Supplementary Information for

A multi-omic map of the lipid-producing yeast *Rhodosporidium toruloides*

Zhiwei Zhu[1, 5], Sufang Zhang[1, 3], Hongwei Liu[2], Hongwei Shen[1, 3], Xinping Lin[1, 5], Fan Yang[1, 4], Yongjin J. Zhou[1, 5], Guojie Jin[1, 5], Mingliang Ye[2], Hanfan Zou[2], Zongbao K. Zhao[1, 3]*

[1] Division of Biotechnology, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 457 Zhongshan Rd., Dalian 116023, China

[2] Key Laboratory of Separation Sciences for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 457 Zhongshan Rd., Dalian 116023, China

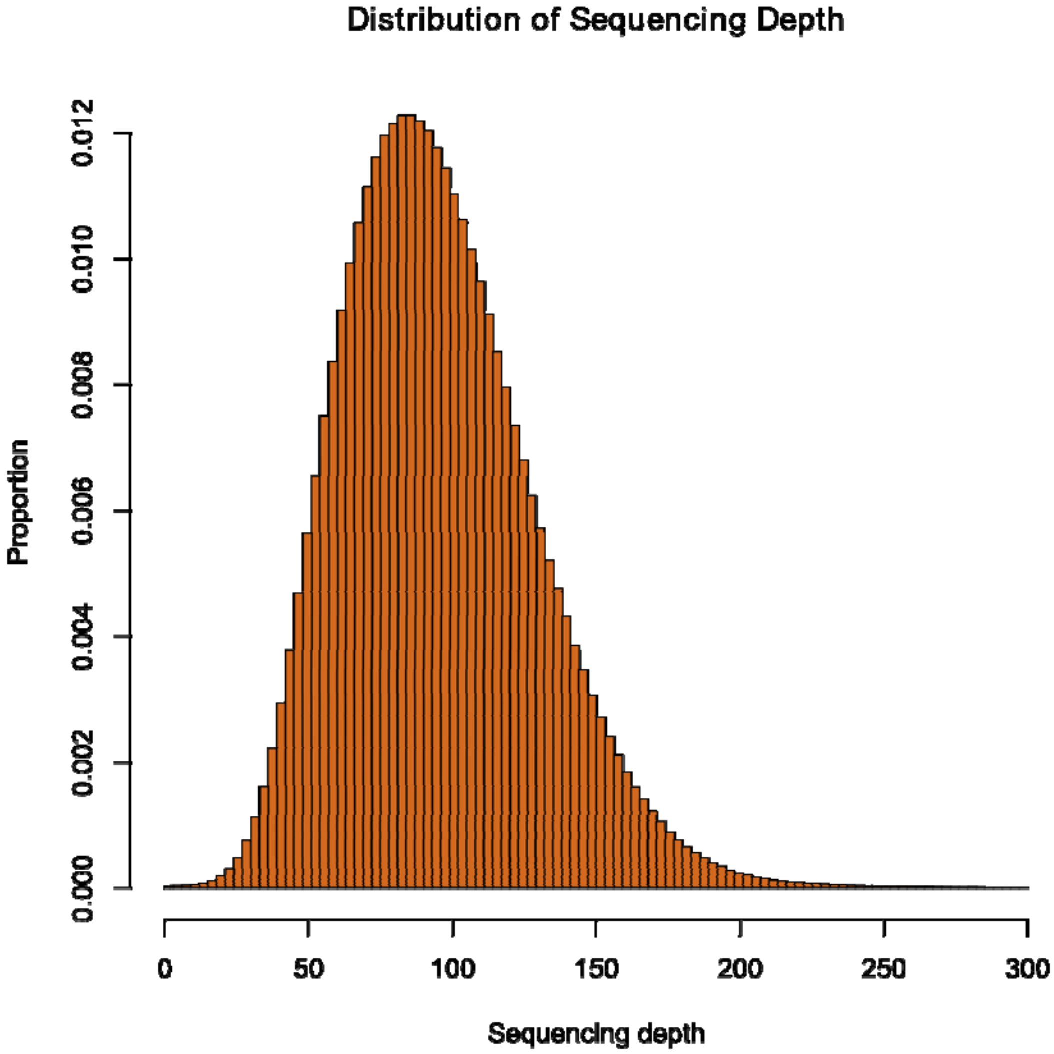
[3] Dalian National Laboratory for Clean Energy, 457 Zhongshan Rd., Dalian 116023, China

[4] School of Biological Engineering, Dalian Polytechnic University, 1 Qinggongyuan, Dalian 116034, China

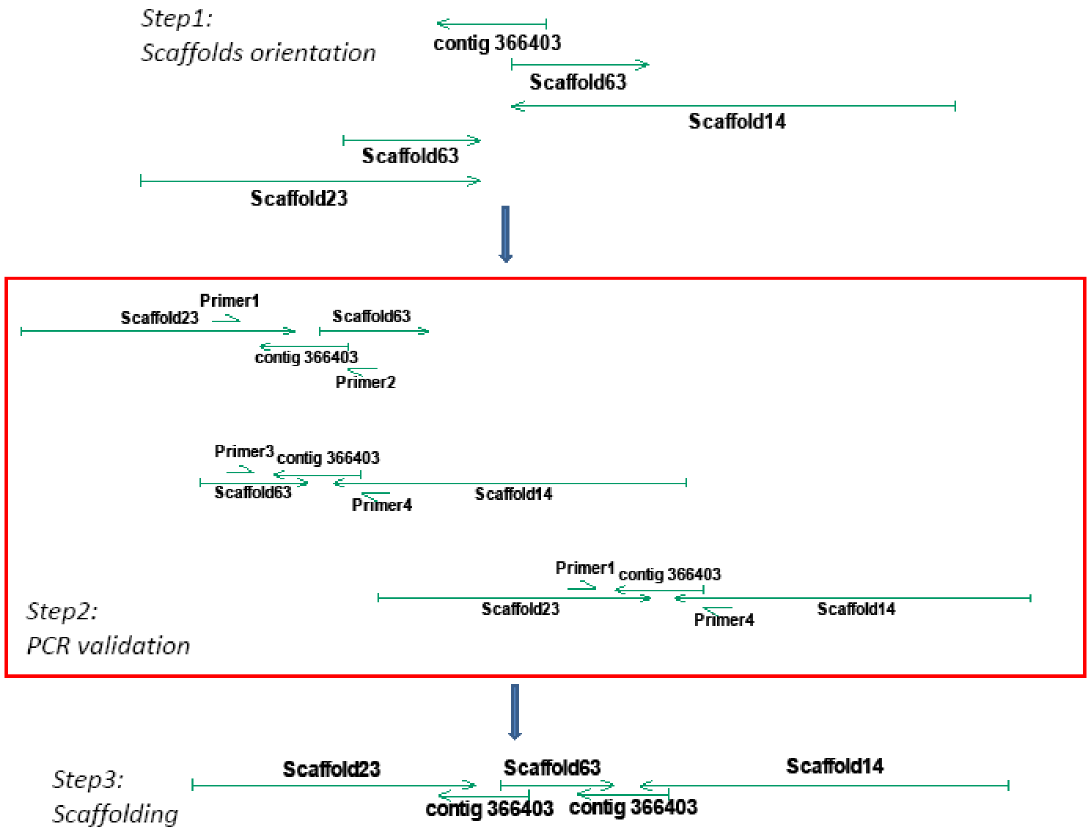
[5] Graduate University of Chinese Academy of Sciences, 19A Yuquanlu, Beijing 100049, China

INDEX

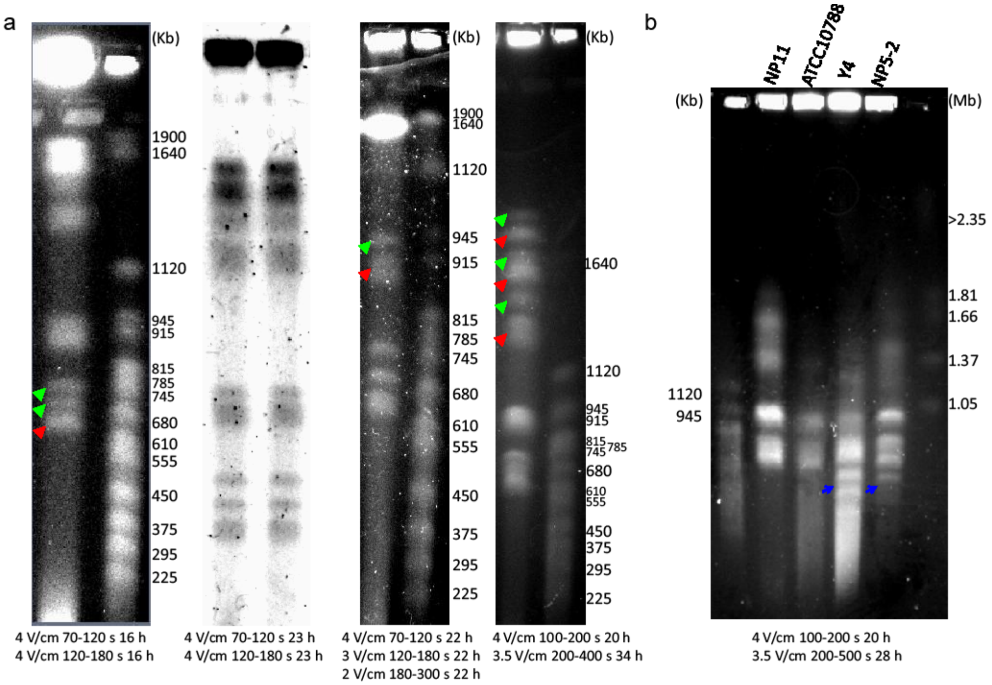
Supplementary Figures



Supplementary Figure S1. Sequencing depth distribution. Raw sequencing reads were trimmed with quality threshold *P* =0.05 (equivalently to Phrep quality score Q≈13) and length sorted to at least 17 bp. Trimmed reads were mapped to the assembled genome with the read aligner Bowtie. The average sequencing depth was 96×.



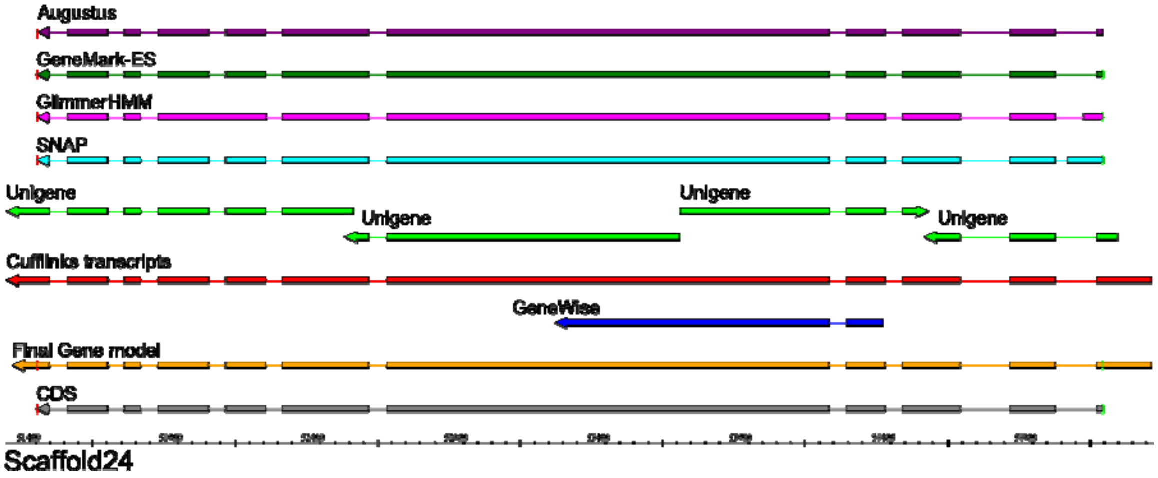
Supplementary Figure S2. Short identical sequences prevent further assembly, but targeted sequencing the PCR fragments solved 15 gaps present in the first version of assembly.



