

Genomics and Transcriptomics Analyses of the Oil-Accumulating Basidiomycete Yeast *Trichosporon oleaginosus*: Insights into Substrate Utilization and Alternative Evolutionary Trajectories of Fungal Mating Systems

Robert Kourist,^a Felix Bracharz,^b Jan Lorenzen,^b Octavia N. Kracht,^a Mansi Chovatia,^c Chris Daum,^c Shweta Deshpande,^c Anna Lipzen,^c Matt Nolan,^c Robin A. Ohm,^{c,d} Igor V. Grigoriev,^c Sheng Sun,^e Joseph Heitman,^e Thomas Brück,^b Minou Nowrousian^f

Junior Research Group for Microbial Biotechnology, Ruhr-Universität Bochum, Bochum, Germany^a; Fachgebiet Industrielle Biokatalyse, Technische Universität München, Garching, Germany^b; U.S. Department of Energy Joint Genome Institute, Walnut Creek, California, USA^c; Department of Microbiology, Utrecht University, Utrecht, The Netherlands^d; Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina, USA^e; Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum, Bochum, Germany^f

ABSTRACT Microbial fermentation of agro-industrial waste holds great potential for reducing the environmental impact associated with the production of lipids for industrial purposes from plant biomass. However, the chemical complexity of many residues currently prevents efficient conversion into lipids, creating a high demand for strains with the ability to utilize all energy-rich components of agricultural residues. Here, we present results of genome and transcriptome analyses of *Trichosporon oleaginosus*. This oil-accumulating yeast is able to grow on a wide variety of substrates, including pentoses and *N*-acetylglucosamine, making it an interesting candidate for biotechnological applications. Transcriptomics shows specific changes in gene expression patterns under lipid-accumulating conditions. Furthermore, gene content and expression analyses indicate that *T. oleaginosus* is well-adapted for the utilization of chitin-rich biomass. We also focused on the *T. oleaginosus* mating type, because this species is a member of the *Tremellomycetes*, a group that has been intensively analyzed as a model for the evolution of sexual development, the best-studied member being *Cryptococcus neoformans*. The structure of the *T. oleaginosus* mating-type regions differs significantly from that of other *Tremellomycetes* and reveals a new evolutionary trajectory paradigm. Comparative analysis shows that recruitment of developmental genes to the ancestral tetrapolar mating-type loci occurred independently in the *Trichosporon* and *Cryptococcus* lineages, supporting the hypothesis of a trend toward larger mating-type regions in fungi.

IMPORTANCE Finite fossil fuel resources pose sustainability challenges to society and industry. Microbial oils are a sustainable feedstock for biofuel and chemical production that does not compete with food production. We describe genome and transcriptome analyses of the oleaginous yeast *Trichosporon oleaginosus*, which can accumulate up to 70% of its dry weight as lipids. In contrast to conventional yeasts, this organism not only shows an absence of diauxic effect while fermenting hexoses and pentoses but also effectively utilizes xylose and *N*-acetylglucosamine, which are building blocks of lignocellulose and chitin, respectively. Transcriptome analysis revealed metabolic networks that govern conversion of xylose or *N*-acetylglucosamine as well as lipid accumulation. These data form the basis for a targeted strain optimization strategy. Furthermore, analysis of the mating type of *T. oleaginosus* supports the hypothesis of a trend toward larger mating-type regions in fungi, similar to the evolution of sex chromosomes in animals and plants.

Received 31 May 2015 Accepted 15 June 2015 Published 21 July 2015

Citation Kourist R, Bracharz F, Lorenzen J, Kracht ON, Chovatia M, Daum C, Deshpande S, Lipzen A, Nolan M, Ohm RA, Grigoriev IV, Sun S, Heitman J, Brück T, Nowrousian M. 2015. Genomics and transcriptomics analyses of the oil-accumulating basidiomycete yeast *Trichosporon oleaginosus*: insights into substrate utilization and alternative evolutionary trajectories of fungal mating systems. *mBio* 6(4):e00918-15. doi:10.1128/mBio.00918-15.

Editor B. Gillian Turgeon, Cornell University

Copyright © 2015 Kourist et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to Robert Kourist, robert.kourist@rub.de, Thomas Brück, brueck@tum.de, or Minou Nowrousian, minou.nowrousian@rub.de.

This article is a direct contribution from a Fellow of the American Academy of Microbiology.

The foreseeable end of fossil resources, associated with significant price fluctuations, drives the development of sustainable biomass-based processes in the chemical industry sector. Plant-derived oils are already essential feedstocks in the production of bio-based specialty and commodity chemical building blocks (1, 2). However, their material utilization competes with food pro-

duction and has an increasing impact on land use changes (3, 4). An emerging alternative is the fermentative conversion of agro-industrial residues for production of microbe-derived lipids.

Oleaginous yeasts are promising microbial factories for sustainable lipid production, as they can accumulate between 40 and 70% of their biomass as intracellular storage triglycerides when

primary nutrients are limited (5). In particular, these organisms can metabolize pentose and hexose sugars from complex biomass hydrolysates and switch on lipogenesis when metabolic stresses, such as nitrogen and phosphorus deprivation, are applied. The metabolic adaptability of oleaginous yeasts to various carbon sources is of interest for second-generation bioprocess engineering approaches (2). While the exact fatty acid profile of accumulated triglycerides depends on cultivation conditions, the principal chemical composition resembles those of plant-derived lipids, such as palm oil. Interestingly, oleaginous yeasts belong to diverse taxonomic groups, indicating that the metabolic capacity for lipid accumulation has evolved independently in basidiomycete (6) and ascomycete fungi (7–9).

At present, nitrogen limitation has been identified as the strongest inducer of lipogenesis in oleaginous organisms. Biochemical and metabolic engineering studies have identified malic enzyme (ME) (10–12), diacylglycerol acyltransferase 1 (DGA1) (13), and acetyl coenzyme A (CoA) carboxylase (13) as metabolic regulators of lipogenesis under N stress conditions for selected fungi and microalgae. However, recent systems biology studies in different oleaginous organisms could not identify a common metabolic mechanism for lipid accumulation under N-limiting conditions (7, 14). Furthermore, metabolic mechanisms governing lipogenesis under P-limiting conditions have not yet been explored. Hence, the overall metabolic networks and regulatory mechanisms that drive lipid accumulation in different oleaginous organisms remain elusive (2).

Here, we present genome and transcriptome analysis results for the oil-accumulating yeast *Trichosporon oleaginosus* strain IBC0246. This *Trichosporon* strain is distinct from conventional yeasts, as it can simultaneously metabolize both pentose and hexose sugars without any catabolic preference. Its capacity to metabolize *N*-acetylglucosamine (NAG) and the ability to accumulate 70% of its dry weight as lipid droplets make it an interesting candidate for biotechnological applications. Sequence identity of the internal transcribed spacer (ITS) sequences shared with two previously characterized *T. oleaginosus* strains (15) identified strain IBC0246 as *Trichosporon oleaginosus*. *Trichosporon* species are basidiomycetes that belong to the order *Tremellales*. While nonpathogenic *Trichosporon* species have been isolated from environmental soil and milk whey samples, some isolates have recently come to attention as pathogens of immunocompromised hosts. *Trichosporon* species mostly grow as yeasts, but many are able to also form pseudohyphae or true hyphae and arthroconidia (16, 17). A sexual cycle and ploidy level have not been described for *Trichosporon* species.

Strain IBC0246 is able to utilize a variety of carbon sources without showing a diauxic lag in growth (unpublished data), making it an interesting candidate for biotechnological applications. Interestingly, it can efficiently utilize pentoses such as xylose (XYL) and NAG, which are major monomeric sugar building blocks of lignocellulosic and chitin-rich biomass residues, respectively (18). Because XYL and NAG utilization is challenging for conventional industrial yeasts (19), *Trichosporon oleaginosus* is poised to be developed as a next-generation microbial production platform for the *de novo* biosynthesis of sustainable oleochemicals. In this study, we therefore focused on lipid accumulation and the possible utilization of the two biomass-derived sugars XYL and NAG. We show that in response to nitrogen starvation this yeast is capable of utilizing xylose as a carbon source for lipid

accumulation, and we highlight several pathways of oleagenicity. Transcriptome analysis of a sample using NAG as the sole carbon source showed that NAG is efficiently channeled toward the primary metabolism.

Furthermore, we analyzed the mating-type structure of *T. oleaginosus*, because this species is a member of the *Tremellomycetes*, which in recent years have become model organisms for the study of the evolution of sexual development. A well-studied member of this group is the pathogenic yeast *Cryptococcus neoformans*. The structure of the *T. oleaginosus* mating-type regions differs significantly from that of the *Cryptococcus* species complex. Comparative analysis suggests that it has undergone independent evolution shaped by a similar trend toward larger mating-type regions, as found in other *Tremellomycetes* (20), but in which orthologous genes functioning in sexual development (*MYO2*, *STE11*, *IKS1*, *STE20*) have been differentially incorporated into either one or the other of the ancestral tetrapolar *MAT* loci in the two lineages. This reveals a heretofore-unknown level of plasticity in expansion of sex-determining regions in fungi, with implications for sex chromosome evolution in animals and plants.

RESULTS AND DISCUSSION

Genome sequencing of *Trichosporon oleaginosus* strain IBC0246 reveals great evolutionary flexibility in mating-type configurations in *Tremellomycetes*. The genome of *T. oleaginosus* strain IBC0246 was sequenced as part of the 1000 Fungal Genomes project (<http://1000.fungalgenomes.org>) (21) and assembled into 180 scaffolds (Table 1). With an assembly size of 19.8 Mb, it is in a size range similar to the genomes of the pathogenic *Tremellomycetes* *Cryptococcus neoformans* and *Cryptococcus gattii* (18 to 19 Mb) and the opportunistic pathogen *Trichosporon asahii* (24 to 25 Mb) (22–26). Evidence-based annotation using transcriptome sequencing (RNA-seq) data (see below) identified about 8,300 protein-coding genes (see Table S1 in the supplemental material), which is similar to *T. asahii* (8,300 to 8,500 protein-coding genes) and somewhat more than the 6,500 to 6,900 predicted genes in *Cryptococcus* species. Seventy-three percent of all predicted proteins showed similarities to proteins from the Swissprot database, and 83% had a KEGG annotation (Table 1). The genome sequence and all annotations are available via the MycoCosm portal (<http://genome.jgi.doe.gov/Triol1>) (27).

Repeat analyses based on similarity to known repeat classes as well as *de novo* repeat finding showed that in *T. oleaginosus* repeats longer than 200 bp constitute only 0.3% of the genome. The total repeat content of the genome, consisting mostly of small repeats, is 2.85%, which is in a range similar to that of *C. neoformans* (5% repeats) (23).

