

Type of the Paper: Article

Development of a transformation method for *Metschnikowia borealis* and other CUG-serine yeasts.

Zachary Gordon^{1,2}, Maximillian P.M. Soltysiak³, Christopher Leichthammer¹, Jasmine A. Therrien¹, Rebecca S. Meaney¹, Carolyn Lauzon¹, Matthew Adams¹, Dong Kyung Lee³, Preetam Janakirama¹, Marc-Andre Lachance³, Bogumil J. Karas^{1,2,*}

¹Designer Microbes Inc., London, ON, N6G 4X8, Canada

²Department of Biochemistry, Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON, N6A 5C1, Canada

³Department of Biology, University of Western Ontario, London, ON, N6A 5B7, Canada

*Correspondence: bkaras@uwo.ca

Abstract: Yeasts belonging to the *Metschnikowia* genus are particularly interesting for the unusual formation of only two needle-shaped ascospores during their mating cycle. Presently, the meiotic process that can lead to only two spores from a diploid zygote is poorly understood. The expression of fluorescent nuclear proteins should allow the meiotic process to be visualized *in vivo*; however, no large-spored species of *Metschnikowia* has ever been transformed. Accordingly, we aimed to develop a transformation method for *Metschnikowia borealis*, a particularly large-spored species of *Metschnikowia*, with the goal of enabling the genetic manipulations required to study biological processes in detail. Genetic analyses confirmed that *M. borealis*, and many other Metchnikowiaceae, are CUG-Ser yeasts. Codon-optimized selectable markers lacking CUG codons were used to successfully transform *M. borealis* by electroporation and lithium acetate, and transformants appeared to be the result of random integration. Mating experiments confirmed that transformed-strains were capable of generating large asci and undergoing recombination. Finally, random integration was used to transform an additional 18 yeast strains, and all attempts successfully generated transformants. The results provide a simple method to transform many yeasts from an array of different clades and can be used to study or develop many species for various applications.

Keywords: genome engineering; synthetic biology; yeasts; *Metschnikowia*; genetic tools; DNA delivery; CUG-Ser

1. Introduction

For decades, yeasts have proven to be tremendously useful in the study of life’s fundamental processes. As single-celled models for eukaryotic life, they have short generation times, can be easily manipulated, and contain thousands of conserved genes—many of which are vitally important across nearly all living organisms[1,2]. While the discovery of new species provides us with new opportunities, genetic tools previously developed for *Saccharomyces cerevisiae* and other model organisms must be re-optimized for use in newly discovered species[3,4]. Altered membrane composition, modified gene expression, and alternative codon usage can prevent the seamless application of established transformation methods onto new species. Such transformation methods are crucial in the development of targeted knockouts, changes in gene expression, and fusion proteins that allow the study of biological processes in detail.

Of particular interest among yeasts are the 81 or more species assigned to the genus *Metschnikowia*, some of which have become popular for their varied applications[5]. For example, *M. pulcherrima* is of importance to winemakers for its generation of pleasant aromas in high-quality wines, and its anti-microbial properties further its potential as an industrial microorganism[6,7].

Similarly, *M. bicuspidata* has been studied extensively for its role as a pathogen, and other species (including *M. agaves* and *M. hawaiiensis*) are being studied as potential anti-aging agents in cosmetics[5].

Metschnikowia species are particularly intriguing due to the unusual formation of only two needle-shaped ascospores during their sexual cycle. A two-spored meiotic product stands in sharp contrast to the four spores arising from meiosis in many other yeast species (like *Saccharomyces cerevisiae*)[5]. Although the occurrence of meiotic recombination has been demonstrated in some *Metschnikowia* species[8] the fate of the nuclei during meiosis remains unclear, and is yet to be elucidated. The expression of fluorescent nuclear fusion proteins should allow the meiotic process to be visualized *in vivo*, but the generation of such fusions is greatly hindered by the lack of genetic tools available for *Metschnikowia* species. *M. pulcherrima* is the only species ever to have been transformed[9]. Unfortunately, the small size of its ascospores and the difficulty of generating abundant asci make *M. pulcherrima* less than ideal for the study of meiotic nuclei.

In contrast, *M. borealis* is one of many *Metschnikowia* species that form ascospores that can reach 20-50 times the size of normal budding cells[10]. The giant, elongated spores would greatly facilitate the task of elucidating the meiotic process in *Metschnikowia* species. The initial selection of *M. borealis* for study is further justified by the fact that the species has the highest maximum growth temperature among the large-spored *Metschnikowia* species (37°C), making it more likely to survive the high-temperature incubations required by many of the conventional yeast transformation methods[11,12]. Therefore, we aimed to develop and optimize a transformation method for *M. borealis* to enable the study of the species in more detail.

Our approach builds on the conventional methods used to transform other yeast species such as *S. cerevisiae* and *M. pulcherrima*. After identifying antibiotics that can be used as selectable markers, we developed various genetic constructs carrying the appropriate genes to allow growth on selective plates. We then tested conventional transformation techniques, such as lithium acetate transformation and electroporation to introduce the various genetic cassettes[11]. It was originally anticipated that differences in membrane composition, antibiotic sensitivity, and origins of replication may pose significant obstacles in our ability to seamlessly apply classical transformation methods from *S. cerevisiae* onto *M. borealis*. Accordingly, we developed a systematic approach entailing a variety of different transformation methods, using a variety of different antibiotic resistance markers to develop a successful transformation method for *Metschnikowia borealis* and other CUG-Ser yeasts.

2. Materials and Methods

Microbial strains and growth conditions

Metschnikowia borealis strains UWOPS 96-101.1 (*MAT α*) and SUB 99-207.1 (*MAT a*), and other strains that were used in transformation experiments (Figure 4) were obtained from the yeast collection of the Department of Biology, University of Western Ontario, where they are kept frozen in liquid nitrogen. *Saccharomyces cerevisiae* VL6-48 (ATCC MYA-3666: *MAT α his3- Δ 200 trp1- Δ 1 ura3-52 lys2 ade2-1 met14 cir⁰*). All strains were grown in YPAD broth (1% yeast extract, 2% peptone, 0.01% adenine hemisulfate, 2% D-glucose) at 30°C with shaking at 225 rpm, with the exception of *M. bicuspidata* and *M. orientalis*, which were grown at 27°C. Stationary cultures were grown on the same media containing 1% agar at the same growth temperatures.

