Doctoral Dissertation

Investigation of non-mevalonate-like pathway for terpene biosynthesis in *Yarrowia lipolytica*

Sivamoke Dissook

June 2021

International Program of Frontier Biotechnology,

Division of Advanced Science and Biotechnology,

Graduate School of Engineering,

Osaka University

Table of Contents

List of figures

List of tables

List of Abbreviations

(*in alphabetical order*)

2OG	2-Oxoglutarate
3PGA	3-Phosphoglycerate
ADP	Adenosine Diphosphate
a-GP	Alpha-Glycerophosphate
AMP	Adenosine Monophosphate
ANDI	Analytical Data Interchange Protocol
ATP	Adenosine Triphosphate
cAMP	Cyclic Adenosine Monophosphate
CMP	Cytosine Monophosphate
ESI	Electrospray Ionization
F1P	Fructose 1-Phosphate
F6P	Fructose 6-Phosphate
FAD	Flavin Adenine Dinucleotide
FBP	Fructose 1,6-bisphosphate
FMN	Flavin Mononucleotide
G1P	Glucose 1-Phosphate
G6P	Glucose 6-Phosphate
GAP	D-Glyceraldehyde 3-phosphate
GC	Gas Chromatography
GC-Q/MS	Gas Chromatography-Quadrupole/Mass Spectrometry
GDP	Guanosine Diphosphate
GMP	Guanosine Monophosphate
GTP	Guanosine Triphosphate
HCA	Hierarchical Cluster Analysis
IMP	Inosine Monophosphate
IP-LC-MS/MS	Ion-pair Liquid Chromatography-tandem Mass Spectrometry
IPP	Isopentenyl pyrophosphate
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Liquid Chromatography
LC/MS	Liquid Chromatography/Mass Spectrometry
MEcPP	2-C-methyl-D-erythritol 2,4-cyclic diphosphate
MEP (pathway)	Non-mevalonate pathway
MEP (metabolite)	2-C-methyl-D-erythritol 4-phosphate
Mn6P	Mannose 6-Phosphate
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MVA (pathway)	Mevalonate pathway
MVA (metabolite)	Mevalonic acid
NAD	Nicotinamide Adenine Dinucleotide

NIST	National Institute of Standards and Technology
OPLS-DA	Orthogonal Projection to Latent Structures-Discriminant Analysis
PC	Principal Component
PCA	Principal Component Analysis
PEP	Phosphoenolpyruvic acid
PLS	Projection to Latent Structures
PLS-DA	Projection to Latent Structures-Discriminant Analysis
PRPP	Phosphoribosyl pyrophosphate
R1P	Ribose 1-Phosphate
R5P	Ribose 5-Phosphate
RMSEE	Root Mean Squared Error of Estimation
RMSEP	Root Mean Squared Error of Prediction
RP-LC-MS/MS	Reversed Phase Liquid Chromatography-tandem Mass Spectrometry
Ru5P	Ribulose 5-Phosphate
S7P	Sedoheptulose 7-Phosphate
SDS-PAGE	Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis
SOMs	Self-Organizing Maps
TAG	Triacylglycerol
TCA cycle	Tricarboxylic Acid cycle
TMP	Thymidine Monophosphate
TPP	Thiamine pyrophosphate
UDP-Glc	UDP-Glucose
UMP	Uridine Monophosphate
UNV	Univariate Analysis
UTP	Uridine Triphosphate
VIP	Variable Importance in Projection
XMP	Xanthosine Monophosphate

Chapter 1

General introduction

1.1 *Yarrowia lipolytica* and *Yarrowia* spp.

Yarrowia is one of the fungal genera in the family Dipodascaceae. For some time, the genus was monotypic, including the single species *Yarrowia lipolytica*. This yeast is an appealing host for converting the various carbon source to high-value chemicals in an economical and environmentally friendly way. Especially for the production of chemicals derived from acetyl-CoA, fatty acids (FA), and lipids (Abghari et al., 2014), *Y. lipolytica* is a hemiascomycetous yeast with a history of industrial applications. It has been used as a model oleaginous organism for a wide range of studies, involving lipid collection and degradation, dimorphism, peroxisome biogenesis, and protein excretion pathways (Beopoulos et al., 2009). Extensive studies show that *Y. lipolytica* is a safe organism for industrial uses (Groenewald et al., 2013). Wild type (WT) *Y. lipolytica* strains can utilize glucose, fructose, glycerol, and hydrophobic substrates as carbon sources; strain PO1g can also use xylose and sugarcane bagasse hydrolysate (Tsigie et al., 2011). In addition, *Y. lipolytica* strains have been engineered to use many others carbon sources (Ledesma-Amaro et al., 2016). As an aerobic organism, *Y. lipolytica* has relatively high flux through the pentose phosphate pathway (PPP) that generates co-factor NADPH to support fatty acid (FA) biosynthesis (Blank et al., 2005; Christen et al., 2011). Most *Y. lipolytica* strains are haploid but can also exist in a diploid form that is more tolerant to various conditions (Flagfeldt et al., 2009). *Y. lipolytica* is one of the most contrasting of the characterized Hemiascomycetes (Dujon et al., n.d.). In spite of a genome approximately twice the size of *Saccharomyces cerevisiae*, *Y. lipolytica* is not believed to have undertaken whole-genome

duplication (Gaillardin et al., 2013). Likewise, *Y. lipolytica* has more characters common with metazoan cells than other yeasts. These involve signal-recognition-particle type 7SL RNA sequence, dispersed 5S genes, and a higher section of the genome composed of introns and intergenic sequences. The *Y. lipolytica* genome also harbors representatives of varied classes of transposable elements, including fragments of a DNA transposon (Neuvéglise et al., 2005), long-terminal repeat (LTR) (Neuvéglise et al., 2002), and non-LTR Long INterspersed Element (LINE) (Casaregola et al., 2002) retrotransposons. Finally, dissimilar to the more well-documented respiro-fermentative *S. cerevisiae*, *Y. lipolytica* stands an obligate aerobe. Molecular phylogenetics study has discovered several other species that have since been added to the genus (Crous et al., 2017; Michely et al., 2013; Nagy et al., 2014).

1.2 Carbon to nitrogen ratio effect on yeast

It is a collective awareness that adjusting the carbon to nitrogen ratio in a culture medium could considerably change the metabolome profile in yeasts, and it might have a distinctive effect on different strains of yeast (Egli et al., 1986; Gao et al., 2007; Jackson et al., 1990).

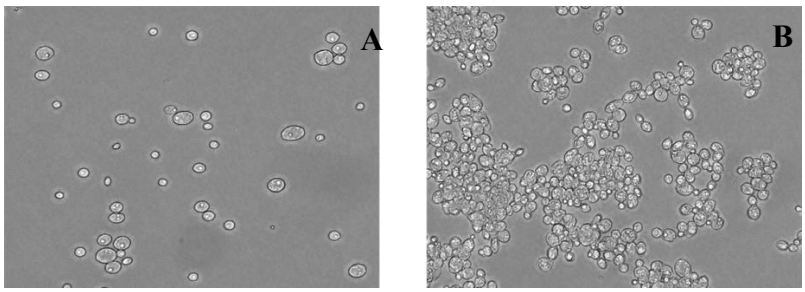


Figure 1- 1 comparison between *Y. lipolytica* (A) cultivated in nitrogen-limiting condition (B) cultivate in normal condition

Substantial work has been made in the past few years to comprehend and modify oleaginous yeast (Silverman et al., 2016). Nonetheless, only narrow work has been made

to illustrate the metabolic reaction of *Yarrowia* spp., especially the recently added species in nitrogen limiting condition. Below are some of the literature on the carbon and nitrogen ratio effect on yeast.

In 2001, Abd-Aziz et al.(Abd-Aziz et al., 2001) studied the effect of C/N ratio, and starch concentration on ethanol production with *Saccharomyces cerevisiae* results from this study show that the limitation of nitrogen inhibited glucoamylase secretion and suppressed starch utilization. For fermentation using a high concentration of nitrogen, the excess nitrogen together with glucose produced from starch hydrolysis was consumed for biomass build-up, and very little glucose remained for conversion to ethanol.

In 2007, Li Gao et al. study the effect of carbon concentration and carbon to nitrogen ratio on the growth and sporulation of various biocontrol fungi. The results suggested that both the carbon concentration and carbon to nitrogen ratio had a significant effect on fungal growth. In general, the carbon to nitrogen ratio had more influence on fungal growth than carbon concentration. Moreover, the results from this study suggested that the effect of carbon concentration and C:N ratio on fungal development and sporulation is strain-dependent; thus, consideration of the complexity of nutrient supplies is crucial for improving yields of fungal biocontrol agents.

In 2010, Manikandan and Viruthagiri (Manikandan et al., 2010) performed the optimization of the C/N ratio of the culture medium and fermentation conditions of ethanol production from Tapioca starch using Co-culture of Aspergillus niger and Saccharomyces cerevisiae. The results show that the C/N ratio of the fermentation medium plays a vital role in the production of ethanol by the SSF process. The biomass was found to increase with the decrease of the C/N ratio. Furthermore, the results indicate that a higher C/N ratio gave an optimized yield of ethanol of 8.85 g/l. The amount of ethanol was found to decrease

drastically with increasing nitrogen concentration.

In 2013, Sitepu et al. (Sitepu et al., 2013) manipulated culture conditions of various oleaginous yeast species. This study shows that lipid accumulation generally takes place when the oleaginous yeasts grow in the presence of high amounts of a carbon source, and the nitrogen source is kept limited. Under these conditions, the excess carbon source is channeled into lipid bodies accumulation. In non-oleaginous yeast cultured under the same metabolic conditions tend to stop growing when the nitrogen source is exhausted; lipid synthesis occurs at a low level (less than 10%), and the left-over carbon source is diverted to form polysaccharides, such as starch, β-glucan, and mannan.

In 2016, Pomraning et al. utilized multi-omics technology including GC-MS-based metabolomics, LC-MS-based proteomics, and genomics to study response to nitrogen limitation in *Y. lipolytica*. This study reveals regulators of the response to nitrogen limitation in *Y. lipolytica* and concluded that the integration of proteome, metabolome and phosphor-proteome data identifies lipid accumulation in response to nitrogen restriction as a two-fold result of improved production of acetyl-CoA from excess citrate and decreased capacity for β-oxidation. Although this study employed extensive omics technology to study the *Y. lipolytica* in different carbon to nitrogen ratios, they did not report MEP pathway or MEP-like pathway-related information. I speculated several reasons why the MEP signal might not be detected or picked up. Firstly, the variation in carbon to nitrogen ratio in this study does not reach the point where no nitrogen source was supplied in contrast to our study, where we completely remove the nitrogen source from the beginning of the cultivation. Secondly, the metabolome platform used in this study is GC-MS based, contrast to our study where widely targeted LC/MS/MS-based metabolomics was utilized, and all chromatogram was manually inspected. In our study using GC-MS analysis, MEP was also

not annotated. Finally, the strains of *Y. lipolytica* are different (ATCC 20460 vs. PO1d); during our study with six additional *Yarrowia* spp., initially, we found that not all *Y. lipolytica* exhibit the same trait.

1.3 MEP pathway and MVA pathway

Isoprenoid compounds are one of the most diverse groups of natural products with a wide range of biological roles (Wanke et al., 2001). Every isoprenoid is a derivative of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (Mcgarvey et al., 1995). In nature, two biological pathways are leading to IPP and DMAPP; the mevalonic acid (MVA) pathway and the methylerythritol phosphate (MEP) pathway (Rohmer, 1999; Vranová et al., 2013) (Fig.1-1). The MVA pathway is scattered among archaebacteria, some eubacteria, and in the cytosol of nearly all eukaryotic cells. In contrast, the MEP pathway is found in numerous bacteria, algae, cyanobacteria, apicomplexan parasites, and plant chloroplasts (Kuzuyama et al., 2003).

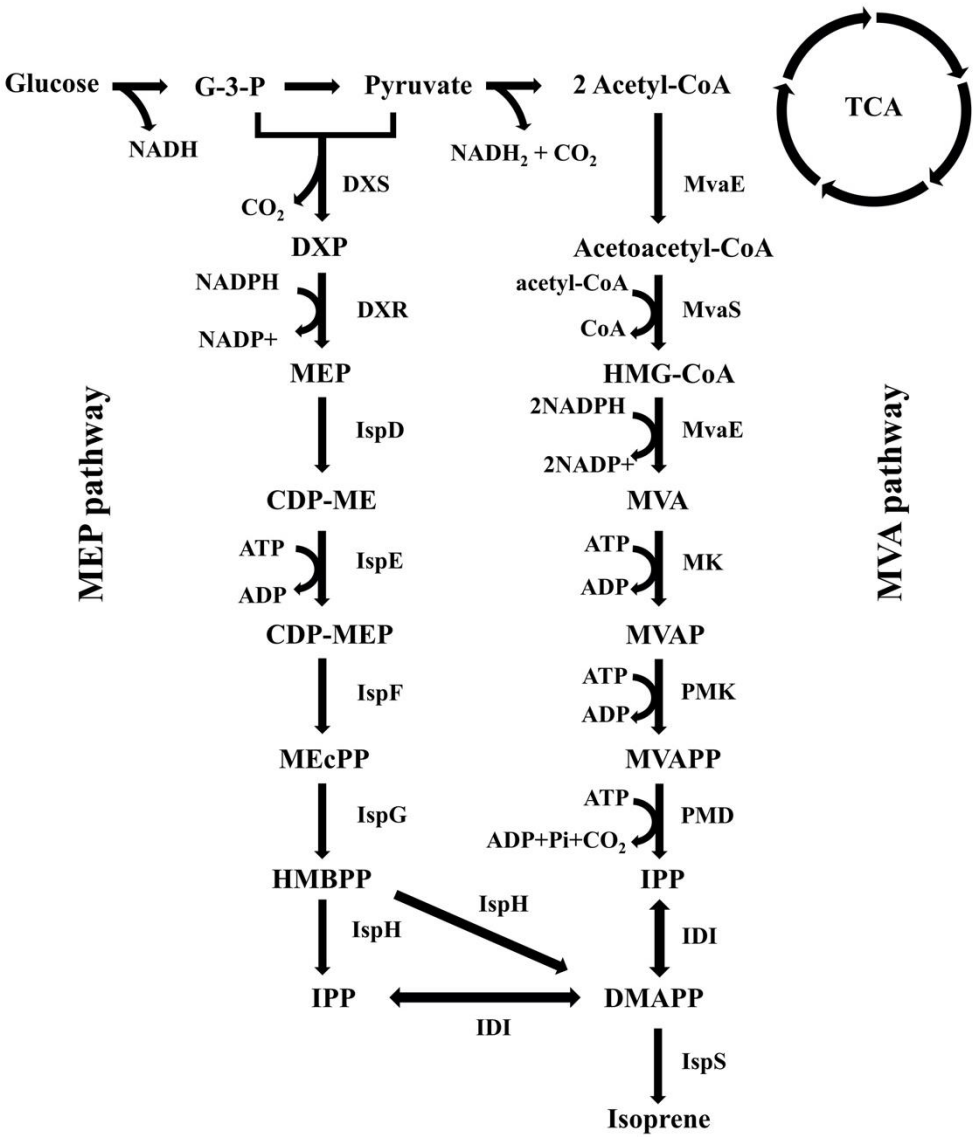


Figure 1- 2 Biosynthetic pathway for isoprenoid precursors.

The mevalonic acid (MVA) pathway in animals, plants (cytosol), fungi, and archaea. The methylerythritol phosphate (MEP) pathway in eubacteria, green algae, and the plastids of higher plants. Abbreviations: CDP-ME, methylerythritol cytidyl diphosphate; DXP, 1-deoxy-D-xylulose 5-phosphate; DXR, DXP reductoisomerase; DXS, DXP synthase; HMBPP, 4-hydroxy-3-methyl-butenyl-1-diphosphate; IPP, Isopentenyl diphosphate; IDI, IPP:DMAPP isomerase; MEcPP, 2-C-methyl-D-erythritol-2,4-cyclodiphosphate; MK, mevalonic acid kinase; PMD, phospho-mevalonate decarboxylase; PMK, phosphomevalonate kinase; DMAPP, dimethylallyl diphosphate; HMG-CoA, hydroxymethylglutaryl-CoA; MVA, mevalonic acid; MVAP, phosphomevalonic acid; MVAPP, diphosphomevalonic acid.

1.4 Metabolomics

Metabolomics can be described as the comprehensive study of all metabolites in a biological system. Unlike proteome and transcriptome that are considered as media concerning genetic information, metabolome, which is related directly to phenotype, plays an essential role as a result of genome information (Putri et al., 201 CE). Therefore, metabolomics is a great tool to reveal changes in phenotype enacted by any perturbations such as genome modification, environmental influences, physical stress, etc. With this capability, metabolomics has been employed in many fields, including an application for biomarker finding of the various disease, drug discovery, food quality assessment, and discovery of new metabolic pathways, and many others.

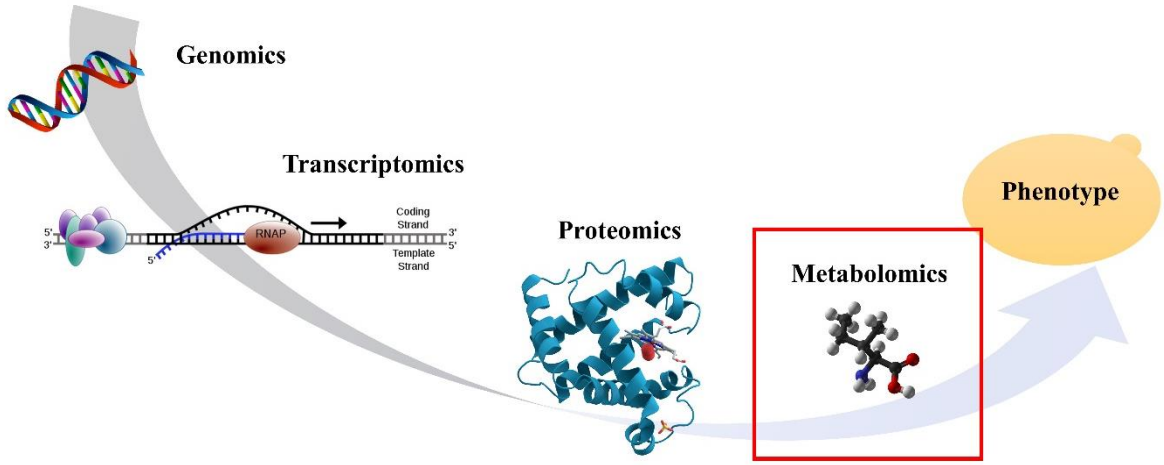


Figure 1- 3 The central dogma of biology showing the flow of information from the genome to the phenotype.

There are two core tactics that can be applied, namely targeted and non-targeted metabolomics. Targeted metabolomics is the method focusing on a specific group of

metabolites related to the pathways of interest. The strategy for targeted metabolomics is usually based on a specific biochemical question or hypothesis that motivates the examination of a particular pathway. This tactic can be applied to pharmacokinetic studies of drug metabolism as well as for determining the influence of therapeutics or genetic modifications on a specific enzyme. On the other hand, non-targeted metabolomics is a technique with the goal of measuring as many metabolites as possible from biological samples without the defined target. This technique can provide a broader coverage that has excellent potential to provide insights into essential biological processes; still, a significant number of metabolites remain uncharacterized with respect to their structure and function. These approaches have been used either discretely or together for the following approaches;

Metabolomics focus on the comprehensive set of metabolites within a biological sample under a particular condition. Due to the connection of metabolic pathways, any factors that cause the deviations in the metabolism are usually not limited to only one pathway. Therefore, it is essential to have a biological system that can cover the identification and quantification of metabolites in all branches of an organism's metabolism.

Metabolic profiling focus on the analysis of the set of metabolites in a pre-defined biochemical pathway or a specific class of compounds. In this technique, it is not compulsory to cover all metabolic pathways of an organism. Instead, this technique emphasizes the identification and quantification of the selected metabolites in biological samples.

Metabolic fingerprinting focuses on sample classification based on the difference of metabolic states due to their biological significance or backgrounds but does not automatically give specific metabolite information (Fiehn, 2002).

Metabolic footprinting (exometabolomics) focuses on extracellular metabolites in

cell culture media. By comparing the metabolic profiles of media before and after culturing, this method provides a reflection of metabolite excretion or uptake by cells (Silva et al., 2015).

The field of metabolomics has made noteworthy progress within the past decade and has employed new tools that offered mechanistic insights by allowing the association of biochemical changes with phenotype. Recent advances in metabolomics technologies are leading to a growing number of mainstream applications such as drug discovery, human disease study, systems biology, plant research, and the food industry. Metabolomics is high-throughput in nature, with it metabolomics scale data analysis remains a significant challenge with all the advances surrounding analytical platforms available for high-throughput and high-resolution data acquisition, ranging from Fourier transform mass spectrometry (FTMS) to ion mobility mass spectrometry (MSI) to imaging mass spectrometry (MSI). There is a constant need for data analysis tools and resources at each step of analytics to meet the ever-growing demand to address challenges associated with diverse data types and more extensive data set. Several databases and identification tools have been published (O'Shea et al., 2020). After the correct metabolite identification, one of the most critical steps is data interpretation. In the last decade, substantial work has been done to develop numerous useful databases on metabolites and comprehensive metabolic pathways such as Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2000), Small Molecule Pathway Database (SMPDB) (Jewison et al., 2014), An atlas of human metabolism (Robinson et al., 2020), WikiPathways (Martens et al., 2021), and MetaCyc (Caspi et al., 2020). The information from the pathway and network analysis approaches will help to characterize the complex relationships in the set of measured metabolites. Recently, many researchers combine metabolomics with other "omics" studies

in order to gain fundamental knowledge from the multiple aspects of biology (Pomraning et al., 2016).

Compound separation, data handling, and statistical analysis play a vital role in metabolomics technique. Robust separation combined with a very sensitive detector was essential to facilitate peak recognition as well as metabolites identification. A recent development in metabolomics has been prominently influenced by the improvements of chromatographic and detection, particularly mass spectrometry, platform. In general, chromatography coupled with a sophisticated detection platform produces a large amount of raw data, including the sample's actual data and instrumental noise. In order to extract important information from these bulk data, it is vital to perform appropriate data preprocessing. Prior to conducting multivariate analysis, the data matrix should be subjected to proper data preprocessing and transformation. Multivariate data analysis is used to extract information from a complicated metabolome data matrix. Various techniques have been employed, such as Principal Component Analysis (PCA), Projection to Latent Structures (PLS), and so on, to assist the investigator in interpreting the data.

1.5 Research background

Initially, the objective of this research was aimed to use metabolomics to improve the production of the fatty acid ethyl ester in *Y. lipolytica* using metabolomics approach; however, during the survey on the *Y. lipolytica* metabolome response to different carbon to nitrogen ratio, a chromatographic peak annotated as MEP was spotted in the intracellular extract of *Y. lipolytica* cultured in a nitrogen limiting condition. This phenomenon is inconsistent with a usual understanding concerning isoprenoid biosynthesis in yeast. This outcome raises a fascinating hypothesis of whether *Y. lipolytica* could use the MEP pathway

for isoprenoid biosynthesis. Therefore, the objective of this study was changed to investigate the existence of MEP pathway in *Y. lipolytica*

1.6 Thesis outline

This thesis includes four chapters that provide the study on metabolome response to different carbon to nitrogen ratio of *Yarrowia* spp. and the more profound investigation of MEP pathway in *Y. lipolytica* using three different approaches. Chapter 1 provides the general introduction regarding *Yarrowia* spp., general yeast response to different carbon to nitrogen ratios, and the MEP pathway. In Chapter 2, widely-targeted LC/MS-based metabolic profiling methods were employed to cover a wide range of metabolites in the central metabolic pathways of time-course *Y. lipolytica* and other *Yarrowia* spp. In Chapter 3, in order to investigate the existence of the MEP pathway in *Y. lipolytica*, three approaches were employed to investigate the existence of the MEP pathway in *Y. lipolytica*; the spiking of the authentic standard, the MEP pathway inhibitor, and the ^{13}C labeling incorporation analysis. Finally, in Chapter 4, the conclusions revealed from this study were summarized, and future perspectives were proposed.

Chapter 2
Metabolome profiling of *Y. lipolytica* in different carbon to nitrogen ratio

2.1 Introduction

In this study, I perform the time-course sampling of one *Y. lipolytica* strain to explore the appropriate time points for multiple strain analysis. Since the condition changes in this study tremendously affect the growth of *Y. lipolytica*, it might not be proper to choose the sampling point grounded on the growth phase only. Therefore, I considered the metabolome profile to aid in understanding the dynamics of metabolic fluctuations happening throughout the cultivation with various C:N ratios and determined the critical sampling time point and C:N ratio for several strains study. I further studied the metabolic reactions of several *Yarrowia* strains to nitrogen-limiting conditions, involving one laboratory strain, *Y. lipolytica* PO1d (Knutsen et al., 2007), three previously described *Candida lipolytica* species from different research groups (currently identified as *Yarrowia* spp.), *C. lipolytica* JCM 2304 (Barth et al., 1996), *C. lipolytica* JCM 21924 (NAKASE et al., 1977), and *C. lipolytica JCM* 8061, and two recently identified species, *Y. deformans* JCM 1694 (Groenewald et al., 2013) and *Y. keelungensis* JCM 14894 (Chang et al., 2013). This report will be valuable for the development and manipulation of *Yarrowia* spp. for potential applications.

2.1 Materials and Methods

2.1.1 Yeast Strains, Cultivation, and Sample Collection

The *Y. lipolytica* laboratory strain PO1d was acquired from the National University of Singapore (NUS), *C. lipolytica* JCM 21924, *C. lipolytica* JCM 2304, *C. lipolytica* JCM

8061, *Y. deformans* JCM 1694, and *Y. keelungensis* JCM 14894 were acquired from the Japan Collection of Microorganisms (JCM). *Yarrowia* strains were cultivated overnight in synthetic complete (SC) medium, all components were prepared according to the producer's direction (20 g/L glucose, 1.7 g/L, yeast nitrogen-based without amino acids, and ammonium sulfate acquired from BD Difco product ID 233520, 5 g/L ammonium sulfate, drop-out medium supplements acquired from Sigma-Aldrich product ID Y1501 consists of all standard amino acids at 76 mg/L except for leucine, which is present at 380 mg/L, inositol (76 mg/L), adenine (18 mg/L), p-aminobenzoic acid (8 mg/L) in 100-mL Erlenmeyer flasks at 30 °C and 200 rpm. The pre-cultured cells were injected into 20 mL of SC medium with distinctive C:N composition (e.g., C:N ratio 4:1 comprises of 20 g/L glucose, 5 g/L ammonium sulfate, all others standard amino acids at 76 mg/L except for leucine, which is present at 380 mg/L, 8 mg/L p-aminobenzoic acids, 18 mg/L adenine, 76 mg/L inositol, 1.7 g/L, yeast nitrogen-based without amino acids and ammonium sulfate. C:N 5:0 consists of 25 g/L glucose, all standard amino acids at 76 mg/L except for leucine, which is present at 380 mg/L, 8 mg/L p-aminobenzoic acid, 1.7 g/L, 18 mg/L adenine, 76 mg/L inositol, yeast nitrogen-based without amino acids and ammonium sulfate) to achieve an initial optical density at 600 nm of 0.5 in 100-mL shake flasks at 30 °C and 200 rpm. For sample collection, cells were collected at an OD_{600} of 5.0 at 12, 24, 36, and 48 h by fast filtration using a 0.45-μm pore size, 47-mm diameter nylon membrane (Millipore, Billerica, MA, USA). The cells were instantly submerged in liquid nitrogen for quenching and stored at −80 °C until extraction.

2.1.2 Extraction of Intracellular Metabolites

For intracellular metabolites extraction, 1800 μL of extraction solvent (methanol/water/chloroform = 5:2:2 v/v/v%, with 20 μg/L camphorsulfonic acids as an

internal standard) was added to each 2 mL Eppendorf tube with the sample, and incubate at −30 °C for 1 h. Next, 700 μL of the solution was transported to a new tube, including 350 μL of water. The mixture was homogenized and centrifuged at 9390× g. for 10 min at 4 °C to isolated polar and non-polar phases. Then, 700 μL of the polar phase was transferred to a new tube by syringe filtration (0.2-μm PTFE hydrophilic membrane; Millipore). The sample was centrifugally condensed for 2 h and freeze-dried, and kept at −80 °C for further uses. After resuspending in 50 μL of ultrapure water, the sample was centrifuged at 16,000× g for 3 min at 4 °C and transported to a glass vial for the LC/MS analysis.

2.1.3 Analysis of Intracellular Metabolites by LC-MS/MS

The samples were investigated based on the method modified from Dempo et al. (Dempo et al., 2014). Briefly, ion-pair reversed-phase LC/MS/MS was done using a Nexera UHPLC system (Shimadzu, Kyoto, Japan) connected with LCMS 8030 Plus (Shimadzu) for the time course analysis and connected with LCMS 8050 (Shimadzu) for analyses of many strains. The column was CERI (Chemicals Evaluation and Research Institute, Tokyo, Japan) L-column 2 metal-free ODS (150 mm 2.1 mm, particle size 3 μm). Mobile phase (A) was 10 mM tributylamine and 15 mM acetate in ultra-pure water, and a mobile phase (B) was 100% methyl alcohol. The flow rate was fixed at 0.2 mL/min. The column oven temperature was kept at 45 °C. The concentration of mobile phase (B) was set to rise from 0% to 15%, 50%, and 100% from 1.0 to 1.5 min, 3.0 to 8.0 min, and 8.0 to 10.0 min, correspondingly, held until 11.5 min, then reduced to 0% from 11.5 min and held at 0% for 20 min. The ion evaluation mode was set to negative ion detection type. The sample injection volume was set to 3 μL, desolvation line temperature was kept at 250 °C, probe position was set at +1.5 mm, heat block temperature was kept at 400 °C, nebulizer gas flow was kept at 2 L/min, and drying gas flow was set at 15 L/min.

2.1.4 Multivariate Analysis

Data analysis was carried out by transforming the unanalyzed data from Shimadzu file type (.lcd) to an analysis base file (.abf) type using an open-source file converter (Reifycs Inc., Tokyo, Japan). After data transformation, MRMPROBS 2.19 (Tsugawa et al., 2013) was used for automatic peak selection and peak area integration. The detected peaks were manually curated by Lab Solution software (Shimadzu Co.). A principal component analysis (PCA) was done by SIMCA-P+ software version 13 (Umetrics, Umeå, Sweden). The metabolome data were standardized using a mean-centered, internal standard and scaled to unit variation. A t-test to determine statistical importance was performed using Microsoft Excel.

2.1.5 Glucose Measurement

The remaining glucose in the medium was measured by GL Science GL-7400 high-performance liquid chromatography system attached to an oven GL-7432 and refractive index detector GL-7454 (HPLC-RI). The column used was Shim-pack SPR-Pb (250 mm length and ID 7.8 mm (Shimadzu). The upper supernatant of the medium was centrifuged at 10,000× g for 300 seconds at 4 °C; 300 µL of the sample was filtered by Whatman syringeless LC vial (Merck, Kenilworth, NJ, USA). The sample injection volume was 10 µL; the oven temperature was set to 80 °C. Ultrapure water was used as the mobile phase with a 0.6 mL/min flow rate and 189 psi of backpressure.

2.2 Results and Discussion

2.2.1 Growth of *Y. lipolytica* under Different Carbon to Nitrogen Ratios

To decide the vital sampling time point and C:N ratio for examines of the effect of

nitrogen-limiting conditions, time course sampling in several C:N compositions were executed using *Y. lipolytica* PO1d. The extracellular glucose measurement and growth of *Y. lipolytica* cultivated in different C:N ratios are presented in Figure 2-1. There were three distinctive, stable physiological states: (1) Carbon-restricted (0:5), (2) in-between C:N ratios (4;1, 2:2, and 1:4), and (3) nitrogen-restricted (5:0); all statistical assessment can be found in Table S1.

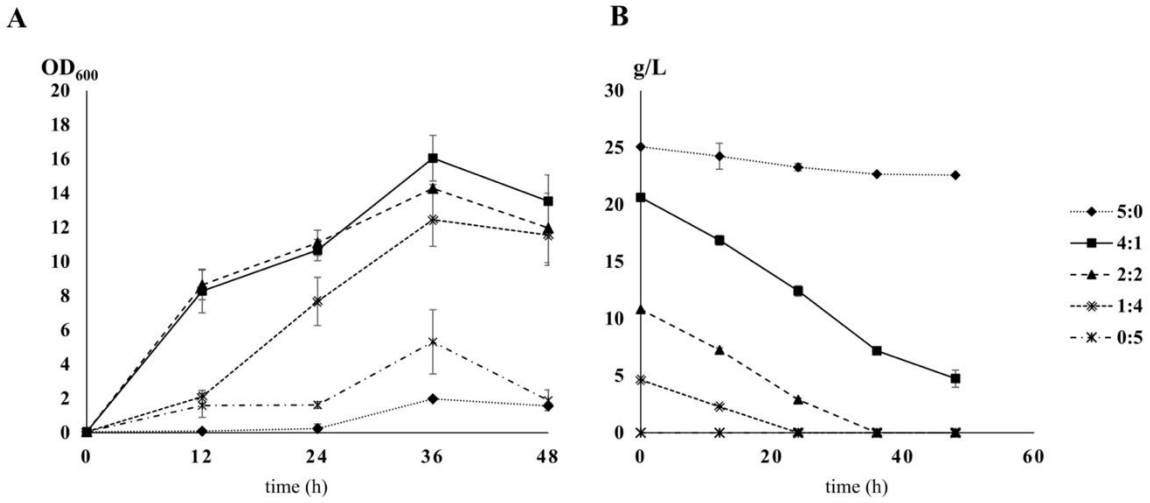


Figure 2- 1 Growth curves of *Y. lipolytica* PO1d and glucose consumption for different carbon-to-nitrogen ratios. *Y. lipolytica* cultures at a starting an OD_{600} of 0.5 for various carbon-to-nitrogen ratios. Error bars represent the standard deviation from three replicates; 5:0 (diamonds with a dashed line), 4:1 (squares with a dark line), 2:2 (triangle), 1:4 (double-crosses with a dashed line), 0:5 (asterisks with a dark line). (A) Growth curve determined by optical density at 600 nm. (B) Glucose (g/L).

Under carbon-restricted conditions, the yeast grows steadily from 0 h to 12 h, remained persistent for 24 h, and cells reached the peak OD_{600} at 36 h. For in-between C:N ratios, growth increased swiftly from 0 h to 12 h, slowed from 12 h to 24 h, and the maximum OD_{600} was reached at 36 h. Under nitrogen-restricted conditions, growth increased gradually from 0 h to 24 h and increased gradually from 24 h until cells reached the peak OD_{600} at 36 h. In this study, all OD_{600} values declined after 36 h of cultivation. In both carbon- and nitrogen-restricted settings, the growth profiles suggested that cells took

roughly 24 h to adapt to the new condition before going to the exponential growth phase. For in-between C:N ratios, particularly 4:1 and 2:2, three stages of cell growth were detected, 0–12 h, 12–24 h, and 24–36 h. This information suggests that the conditions altered during cultivation, and cells had to adjust for growth. Remarkably, this occurrence was not seen for a C:N ratio of 1:4. These results demonstrated that *Y. lipolytica* is extremely sensitive to alterations of the C:N ratio. Astonishingly, *Y. lipolytica* growth did not depend on glucose accessibility, as demonstrated by the continued growth in glucose-depleted settings (C:N 0:5 and 1:4 at 24 h, 2:2 at 36 h). Yeast in the C:N 5:0 settings could not use glucose efficiently, even with plentiful glucose. Moreover, *Y. lipolytica* entered the death stage, regardless of glucose accessibility (C:N 4:1, 5:0).

2.2.2 Intracellular Metabolome Analysis of *Y. lipolytica* for Different Carbon to Nitrogen Ratios

Overall of 93 metabolites were annotated (Table S2). The time-course outcomes for all detected metabolites are brief in Figure 2-2. Annotated compounds mostly involved organic acids, sugars, amino acids, phosphate compounds, and other metabolites. Widely targeted LC-MS/MS results were analyzed by principal component analysis (PCA) to discriminate among cultivation settings. The PCA score plot offers a graphical image of sample dissimilarity from a global view, and, as a non-parametric study, the generated model is objective to the user; therefore, it is unsupervised (Bylesjö et al., 2007; Trygg et al., 2007; Vichi et al., 2009).

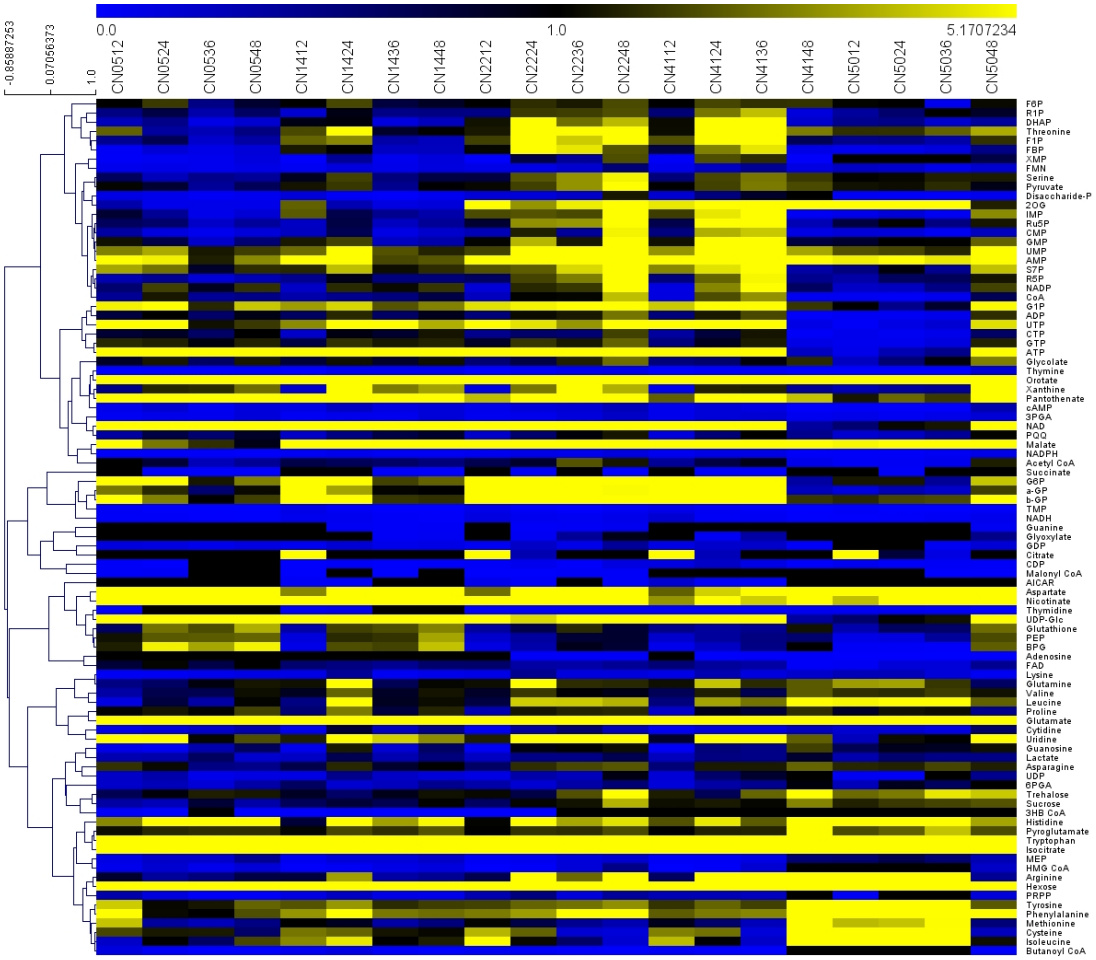


Figure 2- 2. Heat map describes the time-course change of the normalized peak areas metabolome profile of 93 metabolites from *Y. lipolytica* PO1d was clustered hierarchically by a complete linkage method using the MeV software ver. 4.90.

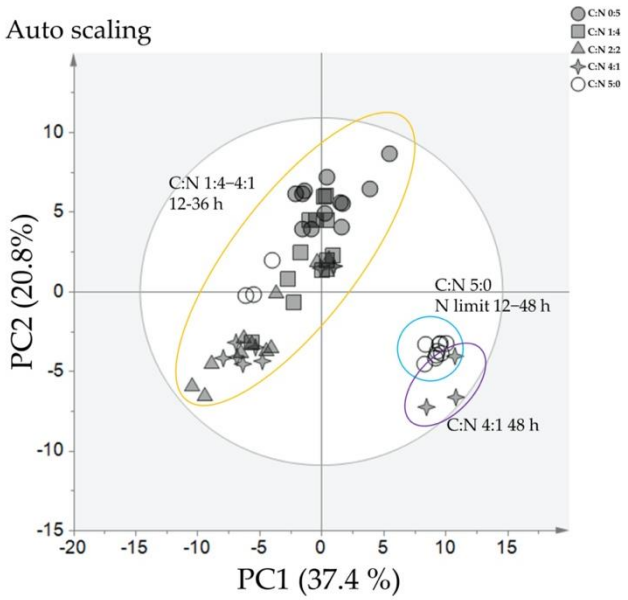


Figure 2- 3 Summary of principal component analysis (PCA) of *Y. lipolytica* cultivated under different carbon-to-nitrogen ratios. The PCA score plot indicates differences in metabolite profiles for different carbon-to-nitrogen ratios based on 93 metabolites. The grey ellipse indicates a 95% confidence border based on Hotelling's T2. Each point represents one replicate for each C:N ratio; 0:5 (grey circles), 1:4 (grey squares), 2:2 (grey triangles), 4:1 (grey stars), 5:0 (white circles). The yellow ellipse indicates the samples from C:N 1:4−4:1 at 12−36 h, The cyan ellipse indicates the samples from C:N 5:0 at 12−48 h, The purple ellipse indicates the samples from C:N 4:1 at 48 h. The complete numerical PCA loading scores are presented in Table S3.

As indicated in Figure 2-3, when all samples at every time point were evaluated by PCA, two major clusters were separated by PC1. Cluster one comprised of samples with C:N ratios of 4:1 and 0:5, 1:4, 2:2 at 12 h to 36 h cultivation. Cluster two split samples into two sub-clusters; sub-cluster one comprised of samples from a C:N ratio of 5:0 at 12 h to 48 h, and sub-cluster two included samples acquired under a C:N ratio of 4:1 at 48 h. These results indicated that the most vital factor dividing the samples was the nitrogen-limiting condition; the C:N ratio of 4:1 at 48 h grouped with the C:N ratio of 5:0, proposing that at 48 h of cultivation, the nitrogen source with a C:N ratio of 4:1 was exhausted. To systematically examine the data, I assessed each time point independently, as summarized in Figure 2-4.