To analyze synteny between *T. oleaginosus* and other *Tremellomycetes*, we used the PROMer algorithm from the MUMmer package (28); however, even a comparison with the closest sequenced relative, *T. asahii*, yielded only a few aligned regions (data not shown), probably due to large evolutionary distances even within the genus *Trichosporon*. Therefore, an orthology-based approach was chosen, which entailed identification of putative orthologous proteins between *T. oleaginosus* and other species by reciprocal BLAST analyses and plotting the genomic positions of orthologous genes. Between 4,587 and 5,048 putative orthologs were identified in the comparisons of *T. oleaginosus* with *T. asahii*, *Tremella mesenterica*, and the two *Cryptococcus* species, whereas the comparison of *C. neoformans* and *C. gattii* that was included as a

TABLE 1 Main features of the *T. oleaginosus* genome^a

Assembly statistic	Value
Assembly length (Mbp)	19.8
Contig length total (Mbp)	19.8
No. of contigs	196
Contig N50 (bp)	34
Contig L50 (bp)	190,830
No. of scaffolds	180
Scaffold N50 (bp)	29
Scaffold L50 (bp)	216,041
No. of scaffold gaps	16
Length of scaffold gaps (bp)	4,451
% of scaffolds in gaps	0.02
No. of repeat-covered regions	4,916
Length of repeat-covered regions (bp)	564,416
% of assembly covered by repeats	2.85
GC content (%)	60.75
Gene statistics	
No. of genes	8,322
Protein length (median no. of amino acids)	367
Exon length (median bp)	238
Gene length (median bp)	1,500
Transcript length (median bp)	1,309
Intron length (median bp)	39
No. of genes with intron	7,010
% of genes with intron	84.23
No. of introns/gapped gene (median)	3
Functional annotations	
No. (%) of genes with KEGG annotation	6,869 (82.54)
No. (%) of genes with KOG annotation	5,937 (71.34)
No. (%) of genes with Swissprot hit	6,088 (73.16)
No. (%) of genes with Pfam domain	4,833 (58.07)
No. of genes with transmembrane domain	1,519 (18.25)
No. of genes with signal peptide	1,873 (22.51)
No. of unique Pfam domains	2,248
Annotation completeness (CEGMA) (%)	99.13

^a Assembly statistics are based on 1 N to denote a gap.

control yielded 5,604 orthologs. Plotting of genomic positions of orthologous pairs showed a high degree of synteny between the two *Cryptococcus* species, as expected, but no large-scale synteny for any of the *T. oleaginosus* comparisons (Fig. 1). This also included the comparison to the sister species *T. asahii*, confirming the preliminary PROmer-based results and indicating a larger evolutionary distance between the two *Trichosporon* species than the distance between *C. neoformans* and *C. gattii*.

While our analysis did not show any macrosynteny between *T. oleaginosus* and other sequenced genomes, an analysis of syntenic gene pairs and triplets showed considerable microsynteny (Fig. 1). Nearly 1,100 syntenic gene pairs and 500 syntenic gene triplets were identified between *T. oleaginosus* and *T. asahii*, and about 700 and 400 syntenic pairs and triplets, respectively, were identified in comparisons of *T. oleaginosus* with *T. mesenterica* or the *Cryptococcus* species. However, these results are still much lower than those from a comparison of *C. neoformans* and *C. gattii*, again indicating a high degree of chromosomal rearrangements in *T. oleaginosus* and thus a large evolutionary distance compared to other sequenced genomes.

One major feature of interest in *Tremellomycete* genomes is the organization and evolution of mating-type regions (20). Sexual development in basidiomycetes is governed by mating-type genes encoding homeodomain transcription factors, pheromones, and pheromone receptors. These mating-type genes can be organized

as one mating-type locus containing all genes or as two mating-type loci, one of which contains the transcription factor genes (HD locus) and the other that contains the pheromone and pheromone receptor genes (P/R locus). In the majority of basidiomycetes, the HD locus comprises at least two genes coding for homeodomain proteins of class 1 (HD1) and class 2 (HD2) (29, 30). This is the case in the *Tremellomycetes* *Cryptococcus amyloletus*, *Kwoniella heveanensis*, *Kwoniella mangrovensis*, and *Tremella mesenterica*, which harbor the basal arrangement of two separate mating-type loci, with two HD genes of classes HD1 and HD2 within the transcription factor-containing locus (31–33). An exception to this rule is found in *C. neoformans*, where the single mating-type locus contains only one or the other HD gene (34, 35). This HD transcription factor gene is clustered with pheromone precursor and receptor genes and additional genes involved in mating within a genomic region of >100 kb. With the exception of a gene conversion hot spot (36), this large mating-type region of *C. neoformans* displays restricted meiotic recombination and might represent an evolutionary trend from small mating regions to mating/sex chromosomes that is also seen in other groups, e.g., animals, plants, and algae (20, 34, 37, 38).

Trichosporon is a sister lineage to the *Tremella/Kwoniella/Cryptococcus* lineage within the *Tremellomycetes* (Fig. 1). To learn more about the mating system of *T. oleaginosus*, we searched for homologs to proteins encoded by the well-studied mating-type locus of *C. neoformans*. In this fungus, strains harboring the HD1 gene *SXII* at the mating-type locus are designated *MAT α* , whereas strains harboring the HD2 gene *SXII* are *MAT α* (35). While for almost all *MAT*-encoded proteins, including *Sxi1*, putative homologs could be found in *T. oleaginosus*, no homologs were found for the HD2 transcription factor *Sxi2* (see Table S2 and Fig. S1 and S2 in the supplemental material). Searches with HD proteins from other *Tremellomycetes* in the predicted *T. oleaginosus* proteins as well as tBLASTn searches in the genome assembly did not yield any putative *Sxi2* homologs either. This was further confirmed by searches in the genomes of the *T. asahii* type strain (26) (see Table S2) and a *T. asahii* environmental strain (25) (data not shown), indicating that the currently available *Trichosporon* genomes do not encode an HD2 transcription factor. Homologs for the pheromone genes were also not found; however, this is not unusual, as pheromones are small peptides that harbor only short conserved motifs. Searches specifically for conserved pheromone motifs within the genome assemblies of *T. oleaginosus* and *T. asahii* as well as in the unmapped *T. oleaginosus* Illumina reads yielded several putative pheromone gene candidates (see Fig. S3 in the supplemental material). Whether these candidates actually encode pheromone precursors remains to be elucidated.

The *T. oleaginosus* homologs encoding putative mating-associated proteins are located on several different scaffolds, making a mating chromosome-like arrangement as found in *C. neoformans* unlikely (see Table S2 in the supplemental material). Importantly, the HD transcription factor homolog *SXI1* and the pheromone receptor gene *STE3* are found on scaffolds 70 and 68, respectively, suggesting two independent mating-type loci harboring transcription factor and pheromone receptor genes, respectively (Fig. 2). Thus, with respect to the genomic organization of the mating-type genes at two different loci, *T. oleaginosus* is similar to other *Tremellomycetes*, with the exception of the pathogenic cryptococci, whereas the presence of only one HD transcription factor gene in *T. oleaginosus* resembles the derived organiza-

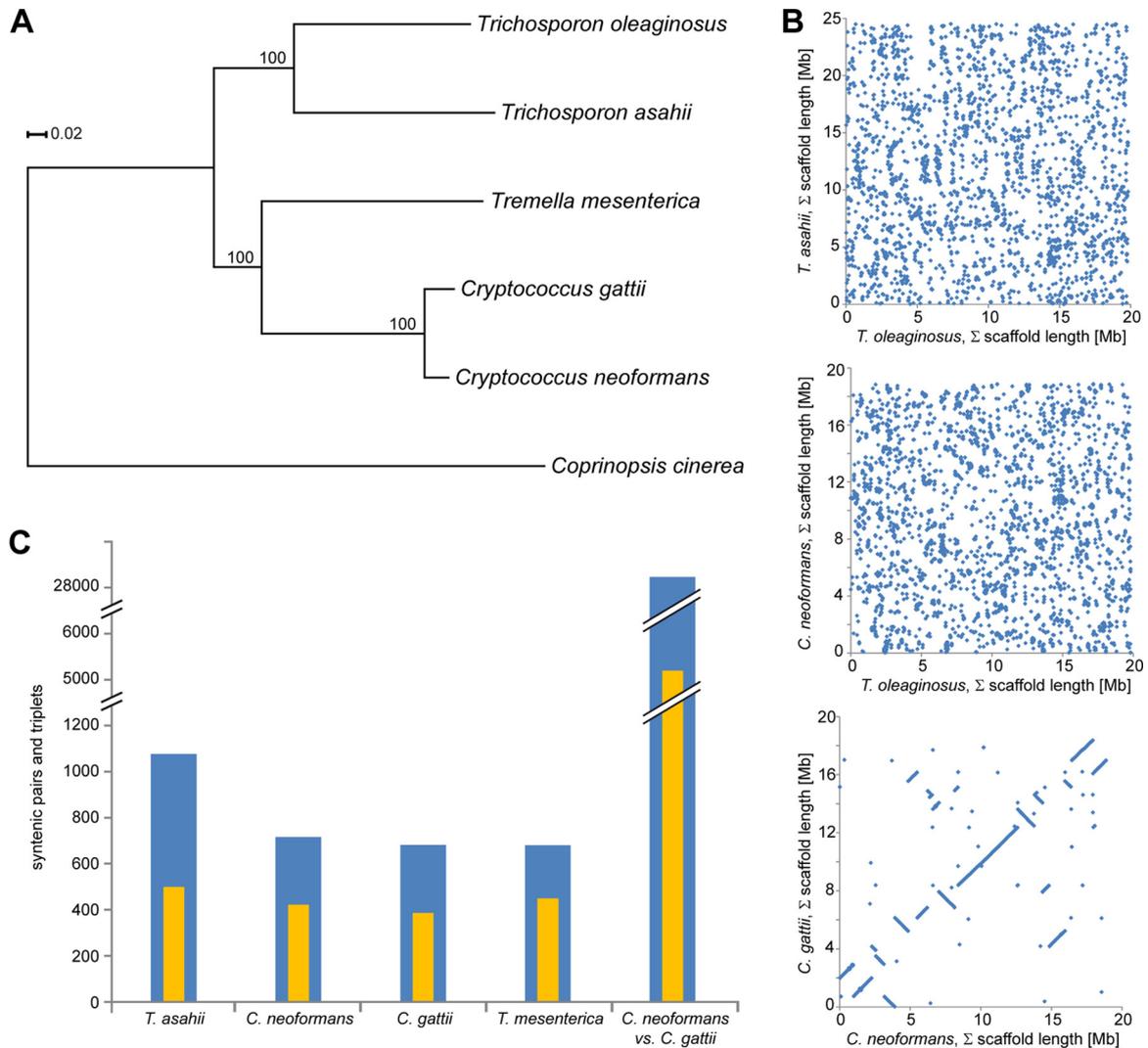


FIG 1 Phylogenetic and synteny analyses. (A) Phylogenetic tree of species in the class *Tremellomycetes* based on analysis of 200 conserved single-copy genes. Bootstrap values (percentages) are indicated at the corresponding branches. *Coprinopsis cinerea* was used as an outgroup. The scale bar indicates the number of amino acid substitutions per site. (B) Genome-wide synteny between *T. oleaginosus* and other *Tremellomycetes*. Positions of orthologous genes were determined along the concatenated scaffolds from each species and are visualized as dot plots for the pairs *T. oleaginosus*/*T. asahii* and *T. oleaginosus*/*C. neoformans*. The analysis of *C. neoformans* versus *C. gattii* was included for comparison. Scaffolds were not specifically ordered for this analysis; therefore, there is a seemingly somewhat-random organization for the *C. neoformans*/*C. gattii* plot. (C) Syntenic gene pairs and triplets in *T. oleaginosus* compared to four other *Tremellomycetes*. Numbers of pairs and triplets of orthologous genes within 20 kb (pairs) and 40 kb (triplets) are shown as blue and orange bars, respectively. On the right, the same comparison for the two *Cryptococcus* species, *C. neoformans* and *C. gattii*, is shown. Note that the y axis is interrupted in two places for better visualization. BioProject numbers and references for genomes used for comparison are as follows: *T. asahii* PRJNA164647 (26), *C. neoformans* PRJNA411 (24), *C. gattii* PRJNA62089 (22), *T. mesenterica* PRJNA225529 (75).

tion of *C. neoformans*/*C. gattii*. However, as the last common ancestor of the *Tremellomycetes* most likely contained the “standard” two-HD gene arrangement (20, 31) and the genus *Trichosporon* branches at the base of the *Tremellomycetes* lineage (Fig. 1), the loss of one HD gene might represent a case of parallel evolution in *Trichosporon* and *C. neoformans*/*C. gattii*.

