BMC Microbiology

Identification of novel genes in the carotenogenic and oleaginous yeast Rhodotorula toruloides through genome-wide insertional mutagenesis

--Manuscript Draft--

Manuscript Number:	MCRO-D-17-00346R1
Full Title:	Identification of novel genes in the carotenogenic and oleaginous yeast Rhodotorula toruloides through genome-wide insertional mutagenesis
Article Type:	Research article
Section/Category:	Microbial genetics, genomics and proteomics
Funding Information:	Singapore National Research Fundation (CRP-8-2011-02) — Dr Lianghui Ji Economic Development Board - Singapore — Dr Lianghui Ji Temasek Trust — Dr Lianghui Ji
Abstract:	Background Rhodotorula toruloides is an outstanding producer of lipids and carotenoids. Currently, information on the key metabolic pathways and their molecular basis of regulation remains scarce, severely limiting efforts to engineer it as an industrial host. Results We have adapted Agrobacterium tumefaciens-mediated transformation (ATMT) as a gene-tagging tool for the identification of novel genes in R. toruloides. Multiple factors affecting transformation efficiency in various species and several strains from the Pucciniomycotina subphylum were optimized. The Agrobacterium transfer DNA (T-DNA) showed predominantly single-copy chromosomal integrations in R. toruloides, which were trackable by high efficiency thermal asymmetric interlaced PCR (hiTAIL-PCR). To demonstrate the application of random T-DNA insertions for strain improvement and gene hunting, 3 T-DNA insertional libraries were screened against cerulenin, nile red and tetrazolium violet respectively, resulting in the identification of 22 mutants with obvious phenotypes in fatty acid or lipid metabolism. Similarly, 5 carotenoid biosynthetic mutants were obtained through visual screening of the transformants. To further validate the gene tagging strategy, the carotenoid mutant RAM5 was analyzed in detail. The mutant had a T-DNA inserted at the putative phytoene desaturase gene CAR1. Deletion of CAR1 by homologous recombination led to a phenotype similar to RAM5 and it could be genetically complemented by re-introduction of the wild-type CAR1 genomic sequence. Conclusions T-DNA insertional mutagenesis is an efficient forward genetic tool for gene discovery in R. toruloides and related oleaginous yeast species. It is also valuable for metabolic engineering in these hosts. Further analysis of the 27 mutants identified in this study should augment our knowledge of the lipid and carotenoid biosynthesis, which may be exploited for oil and isoprenoid metabolic engineering.
Corresponding Author:	Yanbin Liu, Ph.D. Temasek Life Sciences Laboratory Ltd Singapore, SINGAPORE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Temasek Life Sciences Laboratory Ltd
Corresponding Author's Secondary Institution:	
First Author:	Yanbin Liu, Ph.D.
First Author Secondary Information:	
Order of Authors:	Yanbin Liu, Ph.D. Chong Mei John Koh

	Sihui Amy Yap
	Minge Du
	Mya Myintzu Hlaing
	Lianghui Ji
Order of Authors Secondary Information:	
Response to Reviewers:	Response to editor and reviewers Dear Editor and Reviewers, Thank you for all your constructive comments and suggestions. Accordingly, we have completed the suggested experiments and modified extensively our manuscript. Kindly review our point-to-point responses to comments below. 1.To Editor I recommend carefully regarding the comments and re-writing of the manuscript. Referee 1 also requested language editing, I strongly recommend to contact a native speaker or another expert in English to get a correctly written manuscript. -Reply 1: Thank you for the advice. We have made extensive revisions and corrections to the manuscript and the final version has also been proofread by native speakers. Referee 1 already raised some concerns about Figure 5A. I agree with that and I even think that this Figure should be removed. When I understand it correctly, this curve is not reflecting experimental results but theoretical calculations. However, the dots in the curve on the first view provide the impression that there was a kind of testing with these numbers of transformants. To my opinion it should be stated in the text that an acceptable coverage of the genome (more than 80%) would only be achieved if there are more than 20,000 transformants, and that, with the number of obtained transformants, you are still far away from reaching this coverage. This is not a problem for now, since covering of almost the whole genome was not the purpose of the study. -Reply 2: The previous Fig. 5A and related presentation of the data have been removed. 2.To Reviewer #1 The study by Liu et al., describes an insertional mutagenesis method for Rhodotorula toruloides using Agrobacterium tumefaciens. The method was used to identify genes involved in lipid and carotenoid production. This study is the first report of A. tumefaciens successfully transforming Rhodotorula toruloides. The study is very detailed and once the system was developed, they used it to discover genes important in this organism. Most people that use this system to study fungi should find it useful. There are a number of minor comments that need to be addressed. Results: Application of ATMT in Puicciniomycotina species. This section has no context. It just reports a figure. It would be useful to know the number of transformants per (something like Rhodotorula CFU). The small colonies on the SR panel need to be explained. Are they spontaneous drug resistant mutants and not transformants? If they are transformants why are half of them small colonies? Rg2 needs a higher concentration of drug. This level should be confirmed by doing a does response. -Reply 1: The transformation efficiency was represented using as CFU (colony forming unit per 106 fungal cells) and Fig. 1B has been added for the comparison transformation on several Pucciniomycotina species. Please find the relevant description in the result section (Lines 96-106). The hygromycin concentration had been optimized before this study and we found 100 g/ml was sufficient for the selection. Kindly take note that we chose a higher selection pressure (150 µg/ml) to reduce false positives. The resultant mutants have also been confirmed to have true T-DNA integrations by colony PCR and Southern blotting (Fig. 3). Under the conditions and using our selection marker, false positives was rare for Rt, Rg3 and Sr. Results: Figure 2C. The upper limit of transformation efficiency based on ammonium sulfate concentration has not been reached. The highest level of transformants is reached at 0.5 g/L. Why was a concentration lower than 0.5 not used? -Reply 2: We agree that 0.5 g/L ammonium sulfate may not be the optimum. However, the negative effect of high nitrogen source concentration has been well demonstrated (Lines 116-121). Results 182-185: "Regarding to the genome size (s) of 21,490 kb, average gene length

(x) of 1.882 kb [59, 60], and a random distribution of insertion events in R. toruloides, the library coverage should be improved by approaching 100% by increasing T-DNA integration events (library content) (Fig. 5A)". This statement doesn't really say anything, but what it is trying to say is incorrect. As it reads, it conveys that library coverage could be improved by increasing transformation frequency. This outcome is obvious. However, the coverage will never be 100% (at least one insertion per gene) because some genes are essential and cannot be disrupted because a non-function essential gene will result in an inviable cell that will not grow into a colony.
-Reply 3: Thank you for your advice. We have deleted this paragraph.

Substantial editing for grammar is needed.
-Reply 4: Done.

Line 121. Concentration is spelled wrong
-Reply 5: Corrected.

3.To Reviewer #2
Because the red yeast Rhodosporidium toruloides has some unique features for potential biotechnological applications, it is highly attractive to develop more efficient tools and strategies for engineering this yeast. This manuscript describes the results of insertional mutagenesis of R. toruloides and subsequent identification of a few functional genes. Key information for insertional mutagenesis was disclosed. The function of a putative phytoene desaturase gene (CAR1) was confirmed. The work was well-designed and the data were solid.
Comments,
1) Figure 7E: the retention time (X-axis) was missing.
-Reply 1: Corrected.

2) Line 316-318: Please elaborate in more details regarding the membrane-free transformation method.
-Reply 2: Details were added in Materials and Methods section (Line 366-371) and Discussion section (Lines 257-261).

4.To Reviewer #3
The authors of the manuscript describe a known general method that is optimized and applied to R. toruloides. The authors found interesting phenotypes that could be linked to metabolic pathways or regulation. the list of genes presented could be useful for further studies in the field as well as the technique they applied. I have some comments or questions I would like the authors to respond.
In the parts 'Optimization of ATMT protocol for large-scale screening' and 'Characterization of genome-wide T-DNA insertion patterns' the authors should compare their results and values with previous reports for the same technique to see how well is it performing in this yeast.
-Reply 1: More discussions have added (Lines 155-159, 249-256).

according to 'Characterization of genome-wide T-DNA insertion patterns' only 75% of the transformants are monogenic. How this influence the results obtained later? how many of the studied clones were mongenics? are the selected ones for lipid and carotenoids metabolism all mongenics?
-Reply 2: This lower ratio of single-copy integration events was caused by the use of lower strength of promoter (U. maydis gpd1 promoter). We have shown that it could be improved to ~100% if stronger promoter for the expression of hpt-3 was used (Lines 277-291). However, the multiple-copy integrations of T-DNA could result in the higher expression of target protein and would be useful for metabolic engineering work.

In the part 'Chemicals-assisted screening for T-DNA tagging mutants with modified fatty acid profiles' the insertion sites should be more discussed, trying (if possible) to find more information on for example how the resistance to cerulenin is achieved in the mutants. The same for the other two experiments.
-Reply 3: It would be difficult to speculate on how the chemicals and genes/proteins identified interact or function. As you can see, we had identified different categories of enzymes. Answers to the question will take more time and we think it would be beyond the scope of this manuscript.

	In addition, it is not clear in the text how the authors determined the selected mutants when they say: transformants showing larger sizes on cerulenin-containing media, darker purple-color pigmentation on tetrazolium violet-containing media or higher fluorescent intensity on nile red-containing media were transferred to YPD broth. How different size or intensities were measured? Nile red is specified in the methods but the results are not shown in any figure. -Reply 4: The relevant paragraphs were revised by adding more details (lines 170-201). The additional file 4: Fig. S2A-C were also modified to show how these mutants were selected. Figure 6 ABC should show the values also of individual mutants (linked to the insertion sites) to identify which ones present more different phenotypes. -Reply 5: This figure was re-generated to show the relative ALA or lipid yields of individual mutants. line 236-237 the authors discuss the effect of insertions in regulatory elements. do they think in this particular case the insertion is causing activation or repression in the expression? this gene is already known to affect lipid accumulation. Did the authors did q-PCR of this gene? -Reply 6: We did not perform full characterization of the mutants except RAM5. We hope that we can answer this question in our future publications. line 252, can the author comment how this activities could be related to carotenoid production? -Reply 7: The relevant discussion was added (lines 264-269). Did the authors find any mutant producing higher amount of carotenoid content? or different colors? -Reply 8: The color changes were distinguished only if the differences were high enough for human sight. However, the deeper colors as considered to be elevated carotenoid production, were hardly achieved. There are several factors affecting the color, including the storage time and temperature. Currently we have found one IM-mutant with different color, more yellowish as compared to the pinkish color of wildtype strain. The relevant gene in the mutant has studied in details and will be reported soon. Table S1 should look more like table S2, with the integration names, etc. -Reply 9: Updated (Additional file 2: Table S1)

Identification of novel genes in the carotenogenic and oleaginous yeast *Rhodotorula toruloides* through genome-wide insertional mutagenesis

Yanbin Liu[1]*

Email: yanbin@tll.org.sg

Chong Mei John Koh[1]

Current email: john_koh@np.edu.sg

Sihui Amy Yap[1]

Current email: replish@hotmail.com

Minge Du[1]

Current email: minge.du@vai.org

Mya Myintzu Hlaing[1]

Current email: myintzu.hlaing@csiro.au

Lianghui Ji[1,2] *

Email: jilh@tll.org.sg

[1] Biomaterials and Biocatalysts Group, Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604

[2] School of Biological Sciences, Nanyang Technological University, Singapore, 60 Nanyang Drive, 637551, Singapore

* Corresponding authors

Email: yanbin@tll.org.sg (YL), jilh@tll.org.sg (LJ)

Tel: +65 6872 7484 (YL), +65 6872 7483 (LJ)

Fax: +65 6872 7007

Running title: Genome-wide insertional mutagenesis in *R. toruloides*

Abstract

Background

Rhodotorula toruloides is an outstanding producer of lipids and carotenoids. Currently, information on the key metabolic pathways and their molecular basis of regulation remains scarce, severely limiting efforts to engineer it as an industrial host.