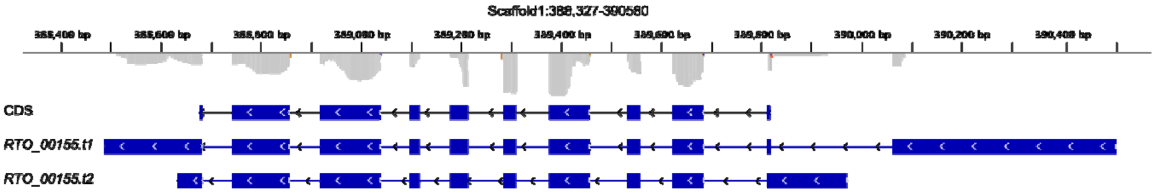
c

Band	Chromosome	Estimated length (Kb)
1	1/2	650
2	3	720
3	4	780
4	5/6	875/915
5	7	945
6	8/9	1370
7	10	1500
8	11/12	1650
9	13	1700
10	14/15	1800
11	16	1900
total	16	20 Mb

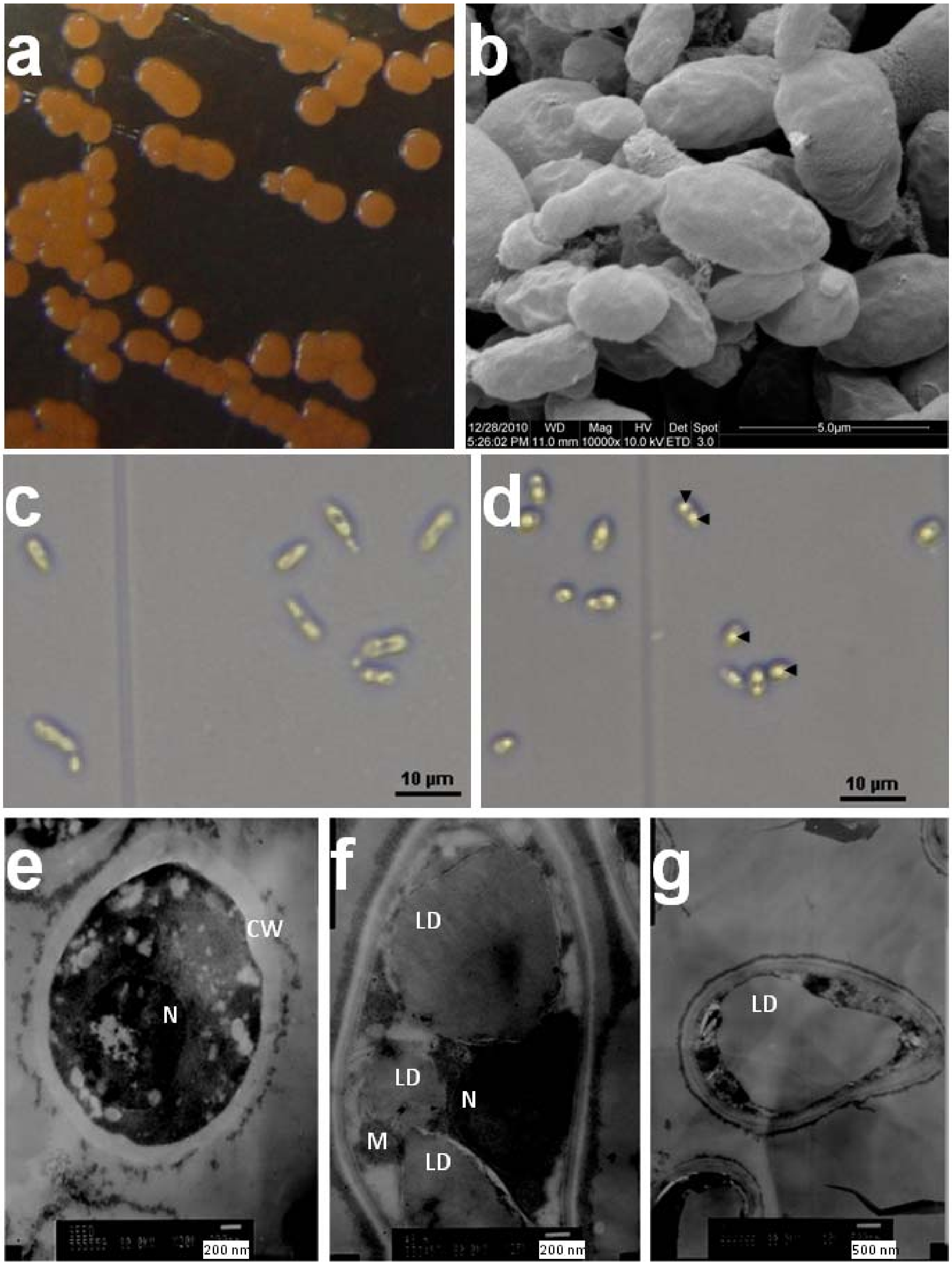
Supplementary Figure S3. Electrophoretic karyotype of *R. toruloides*. (**a**) Pulsed field gel electrophoresis of *R. toruloides* NP11 using contour-clamped homogeneous electric field technology. Total 11 bands (triangle), corresponding to DNA size ranging from 0.65-1.9 Mb, were distinguished with different parameters (voltage, switch time and period as indicated below the figure). And 5 co-migrated bands were detected for their incongruous staining intensity (red triangle). (**b**) Chromosomal DNA patterns of *R. toruloides* NP11 (haploid, *MAT A1*), ATCC10788 (haploid, *MAT A1* and type strain), Y4 (diploid, parent strain) and NP5-2 (haploid, *MAT A2*). The electrophoretic bands of *R. toruloides* NP11 are similar to those of type stain ATCC 10788. Smaller chromosomes (blue arrows) are present in the parental stain Y4 and sister strain NP5-2. (**c**) Estimated Chromosomal length of *R. toruloides* NP11 by PFGE. Two chromosomes were assigned to co-migrated bands, and the total number of chromosomes was estimated to 16. Indeed, 31 canonical telomere repeats (TTAGGG) were present in current assembly and consistent with the estimation.



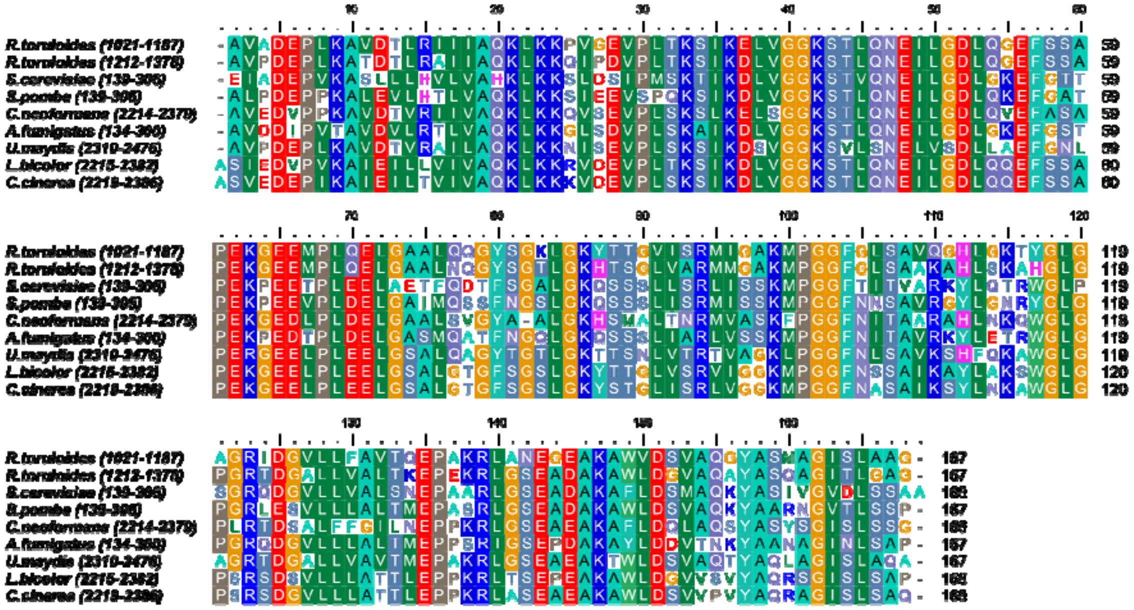
Supplementary Figure S4. Manually merged gene model from *de novo* prediction, homology search and RNA-seq alignment. The presented gene was *RHTO_03906* on the Scaffold24 of the genome assembly (v1).



Supplementary Figure S5. Alternative splicing in *R. toruloides*. Alternative 5’ exons were present in the transcripts of gene *RHTO_00155*.

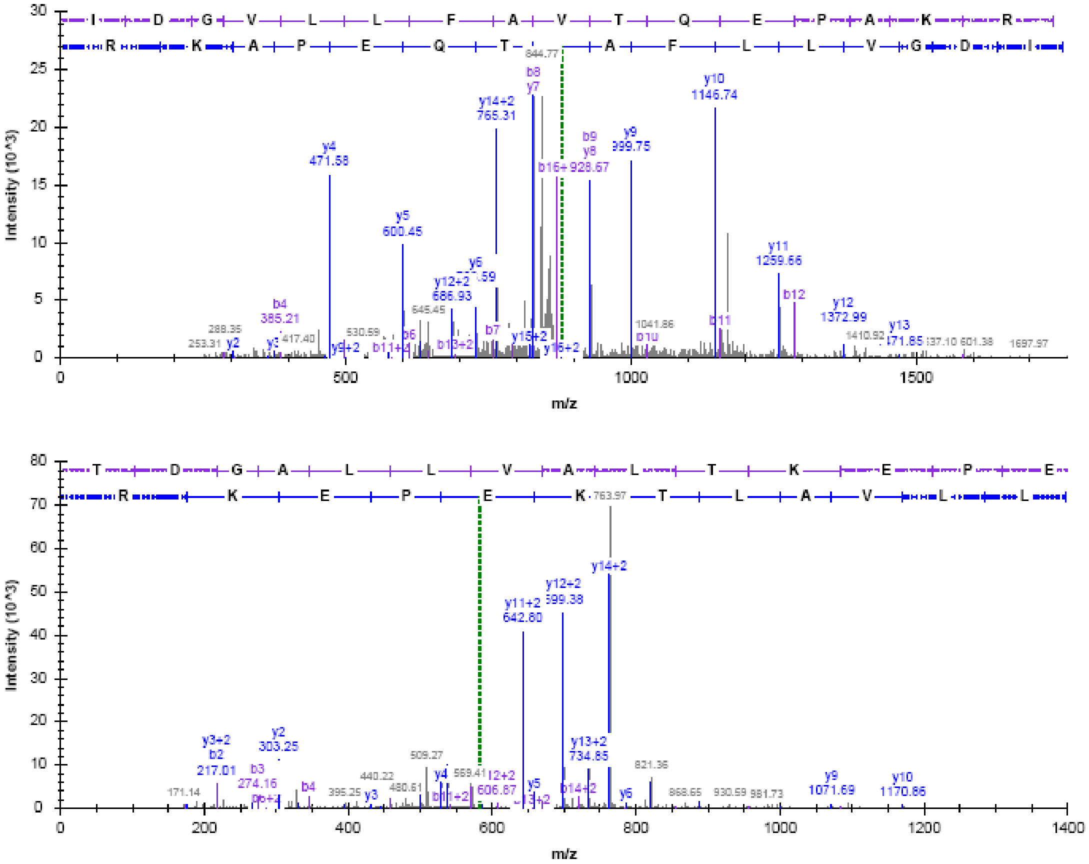


Supplementary Figure S6. Morphology of *R. toruloides*. (**a**) Pink-colored colony. (**b**) Yeast-like growth by budding. Submerged culture in YEPD (**c**) and nitrogen-limited medium (**d**). Ultrastructure of *R. toruloides* in YEPD (**e**) and nitrogen-limited medium (**f** and **g**). N, nucleus; CW, cell wall; M, mitochondrion; LD, lipid droplet; arrows, lipids accumulated in cell.

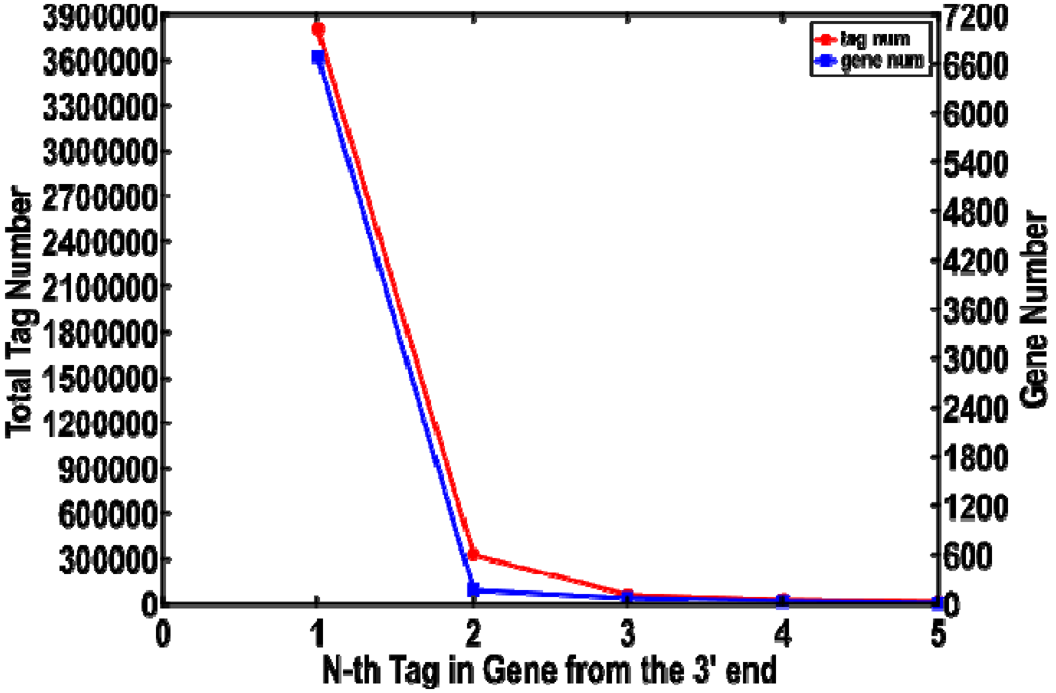


Supplementary Figure S7. Sequence alignment of ACP domains from selected fatty acid synthase (FAS). These include alpha subunit of FAS complex from *R. toruloides* (Fas2), *S. cerevisiae* (Genebank Accession number: P19097.2, the same as below), *S. pombe* (BAA11913.1), *C. neoformans* (XP_571100.1), *Aspergillus fumigatus* (XP_748738.1), and single polypeptide FAS from *U. maydis* (XP_759118.1), *L. bicolor* (XP_001880844.1) and *C. cinerea* (XP_001836417.1). Numbers in parenthesis indicate the boundary of ACP domains. Two ACP domains of *R. toruloides* are highly similar and separated by a spacer with the length of 25 aa.