In *C. neoformans*, *C. gattii*, *C. amyloleptus*, *K. heveanensis*, *K. mangrovensis*, and *T. mesenterica*, several genes besides HD transcription factors, pheromones, and receptors were recruited into the mating-type loci, representing a trend toward larger mating regions culminating in the fusion of both mating-type loci into one large locus in *C. neoformans*/*C. gattii* (20, 31–33). Homologs of several of these genes are also found associated with the HD

transcription factor or *STE3*, respectively, in *T. oleaginosus* (Fig. 2); however, the distribution of these genes between the two loci is distinct from that of the other *Tremellomycetes*. In *T. mesenterica* and *K. heveanensis*, the developmental genes *STE11*, *STE12*, and *STE20*, as well as several other genes (*MYO2*, *IKS1*), are linked to the P/R locus, whereas the essential gene *RPL22* is linked to the HD locus (33). In *T. oleaginosus*, a similar linkage of *RPL22* with the HD locus can be found, but *STE11*, *STE20*, *MYO2*, and *IKS1* are also linked to the HD locus instead of the P/R locus (Fig. 2). *STE12* of *T. oleaginosus* is located on a scaffold different from both the HD and P/R loci, but in the genome assembly of *T. asahii* (26) *STE12* is present on the same scaffold as the putative pheromone receptor gene *STE3* (Fig. 2). A relatively high degree of

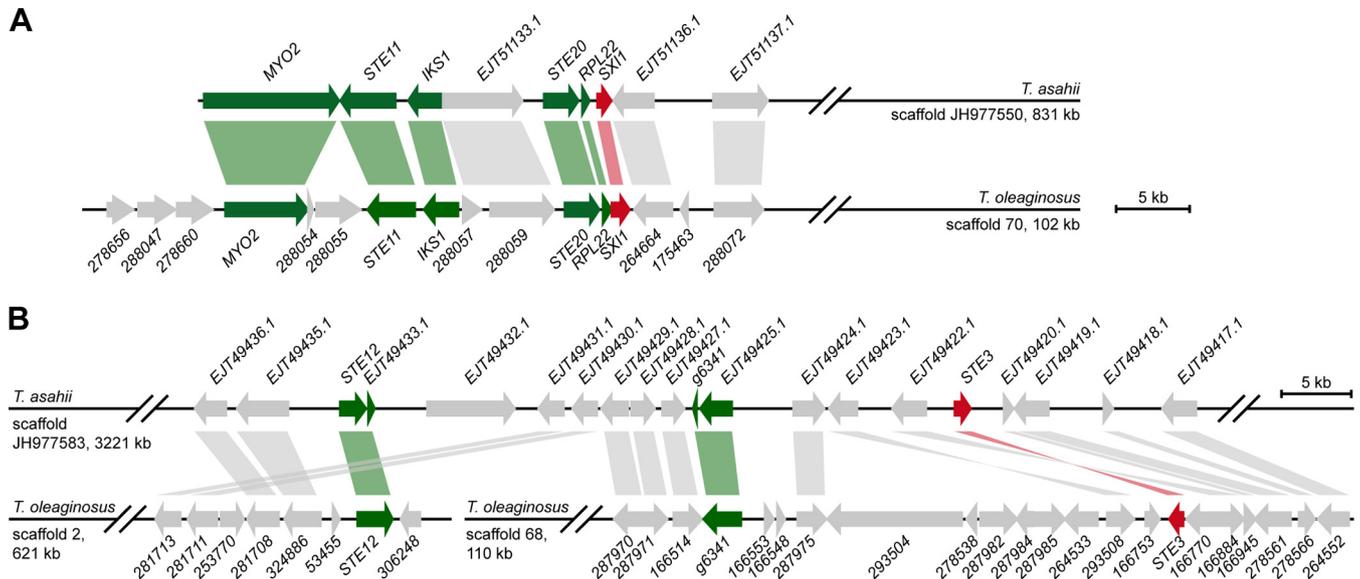


FIG 2 *T. oleaginosus* scaffolds 2, 68, and 70 contain homologs to genes from the *C. neoformans* mating-type locus. Characteristic mating-type genes are shown in red, and other genes that are part of the mating-type locus or flank the mating-type locus (gene *g6341*) of *C. neoformans* are shown in green. Genes unrelated to the mating-type locus are shown in gray. Genes are indicated only in those parts of the scaffolds that contained putative mating-type genes (the left part of scaffold 70 and the right parts of scaffolds 2 and 68). For comparison, homologous regions from the *T. asahii* type strain (26) are shown. In this species, the corresponding regions are present on two scaffolds (left part of scaffold JH977550 and middle part of scaffold 977583). Genes associated with *SXI1* (HD locus) in panel A, genes associated with *STE3* (P/R locus) in panel B.

synteny of the genomic region(s) containing *STE3* and *STE12* in *T. asahii* and *T. oleaginosus* and the finding that the *STE3* and *STE12* regions are located at the ends of their respective scaffolds in *T. oleaginosus*, suggest that these scaffolds might be linked in *T. oleaginosus*, too (Fig. 2B). Furthermore, the linkage of *STE11*, *STE20*, *MYO2*, and *IKS1* to the HD locus was also found in *T. asahii*, indicating that this genomic arrangement is not an assembly artifact in *T. oleaginosus* (Fig. 2A).

Based on the *Tremellomycetes* phylogeny, the last common ancestor of *Trichosporon* and the *Cryptococcus/Kwoniella/Tremella* group is inferred to have been tetrapolar (20) (Fig. 1). Therefore, one hypothesis to explain the different distributions of mating-associated genes around *SXI1* and *STE3* would be that the recruiting of developmental genes like *STE11*, *STE12*, and *STE20* into the mating-type regions occurred independently in *Trichosporon* and the sister groups, leading to different sets of genes becoming linked to the HD locus and the P/R locus, respectively (Fig. 3). Genes without a developmental function might have ended up linked to the mating-type loci as a consequence of recombination events involving larger genomic regions. Another possible hypothesis would be that the last common ancestor already had the *Cryptococcus*-like single mating region and that this region was broken up several times independently in the descendants. However, this hypothesis is more complex and less parsimonious than the first, because it would require (independent, but similar) genomic rearrangements in essentially every lineage except *Cryptococcus*. The most parsimonious assumption of an independent recruitment event for one set of developmental genes into the HD locus or P/R locus would be consistent with an evolutionary trend toward larger mating-type regions (20). For a detailed discussion about possible causes/consequences, see Text S1 in the supplemental material.

The presence of mating-type genes, and thus the possibility of

a sexual cycle, is not only of interest for the analysis of the evolution of sex but also for potential development of *T. oleaginosus* for biotechnological applications, because genetically tractable organisms allow for easier development of molecular tools. In addition, detailed knowledge about cellular adaptations to growth conditions relevant for biotechnological applications also facilitates strain development, and therefore we analyzed several *T. oleaginosus* transcriptomes, as described in the next sections.

Transcriptome analysis in the presence of different carbon sources. Transcriptome analyses were performed under different cultivation conditions (see Text S1 in the supplemental material). Cultivation in complete medium (YPD) was compared with growth on the two alternative carbon sources, xylose (XYL) and NAG. To study the oil accumulation under relevant conditions for biomass utilization, we used xylose as carbon source under conditions of nitrogen limitation (NLM). While nitrogen limitation efficiently induces lipid synthesis, utilization of N-rich biomass, such as that derived from chitin, poses a challenge with this strategy (39). Therefore, we also investigated the possibility of inducing lipid synthesis with phosphate limitation (PLM).

Between 26 million and 100 million RNA-seq reads were obtained for each independent biological replicate (see Text S1 in the supplemental material). Of the predicted 8,322 genes, 8,256 (99.2%) were expressed under at least one condition (see Table S1 in the supplemental material). An analysis of the genes that were among the 500 most strongly expressed under each of the five growth conditions (top 500 analysis) (see Table S1) revealed several genes that were highly expressed under all cultivation conditions. These encoded mostly ribosomal proteins, translation initiation and elongation factors, histones, actin, thioredoxin, proteases, and enzymes from basic metabolism. The latter included enzymes from the tricarboxylic acid (TCA) cycle, mitochondrial respiration, the pentose-phosphate pathway, glycolysis,