Results

We have adapted *Agrobacterium tumefaciens*-mediated transformation (ATMT) as a gene-tagging tool for the identification of novel genes in *R. toruloides*. Multiple factors affecting transformation efficiency in various species and several strains from the *Pucciniomycotina* subphylum were optimized. The *Agrobacterium* transfer DNA (T-DNA) showed predominantly single-copy chromosomal integrations in *R. toruloides*, which were trackable by high efficiency thermal asymmetric interlaced PCR (hiTAIL-PCR). To demonstrate the application of random T-DNA insertions for strain improvement and gene hunting, 3 T-DNA insertional libraries were screened against cerulenin, nile red and tetrazolium violet respectively, resulting in the identification of 22 mutants with obvious phenotypes in fatty acid or lipid metabolism. Similarly, 5 carotenoid biosynthetic mutants were obtained through visual screening of the transformants. To further validate the gene tagging strategy, the carotenoid mutant RAM5 was analyzed in detail. The mutant had a T-DNA inserted at the putative phytoene desaturase gene *CAR1*. Deletion of *CAR1* by homologous recombination led to a phenotype similar to RAM5 and it could be genetically complemented by re-introduction of the wild-type *CAR1* genomic sequence.

Conclusions

T-DNA insertional mutagenesis is an efficient forward genetic tool for gene discovery in *R. toruloides* and related oleaginous yeast species. It is also valuable for metabolic engineering in these hosts. Further analysis of the 27 mutants identified in this study should augment our knowledge of the lipid and carotenoid biosynthesis, which may be exploited for oil and isoprenoid metabolic engineering.

Keywords

Agrobacterium tumefaciens-mediated transformation, *Pucciniomycotina*, insertional mutagenesis, metabolic engineering, carotenoid and lipid biosynthesis

Background

A large number of oleaginous microorganisms capable of producing more than 20% of their dry biomass as lipids have been reported to date [1-3]. They are potential alternative hosts to plants for the production of lipid and fatty acid derivatives, such as biodiesel, alkane, fatty alcohol and wax [1, 4-7]. On the other hand, only limited number of non-photosynthetic microorganisms can naturally produce carotenoids, which are protective agents against UV radiation and oxidative stress (for review, see [8]). *Rhodotorula toruloides* (syn. *Rhodosporidium toruloides* [9]), a species of the *Pucciniomycotina* subphylum, has gained increasing attention due to its outstanding cell growth rate in high density fermentation, high lipid and carotenoid productivity, and the capability to utilize cheap feedstocks [10-15].

Genetic tools for *R. toruloides* have increased steadily over recent years since the first report of stable genetic transformation [16]. *R. toruloides* is being developed as a new synthetic biology platform [16-26]. To date, information regarding the molecular control of metabolism and catabolism remains rare in this host, severely limiting the development of *R. toruloides* as an industrial workhorse.

Microbial adaptive laboratory evolution (ALE) is a useful tool for metabolic engineering [27]. Chemical mutagens or ultraviolet radiation were often used to improve strains or populations of interests using a specific selection pressure. Such techniques usually produce mutants with point mutations. Despite the fast advancement of genome sequencing technology, the identification of point mutations remains a tedious task [28, 29]. DNA insertional mutagenesis (IM) has become a versatile forward genetic tool in diverse species, including fishes [30], plants [31], animals [32], algae [33], bacteria [34] and fungi [35]. Due to the high efficiency in genetic diversity, IM could be exploited for fast strain improvement, particularly for

microcrobes [36]. A good gene-tagging tools should be: i). DNA is randomly integrated into the nuclear genome [35]; ii). The disrupted gene targets can be easily identified [37, 38]; iii). The host contains a haploid genome [39-41]. *Agrobacterium tumefaciens*-mediated transformation (ATMT) delivers the T-DNA into the host's nuclear genome and has been widely used as an IM tool, particularly in fungi and plants [42-44].

Here, we demonstrate the application of ATMT for gene discovery and modifications of the metabolic pathway in *R. toruloides*.

Results

Application of ATMT in *Puicciniomycotina* subphylum

We have reported a reliable transformation protocol for *R. toruloides* ATCC 10657 using the dominant selection conferred by a codon-optimized hygromycin resitance gene (hygromycin-B-phosphotransferase gene *hpt-3*) [16]. While the method was generally applicable in several other *Puccinimycotina* species or strains, *e.g. R. toruloides* ATCC 10788 (Rt2), *R. glutinis* ATCC 90781 (Rg1), *Rhodotorula glutinis* ATCC 204091 (Rg2), *R. graminis* WP1 (Rg3), and *Sporobolomyces roseus* FGSC 10293 (IAM13481, Sr) (Fig. 1A), large variations in the transformation efficiency (TFE, or Colony Forming Unit per 10^6 fungal cells), were observed (Fig. 1B). The average of CFU for Rg3, Rg1, Rt2 and Sr was 985, 409, 227 and 197, respectively. Notably, strains Rt1 and Rg2 showed much lower TFE, producing only 18 and 10 CFU, respectively (Fig. 1B). Colony PCR and Southern blot analysis confirmed that more than 90% transformants contained T-DNA integration (See Fig. 4. Data on Sr and Rg3 are not shown).

Optimization of ATMT protocol for large-scale screening

The low TFE for some strains prompted us to investigate the effects of various co-culture parameters. Similar to other reports, virulence inducer (acetosyringone) for agrobacteria, co-culture time, cell ratio between T-DNA recipient and donor, and promoters used to drive *hpt-3* expression drastically influenced the TFE in *R. toruloides* (Additional file 1: Fig. S1A-D). Notably, TFE was highly sensitive to the pH of the induction medium (Fig. 2A), where even a slight increase of pH from the optimum (pH5.5) resulted in a dramatic decrease in TFE (Fig. 2A). The hardness of co-culture medium (agar concentration) also influenced TFE, with the optimal agar concentration of 2.0% (w/v) (Fig. 2B). The role of nitrogen concentration on TFE was also investigated due to its multiple roles in energy metabolism, cell growth and differentiation [45, 46]. In our standard ATMT protocol, 0.5 g/L ammonium sulfate is used as the sole nitrogen source in the induction medium [16]. Increasing its concentration led to severe reduction in TFE and transformation was completely abolished at 50 g/L of ammonium sulfate (Fig. 2C).

Furthermore, the effect of membrane types on TFE was investigated. Results showed that the different supporting membranes dramatically affected TFE, where the positively charged (nylon Hybond N^+) and neutral membrane (cellulose acetate) supported higher TFE. Interestingly, co-culturing cells directly on the surface of agar medium (without the support of any membrane, membrane-free) led to a high TFE, comparable to that with cellulose acetate membrane (Fig. 2D).

Characterization of genome-wide T-DNA insertion patterns

Southern blot analysis of 64 T-DNA mutants from *R. toruloides* ATCC 10657 showed that 75% of transformants (48/64) contained a single copy of T-DNA, 20% (13/64)

with two copies, and 5% (3/64) with three copies or more (Fig. 3). The average copy number of T-DNA in the genome was 1.36.

The genome sequences adjacent to T-DNA tagging positions were analyzed by high efficiency thermal asymmetric interlaced PCR (hiTAIL PCR) [37, 38]. A total of 480 samples were analyzed, including 192 transformants of *R. toruloides* ATCC 10657 analyzed for both left border (LB) and right border (RB) flanking sequences, and 96 transformants of *R. glutinis* ATCC 90781 analyzed for LB flanking sequences only. HiTAIL PCR success rate was 72.5% (346/480), which led to 268 high-quality sequencing results (77.5%). To identify the chromosomal positions of T-DNA insertion, 61 LB flanking sequences were analyzed by BLASTn searches against the *R. glutinis* ATCC 204091 genome database (Additional file 2: Table S1). As expected, T-DNAs were mapped to the majority of scaffolds (21 out of 29), with scaffold No. 2, 13, 18 and 26 showing the highest number of hits (Fig. 4A). Scaffolds that missed the analysis were all small in size.

T-DNA integration is known to be initiated at the RB, with DNA nicks generated between the 3rd and 4th nucleotide of the 25 bp border repeat sequence (RB canonical insertion). Generally, the inserted T-DNAs contain deletion of various lengths in the RB region [47]. Our results showed that T-DNA integration exhibited much higher accuracy at RB than LB (Fig. 4B), with 30.8% (79 of 172) of the inserted T-DNAs containing deletions (ranging from 1 to 80 bp) at RB end compared to 79.1% deletions (68 of 86) at the LB end.

Microhomology at the insertion junctions was also examined. 73.7% (101 of 137) of LB insertions showed homology of 4 bp or more compared with 42.0% (55 of 131) at RB. Microhomology of up to 10 bp was found, although it was usually less than 4 bp (Fig. 4C). Collectively, T-DNA integration in *Rhodotorula* species does not

require long stretch of sequence homology at the cross-over position, which is similar to the illegitimate recombination in other species, such as budding yeast *Saccharomyces cerevisiae* [48], corn smut fungi *Ustilago maydis* [39], rice blast fungi *Magnaporthe oryzae* [49], plants [50] and mammals [51].

The DNA sequences within 1.0 kb from the insertion sites were annotated. As shown in Additional file 3: Table S2, most mutants could be functionally assigned: 15.8% are likely to be involved in glycan metabolism, 14.3% in transcription, 7.1% in protein folding and trafficking, 4.1% in carbohydrate metabolism, 4.1% in DNA replication and repair, and 4.1% in transport (Fig. 4D). 26.0% of the mutants could not be functionally assigned (presented as "Unknown" in Fig. 4D). It was obvious that T-DNA insertions preferred gene coding regions over gene regulatory regions, such as promoters and terminators (Table 1 and Additional file 3: Table S2). Taken together, T-DNA insertions could be exploited to tag a wide range of genes.

Direct identification of lipogenic mutants

Cerulenin, (2S)(3R)2,3-epoxy-4-oxo-7,10-dodecadienoylamide, was discovered from the culture broth of *Cephalosporium caerulens* [52, 53]. It has been used as a fungicide due to the inhibition effect on the biosynthesis of fatty acids and steroids [54]. Mutants that survived cerulenin treatment were expected to produce higher level of lipid or polyunsaturated fatty acids (PUFAs), and the relevant genes would be useful for lipid metabolic engineering [55-57]. Approximately 10,000 transformants were screened against 50 μg/mL cerulenin in the selection plates, a level that fully blocked the growth of wild-type (WT) cells (Additional file 4: Fig. S2A). A total of 12 T-DNA tagged strains that suvvived the treatment were collected and termed as *Rhodotorula* Cerulenin Mutants, RCM1 to RCM12 respectively. Notably, the average α-linolenic acid (ALA, C18:3$\Delta^{9,12,15}$) level, an omega-3 PUFA naturally produced in

R. toruloides, was significantly improved in the 12-member mutant population (Fig. 6A). In particular, RCM6 produced ~3 folds higher ALA levels than WT (Fig. 6A).

Secondly, Nile red, a fluorescent lipid indicator used extensively for intracellular tracking [58] and rapid estimation of intracellular lipid content [59], was used as a selection marker. A T-DNA mutant library with ~10,000 transformants was selected on nile red-containing YPD agar plates. Visual examination under a fluorescent dissecting microscope (Additional file 4: Fig. S2B) yielded 4 mutants exhibiting stronger fluorescence intensities. The mutants were named *Rhodotorula* Nile red Mutants, RNM1 to RNM4. RNMs had significantly higher lipid contents than WT although they showed little differences in fatty acid profiles (Fig. 6B).. Interestingly, RNM1 had its T-DNA inserted into the 3' UTR region of a putative omega-3 fatty acid desaturase gene (Table 2), resulting in a 47% increase in lipid accumulation (Fig. 6B).

Thirdly, tetrazolium violet, a redox indicator that gives colonies a distinct violet color if the cells accumulate lipids [60, 61], was tested as a selection marker for lipogenic mutants. As expected, supplementation of 10 μg/mL of tetrazolium violet in selection medium resulted in pigmented transformants (Additional file 4: Fig. S2C). Screening of ~3,000 transformants yielded 6 mutants with deeper pigmentation. The strains were named *Rhodotorula* Tetrazolium violet Mutants, RTM1 to RTM6 respectively. Again, RTMs had higher lipid content than WT (Fig. 6C), with little changes in the fatty acid compositions (data not shown).