Codon-optimization of KanMX and Sh *ble*

To codon-optimize the KanMX and Sh *ble* selectable markers, each CTG codon was changed to a TTG, and the resulting gene sequence was uploaded to the OPTIMIZER tool

(<http://genomes.urv.es/OPTIMIZER/>)[13]. The codon frequencies from *M. borealis* in three highly expressed yeast genes—*ADH1*, *TDH3*, and *ENO1*—were used to guide optimization. A table was generated using the Codon Usage Database (<https://www.kazusa.or.jp/codon/>) and entered into OPTIMIZER to guide codon-optimization. The method used for optimization was one amino acid to one codon, with a maximum of 25 allowed codon changes. Optimized genes were called MbKanMX and MbShBle.

Yeast DNA isolation

Genomic DNA was isolated from *M. borealis* using a modified alkaline lysis method. Individual colonies were suspended in 1.0 mL of sterile double-deionized water (sddH₂O) and pelleted at 3000 × g for 5 minutes. Cells were resuspended in 250 µL P1 Buffer (Qiagen), 0.25 µL of 14 M β-mercaptoethanol, and 12.5 µL zymolyase solution (10 mg mL⁻¹ zymolyase 20T in 25% glycerol), and incubated at 37°C for 1 hour. Following incubation, 250 µL of P2 Buffer (Qiagen) was added, gently mixed, and the suspension was incubated at room temperature for 5 minutes. Subsequently, 250 µL P3 Buffer (Qiagen) was added, gently mixed, and the suspension was incubated on ice for 10 minutes. Cellular debris was pelleted at max speed for 10 minutes at 4°C, and the supernatant was transferred to a separate tube. Next, 750 µL ice-cold isopropanol was added to the supernatant, and the mixture was frozen at -80°C for 15 minutes. The DNA was pelleted at top speed for 10 minutes at 4°C, washed with 750 µL ice-cold 70% ethanol, and dissolved in 50 µL Tris-EDTA (TE) pH 8.0.

Lithium acetate transformations

The lithium acetate transformations were conducted as described previously, with some modifications[14]. Briefly, a 50-mL cell culture was grown to an OD₆₀₀ of 1.0 and pelleted at 3000 × g for 5 minutes. The pellet was washed 3 times with and resuspended in 1 mL sterile water. Next, 100 µL of cells were transferred to a sterile tube and pelleted to remove the supernatant. The pellet was resuspended in 360 µL of transformation mix (33% PEG 3350 with 0.1 M LiOAc, 100 µg salmon sperm carrier DNA, and 1 µg DNA), and vortexed briefly. The mixture was incubated with shaking at 39°C and 300 rpm for 30 minutes, and then pelleted to remove the transformation mix. The pellet was resuspended in 1 mL of YPAD in 1.5 mL Eppendorf tube, recovered for 2 hours at 30°C and 225 rpm, plated on the appropriate selective media, and incubated at 30°C for 2 days.

Electroporation conditions

Electroporation experiments were conducted as previously described[15], with slight modifications. Briefly, cells were grown to an OD₆₀₀ of 1.5 in 50 mL YPAD and pelleted at 5000 × g for 5 minutes. The pellet was resuspended in 10 mL 0.1 M LiOAc in 1×TE, pH 7.5, and incubated with shaking at 150 rpm and 30°C for 1 hour. Next, 250 µL 1 M DTT was added, and cells were returned to the incubator for 30 minutes. After adding 40 mL ice-cold sterile water, cells were pelleted, washed with 25 mL ice-cold sterile water, washed again with 5 mL ice-cold 1 M sorbitol, and resuspended in 250 µL 1 M sorbitol. Electro-competent cells were placed at -80°C in 100-µL aliquots. When needed, cells were quickly thawed in a 37°C water bath and immediately placed on ice. Electroporation reactions were prepared in 0.2 cm cuvettes with 40 µL electro-competent cells, 5 µL DNA (1 µg), and 20 µg single-stranded carrier DNA, and conditions were 1.8 kV, 200 Ω, and 25 µF. Cells were collected from each cuvette in 1 mL of YPAD and allowed to recover in Eppendorf tubes for 2 hours at 30°C and 225 rpm. For each transformation, 20 µL were plated on the appropriate selective media and incubated at 30°C for 2 days.

Locating the insertion site

Metschnikowia borealis cells were transformed with PCR-linearized CaNAT1 flanked by 60 base-pair sequences of homology to the HIS3 promoter and terminator, and DNA was isolated from 5

colonies by Alkaline Lysis as described above. DNA samples were digested with *CfoI* (Promega), with expected cut sites every ~600 base-pairs in the *M. borealis* genome and no cut sites within CaNAT1. Digested samples were then ligated using T4 DNA Ligase (New England BioLabs). Qiagen Multiplex (Cat No. 206143) was used to PCR-amplify the unknown sequences surrounding the insertion site, using primers that bind within CaNAT1 and that were oriented to amplify the surrounding sequence. PCR conditions were 95°C melting for 30s, 60°C annealing for 90s, and 72°C extension for 90s, cycled 35 times. PCR products were sent for sequencing at the DNA Sequencing Facility and Robarts Research Institute (London, Ontario). To identify insertion sites, sequences adjacent to the inserted marker were analyzed using BLASTn with Geneious 11.1.5 to search the *M. borealis* genome[16]. For secondary verification, primers were designed to anneal ~150 base-pairs upstream and downstream of the insertion site, and PCR was used to observe the addition of the 693-base-pair cassette (Supplementary Figure S3).

Mating Experiments

M. borealis MATa was transformed with MbShBle, and used in mating experiments with a MATα strain that had been transformed with CaNAT1. MbShBle-transformed MATa was thinly spread on 2% agar plates containing ½ strength glucose media lacking histidine and uracil (Teknova). CaNAT1-transformed MATα was then spread perpendicular to MATa, creating a grid with many sections where both strains were mixed. The cells were incubated for 2 days at room temperature. Approximately 20% of the cells were subsequently collected and resuspended in 25 mL of YPAD, and then grown at 30°C for 6 hrs with shaking at 225 rpm. Next, 1-mL aliquots of culture were plated on YPAD containing 200 mg L⁻¹ zeocin and 75 mg L⁻¹ nourseothricin and incubated at 30°C. After two days, resulting colonies were collected, serially diluted, and plated on YPAD with 200 mg L⁻¹ zeocin and 75 mg L⁻¹ nourseothricin. Purified recombinant colonies were grown overnight in YPAD, serially diluted, and plated alongside wildtype MATa, wildtype MATα, MbShBle-transformed MATa, and CaNAT1-transformed MATα (Figure 3C).