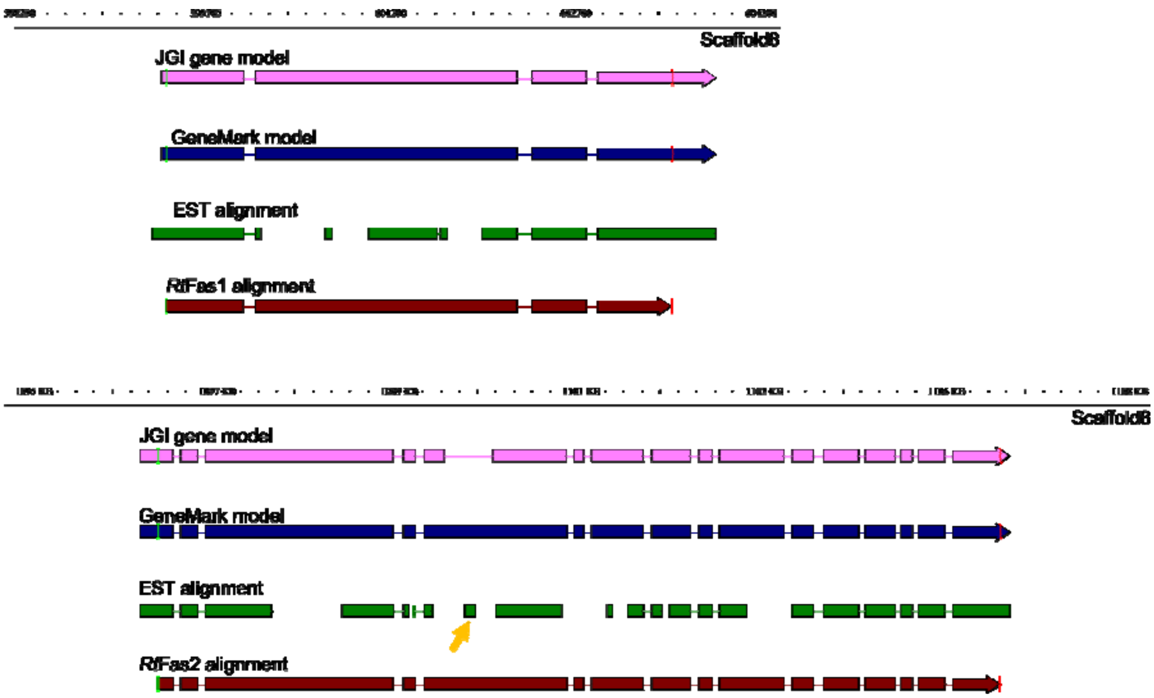
1000-KAAPVAAAPPPPAPVAAAPAAAVADEPLKAVDTLRIIAQKLKKPVGEVPLTKSIKELVGGKSTLQNEILGDLQGEFSSAPEKGEEMPLQELGAALQQGYSGKLGKYTTGVISRMIGAKMPGGFGLSAVQGHLGKTYGLGAGRIDGVLLFAVTQEPAKRLANEGEAKAWVDSVAQGYASMAGISLAAGGGAAAAAPAMAFAAPAAAGGGAPAAVPDEPLKATDTLRAIIAQKLKKQIPDVPLTKSIKDLVGGKSTLQNEILGDLQGEFSSAPEKGEEMPLQELGAALNQGYSGTLGKHTSGLVARMMGAKMPGGFGLSAAKAHLSKAHGLGPGRTDGALLVALTKEPEKRLGSEADAKAWLDGVAQAYASQAGITLGAGGGGGGAAVGGAGFMINTEQLDK-1400



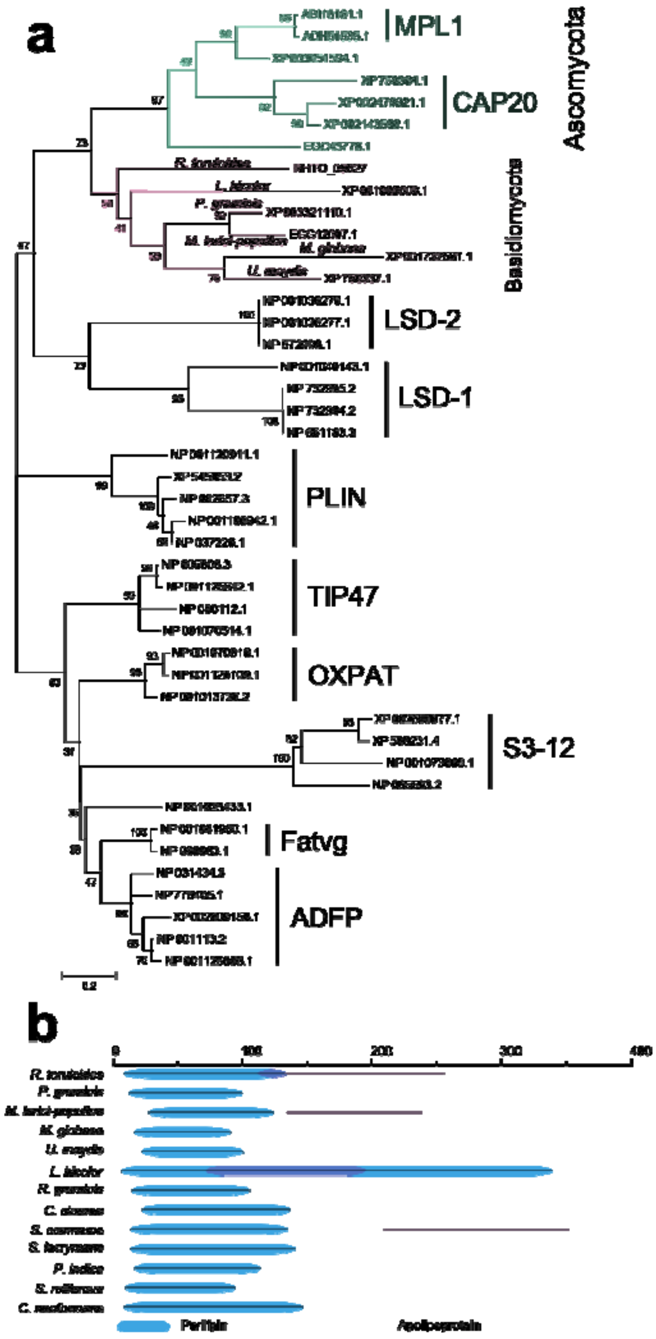
Supplementary Figure S8. Sequence with ACP domains was identified by mass spectrometry. All peptides identified by LC-MS/MS were highlighted in red, and ACP domains were masked in cyan. Annotated MS/MS spectra are that of peptides “IDGVLLFAVTQEPAKR” and “TDGALLVALTKEPEKR” which are underlined. Corresponding amino acids of the identified peptides were informed from labeled b- and y-ions.



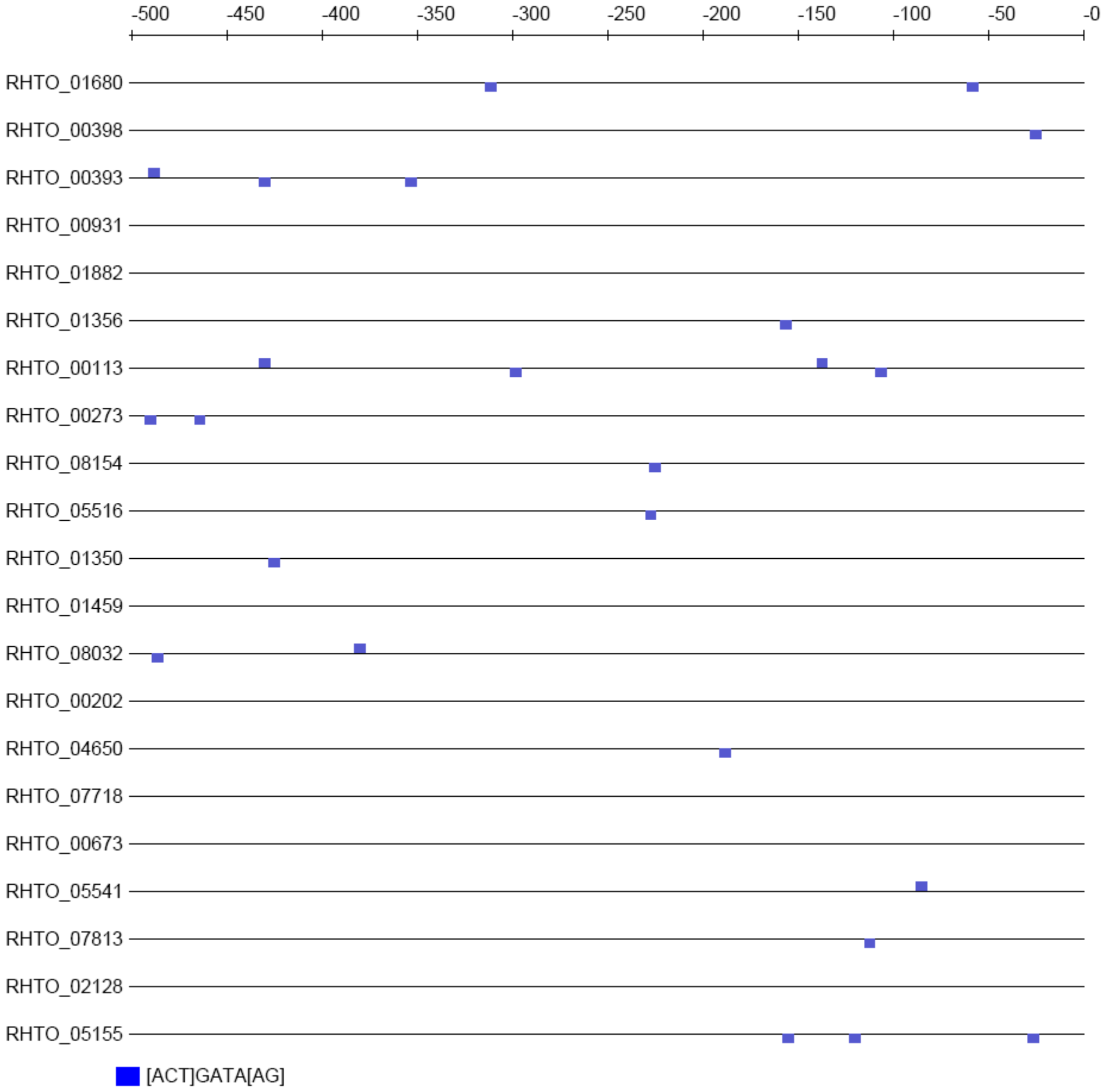
Supplementary Figure S9. Tag position in the transcripts. Most of the mapped tags were aligned to the 3’-most NlaIII restriction sites of annotated transcripts.



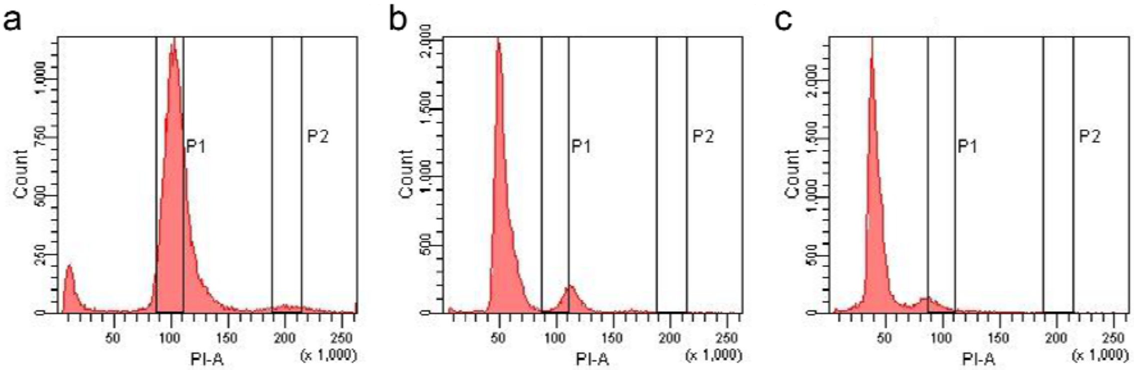
Supplementary Figure S10. Similar FAS in *R. graminis* and *R. toruloides*. Despite the annotated model of *FAS2* in *R. graminis* without two ACPs, model predicted by GeneMark and the EST (yellow arrow) at the position of the second ACP may indicate two ACPs in *R. graminis*. JGI gene model, the "Filtered Model" of JGI; GeneMark model, GeneMark predicted gene model; EST alignment, EST aligned using BLAT; *Rt*Fas1/2 alignment, GeneWise predicted gene model with *Rt*Fas1&2. Data obtained from JGI genome browser (http://genome.jgi-psf.org/Rhoba1_1/Rhoba1_1.home.html). Similar FAS is also found in *S. roseus* and *Rhodotorula glutinis* (data not shown). Noting that sequence data of *R. graminis* were produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) in collaboration with the user community. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.