The T-DNA insertion sites were identified by hiTAIL PCR in 19 of the 22 mutants (Table 2). The distribution of T-DNA insertion sites appeared similar to previous results (Table 1). All T-DNAs were integrated within gene coding and regulatory regions, ranging from the 1.0 kb upstream to the 0.3 kb downstream of the

coding sequence (Table 2). The affected gene products of RCMs showed high correlation with the lipogenic bioprocess (Table 2). Similarily, the affected gene products of RNMs and RTMs were predicted to be involved in the metabolism of lipids (RNM1, RTM5), amino acids (RNM3), energy (RTM2), signal transduction (RTM1) and transcription (RTM3, 4 and 6).

Direct identification of carotenoid production mutants

To discover novel genes that are involved in the regulation of carotenoid biosynthesis, we designed a simple screening strategy based on the changes of colony color (Fig. 5B). From a population of ~20,000 T-DNA tagged mutants, 1 yellowish and 4 albino transformants were found and named *Rhodotorula* Albino Mutants, RAM1 to RAM5 respectively (Fig. 7A and Additional file 4: Fig. S2D). Sequence analysis revealed that T-DNA was inserted into the DNA sequence encoding a putative riboflavin transporter, resolvase, hexose transporter, TATA-binding protein associated factor and phytoene desaturase, respectively (Table 2). These data suggest that new factors for carotenoid biosynthesis could be identified through the IM approach.

Validation of RAM5

To validate the gene tagging strategies used, the albino mutant RAM5 was analyzed in detail (Fig. 7A). BLAST search of the hiTAIL PCR product showed that the T-DNA was inserted between nucleotide 391802 and 391803 in scaffold No.18 (AEVR02000018), disrupting the phytoene desaturase gene (*CAR1*, genome locus RTG_00274) at the 3rd exon. To validate the result, *CAR1* was deleted in WT through homologous recombination, which led to the replacement of the genome sequence between +948 and +2097 (from the translational start of *CAR1*) by the hygromycin resistance gene cassette (P_{GPD1}*::hpt-3::*T*nos*, Fig. 7B). Indeed, the resulting *car1Δ*

mutant, which was confirmed by Southern blot analysis (Fig. 7C), showed similar creamy color as the T-DNA tagged mutant, RAM5 (Fig. 7D). Furthermore, re-introduction of the wild-type *CAR1* sequence (-662 to +2928, Fig. 7B) into the genome of *car1Δ* (the resulting mutant *car1C*) fully restored the cell color (Fig. 7D). HPLC analysis of carotenoids showed that the main carotenoid species produced in WT *R. toruloides*, such as torulene, torularhodin, γ-carotene and β-carotene [62-64], were totally absent in *car1Δ*, whereas the production was fully restored in *car1C* (Fig. 7E and 7F). qRT-PCR analysis confirmed that the transcripts of *CAR1* were undetectable in *car1Δ* and restored in *car1C* (Fig. 7G). These data confirmed that *CAR1* encodes a key enzyme in carotenoid biosynthesis. Thus, the successful identification of *CAR1* further demonstrates that gene identification and strain improvement strategy based on T-DNA insertional mutagenesis is effective and reliable.

Discussion

R. toruloides is a rare yeast species with highly efficient oil and carotenoid production capacity. However, its potential as an industrial host remains largely unexploited, in part because of the lag in the development of genetic tools. In this study, we report comprehensive studies on factors affecting ATMT efficiency, complementing our previous report on the transformation method for this yeast [16]. Amongst the surprises were the choice of membrane, pH value and agar concentration used for co-culture. Even a minor change of the co-culture pH could be fatal for transformation (Fig. 1-2). Therefore, it is advisable to optimize medium pH, co-culture time and donor/recipient ratio when a new strain or taxa of yeast is used for ATMT. As the

Um*gpd*::*hpt-3* selection cassette has also been used successfully in the transformation of *U. maydis* and *U. scitaminea* [16, 65], it should be broadly useful for dominant selection in both *Ustilaginomycotina* and *Puicciniomycotina* subphyla.

In addition, it was feasible to perform ATMT without the use of supporting membrane for co-culture. The transformed cells could be transferred to selection plates by spreading (as was used herein), "wash and plate", replica printing or medium over-lay. Avoiding the use of membranes could also be appealing to researchers in under-developed countries.

Four screening strategies had been tested for direct identification of genes of interest, each with a specific focus. It is encouraging that mutants can be identified in all cases, leading to the discovery of 27 mutants in total. Importantly, many of the genes appeared to be consistent with their expected roles. For example, riboflavin transporter is involved in the uptake of riboflavin (vitamin B2) and flavin adenine dinucleotide (FAD), which are co-factors for many biocatalytic reactions [66]. Hexose transporter is involved in the uptake of monosaccharides, which is regarded as the first and rate-limiting step of glucose metabolism [67]. Most notably, RCM6 mutant, which is inserted into a putative aldehyde dehydrogenase gene produced ~3 folds higher ALA than WT (Fig. 6A). As a proof of concept, one of the mutants (RAM5) was validated by reverse genetics and the results were in full agreement with the prediction that the gene was involved in the biosynthesis of carotenoids. The latter has been confirmed by another laboratory recently [68]. A full characterization of the 27 mutants is expected to yield valuable information on novel strategies to improve lipid and carotenoid production in this yeast.

Single-copy integration rate, the percentage of transformants with a single copy of transgene integrated into the genome, is an important parameter for IM studies

although it is not fully understood how this is controlled. This is particularly relevant for gene tagging work. The non-homologous end-joining (NHEJ) DNA recombination pathway is believed essential for ATMT [69]. Previously, ATMT of *U. maydis* using the same method yielded a single-copy integration rate of 96% [39]. A significantly lower single-copy integration rate of 75% was observed in *R. toruloides*. This could be attributed to the weaker activity of the *U. maydis gpd1* promoter that was used to drive the expression of hygromycin resistance gene in *R. toruloides* [16]. Indeed, single-copy integration rate was almost 100% when it was replaced with the endogenous Rt*GPD1* promoter (our unpublished data). The average T-DNA copy number (1.36) in *Rhodotorula* species remains lower than in plants, such as 1.5 in *Arabidopsis* [47] and 1.76-2.0 in rice [70, 71]. Furthermore, multiple-copy T-DNA integrations could result in the higher expression of target protein. This feature could be exploited for metabolic engineering in this yeast.

Conclusions

We have established a trackable and reliable mutagenesis method for *R. toruloides* using T-DNA as the mutagen. This method will be valuable for gene discovery as well as strain improvement in *Pucciniomycotina* subphylum and beyond. The 27 mutants identified in this study should yield significant novel information on the lipid and carotenoid biosynthetic pathways.

Materials and Methods

Strains, chemicals, media and culture conditions

R. toruloides strain ATCC 10657, ATCC 10788, ATCC 204091 and *R. glutinis* strain ATCC 90781 were obtained from ATCC (USA). *R. graminis* strain WP1 and

Sporobolomyces roseus FGSC 10293 (IAM13481) were obtained from Fungal Genetics Stock Center, University of Missouri, USA. *R. toruloides* strain ku70e [26] and *Agrobacterium tumefaciens* strain AGL1 [72] have been described previously.

Hygromycin B was purchased from Roche Diagnostics (USA). Nylon N and N+ membranes (Φ 82 mm, 0.45 μm) were from GE Healthcare (Uppsala, Sweden), cellulose acetate membrane (47 mm, Φ0.45 μm) from Grace (Deerfield, IL, USA), cellulose nitrate (87 mm, Φ0.45 μm) from Schleicher & Schuell (Dassel, Germany) and filter paper (Grade 4, Φ90 mm, 20-25 μm in thickness) from Whatman (USA). Cerulenin was obtained from Sigma-Aldrich (USA) and prepared as a 5 mg/mL stock in DMSO. All other chemicals were obtained from Sigma-Aldrich.

Rhodotorula strains were cultured at 28°C in YPD broth (1% yeast extract, 2% peptone, 2% glucose) or on solid potato-dextrose agar (PDA). *A. tumefaciens* was grown at 28°C in either liquid or solid 2YT medium (1.6% tryptone, 1% yeast extract, 0.5% NaCl). Carotenoid production medium B2001 was prepared as described previously [73]. It contains (per litre) 46 g glucose, 11.74 g yeast extract, 2 g K_2HPO_4, 2 g $KH2PO_4$, 0.1 g $MgSO_4 \cdot 7H_2O$, 18 g threonoine, 10 mL trace element (TE) solution, pH6.0. TE solution (per litre) contains 4.0 g $CaCl_2 \cdot 2H_2O$, 0.55 g $FeSO_4 \cdot 7H_2O$, 0.52 g citric acid$\cdot H_2O$, 0.1 g $ZnSO_4 \cdot 7H_2O$, 0.076 g $MnSO_4 \cdot H_2O$ and 0.1 mL smoked H_2SO_4 [74]. Lipid production medium MinLG was prepared as previously described [75] with some modification. It contains (per litre) 30 g glucose, 1.5 g yeast extract, 0.5 g $(NH_4)_2SO_4$, 2.05 g K_2HPO_4, 1.45 g KH_2PO_4, 0.6 g $MgSO_4 \cdot 7H_2O$, 0.3 g NaCl, 10 mg $CaCl_2$, 1 mg $FeSO_4$, 0.5 mg $ZnSO_4$, 0.5 mg $CuSO_4$, 0.5 mg H_3BO_4, 0.5 mg $MnSO_4$ and 0.5 mg $NaMoO_4$, pH 6.1. A seed culture in YPD broth was inoculated in medium B2001 and MinLG to initiate the production of

carotenoids and lipids, respectively, and continued culturing at 28°C for 4 days with constant agitation (250 rpm).

DNA constructs

Oligonucleotides used are listed in Additional file 5: Table S3. All restriction and modification enzymes were from New England Biolabs (NEB, Massachusetts, USA). Binary vectors pEX2 [16] are pPZP200 derivatives used for dominant selection against hygromycin B.

Various promoters, such as promoter of *Ashbya gossypii* translational elongation factor 1-α gene (P_{tef}, 348 bp) [76], *Ustilago maydis gpd1* (P_{gpd}, 595 bp in length) [39, 77], *Aspergillus nidulans gpdA* (P_{gpdA}, 884 bp) [78] and *R. toruloides GPD1* (P_{GPD1}, 1429 bp) [16], were amplified from plasmid pTHPR1 [39], genomic DNA of *U. maydis*, *A. nidulans* and *R. toruloides*, respectively. The primer pair used for the amplication of P_{gpd}, P_{gpdA}, P_{tef} and P_{GPD1} were Pgap-Sf/Pgap-Nr, PgpdA-Sf/PgpdA-Nr, Ptef-Sf/Ptef-Nr and Rt011S/Rt012N, respectively. The resulting DNA fragments, P_{gpd}, P_{gpdA}, P_{tef} and P_{GPD1}, were double-digested with SpeI and NcoI and used in a 3-fragment ligation with the 1030-bp BspHI/SmaI DNA fragment of the synthetic *hpt-3* gene casette [16] and the 8855-bp SpeI/SacI (blunt-ended) DNA fragment of pEC3GPD-GUS (Additional file 6: Fig. S3A) to create pEC3UmGPD-HPT3, pEC3GPDA-HPT3, pEC3TEFA-HPT3 and pEC3GPD1-HPT3, respectively (Additional file 6: Fig. S3B).

To delete *CAR1*, the genome sequence ranging from -89 to +2,928 from the translational start of *CAR1* (AEVR02000018) were amplified using oligos Rt127-2 and Rt128-2. The resultant blunt-ended PCR product was ligated to PmeI and SacI (blunt-end) digested pEX2 to create the intemediate vector pEX2CAR1. The partial coding sequence of *CAR1* in pEX2CAR1 was digested by SacII and MfeI (both blunt-

ended) and replaced with the hygromycin resistance cassette P_{GPD1}*::hpt-3::*T*nos* [26] to create the gene deletion plasmid pKOCAR1. To make the complementation plasmid pRHCAR1C, the genomic sequence of *CAR1* ranging from -662 to +2928 from the translational start was amplified using oligos Rt319Sf and Rt128-2, 5'-hydroxyl termini phosphorylated with T4 polynucleotide kinase, digested with SpeI, and inserted to the SpeI and EcoRV sites of pRH201.

Nucleic acid preparations and manipulations

Genomic DNA was isolated using MasterPure Yeast DNA Purification Kit (Epicentre Biotechnologies, Madison, WI, USA). Total RNA was prepared using the RiboPure RNA Purification Kit (ThermoFisher Scientific, Austin, TX, USA). The resulting nucleotide acids were qualified and quantified by agarose gel electrophoresis and NanoDrop® ND-1000 Spectrophotometer (Nanodrop Technologies, USA), respectively.

