonicus* was conducted using D-Genies (https://dgenies.toulouse.inra.fr/) (Cabanettes & Klopp, 2018).

2.4 | Estimating copy number of tDNA-5S-rDNA arrays in *S. japonicus*

To estimate the copy number of cen2-, *bet5*- and ire1-adjacent tDNA-5S-rDNA arrays in *S. japonicus*, we conducted the following analyses. First, we split each tDNA-5S-rDNA array in the assembly into repeat units, with the first base following the tRNA-AlaAGC gene arbitrarily defined as the starting point of a repeat unit. Examining the repeat unit sequences showed that most of the repeat units within each tDNA-5S-rDNA array are identical. Then, we analysed ONT ultra-long reads that overlap with the tDNA-5S-rDNA arrays and confirmed that repeat units within each tDNA-5S-rDNA array exhibit homogeneity. We used PacBio HiFi reads to generate a consensus repeat unit sequence for each tDNA-5S-rDNA array by running ribotin (https://github.com/maickrau/ribotin). We then masked all sequences of the tDNA-5S-rDNA arrays in the assembly using bedtools maskfasta (Quinlan & Hall, 2010) and appended the consensus repeat unit sequences of the three tDNA-5S-rDNA arrays as three separate contigs. Subsequently, all the fastp-processed Illumina reads were mapped to the resulting genome assembly and the median read depth of the previously defined single-copy orthologous genes conserved across fission yeasts (Jia et al., 2023) was calculated and used as the baseline read depth of single-copy sequences. We then calculated the median read depth of each consensus repeat unit sequence and divided it by the baseline read

depth of single-copy sequences to obtain repeat copy number estimation. The estimation showed that there are approximately 99, 89 and 123 copies of repeat units in the cen2-adjacent, *bet5*-adjacent and ire1-adjacent tDNA-5S-rDNA arrays, respectively.

2.5 | Finalizing the genome assembly of *S. japonicus*

The complete genome assembly of *S. japonicus* was generated by combining the gapless contigs of chromosome 1 and chromosome 3 with a modelled chromosome 2 contig. This chromosome 2 contig was created by filling the gap between ctg30 and ctg00 within the chromosome 2 scaffold using the estimated number of consensus units of cen2-adjacent tDNA-5S-rDNA array. Additionally, we modelled the correct copy number of the *bet5*-adjacent and ire1-adjacent tDNA-5S-rDNA arrays using the consensus unit sequence obtained in the previous section. Specifically, we defined the boundary repeat unit in each type of array and replaced all internal repeat units within the boundary repeat units with tandem copies of consensus repeat units to match the estimated correct copy number. For the boundary between the *bet5*-adjacent and ire1-adjacent tDNA-5S-rDNA arrays, we utilized the sequence of a hifiasm-assembled HiFi-read-based contig (contig ptg12l) spanning the boundary (Figure S4a). We trimmed several outermost repeat units from contig ptg12l and placed the remaining portion of the contig between tandem copies of the consensus repeat unit sequences of the two arrays.

The final *S. japonicus* genome assembly contains three chromosomal contigs and has a total length of 16,829,384 bp (Supporting Information S7: Table 2). The genome assembly has been deposited at GenBank under the accession number GCA_032883015.1.

2.6 | De novo assembly of the *S. versatilis* genome

We used strategies similar to those described above to assemble the genome of the *S. versatilis* type strain. Using Oxford Nanopore reads longer than 200 kb (about 383× coverage based on a genome size of 12 Mb) to run nextdenovo v2.5.0, a total of two contigs with a combined length of 14,637,562 bp were assembled (Figure S2a). Three rounds of polishing using ONT reads >100 kb and three rounds of polishing using PacBio HiFi reads were performed using nextpolish v1.4.1. After further polishing using nextpolish2, we obtained an assembly that contains two contigs (ctg00 and ctg10) and has a total length of 14,705,892 bp. In addition, we assembled the PacBio HiFi reads using hifiasm to generate a total of 146 contigs with a total length of 22,409,863 bp. Protein-coding genes, tRNA genes, and rRNA genes in the polished nextdenovo-based genome assembly were annotated as described above. The genome alignment between *S. versatilis* contigs and *S. japonicus* genome assembly obtained in this study was conducted using D-Genies (https://dgenies.toulouse.inra.fr/) (Cabanettes & Klopp, 2018).

2.7 | Estimating copy number of tDNA-5S-rDNA arrays in *S. versatilis*

To estimate the copy number of cen2-, *bet5*- and ire1-adjacent tDNA-5S-rDNA arrays in *S. versatilis*, we conducted the following analyses. First, we split each tDNA-5S-rDNA array in the assembly into repeat units. Examining the repeat unit sequences showed that most of the repeat units within each tDNA-5S-rDNA array are identical. Then, we analysed ONT ultra-long reads that overlap with the tDNA-5S-rDNA arrays and confirmed that repeat units within each tDNA-5S-rDNA array exhibit homogeneity. We noticed that unlike the situation in the *S. japonicus* assembly, the repeat units in the *bet5*-adjacent tDNA-5S-rDNA array are highly similar to the repeat units in the ire1-adjacent tDNA-5S-rDNA (Figure S3b), differing only by four single nucleotide substitutions and a copy number difference of the 5S rDNAs. Consequently, we generated consensus sequences for different tDNA-5S-rDNA arrays using different strategies. For the cen2-adjacent array, we employed ribotin to generate a consensus repeat unit sequence from PacBio HiFi reads. For the *bet5*-adjacent and ire1-adjacent arrays, the consensus repeat unit sequences were generated by aligning the repeat units extracted from the genome assembly and obtaining a consensus sequence from the alignment. We then masked sequences of the tDNA-5S-rDNA arrays in the assembly and appended the consensus repeat unit sequences of the cen2-adjacent and the ire1-adjacent arrays as two separate contigs. After Illumina reads mapping, we calculated a baseline read depth of single-copy sequences as described above. The median read depth of the consensus repeat unit sequence of the cen2-adjacent array was divided by the baseline read depth of single-copy sequences to obtain the repeat copy number of the cen2-adjacent array. The median read depth of the consensus repeat unit sequence of the ire1-adjacent array, excluding the region that contains 5S rDNAs, was divided by the baseline read depth of single-copy sequences to obtain the combined repeat copy number of the *bet5*-adjacent and ire1-adjacent arrays. Finally, we determined the relative repeat copy number ratio of the *bet5*-adjacent array versus the ire1-adjacent array by calculating the mean allele frequency of the four single nucleotide substitutions. Our estimations showed that there are approximately 96, 133 and 35 copies of repeat units in the cen2-adjacent, *bet5*-adjacent and ire1-adjacent tDNA-5S-rDNA arrays, respectively.

2.8 | Finalizing the genome assembly of *S. versatilis*

To correct the repeat copy number errors of chromosome 2 and generate a complete genome assembly, we modelled all three types of tDNA-5S-rDNA arrays by replacing the nonboundary repeat units in each array with tandem copies of consensus repeat unit sequence so that the total number of repeat units matches the estimated copy number. For the left boundary of the cen2-adjacent array, we kept the first non-boundary repeat unit because it harbours a

retrotransposon. For the boundary between the *bet5*-adjacent and ire1-adjacent arrays, we used the sequence of a hifiasm-assembled HiFi-read-based contig (contig ptg101l) spanning the boundary (Figure S4b). We trimmed the outermost partial repeat units from contig ptg101l and placed the remaining portion of the contig between tandem copies of the consensus repeat unit sequences of the two arrays.

The final *S. versatilis* genome assembly contains three chromosomal contigs and has a total length of 15,482,186 bp (Supporting Information S7: Table 2). The genome assembly has been deposited at GenBank under the accession numbers GCA_032882995.1.

2.9 | Comparison of the consensus repeat unit sequences of the tDNA-5S-rDNA arrays

To assess the sequence similarity between the consensus repeat unit sequences of different tDNA-5S-rDNA arrays, we performed multiple sequence alignment using MAFFT v7.149b with the E-INS-i algorithm (Katoh & Standley, 2013) and performed sliding window analysis using SimPlot++ v1.3 (Samson et al., 2022) with the following parameters: Distance model = identity, Window length = 40, Step = 5, Strip gap = No Strip Gap, and Plot refresh rate = every window.

2.10 | Calculation of the average nucleotide identity (ANI) values

We employed the OrthoANIu (OAU) tool (v1.2, http://www.ezbiocloud.net/tools/orthoaniu) (Yoon et al., 2017) to calculate pairwise ANI values for all combinations of the type strains of fission yeast species. The genome assemblies used as input for the OrthoANIu analysis are: *S. pombe* type strain (CBS 356^T = JB861) assembly downloaded from https://figshare.com/articles/dataset/SGA_de_novo_assemblies_of_all_wild_strains/3978279 (Jeffares et al., 2015), *S. octosporus* type strain (CBS 371^T = ATCC 4206 = yFS286) assembly and *S. cryophilus* type strain (CBS 11777^T = OY26) assembly downloaded from http://bifx-core.bio.ed.ac.uk/~ptong/genome_assembly/ (Tong et al., 2019), *S. osmophilus* type strain (CBS 15793^T) assembly downloaded from the NCBI GenBank database (https://www.ncbi.nlm.nih.gov/assembly/GCA_027921745.1/) (Jia et al., 2023), *S. lindneri* type strain (CBS 18203^T = SZ1623-FG-A) assembly downloaded from the NCBI GenBank database (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_029532395.1/) (Brysch-Herzberg et al., 2023), and *S. japonicus* and *S. versatilis* type strain genome assemblies generated in this study.

2.11 | Species tree inference and divergence time estimation

For genome-based phylogenetic analysis, we employed single-copy BUSCO genes as described previously (Jia et al., 2023). First, the presence/absence status of 1315 predefined single-copy orthologs in *Ascomycota* (the ascomycota_odb9 gene set) was determined by submitting the newly generated *S. japonicus* and *S. versatilis* genome assemblies to BUSCO v3.0.2 (https://busco.ezlab.org/) (Supporting Information S7: Table 3) (Simão et al., 2015). Subsequently, we selected 1,045 'complete and single-copy' BUSCO genes present in *S. versatilis*, the six previously recognized fission yeast species (*S. japonicus*, *S. pombe*, *S. octosporus*, *S. lindneri*, *S. osmophilus*, and *S. cryophilus*), and the outgroup species *Saitoella complicata* for phylogenetic analysis. Alignment for each BUSCO gene was generated separately using MAFFT v7.475 with parameters '--thread 4 --auto --maxiterate 1,000' (Katoh & Standley, 2013), and the resulting multiple sequence alignment was trimmed using trimAL v1.4. rev15 (http://trimal.cgenomics.org/) with parameters '-gappyout' (Capella-Gutiérrez et al., 2009). Alignments were concatenated using catsequences (https://github.com/ChrisCreevey/catsequences) (https://zenodo.org/record/4409153#.ZAHewnZBybg). We employed the IQ-TREE built-in ModelFinder (IQ-TREE v 2.0.3, https://github.com/Cibiv/IQ-TREE) with the options '-m TESTONLY -nt 1' to determine the best-fitting model for each trimmed BUSCO gene alignment (Kalyaanamoorthy et al., 2017). For the majority of BUSCO genes (588 out of 1045 BUSCO genes, 56.3%), the best-fitting model is the 'LG + G4' model. Consequently, we performed concatenation-based species tree inference by running IQ-TREE v2.0.3 with the parameters '-m LG + G4 -alrt 1,000 -bb 1,000' (Minh et al., 2020).