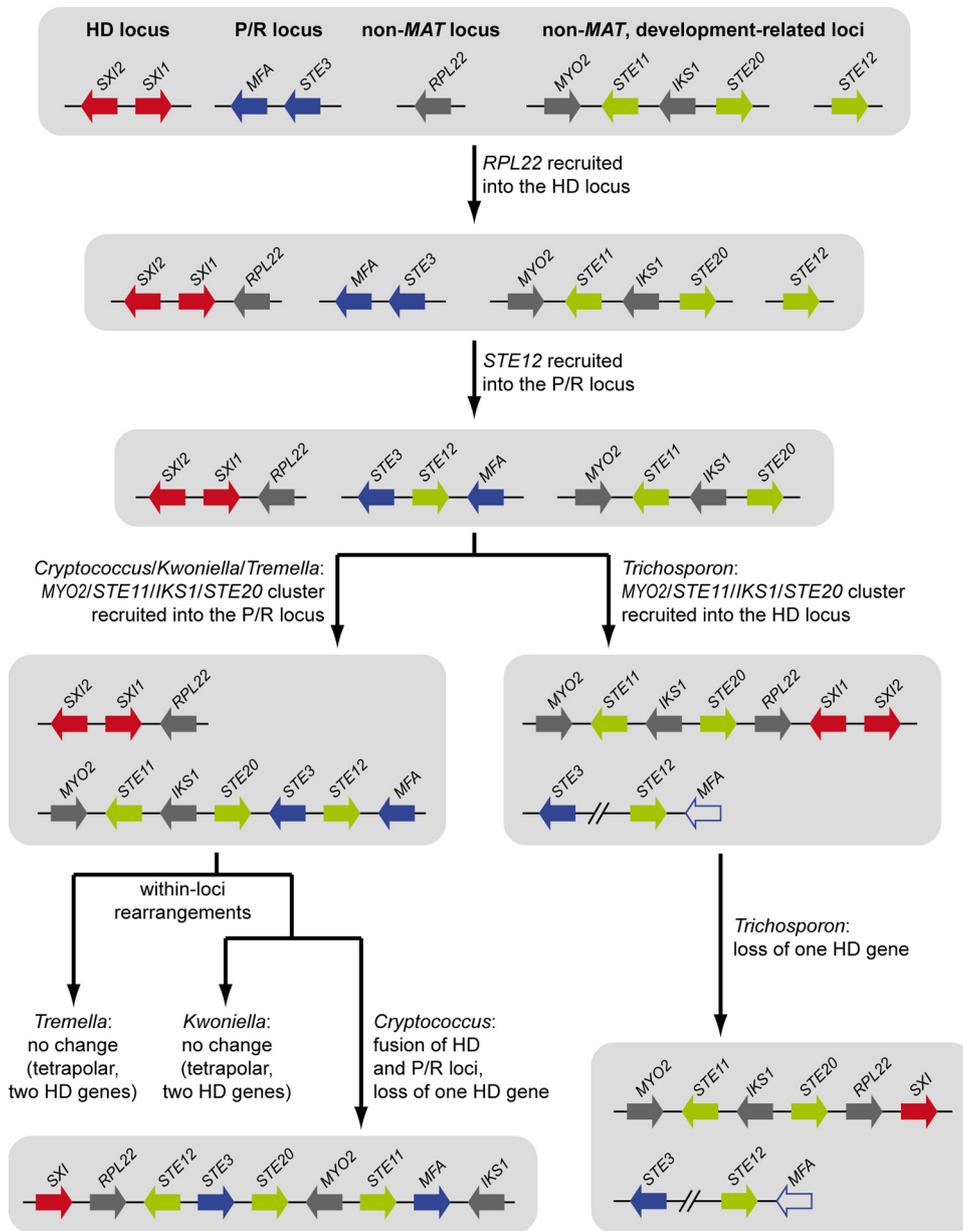


FIG 3 Model for the evolution of mating-type loci in *Tremellomycetes*. Genes at the MAT loci containing homeodomain transcription factor genes (HD locus) or pheromone and receptor genes (P/R locus) are shown in red and blue, respectively. Genes involved in sexual development but not originally part of a MAT locus are shown in green; other genes are shown in gray. Only genes from the *C. neoformans* MAT locus that are also linked to the core MAT genes (*STE3*, HD genes) in the *Trichosporon* lineage are shown (*STE11*, *STE12*, *STE20*, *IKS1*, *MYO2*, and *RPL22*); other genes present at the MAT loci were left out for the sake of clarity. A trend toward integrating other developmental genes into the MAT loci is reflected in the recruitment of the *STE12* gene into the P/R locus and the subsequent recruitment of an ancestral *STE11/STE20* cluster into the P/R locus in the *Cryptococcus/Kwoniella/Tremella* lineage and into the HD locus in the *Trichosporon* lineage. In *Cryptococcus*, both MAT loci were subsequently fused. The loss of one HD gene at the HD-containing locus occurred independently in *Cryptococcus* and *Trichosporon*. Gene names are given according to naming in the *C. neoformans* MAT locus. A putative mating pheromone gene, *MFA*, in the *Trichosporon* lineage is shown in outline only in order to indicate that no such gene has been definitively identified in the *Trichosporon* lineage. The phylogenetic relationships are depicted according to descriptions in reference 20 and the illustrations in Fig. 1.

amino acid metabolism, glycogen metabolism, and cell wall synthesis and degradation.

Transcriptome analysis under lipid-accumulating conditions. Under nonlimiting cultivation conditions, analysis of total lipids showed a maximum lipid formation of 500 mg/liter, or 4% (wt/wt) of the total biomass with any carbon source. The lipid fraction contained palmitic acid ($C_{16:0}$), palmitoleic acid ($C_{16:1}$),

stearic acid ($C_{18:0}$), oleic acid ($C_{18:1}$), linoleic acid ($C_{18:2}$), and linolenic acid ($C_{18:3}$) as their main constituents (see Fig. S4 in the supplemental material). Lowering the concentration of ammonium sulfate from 4 g/liter to 1.2 mg/liter in the NLM sample led to an increased lipid yield of 5 g/liter and a lipid content as high as 50% (wt/wt) (see Fig. S4). In contrast, lowering the phosphate concentration in the PLM sample cohort from 0.4 mM to

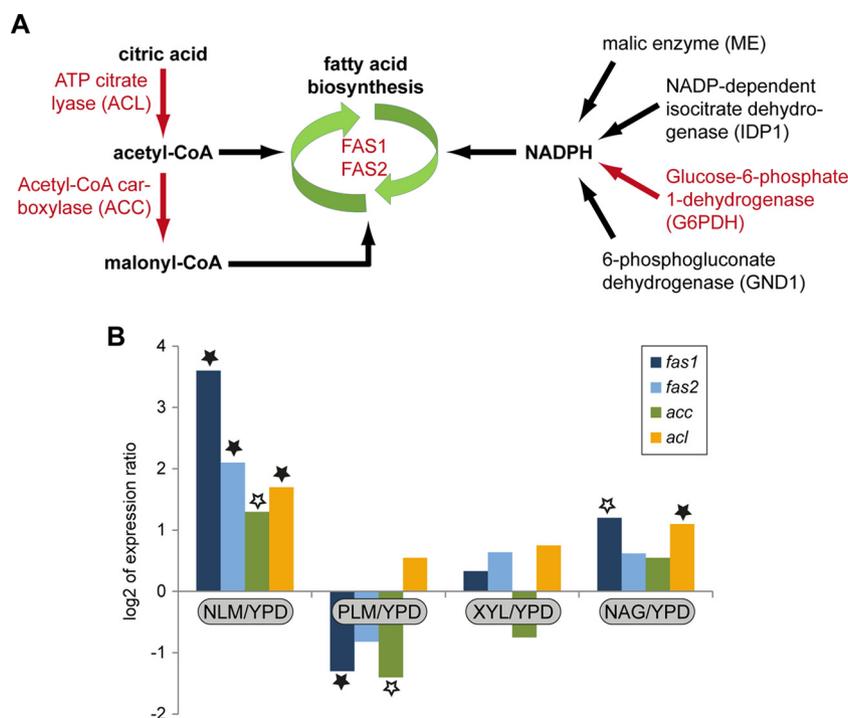


FIG 4 Lipid biosynthesis in *T. oleaginosus*. (A) Core lipid biosynthesis enzymes. Corresponding genes that are upregulated during growth on NLM versus YPD are marked in red (for details, see Table S3 in the supplemental material). FAS, fatty acid synthase. (B) Expression of core lipid biosynthesis genes under different growth conditions. Expression was analyzed by RNA-seq (three independent biological replicates). Genes that were significantly differentially expressed are indicated with a black star (adjusted *P* value, ≤ 0.05) or a white star (adjusted *P* value, ≤ 0.1).

0.14 mM, which significantly increases lipid synthesis in *Cryptococcus curvatus* (39), had only marginal effects in our experiments, resulting in 15% (wt/wt) lipid content with respect to the dried biomass (see Fig. S4).

The transcriptome from cultivation under nitrogen starvation with xylose as sole carbon source (NLM) was compared to that in YPD. A comparison to cultivation using a xylose-based medium with an abundant nitrogen source (XYL) served as an additional control. The automated annotation of core metabolism genes was manually confirmed (see Table S3 in the supplemental material), and we compared expression patterns of selected genes between different growth conditions. The transcriptome response under nitrogen limitation differed from that under all other conditions, indicating a different physiology during lipid assimilation (see Fig. S4 in the supplemental material). Differences in growth conditions also had a significant effect on the fatty acid profiles. For all data sets, oleic acid ($C_{18:1}$) was the main component of the *T. oleaginosus*-derived triglyceride fraction (see Fig. S4). Interestingly, in comparison to other growth conditions, the amount of linoleic acid ($C_{18:2}$) was 60 to 80% decreased under N-limiting growth conditions. The biochemical basis for this shift is yet unclear.

The manually annotated genes were used to reconstruct the central lipid metabolism in *T. oleaginosus* (Fig. 4; see also Table S3 in the supplemental material). Lipid metabolism requires an increased supply of NADPH for fatty acid biosynthesis. The most important producers of NADPH are malic enzyme (ME, Triol1|285398) and NADP⁺-dependent isocitrate dehydrogenase (IDP1, Triol1|288361), as well as glucose-6-phosphate dehydrogenase (G6PDH, Triol1|300435) and 6-phosphogluconate dehydrogenase (GND1, Triol1|307888) from the pentose-phosphate

cycle. ME belongs to the top 500 transcripts under nitrogen and phosphate limitation, but not under control cultivation conditions (see Table S1 in the supplemental material). ME was 2-fold upregulated in the comparison of NLM versus XYL medium (see Table S3). However, in comparison to YPD, no upregulation was found. This is consistent with *R. toruloides* data (14), where ME is downregulated under nitrogen starvation. Interestingly, overexpression of ME increases lipid synthesis in several organisms 3- to 4-fold (10, 11). Similar to *R. toruloides*, G6PDH was upregulated under nitrogen starvation. Transcript levels of IDP1 and GND1 decreased (see Table S3). As *T. oleaginosus* was cultivated on the pentose xylose as the carbon source, involvement of the oxidative part of the pentose-phosphate shunt pathway was somewhat unexpected; however, an explanation might be the need to generate NADPH for lipid biosynthesis.

Fatty acid synthase (FAS) initiates lipid synthesis by forming acyl-CoA from acetyl-CoA, malonyl-CoA, and NADPH. FAS in fungi can consist of one or two subunits; the latter is the case for *Tremellomycetes*, where two genes under the control of a divergent promoter encode the FAS complex (14, 40–43). Both subunits are also present and encoded by divergently transcribed genes in *T. oleaginosus* (*FAS1*, Triol1|330866; *FAS2*, Triol1|279681) (see Fig. S5 in the supplemental material). ATP:citrate lyase (ACL) generates acetyl-CoA from citrate, and acetyl-CoA carboxylase (ACC) further produces malonyl-CoA for fatty acid synthesis. *FAS1*, *FAS2*, *ACC*, and *ACL* are significantly upregulated under nitrogen limitation but not under phosphate limitation, which reflects the differences in lipid accumulation (Fig. 4). While these genes were also significantly overexpressed in *R. toruloides* (14), in

the ascomycete *Y. lipolytica* neither of them displayed a significant change in transcription level under nitrogen starvation (7).

Later-stage biosynthesis of neutral lipids and membranes involves transport of CoA-bound fatty acids into peroxisomes, desaturation and elongation reactions, and finally transfer of the acyl moiety to the glycerol backbone. It is expected that cells grown in limited medium also respond in terms of their lipid composition. The two genes for fatty acid desaturases, Triol1|289557 and Triol1|308253, are upregulated in comparison with NLM/XYL but not NLM/YPD, and both desaturase genes are highly transcribed. The first belongs to the top 500 transcripts under NLM, PLM, and YPD conditions, the second of the top 500 transcripts that are shared under all cultivation conditions. In *Y. lipolytica*, one of the fatty acid desaturases also showed a higher expression level under nitrogen starvation (7). 1-Acyl-*sn*-glycerol-3-phosphate acyltransferase (Triol1|249914), which is involved in neutral lipid synthesis, is upregulated in NLM compared to XYL medium. Serine palmitoyltransferase (Triol1|95840), the first and committed step for sphingolipid synthesis, was not upregulated.