Agrobacterium tumefaciens-mediated transformation

Fungi transformation via ATMT was performed as described previously unless indicated otherwise [16]. For membrane-free ATMT, *Agrobacterium* and fungi cells were mixed and spread on the surface of IM agar without any supporting membranes. After co-culture at 24^{o}C for 2 days, cells were scrapped out using a L-shape spreader and plated on the surface of YPD agar supplemented with appropriate antibiotics as described previously [16], and incubated at 30^{o}C until the appearance of transformants.

Southern blot analysis

Genomic DNA was digested with PstI and separated by electrophoresis on 0.8% agarose gels. DIG-labeled probe of the partial *hpt-3* gene (+375 through +1036) was

amplified using oligos HptRU and HptRSL2 (Additional file 5: Table S3). To verify *CAR1* deletion mutants, genomic DNA was digested with HindIII and the digoxigenin-labeled *CAR1R* fragment (Fig. 7B) was used as the probe. Southern hybridization was carried out according to the manufacture's instructions (DIG-High prime DNA labeling and detection starter Kit II, Roche Diagnostics).

Identification of T-DNA tagging positions

T-DNA tagging positions in the genome were identified using hiTAIL PCR. Specific primers (HRSP1, HRSP2 and HRSP3) and arbitrary primers (LAD1-1 and LAD1-4) were used for LB flanking sequences. Specific primers (HRRSP1, HRRSP2 and HRRSP3) and arbitrary primers (LAD1-1 and LAD1-4) for RB flanking sequences. PCR reactions were carried out with i-Taq DNA polymerase (i-DNA Biotech, Singapore) in a PTC-200^{TM} Programmable Thermal Controller (Bio-Rad, USA). PCR products were purified using gel extraction kit (Qiagen, CA, USA) and sequenced directly using BigDye terminator kit (Applied Biosystems, USA) with oligo HRSP3 (for LB) or HRRSP3 (for RB). In some cases, PCR products were cloned into pGEM-T easy vector (Promega, USA) by TA cloning technique and sequenced using oligos M13FP and M13RP.

Quantitative reverse transcription PCR (qRT-PCR)

qRT-PCR was performed in triplicates as described [21]. Relative gene expression levels were calculated against the reference gene *ACT1* (actin encoding gene, GenBank acc. no. KR138696) [17] using the $2^{-\Delta\Delta Ct}$ method (RQ Manager software v1.2.1, Applied Biosystems). Oligonucleotide pair used for *CAR1* and *ACT1* was qCAR1f/qCAR1r and qACT1f/qACT1r, respectively.

Screening for lipid and carotenoid production mutants

The genome of *R. toruloides* was mutagenized by random insertion of T-DNA of plasmid pRH201. Candidate lipid production mutants were selected by supplementation of various chemicals, such as 50 μg/mL cerulenin, 0.5 μg/mL nile red or 10 μg/mL tetrazolium violet. After incubation at 28°C for 5 days, transformants that survived cerulenin treatment (Additional file 4: Fig. S2A); showing higher fluorescent intensity on nile red-containing media or darker purple-color pigmentation (Additional file 4: Fig. S2B) on tetrazolium violet-containing media (Additional file 4: Fig. S2C), were transferred to YPD broth (300 μg/mL cefotaxime and 150 μg/mL hygromycin) for propagation and cryopreservation. Candidate carotenoid production mutants were selected by visual screening of transformants.

Extraction of lipids and carotenoids

Total lipid was extracted essentially as described previously [79]. Dry cell biomass (10 mg) was mixed with 500 μl of 4 M HCl and incubated in a boiling water bath for 15 min. Subsequently, samples were placed in a -20°C freezer for at least 1 h and the cell lysate was mixed with 1.0 mg pentadecanoic acid (C15:0, the internal standard for the subsequent GC analysis) and 1.0 mL chloroform:methanol (2:1, v/v). After centrifugation, the lower solvent phase was transferred to a new tube. The total lipid mass was determined by weighing after drying in a vacuum concentrator (Eppendorf, USA).

Carotenoids were essentially extracted as described previously [80]. Samples were kept away from light as much as possible during the extraction. Cells from 50 mL cultures was pelleted by centrifugation and washed twice with water. Equal amount of acid-washed glass beads (0.4-0.6 mm in dimeter, Sigma-Aldirch) and 5 mL

DMSO were added and mixed vigorously for 10 min by vortexing. Samples were incubated for 1 h at 65^{o}C and then freezed at -20^{o}C. Supernatant was removed to a new tube after centrifugation (DMSO-dissolved carotenoids). The pellet was mixed with 30 mL of light petrolium ether-ethyl acetate (2:1, v/v) for 10 min by vigorous vortexing. After centrifugation, the supernatant (solvent-dissolved carotenoids) was combined with previous DMSO-dissolved carotenoids and then 2 mL of saturated NaCl solution was added. The upper solvent phase was separated and blow-dried with nitrogen gas and carotenoids was re-dissolved in hexane.

Quantification methods

Cell biomass (dry cell weight) was determined by lyophilizing the cell pellet until constant weight was reached.

Residual glucose was quantified by HPLC in a Prominence ultra fast liquid chromatography (UFLC) system (Shimadzu, Kyoto, Japan). Culture was filtered through a 0.2 μm nylon membrane and run through a 300 x 7.0 mm Aminex 87H column (Bio-Rad) at a constant flow rate of 0.7 mL/min using 5 mM sulfuric acid as the mobile phase. The column was maintained at 50^{o}C and glucose was detected with a Refractive Index Detector (RID, Shimadzu). Concentration of glucose in the cell culture was determined using calibration curves built with the standard glucose aqueous solution.

Quick estimation of lipid content using nile red was performed as described previously [59]. Briefly, 10 μl cell culture and 2 μl nile red stock (50 mM in acetone) were mixed with 200 μl PBS buffer (pH7.4) in a well of a FluoroNunc plate (Thermo Fisher Scientific, Langenselbold, Germany). Each sample was accompanied with a nile red-free well as the background control. Another fraction of the cell culture (10 μl) was mixed with 90 μl PBS buffer (pH7.4) in a 96-well flat-bottom transparent

plate (Nunc, Roskilde, Denmark) to measure cell optical density. The data was acquired and analyzed using the Infinite M200 plate reader (Tecan, Salzburg, Austria) using the iControl^{TM} version 3.0 software (Tecan, Salzburg, Austria). Cell optical density was read at 600 nm after subtracting background while fluorescence intensity was measured with an excitation and emission wavelength of 488 nm and 508 nm, respectively. The relative lipid content is calculated by normalization against its absorptance at 600 nm after deducting the background control. Statistical triplicates were used for all tests.

Fatty acid profiles were determined by gas chromatography mass spectrometry (GCMS). Preparation of fatty acid methyl esters (FAMEs) and subsequent GCMS analysis were performed as described previously [81]. Briefly, lipids were dissolved in 300 µl petroleum ether-benzene (1:1, v/v), mixed with equal volume of methanolic hydrochloride acid (3 M, Sigma) and kept at 80^{o}C for 1 h. FAMEs were extracted with 1 mL of hexane, 1 µl of which was injected to a HP-88 fused silica capillary column (30-m length, 0.25-µm diameter, and 0.25-mm film thickness, Agilent J&W Scientific, Folsom, CA, USA) fitted in a GCMS (QP2010 Ultra, Shimadzu). The running conditions were typically 42.3 mL/min nitrogen flow, 150^{o}C for starting temperature (3 min), a 15-min ramp to 240^{o}C, and holding at 240^{o}C for 7 min. The FAME peaks were identified by searching against Shimadzu NIST08 compound library and quantified as percentages of total fatty acids (%TFA).

The total carotenoid concentration was estimated by spectrophotometry method [82]. Briefly, the total carotenoids were dissolved in 1 mL petroleum ester and the absorbance was measured at 485 nm (A_{485}) and calculated using the following formula: Carotenoids (mg/L) = A_{485} x 1000/2680, where the coefficient of absorbance used was that equivalent to β-carotene: $E^{1\%}_{1\ cm}$ = 2680 for petroleum ether. The carotenoid

profiles were analyzed by HPLC (Shimadzu Prominence UFLC system coupled with photodiode array detector) as previously described [83]. Briefly, carotenoids were filtered through a 0.2 μm nylon membrane and separated through the Kinetex C18 reverse phase column (100 x 3 mm, ϕ 2.6 μm, Phenomenex Inc., CA, USA) at a constant temperature of 35ºC. The mobile phase (acetonitrile:methanol containing 0.1 M ammonium acetate:dichloromethane = 71:22:7, v/v/v) was run at a constant flow rate of 0.3 mL/min. Various carotenoid compositions were quantified using β-carotene (C-4582, Sigma-Aldrich) as the external standard.

Abbreviations (in alphabetical order)

ATCC: American Type Culture Collection, USA

BLAST: Basic Local Alignment Search Tool (National Library of Medicine, National Institutes of Health, USA)

DCW: dry cell weight

CAR1: phytoene desaturase gene

GPD1: glyceraldehyde 3-phosphate dehydrogenase gene

hpt-3: a synthetic *E. coli* hygromycin B phosphotransferase gene optimized according to the codon bias of *R. toruloides*

IM: insertional mutagenesis

RACE: rapid amplification of cDNA ends

PUFA: polyunsaturated fatty acid

TFE: transformation efficiency

UTR: untranslated region

WT: wild-type strain

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

Consent for publication

All authors consent for publication.

Competing interests

The authors declare that they have no competing interests. Temasek Life Sciences Laboratory has an interest in developing *Rhodotorula toruloides* as an industrial biotechnology platform.

Funding

This work was financially supported by Singapore National Research Fundation, Competitive Research Program CRP-8-2011-02; Singapore Economic Development Board and Temasek Foundation.

Authors' contributions

YL and LJ conceived and designed the experiments and drafted the manuscript. YL analyzed the data. YL, CMJK, SAY, MD and MMH carried out the experiments,

contributed the regents and materials. All authors read and approved the final manuscript.

References

1. Ratledge C: **Regulation of lipid accumulation in oleaginous micro-organisms**. *Biochem Soc Trans* 2002, **30**(Pt 6):1047-1050.
2. Turcotte G, Kosaric N: **Lipid biosynthesis in oleaginous yeasts**. In: *Bioprocesses and Engineering.* vol. 40: Springer Berlin / Heidelberg; 1989: 73-92.
3. Ratledge C, Wynn JP: **The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms**. *Adv Appl Microbiol* 2002, **51**:1-51.
4. Rossi M, Amaretti A, Raimondi S, Leonardi A: **Getting Lipids for Biodiesel Production from Oleaginous Fungi**. In: *Biodiesel: Feedstocks and Processing Technologies.* Edited by Stoytcheva M, Montero G: InTech; 2011: 71-92.
5. Papanikolaou S, Aggelis G: **Lipids of oleaginous yeasts. Part II: Technology and potential applications**. *European Journal of Lipid Science and Technology* 2011, **113**(8):1052-1073.
6. Gong Z, Shen H, Q. W, Yang X, Xie H, Zhao ZK: **Efficient conversion of biomass into lipids by using the simultaneous saccharification and enhanced lipid production process**. *Biotechnol Biofuels* 2013, **6**(1):1-12.
7. Meng X, Yang J, Xu X, Zhang L, Nie Q, Xian M: **Biodiesel production from oleaginous microorganisms**. *Renew Energy* 2009, **34**(1):1-5.
8. Vachali P, Bhosale P, Bernstein PS: **Microbial carotenoids**. In: *Microbial carotenoids from fungi: methods and protocols.* Edited by Barredo JL, vol. 898. New York: Springer Science+Business Media; 2012: 41-59.
9. Wang QM, Yurkov AM, Goker M, Lumbsch HT, Leavitt SD, Groenewald M, Theelen B, Liu XZ, Boekhout T, Bai FY: **Phylogenetic classification of yeasts and related taxa within *Pucciniomycotina***. *Stud Mycol* 2015, **81**:149-189.
10. Li Y, Zhao ZK, Bai F: **High-density cultivation of oleaginous yeast *Rhodosporidium toruloides* Y4 in fed-batch culture**. *Enzyme Microb Technol* 2007, **41**(3):312-317.
11. Zhao X, Hu C, Wu S, Shen H, Zhao ZK: **Lipid production by *Rhodosporidium toruloides* Y4 using different substrate feeding strategies**. *J Ind Microbiol Biotechnol* 2011, **38**(5):627-632.
12. Pan JG, Kwak MY, Rhee JS: **High density cell culture of *Rhodotorula glutinis* using oxygen-enriched air**. *Biotechnol Lett* 1986, **8**(10):715-718.
13. Yu X, Zheng Y, Dorgan KM, Chen S: **Oil production by oleaginous yeasts using the hydrolysate from pretreatment of wheat straw with dilute sulfuric acid**. *Bioresour Technol* 2011, **102**(10):6134-6140.
14. Liu H, Zhao X, Wang F, Jiang X, Zhang S, Ye M, Zhao ZK, Zou H: **The proteome analysis of oleaginous yeast *Lipomyces starkeyi***. *FEMS Yeast Res* 2011, **11**(1):42-51.
15. Cheirsilp B, Louhasakul Y: **Industrial wastes as a promising renewable source for production of microbial lipid and direct transesterification of the lipid into biodiesel**. *Bioresour Technol* 2013, **142**:329-337.
16. Liu Y, Koh CM, Sun L, Hlaing MM, Du M, Peng N, Ji L: **Characterization of glyceraldehyde-3-phosphate dehydrogenase gene Rt*GPD1* and development of genetic transformation method by dominant selection in**

oleaginous yeast *Rhodosporidium toruloides*. *Appl Microbiol Biotechnol* 2013, **97**(2):719-729.