To estimate the divergence time between *S. versatilis* and *S. japonicus*, we employed the RelTime method in MEGA11 (Tamura et al., 2021), using *S. complicata* as the outgroup in the analysis. We used the following divergence time calibration nodes: the *S. japonicus*–*S. pombe* split (207.2 million years ago), the *S. pombe*–*S. octosporus* split (108.2 million years ago), the *S. octosporus*–*S. cryophilus* split (29.4 million years ago) (Shen et al., 2020), the *S. octosporus*–*S. osmophilus* split (15.7 million years ago) (Jia et al., 2023), and the *S. octosporus*–*S. lindneri* split (5.1 million years ago) (Brysch-Herzberg et al., 2023). We visualized the resulting time-calibrated phylogenetic tree in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and removed the branch of the outgroup species *S. complicata* manually in Adobe Illustrator.

2.12 | Visualization of genomic features and interspecific chromosomal rearrangements

Protein-coding genes in the newly generated *S. japonicus* and *S. versatilis* genome assemblies were annotated by transferring the annotations of the reference *S. japonicus* genome using liftoff v1.6.3 with the options '-sc 0.95 -exclude_partial -polish'. The tRNA and rRNA genes were annotated using tRNAscan-SE v2.0.12 (Lowe & Chan, 2016) and barrnap v0.9 (Seemann, 2018), respectively. The density of protein-coding sequences, the number of tRNA genes, and the number of 5 S rRNA genes in 20-kb sliding windows were calculated using bedtools coverage while the GC content in 5-kb sliding windows was calculated using bedtools nuc (Quinlan & Hall, 2010). The resulting statistics and the genomic features were

visualized using Circos v0.69 (http://circos.ca/) (Krzywinski et al., 2009). To visualize the syntenic relationships between the newly generated *S. japonicus* and *S. versatilis* genome assemblies in the Circos plot, we selected 4067 protein-coding genes that are annotated in both assemblies by liftoff and are among the previously defined single-copy orthologous genes conserved across fission yeasts (Jia et al., 2023) and then used coloured lines to represent orthologous gene pairs, with each line directly connecting two genes that belong to the same pair.

2.13 | Strains and culture media used for hybridization

The type strains of *S. japonicus* (CBS 354^T) and of *S. versatilis* (CBS103^T) were used as mating partners in this study. The mutants used in the current study are listed in Table 1. The culture media used were: YEA (yeast extract agar), YEL (yeast extract liquid), SML (synthetic minimal liquid) and SMA (synthetic minimal agar) (Sipiczki & Ferenczy, 1977). Edinburgh minimal medium (EMM) was used without (EMM−N) and with NH_4Cl (EMM). EMM was used in solid (EMMA) and liquid form (EMML).

2.14 | Isolation and characterization of auxotrophic mutants

The isolation and characterization mainly followed the methods described by (Sipiczki & Ferenczy, 1978). Cells of the type strains were mutagenized with UV light. Cells of logarithmic-phase cultures were plated on YEA and irradiated with UV light to about 20% survival. After mutagenesis, the plates were incubated at 25°C for at least 7 days. By replica-plaiting of the developing colonies onto minimal medium by means of a velvet stamp the auxotrophic mutants were identified. The colonies whose replicas did not grow on minimal medium were isolated and nutrient requirements investigated on minimal medium using the procedure described by Sipiczki and Ferenczy (1978).

2.15 | Hybridization

Hybridization was done according to the methods described earlier (Bozsik et al., 2002) with minor modifications. One milliliters samples of early log-phase cultures of the partners grown overnight in YEL were mixed. Two milliliters of fresh YEL was added to the mixed culture which was subsequently incubated at room temperature for 2 days. Cells were separated from the medium by centrifugation and washed with sterile water. The pellet was resuspended in 100 µL sterile water. Twenty microliters of the suspension was dropped on EMMA−N and incubated at room temperature for 3 days to induce conjugation, zygote formation and sporulation. After 3 days incubation, the presence of ascospores was examined microscopically. The sporulating culture was suspended in 200 µL of sterile water and the resulting suspension was spread on two minimal medium agar plates. Growing prototrophic colonies were isolated as putative hybrids and isolated onto fresh synthetic minimal medium agar. The same procedure was also performed separately with both partners (without mixing with each other) to test them for back mutation. Sporulation of hybrid cultures was investigated microscopically.

2.16 | Characterization of hybrids

Sporulating hybrid cultures were treated with lysing enzyme (Sigma-Aldrich product no. L1412) overnight to release the spores from the asci and eliminate the nonsporulating cells. Aliquots of the spore suspensions were plated on complete medium. The developing colonies were replica-plated onto SMA to differentiate the prototrophs and the auxotrophs. Auxotrophic colonies were characterized concerning their nutrient requirements.

It was tested if arising prototrophic colonies segregate to determine if they are haploid recombinants or diploid hybrids. To this end, prototrophic colonies of each strain combination were tested for stability by random spore analysis. After the elimination of vegetative cells and the dissolution of the ascus walls by overnight lysing enzyme (Sigma-Aldrich product no. L1412) treatment, samples of the spore suspensions were spread onto YEA plates. The colonies formed were then replica-plated on EMMA to identify auxotrophic segregants. At least 200 colonies were tested after elimination of asexual cells.

2.17 | PCR fingerprinting

PCR fingerprinting was employed to verify the hybrid nature of prototrophic colonies. The oligonucleotide $(GAC)_5$ was used to generate PCR fingerprints. The composition of the 25-µL PCR mixture was as follows: 18.2 µL MilliQ water, 2.5 µL DreamTaq 10X buffer, 0.3 µL DreamTaq polymerase, 1 µL DNTP mix, 2 µL $(GAC)_5$ primer and 1 µL genomic DNA. The PCR programme was as follows: 5 min at 94°C, followed by 40 cycles of 1 min at 94°C, 1 min at 52.6°C and 1 min at 72°C. The fragments of the amplified DNA were separated on a 1.5% agarose gel by electrophoresis at 120 V for 150 min.

3 | RESULTS AND DISCUSSION

3.1 | Analysis of DNA sequences of the rRNA gene

The D1/D2 domain of the 26S LSU gene of the *S. japonicus* type strain CBS 354^T is 590 bp long (Accession no.: U94943). The D1/D2 domain of the *S. versatilis* type strain CBS 103^T is 1 bp shorter (sequence derived from the Westerdijk Fungal Biodiversity Institute database, URL: https://wi.knaw.nl). However, other strains of *S. versatilis* do not differ

in length of the the D1/D2 domain from CBS 354^T (see below). In addition to the indel the sequences of the two strains mentioned above differ by five substitutions. Excluding the indel position, the identity between the two sequences is 99.2%. Among the D1/D2 domain sequences of the 7 *S. versatilis* strains included in this analysis only two differed from the others. The sequence of CBS 103^T is 1 bp shorter than that of the other strains. Compared to the sequences of the other strains the D1/D2 domain of CBS 5679 shows three substitutions. These three substitutions are not located at the same sites that differentiate the type strains of *S. versatilis* and *S. japonicus*. Based on the currently available data the sequence of the D1/D2 domain seems to be well suited to determine unknown strains of both species to the species level.

In their epoch-making study on the identification and phylogeny of ascomycetous yeast species based on the D1/D2 domain, Kurtzman and Robnett (1998) predict that strains showing more than 1% substitutions in their D1/D2 domain are likely to belong to different species, while strains with 0-3 nucleotide differences either belong to the same or to sister species. The authors did not define cut-off values. Strains differing by more than 1% substitutions may still be conspecific while those showing less than 0-3 nucleotide differences may still belong to different but close related species. For both situations several examples have been described (Lachance, 2018). The findings of Kurtzman and Robnett (1998) were later confirmed by Vu et al. (2016) who evaluated the statistical confidence of strain assignment to species in dependence of different taxonomic threshold values on basis of more than 9000 yeast strains. The authors predict a taxonomic threshold for yeast species discrimination of 99.5% in the D1/D2 region. The currently recognized six species in the genus *Schizosaccharomyces* show much lower levels of identity in their D1/D2 domain if compared to each other. The only exception is the pair of *S. octosporus* and *S. lindneri* that differ by three substitutions equating to a 99.5% identity in their D1/D2 domain sequence. Considering the aforementioned studies, the five substitutions and one indel observed between the D1/D2 domain sequences of *S. japonicus* and *S. versatilis* led to the suspicion that they represent separate species but obviously further investigations were necessary to evaluate the situation. The phylogenetic position of *S. versatilis* within the genus *Schizosaccharomyces*, as determined by D1/D2 sequences, is illustrated in Figure 1a and Figure S5a. Inferring trees employing the Neighbor joining method (Figure 1a) or the maximum likelihood method (Figure 1b) resulted in highly similar results. The clades harbouring the *S. japonicus* or the *S. versatilis* strains had bootstrap support values of 93% or above, indicating that the assignment of strains to these clades is highly stable.

The ITS region (ITS1-5.8S-ITS2) of *S. japonicus* (CBS 354^T) is 604 bp long (Accession no.: NR_121199) while that of *S. versatilis* (CBS 103^T) is 601 bp long (sequence derived from the Westerdijk Fungal Biodiversity Institute database, URL: https://wi.knaw.nl). The length difference results from three indels each of which is 1 bp long. The aligned ITS sequences without positions containing gaps differ by 23 substitutions, equating to a 96.2% identity. All substitutions are confined to the ITS2 region. After analysing the D1/D2 domain and ITS region sequences of 450 strains of 242 species of basidiomycetous yeasts, Scorzetti et al. (2002) concluded that general guidance values could not be defined for species definition because of fundamentally different results derived from sequence comparison of closely related species. Based on all strains available at the time at the Westerdijk Fungal Biodiversity Institute, Vu et al. (2016) predicted an ITS taxonomic threshold of 98.3% for discriminating ascomycetous yeast species. Although the identity of 96.2% reported in the current study is considerably below the taxonomic threshold predicted in Vu et al. (2016), this value only constitutes a hint that the type strains of *S. japonicus* and *S. versatilis* represent separate species. Phylogenetic trees inferred on basis of the ITS region sequences are given in Figure 1b and Figure S5b. The trees constructed using the Neighbor joining method or the maximum likelihood method are highly similar at least with respect to the position of *S. versatilis* and *S. japonicus* strains. High bootstrap support values show that the assignment of these strains to the two clades is stable.

Judged by the sequences of the strains included in this study, the intraspecies variation is rather low. In *S. japonicus* the ITS region sequence of PYCC 6842 and CBS 354^T differ by 9 substitutions. Eight substitutions are located in the ITS1, 1 in the 5.8S rRNA gene and 1 in the ITS2 region. NBRC 1646 differs from the type strain in one substitution in the 5.8S rRNA gene and NBRC 1712 is identical with the sequence of CBS 354^T.