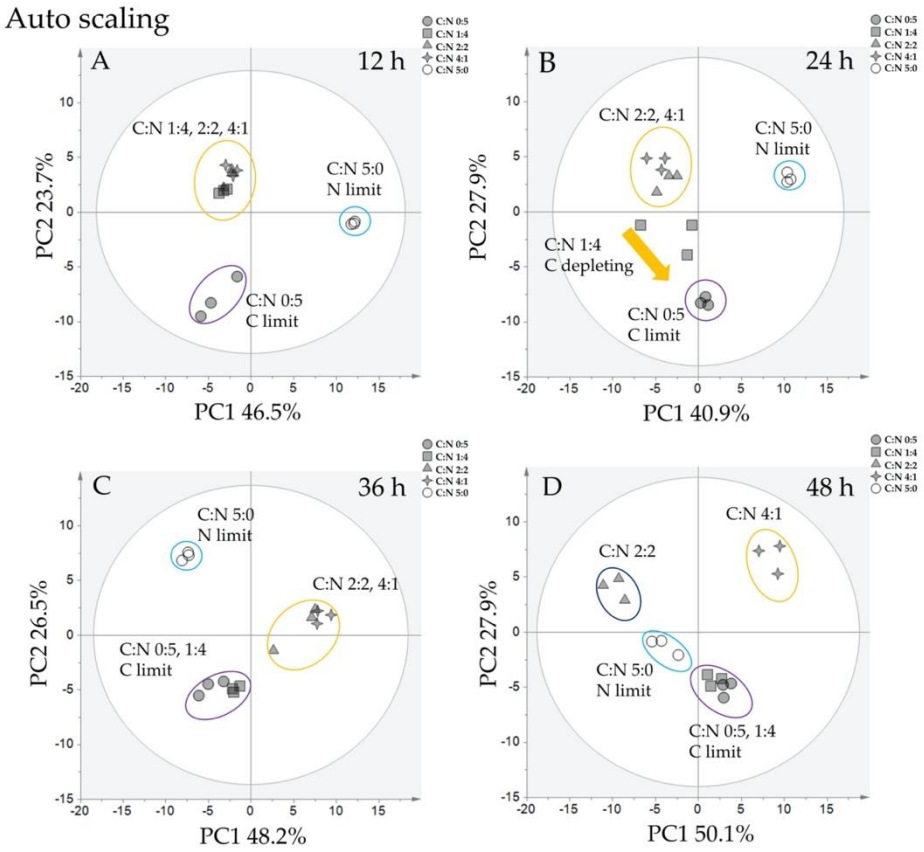


Figure 2- 4 PCA results for *Y. lipolytica* cultivated in different carbon-to-nitrogen ratios. PCA score plot indicating differences in metabolite profiles for different carbon-to-nitrogen ratios based on 93 metabolites. The grey ellipse indicates a 95% confidence border based on Hotelling's T2. Each point represents one replicate in each condition; 0:5 (grey circle), 1:4 (grey square), 2:2 (grey triangle), 4:1 (grey star), 5:0 (white circle). (Scale: Auto-scale; n = 3); (A) profile at 12 h, (B) profile at 24 h, (C) profile at 36 h, and (D) at profile at 48 h. The cyan ellipse indicates the nitrogen-limited samples, the purple ellipse indicates the carbon-limited samples, the yellow and blue ellipse indicates samples in which neither carbon nor nitrogen were limited. The full list of PCA loading scores is presented in Table S4.

As revealed in Figure 2-4A, there were three different clusters at 12 h of cultivation, conforming to the nitrogen-restricted settings, the carbon-restricted settings, and the settings where both carbon and nitrogen were accessible. PC1 divided the nitrogen-limiting settings and other settings, while PC2 divided the carbon-limiting settings and other settings. Almost all metabolites with high PCA scores (upper 10 on the positive and negative edges

of the PC axis) at 12 h were mostly amino acids concerned with the aminoacyl-tRNA biosynthesis pathway, it is substrates for translation and controls the interpretation of the genetic code (Ibba et al., 2004; Soutourina et al., 2000). This suggests that cells in both nitrogen- and carbon-limiting settings adapt to the new environment at this point. At 24 h of cultivation (Figure 2-4B), the grouping results were comparable to those at 12 h, but the C:N 1:4 profile was more parallel to that under the carbon-restricted settings relative to glucose depletion (Figure 2-4B). This perhaps reflects glucose exhaustion in the medium, leading to alterations in cell metabolism to the carbon-restricted settings, comparable to 12 h. PC1 at 24 h divided samples into nitrogen-limiting settings and other settings, while PC2 divided the carbon-limiting settings from other settings. Nonetheless, the metabolites of high PCA score from PC1 are made up of sugar-phosphate from purine and pyrimidine metabolism. These outcomes proposed that purine and pyrimidine metabolism are crucial for the reaction to nitrogen-limiting settings. At 36 h of cultivation (Figure 2-4C), the PCA score plot distinctly shows that the C:N 1:4 profile combined with the C:N 0:5 profile, implying that at this time, the metabolomic profile for the C:N 1:4 settings totally changed to that of the carbon-restricted settings. The C:N 4:1 and 2:2 profiles are grouped together; probably, the metabolomic profile at this time embodies the condition in which carbon and nitrogen are accessible. Bearing in mind the remaining glucose at this time, in the C:N 4:1 settings (Figure 2-1B), around 10 g/L glucose still leftovers, whilst glucose in C:N 2:2 is exhausted; this might elucidate why one of the samples in C:N 2:2 began to move toward the profile of the carbon-restricted settings. Lastly, at 48 h of cultivation (Figure 2-4D), there were four different clusters: C:N 2:2, C:N 4:1, C:N 5:0, and C:N 1:4, 0:5. At this point, the growth profile (Figure 2-1A) proposed that cells entered the death phase, and glucose was exhausted except in C:N 5:0 and 4:1. Fascinatingly, the C:N 2:2 group did not cluster

with the carbon-restricted group, signifying that metabolism during the death phase is unusual, even in the carbon- or nitrogen-restricted settings. High PCA score metabolites at 36 h and 48 h were comparable to those at 24 h, where PC1 comprised typically of sugar-phosphate compounds, and PC2 comprised of metabolites from amino acid and central carbon metabolism. Purine and pyrimidine are intricate in the synthesis of thiamine, riboflavin, folic acid pteridines, and histidine or are parts of cytokinins, purine alkaloids, and other rare nitrogen compounds related to nitrogen collection or nitrogen defecation (O'donovan et al., 1970; Vogels et al., 1976). As purine and pyrimidine are nitrogen sources kept in cells, the deficiency of nitrogen source in the medium might directly or indirectly disturb purine and pyrimidine biosynthesis in yeast. Furthermore, purine and pyrimidine and their byproducts are associated with the synthesis, relocation, and consumption of kept energy in energy metabolism (Henderson et al., 1973).

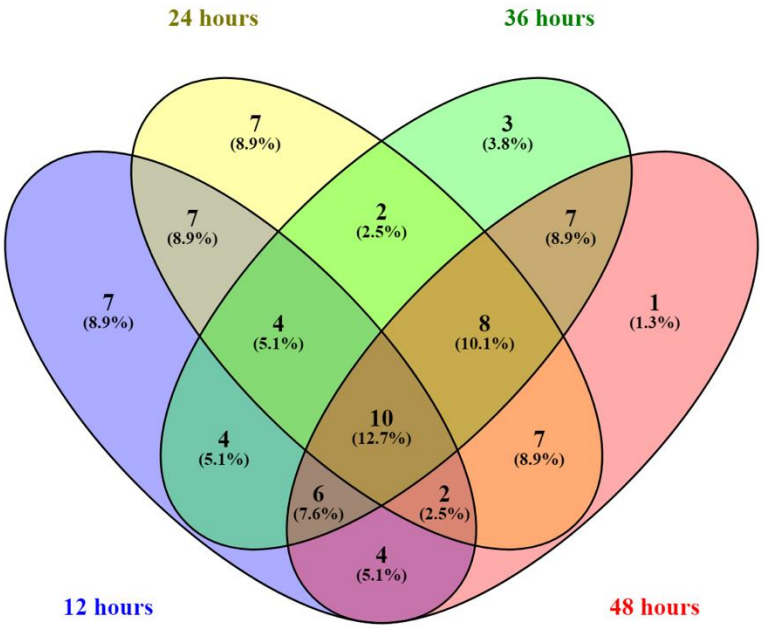


Figure 2- 5. Venn diagram of selected metabolites based on principal component loading scores. The metabolites with high scores from PCA 1 and 2 at each cultivation time were included in the analysis. A list of metabolites in each subset is provided in Table S5.

The principal component loading scores from PC1 and PC2 at each time were transformed to an absolute value to rank their influence on the separation of the data point on the PC scatter scheme. Metabolites with a PC score greater than 75th percentile for the identical cultivation time and principal component were select for Venn diagram building, as displayed in Figure 2-5. I found ten common metabolites at every cultivation time, i.e., ADP, arginine, F6P, lysine, trehalose, IMP, BPG, FAD, PEP, and NADP. These metabolites with high PCA scores through the cultivation period have a broad range of variability with respect to C:N ratios in every phase of growth. Trehalose is a non-reducing disaccharide found universally in fungi and is generally found in bacteria and animals (de Virgilio et al., 1994). Trehalose was first believed to be a reserve carbohydrate in yeast. Nonetheless, its character as a stress protectant in yeast has been acknowledged (Hottiger' et al., 1994). The main role of trehalose is to guard the cytosol against unfavorable conditions, such as dehydration, cold, and heat. In the budding yeast *S. cerevisiae*, trehalose concentrations are low throughout rapid growth periods and rise during periods of gradual or no growth (Leslie et al., 1994). These results coincide with our results for *Y. lipolytica*, excluding the condition where carbon or nitrogen was limited. As reviewed above, purine and pyrimidine metabolism were disturbed. Particularly, IMP and ADP were amid the metabolites with high PCA scores at all times. It is probable that these metabolites were at the heart of the biological reaction in purine metabolism. Therefore, the result of changes in conditions could be detected in these metabolites. Subsequently, FAD, which needs GTP from the purine biosynthesis pathway, was also disturbed. For arginine, carbamoyl phosphate, a precursor metabolite for arginine biosynthesis, was frankly involved in pyrimidine metabolism and nitrogen metabolism. The high PCA score for arginine may suggest the influence of the upstream pathway. PEP, F6P, and BPG are crucial intermediates in the TCA

cycle, proposing that glycolysis/gluconeogenesis or associated metabolites in the pathway were used differently under different C:N ratios; subsequently, lysine, which is linked to central metabolism, was also disturbed. Lastly, NADP is a cofactor in anabolic bio-reactions, like the Calvin cycle and lipid and nucleic acid syntheses, which involve NADPH as a reducing agent. As the carbon and nitrogen sources were interfered, the NADP level altered significantly. The precise mechanisms of how the nitrogen-limiting state affects the glucose utilization in *Y. lipolytica* were ambiguous, but I hypothesized that the lesser energy molecule might cause the ATP-related pathway in *Y. lipolytica* to slow down.

2.2.3 Intracellular Metabolome Analysis of Different *Yarrowia* spp. in Normal and Nitrogen-Limiting Conditions

The time-course investigation of the *Y. lipolytica* PO1d metabolomics profile delivers information about metabolic dynamics regard to altered C:N ratios and time to further study the responses of *Yarrowia* strains to nitrogen-limiting conditions. Based on growth, glucose depletion, and metabolomic profile, I decided that the suitable time for sample collection, for this reason, was 36 h, as the metabolomic data of C:N 4:1 was overlaid with that for C:N 5:0 at 48 h (Figure 2-6), proposing that the nitrogen at 48 h for C:N 4:1 was exhausted. Furthermore, the growth of C:N 5:0 at 24 h was inadequate. To study the reaction to nitrogen-limiting conditions, a normal C:N ratio of 4:1 was chosen for evaluation against the nitrogen-limiting condition of 5:0 as the C:N 5:0 provided the distinct metabolome profile through the cultivation. The growth profiles of distinctive strains established that *Yarrowia* spp. have comparable typical growth profiles (Figure 2-7). In this analysis, I examined the reactions of six *Yarrowia* strains, *Y. lipolytica* PO1d, *Y. deformans* JCM 1694, *Y. keelungensis* JCM 14894, *Y. lipolytica* JCM 2304, *Y. lipolytica* JCM 21924, and *Y.*

lipolytica JCM 8061. The whole arithmetical results of all identified normalized peak areas of six *Yarrowia* spp. in standard and nitrogen-limiting conditions are given in Table S6.

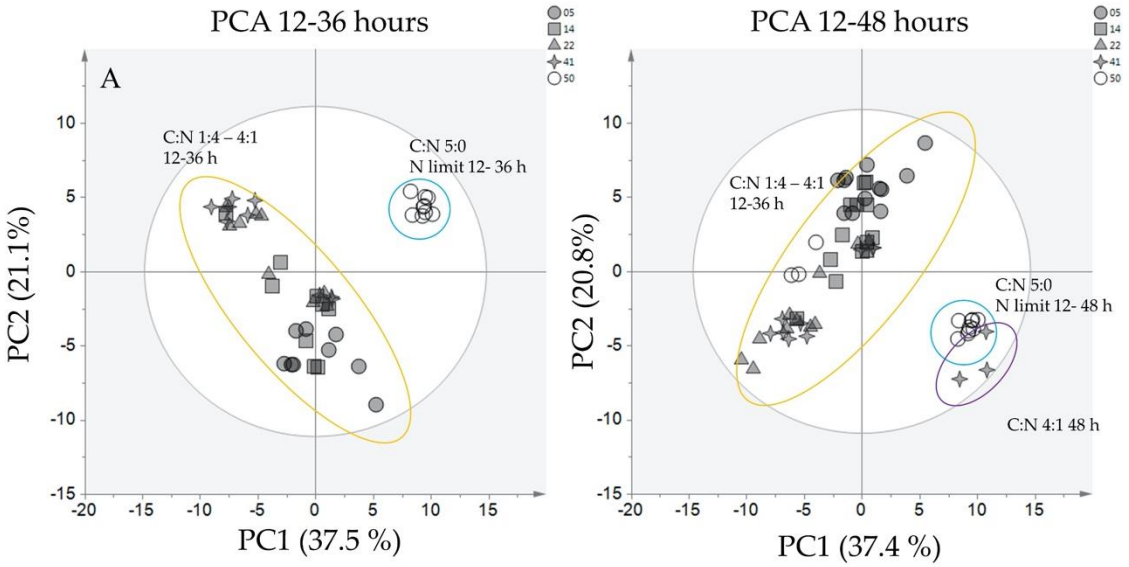


Figure 2- 6 All-time point score plot (A) all samples from 12 – 36 hours of cultivation (B) all samples from 12 – 48 hours of cultivation.

The results demonstrated that the C:N 4:1 cluster with the nitrogen-limiting condition at 48 hours. Thus the 36 hour time point is the latest suitable time point for nitrogen-limiting evaluations. As shown in Figure 2-8, PC1 separated the yeast cultivated under normal and nitrogen-limiting settings, while PC2 divided each strain of *Yarrowia*. While every yeast in this study came from the genus *Yarrowia*, the outcomes proposed that each *Yarrowia* strain has an exclusive metabolomic profile, including differences at the subspecies. The high PC score for PC1 corresponds with the earlier analysis of *Y. lipolytica* PO1d, in which many of the metabolites with high PC scores belonged to purine and pyrimidine metabolism, proposing that *Yarrowia* spp. react to nitrogen-limiting settings in a similar way.

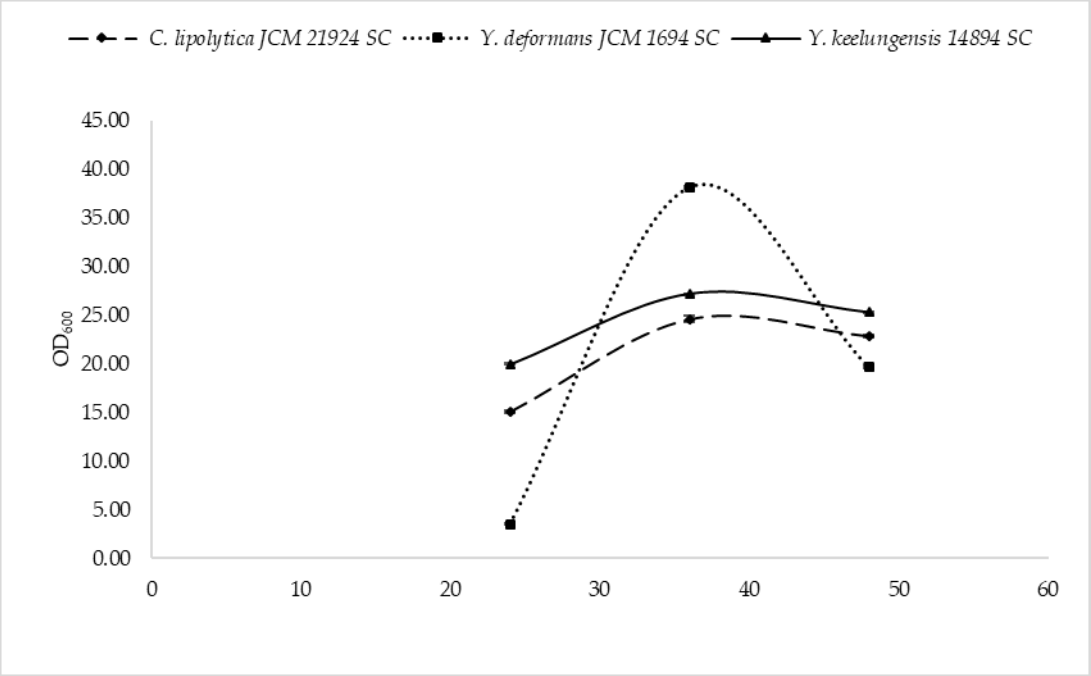


Figure 2- 7. Growth profile comparing *Y. lipolytica* spp. Error-bar represents the standard deviation from three replicates; *Y. deformans* JCM 1694 (square with dot line), *Y. keelungensis* JCM 14894 (triangle with dark line), C. lipolytica JCM 21924 (diamond with dash line).

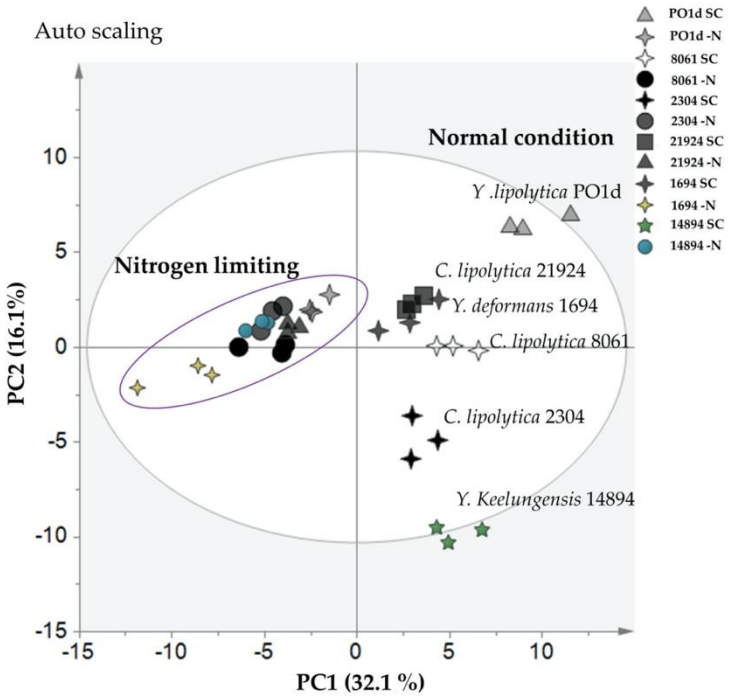


Figure 2- 8 PCA results for six *Yarrowia* spp. cultivated in normal and nitrogen-limiting conditions. PCA score plot indicating differences in metabolite profiles for two carbon-to-nitrogen ratios based on 100 metabolites. The grey ellipse indicates a 95% confidence border based on Hotelling's T2. The purple ellipse indicates the nitrogen-limited samples. Each point represents one replicate; *Y. lipolytica* PO1d in normal conditions (light grey triangles), *Y. lipolytica* PO1d in nitrogen-limiting conditions (light grey four-point stars), *C. lipolytica* JCM 8061 in normal conditions (white four-point stars), *C. lipolytica* JCM 8061 in nitrogen-limiting conditions (black circles), *C. lipolytica* JCM 2304 in normal conditions (black four-point stars), *C. lipolytica* JCM 2304 in nitrogen-limiting conditions (dark grey circles), *Y. lipolytica* JCM 21924 in normal conditions (dark grey squares), *Y. lipolytica* JCM 21924 in nitrogen-limiting conditions (dark grey triangles), *Y. deformans* JCM 1694 in normal conditions (dark grey four-point stars), *Y. deformans* JCM 1694 in nitrogen-limiting conditions (yellow four-point star), *Y. keelungensis* JCM 14894 in normal conditions (green five-point stars), *Y. keelungensis* JCM 14894 in nitrogen-limiting conditions (blue circles). Right, top 25% of metabolites based on PCA scores from each principal component. The full list of PCA loading scores is presented in Table S7.

I further examined variances among *Yarrowia* strains under nitrogen-limiting settings. As displayed in Figure 2-10, PC1 divided *Y. deformans* JCM 1694 and the remaining taxa, whereas PC2 divided the following three clusters: (1) *Y. keelungensis*, JCM 14894 (2) *Y. lipolytica* PO1d, *C. lipolytica* JCM 21924, and *C. lipolytica* JCM 8061, and (3) *C. lipolytica* JCM 2304. The high PC1 score proposed that *Y. deformans* JCM 1694 has a distinctive response to nitrogen-limiting settings than those of other *Yarrowia* spp. Remarkably, PC2 separated not only the diverse *Yarrowia* species but also the subspecies *C. lipolytica* JCM 2304; this outcome indicated that *Yarrowia* is a very varied genus even at the subspecies level. The outcomes for all identified metabolites are brief in Figure 2-9. Amazingly, the metabolites from PC2 contain MEP and MEcPP; (Table S7). These metabolites are found in the MEP pathway, which has not been described in this yeast genus. Though, Soliman et al. (Soliman et al., 2011) noticed MEP-like pathways in fungi. I hypothesized that the nitrogen-limiting condition reduces purine and pyrimidine activity in yeast, which affects the energy availability in the cell. This pressure leads to a cellular alteration to the less ATP-dependent pathway, i.e., a shift from the MVA pathway to the ATP demanding MEP

pathway.

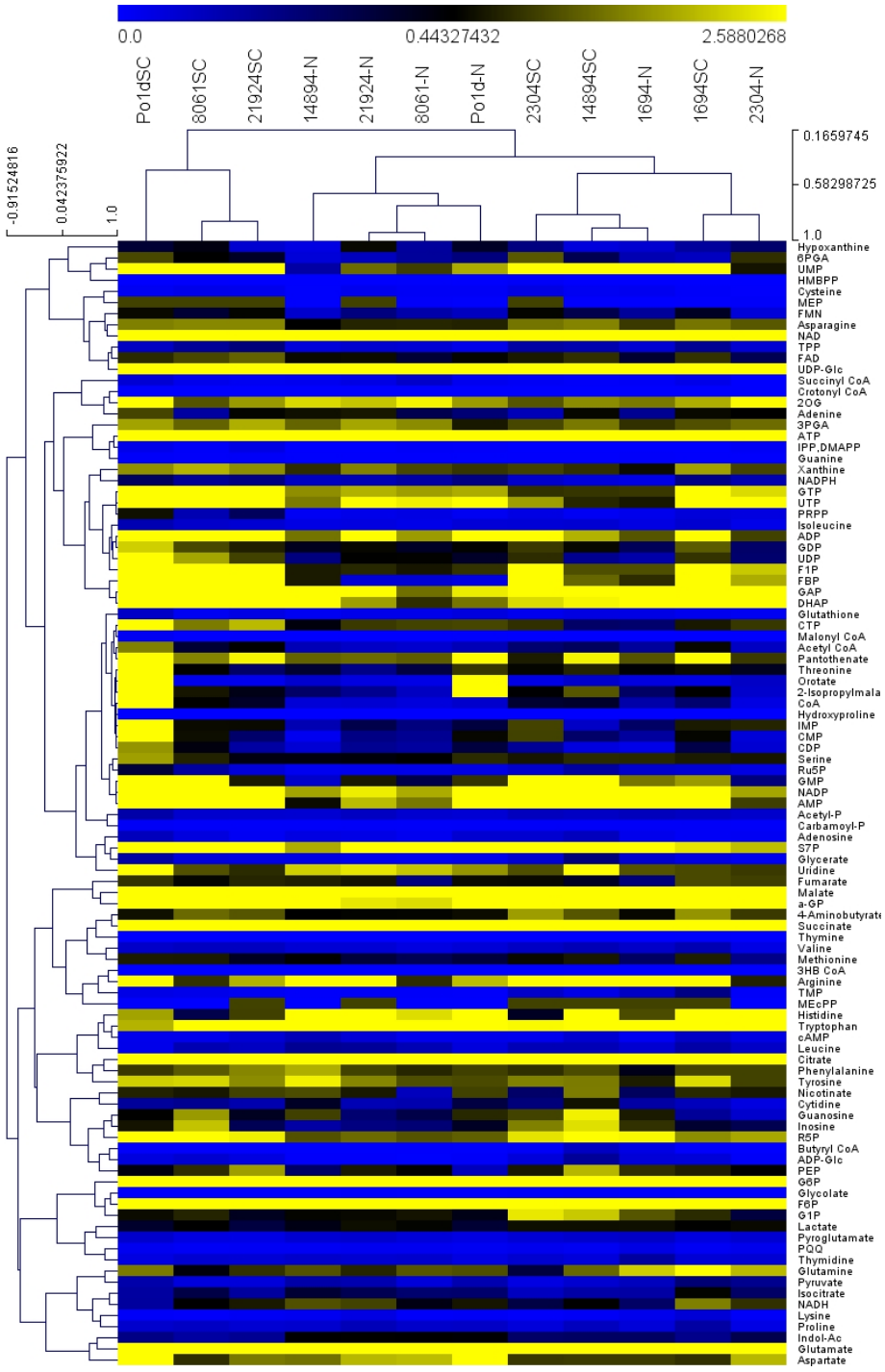


Figure 2- 9. Heat map describes the average normalized peak areas metabolome profile of 101 metabolites from 6 *Yarrowia* spp. cultivated in two conditions sample was clustered hierarchically by a complete linkage method using the MeV software ver. 4.90. SC indicates synthetic complete medium, -N indicates the removal of ammonium sulfate.

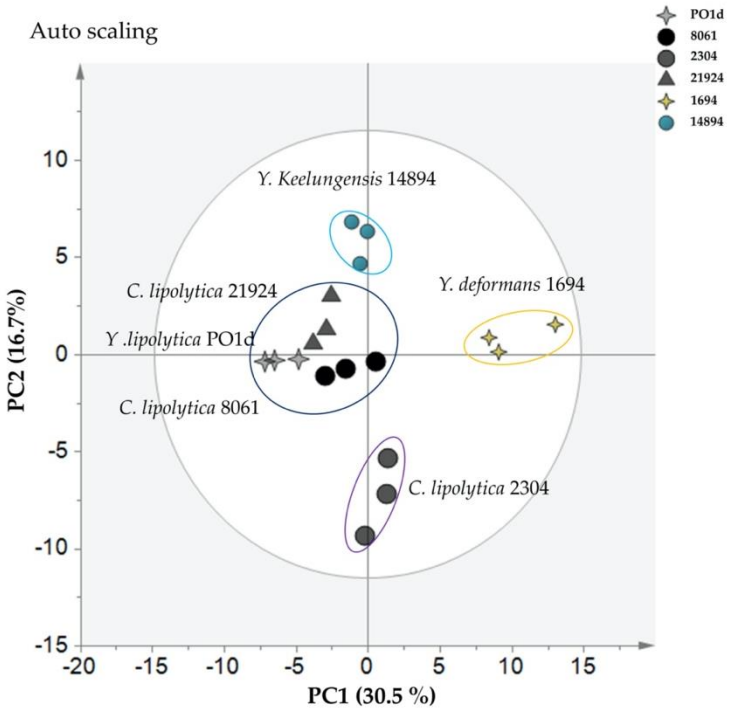


Figure 2- 10. PCA results for six *Yarrowia* spp. cultivated in normal and nitrogen-limiting conditions; (left) PCA score plot indicates differences in metabolite profiles among strains in the nitrogen-limiting condition based on 100 metabolites. The grey ellipse indicates a 95% confidence border based on Hotelling's T2. The cyan ellipse indicates samples from *Y. keelungensis* JCM 14894, the yellow ellipse indicates samples from *Y. deformans* JCM 1694, the purple ellipse indicates samples from *C. lipolytica* JCM 2304, the blue ellipse indicates samples from *C. lipolytica* JCM 21924, 8061, and *Y. lipolytica* PO1d. Each point represents one replicate; *Y. lipolytica* PO1d (light grey four-point stars), *C. lipolytica* JCM 8061 (black circles), *C. lipolytica* JCM 2304 (dark grey circles), *C. lipolytica* JCM 21924 (dark grey triangles), *Y. deformans* JCM 1694 (yellow four-point stars), and *Y. keelungensis* JCM 14894 (blue circles); (right) top 25% of metabolites based on PCA scores from each principal component. The full list of PCA loading scores is presented in Table S7.

Peak areas for all identified metabolites of purine and pyrimidine metabolism are reviewed in Figures 2-11 and 2-12, respectively. These results evidently reveal that *Yarrowia* strains react to the nitrogen-limiting condition in a comparable way; particularly, R5P, ADP, ATP, GDP, and GTP in purine metabolism were comparable among strains. On the other hand, each species showed discrete metabolomic profiles for other metabolites belong to purine metabolism and all other measurable metabolites from pyrimidine

metabolism. The laboratory *Y. lipolytica* PO1d strain had a discrete profile in which the orotate, PRPP, IMP, CTP, CDP, and CMP levels were extremely high compared to other strains. Overall, the levels of metabolites in purine and pyrimidine metabolism were lesser under nitrogen-limiting settings than under normal settings; however, *C. lipolytica* JCM 8061 had a noteworthy pattern in which metabolite profiles in the nitrogen-limiting settings were inverted compared to those of other strains involving glutamine, adenine, and uridine. *C. lipolytica* JCM 2304 had a distinctive profile in which several metabolites in purine and pyrimidine metabolism remained unaffected by nitrogen-limiting settings, comprising of R5P, guanine, cAMP, xanthine, hypoxanthine, UTP, uridine, and PRPP. *C. lipolytica* JCM 21924 metabolites under nitrogen-limiting settings, comprising of adenine and uridine, exhibited the opposite patterns compared to those of other strains; though, unlike JCM 8061, glutamine was unaffected. *Y. deformans* JCM 1694 had the top glutamine content, and intensities of all metabolites in purine and pyrimidine metabolism were lesser in the nitrogen-limiting settings. *Y. keelungensis* JCM 14894 was the one strain in which metabolites from purine metabolism were generally unaffected; however, metabolites of pyrimidine metabolism were fairly comparable to those of other strains. The diverse profiles of *Yarrowia* spp. perceived in this study can be used for metabolic manipulations for different uses in the future. Based on these outcomes, to engineer microorganisms for a specific application, it is essential to contemplate the exact metabolic traits of the strain for optimization. I exhibited that metabolomics is a great tool for this purpose.

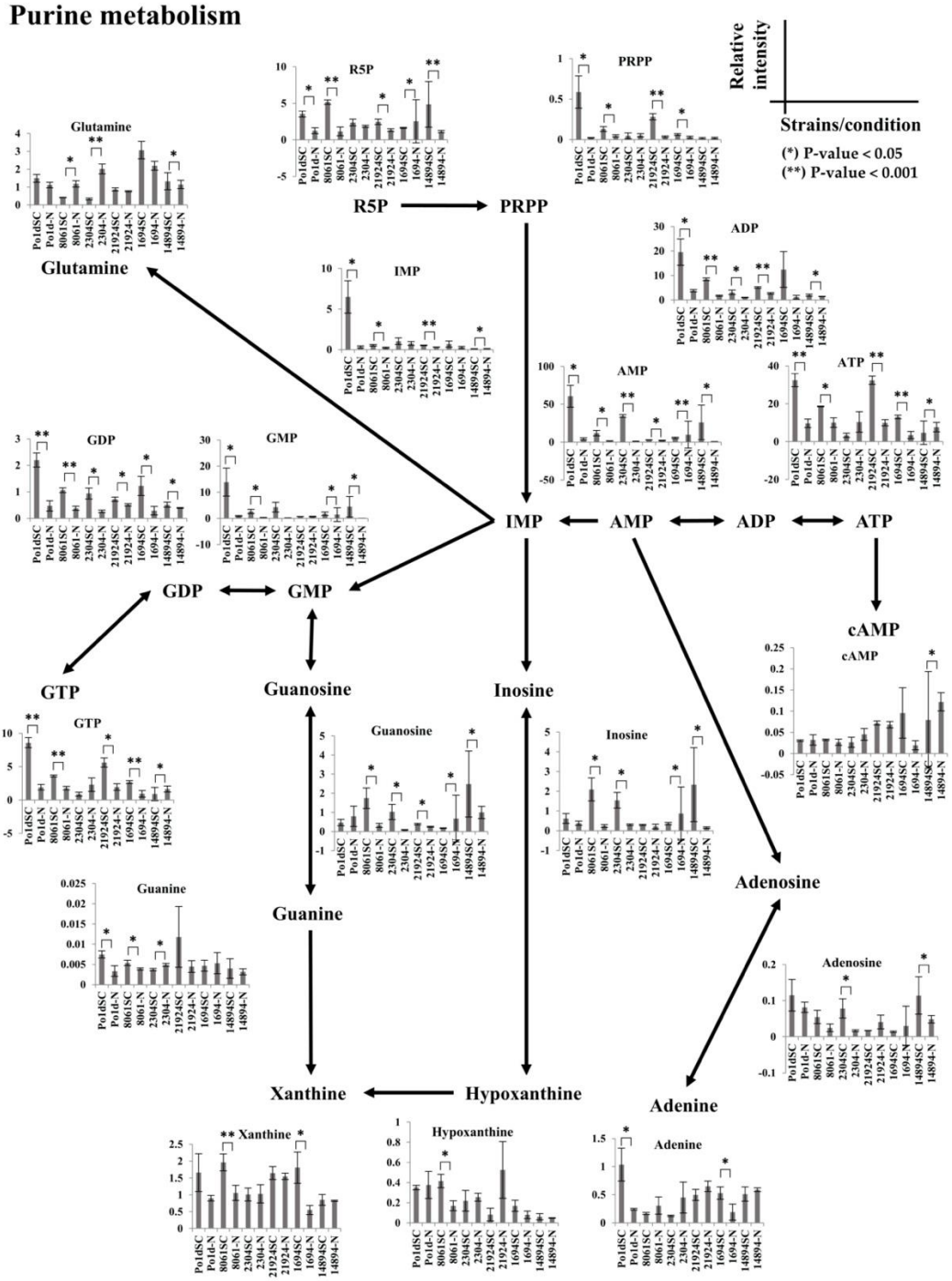


Figure 2- 11 Normalized peak areas for all detected metabolites in purine metabolism. The pathway skeleton was obtained from the KEGG pathway database (Kanehisa et al., 2000). SC indicates C:N 4:1, -N indicates C:N 5:0. Error bars indicates standard deviation, * indicates p-value < 0.05, ** indicates p-value < 0.001.

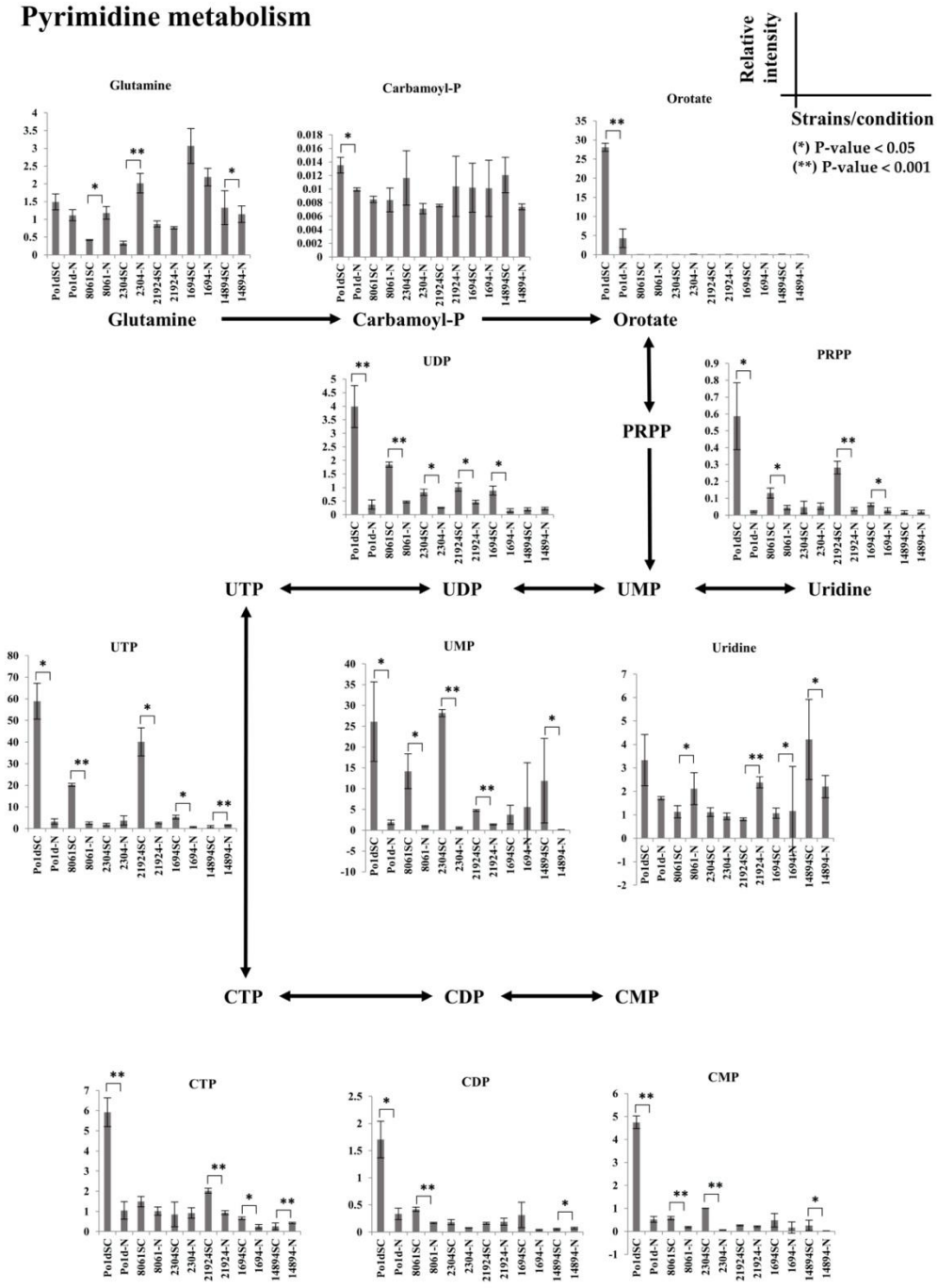


Figure 2- 12 Normalized peak areas for all detected metabolites in pyrimidine metabolism. The pathway skeleton was obtained from the KEGG pathway database. SC indicates C:N 4:1, -N indicates C:N 5:0. Error bars indicates standard deviation, * indicates p-value < 0.05, ** indicates p-value < 0.001.

2.4 Conclusions

Metabolomics analysis has shown that *Yarrowia* is a very varied genus, although the reaction to the nitrogen-limiting settings is comparable among strains. However, despite the resemblance, each strain had a distinctive metabolome profile. Nitrogen-limiting settings cause lower purine and pyrimidine metabolisms in *Yarrowia* spp., which disturbed the accessibility of energy molecules in the cell. This pressure might lead to the regulation of a less ATP-dependent pathway. This data will be useful for the improvement of *Yarrowia* strains for further scientific and engineering uses.

Chapter 3
Investigation of a non-mevalonate-like pathway in *Yarrowia lipolytica*

3.1 Introduction

Methyl erythritol phosphate (MEP) is the metabolite located in the MEP pathway for isoprenoid biosynthesis, which is well-known to be used by plants, algae, and bacteria. In this work, an extraordinary observation was discovered in the oleaginous yeast *Y. lipolytica*, in which one of the chromatographic peaks was interpreted as MEP when cultivated in the nitrogen limiting settings. This phenomenon is conflicting with an established understanding concerning isoprenoid biosynthesis in yeast. This discovery raised a curious hypothesis of whether *Y. lipolytica* exploits the MEP pathway for isoprenoid biosynthesis since there is no information on yeast harboring the MEP pathway. Here, I present a further examination of the presence of the MEP pathway in *Y. lipolytica.* four independent methods were used to explore the presence of the MEP pathway in *Y. lipolytica*; the spiking of the analytical standard, the addition of MEP pathway inhibitor, the ^{13}C labeling assimilation analysis, and the *Y. lipolytica* reference genome examination.

3.2 Methods

3.2.1 Yeast strain, cultivation, and sample collection

Y. lipolytica PO1d acquired from the National University of Singapore were cultivated for 17 hours in synthetic complete medium (2%w/v glucose, 0.19%w/v drop-out mix without uracil, 0.64%w/v yeast nitrogen-based without amino acid, 0.00076 %w/v uracil) in 100 mL Erlenmeyer flasks at the temperature of 30°C and the rotating speed of

200 rpm. After that the cells were injected into 20 mL SC medium deprived of ammonium sulfate (nitrogen limiting settings). For the sample collecting procedure, 5 OD_{600} units of cells were harvested at 36 hours by fast filtration by the 0.45 μm pore size, 47 mm diameter nylon membrane (Millipore). The cells were submerged in liquid nitrogen instantly for metabolism quenching and maintained at -80°C until further use. For extraction, 1.8 mL of extraction solution (methanol/water/chloroform = 5:2:2 v/v/v, with the addition of 20 μg/L of (+)-10 camphorsulfonic acid as an internal standard) was added to 2-mL sample tube and leave at -30°C for 1 h. After that, 700 μL of the mixture was transported to a tube including 350 μL of ultrapure water. The mixture was mixed by means of a vortex mixer and centrifuged at 16000 × g for 3 min at 4°C to isolated polar and non-polar phases. 700 μL of the polar phase was transported to a new cylinder by syringe filtration (0.2 μm PTFE hydrophilic membrane, Millipore). The sample was centrifugally condensed for 2 hours and freeze-dried and kept at -80°C until further use. After reconstituting in 100 μL ultrapure water, the sample was centrifuged at 16000 × g for 3 min at 4°C and moved to an LC glass vial. All experiment was done in triplicates.

3.2.2 LCMS analysis for MEP

The MEP standard was acquired from Sigma-Aldrich (product number 52131). The samples were examined grounded on the method adopted from Dempo et al. using a Nexera UHPLC system (Shimadzu, Kyoto, Japan) connected to LC/QqQ/MS 8030 Plus (Shimadzu). The column was CERI ODS L-column 2 metal free type (150 mm × 2.1 mm, particle size 3 μm), the mobile phase was (A): 10 mM tributylamine and 15 mM acetate in ultra-pure water, (B): methanol. The flow rate was kept at 0.2 mL/min, and the column oven temperature was kept at 45°C. The mobile phase (B) was programed to increase from 0% to 15%, 50%, and 100% from 1.0 to 1.5 min, 3.0 to 8.0 min, and 8.0 to 10.0 min,

correspondingly; held until 11.5 min, decreased to 0% from 11.5 minutes and held at 0% until 20 minutes. The analysis mode was set to negative ion detection. The injection volume was set to 3 μL, probe position was kept at +1.5 mm, desolvation line temperature was kept at 250°C, heat block temperature was kept at 400°C, nebulizer gas flow was set to 2 L/min, and drying gas flow was set at 15 L/min. All analysis was performed in triplicates.

3.2.3 Effect of MEP pathway inhibitor analysis

To evaluate the DXR activity, the enzyme specific to the MEP pathway in *Y. lipolytica*, DXP reductoisomerase inhibitor (Fosmidomycin) was utilized. Fosmidomycin sodium salt hydrate ≥95% (NMR) was acquired from Sigma-Aldrich (product number F8682). *Y. lipolytica,* the main culture, was cultivated and harvested as stated above, using a culture medium with a limited nitrogen source with the supplement of 100 μM inhibitor. The samples were analyzed using the LCMS method above. All analysis was done in triplicates.