3. Results

3.1. Identification of antibiotics that inhibit growth of *M. borealis*

Before designing genetic cassettes (containing selection marker flanked by homology/promoter/terminator sequences) for transformation, it was necessary to identify antibiotics that inhibit the growth of *M. borealis* to enable us to choose appropriate antibiotic resistance markers. Cultures of *M. borealis* MATa and MATα were grown to OD₆₀₀ of 1.5, pelleted, and resuspended to an OD₆₀₀ of 3.0. Cultures were spot-diluted in 5-μL aliquots on YPAD agar plates with various concentrations of geneticin, zeocin, or nourseothricin (Supplementary Table S1); and it was determined that each antibiotic was effective at inhibiting growth at 400, 125, and 50 mg L⁻¹, respectively. To further prevent the occurrence of spontaneous colonies, higher concentrations of nourseothricin (75 mg L⁻¹ or 100 mg L⁻¹) and zeocin (200 mg L⁻¹) were used for selection.

3.2. Growth rate of *M. borealis*

Many conventional yeast transformation methods require that a culture be grown to a specific optical density at 600 nm; accordingly, it was necessary to determine the growth rate of *M. borealis*, to enable us to accurately calculate the growth time to the required optical density. Three cultures of each mating type of *M. borealis* (MATa and MATα) were grown to mid-log phase (OD₆₀₀ = 1.0), and diluted in 50 mL of YPAD to an optical density of 0.065. Each culture was grown at 30°C with shaking at 225 rpm, and OD₆₀₀ was recorded at 1-hr time points for 8 hours. The average OD₆₀₀ of the three cultures for each mating type was recorded, and the doubling time was calculated as 76 minutes for mating type a and 74 minutes for mating type α (Supplementary Figure S1).

3.3. *M. borealis* codon usage

A series of transformation attempts were made with *M. borealis* using genetic cassettes containing the NAT, KanMX, and Sh *ble* genes; however, no colonies were ever obtained for any of the transformation attempts. While investigating factors that would prevent us from obtaining colonies of *M. borealis*, the question of genetic code became significant. Some species (including *M. bicuspidata* and *M. fructicola*) have been reported to recognize a CUG codon as a serine (yeast alternative nuclear code)[17,18], which fits within the current view that the Metschnikowiaceae belong to the CUG-Ser 1 clade of budding yeasts[19]. However, the only published literature regarding the genetic code of *M. pulcherrima* indicated that the species translates the CUG codon to leucine (standard code)[20]. The NAT, KanMX, and Sh *ble* cassettes used in the transformation attempts contain 8, 4, and 5 CUG codons in their open reading frames, respectively.

In an attempt to determine the genetic code of *M. borealis*, the coding sequence of *URA3* from *M. borealis* was translated with the ExPASy translate tool, using both the standard code and yeast alternative nuclear code, and aligned with the 30 top hits from a *URA3*-BLASTx search using MUSCLE (Figure 1A). A leucine aligned at a highly conserved serine or threonine position when *URA3* was translated with the standard genetic code; however, a serine aligned at that position when translated with the yeast alternative nuclear code.

An examination of available Metschnikowiaceae genome sequences by tRNAscan-SE v2.0 [21] revealed that most species encoded a single tRNA with a CAG anticodon. In exception, *M. pulcherrima* had two copies, which is consistent with its diploid genome. The retrieved tRNA sequences were analyzed through the tRNAscan-SE on-line web server, aimed at estimating the most probable translated amino acid. All *Metschnikowia* species that were examined had serine as the most probable translation with scores of 85 or more. Genome sequences for *Candida bentonensis* or relatives are not available. Further evidence arose from aligning the sequence for tRNA_{CAG} from *M. borealis* with the tRNA_{CAG}^{Ser} of related CUG-Ser clade yeast species (Figure 1B). The tRNA_{CAG} from *M. borealis* and other Metschnikowiaceae contain the serine discriminator base and identity elements, indicating that the tRNAs identified are in fact tRNA_{CAG}^{Ser}. No tRNA_{CAG}^{Leu} could be identified in any of the *Metschnikowia* species that were analyzed, including *M. pulcherrima*.

3.4. Transformation of *M. borealis* with codon-optimized cassettes

OPTIMIZER was used to codon-optimize the NAT, KanMX, and Sh *ble* genes for expression in *M. borealis* based on the codon frequencies in three highly expressed yeast genes: *ADH1*, *TDH3*, and *ENO1*. The optimized genes were translated with the ExPASy translate tool, using both the standard code and yeast alternative nuclear code, and aligned with BLAST to ensure that the same protein product would be produced. No differences were found between the wildtype gene products and the optimized gene products, regardless of whether the optimized genes were translated with the standard code or yeast alternative nuclear code. The optimized NAT gene was substantially similar to CaNAT1, a gene optimized for expression in *Candida albicans*[22]. Due to its previous success in transforming *C. albicans*, another CUG-Ser species, CaNAT1 was used in this experiment. Codon-optimized KanMX and Sh *ble* genes were called MbKanMX and MbShble.

A second set of antibiotic resistance genes were created by reintroducing the same 8, 4, and 5 CUG codons found in the wildtype NAT, KanMX, and Sh *ble* genes, respectively, as a control for codon usage. All six codon-optimized genes (three in standard code and three in yeast alternative nuclear code) were ordered as separate G-block fragments from Integrated DNA Technologies. All genes were amplified with primers containing homologous hooks to the *M. borealis* *ADH1* promoter and terminator and were used to transform the species by electroporation. Transformants were plated on YPAD with appropriate concentrations of each antibiotic (75 mg L⁻¹ neorseothricin for CaNAT1, 200 mg L⁻¹ zeocin for MbShBle, and 400 mg L⁻¹ geneticin for MbKanMX) alongside negative

controls. Several colonies appeared for all samples that were transformed using the yeast alternative nuclear code, but no colonies appeared on any of the other plates (Figure 2). Similar results were seen using our best lithium acetate protocol. Selected transformants from the electroporation experiment were verified by genotyping (Supplementary Figure S2). Transformation efficiencies were determined for each method by transforming *MATa* and *MATα* by each method in triplicate, and were calculated to be approximately 1600 and 817 CFUs per μg of DNA for electroporation, and 9 and 4 CFUs per μg of DNA for lithium acetate transformation, respectively (Supplementary Table S2).