The ITS region sequence of the type strain of *S. versatilis* differs from all other *S. versatilis* strains in the current study in one gap which is located in the ITS2. Strain CBS 8543 differs from all other strains in three substitutions and three gaps that are located in the ITS1. The rest of the strains is identical in its ITS region sequence.

All four sequences derived from strains of *S. japonicus* agree in the sites in which they differ from the type strain of *S. versatilis*. The other way around all seven sequences derived from *S. versatilis* strains agree in the sites in which they differ from the type strain of *S. japonicus*. Therefore, the ITS region sequence constitutes a sound base for reliable assignment of new strains to one or the other species. Especially, this is true if ITS region sequence data are combined with those derived from the D1/D2 domain. In view of the fact that no intraspecies variability was observed at the sites of the alignment which distinguish the type strains of the two species and that a certain level of intraspecies variation was detected at other sites of the ITS region, it seems advisable to align the sequence of any strain that is to be determined with those of the two type strains before a final assessment of its taxonomic status is made and not to rely on identity values only.

The borderline value derived from the D1/D2 domain analysis, combined with the rather low identity value of the ITS regions, suggests that the two type strains represent separate species. Nonetheless, it needs to be emphasized that while sequences of the rRNA gene may raise suspicions that strains are not conspecific, a final taxonomic decision should consider all available data (Boekhout et al., 2021).

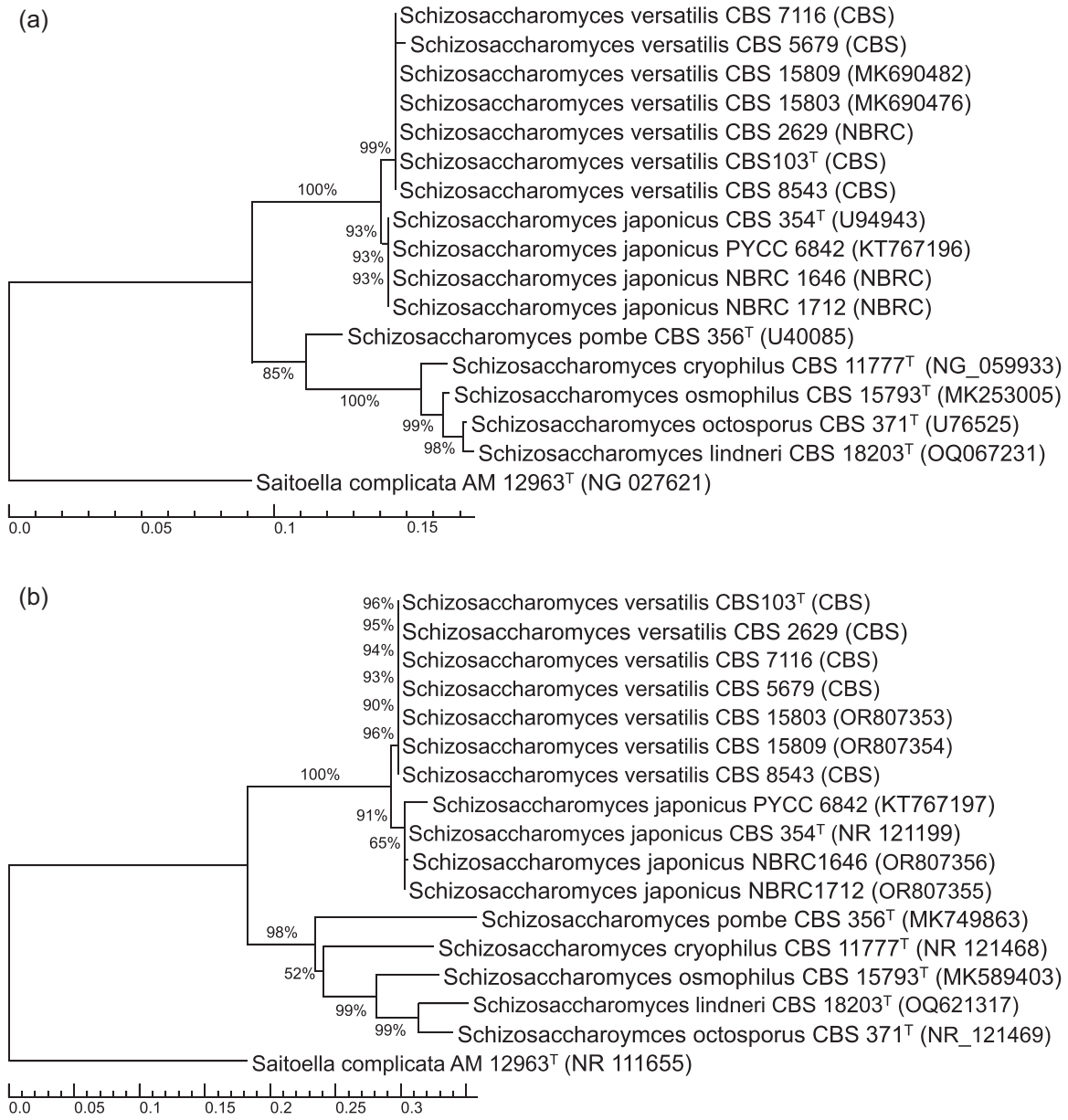


FIGURE 1 (a) Neighbor-joining tree based on the D1/D2 domain of the large subunit rRNA gene as inferred by the Neighbor-joining method (Felsenstein, 1981) and based on the Tamura Nei substitution model (Tamura & Nei, 1993). Distances are in the units of the number of base substitutions per site. Percentage bootstrap values of 1000 replicates were calculated. Bootstrap values lower than 50% are not shown. Alignment positions containing gaps and missing data were eliminated before analysis. Five hundred and five positions were in the final data set. *Saitoella complicata* was used as outgroup. (b) Neighbor-joining tree based on the ITS region (ITS2-5.8S-ITS2) as inferred by the Neighbor-joining method (Felsenstein, 1981) and based on the Tamura Nei substitution model (Tamura & Nei, 1993). Distances are in the units of the number of base substitutions per site. Percentage bootstrap values of 1000 replicates were calculated. Bootstrap values lower than 50% are not shown. Alignment positions containing gaps and missing data were eliminated before analysis. Three hundred and eighty-one positions were in the final data set. *S. complicata* was used as outgroup.

3.2 | Assembly of a high-quality genome of the *S. japonicus* type strain

The current reference genome of *S. japonicus* (SJ5 assembly), which is the genome of the type strain of *S. japonicus*, is not highly contiguous, consisting of 32 scaffolds (called supercontigs 5.1–5.32), with the three chromosome-scale scaffolds (supercontigs 5.1–5.3) containing 15, 11 and 1 gap, respectively (https://www.ncbi.nlm.nih.gov/assembly/GCF_000149845.2/) (Rhind et al., 2011). To generate a more complete assembly, we sequenced the *S. japonicus* type strain using Oxford Nanopore ultra-long read sequencing, PacBio HiFi sequencing and Illumina paired-end sequencing (Figure S1 and Supporting Information S7: Table 1). De novo assembly and polishing using these sequencing data yielded four large contigs totalling

~15.4 Mb. Telomeric repeats, most commonly TTAGGG and TTAGTC, were found at four contig ends (Figure S1b). tDNA-5S-rDNA arrays, described in details below, were present at two other ends, while 18S-5.8S-28S rDNA arrays were at the two remaining ends, though incorrectly oriented and flanked internally by telomeric repeats in one case (Figure S1b). We trimmed this misoriented rDNA array and aligned the contigs to the reference genome. Two contigs matched chromosomes 1 and 3 of the reference genome. The other two contigs together comprised chromosome 2, therefore we merged them into a chromosome 2 scaffold (Figure S1b). We adjusted the orientation of the contigs to match the orientation of the chromosomes in the reference genome.

The nontelomere ends of the two chromosome 2 contigs contain tandem repeats of the same 6.3-kb sequence. This sequence consists of 18 tRNA genes (tDNAs) and three 5S rRNA genes (5S rDNAs). Inspection of Nanopore ultra-long reads containing this type of repeat sequence revealed no additional repeat-nonrepeat junctions other than the two present in the two chromosome 2 contigs, indicating that the gap between these two contigs lies within a single long repeat array. This repeat array is the most centromere-proximal sequence on the right arm of chromosome 2, and is therefore named the cen2-adjacent tDNA-5S-rDNA array.

On the right arm of chromosome 2, approximately 400 kb away from the cen2-adjacent tDNA-5S-rDNA array, there is another region that contains tDNA-5S-rDNA repeats. This region lies between *SJAG_04837* (ortholog of *S. pombe bet5*) on the centromere-proximal side and *SJAG_04436* (ortholog of *S. pombe SPAC1705.02*) on the centromere-distal side. The gene next to *SJAG_04436* on the centromere-distal side, *SJAG_04435*, shares an orthologous relationship with *S. pombe ire1*. Within this region are two long tandem repeat arrays situated next to each other. These two arrays consist of tDNAs and 5S rDNAs arranged in the same order as in the cen2-adjacent tDNA-5S-rDNA array, but they differ from each other and from the cen2-adjacent array in the intervening sequences between tDNAs (Figure S3). We refer to these two arrays as the *bet5*-adjacent tDNA-5S-rDNA array and the ire1-adjacent tDNA-5S-rDNA array.

Illumina read depth analysis revealed that the copy numbers of the repeats in the *bet5*-adjacent and ire1-adjacent tDNA-5S-rDNA arrays are underrepresented in the genome assembly. To address this issue and to fill the gap inside the cen2-adjacent tDNA-5S-rDNA array, we estimated the copy numbers of the repeats in the three arrays and then replaced the nonboundary repeat units in each array with tandem copies of the consensus repeat unit sequence so that the total number of repeat units matched the estimated copy number.

The final genome assembly of *S. japonicus* consists of three gapless chromosomal contigs, totalling ~16.8 Mb (Figure 2 and Supporting Information S7: Table 2). Chromosome 1 is ~5.2 Mb, chromosome 2 is ~7.6 Mb and chromosome 3 is ~4.0 Mb. All contigs span telomere-to-telomere, except for the left end of chromosome 3, which contains the 18S-5.8S-28S rDNA array. Our *S. japonicus* assembly is 51% longer than the ungapped length of the current reference genome (~11.1 Mb). This length difference is primarily due to several large repeat-rich regions, including the three tDNA-5S-rDNA arrays. These repeat-rich regions correspond to gaps in the reference genome.

3.3 | Assembly of a high-quality genome of the *S. versatilis* type strain

The genome of the *S. versatilis* type strain was sequenced and assembled following the same approach used for the *S. japonicus* type strain (Figure S2 and Supporting Information S7: Table 1). The initial assembly consists of two contigs, one ~10.3 Mb in length and the other ~4.5 Mb. Given that a previous electron microscopy study has shown that *S. japonicus* and *S. versatilis* both have three chromosomes (Robinow & Hyams, 1989), we hypothesized that two chromosomes were mis-joined into the longer contig. Upon inspection, telomeric repeats—most commonly TTAGGG and TTAGTC—were found not only at all four contig ends, but also internally in the long contig (Figure S2b). This confirmed the erroneous joining of two chromosomes. The long contig was therefore manually split into two separate contigs. The orientation of the three contigs was rearranged based on alignment with the *S. japonicus* genome assembly and chromosome numbers were assigned to them (Figure S2b).

