Isolation, characterisation and genome assembly of Barnettozyma botsteinii sp. nov. and novel strains of Kurtzmaniella quercitrusa isolated from the intestinal tract of the termite Macrotermes bellicosus

Arrey, Gerard; Li, Guangshuo; Murphy, Robert; Guimaraes, Leandro; Alizadeh, Sefa; Poulsen, Michael; Regenberg, Birgitte

Published in:
G3: Genes, Genomes, Genetics

DOI:
10.1093/g3journal/jkab342

Publication date:
2021

Document version
Peer reviewed version

Document license:
CC BY

Citation for published version (APA):
Arrey, G., Li, G., Murphy, R., Guimaraes, L., Alizadeh, S., Poulsen, M., & Regenberg, B. (Accepted/In press). Isolation, characterisation and genome assembly of Barnettozyma botsteinii sp. nov. and novel strains of Kurtzmaniella quercitrusa isolated from the intestinal tract of the termite Macrotermes bellicosus. G3: Genes, Genomes, Genetics, [jkab342]. https://doi.org/10.1093/g3journal/jkab342
Isolation, characterisation and genome assembly of *Barnettozyma botsteinii* sp. nov. and novel strains of *Kurtzmaniella quercitrusa* isolated from the intestinal tract of the termite *Macrotermes bellicosus*

Gerard Arrey#, Guangshuo Li#, Robert Murphy, Leandro Guimaraes, Sefa Alizadeh, Michael Poulsen and Birgitte Regenberg*

---

© The Author(s) (2021). Published by Oxford University Press on behalf of the Genetics Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
ABSTRACT

Bioconversion of hemicelluloses into simpler sugars leads to production of a significant amount of pentose sugars, such as D-xylose. However, efficient utilization of pentoses by conventional yeast production strains remains challenging. Wild yeast strains can provide new industrially relevant characteristics and efficiently utilize pentose sugars. To explore this strategy, we isolated gut-residing yeasts from the termite *Macrotermes bellicosus* collected in Comoé National Park, Côte d’Ivoire. The yeasts were classified through their ITS/LSU sequence, their genomes were sequenced and annotated. We identified a novel yeast species, which we name *Barnettozyma botsteinii* sp. nov. T (MycoBank: 833563, CBS 16679T and IBT 710) and two new strains of *Kurtzmaniella quercitrusa*: var. *comoensis* (CBS 16678, IBT 709) and var. *filamentosus* (CBS 16680, IBT 711). The two *K. quercitrusa* strains grow 15% faster on synthetic glucose medium than *Saccharomyces cerevisiae* CEN.PK in acidic conditions (pH = 3.2) and both strains grow on D-xylose as the sole carbon source at a rate of 0.35 h⁻¹. At neutral pH, the yeast form of *K. quercitrusa* var. *filamentosus*, but not var. *comoensis*, switched to filamentous growth in a carbon source dependent manner. Their genomes are 11.0-13.2 Mb in size and contain between 4888 and 5475 predicted genes. Together with closely related species, we did not find any relationship between gene content and ability to grow on xylose. Besides its metabolism, *K. quercitrusa* var. *filamentosus* also has a large potential as a production organism, because of its capacity to grow at low pH and to undergo a dimorphic shift.
Keywords: Barnettozyma botsteinii, D-xylose, insect-associated yeasts, Kurtzmaniella quercitrusa, new yeast species.

INTRODUCTION

The biotechnology industry has historically relied on yeast species to carry out fermentation processes (Demain 2000). A major challenge today is the conversion of raw plant biomass directly into organic compounds of added value, such as organic acids or biogas (Lee et al. 2019; Qin et al. 2017; Martínez-Gutiérrez 2018). Lignocellulose material is mainly composed of cellulose (30-50%), hemicellulose (20-30%) and lignin (15-25%) (Kricka et al. 2014). In order to break down these complex polysaccharides, filamentous fungi, such as Trichoderma reesi, Aspergillus niger and Aspergillus oryzae, and their secreted enzymatic cocktails, have been used with success (Lange 2017; Mäkelä et al. 2014). Because these fungi do not efficiently ferment pentose sugars such as D-xylose and L-arabinose produced during depolymerization of hemicelluloses, yeast species or genetically engineered Saccharomyces cerevisiae strains have been suggested for downstream processes (Bettiga et al. 2009). However, co-fermentation of hexoses and pentoses is still a major challenge during efficient conversion of lignocellulosic plant biomass, mainly due to difficulties in integrating a new catabolic pathway within an existing metabolic network (Young et al. 2010).

There is therefore a need to identify and optimize yeasts that can efficiently ferment both pentoses and hexoses. One strategy is to isolate novel wild yeast strains that can metabolize these sugars, which can either serve as a production organism on their own, or as a genetic resource for metabolic pathways that can be engineered into industrial strains (Steensels et al. 2014).

The unicellular yeasts are commonly found in soil, on plants and flowers, and within the microbiome of insects (bees, beetles, termites and ants), where they establish a variety of
relationships, ranging from obligate to facultative mutualisms (Ganter 2006; Greig and Leu; 2009; Kurtzman et al. 2011; Buzzini et al. 2017). Described yeast species are found both in the Basidiomycota and Ascomycota phyla and are thought to represent only one percentage of the expected yeast diversity worldwide (Robert 2006). In addition, yeast species from tropical rainforest, desert or tundra ecosystems are underrepresented among characterised species (Kurtzman et al. 2015; Groenewald et al. 2017), and therefore serve as a reservoir for novel species with new characteristics. One example are the gut-residing yeasts from the genus of Candida, Pichia, Sporothrix, Debratomyces, which have been isolated from a variety of termite species (Schäfer 1996; Stefanini 2018). Here we explored yeasts within Macrotermes bellicosus termite guts to find novel yeast strains attractive from a biotechnological perspective and to shed light on insect-yeast relationships. M. bellicosus is a fungus-growing termite species (sub-family Macrotermitinae) that has an obligate symbiosis with the basidiomycete fungus Termitomyces (Basidiomycota: Lyophyllaceae) maintained on plant material harvested by the termites from outside the nest (Rouland-Lefèvre 2000). During a first passage through the termite gut, this plant biomass is mixed with asexual fungal spores and the resulting faeces is used to build a so-called fungus comb (Leuthold et al. 1989). The maturation of the comb allows the fungus to grow and the plant material to be degraded. Termites will ingest the mature comb and during a second gut passage obtain nutrition (Poulsen et al. 2014; da Costa et al. 2019). The combined enzymatic activities from the termite, the gut microbiome and the fungus accounts for the degradation of up to 90% of all dead wood in Kenya (Buxton 1981). Altogether, their foraging and digestive capabilities substantially contribute to carbon cycling in these ecosystems (Sugimoto et al. 2000) converting wood, leaves and roots (Collins 1981; Buxton, 1981; Smith et al. 2019).