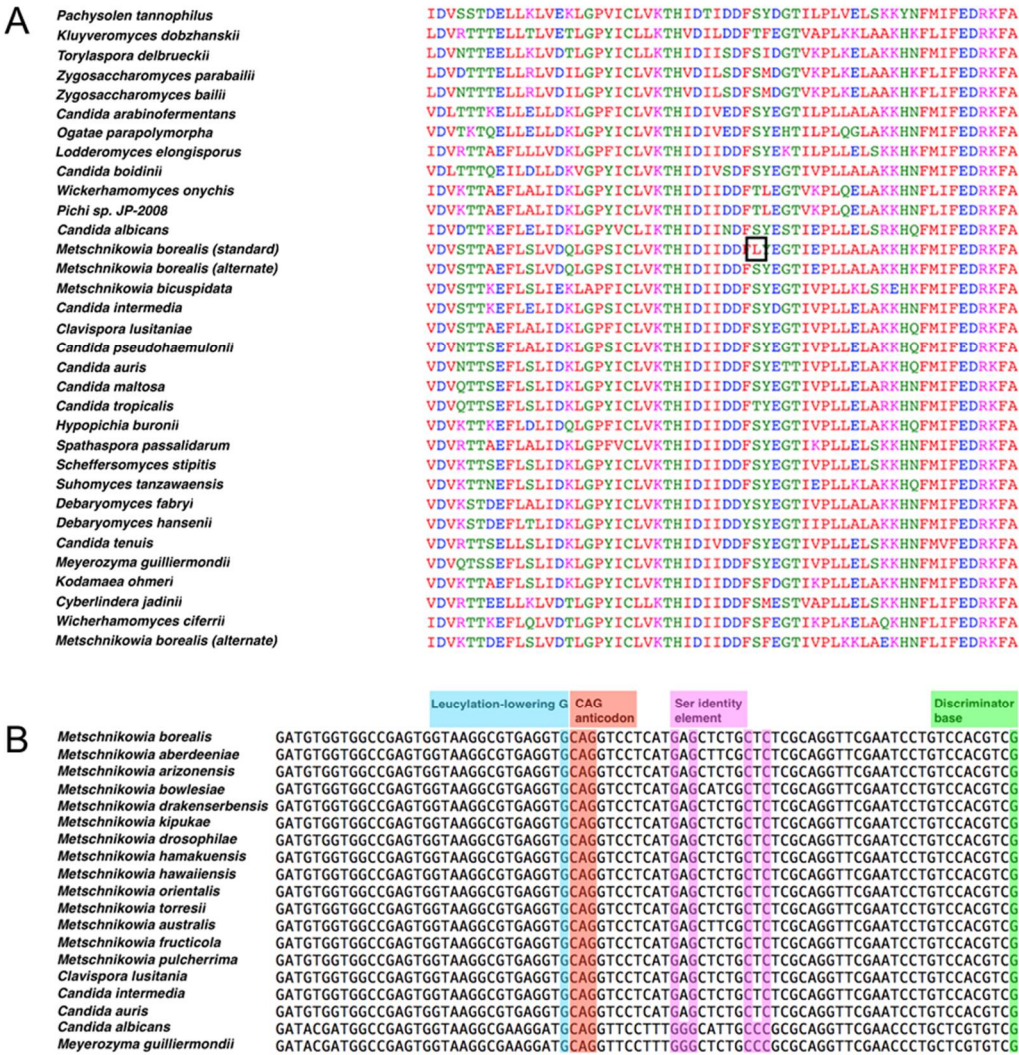


Figure 1. Bioinformatic analyses of the genetic code of *M. borealis*. (A) Sequence alignment of *URA3*. The coding sequence of *URA3* from *M. borealis* was submitted as a query sequence to BLASTx, using the standard code for the predicted protein sequence. The protein sequences of the top 30 hits were aligned in MUSCLE with the *M. borealis* *URA3* gene, translated in both the standard code and yeast alternative code using the ExPASy translate tool. A CUG codon aligned at a highly conserved serine or threonine amino acid position, producing a leucine when *URA3* was translated using the standard code (black square). (B) *tRNA^{CAG}_{Ser}* sequences aligned for many species of *Metschnikowia* and various CUG-Ser yeasts. Sequences were aligned in MEGA6 and annotated as described previously[18].

The relatively high efficiency of electroporation was believed to be attributable to homologous recombination, because each cassette contained by 60 base-pairs of homology to the *ADH1* promoter and terminator. Based on this assumption, we attempted to generate histidine auxotrophs by

transforming *M. borealis* with CaNAT1 flanked by 60 base-pair sequences of the *HIS3* promoter and terminator. After transformation, ten colonies of each making type were spot-diluted and re-isolated, but all grew on media lacking histidine. Subsequent genotyping by PCR identified the presence of both *HIS3* and CaNAT1, but failed to obtain bands across the *HIS3* promoter and terminator junctions with CaNAT1: bands that would be expected in the case of a successful targeted knockout (Supplementary Figure S3).