Three tDNA-5S-rDNA arrays are present in chromosome 2 of the *S. versatilis* genome assembly at positions syntenic to the three tDNA-5S-rDNA arrays in chromosome 2 of the *S. japonicus* genome assembly. Illumina read depth analysis indicated that the copy numbers of the repeats in all three arrays are underrepresented. To correct this, we replaced the nonboundary repeat units in each array with tandem copies of the consensus repeat unit, matching the estimated copy number. The final *S. versatilis* genome assembly comprises three gapless chromosomal contigs, totalling ~15.5 Mb (Figure 2 and Supporting Information S7: Table 2).

3.4 | Assessment of the gene content completeness of the genome assemblies

We evaluated the gene content completeness of our newly generated *S. japonicus* and *S. versatilis* genome assemblies by analysing the presence/absence status of a set of 1315 Benchmarking Universal Single Copy Orthologs (BUSCO) genes from the ascomycota_odb9 gene set (Supporting Information S7: Table 3) (Simão et al., 2015). This analysis showed that 1217 (92.5%) and 1220 (92.8%) of the BUSCO genes are 'complete and single-copy' in the *S. japonicus* and *S. versatilis* assemblies, respectively, and 32 (2.4%) and 29 (2.2%) BUSCO genes are 'complete and duplicated' in the *S. japonicus* and *S. versatilis* assemblies, respectively. As a comparison, 1209 (91.9%) BUSCO genes are 'complete and single-copy' and 28 (2.1%) BUSCO genes are 'complete and duplicated' in the *S. japonicus* reference genome (Supporting Information S7: Table 3). Thus, our genome assemblies of *S. japonicus* and *S. versatilis* have higher BUSCO completeness than the current reference genome of *S. japonicus*.

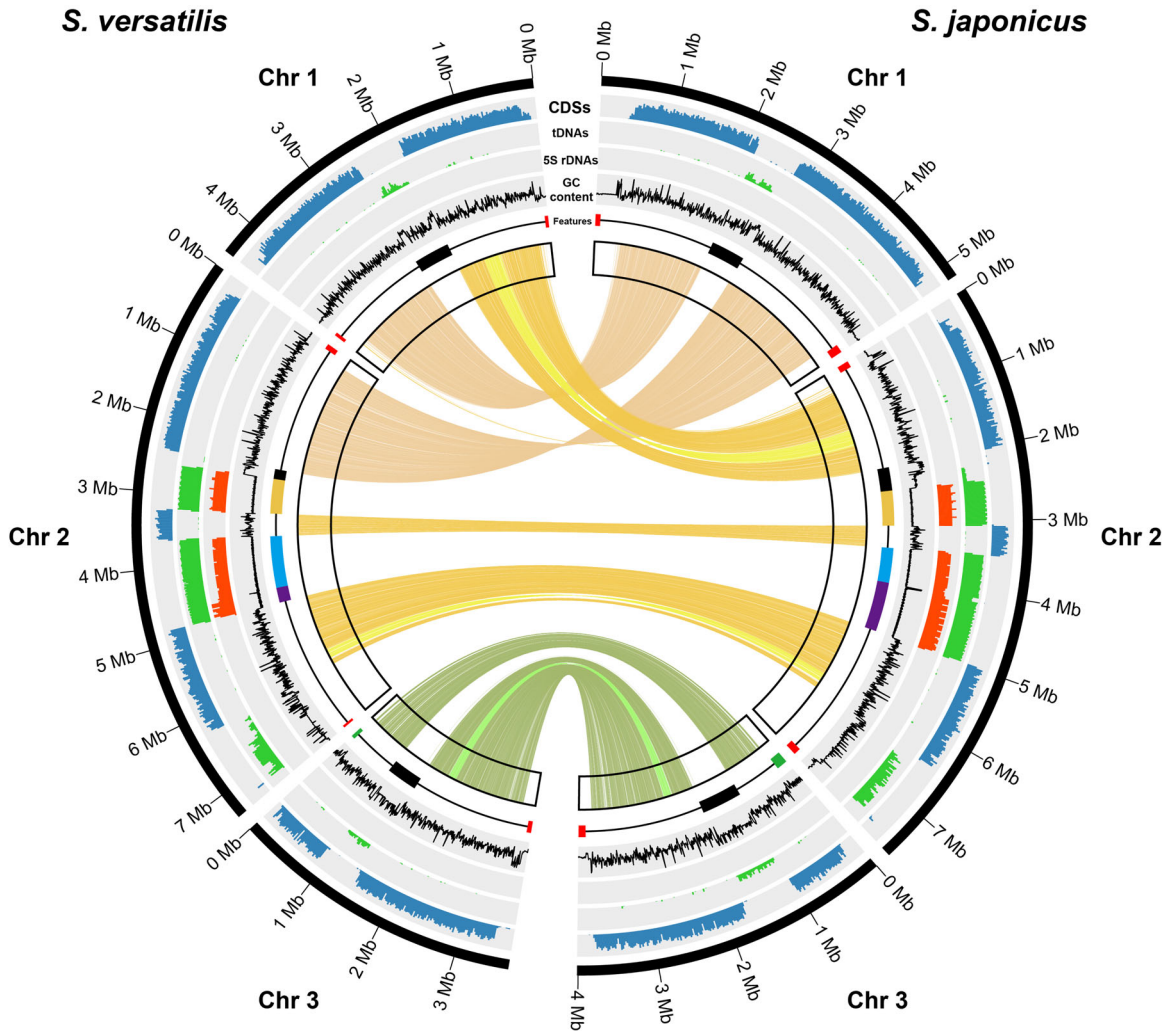


FIGURE 2 A Circos diagram showing the genomic features and syntenic relationships between the genome assemblies of *Schizosaccharomyces japonicus* and *Schizosaccharomyces versatilis*. The orientations of the chromosomes of *S. japonicus* are clockwise and the orientations of the chromosomes of *S. versatilis* are counterclockwise. Tracks from outer to inner include: protein-coding sequence density (bin size = 20 kb, range = 0–1); number of tRNA genes (tDNAs) per bin (bin size = 20 kb, range = 0–60); number of 5S rRNA genes (5S rDNAs) per bin (bin size = 20 kb, range = 0–15); GC content (bin size = 5 kb, range = 30%–60%); genomic features including centromeres (black), telomeres (red), 18S-5.8S-28S rDNA arrays (green), cen2-adjacent tDNA-5S-rDNA arrays (orange), *bet5*-adjacent tDNA-5S-rDNA arrays (blue) and ire1-adjacent tDNA-5S-rDNA arrays (purple). The innermost track is used for depicting syntenic relationships represented by coloured lines connecting orthologous protein coding genes in the two species. The colours of lines are based on which chromosome the *S. japonicus* genes are on. Three intrachromosomal inversions are highlighted using brighter colours.

3.5 | Genome-based comparison between *S. japonicus* and *S. versatilis*

To estimate the divergence level between *S. japonicus* and *S. versatilis* using whole genome comparison, we calculated the genome-wide ANI values between all *Schizosaccharomyces* species, including *S. versatilis* (Table 2). The ANI between *S. versatilis* and *S. japonicus* is 86.35%, slightly higher than the ANI between *S. osmophilus* and *S. octosporus* (83.57%) and the ANI between *S. osmophilus* and *S. lindneri* (83.88%), but lower than the ANI between *S. octosporus* and *S. lindneri* (90.43%). As an ANI value of 95% has been considered a suitable threshold for delineating yeast species (Lachance et al., 2020), the ANI between *S. japonicus* and *S. versatilis* indicates that the reinstatement of *S. versatilis* is justified.

Using 1045 'complete and single-copy' BUSCO genes, we constructed a maximum likelihood species tree encompassing *S. versatilis* and the six previously recognized fission yeast species and estimated the divergence time between *S. japonicus* and *S. versatilis* using the RelTime method (Tamura et al., 2012, 2018). This estimation showed that *S. japonicus* and *S. versatilis* diverged approximately 13.3 million years ago, slightly younger than the divergence time between *S. osmophilus* and the ancestor of *S. octosporus* and *S. lindneri* (15.7 million years ago) (Figure 3).

Using our chromosome-level genome assemblies, we visualized the syntenic relationship between *S. japonicus* and *S. versatilis* (Figure 2). There is a large-scale reciprocal interchromosomal translocation that results in the right arm of chromosome 1 and the

TABLE 2 Average nucleotide identities (ANIs) for species pairs in the genus *Schizosaccharomyces*.

	Schizosaccharomyces cryophilus	*Schizosaccharomyces lindneri*	*Schizosaccharomyces octosporus*	*Schizosaccharomyces osmophilus*	*Schizosaccharomyces pombe*	*Schizosaccharomyces japonicus*	*Schizosaccharomyces versatilis*
S. cryophilus	100.00	79.15	79.12	81.29	70.76	69.39	69.25
S. lindneri		100.00	90.43	83.88	70.69	69.36	69.10
S. octosporus			100.00	83.57	70.78	69.29	69.05
S. osmophilus				100.00	70.76	69.48	69.22
S. pombe					100.00	69.28	69.00
S. japonicus						100.00	86.35
S. versatilis							100.00

left arm of chromosome 2 in *S. japonicus* swapping their locations in *S. versatilis*. Additionally, there are three small-scale intrachromosomal inversions.

3.6 | Hybridization analysis

Like other species of the genus *Schizosacccharomyces*, *S. japonicus* has a haplontic life cycle and is homothallic. After mating type switching (Klar et al., 2015), clonal haploid cells of opposite mating type may conjugate and form a zygote. The diploid phase is restricted to the zygote, which promptly undergoes meiosis upon its formation. Eight haploid spores are formed (Bozsik et al., 2002).

For the sake of clarity, in the following the term 'hybrid' is solely used in the sense of Sipiczki (2018) exclusively for cells containing two complete parental genomes.

Auxotrophic mutants of the *S. japonicus* type strain and the *S. versatilis* type strain were isolated. Those used for the crossing experiments are listed in Table 1. Prototrophic colonies were obtained in all combinations of strains after mating (Figure 4; Table 3). The efficiency was by about two orders of magnitude higher in intraspecies mating than in interspecies mating. The results indicate that the prezygotic reproductive barrier between the two type strains is strong but incomplete.