3.2.4 1-13C carbon isotope glucose labeled analysis

To further evaluate if the MEP pathway was working in *Y. lipolytica*, I analyze ergosterol in *Y. lipolytica*, which is the end product derived from DMAPP. The 1-13C D-glucose (98-99%) was acquired from Cambridge Isotope Laboratories (Item Number CLM-420-PK). *Y. lipolytica* PO1d were cultivated for 17 hours in synthetic complete (SC) medium (2%w/v glucose, 0.19%w/v drop-out mix without uracil, 0.64%w/v yeast nitrogen-based without amino acid, 0.00076 %w/v uracil) in 100 mL shake flasks at 30°C and 200 rpm. The cultured cells were injected into 20 mL 10% 1-^{13}C D-glucose synthetic-complete medium without ammonium sulfate to obtain the start OD_{600} of 0.1 in 100 mL flasks at 30°C and 200 rpm. For sample collection, the cell pellet was harvested at 36 hours by centrifugation at 4°C, 10,000 rpm, 10 minutes. The cell pellet was then immersed in liquid

nitrogen and subjected to lyophilization overnight before further use. For extraction, 1g of cell pellet was extracted with 20 mL of extraction solution (chloroform: methanol with the ratio of 2:1 v/v) for 2 hours with stirring. After that, the mixture was filtered with filter paper (Advantech No.6), the extraction was done three times. The mixture was evaporated until dry using a rotary evaporator. The ergosterol was separated from the crude extract by column chromatography; the fraction holding ergosterol was confirmed by GCMS. 0.1 mg. of lyophilized ergosterol fraction was suspended in 1mL methanol and analyzed using a Nexera UHPLC system (Shimadzu, Kyoto, Japan) connected to LC/QqQ/MS 8050 (Shimadzu) in positive ion detection mode. The column used was AQ-C18 inertSustain 3 µm 2.1 x 150 mm (GL science); the mobile phase was methanol: acetonitrile; ratio = 80:20, the flow rate was set to 0.2 mL/min, the oven temperature was set to 40°C, injection volume was set to 3 µL, LC program was set as isocratic flow. The collision energy for ergosterol fragmentation at q2 was set at -25 eV. All analysis was performed in triplicates. For qualitative analysis, the representative analysis was nominated for discussion.

3.3 Results

3.3.1 MEP detection and verification

The interpreted MEP from the liquid chromatography with tandem mass spectrometry (LC/QqQ/MS) was discovered in *Y. lipolytica* in which the nitrogen source was restricted. Remarkably, the chromatographic peak was comparable to the previous publication on the MEP study in *E. coli*, where a bigger peak with the identical MRM was found next to the MEP peak (González-Cabanelas et al., 2016). To verify that the potential MEP peak discovered in *Y. lipolytica* is equivalent to the analytical standard, the addition of the authentic MEP analytical standard was done. The outcomes indicated that the

interpreted MEP peak detected in *Y. lipolytica* was indistinguishable from the analytical MEP standard, thus meeting the standards for metabolite identification grounded on The Metabolomics Standards Initiative (MSI) for level 1 identification (Figure 3-1).

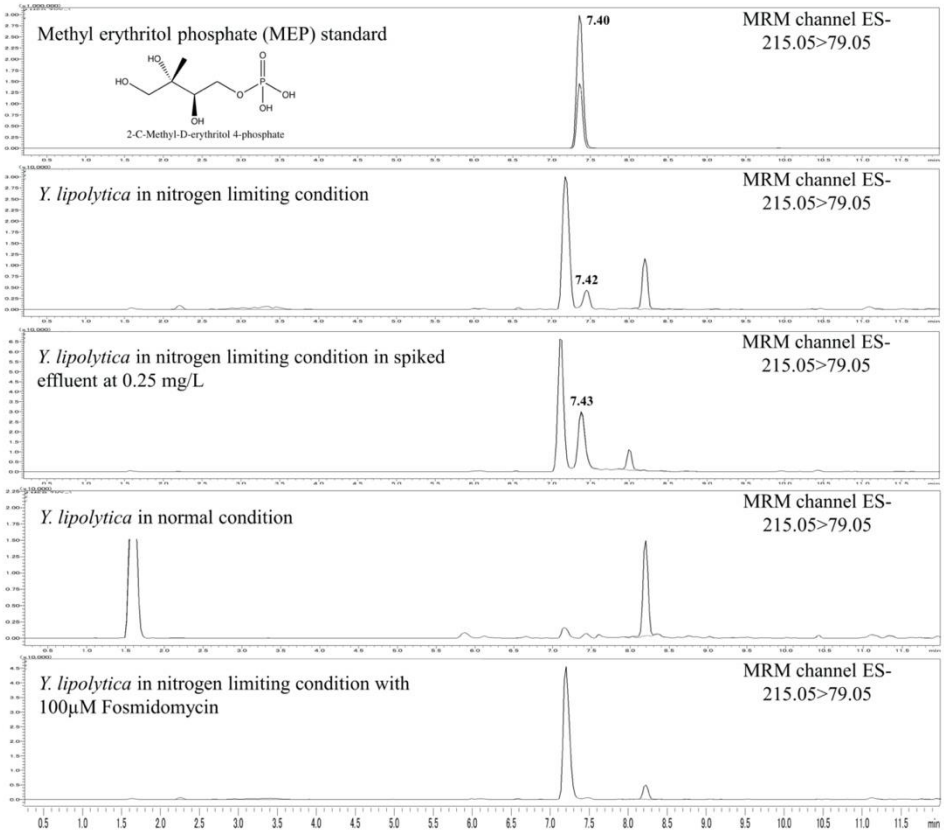


Figure 3- 1 LC/QqQ/MS analysis results; Multiple reaction monitoring (MRM) chromatograms of methyl erythritol phosphate (MEP) standard, *Y. lipolytica* in nitrogen limiting condition, *Y. lipolytica* in nitrogen limiting condition in spiked effluent at 0.25 g/L, *Y. lipolytica* in normal condition and *Y. lipolytica* in nitrogen limiting condition cultivated with 100 µM fosmidomycin.

3.3.2 Effect of MEP pathway inhibitor

The antibiotic fosmidomycin initially purifies from *Streptomyces lavendulae*, has been revealed to inhibit bacterial isoprenoid biosynthesis (Dhiman et al., 2005). It was later verified that fosmidomycin explicitly inhibits 1-deoxyxylulose 5-phosphate reductoisomerase (DXR) (Mueller et al., 2000; Shimizu et al., 1998), which catalyzes the

transformation of the DXP to MEP in the explicit step of the MEP pathway. In this work, *Y. lipolytica* cultivated in nitrogen limiting condition was cultivated with fosmidomycin and investigated with LC/QqQ/MS. The data show that the MEP peak was reduced with the addition of fosmidomycin. The result proposed that DXR-like activity exists in *Y. lipolytica* (Figure 3-1).

3.3.3 1-^{13}C carbon isotope glucose labeled experiment

To further evaluate whether the MEP-like pathway is workings in *Y. lipolytica*, I inspected the natural compound derived from DMAPP, which is a known compound from both the MVA and the MEP pathway. Ergosterol, which acts as a cell membrane stabilizer in yeast, is one of the richest natural products synthesized from IPP and DMAPP(Chiocchio et al., 2011; Fryberg et al., 1973). It has been recognized that assimilation of the 1-^{13}C-labeled glucose throughout the MVA pathway and the MEP pathway resulted in distinctive labeling configurations (Figure 3-2) (Rohmer, 1999). With this knowledge, I am able to decide the hypothetically labeled carbon location of the ergosterol from the MVA and MEP pathways (Figure 3-3).

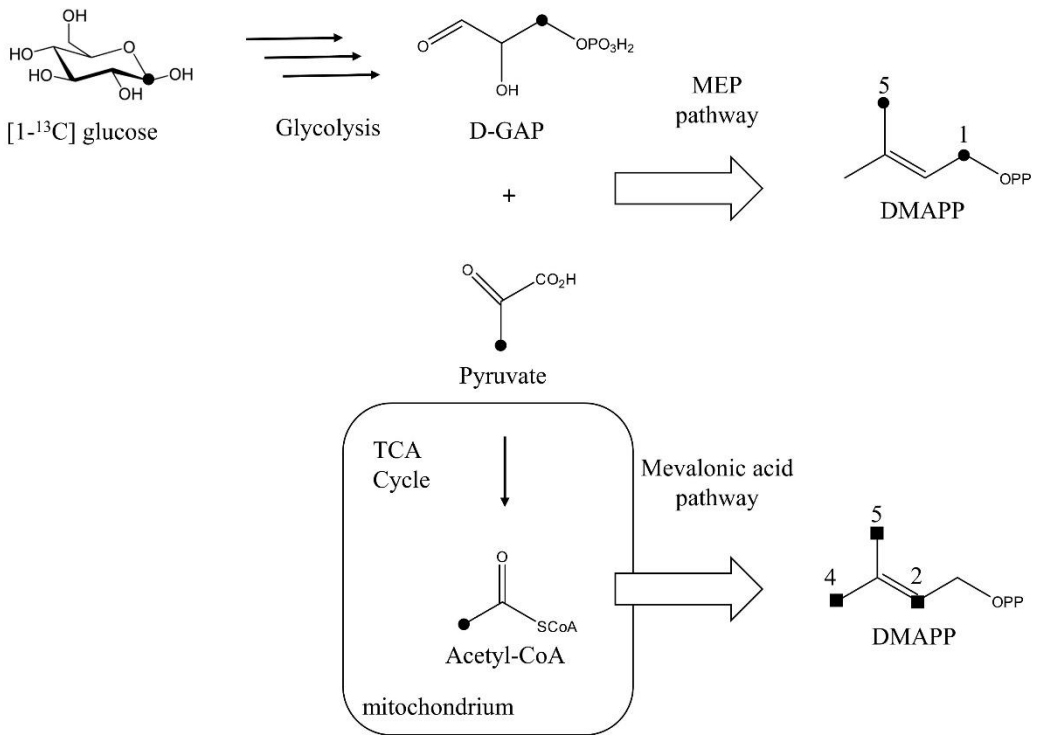


Figure 3- 2 Incorporation of [1-^{13}C] glucose into isoprenoids via the MEP pathway and the MVA pathway. In the MEP pathway, DMAPP is biosynthesized from one molecule of D-GAP and one molecule of pyruvate, each of which is derived from one molecule of glucose; one molecule of DMAPP is derived from two molecules of glucose. Therefore, the ^{13}C atoms derived from [1-^{13}C] glucose is incorporated at C1 and C5 of DMAPP. In the MVA pathway, DMAPP is biosynthesized from three molecules of acetyl-CoA, which are derived from three molecules of glucose; one molecule of DMAPP is derived from three molecules of glucose. Therefore, the ^{13}C atoms derived from [1-^{13}C] glucose is incorporated at C2, C4, and C5 of DMAPP.

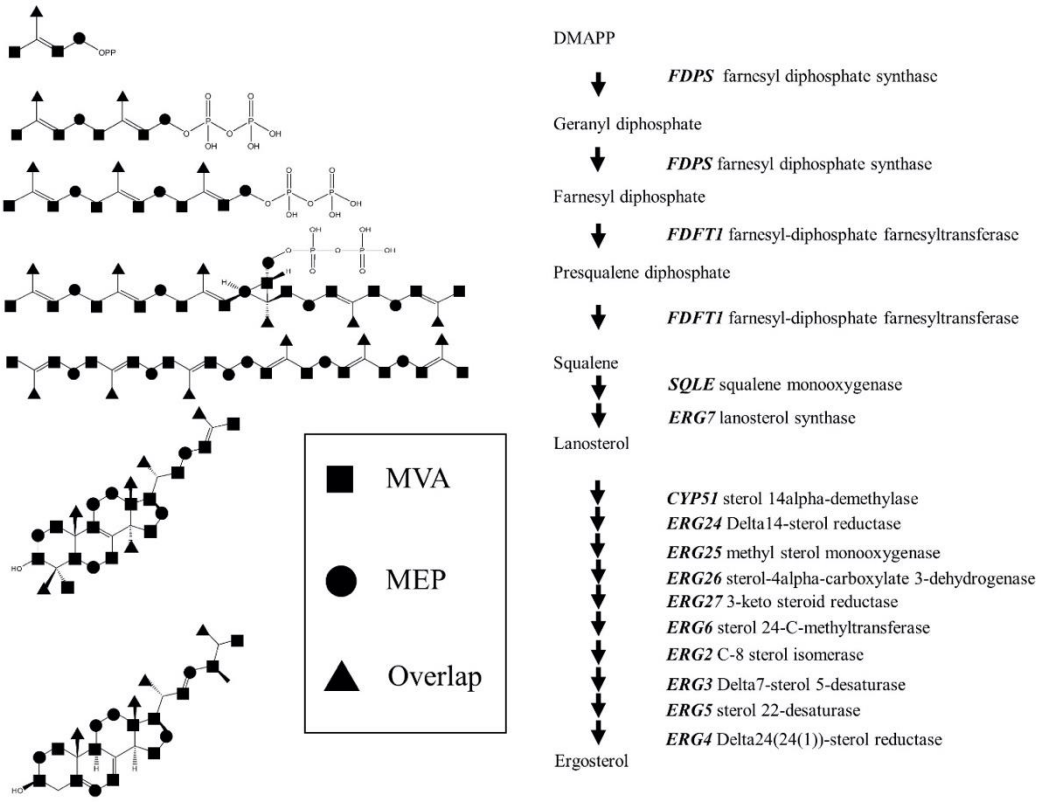


Figure 3- 3 ^{13}C incorporation into ergosterol; the circles indicate the theoretically labeled position by the MEP pathway, the squares indicate the theoretically labeled position by the MVA pathway, the triangles indicate the theoretically labeled position by both pathway

In this analysis, LC/QqQ/MS was used to examine the biological pathway involved in the synthesis of the ergosterol from *Y. lipolytica* cultivated in several conditions. All mass spectral information was obtained at 7.7 minutes, where ergosterol eluted. Typically, the main fragment of ergosterol identified in LC/QqQ/MS has a mass to charge ratio (m/z) of 379 (Headley et al., 2002). Once 1-^{13}C-labeled glucose was supplement to the medium of *Y. lipolytica*, the assimilation of the ^{13}C carbon isotope added the m/z of the fragment to the highest of 383. Therefore, this fragment was chosen as the parent ion for the product ion examination (Figure 3-4A). By means of CFMID 3.0, a mass spectra prediction program (Djoumbou-Feunang et al., 2019), I am able to computationally appoint the chemical structure to the product ion produced by fragmentation of the parent ion. The complete product ions are registered in Table S8. By relating the neutral fragment m/z with the labeled fragment m/z, it was feasible to calculate if the ergosterol were produced from the MVA pathway or the MEP pathway (Figure 3-4B).

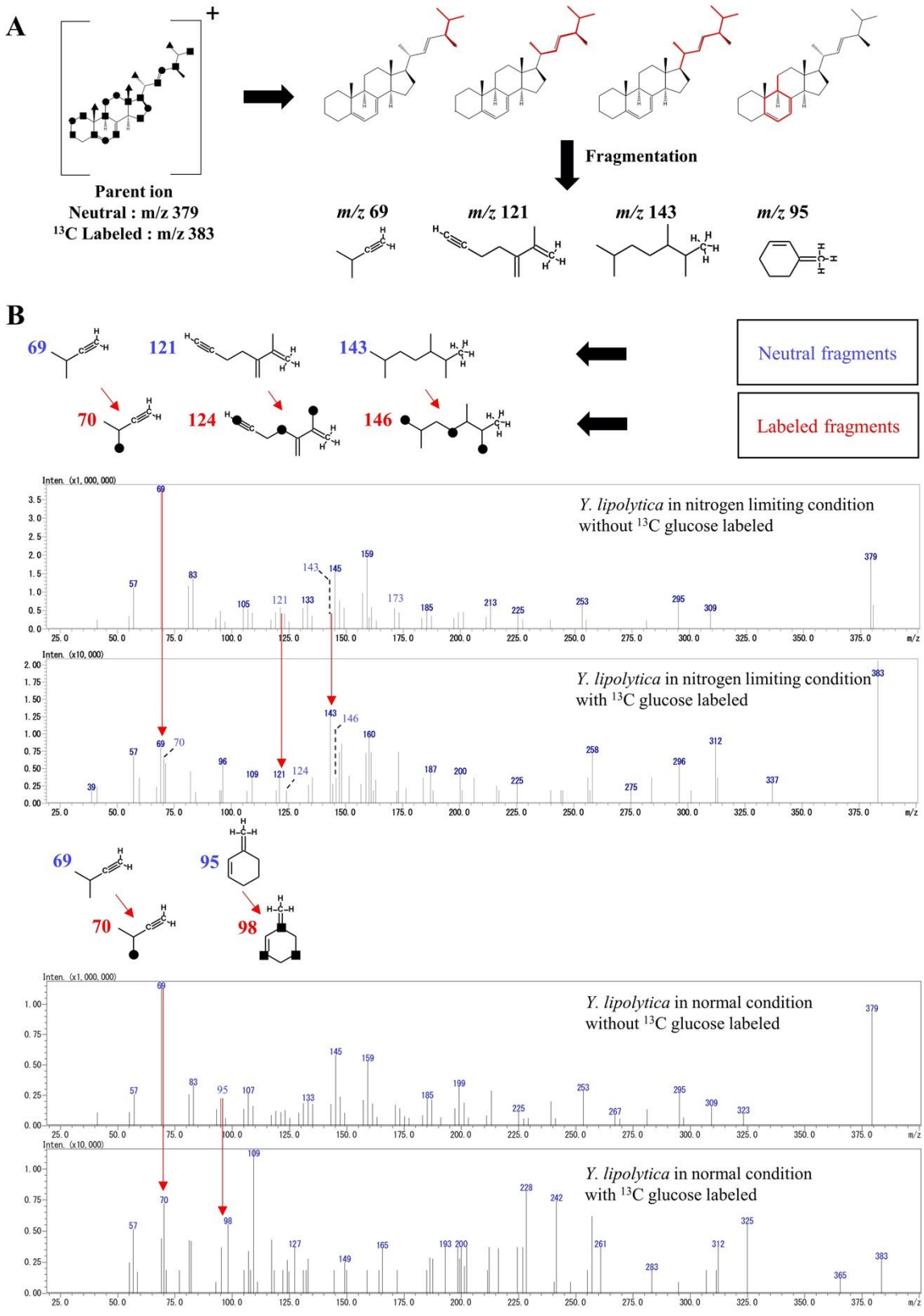


Figure 3- 4: Mass spectroscopic analysis of 13C labeled ergosterol. (A), the predicted fragment from parent ion; (B), the mass spectra comparing the neutral and the labeled ergosterol product ion. The assign fragment structure elucidates the m/z increased specifically to the MVA pathway or the MEP pathway; circles indicate the theoretically

labeled position from the MEP pathway, squares indicate the theoretically labeled position from the MVA pathway. All fragment structure assignment was done using CFMID 3.0 software. The predicted m/z and the structure of fragments resulted from the MEP pathway and the MVA pathway are found in Figure 3-5.

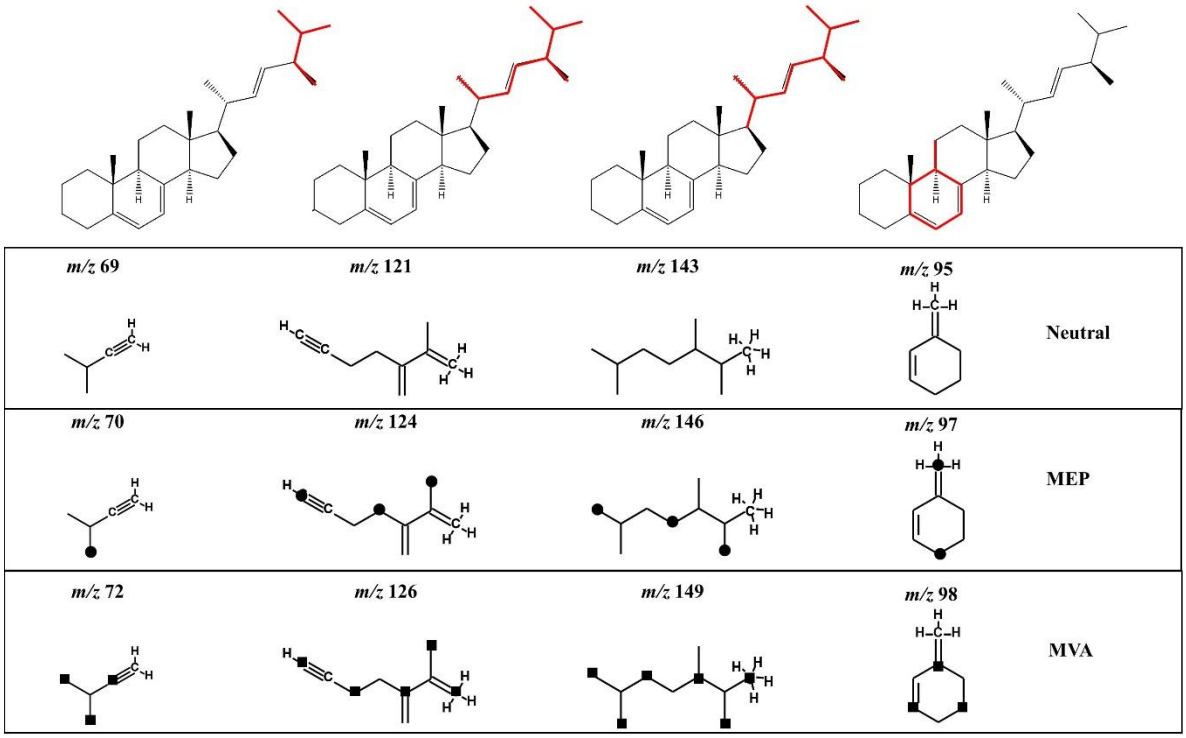


Figure 3- 5 The predicted fragment *m/z* and labeled positions.

In the settings where nitrogen source was restricted, three fragments mass alteration were discovered to match those of the peaks probably resulting from the MEP pathway; m/z 121 to 124, m/z 69 to 70, and m/z 143 to 146, while, in the standard settings, one fragment was found to be hypothetically originated from each pathway; m/z 69 to 70 of the MEP pathway then m/z 95 to 98 of the MVA pathway. The projected m/z and structure of fragments from the MEP pathway and the MVA pathway can be found in Figure 3-5. The fragment information shows that ergosterol from the two settings has a distinctive m/z alteration; this could be the consequence of the biosynthesis from distinct pathways. Combined with other data mentioned above, I concluded that ergosterol in *Y. lipolytica*

might be originated from both the MVA pathway and MEP pathway.

3.3.4 Genome analysis of *Y. lipolytica*

To evaluate the genetic potential for the MEP pathway in *Y. lipolytica*, a custom bioinformatics software was developed. In this analysis, a custom-made database for each enzyme in the MEP pathway was constructed. The conserved protein sequence of genes in the MEP pathway was download from the NCBI. The sequence was compared against the NCBI non-redundant protein databank using NCBI blast+ v.2.7.1. The genes that produce a meaningful alignment with an e-value of no more than 10e-6 were isolated to construct the custom MEP pathway genes database. The reference genome of *Y. lipolytica* was download from the NCBI FTP location in GenBank format. The genome data was handled using an in-house processing software created in Perl programming linguistic, resulting in 6,448 protein-coding sequences, 6,950 DNA sequences, and 6 chromosomes. All of the genes from the *Y. lipolytica* genome were compared with the MEP pathway enzyme custom database created prior using NCBI blast+ v.2.7.1. The genes that make a significant alignment with e-value less than 10e-6 and percent sequence identity more than or equal to 25 percent (Rost, 1999) were obtained for further analyzed by global alignment algorithm using CLUSTALW 2.0 (Larkin et al., 2007). Top *Y. lipolytica* genes that have significant alignment with genes from plants or procaryotes were chosen as candidate genes for the MEP pathway in *Y. lipolytica*. Though I discovered some candidate genes, none of the candidates genes have NADPH binding site (GSTGSIG), which is crucial for DXR activity. Hence the gene accountable for the DXR-like activity in *Y. lipolytica* is still ambiguous. The results could be found in the supplementary table S9-S14 (the first column shows the gene id in *Y. lipolytica*, the second column shows the gene in from the MEP pathway, the third column show the pair score from clustal W sequence alignment software, the fourth column indicate annotation information [indicate the origin of the gene] the fifth column

indicate taxon classification of the original organism).

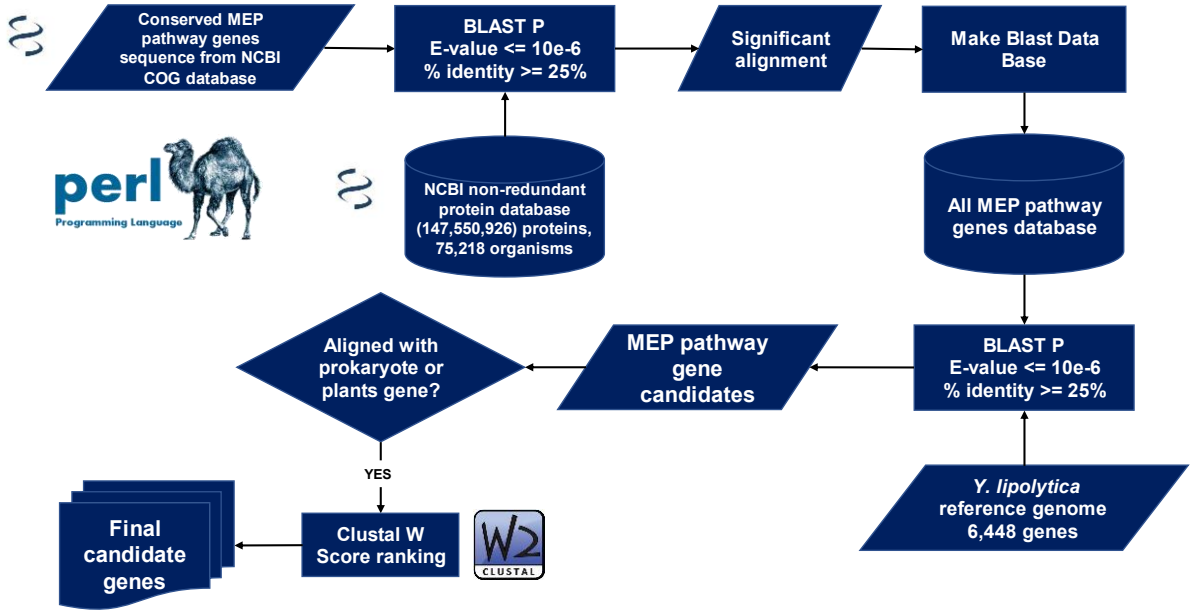


Figure 3- 6 Illustration of the developed in-house bioinformatics workflow for MEP pathway gene candidate based on sequence similarity.

With the bioinformatics workflow mention above the top gene candidate sequence are as follow;

DXS (1-deoxy-D-xylulose-5-phosphate synthase): *Y. lipolytica* :YALI0_E06479g

(XP_503628) aligned with *Sphingobium yanoikuyae* (WP_010337234.1)

```
CLUSTAL 2.1 multiple sequence alignment

WP_010337234.1      ----MTVSDTQLANAIRALSMDAVQAANSGHPGMPMGMADVATVLFKDYVKFDPAAPKWA
XP_503628.1         MAPQFSKTDETAINTIRTLAIDAVAKANSGHPGAPMGLAPVAHVLWNYYMNFTSSNPEWI
                        :: :*    *:**:*::***  ******* ***:* ** **:: *::* .: *:*

WP_010337234.1      DRDRFVLSAGHGSMLIYSLLNLTGYARPTMDDIRNFRQLGSPCAGHPENFELAGVEATTG
XP_503628.1         NRDRFILSNGHACMLHYSLLHLFGYD-ITIDDLKNFRQLNSKTPGHPEAETPG-IEVTTG
                    :****:** **..** ****:* **   *:**::*****.*  .****    . :*.***

WP_010337234.1      PLGSGLATAVGMAIAERHLNAEFG---DDLVDHRTWVIAGDGCLMEGINHEAIGLAGHLN
XP_503628.1         PLGQGVSNAVGFAIAQAHLGATYNKPGYDIINNYTYCIFGDGCMMEGVASEAMSLAGHLQ
                    ***.*::.***:***: **.* :.    *:::: *: * ****:***:  **:.*****:

WP_010337234.1      LSRLIVLWDDNKITIDGAVDLSSNEDVRARYAATGWHVVSCDG--HDVADVRRAIDEALA
XP_503628.1         LGNLITFYDDNHISIDGDTNVAFTEDVSQRLEAYGWEVIWVKDGNNDLAGMAAAIEQAKK
                    *..**.::***:*:*** .::: .***  *  * **.*:  ..  :*:*.:  **::*

WP_010337234.1      DPRPSLIACATKIGYGAPNKAGTSGVHGSALGEAEVAAAREFLGWT-AEPFVVPADIDAA
XP_503628.1         SKDKPTCIRLTTIIGYGSLQQGTHGVHGSPLKPDDIKQFKEKVGFNPEETFAVPKETTDL
                    .   .     *.*   .. : ** *****.*   ::   :* :*:.  *.*.** :
```

```
WP_010337234.1      WKAIGAKGGDVRKAWEGRLAN-----ASEGAEFSRRMAGELPAGFSLDAYIDSLIANPQK
XP_503628.1         YAKTIDRGANAEKEWNELFAKYGKEYPKEHSEIIRRFKRELPEGWEKALPTYTPADN--A
                    :     :*.:..* *:  :*:     ..* :*: **:  *** *:.      :   *

WP_010337234.1      VATRKASELALGAINDLLPETVGGSADLTGSNNTKTKSTGPLTR-----DDYSGRYVYYG
XP_503628.1         VASRKLSEIVLTKIHEVLPELVGGSADLTGSNLTRWKDAVDFQPPVTHLGDYSGRYIRYG
                    **:** **:.*  *:::*** ********** *: *.:  :       .******: **

WP_010337234.1      IREFGMACAMNGMALHGGVIPYGGTFLVFSDYMRAGIRLAALQEQRVIHVLTHDSIGLGE
XP_503628.1         VREHGMGAIMNGMNAYGGIIPYGGTFLNFVSYAAGAVRLSALSGHHVIWVATHDSIGLGE
                    :**.**.. ****  :**:******** * .*  ..:**:**. ::** * *********

WP_010337234.1      DGPTHQPIEHVMSMRMIPNLDVYRPADIVETAECWELALKDATGPSVLALTRQNLPQLRL
XP_503628.1         DGPTHQPIETVAWLRATPNLSVWRPADGNETSAAYYKAITNYHTPSVLSLTRQNLPQLEG
                    ********* *  :*  ***.*:****  **: .:  *:.:   ****:*********.

WP_010337234.1      EKAENLSAKGAYRLVAATADRKVVLLATGSEVEIAVATAKALEEQ-GIGADVVSMPSWAH
XP_503628.1         SSIEKASKGGYQLISEDKGD--IYLVSTGSEVAICVAAAKLLKEKKGITAGVISLPDWFT
                    .. *: *  *   :   ..*  : *::***** *.**:** *:*: ** *.*:*:*.*

WP_010337234.1      FDAQDAAYKADILPAGVLRASIEAGTTFGWERYTGLDGLRFGIDSFGASAPAEVLYDHFG
XP_503628.1         FEQQSLEYRKSVFPDGIPMLSVEVYSDFGWSRYS---HQQFGLDRFGASAPFQQVYDAFE
                    *: *.  *: .::* *:   *:*. : ***.**:     :**:* ****** : :** *

WP_010337234.1      LTAAKIAPQIKAALGA----------------------------
XP_503628.1         FNAEGVAKRAEATINYYKGQTVKSPIQRAFDPIDVNTRPGHGV
                    :.*  :* : :*::.
```

DXR(1-deoxy-D-xylulose-5-phosphate reductoisomerase): *Y. lipolytica* :

YALI0_B14267g (XP_500876.1) aligned with *Daucus carota* subsp. Sativus

(KZM93394.1)

CLUSTAL 2.1 multiple sequence alignment

```
XP_500876.1     MSGYDRALSVFSPDGHVFQVEYALEAVKRGTAAVGVKGKDVVILGCEKRSALKLQDPRIT
KZM93394.1      ----------------MHPHCCRSVR-------------------------------
                      :. .:*:

XP_500876.1     PSKICKLDTHVSLAFAGLNADARILVDKARVEAQSHRLTVEDPVSVEYISRYIAGVQQRY
KZM93394.1      --KIVNLDDHIALACAGLKADARVLINRARIECQSHKLTVEDPVTVEYITRYIACLQQKY
           ** :** *:.** ***:****:*:::**:*.***:*******:****:**** :**:*

XP_500876.1     TQSGG-VRPFGISTLIVGFDPNDKVPKLYQTEPSGIYNAWKANAIGKSSKTVRDFLEKNY
KZM93394.1      TQSGGTVRPVGLSTLIVGFDPHTGVPALYQTDPSGTFSAWKANATGRNSNSMREFLEKNY
          ***** ***.*:*********:  ** ****:*** :.****** *:.*:::*:******

XP_500876.1     EEDMDKAATIRLAIKSLLEVVQTGAKNIEIAILSADNPT-------EVLSTETIAEYVAE
KZM93394.1      TETSG-QETLKLAIRALLEVIAAIEAGKDIALSNKDTLIDGGPFVLPLAHKHKVKILPAD
          * .  *::***::****:: :  .:**:. *.      : ...:   *:

XP_500876.1     IEEEKEAEAKKQKPSRD------------------------------
KZM93394.1      SEHFAIFQCIQGLPEGAFRRIILTASGGAFRYNYEFVLICSYPYFIYF
           *.   :. : *.
```

ispD (2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase) : *Y. lipolytica* :

YALI0_E20031g(XP_504170.2) aligned with *Gonium pectoral* (KXZ55794.1)

CLUSTAL 2.1 multiple sequence alignment

```
XP_504170.2     ----------------------MSQNVLLVLRGTLEGHNGWVTSLATS-----------
KXZ55794.1      MQRWLNTRAAHNPEHQQQRPPACKTESVATLKHKLHNGHNYPVTCVTFSPRGDLLATASR
                        ::.*  : :   :***  **.:: *

XP_504170.2     --------------------------SNNPD--ILLSGSRDKSLIVWSLTR---------
KXZ55794.1      DIRLYDVPTGRLRAILDGHGGLVRCVAFTPDSRLLCSGGDDCSAMVWSLEQCLALADQTP
                          : .**  :* **. * * :**** :

XP_504170.2     ------------------------------------------------------------
KXZ55794.1      HAPPTLTPPGGFSSTNLPTNDQRPTSQSKVAAQADRSMGFGICTGAGSYWSHGAQEPPPS


XP_504170.2     --------------DDTNYGVPRKSLKG--------------------------------
KXZ55794.1      PGGGPKRIIRAAGADNVALTAQRSKLRASNSLLSLAGGTAASAGAGPTHGSHPMARVVQQ
                  *:.  . *..*:.

XP_504170.2     ----HSHIVQDCAISHDGAYAISGSWDKTLRVWDLKTGVSKDRFVGHTGDVLSVSFSPDN
KXZ55794.1      VLRRHHKPVLCLAVTPDSKYLVTGSADATVRVWDVATGYCMETLRGHAKEVNAVAVTPNA
              * : *   *:: *. * ::** * *:****: ** . : : **: :* :*:.:*:

XP_504170.2     RQIVSAGRDRTIKIWNTIGECKYTITDKGHEDWVSTVRFNPATKDEEEP-----------
KXZ55794.1      KRILSAGADGAVLVWDFTTGERK-RTLRGHKGAVSCMAVSPDGTWMVTGGHDKSLRIWDL
            ::*:*** * :: :*:    :  * :**:. ** : ..*  .

XP_504170.2     ----------------------SPNANTIISAGWDKKVKVWDLDTCTLKADHLGHTGYV
KXZ55794.1      LNGNQLAAIPAHGGAFGVLGCAVSPDGSTIMSGGYDHLARLWDIGSGLPLASLQGHREMV
                        **:..**:*.*:*: .::**:.:   *.  **   *

XP_504170.2     STITISPDGSLCASAGKDGSILLWDLT---------SSTALYTLPAGDEVHAVCFS----
KXZ55794.1      SAVAFSPDCTTLATAGRDDEVRMWIWVPHYSHSPVAKSTIKWEQAHPGEARALREGRVLY
            *::::*** :  *:**:*..: :*  .       .**  :  . .*.:*:  .

XP_504170.2     -------PNRYWLCAATSSSIKVFDLEKRVLIDELKTQTTS--------------------
KXZ55794.1      PTPDEQRGNRHWACLFGEQGGRTAEEEEQQAAMAARARARTGSVGRRPQTAPAARRWRAT
                **:* *   ... :. : *::     :::: :

XP_504170.2     ------------------------------------------------------------E
KXZ55794.1      LVVRASSNAPGGLEKGSVSVVLLAGGVGKRMGAAIPKQYLDLRGQPIATYSLETFARMPE
                                                                *

XP_504170.2     DGKEPEALSLTWSADGQNLFAGYTDNVIRVWQIMQSS----------------------
KXZ55794.1      VGEIVIVCDPSWRDVFEKRFPGLPPHLKLKWALPGPERQDSVFNGLQQVDAGAAIVAVHD
             *:  . . :*   :: *.* . ::   * :  ..

XP_504170.2     ------------------------------------------------------------
KXZ55794.1      SARPLVTAEDAAKCMADGLAVGAAVLGVQVKPTIKEVDKDLTVIKTLQRSKLWEVQTPQC


XP_504170.2     ------------------------------------------------------------
KXZ55794.1      IRPELLRAGFDLVKRENLEVTDDVSIIEAMGKPVKITPGAYTNIKVTTPDDMAVAEKFLE


XP_504170.2     -----------------
KXZ55794.1      ERAEAAAAAAAKQAVAA
```

ispE (4-diphosphocytidyl-2-C-methyl-D-erythritol kinase) *Y. lipolytica* : YALI0_D01089g (XP_502270.1) aligned with *Vibrio cholerae* (CSB35022.1)

CLUSTAL 2.1 multiple sequence alignment

```
XP_502270.1     ------------------------------------------------------------
CSB35022.1      MAVIESALRQGEVEAQIVELESSFYALLDGIKAQLPNLDDSAYQQQVHSSMTQLRQFVHG

XP_502270.1     ------------------------------------------------------------
CSB35022.1      ITLLGMCPDNVNARIISKGERVSIQLMKAVMEAKGLPANLVDPVKYLLAKGDHLEAMVDV

XP_502270.1     ------------------------------------------------------------
CSB35022.1      EISTQRFRQAPLPQQHVNIMPGFTAGNAQGELVCLGRNGSDYSAAVLAACLRADCCEIWT

XP_502270.1     ------------------------------------------------------------
CSB35022.1      DVDGVYNCDPRLVDDARLLKSLSYQEAMELSYFGASVLHPKTIAPIAQFQIPCLIKNSFN

XP_502270.1     ------------------------------------------------------------
CSB35022.1      PQGAGTLIGQDTGEDKLAIKGITTLSNLTMVNVSGPGMKGMVGMASRVFGAMSAADVSIV

XP_502270.1     ------------------------------------------------------------
CSB35022.1      LITQSSSEYSISFCIEAQHKALAQQALADAFELELKDGLLEPVEFVDNVAIITLVGDGMR

XP_502270.1     ----------------------------------------------------MSKAVN
CSB35022.1      TSRGVASQFFSSLAEVHVNVIAIAQGSSERAISAVIPDDKISEAIKACHENLFNSKHFLD
                                                .: ::

XP_502270.1     IAIIGTGLVGKAFINQLAAVKSSIAYNVVLIARS---SKTLISKDFKPLSLTNWESELNS
CSB35022.1      VFVVGVGGVGGELVDQIQRQQAKLAEKGIMMRVCGLANSKGMLLDSQGLPLEQWRDRMIN
                : ::*.* **  :::*:   ::.:* : :::  .   ... :  * : *.* :*...: .

XP_502270.1     SPVGAMSFTEINDFLKKSPLPVILVDNTSNEALANEYPTFVNSGISIATPNKKAFSSDLK
CSB35022.1      ADQAFSLENLVALVQRNHIINPVLVDCTSSDEIANQYADFLAAGFHVVTPNKKANTASMS
                :  .   . :  . ::  :  :*** **.: :**:*. *: :*: :.****** ::.:.

XP_502270.1     TWEAIFGGAEKSGALVYHEATVGAGLPVISTLNDLIATGDEVETVEGIVSGTLSYIFNEF
CSB35022.1      YYHQLRNVARHSRRKLMYETTVGAGLPVIENLQNLIAAGDELEKFSGILSGSLSFIFGKL
                :. : . *.:*   : :*:*********..*::***:***:*...**:**:**:**.::

XP_502270.1     STLSGSDVKFSDVVTKAKQLGYTEPDPREDLNGLDVARKVTILARLSGFDVESPTAFPVQ
CSB35022.1      DEG----MTLSQATQLAKEKGFTEPDPRDDLSGMDVARKLLILAREAGYELELS-DVDVE
                .    :.:*:..  **: *:******:**.*:*****: **** .*:.:* .  . *:

XP_502270.1     SLIPKPLETASSADEFLQKLPEYDADLAKLRDEAFAEKKVLRFVGSINKGTGKVEVGIQK
CSB35022.1      QALPAGFDASGSVEEFMARLAQADAAFAERVAQAKAEGKVLRYVAQIVDG--QCQVRIVA
                . :*  :::.*.:**: :*.: ** :*:  .* ** ****:*..* .*  : :* *

XP_502270.1     YDASHPFASLKGSDNIIAFKTKRYP-NPLVIQGAGAGDEVTAAGVLADVFKAAQRLAK--
CSB35022.1      VDENDPMFKVKEGENALAFYSRYYQPIPLVLRGYGAGSEVTAAGVFSDVMRTXXXXXXXX
                * ..*: .:* .:* :** :: *   ***::* *** .*******::**:::

XP_502270.1     ------------------------------------------------------------
CSB35022.1      XXXXXGLIEMSSGDMSVVVYAPASIGNVSVGFDVLGAAVSPIDGTLLGDRVKVEAGAEAF

XP_502270.1     ------------------------------------------------------------
CSB35022.1      TLKTAGRFVDKLPANPQENIVYDCWQVFARELEKKSVVLKPLTMTLEKNMPIGSGLGSSA

XP_502270.1     ------------------------------------------------------------
CSB35022.1      CSIVAALDALNQFHASPLDETELLALMGEMEGKISGSIHYDNVAPCYLGGVQLMLEELGI
```

```
XP_502270.1     ------------------------------------------------------------
CSB35022.1      ISQSVPSFDDWYWVMAYPGIKVSTAEARAILPAQYRRQDIVAHGRYLAGFIHACHTQQPE

XP_502270.1     ------------------------------------------------------------
CSB35022.1      LAAKMIKDVIAEPYREKLLPGFAKARNYAASAGALATGISGSGPTLFSVCKEQAVAERVA

XP_502270.1     ---------------------------------
CSB35022.1      RWLEQNYVQNEEGFVHICRLDKQGSKVTGSEL
```

ipsF (2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase): *Y. lipolytica* : YALI0_B14509g (XP_500887.1) aligned with *Parvimonas* sp. oral taxon 393(WP_009448629.1)

```
CLUSTAL 2.1 multiple sequence alignment


XP_500887.1         ------------------------------------------------------------
WP_009448629.1      MRIGLGFDVHELVVGRDLIIGGVKIEHSKGLLGHSDADVLVHAINDAIIGALCLGDIGKL

XP_500887.1         ------------------------------------------------------------
WP_009448629.1      FPDNDPKYKDISSLIMTKEVVRLMKEKGYKIGNVDSVICAQKPKLSGYILEMRKNIAKVL

XP_500887.1         -----------------------------------MTAAAKETFLFTSESVGEGHPDKI
WP_009448629.1      ETDIENISVKATTTEHLGFEGREEGISSHAVVLLEKFKEVKMKKILFTSESVTEGHPDKI
                                                       :. .  :.:******* *******

XP_500887.1         CDQVSDAILDACLKEDPLSKVACETAAKTGMIMVFGEITTKATLDYQKIVRDAIKHIGYD
WP_009448629.1      CDQVSDAILDEILKVDPVARVACETLTTTGIVMVVGEITTSKYVDIQSIVRNVLKEIGYT
                    **********  ** **:::***** :.**::**.*****.  :* *.***:.:*.***

XP_500887.1         DSEKGFDYKTCNVLVAIEQQSPDIAQGLH--------YEKALEELGAGDQGIMFGYATDE
WP_009448629.1      RAKFGFDASTCSVLTSINEQSRDIALGVDSALEYKKGNEDKYNSVGAGDQGMMFGYACTE
                     :: *** .**.**.:*::** *** *:.         *.  :.:******:*****  *

XP_500887.1         TEELLPLTVILSHKLNAALAAARRDGSLGWLRPDTKTQVTVEYANDNGATVPQRVDTVVI
WP_009448629.1      TDEFMPLPITLAHKLTKRLAYVRKNDILAYLRPDGKSQVTVEYIDG----TPTRIEAIVV
                    *:*::**.: *:***.  ** .*::. *.:**** *:****** :.    .* *::::*:

XP_500887.1         STQHSDDISTEDLRSEILEKIIKKVIPAKLLDDKTVYHIQPSGRFIIGGPQGDAGLTGRK
WP_009448629.1      STQHGPSVKTEQIEKDIKELVIDEVIPEHLIDENTKIYVNPTGRFEIGGPMGDSGLTGRK
                    ****. .:.**::..:* * :*.:*** :*:*::*  :::*:*** **** **:******

XP_500887.1         IIVDTYGGWGAHGGGAFSGKDFSKVDRSAAYAARWVAKSLVKAGLARRALVQFSYAIGVA
WP_009448629.1      LIVDTYGGFGRHGGGAFSGKDPTKVDRSGAYMARYIAKNIVASGICEKLEIGISYAIGVA
                    :*******:* ********** :*****.** **::**.:* :*:..:  : :*******

XP_500887.1         EPLSIFVDTYGTSKHSSAELVEIIKKNFDLRPGVIVKDLDLARPIYFKTASYGHFTNQ--
WP_009448629.1      KPLSIYVDTFGTGKISDDKIIEIINKVFDLRPAAIIDILDLRRPIYRQIAAYGHFGRNDL
                    :****:***:**.* *. :::***:* *****..*:. *** **** : *:**** .:

XP_500887.1         ENPWEQVRELKY-----
WP_009448629.1      DLPWEKLDKVDEIKELI
                    : ***:: ::.
```

ipsG (4-hydroxy-3-methylbut-2-enyl-diphosphate synthase): *Y. lipolytica* : YALI0_F09834g (XP_505224.1) aligned with *Auxenochlorella protothecoides* (XP_011400578.1)