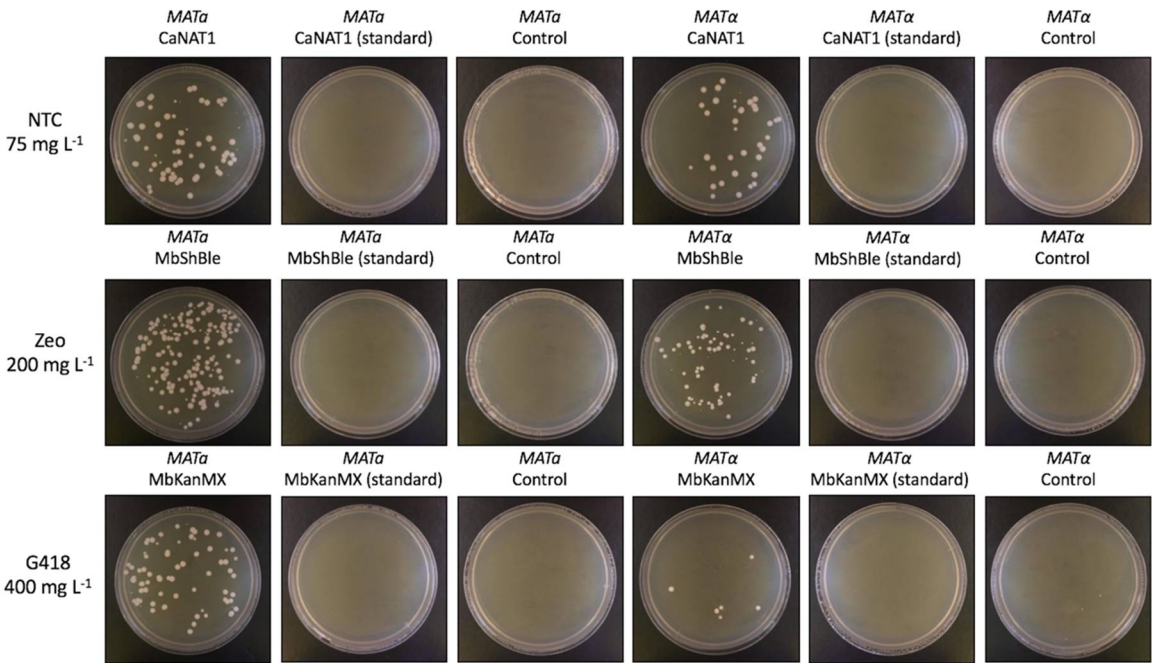


Figure 2. Electroporation of *M. borealis* with CaNAT1, MbKanMX, and MbShBle in standard and yeast alternative nuclear code. Electro-competent cells of *M. borealis* (*MATa* and *MATα*) were transformed with 1 μ g of PCR-amplified CaNAT1, MbShBle, or MbKanMX, in standard or alternative code, flanked by 60 base-pair sequences of the *M. borealis* *ADH1* promoter and terminator, using the electroporation protocol outlined in this study. After recovery, 20 μ L of each transformation was plated on YPAD with the appropriate antibiotic and incubated at 30°C for 2 days. NTC = nourseothricin; Zeo = zeocin; G418 = geneticin.

3.5. Location of Insertion Site

Subsequent screening of more than 30 *MATa* and 30 *MATα* transformants for *HIS3* knockouts failed to identify any successful insertions in the *HIS3* gene. To identify the correct insertion sites, DNA was isolated from 6 *MATα* transformants, digested with *CfoI*, circularized with T4 ligase, and the surrounding gDNA was PCR-amplified with primers binding within the inserted CaNAT1 cassette (oriented to amplify the surrounding sequence) (Supplementary Figure S3). We were successful in obtaining a PCR-fragment from the third *MATα* transformant (clone 3), and the amplified gDNA was sequenced. Analysis with BLASTn indicated that the CaNAT1 cassette was inserted in scaffold bor_s886 of *M. borealis* UWOPS 96-101.1, as annotated in GenBank (PRJNA312754), just upstream of the gene for peroxin-14. PCR was used to amplify this site in wildtype *M. borealis*, as well as all 6 isolated transformants, and the insertion was only observed at this site in clone 3.

3.6. Mating experiments

To confirm that transformed-strains of *M. borealis* *MATa* and *MATα* were capable of successfully mating, MbShBle-transformed *MATa* and CaNAT1-transformed *MATα* were mated to develop ascospores (Figure 3A). Recombinant cells that contain both markers were isolated by serial dilution

and genotyped to confirm that they contain both MbShBle and CaNAT1 (Figure 3B). Mated cells were then plated alongside wildtype *MATa* and *MATα*, MbShBle-transformed *MATa*, and CaNAT1-transformed *MATα* on YPAD, YPAD with zeocin, YPAD with nourseothricin, and YPAD with nourseothricin and zeocin for comparison (Figure 3C).