In this work, we present and characterize two novel strains and one new species of yeast isolated from the gut of M. bellicosus. The two novel strains belong to the species Kurtzmaniella quercitrusa, phylum Ascomycota, and they were thus named K. quercitrusa var. comoensis (strain...
number in this work 1112) and K. quercitrusa var. filamentosus (strain number in this work 1120). We also identified a previously unknown species from the Ascomycota, which we name Barnettozyma botsteinii sp. nov. 1118T (strain number in this work 1118). All three strains assimilate pentose sugars, which makes them potentially suitable for biotechnological applications and interesting for further genomic and metabolic studies. Their degradation of plant derived polymers suggests a potential role in assisting termite and Termitomyces metabolism and support their potential for biotechnological production.

METHODS

Sample collection

We collected M. bellicosus minor workers from the inside of wild nests in Comoé National Park, Côte d’Ivoire during the month of November 2018. Specifically, the termite nests were located in the proximity of the Comoé National Park Research Station (8° 46’ 11″ N, 3° 47’ 21″ W), in the southwest end of the park next to the Comoé River. The gut of small workers (25 termites per nest) was dissected in situ and divided into midgut, hindgut and foregut (Fig. 1) (Bos et al., 2020). Five to ten replicates were pooled together in order to increase the starting material and manual tissue dissociation was performed with a pestle in sterile 1X phosphate buffered saline (PBS) supplemented with Tween 80 (5 mg/ml). The resulting lysates were diluted (1/100 and 1/1000) and spread on Potato Dextrose agar plates. The medium was prepared by diluting 39 g/l of PDA (Sigma Aldrich, 70139) in water, autoclaved and supplemented with 0.1 g/l of chloramphenicol, 0.05 g/l of streptomycin sulphate and 0.035 g/l of ampicillin before pouring it into plates. Plates were then incubated at 30°C until growth was observed. In order to screen for yeast species, single colonies were first streak out and subsequent incubated with YPD medium supplemented with 6% ethanol at room temperature until growth was observed. Isolates with a yeast-like morphology under the microscope were then subjected to genomic DNA extraction and rDNA sequencing.
ITS and D1/D2 sequencing and phylogeny

For genomic DNA (gDNA) extraction, a standardized protocol for yeasts was applied to all the isolates. Briefly, yeast cell wall was broken using glass beads and vortexing. Buffers P1, P2 and N3 (Qiagen, Cat. No.: 19051, 19052 and 19064) were used to isolate the DNA fraction, which was further precipitated by the addition of isopropanol. Finally, after an ethanol wash, pellets were resuspended in TE buffer (10 mM Tris-HCl pH=8, 1mM EDTA) and DNA concentration was assessed by Nanodrop. Primers ITS4 (5’-TCCTCCGCTTATTGATATGC) and ITS1 (5’-CTTGGTCATTAGAGGAGTAA) (Kuiper-Goodman & Scott, 1989) were used to amplify Internal Transcribed Spacer (ITS) regions while primers NL1 (5’-GCATATCAATAAGCGGAGGAAAAG) and NL4 (5’-GGTCCGTGTTTCAAGACGG) were used to amplify the Large Subunit of the nuclear ribosomal RNA (LSU rRNA) gene region (Kurtzman and Robnett 1997). Both PCR products were purified, Sanger sequenced and results aligned to NCBI database using BLASTn (Altschul et al. 1990). Nucleotide sequences for a set of the most closely yeast species were retrieved from GenBank (Table S1). Alignment was performed with Mafft v.7.471 (http://mafft.cbrc.jp/alignment/software/) and a Maximum-Likelihood phylogeny based on ITS and LSU was constructed in MEGA v. X (Kumar et al. 2018). Branch support was obtained from bootstrap analysis with 1000 repetitions.

Genome sequencing and analyses

We sequenced and assembled the genomes of isolates 1118, 1112 and 1120 in the following steps. Code used is provided in File S1.

DNA extraction, sequencing and assembly

Whole-genome sequencing of isolate 1118, 1112 and 1120 was done by firstly extracting DNA using a scaled-up CTAB extraction (Poulsen et al., 2014). Whole-genome sequencing was performed using a combination of 100bp/150bp paired-end shotgun (BGIseq/DNBseq) and long-
read (PacBio sequel) sequencing by BGI. The resulting sequences were checked for quality using FastQC v0.11.9 (Andrews et al., 2012). Genomes were assembled using three different approaches and the best result was chosen for subsequent analyses. First, short and long reads were assembled together and used to generate scaffolds using SPAdes 3.14.1 (Nurk et al., 2013). Second, long reads were assembled into scaffolds with Canu 2.0 (Koren et al., 2017) and then polished by the short reads using NextPolish 1.3.1 (Hu et al., 2020). Lastly, short reads were assembled into scaffolds alone, also using SPAdes 3.14.1 (Nurk et al., 2013). Contigs that were shorter than 500bp were discarded in the final assemblies, and assembly quality was assessed by quantifying genome completeness based on the expected gene content of the Benchmarking Universal Single-Copy Orthologs (BUSCO), version 4.4.1, against the database for fungal genomes (Seppey et al., 2019).

**Genome annotation**

Genomes were annotated using MAKER v3.01.03 (Holt & Yandell, 2011). Two *ab initio* gene predictors were used with the MAKER pipeline: SNAP v2013-11-29 (Leskovec & Sosič, 2016) and AUGUSTUS v3.3.3 (Stanke et al., 2004), each of which was trained for individual genomes. The training of both SNAP and AUGUSTUS requires pre-existing gene models as training data. Therefore, an initial MAKER analysis was carried out where gene annotations were generated directly from homology evidence without using the *ab initio* gene predictors. The resulting gene annotations supported by homology evidence were then used to train SNAP and AUGUSTUS. Once both *ab initio* gene predictors were trained, they were used together with homology evidence in a full MAKER analysis. Homology evidence was only used to inform gene predictions. Resulting gene models supported by homology evidence were used to re-train SNAP and AUGUSTUS. A second round analysis was conducted using the newly trained SNAP and AUGUSTUS parameters, and once again the resulting gene models with homology supports were used to re-train SNAP and AUGUSTUS. Finally, a third round of MAKER analysis was performed using the new SNAP and
AUGUSTUS parameters. All resulting gene models are reported, and these comprise the final set of annotated genes.

Functional gene annotation was performed using InterProScan version 5.48-83.0 (Jones et al., 2014) with annotation of Gene ontology (GO) terms (Ashburner et al., 2000), KEGG pathways (Kanehisa et al., 2021), Pfam 33. (Mistry et al., 2021). The predicted proteins sequences were blasted against NCBI-NR (O’Leary et al., 2016) and Swiss-Prot databases (Bateman, 2019) with an E value cutoff of 1e-5. Non-coding RNA (ncRNA) genes were identified by BAsic Rapid Ribosomal RNA Predictor v0.9, Barrnap (Seemann, 2018) and tRNA genes were predicted using the tRNAScan-SE v2.0.5 (Lowe & Chan, 2016) algorithm with default parameters. Repeat sequences were identified and classified using RepeatMasker v4.1.0 (Smit et al. 2019). The criterion used for LTR family classification was that the 5’LTR sequences should share at least 80% identity over at least 80% of their length for the same family.