17. Liu Y, Koh CMJ, Ngoh ST, Ji L: **Engineering an efficient and tight D-amino acid-inducible gene expression system in *Rhodosporidium/Rhodotorula species***. *Microb Cell Fact* 2015, **14**(1):170-185.

18. Lin X, Wang Y, Zhang S, Zhu Z, Zhou YJ, Yang F, Sun W, Wang X, Zhao ZK: **Functional integration of multiple genes into the genome of the oleaginous yeast *Rhodosporidium toruloides***. *FEMS Yeast Res* 2014, **14**(4):547-555.

19. Wang Y, Lin X, Zhang S, Sun W, Ma S, Zhao ZK: **Cloning and evaluation of different constitutive promoters in the oleaginous yeast *Rhodosporidium toruloides***. *Yeast* 2016, **33**(3):99-106.

20. Johns AM, Love J, Aves SJ: **Four inducible promoters for controlled gene expression in the oleaginous yeast *Rhodotorula toruloides***. *Front Microbiol* 2016, **7**:1666.

21. Liu Y, Yap SA, Koh CM, Ji L: **Developing a set of strong intronic promoters for robust metabolic engineering in oleaginous *Rhodotorula* (*Rhodosporidium*) yeast species**. *Microb Cell Fact* 2016, **15**(1):200.

22. Zhang S, Skerker JM, Rutter CD, Maurer MJ, Arkin AP, Rao CV: **Engineering *Rhodosporidium toruloides* for increased lipid production**. *Biotechnol Bioeng* 2016, **113**(5):1056-1066.

23. Zhang S, Ito M, Skerker JM, Arkin AP, Rao CV: **Metabolic engineering of the oleaginous yeast *Rhodosporidium toruloides* IFO0880 for lipid overproduction during high-density fermentation**. *Appl Microbiol Biotechnol* 2016.

24. Wang Y, Zhang S, Potter M, Sun W, Li L, Yang X, Jiao X, Zhao ZK: **Overexpression of Δ12-fatty acid desaturase in the oleaginous yeast *Rhodosporidium toruloides* for production of linoleic acid-rich lipids**. *Appl Biochem Biotechnol* 2016, **180**(8):1497-1507.

25. Lee JJ, Chen L, Cao B, Chen WN: **Engineering *Rhodosporidium toruloides* with a membrane transporter facilitates production and separation of carotenoids and lipids in a bi-phasic culture**. *Appl Microbiol Biotechnol* 2016, **100**(2):869-877.

26. Koh CM, Liu Y, Du M, Ji L: **Molecular characterization of *KU70* and *KU80* homologues and exploitation of a *KU70*-deficient mutant for improving gene deletion frequency in *Rhodosporidium toruloides***. *BMC Microbiol* 2014, **14**(1):50.

27. Conrad TM, Lewis NE, Palsson BO: **Microbial laboratory evolution in the era of genome-scale science**. *Mol Syst Biol* 2011, **7**:509.

28. Brenner S: **The genetics of *Caenorhabditis elegans***. *Genetics* 1974, **77**(1):71-94.

29. Schneeberger K: **Using next-generation sequencing to isolate mutant genes from forward genetic screens**. *Nat Rev Genet* 2014, **15**(10):662-676.

30. Amsterdam A, Burgess S, Golling G, Chen W, Sun Z, Townsend K, Farrington S, Haldi M, Hopkins N: **A large-scale insertional mutagenesis screen in zebrafish**. *Genes Dev* 1999, **13**(20):2713-2724.

31. Parinov S, Sundaresan V: **Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project**. *Curr Opin Biotechnol* 2000, **11**(2):157-161.

32. Gragerov A, Horie K, Pavlova M, Madisen L, Zeng H, Gragerova G, Rhode A, Dolka I, Roth P, Ebbert A *et al*: **Large-scale, saturating insertional mutagenesis of the mouse genome**. *Proc Natl Acad Sci U S A* 2007, **104**(36):14406-14411.
33. Dent RM, Sharifi MN, Malnoe A, Haglund C, Calderon RH, Wakao S, Niyogi KK: **Large-scale insertional mutagenesis of *Chlamydomonas* supports phylogenomic functional prediction of photosynthetic genes and analysis of classical acetate-requiring mutants**. *Plant J* 2015, **82**(2):337-351.
34. Hensel M, Holden DW: **Molecular genetic approaches for the study of virulence in both pathogenic bacteria and fungi**. *Microbiology* 1996, **142 (Pt 5)**:1049-1058.
35. Brown JS, Holden DW: **Insertional mutagenesis of pathogenic fungi**. *Curr Opin Microbiol* 1998, **1**(4):390-394.
36. Liu Y, Koh CM, Sun L, Ji L: **Tartronate semialdehyde reductase defines a novel rate-limiting step in assimilation and bioconversion of glycerol in *Ustilago maydis***. *PLoS One* 2011, **6**(1):e16438.
37. Liu YG, Chen Y: **High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences**. *Biotechniques* 2007, **43**(5):649-650, 652, 654 passim.
38. Liu YG, Whittier RF: **Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking**. *Genomics* 1995, **25**(3):674-681.
39. Ji L, Jiang ZD, Liu Y, Koh CM, Zhang LH: **A Simplified and efficient method for transformation and gene tagging of *Ustilago maydis* using frozen cells**. *Fungal Genet Biol* 2010, **47**(4):279-287.
40. Koncz C, Martini N, Mayerhofer R, Koncz-Kalman Z, Korber H, Redei GP, Schell J: **High-frequency T-DNA-mediated gene tagging in plants**. *Proc Natl Acad Sci U S A* 1989, **86**(21):8467-8471.
41. Ross-Macdonald P, Coelho PS, Roemer T, Agarwal S, Kumar A, Jansen R, Cheung KH, Sheehan A, Symoniatis D, Umansky L *et al*: **Large-scale analysis of the yeast genome by transposon tagging and gene disruption**. *Nature* 1999, **402**(6760):413-418.
42. Frandsen RJ: **A guide to binary vectors and strategies for targeted genome modification in fungi using *Agrobacterium tumefaciens*-mediated transformation**. *J Microbiol Methods* 2011, **87**(3):247-262.
43. Ziemienowicz A, Tzfira T, Hohn B: **Mechanisms of T-DNA integration *Agrobacterium*: From Biology to Biotechnology**. In. Edited by Tzfira T, Citovsky V: Springer New York; 2008: 395-440.
44. Michielse CB, Hooykaas PJ, van den Hondel CA, Ram AF: ***Agrobacterium*-mediated transformation as a tool for functional genomics in fungi**. *Curr Genet* 2005, **48**(1):1-17.
45. Donofrio NM, Oh Y, Lundy R, Pan H, Brown DE, Jeong JS, Coughlan S, Mitchell TK, Dean RA: **Global gene expression during nitrogen starvation in the rice blast fungus, *Magnaporthe grisea***. *Fungal Genet Biol* 2006, **43**(9):605-617.
46. Johnston GC, Singer RA, McFarlane S: **Growth and cell division during nitrogen starvation of the yeast *Saccharomyces cerevisiae***. *J Bacteriol* 1977, **132**(2):723-730.

47. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R *et al*: **Genome-wide insertional mutagenesis of *Arabidopsis thaliana***. *Science* 2003, **301**(5633):653-657.
48. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJJ: **Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae***. *Embo J* 1995, **14**(13):3206-3214.
49. Choi J, Park J, Jeon J, Chi MH, Goh J, Yoo SY, Park J, Jung K, Kim H, Park SY *et al*: **Genome-wide analysis of T-DNA integration into the chromosomes of *Magnaporthe oryzae***. *Mol Microbiol* 2007, **66**(2):371-382.
50. Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Crameri A, Angelis K, Redei GP, Schell J, Hohn B, Koncz C: **T-DNA integration: a mode of illegitimate recombination in plants**. *Embo J* 1991, **10**(3):697-704.
51. Kunik T, Tzfira T, Kapulnik Y, Gafni Y, Dingwall C, Citovsky V: **Genetic transformation of HeLa cells by *Agrobacterium***. *Proc Natl Acad Sci U S A* 2001, **98**(4):1871-1876.
52. Omura S: **The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis**. *Bacteriol Rev* 1976, **40**(3):681-697.
53. Vance D, Goldberg I, Mitsuhashi O, Bloch K: **Inhibition of fatty acid synthetases by the antibiotic cerulenin**. *Biochem Biophys Res Commun* 1972, **48**(3):649-656.
54. Nguyen LN, Nosanchuk JD: **The inhibitory effect of cerulenin to yeasts is fungicidal**. *Commun Integr Biol* 2011, **4**(6):631-632.
55. Morita N, Nishida T, Tanaka M, Yano Y, Okuyama H: **Enhancement of polyunsaturated fatty acid production by cerulenin treatment in polyunsaturated fatty acid-producing bacteria**. *Biotechnol Lett* 2005, **27**(6):389-393.
56. Wang J, Li R, Lu D, Ma S, Yan Y, Li W: **A quick isolation method for mutants with high lipid yield in oleaginous yeast**. *World J Microbiol Biotechnol* 2009, **25**(5):921-925.
57. Zhong Y, Wang X, Yu H, Liang S, Wang T: **Application of T-DNA insertional mutagenesis for improving cellulase production in the filamentous fungus *Trichoderma reesei***. *Bioresour Technol* 2012, **110**:572-577.
58. Greenspan P, Mayer EP, Fowler SD: **Nile red: a selective fluorescent stain for intracellular lipid droplets**. *J Cell Biol* 1985, **100**(3):965-973.
59. Kimura K, Yamaoka M, Kamisaka Y: **Rapid estimation of lipids in oleaginous fungi and yeasts using Nile red fluorescence**. *J Microbiol Methods* 2004, **56**(3):331-338.
60. Spiekermann P, Rehm BH, Kalscheuer R, Baumeister D, Steinbuchel A: **A sensitive, viable-colony staining method using Nile red for direct screening of bacteria that accumulate polyhydroxyalkanoic acids and other lipid storage compounds**. *Arch Microbiol* 1999, **171**(2):73-80.
61. Zhao YF, Kong QZ: **Tetrazolium violet inhibits cell growth and induces cell death in C127 mouse breast tumor cells**. *Chem Biol Interact* 2008, **174**(1):19-26.
62. Peterson WJ, Evans WR, Lecce E, Bell TA, Etchells JL: **Quantitative determination of the carotenoids in yeasts of the genus *Rhodotorula***. *J Bacteriol* 1958, **75**(5):586-591.
63. Perrier V, Dubreucq E, Galzy P: **Fatty acid and carotenoid composition of *Rhodotorula* strains**. *Arch Microbiol* 1995, **164**(3):173-179.