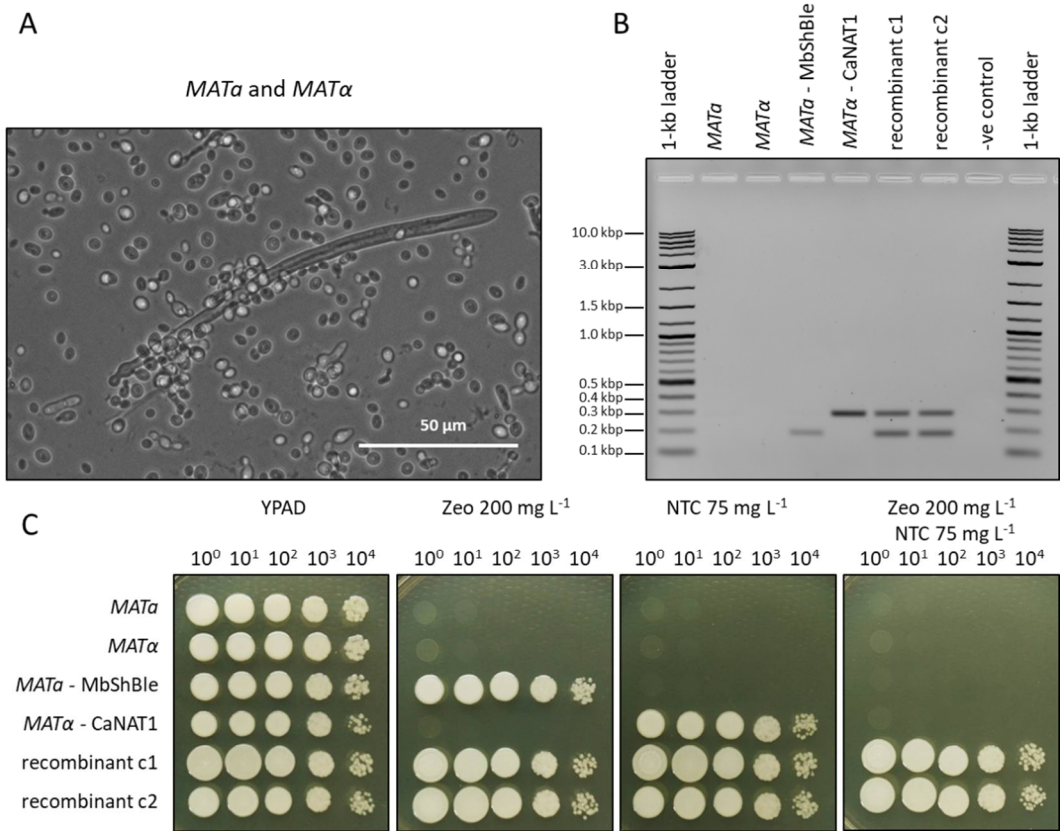


Figure 3. Mating experiments with marker-transformed *M. borealis*. (A) Photograph of a large mating ascospore. MbShBle-transformed *MATa* and CaNAT1-transformed *MATα* were mated for two days on ½-strength glucose broth lacking histidine and uracil (Teknova), collected in YPAD, and photographed under a light microscope. (B) Genotyping results of marker-transformed diploids. Recombinant colonies harbouring both selectable markers were purified by spot dilution on YPAD with 200 mg L⁻¹ zeocin and 75 mg L⁻¹ nourseothricin. Recombinant colonies were genotyped by Quiagen multiplex with primers that amplify the CaNAT1 and MbShBle genes, alongside marker-transformed haploid strains and untransformed controls. MbShBle product size (C) Spot-dilutions of mated-colonies alongside transformed haploid strains and untransformed controls on YPAD, YPAD with 200 mg L⁻¹ zeocin, YPAD with 75 mg L⁻¹ nourseothricin, and YPAD with 200 mg L⁻¹ zeocin and 75 mg L⁻¹ nourseothricin. Plates were incubated at 30°C for 2 days. NTC = nourseothricin; Zeo = zeocin; G418 = geneticin.

3.7. Transformation of other CUG-Ser yeasts

To determine if our electroporation method would yield transformants of other yeasts by random integration, we transformed strains of 19 yeast species with CaNAT1 cassettes (in standard and yeast alternative nuclear code) flanked by the same 60-base-pair minimal *ADH1* promoter and terminator from *M. borealis* (Figure 4; Supplementary Table S3). Transformants were plated on YPAD with 75-200 mg L⁻¹ nourseotricin and incubated at 27-30°C for 24-72 hrs until colonies appeared. All strains yielded colonies for the alternative-coded CaNAT1, but only *S. cerevisiae* and *Candida* aff. *bentonensis* yielded colonies when transformed with standard-coded CaNAT1 (Figure 4; Supplementary Table S3). It's important to note that the alternative-coded CaNAT1 lacks CUG

codons, and is therefore degenerate between the standard code and yeast alternative code; however, the standard-coded CaNAT1 uses CUG codons to translate leucine and is only translatable to CUG-Leu yeasts.

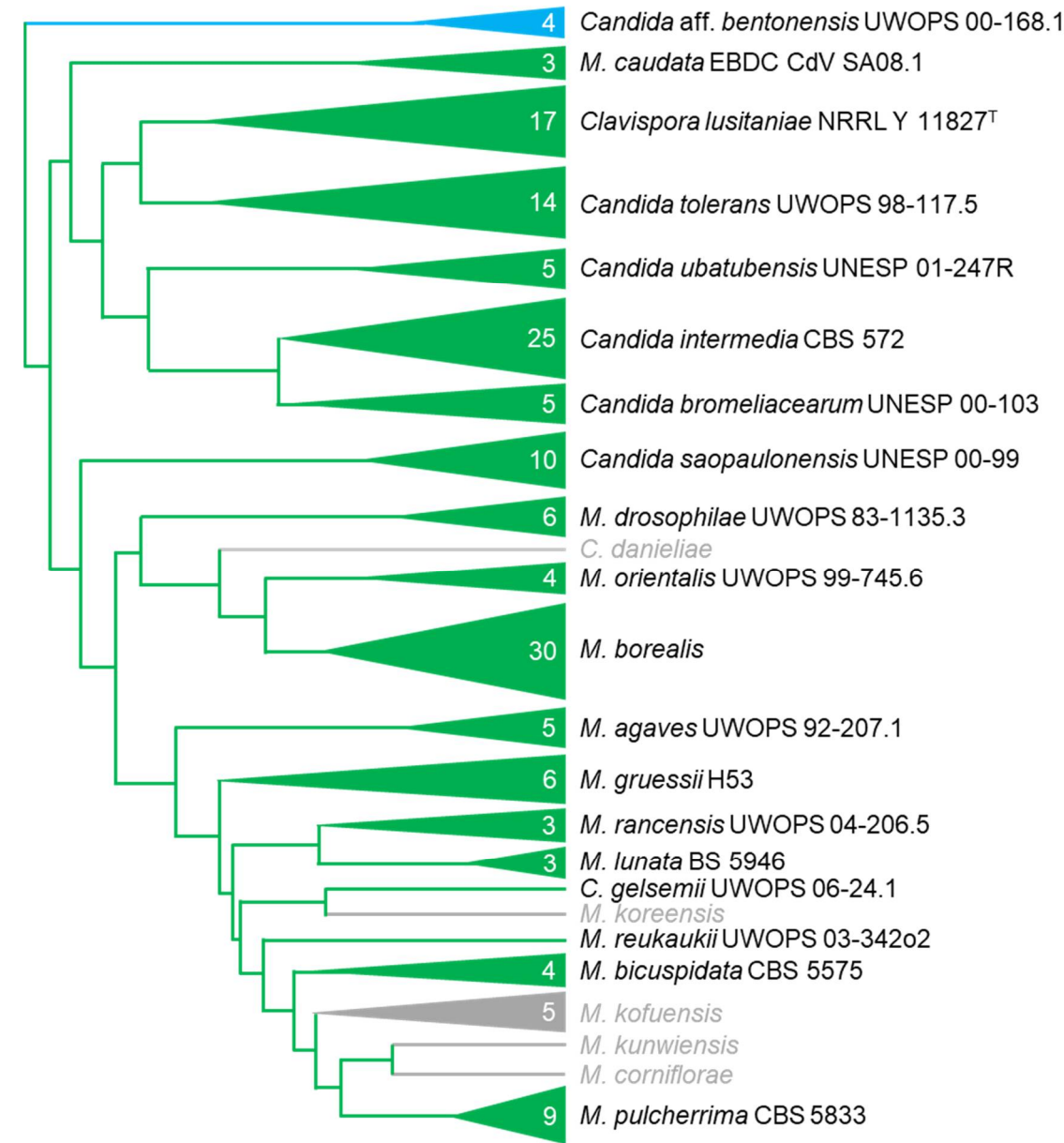


Figure 4. A phylogeny of the Metschnikowiaceae, modified from Lachance (2016), showing the taxonomic range of successful transformations in the family. Representative strains were successfully transformed with only the alternative-coded CaNAT1 construct (green) or with both the alternative-coded and the standard-coded CaNAT1 (blue) construct. Species in grey were not examined.