**MLST analysis**

We performed Multi-Locus-Sequence-Typing (MLST) using a set of Universal Single-Copy Orthologous genes for all genomes that were identified with BUSCO v4 (Seppey et al., 2019) with default settings and using the fungi_odb10.2019-11-20 dataset. We collated 344 (of 758 potential) genes present in all 48 genomes and generated multi-fasta files of the orthologous genes (Table S2). These multi-fasta files were aligned with Clustal Omega v1.2.4 (Sievers et al., 2011). Phylogenies were generated using RAxML-NG v0.9.9 (Kozlov et al., 2019), employing the –all mode, GTR+G model and a seed of 2. Branch support based on bootstrapping and transfer distance were obtained, before generating a randomly rooted consensus tree using ASTRAL-Pro v1.12 (Zhang et al., 2020).

**Yeast media and growth conditions**
Strains were grown and kept on YPD medium (10 g/L of Yeast Extract, 20 g/L of Bacto-Peptone and 20 g/L of Glucose) at 30°C. To study growth dynamics and carbon source assimilation, strains were grown and kept on minimal medium (1.6 g/L of Yeast Nitrogen Base without amino acids, 5 g/L of Ammonium Sulphate with a pH of 6, supplemented with 1% glucose or 1% D-xylose). To check D-xylose assimilation in different concentrations and aeration conditions, single colonies from minimal glucose agar plates were used to inoculate glass-test tubes (max. vol. 35 mL) at an OD$_{600}$ = 0.15, containing minimal medium with 2% or 10% D-xylose at different volumes: 3 mL (high aeration), 7.5 mL (low aeration) or 20 mL (microaeration). The tubes were then kept at 30°C with shaking (150 rpm) for 14 days. To calculate the growth rate in D-xylose and D-glucose medium, starter cultures were prepared by growing single colonies from glucose or xylose minimal medium agar plates into corresponding liquid medium buffered to a pH = 3.2 or pH = 6 overnight at 30°C. The next day, the overnight cultures were used to inoculate 50 mL medium to an OD$_{600}$ ~ 0.1 in 250 mL baffled Erlenmeyer flasks containing the same buffered medium. For both carbon sources, the medium was buffered to a pH = 3.2 with 0.1M final concentration of phosphoric buffer (1M KH$_2$PO$_4$ adjusted to pH = 3.2 with phosphoric acid) or to a pH = 6 with 0.1M final concentration of potassium phosphate buffer (pH = 6). Cultures were incubated at 30°C and 150 rpm. To calculate each $\mu$, OD$_{600}$ measurements from two different single colonies were used. Each individual OD$_{600}$ measurement was obtained by the average of triplicates. As previously described, colonization patterns were made on semi-solid YPD 0.3% agar plates and cultures were grown for 14 days at room temperature (C. P. Kurtzman et al., 1991).

**Carbon assimilation**

Carbon-source assimilation was assessed using microplates (Cletus P. Kurtzman, Fell, Boekhout, et al., 2011) using Phenotypic Microarray plates (Biolog, PM1 #12111 and PM2 #12112). Briefly, the plates were inoculated with strains 1112, 1118 and 1120 in buffer PM containing Dye mix E to facilitate and improve detection, incubated at 30°C for 144h and the absorbance measured at 550
nm, as described by the manufacturer. Values at +144h were used to assess growth on each of the carbon sources. For each time point, background absorbance was subtracted and the fold change (FC) over negative control (no carbon source) was calculated. Signals were assigned to a strong growth (FC >3 at 144h, +++), intermediate growth (FC >2 and ≤3 at 144h, ++), weak growth (FC >1 and ≤2) and no growth (FC ≤1). Strains were seeded in duplicate plates in two independent experiments. Mean of the fold change was calculated.

**Microscopy**

Yeast cells for microscopy were grown overnight in YPD medium at 30°C at 150 rpm. Cells were washed once with distilled sterile water and pellets were re-suspended in water. 1-5 μL were mounted in a glass slide with a 0.17 mm thin cover slip. Samples were visualized using a Nikon Eclipse E600 and pictures taken with an Optronics MagnaFire CCD Microscope Camera.

**Xylose metabolism gene analysis**

We performed detailed analyses of the genes putatively involved in xylose metabolism by comparing the newly assembled genomes to 27 genomes available from GenBank. We inferred the potential for D-xylose utilization if the following orthologues were present: Xylose reductase (XYL1), xylitol dehydrogenase (XYL2), and D-xylulokinase (XYL3), which translate D-xylose to xylulose-5P, where the xylulose-5P enters the pentose phosphate pathway (PPP). Furthermore, we included genes from the pentose phosphate pathway oxidative phase involving the enzymes glucose 6-phosphate dehydrogenase (ZWFI), 6-phosphogluconate dehydrogenase (GND1), and a non-oxidative phase carried out by D-ribulose-5-phosphate 3-epimerase (RPE1), ribose-5-phosphate ketol-isomerase (RKII), transketolase (TKLI) and transaldolase (TALI), while phosphoglycerate isomerase (PGII) completes the cycle. We counted the number of copies of the genes encoding these enzymes in the genome and visualised differences across yeasts in a heatmap generated in Excel. We obtained and aligned genes coding for these enzymes from all 27 genomes using Mafft.
v.7.471 (http://mafft.cbrc.jp/alignment/software/) and subsequently predicted the secondary protein structure using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/).

RESULTS

Phylogenetic analysis of yeast isolates

We collected worker termites of the species *M. bellicosus* in Comoé National Park (Côte d’Ivoire) and dissected their intestines in foregut, midgut and hindgut for yeast isolation (Bos et al., 2020). We identified a total of 20 individual yeast colonies spanning all three different gut compartments and four independent termite colonies (Fig. 1). To identify novel species and subspecies, we next characterised the isolates taxonomically by sequencing the ITS and D1/D2 rDNA regions of LSU and aligning the sequences to GenBank repository using BLASTN (Altschul et al. 1990) (Fig. 1).

From the set of 20 isolates, 11/20 were assigned to *Meyerozyma caribbica*. Next, 3/20 isolates were assigned to *Candida* sp. TMS-2011 and the rest (6/20 isolates) did not have a clear match with any NCBI sequence, showing similarities between 84.4% and 99% to their respective best matches (Fig. 1). From the six uncategorized isolates, four (IDs: 1112, 1113, 1105 and 1106) showed three gaps with *K. quercitrusa* KMC-Y23, isolate 1120 showed one mismatch with *K. quercitrusa* WHFS-3 and isolate 1118 showed 12 gaps and 37 mismatches with *Barnettozyma* sp. EN30S02, leading to a similarity of 91%. Because 1112, 1113, 1105 and 1106 had identical rDNA D1/D2 sequences, we grouped them and selected 1112 as a representative of this group (Fig. 1).