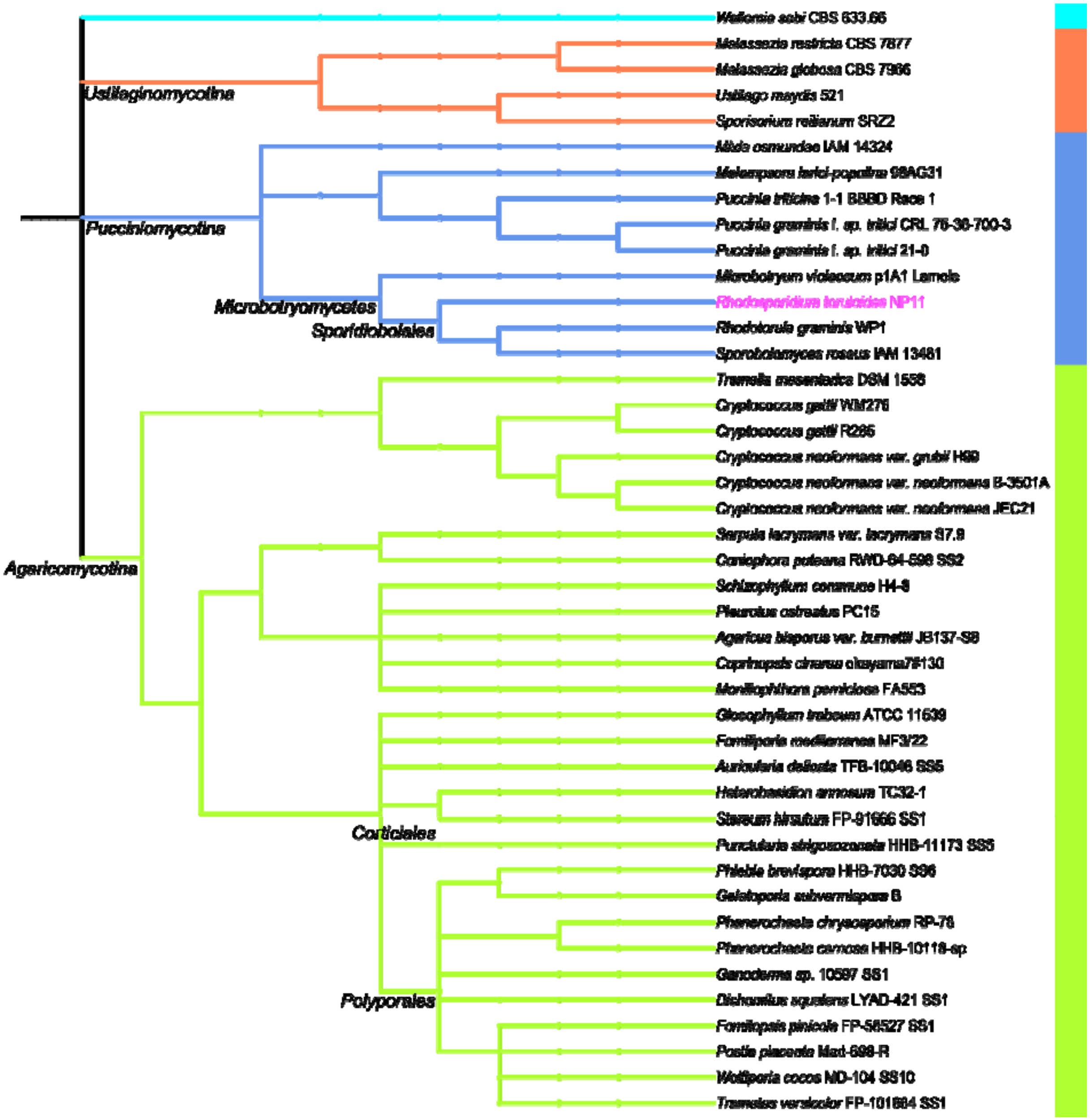
Supplementary Figure S11. Perilipins in Basidiomycetes. (**a**) Neighbor-joining phylogram of perilipin proteins. Perilipin proteins from Basidiomycetes are distinct to CAP20-like proteins from Ascomycetes. Numbers at branches indicate bootstrap support percent with 1000 bootstrap replications. (**b**) Schematic representation of the perilipin (PF03036) and apolipoprotein (PF01442) domains in perilipin family proteins in Basidiomycetes. These include RHTO_05627 (*R. toruloides*), XP_003321110.1 (Genebak accession number, the same below, *Puccinia graminis*), EGG12007.1 (*Melampsora larici-populina*), XP_001732561.1 (*Malassezia globosa*), XP_760337.1 (*Ustilago maydis*), XP_001885608.1 (*Laccaria bicolor*), 64920 (JGI protein ID, *Rhodotorula graminis*), XP_001837169.1 (*Coprinopsis cinerea*), XP_003032950.1 (*Schizophyllum commune*), EGO03996.1 (*Serpula lacrymans*), CCA75851.1 (*Piriformospora indica*), CBQ68792.1 (*Sporisorium reilianum*), XP_568844.1 (*Cryptococcus neoformans*).



Supplementary Figure S12. Putative binding sites of GATA transcription factor in the upstream of genes up-regulated in response to nitrogen limitation (**Supplementary Table S7**). "-0" represent the "A" of the initial codon.



Supplementary Figure S13. Ploidy determination of *R. toruloides* with Flow cytometry. Histogram of propidium iodide (PI) fluorescence of ethanol-fixed cells of diploid *R. toruloides* Y4 (**a**), haploid NP11 (**b**) and NP5-2 (**c**) was shown.



Supplementary Figure S14. Taxonomic tree of selected basidiomycete fungi whose genome sequences are available. *R. toruloides* is the member of the order *Sporidiobolales* and the class of *Microbotryomycetes* in the clade of *Pucciniomycotina*, of which few genomes have been sequenced. The genome size of *R. toruloides* (20.2 Mb) is approximately close to the draft assemblies of closely relative species *Rhodotorula graminis* (21.0 Mb) and *Sporobolomycetes roseus* (21.2 Mb) released by the DOE Joint Genome Institute. Sequencing status was retrieved from NCBI genome project (http://www.ncbi.nlm.nih.gov/bioproject) and JGI fungal Genomics Program (http://genome.jgi-psf.org/programs/fungi/index.jsf). The tree was constructed from taxonomy ID in NCBI (http://www.ncbi.nlm.nih.gov/taxonomy). Branch lengths are not proportional to genetic distances.

Supplementary Tables

Supplementary Table S1. Primary sequencing data obtained from Illumina sequencing.

Genome sequencing[a]				
Insert size	200 bp	200 bp	2 kb	6 kb
Sequencing platform	GAII	GAIIx	GAII	GAIIx
Number of reads pair	9560155	12269060	10293853	13407834
Reads length (nt)	44	100/60[b]	44	44
Total base (Mb)	841.3	1963.0	905.9	1179.9
RNA-seq				
Insert size	200 bp			
Sequencing platform	HiSeq2000			
Number of reads pair	6666667			
Reads length (nt)	90			
Total base (Mb)	1200.0			
3’ tag digital gene expression profiling sequencing				
Sample	MM	MM-N		
Sequencing platform	GAIIx	GAIIx		
Tag length (nt)	35	35		
Number of tags	3531500	3703000		

[a] one lane per genomic library.

[b] 100 nt for forward strand and 60 nt for reverse strand.

Supplementary Table S2. *R. toruloides* genome assembly statistics. N50 and L50 are the number and minimum length, respectively, of all sequences that together account for 50% of the assembly.

	Assembly version I			**Assembly version II**
	Scaffold	Contig	Gap	Scaffold
N50	11	43	-	7
L50 (bp)	574924	163970	-	1297466
Total Number	94	334	240	34
Size (bp)	20225222	20181323	43899	20254837
Max (bp)	1534814	480266	1676	1723588

Supplementary Table S3. Repetitive sequences in *R. toruloides*. 1.66% of the current version of genome assembly consists of repetitive sequences, and most of which are low complexity sequences.

Elements	Total number	Length (bp)	% assembled genome
SINEs	35	4510	0.02
LINEs	46	14585	0.07
LTR elements	175	73022	0.36
DNA elements	21	927	0.00
Unclassified	191	66428	0.33
Total interspersed repeats		159472	0.79
Simple repeats	811	36918	0.18
Low complexity	1699	140162	0.69
Total repetitive sequence		336002	1.66

Supplementary Table S4. Genome feature of *R. toruloides*.

GC content of genome	62.03%
GC content of CDS	63.12%
Coding genes	8171
Incomplete Genes	95
Single exon genes	226
Mead Gene density	384 per Mb
Median Transcript length	1882 bp
Median Coding length	1337 bp
Median Coding exons per gene	5
Median Exon length*	146 bp
Median Intron length	62 bp
Genes with 5' UTR	6408
Median Length of 5' UTR	89 bp
Genes with 3' UTR	6500
Median Length of 3' UTR	92 bp
tRNA genes	117
Spliceosomal RNA	6

*** 168, 134 and 199 bp for initial, internal and terminal exons, respectively**

Supplementary Table S5. Functional annotation of genes in *R. toruloides*

Total Genes	**8171**	**100.0%**
Homology	6454	79.0%
NR alignment[a]	6120	74.9%
Trembl alignment[a]	6171	75.5%
SwissProt alignment[a]	5096	62.4%
KEGG alignment[a]	6109	74.8%
InterPro[a]	5818	71.2%
GO annotation	4910	60.1%
KOG[b]	4403	53.9%
Expression	7759	95.0%
RNA-seq[c]	7357	90.0%
3' tag-seq[d]	7045	86.2%
Proteomics[e]	3108	38.0%

a. with a cutoff of 10E-5

b. with a cutoff of 10E-10

c. >30% length coverage and at least 3 reads

d. with at least two tags

e. with at least one peptide (FDR<1%)

Supplementary Table S6. Tag statistics for 3’-tag digital gene expression profiling sequencing of *R. toruloides*.

			MM		**MM-N**		**Total**	
Tag statistics								
Raw data		**Distinct tag**	100146		127914		181443	
		Total tag	3531500		3703000		7234500	
Clean tags[a]		**Distinct tag**	92643		119191		168705	
		Total tag	3490289		3659813		7150102	
Sampled tags[b]		**Distinct tag**	56846		60010		73727	
		Total tag	3454492		3600632		7055124	
Copy number statistics								
Copy_Num >0		**Distinct tag**	56846	100.00%	60010	100.00%	-	-
		Total tag	3454492	100.00%	3600632	100.00%	-	-
Copy_Num >1		**Distinct tag**	46312	81.47%	48636	81.05%	73727	100.00%
		Total tag	3443958	99.70%	3589258	99.68%	7055124	100.00%
Copy_Num >5		**Distinct tag**	18929	33.30%	20871	34.78%	31459	42.67%
		Total tag	3367163	97.47%	3510318	97.49%	6935633	98.31%
Copy_Num >10		**Distinct tag**	12613	22.19%	14205	23.67%	21131	28.66%
		Total tag	3319315	96.09%	3460123	96.10%	6857248	97.20%
Copy_Num >20		**Distinct tag**	8646	15.21%	9814	16.35%	14172	19.22%
		Total tag	3261145	94.40%	3395696	94.31%	6756093	95.76%
Copy_Num >50		**Distinct tag**	5167	9.09%	5955	9.92%	8715	11.82%
		Total tag	3149843	91.18%	3269879	90.81%	6581129	93.28%
Copy_Num >100		**Distinct tag**	3378	5.94%	3843	6.40%	5887	7.98%
		Total tag	3021558	87.47%	3118755	86.62%	6378458	90.41%
Mapping statistics[c]								
Map to transcripts	**Sense (Unique)**	**Distinct tag**	24750	43.54%	24497	40.82%	28616	38.81%
		Total tag	2314492	67.00%	2000070	55.55%	4314562	61.16%
		Gene	6887	84.29%	6916	84.64%	7045	86.22%
	Sense (Multiple)	**Distinct tag**	380	0.67%	401	0.67%	440	0.60%
		Total tag	137938	3.99%	117168	3.25%	254106	3.60%
		Gene	165	2.02%	163	1.99%	184	2.25%
	Antisense (Unique)	**Distinct tag**	13544	23.83%	12894	21.49%	15648	21.22%
		Total tag	238570	6.91%	163347	4.54%	401917	5.70%
		Gene	5746	70.32%	5595	68.47%	6069	74.27%
	Antisense (Multiple)	**Distinct tag**	101	0.18%	95	0.16%	116	0.16%
		Total tag	2634	0.08%	1616	0.05%	4299	0.06%
		Gene	141	1.73%	134	1.64%	153	1.87%
	Total	**Distinct tag**	38775	68.21%	37887	63.13%	44820	60.79%
		Total tag	2693634	77.97%	2282250	63.38%	4975884	70.53%
		Gene	7121	87.15%	7135	87.32%	7251	88.74%
Map to genome	**Total**	**Distinct tag**	43670	76.82%	42911	71.51%	51387	69.70%
		Total tag	3084829	89.30%	2664186	73.99%	5890651	83.49%
Unmaped tag	**Total**	**Distinct tag**	13176	23.18%	17099	28.49%	22340	30.30%
		Total tag	369663	10.70%	936446	26.01%	1164473	16.51%

a. Removing adaptor tags and tags of low quality.

b. Only tags with occurrence ≥ 2 in total.

c. No more than 1 bp mismatch.