In *R. toruloides*, several lipid-degrading enzymes were upregulated under lipid-accumulating conditions, including lipases and the glyoxylate pathway enzymes isocitrate lyase and malate synthase. This was attributed to the presence of free fatty acids due to an elevated autophagy process (14). In *T. oleaginosus*, isocitrate lyase (Triol1|282916) is upregulated in NLM compared to XYL medium, and malate synthase (Triol1|288256) is upregulated in NLM compared to both YPD and XYL media (see Table S3 in the supplemental material), suggesting a similar mechanism. Acyl-CoA synthetases initiate the degradation of free fatty acids. The 13 putative acyl-CoA synthetases in *T. oleaginosus* show distinct patterns of regulation. Under nitrogen starvation, two are upregulated while three are downregulated (see Table S3), indicating a highly differentiated lipid metabolism in this species.

It is interesting that in several fungi, including *C. neoformans*, sexual development is induced by starvation, and specifically by nitrogen limitation (44–49). In *T. oleaginosus*, the putative mating-type transcription factor gene *SXII* is upregulated under nitrogen limitation and growth with NAG as carbon source compared to YPD. This is also the case for the two putative developmental genes *STE11* and *STE20*, which are linked to *SXII* on the same scaffold (see Fig. S1 in the supplemental material), suggesting that a similar connection between nitrogen starvation and sexual development might be present in *T. oleaginosus*. Mating-type genes of other species have already been shown to regulate not only sexual development but also other processes, e.g., pathogenicity in the basidiomycetes *C. neoformans* and *Ustilago maydis* and antibiotic production in the ascomycete *Penicillium chrysogenum* (50–55). Sexual development has not yet been observed in *T. oleaginosus*, because currently only one mating-type configuration is known (see Text S1 in the supplemental material); however, relationships between mating-type-dependent regulation, developmental events, and lipid accumulation will be of interest for future applications.

Nitrogen metabolism. As described above, *T. oleaginosus* accumulates lipids under nitrogen starvation. Besides metabolic pathways directly involved in fatty acid biosynthesis, this condition also leads to changes in nitrogen metabolism itself that can be important factors in biotechnological applications. Amino acid transporters and permeases play an important role for nitrogen metabolism, and a total of 66 corresponding genes were identified

in *T. oleaginosus* (see Table S3 in the supplemental material). Under nitrogen starvation, 32 were found to be upregulated, 11 were downregulated, and 23 did not change significantly. Under phosphate-limiting conditions, 19 were upregulated and 17 were downregulated, whereas the majority (30) were not significantly changed compared to growth in YPD. Zhu et al. also found upregulation of many amino acid transporters in *R. toruloides* under nitrogen-limiting conditions (14). This was confirmed for *Y. lipolytica* by Morin et al. (7), who compared expression levels between stages of biomass production and lipid accumulation. The downregulation of some amino acid transporters in minimal medium compared to YPD in *T. oleaginosus* might be a reaction to the smaller amount of peptides and amino acids in the minimal medium.

An important strategy for cells to adapt to nitrogen scarcity is the degradation of nonessential proteins to peptides and ultimately amino acids, which are then again available for protein biosynthesis. We found that six proteases were upregulated and another six were downregulated during nitrogen limitation.

The source of nitrogen plays an important role for the accumulation of lipids. In *R. toruloides*, it was found that organic nitrogen compounds such as urate or urea led to a higher accumulation of lipids than observed with inorganic sources (e.g., NH₄Cl) (56). Based on data from *C. neoformans* (57), the urate catabolic pathway was annotated in *T. oleaginosus* (Fig. 5; see also Table S3 in the supplemental material). In *C. neoformans*, the gene encoding allantoinase, *DAL1*, and the urate oxidase gene *URO1* are induced in the presence of uric acid (58). In *T. oleaginosus*, we found a 5.6-fold upregulation of *URO1* (Triol1|306408) but a downregulation of all subsequent steps of urate utilization, including *URO2* (Triol1|317784), *URO3* (Triol1|309759), and *DAL1* (Triol1|311323) under nitrogen-limited conditions compared to full medium. In contrast, the majority of allantoin permeases were found to be upregulated, presumably to facilitate the import of remaining nitrogen sources. We also found a strong upregulation of the urea transporter gene *DUR3* (Triol1|161533) and the putative ammonia transporter genes (Triol1|302913 and Triol1|69276), all of which are also among the 500 most strongly expressed genes under nitrogen-limited conditions (see Tables S1 and S3 in the supplemental material). Furthermore, the gene for the urease *URE1* was upregulated. Urease acts downstream of *DUR3* to convert urea into ammonium for central nitrogen metabolism (Fig. 5).

A central element of the response to nitrogen limitation is the conversion of ammonium to glutamate and glutamine (Fig. 5). We found a strong upregulation of glutamate synthase (GOGAT; Triol1|329209), while glutamine synthetase 2 (GLN2; Triol1|283213), which catalyzes the reverse reaction, was downregulated. No change was found for glutamate dehydrogenase 1 (GDH1; Triol1|286063), while the enzyme for the reverse reaction, GDH2 (Triol1|299402), was downregulated. These transcript changes suggest a shift toward the synthesis of glutamate, presumably in an effort to make the remaining nitrogen available for protein biosynthesis (Fig. 5).

The response in *T. oleaginosus* differs from that observed in *R. toruloides*, where GDH1, GDH2, and GLN1 were upregulated under nitrogen starvation (14), and from *Y. lipolytica*, where GLN1 was also upregulated (7). However, in both fungi, genes encoding transporters for nitrogen-containing compounds like ammonia and amino acids were upregulated, similar to *T. oleagi-*

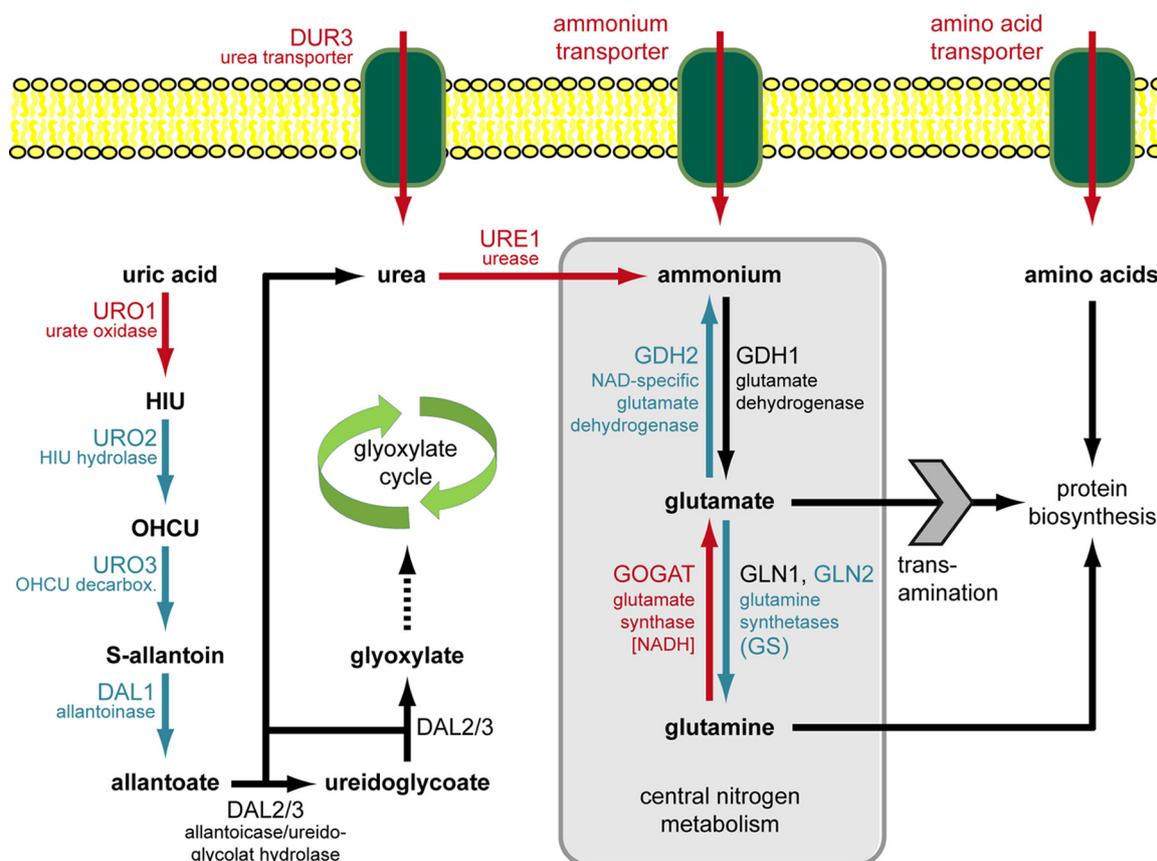


FIG 5 Key enzymes of central and peripheral nitrogen metabolism in *T. oleaginosus*. Corresponding genes that are upregulated during growth on NLM versus YPD are marked in red, and downregulated genes are shown in blue (for details, see Table S3 in the supplemental material). HIU, hydroxyisourate; OHCU, 2-2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazole.

nosus, suggesting that parts of the response to nitrogen limitation are similar in these fungi but distribution of nitrogen-based metabolites toward downstream pathways may differ.

Transcriptome analysis using alternative carbon sources. Beside cellulose, chitin is the most abundant polysaccharide in nature. It is a major component in the cell walls of fungi and widespread in other groups, such as insects and crustaceans. Here, we have used the chitin monomer NAG as the sole carbon source to analyze the resulting changes in transcription in *T. oleaginosus*. NAG can be channeled into glycolysis via fructose 6-phosphate (Fru-6P) or into chitin biosynthesis, with both pathways requiring phosphorylation of NAG first (Fig. 6). A predicted NAG kinase with homology to the *C. albicans* NAG kinase CaNAG5 (59) is present in *T. oleaginosus* (Triol1|249892) and upregulated during growth on NAG (Fig. 6; see also Table S3 in the supplemental material). Enzymes involved in chitin biosynthesis downstream of phosphorylated NAG are present in *T. oleaginosus*, including seven chitin synthase genes that represent the chitin synthase classes I to V, similar to the situation in *C. neoformans* (60), but these genes are not consistently upregulated during growth on NAG (see Table S3). In contrast, genes encoding proteins required for conversion to Fru-6P (*N*-acetyl-glucosamine-6-phosphate deacetylase [Triol1|281629] and glucosamine-6-phosphate deaminase [Triol1|281628]) are 180- and 1,000-fold upregulated, respectively. Both genes were among the 500 most strongly expressed genes in the NAG samples, but not in other samples

(Fig. 6; see also Tables S1 and S3), reflecting the large extent to which *T. oleaginosus* channels NAG into its primary metabolism. Interestingly, genes for glycolysis are not strongly regulated, while genes playing a role in the tricarboxylic acid cycle and respiration are mostly downregulated during growth on NAG (see Table S3).

For use of chitin as a carbon source, chitinolytic enzymes are required for breakdown of the polymer into oligomeric and monomeric units (61). *T. oleaginosus* encodes three putative chitinases and two putative β -*N*-acetylhexosaminidases (see Table S3 in the supplemental material), similar to *C. neoformans*, which encodes four and one enzyme, respectively (62). Two of the chitinase genes are downregulated during growth on NAG, probably reflecting feedback inhibition, while one β -*N*-acetylhexosaminidase gene is upregulated and the other downregulated (see Table S3). In summary, the pronounced upregulation of genes involved in channeling NAG into Fru-6P, and thus into primary metabolism, together with the large number of proteases and amino acid transporters (see above) suggest that *T. oleaginosus* is well-adapted to protein- and chitin-rich environments.