The best Maximum-Likelihood (ML) model based on ITS and LSU was TN93+G. The ML tree allowed us to include all closely related yeast species within the genera (Fig. 2). *B. botsteinii* sp. nov. (1118) clusters as an independent branch with *Barnettozyma salicaria* NRRL Y-6780T (Fig. 2a) while *K. quercitrusa* var. *comoensis* (1112) and var. *filamentosus* (1120) cluster with *K. quercitrusa* type strain CBS 4412T (Fig. 2b).
To gain more confidence on its phylogenetic placement, we sequenced the genomes of isolates 1118, 1112 and 1120. We generated a Multiple-Locus-Sequence-Typing (MLST) analysis based on a set of 1706 BUSCO genes inferred to be Ascomycota-specific and single-copy in at least 85% of the available yeast genomes with high (<80%) genome completeness (for general genome features and completeness, see Table 1). The genome sizes and gene content of 1112 and 1120 are similar to Kurtzmaniella cleridarum, which has a 12.1Mb large genome and contains 5478 genes, while 1118 is 11.1Mb with 4888 genes; comparable to the closest relative with a genome available (Barnettozyma salicaria; 11.01Mb, 5586 genes). The MLST phylogeny thus confirmed the placement of these new isolates in their respective genera and clades (Fig. 2).

**Metabolic profile characterization**

To characterize their carbon metabolism, we grew strains 1112, 1120 and 1118 in an array of carbon sources (Table S4). Metabolism was detected as a function of respiratory activity (production of NADH) by monitoring the colour change in the medium (reduction of tetrazolium dye). From the 190 carbon sources tested, *K. quercitrusa* var. comoensis (1112) metabolized among others pentose sugars (2-reoxy-D-ribose, D-arabinose, L-arabinose, D-mannose, D-ribose, L-lyxose and D-xylose), hexose sugars (α-D-glucose, D-fructose, D-galactose and D-tagatose), disaccharides (isomaltulose, maltose, sucrose, gentiobiose, palatinose and turanose), trisaccharides (D-melezitose and maltotriose), organic acids (5-keto-D-gluconic acid, citramalic acid, D,L-malic acid, δ-amino-α-lactic acid, γ-amino butyric acid, L-malic acid, N-acetyl-neuraminic acid succinic acid), Tween (Tween 20, Tween 40 and Tween 80), etc. (See complete list in Table S4).

*K. quercitrusa* var. filamentosus (1120) metabolized among others pentose sugars (2-reoxy-D-ribose, D-arabinose, L-arabinose, D-ribose and D-mannose and L-lyxose), (hexose sugars (α-D-glucose, D-fructose, D-galactose), disaccharides (D-trehalose and gentiobiose), amino acids (L-
alanine, L-asparagine, L-aspartic acid, L-glutamine, L-isoleucine, L-lysine, L-proline and L-pyroglutamic acid), etc. (See complete list in Table S4).

*B. botsteinii* sp. nov. 1118\(^T\) metabolized among others pentose sugars (2-reoxy-D-ribose, L-arabinose, D-ribose, D-xylose, L-lyxose), methyl pentose (L-rhamnose), hexose sugars (α-D-glucose, D-glucose-1-Phosphate, D-fructose, D-mannose, D-psicose), disaccharides (D-cellobiose, gentiobiose, maltose), trisaccharides (maltotriose), sugar alcohols (D-mannitol, D-sorbitol, glycerol), organic acids (5-keto-D-gluconic acid, α-keto-glutaric acid, acetic acid, citric acid, fumaric acid, pyruvic acid, L-malic acid, succinic acid), Tween (Tween 20, Tween 40 and Tween 80), etc. (See complete list in Table S4).

We then compared the metabolic profile characterization of the 1118 with other known and closely related *Barnettozyma* and *Candida* species such as *B. salicaria, B. siamensis* and *C. montana*. The result revealed that it has different metabolic characteristics, as it can assimilate D-ribose, L-arabinose, gelatine and maltose, but not salicin and sucrose (Table S5). Similarly, the comparison of the physiology of 1112 and *K. quercitrusa* CBS 4412 showed that 1112 differed by being able to assimilate D-arabinose, D-xylose, inulin, L-arabinose, N-acetyl-D-glucosamine and salicin, but not D-glucosamine, D-mannitol, D-trehalose, glycerol and xylitol. In contrast to *K. quercitrusa* CBS 4412, 1120 grows on D-arabinose, gelatine, L-arabinose and gelatine but not D-mannitol, maltose, sucrose and L-sorbose (Table S5).

**Morphology description**

*Kurtzmaniella quercitrusa* var. *comoensis*

Strain ID in this work: 1112. Microscopic examination of this isolate confirmed that it grew as yeast. Cells appear as black, oval shaped and were 3-4 \(\mu\)m long and 2-3 \(\mu\)m wide during exponential growth in rich (YPD) and minimal glucose medium (see Methods section). Cells divide
mitotically by a small bud that fully separates from the mother cell (Fig. 3a). Colonies on YPD agar are white, bright and round, while they appear as a flat and round with an even rim on medium with a low concentration of agar, 0.3% YPD agar (Fig. 3d).

*Kurtzmaniella quercitrusa* var. *filamentosus*

Strain ID in this work: 1120. Yeast cells of this strain are clearly distinguishable from *K. quercitrusa* var. *comoensis*. Being 5-6 μm in length and 1-2 μm wide, cells appear as longer and thinner than 1112 (Fig. 3b). Moreover, during exponential growth in glucose based medium, YPD and minimal glucose medium (see Methods section), some cells do not fully separate after division leading to formation of 5-6 cell long pseudohyphal structures. After 5 days of growth in liquid 1% xylose medium, formation of pseudohyphae is strongly enhanced, leading to truly septated hyphae. Hyphae first appear as germ tubes (Fig. 4a), which are 2-7 cell long chains growing from a single round cell. They later elongate and mature to true hyphal structures showing bifurcations (Fig. 4c and 4c’). Colonies on rich glucose based medium (YPD) and minimal glucose medium show a rough, butyrous and ivory coloured phenotype. In addition, their colonization pattern on 0.3% agar plates is non-even with visible edges and with some growth elevated from the agar (see Methods section) (Fig. 3e).

*Barnettozyma botsteinii* sp. nov.

*Barnettozyma botsteinii* sp. nov. 1118T yeast cells are mostly spherical and small (3-4 μm x 2-3 μm) during exponential growth in liquid rich medium (YPD) and minimal glucose medium (see Methods section). They are commonly observed as clusters of 3-5 cells forming pseudohyphal structures (Fig. 3c). No sign of true hyphae formation was observed in any condition. Although colonies on YPD or minimal glucose medium appear as smooth, the colonization on glucose 0.3% YPD agar leads to an edgy pattern (Fig. 3f). The colonies are ivory coloured and flat. The relatively
low growth rate observed on liquid medium (Table 2) was also reflected in smaller colonies on solid medium (both on YPD and on SC-1% glucose).