64. Buzzini P, Innocenti M, Turchetti B, Libkind D, van Broock M, Mulinacci N: **Carotenoid profiles of yeasts belonging to the genera *Rhodotorula*, *Rhodosporidium*, *Sporobolomyces*, and *Sporidiobolus***. *Can J Microbiol* 2007, **53**(8):1024-1031.
65. Sun L, Yan M, Ding Z, Liu Y, Du M, Xi P, Liao J, Ji L, Jiang Z: **Improved dominant selection markers and co-culturing conditions for efficient *Agrobacterium tumefaciens*-mediated transformation of *Ustilago scitaminea***. *Biotechnol Lett* 2014, **36**(6):1309-1314.
66. Reihl P, Stolz J: **The monocarboxylate transporter homolog Mch5p catalyzes riboflavin (vitamin B2) uptake in *Saccharomyces cerevisiae***. *J Biol Chem* 2005, **280**(48):39809-39817.
67. Ozcan S, Johnston M: **Function and regulation of yeast hexose transporters**. *Microbiol Mol Biol Rev* 1999, **63**(3):554-569.
68. Sun W, Yang X, Wang X, Lin X, Wang Y, Zhang S, Luan Y, Zhao ZK: **Homologous gene targeting of a carotenoids biosynthetic gene in *Rhodosporidium toruloides* by *Agrobacterium*-mediated transformation**. *Biotechnol Lett* 2017.
69. van Attikum H, Bundock P, Hooykaas PJJ: **Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration**. *Embo J* 2001, **20**(22):6550-6558.
70. Sallaud C, Meynard D, van Boxtel J, Gay C, Bes M, Brizard JP, Larmande P, Ortega D, Raynal M, Portefaix M *et al*: **Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics**. *Theor Appl Genet* 2003, **106**(8):1396-1408.
71. Sha Y, Li S, Pei Z, Luo L, Tian Y, He C: **Generation and flanking sequence analysis of a rice T-DNA tagged population**. *Theor Appl Genet* 2004, **108**(2):306-314.
72. Lazo GR, Stein PA, Ludwig RA: **A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium***. *Biotechnology (N Y)* 1991, **9**(10):963-967.
73. Bhosale P, Gadre RV: **Optimization of carotenoid production from hyper-producing *Rhodotorula glutinis* mutant 32 by a factorial approach**. *Lett Appl Microbiol* 2001, **33**(1):12-16.
74. Meesters PAEP, Huijberts GNM, Eggink G: **High-cell-density cultivation of the lipid accumulating yeast *Cryptococcus curvatus* using glycerol as a carbon source**. *Appl Microbiol Biotechnol* 1996, **45**(5):575-579.
75. Wu S, Hu C, Jin G, Zhao X, Zhao ZK: **Phosphate-limitation mediated lipid production by *Rhodosporidium toruloides***. *Bioresour Technol* 2010, **101**(15):6124-6129.
76. Steiner S, Philippsen P: **Sequence and promoter analysis of the highly expressed *TEF* gene of the filamentous fungus *Ashbya gossypii***. *Mol Gen Genet* 1994, **242**(3):263-271.
77. Smith TL, Leong SA: **Isolation and characterization of a *Ustilago maydis* glyceraldehyde-3-phosphate dehydrogenase-encoding gene**. *Gene* 1990, **93**(1):111-117.
78. Punt PJ, Dingemanse MA, Kuyvenhoven A, Soede RD, Pouwels PH, van den Hondel CA: **Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase**. *Gene* 1990, **93**(1):101-109.

79. Elsey D, Jameson D, Raleigh B, Cooney MJ: **Fluorescent measurement of microalgal neutral lipids**. *J Microbiol Methods* 2007, **68**(3):639-642.
80. Saelices L, Youssar L, Holdermann I, Al-Babili S, Avalos J: **Identification of the gene responsible for torulene cleavage in the *Neurospora* carotenoid pathway**. *Mol Genet Genomics* 2007, **278**(5):527-537.
81. Voelker TA, Davies HM: **Alteration of the specificity and regulation of fatty acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase**. *J Bacteriol* 1994, **176**(23):7320-7327.
82. Machmudah S, Goto M: **Methods for Extraction and Analysis of Carotenoids**. In: *Natural Products.* Edited by Ramawat KG, Mérillon J-M: Springer Berlin Heidelberg; 2013: 3367-3411.
83. Razavi SH, Blanchard F, Marc I: **UV-HPLC/APCI-MS method for separation and identification of the carotenoids produced by *Sporobolomyces ruberrimus* H110**. *Iran J Chem Chem Eng* 2006, **25**.

Figure Captions

Fig. 1. ATMT in *Pucciniomycotina* species. *Agrobacterium* culture harboring binary vector pRH201 [16] was used as the donor in all trials. (A) ATMT without the use of membrane support. (B) Transformation efficiency (TFE) expressed as CFU, colony forming unit per 10^6 transformed fungal cells. Rt1, *R. toruloides* ATCC 10657; Rt2, *R. toruloides* ATCC 10788; Rg1, *R. glutinis* ATCC 90781; Rg2, *R. glutinis* ATCC 204091; Rg3, *R. graminis* WP1; Sr, *S. roseus*.

Fig. 2. Factors affecting transformation efficiency. *R. toruloides* ATCC 10657 was used as the T-DNA recipient. TFE is presented as the relative transformation efficiency, where the TFE using the original standard protocol was set as 100%. Each condition was done with 3 repeats. Error bars represent standard derivations. (A) pH of induction medium (IM). (B) Agar concentration of IM. (C) Ammonium sulfate concentration in IM. (D) Different membrane types. (E) Representative transformation on selected membranes. Membrane-free: transformation conducted without supporting membrane. The induction medium was pH5.5. N: Nylon Hybond N membrane; CA: Cellulose acetate membrane; N^+: Nylon Hybond N^+ membrane; MF, membrane-free.

Fig. 3. Southern blot analysis of *R. toruloides* transformants. Genomic DNA samples of 64 randomly selected transformants of pRH201 and wildtype strain *R. toruloides* ATCC 10657 (5 µg) were digested with PstI and separated by electrophoresis in 0.8% agarose gel. The 581 bp *hpt-3* DNA fragment was amplified using oligos HptRU and HptRSL2 and used the probe (Additional file 5: Table S3). Lane M, DIG-labeled DNA molecular size marker III (Roche Diagnosis, USA).

Fragment sizes (bp): 564, 831, 947, 1375, 1584, 1904, 2027, 3530, 4268, 4973, 5148 and 21226. WT: wild-type strain ATCC 10657.

Fig. 4. Identification of T-DNA tagging sites. (A) Genomic distribution of T-DNAs. hiTAIL PCR product sequences were searched against the genome database of *R. glutinis* ATCC 204091. (B) Nicking positions of the integrated T-DNAs. Histogram of nicking site is presented based on the nicking frequency calculated from 61 LB and 196 RB flanking sequences. T-DNA border repeats were capitalized. Arrows indicate the positions of T-DNA nicking by Agrobacterium. (C) Size distribution of microhomology. Columns show the distribution of microhomology sizes found in this study; Black line shows the expected sizes of microhomology based on calculation. (D) Classification of proteins affected by the 132 T-DNA tags.

Fig. 5. Schematic diagram for gene identification based on IM. (A) Screening for lipid/fatty acid production mutants. CL+, NR+ and TV+ represents the cerulenin (50 μg/mL), nile red (0.5 μg/mL) and tetrazolium violet (10 μg/mL), respectively. (B) Visual screening for carotenoid production mutants. Cef+: cefotaxime (300 μg/mL); Hyg+: hygromycin B (150 μg/mL).

Fig. 6. Characterization of lipogenic mutants. (A) Relative α-Linolenic acid (C18:3$\Delta^{9,12,15}$, ALA) yields of RCM mutants. (B) Relative lipid contents of RNM mutants. (C) Relative lipid contents of RTM mutants. (D) Mutant classification and statistical analysis (Chi-square). 22 lipogenic mutants (RCMs, RNMs and RTMs) were plotted according to their predicted protein functions (Dark bars). The predicted percentage of protein fucntions affected in shown in grey bars (Probability: * P,<0.05; ** P < 0.01).

Fig. 7. Identification of carotenogenic mutants and functional validation of RAM5. (A) Colony color phenotypes of RAMs. All strains were streaked on PDA plate and incubated at 28ºC for 2 days. (B) Schematic diagram of *CAR1* structure and its deletion and complementation strategies. *CAR1* genomic sequence (Dark red line) ranging from -662 to +2928 was used to complement the *car1* mutant. (C) Southern blot hybridization of candidate *CAR1* null mutant (*car1Δ*). (D) Pigment colors of WT, null mutant (*car1Δ*) and complementation strain (*car1C*) in either liquid broth or agar medium. (E) HPLC separation of carotenoids from WT, *car1Δ* and *car1C*. All strains were cultured in carotenoid production media B2001 for 5 days. (F) Quantitative analysis of 4 different carotenoid species (microgram of carotenoid per dry cell weight, μg/g DCW) in WT, *car1Δ* and *car1C* strains. (G) Changes of *CAR1* mRNA levels in WT, *car1Δ*, and *car1C*. All cells were cultured in carotenoid production media B2001 for 1 day before RNA extraction. Actin encoding gene (*ACT1*) was used as the reference.

Tables

Table 1. Distrubution of T-DNA insertion positions

Locations	Tagged loci	Percentage (%)
Upstream 0.5-1.0 kb	13	10.7
Upstream 0.5 kb	18	14.9
Downstream 0.3 kb	5	4.1
Coding sequence	75	62.0
Intergenic sequence	10	8.3
Total	121	100.0

Table 2. T-DNA insertion sites of various mutants

Sequence number[a]	Genic site[b,c]	Best hit[d]	Annotation[e]	*Organism*[f]	Identity[g]
RB sequences					
RCM1	Upstream-0.5 kb	XP_001549261	cellulose-binding GDSL lipase/acylhydrolase	*Botryotinia fuckeliana*	26%
RCM2	Upstream-1.0 kb	GENE ID: 5545759 Kpol_534p16	mannosyltransferase	*Vanderwaltozyma polyspora*	77%
RCM3	Upstream-0.5 kb	XP_001789963	regulator of nonsense transcripts; NADP-dependent isocitrate dehydrogenase	*Bos taurus*	40%
RCM4	Genic sequence	XP_001607008.2	protein bric-a-brac 2 isoform X1	*Nasonia vitripennis*	29%
RCM5	Upstream-0.5 kb	XP_003174510	C6 zinc finger domain-containing protein	*Arthroderma gypseum*	32%
RCM6	Genic sequence	XP_001629556	Fatty aldehyde dehydrogenase	*Nematostella vectensis*	51%
RCM7	Genic sequence	XP_571856	Hexose transport-related protein	*Cryptococcus neoformans*	53%
RCM8	Genic sequence	XP_001731990	Transcription initiation factor TFIID subunit 2	*Malassezia globosa*	40%
RCM9	Genic sequence	NP_001125572	stAR-related lipid transfer protein 3	*Pongo abelii*	29%
RCM10	Upstream-1.0 kb	XP_001645395	GPI mannosyltransferase 3	*Vanderwaltozyma polyspora*	77%
RCM11	Upstream-0.5 kb	XP_662119	Cell wall protein that functions in the transfer of chitin to beta (1-6) glucan	*Aspergillus nidulans*	36%
RCM12	NA[h]	-	-	-	-
RNM1	Downstream 0.3 kb	AAC98967.2	omega-3 fatty acid desaturase	*Vernicia fordii*	41%
RNM2	NA				
RNM3	Upstream 0.5 kb	ZP_03104366	amino acid permease	*Bacillus cereus W*	87%
RNM4	NA				
RTM1	Genic sequence	XP_016273537	SH3 domain-containing protein	*Rhodotorula glutinis NP11*	93%