Conclusions. Here, we have analyzed the genome of the lipid-accumulating basidiomycete yeast *T. oleaginosus* as well as transcriptomes under various physiological conditions. A key finding was the lineage-specific arrangement of mating-type genes, indicating that recruitment of developmental genes to the ancestral tetrapolar mating-type loci occurred independently in the *Trichosporon* and *Cryptococcus* lineages, supporting the hypothesis of a

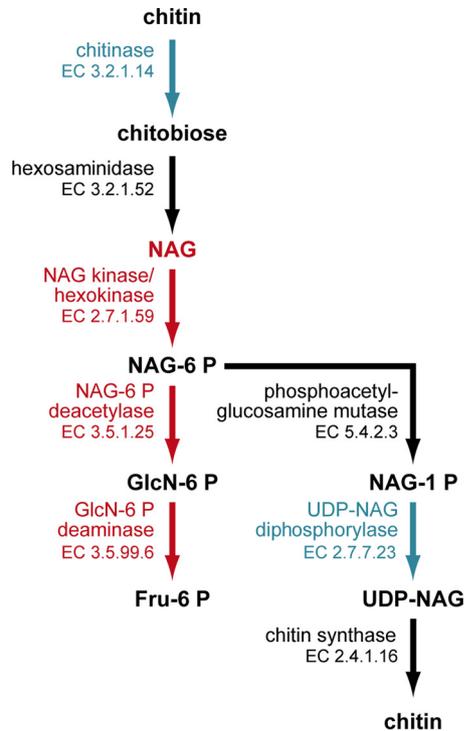


FIG 6 NAG is mainly channeled into primary metabolism in *T. oleaginosus*. Pathways were constructed according to the predicted *C. neoformans* pathways in the KEGG database (77) and the published data for *C. albicans* and *C. neoformans* (59, 60, 62). Genes that are upregulated during growth on NAG are marked in red, and downregulated genes are shown in blue (for details, see Table S3 in the supplemental material).

general trend toward larger mating-type regions in fungi, similar to the evolution of sex chromosomes in animals and plants. Furthermore, the transcriptome data showed specific metabolic shifts within lipid and nitrogen metabolism under nitrogen limitation, concurrent with the accumulation of fatty acids under this condition. We also showed that *N*-acetylglucosamine is efficiently shuttled into primary metabolism, supporting the ability of *T. oleaginosus* to grow on a variety of substrates, including chitin-rich ones. The genome and transcriptome data are valuable resources for future biotechnological applications using *T. oleaginosus* and for the analysis of the evolution of sex in eukaryotes.

MATERIALS AND METHODS

Strains and growth conditions. *T. oleaginosus* strains IBC0246 (from the laboratory collection of the culture collection of the research group Industrial Biocatalysis, TU, Munich), ATCC 20508 and ATCC 20509 were maintained on YPD (see Text S1 in the supplemental material). For transcriptome analysis, *T. oleaginosus* was grown on YPD (complete medium), NLM (nitrogen limitation), PLM (phosphate limitation), XYL (xylose as carbon source), or NAG (*N*-acetyl glucosamine as carbon source); media compositions are provided in Text S1.

Analysis of total lipid content. For the determination of lipid content, cells were lysed using high-pressure homogenization and lipids were extracted according to the methods of Bligh and Dyer (63). The solvent phase was then evaporated under nitrogen, and the dry weight of the remaining lipids was determined.

Analysis of lipid profiles under different growth conditions. *T. oleaginosus* was grown in shake flask cultures (50 ml) for 6 days at 28°C and 250 rpm in the different media (see Text S1 in the supplemental material).

The biomass was subsequently harvested and prepared for lipid extraction. The direct transesterification of wet biomass was performed according to a modified protocol of Griffiths et al. (64) with the following modifications: we replaced the C₁₇ tag with a C₁₂ tag, we replaced BF3 methanol with an HCl-methanol solution, and we omitted the C₁₉-ME. Subsequently, the resulting fatty acid methyl esters (FAME) extract was injected into a Shimadzu GC-2010 Plus system equipped with a flame ionization detector and Zebron ZB-WAX column (30 m by 0.32 mm [inner diameter], 0.25- μ m film thickness; Phenomenex, USA). Standard split/splitless injection was used, with a split ratio of 10 and an injector temperature of 240°C. The column temperature was increased from 150°C to 240°C at 5°C/min. Nitrogen (3 ml/min) was used as the carrier gas, and the detector temperature was 245°C. Peaks were identified by retention time using the FAMEs marine oil standard (Restek). Peak areas were used to quantify each FAME relative to the internal standards.

Genome sequencing and assembly. Genomic DNA was prepared as described in Text S1 in the supplemental material. One hundred nanograms of genomic DNA was sheared to 270 bp by using the Covaris E210 apparatus and size selected by using SPRI beads (Beckman Coulter). The fragments were treated with end repair, A-tailing, and ligation of Illumina compatible adapters (IDT, Inc.) by using the KAPA Illumina library creation kit (KAPA Biosystems). Quantitative PCR was used to determine the concentrations of the libraries. Libraries were sequenced as paired ends with a read length of 150 bp on the Illumina HiSeq system. The resulting fastq files were quality-control filtered to separate mitochondrial data and remove artifact/process contamination, and the information was initially assembled using the Velvet assembler (65). The Velvet assembly was used to simulate long mate-pair libraries, the first with inserts of 2,000 \pm 50 bp, the second with inserts of 5,000 \pm 50 bp. These two software-constructed libraries were then assembled together with the original Illumina library by using AllPathsLG release version R47710 (66) to produce a 19.8-Mb assembly in 221 contigs and 180 scaffolds with a 115.8 \times read depth coverage (Table 1). K-mer analysis of the filtered sequence reads suggested that the genome is haploid (data not shown).

Transcriptome sequencing and quantitative analysis of gene expression. Five different growth conditions (YPD, NLM, PLIM, XYL, and NAG) were analyzed with two (YPD) or three (all others) independent biological replicates (see Text S1 in the supplemental material). Total RNA was prepared as described in Text S1. Stranded cDNA libraries were generated using the Illumina Truseq stranded RNA LT kit (for details, see Text S1). Paired-end sequencing was performed using an Illumina HiSeq2000 instrument, generating 2 150-bp or 2 157-bp reads from each library (see Text S1).

Raw fastq file reads were filtered and trimmed using the JGI QC pipeline, resulting in the filtered fastq file. Using BBduk (<https://sourceforge.net/projects/bbmap/>), raw reads were evaluated for artifact sequence by kmer matching (kmer value, 25), allowing 1 mismatch, and detected artifacts were trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads, and reads containing any Ns were removed. Quality trimming was performed using the Phred trimming method set at Q6. Finally, following trimming, reads under the length threshold were removed (minimum length, 1/3 of the original read length). For *de novo* transcriptome assembly, which was used for annotation (see below), 432,000,000 paired-end 150-bp Illumina HiSeq-2000 reads of stranded RNA-seq data (10 libraries) were used as input for Rnnotator v3.3.1 (67). For quantitative analysis of gene expression, reads from each library were aligned to the reference genome by using TopHat (68). DESeq2 (69) was used to determine which genes were differentially expressed between pairs of conditions. For an analysis of the 500 most strongly expressed genes under each condition (Top 500 analysis), RPKM (reads per kilobase per million mapped reads) values were calculated for each gene and condition. For details, see Text S1 in the supplemental material.

Genome annotation. The genome assembly of *T. oleaginosus* was annotated using the JGI annotation pipeline (70), which combines several gene prediction and annotation methods and integrates the annotated

genome into the Web-based fungal resource MycoCosm (27) for comparative genomics. Details are provided in Text S1 in the supplemental material.

Synten analysis. An orthology-based analysis of synten was performed by determining orthologs for all *T. oleaginosus* proteins in the predicted proteomes of three other *Tremellomycetes* by reciprocal BLAST analysis (71) and using custom-made Perl scripts based on BioPerl modules (72) to determine the positions of orthologous proteins on sequenced scaffolds or chromosomes.

Phylogenetic analyses. Multiple alignments were created in ClustalX (73) and trimmed with Jalview (74), and the same alignment was used for analysis by distance matrix (DM), maximum parsimony (MP), or Bayesian methods (for details, see Text S1 in the supplemental material). For generating a species tree of *Tremellomycetes*, the genomes of *T. asahii* (26), *T. mesenterica* (75), *C. neoformans* (24), *C. gattii* (22), and *Coprinopsis cinerea* (76) were used for comparative purposes (details are given in Text S1 in the supplemental material).

Accession numbers. The genome sequence and annotation are publicly available via the JGI MycoCosm portal (<http://genome.jgi.doe.gov/Triol1>) (27). The BioProject accession number for the genome is PRJNA239490. The sequence reads that were used for the assembly of the *T. oleaginosus* genome were submitted to the NCBI sequence read archive (accession number SRX873336) and the assembly was submitted to GenBank (accession number JZUH000000001). The BioProject accession number for the transcriptome is PRJNA250808. Sequences for the *sx11* and *ste3* regions of strains ATCC 20508 and ATCC 20509 and the ITS region of strain IBC0246 were submitted to GenBank (NCBI) under the following accession numbers, respectively: KM821409 (*sx11* ATCC 20508), KM821410 (*sx11* ATCC 20509), KM821407 (*ste3* ATCC 20508), KM821408 (*ste3* ATCC 20509), and KM821406 (ITS region IBC0246).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00918-15/-/DCSupplemental>.

- Table S1, XLSX file, 2.9 MB.
- Table S2, XLSX file, 0.02 MB.
- Table S3, XLSX file, 0.1 MB.
- Text S1, PDF file, 0.1 MB.
- Figure S1, PDF file, 0.2 MB.
- Figure S2, PDF file, 0.3 MB.
- Figure S3, PDF file, 0.2 MB.
- Figure S4, PDF file, 0.2 MB.
- Figure S5, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

We thank Ann-Christin Müller and Swenja Ellßel for excellent technical assistance, Stephen Mondo for GenBank submissions, and Joseph Spatafora and the 1000 Fungal Genomes Project for making the sequencing of the *T. oleaginosus* genome and transcriptomes possible.

The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231. Minou Nowrousian acknowledges funding by the German Research Foundation (DFG) and would like to thank Ulrich Kück for his support at the Department of General and Molecular Botany. Sheng Sun and Joseph Heitman were supported by NIAID/NIH R37 merit award AI39115-17. Thomas Brück, Felix Bracharz, and Jan Lorenzen acknowledge funding through the Federal Ministry for Education and Research (BMBF) project “Advance Biomass Value” (award 03SF0446A). Robert Kourist gratefully acknowledges financial support from the Ministry for Innovation, Science and Investigation of the State of North Rhine-Westphalia (grant 005-1503-0006).