D-xylose metabolism

After observing that the novel isolates grow on pentose sugars, we next characterised their D-xylose utilization, using different aeration and concentrations of sugar in the growth medium. Strains *K. quercitrusa* var. *comoensis* (strain 1112) and *K. quercitrusa* var. *filamentosus* (strain 1120) show an aeration-dependent growth preference on D-xylose (Fig. 5). From the conditions tested, the highest biomass is reached with high aeration in combination with 2% D-xylose, where the OD\(_{600}\) after 14 days of growth reaches 16.6 ± 0.6 and 14.9 ± 1.0, respectively. For both *K. quercitrusa* strains, we observed a correlation between aeration levels and OD\(_{600}\). A decrease in aeration level correlates with a decrease in OD\(_{600}\) both with 2% D-xylose and 10% D-xylose. For example, in low aeration conditions 1112 grows to OD\(_{600}\) of 5.8 ± 0.2 with 2% D-xylose and 5.8 ± 0.2 with 10% D-xylose.

Similarly, 1120 grows to an OD\(_{600}\) of 5.4 ± 0.8 with 2% D-xylose and 2.3 ± 0.6 with 10% D-xylose. We observed a further reduction in the OD\(_{600}\) in microaeration conditions, especially in 1112 strain, where it reaches only 3.2 ± 0.29 in 2% D-xylose and 2.4 ± 0.1 in 10% D-xylose.

Strain *B. botsteinii* sp. nov. 1118\(^\text{T}\), on the other hand, did not efficiently grow on D-xylose. Only marginal to very limited growth was observed with microaeration and high D-xylose concentration (Fig. 5). Interestingly, *K. quercitrusa* var. *filamentosus* (strain 1120) undergoes a dimorphic shift and forms hyphae in 2% D-xylose medium combined with high or low aeration (Fig. 5). The dimorphic shift consisted of formation of pseudohyphae and hyphae of various lengths. No changes in morphology were observed with the others strains.

We next compared the growth rates of *K. quercitrusa* when propagate in D-xylose and D-glucose at low and neutral pH (Table 2). *K. quercitrusa* var. *comoensis* (strain 1112) cells grew in 1%
synthetic glucose medium with a doubling time of 2.0 hours at pH = 3.2. At pH = 6, the doubling time was similar at 1.9 hours. In 1% xylose medium, cells grew at a doubling time of 19.3 hours, while they exhibited no growth at pH=3.2. Cells grew as single yeast cells and divided by budding through exponential phase at both carbon source medium (Table 2). For comparison, a S. cerevisiae CEN.PK type strain grew only in synthetic glucose medium at a doubling time of 2.3 hours at pH = 3.2 while at pH = 6 it grew with a doubling time of 2.1 hours.

K. quercitrusa var. filamentosus (strain 1120) grew on 1% glucose with a doubling time of 2.0 hours at both tested pH (3.2 and 6.0). In xylose, the strain showed pH-dependent morphologic differences (Table 2). At pH = 3.2, the strain grew as single yeast cells with a doubling time of 19.6 hours, while at pH = 6.0 medium, pseudohyphae accumulated and matured into true hyphae (Table 2 and Fig. 5). Both strains of K. quercitrusa (strains 1112 and 1120) showed very similar growth rates in 1% glucose medium and slightly lower than S. cerevisiae CEN.PK<sup>T</sup> strain (Table 2).

B. botsteinii sp. nov. 1118<sup>T</sup> grew in 1% glucose medium at pH = 6 with doubling time of 2.8 hours (Table 2). Interestingly, the strain was sensitive to acidic pH that drastically increased its doubling time of 25.7 hours at pH = 3.2. Compared to the K. quercitrusa strains and S. cerevisiae CEN.PK<sup>T</sup>, this species showed the slowest growth rate at both pHs (Table 2). Under this experimental growth condition, the strain was not able to grow on 1% xylose medium at either of the tested pHs (Table 2).

Copy number changes of genes coding for enzymes in metabolic pathways are known to affect metabolic pathway efflux (Steenwyk & Rokas, 2018). Therefore, we investigated the copy number of genes involved in xylose and pentose metabolism. Our comparative genomics analysis for XYL and PPP genes revealed that the vast majority of yeasts contained all the orthologues necessary for
xylose metabolism (Fig. 6). The lack of growth of some of the species on xylose thus does not appear to be as a result of the absence of xylose-metabolizing genes (Fig. 6).

**Taxonomy**

Description of *Barnettozyma botsteinii* sp. nov. 1118

MycoBank: MB 833563

*Barnettozyma botsteinii* (bɔtʃtani, N.L. gen. n. botsteinii, named after scientist David Botstein) isolated from the hindgut of termite *Macrotermes bellicosus* in Côte d’Ivoire. Holotype is preserved at the Westerdijk Institute as CBS 16679^T^ and an isotype at the IBT Culture Collection of Fungi as IBT 710. Cells appear as spherical and small (3-4 µm x 2-3 µm), typically forming clusters of 3-5 cells in pseudohyphal structures. No sign of true hyphae formation was observed in the conditions tested. The colonies appear ivory coloured, flat and circular in YPD or minimal 1% glucose medium. However, at low agar (0.3% agar) glucose medium the edge becomes undulate (Fig. 3).

Cells are able to grow in liquid or solid YPD at 22-30°C. The cells grow in 1% glucose medium at pH = 6 with a growth rate of 0.24 h^-1±0.02) and doubling time of 2.8 hours. Their growth rate is decreased to 0.03 h^-1±0.005) corresponding to a doubling time of 25.7 hours at pH = 3.2 (Table 2).

Based on its genome, phylogeny and physiology analysis, we confirm that isolate 1118 is clearly distinct from any other known *Barnettozyma* species, we assign the status of sp. nov. to isolate 1118 and thus, propose *Barnettozyma botsteinii* sp. nov. 1118^T^ as the founder and type strain of a novel yeast species. We name the species after Professor David Botstein, a true inspirer, great geneticist and important figure for the yeast research community (D. Botstein et al., 1980) (David Botstein & Fink, 1988).

**DISCUSSION**

In this study, we isolated novel and biotechnologically relevant yeast strains from termite guts. We describe two novel strains of *Kurtzmaniella quercitrusa* and a new species of yeast, which we name
*Barnettozyma botsteinii* based on genomic information. The criteria to classify these three novel taxa follows Kurtzman et al. 2011, in which conspecific strains differ by no more than three nucleotides, whereas distinct species show six or more nucleotide differences (Kurtzman et al. 2011b).

Here, the novel yeast species *B. botsteinii* sp. nov. 1118\(^T\) is described based on a single isolate. Novel single-strain fungal taxa descriptions are debated as complete intraspecific variability (including genetic and phenotypic) and its ecological role, cannot be fully captured using one strain. However, nearly one-third of described yeast species are based on single isolates (Kurtzman et al. 2011a). The documentation and publishing of single strains novel taxa serves as a nucleation point for further studies and the possibility of combining information from strains independently isolated in time. While it was not possible for us to collect new material nor find any other available strain with a complete match in GenBank, it is possible that more isolates of *B. botsteinii* sp. nov. are found in the future. This debate was recently reviewed in (Brysch-Herzberg et al., 2020).