CLUSTAL 2.1 multiple sequence alignment

```
XP_505224.1       ------MATTAAPILALLDEPNYDLKTHALESLDENITQLWAEVSDHVSQIEELYEDKDF
XP_011400578.1    MAQAIPAINTAEGVLALLAEDDDALKVYALQQINASIHQFWHQASGSIATIEALFEDEEF
                  .**  :**** * :  **.:**:.:: .* *:* :.*. :: ** *:**::*

XP_505224.1       PKPELAALVAS----KIYYNLGDYESSMKFALAAGKLFNFDSSSDYVETIVSKCIREYIE
XP_011400578.1    SHRELAALVASKARRLIFYYLGDLQEAQAYALGAGALFDVDSPSAFERTILDSCLDAYVA
                  .: ********    *:* *** :.:  :**.** **:.**.* : .**:..*:  *:

XP_505224.1       VSQQHYEDP---SAPVPDARLATVLDQMLQRSYAQGHLKLALGVALEARRLDIITKILET
XP_011400578.1    ARAGLAGDEPGDAPASLDPRLAATVQRVLDRCLADGQHEQAVGIALEARRLDVLAAAVLA
                  .    *   :..  *.***:.::::*:*. *:*: : *:*:********:::  : :

XP_505224.1       ADTETTVELVHYVLDCVVTVVS--SRRFKDEVLLELNNVILGLPQPDYHAICKIIVQLNS
XP_011400578.1    AADPPTLLAHALDAGARTVVERGFRRRALATLALLAYHIGFDLVDAQLHEFTRDAREGEG
                  *  .*:     .. ..*    **   : *  :: :.* :.: * : :  : :.

XP_505224.1       SEGASRLFKSLLASKDPDSTLIAYQCAFDLVASGSQQLLNEVAAEFKDNETLAKILSGVP
XP_011400578.1    DAAEEAGLETPLLGEDGKTDAGGATPAPELAP----------ELDAFLRILSGSV
                  . . . ::: * .:* .:   .  * :*..           : ::: :****

XP_505224.1       TCDYDLTFLNQNNNTDTQMLNRTKNSLDPKSSMNHSAVTFANAFLHAGTAQDSFIRSNLD
XP_011400578.1    SVELERQFLARACHADLQILKNIKAGVEARNSVCHSATVWANAIMHAGTGVDAFLRTNLE
                  : : :  ** :  ::* *:*:. * .::.:.*: ***..:***.:****. *:*:*:**:

XP_505224.1       WLSKASNWTKLSATAALGVIHKGNLSQGRR--ILAPYLPQGDNAASGSPYTRGGSLYALG
XP_011400578.1    WLAKATNWAKFGATAGLGVIQRGSVADGRSRSLLSPYLPAAPGSSTGSPYSEGGALYALG
                  **:**:**:*:.***.****.:*.:::**  :*:**** . .:::****:.**:*****

XP_505224.1       LIFAGHGNEVLDYLQQQIEDTASDDGGDVVLHGACLGIGVAGMGSGKASLYDKLRDVLFS
XP_011400578.1    LIHVNSG----AGAAQLLTDSLRATQHEVIQHGACLGLGIACLGSENAEAFEDVKNVLYT
                  **... *       * : *:     :*: ******.*:* .** :*. :.:::**::

XP_505224.1       DSAVAGEAAGLAMGLVMLGSADDDAQKEMFLYAHETQHEKIIRGLALGLALVNYAREESA
XP_011400578.1    DSAVAGEAAGYALGLLCAGSGREERSDEMLAYAHDTQHEKIVRALAVGVAFVHYGREEAA
                  ********** *:**:  **. :: ..**. ***.******:*.**:*:*:*:*.***:*

XP_505224.1       EPMIEQLLAEQDPILRYGGAYTIALAYAGTGNNKAIQRLLHIAVSDASDDVRRAAVISLG
XP_011400578.1
DGVIDAMLQDQDAIIRYGGQFAMGMAYRGTRNNAAIQKLLHHAVTDVSYDVRRAAVMNLG
                  : :*: :* :**.*:**** :::::** ** ** ***:*** **:*.* *******:.**

XP_505224.1       FILLRNYTAVPRMVELLSESHNPHVRYGTALALGISCAGTGLKEAVEVLEPLTKDPADFV
XP_011400578.1    FVLLSAPEQCPRLVALLAASYNPHVRYGSAMAVGIACAGTGSKEAVALLEPMLTDPTDFG
                  *:**    **:* **: *:*******:*:*:**:***** **** :***: .**:**

XP_505224.1       RQGALVALAMILIQQNEKMSPRVKTARETFAHAIASKHEDPMVKFGAALAQGIIDAGGRN
XP_011400578.1    ---ALIALAMVLMEQP---ASKTETLRAKIDKLHGNKGVEIMVRMGAIMAAGILDAGGRN
                    **:****:*::*   :.:::* * .: :  ..*  : **::** :* **:******

XP_505224.1       VTIALENPQTGSLNMKAIVGLVVFSQFWYWFPYTHFLSLSFSTTAVTGVNEDLKIP-QFD
XP_011400578.1    ATLGLQS-RTGPYRRTAVVGLALFLQYWYWHPLAYSLSLALQPTAIIGVDASLRAPADFK
                  .*:.*:. :**. . .*:***.:* *:***.* :: ***:..**: **: .*: * :*.
```

```
XP_505224.1       FVTSHKPSYFGYPPPAEEVVDKAPEKVATAVLSTTARAKARAKKNEKDKADKMEVE----
XP_011400578.1    VISRCKPSLFAYPPPVAMEDKRSKEKVPTAVLSTTVRSKARAAAKKEEQGKEMGALPSSY
                  .::  *** *.****.    .:: ***.*******.*:**** ::::..:* .

XP_505224.1       ------------------------------------------------------------
XP_011400578.1    VLSNPARVVPAQRRFVAFERDQRWQPVKAVTGDSTAFTHSGAPRQSPSALGTAVHPLAAV


XP_505224.1       ------------------------------------------------------------
XP_011400578.1    AAQPSAASAEKELWLPTPHYCESIYQTRRRPTRTIYVGKVPVGSEHPLARQTMTTTDTRD


XP_505224.1       ------------------------------------------------------------
XP_011400578.1    VEATVEQVKKCADAGADIVRITVQGKREATACLKIRDRLFQDRYDTPLVADIHFQPTVAM


XP_505224.1       ------------------------------------------------------------
XP_011400578.1    MVADAFEKIRINPGNFADGRKSFEEINYDDPAQFARERESFRETFMPLVEKCKALGRAMR


XP_505224.1       ------------------------------------------------------------
XP_011400578.1    IGTNHGSLSARILSYYGDTPRGMVESAFEFADMCRDADYHNFLFSMKASNPLVMVQAYRL


XP_505224.1       ------------------------------------------------------------EKPEEPSS
XP_011400578.1    LSEEMYAKGWDYPLHLGVTEAGEGEDGRMKSAIGIGALLMDGLGDTIRVSLTEDPEFEIV
                                                 *.**

XP_505224.1       AAAASSKPAKEAKVDEDADKRIPGQKKWRLANLTRVLPSQLKYISFPEENRFVPIRKLHS
XP_011400578.1
PCGALAALGSRAQTERWGVEPFEEHRDTHLFNRRRALMAGVGELPGQKEGDVLDMRSILH
                  ...* :  ...*:.:. . : :  ::. :* *  *.* : :  :.  :*. .: :*.:

XP_505224.1       VGGVVVLRD----------QEKGKGGKKPQFIKTVRQINNPFDDEDDEAPLPEPF
XP_011400578.1    RDGSVISSLSLEDLRTPETTYRNLGCKMAVGMPFKDIATSDSIFLREVPAEDDAQARRAL
                  .* *:                *   *   : * * .*      :*:*   ..:

XP_505224.1       KYVEETE-------------------------------------------------------
XP_011400578.1    RRLQEVRGEGWEFTCWRRPSALAASPLPDSVAVISLAAALAARAEGRSPLPAGAVRLALA
                  : ::*..

XP_505224.1       ------------------------------------------------------------
XP_011400578.1    VDGTEPDADVEAAGGLDATLVLLRVAPGVSRVHASRRVFEALTRARCHLPVVHALEFAPG


XP_505224.1       ------------------------------------------------------------
XP_011400578.1
TGREELVLSAGALVGALMVDGRGDGVMLSAPGQDLAYLRTTSFGLLQGVRMRNTKTDYVS


XP_505224.1       ------------------------------------------------------------
XP_011400578.1    CPSCGRTLFDLQEVTDQIRQKTGHLPGVAIAIMGCIVNGPGEMADADFGYVGGAPGKIDL


XP_505224.1       ----------------------------------------------------
XP_011400578.1    YVGKEVVRRAIPMESACDQLVELIKEHGRWKDPEPEEEGEGQEEKQAVMA
```

IpsH (4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase): *Y. lipolytica* :

YALI0_D16357g (XP_502897.1) aligned with *Marinifilum fragile* (WP_083464553.1)

CLUSTAL 2.1 multiple sequence alignment

```
XP_502897.1       MIEGISFASFVTHEKPKFVRALDFYKALGFLPTKEYKHGTDHHATDEEGAGSIQEVWLTS
WP_083464553.1      ------------------MKIEIDPNAGFCFGVVNAINKAEENLSKDGQ-------MYC
                          :::   **       .. ..  .::*       : .

XP_502897.1       SRAGVPSVTVKLRLSRHGNEHVSLPNLKHDWRSLVPSLVYYAPDLDAVRAAITPFLHEDH
WP_083464553.1      IGDIVHNSIEVNRLNDLG-----LQTIEHDEFKGIK-----------------------
                  * .    **. *    * .::** . :

XP_502897.1       STLLERPSHTNFIELYAIDPMGNLVGFSRRENPYSSAMQKPFSADDIGPQNFSKPNETKI
WP_083464553.1      ----------------------------ERKVLFRAHGEPPSSYRIAK-----KNKMEI
                                 ..:   *  :* **:  *.     *: :*

XP_502897.1       KGKKRIGVMTSGGDAPGMCAAVRAVVRAGIARGCEVYAVREGYEGLVKGGDLIEPLSWED
WP_083464553.1      VDATCPVVLNLQKRIKKSFRELKKVNGQILIYGKKGHAEVNGLVGQTSGKAIVIET----
                . .  *:.         :: *   :  * : :*  :*  * ..* ::

XP_502897.1       VRGWLSLGGTLIGTARCKEFREREGRLAGALNMVKNGIDALIVIGGDGSLTGADLFREEW
WP_083464553.1      ----------------------------AKDLKSVNPDLPAELFSQTTKTLEGFNLLREEL
                           *  *: *: .: * :.    :* * :*:***

XP_502897.1       PSLIEELVTNGSITAEQAERHRHLDICGMVGSIDNDMATTDVTIGAYSSLDRICELVDFI
WP_083464553.1      QKIMKAPFK------------AHDTICRKVANR-----------------------VPLI
                .::: ..          *  **  *..                 * :*

XP_502897.1       DATAQSHSRAFVVEVMGRHCGWLALMAGTATGADYIFIPEAAPDATQWAEKMTRVVKRHR
WP_083464553.1      KEFAQKHDVIVFVSGKKSSNGKLLFDVCKGINRNSYFISEPQDIDISWFVNAKSVG----
              . **.*.  ..*.     * * : .. . :  **.*.    .*  : . *

XP_502897.1       SQGKRKTVVIVAEGAIDSDLNPITAKMVKDVLDGIGLDTRISTLGHVQRGGPPVAADRVL
WP_083464553.1      ------------------------------------------------------------


XP_502897.1       ASLQGVEAIDAILSLTPETPSPMIALNENKITRKPLVESVALTKKVADAIGNKDFAEAMR
WP_083464553.1      -------------VSGATSTPGWLMN-------------------------EITEKIK
                      ::  *.:*   :*                    :::* ::

XP_502897.1       LRNPEFVEQLQGFLLTNSADKDRPQEPAKDPLRVAIVCTGAPAGGMNAAIRSAVLYGLAR
WP_083464553.1      LENKINISMSR-------------------IKKIGVLTSGGDAPGMNAAIRAVVRTAIFN
              *.*   :.  :               ::.:: :*. * *******:.*  .: .

XP_502897.1       GHQMFAIHNGWSGLVKNGDDAVRELTWLEVEPLCQKGGCEIGTNRSLPECDLGMIAYHFQ
WP_083464553.1      KLEVVGIKQGYEGLING---DFKKMKSHDVSNILQKGGTILRSARSEEFRTVEGRQKAYQ
               ::..*::*:.**::.   .:::.  :*. : **** : : **    :     :*

XP_502897.1       RQRFDGLIVIGGFEAFRALNQLDDARHAYPALRIPMVGIPATISNNVPGTDYSLGADTCL
WP_083464553.1      QMLDN---KIDALVVIGGDGTFSGARIFTQEHDIPVVGIPGTIDNDLFGTDYTIGYDTAI
              :  :   *..: .: . :..**      **.****.**.*:: ****::* **.:

XP_502897.1       NSLVQYCDVLKTSASATRLRLFVVEVQGGNSGYIATVAGLITGAYVVYTPESGINLRLLQ
WP_083464553.1      NTVIDAVDKIRDTASAH-NRLFFIEVMGRDAGFIALRSGIATGAEAILIPEKETHAKELQ
              *:::: * :: :***   ***.:** * ::*:**  :*: *** .:  **. : : **

XP_502897.1       HDISYLKDTFAHQADVNRTGKLLLRNERSSNVFTTDVITGIINEEAKGSFDARTAIPGHV
WP_083464553.1      AYLEKGYKEHKSSGIVIVAEGDKSGG---------AYTIAKEIEKEHPEYDIRVSVLGHM
               :.  . .  .. * :     .          .  *: : .:* *.:: **:

XP_502897.1       QQGGHPSPTDRVRAQRFAIKAVQFIEEHHGSKNNADHCVILGVRGSKFKYTSVSHLYAHK
WP_083464553.1      QRGGSPSAFDRVTASTLGVAAVEALLDDQKS-------IMVGMVHG-----EVTHVPFNK
              *:** **. *** *. .:: **: : :.: *      :::*:  .    .*:*:  :*

XP_502897.1       TEHGARRPKHSYWHAIGDIANMLVGRKAPPLPETLNDEIEKNIAKEQGIIDPC
WP_083464553.1      TIKNKKKVKDTYLRINEILSI--------------------------------
              * :. :: *.:* :    ::
```

Additionally, using standard protein BLAST search I discovered the putative IDI (isopentenyl-diphosphate Delta-isomerase) type 1 and type 2 in *Y. lipolytica* genome. According to Kuzuyama et. al. (2003) (Figure 3- 7) the one type of organisms that harbor both IDI type 1 and type 2 is *Streptomyces* sp. and this bacterium also harbor both MEP and MVA pathway. I hypothesized that *Y. lipolytica* might also have similar systems owing to the present of both type IDI gene.

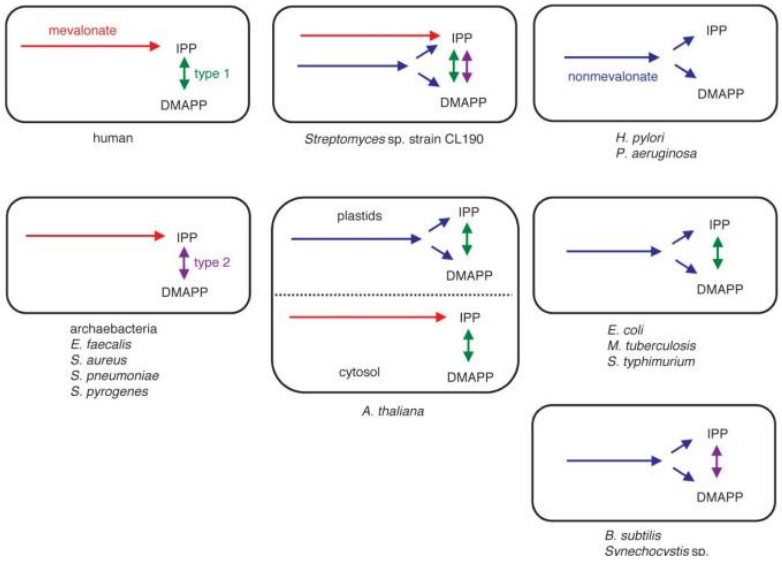


Figure 3- 7 Summary of the different type of the known MEP and MVA pathway systems

Putative IDI type 1 found in *Y. lipolytica* reference genome

```
sp|P15496.2|IDI1_YEAST      MTADNNSMPHGAVSSYAKLVQNQTPEDILEEFPEIIPLQQRPNTRSSETS
XP_504974.1                 ----------MTTSYSDKIKSISVSSVAQQFPEVAPIADVSKASRPSTE
                                       .:**:. ::. : ..: ::***: *: : .::  ..*.

sp|P15496.2|IDI1_YEAST      NDESGETCFSGHDEEQIKLMNENCIVLDWDDNAIGAGTKKVCHLMENIEK
XP_504974.1                 SSDSSAKLFDGHDEEQIKLMDEICVVLDWDDKPIGGASKKCCHLMDNIND
                            ..:*. . *.**********:* *:******:.**..:** ****:**:.

sp|P15496.2|IDI1_YEAST      GLLHRAFSVFIFNEQGELLLQQRATEKITFPDLWTNTCCSHPLCIDDELG
XP_504974.1                 GLVHRAFSVFMFNDRGELLLQQRAAEKITFANMWTNTCCSHPLAVPSEMG
                            **:*******:**::*********:*****.::**********.: .*:*

sp|P15496.2|IDI1_YEAST      LKGKLDDKIKGAITAAVRKLDHELGIPEDETKTRGKFHFLNRIHYMAPSN
XP_504974.1                 G-LDLESRIQGAKNAAVRKLEHELGIDP-KAVPADKFHFLTRIHYAAPSS
                               .*:.:*:** .******:*****   :: . .*****.**** ***.

sp|P15496.2|IDI1_YEAST      EPWGEHEIDYILFYKINAKENLTVNPNVNEVRDFKWVSPNDLKTMFADPS
XP_504974.1                 GPWGEHEIDYILFVRGDPELKVVAN----EVRDTVWVSQQGLKDMMADPK
                             ************ : :.: ::..*    ****  *** :.** *:***.

sp|P15496.2|IDI1_YEAST      YKFTPWFKIICENYLFNWWEQLDDLSEVENDRQIHRML
XP_504974.1                 LVFTPWFRLICEQALFPWWDQLDNLPAGDDEIRRWIK-
                              *****::***: ** **:***:*.  ::: :
```

Putative IDI type 2 found in *Y. lipolytica* reference genome

CLUSTAL 2.1 multiple sequence alignment

```
sp|Q9KWG2.1|IDI2_STRC1      MTSAQRKDDHVRLAIEQHNAHSGRNQFDDVSFVHHALAGIDRPDVSLATS
XP_504224.1                 MVSVEEVAKHNSKDSCWVILHGKAYDLTEFLPEHPGGQAIILKYAGKDAT
                          *.*.:. .*       *.  :: :.  * . .*   .. ::

sp|Q9KWG2.1|IDI2_STRC1      FAGISWQVPIYINAMTGGSEKTGLINRDLATAARETGVPIASGSMNAYIK
XP_504224.1                 KAFDPIHPRDVVDKFLDKDKHLGEVTGEIAAEEEEEVTPEEKARLERYEN
                           *  . :   :: : . .:: * :. ::*:  .*  .*  .. :: * :

sp|Q9KWG2.1|IDI2_STRC1      DPSCADTFR---------------------------------VLRDENPNG
XP_504224.1                 RPPLSQIFNSFDFEYVARHTMSPNAWAYYSSGSDDEITVRENHRAFHKIW
                           *. :: *.                        *  :

sp|Q9KWG2.1|IDI2_STRC1      FVIANINATTTVDNAQRAIDLIEANALQIHINTAQETPMPEGDRSFAS--
XP_504224.1                 FRPRVLVDVKNVDISTTMLGTKSSVPFYITATALGKLGHPEGEVVLTRGA
                          *   : ...** :  :.  .: .: *  .: :  ***:  ::

sp|Q9KWG2.1|IDI2_STRC1      ------------------------------WVPQIEKIAAAVDIPVIVKE
XP_504224.1                 DKMDVIQMIPTLASCSFDEIVDAATDKQTQWMQLYVNMDREVTKKIVQHA
                                       *:   ::  *  :: :

sp|Q9KWG2.1|IDI2_STRC1      VGNGLSRQTILLLA-DLGVQAADVSGRGGTDFARIE----------
XP_504224.1                 EKRGVKGLFITVDAPQLGRREKDMRTKFGDPGAQVQQSDDSVDRSQGAAR
                           .*:.  * : * .** :  *:  : *  *:::

sp|Q9KWG2.1|IDI2_STRC1      ----------------------------------------------------
XP_504224.1                 AISSFIDPSLSWKDIPWFQSITKMPIILKGVQCAEDALKAVEYKVDGILL


sp|Q9KWG2.1|IDI2_STRC1      ---NGRRELGDYAFLHGWGQSTAACLLDAQDISLPVLASGGVRHPLDVVR
XP_504224.1                 SNHGGRQLEFARPSIEVLVEVMAALRAKGWQDYIEVYIDGGIRRATDVIK
                           .**:    . :.  :  **   .. : : *  .**.*:. **::

sp|Q9KWG2.1|IDI2_STRC1      ALALGARAVGSSAGFLRTLMDDGVDALITKLTTWLDQLAALQTMLGARTP
XP_504224.1                 ALCLGAKGVGIGRPFLYAMSTYGEDGVCHLIQLLKDEMEMNMRLIGATKI
                          **.***:.** .  ** ::   * *.:  :   *::    ::** .

sp|Q9KWG2.1|IDI2_STRC1      ADLTRCDVLLHGELRDFCADRGIDTRRLAQRSSSIEALQTTGSTR
XP_504224.1                 EDLNPSMVDLKSIFTHSADTARDILGENVYQHMEMPLFKGSKL--
                           **. . * *:. : . .       . . :  .:  :: :
```

Although I could not identify significant alignment with the protein BLAST search for the MEP pathway genes, with my bioinformatics workflow, I could suggest the potential candidate for further validation. The advantages and disadvantages of my in-house bioinformatics workflow are up for debate. On the one hand, my protocol is very flexible for capturing the gene sequence, which might have the MEP pathway genes function. On the other hand, by using the bare minimum requirement for two protein sequences to

perform a similar function, the false positive rate might be high. Moreover, this approach might not be able to capture isozyme with no homology relation. Therefore I devised alternative gene candidate finding workflow grounded on previously published RNA sequencing data analysis.

To search for the potential MEP pathway gene, I review the publicly available RNA sequencing records in which the *Y. lipolytica* was subjected to the nitrogen limiting condition. The closest data set we found was that of Kerkhoven, Eduard J., et al. (2017) (Kerkhoven et al., 2017). This study compares the nitrogen limiting condition with the carbon limiting condition, ideally the comparison between the nitrogen limiting condition and standard condition; however, for the purpose of extracting the differentially expressed genes during the nitrogen limiting condition, any other different condition would be useful. We start by mapping the raw reads from the array express archive number E-MTAB-5284 to the reference *Y. lipolytica* genome assembly accession: GCF_000002525.2 using HISAT2 (Kim et al., 2019). Then we use StringTie (Pertea et al., 2016) to assemble the alignment into the potential transcripts. Finally, we calculate the significantly different gene expressions using the DESeq2 package (Love et al., 2014). To narrow down the significantly expressed genes, we compare the protein motif from the significantly expressed genes in nitrogen limiting condition to that of the genes from the MEP pathway. We utilized the MEME (Multiple EM for Motif Elicitation) (Bailey et al., 2006) software to extract the motif information from six gene clusters in the MEP pathway listed in the NCBI COG database, namely DXS, DXR, *ispD*, *ispE*, *ispF*, *ipsG*, and *ipsH*. Then we scan the significantly expressed genes in nitrogen limiting condition with the motif information obtained from the earlier step using FIMO software (Grant et al., 2011). The candidate gene sequence and the subject sequence could be found in the supplementary table S15- S20 . (the first column indicates the gene id in *Y. lipolytica* reference genome, the second and

third column shows the p-value and q-value from FIMO software, respectively, the fourth column indicates the protein motif sequence).

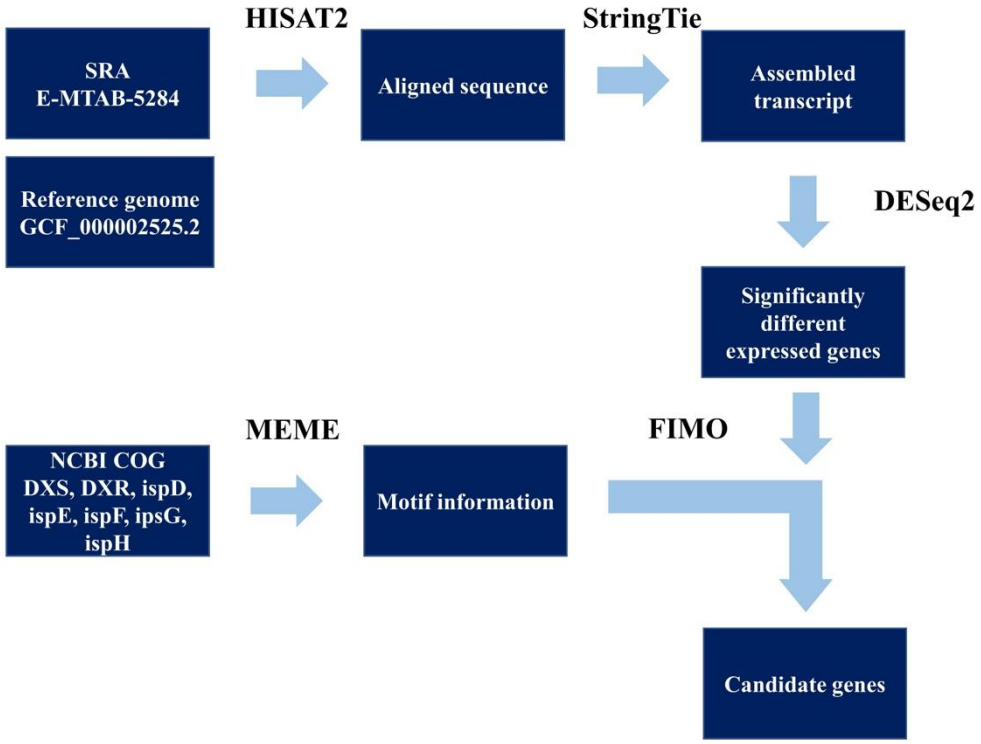


Figure 3- 8: Illustration of the bioinformatics workflow for MEP pathway gene candidate based on previously published RNA sequencing data.

3.4 Discussion

It is shared knowledge that changing the carbon to nitrogen ratio could significantly change the quantity of lipid stored in the oleaginous yeasts, and it may have a distinctive effect on distinctive strains of yeast (Egli et al., 1986; Gao et al., 2007; Jackson et al., 1990; Manikandan et al., 2010). Huge labor has been made in the past few years to comprehend and engineer oleaginous yeast (Sheng et al., 2015). Nevertheless, only narrow work has been completed to describe the metabolic reaction of *Y. lipolytica* in nitrogen limiting settings. One of the best approaches to examine the cell's response is through metabolomics. Metabolites are final products of a complex system of reactions in which several regulatory

routes (genes, mRNAs, enzymes, metabolites, etc.) are involved; these information are recognized to have a high association to a physical phenomenon (Ryan et al., 2006). Upon data examination, I tripped upon very surprising data where one of the chromatographic data was interpreted as MEP. Though the LC/QqQ/MS data are generally accepted in the field of metabolomics owing to its selectivity and sensitivity (Xiao et al., 2012), I took additional steps to confirm the outcomes by further tested the peak in question with the authentic standard spiking. The authentic standard provided the analytical chemistry initial proof of the presence of the MEP pathway enzyme(s) within *Y. lipolytica*. Note that in the nitrogen limiting settings, another peak near 7.40 minutes also only appears in the nitrogen limiting settings. This could be a chemical that is vital for adapting to the nitrogen limiting settings or associated to the MEP-like pathway. It would be exciting for future studies. The next matter I address is the confirmation from the biological standpoint. One of the finest techniques to evaluate the MEP pathway activity is to assess the enzyme explicit to the MEP pathway, the DXP reductoisomerase (Brown et al., 2008). Nevertheless, no DXR gene has been annotated in the reference genome of *Y. lipolytica,* and the NCBI BLAST search yields no significant alignment. I further investigated this issue with an in-house bioinformatics pipeline and obtained several gene candidates. However, a vital element of the known DXR gene was still lacking. The absence of the DXR gene information might be owing to the shortage of data on the wet-lab tested gene function, as more than half of the *Y. lipolytica* gene was hypothetical proteins (Magnan et al., 2016). I also hypothesized that *Y. lipolytica* may use an isozyme that catalyzes the DXR-like activity (Carretero-Paulet et al., 2013; Soliman et al., 2011) or that the MEP discovered in *Y. lipolytica* was produced by an enzyme with very low or no homology with the enzyme in the MEP pathway. Furthermore, I make an experiment using HMG-CoA reductase inhibitors. However, the HMG-CoA reductase inhibitors do not significantly affect the MEP level (Supplementary

Figure S2 – S4). Furthermore, to provide additional information about the potential MEP pathway gene target in *Y. lipolytica*, the RNA sequencing data from a previous publication was analyzed to create a list of candidate MEP pathway enzymes in *Y. lipolytica*. The discoveries that *Y. lipolytica* is able to synthesize MEP and the strain contain the DXR-like activity could lead to a assumption that a pathway comparable to the MEP pathway is present in *Y. lipolytica*. To evaluate this hypothesis, I performed an assimilation analysis of 1-^{13}C-labeled glucose into ergosterol. There are several reasons why I picked ergosterol; 1. ergosterol was located in the yeast cell membrane with high abundance (1-3 % dried cell weight), 2. the ergosterol molecule offer several fragments that can be examined for the pathway dissimilarities; therefore, the influence of the MEP or MVA pathway can be noticeably observed, 3. the extraction and refinement method for ergosterol was well-known. This analysis was based on the information that the building blocks for ergosterol are synthesized by entirely different routes, the MVA pathway and the MEP pathway (Rohmer, 1999). Though this test was done with the hypothesis that the known MEP pathway was used in *Y. lipolytica*, the fragments discovered from the experiment match with the projected mass change. Furthermore, the fact that the fragment mass change from different settings changed differently might be a clue that distinctive biosynthesis pathways were used. Combined with other outcomes stated above, I concluded that ergosterol in *Y. lipolytica* may be resulting from the MVA pathway and the MEP-like pathway. While more straight proof such as the identification of all intermediates and identification of crucial enzymes are still indispensable, the remarks in this report offer important information for researchers to further validate the existence of the MEP-like pathway in yeast.

3.5 Conclusion

Overall, I proposed that both the MVA pathway and the MEP-like pathway might exist in *Y. lipolytica*. Remarkably, the nitrogen limiting settings initiated the use of the MEP-like pathway. Four types of data backing the conclusion that *Y. lipolytica* may have an enzyme(s) in the MEP-like pathway. Firstly, the authentic standard established that the MEP was identified in *Y. lipolytica*. Secondly, the MEP peak was disturbed by fosmidomycin. Thirdly, the ^{13}C labeled glucose showed that the ergosterol fragments matched with the hypothetical mass shifted from the MVA pathway and the MEP pathway. Finally candidate genes was found and putative IDI type I and II was located in the *Y. lipolytica* reference genome. The exact genes involved in the MEP production in *Y. lipolytica* and the mechanism on how nitrogen limiting settings activate this occurrence remain ambiguous. My findings stated here may stimulus new explorations on the probability of the MEP pathway-like presence in yeasts that may be triggered in specific environmental settings. Furthermore, this finding may open a new way for oleaginous yeast control and alteration for further industrial purposes.

Chapter 4
Conclusions and perspectives

The elucidation of cellular functions is an enormous effort and requires various strategies to capture the entire system. The metabolomics-based analysis offers a holistic view of the metabolic shifts under genetic or environmental perturbations, both qualitatively and quantitatively. Here, I demonstrated that the metabolomics analysis of oleaginous yeast *Yarrowia* spp. can yield known and unknown metabolic behavior, which raised an interesting follow-up experiment. The metabolome dataset presented in this study does not only provide information about key metabolites but also represents a useful resource for future research regarding non-mevalonate-like pathway in yeast.

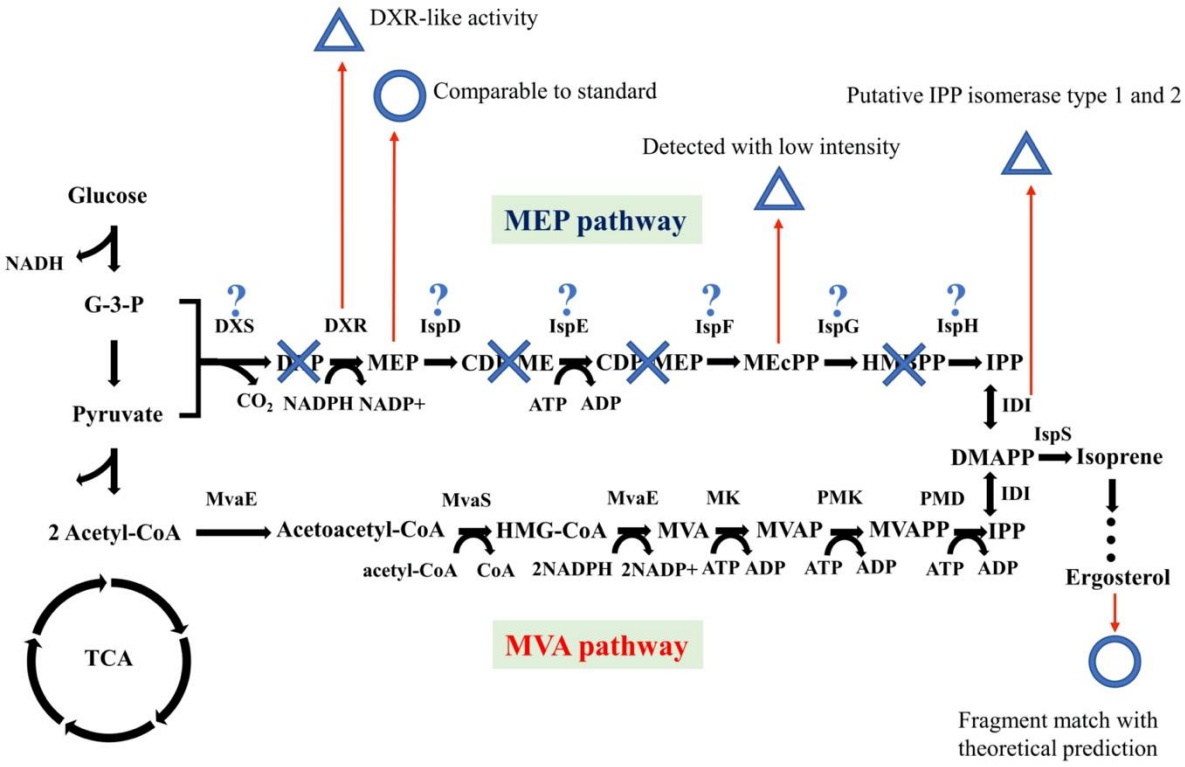


Figure 4- 1: summary of all evidence regarding MEP-like pathway in *Y. lipolytica* in this study.

As described in figure 4-1, several pieces of evidence were found that support the existence of the MEP-like pathway in *Y. lipolytica*. 1: the MEP peak was comparable to the authentic standard. 2: the chemical inhibition suggested the existence of DXR-like activity. 3: the 2-C-Methyl-D-erythritol 2,4-cyclodiphosphate (MEcPP) was detected with low intensity (Figure S1). 4: both type 1 and 2 putative isopentenyl diphosphate isomerase similar to the organism, which has both MVA and MEP pathways in the same organism. 5: the carbon isotope label ergosterol fragment analysis matched with the theoretical prediction of the MEP pathway mass shift. This study also provides MEP pathway gene candidates from two different approaches based on sequence similarity search and RNA sequencing data combined with protein motif analysis.

With the current data, I proposed that the mevalonate pathway and the non-mevalonate-like pathway co-exist in *Y. lipolytica*. Moreover, I speculated that because nitrogen is a building block of purine and pyrimidine compounds, which play a crucial role in the biosynthesis of bio-energy carrier compounds such as ATP. The lack of nitrogen creates a stress condition for the cell. This stress might lead to the metabolic adjustment to a less ATP-dependent pathway. This study further raised the question of how eukaryotic cells can adopt the non-mevalonate pathway in this first place. According to the literature review, some eukaryotic cells can harbor the non-mevalonate-like pathway; however, in each case, the biological systems of the said non-mevalonate pathway are different. It would be interesting to study this adaptation of the non-mevalonate pathway in yeast from the evolutionary pathway in the future.

To further elucidate how the non-mevalonate-like pathway is working in *Y. lipolytica*, more direct evidence like a crucial enzyme(s) involved in the non-mevalonate-like pathway has to be discovered to understand this phenomenon; this will enable a wider array of analyses, including in-vitro enzymatic reaction analysis, gene overexpression, and

gene knockout, etc. Moreover, a multi-omic approach should be utilized to uncover mechanisms behind the utilization of the MEP-like pathway. The working model present here is based on the assumption that the MEP observed in this study came from the know MEP pathway. However, it is possible that the modified version of the known MEP pathway or the entirely different pathway might also be at play here.

References

Abd-Aziz, S., Ang, D. C., Yusof, H. M., Karim, M. I. A., Ariff, A. B., Uchiyama, K., & Shioya, S. (2001). Effect of C/N ratio and starch concentration on ethanol production from sago starch using recombinant yeast. *World Journal of Microbiology and Biotechnology*, *17*(7), 713–719.

Abghari, A., & Chen, S. (2014). Yarrowia lipolytica as an oleaginous cell factory platform for production of fatty acid-based biofuel and bioproducts. *Frontiers in Energy Research*, *2*, 21.

Leslie, S. B., Teter, S. A., Crowe, L. M., & Crowe, J. H. (1994). Bioehi f i e~a Trehalose lowers membrane phase transitions in dry yeast cells. In Biochimica et Biophysica Acta (Vol. 1192).

Bailey, T. L., Williams, N., Misleh, C., & Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research*, *34*(suppl_2), W369–W373.

Barth, G., & Gaillardin, C. (1996). Yarrowia lipolytica. In Nonconventional yeasts in biotechnology (pp. 313–388). Springer.

Beopoulos, A., Chardot, T., & Nicaud, J. M. (2009). Yarrowia lipolytica: A model and a tool to understand the mechanisms implicated in lipid accumulation. *Biochimie*, *91*(6), 692–696. doi: 10.1016/j.biochi.2009.02.004

Blank, L. M., Lehmbeck, F., & Sauer, U. (2005). Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts. *FEMS Yeast Research*, *5*(6–7), 545–558.

Brown, A. C., & Parish, T. (2008). DXR is essential in Mycobacterium tuberculosis and fosmidomycin resistance is due to a lack of uptake. *BMC Microbiology*, *8*, 1–9. doi: 10.1186/1471-2180-8-78

Bylesjö, M., Eriksson, D., Sjödin, A., Jansson, S., Moritz, T., & Trygg, J. (2007). Orthogonal projections to latent structures as a strategy for microarray data normalization. *BMC Bioinformatics*, *8*(1), 207.

Carretero-Paulet, L., Lipska, A., Pérez-Gil, J., Sangari, F. J., Albert, V. A., & Rodríguez-Concepción, M. (2013). Evolutionary diversification and characterization of the eubacterial gene family encoding DXR type II, an alternative isoprenoid biosynthetic enzyme. *BMC Evolutionary Biology*, *13*(1). doi: 10.1186/1471-2148-13-180

Casaregola, S., Neuvéglise, C., Bon, E., & Gaillardin, C. (2002). Ylli, a Non-LTR Retrotransposon L1 Family in the Dimorphic Yeast Yarrowia lipolytica. *Mol. Biol. Evol*, *19*(5), 664–677. Retrieved from https://academic.oup.com/mbe/article-abstract/19/5/664/1067799

Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., Ong, W. K., Paley, S., Subhraveti, P., & Karp, P. D. (2020). The MetaCyc database of metabolic pathways and enzymes-a 2019 update. *Nucleic Acids Research*, *48*(D1), D445–D453.

Chang, C.-F., Chen, C.-C., Lee, C.-F., & Liu, S.-M. (2013). Identifying and characterizing Yarrowia keelungensis sp. nov., an oil-degrading yeast isolated from the sea surface microlayer. *Antonie Van Leeuwenhoek*, *104*(6), 1117–1123.

Chiocchio, V. M., & Matkovi??, L. (2011). Determination of ergosterol in cellular fungi by HPLC. A modified technique. *Journal of the Argentine Chemical Society*, *98*(December), 10–15.