Supplementary Table S7. Genes related to nitrogen metabolite repression and their expression during N starvation.

Gene symbol	ORF	Tag counts		TPM (tag per million)		Ratio MM-N/MM	*P*-value*	Annotation
		MM	MM-N	MM	MM-N			
MEP2	RHTO_01680	63	663	18.24	184.13	10.1	0	ammonium transmembrane transporter
MEP1	RHTO_01846	117	87	33.87	24.16	0.7	1.7E-02	ammonium transmembrane transporter
	RHTO_00398	62	149	17.95	41.38	2.3	8.2E-09	amino acid transmembrane transporter
	RHTO_00393	0	3	0.01	0.83	83	1.4E-01	amino acid transmembrane transporter
	RHTO_00931	42	398	12.16	110.54	9.1	9.7E-14	amino acid transmembrane transporter
	RHTO_01882	421	1345	121.87	373.55	3.1	4.1E-13	amino acid transporter
	RHTO_03593	1151	1551	333.19	430.76	1.3	3.7E-11	amino acid transmembrane transporter
	RHTO_07825	168	52	48.63	14.44	0.3	1.0E-16	dicarboxylic amino acid permease
	RHTO_07915	1050	396	303.95	109.98	0.4	1.1E-74	amino acid transmembrane transporter
	RHTO_01356	177	1704	51.24	473.25	9.2	0	oligopeptide transporter
DUR3	RHTO_00113	18	36	5.21	10	1.9	2.2E-02	urea transporter
URO1	RHTO_03331	161	167	46.61	46.48	1.0	9.6E-01	urate oxidase
	RHTO_00273	27	137	7.82	38.05	4.9	2.1E-11	Transthyretin/hydroxyisourate hydrolase family protein
	RHTO_02662	681	789	197.13	219.13	1.1	4.3E-02	OHCU decarboxylase family protein
DAL2/3	RHTO_08154	39	191	11.29	53.05	4.7	9.7E-14	allantoicase
URE1	RHTO_05516	96	198	27.79	54.99	2.0	1.6E-08	urease
DUR1,2	RHTO_01350	0	6	0.01	1.67	167.0	1.8E-02	urea carboxylase / allophanate hydrolase
DAL1	RHTO_00639	-	-	-	-	-	-	allantoinase
DAL7/MLS1	RHTO_01459	73	515	21.13	143.03	6.8	0.0E+00	malate synthase
PUT1	RHTO_08032	39	63	11.29	17.5	1.6	3.0E-02	proline dehydrogenase
PUT2	RHTO_07378	805	61	233.03	16.94	0.1	1.4E-173	delta-1-pyrroline-5-carboxylate dehydrogenase
NR	RHTO_02415	27	20	7.82	5.55	0.7	2.5E-01	nitrate reductase (NADH)
NIR	RHTO_02416	16	9	4.63	2.5	0.5	1.4E-01	nitrite reductase
NRT2	RHTO_02417	-	-	-	-	-	-	MFS nitrate transporter
NRT1	RHTO_04837	-	-	-	-	-	-	nitrate transporter
CAR1	RHTO_05160	1026	358	297	99.43	0.3	3.4E-81	arginase
UGA4	RHTO_05974	64	91	18.53	25.27	1.4	5.6E-02	GABA permease
UGA1	RHTO_00202	255	1492	73.82	414.37	5.6	0	4-aminobutyrate transaminase
GDH1	RHTO_04650	701	2437	202.92	676.83	3.3	1.4E-13	glutamate dehydrogenase (NADP+)
GDH2	RHTO_07718	88	196	25.47	54.43	2.1	8.7E-10	glutamate dehydrogenase (NAD+)
GLN1	RHTO_00673	2268	5681	656.54	1577.78	2.4	0	glutamine synthetase
GLN2	RHTO_00401	107	117	30.97	32.49	1.05	7.2E-01	glutamine synthetase
GLT1	RHTO_00025	362	234	104.79	64.99	0.62	8.3E-09	GOGAT, glutamate synthase
NPR1	RHTO_05632	29	48	8.39	13.33	1.6	8.1E-02	protein serine/threonine kinase
NPR2	RHTO_03203	77	38	22.29	10.55	0.5	2.9E-04	nitrogen permease regulator Npr2
NPR3	RHTO_03073	22	24	6.37	6.67	1.0	9.2E-01	Nitrogen permease regulator 3
NPI1	RHTO_02946	589	299	170.50	83.04	0.5	1.9E-24	E3 ubiquitin-protein ligase NEDD4
PRB1	RHTO_05541	148	325	42.84	90.26	2.1	0.0E+00	autophagic serine protease Alp2
LAP1	RHTO_00296	48	38	13.89	10.55	0.8	2.9E-01	aminopeptidase 2
PEP4	RHTO_07813	202	920	58.47	255.51	4.4	0	saccharopepsin
CPS1	RHTO_03874	81	25	23.45	6.96	0.3	3.5E-08	Gly-Xaa carboxypeptidase
	RHTO_02128	40	91	11.58	25.27	2.2	2.1E-05	amino acid transporter
	RHTO_03777	28	43	8.11	11.94	1.5	1.7E-01	vacuolar amino acid transporter 4
	RHTO_05155	2	18	0.58	5	8.6	1.1E-03	vacuolar amino acid transporter 1
ATG1	RHTO_06351	57	451	16.50	125.26	7.6	3.6E-14	Serine/threonine protein kinase ATG1
ATG2	RHTO_01885	20	59	5.79	16.39	2.8	6.0E-05	autophagy regulatory protein Atg2
ATG20	RHTO_04181	19	70	5.50	19.44	3.5	3.6E-07	sorting nexin-41
ATG8	RHTO_06526	181	402	52.40	111.65	2.1	6.6E-13	GABA(A) receptor-associated protein
VMA5	RHTO_04163	23	220	6.66	61.10	9.2	1.3E-09	V-type H+-transporting ATPase subunit C
VMA9	RHTO_07443	180	500	52.11	138.86	2.7	1.6E-12	ATPase, V0 complex, subunit E
VMA11	RHTO_01541	28	109	8.11	30.27	3.7	8.4E-11	V-type H+-transporting ATPase 21kDa proteolipid subunit
AMS1	RHTO_00613	28	83	8.11	23.05	2.8	1.3E-06	alpha-mannosidase, glycoside hydrolase family 38 protein
ATG15	RHTO_00361	25	57	7.24	15.83	2.2	7.6E-04	triacylglycerol lipase, homolog of ATG15
AMD1	RHTO_00810	98	152	28.37	42.21	1.5	2.0E-03	AMP deaminase

* Audic&Claverie's test (ref 58).

Supplementary Table S8. Genes related to central and lipid metabolisms and their expression during N starvation.