We find that both novel strains of *K. quercitrusa* are able to grow on D-xylose, and their overall metabolic profile makes them interesting for further applied studies. The complex phenotypic change that *K. quercitrusa* var. *filamentosus* (strain 1120) undergoes (filamentation when growing on certain carbon sources), reveal the existence of a metabolic shift that can control alterations in protein secretion and enzyme production similar to that of other yeasts with a dimorphic shift (Gauthier 2017; Celińska and Nicaud 2019).

We identified two novel strains and a new species of yeast by studying a relatively small number of individuals (around 100 termites from only one termite species, *M. bellicosus*). Other yeast strains, without a clear match to described species, have previously been isolated from *Macrotermes subhyalinus* (Yoro et al. 2014) and other fungus-growing termites (*Odontotermes formosanus*)
(Mathew et al. 2012). Therefore, the Macrotermitinae termite sub-family, comprising 330 described species, appears as a potentially extensive biological reservoir for new yeast species and strains of biotechnological relevance.

*K. quercitrusa* was first isolated from the insect frass on an oak tree (*Quercus* sp.) (Meyer and Yarrow 1998), type strain CBS 4412^T^.

Other isolates of the same species have been collected from other insects (beetles) and from flowers around the world (*e.g.*, Oceania, South America and Southeast Asia) (Crous et al. 2004; Molnár et al. 2008). Although generally considered non-pathogenic, adult and paediatric human infections have been reported in immunocompromised patients, widening the spectrum of non-albicans candidiasis (Xiao et al. 2014; Westblade et al. 2015).

Yeasts are widely used in biotechnological production of metabolites and protein. However, most do not metabolize pentose sugars well. The two new *K. quercitrusa* strains (var. *comoensis* 1112 and var. *filamentosus* 1120) grow well on hexose sugars: glucose, fructose, mannose and galactose and are also able to assimilate pentose sugars such as ribose, xylose, arabinose and lyxose. They show differences in growth on short oligosaccharides in which var. *comoensis* grows on maltose and sucrose while var. *filamentosus* grow on trehalose and adonitol. Furthermore, they both can grow in the presence of complex plant-derived polysaccharides such as mannan, dextrin and laminarin. The metabolic profile resemble that of *K. quercitrusa* type strain CBS 4412^T^, including xylose assimilation (Kurtzman et al. 2011). Strain var. *filamentosus* is different to var. *comoensis* in its ability to differentially metabolize both monosaccharides (adonitol) and disaccharides (D-trehalose). The filamentous growth of var. *filamentosus* has been reported in another member of *K. quercitrusa*, strain NRRL Y-5392, although described as pseudohyphae (Kurtzman et al. 2011). In addition, other members of the *Kurtzmaniella* clade, such as *Candida railenensis*, also show hyphae formation under certain conditions (Ramírez and González 1984; Kurtzman et al. 2011). We observed that this phenotype is highly variable and dependent on factors which are common, for
example, among _Candida_ species, namely: carbon source and concentration, pH and aeration level (Caplice and Moran 2015).

We also describe a new yeast species, named _Barnettozyma botsteinii_ sp. nov. 1118, closely related to _Barnettozyma salicaria_ NRRL Y-6780\(^T\). Metabolically, _B. botsteinii_ grows on an array of hexose and pentose sugars. Notably, the strain showed positive growth on short organic acids: acetic and glyoxylic acids (2 carbons), pyruvic acid (3C), fumaric and malic acids (4C), \(\alpha\)-ketoglutaric acid (5C) and citric acid (6C). Because many of these are part of the Krebs cycle, this metabolic profile suggests that this species could be classified as Krebs-positive (Barnett and Kornberg 1960; Casal et al. 2008). The isolated strain grew in culture conditions although at relatively low rates, suggesting that we did not find the best growth conditions for this species. _B. botsteinii_ could be dependent on the environment provided by the termite for optimal growth like some other known gut-associated yeasts, such as _Coccidiascus legeri_, which have an obligate relationship with their host insect and are unable to grow on any tested laboratory conditions (Kurtzman et al. 2011) (Stefanini, 2018).

The annotation and analyses of genes involved in xylose metabolism did not spur differences in the novel strains and other known yeasts, suggesting that the phenotypic variation in growth in this group of yeasts is unlikely to be driven by differences in the gene content. Although no effect of gene copy number was observed, other genetic factors could influence gene expression and explain the observed phenotypic diversity such as single-nucleotide polymorphisms (SNPs), small inversions-deletions or translocations. Additionally, differences in balance of key metabolic components and co-factors can influence enzyme activities, for example NADH/NAD+, which important for xylose metabolism (Riley et al., 2016). This suggests that the capacity to assimilate and metabolise xylose may be ancient and conserved in yeasts.
A question that remains open is whether these yeast species stably colonize the termite gut or are ingested during foraging. *K. quercitrusa* strain CBS 4412^T^ was isolated from insect frass and we isolated *K. quercitrusa* var. *filamentosus* from the termite hindgut (*i.e.*, last compartment before faeces deposition), suggesting that at least this particular yeast species can survive gut passage, a feature not common to all yeasts nor insect guts (Madden et al. 2018). Gut passage is exploited by other yeast species for their dispersal and breeding (Reuter et al. 2007; Stefanini et al. 2012; Madden et al. 2018) and it is possible that *K. quercitrusa* and *B. botsteinii* are also dependent on gut passage for growth, mating and dispersal.

Biotechnologically, the strains and species characterised here show interesting features that warrant further exploration. *K. quercitrusa* growth on xylose makes them interesting for industrial applications, especially if they are able to co-ferment xylose in the presence of hexoses (Hahn-Hägerdal et al. 2007; Ceccato-Antonini et al. 2017). Especially, the characteristics of *K. quercitrusa* var. *filamentosus* make it suited for biotechnologically production, since growth at low pH, metabolism of pentose sugars and a controllable filamentous state are particularly desired traits. In addition, all three strains appear to be able to metabolize polysorbate, also known as Tween. This is a complex molecule formed by polyoxyethyleneated sorbitol molecules with lauric acid and used as a detergent and surfactant. Other yeasts such as *Candida albicans* or *Candida parapsilosis* sustain growth on Tween as a sole carbon source, a trait considered as a virulence factor as its lipid content triggers expression of extracellular phospholipases, including *LIP1* (Fu et al. 1997; Slifkin 2000).

In addition, other metabolic processes other than pentose consumption are also biotechnologically attractive. The cellobiose respiration by *B. botsteinii* sp. nov. 1118^T^ suggests the presence of β-glucosidases enzymes (EC 3.2.1.21), which are able to break down the cellobiose molecule into single glucose molecules. During depolymerization of cellulose, cellobiose is produced which inhibits cellulose-degrading enzymes (cellobiose 1,4-β-cellobiosidase and endoglucanases), reducing...
the overall efficiency (Murphy et al. 2013; Sørensen et al. 2013). Thus, the addition of β-glucosidases to the process can greatly improve its efficiency (Woodward and Wiseman 1982).