RTM2	Upstream 0.5 kb	EGU13095.1	salicylate hydroxylase	*Rhodotorula glutinis ATCC 204091*	73%
RTM3	Genic sequence	XP_501740.1	nitrogen assimilation transcription factor	*Yarrowia lipolytica*	71%
RTM4	Upstream 0.5 kb	ZP_08453184.1	putative zinc-binding oxidoreductase	*Streptomyces sp.*	47%
RTM5	Genic sequence	ZP_07628725.1	putative lipoprotein	*Prevotella amnii*	45%
RTM6	Genic sequence	YP_001220603.1	resolvase site-specific recombinase	*Aeromonas bestiarum*	94%
RAM1	Genic sequence	XP_003032296	Riboflavin transporter MCH5	*Schizophyllum commune*	52%
RAM2	Upstream-0.5 kb	YP_001220603	resolvase	*Aeromonas bestiarum*	95%
RAM3	Genic sequence	XP_571856	hexose transport-related protein	*Cryptococcus neoformans*	36%
RAM4	Genic sequence	XP_758766	TATA-binding protein associated factor	*Ustilago maydis*	35%
RAM5	Genic sequence	AHB14354	phytoene synthase	*Rhodosporidium diobovatum*	98%

[a] Flanking sequence obtained from corresponding to number of T-DNA transformant

[b] T-DNA tagged genes were determined according to the BLASTx results

[c] Upstream 1.0 kb, Upstream 0.5 kb and downstream 0.3 kb denotes T-DNA insertions within upstream 501~1000 bp, 500 bp and downstream 300 bp of the corresponding tagged gene, respectively