Gene symbol	ORF	Tag counts		TPM (tag per million)		Ratio MM-N/MM	*P*-value*	Annotation
		MM	MM-N	MM	MM-N			
HXK1	RHTO_06072	270	515	78.16	143.03	1.8	3.8E-13	hexokinase
HXK2	RHTO_06870	8	12	2.32	3.33	1.4	4.4E-01	hexokinase
GPI1	RHTO_04058	31	22	8.97	6.11	0.7	1.7E-01	glucose-6-phosphate isomerase
PFK1	RHTO_00494	46	61	13.32	16.94	1.3	2.2E-01	6-phosphofructokinase
FBP1	RHTO_03046	1149	1526	332.61	423.81	1.3	4.7E-10	fructose-1,6-bisphosphatase I
FBA1	RHTO_03043	2064	7776	597.48	2159.62	3.6	0.0E+00	fructose-bisphosphate aldolase, class II
TPI1	RHTO_01329	680	2020	196.85	561.01	2.8	2.5E-13	triosephosphate isomerase (TIM)
GAPDH	RHTO_03746	10729	7128	3105.81	1979.65	0.6	4.0E-195	glyceraldehyde 3-phosphate dehydrogenase
PGK1	RHTO_00033	787	1381	227.82	383.54	1.7	1.0E-12	phosphoglycerate kinase
GPM1	RHTO_03049	92	62	26.63	17.22	0.6	7.5E-03	phosphoglycerate mutase
GPM2	RHTO_04793	45	39	13.03	10.83	0.8	4.0E-01	phosphoglycerate mutase
GPM3	RHTO_05036	312	310	90.32	86.1	1.0	5.5E-01	phosphoglycerate mutase
GPM4	RHTO_07773	-	-	-	-	-	-	protein of phosphoglycerate mutase 1 family
ENO1	RHTO_00323	3848	4669	1113.91	1296.72	1.2	6.3E-12	enolase
PYK1	RHTO_01610	218	274	63.11	76.1	1.2	3.9E-02	pyruvate kinase
PDA1	RHTO_01852	160	96	46.32	26.66	0.6	1.4E-05	pyruvate dehydrogenase E1 component subunit alpha
PDB1	RHTO_07250	282	533	81.63	148.03	1.8	5.9E-13	pyruvate dehydrogenase E1 component subunit beta
LAT1	RHTO_03543	3272	3195	947.17	887.34	0.9	8.7E-03	pyruvate dehydrogenase E2 component
CIT1	RHTO_06406	122	90	35.32	25	0.7	1.2E-02	citrate synthase, mitochondrial
CIT2	RHTO_07345	128	120	37.05	33.33	0.9	4.0E-01	citrate synthase
ACO1	RHTO_00539	76	85	22	23.61	1.1	6.6E-01	aconitate hydratase
ACO2	RHTO_08030	88	152	25.47	42.21	1.7	1.3E-04	aconitate hydratase
IDH1	RHTO_01289	2566	3704	742.8	1028.71	1.4	0.0E+00	isocitrate dehydrogenase (NAD+)
IDH2	RHTO_01290	3792	2293	1097.7	636.83	0.6	3.4E-97	isocitrate dehydrogenase (NAD+)
IDP1	RHTO_04315	326	1649	94.37	457.98	4.9	5.6E-13	isocitrate dehydrogenase (NADP+), mitochondrial precursor
KGD1	RHTO_07860	1023	1261	296.14	350.22	1.2	6.5E-05	2-oxoglutarate dehydrogenase E1 component (succinyl-transferring)
KGD2	RHTO_02312	64	57	18.53	15.83	0.9	3.9E-01	2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase)
LPD1	RHTO_07893	3274	2664	947.75	739.87	0.8	1.8E-21	dihydrolipoyl dehydrogenase
LSC1	RHTO_01205	24	32	6.95	8.89	1.3	3.7E-01	succinyl-CoA synthetase alpha subunit
LSC2	RHTO_05264	121	135	35.03	37.49	1.1	5.9E-01	succinyl-CoA synthetase beta subunit
SDH1	RHTO_05714	443	324	128.24	89.98	0.7	1.1E-06	succinate dehydrogenase (ubiquinone) flavoprotein subunit
SDH2	RHTO_06068	1058	930	306.27	258.29	0.8	1.5E-04	succinate dehydrogenase (ubiquinone) iron-sulfur protein
SDH3	RHTO_00534	2566	206	742.8	57.21	0.1	0.0E+00	succinate dehydrogenase (ubiquinone) cytochrome b subunit
SDH4	RHTO_00723	2008	2890	581.27	802.64	1.4	0.0E+00	succinate dehydrogenase (ubiquinone) membrane anchor subunit
FUM1	RHTO_05746	378	1726	109.42	479.36	4.4	5.7E-13	fumarate hydratase, class II
MDH1	RHTO_04363	6953	5634	2012.74	1564.73	0.8	4.4E-45	L-malate dehydrogenase, mitochondrial
MDH2	RHTO_03745	1565	1301	453.03	361.33	0.8	1.5E-09	L-malate dehydrogenase
CTP1	RHTO_03384	4	8	1.16	2.22	1.9	3.0E-01	mitochondrial tricarboxylate transporter, citrate transporter
ACL1	RHTO_03915	20829	19480	6029.54	5410.16	0.9	1.5E-27	ATP citrate synthase
MAE1	RHTO_07218	165	204	47.76	56.66	1.2	1.0E-01	malic enzyme, mitochondrial
ME1	RHTO_03795	249	58	72.08	16.11	0.2	2.6E-31	malic enzyme
PYC1	RHTO_02628	473	2903	136.92	806.25	5.9	3.3E-13	pyruvate carboxylase
PDC1	RHTO_00098	193	285	55.87	79.15	1.4	1.7E-04	pyruvate decarboxylase
ALD1	RHTO_05838	364	486	105.37	134.98	1.3	3.4E-04	aldehyde dehydrogenase (NAD)
ALD2	RHTO_04425	42	53	12.16	14.72	1.2	3.6E-01	aldehyde dehydrogenase, mitochondrial
ALD3	RHTO_04310	12	21	3.47	5.83	1.7	1.5E-01	aldehyde dehydrogenase
ALD4	RHTO_06724	-	-	-	-	-	-	aldehyde dehydrogenase
ALD5	RHTO_04543	4	6	1.16	1.67	1.4	6.0E-01	aldehyde dehydrogenase
ACS1	RHTO_08027	227	329	65.71	91.37	1.4	1.2E-04	acetyl-CoA synthetase
ACH1	RHTO_07106	114	148	33	41.1	1.2	7.8E-02	acetyl-CoA hydrolase
ACC1	RHTO_02004	125	120	36.18	33.33	0.9	5.2E-01	acetyl-CoA carboxylase
FAS1	RHTO_02032	955	4697	276.45	1304.49	4.7	1.2E-12	fatty acid synthase subunit beta, fungi type
FAS2	RHTO_02139	1641	3123	475.03	867.35	1.8	1.1E-12	fatty acid synthase subunit alpha, fungi type
ERG10	RHTO_02048	88	1433	25.47	397.99	15.6	6.7E-16	acetyl-CoA C-acetyltransferase
ERG13	RHTO_02305	792	762	229.27	211.63	0.9	1.1E-01	hydroxymethylglutaryl-CoA synthase
HMG1	RHTO_04045	31	21	8.97	5.83	0.6	1.3E-01	hydroxymethylglutaryl-CoA reductase (NADPH)
ERG12?	RHTO_02122	73	398	21.13	110.54	5.2	0.0E+00	cystathionine beta-lyase
ERG8	RHTO_02073	65	26	18.82	7.22	0.4	1.5E-05	phosphomevalonate kinase
MVD1	RHTO_06005	69	58	19.97	16.11	0.8	2.3E-01	diphosphomevalonate decarboxylase
IDI1	RHTO_05138	205	280	59.34	77.76	1.3	3.2E-03	isopentenyl-diphosphate delta-isomerase
ERG20	RHTO_01660	108	234	31.26	64.99	2.1	7.9E-11	farnesyl diphosphate synthase
BTS1	RHTO_02504	97	212	28.08	58.88	2.1	4.2E-10	geranylgeranyl diphosphate synthase, type III
ERG9	RHTO_02590	165	136	47.76	37.77	0.8	4.2E-02	farnesyl-diphosphate farnesyltransferase
PSY1	RHTO_03752	45	44	13.03	12.22	0.9	7.6E-01	protein of squalene/phytoene synthase family
CRTI	RHTO_04602	2905	1632	840.93	453.25	0.5	1.3E-92	Phytoene dehydrogenase
ICL1	RHTO_03507	148	241	42.84	66.93	1.6	1.5E-05	isocitrate lyase, mitochondrial
ICL2	RHTO_05768	96	408	27.79	113.31	4.1	0.0E+00	isocitrate lyase
MLS1/DAL7	RHTO_01459	73	515	21.13	143.03	6.8	0.0E+00	malate synthase
ZWF1	RHTO_07853	731	1433	211.61	397.99	1.9	0.0E+00	glucose-6-phosphate 1-dehydrogenase
SOL3	RHTO_07939	3518	3131	1018.38	869.57	0.9	1.2E-10	6-phosphogluconolactonase
GND1	RHTO_02788	6120	935	1771.61	259.68	0.1	0.0E+00	6-phosphogluconate dehydrogenase
RPE1	RHTO_05984	117	170	33.87	47.21	1.4	5.4E-03	ribulose-phosphate 3-epimerase
RKI1	RHTO_06311	24	3	6.95	0.83	0.1	1.7E-05	ribose 5-phosphate isomerase A
TAL1	RHTO_06955	3210	7771	929.22	2158.23	2.3	0.0E+00	transaldolase
TKL1	RHTO_03248	3825	2593	1107.25	720.15	0.7	2.1E-65	transketolase
PGM1	RHTO_07820	141	84	40.82	23.33	0.6	3.8E-05	phosphoglucomutase
PGM2	RHTO_01766	6	21	1.74	5.83	3.4	5.2E-03	phosphoglucomutase
PCK1	RHTO_07480	12	22	3.47	6.11	1.8	1.2E-01	phosphoenolpyruvate carboxykinase (ATP)
OAC1	RHTO_00349	202	221	58.47	61.38	1.0	6.2E-01	oxaloacetate carrier
ODC1	RHTO_03731	50	359	14.47	99.7	6.9	2.2E-14	mitochondrial 2-oxodicarboxylate carrier
GDH1	RHTO_04650	701	2437	202.92	676.83	3.3	1.4E-13	glutamate dehydrogenase (NADP+)
GDH2	RHTO_07718	88	196	25.47	54.43	2.1	8.7E-10	glutamate dehydrogenase (NAD+)
GLN1	RHTO_00673	2268	5681	656.54	1577.78	2.4	0.0E+00	glutamine synthetase
GLT2	RHTO_00025	362	234	104.79	64.99	0.6	8.3E-09	GOGAT, glutamate synthase
GLN2	RHTO_00401	107	117	30.97	32.49	1.0	7.2E-01	glutamine synthetase
ACR1	RHTO_05007	60	333	17.37	92.48	5.3	0.0E+00	mitochondrial carrier protein, succinate:fumarate antiporter
GPD1	RHTO_08013	16	12	4.63	3.33	0.7	3.9E-01	glycerol-3-phosphate dehydrogenase (NAD+) (partial sequence c terminus)
GPD2	RHTO_02273	66	115	19.11	31.94	1.7	7.4E-04	glycerol-3-phosphate dehydrogenase (NAD+)
GNPAT1	RHTO_06391	64	28	18.53	7.78	0.4	7.0E-05	glyceronephosphate O-acyltransferase
GAT1	RHTO_03058	422	446	122.16	123.87	1.0	8.4E-01	glycerol-3-phosphate O-acyltransferase / dihydroxyacetone phosphate acyltransferase
AYR1	RHTO_03193	24	44	6.95	12.22	1.8	2.4E-02	1-acylglycerone phosphate reductase
SLC1	RHTO_06718	181	105	52.4	29.16	0.6	1.2E-06	1-acylglycerol-3-phosphate acyltransferase
PAH1	RHTO_04894	48	93	13.89	25.83	1.9	3.7E-04	lipin 1
CDS1	RHTO_01718	289	169	83.66	46.94	0.6	1.2E-09	phosphatidate cytidylyltransferase
CDK1	RHTO_06970	100	131	28.95	36.38	1.3	8.5E-02	diacylglycerol kinase
DGA1	RHTO_01962	5	1	1.45	0.28	0.2	1.1E-01	2-acylglycerol O-acyltransferase 2
LRO1	RHTO_01945	71	103	20.55	28.61	1.4	3.1E-02	phospholipid:diacylglycerol acyltransferase
ATG15	RHTO_00361	25	57	7.24	15.83	2.2	7.6E-04	triacylglycerol lipase, homolog of ATG15

TGL1	RHTO_00993	6	5	1.74	1.39	0.8	7.2E-01	triacylglycerol lipase
TGL2	RHTO_06751	48	43	13.89	11.94	0.9	4.7E-01	lipase 2
ARE1	RHTO_00726	8	5	2.32	1.39	0.6	3.8E-01	sterol O-acyltransferase
YEH1	RHTO_08052	5	1	1.45	0.28	0.2	1.1E-01	lysosomal acid lipase/cholesteryl ester hydrolase
YEH2	RHTO_01062	20	43	5.79	11.94	2.1	6.2E-03	ab-hydrolase associated lipase
FAA1	RHTO_00058	50	73	14.47	20.27	1.4	6.6E-02	long-chain acyl-CoA synthetase
FAA2	RHTO_04350	10	99	2.89	27.5	9.5	1.3E-06	long-chain acyl-CoA synthetase
FAA3	RHTO_04957	26	85	7.53	23.61	3.1	3.8E-08	long-chain acyl-CoA synthetase
FAA4	RHTO_04941	-	-	-	-	-	-	long-chain acyl-CoA synthetase, putative
FAA5	RHTO_00057	3	8	0.87	2.22	2.6	1.7E-01	long-chain acyl-CoA synthetase, putative
FAA6	RHTO_01192	2	0	0.58	0.01	0.0	2.3E-01	long-chain acyl-CoA synthetase
FAT1	RHTO_07775	0	4	0.01	1.11	111.0	6.9E-02	bifunctional fatty acid transporter/acyl-CoA synthetase
AOX1	RHTO_03890	4	29	1.16	8.05	6.9	5.3E-05	acyl-CoA oxidase
AOX2	RHTO_03776	64	154	18.53	42.77	2.3	4.4E-09	acyl-CoA oxidase
AOX3	RHTO_03787	1	3	0.29	0.83	2.9	4.0E-01	acyl-CoA oxidase
ACD1	RHTO_04971	20	296	5.79	82.21	14.2	2.1E-09	medium-chain-acyl-CoA dehydrogenase, mitochondrial
ACD2	RHTO_06738	106	270	30.68	74.99	2.4	0.0E+00	acyl-CoA dehydrogenase
ACD3	RHTO_05797	316	66	91.48	18.33	0.2	9.5E-43	acyl-CoA dehydrogenase
ACD4	RHTO_03594	142	448	41.11	124.42	3.0	2.7E-13	short/branched chain acyl-CoA dehydrogenase, mitochondrial
ACD5	RHTO_01625	7	86	2.03	23.88	11.8	8.0E-06	acyl-CoA dehydrogenase
ACD6	RHTO_00397	201	356	58.19	98.87	1.7	9.5E-10	acyl-CoA dehydrogenase
ACD7	RHTO_05970	215	161	62.24	44.71	0.7	1.4E-03	acyl-CoA dehydrogenase
ACD8	RHTO_04579	17	67	4.92	18.61	3.8	7.7E-08	acyl-CoA dehydrogenase
FOX2	RHTO_06581	20	132	5.79	36.66	6.3	2.1E-09	multifunctional beta-oxidation protein
ECH1	RHTO_05407	45	78	13.03	21.66	1.7	6.0E-03	enoyl-CoA hydratase, mitochondrial
ECH2	RHTO_02517	1	12	0.29	3.33	11.5	2.9E-03	enoyl-CoA hydratase
ECH3	RHTO_04298	51	42	14.76	11.66	0.8	2.6E-01	enoyl-CoA hydratase/isomerase family protein
HCD1	RHTO_05520	26	36	7.53	10	1.3	2.7E-01	3-hydroxyacyl-CoA dehyrogenase, mitochondrial
POT1	RHTO_02848	142	394	41.11	109.43	2.7	2.7E-13	3-ketoacyl-CoA thiolase (POT1)
POT2	RHTO_07118	72	208	20.84	57.77	2.8	0.0E+00	3-ketoacyl-coA thiolase peroxisomal A precursor
POT3	RHTO_00300	115	177	33.29	49.16	1.5	1.0E-03	peroxisomal 3-ketoacyl-CoA thiolase
ACA1	RHTO_00476	320	327	92.63	90.82	1.0	8.0E-01	3-ketoacyl-CoA thiolase, mitochondrial
DIC1	RHTO_06493	403	119	116.66	33.05	0.3	6.6E-40	mitochondrial dicarboxylate carrier

* Audic&Claverie's test (ref 58).

Supplementary Table S9. Abbreviations in Figure 1.

