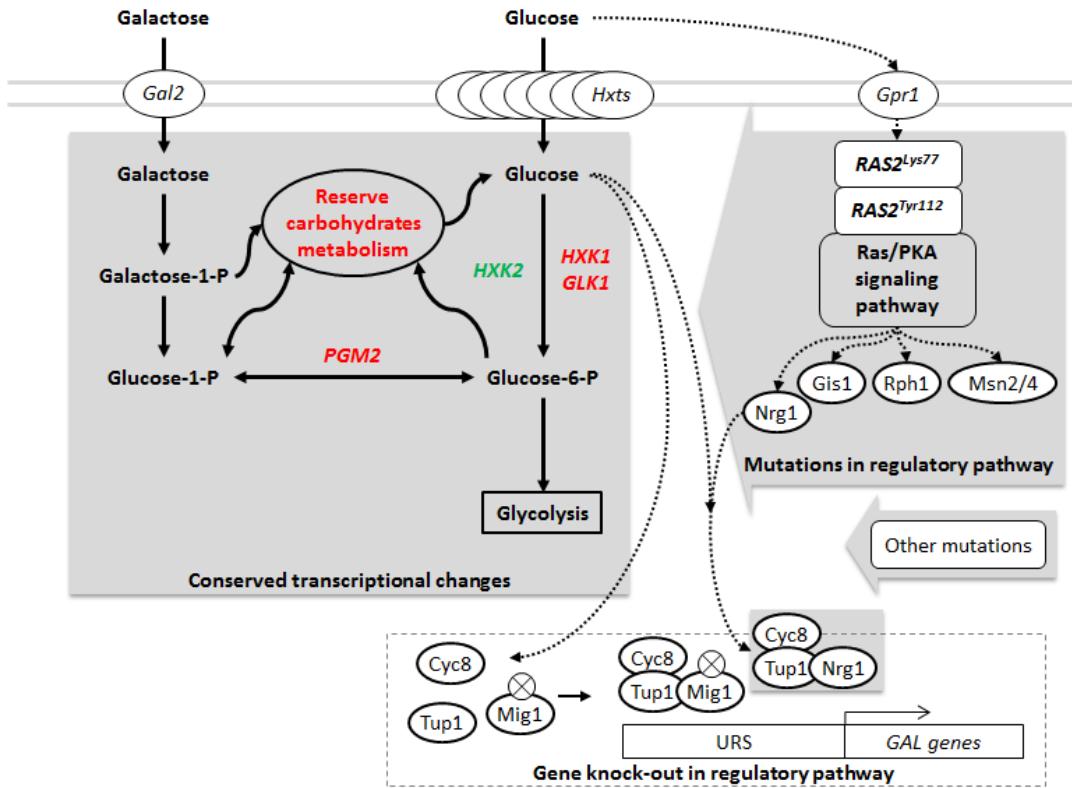


CHALMERS



Advancing Metabolic Engineering through combination
of Systems Biology and Adaptive Evolution

KUK-KI HONG

Department of Chemical and Biological Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2012

Thesis for the degree of doctor of philosophy

**Advancing Metabolic Engineering through
Combination of Systems Biology and Adaptive Evolution**

KUK-KI HONG



Systems and Synthetic Biology
Department of Chemical and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2012

Advancing Metabolic Engineering through combination of Systems Biology and Adaptive Evolution

KUK-KI HONG

© KUK-KI HONG, 2012.

ISBN: 978-91-7385-735-2

Doktorsavhandlingar vid Chalmers tekniska högskola

Ny serie nr 3416

ISSN: 0346-718X

PhD Thesis

Systems and Synthetic Biology

Department of Chemical and Biological Engineering

Chalmers University of Technology

SE-412 96 Göteborg

Sweden

Telephone +46 (0) 31-772 1000

Thesis Supervisor: Prof. Jens B Nielsen

Primary Funding Source:

Bioscience Research Center, CJ CheilJedang Corp.

92-1 Gayang-dong, Gangseo-gu, Seoul

Republic of Korea

Cover illustration: evolutionary strategies of yeast for improving galactose utilization; for more details, refer to Fig. 3-13.