Christen, S., & Sauer, U. (2011). Intracellular characterization of aerobic glucose metabolism in seven yeast species by 13C flux analysis and metabolomics. *FEMS Yeast Research*, *11*(3), 263–272. doi: 10.1111/j.1567-1364.2010.00713.x

Crous, P. W., Wingfield, M. J., Burgess, T. I., Carnegie, A. J., Hardy, G. E. S. J., Smith, D., Summerell, B. A., Guarro, J., Houbraken, J., Lombard, L., Martín, M. P., Alexandrova, A. v, Barnes, C. W., Baseia, I. G., Miller, A. N., Ordoñez, M. E., Abreu, V. P., Accioly, T.,

Agnello, C., & Colmán, A. A. (2017). *Fungal Planet description sheets : 625 – 715*.
Dempo, Y., Ohta, E., Nakayama, Y., Bamba, T., & Fukusaki, E. (2014). Molar-based targeted metabolic profiling of cyanobacterial strains with potential for biological production. *Metabolites*, *4*(2), 499–516.
de Virgilio, C., Hottiger, T., Dominguez, J., Boller, T., & Wiemken, A. (1994). The role of trehalose synthesis for the acquisition of thermotolerance in yeast I. Genetic evidence that trehalose is a thermoprotectant. In Eur. J. Biochem (Vol. 219).
Dhiman, R. K., Schaeffer, M. L., Bailey, A. M., Testa, C. A., Scherman, H., & Crick, D. C. (2005). 1-Deoxy-D-xylulose 5-phosphate reductoisomerase (IspC) from Mycobacterium tuberculosis: Towards understanding mycobacterial resistance to fosmidomycin. *Journal of Bacteriology*, *187*(24), 8395–8402. doi: 10.1128/JB.187.24.8395-8402.2005
Djoumbou-Feunang, Y., Pon, A., Karu, N., Zheng, J., Li, C., Arndt, D., Gautam, M., Allen, F., & Wishart, D. S. (2019). Cfm-id 3.0: Significantly improved esi-ms/ms prediction and compound identification. *Metabolites*, *9*(4), 1–23. doi: 10.3390/metabo9040072
Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., Montigny, J. de, Blanchin, S., Beckerich, J., Beyne, E., Bleykasten, C., Babour, A., Boyer, J., Cattolico, L., Confanioleri, F., Daruvar, A. de, Despons, L., & Fabre, E. (n.d.). *Genome evolution in yeasts*. 35–44.
Egli, T., & Quayle, J. R. (1986). Influence of the carbon: Nitrogen ratio of the growth medium on the cellular composition and the ability of the methylotrophic yeast Hansenula polymorpha to utilize mixed carbon sources. *Journal of General Microbiology*, *132*(1986), 1779–1788. doi: 10.1099/00221287-132-7-1779
Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. *Functional Genomics*, 155–171.
Flagfeldt, D. B., Siewers, V., Huang, L., & Nielsen, J. (2009). Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast (Chichester, England)*, *26*(10), 545–551. doi: 10.1002/yea
Fryberg, M., Oehlschlager, A. C., & Unrau, A. M. (1973). Biosynthesis of Ergosterol in Yeast. Evidence for Multiple Pathways. *Journal of the American Chemical Society*, *95*(17), 5747–5757. doi: 10.1021/ja00798a051
Gaillardin, C., Mekouar, M., & Neuvéglise, C. (2013). *Comparative Genomics of Yarrowia lipolytica* (pp. 1–30). doi: 10.1007/978-3-642-38320-5_1
Gao, L., Sun, M. H., Liu, X. Z., & Che, Y. S. (2007). Effects of carbon concentration and carbon to nitrogen ratio on the growth and sporulation of several biocontrol fungi. *Mycological Research*, *111*(1), 87–92. doi: 10.1016/j.mycres.2006.07.019
González-Cabanelas, D., Hammerbacher, A., Raguschke, B., Gershenzon, J., & Wright, L. P. (2016). Quantifying the Metabolites of the Methylerythritol 4-Phosphate (MEP) Pathway in Plants and Bacteria by Liquid Chromatography–Triple Quadrupole Mass Spectrometry. *Methods in Enzymology*, *576*, 225–249. doi: 10.1016/bs.mie.2016.02.025
Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics*, *27*(7), 1017–1018.
Groenewald, M., & Smith, M. T. (2013). The teleomorph state of Candida deformans Langeron & Guerra and description of Yarrowia yakushimensis comb. nov. *Antonie van Leeuwenhoek*, *103*(5), 1023–1028.
Headley, J. v, Peru, K. M., Verma, B., & Robarts, R. D. (2002). Mass spectrometric determination of ergosterol in a prairie natural wetland. *Journal of Chromatography A*, *958*(1–2), 149–156. doi: 10.1016/S0021-9673(02)00326-6
Henderson, J. F., & Paterson, R. P. (1973). Catabolism of purine nucleotides. *Nucleotide Metabolism: An Introduction*.
Hottiger', T., de Virgilio', C., Hall2, M. N., Boller', T., & Wiemken', A. (1994). The role of

trehalose synthesis for the acquisition of thermotolerance in yeast 11. Physiological concentrations of trehalose increase the thermal stability of proteins in vitro. In Eur. J. Biochem (Vol. 219).

Ibba, M., & Söll, D. (2004). Aminoacyl-tRNAs: setting the limits of the genetic code. *Genes & Development*, *18*(7), 731–738.

Jackson, M. A., & Bothast, R. J. (1990). Carbon concentration and carbon-to-nitrogen ratio influence submerged-culture conidiation by the potential bioherbicide Colletotrichum truncatum NRRL 13737. *Applied and Environmental Microbiology*, *56*(11), 3435–3438.

Jewison, T., Su, Y., Disfany, F. M., Liang, Y., Knox, C., Maciejewski, A., Poelzer, J., Huynh, J., Zhou, Y., Arndt, D., & others. (2014). SMPDB 2.0: big improvements to the Small Molecule Pathway Database. *Nucleic Acids Research*, *42*(D1), D478–D484.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Kerkhoven, E. J., Kim, Y.-M., Wei, S., Nicora, C. D., Fillmore, T. L., Purvine, S. O., Webb-Robertson, B.-J., Smith, R. D., Baker, S. E., Metz, T. O., & Nielsen, J. (2017). Leucine Biosynthesis Is Involved in Regulating High Lipid Accumulation in Yarrowia lipolytica *MBio*, *8*(3), e00857-17. doi: 10.1128/mBio.00857-17

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915.

Knutsen, A. K., Robert, V., Poot, G. A., Epping, W., Figge, M., Holst-Jensen, A., Skaar, I., & Smith, M. T. (2007). Polyphasic re-examination of Yarrowia lipolytica strains and the description of three novel Candida species: Candida oslonensis sp. nov., Candida alimentaria sp. nov. and Candida hollandica sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *57*(10), 2426–2435. doi: 10.1099/ijs.0.65200-0

Kuzuyama, T., & Seto, H. (2003). Diversity of the biosynthesis of the isoprene units. *Natural Product Reports*, *20*(2), 171–183. doi: 10.1039/b109860h

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947–2948. doi: 10.1093/bioinformatics/btm404

Ledesma-Amaro, R., & Nicaud, J.-M. (2016). Metabolic Engineering for Expanding the Substrate Range of Yarrowia lipolytica. *Trends in Biotechnology*, *xx*, 1–12. doi: 10.1016/j.tibtech.2016.04.010

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 1–21.

Magnan, C., Yu, J., Chang, I., Jahn, E., Kanomata, Y., Wu, J., Zeller, M., Oakes, M., Baldi, P., & Sandmeyer, S. (2016). Sequence assembly of yarrowia lipolytica strain W29/CLIB89 shows transposable element diversity. *PLoS ONE*, *11*(9), 1–28. doi: 10.1371/journal.pone.0162363

Manikandan, K., & Viruthagiri, T. (2010). Optimization of C/N ratio of the medium and fermentation conditions of ethanol production from tapioca starch using co - Culture of aspergillus niger and sachormyces cerevisiae. *International Journal of ChemTech Research*, *2*(2), 947–955.

Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D. N., Hanspers, K., A. Miller, R., Digles, D., Lopes, E. N., Ehrhart, F., & others. (2021). WikiPathways: connecting communities. *Nucleic Acids Research*, *49*(D1), D613–D621.

Mcgarvey, D. J., & Croteau', R. (1995). Terpenoid Metabolism. In The Plant Cell (Vol. 7). American Society of Plant Physiologists.

Michely, S., Gaillardin, C., Nicaud, J. M., & Neuvéglise, C. (2013). Comparative Physiology of

Oleaginous Species from the Yarrowia Clade. *PLoS ONE*, *8*(5), 1–10. doi: 10.1371/journal.pone.0063356

Mueller, C., Schwender, J., Zeidler, J., & Lichtenthaler, H. K. (2000). Properties and inhibition of the first two enzymes of the non-mevalonate pathway of isoprenoid biosynthesis. *Biochemical Society Transactions*, *28*(6), 792–793. doi: 10.1042/bst0280792

Nagy, E., Dlauchy, D., Medeiros, A. O., Péter, G., & Rosa, C. A. (2014). Yarrowia porcina sp. nov. and Yarrowia bubula f.a. sp. nov., two yeast species from meat and river sediment. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *105*(4), 697–707. doi: 10.1007/s10482-014-0125-4

NAKASE, T., GOTO, S., KOYAMA, Y., KOMAGATA, K., & IIZUKA, H. (1977). Microbiological Studies on Cheese (II). *Food Hygiene and Safety Science (Shokuhin Eiseigaku Zasshi)*, *18*(4), 353–361_1.

Neuvéglise, C., Chalvet, F., Wincker, P., Gaillardin, C., & Casaregola, S. (2005). Mutator-like element in the yeast Yarrowia lipolytica displays multiple alternative splicings. *Eukaryotic Cell*, *4*(3), 615–624. doi: 10.1128/EC.4.3.615-624.2005

Neuvéglise, C., Feldmann, H., Bon, E., Gaillardin, C., & Casaregola, S. (2002). Genomic evolution of the long terminal repeat retrotransposons in hemiascomycetous yeasts. *Genome Research*, *12*(6), 930–943. doi: 10.1101/gr.219202

O'donovan, G. A., & Neuhard, J. (1970). Pyrimidine Metabolism in Microorganisms. In BACTERIOLOGICAL REVIEWS (Vol. 34, Issue 3).

O'Shea, K., & Misra, B. B. (2020). Software tools, databases and resources in metabolomics: updates from 2018 to 2019. In Metabolomics (Vol. 16, Issue 3). Springer. doi: 10.1007/s11306-020-01657-3

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, *11*(9), 1650.

Pomraning, K. R., Kim, Y.-M., Nicora, C. D., Chu, R. K., Bredeweg, E. L., Purvine, S. O., Hu, D., Metz, T. O., & Baker, S. E. (2016). Multi-omics analysis reveals regulators of the response to nitrogen limitation in Yarrowia lipolytica. *BMC Genomics*, *17*(1), 138. doi: 10.1186/s12864-016-2471-2

Putri, S. P., Fukusaki, E., Shriver, L. P., & Shriver, L. (201 C.E.). Mass Spectrometry-Based Metabolomics: A Practical Guide. *J. Am. Soc. Mass Spectrom*, *27*, 1–2. doi: 10.1007/s13361-015-1246-3

Robinson, J. L., Kocabaş, P., Wang, H., Cholley, P.-E., Cook, D., Nilsson, A., Anton, M., Ferreira, R., Domenzain, I., Billa, V., & others. (2020). An atlas of human metabolism. *Science Signaling*, *13*(624).

Rohmer, M. (1999). The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Natural Product Reports*, *16*(5), 565–574. doi: 10.1039/a709175c

Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Engineering Design and Selection*, *12*(2), 85–94. doi: 10.1093/protein/12.2.85

Ryan, D., & Robards, K. (2006). Metabolomics: The greatest omics of them all? *Analytical Chemistry*, *78*(23), 7954–7958. doi: 10.1021/ac0614341

Sheng, J., & Feng, X. (2015). Metabolic engineering of yeast to produce fatty acid-derived biofuels: bottlenecks and solutions. *Frontiers in Microbiology*, *6*(June), 554. doi: 10.3389/fmicb.2015.00554

Shimizu, T., & Takahashi, S. (1998). Fosmidomycin, a Specific Inhibitor of 1-Deoxy-D-Xylulose 5-Phosphate Reductoisomerase in the Nonmevalonate Pathway for Terpenoid Biosynthesis. *Tetrahedron Letters 39*, 7913–7916.

Silva, L. P., & Northen, T. R. (2015). Exometabolomics and MSI: deconstructing how cells

interact to transform their small molecule environment. *Current Opinion in Biotechnology*, *34*, 209–216.

Silverman, A. M., Qiao, K., Xu, P., & Stephanopoulos, G. (2016). Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *100*(8), 3781–3798. doi: 10.1007/s00253-016-7376-0

Sitepu, I. R., Sestric, R., Ignatia, L., Levin, D., German, J. B., Gillies, L. A., Almada, L. A. G., & Boundy-Mills, K. L. (2013). Manipulation of culture conditions alters lipid content and fatty acid profiles of a wide variety of known and new oleaginous yeast species. *Bioresource Technology*, *144*, 360–369.

Soliman, S. S. M., Tsao, R., & Raizada, M. N. (2011). Chemical inhibitors suggest endophytic fungal paclitaxel is derived from both mevalonate and non-mevalonate-like pathways. *Journal of Natural Products*, *74*(12), 2497–2504. doi: 10.1021/np200303v

Soutourina, J., Plateau, P., & Blanquet, S. (2000). Metabolism of d-Aminoacyl-tRNAs inEscherichia coli and Saccharomyces cerevisiae Cells. *Journal of Biological Chemistry*, *275*(42), 32535–32542.

Trygg, J., Holmes, E., & Lundstedt, T. (2007). Chemometrics in metabonomics. *Journal of Proteome Research*, *6*(2), 469–479.

Tsigie, Y. A., Wang, C.-Y., Truong, C.-T., & Ju, Y.-H. (2011). Lipid production from Yarrowia lipolytica Po1g grown in sugarcane bagasse hydrolysate. *Bioresource Technology*, *102*(19), 9216–9222.

Tsugawa, H., Arita, M., Kanazawa, M., Ogiwara, A., Bamba, T., & Fukusaki, E. (2013). MRMPROBS: A data assessment and metabolite identification tool for large-scale multiple reaction monitoring based widely targeted metabolomics. *Analytical Chemistry*, *85*(10), 5191–5199. doi: 10.1021/ac400515s

Vichi, M., & Saporta, G. (2009). Clustering and disjoint principal component analysis. *Computational Statistics & Data Analysis*, *53*(8), 3194–3208.

Vogels, G. D., & van der Drift, C. (1976). *Degradation of Purines and Pyrimidines by Microorganisms* (Vol. 40, Issue 2).

Vranová, E., Coman, D., & Gruissem, W. (2013). Network Analysis of the MVA and MEP Pathways for Isoprenoid Synthesis. *Annual Review of Plant Biology*, *64*(1), 665–700. doi: 10.1146/annurev-arplant-050312-120116

Wanke, M., Skorupinska-tudek, K., & Swiezewska, E. (2001). Isoprenoid biosynthesis. *Acta Biochimica Polonica*, *48*(3), 663–671. Retrieved from http://www.actabp.pl/pdf/3_2001/663.pdf

Xiao, J. F., Zhou, B., & Ressom, H. W. (2012). Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC - Trends in Analytical Chemistry*, *32*, 1–14. doi: 10.1016/j.trac.2011.08.009

Supplementary data

Table S 1: p-value from t-test of growth and glucose consumption profile

Growth										
time/condition	5:0vs4:1	5:0vs2:2	5:0vs1:4	5:0vs0:5	4:1vs2:2	4:1vs1:4	4:1vs0:5	2:2vs1:4	2:2vs0:5	1:4vs0:5
12h	0.012017	0.005095	0.017147	0.092558	0.758542	0.002819	0.002904	0.000611	0.000852	0.413077
24h	2.03E-05	0.001927	0.017047	0.000964	0.57386	0.051522	4.02E-05	0.038758	6.33E-05	0.024312
36h	0.000129	0.000001	0.000725	0.126107	0.139138	0.06847	0.002749	0.237235	0.019965	0.014392
48h	0.007939	0.018235	0.015151	0.488133	0.435604	0.299708	0.000574	0.836493	0.002535	0.001888
Glucose consumption										
time/condition	5:0vs4:1	5:0vs2:2	5:0vs1:4	5:0vs0:5	4:1vs2:2	4:1vs1:4	4:1vs0:5	2:2vs1:4	2:2vs0:5	1:4vs0:5
12h	0.001011	3.17E-05	0.001232	0.001103	7.8E-06	1.1E-06	0.000305	6.3E-06	0.000369	0.001135
24h	1.06E-05	1E-07	0.000096	0.000096	8.8E-06	0.000676	0.000676	0.000672	0.000672	No info.
36h	0	8.9E-06	8.9E-06	8.9E-06	0.000118	0.000118	0.000118	No info.	No info.	No info.
48h	0.000849	2.5E-06	2.5E-06	2.5E-06	0.012375	0.012375	0.012375	No info.	No info.	No info.

Table S 2: complete numerical results of normalized peak area in the time-course sampling of *Y. lipolytica* PO1d in different C:N ratio

	C:N 0:5 12 h	C:N 0:5 24 h	C:N 0:5 36 h	C:N 0:5 48 h	C:N 1:4 12 h	C:N 1:4 24 h	C:N 1:4 36 h	C:N 1:4 48 h	C:N 2:2 12 h	CN: 2:2 24 h	CN 2:2 36 h	C:N 2:2 48 h	C:N 4:1 12 h	C:N 4:1 24 h	C:N 4:1 36 h	C:N 4:1 48 h	C:N 5 0 12 h	C:N 5:0 24 h	C:N 5:0 36 h	C:N 5:0 48 h
Arginine	0.82	0.39	0.51	0.73	0.75	3.62	0.35	0.43	0.69	7.50	2.74	6.31	0.62	7.64	5.42	42.04	7.47	15.67	27.24	0.42
Lysine	0.04	0.04	0.07	0.10	0.10	0.41	0.05	0.08	0.09	0.44	0.13	0.13	0.10	0.31	0.13	0.19	0.27	0.28	0.15	0.05
Histidine	3.25	15.00	6.48	10.46	0.78	7.69	3.64	5.08	1.08	6.76	3.76	4.71	2.34	4.84	2.52	21.79	5.27	9.30	13.61	3.73
Serine	0.63	0.27	0.45	0.40	0.87	2.08	0.51	0.73	0.74	2.42	3.43	5.00	0.52	2.17	2.95	1.96	1.06	1.13	1.52	1.46
Asparagine	1.84	1.17	0.49	0.46	0.71	1.67	0.82	1.04	0.60	1.75	2.18	2.44	0.53	1.46	1.44	2.65	1.65	1.57	2.16	1.53
Glutamine	0.49	0.61	0.94	1.37	1.29	5.36	0.67	1.12	1.26	5.93	1.74	1.66	1.27	4.24	1.72	2.49	3.60	3.67	1.91	0.62
Threonine	2.63	0.38	0.28	0.50	2.22	5.47	0.83	1.04	1.41	9.49	11.76	8.91	1.20	8.05	6.85	3.06	1.79	1.82	2.63	3.84
Hexose	8.03	37.84	40.52	38.63	10.33	14.13	26.59	35.18	25.47	30.64	13.07	47.29	40.68	56.60	53.51	254.06	131.16	159.95	257.86	36.11
Cysteine	2.17	1.42	1.45	0.59	2.93	2.68	1.33	1.14	4.00	2.62	0.17	0.15	3.14	2.89	0.15	21.44	16.86	13.52	18.98	0.26
Trehalose	0.68	0.95	1.52	1.40	0.86	0.59	0.89	1.31	1.16	0.72	2.33	5.92	1.49	0.66	2.70	5.50	2.77	3.03	4.80	4.33
Proline	1.09	1.39	1.11	2.11	0.55	3.03	0.76	1.61	0.31	1.37	1.59	1.68	0.22	0.90	0.78	1.77	0.84	1.31	1.23	0.46
Sucrose	0.36	0.28	0.78	0.33	0.65	0.78	0.67	1.05	0.61	1.08	1.32	3.96	1.25	1.44	1.12	3.19	1.56	1.94	2.14	2.37
Valine	0.35	0.69	0.68	1.24	0.87	2.67	0.85	1.33	0.84	1.94	0.98	0.90	0.66	1.53	0.68	2.50	1.79	1.74	2.03	1.31
Cytidine	0.10	0.15	0.22	0.33	0.08	0.79	0.26	0.48	0.06	0.39	0.58	0.29	0.05	0.15	0.13	0.26	0.18	0.16	0.12	0.60
Methionine	4.24	0.39	0.27	0.49	1.05	0.87	0.35	0.52	1.09	0.72	0.43	0.31	0.88	0.58	0.27	6.76	4.07	4.22	5.17	0.53
Guanine	0.00	0.00	0.00	0.00	0.00	0.04	0.03	0.03	0.00	0.02	0.03	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
Tyrosine	4.35	1.14	1.37	2.67	2.38	3.49	1.73	2.04	2.12	2.66	2.89	3.37	1.86	2.66	2.07	10.53	6.27	6.36	7.72	2.47
Isoleucine	0.21	1.03	0.61	2.17	3.25	4.83	1.01	1.50	6.28	1.15	0.59	0.18	4.09	0.98	0.21	71.01	54.08	47.15	55.17	1.23
Leucine	0.12	0.75	0.35	1.07	0.49	5.80	0.87	1.21	0.74	4.40	4.35	3.98	0.51	3.68	2.92	10.66	5.06	5.71	8.01	2.61
Xanthine	0.47	1.63	1.71	2.80	0.21	6.78	3.04	3.64	0.13	2.97	5.34	3.89	0.08	1.58	1.51	0.43	0.31	0.39	0.42	7.18
Glutamate	16.43	23.79	15.79	25.17	10.67	28.24	17.58	24.15	9.41	18.38	19.19	19.68	8.82	15.22	15.26	25.04	15.25	18.90	16.10	21.63
Uridine	5.43	5.16	1.12	2.30	1.51	32.03	4.44	3.30	1.50	12.83	29.04	17.08	0.83	6.03	13.08	2.60	0.26	1.22	1.05	14.99
Aspartate	10.02	30.25	17.50	30.75	3.23	22.59	19.22	30.96	2.86	6.73	18.97	13.89	2.61	4.44	5.82	19.67	10.80	13.62	14.54	21.39
Thymine	0.01	0.01	0.02	0.02	0.01	0.06	0.06	0.07	0.01	0.04	0.08	0.09	0.01	0.02	0.04	0.01	0.00	0.00	0.00	0.17
Guanosine	0.08	0.04	0.31	0.64	0.07	1.46	0.17	0.61	0.05	0.95	1.04	1.25	0.03	0.53	0.47	2.07	0.66	0.89	0.86	1.28
Adenosine	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.02	0.00	0.01	0.01	0.02	0.01	0.01	0.01	0.01
Thymidine	0.03	0.00	0.00	0.00	0.07	0.09	0.00	0.00	0.05	0.05	0.05	0.04	0.04	0.04	0.03	0.06	0.05	0.05	0.05	0.01
Phenylalanine	5.90	1.17	1.02	2.32	3.46	5.38	2.98	2.81	3.03	3.59	5.51	6.75	2.55	3.03	3.21	16.91	10.23	10.22	13.03	8.05
Glycolate	0.85	1.39	0.61	1.41	0.78	1.26	0.83	1.26	0.56	0.73	1.54	2.81	0.53	0.57	0.95	1.70	0.24	0.55	0.95	3.12
Glyoxylate	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.05	0.00	0.04	0.41	0.91	0.00	0.04	0.35	0.00	0.00	0.00	0.00	0.46
G6P	7.05	6.79	1.40	3.08	6.00	7.30	2.10	2.63	7.41	7.92	5.24	6.32	9.37	12.60	5.19	0.37	1.38	1.36	0.40	4.25
Disaccharide-P	0.10	0.07	0.03	0.05	0.15	0.03	0.01	0.06	0.16	0.15	0.21	1.40	0.18	0.51	0.83	0.00	0.04	0.03	0.01	0.06
Pyroglutamate	1.24	1.68	1.76	1.75	1.12	2.17	1.77	2.39	1.05	1.85	1.74	2.25	1.03	1.57	1.60	5.31	2.24	2.56	4.20	2.27
Tryptophan	9.63	6.24	7.56	16.47	8.22	7.46	10.69	12.22	7.28	5.80	11.35	13.69	6.30	5.75	7.95	52.57	14.45	23.25	43.70	15.48
RSP	0.31	0.43	0.16	0.30	0.45	0.89	0.51	0.52	0.62	2.14	3.08	12.32	0.51	2.64	5.07	0.44	0.35	0.62	0.65	1.38
Lactate	0.31	0.33	0.63	0.19	0.23	0.68	0.31	0.51	0.24	0.47	0.47	0.68	0.26	0.49	0.47	0.81	0.31	0.56	0.47	0.47
S7P	3.74	2.90	0.82	1.77	1.66	4.13	1.24	1.74	2.35	2.56	4.57	13.39	3.23	4.72	7.42	0.35	0.50	0.95	0.44	4.18
F6P	0.00	1.96	0.48	0.81	0.00	2.18	0.75	0.86	0.00	1.76	1.44	2.24	0.00	2.16	1.81	1.86	1.07	0.00	0.07	1.17
G1P	5.00	6.28	1.64	4.30	3.59	4.54	2.38	3.15	4.73	5.01	6.61	11.04	4.33	7.44	7.65	1.91	1.00	0.45	0.83	5.11
a-GP	2.96	1.70	0.56	1.18	30.60	3.68	1.09	1.01	23.44	24.77	6.05	5.13	17.88	22.23	15.95	0.27	0.13	0.30	0.21	2.05
NAD	8.04	17.02	13.60	16.67	7.36	15.47	16.26	19.50	7.02	11.05	21.70	32.58	6.42	12.03	20.14	0.38	0.56	1.14	1.37	25.22
Orotate	84.84	117.30	118.87	136.60	115.41	262.67	200.14	250.32	97.67	205.32	287.98	301.28	84.20	152.27	191.95	39.49	29.45	26.10	33.94	381.45
Ru5P	0.64	0.58	0.21	0.39	0.23	0.73	0.25	0.40	0.80	3.37	3.49	8.29	0.70	5.57	7.91	0.29	0.74	1.05	0.49	1.38
CMP	0.37	0.13	0.06	0.14	0.24	0.68	0.10	0.19	0.28	1.52	0.40	5.02	0.50	3.90	4.29	0.10	0.15	0.14	0.04	0.25
b-GP	5.54	3.14	1.03	2.06	56.48	6.42	1.90	1.80	42.44	45.59	10.87	9.48	32.10	39.95	28.83	1.91	1.70	2.16	2.25	10.33
MEP	0.10	0.22	0.20	0.43	0.05	0.15	0.15	0.21	0.04	0.06	0.14	0.18	0.03	0.04	0.10	0.56	0.57	0.70	0.56	0.30
F1P	0.43	0.68	0.19	0.36	2.68	3.25	0.34	0.26	2.07	5.80	4.34	6.18	2.33	5.56	6.29	0.67	0.46	0.47	0.33	1.81
Pyruvate	1.10	0.81	0.40	0.65	1.33	1.93	0.44	0.99	1.13	2.08	3.43	5.18	1.03	2.11	3.01	2.25	1.30	1.23	1.73	0.91
R1P	0.49	0.72	0.33	0.61	0.18	0.92	0.47	0.48	0.48	1.97	2.03	2.32	0.54	2.96	4.17	0.13	0.39	0.57	0.00	0.84
UMP	3.14	3.78	1.41	2.03	2.78	6.89	2.37	1.60	2.12	8.53	5.66	25.47	3.45	15.43	19.34	3.74	2.34	2.18	1.62	6.90
AICAR	0.00	0.00	0.00	0.00	0.03	0.04	0.00	0.00	0.08	0.02	0.00	0.02	0.19	0.03	0.02	0.00	0.00	0.00	0.00	0.00
GMP	1.30	0.72	0.21	0.54	1.44	2.13	0.24	0.33	1.09	4.00	1.41	13.40	1.31	5.92	8.58	0.75	0.82	1.02	1.06	2.62
IMP	0.80	0.41	0.08	0.16	2.50	0.71	0.44	0.33	2.24	2.40	2.22	7.60	2.03	4.80	6.36	0.04	0.08	0.15	0.04	3.29
DHAP	0.27	0.46	0.10	0.18	0.92	0.96	0.13	0.26	1.31	11.37	2.83	4.06	1.19	14.70	7.94	0.20	0.44	0.47	0.20	0.39
TMP	0.03	0.01	0.01	0.01	0.09	0.03	0.01	0.01	0.09	0.02	0.01	0.04	0.08	0.04	0.03	0.00	0.02	0.01	0.00	0.03
AMP	8.47	4.99	1.51	3.30	9.48	15.87	2.21	2.45	7.13	19.19	11.08	46.48	8.25	28.95	35.76	5.62	4.96	6.45	4.83	15.34
Pantothenate	14.78	14.82	10.40	20.58	7.77	40.98	17.98	21.91	4.14	18.07	44.74	33.88	2.58	10.09	18.86	4.20	1.37	2.73	1.97	63.74
Nicotinate	11.18	23.64	23.23	29.73	8.08	26.19	22.37	34.52	5.18	16.27	19.35	7.14	3.44	8.81	4.41	27.51	4.14	7.58	16.54	31.08
cAMP	0.08	0.18	0.05	0.14	0.10	0.23	0.12	0.15	0.08	0.14	0.16	0.26	0.07	0.15	0.11	0.01	0.02	0.01	0.01	0.32
Succinate	0.00	0.03	0.01	0.01	0.00	0.00	0.02	0.01	0.00	0.05	0.00	0.01	0.00	0.03	0.02	0.00	0.00	0.02	0.00	0.00
Glutathione	0.62	2.89	2.26	3.89	0.35	2.08	2.28	3.14	0.39	0.71	1.73	0.81	0.47	0.39	0.44	1.32	0.32	0.63	0.59	2.79
Malate	10.66	2.98	1.78	0.91	15.65	21.65	9.87	8.03	13.28	18.14	23.28	25.54	10.82	18.89	17.56	9.72	5.06	5.94	6.17	22.99
UDP-Glc	7.93	59.04	53.31	66.48	6.03	23.87	62.14	75.60	10.34	4.59	12.11	28.28	14.22	7.38	11.88	0.38	0.58	1.08	1.31	25.63
XMP	0.02	0.05	0.06	0.13	0.01	0.40	0.07	0.13	0.01	0.73	0.40	2.30	0.02	2.30	1.74	0.02	0.00	0.00	0.00	0.73
CDP	0.02	0.01	0.00	0.00	0.01	0.02	0.01	0.02	0.02	0.06	0.02	0.11	0.11	0.22	0.09	0.01	0.05	0.02	0.03	0.02
2OG	0.33	0.08	0.08	0.08	2.60	0.30	0.11	0.14	5.95	3.44	4.90	35.46	4.82	12.17	21.38	11.59	5.77	6.75	7.76	1.56
GDP	0.12	0.12	0.07	0.11	0.06	0.16	0.08	0.10	0.05	0.22	0.08	0.40	0.10	0.20	0.21	0.02	0.00	0.00	0.05	0.16
6PGA	0.20	0.28	0.15	0.34	0.62	0.18	0.26	0.44	0.34	0.17	0.35	0.55	0.15	0.35	0.44	0.95	0.27	0.88	0.63	0.00
UDP	0.22	0.33	0.08	0.11	0.11	0.45	0.22	0.15	0.10	0.33	0.27	0.92	0.16	0.58	0.82	0.00	0.05	0.03	0.00	0.44
3PGA	0.05	0.08	0.06	0.14	0.04	0.07	0.06	0.15	0.05	0.07	0.13	0.22	0.07	0.15	0.16	0.06	0.08	0.05	0.08	0.16
NADH	0.02	0.02	0.01	0.00	0.06	0.05	0.02	0.01	0.12	0.06	0.04	0.08	0.17	0.03	0.04	0.01	0.01	0.01	0.03	0.07
NADP	0.57	2.07	0.87	1.84	0.22	1.72	1.17	1.89	0.16	1.68	1.92	16.54	0.11	3.22	9.05	0.62	0.22	0.25	0.47	2.12
ADP	0.75	0.99	0.59	0.91	0.62	1.66	0.61	0.88	0.52	1.19	1.42	2.93	0.51	1.42	2.21	0.01	0.04	0.04	0.07	1.77
Citrate	0.00	0.00	0.00	0.00	11.47	0.00	0.00	0.00	9.82	0.31	0.00	0.00	10.97	0.27	0.00	0.00	64.82	0.77	0.11	0.00
FBP	0.03	0.16	0.09	0.15	1.44	1.03	0.24	0.25	1.12	5.34	4.81	2.38	0.73	3.14	4.78	0.10	0.06	0.08	0.13	0.48
PEP	1.31	2.48	2.51	2.68	0.14	1.71	1.85	3.73	0.20	0.38	0.79	0.81	0.20	0.38	0.49	0.57	0.09	0.12	0.39	2.24
Isocitrate	38.94	108.09	111.14	163.16	11.69	56.47	68.00	100.72	9.22	21.05	51.33	72.13	10.35	22.46	40.68	320.83	65.53	132.31	226.06	110.87
FMN	0.07	0.09	0.06	0.10	0.05	0.12	0.08	0.12	0.09	0.10	0.18	0.54	0.08	0.23	0.42	0.12	0.23	0.20	0.09	0.19
GTP	1.70	1.54	0.94	1.57	0.74	1.81	1.42	1.59	0.85	1.90	1.42	2.65	0.51	0.88	1.47	0.09	0.06	0.06	0.13	1.57
CTP	0.79	0.85	0.60	0.99	0.21	0.93	0.81	1.35	0.45	0.59	0.55	2.11	0.38	0.76	1.34	0.03	0.07	0.05	0.08	0.60
UTP	5.91	8.55	1.29	1.94	3.30	7.86	6.24	3.89	5.17	4.76	3.47	10.36	5.40	8.95	9.17	0.09	0.02	0.10	0.17	4.69
ATP	8.59	11.22	6.14	9.74	6.23	13.62	8.30	10.83	6.79	9.89	10.61	13.29	5.17	6.84	7.70	0.24	0.07	0.30	0.45	9.43
BPG	1.51	4.96	3.74	5.02	0.09	2.30	2.12	4.20	0.11	0.33	0.89	0.80	0.07	0.22	0.48	0.00	0.03	0.02	0.01	2.56
FAD	0.78	1.09	0.72	0.99	0.57	0.50	0.42	0.57	0.58	0.36	0.35	0.42	0.54	0.39	0.36	0.02	0.04	0.15	0.15	0.43
PRPP	0.33	0.12	0.02	0.02	0.11	0.24	0.09	0.08	0.31	0.51	0.14	0.39	0.29	0.28	0.25	0.00	0.04	0.00	0.00	0.14
NADPH	0.02	0.02	0.02	0.01	0.04	0.07	0.05	0.05	0.06	0.09	0.08	0.14	0.05	0.10	0.10	0.02	0.02	0.02	0.04	0.09
PQQ	0.37	0.73	0.56	0.84	0.22	0.48	0.73	0.82	0.20	0.57	0.96	1.39	0.14	0.54	1.03	0.44	0.18	0.24	0.40	1.07
CoA	0.42	1.41	0.41	0.47	0.33	0.45	0.51	0.72	0.25	1.06	1.10	4.14	0.12	2.28	3.38	0.05	0.04	0.02	0.06	0.66
3HB CoA	0.03	0.01	0.00	0.00	0.01	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Malonyl CoA	0.02	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Acetyl CoA	1.02	0.79	0.27	0.41	0.72	0.65	0.59	0.52	0.69	0.83	2.38	1.39	0.41	0.74	0.90	0.08	0.03	0.04	0.07	1.51
HMG CoA	0.07	0.18	0.07	0.15	0.01	0.04	0.09	0.13	0.01	0.02	0.15	0.47	0.00	0.03	0.17	0.00	0.00	0.00	0.00	0.22
Butanoyl CoA	0.13	0.04	0.00	0.01	0.06	0.11	0.02	0.01	0.05	0.03	0.02	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.03

Table S 3: complete numerical results of PCA loading score from all samples of *Y. lipolytica* PO1d in different C:N ratio

Metabolites	PC1	Metabolites	PC2
Isoleucine	0.145755	FAD	0.165061
Cysteine	0.142235	UDP-Glc	0.155623
Methionine	0.126357	BPG	0.147082
Tryptophan	0.100284	PEP	0.120143
Hexose	0.09271	Glutathione	0.10866
MEP	0.089162	HMG CoA	0.095616
Pyroglutamate	0.084021	NAD	0.095446
Isocitrate	0.066814	Butanoyl CoA	0.086196
Phenylalanine	0.061142	CTP	0.081991
6PGA	0.059523	Aspartate	0.071386
Histidine	0.056921	Nicotinate	0.068867
Valine	0.055293	ADP	0.060991
Arginine	0.053005	ATP	0.060981
Tyrosine	0.045344	cAMP	0.06056
Trehalose	0.042384	Acetyl CoA	0.05647
Leucine	0.019327	Malonyl CoA	0.041964
Lysine	0.018816	Pantothenate	0.040872
Glutamine	0.01623	Orotate	0.038411
Sucrose	0.014742	GTP	0.037415
Thymidine	0.012312	GDP	0.029493
Guanosine	0.002933	CoA	0.028167
Aspartate	0.001326	PRPP	0.02448
Lactate	-0.00261	UDP	0.019263
Proline	-0.00268	Thymine	0.016133
Nicotinate	-0.00472	G1P	0.008184
Asparagine	-0.00514	Xanthine	0.007121
2OG	-0.00804	PQQ	0.007069
Glutamate	-0.01898	a-GP	0.004156
Glutathione	-0.02537	TMP	0.003011
Adenosine	-0.03873	UTP	0.001817
Cytidine	-0.05024	S7P	-0.0144
CDP	-0.05691	G6P	-0.01464
PEP	-0.0603	Isocitrate	-0.01542
Serine	-0.06132	Glutamate	-0.01672
Pyruvate	-0.06164	Glycolate	-0.0177
FAD	-0.06364	Guanine	-0.01915
FMN	-0.06573	Cytidine	-0.02947
Malonyl CoA	-0.07867	NADH	-0.0304
Metabolites	PC1	Metabolites	PC2
Threonine	-0.07893	MEP	-0.03586
Glycolate	-0.07969	IMP	-0.03662
Guanine	-0.08285	XMP	-0.03877

F6P	-0.0871	NADP	-0.04087
3PGA	-0.09106	Disaccharide-P	-0.04182
b-GP	-0.09619	Uridine	-0.04461
NADH	-0.09683	Proline	-0.04518
TMP	-0.09774	R1P	-0.04551
PQQ	-0.09795	Cysteine	-0.04687
Xanthine	-0.10141	Histidine	-0.04719
GMP	-0.10616	b-GP	-0.05129
BPG	-0.1064	3PGA	-0.05354
UDP-Glc	-0.10683	Isoleucine	-0.07026
Butanoyl CoA	-0.10789	Glyoxylate	-0.0783
DHAP	-0.11173	FBP	-0.08272
AMP	-0.1119	6PGA	-0.08335
G6P	-0.11365	F6P	-0.09061
Ru5P	-0.11721	Tryptophan	-0.09124
Glyoxylate	-0.12274	Methionine	-0.09148
F1P	-0.1237	NADPH	-0.09714
Malate	-0.12388	Lactate	-0.1027
FBP	-0.12425	CMP	-0.10838
NADP	-0.12436	Pyroglutamate	-0.11502
R5P	-0.12499	Malate	-0.1174
a-GP	-0.12722	F1P	-0.12053
UMP	-0.12779	Hexose	-0.12395
NADPH	-0.1288	Tyrosine	-0.12416
CMP	-0.13233	DHAP	-0.12422
Disaccharide-P	-0.13498	Valine	-0.12681
Uridine	-0.13546	UMP	-0.12897
GTP	-0.13847	Trehalose	-0.12926
Pantothenate	-0.13864	Ru5P	-0.13552
Thymine	-0.13922	R5P	-0.14057
R1P	-0.1412	Guanosine	-0.14314
HMG CoA	-0.1421	FMN	-0.14863
UTP	-0.14263	Thymidine	-0.15028
NAD	-0.14501	AMP	-0.15247
ATP	-0.14515	GMP	-0.15424
CTP	-0.1469	CDP	-0.16103
PRPP	-0.14924	Glutamine	-0.16724
GDP	-0.14924	Phenylalanine	-0.16869
IMP	-0.15142	Asparagine	-0.16937
Metabolites	PC1	Metabolites	PC2
XMP	-0.15358	Lysine	-0.16992
Orotate	-0.15416	Adenosine	-0.17577
G1P	-0.15426	Sucrose	-0.18202
Acetyl CoA	-0.15508	Threonine	-0.18593
cAMP	-0.15751	Leucine	-0.19259
CoA	-0.15761	Pyruvate	-0.19409

S7P	-0.15953	Serine	-0.20211
UDP	-0.16217	2OG	-0.2043
ADP	-0.16304	Arginine	-0.20713

Table S 4: complete numerical results of PCA loading score from each time point of *Y. lipolytica* PO1d in different C:N ratio

time	12 Hours		24 Hours		36 Hours		48 Hours	
Metabolite/PC	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
2OG	0.066496	0.138707	-0.00339	0.154354	0.092816	0.113785	-0.05301	0.187991
3PGA	0.105384	0.008952	-0.09676	0.046651	0.116266	0.087271	-0.11285	-0.03149
6PGA	-0.01925	0.055817	0.142049	0.069062	-0.00605	0.172311	0.067418	0.125478
Acetyl CoA	-0.1161	-0.07423	-0.13855	-0.05675	0.133268	-0.06677	-0.13869	0.000324
Adenosine	0.153168	-0.02307	0.000866	0.12965	0.118453	0.096184	-0.05264	0.188149
ADP	-0.14799	-0.05492	-0.14888	-0.01934	0.147741	0.006244	-0.14554	0.012393
a-GP	-0.08448	0.146008	-0.09464	0.112821	0.151623	-0.00355	-0.10955	-0.07318
AMP	-0.09151	-0.00591	-0.1203	0.117303	0.123982	0.113194	-0.12489	0.105359
Arginine	0.151969	-0.03329	0.095614	0.153492	0.00048	0.205324	0.039352	0.187493
Asparagine	0.04964	-0.18639	-0.03615	0.098747	0.038073	0.176946	-0.01897	0.187377
Aspartate	0.084791	-0.17995	0.015921	-0.17421	-0.04192	-0.05702	0.056735	-0.13198
ATP	-0.14629	-0.05639	-0.12606	-0.10949	0.11541	-0.09522	-0.11923	-0.09446
b-GP	-0.08255	0.146578	-0.11991	0.113437	0.13584	0.083317	-0.09088	-0.01395
BPG	-0.09277	-0.16219	-0.02221	-0.19435	0.051882	-0.18643	0.010867	-0.19189
Butanoyl CoA	-0.0835	-0.11515	-0.12594	-0.10801	0.049757	-0.09855	-0.07447	-0.04167
cAMP	-0.11802	0.008675	-0.13367	-0.08369	0.124124	-0.05139	-0.12509	-0.03621
CDP	0.050522	0.050485	-0.07625	0.158885	0.075534	0.153035	-0.1145	0.107544
CMP	-0.07231	-0.01037	-0.12368	0.128332	0.143512	0.02682	-0.13485	0.058023
CoA	-0.10674	-0.10207	-0.09892	0.011545	0.140275	-0.06944	-0.13956	-0.05654
CTP	-0.09735	-0.12808	-0.12941	-0.09325	0.117724	-0.0671	-0.1159	-0.03633
Cysteine	0.147801	0.026537	0.121171	0.115777	-0.13884	0.070055	0.133739	0.085457
Cytidine	0.114003	-0.09865	-0.06363	0.004289	0.063286	-0.02767	-0.03694	-0.0779
DHAP	-0.04783	0.154345	-0.09843	0.131686	0.140505	0.078207	-0.1317	0.063351
Disaccharide-P	-0.12057	0.091327	-0.10585	0.074495	0.141012	0.038845	-0.13814	0.014808
F1P	-0.07661	0.152632	-0.13389	0.11242	0.137941	0.057105	-0.12198	0.104149
F6P	0.153474	-0.02343	-0.14944	-0.05561	0.138813	-0.00167	-0.04766	0.145281
FAD	-0.14841	-0.05207	-0.04713	-0.17369	0.035423	-0.15376	-0.0242	-0.16205
FBP	-0.07265	0.15215	-0.08811	0.109311	0.140465	0.055881	-0.14503	0.030018
FMN	0.140094	-0.01125	0.014376	0.116597	0.121937	0.082053	-0.12462	0.078963
G1P	-0.12502	-0.02958	-0.14339	-0.04377	0.147641	0.02916	-0.13863	-0.00415
G6P	-0.1301	0.033946	-0.14657	0.024315	0.143425	0.011387	-0.14045	-0.01663
GDP	-0.11964	-0.08175	-0.15943	0.019804	0.116912	0.007459	-0.1409	-0.05391
Glutamate	0.063133	-0.19169	-0.02584	-0.10984	0.035643	0.017533	0.06626	-0.01166
Glutamine	0.150879	0.040505	-0.06202	0.136884	0.048507	0.178358	0.063553	0.143512
Glutathione	-0.07148	-0.12258	-0.02147	-0.18426	-0.02029	-0.14781	0.036876	-0.1733
Glycolate	-0.13118	-0.08173	-0.04862	-0.16176	0.080606	0.082839	-0.09686	0.07031
time	12 Hours		24 Hours		36 Hours		48 Hours	
Metabolite/PC	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
GMP	-0.08397	-0.01143	-0.1132	0.130324	0.107709	0.130527	-0.12925	0.094768
GTP	-0.108	-0.13891	-0.11741	-0.08206	0.114678	-0.0966	-0.12997	-0.0541