On EMMA-N all parental strains formed zygotes (Figure 5a) which then converted into zygotic asci usually containing eight ascospores (Figure 5b). The intraspecies crosses also formed only zygotic asci with normal spore number (eight) and morphology, indicating that the spores have a haploid genome as it is typical for the species in *Schizosaccharomyces*. In contrast, the interspecies crosses produced mixtures of azygotic and giant zygotic asci (Figure 5c). The former usually contained spores of aberrant morphology (Figure 5c,d,f) or of low number (Figure 5c,e). Prerequisite for the formation of azygotic asci is the presence of a diploid or polyploid genome. Most of the giant asci had the expected number (eight) of spores (Figure 5c), but the spores were larger than those of the zygotic asci of the parental cultures. Because for *S. pombe* it was shown that diploid spores are significantly larger than the usual haploid spores (Molnar & Sipiczki, 1993) it seems reasonable to assume that the spores in the giant zygotic asci are allodiploid hybrids. Occasionally, giant zygotic asci with abnormal spore numbers were also found (Figure 5g,h). The formation of giant zygotic asci has been described before in *S. pombe* by Sipiczki and Kucsera (1983). The formation of azygotic asci has been observed in diploid cultures of *S. japonicus* which were produced by protoplast fusion (Bozsik et al., 2002). In the current study, the presence of both asci types strongly suggest that the prototrophic colonies derived by mating of the two type strains were allodiploid hybrids.

The results of the $(GAC)_5$ PCR fingerprinting show that the banding pattern of all prototrophic colonies contain all major parental bands (Figure 6). These results support the findings of the microscopic investigation of the sporulation behaviour of the prototrophic colonies and indicate that the colonies are allodiploid hybrids containing both parental genomes.

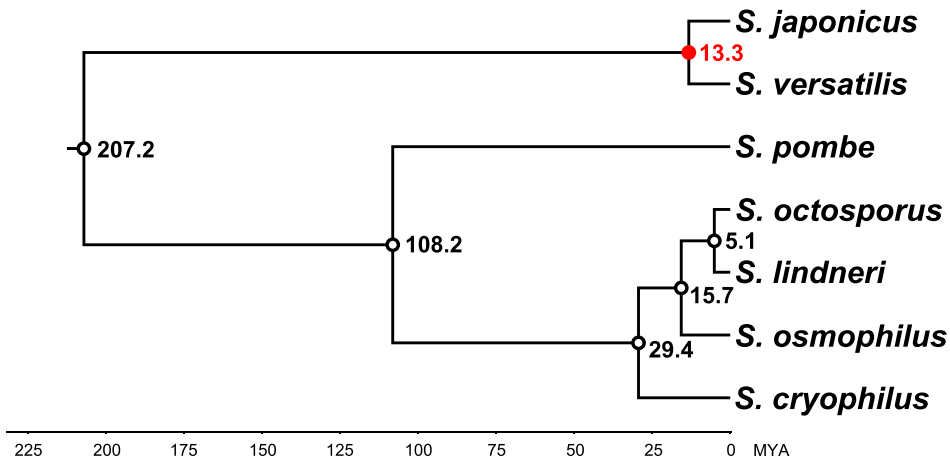


FIGURE 3 Time-calibrated phylogeny of the *Schizosaccharomyces* genus. A concatenated matrix comprising 1045 'complete and single-copy' Benchmarking Universal Single-Copy Orthologs (BUSCO) genes from the seven fission yeast species and the outgroup species *Saitoella complicate* was constructed. Maximum likelihood phylogeny inference was performed on this concatenated matrix. The RelTime method was employed to estimate the divergence time between *Schizosaccharomyces japonicus* and *Schizosaccharomyces versatilis* (filled red circle) with the aid of the five time calibration nodes (empty black circles).

The intraspecies mating of two mating partners with different auxotrophic markers resulted in the development of prototrophic colonies. Spore clones derived from these prototrophic colonies were all prototrophic. As they do not segregate and only form zygotic asci, everything indicates that they are haploid recombinants (Table 3) as it is typical for the life cycle in the genus *Schizosaccharomyces*. The majority of colonies derived from spores of the prototrophic interspecies hybrids were also prototrophic, but a few of them exhibited the auxotrophy of one or the other parent (Table 3). Because the hybrids segregate, it can be concluded that they are allodiploid. The PCR banding patterns of representatives of prototrophic and auxotrophic spore clones were compared with the patterns of the hybrids and the parental strains. All spore clones turned out to have recombinant patterns (Figure 6; Table 3). Some spore clones only produce normal zygotic asci whereas others produce azygotic and giant zygotic asci, indicating that these spore clones are diploid. Most likely these diploid spore clones originate from diploid spores of giant zygotic asci developing from allotetraploid zygotes which are the result of conjugation of two allodiploid hybrid cells. About half of the spore clones examined were prototrophic while the rest exhibited one or the other parental auxotrophic marker. Double auxotrophs were not found. This can be interpreted in a way that the two subgenomes present in the hybrids are not fully compatible.

The mating experiments show that *S. versatilis* and *S. japonicus* are separated by an almost complete biological isolation but both the prezygotic and the postzygotic sterility barriers are incomplete, a situation reminiscent of the incomplete biological isolation between the *Saccharomyces* species (Sipiczki, 2018). Reproductive interaction between *S. versatilis* and *S. japonicus* is not only less efficient. The development of the progeny after interspecies mating is fundamentally different compared to what was observed during intraspecies mating. During interspecies mating only allodiploid hybrids which contain both parental genomes are formed while intraspecies mating leads directly to the formation of asci and to haploid recombinant progeny. In contrast to this strictly haplontic life cycle being typical for the genus *Schizosaccharomyces*, the allodiploid hybrids reproduce vegetatively, by fission. Their cells can also sporulate, but the formation of asci is aberrant. Either azygotic asci producing aberrant haploid (and probably alloaneuploid) spores are formed directly from the hybrid cells or giant zygotic asci producing diploid spores are formed after conjugation of the hybrid cells. A life cycle that is characteristic of *S. japonicus* or *S. versatilis* is not observed in cultures arising from crosses between the two species.

3.7 | Species concepts

Phenotypically *S. japonicus* and *S. versatilis* are undistinguishable. Nonetheless, they are not typologically indistinguishable. In this context it seems important to note that the genetic species concept can be understood as variant of the phenotypic species concept (Boekhout et al., 2021). Comparing sequences of several strains of each species with those of the other species, fixed differences in the ITS region and the D1/D2 domain were identified in the present study. Intermediate types of sequences that would bridge this discontinuity gap between *S. japonicus* and *S. versatilis* were not found at all. The relatively low ANI value found by comparing the genome sequences of the type strains and the genome rearrangements detected by comparing the synteny of the genomes additionally show that strains of *S. japonicus* and *S. versatilis* can be distinguished on a typological basis. However, all these differences tell little about the nature of the two distinct groups of strains because discontinuities in the

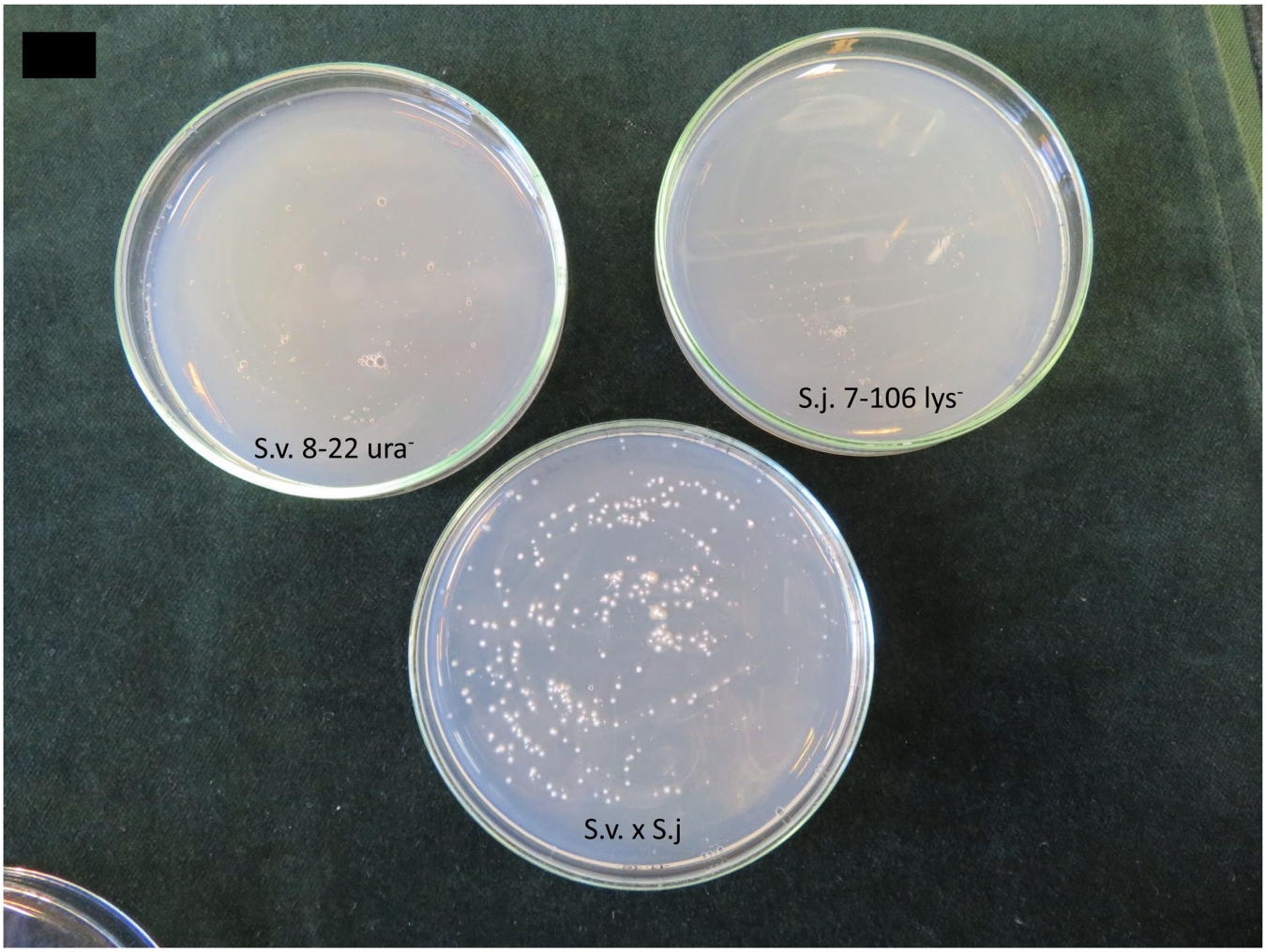


FIGURE 4 Hybridisation of auxotrophic *Schizosaccharomyces japonicus* and *Schizosaccharomyces versatilis* mutants. On the minimal medium EMMA, the parental strains do not form colonies.

variability of certain characteristics of individuals are the rule and not the exception in nature (Dobzhansky, 1935). The discontinuity alone cannot answer the question if the two groups of strains represent different species or just subpopulations within the same species, a question that hardly can be answered by a typological (phenotypic and genetic) species concept.

According to Dobzhansky (1935) the term 'species' should be used for groups of individuals that are reproductively isolated to a degree that leads to fixed and irreversible phenotypic differences between these groups. In this sense taxonomists applying a typological species concept often indirectly and unintentionally detect reproductive incompatibility between the species they separate (Dobzhansky, 1935). Although Dobzhansky (1935) did not use the term 'biological species' or 'biological species concept', later authors used these terms to denominate his concept. Consequently, Mayr (2000) defined biological species 'as groups of interbreeding natural populations that are reproductively isolated from other such groups'. Additionally he clearly states that 'A population does not lose its species status when an individual belonging to it makes a mistake and hybridizes'. Both statements make clear that species delimitation by this concept alone would require the investigation of the reproductive compatibility between numerous yeast strains of various populations, a task that is hardly feasible in view of the relatively few strains available in the current case and the immense time and effort required. The mating trials performed in the current study between the type strains provide strong evidence that the type strains actually represent separate biological species but because of the restriction to one individual of each species it cannot serve as a standalone proof. The diagnose of a strong reproductive incompatibility inferred from the mating trials conducted with the type strains is supported by major genome rearrangements described above. Particularly, the large scale reciprocal interchromosomal translocation affecting chromosomes 1 and 2 suggests that genetic recombination is strongly impaired in zygotes formed by the two species.