Abbreviations		Abbreviations	
Glc	glucose	DMAPP	dimethylallyl diphosphate
G-6-P	glucose-6-phosphate	GPP	geranyl pyrophosphate
F-6-P	fructose-6-phosphate	FPP	farnesyl pyrophosphate
G-1-P	glucose-1-phosphate	GGPP	geranylgeranyl pyrophosphate
F-1,6-BP	fructose-1,6-bisphosphate	AT	acetyltransferase
G-3-P	glyceraldehyde 3-phosphate	ER	enoyl reductase
G-1,3-BP	glycerate-1,3-bisphosphate	MPT	malonyl/palmitoyl transferase
3-PGA	3-phosphoglyceric acid	PPT	phosphopantetheinyl transferase
2-PGA	2-phosphoglyceric acid	ACP	acyl carrierprotein protein
PEP	phosphoenolpyruvate	DH	dehydratase
Pyr	pyruvate	KR	ketoacyl reductase
AcCoA	acetyl coenzyme A	KS	ketoacyl synthase
OAA	oxaloacetate	DHAP	dihydroxyacetonephosphate
6-PGL	6-phosphogluconolactone	Glycerol-3-P	glycerol 3-phosphate
6-PG	6-phosphogluconate	Lyso-PA	lysophosphatidic acid
Ru-5-P	ribulose-5-phosphate	Acyl-DHAP	acyl dihydroxyacetonephosphate
X-5-P	xylulose-5-phosphate	PA	phosphatidic acid
Ri-5-P	ribose-5-phosphate	DAG	diacylglycerol
E-4-P	erythrose-4-phosphate	TAG	triacylglycerol
S-7-P	sedoheptulose-7-phosphate	CDP-DAG	cytidine diphosphate-diacylglycerol
AKG	α-ketoglutarate	PL	phospholipid
Glu	glutamate	SE	steryl ester
Gln	glutamine	FFA	free fatty acid
Malonyl-CoA	malonyl coenzyme A	AT	acyltransferase
Palmitoyl-CoA	palmitoyl coenzyme A	ER	enoyl reductase
Acyl-CoA	acyl coenzyme A	DH	dehydratase
AcAcCoA	acetoacetyl coenzyme A	MPT	malonyl/palmitoyl transferase;
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A	ACP	acyl carrier protein;
MVA	mevalonic acid	KR	ketoacyl reductase;
MVA-5P	mevalonate-5-phosphate	KS	ketoacyl synthase;
MVA-5PP	mevalonate-5-pyrophosphate	PPT	phosphopantetheinyl transferase
IPP	3-isopentenyl pyrophosphate		

Supplementary Table S10. Highly expressed genes described in Figure 4 and their annotations.

Group	ORF	Gene symbol	Annotation
A	RHTO_00124		hypothetical protein
	RHTO_00177		inorganic pyrophosphatase
	RHTO_00539	*ACO1*	aconitate hydratase
	RHTO_00545		ubiquinol-cytochrome c reductase core subunit 2
	RHTO_00857		histone H3B (partial sequence n rerminus)
	RHTO_01414		small ubiquitin-related modifier
	RHTO_01515		hypothetical protein
	RHTO_01776		rRNA 2'-O-methyltransferase fibrillarin
	RHTO_01875		Cytochrome b-c1 complex, subunit 10
	RHTO_01933		60s ribosomal protein l39
	RHTO_02067		40s ribosomal protein s4
	RHTO_02175		60s acidic ribosomal protein p2
	RHTO_02232		60s acidic ribosomal protein p0
	RHTO_02290		60s ribosomal protein l10
	RHTO_02584		60s ribosomal protein l29
	RHTO_02614		Hyaluronan/mRNA-binding protein
	RHTO_02750		60s acidic ribosomal protein p1
	RHTO_02851		ubiquinol-cytochrome c reductase subunit 9
	RHTO_03081		40s ribosomal protein s30
	RHTO_03117		adenylate kinase
	RHTO_03206		ATP synthase subunit beta
	RHTO_03654		F-type H+-transporting ATPase subunit h
	RHTO_03655		mitochondrial ATPase inhibitor, IATP
	RHTO_03666		cytochrome c oxidase subunit Va
	RHTO_03705		large subunit ribosomal protein L40e
	RHTO_03737		60s ribosomal protein l20
	RHTO_03921		elongation factor 1-gamma
	RHTO_03922		elongation factor 1-gamma
	RHTO_03968		large subunit ribosomal protein L7Ae
	RHTO_04030		calmodulin
	RHTO_04043		Histone H3
	RHTO_04088		Protein of unknown function DUF2611
	RHTO_04334		respiratory chain complex I component
	RHTO_04427		ATP synthase oligomycin sensitivity conferral
	RHTO_04566		ketol-acid reductoisomerase (partial sequence c terminus)
	RHTO_04839		60s ribosomal protein l15
	RHTO_04874		ran-specific GTPase-activating protein
	RHTO_04891		NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4
	RHTO_05035		60s ribosomal protein l43
	RHTO_05096		translation elongation factor EF-1 alpha subunit
	RHTO_05097		elongation factor EF-1 alpha subunit
	RHTO_05103		60s ribosomal protein l32
	RHTO_05117		peroxiredoxin (alkyl hydroperoxide reductase subunit C)
	RHTO_05208		cytochrome c
	RHTO_05390		F-type H+-transporting ATPase subunit b
	RHTO_05440		histone H2A (partial sequence c terminus)
	RHTO_05517		60s ribosomal protein l6
	RHTO_05647		mitochondrial F1 atpase subunit
	RHTO_05648		ATP synthase subunit alpha
	RHTO_05796		molecular chaperone DnaK
	RHTO_06010		40s ribosomal protein s23
	RHTO_06057		ATP-synthase delta chain
	RHTO_06114		hypothetical protein
	RHTO_06190		60s ribosomal protein l27
	RHTO_06329		single-strand binding protein
	RHTO_06406	*CIT1*	citrate synthase
	RHTO_06524		conserved hypothetical protein
	RHTO_06642		peptidylprolyl isomerase
	RHTO_06766		60s ribosomal protein l44
	RHTO_06972		chaperonin GroES
	RHTO_07038		ATP synthase D chain, mitochondrial
	RHTO_07236		phenazine biosynthesis PhzC/PhzF family protein
	RHTO_07504		secreted protein
	RHTO_07620		GDP-mannose 4,6-dehydratase
	RHTO_07648		voltage-dependent ion-selective channel
	RHTO_07756		large subunit ribosomal protein L17e
	RHTO_07772		G-protein complex beta subunit
	RHTO_07773	*GPM4*	Protein of phosphoglycerate mutase 1 family
	RHTO_07842		heat shock 70kDa protein 1/8
B	RHTO_00009		Peptidase A1 family protein
	RHTO_00066		GMP synthase (glutamine-hydrolysing)
	RHTO_00085		UPF0041 domain protein
	RHTO_00152		threonyl-trna synthetase
	RHTO_00216		H/ACA ribonucleoprotein complex subunit 3
	RHTO_00443		thioredoxin
	RHTO_00453		ribosomal biogenesis protein Gar2
	RHTO_00534	*SDH3*	succinate dehydrogenase (ubiquinone) cytochrome b subunit

	RHTO_00755		Cytochrome c oxidase, subunit VIIa
	RHTO_00984		MFS transporter
	RHTO_01082		Transcription factor
	RHTO_01344		amino acid transmembrane transporter
	RHTO_01398		ribosome assembly protein Noc2
	RHTO_01403		beta subunit of the sec61p er translocation complex (sec61p-sss1p-sbh1p)
	RHTO_01637		pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15/PRP43
	RHTO_01638		ubiquinol-cytochrome c reductase subunit 8
	RHTO_02144		37s ribosomal protein s5
	RHTO_02297		phosphoribosylaminoimidazole carboxylase
	RHTO_02453		mitochondrial import receptor subunit TOM40
	RHTO_03307		hypothetical protein
	RHTO_03555		SNF2 family helicase Ino80
	RHTO_03677		hypothetical protein
	RHTO_03952		translation initiation factor eIF-2 gamma subunit
	RHTO_04104		diamine N-acetyltransferase
	RHTO_04193		phosphomannomutase
	RHTO_04426		translation initiation factor eIF-3 subunit 8
	RHTO_04602	*CRTI*	Phytoene dehydrogenase
	RHTO_04815		expansin family protein
	RHTO_04819		phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP cyclohydrolase
	RHTO_04946		NADH-ubiquinone oxidoreductase
	RHTO_05437		ABC bile acid transporter
	RHTO_05727		Xaa-Pro dipeptidase
	RHTO_05854		protein of heat shock protein 70 family
	RHTO_05866		hypothetical protein (partial sequence c terminus)
	RHTO_05883		T-complex protein 1 subunit eta
	RHTO_06035		lysine-trna ligase
	RHTO_06236		Ribosomal protein S8e/ribosomal biogenesis NSA2 family protein
	RHTO_06293		DNA topoisomerase I
	RHTO_06431		N-acetyltransferase 10, KRE33
	RHTO_06538		delta-12 fatty acid desaturase
	RHTO_06598		translation initiation factor eIF-3 subunit 12
	RHTO_06625		ribonucleoprotein
	RHTO_06679		protein farnesyltransferase subunit beta
	RHTO_06807		zuotin
	RHTO_06818		nucleolar protein 56
	RHTO_07244		mitochondrial small subunit ribosomal protein S28
	RHTO_07729		H/ACA ribonucleoprotein complex subunit 2
	RHTO_07926		translation initiation factor eIF-6
	RHTO_07988		polyadenylate-binding protein
C	RHTO_00225		glutathione peroxidase
	RHTO_00229		alpha-glucosidase, glycoside hydrolase family 13 protein
	RHTO_00424		AN1-type zinc finger protein
	RHTO_00521		7alpha-cephem-methoxylase P8 chain related protein
	RHTO_00693		bcl-2 associated athanogene 3-like protein
	RHTO_01036		conserved hypothetical protein
	RHTO_01102		Ras-related protein Rap-11A
	RHTO_01320		riboflavin kinase
	RHTO_01329	*TPI1*	triosephosphate isomerase (TIM)
	RHTO_01466		Short-chain dehydrogenase/reductase SDR family protein
	RHTO_01497		stress response RCI peptide
	RHTO_01538		protein of heat shock protein Hsp70 family
	RHTO_01597		NADH dehydrogenase (ubiquinone) Fe-S protein 1
	RHTO_01666		Transcription factor
	RHTO_01949		hypothetical protein
	RHTO_02032	*FAS1*	fatty acid synthase subunit beta, fungi type
	RHTO_02103		Thioredoxin fold domain protein
	RHTO_02113		sulfite reductase (NADPH) hemoprotein beta-component
	RHTO_02139	*FAS2*	fatty acid synthase subunit alpha, fungi type
	RHTO_02260		protein of short-chain dehydrogenase/reductase SDR family
	RHTO_02346		Barwin-related endoglucanase domain protein
	RHTO_02448		GTP-binding protein SAR1
	RHTO_02628	*PYC1*	pyruvate carboxylase
	RHTO_02670		hypothetical protein
	RHTO_02686		conserved hypothetical protein
	RHTO_02888		conserved hypothetical protein
	RHTO_03100		hypothetical protein
	RHTO_03201		DOPA 4,5-dioxygenase
	RHTO_03853		Acyl-CoA-binding protein, ACBP
	RHTO_04032		phenazine biosynthesis PhzC/PhzF family protein
	RHTO_04650	*GDH1*	glutamate dehydrogenase (NADP+)
	RHTO_05032		endoglucanase
	RHTO_05324		pirin family protein
	RHTO_05438		aromatic ring-opening dioxygenase
	RHTO_06037		protein of GFA famliy
	RHTO_06292		NADH:ubiquinone oxidoreductase, B12 subunit
	RHTO_06575		NmrA-like domain containing protein
	RHTO_06626		Arrestin-like, N-terminal and Immunoglobulin E-set domain protein
	RHTO_06652		SNF2-related domain containing protein