In conclusion, the intestine of fungus-growing termites appears to be a promising niche for the identification of new yeasts strains and species. The metabolic profiles of the yeasts isolated from termite gut reflect the adaptations to a plant-based diet found either within the wild or within the gut of the termite, and these gut-residing yeasts of *Macrotelmes* termites have interesting traits of relevance to the biotechnology industry.

**Data availability**

Sanger sequences of D1/D2 rDNA regions of LSU and ITS for isolate 1112, 1120 and 1118 are deposited in the GenBank under the accessions MN497060.1, MN497056.1, MN509221.1, MN497061.1, MN497057.1 and MN509222.1. Genomes of three strains are deposited in GenBank under the BioProject number PRJNA695125 and accession number are JAFIDI00000000, JAFIU00000000 and JAFIDH00000000. The strains are available as follows: *Barnettozyma botsteinii* sp. nov. 1118\(^T\) (MycoBank: 833563, CBS 16679\(^T\) and IBT 710), *Kurtzmaniella quercitrusa* var. *comoensis* 1112 (CBS 16678, IBT 709) and *Kurtzmaniella quercitrusa* var. *filamentosus* 1120 (CBS 16680, IBT 711). Supplemental Material available at figshare: [https://doi.org/10.25387/g3.14899278](https://doi.org/10.25387/g3.14899278).

**ACKNOWLEDGMENTS**

We would like to thank N’golo A. Kone, Kolotchéléma Simon Silue and the staff at the Comoé National Park Research Station in Côte d’Ivoire for assistance with collections. We thank Josefin Stiller, Thomas Læssøe and Bolette Rud. Pallesen for their advice on phylogenetics, morphological description and phonetics.
DECLARATIONS

Funding
This work was supported by an ERC Consolidator grant (771349) to MP.

Conflict of interest
The authors declare that they have no conflict of interest.

Authors’ contributions
Termite collection and dissection was performed by Leandro Guimaraes. Yeast isolation, identification and characterization was performed by Gerard Arrey and Sefa Alizadeh. Guangshuo Li performed genome analyses and Robert Murphy performed MLST. Michael Poulsen and Birgitte Regenberg raised funding for the project and designed the study, with input from all authors. The first draft of the manuscript was written by Gerard Arrey and Birgitte Regenberg. All authors commented on previous versions of the manuscript, and read and approved the final manuscript.

Ethical statement
All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. Termites were collected under permit 189/UNA/CRE/SREC issued by Universite Nangui Abrogoua and Research Station Comoé on the 12.12.2018. M. bellicosus is not an endangered species of insect. This article does not contain any studies with human participants performed by any of the authors.

REFERENCES


https://doi.org/10.1186/s12862-020-01727-z


https://doi.org/10.1155/2017/8491383


https://doi.org/10.1186/1471-2105-12-491


Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J.,

Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M.,

Yong, S. Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function


Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat

scalable and user-friendly tool for maximum likelihood phylogenetic inference. *35*(21), 4453–
4455. https://doi.org/10.1093/bioinformatics/btz305

enzyme production for degradation of cellulose and hemicellulose from biomass: A

Kuiper-Goodman, T., & Scott, P. M. (1989). Risk assessment of the mycotoxin ochratoxin A. In


https://doi.org/10.1016/B978-0-444-52149-1.00010-0


https://doi.org/10.1093/femsyr/fov050


the evergreen rainy Valdivian forest of southern Chile. Mycopathologia, 88(2–3), 99–103.

Reuter, M., Bell, G., & Greig, D. (2007). Increased outbreeding in yeast in response to dispersal by
an insect vector. In Current Biology (Vol. 17, Issue 3).

Coughlan, A. Y., Deshpande, S., Douglass, A. P., Hanson, S. J., Klenk, H. P., … Jeffries, T.
Sci. U. S. A. https://doi.org/10.1073/pnas.1603941113

In Biodiversity and Ecophysiology of Yeasts (pp. 31–44). Springer-Verlag.

Ecology (pp. 289–306). https://doi.org/10.1007/978-94-017-3223-9_14


Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H.,
of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol.,
7(1), 539. https://doi.org/10.1038/msb.2011.75


http://www.repeatmasker.org/


Torsten, S. (2013). BarNap 0.2 - Bacterial ribosomal RNA predictor.


https://doi.org/10.1016/0141-0229(82)90084-9


https://doi.org/10.1128/JCM.00246-14


https://doi.org/10.1093/molbev/msaa139
**LIST OF FIGURES AND TABLES**

**Fig. 1.** Origin and identification of yeast isolates. Small workers termites were collected in Comoé National Park, in the north east of Côte d’Ivoire (highlighted area in the map) from four *Macrotermes bellicosus* termite colonies (IC0021, IC0027, IC0031 and IC0034) (see Methods). Guts were dissected into foregut, midgut and hindgut and yeast species were identified by comparing their ITS rDNA sequences to strains in the NCBI GenBank repository. According to their classification and origin, we grouped the isolates and assigned an internal ID. GenBank sequences IDs denoted with * were generated in this study. Otherwise, sequences from the reported isolates were found to be identical to the GenBank IDs provided. Map image modified from © OpenStreetMap contributors.

**Fig. 2** Phylogenetic analyses based on rDNA LSU and ITS sequence alignments. Maximum-likelihood phylogenies of the new yeast isolates and their closest species in GenBank (Table S1 for sequence IDs) for *Barnettozyma botsteinii* CBS 16679T (a) and novel strains of *Kurtzmaniella quercitrusa* (b). Branches show result of bootstrap analysis (1000 replicates). The evolutionary distances were calculated using Kimura 2-parameter method and scale bar is set at 0.08 and 0.06 substitutions per site, respectively. The analysis was performed using MEGA v. X (Kumar et al. 2018).

**Fig. 3** Morphology description. Bright field microscopy pictures (left images) of cells grown in YPD at exponential phase and colonization pattern on YPD 0.3% agar (right images) of *K. quercitrusa* var. *comoensis* (a and d), *K. quercitrusa* var. *filamentosus* (b and e) and *B. botsteinii* sp. nov. 1118T (c and f). Colonies were grown at room temperature for 14 days. Bars show respective scales.

**Fig. 4** Filamentous growth of *K. quercitrusa* var. *filamentosus* (1120). Cells grown in synthetic liquid medium 1% xylose (pH = 6) grow as filaments. Single yeast colonies from xylose agar plates
were grown in 1% xylose liquid medium buffered at pH = 6 (see methods). After 48-72 hours at 30°C at 150 rpm, small cell chains are visible (a and b) elongating either from single cells (*) or a cluster of cells (**). Eventually, after 96h – 120 h, these structures develop to longer septate filaments (c and c’) with bifurcations (arrowhead in c’). Bright-field microscopy images at indicated scale.

**Fig. 5** Effect of aeration and sugar concentration on D-xylose growth. OD<sub>600</sub> measurements at selected times in microaeration (gray), low aeration (orange), high aeration (blue) in 2% D-xylose (continuous line) or 10% D-xylose (discontinuous line) conditions of *K. quercitrusa* var. *comoensis* (1112) (top), *K. quercitrusa* var. *filamentosus* (1120) (middle) and *B. botsteinii* sp. nov. 1118<sup>T</sup> (bottom). Representative bright-field microscopy images at indicated scale of cells growing at low aeration (2% or 10% D-xylose) are shown below each graph. Error bars show SD of three OD<sub>600</sub> measurements. Growth experiments were repeated twice.