Guanosine	0.13989	-0.06867	-0.00803	0.10584	0.029055	0.154997	0.027955	0.179024
Hexose	0.131657	0.071984	0.119435	0.080423	-0.08117	0.13795	0.058104	0.175184
Histidine	0.108843	-0.10887	0.04861	-0.15345	-0.10539	0.12016	0.116773	0.110907
HMG CoA	-0.12761	-0.11655	-0.08793	-0.16652	0.124508	-0.05262	-0.14399	0.016698
IMP	-0.1081	0.140406	-0.14295	0.077357	0.151843	-0.00347	-0.14859	-0.00427
Isocitrate	0.124492	-0.12922	0.134219	-0.08394	-0.12057	0.115449	0.121844	0.103753
Isoleucine	0.127742	0.079085	0.134516	0.052981	-0.11251	0.129441	0.135116	0.072208
Lactate	0.060211	-0.08422	-0.01175	0.053403	-0.00279	0.022749	0.000994	0.129723
Leucine	0.13953	0.036248	0.012454	0.119672	-0.02357	0.198814	0.018723	0.195281
Lysine	0.150893	0.033515	-0.05817	0.13349	0.03665	0.181455	0.062513	0.15167
Malate	-0.12895	0.055619	-0.12249	0.093492	0.132896	0.052596	-0.11886	0.092604
Malonyl CoA	-0.09562	-0.12722	-0.02018	-0.14351	0.044296	-0.04252	-0.13583	0.062413
MEP	0.151075	-0.04808	0.155346	0.01386	-0.09698	0.128336	0.081931	0.05546
Methionine	0.076428	-0.19033	0.126027	0.097491	-0.08425	0.163864	0.112682	0.125376
NAD	-0.15553	0.008067	-0.12589	-0.12789	0.13098	-0.08122	-0.13809	-0.04443
NADH	-0.07371	0.14968	-0.12319	0.030304	0.089421	0.099416	-0.13198	0.048351
NADP	-0.03446	-0.20938	-0.13536	-0.00478	0.142591	0.001198	-0.14033	0.010067
NADPH	-0.08262	0.160742	-0.13995	0.09939	0.13855	0.08111	-0.13682	0.055809
Nicotinate	-0.07389	-0.16814	-0.06687	-0.14932	-0.05894	-0.06867	0.081488	-0.0862
Orotate	-0.13833	0.048224	-0.13774	-0.01867	0.121323	-0.05547	-0.11962	-0.04243
Pantothenate	-0.09026	-0.17303	-0.1299	-0.08289	0.113476	-0.01997	-0.10635	-0.04434
PEP	-0.08241	-0.16455	-0.03713	-0.18376	-0.04356	-0.16298	0.007719	-0.16865
Phenylalanine	0.13495	-0.09982	0.112915	0.092784	-0.00611	0.177089	0.051608	0.173775
PQQ	-0.0455	-0.16443	-0.09325	-0.11234	0.138464	-0.03669	-0.14312	-0.0129
Proline	0.034032	-0.20158	-0.03935	-0.05024	0.015315	0.073569	0.046221	0.015138
PRPP	-0.0971	-0.04116	-0.12504	0.046178	0.146109	0.009052	-0.14509	-0.03295
Pyroglutamate	0.147527	-0.05215	0.087527	0.0325	-0.08949	0.154144	0.082241	0.149543
Pyruvate	0.014063	-0.02705	-0.1044	0.138201	0.115439	0.123797	-0.06238	0.166927
R1P	-0.02149	-0.03587	-0.11132	0.122399	0.143578	0.036296	-0.14282	-0.03818
R5P	-0.0423	0.133976	-0.10402	0.141345	0.139963	0.067332	-0.13793	0.072566
Ru5P	0.035917	-0.00747	-0.08843	0.140406	0.136502	0.060009	-0.14343	0.042253
S7P	-0.10676	-0.07023	-0.12743	-0.00294	0.146905	0.042107	-0.14595	-0.02831
Serine	0.108533	0.012616	-0.09941	0.139905	0.117422	0.106726	-0.07915	0.166957
Sucrose	0.11533	0.075644	0.072667	0.163754	-0.03553	0.176514	-0.06031	0.176308
Threonine	-0.02003	-0.16895	-0.11847	0.132678	0.117962	0.080514	-0.0885	0.155568
Thymidine	0.005048	0.107771	-0.02171	0.078557	0.05308	0.17812	0.012628	0.189308
Thymine	-0.12452	0.005223	-0.10781	-0.01193	0.103392	-0.05107	-0.09902	-0.0237
TMP	-0.08674	0.132182	-0.12578	0.073868	0.122867	-0.06402	-0.14282	0.045285
Trehalose	0.147143	0.039696	0.154418	0.04489	-0.0287	0.198439	-0.04102	0.169008
Tryptophan	0.132588	-0.09413	0.14768	0.056162	-0.07386	0.159227	0.092932	0.139648
Tyrosine	0.12448	-0.1259	0.104294	0.106071	-0.07011	0.179926	0.00445	0.113202
UDP	-0.09779	-0.11129	-0.15098	-0.00141	0.140307	-0.07326	-0.1495	-0.02365
UDP-Glc	-0.11451	0.069775	-0.08906	-0.16423	-0.03129	-0.1764	-0.0084	-0.17409
UMP	-0.04249	-0.03843	-0.12684	0.095971	0.142307	0.043522	-0.12015	0.116807
Uridine	-0.11792	-0.10775	-0.12887	-0.02666	0.142206	0.005085	-0.13808	0.057317
UTP	-0.133	-0.03191	-0.1299	-0.08484	0.12032	-0.04184	-0.14252	0.021744

Valine	0.137499	-0.00945	-0.04219	0.079492	-0.07681	0.162085	0.110542	0.106998
Xanthine	0.014157	-0.20757	-0.134	-0.05062	0.082908	-0.05248	-0.09772	-0.06589
XMP	-0.12548	-0.08946	-0.15003	0.079983	0.1497	-0.01665	-0.15021	-0.0234

Table S 5: complete results of Venn diagram analysis from Figure 2-5

a	ATP	FMN	GTP	Glutamate	NADH	Proline	b-GP			
b	CTP	Glycolate	Histidine	Isocitrate	Isoleucine	Malonyl	cAMP			
c	CMP	Tyrosine	UMP							
d	TMP									
e	Aspartate	F1P	MEP	Nicotinate	Orotate	Pantothenate	Xanthine			
f	Cysteine	Methionine	Valine	a-GP						
g	6PGA	CDP								
h	Adenosine	Hexose	NAD	PQQ						
i	2OG	Acetyl	GDP	HMG	Pyruvate	Ru5P	Serine			
j	Disaccharide	PRPP	R1P	S7P	Thymidine	Uridine				
k	DHAP	Glutamine	NADPH	Tryptophan						
l	Threonine	UTP								
m	Asparagine	FBP	Guanosine	Leucine	Phenylalanin	Pyroglutamate				
n	CoA	G1P	G6P	Glutathione	R5P	Sucrose	UDP	UDP-Glc	XMP	
o	ADP	Arginine	BPG	F6P	FAD	IMP	Lysine	NADP	PEP	Trehalose

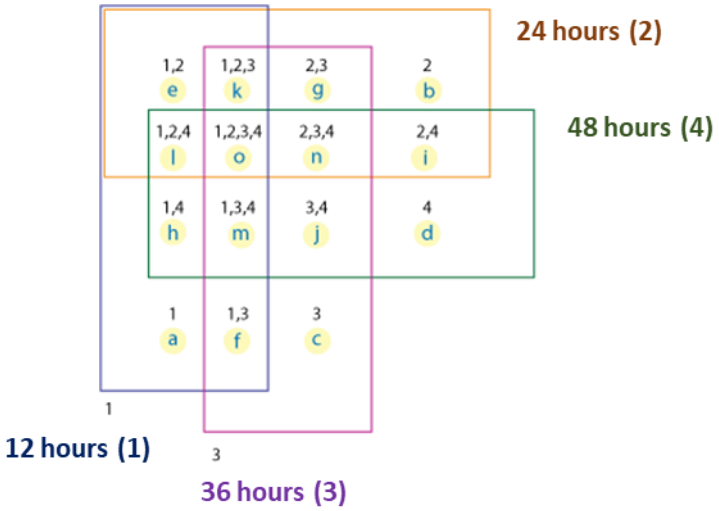


Table S 6: complete numerical results of all detected normalized peak area of six *Yarrowia* spp. in normal and nitrogen-limiting conditions

	Po1dSC	Po1d-N	8061SC	8061-N	2304SC	2304-N	21924SC	21924-N	1694SC	1694-N	14894SC	14894-N
Arginine	6.990217	2.057327	0.86584	0.826636	2.58607	0.73301	1.957127	3.269987	19.1666	8.97218	8.825213	9.147495
Lysine	0.008787	0.013415	0.006504	0.006956	0.006579	0.011373	0.018351	0.049247	0.070463	0.031458	0.013839	0.023743
Histidine	1.83024	3.053487	0.31203	2.311567	0.381482	3.082547	0.990264	9.091963	4.016907	1.095249	5.18671	9.98912
4-Aminobutyrate	0.591193	0.558017	1.337263	0.488623	1.775163	0.93119	1.18162	0.480934	1.698447	0.513709	1.178647	0.454819
Serine	1.796757	0.813464	0.740169	0.441401	0.630697	0.643726	0.415261	0.424523	0.680103	0.794294	0.813314	0.415979
Asparagine	1.527447	0.744805	1.60182	0.772436	1.43528	1.19235	1.598907	0.75966	1.440667	0.942645	1.569797	0.446943
Glutamine	1.491023	1.117029	0.420871	1.182996	0.332848	2.019887	0.870966	0.763383	3.067417	2.191208	1.327446	1.145573
Threonine	3.443293	0.91151	0.464582	0.326064	0.49914	0.385837	0.262922	0.208197	0.542017	0.436281	0.756715	0.357158
Hydroxyproline	0.000952	8.89E-07	8.52E-07	1.20E-06	1.28E-06	1.21E-06	8.77E-07	1.13E-06	9.42E-07	1.63E-06	1.09E-06	1.05E-06
Cysteine	0.021209	0.015137	0.034437	0.0139	0.049858	0.004765	0.031811	0.027181	0.008127	0.025529	0.020408	0.003285
Proline	0.086083	0.098385	0.100605	0.089956	0.060933	0.046684	0.075646	0.120156	0.158439	0.064678	0.111134	0.083541
Valine	0.093832	0.070761	0.106772	0.050674	0.100268	0.044336	0.107016	0.07301	0.132735	0.092175	0.12504	0.074028
Cytidine	0.174548	0.325382	0.192182	0.140425	0.230735	0.051166	0.138755	0.12882	0.11504	0.156886	0.607036	0.374884
Methionine	0.68339	0.338146	0.672949	0.296875	0.519509	0.206432	0.383593	0.341342	0.703357	0.270619	0.650745	0.466632
Guanine	0.007459	0.003333	0.005351	0.003864	0.003683	0.004873	0.011788	0.004477	0.004667	0.005309	0.003977	0.003165
Hypoxanthine	0.349463	0.375544	0.415212	0.169524	0.218559	0.254514	0.081608	0.524693	0.168912	0.079295	0.060413	0.047529
Tyrosine	2.209493	1.076192	2.26193	1.116729	1.538433	1.011535	1.629737	1.582887	2.303423	0.693596	1.494017	2.493385
Adenine	1.035177	0.238766	0.163076	0.30687	0.123363	0.449411	0.496729	0.649159	0.528562	0.189993	0.511009	0.587709
Isoleucine	0.132252	0.069432	0.091035	0.039098	0.069824	0.037243	0.075281	0.05187	0.079059	0.047196	0.074933	0.043605
Leucine	0.034927	0.059565	0.133432	0.130287	0.124403	0.0964	0.102786	0.178926	0.153691	0.080928	0.126893	0.186135
Xanthine	1.65877	0.899488	1.960703	1.057942	1.010332	1.02923	1.644857	1.548547	1.806703	0.549034	0.851608	0.823211
Glutamate	14.2577	15.77927	3.432253	11.55607	3.126963	6.803647	7.434977	14.31803	12.21463	9.946868	6.167623	19.74455
Uridine	3.332607	1.71094	1.133098	2.117443	1.11644	0.937811	0.8142	2.38497	1.071222	1.160466	4.20998	2.203415
Aspartate	2.60894	3.04952	0.846148	2.041447	0.92287	1.994583	1.471	1.996263	0.856281	0.955867	0.953755	1.415605
Inosine	0.617533	0.378054	2.094333	0.239278	1.53921	0.303319	0.296282	0.218212	0.359567	0.86996	2.335605	0.156216
Thymine	7.31E-07	8.89E-07	0.001626	1.20E-06	1.28E-06	1.21E-06	0.002389	1.13E-06	0.003228	0.000774	0.001304	1.05E-06
Guanosine	0.463338	0.795167	1.74565	0.317803	1.020704	0.087139	0.385593	0.24814	0.176598	0.678333	2.486064	0.99625
Adenosine	0.114785	0.081455	0.054675	0.024691	0.078318	0.016733	0.017172	0.040786	0.014186	0.029961	0.114087	0.048418
Glycerate	0.159151	0.035607	0.063238	0.024622	0.078297	0.023567	0.048612	0.029773	0.047847	0.084244	0.231447	0.045943
Thymidine	0.051163	0.05236	0.081494	0.088182	0.083055	0.07608	0.072125	0.079271	0.073763	0.167831	0.062631	0.091898
Phenylalanine	0.928937	0.994976	1.206053	0.778671	0.839576	1.004104	1.578773	0.999544	1.031463	0.401016	1.096501	1.92249
Glycolate	0.005977	0.005272	0.00604	0.006695	0.008965	0.008107	0.004946	0.007002	0.00535	0.007612	0.012236	0.005017
G6P	6.9426	13.53147	6.38937	12.18824	16.5503	12.08383	8.91853	10.9618	8.963797	11.41109	13.91566	5.48898
Pyroglutamate	0.087933	0.049677	0.051684	0.064413	0.072213	0.078787	0.044204	0.084846	0.056364	0.097625	0.086357	0.076612
Tryptophan	1.964933	2.722183	3.057733	2.895087	2.552653	3.60287	2.83756	4.98054	4.908697	2.780035	5.222473	9.374045
R5P	3.541947	1.253443	5.16423	1.151501	2.37117	1.86392	2.468273	1.34391	1.645917	2.544572	4.858323	1.14109
Lactate	0.366541	0.352442	0.456389	0.480293	0.57466	0.535374	0.33051	0.563496	0.460603	0.584044	0.573397	0.405963
S7P	13.04073	3.945163	6.247427	5.086577	11.61135	2.05354	3.996377	4.186267	2.38082	6.692695	13.90025	1.90049
F6P	11.03943	7.97161	6.387263	7.57979	16.58243	10.40454	7.74898	8.316653	7.39666	10.33519	16.62076	6.165105
G1P	0.550307	0.403324	0.756985	0.615386	2.405723	0.333087	0.387719	0.527379	0.818153	1.178201	2.147806	0.501968
a-GP	11.09317	7.903003	20.7275	2.311287	32.4254	59.9928	5.049117	2.394863	30.2734	5.22131	11.08104	10.66161
TPP	0.106016	0.035349	0.177849	0.072278	0.156472	0.07667	0.230571	0.082003	0.186642	0.075168	0.139961	0.061664
NAD	23.05877	13.72913	18.22087	12.52943	27.57793	21.8871	22.96117	15.64657	20.87057	11.69544	19.01607	10.43412
GAP	8.935653	2.438403	10.05219	1.39911	3.78939	6.971373	6.76785	3.002167	7.359567	4.916863	4.183253	4.530015
Orotate	28.0978	4.312853	0.033002	0.043015	0.044907	0.096407	0.029554	0.051211	0.045572	0.053204	0.09013	0.085157
Ru5P	0.343219	0.047062	0.171217	0.036852	0.090628	0.047961	0.068553	0.030012	0.075032	0.085245	0.16543	0.024773
CMP	4.747967	0.518565	0.576828	0.18686	1.00816	0.06502	0.258872	0.212475	0.487232	0.168269	0.261802	0.024836
	Po1dSC	Po1d-N	8061SC	8061-N	2304SC	2304-N	21924SC	21924-N	1694SC	1694-N	14894SC	14894-N
MEP	0	0.002707	0	0.005572	0	0.008602	0	0	0.00436	0.001644	0.00496	0.002885
F1P	7.327443	0.88852	5.07939	0.632242	4.624333	2.14983	3.55022	0.78828	4.256143	1.117297	1.11996	0.677826
Pyruvate	0.144751	0.136442	0.058872	0.063421	0.049446	0.194048	0.06148	0.137036	0.23847	0.109597	0.07826	0.150744
UMP	26.10797	1.893067	14.17237	0.979963	28.1746	0.618966	4.740457	1.356643	3.753293	5.579135	11.88247	0.155948

GMP	13.89191	0.897322	2.659233	0.316311	4.307187	0.228147	0.673045	0.670827	1.763117	1.448056	4.465913	0.091032
IMP	6.49949	0.335154	0.542801	0.220824	1.030447	0.755439	0.527859	0.313788	0.678568	0.268533	0.097789	0.133427
DHAP	5.50123	1.44374	6.081773	0.835612	2.30068	4.192527	4.08489	1.79071	4.437357	2.991663	2.525277	2.69218
TMP	0.041271	0.006801	0.036164	0.007395	0.092083	0.007597	0.022528	0.008847	0.211421	0.101297	0.041781	0.006228
AMP	60.41573	4.12808	11.86978	1.467187	34.48917	0.987267	2.99138	2.02601	5.593697	9.811675	26.04729	0.546903
Pantothenate	14.39177	2.74954	1.56617	1.241652	0.689389	0.937545	2.520133	1.344077	3.690717	1.150794	2.50627	1.282935
Nicotinate	0.743991	1.011003	0.648922	0.113007	0.26075	0.662469	1.001501	0.667152	0.799677	0.279982	1.497646	1.10193
cAMP	0.030397	0.031988	0.032475	0.026674	0.026466	0.045268	0.071909	0.068005	0.095683	0.019367	0.079144	0.122071
Succinate	10.91062	8.783063	12.1043	10.78891	13.90387	11.52747	13.379	9.743537	16.60107	9.58821	12.00063	10.8937
Carbamoyl-P	0.013527	0.009922	0.008467	0.008402	0.011638	0.007114	0.007576	0.010398	0.010202	0.010108	0.012061	0.007363
Glutathione	0.096585	0.054228	0.021224	0.034408	0.026878	0.022431	0.0456	0.03424	0.026739	0.043577	0.049691	0.022332
Malate	10.23554	9.426563	9.470647	5.21136	18.8429	14.88477	10.71238	11.14603	15.88797	6.425098	8.473	13.0795
UDP-Glc	18.69993	19.52337	26.9317	15.49137	24.42163	15.7929	37.46737	21.79503	22.25407	12.7092	25.34914	7.590075
CDP	1.703567	0.334731	0.41732	0.168186	0.181896	0.07628	0.15996	0.1896	0.315531	0.038191	0.055399	0.073013
Acetyl-P	0.150687	0.07252	0.084459	0.079136	0.111874	0.08551	0.079884	0.104633	0.104544	0.115732	0.099669	0.077269
2OG	6.067143	1.75611	1.161763	2.51734	1.169717	2.573667	1.744843	2.110243	1.89183	1.47703	1.631241	2.279715
ADP-Glc	0.057048	0.011555	0.080102	0.003219	0.063505	0.009195	0.050605	0.006149	0.066217	0.069125	0.177975	0.00523
Fumarate	0.873169	0.53978	0.508855	0.223331	0.529315	0.982064	0.764655	0.612708	1.066285	0.234228	0.382584	0.727272
GDP	2.200443	0.46891	1.05975	0.376941	0.938609	0.261762	0.724182	0.51162	1.221521	0.28871	0.515495	0.390708
6PGA	1.060462	0.227213	0.449624	0.175083	1.198807	0.849448	0.373906	0.110144	0.107001	0.131021	0.300933	0.06594
UDP	3.987597	0.373689	1.847063	0.47077	0.823361	0.253041	1.013934	0.46768	0.884647	0.149347	0.195412	0.221698
3PGA	1.821127	0.654616	1.28813	1.63741	1.055888	1.361313	1.92953	1.84849	1.128559	0.85802	1.473123	1.307385
NADH	0.165843	0.635168	0.451266	0.411518	0.322615	0.860051	0.669742	1.021383	1.529641	0.267259	0.465977	1.126655
NADP	28.46773	4.748447	8.224783	1.90292	14.28757	1.850653	5.604357	2.45762	6.2739	4.274754	11.70693	1.83499
ADP	19.55597	3.7822	8.495397	1.762167	3.09936	1.000426	5.024867	2.701393	12.43014	1.167927	1.943177	1.42219
Citrate	14.87477	17.28883	15.6221	11.89995	6.427083	4.357153	13.93483	5.965757	12.174	5.254168	22.44233	37.03595
FBP	8.764047	0.051985	15.69503	0.071507	4.906247	1.890217	5.78075	0.072569	11.92503	0.827974	1.319511	0.673234
PEP	0.418062	0.119717	0.83635	0.47462	0.655957	0.449212	1.78338	0.667934	0.720238	0.850914	1.985278	0.268236
HMBPP	7.31E-07	0.000732	0.003493	0.002115	0.007229	1.21E-06	8.77E-07	1.13E-06	9.42E-07	1.63E-06	1.09E-06	1.05E-06
Isocitrate	0.162519	0.253545	0.321551	0.248893	0.106316	0.264977	0.166748	0.299862	0.482156	0.15716	0.154602	0.351336
FMN	0.528315	0.109482	0.351521	0.147505	0.493489	0.06651	0.490531	0.222704	0.378508	0.172203	0.309132	0.087885
2-Isopropylmalate	100.0869	27.7249	0.636916	0.105719	0.437704	0.086282	0.404915	0.191531	0.445147	0.268178	1.185701	0.26456
GTP	8.588683	1.909377	3.59656	1.78128	0.865646	2.301743	5.624823	1.943423	2.692283	0.929851	0.89751	1.65139
CTP	5.91564	1.045425	1.486963	1.019714	0.852291	0.918145	2.02066	0.935254	0.654714	0.252449	0.263496	0.419985
UTP	58.90753	3.279987	20.17783	2.47695	1.78802	3.72042	40.1054	2.588027	5.28886	0.649735	0.779687	1.489275
ATP	32.48073	9.622743	18.57323	10.08215	3.383257	10.40525	32.41133	9.923943	13.09647	3.488828	4.685064	7.670585
Indol-Ac	0.211005	0.436092	0.162424	0.448126	0.263363	0.309946	0.186475	0.462072	0.202871	0.269968	0.270457	0.448546
FAD	0.811839	0.499554	1.056922	0.347592	0.666596	0.299285	1.28115	0.548803	0.877539	0.353363	0.850332	0.533453
PRPP	0.586736	0.022082	0.130991	0.045226	0.046818	0.052553	0.281835	0.034574	0.061807	0.029727	0.016299	0.018847
NADPH	0.281784	0.218432	0.18468	0.158794	0.091801	0.138971	0.220287	0.129584	0.212501	0.043935	0.061943	0.10889
PQQ	0.023828	0.019536	0.01558	0.02588	0.027214	0.030949	0.017689	0.023917	0.017804	0.046518	0.020382	0.017585
CoA	4.53332	0.086665	0.461193	0.059019	0.356818	0.047325	0.372446	0.075093	0.251333	0.150613	0.364955	0.071201
3HB CoA	0.00055	8.89E-07	0.000858	1.16E-06	0.000751	1.19E-06	0.001021	1.11E-06	0.004263	0.001491	0.004072	0.001723
IPP,DMAPP	0.034177	0.005424	0.007894	0.007838	1.28E-06	0.01256	0.046341	1.13E-06	1.03E-06	1.63E-06	1.09E-06	1.05E-06
Malonyl CoA	0.02124	0.003026	8.52E-07	1.20E-06	1.28E-06	1.21E-06	0.004985	1.13E-06	9.42E-07	1.63E-06	1.09E-06	1.05E-06
Acetyl CoA	1.537117	0.123501	0.351916	0.084188	0.249449	0.104853	0.468097	0.09738	0.511202	0.174494	0.280952	0.123467
Succinyl CoA	0.073397	0.025799	0.030907	0.090449	0.015067	1.21E-06	0.022527	0.046253	0.035057	0.010584	0.005904	0.025064
	Po1dSC	Po1d-N	8061SC	8061-N	2304SC	2304-N	21924SC	21924-N	1694SC	1694-N	14894SC	14894-N
Crotonyl CoA	0.002576	9.22E-07	0.000523	0.000756	1.25E-06	0.000736	0.000115	0.000767	0.00128	0.000268	0.000712	1.05E-06
Butyryl CoA	0.005829	0.002733	0.00486	0.004684	0.009471	0.001669	0.004176	0.002474	0.004864	0.044237	0.103842	0.003233
MEcPP	0.00371	0.000993	0.00108	0.001236	0	0.000643	0	0	0	0	0	0.000431

Table S 7: complete numerical results of PCA loading score of six *Yarrowia* spp. in normal and nitrogen-limiting conditions

Condition	Normal and N limiting		N limiting only	
Metabolite/PC	PC1	PC2	PC1	PC2
2-Isopropylmalate	0.103402	0.0930917	-0.09242	0.0418765
2OG	-0.00258	0.20104	-0.01166	0.0183392
3HB CoA	0.123618	-0.066005	0.005494	0.159335
3PGA	0.065768	0.03277	-0.07507	-0.0126478
4-Aminobutyrate	0.102806	-0.100622	-0.06601	-0.146889
6PGA	0.119557	-0.007985	-0.08935	-0.194806
Acetyl CoA	0.154746	0.0453931	0.025702	0.0019639
Acetyl-P	0.071517	-0.001041	0.097363	0.0244836
Adenine	0.05305	0.0983977	-0.10117	0.0454338
Adenosine	0.110084	-0.035969	-0.17512	0.0464126
ADP	0.148762	0.0946601	-0.16253	0.0398319
ADP-Glc	0.159098	-0.080085	-0.14383	-0.057401
a-GP	0.06428	-0.017255	-0.01566	-0.12834
AMP	0.15918	-0.059241	-0.11027	-0.0484878
Arginine	0.019608	-0.021793	0.082657	0.18666
Asparagine	0.151087	-0.082738	-0.0316	-0.208626
Aspartate	-0.02328	0.160939	-0.12511	-0.0360241
ATP	0.039215	0.21446	-0.14238	-0.0586934
Butyryl CoA	0.093391	-0.178241	-0.01201	0.0894121
cAMP	-0.03413	0.130863	-0.05269	0.109054
Carbamoyl-P	0.045674	-0.008858	0.008801	0.0856227
CDP	0.115023	0.151404	-0.163	-0.0073873
Citrate	0.075415	0.0315059	-0.13479	0.108801
CMP	0.152437	0.0290996	-0.11883	-0.0573123
CoA	0.168623	0.0256591	-0.09242	0.0556405
Crotonyl CoA	0.065657	0.0164996	-0.03065	-0.130453
CTP	0.080466	0.191623	-0.14773	-0.110895
Cysteine	0.060433	-0.079802	0.042266	0.0319102
Cytidine	0.084608	-0.087531	-0.14477	0.122284
DHAP	0.091626	0.0717803	0.069219	-0.0599459
F1P	0.134351	0.0588204	0.045643	-0.162643
F6P	0.076177	-0.161603	-0.04608	-0.14694
FAD	0.14013	-0.019495	-0.14708	0.104453
FBP	0.121541	-0.002224	0.111286	-0.0631346
FMN	0.157448	-0.021727	-0.04558	0.063661
Fumarate	0.029793	0.157106	-0.05262	-0.0620273
G1P	0.067992	-0.195991	0.043243	0.110832
Condition	Normal and N limiting		N limiting only	
Metabolite/PC	PC1	PC2	PC1	PC2
G6P	0.005826	-0.160854	-0.07885	-0.154028
GAP	0.089846	0.0716688	0.067641	-0.0592601

GDP	0.15809	0.0561511	-0.14627	0.0554731
Glutamate	-0.06959	0.162177	-0.04987	0.212072
Glutamine	-0.03673	0.0370184	0.129256	-0.12324
Glutathione	0.09229	0.0580779	-0.04629	0.0007241
Glycerate	0.138266	-0.090403	-0.04789	0.156374
Glycolate	0.024393	-0.138418	0.00334	-0.0139832
GMP	0.162537	-0.036996	-0.11613	-0.0649017
GTP	0.051263	0.217445	-0.11257	-0.0928389
Guanine	0.018204	0.133898	0.107324	-0.0399853
Guanosine	0.097963	-0.10912	-0.12573	0.138445
Histidine	-0.0518	0.0386425	-0.12493	0.0807208
Hydroxyproline	0.089489	0.126482	0.175721	-0.0170831
Hypoxanthine	0.013869	0.111633	-0.09832	-0.115445
IMP	0.096764	0.144872	0.01672	-0.202985
Indol-Ac	-0.12216	0.0444772	-0.13322	0.116236
Inosine	0.124415	-0.144116	-0.06965	-0.122786
IPP,DMAPP	0.044929	0.135678	-0.06966	-0.176349
Isocitrate	-0.01353	0.0563529	-0.0612	0.0459708
Isoleucine	0.155382	0.0610317	-0.08802	0.0516704
Lactate	-0.08078	-0.076254	0.110884	0.0138871
Leucine	-0.00415	-0.102412	-0.0598	0.0974945
Lysine	-0.06935	0.0676719	0.06232	0.126621
Malate	0.01372	0.0459102	-0.03992	-0.0311263
Malonyl CoA	0.067771	0.136773	-0.08963	-0.0100576
MEcPP	0.076149	0.143813	-0.06925	-0.070987
MEP	-0.04145	-0.05511	-0.0337	-0.157402
Methionine	0.14396	-0.040276	-0.11833	0.158333
NAD	0.13176	-0.029099	-0.07573	-0.16968
NADH	-0.06847	0.102712	-0.10088	0.0220534
NADP	0.171188	-0.027788	-0.18122	-0.0203405
NADPH	0.088082	0.168881	-0.14694	-0.0650498
Nicotinate	0.040694	0.0196148	-0.08942	0.0705267
Orotate	0.057965	0.141514	-0.08997	-0.0026791
Pantothenate	0.117087	0.123351	-0.14117	0.0507348
PEP	0.094228	-0.127373	-0.00163	-0.0445573
Phenylalanine	0.029591	0.0588961	-0.08725	0.10011
PQQ	-0.09531	-0.006068	0.167137	-0.0294234
Proline	0.060071	-0.012354	-0.13274	0.102127

Condition	Normal and N limiting		N limiting only	
Metabolite/PC	PC1	PC2	PC1	PC2
PRPP	0.096457	0.16166	0.037328	-0.156634
Pyroglutamate	-0.05692	0.0280922	0.096158	0.0463282
Pyruvate	-0.04943	0.154706	-0.00307	-0.0113046
R5P	0.141562	-0.087383	-0.03236	-0.110391
Ru5P	0.164235	-0.029956	-0.049	-0.160393

S7P	0.130593	-0.097139	-0.11294	-0.0105482
Serine	0.122108	0.0349074	-0.02389	-0.09042
Succinate	0.070119	-0.007672	-0.00619	-0.0399379
Succinyl CoA	0.045155	0.0249999	-0.0272	0.1842
Threonine	0.111067	0.0657712	-0.04719	-0.0137839
Thymidine	-0.09941	-0.020118	0.166576	0.0423251
Thymine	0.085673	-0.086634	0.175721	-0.0170831
TMP	0.075868	-0.077044	0.16	0.0176543
TPP	0.126286	-0.078332	-0.04588	-0.0670686
Tryptophan	-0.05311	0.0092748	-0.03406	0.144703
Tyrosine	0.108693	0.0354409	-0.08242	0.124355
UDP	0.142493	0.106264	-0.13886	-0.0355074
UDP-Glc	0.114452	-0.075344	-0.13299	-0.0743309
UMP	0.163604	-0.055885	-0.12475	-0.0877905
Uridine	0.080108	-0.019742	-0.16627	0.0504473
UTP	0.102091	0.172789	-0.14694	-0.12234
Valine	0.118432	-0.099294	-0.00094	0.135429
Xanthine	0.107323	0.0996565	-0.11556	-0.0185456

Table S 8: The product ion m/z from fragment analysis of ergosterol.
The list of product ion derived from ergosterol fragmentation. “Nitrogen limiting C12”, ergosterol from the yeast cells cultivated in nitrogen limiting condition; “Nitrogen limiting C13”, ergosterol from the yeast cells cultivated in the presence of [1-^{13}C] glucose in nitrogen limiting condition; “Normal C12”, ergosterol from the yeast cells cultivated in normal condition; “Normal C13”, ergosterol from the yeast cells cultivated in the presence of [1-^{13}C] glucose in normal condition.

Nitrogen limiting C12		Nitrogen limiting C13		Normal C12		Normal C13	
m/z	intensity	m/z	intensity	m/z	intensity	m/z	intensity
41	256554	39	1855	41	110237	55	2491
55	353476	41	2378	55	112282	57	5178
57	1136353	57	6959	57	260743	59	1740
69	3868992	60	3711	69	1177071	69	4425
81	1171627	67	2378	81	256293	70	7226
83	1379537	69	8117	83	332143	71	1855
93	290600	70	6444	93	137154	77	1855
95	488329	71	5757	95	224395	81	4302
97	198103	82	4640	97	62890	82	4222
105	586472	84	1582	105	138456	93	927
107	547729	95	1855	107	263943	95	3711
109	451180	96	1855	109	162104	98	5568
117	255809	96	5568	117	87145	105	1855
119	459903	107	1771	119	121133	107	3401
121	587729	109	3711	121	112940	108	1855
123	420763	120	1855	123	127666	109	11489
125	206473	121	3711	125	64215	111	927
131	569455	124	1855	129	105519	117	4306
133	696790	134	2658	131	185720	118	1855
135	369252	136	3711	133	201354	122	1855
143	452733	143	12547	135	176848	124	2699
145	1553569	145	2783	143	176650	125	1771
147	775393	146	3711	145	582841	127	3711
149	582122	147	7426	147	238894	131	1855
157	979845	148	8616	149	106869	132	1855
159	1943788	152	3932	157	211179	133	2783
160	319400	157	2783	159	530682	145	1855
161	591147	159	7278	161	181820	149	2491
163	231201	160	9501	163	71896	150	1855
171	567326	161	7363	171	174409	159	1855
173	446233	162	1855	173	144025	164	1855
Nitrogen limiting C12		Nitrogen limiting C13		Normal C12		Normal C13	
m/z	intensity	m/z	intensity	m/z	intensity	m/z	intensity
183	297834	163	3377	175	77528	166	3597

185	490710	172	1771	177	65388	172	1855
187	361713	173	7426	183	85516	185	1855
197	300844	177	2158	185	222970	186	2878
199	445516	184	3711	187	214157	188	2783
201	463306	187	4431	191	69213	193	3711
211	327388	188	1839	197	144018	198	3711
213	628305	200	4139	199	322460	199	2783
225	418192	201	1855	201	188933	200	3711
227	260777	206	3711	203	65473	201	2158
239	250317	216	2491	211	83486	202	3711
253	657300	217	1855	213	289662	211	1855
255	242392	225	2783	225	114123	212	3711
281	238069	240	1855	227	59420	216	3628
295	724590	244	1855	229	65283	225	3711
309	481630	245	1855	239	197344	227	3711
379	1873541	256	3723	241	61220	228	8253
380	642199	257	1855	253	287472	241	927
		258	7258	267	88930	242	7426
		275	1855	269	58872	248	927
		284	3711	281	134583	255	1855
		296	5568	295	269126	257	6204
		301	1855	297	67253	261	3711
		312	8525	309	160762	283	1855
		313	3711	323	98457	295	927
		337	2878	379	940153	307	1855
		383	20654			311	1855
						312	3711
						325	5568
						366	1188
						383	2726

Table S 9 DXR gene candidate from similarity search

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382128|YALI0_B14267g	KZM93394.1	39	hypothetical protein DCAR_016639 [Daucus carota subsp. sativus] carrot	plantae
CR382132|YALI0_F09185g	GAU48374.1	32	hypothetical protein TSUD_405360 [Trifolium subterraneum]	plantae
CR382132|YALI0_F23749g	WP_086172692.1	26	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Streptomyces pharetrae]	bacteria
CR382131|YALI0_E12573g	WP_086172692.1	25	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Streptomyces pharetrae]	bacteria
CR382130|YALI0_D25190g	OIW18865.1	24	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D19492g	OIW18865.1	24	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F06314g	KZM93394.1	23	hypothetical protein DCAR_016639 [Daucus carota subsp. sativus] carrot	plantae
CR382128|YALI0_B10758g	OIW18865.1	23	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E23496g	OIW18865.1	23	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D15708g	WP_086172692.1	23	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Streptomyces pharetrae]	bacteria
CR382131|YALI0_E02794g	KZM93394.1	22	hypothetical protein DCAR_016639 [Daucus carota subsp. sativus] carrot	plantae
CR382129|YALI0_C17325g	KZM93394.1	22	hypothetical protein DCAR_016639 [Daucus carota subsp. sativus] carrot	plantae
CR382129|YALI0_C10967g	OIW18865.1	22	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D02101g	OIW18865.1	21	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D04334g	OIW18865.1	21	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E31361g	OIW18865.1	21	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C06039g	KZM93394.1	20	hypothetical protein DCAR_016639 [Daucus carota subsp. sativus] carrot	plantae
CR382131|YALI0_E25135g	OIW18865.1	20	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D20966g	OIW18865.1	20	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E27632g	OIW18865.1	19	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E33803g	OIW18865.1	19	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E34375g	OIW18865.1	19	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B02816g	OIW18865.1	18	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C19382g	KZM93394.1	17	hypothetical protein DCAR_016639 [Daucus carota subsp. sativus] carrot	Plantae

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382132|YALI0_F25047g	KZV42249.1	17	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplastic [Dorcoceras hygrometricum]	plantae
CR382131|YALI0_E06519g	OIW18865.1	17	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C16797g	WP_086172692.1	17	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Streptomyces pharetrae]	bacteria
CR382132|YALI0_F02805g	KZV32814.1	16	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382128|YALI0_B00836g	KZV42249.1	16	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplastic [Dorcoceras hygrometricum]	plantae
CR382129|YALI0_C08305g	OIW18865.1	16	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D16863g	OIW18865.1	16	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B18700g	OIW18865.1	16	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E32351g	KZV42249.1	15	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplastic [Dorcoceras hygrometricum]	plantae
CR382132|YALI0_F06820g	KZV42249.1	15	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplastic [Dorcoceras hygrometricum]	plantae
CR382128|YALI0_B13552g	KZV42249.1	15	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplastic [Dorcoceras hygrometricum]	plantae
CR382132|YALI0_F08165g	OIW18865.1	15	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C21758g	OIW18865.1	15	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B08558g	OIW18865.1	15	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F08305g	OIW18865.1	15	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B13156g	KZV32814.1	14	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382132|YALI0_F18106g	KZV32814.1	14	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382132|YALI0_F03113g	OIW18865.1	14	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E24563g	OIW18865.1	14	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B17556g	OIW18865.1	14	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D25388g	OIW18865.1	14	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B11946g	KZV32814.1	13	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382128|YALI0_B22792g	KZV32814.1	13	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382128|YALI0_B05566g	OIW18865.1	13	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D22935g	OIW18865.1	13	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D07502g	OIW18865.1	13	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae

CR382130|YALI0_D08822g	OIW18865.1	12	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
Y. lipolytica	Database	Pair score	Annotation	Classification
CR382132|YALI0_F27093g	OIW18865.1	12	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F11385g	OIW18865.1	12	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B14201g	OIW18865.1	12	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F27159g	OIW18865.1	12	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E32835g	WP_086172692.1	12	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Streptomyces pharetrae]	bacteria
CR382132|YALI0_F10857g	WP_086172692.1	12	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Streptomyces pharetrae]	bacteria
CR382127|YALI0_A02310g	GAU48374.1	11	hypothetical protein TSUD_405360 [Trifolium subterraneum]	plantae
CR382127|YALI0_A18700g	KZV32814.1	11	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382130|YALI0_D19470g	OIW18865.1	11	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B14949g	OIW18865.1	11	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F14707g	OIW18865.1	11	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C15444g	OIW18865.1	11	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D26015g	OIW18865.1	11	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C09328g	WP_094672305.1	11	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Hydrocoleum sp. CS-953]	bacteria (cyano)
CR382128|YALI0_B01628g	KZV32814.1	10	rab GTPase-activating protein 1-like [Dorcoceras hygrometricum]	plantae
CR382128|YALI0_B13178g	OIW18865.1	10	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D26213g	OIW18865.1	10	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382127|YALI0_A04697g	OIW18865.1	10	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B22528g	OIW18865.1	10	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F09746g	OIW18865.1	9	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E13750g	OIW18865.1	9	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C16665g	OIW18865.1	9	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F00572g	OIW18865.1	8	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E06501g	OIW18865.1	8	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B04268g	OIW18865.1	8	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382131|YALI0_E17743g	OIW18865.1	8	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae

CR382127|YALI0_A10230g	OIW18865.1	8	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F07557g	OIW18865.1	8	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
Y. lipolytica	Database	Pair score	Annotation	Classification
CR382129|YALI0_C09328g	WP_047156531.1	8	1-deoxy-D-xylulose-5-phosphate reductoisomerase [Trichodesmium erythraeum]	bacteria (cyano)
CR382132|YALI0_F08855g	OIW18865.1	7	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F13629g	OIW18865.1	7	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D04114g	OIW18865.1	7	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F16159g	OIW18865.1	7	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382127|YALI0_A05247g	OIW18865.1	6	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F23287g	OIW18865.1	6	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C22770g	OIW18865.1	6	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C00891g	OIW18865.1	6	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382130|YALI0_D14542g	OIW18865.1	6	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382132|YALI0_F03542g	OIW18865.1	5	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382128|YALI0_B00880g	OIW18865.1	5	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C04587g	OIW18865.1	5	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae
CR382129|YALI0_C04158g	OIW18865.1	4	hypothetical protein TanjilG_25308 [Lupinus angustifolius]	plantae