As in many cases the current one requires a polyphasic approach. Therefore, the decision to propose the reinstatement of *S. versatilis* was drawn from results of rRNA gene sequence analysis, whole

TABLE 3 Hybridisation and hybrid analysis.

Strain	Phenotype		GAC pattern	Sporulation	Spore clones (segregants)
	Auxotrophy	Mycelium			
Parental strains					
S.j. 7-83 ade$^-$	ade$^-$	+++	S.j.	zygotic	
S.j. 7-106 lys$^-$	lys$^-$	+++	S.j.	zygotic	
S.v. 8-5 ura$^-$	ura$^-$	+	S.v.	zygotic	
S.v. 8-8 ade$^-$	ade$^-$	+	S.v.	zygotic	
S.v. 8-28 ura$^-$	ura$^-$	+	S.v	zygotic	
S.v. 8-22 ura$^-$	ura$^-$	+	S.v.	zygotic	
Intraspecies hybridisation					
S.j. 7-83 × S.j. 7-106 (five hybrids)	prototroph	+++	S.j.	zygotic	prototroph
S.v. 8-5 × S.v. 8-8 (five hybrids)	prototroph	+	S.v.	zygotic	prototroph
Interspecies hybridisation					
S.j. 7-83 × S.v. 8-5 (five hybrids)	prototroph	+++	Hybrid	azygotic, giant zygotic	prototroph, auxotroph
S.j. 7-106 × S.v. 8-8 (five hybrids)	prototroph	+++	Hybrid	azygotic, giant zygotic	prototroph, auxotroph
S.j. 7-106 × S.v. 8-22 (five hybrids)	prototroph	+++	Hybrid	azygotic, giant zygotic	prototroph, auxotroph
S.j. 7-83 × S.v. 8-28$^{\times}$ (five hybrids)	prototroph	+++	Hybrid	azygotic, giant zygotic	prototroph, auxotroph
Segregants (spore clones)					
From 7-106 × 8-8/1					
a	prototrophic	+++	Recombinant	azygotic, giant zygotic	
b, 2, 3, 4 (four clones)	prototrophic	+++		azygotic, giant zygotic	
c	prototrophic	+++	Recombinant	zygotic	
d	lys$^-$	+++	Recombinant	zygotic	
e	lys$^-$	+++		azygotic, giant zygotic	
f	ade$^-$	+	Recombinant	zygotic	
g, h, i, j, k (five clones)	ade$^-$	+		zygotic	
l	prototrophic	+	Recombinant	azygotic, giant zygotic	
From 7-106 × 8-22/1					
a	ura$^-$	+++	Recombinant	zygotic	
b, d (two clones)	ura$^-$	+++		zygotic	
c	prototrophic	+++	Recombinant	azygotic, giant zygotic	
e	prototrophic	+++		azygotic, giant zygotic	
1,2,3,4 (four clones)	prototrophic	+		azygotic, giant zygotic	

Abbreviations: S.j., *Schizosaccharomyces japonicus*; S.v., *Schizosaccharomyces versatilis*.

genome sequence and synteny analysis, as well as mating trials and offspring analysis.

3.8 | Comments on the species description

Wickerham and Duprat (1945) did not define a holotype strain as required by the International Code of Nomenclature for algae, fungi, and plants (2018). Consequently, the name *Schizosaccharomyces versatilis* is not valid. Below a complete species description is given according to the current knowledge. Wickerham and Duprat (1945) did not explicitly explain why they chose the epithet *versatilis* but it likely refers to the dimorphic growth of the species. Data on phenotypic characteristic of the species were taken from the database of the Westerdijk Fungal Biodiversity Institute, from the original species description (Wickerham & Duprat, 1945) and from our own observations.

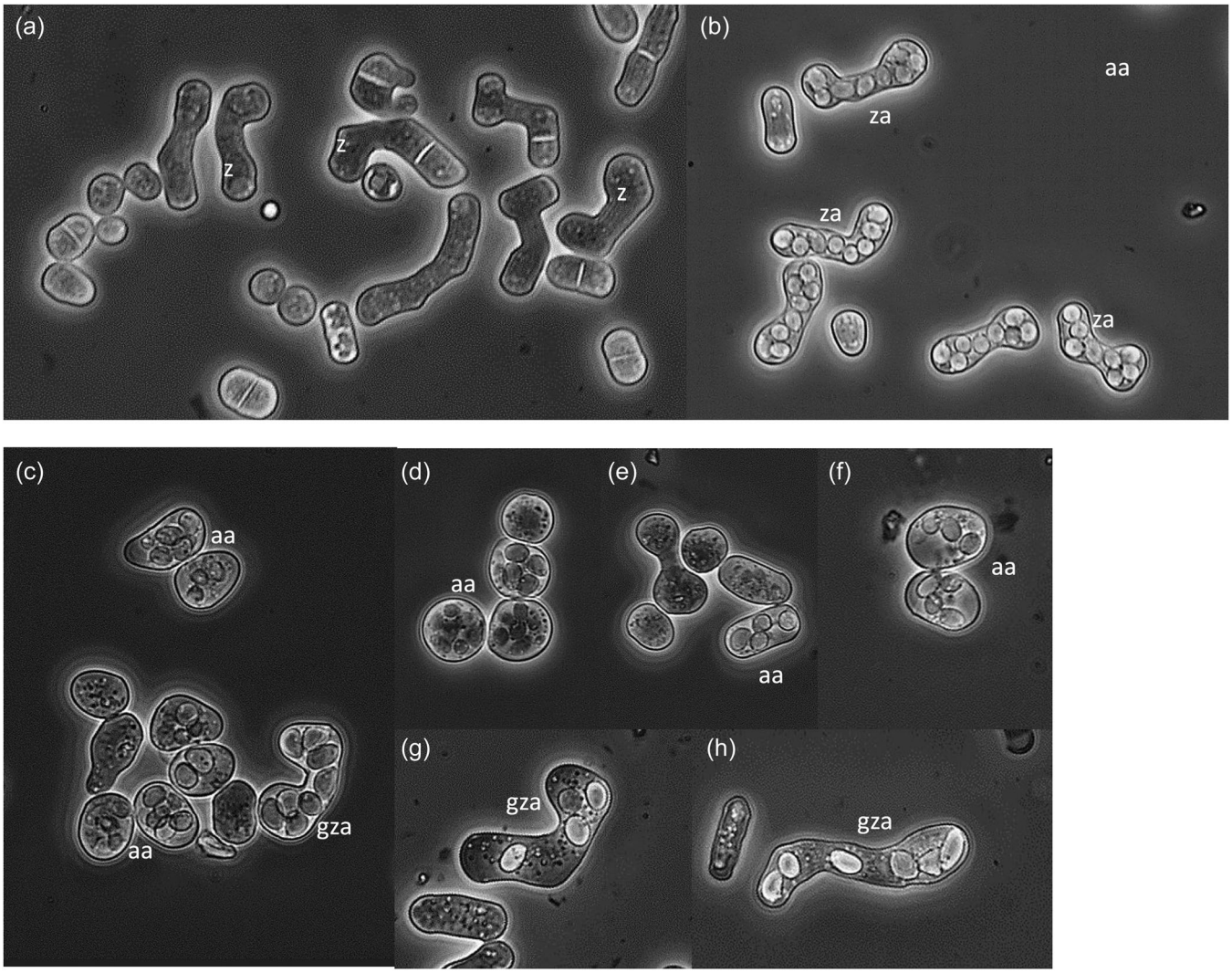
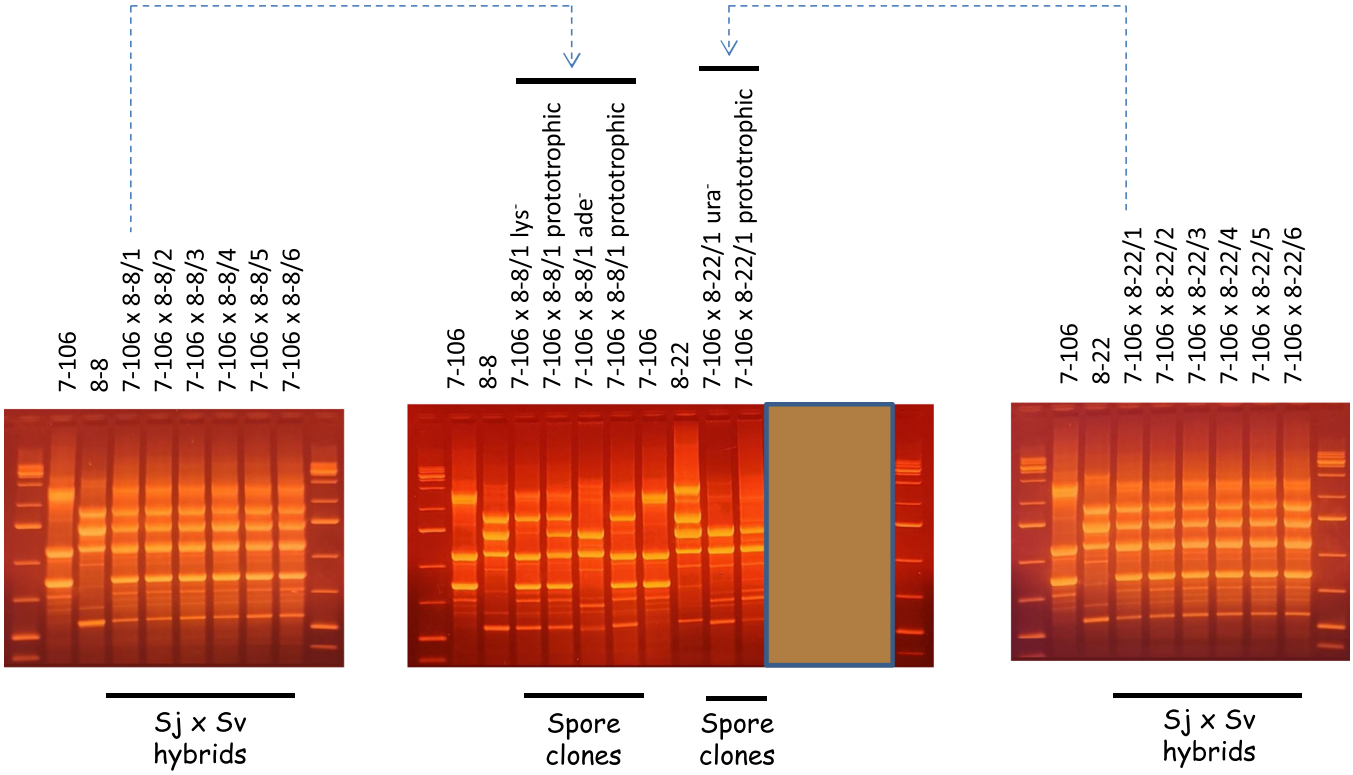


FIGURE 5 Mating and sporulation. (a) mating (conjugation) of haploid cells of *Schizosaccharomyces japonicus* 7-106. (b) sporulation of *S. japonicus* 7-106. (c–h) Sporulation of a *S. japonicus* 7-106 × *Schizosaccharomyces versatilis* 8-22 hybrid. Note the aberrant morphology and number of spores in the azygotic asci (c–f) and in the giant zygotic asci (g and h). aa, azygotic ascus of a diploid cell formed without conjugation with another cell; gza, 'giant zygotic ascus' of a tetraploid zygote; z, zygote of haploid cells; za, zygotic ascus of a diploid zygote.