4. Discussion

Through pioneering experiments, we were successful in determining the necessary conditions to transform *Metschnikowia borealis*, and the method was effective in many other yeast species. We were able to identify the tRNA_{CAG^{Ser}} in *M. borealis*, as well as many other *Metschnikowia* species, and transformation experiments confirmed that many other—if not all—*Metschnikowia* are also CUG-Ser 1 clade yeasts (Figure 2B; Figure 4). It is unclear why *M. pulcherrima* was previously reported to be a CUG-Leu yeast[9].

Once we determined that *M. borealis* utilizes the yeast alternative nuclear code, the species was successfully transformed with selectable markers that contain homologous hooks to the HIS3 promoter and terminator, with the high efficiency. Transformants were first believed to have been obtained through a targeted knockout of *HIS3* by homologous recombination; however, all colonies grew on medium lacking histidine, and genotyping confirmed the presence of both the selectable marker and the wildtype *HIS3* gene. Additionally, we failed to observe any of the expected genotyping bands associated with the knockout junctions. This indicated that the cassettes were not incorporated into the genome by targeted integration, but by another mechanism with high efficiency.

Although we only identified and confirm one insertion site (in clone 3), the HIS3-targeted cassette was found to insert in the site just upstream of the peroxin-14 gene in clone 3, and not in any of the other transformants tested (Supplementary Figure 2C). Therefore, the cassette must have inserted in different locations in the other transformants, indicating that *M. borealis* appears to integrate DNA cassettes randomly into its genome. Even though the species is not efficient at targeted integration, its high efficiency at random integration makes it simple to transform and incorporate exogenous sequences into the genome.

Finally, our transformation method was successful across all CUG-Ser yeasts that were attempted, as well as *S. cerevisiae* and *Candida* aff. *bentonensis*. This confirms the suspicion[5] that *C. bentonensis* is not a member of the Metschnikowiaceae. The results of this study indicate that the species is in fact a CUG-Leu yeast. Fortunately, the CaNAT1, MbShBle, and MbKanMX genes used in our experiments lack the use of a CUG codon and are degenerate to the standard and yeast alternative nuclear code. Therefore, the markers would likely be universal to all CUG-Leu, CUG-Ser, and CUG-Ala yeasts[19].

Through the use of random integration, the methods outlined in this study make it possible to transform a large range of yeasts, and would likely be successful in numerous additional species in a wide array of clades. Random integration can be used to introduce elements to many understudied yeast species to enable their study in more detail. In particular, it is now possible to design codon-optimized fluorescent nuclear-fusion proteins that can be used to elucidate the unusual meiotic process in *Metschnikowiae* that leads to the production of only two ascospores from a diploid zygote.

Supplementary Materials: The following are available online at www.mdpi.com: Figures S1, S2, S3; Tables S1, S2, S3

Author Contributions:

Conceptualization, Z. G., M.P.M.S., P.J., M-A.L., B.J.K; methodology, Z. G., M.P.M.S., C.L., J.A.T., R.S.M., C.L., M.A., D.K.L., P.J., M-A.L., B.J.K; validation, Z. G., M.P.M.S., C.L., J.A.T., R.S.M., C.L., M.A., D.K.L., P.J., M-A.L., B.J.K; formal analysis, Z. G., M.P.M.S., C.L., J.A.T., R.S.M., C.L., M.A., D.K.L., P.J., M-A.L., B.J.K; investigation, Z. G., M.P.M.S., C.L., J.A.T., R.S.M., C.L., M.A., D.K.L., P.J., M-A.L., B.J.K; resources, M-A.L., B.J.K; data curation, Z. G., M.P.M.S., C.L., J.A.T., R.S.M., C.L., M.A., D.K.L., P.J., M-A.L., B.J.K; writing—original draft preparation, Z. G., M.P.M.S., R.S.M., M.A., M-A.L., B.J.K; writing—review and editing, Z. G., M.P.M.S., R.S.M., M.A., M-A.L., B.J.K; visualization, Z. G., M.P.M.S., C.L., J.A.T., R.S.M., C.L., M.A., D.K.L., P.J., M-A.L., B.J.K; supervision, Z. G., P.J., M-A.L., B.J.K; project administration M-A.L., B.J.K; funding acquisition, M-A.L., B.J.K;

Funding:

This research was funded by: Designer Microbes Inc.; NSERC Experience grants to Z.G., C. L., C.L., R.S.M. The B.J.K. lab is also supported by Natural Sciences and Engineering Research Council of Canada (NSERC), grant number: RGPIN-2018-06172.

Conflicts of Interest:

B.J.K. is Chief Executive Officer of Designer Microbes Inc. B.J.K and M-A.L. hold Designer Microbes Inc. stock.