Table S 10 ipsD gene candidate from similarity search

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382131|YALI0_E20031g	KXZ55794.1	25	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E13530g	KXZ55794.1	25	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B21406g	KXZ55794.1	21	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E04697g	KXZ55794.1	21	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C18381g	KXZ55794.1	21	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382132|YALI0_F26389g	KXZ55794.1	20	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382132|YALI0_F03971g	KXZ55794.1	20	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E01364g	KXZ55794.1	19	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E28226g	KXZ55794.1	19	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B12430g	KXZ55794.1	18	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382132|YALI0_F07887g	KXZ55794.1	18	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B00902g	KXZ55794.1	18	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E21439g	KXZ55794.1	17	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C05555g	KXZ55794.1	17	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382127|YALI0_A06479g	KXZ55794.1	16	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382127|YALI0_A01331g	KXZ55794.1	16	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382127|YALI0_A14542g	KXZ55794.1	16	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382127|YALI0_A17424g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382127|YALI0_A09658g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B01078g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E19767g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C20339g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E20933g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C22704g	KXZ55794.1	15	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382132|YALI0_F08657g	KXZ55794.1	14	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382130|YALI0_D11132g	KXZ55794.1	14	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B14245g	KXZ55794.1	14	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382130|YALI0_D14080g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E32043g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382130|YALI0_D27346g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C03520g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382132|YALI0_F21901g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C03377g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E32879g	KXZ55794.1	13	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E15290g	KXZ55794.1	12	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382127|YALI0_A19470g	KXZ55794.1	12	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C21802g	KXZ55794.1	12	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382130|YALI0_D06435g	KXZ55794.1	12	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382132|YALI0_F12661g	KXZ55794.1	11	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382130|YALI0_D26279g	KXZ55794.1	11	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B13442g	KXZ55794.1	11	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382132|YALI0_F11165g	KXZ55794.1	11	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B09977g	KXZ55794.1	11	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E01100g	KXZ55794.1	10	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E26389g	KXZ55794.1	10	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C18469g	KXZ55794.1	9	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C23386g	KXZ55794.1	9	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382129|YALI0_C00913g	KXZ55794.1	8	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382131|YALI0_E01870g	KXZ55794.1	7	lcl|KXZ55794.1 hypothetical protein GPECTOR_2g1344 [Gonium pectorale]	eukaryote (algae)
CR382128|YALI0_B09317g	ODN71369.1	24	lcl|ODN71369.1 Bifunctional enzyme IspD/IspF [Methylobrevis pamukkalensis]	bacteria
CR382130|YALI0_D05929g	PIS01726.1	26	lcl|PIS01726.1 hypothetical protein COT84_00800 [Chlamydiae bacterium CG10_big_fil_rev_8_21_14_0_10_35_9]	bacteria (Chlamydiae)
CR382132|YALI0_F29975g	PIS01726.1	24	lcl|PIS01726.1 hypothetical protein COT84_00800 [Chlamydiae bacterium CG10_big_fil_rev_8_21_14_0_10_35_9]	bacteria (Chlamydiae)
CR382130|YALI0_D01694g	PLY02883.1	23	lcl|PLY02883.1 hypothetical protein C0624_07440, partial [Desulfuromonas sp.]	bacteria
CR382131|YALI0_E05643g	SDK52075.1	23	lcl|SDK52075.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [Cryobacterium psychrotolerans]	bacteria
CR382127|YALI0_A16148g	SHJ76550.1	25	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria
CR382127|YALI0_A12705g	SHJ76550.1	24	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria
CR382127|YALI0_A15059g	SHJ76550.1	20	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria
CR382127|YALI0_A09878g	SHJ76550.1	18	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria
CR382132|YALI0_F26543g	SHJ76550.1	18	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria

CR382127|YALI0_A12683g	SHJ76550.1	15	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382127|YALI0_A06523g	SHJ76550.1	14	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria
CR382131|YALI0_E10725g	SHJ76550.1	13	lcl|SHJ76550.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [[Clostridium] lactatifermentans DSM 14214]	bacteria
CR382129|YALI0_C19965g	WP_013942589.1	25	lcl|WP_013942589.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [Simkania negevensis]	bacteria (Chlamydiae)
CR382127|YALI0_A02046g	WP_015850282.1	14	lcl|WP_015850282.1 ACP S-acetyltransferase [Thermococcus sibiricus]	archaea
CR382127|YALI0_A19536g	WP_023362146.1	18	lcl|WP_023362146.1 MULTISPECIES: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [Campylobacter]	bacteria
CR382131|YALI0_E19789g	WP_026470000.1	17	lcl|WP_026470000.1 D-ribitol-5-phosphate cytidylyltransferase [Alkanindiges illinoisensis]	bacteria
CR382128|YALI0_B03190g	WP_041287897.1	20	lcl|WP_041287897.1 D-ribitol-5-phosphate cytidylyltransferase [Jonesia denitrificans]	bacteria
CR382128|YALI0_B11176g	WP_044276282.1	17	lcl|WP_044276282.1 UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase [Caldilinea aerophila]	bacteria
CR382128|YALI0_B17688g	WP_057324285.1	14	lcl|WP_057324285.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [Nocardioides sp. Soil797]	archaea
CR382127|YALI0_A15081g	WP_083498599.1	15	lcl|WP_083498599.1 UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase [Thermosulfidibacter takaii]	bacteria
CR382129|YALI0_C06490g	WP_088553157.1	23	lcl|WP_088553157.1 bifunctional UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase GlmU [Calderihabitans maritimus]	bacteria
CR382131|YALI0_E15125g	WP_088553157.1	18	lcl|WP_088553157.1 bifunctional UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase GlmU [Calderihabitans maritimus]	bacteria
CR382129|YALI0_C03003g	WP_088962956.1	25	lcl|WP_088962956.1 D-ribitol-5-phosphate cytidylyltransferase [Micromonospora purpureochromogenes]	bacteria
CR382130|YALI0_D01980g	WP_089001185.1	24	lcl|WP_089001185.1 D-ribitol-5-phosphate cytidylyltransferase [Micromonospora echinofusca]	bacteria
CR382132|YALI0_F01650g	WP_089397188.1	22	lcl|WP_089397188.1 D-ribitol-5-phosphate cytidylyltransferase [Micrococcales bacterium KH10]	bacteria
CR382130|YALI0_D18964g	WP_091263343.1	23	lcl|WP_091263343.1 D-ribitol-5-phosphate cytidylyltransferase [Micromonospora chaiyaphumensis]	bacteria
CR382128|YALI0_B16192g	WP_091561492.1	21	lcl|WP_091561492.1 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [Micromonospora pattaloongensis]	bacteria
CR382132|YALI0_F02211g	WP_092323132.1	19	lcl|WP_092323132.1 KR domain-containing protein [Cryobacterium psychrotolerans]	bacteria
CR382131|YALI0_E15378g	WP_099851926.1	20	lcl|WP_099851926.1 D-ribitol-5-phosphate cytidylyltransferase [Micromonospora sp. CNZ299]	bacteria
CR382131|YALI0_E14322g	WP_103528353.1	22	lcl|WP_103528353.1 KR domain-containing protein [Streptomyces sp. SM12]	bacteria

Table S 11 ispE gene candidate from similarity search

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382130|YALI0_D01089g	CSB35022.1	34	lcl|CSB35022.1 bifunctional aspartokinase I/homoserine dehydrogenase I [Vibrio cholerae]	bacteria
CR382132|YALI0_F22517g	CVH74888.1	27	lcl|CVH74888.1 tRNA-specific adenosine deaminase [Coriobacteriaceae bacterium CHKCI002]	bacteria
CR382131|YALI0_E13728g	CVH74888.1	21	lcl|CVH74888.1 tRNA-specific adenosine deaminase [Coriobacteriaceae bacterium CHKCI002]	bacteria
CR382131|YALI0_E33869g	EMS45141.1	17	lcl|EMS45141.1 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, chloroplastic/chromoplastic [Triticum urartu]	plants
CR382132|YALI0_F20834g	EMS45141.1	14	lcl|EMS45141.1 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, chloroplastic/chromoplastic [Triticum urartu]	plants
CR382132|YALI0_F05434g	EMS45141.1	12	lcl|EMS45141.1 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, chloroplastic/chromoplastic [Triticum urartu]	plants
CR382130|YALI0_D09735g	EMS45141.1	7	lcl|EMS45141.1 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, chloroplastic/chromoplastic [Triticum urartu]	plants
CR382132|YALI0_F24915g	EMS45141.1	5	lcl|EMS45141.1 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, chloroplastic/chromoplastic [Triticum urartu]	plants
CR382131|YALI0_E00770g	ERH17114.1	24	lcl|ERH17114.1 dimethyladenosine transferase [Actinomyces graevenitzii F0530]	bacteria
CR382132|YALI0_F23221g	KEL87537.1	26	lcl|KEL87537.1 threonine synthase [Escherichia coli 5-366-08_S3_C2]	bacteria
CR382132|YALI0_F05632g	OUX39673.1	35	lcl|OUX39673.1 diphosphomevalonate decarboxylase [Proteobacteria bacterium TMED261]	bacteria
CR382127|YALI0_A02310g	OYV88615.1	27	lcl|OYV88615.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase, partial [Ignavibacteriae bacterium 37-53-5]	bacteria
CR382132|YALI0_F13453g	WP_018086699.1	40	lcl|WP_018086699.1 homoserine kinase [Desulfurispora thermophila]	bacteria
CR382129|YALI0_C23705g	WP_026636686.1	10	lcl|WP_026636686.1 MULTISPECIES: 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Dyella]	bacteria
CR382128|YALI0_B16038g	WP_048051657.1	27	lcl|WP_048051657.1 mevalonate kinase [Methanosarcina soligelidi]	archaea
CR382131|YALI0_E12859g	WP_096866185.1	20	lcl|WP_096866185.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Mycobacterium interjectum]	bacteria
CR382127|YALI0_A15103g	WP_096866185.1	20	lcl|WP_096866185.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Mycobacterium interjectum]	bacteria
CR382127|YALI0_A14234g	WP_096866185.1	16	lcl|WP_096866185.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Mycobacterium interjectum]	bacteria

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382128|YALI0_B07755g	WP_096866185.1	16	lcl|WP_096866185.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Mycobacterium interjectum]	bacteria
CR382132|YALI0_F06556g	WP_096866185.1	15	lcl|WP_096866185.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Mycobacterium interjectum]	bacteria
CR382130|YALI0_D17314g	WP_096866185.1	15	lcl|WP_096866185.1 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase [Mycobacterium interjectum]	bacteria
CR382129|YALI0_C13090g	WP_104252492.1	19	lcl|WP_104252492.1 galactokinase [Pseudoclavibacter sp. RFBA6]	bacteria
CR382131|YALI0_E32351g	XP_011014724.1	45	lcl|XP_011014724.1 PREDICTED: ribose-phosphate pyrophosphokinase 2, chloroplastic-like [Populus euphratica]	plants
CR382132|YALI0_F25047g	XP_011014724.1	41	lcl|XP_011014724.1 PREDICTED: ribose-phosphate pyrophosphokinase 2, chloroplastic-like [Populus euphratica]	plants
CR382128|YALI0_B00836g	XP_011014724.1	32	lcl|XP_011014724.1 PREDICTED: ribose-phosphate pyrophosphokinase 2, chloroplastic-like [Populus euphratica]	plants
CR382128|YALI0_B13552g	XP_011014724.1	21	lcl|XP_011014724.1 PREDICTED: ribose-phosphate pyrophosphokinase 2, chloroplastic-like [Populus euphratica]	plants

Table S 12 ipsF gene candidate from similarity search

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382128|YALI0_B14509g	WP_009448629.1	54	lcl|WP_009448629.1 methionine adenosyltransferase [Parvimonas sp. oral taxon 393]	bacteria
CR382129|YALI0_C11473g	OIV94920.1	36	lcl|OIV94920.1 hypothetical protein TanjilG_22117 [Lupinus angustifolius]	plants
CR382130|YALI0_D15268g	CKN33129.1	24	lcl|CKN33129.1 cysteinyl-tRNA synthetase [Mycobacterium tuberculosis]	bacteria
CR382128|YALI0_B09317g	ODN71369.1	24	lcl|ODN71369.1 Bifunctional enzyme IspD/IspF [Methylobrevis pamukkalensis]	bacteria
CR382130|YALI0_D01628g	WP_005961752.1	19	lcl|WP_005961752.1 tRNA pseudouridine(13) synthase TruD [endosymbiont of Riftia pachyptila]	bacteria

Table S 13 ipsF gene candidate from similarity search

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382132|YALI0_F09834g	XP_011400578.1	37	lcl|XP_011400578.1 26S proteasome non-ATPase regulatory subunit 1 [Auxenochlorella protothecoides]	eukaryote(algae)
CR382128|YALI0_B18898g	EEE52320.1	34	lcl|EEE52320.1 hypothetical protein OsJ_34339 [Oryza sativa Japonica Group]	plants
CR382132|YALI0_F07711g	EEE52320.1	32	lcl|EEE52320.1 hypothetical protein OsJ_34339 [Oryza sativa Japonica Group]	plants
CR382129|YALI0_C02805g	XP_009035331.1	28	lcl|XP_009035331.1 hypothetical protein AURANDRAFT_71112 [Aureococcus anophagefferens]	eukaryote(algae)
CR382128|YALI0_B06006g	EEE52320.1	27	lcl|EEE52320.1 hypothetical protein OsJ_34339 [Oryza sativa Japonica Group]	plants
CR382127|YALI0_A01023g	OYV46097.1	25	lcl|OYV46097.1 hypothetical protein B7X10_05195, partial [Burkholderiales bacterium 21-58-4]	bacteria
CR382131|YALI0_E05643g	XP_009035331.1	25	lcl|XP_009035331.1 hypothetical protein AURANDRAFT_71112 [Aureococcus anophagefferens]	eukaryote(algae)
CR382131|YALI0_E19338g	OYV46097.1	24	lcl|OYV46097.1 hypothetical protein B7X10_05195, partial [Burkholderiales bacterium 21-58-4]	bacteria
CR382131|YALI0_E09471g	OYV46097.1	23	lcl|OYV46097.1 hypothetical protein B7X10_05195, partial [Burkholderiales bacterium 21-58-4]	bacteria
CR382132|YALI0_F01650g	XP_009035331.1	20	lcl|XP_009035331.1 hypothetical protein AURANDRAFT_71112 [Aureococcus anophagefferens]	eukaryote(algae)

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382131|YALI0_E09449g	OMO67493.1	19	lcl|OMO67493.1 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, bacterial-type [Corchorus capsularis]	plants
CR382128|YALI0_B22066g	OYV46097.1	18	lcl|OYV46097.1 hypothetical protein B7X10_05195, partial [Burkholderiales bacterium 21-58-4]	bacteria
CR382132|YALI0_F05038g	WP_078492500.1	14	lcl|WP_078492500.1 1-deoxy-D-xylulose-5-phosphate synthase [Streptomyces yerevanensis]	bacteria
CR382131|YALI0_E06479g	WP_078492500.1	13	lcl|WP_078492500.1 1-deoxy-D-xylulose-5-phosphate synthase [Streptomyces yerevanensis]	bacteria
CR382127|YALI0_A02541g	EEE52320.1	9	lcl|EEE52320.1 hypothetical protein OsJ_34339 [Oryza sativa Japonica Group]	plants

Table S 14 ipsH gene candidate from similarity search

Y. lipolytica	Database	Pair score	Annotation	Classification
CR382130|YALI0_D16357g	WP_083464553.1	22	lcl|WP_083464553.1 6-phosphofructokinase [Marinifilum fragile]	bacteria
CR382131|YALI0_E30217g	ERT56795.1	21	lcl|ERT56795.1 LytB-like protein [Peptoniphilus sp. BV3C26]	bacteria
CR382131|YALI0_E05753g	WP_078627944.1	21	lcl|WP_078627944.1 4-hydroxy-3-methylbut-2-enyl diphosphate reductase [Streptomyces sp. CNH099]	bacteria
CR382127|YALI0_A19602g	WP_100097079.1	21	lcl|WP_100097079.1 4-hydroxy-3-methylbut-2-enyl diphosphate reductase [Candidatus Hamiltonella defensa]	bacteria
CR382129|YALI0_C18755g	WP_073501779.1	20	lcl|WP_073501779.1 4-hydroxy-3-methylbut-2-enyl diphosphate reductase [Streptomyces paucisporeus]	bacteria
CR382132|YALI0_F04015g	SFC70576.1	19	lcl|SFC70576.1 4-hydroxy-3-methylbut-2-enyl diphosphate reductase [Flexibacter flexilis DSM 6793]	bacteria
CR382128|YALI0_B21516g	APD09172.1	18	lcl|APD09172.1 hydroxy-3-methylbut-2-enyl diphosphate reductase [Thermus brockianus]	bacteria
CR382128|YALI0_B01034g	WP_014255039.1	18	lcl|WP_014255039.1 bifunctional 4-hydroxy-3-methylbut-2-enyl diphosphate reductase/30S ribosomal protein S1 [[Clostridium] clariflavum]	bacteria
CR382131|YALI0_E26059g	APD09172.1	17	lcl|APD09172.1 hydroxy-3-methylbut-2-enyl diphosphate reductase [Thermus brockianus]	bacteria
CR382132|YALI0_F29205g	APD09172.1	14	lcl|APD09172.1 hydroxy-3-methylbut-2-enyl diphosphate reductase [Thermus brockianus]	bacteria
CR382128|YALI0_B11352g	WP_004813195.1	13	lcl|WP_004813195.1 bifunctional 4-hydroxy-3-methylbut-2-enyl diphosphate reductase/30S ribosomal protein S1 [Anaerococcus hydrogenalis]	bacteria
CR382131|YALI0_E26048g	APD09172.1	10	lcl|APD09172.1 hydroxy-3-methylbut-2-enyl diphosphate reductase [Thermus brockianus]	bacteria

Table S 15 DXR gene candidate based on transcriptomics analysis

Gene name	P-value	Q - value	Sequence
YALI0_C05709g	pvalue=4.22e-05	qvalue= 0.56	sequence=DLAELILSKSNGAFSRALFLNSGSEANETAMKIATQFFYEQ
YALI0_A03949g	pvalue=2.76e-05	qvalue= 0.546	sequence=VNQFFAFRFPSISLPSNILQLLSYPSGKLLEFILPDWGFTF
YALI0_B04202g	pvalue=1.11e-05	qvalue= 0.483	sequence=INVDTAGSDLPMLAGNVVALLSPLMTIPVLTLIFGLDKFDF
YALI0_C09889g	pvalue=1.79e-05	qvalue= 0.517	sequence=ELVAAAMTVNFWNEINGTSINTAAWIIIFWTLIVCINLFGV
YALI0_F27005g	pvalue=3.16e-05	qvalue= 0.546	sequence=KTLTQFARSPQWIIPRIEKILHPGFMKFLSYIPGAVQLTRL
YALI0_D22847g	pvalue=1.17e-05	qvalue= 0.483	sequence=PRLQESLVEETLKSGGMKGARKVFLANSGTEANEAALKFAR
YALI0_D05819g	pvalue=3.08e-05	qvalue= 0.546	sequence=ICVLSLVASNTAGHTKKISVNAILLIGYCLGNLLGPQTFRA
YALI0_D05819g	pvalue=8.35e-05	qvalue= 0.672	sequence=MSLSPSTSDVEKNSVRKPSLDSDAEVEKTPGVMLDTQIIDE
YALI0_E06567g	pvalue=3.72e-05	qvalue= 0.557	sequence=CTADKVRDQFVKVTDMLLNLDVSDQEAVPKAIAEAKELFSV
YALI0_E00704g	pvalue=4.38e-05	qvalue= 0.56	sequence=FGLSSDGAVDYTVLGLKIAILIASQFWATLITKRLRYLPGF
YALI0_F16731g	pvalue=8.18e-05	qvalue= 0.672	sequence=ATPPSSRRSSIDSLDSMDSMGSISSGETLLGSISESETEGR
YALI0_E30789g	pvalue=1.83e-05	qvalue= 0.517	sequence=ERLMIMIRAIIESSINPSTVSSVSLFRAHPMILEGHILTIL
YALI0_E30789g	pvalue=3.45e-05	qvalue= 0.546	sequence=YSQTVPPQTFRSSLEALCGIGDTSSSHYGLRILGTALFFEV
YALI0_C07128g	pvalue=2.03e-05	qvalue= 0.517	sequence=QTSKTVFVPLTVGGGIRDVTDSSGNVTTALQVATQYFMSGA
YALI0_F10857g	pvalue=5.38e-06	qvalue= 0.483	sequence=MTADEVGALFSRTMLKMDDLKPDDKAGLNEVVSDVKELFSV
YALI0_F00484g	pvalue=6.47e-05	qvalue= 0.649	sequence=VTALSGLFAPYWRTDARGTILGLTQFTTKAHICRAALEATC
YALI0_C23991g	pvalue=2.01e-05	qvalue= 0.517	sequence=VNVFETVNTYKFLEAAGITPDSSKTYAKTDFLNALNSNFDG
YALI0_F23023g	pvalue=2.61e-05	qvalue= 0.546	sequence=RRPSILIPPSVPNMNEGLRHQSAELANMALEEGMGTSPSNN
YALI0_E07744g	pvalue=6.96e-05	qvalue= 0.649	sequence=LNVSTLPAFNWTALGANLTVSNFTVNNTSFGYFNFSSLQNV
YALI0_E07744g	pvalue=7.67e-05	qvalue= 0.649	sequence=APLANTSATAASRGPVFLPVPGPVNVTKGTEYNETALEYNN
YALI0_B19008g	pvalue=9.35e-05	qvalue= 0.701	sequence=VFAMIAAGMPNLAGYCVLVAMSAAAAGGNLVLDTTVFLEFL
YALI0_C23859g	pvalue=9.84e-05	qvalue= 0.706	sequence=QLTYGALLSGRVTMIAESHLLSARFLTIALRYACIRRQFGA
YALI0_B12188g	pvalue=5.03e-05	qvalue= 0.615	sequence=SLVQQEPVLFSGTVRENILLGSLRDDVTEEEMIEAAEMANI
YALI0_C22990g	pvalue=3.18e-05	qvalue= 0.546	sequence=QIRSDAELGKNALMILKDSSISAARTYKLLNSLFEQLNRRT
YALI0_D22352g	pvalue=8.24e-07	qvalue= 0.147	sequence=INPDEAVAYGAAVQAAILTGDTSSKTQDLLLLDVAPLSLGI
YALI0_C12122g	pvalue=7.29e-06	qvalue= 0.483	sequence=TSLQQLMAIVKRRVALKGEVDIQELFFMLTMDTATNLLYGE
YALI0_E15378g	pvalue=4.41e-05	qvalue= 0.56	sequence=KRVADVVVDEIVSKGGKAVANYDSVENGDKIVETAVKAFGS

Gene name	P-value	Q - value	Sequence
YALI0_E28153g	pvalue=7.59e-05	qvalue= 0.649	sequence=KRFQGAARQWVQFAMTKLTADMKSLLQSYLVDMEGSRDFEH
YALI0_C16753g	pvalue=5.55e-05	qvalue= 0.632	sequence=IGQSQVLSIDVDKEGRMIASDLAAKIETAKTKGYTPLYINA
YALI0_F27159g	pvalue=8.92e-05	qvalue= 0.687	sequence=VTLLKQLNHPNLIRYNHVWLETSQWSPFGPAVPCAFILQEY
YALI0_F09251g	pvalue=6.91e-05	qvalue= 0.649	sequence=ATENSASVQATMYNMSTDILDKASLVETVDYELPNKHYFEL
YALI0_B08118g	pvalue=3.48e-05	qvalue= 0.546	sequence=QTATAAGATATLTRVRLITSSAPTLAQGDTSTIDFCKLPHT
YALI0_D08184g	pvalue=8.9e-07	qvalue= 0.147	sequence=INPDEAVAYGAAVQAAILSGDTSSKTQDLLLLDVAPLSLGI
YALI0_F10241g	pvalue=8.04e-06	qvalue= 0.483	sequence=ILPTLAVNRALLAPGRKENINGWTISRYLFFLIVTVGAFLY
YALI0_A10637g	pvalue=7.16e-05	qvalue= 0.649	sequence=VVEECKVASACFTRESILTLRKAALANSIFFYTRIMSLGGS
YALI0_E32901g	pvalue=1.9e-05	qvalue= 0.517	sequence=ASQEEAGIEKKFWNGIKVTFKNYWLMFIYLVILMAGFNFMS
YALI0_F02959g	pvalue=6.81e-05	qvalue= 0.649	sequence=MILAAVATFVADPMIRGLTVGWCFTLWALLSIVVAFSLVAI
YALI0_C01199g	pvalue=7.16e-05	qvalue= 0.649	sequence=MTSLDDAIAMAYVSGGTSQLDKDSAVQKAERMAQAKRLWAT
YALI0_B23188g	pvalue=9.1e-06	qvalue= 0.483	sequence=SIVQMPRGIPVATMGVNNSTNAALLAIRILAAYDATLHIRL
YALI0_A09020g	pvalue=7.31e-05	qvalue= 0.649	sequence=PTPEQLAELERKDSQDVLKVSRRRLVTWVSESIFTFLLTAL
YALI0_C08473g	pvalue=6.8e-05	qvalue= 0.649	sequence=VETEVVTVSGTESTVTKNSCITSTPVADPCVVTETISSGTI
YALI0_F09559g	pvalue=3e-05	qvalue= 0.546	sequence=VKLAQDVRACKEARPKKVELLKSMIADPKLELKELPYPVIL
YALI0_F08899g	pvalue=9.76e-05	qvalue= 0.706	sequence=IYPSDFIDIQLPSMTDVITYDYAGKVKSLLGLENLTKSETA
YALI0_D23419g	pvalue=8.95e-05	qvalue= 0.687	sequence=IRLATYAAETTVESARVSTVTSLALSVKAIEGLISKANQDT
YALI0_C18557g	pvalue=5.22e-05	qvalue= 0.616	sequence=ITIASAATSFNWNNHTVNLIDTPGHADFTFEVIRSIRVLDG
YALI0_B05522g	pvalue=4.31e-05	qvalue= 0.56	sequence=AVLSRLAGAPGQNDGSDVSINYQSLASSFLLSADQAHAFHP

Table S 16 ipsD gene candidate based on transcriptomics analysis

Gene name	P-value	Q - value	Sequence
YALI0_C05709g	pvalue=4.22e-05	qvalue= 0.56	sequence=DLAELILSKSNGAFSRALFLNSGSEANETAMKIATQFFYEQ
YALI0_A03949g	pvalue=2.76e-05	qvalue= 0.546	sequence=VNQFFAFRFPSISLPSNILQLLSYPSGKLLEFILPDWGFTF
YALI0_B04202g	pvalue=1.11e-05	qvalue= 0.483	sequence=INVDTAGSDLPMLAGNVVALLSPLMTIPVLTLIFGLDKFDF
YALI0_C09889g	pvalue=1.79e-05	qvalue= 0.517	sequence=ELVAAAMTVNFWNEINGTSINTAAWIIIFWTLIVCINLFGV
YALI0_F27005g	pvalue=3.16e-05	qvalue= 0.546	sequence=KTLTQFARSPQWIIPRIEKILHPGFMKFLSYIPGAVQLTRL
YALI0_D22847g	pvalue=1.17e-05	qvalue= 0.483	sequence=PRLQESLVEETLKSGGMKGARKVFLANSGTEANEAALKFAR
YALI0_D05819g	pvalue=3.08e-05	qvalue= 0.546	sequence=ICVLSLVASNTAGHTKKISVNAILLIGYCLGNLLGPQTFRA
YALI0_D05819g	pvalue=8.35e-05	qvalue= 0.672	sequence=MSLSPSTSDVEKNSVRKPSLDSDAEVEKTPGVMLDTQIIDE
YALI0_E06567g	pvalue=3.72e-05	qvalue= 0.557	sequence=CTADKVRDQFVKVTDMLLNLDVSDQEAVPKAIAEAKELFSV
YALI0_E00704g	pvalue=4.38e-05	qvalue= 0.56	sequence=FGLSSDGAVDYTVLGLKIAILIASQFWATLITKRLRYLPGF
YALI0_F16731g	pvalue=8.18e-05	qvalue= 0.672	sequence=ATPPSSRRSSIDSLDSMDSMGSISSGETLLGSISESETEGR
YALI0_E30789g	pvalue=1.83e-05	qvalue= 0.517	sequence=ERLMIMIRAIIESSINPSTVSSVSLFRAHPMILEGHILTIL
YALI0_E30789g	pvalue=3.45e-05	qvalue= 0.546	sequence=YSQTVPPQTFRSSLEALCGIGDTSSSHYGLRILGTALFFEV
YALI0_C07128g	pvalue=2.03e-05	qvalue= 0.517	sequence=QTSKTVFVPLTVGGGIRDVTDSSGNVTTALQVATQYFMSGA
YALI0_F10857g	pvalue=5.38e-06	qvalue= 0.483	sequence=MTADEVGALFSRTMLKMDDLKPDDKAGLNEVVSDVKELFSV
YALI0_F00484g	pvalue=6.47e-05	qvalue= 0.649	sequence=VTALSGLFAPYWRTDARGTILGLTQFTTKAHICRAALEATC
YALI0_C23991g	pvalue=2.01e-05	qvalue= 0.517	sequence=VNVFETVNTYKFLEAAGITPDSSKTYAKTDFLNALNSNFDG
YALI0_F23023g	pvalue=2.61e-05	qvalue= 0.546	sequence=RRPSILIPPSVPNMNEGLRHQSAELANMALEEGMGTSPSNN
YALI0_E07744g	pvalue=6.96e-05	qvalue= 0.649	sequence=LNVSTLPAFNWTALGANLTVSNFTVNNTSFGYFNFSSLQNV
YALI0_E07744g	pvalue=7.67e-05	qvalue= 0.649	sequence=APLANTSATAASRGPVFLPVPGPVNVTKGTEYNETALEYNN
YALI0_B19008g	pvalue=9.35e-05	qvalue= 0.701	sequence=VFAMIAAGMPNLAGYCVLVAMSAAAAGGNLVLDTTVFLEFL
YALI0_C23859g	pvalue=9.84e-05	qvalue= 0.706	sequence=QLTYGALLSGRVTMIAESHLLSARFLTIALRYACIRRQFGA
YALI0_B12188g	pvalue=5.03e-05	qvalue= 0.615	sequence=SLVQQEPVLFSGTVRENILLGSLRDDVTEEEMIEAAEMANI
YALI0_C22990g	pvalue=3.18e-05	qvalue= 0.546	sequence=QIRSDAELGKNALMILKDSSISAARTYKLLNSLFEQLNRRT
YALI0_D22352g	pvalue=8.24e-07	qvalue= 0.147	sequence=INPDEAVAYGAAVQAAILTGDTSSKTQDLLLLDVAPLSLGI

Gene name	P-value	Q - value	Sequence
YALI0_C12122g	pvalue=7.29e-06	qvalue= 0.483	sequence=TSLQQLMAIVKRRVALKGEVDIQELFFMLTMDTATNLLYGE
YALI0_E15378g	pvalue=4.41e-05	qvalue= 0.56	sequence=KRVADVVVDEIVSKGGKAVANYDSVENGDKIVETAVKAFGS
YALI0_E28153g	pvalue=7.59e-05	qvalue= 0.649	sequence=KRFQGAARQWVQFAMTKLTADMKSLLQSYLVDMEGSRDFEH
YALI0_C16753g	pvalue=5.55e-05	qvalue= 0.632	sequence=IGQSQVLSIDVDKEGRMIASDLAAKIETAKTKGYTPLYINA
YALI0_F27159g	pvalue=8.92e-05	qvalue= 0.687	sequence=VTLLKQLNHPNLIRYNHVWLETSQWSPFGPAVPCAFILQEY
YALI0_F09251g	pvalue=6.91e-05	qvalue= 0.649	sequence=ATENSASVQATMYNMSTDILDKASLVETVDYELPNKHYFEL
YALI0_B08118g	pvalue=3.48e-05	qvalue= 0.546	sequence=QTATAAGATATLTRVRLITSSAPTLAQGDTSTIDFCKLPHT
YALI0_D08184g	pvalue=8.9e-07	qvalue= 0.147	sequence=INPDEAVAYGAAVQAAILSGDTSSKTQDLLLLDVAPLSLGI
YALI0_F10241g	pvalue=8.04e-06	qvalue= 0.483	sequence=ILPTLAVNRALLAPGRKENINGWTISRYLFFLIVTVGAFLY
YALI0_A10637g	pvalue=7.16e-05	qvalue= 0.649	sequence=VVEECKVASACFTRESILTLRKAALANSIFFYTRIMSLGGS
YALI0_E32901g	pvalue=1.9e-05	qvalue= 0.517	sequence=ASQEEAGIEKKFWNGIKVTFKNYWLMFIYLVILMAGFNFMS
YALI0_F02959g	pvalue=6.81e-05	qvalue= 0.649	sequence=MILAAVATFVADPMIRGLTVGWCFTLWALLSIVVAFSLVAI
YALI0_C01199g	pvalue=7.16e-05	qvalue= 0.649	sequence=MTSLDDAIAMAYVSGGTSQLDKDSAVQKAERMAQAKRLWAT
YALI0_B23188g	pvalue=9.1e-06	qvalue= 0.483	sequence=SIVQMPRGIPVATMGVNNSTNAALLAIRILAAYDATLHIRL
YALI0_A09020g	pvalue=7.31e-05	qvalue= 0.649	sequence=PTPEQLAELERKDSQDVLKVSRRRLVTWVSESIFTFLLTAL
YALI0_C08473g	pvalue=6.8e-05	qvalue= 0.649	sequence=VETEVVTVSGTESTVTKNSCITSTPVADPCVVTETISSGTI
YALI0_F09559g	pvalue=3e-05	qvalue= 0.546	sequence=VKLAQDVRACKEARPKKVELLKSMIADPKLELKELPYPVIL
YALI0_F08899g	pvalue=9.76e-05	qvalue= 0.706	sequence=IYPSDFIDIQLPSMTDVITYDYAGKVKSLLGLENLTKSETA
YALI0_D23419g	pvalue=8.95e-05	qvalue= 0.687	sequence=IRLATYAAETTVESARVSTVTSLALSVKAIEGLISKANQDT
YALI0_C18557g	pvalue=5.22e-05	qvalue= 0.616	sequence=ITIASAATSFNWNNHTVNLIDTPGHADFTFEVIRSIRVLDG
YALI0_B05522g	pvalue=4.31e-05	qvalue= 0.56	sequence=AVLSRLAGAPGQNDGSDVSINYQSLASSFLLSADQAHAFHP

Table S 17 ipsE gene candidate based on transcriptomics analysis

Gene name	P-value	Q - value	Sequence
YALI0_B06391g	pvalue=6.09e-05	qvalue= 0.754	sequence=SFGLSNIHGVAGFRVAWGLQIIPGLLMSL
YALI0_B07898g	pvalue=1.2e-05	qvalue= 0.643	sequence=SHPPTMLRLKLFGKEKSSEEEVETEIGEV
YALI0_F07535g	pvalue=6.49e-05	qvalue= 0.754	sequence=APQQANGASAFNVQLAAGVAIVGAVAALI
YALI0_B06248g	pvalue=8.73e-05	qvalue= 0.763	sequence=SFPRSLWHFYVFLRESRKHKLLEGFQRSF
YALI0_A19536g	pvalue=4.71e-06	qvalue= 0.588	sequence=ALAKIIVDNIANGQRGEVIEPLYARLTPL
YALI0_A19536g	pvalue=5.24e-05	qvalue= 0.754	sequence=KHAVVGLHDSITHERGLAGKVGTTLVCPG
YALI0_E14234g	pvalue=5.2e-06	qvalue= 0.588	sequence=SVAKVVQSSPVLQSLSVFYHNVTSNFVNV
YALI0_E17655g	pvalue=1.74e-05	qvalue= 0.643	sequence=LLPKINIYKDVFVKRPAHSEYEEPVECLL
YALI0_E34771g	pvalue=2.36e-05	qvalue= 0.695	sequence=TYANAPVGLQIVTPRWNDELALKMMEVFS
YALI0_A01716g	pvalue=8.71e-05	qvalue= 0.763	sequence=VVSKANTRKDVIKVYDSYCGKLDKVVTFL
YALI0_A19096g	pvalue=6.8e-05	qvalue= 0.754	sequence=RYSALTFNGHKIHYDPDYAREVESLPAVI
YALI0_D10043g	pvalue=5e-05	qvalue= 0.754	sequence=VPAYGYAYFATEIVKSLGYSPIQTQFRSV
YALI0_E07744g	pvalue=6.37e-05	qvalue= 0.754	sequence=SRALANFHHYSKQVEKKKLSPVDASLLAI
YALI0_B19008g	pvalue=5.23e-05	qvalue= 0.754	sequence=AAAGGNLVLDTTVFLEFLPSSQQWLLTLL
YALI0_A07623g	pvalue=8.15e-05	qvalue= 0.763	sequence=SYLDGSISDQIHGPDSNRVKVLKVVYASA
YALI0_D02739g	pvalue=9.35e-05	qvalue= 0.763	sequence=VWAWAFFAFHVLKCLTLYLVLWSGTMNPL
YALI0_E19536g	pvalue=6.27e-05	qvalue= 0.754	sequence=SLQLINVVFEVAGKELTKHHSLLQITTDT
YALI0_C18645g	pvalue=1.79e-05	qvalue= 0.643	sequence=VNEITPSTLKFTPFEAAKLGLLQIMALSH
YALI0_E33957g	pvalue=6.57e-05	qvalue= 0.754	sequence=HSGWSLLVECIGDVHVDDHHTVEDTAIAL
YALI0_C23969g	pvalue=1.89e-05	qvalue= 0.643	sequence=SIAVENVDDALKAAESVGYPIIVRSAYAL
YALI0_E15378g	pvalue=4.78e-05	qvalue= 0.754	sequence=AKAGILGFSRALAREGEKYNILVNTIAPN
YALI0_E15378g	pvalue=9.28e-05	qvalue= 0.763	sequence=AKLALVGFGETLAKEGAKYNITSNVIAPL
YALI0_F31845g	pvalue=1.79e-05	qvalue= 0.643	sequence=SAASAIPALNLFGRLLEPFINIPYIIQFI
YALI0_F31845g	pvalue=5.86e-05	qvalue= 0.754	sequence=AIPALNLFGRLLEPFINIPYIIQFIIQAW
YALI0_E07601g	pvalue=8.85e-05	qvalue= 0.763	sequence=APDHISLHQKTDSMSSNDLDTLEQLVQPD
YALI0_E18568g	pvalue=9.22e-05	qvalue= 0.763	sequence=VAAKYNVSRKAQDAFAAKSYEKAAAAQAA
YALI0_C00649g	pvalue=9.9e-06	qvalue= 0.643	sequence=KKKKEGLDIDIVLRPQDGHKLLQGKTLAA

Gene name	P-value	Q - value	Sequence
YALI0_C10780g	pvalue=5.47e-07	qvalue= 0.186	sequence=GPPYANGALHVGHVLNKVTKDIINRYQLI
YALI0_B17160g	pvalue=3.46e-05	qvalue= 0.754	sequence=DNATENAFRDLNAKVKSKNDEVQRLEMTI
YALI0_E14146g	pvalue=7.11e-05	qvalue= 0.754	sequence=SYGGKYQYLTILGLCGSYLAFLLALLADA
YALI0_E19679g	pvalue=9.61e-05	qvalue= 0.763	sequence=ALPQKTITLDELVKFPRDLETKERNEILI
YALI0_D15114g	pvalue=6.9e-05	qvalue= 0.754	sequence=IDAAVVLGSAMLGARAAHKKDLWEIMGEL
YALI0_B07535g	pvalue=5.84e-05	qvalue= 0.754	sequence=SPAVQGRALVLLGCLAKTEVEDDMVRRVV
YALI0_C23815g	pvalue=3.66e-05	qvalue= 0.754	sequence=VIAKVAIDQAVFAPSSIGYYFSVMGLLEG
YALI0_F13321g	pvalue=9.28e-05	qvalue= 0.763	sequence=DNKRLNEDLDRLKRLMDDEHVLEESMPSL
YALI0_A19888g	pvalue=5.69e-05	qvalue= 0.754	sequence=CDPNTTIFDRDTQPEFRYYRDNVYFYQLL
YALI0_A19580g	pvalue=1.43e-05	qvalue= 0.643	sequence=AISRANVFYYLMFGLVAGLSEIISNLIFL
YALI0_A19624g	pvalue=5.28e-05	qvalue= 0.754	sequence=IDAEVVLGLTKEPYFVDKYKVIKRLLASH
YALI0_F27093g	pvalue=7.58e-05	qvalue= 0.763	sequence=SVVKIPSASKTRRARPNTWLNLKSLPNTI
YALI0_E34815g	pvalue=2.66e-05	qvalue= 0.695	sequence=VIALDLLSFQPTKVEKDDFQDVETPDVTL
YalifMp18	pvalue=2.61e-05	qvalue= 0.695	sequence=TPGRLNQVSALLQRESVYYGQCSELCGVL
YALI0_B06754g	pvalue=9.67e-05	qvalue= 0.763	sequence=SDAFTKLQNELRAKEVERKKELQRLYADP
YALI0_F12661g	pvalue=4.36e-05	qvalue= 0.754	sequence=VHPTGRLALSVGQDRKLILWNLMTAKKAS

Table S 18 ipsF gene candidate based on transcriptomics analysis

Gene name	P-value	Q - value	Sequence
YALI0_A00374g	pvalue=2.3e-05	qvalue= 0.294	sequence=DSSADASASASASASSSGASASADSSATSASAD
YALI0_A00374g	pvalue=7.01e-05	qvalue= 0.384	sequence=DSSATSASADATSAASGSAAASGSAAASGSTAA
YALI0_E22396g	pvalue=6.09e-05	qvalue= 0.384	sequence=VQAVDGLLKSVSDLLQLLGGLPGLTDLAALTGA
YALI0_E35156g	pvalue=1.99e-05	qvalue= 0.291	sequence=TCDGPSLGGFVCAAVVAEAELWKIGQVKPGDTV
YALI0_F04169g	pvalue=7.76e-05	qvalue= 0.384	sequence=GFRLEDLIEQEPDAALGNGGLGRLAACFVDSLA
YALI0_C09889g	pvalue=5.17e-05	qvalue= 0.384	sequence=LSVGNAGVYAFSRTIAGMAEAGQAPKMFAYIDR
YALI0_D01331g	pvalue=2.09e-05	qvalue= 0.291	sequence=VYDIGADKIAVSQAKFGSGSAEDLADIQIEDKG
YALI0_E07271g	pvalue=6.98e-05	qvalue= 0.384	sequence=HGDPAWITVASTEHINKQWKELQLMKKNPESLP
YALI0_E07271g	pvalue=8e-05	qvalue= 0.384	sequence=TCDGPSLGGFVCAAVVVEAELWKIGQVKPGDTV
YALI0_D00495g	pvalue=8.28e-05	qvalue= 0.384	sequence=WGNSILALAIGCLCMIDAAAAAALFSLSAGGNA
YALI0_B19800g	pvalue=4.45e-05	qvalue= 0.355	sequence=LSVGNSAVYGFSRTIAGLAAMGHAPKWFAYIDR
YALI0_D02046g	pvalue=8.02e-05	qvalue= 0.384	sequence=VDNTDAHMSTAAIPLLSALTLGIFVSYTSSLLY
YALI0_C06105g	pvalue=8.73e-05	qvalue= 0.384	sequence=GGVGQVIGDSISGAIYRTLYPKYIHKYAPDIPA
YALI0_E14234g	pvalue=6.02e-05	qvalue= 0.384	sequence=LPDAAEAAAADAESDLLTKVASRIAEETPSGAG
YALI0_E14234g	pvalue=9.76e-05	qvalue= 0.384	sequence=KIDSGLAGVSLTYAILFSENVLWAVRLYAVNEM
YALI0_B16214g	pvalue=1.05e-05	qvalue= 0.291	sequence=VADHKVLVEAACGTALAPIYNGTLKDIVPGLGP
YALI0_C07128g	pvalue=8.08e-05	qvalue= 0.384	sequence=LELIDHVKSAVKIPVIASSGAGNPGHFEEVFRE
YALI0_C15488g	pvalue=2.03e-06	qvalue= 0.162	sequence=GFSPDRSTVTNCVANLGAIAGGVIIGHFSSVLG
YALI0_C04730g	pvalue=9.41e-05	qvalue= 0.384	sequence=INAAMVPVVYFFYPETAGRSLEEIDQIFADSNP
YALI0_E07051g	pvalue=6.73e-05	qvalue= 0.384	sequence=RICWIVGPKSYIKALASAGSFLDGGSNAPFQKA
YALI0_D21890g	pvalue=3.82e-05	qvalue= 0.349	sequence=KIHGRVIDTASHLIDKAASALGLSGHPSKDSFD
YALI0_C23379g	pvalue=9.19e-05	qvalue= 0.384	sequence=SSNGVLSTSEGTDADFAFLISGLLNLYDVTGDR
YALI0_A01716g	pvalue=9.02e-05	qvalue= 0.384	sequence=WSEASVRIDTNLRSLIASAGGVVVDEPNKDTFN
YALI0_E07744g	pvalue=3.71e-05	qvalue= 0.349	sequence=LSPVDASLLAIDSAALSESGSGSYTYLLQASQP
YALI0_C23859g	pvalue=1.39e-05	qvalue= 0.291	sequence=SFGALLKLNPVKDAEKFAVATADLKALFASSAG
YALI0_A00396g	pvalue=1.73e-05	qvalue= 0.291	sequence=QHGGYGAAGAAAGAAAGAAAGYGASQYNRPPAG
YALI0_C22616g	pvalue=9.25e-05	qvalue= 0.384	sequence=VWETDLPIWALIIALLIAGGLLIPAGMIAALTN