3.9 | Conclusions

According to the analysis of sequences of the rRNA gene, the type strains of *S. japonicus* and *S. versatilis* are closely related. The distance between their D1/D2 domain sequences is just at the limit of what generally is expected within a species. The distance between their ITS regions supports the reinstatement of *S. versatilis*. Importantly, no intraspecies variation was observed at the sites that distinguish these two species. Thus, these sites represent a reliable base to assign new strains to one or the other species. The ANI calculated from comparison of whole genome sequences of the two type strains is a strong argument that they represent separate species. Their taxonomic separation is supported by the very low mating efficiency of the type strains and the finding that the mating products are allodiploid hybrids rather than haploid recombinants. This indicates that the two type strains are separated by almost complete sterility barriers. Collectively, our findings show that *S. versatilis* is a distinct species and its reinstatement is warranted.

The genome of *S. versatilis* strain CBS 5679 has been sequenced and assembled (Bouwknegt et al., 2021). Recently, an improved genome assembly of CBS 5679 was generated by Etherington et al. (2023). By comparing the CBS 5679 genome with the *S. japonicus* type strain genome and conducting experiments that demonstrated reproductive isolation between the two strains, Etherington et al. concluded that CBS 5679 belongs to a species distinct from *S. japonicus*, which corroborates our findings.



7-106: *S. japonicus* lys-
8-8: *S. versatilis* ade-
8-22: *S. versatilis* ura-

FIGURE 6 Fingerprinting of the parental strains, their hybrids and certain spore clones of two hybrids.

4 | *S. VERSATILIS* WICKERHAM AND DUPRAT EX BRYSCH-HERZBERG, JIA, SIPICZKI, SEIDEL, ZHANG, DU

MycoBank number 847838: The specific epithet *versatilis* (L. masc. adj. *versatilis* revolving, moving, versatile) refers to the species' ability to grow dimorphically as yeast and with true hyphae.

After 48 h in yeast extract broth at 25°C, cells are spheroidal to cylindrical, measuring 5.4 µm for the spheroidal cells and 14.0 × 5.7 µm for the cylindrical cells. The average is 10.2 × 5.1 µm. No ring or pellicle is formed.

After 6 days on yeast extract agar, colonies are smooth, glistening and circular. The consistency is creamy. The colonies develop a margin of hyphae that grow into the agar. Cell sizes vary from 10 to 800 µm, with the extreme long cells forming the mycelium.

After 1 week at 25°C on 2% malt extract agar, cells conjugate and form asci. Asci measure 15.6 × 3.5–18.5 × 5.7 µm with an average of 16 × 5 µm. The standard number of spores per ascus is 8. The ascospores are reniform in shape and measure 4.3 × 2.8–7.1 × 2.5 with an average of 5.5 × 2.8 µm. Ascospores may have hemispheroidal protuberances.

Glucose, maltose, sucrose, raffinose and melibiose are fermented. Inulin, galactose, lactose, trehalose, melezitose, methyl-α-D-glucoside, soluble starch, cellobiose and xylose are not fermented.

Glucose, inulin (variable, delayed, weak), sucrose, raffinose, melibiose (variable, delayed, weak) maltose (variable, delayed, week) are assimilated. Inulin, galactose, lactose, trehalose, melezitose, Methyl-α-D-glucoside, soluble starch, cellobiose, salicin, L-sorbose, L-rhamnose, D-xylose, L-arabinose, D-arabinose, D-ribose, methanol, ethanol, glycerol, erythritol, ribitol, galactitol, D-mannitol, D-glucitol, myo-inositol, DL-lactate, succinate, citrate, D-gluconate, D-glucosamine, N-acetyl-D-glucosamin, saccharate, xylitol, L-arabinitol, arbutin, propane 1,2 diol, and butane 2,3 diol are not assimilated.

L-lysine is assimilated (weak) as the sole nitrogen source whereas nitrate, nitrite, ethylamine, cadaverine, creatine, glucosamine, imidazole and D-tryptophane are not assimilated. Growth without vitamins does not take place.

Urease test is positive. Diazonium blue B reaction is negative. No starch is formed and no acetic acid is produced.

No growth occurs on 50% glucose. Ten percent NaCl, 0.001% cycloheximide and 1% acetic acid are not tolerated.

Holotype strain: CBS 103^T and ex-types: NRRL Y-1026, NBRC 1607, ATCC 9987, PYCC 7100; the strains are permanently

preserved in a metabolically inactive state. The type culture was isolated from home-canned grape juice in Michigan, USA.

AUTHOR CONTRIBUTIONS

Writing of the manuscript: Michael Brysch-Herzberg and Guo-Song Jia. *DNA sequence analysis*: Michael Brysch-Herzberg. *Mating of strains*: Matthias Sipiczki. *Microphotography*: Matthias Sipiczki. *Writing of the paper*: Matthias Sipiczki, Martin Seidel, Wen-Cai Zhang, and Li-Lin Du. *Genome data analysis*: Guo-Song Jia. *Molecular genetic work*: Martin Seidel. *Data analysis*: Wen-Cai Zhang and Li-Lin Du.

ACKNOWLEDGEMENTS

We thank Graham Etherington and Conrad Nieduszynski for sharing results before publication. Open Access funding enabled and organized by Projekt DEAL. This work was supported by The Deutsche Bundesstiftung Umwelt DBU (German Federal Environmental Foundation) [34053/01-32 to M. B.-H.] and grants from the Ministry of Science and Technology of China to L.-L. D.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All data relevant for the study are accessible.

ORCID

Michael Brysch-Herzberg http://orcid.org/0000-0002-3166-2010

REFERENCES

Beijerinck, M. (1894). Schizosaccharomyces octosporus, eine achtsporige Alkoholhefe. *Zentralblatt Bakteriologie Parasitenkunde*, *16*, 49–58.

Boekhout, T., Aime, M. C., Begerow, D., Gabaldón, T., Heitman, J., Kemler, M., Khayhan, K., Lachance, M. A., Louis, E. J., Sun, S., Vu, D., & Yurkov, A. (2021). The evolving species concepts used for yeasts: From phenotypes and genomes to speciation networks. *Fungal Diversity*, *109*(1), 27–55.

Bouwknegt, J., Wiersma, S. J., Ortiz-Merino, R. A., Doornenbal, E. S. R., Buitenhuis, P., Giera, M., Müller, C., & Pronk, J. T. (2021). A squalene-hopene cyclase in *Schizosaccharomyces japonicus* represents a eukaryotic adaptation to sterol-limited anaerobic environments. *Proceedings of the National Academy of Sciences*, *118*(32), 1–9. https://doi.org/10.1073/pnas.2105225118

Bozsik, A., Szilagyi, Z., Benko, Z., & Sipiczki, M. (2002). Marker construction and cloning of a cut1-like sequence with ARS activity in the fission yeast *Schizosaccharomyces japonicus*. *Yeast*, *19*(6), 485–498.

Brysch-Herzberg, M., Jia, G. S., Sipiczki, M., Seidel, M., Li, W., Assali, I., Du, L. L. (2023). Schizosaccharomyces lindneri sp. nov., a fission yeast occurring in honey. *Yeast*, *40*(15), 237–253.

Brysch-Herzberg, M., Tobias, A., Seidel, M., Wittmann, R., Wohlmann, E., Fischer, R., Dlauchy, D., & Peter, G. (2019). Schizosaccharomyces osmophilus sp. nov., an osmophilic fission yeast occurring in bee bread of different solitary bee species. *FEMS Yeast Research*, *19*(4), foz038.

Cabanettes, F., & Klopp, C. (2018). D-GENIES: Dot plot large genomes in an interactive, efficient and simple way. *PeerJ*, *6*, e4958. https://doi.org/10.7717/peerj.4958

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*, 1972–1973.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890.

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, *18*(2), 170–175.

Delaluna, J. O., Kang, H., Chang, Y. Y., Kim, M., Choi, M.-H., Kim, J., & Song, H. B. (2023). De novo assembled mitogenome analysis of Trichuris trichiura from Korean individuals using nanopore-based long-read sequencing technology. *PLoS Neglected Tropical Diseases*, *17*(8), e0011586.

Delfini, C., Torreggiani, M., & Ottina, R. (1983). Fermentazioni maloalcoholiche spontanee di Schizosaccharomyces nei mosti di uve a maturazione precoce de Nord Italia. *Rivista di viticoltura ed enologia*, *12*, 29–38.

Dobzhansky, T. (1935). A critique of the species concept in biology. *Philosophy of Science*, *2*(3), 344–355.

Ebersberger, I., de Matos Simoes, R., Kupczok, A., Gube, M., Kothe, E., Voigt, K., & Von Haeseler, A. (2012). A consistent phylogenetic backbone for the fungi. *Molecular Biology and Evolution*, *29*(5), 1319–1334.

Etherington, G. J., Gil, E. G., Haerty, W., Oliferenko, S., & Nieduszynski, C. A. (2023). Schizosaccharomyces versatilis represents a distinct evolutionary lineage of fission yeast. Yeast. Advance online publication. https://doi.org/10.1002/yea.3919

Fantes, P. A., & Hoffman, C. S. (2016). A brief history of *Schizosaccharomyces pombe* research: A perspective over the past 70 years. *Genetics*, *203*(2), 621–629.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*(6), 368–376.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, *39*(4), 783–791.

Hayashibe, M., Sando, N., Awano, S., Kurasawa, J., & Takahashi, S. (1974). Yeasts associated with living plants in the sandy coastal area of the Sea of Japan in North-eastern provinces of Japan. *Bulletin of the Yamagata University. Agricultural Science*, *7*, 185.

Helston, R. M., Box, J. A., Tang, W., & Baumann, P. (2010). Schizosaccharomyces cryophilus sp. nov., a new species of fission yeast. *FEMS Yeast Research*, *10*(6), 779–786.

Hoffman, C. S., Wood, V., & Fantes, P. A. (2015). An ancient yeast for young geneticists: A primer on the *Schizosaccharomyces pombe* model system. *Genetics*, *201*(2), 403–423.

Hu, J., Fan, J., Sun, Z., & Liu, S. (2020). NextPolish: A fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*, *36*(7), 2253–2255.