[d] Best hit denotes the BLASTx result with the highest E-score

[e] Annotations were determined according to the BLASTx results

[f] Microorganism denotes the host of Best hit

[g] Identity values were from BLASTx results

[h] Not available due to the bad sequencing result

Fig. 1

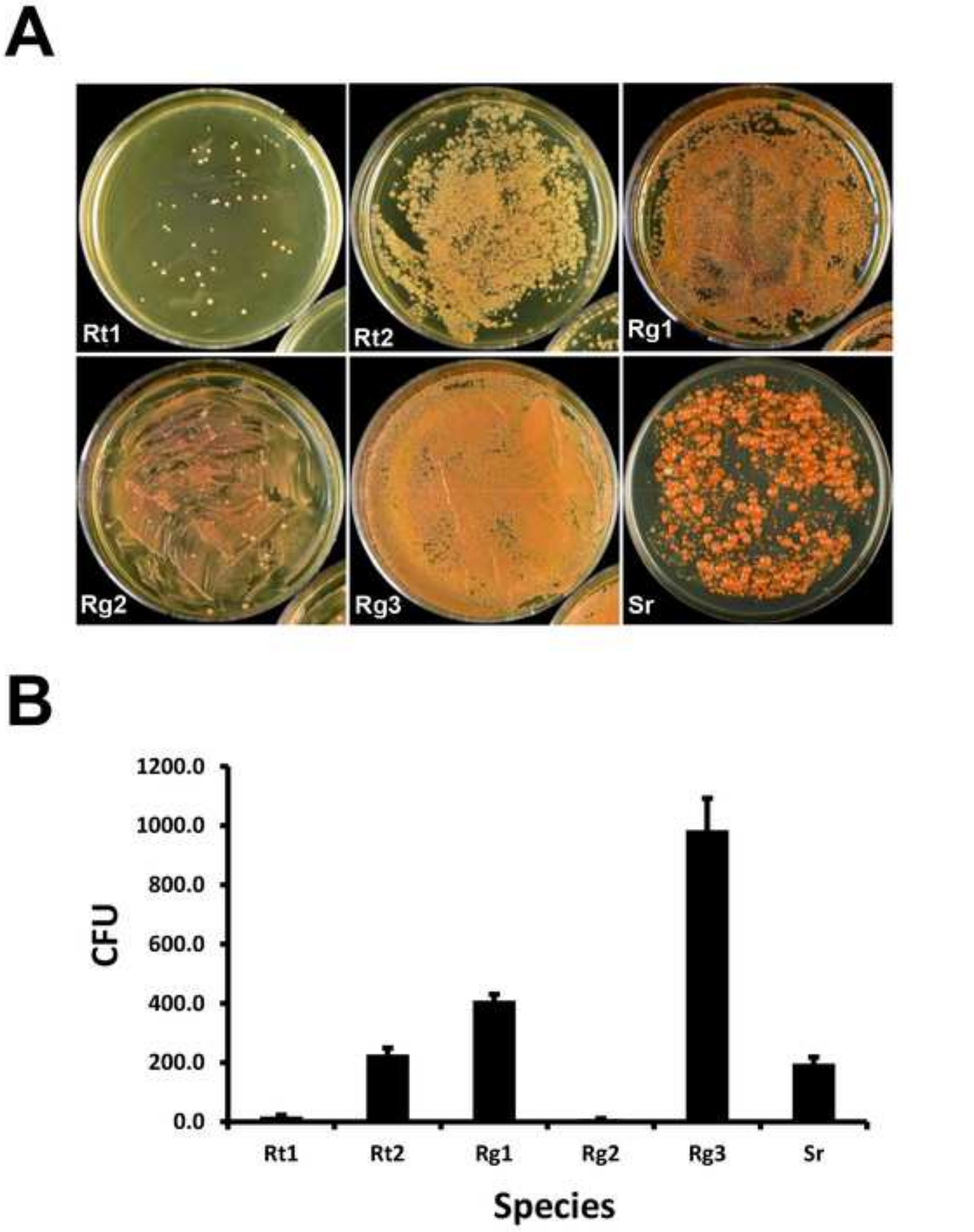

Fig. 2

Click here to download Figure Fig 2.tif

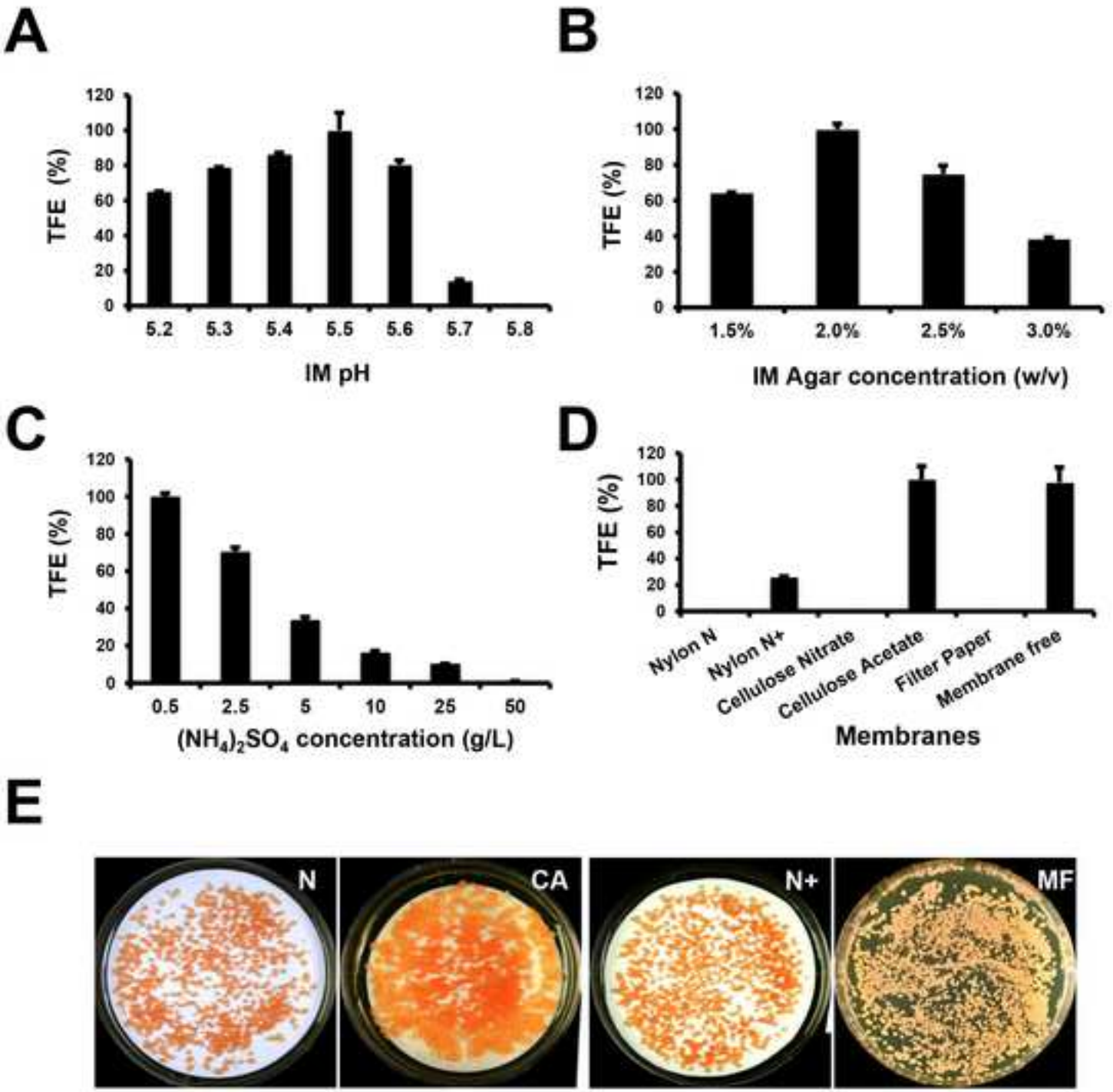

Fig. 3

Click here to download Figure Fig 3.tif

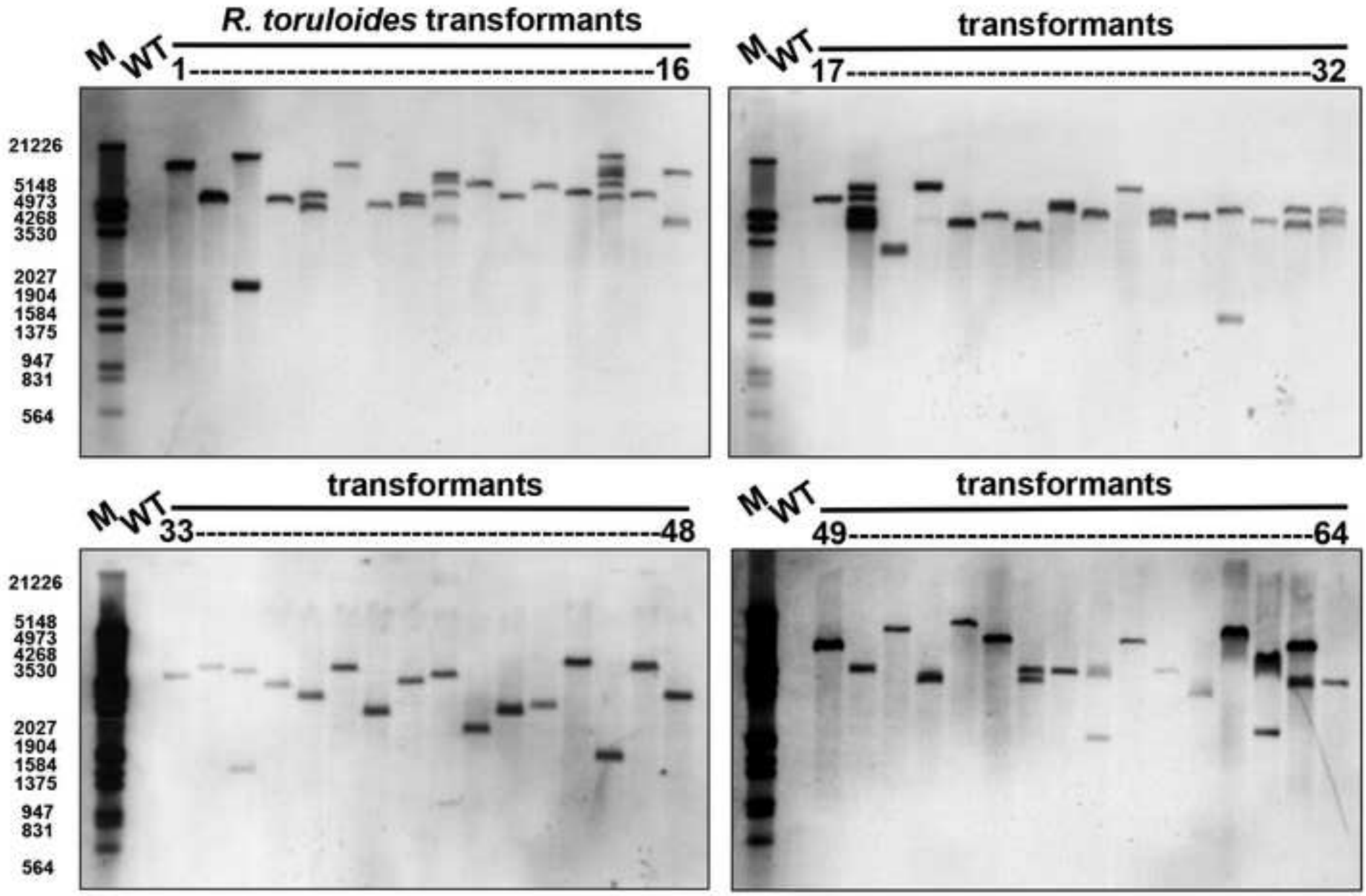

Fig. 4

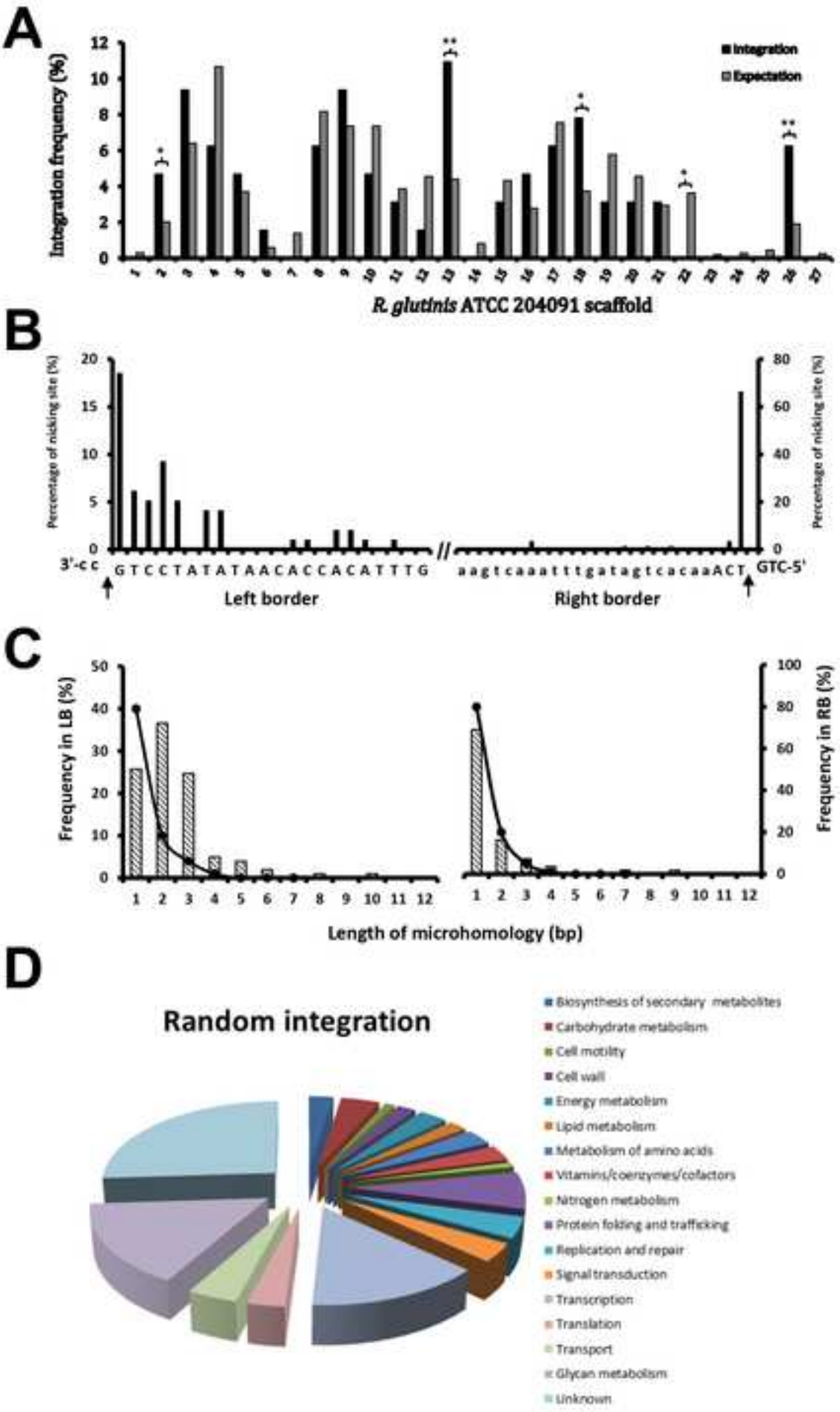

Fig. 5

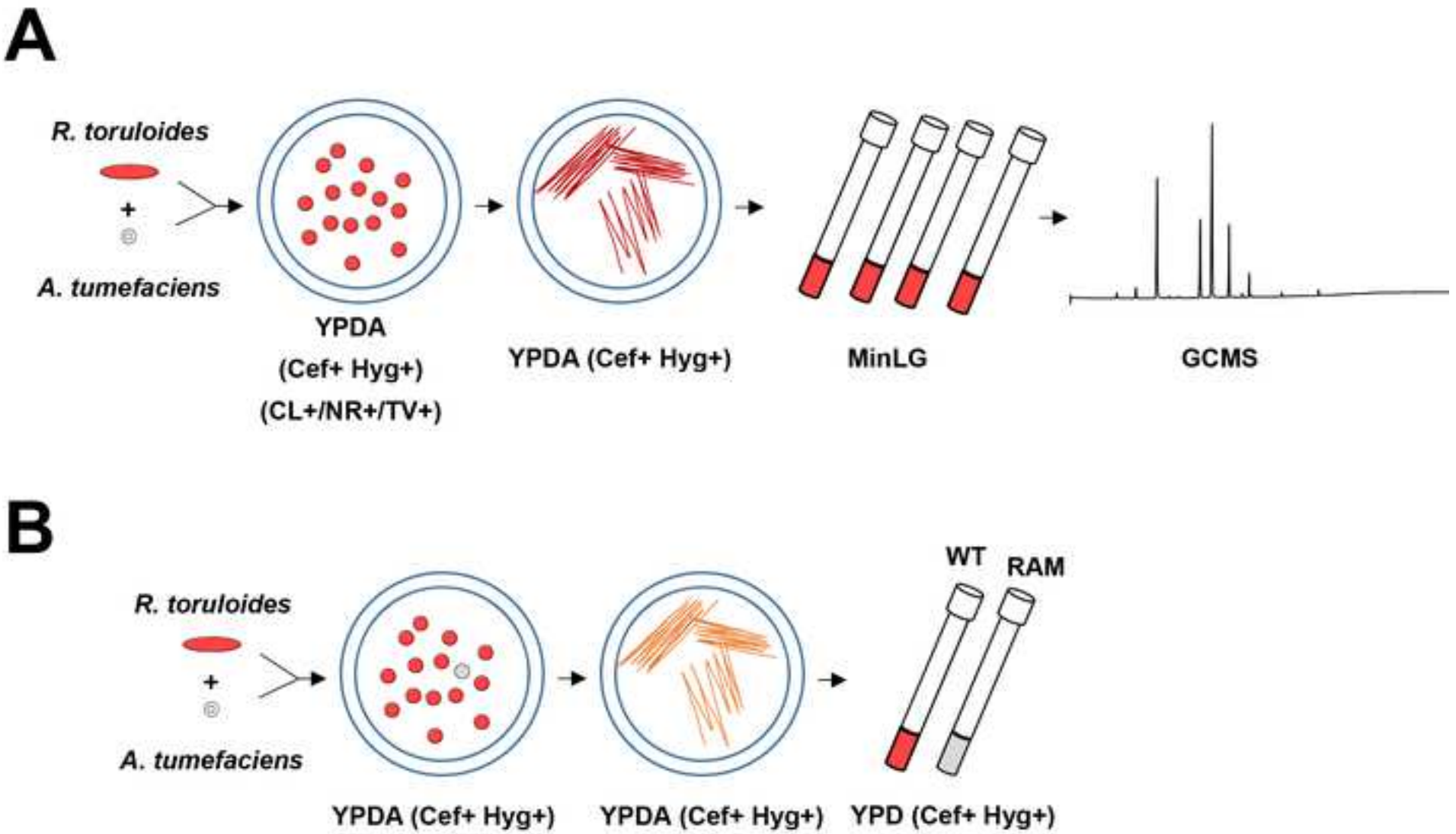

Fig. 6

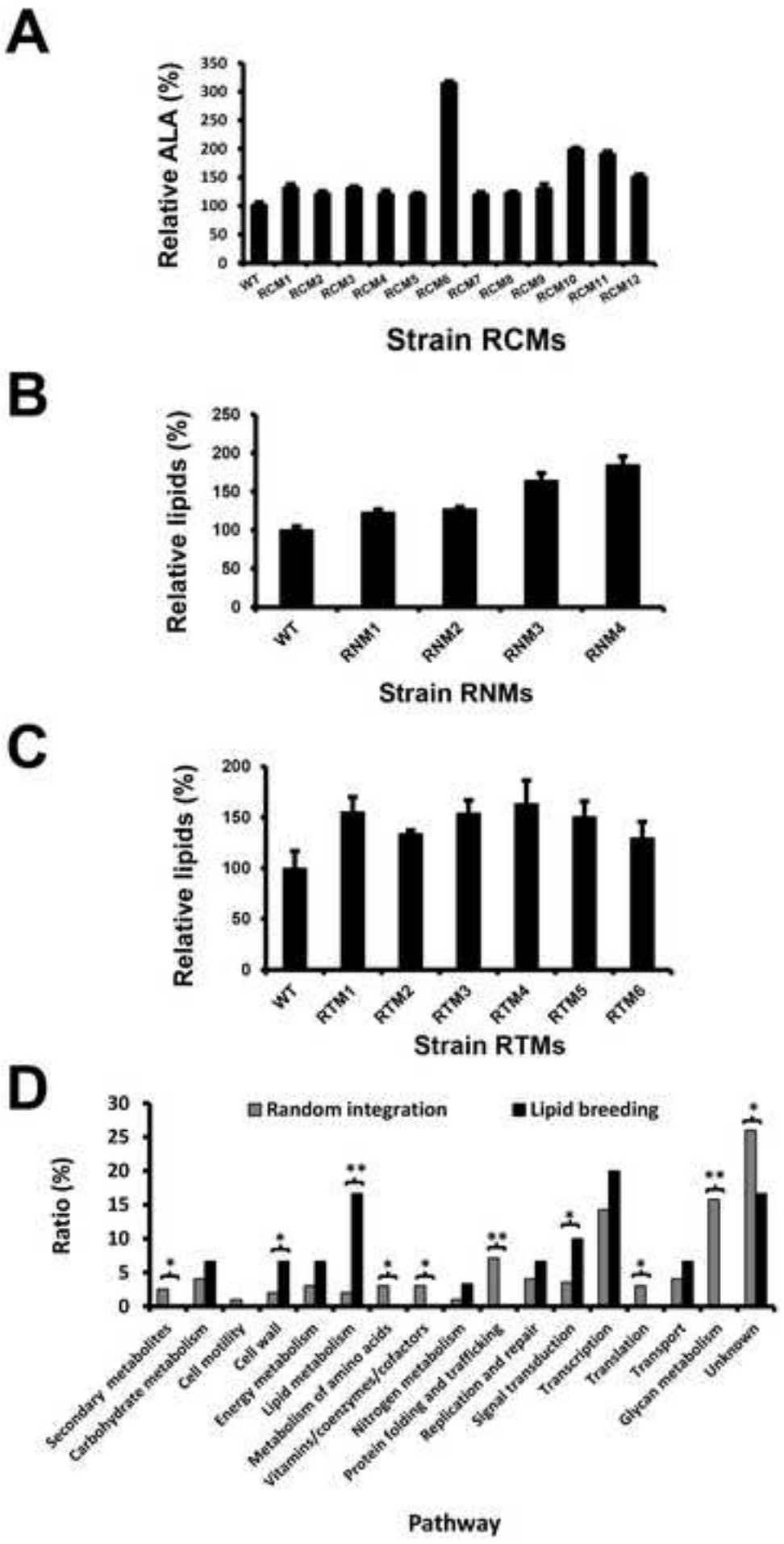

Fig. 7

Click here to download Figure Fig 7.tif