REFERENCES

1. Biermann U, Bornscheuer U, Meier MA, Metzger JO, Schäfer HJ. 2011. Oils and fats as renewable raw materials in chemistry. *Ang Chem Int Ed* 50:3854–3871. <http://dx.doi.org/10.1002/anie.201002767>.
2. Garbe D, Reißer S, Brück T. 2014. Nachhaltige Bausteine für die Kunststoff-Herstellung. *Chemie* 48:284–295. <http://dx.doi.org/10.1002/ciuz.201400673>.
3. Costanza R, de Groot R, Sutton P, van der Ploeg S, Anderson SJ, Kubiszewski I, Farber S, Turner RK. 2014. Changes in the global value of ecosystem services. *Glob Environ Change* 26:152–158. <http://dx.doi.org/10.1016/j.gloenvcha.2014.04.002>.
4. Sodhi NS, Koh LP, Brook BW, Ng PK. 2004. Southeast Asian biodiversity: an impending disaster. *Trends Ecol Evol* 19:654–660. <http://dx.doi.org/10.1016/j.tree.2004.09.006>.
5. Ageitos JM, Vallejo JA, Veiga-Crespo P, Villa TG. 2011. Oily yeasts as oleaginous cell factories. *Appl Microbiol Biotechnol* 90:1219–1227. <http://dx.doi.org/10.1007/s00253-011-3200-z>.
6. Kumar S, Kushwaha H, Bachhawat AK, Raghava GP, Ganesan K. 2012. Genome sequence of the oleaginous red yeast *Rhodospiridium toruloides* MTCC 457. *Eukaryot Cell* 11:1083–1084. <http://dx.doi.org/10.1128/EC.00156-12>.
7. Morin N, Cescut J, Beopoulos A, Lelandais G, Le Berre V, Uribelarrea J-L, Molina-Jouve C, Nicaud J-M. 2011. Transcriptomic analyses during the transition from biomass production to lipid accumulation in the oleaginous yeast *Yarrowia lipolytica*. *PLoS One* 6:e27966. <http://dx.doi.org/10.1371/journal.pone.0027966>.
8. Papanikolaou S, Aggelis G. 2010. *Yarrowia lipolytica*: a model microorganism used for the production of tailor-made lipids. *Eur J Lipid Sci Technol* 112:639–654. <http://dx.doi.org/10.1002/ejlt.200900197>.
9. Beopoulos A, Cescut J, Haddouche R, Uribelarrea J-L, Molina-Jouve C, Nicaud J-M. 2009. *Yarrowia lipolytica* as a model for bio-oil production. *Prog Lipid Res* 48:375–387. <http://dx.doi.org/10.1016/j.plipres.2009.08.005>.
10. Brown S, Luttringer S, Yaver D, Berry A. 3 March 2011. Methods for improving malic acid production in filamentous fungi. U.S. patent 20110053233 A1.
11. Xue J, Niu Y-F, Huang T, Yang W-D, Liu J-S, Li H-Y. 2015. Genetic improvement of the microalga *Phaeodactylum tricornutum* for boosting neutral lipid accumulation. *Metab Eng* 27:1–9. <http://dx.doi.org/10.1016/j.ymben.2014.10.002>.
12. Zhang Y, Adams IP, Ratledge C. 2007. Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiology* 153:2013–2025. <http://dx.doi.org/10.1099/mic.0.2006/002683-0>.
13. Tai M, Stephanopoulos G. 2013. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metab Eng* 15:1–9. <http://dx.doi.org/10.1016/j.ymben.2012.08.007>.
14. Zhu Z, Zhang S, Liu H, Shen H, Lin X, Yang F, Zhou YJ, Jin G, Ye M, Zou H, Zhao ZK. 2012. A multi-omic map of the lipid-producing yeast *Rhodospiridium toruloides*. *Nat Commun* 3:1112. <http://dx.doi.org/10.1038/ncomms2112>.
15. Gujjari P, Suh SO, Coumes K, Zhou JJ. 2011. Characterization of oleaginous yeasts revealed two novel species: *Trichosporon cacaoliposimilis* sp. nov. and *Trichosporon oleaginosus* sp. nov. *Mycologia* 103:1110–1118. <http://dx.doi.org/10.3852/10-403>.
16. Colombo AL, Padovan AC, Chaves GM. 2011. Current knowledge of *Trichosporon* spp. and trichosporonosis. *Clin Microbiol Rev* 24:682–700. <http://dx.doi.org/10.1128/CMR.00003-11>.
17. Middelhoven WJ, Scorzetti G, Fell JW. 2004. Systematics of the anamorphic basidiomycetous yeast genus *Trichosporon* behrend with the description of five novel species: *Trichosporon vadense*, *T. smithiae*, *T. dehoogii*, *T. scarabaeorum* and *T. gamsii*. *Int J Syst Evol Microbiol* 54:975–986. <http://dx.doi.org/10.1099/ijs.0.02859-0>.
18. Fernandez JG, Ingber DE. 2013. Bioinspired chitinous material solutions for environmental sustainability and medicine. *Adv Funct Mater* 23:4454–4466. <http://dx.doi.org/10.1002/adfm.201300053>.
19. Zhou H, Cheng JS, Wang BL, Fink GR, Stephanopoulos G. 2012. Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*. *Metab Eng* 14:611–622. <http://dx.doi.org/10.1016/j.ymben.2012.07.011>.
20. Heitman J, Sun S, James TY. 2013. Evolution of fungal sexual reproduction. *Mycologia* 105:1–27. <http://dx.doi.org/10.3852/12-253>.
21. Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, Kubicek CP, Kuske C, Magnuson JK, Martin F, Spatafora JW, Tsang A, Baker SE. 2011. Fueling the future with fungal genomics. *Mycology* 2:192–209. <http://dx.doi.org/10.1080/21501203.2011.584577>.

22. D'Souza CA, Kronstad JW, Taylor G, Warren R, Yuen M, Hu G, Jung WH, Sham A, Kidd SE, Tangen K, Lee N, Zeilmaker T, Sawkins J, McVicker G, Shah S, Gnerre S, Griggs A, Zeng Q, Bartlett K, Li W, Wang X, Heitman J, Stajich JE, Fraser JA, Meyer W, Carter D, Schein J, Krzywinski M, Kwon-Chung KJ, Varma A, Wang J, Brunham R, Fyfe M, Ouellette BFF, Siddiqui A, Marra M, Jones S, Holt R, Birren BW, Galagan JE, Cuomo CA. 2011. Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *mBio* 2(1):e00342-10. <http://dx.doi.org/10.1128/mBio.00342-10>.
23. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, Allen JE, Bosdet IE, Brent MR, Chiu R, Doering TL, Donlin MJ, D'Souza CA, Fox DS, Grinberg V, Fu J, Fukushima M, Haas BJ, Huang JC, Janbon G, Jones SJM, Koo HL, Krzywinski MI, Kwon-Chung JK, Lengeler KB, Maiti R, Marra MA, Marra RE, Mathewson CA, Mitchell TG, Perteza M, Riggs FR, Salzberg SL, Schein JE, Shvartsbeyn A, Shin H, Shumway M, Specht CA, Suh BB, Tenney A, Utterback TR, Wickes BL, Wortman JR, Wye NH, Kronstad JW, Lodge. 2005. The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* 307:1321-1324. <http://dx.doi.org/10.1126/science.1103773>.
24. Janbon G, Ormerod KL, Paulet D, Byrnes EJ, Jr, Yadav V, Chatterjee G, Mullanpudi N, Hon CN, Billmyre RB, Brunel F, Bahn YS, Chen W, Chen Y, Chow EW, Coppée JY, Floyd-Averette A, Gaillardin C, Gerik KJ, Goldberg J, Gonzalez-Hilarion S, Gujja S, Hamlin JL, Hsueh YP, Ianiri G, Jones S, Kodira CD, Kozubowski L, Lam W, Marra M, Mesner LD, Mieczkowski PA, Moyrand F, Nielsen K, Proux C, Rossignol T, Schein JE, Sun S, Wollschlaeger C, Wood IA, Zeng Q, Neuvéglise C, Newlon CS, Perfect JR, Lodge. 2014. Analysis of the genome and transcriptome of *Cryptococcus neoformans* var. *grubii* reveals complex RNA expression and microevolution leading to virulence attenuation. *PLoS Genet* 10:e1004261. <http://dx.doi.org/10.1371/journal.pgen.1004261>.
25. Yang RY, Li HT, Zhu H, Zhou GP, Wang M, Wang L. 2012. Genome sequence of the *Trichosporon asahii* environmental strain CBS 8904. *Eukaryot Cell* 11:1586-1587. <http://dx.doi.org/10.1128/EC.00264-12>.
26. Yang RY, Li HT, Zhu H, Zhou GP, Wang M, Wang L. 2012. Draft genome sequence of CBS 2479, the standard type strain of *Trichosporon asahii*. *Eukaryot Cell* 11:1415-1416. <http://dx.doi.org/10.1128/EC.00237-12>.
27. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 42:D699-D704. <http://dx.doi.org/10.1093/nar/gkt1183>.
28. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. <http://dx.doi.org/10.1186/gb-2004-5-2-r12>.
29. Kües U, James TY, Heitman J. 2011. Mating type in basidiomycetes: unipolar, bipolar, and tetrapolar patterns of sexuality, p 97-160. In Pöggeler S, Wöstemeyer J (ed), *The Mycota*, XIV. Evolution of fungi and fungal-like organisms. Springer, Berlin, Germany.
30. Kües U, Casselton LA. 1992. Homeodomains and regulation of sexual development in basidiomycetes. *Trends Genet* 8:154-155.
31. Findley K, Sun S, Fraser JA, Hsueh YP, Averette AF, Li W, Dietrich FS, Heitman J. 2012. Discovery of a modified tetrapolar sexual cycle in *Cryptococcus amyloletus* and the evolution of *MAT* in the *Cryptococcus* species complex. *PLoS Genet* 8:e1002528. <http://dx.doi.org/10.1371/journal.pgen.1002528>.
32. Guerreiro MA, Springer DJ, Rodrigues JA, Rusche LN, Findley K, Heitman J, Fonseca A. 2013. Molecular and genetic evidence for a tetrapolar mating system in the basidiomycetous yeast *Kwoniella mangrovensis* and two novel sibling species. *Eukaryot Cell* 12:746-760. <http://dx.doi.org/10.1128/EC.00065-13>.
33. Metin B, Findley K, Heitman J. 2010. The mating type locus (*MAT*) and sexual reproduction of *Cryptococcus heveanensis*: insights into the evolution of sex and sex-determining chromosomal regions in fungi. *PLoS Genet* 6:e1000961. <http://dx.doi.org/10.1371/journal.pgen.1000961>.
34. Lengeler KB, Fox DS, Fraser JA, Allen A, Forrester K, Dietrich FS, Heitman J. 2002. Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. *Eukaryot Cell* 1:704-718. <http://dx.doi.org/10.1128/EC.1.5.704-718.2002>.
35. Hsueh YP, Heitman J. 2008. Orchestration of sexual reproduction and virulence by the fungal mating-type locus. *Curr Opin Microbiol* 11:517-524. <http://dx.doi.org/10.1016/j.mib.2008.09.014>.
36. Sun S, Hsueh YP, Heitman J. 2012. Gene conversion occurs within the mating-type locus of *Cryptococcus neoformans* during sexual reproduction. *PLoS Genet* 8:e1002810. <http://dx.doi.org/10.1371/journal.pgen.1002810>.
37. Karos M, Chang YC, McClelland CM, Clarke DL, Fu J, Wickes BL, Kwon-Chung KJ. 2000. Mapping of the *Cryptococcus neoformans* *MAT α* locus: presence of mating type-specific mitogen-activated protein kinase cascade homologs. *J Bacteriol* 182:6222-6227. <http://dx.doi.org/10.1128/JB.182.21.6222-6227.2000>.
38. Fraser JA, Diezmann S, Subaran RL, Allen A, Lengeler KB, Dietrich FS, Heitman J. 2004. Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biol* 2:e384. <http://dx.doi.org/10.1371/journal.pbio.0020384>.
39. Zhang G, French WT, Hernandez R, Hall J, Sparks D, Holmes WE. 2011. Microbial lipid production as biodiesel feedstock from N-acetylglucosamine by oleaginous microorganisms. *J Chem Technol Biotechnol* 86:642-650. <http://dx.doi.org/10.1002/jctb.2592>.
40. Reich M, Göbel C, Kohler A, Buée M, Martin F, Feussner I, Polle A. 2009. Fatty acid metabolism in the ectomycorrhizal fungus *Laccaria bicolor*. *New Phytol* 182:950-964. <http://dx.doi.org/10.1111/j.1469-8137.2009.02819.x>.
41. Lomakin IB, Xiong Y, Steitz TA. 2007. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell* 129:319-332. <http://dx.doi.org/10.1016/j.cell.2007.03.013>.
42. Jenni S, Leibundgut M, Boehringer D, Frick C, Mikolásek B, Ban N. 2007. Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science* 316:254-261. <http://dx.doi.org/10.1126/science.1138248>.
43. Wang L, Chen W, Feng Y, Ren Y, Gu Z, Chen H, Wang H, Thomas MJ, Zhang B, Berquin IM, Li Y, Wu J, Zhang H, Song Y, Liu X, Norris JS, Wang S, Du P, Shen J, Wang N, Yang Y, Wang W, Feng L, Ratledge C, Zhang H, Chen YQ. 2011. Genome characterization of the oleaginous fungus *Mortierella alpina*. *PLoS One* 6:e28319. <http://dx.doi.org/10.1371/journal.pone.0028319>.
44. McClelland CM, Fu J, Woodlee GL, Seymour TS, Wickes BL. 2002. Isolation and characterization of the *Cryptococcus neoformans* *MAT α* pheromone gene. *Genetics* 160:935-947.
45. Wickes BL, Mayorga ME, Edman U, Edman JC. 1996. Dimorphism and haploid fruiting in *Cryptococcus neoformans*: association with the α -mating type. *Proc Natl Acad Sci U S A* 93:7327-7331. <http://dx.doi.org/10.1073/pnas.93.14.7327>.
46. Wang P, Perfect JR, Heitman J. 2000. The G-protein β subunit GPB1 is required for mating and haploid fruiting in *Cryptococcus neoformans*. *Mol Cell Biol* 20:352-362. <http://dx.doi.org/10.1128/MCB.20.1.352-362.2000>.
47. Shen WC, Davidson RC, Cox GM, Heitman J. 2002. Pheromones stimulate mating and differentiation via paracrine and autocrine signaling in *Cryptococcus neoformans*. *Eukaryot Cell* 1:366-377. <http://dx.doi.org/10.1128/EC.1.3.366-377.2002>.
48. Davis RL, deSerres D. 1970. Genetic and microbial research techniques for *Neurospora crassa*. *Methods Enzymol* 27A:79-143.
49. Pöggeler S, Nowrousian M, Kück U. 2006. Fruiting body development in ascomycetes, p 325-355. In Kües U, Fischer R (ed), *The Mycota*. I. Growth, differentiation and sexuality. Springer, Berlin, Germany.
50. Böhm J, Hoff B, O'Gorman CM, Wolfers S, Klux V, Binger D, Zadra I, Kürsteiner H, Pöggeler S, Dyer PS, Kück U. 2013. Sexual reproduction and mating-type mediated strain development in the penicillin-producing fungus *Penicillium chrysogenum*. *Proc Natl Acad Sci U S A* 110:1476-1481. <http://dx.doi.org/10.1073/pnas.1217943110>.
51. Bölker M, Genin S, Lehmler C, Kahmann R. 1995. Genetic regulation of mating and dimorphism in *Ustilago maydis*. *Can J Bot* 73:320-325. <http://dx.doi.org/10.1139/b95-262>.
52. Wahl R, Zairi A, Kämper J. 2010. The *Ustilago maydis* b mating type locus controls hyphal proliferation and expression of secreted virulence factors in *planta*. *Mol Microbiol* 75:208-220. <http://dx.doi.org/10.1111/j.1365-2958.2009.06984.x>.
53. Lin X, Huang JC, Mitchell TG, Heitman J. 2006. Virulence attributes and hyphal growth of *C. neoformans* are quantitative traits and the *MAT α* allele enhances filamentation. *PLoS Genet* 2:e187. <http://dx.doi.org/10.1371/journal.pgen.0020187>.
54. Lin X, Nielsen K, Patel S, Heitman J. 2008. Impact of mating type, serotype, and ploidy on the virulence of *Cryptococcus neoformans*. *Infect Immun* 76:2923-2938. <http://dx.doi.org/10.1128/IAI.00168-08>.
55. Ren P, Springer DJ, Behr MJ, Samsonoff WA, Chaturvedi S, Chaturvedi