Printed by Chalmers Reproservice

Göteborg, Sweden 2012

To my family

My wife, Min-Jin

My angels, Jin-Seo and Jin-Ha

A thunderstorm can be viewed as a consequence of Zeus' anger or of a difference of potential between the clouds and the earth. A disease can be seen as the result of a spell cast on the patient or of an infection by a virus. In all cases, however, one watches the visible effect of some hidden cause related to the whole set of invisible forces that are supposed to run the world.

- François Jacob (Evolution and Tinkering, Science, 1977)

PREFACE

This dissertation is submitted for the partial fulfillment of the degree of doctor of philosophy at the department of chemical and biological engineering, Chalmers University of Technology, Sweden. The doctoral research is on the application of systems biology for the characterization of adaptively evolved mutants. The process of systems biology approach enhances the understanding of evolutionary strategies that may contribute to advance metabolic engineering. This advance is likely useful to improve biological engineering, which provides one of the possible solutions to substitute petroleum based chemical production. This research was funded by the doctoral fellowship program of CJ CheilJedang (Korea), the Chalmers Foundation, the Knut and Alice Wallenberg Foundation, the European Union funded projects UNICELLSYS (Contract 201142), SYSINBIO (Contract 212766), European Research Council Grant 247013 and the Novo Nordisk Foundation.

Kuk-Ki Hong

August 2012

ABSTRACT

Understanding evolutionary strategies of microorganisms may provide opportunities for advanced strain development with the aim to produce valuable bio-products from renewable biomass resources. Through evolutionary processes, microorganisms can attain new traits associated with genetic changes that may be useful for the construction of improved strains. Therefore, the characterization of evolutionary strategies may result in identification of the molecular and genetic changes underlying newly obtained traits, and can hereby become an essential step in strain development. However, so far the depth of analysis has limited the range of comprehension. This thesis applied genome-wide analyses such as transcriptome, metabolome and whole-genome sequencing to investigate the evolutionary strategies of the yeast *Saccharomyces cerevisiae*. Three evolved mutants were independently generated by adaptive evolution on galactose minimal media to obtain the trait of improved galactose utilization by yeast. Those strains expressed higher galactose utilization rates than a reference strain in terms of both maximum specific growth rate and specific galactose uptake rate. Application of the genome-scale comparative analyses employing engineered strains as controls elucidated unique changes obtained by adaptive evolution. Molecular bases referred from the changes of transcriptome and metabolome were located around galactose metabolism, while genetic bases from whole-genome sequencing showed no mutations in those changes. Common mutations among the evolved mutants were identified in the Ras/PKA signaling pathway. Those mutations were placed on the reference strain background and their effects were evaluated by comparison with the evolved mutants. One of the site-directed mutants showed even higher specific galactose uptake rate than the evolved mutants, and just few number of genetic and molecular changes were enough to recover complete the adaptive phenotype. These results indicate that identification of key mutations provide new strategies for further metabolic engineering of strains. In addition, the pleiotropy of obtained phenotype that is improved galactose availability was tested. When the galactose-evolved mutants were cultured on glucose that is the most favorite carbon source of yeast, those mutants showed reduction of glucose utilization. Genome-wide analyses and site-directed mutagenesis were applied again to understand underlying molecular and genetic bases of this trade-off in carbon utilization. The results indicated that loosening of tight glucose regulation was likely the reason of increased galactose availability. The implications of evolutionary strategies and the impact of genome-scale analyses on characterization of evolved mutants are discussed.

Key words: metabolic engineering, evolutionary engineering, systems biology, galactose utilization, Ras/PKA signaling pathway, pleiotropy of evolutionary strategies

LIST OF PUBLICATIONS

This thesis is based on the following publications, referred to as Paper I to IV in the text:

- I. **Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis**
Kuk-Ki Hong, Wanwipa Vongsangnak, Goutham N. Vemuri and Jens Nielsen
Proc. Natl. Acad. Sci. USA. 2011 Jul 19; 108(29):12179–84.
- II. **Recovery of phenotypes obtained by adaptive evolution through inverse metabolic engineering**
Kuk-Ki Hong and Jens Nielsen
Accepted in Appl. Environ. Microbiol. 2012
- III. **Adaptively evolved yeast mutants on galactose show trade-offs in carbon utilization on glucose**
Kuk-Ki Hong and Jens Nielsen
Submitted for publication
- IV. **Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries (Review)**
Kuk-Ki Hong and Jens Nielsen
Cell Mol Life Sci. 2012 Aug; 69(16):2671-90. Epub 2012 Mar 3.