	RHTO_07076		nitroreductase
	RHTO_07456		hypothetical protein
	RHTO_07647		hypothetical protein
D	RHTO_00155		cyclophilin, peptidyl-prolyl cis-trans isomerase
	RHTO_00206		superoxide dismutase, Fe-Mn family
	RHTO_01358		translation initiation factor eIF-1A
	RHTO_02788	*GND1*	6-phosphogluconate dehydrogenase
	RHTO_03087		U4/U6 small nuclear ribonucleoprotein SNU13
	RHTO_03113		60s ribosomal protein l16
	RHTO_04131		glycine hydroxymethyltransferase
	RHTO_04592		60s ribosomal protein l33
	RHTO_04740		40s ribosomal protein s18
	RHTO_05162		elongation factor Tu
	RHTO_05630		regulatory protein suaprga1
	RHTO_06701		hypothetical protein
	RHTO_07146		mitochondrial ATPase, F1 complex, epsilon subunit
	RHTO_07151		cellular nucleic acid-binding protein
	RHTO_07734		histone H2A
	RHTO_07843		elongation factor EF-1 beta subunit
E	RHTO_00314		S-phase kinase-associated protein 1
	RHTO_00340		GCN5-related N-acetyltransferase
	RHTO_00455		NADH dehydrogenase (ubiquinone) 1 beta subcomplex 8
	RHTO_00488		UbiE/COQ5 family methyltransferase
	RHTO_00627		hypothetical protein
	RHTO_00723	*SDH4*	succinate dehydrogenase (ubiquinone) membrane anchor subunit
	RHTO_01101		proton-dependent oligopeptide transporter, POT family
	RHTO_01139		Protein of 14-3-3 protein family
	RHTO_01289	*IDH1*	isocitrate dehydrogenase (NAD+)
	RHTO_02257		adenylosuccinate synthetase
	RHTO_02398		Cyclin PHO80-like family protein
	RHTO_03155		CCCH finger DNA binding protein
	RHTO_03248	*TKL1*	transketolase
	RHTO_03543	*LAT1*	pyruvate dehydrogenase E2 component
	RHTO_03588		NADH dehydrogenase (ubiquinone) 1 beta subcomplex 7
	RHTO_03741		conserved hypothetical protein
	RHTO_03852		zip-like iron-zinc transporter
	RHTO_03915	*ACL1*	ATP citrate synthase
	RHTO_04266		MFS transporter, sugar:H+ symporter
	RHTO_04276		pyridoxamine phosphate oxidase family protein
	RHTO_04330		thiamine biosynthesis protein Nmt1
	RHTO_04448		protein of ornithine decarboxylase antizyme family
	RHTO_04628		hypothetical protein
	RHTO_04861		conserved hypothetical protein
	RHTO_04910		Cytochrome c oxidase, subunit VIIa
	RHTO_05679		thiamine biosynthetic enzyme
	RHTO_05829		heterogeneous nuclear ribonucleoprotein M
	RHTO_05912		IMP dehydrogenase
	RHTO_06048		ATP synthase gamma chain
	RHTO_06080		MFS monosaccharide transporter
	RHTO_06449		siderophore biosynthesis protein
	RHTO_06773		UDP-galactopyranose mutase
	RHTO_06930		mitochondrial peroxiredoxin 6, 1-Cys peroxiredoxin
	RHTO_07025		Trp repressor binding protein
	RHTO_07616		EXTL3, alpha-1,4-N-acetylglucosaminyltransferase, glycosyltransferase family 64 protein
	RHTO_07708		alpha tubulin
	RHTO_07812		phosphoglycerate mutase family protein
	RHTO_07893	*LPD1*	dihydrolipoyl dehydrogenase
	RHTO_07939	*SOL3*	6-phosphogluconolactonase
	RHTO_08148		conserved hypothetical protein
F	RHTO_00600		conserved hypothetical protein
	RHTO_00789		molecular chaperone HtpG
	RHTO_02321		ubiquitin-conjugating enzyme E2 D/E
	RHTO_03987		glycine-rich RNA binding protein
	RHTO_04325		40s ribosomal protein s6
	RHTO_04577		Cytochrome c oxidase, subunit Vb
	RHTO_04890		40s ribosomal protein s7
	RHTO_05418		gpi anchored protein
	RHTO_06539		40s ribosomal protein s12
	RHTO_06973		Protein of chaperonin Cpn60 family
	RHTO_06994		40s ribosomal protein s21
	RHTO_07359		GTP-binding nuclear protein Ran
G	RHTO_00070		ubiquinol-cytochrome c reductase complex 17 kd protein
	RHTO_00323	*ENO1*	enolase
	RHTO_00354		40s ribosomal protein s10
	RHTO_00469		60s ribosomal protein l34
	RHTO_00474		mitochondrial phosphate carrier protein
	RHTO_00632		large subunit ribosomal protein L27Ae
	RHTO_00673	*GLN1*	glutamine synthetase
	RHTO_01016		40s ribosomal protein s28
	RHTO_01017		40s ribosomal protein s5

RHTO_01048		acyl-carrier protein
RHTO_01290	*IDH2*	isocitrate dehydrogenase (NAD+)
RHTO_01367		37s ribosomal protein s24
RHTO_01695		adenosylhomocysteinase
RHTO_01866		plasma membrane H+-transporting ATPase
RHTO_01869		ubiquinol-cytochrome c reductase iron-sulfur subunit
RHTO_01873		60s ribosomal protein l5
RHTO_01907		translation initiation factor eIF-4A
RHTO_02009		60s ribosomal protein l30
RHTO_02115		small subunit ribosomal protein S27Ae
RHTO_02116		60s ribosomal protein l37
RHTO_02176		40s ribosomal protein s15
RHTO_02373		large subunit ribosomal protein L14e
RHTO_02403		40s ribosomal protein s25
RHTO_02455		60s ribosomal protein l26
RHTO_02483		40s ribosomal protein s27
RHTO_02569		heat shock 70kDa protein 1/8
RHTO_02577		60s ribosomal protein l23
RHTO_02646		Protein of 14-3-4 protein family
RHTO_02697		elongation factor 3
RHTO_03043	*FBA1*	fructose-bisphosphate aldolase, class II
RHTO_03161		60s ribosomal protein l22
RHTO_03292		nucleoside-diphosphate kinase
RHTO_03294		ubiquinol-cytochrome c reductase subunit 7
RHTO_03373		translation initiation factor 5A
RHTO_03560		actin
RHTO_03573		translation elongation factor 2
RHTO_03637		40s ribosomal protein s9
RHTO_03665		60s ribosomal protein l21
RHTO_03746	*GAPDH*	glyceraldehyde 3-phosphate dehydrogenase
RHTO_03896		mitochondrial processing peptidase beta subunit
RHTO_03965		60s ribosomal protein l7
RHTO_03967		60s ribosomal protein l36
RHTO_04048		60s ribosomal protein l24
RHTO_04114		60s ribosomal protein l18
RHTO_04152		cytochrome c1
RHTO_04184		histone H4
RHTO_04185		histone H3
RHTO_04216		40s ribosomal protein s14
RHTO_04326		40s ribosomal protein s13
RHTO_04363	*MDH1*	L-malate dehydrogenase
RHTO_04421		SERF family protein
RHTO_04531		hypothetical protein
RHTO_04593		large subunit ribosomal protein L31e
RHTO_04816		60s ribosomal protein l19
RHTO_05011		40s ribosomal protein s3
RHTO_05106		solute carrier family 25 (mitochondrial carrier)
RHTO_05119		60s ribosomal protein l13
RHTO_05130		40s ribosomal protein s19
RHTO_05147		argininosuccinate synthase
RHTO_05301		60s ribosomal protein l10a
RHTO_05368		40s ribosomal protein s26
RHTO_05447		40s ribosomal protein s16
RHTO_05617		60s ribosomal protein l11
RHTO_05740		protein of short-chain dehydrogenase/reductase SDR family
RHTO_05917		60s ribosomal protein l2
RHTO_05964		60s ribosomal protein l9
RHTO_06133		ubiquitin C
RHTO_06166		40s ribosomal protein s0
RHTO_06248		40s ribosomal protein s22
RHTO_06249		40s ribosomal protein s17
RHTO_06384		Peptidyl-prolyl cis-trans isomerase, FKBP-type
RHTO_06491		protein of cation efflux protein family, zinc transporter
RHTO_06792		60s ribosomal protein l35
RHTO_06815		40s ribosomal protein s1
RHTO_06955	*TAL1*	transaldolase
RHTO_07080		S-adenosylmethionine synthetase
RHTO_07200		hypothetical protein
RHTO_07334		translationally controlled tumor protein homolog
RHTO_07683		40s ribosomal protein s2 (partial sequence n rerminus)
RHTO_07691		5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase
RHTO_07771		60s ribosomal protein l12
RHTO_07779		ADP-ribosylation factor
RHTO_07804		F-type H+-transporting ATPase subunit f
RHTO_07868		60s ribosomal protein l2
RHTO_07886		60s ribosomal protein l23a
RHTO_07911		40s ribosomal protein s20
RHTO_07954		60s ribosomal protein l3
RHTO_08150		40s ribosomal protein s8
RHTO_08151		40s ribosomal protein s29

Supplementary Table S11. Total spectral counts and proteins identified from the seed, 24 h and 96 h[a]

Sample name	Total spectral counts	Unique peptides	Total proteins found[b]	Accumulative proteins found[b]	Identified proteins[c]
Seed 1	1929 (35949)	186 (7280)	67 (1638)	105 (2088)	50 (1473)
Seed 2	1721 (36160)	184 (8313)	63 (1755)		
Seed 3	1717 (33679)	188 (8003)	65 (1630)		
24 h 1	699 (18288)	116 (5692)	38 (1418)	78 (2029)	29 (1353)
24 h 2	1009 (28171)	131 (6569)	47 (1555)		
24 h 3	819 (28923)	119 (7782)	43 (1675)		
96 h 1	818 (28711)	146 (8384)	56 (1901)	84 (2270)	35 (1522)
96 h 2	722 (28992)	124 (6975)	46 (1613)		
96 h 3	474 (21697)	103 (6409)	38 (1669)		
Total	9908 (260570)	-	-	184 (3108)	63 (2057)

a. data from Liu et al (ref 32) and this study (in parentheses) were presented.

b. ≥ 2 unique peptides and false-positive-rate < 5% for previous data and optimal criteria obtained from SFOER and false-positive-rate < 1% for this study.

c. at least two occurrences within three replicates.

Supplementary Methods

Electrophoretic karyotyping. The genome size and chromosome number were estimated by contour-clamped homogeneous electric field (CHEF) technology according to known procedures[1], excepting that MAN5C[2] was used for cell lysis. Results were shown in **Supplementary Figure S5**. Cells of stationary phase were harvested by centrifugation (1000 *g*, 15 min and 4 °C), then washed twice with 0.05 M EDTA (pH 7.6) at 4 °C. Approximate 10^9 cells suspended in 500 ul citrate buffer (pH 4.5, 0.3 M) with EDTA (0.1 M) and MAN5C (10 U/ml) were mixed with 500 ul 1.2% low melting point agarose (Bio Basic Inc.), which was pre-melted in citrate buffer (pH 4.5, 0.3 M) with EDTA (0.1 M) and equilibrated to 45 °C. The cell:agarose suspension was pipetted into casting molds and cooled at 4 °C for 20 min. Incubate the cell:agarose plugins in 2× volume of lysis buffer (0.3 M citrate, 0.1 M EDTA, 5 U/ml MAN5C, pH 4.5) at 50 °C for 2 hours, and then replace with 2× volume of digestion buffer (0.45 M EDTA, 1% Sarkosyl, 1 mg/ml protease K, pH 8.0) and refresh the digestion buffer after 24 hours. After incubation at 50 °C for 48 h, the plugins were directly stored at 4 °C until used. Plugins of appropriate volume were cut and inserted into loading wells of 1% agarose gel (D5, Takara). The electrophoresis was run at 14 °C in 0.5× TBE buffer (45 mM Tris-borate, 1 mM EDTA) with a CHEF apparatus (Liuyi Instrument Factory, Beijing). The chromosomes of *S. cerevisiae* (NEB) and *Hansenula wingei* (Bio-Rad) were used as molecular size standards. The setup of electrophoresis was constant in reorientation angle of 120° but variant in voltage and switch time. The optimal parameters obtained was ramped switch time of 100-200 s at 4 V/cm for 20 h and then 200-400 s at 3.5 V/cm for 34 h. Similar karyotypes of *R. toruloides* ATCC 10788 (type strain) and NP11 were found, and smaller chromosome bands were found in *R. toruloides* Y4 and NP5-2.

Initial gene prediction. Repetitive sequences were predicted *de novo* by RepeatModeler (v 1.0.3, http://www.repeatmasker.org/RepeatModeler.html). 25 repeat families were classified with TEclase[3] and used as the search library for RepeatMasker (v 3.2.9, http://www.repeatmasker.org/). For gene prediction, 14830 Unigenes from *de novo* assembly and 6602 transcription unit (8234 transcripts) from *ab initio* assembly were aligned to genome sequence. Gene models were predicted with Augustus[4], GlimmerHMM[5], SNAP[6], and GeneMark-ES[7]. The programs were trained on the species of *Cryptococcus neoformans*, excepting for GeneMark-ES who was self-trained. Proteomes of relative species were used for homology-based gene prediction with GeneWise[8]. For non-coding sequence prediction, putative tRNA coding genes were identified with tRNAScan-SE[9]. The rDNA locus including 5S, 18S and 26S rRNA coding genes was PCR amplified and sequenced. The length of rDNA repeat unit is 8.3 kb, and it's attached to Scaffold 22. Spliceosomal RNAs (SnRNA) were scanned with cmsearch from Infernal package[10] and classified according to RFAM database[11], and U2 (2 members), U4 (2 members) and U6 (2 members) SnRNAs were identified.

Supplementary References

1. Herschleb, J., Ananiev, G. & Schwartz, D.C. Pulsed-field gel electrophoresis. *Nat. Protoc.* **2**, 677—684 (2007).
2. Yang, F., Zhang, S., Jin, G., Lin, X. & Zhao, Z.K. Purification and characterization of a β-1,3-glucomannanase expressed in *Pichia pastoris*. *Enzyme Microb. Technol.* **49**, 223—228 (2011).
3. Abrusan, G., Grundmann, N., DeMester, L. & Makalowski, W. TEclass——a tool for automated classification of unknown eukaryotic transposable elements. *Bioinformatics* **25**, 1329—1330 (2009).
4. Stanke, M., Tzvetkova, A. & Morgenstern, B. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biol.* **7**, S11 (2006).
5. Majoros, W. H., Pertea, M. & Salzberg, S. L. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878—2879 (2004).
6. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).

7. Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O. & Borodovsky, M. Gene prediction in novel fungal genomes using an *ab initio* algorithm with unsupervised training. *Genome Res.* **18**, 1979—1990 (2008).
8. Birney, E. & Durbin, R. Using GeneWise in the Drosophila annotation experiment. *Genome Res.* **10**, 547—548 (2000).
9. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 0955—0964 (1997).
10. Nawrocki, E. P., Kolbe, D. L. & Eddy, S. R. Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25**, 1335—1337 (2009).
11. Griffiths-Jones, S. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **33**, D121—D124 (2004).