Gene name	P-value	Q - value	Sequence
YALI0_C05170g	pvalue=4.61e-06	qvalue= 0.291	sequence=HNRLDIALSTNLPLALVFLPAADIPRFVGEGRV
YALI0_D16423g	pvalue=9.76e-05	qvalue= 0.384	sequence=HDAAAAAEKVKSDALKAAHVVGEKAQEAKESAS
YALI0_A00594g	pvalue=1.55e-05	qvalue= 0.291	sequence=DIDLDALLNRIDDLALGDLAAALDGADAAKAKA
YALI0_A01441g	pvalue=4.11e-05	qvalue= 0.354	sequence=WHYLDIPPTSPSPALLGSGKTLMLCHGFPDSWY
YALI0_C23969g	pvalue=4.67e-05	qvalue= 0.356	sequence=ASNGKMLVHAISEHIENAGIHSGDASLVLPPQD
YALI0_A18183g	pvalue=2.86e-05	qvalue= 0.305	sequence=GSFAWAGSIFSSSSVTLSIAQGIVGMYALDHDG
YALI0_B10626g	pvalue=1.42e-05	qvalue= 0.291	sequence=TPETVAPTSAPSSASISSAAGVTNGTVFGSSEP
YALI0_F31845g	pvalue=9.33e-05	qvalue= 0.384	sequence=SQNGLFILPGAIAILFGAPGLYNLSKLPGKRPA
YALI0_E13057g	pvalue=1.73e-05	qvalue= 0.291	sequence=PRMATLLGFFATDAPVSASALGVALKHAADRSF
YALI0_C16753g	pvalue=8.96e-05	qvalue= 0.384	sequence=GSLASEASANSDATRFMAKKLYDSGKFLVDYAP
YALI0_F28303g	pvalue=3.2e-05	qvalue= 0.32	sequence=PIPDDSLLFSVCIVGFHHSLGPQIEEWFGPEDY
YALI0_F09669g	pvalue=2.48e-05	qvalue= 0.294	sequence=ASQADVAVYKAVASAPDAEKYPNAARWFKHIAA
YALI0_C16148g	pvalue=1.08e-06	qvalue= 0.115	sequence=PTDATGALLQGIAGLLGGGALGNLGNLVGGSSI
YALI0_D22396g	pvalue=8.46e-05	qvalue= 0.384	sequence=SENAASGTSSVAGAPSSAAGPALSASPAPSSSA
YALI0_B08118g	pvalue=6.91e-05	qvalue= 0.384	sequence=STPNPVAAPAQTATAAGATATLTRVRLITSSAP
YALI0_D20196g	pvalue=2.31e-05	qvalue= 0.294	sequence=RSFGGAVGLAVCSAILANSLKADLKTKNLPSEL
YALI0_C14938g	pvalue=9.82e-07	qvalue= 0.115	sequence=KDTIDVAGASIPDFVFAAAKKTDTDPVFGVGYP
YALI0_B03080g	pvalue=1.68e-05	qvalue= 0.291	sequence=HKVADVVSFYSQGSTLMPGDLIFTGTPFGVAAG
YALI0_B00814g	pvalue=2.99e-05	qvalue= 0.309	sequence=HGHGHTLGTAAADAHNLASSTHNLSSTAPSTAQ
YALI0_F06776g	pvalue=3.67e-05	qvalue= 0.349	sequence=HTPALVVISVVIFNAAFGYSWGPIPWLYPPEIM
YALI0_E32901g	pvalue=5.57e-07	qvalue= 0.115	sequence=HFSADRSTVTNCVANFGAIAGGMLIGHFSTALG
YALI0_C04136g	pvalue=7.52e-06	qvalue= 0.291	sequence=IPTGDVSSATPTGDVSSATPTGDVSSVTPTGDV
YALI0_C04136g	pvalue=1.06e-05	qvalue= 0.291	sequence=IPTGDVSSATPTDDVSSSIPTGDVSSATPTGDA
YALI0_C04136g	pvalue=1.28e-05	qvalue= 0.291	sequence=TPTGDVSSATPTDDASSSIPTGDVSSATPTGDA
YALI0_C04136g	pvalue=2.43e-05	qvalue= 0.294	sequence=TPTGDVSSATPTGDVSSATPTGDASSSIPTGDV
YALI0_C04136g	pvalue=2.69e-05	qvalue= 0.3	sequence=SSSAVVVPSSSVPTIVSSTATTDIVSPSSSTAD
YALI0_C04136g	pvalue=8.85e-05	qvalue= 0.384	sequence=TPTGDASSSIPTGDVSSATPTGDVSSSTPTGDA
YALI0_C04136g	pvalue=8.85e-05	qvalue= 0.384	sequence=TPTGDASSSIPTGDVSSATPTGDVSSSTPTGDA
YALI0_C04136g	pvalue=9.15e-05	qvalue= 0.384	sequence=TPTGDVSSSTPTGDASSSIPTGDVSSATPTGDA

Gene name	P-value	Q - value	Sequence
YALI0_C04136g	pvalue=9.15e-05	qvalue= 0.384	sequence=IPTGDVSSATPTGDASSSIPTGDVSSATPTGDA
YALI0_C04136g	pvalue=9.15e-05	qvalue= 0.384	sequence=TPTGDVSSATPTGDASSSIPTGDVSSATPTGDA
YALI0_C04136g	pvalue=9.15e-05	qvalue= 0.384	sequence=IPTGDVSSATPTGDASSSIPTGDVSSATPTGDA
YALI0_A15950g	pvalue=9.53e-05	qvalue= 0.384	sequence=PALEDVATLASTFAKQLLPSGGAPVVYVNVEDP
YALI0_E11429g	pvalue=7.17e-05	qvalue= 0.384	sequence=APFGVELQATYIAKAIRKAQLEDIKSLVPSVRA
YALI0_F02959g	pvalue=4.2e-05	qvalue= 0.354	sequence=WFLTGVVSLAAMAAPLGTSIMLPGIDDVAEGLN
YALI0_C00407g	pvalue=6.5e-05	qvalue= 0.384	sequence=KPAGDAVEAAVKESVEAGITTADIGGSSSTSEV
YALI0_F05874g	pvalue=2.72e-05	qvalue= 0.3	sequence=FSSGSAATATILQSLAQGSHVVSIGDVYGGTHR
YALI0_D21802g	pvalue=9.92e-05	qvalue= 0.384	sequence=ESDLTAISDAAESAKIAEATRIAEARAAADTKA
YALI0_B22418g	pvalue=8.21e-05	qvalue= 0.384	sequence=GEDATLLLEEEAAALQAQVEQNNVPGAFPRGNN
YALI0_A03135g	pvalue=1.83e-05	qvalue= 0.291	sequence=WGNALLALALGSICMIDATAAAAIFSLSAGGNY
YALI0_D21593g	pvalue=6.47e-06	qvalue= 0.291	sequence=HFTGTSLVVAIIIYALSTGKYKFLLGLGPAGYG
YALI0_D22286g	pvalue=1.81e-05	qvalue= 0.291	sequence=HLCSDKALNMLFRQLQTSVGQRTIGMFAGDGEG
YALI0_B14014g	pvalue=4.52e-05	qvalue= 0.355	sequence=GTAGPTLARRLADAWISGKKLKVLLLESGPSSE
YALI0_B07535g	pvalue=7.48e-05	qvalue= 0.384	sequence=VSGMSVTDSASVASMGSVGSRLSAGSRVPDDNF
YALI0_B11066g	pvalue=2e-05	qvalue= 0.291	sequence=IAERKAAREATAKAKESGVGGGDIASESSDTMD
YALI0_B12210g	pvalue=9.76e-05	qvalue= 0.384	sequence=SQWPELVELGISDCLLSGKGSVALGEALRDGKP
YALI0_F08899g	pvalue=8.71e-05	qvalue= 0.384	sequence=CEDKAKAHVISAVDVINTVGAEKIGSEFMANKV
YALI0_C18557g	pvalue=1.18e-05	qvalue= 0.291	sequence=ILDGVAGVEAQTEKVWKQASEMGIPKIAFVNKM
YALI0_C18557g	pvalue=8.68e-05	qvalue= 0.384	sequence=LSTANATLMEPIMNVYVSVNEADIGLVVKDLSG
YALI0_F05390g	pvalue=9.97e-05	qvalue= 0.384	sequence=YNDIVVGIELVNEPLGPAIGMEVIEKYFQEGFW
YALI0_D00132g	pvalue=5.91e-05	qvalue= 0.384	sequence=HSHARVGLIALGIYLFSAVYSCGEGPVPFTYSA
YALI0_B22528g	pvalue=4.11e-05	qvalue= 0.354	sequence=QSNRGAGSSRNDKSIRGGRGGRGGGSRNNSSRG
YALI0_B05676g	pvalue=4.55e-05	qvalue= 0.355	sequence=KSGGTAAASAATAAASSAGNSSETASSRTPQKP
YALI0_E08954g	pvalue=1.74e-05	qvalue= 0.291	sequence=SRNAFVALTYLLPHFDDALGLLDSAEFKKLSDT
YALI0_C18293g	pvalue=9.09e-05	qvalue= 0.384	sequence=EPTASATTQTGSESKPSPTDSGDITSYVPSVTP

Table S 19 ipsG gene candidate based on transcriptomics analysis

Gene name	P-value	Q - value	Sequence
YALI0_E27203g	pvalue=4.5e-05	qvalue= 0.492	sequence=IAGLVGITPAAGFVPIWSAVPIGIITAVFANISGDLKNLLRIDDGLDVFS
YALI0_E10219g	pvalue=4.2e-05	qvalue= 0.492	sequence=SPFVIAIRVAGVKGLPSVMNVVIMISVLSVGNAAVYGFSRTIAAMGESGQ
YALI0_E22396g	pvalue=5.92e-05	qvalue= 0.492	sequence=TEIQVLTTTFNSLNAGVDKLLADLGKIPLVGSLVTIIAKIAANPVIHGLL
YALI0_C15004g	pvalue=8.76e-05	qvalue= 0.54	sequence=IFLAYAKIIPSLAAANIVTNAGALALKMGMKIPFVNSRDLSRPPTHEELL
YALI0_E35156g	pvalue=3.78e-05	qvalue= 0.492	sequence=YRIHRLIEEVQQSIKGIVEMSRGVRSVLIEFHASASRSTLMKALVDFEKR
YALI0_B07898g	pvalue=7.35e-05	qvalue= 0.531	sequence=FGLSIALVEHWPTDTPVWVIVLCLGMCLVFLLPFTIFYSISGVLLTLNVL
YALI0_C09889g	pvalue=3.39e-05	qvalue= 0.492	sequence=SPFVIAIRLAGVKGLPSVMNVVIMIAVLSVGNAGVYAFSRTIAGMAEAGQ
YALI0_F12925g	pvalue=3.05e-05	qvalue= 0.492	sequence=IVGLVGVTPAAGFVPAWSAVVIGIVSGVAANLGIGIKGVVYIDDGLDVFA
YALI0_F18964g	pvalue=5.78e-05	qvalue= 0.492	sequence=IVLILCVTLVFLFPFTIVMSATGAQLTLNVILELIMGYTIAGRPIGINVA
YALI0_E07271g	pvalue=7.82e-06	qvalue= 0.367	sequence=YRIHRLIEEAQQSIKGIVEMSRGVRSVLIEFHPSASRSTLMQALVDFEKR
YALI0_E18139g	pvalue=5.44e-06	qvalue= 0.367	sequence=SPFHLAEECSNASVAAPRAIVMTSGVGGLMGWAFQIAIAYTVRDVAGVTQ
YALI0_E18139g	pvalue=2.66e-05	qvalue= 0.492	sequence=GQLLLLLMFAGDTAIGAIFSVGAISGYVAFTMPIGIKVFWSSDKFKPGPW
YALI0_E14729g	pvalue=9.33e-05	qvalue= 0.552	sequence=ARHTRDVTMATLGLSHTINTKVGSDLVRGVSGGERKRVSIAESVVCGAPL
YALI0_F12089g	pvalue=1.74e-05	qvalue= 0.457	sequence=NKPLDPDSESGLITKGIMVVTAMLALSLVQIVSSNKFIIGCFNMVLEVTG
YALI0_F17446g	pvalue=9.48e-06	qvalue= 0.373	sequence=FLLAQGICTSLGNSMLYYASLLGVSTWFKAKRGFTLGIAVAGSSIGGLTM
YALI0_D19558g	pvalue=1.3e-05	qvalue= 0.416	sequence=FASITGLEYAYTKAPTSMKSFIMSIFLFTNAIGAAIGIALSPTAKDPKLI
YALI0_F21197g	pvalue=5.16e-06	qvalue= 0.367	sequence=NPFFLIEDDPTPVQQSQVDRAATLTTSSLRFIRALRREELPPDNVRGTKL
YALI0_C01067g	pvalue=5.88e-05	qvalue= 0.492	sequence=MALKKSVSAAKSPTHSIEKKVAQESVSISLKTGFTYSFKYADCVTEIISI
YALI0_F16731g	pvalue=2.89e-06	qvalue= 0.367	sequence=DTLADSFATPPSSRRSSIDSLDSMDSMGSISSGETLLGSISESETEGRPV
YALI0_D15422g	pvalue=4.79e-05	qvalue= 0.492	sequence=IALGGCIGTGLFIGSGTILSSSGPGSMLCAYIVMSFVIWTVMNALGEMTT
YALI0_C04455g	pvalue=6.01e-05	qvalue= 0.492	sequence=NAPVLVGGTLGAAVASDPEDPEAAERDLEGGLQPEETVHVSDEAEKEITV
YALI0_F19866g	pvalue=5.82e-05	qvalue= 0.492	sequence=SPFVIAIKNAGVKGLPSVMNVVIMVAVLSVGNSAIYGCSRTVAALAAQGM
YALI0_A02607g	pvalue=5.84e-05	qvalue= 0.492	sequence=NRRCDALVAAAAMIQAIEETAVSLGGLGTVGIINSLPQSTNTIPGEVNFS
YALI0_F00484g	pvalue=5.73e-05	qvalue= 0.492	sequence=QADILGIEVERPAMRETTALGAAIAAGFAVGVWKSIEDLKDINTEGMTEF
YALI0_E07205g	pvalue=4.76e-05	qvalue= 0.492	sequence=KYTVYGVFAACIISHAIVASLASDGMSKLQTGCIVLNIAIIIIAIIALPI
YALI0_E07744g	pvalue=7.71e-05	qvalue= 0.538	sequence=SSDASPATFATARKAALVARALGFDTLLYEHSNAWNNIWNDGDIVVPGDE
YALI0_C08393g	pvalue=5.48e-05	qvalue= 0.492	sequence=GPSVTPLTYYNPAWRYYEVDTKTFEVMNSHTYYAKLNESFAFDFEHPVEK

Gene name	P-value	Q - value	Sequence
YALI0_A07623g	pvalue=6.25e-05	qvalue= 0.492	sequence=QKLALCVTFSLGAITIIVCIARAVAIAVSGNIAQVAILTAFECSTAMLVA
YALI0_C04026g	pvalue=8.58e-05	qvalue= 0.539	sequence=VLFQAGMSSAGRSFAAKHAETVFVSAPSPQLVRQTVDALRAAAGDRQIFV
YALI0_A08734g	pvalue=5.89e-05	qvalue= 0.492	sequence=GAAVAALGHGNERMIAKIQQAAGLPYILSHGLGSKHADDLAKYLVEQSPE
YALI0_B10626g	pvalue=7.43e-05	qvalue= 0.531	sequence=APSSASISSAAGVTNGTVFGSSEPPTPTVTNNTNSSVESPATALVNSSLT
YALI0_F20174g	pvalue=8.18e-05	qvalue= 0.539	sequence=HLLNQDLRFFDTTTTGKLTSSLSKDTQNVQGLGGATFGQILSSIVTVIIS
YALI0_E28153g	pvalue=6.24e-05	qvalue= 0.492	sequence=RALATSLVIVGDASESSVLDSIGGLSTEISNPGNNLLNQLNEYAIREIHS
YALI0_D17182g	pvalue=5.96e-05	qvalue= 0.492	sequence=NQWGLVYGGGTTGIMGSVARAVATRGGYVNGIIPEALIHKEQNKINPPTV
YALI0_D27170g	pvalue=1.55e-05	qvalue= 0.444	sequence=QLLSNELPDGHPISDYVIMNAAALAVIDGHAKDWKHGVELARESIKSGKA
YALI0_A03179g	pvalue=6.62e-06	qvalue= 0.367	sequence=CRCRNCLSEFKYSNNVSHKSSHERSPLLNEDSGIISFNSPSSELECPVLR
YALI0_C04048g	pvalue=4.17e-05	qvalue= 0.492	sequence=YIDTMCIFSVGDRTYVVRLKSSTAKCLSAQMSGDTLYCIISDHQWTVVRV
YALI0_C16148g	pvalue=8.38e-05	qvalue= 0.539	sequence=QGPTSTLTPSSTPPAGNIAGTRSGLIVLQNTNTASSSASSRPPAPSTVTV
YALI0_B02024g	pvalue=7.87e-05	qvalue= 0.538	sequence=QPRHIVYVYGTSFKNGNELPAAGCGVSFGKIAENNLSYGLEDDMNPEQAL
YALI0_F07194g	pvalue=2.64e-05	qvalue= 0.492	sequence=RRTTIGHCLAVALSHQTVTLALGLLAKSQHVGDVSTIISAAANIVSIGVA
YALI0_C04136g	pvalue=3.45e-05	qvalue= 0.492	sequence=VPTATSATGTVSVESQITSSSGGGGGLASEDIPGSATATGTATVTGLVPD
YALI0_C04136g	pvalue=6.65e-05	qvalue= 0.51	sequence=STSSDGITSATSESSSTITATISDSGSTTQLSSSGISSDISGSVSQTSES
YALI0_D25828g	pvalue=7.36e-05	qvalue= 0.531	sequence=VPFHCTPDTGGYDAMGNCIFTNGTQVLEAYGFKNHVRVYLGVAAVCLVLY
YALI0_C07436g	pvalue=3.21e-05	qvalue= 0.492	sequence=TPPTKGITKPKAPSKSTAKKQAAISSFFQKPDPEAAVVSTCVSDVHTKPL
YALI0_A14861g	pvalue=9.47e-05	qvalue= 0.552	sequence=SWVVDHLDFASFKVASRTWVGIWFGLMLMMIPGYVKWLGRAPYLTAIIGV
YALI0_D22286g	pvalue=9.91e-05	qvalue= 0.567	sequence=TGQRKGSVVSYSSTSGRSLSPVIDSLDSAQSTGSVSINSLDPAQGVLVTE
YALI0_C19844g	pvalue=8.99e-05	qvalue= 0.544	sequence=EPLTIELVQKCPLGSGTQSTSIWSTQVLSPGTHMPDMVSSTLRDVCIPIQ
YALI0_B18568g	pvalue=4.01e-05	qvalue= 0.492	sequence=ANVKSFTTDAGVASAVSSGSARVSKLFSDLGIFATPTGDTSATPTSAVPT
YALI0_A08360g	pvalue=1.32e-05	qvalue= 0.416	sequence=ILSADYQTDNGILSNSGTSMATPHVAGLAAQLMATSSELLTPAQVEEQII
YALI0_B19338g	pvalue=8.47e-05	qvalue= 0.539	sequence=SPFVIAIINSGTKVLPHIFNAVILITLISAGNSNVYIGSRVVYALADSGT
YALI0_C22231g	pvalue=4.77e-06	qvalue= 0.367	sequence=QLIHRTKVASVPATHNPIRESLSGAFKLYHNPVANPRRSLGNEDKSPHRV
YALI0_D13882g	pvalue=6.05e-05	qvalue= 0.492	sequence=KIIYLRAMGGEVGASSALAPKIGPLGLSPKKIGEDIAKATKAHKGVRVTV
YALI0_E22627g	pvalue=8.16e-06	qvalue= 0.367	sequence=FGTGLGVLELHQADTSLVNERIAQATLAALDKGATVICLGCAGMAGMEDV
YALI0_B17028g	pvalue=5.57e-05	qvalue= 0.492	sequence=DGLPTLMTTMPPKSTSTVLGPIVGTEVTLDGNGDPVKIVTATPTGPVTTQ
YALI0_F13673g	pvalue=4.94e-05	qvalue= 0.492	sequence=IPPSIYVPETLPLTLNKNSSSNTSSPQLDCNIEDPFAIPQHTPYNPDRTQ

Table S 20 ipsH gene candidate based on transcriptomics analysis

Gene name	P-value	Q - value	Sequence
YALI0_E20427g	pvalue=9.54e-05	qvalue= 0.708	sequence=RWLMKIGKEEEAVETLCKIRKLPADDLYVQGEILMVREQIEREKQAMTGA
YALI0_C11165g	pvalue=3.77e-05	qvalue= 0.603	sequence=VSRRTVFNMYMAQTIARHAENALNAKIVINDPNVTVTVVISPAPASTAPA
YALI0_D18876g	pvalue=4.89e-05	qvalue= 0.603	sequence=SGRSLESIDEMFSLPWYVIGRRGHKMIPETSATVHLRQEEDSESKKGGVV
YALI0_F24937g	pvalue=7.15e-05	qvalue= 0.603	sequence=HAPNVKKIVITSSLAAIMDVSKYNDPTHTLTEKDWNPMTREQAAEKGNPF
YALI0_E07271g	pvalue=5.92e-05	qvalue= 0.603	sequence=NPALFDDAVERLRQQGVKVVKIDFTPLFDLAKCLYEGPWVAERYAAIKDF
YALI0_A15147g	pvalue=7.01e-05	qvalue= 0.603	sequence=SPVGNRADTQEALSFFTRGLVHSPIHVVGLSELQKVFTLMEEGKIAGRYV
YALI0_B21868g	pvalue=7.11e-05	qvalue= 0.603	sequence=DKDDSVANVKRAIEKLKEIIVEAPKYASVYNNLAQLERILDQKGEKSASK
YALI0_E28548g	pvalue=9.41e-05	qvalue= 0.708	sequence=NSREPRSAFDREIERVDLDLREWNDKVVHMWIEAKIELKMRALEQRRIMK
YALI0_F08129g	pvalue=4.35e-05	qvalue= 0.603	sequence=GVDEFKAGDRVYSRIGERQDGTMAEFVSVKSSLVARAPSNISLQDAAGIP
YALI0_E20625g	pvalue=3.4e-05	qvalue= 0.603	sequence=FVLDPVIATGGTAKAVVEILKEWGVEHIVIVSLIASKQGIEKLSKLEGVE
YALI0_D00176g	pvalue=7.13e-05	qvalue= 0.603	sequence=KPRDLMGYEEMGLEAAVKALLDAGITYDKVEQCVAGYVYGDSTSGQRVVY
YALI0_F06468g	pvalue=2.62e-05	qvalue= 0.603	sequence=EQLIKKGGLKVQQKEAEEKAASLYAALDAHKKLYNLPVAVADRSKMNIVF
YALI0_E34771g	pvalue=2.38e-05	qvalue= 0.603	sequence=GPFHGFPLSIKELYALEGRAAHWENPDLINNMQTDNAVLVNGLLELGCVP
YALI0_C20251g	pvalue=9.2e-05	qvalue= 0.708	sequence=HAPNVKKIVITSSVASNLDVATFYDPNTTYTEKSWNPMTWEQAAEKGNPF
YALI0_B08536g	pvalue=1.23e-05	qvalue= 0.603	sequence=SSLSGQTYVDLGAHAVKAALAKTDIKPDQVDEIIFGNVLSAGVGQAPARQ
YALI0_E07744g	pvalue=1.7e-05	qvalue= 0.603	sequence=NLQGLPAAIDSTVSPALLALYPQFAKNVVNYRVGLANKVSSKSKRSLVEP
YALI0_A20130g	pvalue=2.99e-05	qvalue= 0.603	sequence=GEYFDIQKLFFAQSMDTSTDFLLGESVGCLKELLETRDGDESSKKFGANF
YALI0_A21076g	pvalue=2.15e-05	qvalue= 0.603	sequence=STLKMLESVRRGIEKTYHSRVLMRSLVAILAALGHDREALASFEVFMDYE
YALI0_B10626g	pvalue=6.2e-05	qvalue= 0.603	sequence=EPAVSTSSSEPAVSTPSSELVVSTPTIESEVPVSTSEPVVPASTSESNAP
YALI0_B22814g	pvalue=5.55e-05	qvalue= 0.603	sequence=WMISFVMPHKRPRRPLLLRLGVLAGTIYLTIILAHSTQLINEQYNKGYLD
YALI0_E28153g	pvalue=1.27e-05	qvalue= 0.603	sequence=SDLRYSFYSRVASLTEKLDNKSYVAPVEVKNLLFAGASYLSRVGHKGVDA
YALI0_E13057g	pvalue=5.75e-05	qvalue= 0.603	sequence=NAVTGEGGIHDAQQMVKKADQVFGSESDLPASLVMSTGVIGQRLKIDKIL
YALI0_C08965g	pvalue=6.49e-05	qvalue= 0.603	sequence=APRYYYVAIIVCVVFLVVLLAVYFTRHFVQKRRLRNQSARRDLEANQMVQ
YALI0_B02662g	pvalue=1.5e-05	qvalue= 0.603	sequence=APRTEPSRDGRGSHDLWRPRHAPGAHAYHPGSCRHNHGLSDEPKRRGDDF
YALI0_B17666g	pvalue=5.03e-05	qvalue= 0.603	sequence=VCASVAAPQIAAPATKFVALKQSIRPISTRSTVVQLLNNIGSKREVEQYL
YALI0_D12474g	pvalue=9.72e-06	qvalue= 0.603	sequence=EEDGVAPATPTSNETVENSLLVDNTSIDHIVETVQLPDLAAELKRARRQV

Gene name	P-value	Q - value	Sequence
YALI0_E07601g	pvalue=4.15e-05	qvalue= 0.603	sequence=ASSLGVARKKDTPVVVLSVSDRATSKVLWRCSKEYNAFVALDNAIAPQIP
YALI0_E10659g	pvalue=9.24e-05	qvalue= 0.708	sequence=DPSQITIGGESAGSIGLHALMVHESMKPKEECIIHNVILSSGTMDRMGTG
YALI0_D15444g	pvalue=6.27e-05	qvalue= 0.603	sequence=GLALYNSSMLDLPLPPVVFKKLLGCPVGLEDFTVVNPTVGRSLSQLLKFT
YALI0_A10637g	pvalue=5.55e-05	qvalue= 0.603	sequence=NSRSSKSRLDRIESLLDKVLRRLGDDRLGDAADIILRHNGSEFVSRRASP
YALI0_C10780g	pvalue=6.69e-05	qvalue= 0.603	sequence=EKLGELEIVGDGIFGSELTSSTYTNPLTDTHHPIYHADYVSDESGTGMVH
YALI0_D20460g	pvalue=3.78e-05	qvalue= 0.603	sequence=FHKRYSHLTESLFETGLNFLRNWRPPVLPSRPICMNEDVLRFLARTFLEA
YALI0_D21802g	pvalue=6.53e-05	qvalue= 0.603	sequence=TKSGSPFGSVFSSNVVKKALANHYPTTPRIEEKGHRKTQSAGQQYSSMII
YALI0_B08382g	pvalue=4.09e-05	qvalue= 0.603	sequence=QPELKDARKVTKKFLEDNSLVHFDKSTLVYGNCISETYADMEDWKLVAEN
YALI0_A04521g	pvalue=6.07e-05	qvalue= 0.603	sequence=PGRDRCARLEGAVQLMKQNGAGNMARLVDYDEIGYERDLVYITACQGAEP
YALI0_F19888g	pvalue=4.23e-05	qvalue= 0.603	sequence=PTYGASYGPISYAGLSPEVLAAANAMNTTPGAIQQTEEEAERAAEKAANP
YALI0_B07535g	pvalue=1.41e-05	qvalue= 0.603	sequence=GYLRETLRGVLSRDVSDLSRDVSGMSVTDSASVASMGSVGSRLSAGSRVP
YALI0_A09020g	pvalue=1.71e-05	qvalue= 0.603	sequence=SFYQSMRFVITANDTGGSRLCIWDDLVYTEKKRSLFKSVVETVAMRYMEH
YALI0_F30085g	pvalue=9.29e-05	qvalue= 0.708	sequence=GKKSPITGKQNARKTTGLKAASTRSIINKYHQLNKRKLQLQGLKKRMKKA
YALI0_F08899g	pvalue=5.5e-05	qvalue= 0.603	sequence=ENRISLGRSISAVMGILRELQDVNRDRPVYYPPVANPQPPTATSSRHSSP
YALI0_A19580g	pvalue=6.02e-05	qvalue= 0.603	sequence=QQQLLCLGRAVLKKAAVVCLDEATSSVDKDTEKVVGSVLQNAFANSTVIT
YALI0_A19580g	pvalue=9.72e-05	qvalue= 0.708	sequence=QNAFANSTVITIAHRLETVIDSDKICVLDKGEVVEYGSPSELLAKKGLFY
YALI0_C22231g	pvalue=3.61e-06	qvalue= 0.603	sequence=ETTEPEPAPGRLTSRGVVSFAAMERKSRILKLLSANNGVLEGGKQLEIMF
YALI0_F09185g	pvalue=2.07e-05	qvalue= 0.603	sequence=LKFGVEHGVDMIFASFVRTANDVQAIRDVLGEKGKGIQVISKIENQQGVN

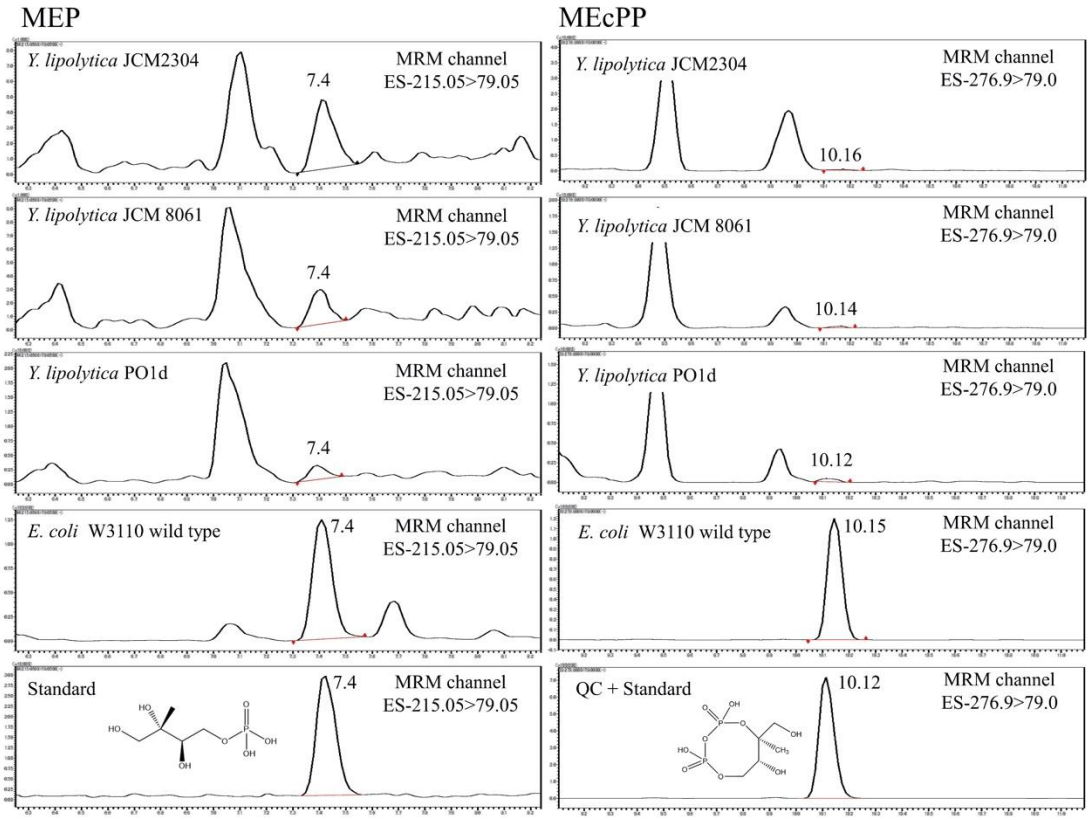


Figure S 1 analysis of MEP and MEcPP in different *Y. lipolytica* strains

In another analysis of different *Y. lipolytica* strains suggested that MEcPP; another MEP pathway metabolite was also found in *Y. lipolytica*. Note that we use the *E. coli* W3110 wild type as a biological standard. For MEcPP, we found that the retention time of the E. coli W3110 wild type was a little bit different; therefore, we mixed the MEcPP with yeast and E. coli to make a QC sample for retention time comparison. Although the MRM transition and retention time of the detected MEcPP is comparable to the standard, the intensity of detected MEcPP in yeast sample was still very low.

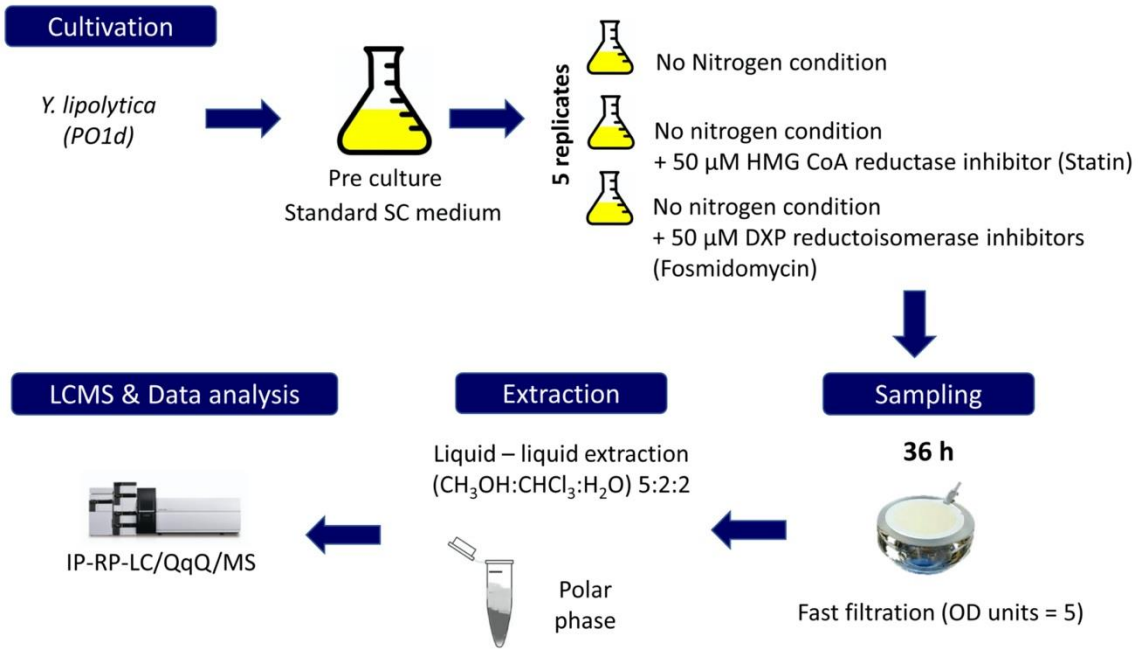
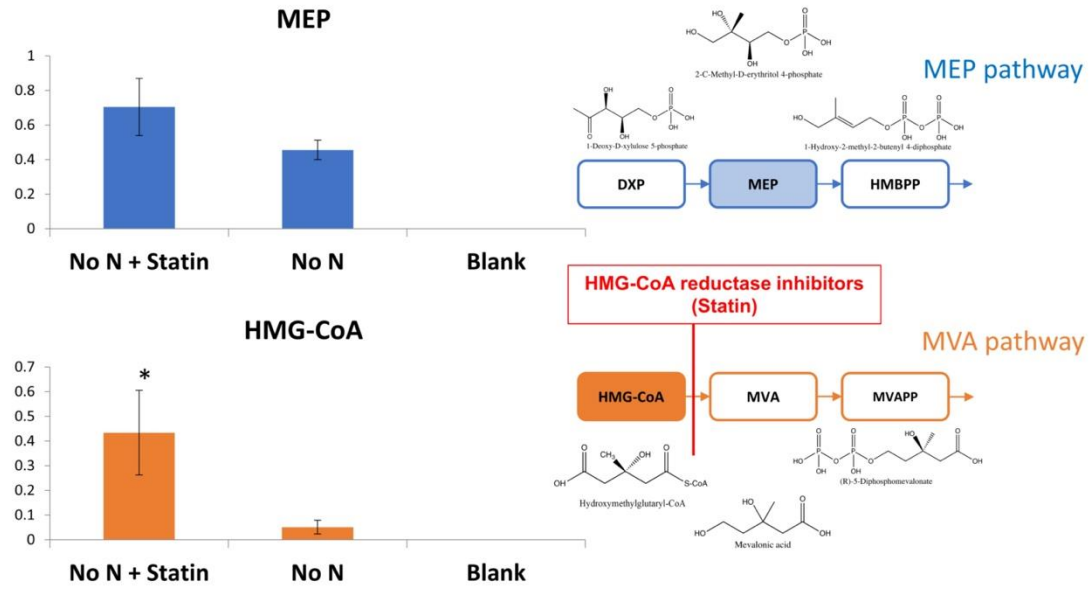


Figure S 2 Additional chemical inhibitor experimental design



- MEP relative intensity was increased, however it is not statistically significant (P-value = 0.068)
- HMG-CoA increased significantly, confirming the effect of HMG-CoA reductase inhibitors (P-value = 0.018)

Figure S 3 relative intensity of MEP and HMG-CoA subjected to HMG-CoA reductase inhibitor

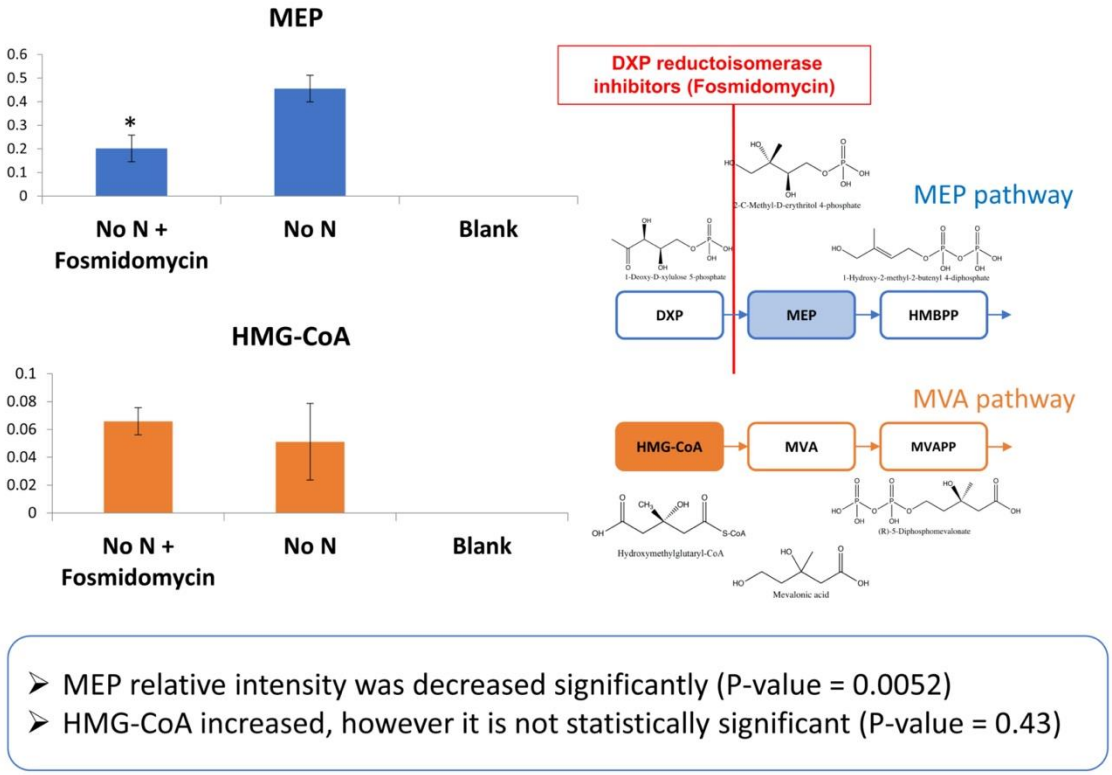


Figure S 4 relative intensity of MEP and HMG-CoA subjected to DXP reductoisomerase inhibitors

Acknowledgement

First of all, I would like to express my great gratitude to Prof. Eiichiro Fukusaki, for his support and excellent collaborative effort arrangement. I would like to express my deep appreciation to Assoc. Prof. Shuichi Shimma, for his kindness in process my chemical, materials and consumables order during my study, moreover I thank you for the opportunity to work collaboratively with your research group. And most importantly I would like to express my sincere gratitude to Assist. Prof. Sastia Prama Putri for valuable and constructive suggestions helped me conclude the thesis topic and guided me over these years to achieve the final goals. During the most difficult times she gave me the moral support I needed to move on. Besides my supervisors, I would like to deliver my gratitude to Prof. Tomohisa Kuzuyama for his insightful comment on my research and publication.

I would like to give special appreciation to Assoc. Shigeru Kitani and Ms. Yuri Nishimoto for their help in my research. I would also like to thank the rest of my thesis committee: Prof. Kazuhito Fujiyama and Prof. Prof. Toshiya Muranaka for their insightful comments which incented me to improve my doctoral dissertation from various perspectives. I would also like to thank Ministry of Education, Culture, Sports, Science and Technology (MEXT) for financial support throughout my study.

Dozens of people have helped and taught me immensely at Osaka University. Special thanks to Adinda Putri Wisman (Wissy), for being there for me from the beginning until the end of my Ph.D. journey. I can’t thank you enough for encouraging me throughout this experience. I would like to also thank you Arakawa (Tom kun) for the technical support during my study. I also wish to thank to all Fukusaki Lab members especially to my beloved friends

Adinda Kadar, Safira, Harada, Kana, Nakano, Nakatani, Abi, Malik, Hadi, and Mega for your supports not only in my research but also in my life here. My gratitude also goes to my Thai friends especially Nanny, Boss, and Anna for their warm friendship. All of you are my second family here in Japan. Finally, I would like to thank my family, for their endless love, support and encouragement in my whole life. Words cannot express how grateful I am.

List of publications

Dissook, S., Putri, S. P., & Fukusaki, E. (2021). Metabolomic Analysis of Response to Nitrogen-Limiting Conditions in *Yarrowia* spp. *Metabolites*, 11(1), 16.

Dissook, S., Kuzuyama ,T., Nishimoto, Y., Kitani, S., Putri, S. P., & Fukusaki, E. (2021). Stable isotope and chemical inhibition analyses suggested the existence of a non-mevalonate-like pathway in the yeast *Yarrowia lipolytica*. *Scientific reports*, (Accepted for publication).