**Fig. 6** Multiple-Locus-Sequence-Typing (MLST) analysis. In bold, species and strains described in this study. Consensus maximum-likelihood distance tree based on 344 single copy orthologous genes with non-parametric bootstrapping with individual gene trees built in RAxML-NG using a GTR+G model. Heatmap of identified number of genes involved in the assimilation and metabolism of xylose across selected yeast species. The heat map indicates that the novel strains and species are comparable to other yeast species (see text for details). Color code indicates number of orthologous genes (top gene panel) from four to none. Positive (+), negative (-), weak (W), variable (V) and slow (S). Details and references of strains used can be found in Table S6. *The species was transferred from *Candida* to *Kurtzmaniella* genus (Lopes et al., 2019).
Table 1. Genome characteristics of the three new genomes. BUSCO (Benchmarking Universal Single-Copy Orthologs), see methods.

<table>
<thead>
<tr>
<th>Species</th>
<th>Assembly size (Mb)</th>
<th>#contigs</th>
<th>N50 (Mb)</th>
<th>%GC</th>
<th>#genes</th>
<th>#ncRNAs</th>
<th>#Repetitive Elements</th>
<th>BUSCO complete</th>
<th>BUSCO duplicated</th>
<th>BUSCO fragmented</th>
<th>BUSCO missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. botsteinii sp. nov. 1118</td>
<td>11.1</td>
<td>37</td>
<td>1.263</td>
<td>35.0</td>
<td>4888</td>
<td>277</td>
<td>3871</td>
<td>1464</td>
<td>1</td>
<td>55</td>
<td>186</td>
</tr>
<tr>
<td>K. quercitrusa var. comoensis</td>
<td>13.0</td>
<td>12</td>
<td>1.732</td>
<td>39.9</td>
<td>5439</td>
<td>361</td>
<td>2337</td>
<td>1466</td>
<td>13</td>
<td>18</td>
<td>209</td>
</tr>
<tr>
<td>K. quercitrusa var. filamentosus</td>
<td>13.2</td>
<td>17</td>
<td>1.456</td>
<td>45.5</td>
<td>5475</td>
<td>213</td>
<td>3094</td>
<td>1471</td>
<td>4</td>
<td>26</td>
<td>205</td>
</tr>
</tbody>
</table>

Table 2. Growth rates in defined liquid media. Growth rate ($\mu_{\text{max}}$, hours$^{-1}$) and doubling times (DT, hours) of CEN.PK, B. botsteinii sp. nov. 1118, K. quercitrusa var. nov. 1112 and K. quercitrusa var. nov. 1120 in liquid SC-1% glucose and SC-1% xylose at pH 3.2 or 6. All growth rates are averages of two independent experiments from two independent clones.

<table>
<thead>
<tr>
<th></th>
<th>SC-1% GLUCOSE</th>
<th>SC-1% XYLOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pH = 3.2</td>
<td>pH = 6</td>
</tr>
<tr>
<td></td>
<td>$\mu_{\text{max}}$ ($\pm$ST)</td>
<td>DT</td>
</tr>
<tr>
<td>CEN.PK</td>
<td>0.30 ($\pm$0.03)</td>
<td>2.3</td>
</tr>
<tr>
<td>B. botsteinii sp. nov. 1118</td>
<td>0.03 ($\pm$0.005)</td>
<td>25.7</td>
</tr>
<tr>
<td>K. quercitrusa var. nov. 1112</td>
<td>0.35 ($\pm$0.04)</td>
<td>2.0</td>
</tr>
<tr>
<td>K. quercitrusa var. nov. 1120</td>
<td>0.35 ($\pm$0.02)</td>
<td>2.0</td>
</tr>
</tbody>
</table>
### Yeast isolation

**M. bellicosus**

![Map of Côte d’Ivoire with termite colony illustration](https://example.com)

<table>
<thead>
<tr>
<th>Termite colony</th>
<th>Gut compartment</th>
<th>Species</th>
<th>Ref. Strain</th>
<th>GenBank Seq. ID</th>
<th>N° of isolates</th>
<th>Isolate ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC0021</td>
<td>Foregut</td>
<td>Meyerozyma caribbica</td>
<td>CBS 9966</td>
<td>NR_149348.1</td>
<td>1</td>
<td>1102</td>
</tr>
<tr>
<td></td>
<td>Foregut</td>
<td>Meyerozyma caribbica</td>
<td>IMUFRJ 51968</td>
<td>FN428887.1</td>
<td>1</td>
<td>1103</td>
</tr>
<tr>
<td>IC0027</td>
<td>Foregut</td>
<td>Meyerozyma caribbica</td>
<td>CBS 9966</td>
<td>MK394109.1</td>
<td>1</td>
<td>1107</td>
</tr>
<tr>
<td></td>
<td>Midgut</td>
<td>Meyerozyma caribbica</td>
<td>CBS 9966</td>
<td>MK394109.1</td>
<td>4</td>
<td>1114, 1115, 1116, 1117</td>
</tr>
<tr>
<td></td>
<td>Midgut</td>
<td>Kurtzmaniella quercitrusa var. comoensis</td>
<td>CBS 16678</td>
<td>MN497061.1*</td>
<td>4</td>
<td>1112, 1105, 1106, 1113</td>
</tr>
<tr>
<td></td>
<td>Hindgut</td>
<td>Meyerozyma caribbica</td>
<td>IMUFRJ 51968</td>
<td>FN428887.1</td>
<td>2</td>
<td>1101, 1104</td>
</tr>
<tr>
<td></td>
<td>Hindgut</td>
<td>Candida sp.</td>
<td>3TSM-2011</td>
<td>HQ631010.1</td>
<td>3</td>
<td>1111, 1119, 1109</td>
</tr>
<tr>
<td></td>
<td>Hindgut</td>
<td>Meyerozyma caribbica</td>
<td>CBS 9966</td>
<td>MK394109.1</td>
<td>2</td>
<td>1108, 1110</td>
</tr>
<tr>
<td>IC0031</td>
<td>Hindgut</td>
<td>Barnettzyma botsteinii sp. nov.</td>
<td>CBS 16679</td>
<td>MN509221.1*</td>
<td>1</td>
<td>1118</td>
</tr>
<tr>
<td>IC0034</td>
<td>Hindgut</td>
<td>Kurtzmaniella quercitrusa var. filamentosus</td>
<td>CBS 16680</td>
<td>MN497056.1*</td>
<td>1</td>
<td>1120</td>
</tr>
</tbody>
</table>

**Figure 1**

Yeast isolation diagram showing the movement of yeasts through the termite gut compartments.
Figure 2
Figure 5

K. quercitrusa var. comoensis (1112)

K. quercitrusa var. filamentosus (1120)

B. botsteinii sp. nov. 1118

Figure 5