Bibliography

1. Deng, J.; Deng, L.; Su, S.; Zhang, M.; Lin, X.; Wei, L.; Minai, A.A.; Hassett, D.J.; Lu, L.J. Investigating the predictability of essential genes across distantly related organisms using an integrative approach. *Nucleic Acids Res.* **2011**, *39*, 795–807, doi:10.1093/nar/gkq784.
2. Zhang, R.; Ou, H.-Y.; Zhang, C.-T. DEG: a database of essential genes. *Nucleic Acids Res.* **2004**, *32*, D271–D272, doi:10.1093/nar/gkh024.
3. De Backer, M.D.; Maes, D.; Vandoninck, S.; Logghe, M.; Contreras, R.; Luyten, W.H. Transformation of *Candida albicans* by electroporation. *Yeast Chichester Engl.* **1999**, *15*, 1609–1618, doi:10.1002/(SICI)1097-0061(199911)15:15<1609::AID-YEA485>3.0.CO;2-Y.
4. Yildirim, S.; Thompson, M.G.; Jacobs, A.C.; Zurawski, D. V.; Kirkup, B.C. Evaluation of Parameters for High Efficiency Transformation of *Acinetobacter baumannii*. *Sci. Rep.* **2016**, *6*, 22110, doi:10.1038/srep22110.
5. Lachance, M.-A. Metschnikowia: half tetrads, a regicide and the fountain of youth. *Yeast Chichester Engl.* **2016**, doi:10.1002/yea.3208.
6. González-Pombo, P.; Pérez, G.; Carrau, F.; Guisán, J.M.; Batista-Viera, F.; Brena, B.M. One-step purification and characterization of an intracellular beta-glucosidase from *Metschnikowia pulcherrima*. *Biotechnol. Lett.* **2008**, *30*, 1469–1475, doi:10.1007/s10529-008-9708-3.
7. Macwilliam, I.C. A survey of the antibiotic powers of yeasts. *J. Gen. Microbiol.* **1959**, *21*, 410–414, doi:10.1099/00221287-21-2-410.
8. Marinoni, G.; Piškur, J.; research, M.L.-F. yeast; 2003, undefined Ascospores of large-spored *Metschnikowia* species are genuine meiotic products of these yeasts. *academic.oup.com*.
9. Nigro, F.; Sialer, M.M.F.; Gallitelli, D. TRANSFORMATION OF METSCHNIKOWIA PULCHERRIMA 320, BIOCONTROL AGENT OF STORAGE ROT, WITH THE GREEN FLUORESCENT PROTEIN GENE. *J. Plant Pathol.* **1999**, *81*, 205–208.
10. Lachance, M.-A.; Hurtado, E.; Hsiang, T. A stable phylogeny of the large-spored *Metschnikowia* clade. *Yeast Chichester Engl.* **2016**, *33*, 261–275, doi:10.1002/yea.3163.
11. Kawai, S.; Hashimoto, W.; Murata, K. Transformation of *Saccharomyces cerevisiae* and other fungi. *Bioeng. Bugs* **2010**, *1*, 395–403, doi:10.4161/bbug.1.6.13257.
12. Lachance, M.-A.; Rosa, C.A.; Starmer, W.T.; Schlag-Edler, B.; Baker, J.S.F.; Bowles, J.M. *Metschnikowia continentalis* var. *borealis*, *Metschnikowia continentalis* var. *continentalis*, and *Metschnikowia hibisci*, new heterothallic haploid yeasts from ephemeral flowers and associated insects. *Can. J. Microbiol.* **1998**, *44*, 279–288, doi:10.1139/w97-148.
13. Puigbò, P.; Guzmán, E.; Romeu, A.; Garcia-Vallvé, S. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.* **2007**, *35*, W126–131, doi:10.1093/nar/gkm219.
14. Gietz, R.D. Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol. Biol. Clifton NJ* **2014**, *1205*, 1–12, doi:10.1007/978-1-4939-1363-3_1.
15. Froyd, C.A.; Kapoor, S.; Dietrich, F.; Rusche, L.N. The deacetylase Sir2 from the yeast *Clavispora lusitaniae* lacks the evolutionarily conserved capacity to generate subtelomeric heterochromatin. *PLoS Genet.* **2013**, *9*, e1003935, doi:10.1371/journal.pgen.1003935.
16. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; Thierer, T.; Ashton, B.; Meintjes, P.; Drummond, A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinforma. Oxf. Engl.* **2012**, *28*, 1647–1649, doi:10.1093/bioinformatics/bts199.

17. Mühlhausen, S.; Kollmar, M. Molecular Phylogeny of Sequenced Saccharomycetes Reveals Polyphyly of the Alternative Yeast Codon Usage. *Genome Biol. Evol.* **2014**, *6*, 3222–3237, doi:10.1093/gbe/evu152.

18. Riley, R.; Haridas, S.; Wolfe, K.H.; Lopes, M.R.; Hittinger, C.T.; Göker, M.; Salamov, A.A.; Wisecaver, J.H.; Long, T.M.; Calvey, C.H.; Aerts, A.L.; Barry, K.W.; Choi, C.; Clum, A.; Coughlan, A.Y.; Deshpande, S.; Douglass, A.P.; Hanson, S.J.; Klenk, H.-P.; LaButti, K.M.; Lapidus, A.; Lindquist, E.A.; Lipzen, A.M.; Meier-Kolthoff, J.P.; Ohm, R.A.; Otilar, R.P.; Pangilinan, J.L.; Peng, Y.; Rokas, A.; Rosa, C.A.; Scheuner, C.; Sibirny, A.A.; Slot, J.C.; Stielow, J.B.; Sun, H.; Kurtzman, C.P.; Blackwell, M.; Grigoriev, I.V.; Jeffries, T.W. Comparative genomics of biotechnologically important yeasts. *Proc. Natl. Acad. Sci.* **2016**, *113*, 9882–9887, doi:10.1073/pnas.1603941113.

19. Krassowski, T.; Coughlan, A.Y.; Shen, X.-X.; Zhou, X.; Kominek, J.; Opulente, D.A.; Riley, R.; Grigoriev, I. V.; Maheshwari, N.; Shields, D.C.; Kurtzman, C.P.; Hittinger, C.T.; Rokas, A.; Wolfe, K.H. Evolutionary instability of CUG-Leu in the genetic code of budding yeasts. *Nat. Commun.* **2018**, *9*, 1887, doi:10.1038/s41467-018-04374-7.

20. Sugita, T.; Nakase, T. Non-universal usage of the leucine CUG codon and the molecular phylogeny of the genus *Candida*. *Syst. Appl. Microbiol.* **1999**, *22*, 79–86, doi:10.1016/S0723-2020(99)80030-7.

21. Lowe, T.M.; Chan, P.P. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* **2016**, *44*, W54–57, doi:10.1093/nar/gkw413.

22. Shen, J.; Guo, W.; Köhler, J.R. CaNAT1, a Heterologous Dominant Selectable Marker for Transformation of *Candida albicans* and Other Pathogenic *Candida* Species. *Infect. Immun.* **2005**, *73*, 1239–1242, doi:10.1128/IAI.73.2.1239-1242.2005.