- V. 2006. Transcription factor STE12 α has distinct roles in morphogenesis, virulence, and ecological fitness of the primary pathogenic yeast *Cryptococcus gattii*. *Eukaryot Cell* 5:1065–1080. <http://dx.doi.org/10.1128/EC.00009-06>.
56. Evans CT, Ratledge C. 1984. Effect of nitrogen source on lipid accumulation in oleaginous yeasts. *J Gen Microbiol* 130:1693–1704.
57. Lee IR, Yang L, Sebetso G, Allen R, Doan TH, Blundell R, Lui EY, Morrow CA, Fraser JA. 2013. Characterization of the complete uric acid degradation pathway in the fungal pathogen *Cryptococcus neoformans*. *PLoS One* 8:e64292. <http://dx.doi.org/10.1371/journal.pone.0064292>.
58. Lee IR, Chow EW, Morrow CA, Djordjevic JT, Fraser JA. 2011. Nitrogen metabolite repression of metabolism and virulence in the human fungal pathogen *Cryptococcus neoformans*. *Genetics* 188:309–323. <http://dx.doi.org/10.1534/genetics.111.128538>.
59. Yamada-Okabe T, Sakamori Y, Mio T, Yamada-Okabe H. 2001. Identification and characterization of the genes for N-acetylglucosamine kinase and N-acetylglucosamine-phosphate deacetylase in the pathogenic fungus *Candida albicans*. *Eur J Biochem* 268:2498–2505. <http://dx.doi.org/10.1046/j.1432-1327.2001.02135.x>.
60. Banks IR, Specht CA, Donlin MJ, Gerik KJ, Levitz SM, Lodge JK. 2005. A chitin synthase and its regulator protein are critical for chitosan production and growth of the fungal pathogen *Cryptococcus neoformans*. *Eukaryot Cell* 4:1902–1912. <http://dx.doi.org/10.1128/EC.4.11.1902-1912.2005>.
61. Adrangi S, Faramarzi MA. 2013. From bacteria to human: a journey into the world of chitinases. *Biotechnol Adv* 31:1786–1795. <http://dx.doi.org/10.1016/j.biotechadv.2013.09.012>.
62. Baker LG, Specht CA, Lodge JK. 2009. Chitinases are essential for sexual development but not vegetative growth in *Cryptococcus neoformans*. *Eukaryot Cell* 8:1692–1705. <http://dx.doi.org/10.1128/EC.00227-09>.
63. Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 37:911–917. <http://dx.doi.org/10.1139/o59-099>.
64. Griffiths MJ, van Hille RP, Harrison ST. 2010. Selection of direct transesterification as the preferred method for assay of fatty acid content of microalgae. *Lipids* 45:1053–1060. <http://dx.doi.org/10.1007/s11745-010-3468-2>.
65. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
66. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* 108:1513–1518. <http://dx.doi.org/10.1073/pnas.1017351108>.
67. Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, Sherlock G, Snyder M, Wang Z. 2010. Rnnotator: an automated *de novo* transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics* 11:663. <http://dx.doi.org/10.1186/1471-2164-11-663>.
68. Kim D, Pertege A, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36. <http://dx.doi.org/10.1186/gb-2013-14-4-r36>.
69. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <http://dx.doi.org/10.1101/002832>.
70. Grigoriev IV, Martinez DA, Salamov AA. 2006. Fungal genomic annotation. In Arora DK, Berka RM, Singh GB (ed), *Applied mycology and biotechnology*, vol 6, p 123–142. Elsevier, Amsterdam, The Netherlands. [http://dx.doi.org/10.1016/S1874-5334\(06\)80008-0](http://dx.doi.org/10.1016/S1874-5334(06)80008-0).
71. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and psi-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <http://dx.doi.org/10.1093/nar/25.17.3389>.
72. Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, Fuellen G, Gilbert JG, Korf I, Lapp H, Lehväslaiho H, Matsalla C, Mungall CJ, Osborne BI, Pocock MR, Schattner P, Senger M, Stein LD, Stupka E, Wilkinson MD, Birney E. 2002. The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* 12:1611–1618. <http://dx.doi.org/10.1101/gr.361602>.
73. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882.
74. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2: a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191. <http://dx.doi.org/10.1093/bioinformatics/btp033>.
75. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otilar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM, de Vries RP, Ferreira P, Findley K, Foster B, Gaskell J, Glotzer D, Görecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, Kohler A, Kües U, Kumar TKA, Kuo A, LaButti K, Larrondo LF, Lindquist E, Ling A, Lombard V, Lucas S, Lundell T, Martin R, McLaughlin DJ, Morgenstern I, Morin E, Murat C, Nagy LG, Nolan M, Ohm RA, Patyshakuliyeva A. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336:1715–1719. <http://dx.doi.org/10.1126/science.1221748>.
76. Stajich JE, Wilke SK, Ahrén D, Au CH, Birren BW, Borodovsky M, Burns C, Canbäck B, Casselton LA, Cheng CK, Deng J, Dietrich FS, Fargo DC, Farman ML, Gathman AC, Goldberg J, Guigó R, Hoegger PJ, Hooker JB, Huggins A, James TY, Kamada T, Kilaru S, Kodira C, Kües U, Kupfer D, Kwan HS, Lomsadze A, Li W, Lilly WW, Ma L-J, Mackey AJ, Manning G, Martin F, Muraguchi H, Natvig DO, Palmerini H, Ramesh MA, Rehmeier CJ, Roe BA, Shenoy N, Stanke M, Ter-Hovhannisyan V, Tunlid A, Velagapudi R, Vision TJ, Zeng Q, Zolan ME, Pukkila PJ. 2010. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proc Natl Acad Sci U S A* 107:11889–11894. <http://dx.doi.org/10.1073/pnas.1003391107>.
77. Tanabe M, Kanehisa M. 2012. Using the KEGG database resource. *Curr Protoc Bioinformatics Chapter 1:Unit1.12*. <http://dx.doi.org/10.1002/0471250953.bi0112s38>.