Hu, J., Wang, Z., Liang, F., Liu, S., Ye, K., & Wang, D.-P. (2023). *NextPolish2: A repeat-aware polishing tool for genomes assembled using HiFi long reads*. bioRxiv. https://doi.org/10.1101/2023.04.26.538352

Hu, J., Wang, Z., Sun, Z., Hu, B., Ayoola, A. O., Liang, F., & Wang, S. (2023). *An efficient error correction and accurate assembly tool for noisy long reads*. bioRxiv. https://doi.org/10.1101/2023.03.09.531669

Jeffares, D. C., Rallis, C., Rieux, A., Speed, D., Převorovský, M., Mourier, T., Marsellach, F. X., Iqbal, Z., Lau, W., Cheng, T. M. K., Pracana, R., Mülleder, M., Lawson, J. L. D., Chessel, A., Bala, S., Hellenthal, G., O'Fallon, B., Keane, T., Simpson, J. T., … Bähler, J. (2015). The genomic and phenotypic diversity of *Schizosaccharomyces pombe*. *Nature Genetics*, *47*(3), 235–241. https://doi.org/10.1038/ng.3215

Jia, G. S., Zhang, W. C., Liang, Y., Liu, X. H., Rhind, N., Pidoux, A., Brysch-Herzberg, M., & Du, L. L. (2023). A high-quality reference genome for the fission yeast *Schizosaccharomyces osmophilus*. *G3: Genes|Genomes|Genetics*, *13*(4), 1–31. https://doi.org/10.1093/g3journal/jkad028

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589.

Kashyap, R. L., & Subas, S. (1974). Statistical estimation of parameters in a phylogenetic tree using a dynamic model of the substitutional process. *Journal of Theoretical Biology*, *47*(1), 75–101.

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, *20*(4), 1160–1166.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.

Klar, A. J., Ishikawa, K., & Moore, S. (2015). A unique DNA recombination mechanism of the mating/cell-type switching of fission yeasts: A review. *Mobile DNA III*, 515–528.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., & Marra, M. A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*, *19*(9), 1639–1645. https://doi.org/10.1101/gr.092759.109

Kudrjawzew, V. (1960). *Systematics of the Yeasts*. Akademie-Verlag.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549.

Kurtzman, C. P., & Robnett, C. J. (1998). Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenhoek*, *73*(4), 331–371.

Lachance, M.-A. (2018). CP Kurtzman's evolving concepts of species, genus and higher categories. *FEMS Yeast Research*, *18*(8), foy103.

Lachance, M.-A., Lee, D. K., & Hsiang, T. (2020). Delineating yeast species with genome average nucleotide identity: A calibration of ANI with haplontic, heterothallic Metschnikowia species. *Antonie Van Leeuwenhoek*, *113*, 2097–2106.

Lindner, P. (1893). Schizosaccharomyces pombe n. sp., a new agent causing fementation. *Wochenschrift für Brauerei*, *10*, 1298–1300.

Liu, Y., Leigh, J. W., Brinkmann, H., Cushion, M. T., Rodriguez-Ezpeleta, N., Philippe, H., & Lang, B. F. (2009). Phylogenomic analyses support the monophyly of Taphrinomycotina, including Schizosaccharomyces fission yeasts. *Molecular Biology and Evolution*, *26*(1), 27–34.

Lodder, J., & Kreger-van Rij, N. J. W. (1952). *The yeasts-a taxonomic study*. North Holland Publishing Company.

Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, *44*(W1), W54–W57. https://doi.org/10.1093/nar/gkw413

Martini, A. V. (1991). Evaluation of phylogenetic relationships among fission yeast by nDNA/nDNA reassociation and conventional taxonomic criteria. *Yeast*, *7*(1), 73–78.

Mayr, E. (2000). The biological species concept. In Q. D. Quentin, D. Wheeler, & R. Meier (Eds.), *Species concepts and phylogenetic theory: A debate* (pp. 17–29). Columbia university press.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534.

Molnar, M., & Sipiczki, M. (1993). Polyploidy in the haplontic yeast *Schizosaccharomyces pombe*: Construction and analysis of strains. *Current Genetics*, *24*, 45–52.

Neyman, J. (1971). Molecular studies of evolution: A source of novel statistical problems. In S. S. Gupta, & J. Yackel (Eds.), *Statistical decision theory and related topics* (pp. 1–27). Academic Press.

Phaff, H., Yoneyama, M., & do Carmo-Sousa, L. D. (1964). A one-year, quantitative study of the yeast flora in a single slime flux of Ulmuscarpinifolia Gled. *Rivista di Patologia Vegetale*, *4*, 485–497.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rhind, N., Chen, Z., Yassour, M., Thompson, D. A., Haas, B. J., Habib, N., Wapinski, I., Roy, S., Lin, M. F., Heiman, D. I., Young, S. K., Furuya, K., Guo, Y., Pidoux, A., Chen, H. M., Robbertse, B., Goldberg, J. M., Aoki, K., Bayne, E. H., … Nusbaum, C. (2011). Comparative functional genomics of the fission yeasts. *Science*, *332*(6032), 930–936.

Robinow, C. F., & Hyams, J. S. (1989). General cytology of fission yeast. In A. Nasim, P. Young, & B. F. Johnson (Eds.), *Molecular biology of the fission yeast* (pp. 273–330). Academic Press.

Rutherford, K. M., Harris, M. A., Oliferenko, S., & Wood, V. (2022). JaponicusDB: Rapid deployment of a model organism database for an emerging model species. *Genetics*, *220*(4), iyab223.

Samson, S., Lord, É., & Makarenkov, V. (2022). SimPlot++: A Python application for representing sequence similarity and detecting recombination. *Bioinformatics*, *38*(11), 3118–3120. https://doi.org/10.1093/bioinformatics/btac287

Scorzetti, G., Fell, J., Fonseca, A., & Statzell-Tallman, A. (2002). Systematics of basidiomycetous yeasts: A comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. *FEMS Yeast Research*, *2*(4), 495–517.

Seemann, T. (2018). *Barrnap 0.9: rapid ribosomal RNA prediction*. https://github.com/tseemann/barrnap

Shen, X.-X., Steenwyk, J. L., LaBella, A. L., Opulente, D. A., Zhou, X., Kominek, J., Li, Y., Groenewald, M., Hittinger, C. T., & Rokas, A. (2020). Genome-scale phylogeny and contrasting modes of genome evolution in the fungal phylum Ascomycota. *Science Advances*, *6*(45), eabd0079.

Shumate, A., & Salzberg, S. L. (2021). Liftoff: Accurate mapping of gene annotations. *Bioinformatics*, *37*(12), 1639–1643. https://doi.org/10.1093/bioinformatics/btaa1016

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212.

Sipiczki, M. (1995). Phylogenesis of fission yeasts. Contradictions surrounding the origin of a century old genus. *Antonie Van Leeuwenhoek*, *68*(2), 119–149.

Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? *Genome Biology*, *1*(2), 1–4.

Sipiczki, M. (2018). Interspecies hybridisation and genome chimerisation in Saccharomyces: Combining of gene pools of species and its biotechnological perspectives. *Frontiers in Microbiology*, *9*, 3071.

Sipiczki, M., & Ferenczy, L. (1977). Protoplast fusion of *Schizosaccharomyces pombe* auxotrophic mutants of identical mating-type. *Molecular and General Genetics MGG*, *151*(1), 77–81.

Sipiczki, M., & Ferenczy, L. (1978). Enzymic methods for enrichment of fungal mutants I. Enrichment of *Schizosaccharomyces pombe* mutants. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *50*(2), 163–173.

Sipiczki, M., & Kucsera, J. (1983). A possible control mechanism over conjugation and meiotic division in diploid cells of the fission yeast *Schizosaecharomyces pombe* var. malidevorans. Paper presented at the 6th European Cell Cycle Workshop, Prague.

Sipiczki, M., Kucsera, J., Ulaszewski, S., & Zsolt, J. (1982). Hybridization studies by crossing and protoplast fusion within the genus *Schizosaccharomyces lindner*. *Microbiology*, *128*(9), 1989–2000.

Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, *11*(1), 7.

Tamura, K., Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences*, *109*(47), 19333–19338.

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, *10*(3), 512–526.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027.

Tamura, K., Tao, Q., & Kumar, S. (2018). Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. *Molecular Biology and Evolution*, *35*(7), 1770–1782.

Tao, Y.-T., Suo, F., Tusso, S., Wang, Y.-K., Huang, S., Wolf, J. B. W., & Du, L.-L. (2019). Intraspecific diversity of fission yeast mitochondrial genomes. *Genome Biology and Evolution*, *11*(8), 2312–2329.

Tong, P., Pidoux, A. L., Toda, N. R. T., Ard, R., Berger, H., Shukla, M., Torres-Garcia, J., Müller, C. A., Nieduszynski, C. A., & Allshire, R. C. (2019). Interspecies conservation of organisation and function between nonhomologous regional centromeres. *Nature Communications*, *10*(1), 2343.

Vaughan-Martini, A., & Martini, A. (2011). Schizosaccharomyces Lindner (1893). In C. Kurtzman, J. Fell, & T. Boekhout (Eds.), *The yeasts (fifth edition)* (Vol. 2, pp. 779–784). Elsevier.

Vu, D., Groenewald, M., Szöke, S., Cardinali, G., Eberhardt, U., Stielow, B., de Vries, M., Verkleij, G. J. M., Crous, P. W., Boekhout, T., & Robert, V. (2016). DNA barcoding analysis of more than 9 000 yeast isolates contributes to quantitative thresholds for yeast species and genera delimitation. *Studies in Mycology*, *85*, 91–105.

Waterhouse, A. M., Procter, J. B., Martin, D.M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2: A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Wickerham, L. J., & Duprat, E. (1945). A remarkable fission yeast, Schizosaccharomyces versatilis nov. sp. *Journal of Bacteriology*, *50*(5), 597–607.

Yamada, Y., & Banno, I. (1987). Hasegawaea gen. nov., an ascosporogenous yeast genus for the organisms whose asexual reproduction is by fission and whose ascospores have smooth surfaces without papillae and which are characterized by the absence of coenzyme Q and by the presence of linoleic acid in cellular fatty acid composition. *The Journal of General and Applied Microbiology*, *33*(3), 295–298.

Yarrow, D. (1984). Schizosaccharomyces Lindner. In N. J. W. Kreger-van Rij (Ed.), *The yeast, a taxonomic study* (pp. 429–449). Elsevier Science Publishers B. V.

Yoon, S.-H., Ha, S., Lim, J., Kwon, S., & Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek*, *110*, 1281–1286.

Yukawa, M., & Maki, T. (1931). Schizosaccharomyces japonicus nov. spec. *La Bul. Sci. Falkultato Terkultura. Kjusu Imp Univ, Fukuoka, Japan*, *4*, 218–226.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brysch-Herzberg, M., Jia, G.-S., Sipiczki, M., Seidel, M., Zhang, W.-C., & Du, L.-L. (2024). Reinstatement of the fission yeast species *Schizosaccharomyces versatilis* Wickerham et Duprat, a sibling species of *Schizosaccharomyces japonicus*. *Yeast*, *41*, 108–127. https://doi.org/10.1002/yea.3922