Additional publications during doctoral research not included in this thesis

- V. **Dynamic (13) C-labeling experiments prove important differences in protein turnover rate between two *Saccharomyces cerevisiae* strains.**
Kuk-Ki Hong, Jin Hou, Saeed Shoaie, Jens Nielsen and Sergio Bordel
FEMS Yeast Res. 2012 Jun 20. doi: 10.1111/j.1567-1364.2012.00823
- VI. **Quantitative analysis of glycerol accumulation under hyper-osmotic stress and its various links to glycolysis**
Elzbieta Petelenz-Kurdziel, Clemens Kuehn, Bodil Nordlander, Dagmara Klein, Kuk-Ki Hong, Therese Jacobson, Peter Dahl, Joerg Schaber, Jens Nielsen, Stefan Hohmann, Edda Klipp.
Submitted for publication

CONSTRIBUTION SUMMARY

A summary of contribution of Kuk-Ki Hong to each of the publications

- I. Designed research; performed research; analyzed data; wrote the paper.
- II. Designed research; performed research; analyzed data; wrote the paper.
- III. Designed research; performed research; analyzed data; wrote the paper.
- IV. Designed review; analyzed data; wrote the paper.

- V. Designed research; performed research; analyzed data; wrote the paper.
- VI. Performed research; analyzed data.

TABLE OF CONTENTS

Preface.....	5
Abstract.....	6
List of publications / Contribution summary.....	7
List of figures and tables.....	10
Abbreviations and Symbols.....	11
1. INTRODUCTION	
1.1. Yeast <i>Saccharomyces cerevisiae</i> for future biorefineries.....	13
1.2. Evolutionary approaches in strain development.....	18
1.2.1. Evolutionary engineering.....	18
1.2.2. Inverse metabolic engineering.....	18
1.2.3. Adaptive evolution.....	19
1.3. Characterization of evolved mutants by genome-scale analysis.....	20
1.3.1. Transcriptome analysis.....	20
1.3.2. Metabolome analysis.....	21
1.3.3. Whole-genome sequencing.....	21
1.4. Improving galactose utilization in <i>S. cerevisiae</i>	24
1.4.1. Galactose metabolism in <i>S. cerevisiae</i>	24
1.4.1.1. Leloir pathway.....	24
1.4.1.2. Regulation of <i>GAL</i> genes.....	25
1.4.2. Galactose as a feedstock in industrial biotechnology.....	27
1.4.2.1. Galactose content of biomass.....	27
1.4.2.2. Metabolic engineering for improved galactose utilization.....	29
2. OVERVIEW OF THE THESIS.....	31
3. RESULTS AND DISCUSSION.....	35
3.1. Molecular and genetic basis of evolutionary strategies of the galactose-evolved mutants.....	35
3.2. Complete recovery of adaptive phenotype through inverse metabolic engineering.....	42
3.3. Characterization of molecular mechanism of trade-offs in carbon utilization.....	48
4. CONCLUSIONS.....	53
5. PERSPECTIVE.....	56
6. ACKNOWLEDGEMENTS.....	59
7. REFERENCES.....	61
[Publications]	

LIST OF FIGURES AND TABLES

Fig. 1-1.	Overview of relevant carbon sources for yeast fermentation.....	17
Fig. 1-2.	Galactose pathway (Leloir pathway) and regulation in <i>S. cerevisiae</i>	26
Fig. 1-3.	Composition of non-food biomass.....	28
Fig. 3-1.	Yeast strains used in this study.....	35
Fig. 3-2.	Overall flow of the experiments in this study.....	36
Fig. 3-3.	Phenotypic changes of evolved mutant strains.....	37
Fig. 3-4.	Changes in the galactose, reserve carbohydrates, and ergosterol metabolism in the evolved mutants.....	38
Fig. 3-5.	Summary of evolution changes in the evolved mutants.....	41
Fig. 3-6.	Overall fermentation physiology of the site-directed mutants and the combined mutants.....	44
Fig. 3-7.	Effect of reconstructed strains compared to the evolved strains by differentially expressed genes.....	46
Fig. 3-8.	Changes in the galactose and reserve carbohydrates metabolisms in the reconstructed strains.....	47
Fig. 3-9.	Comparison of fermentation physiology of the evolved mutants and the engineered mutants in galactose (gal) and glucose (glu).....	48
Fig. 3-10.	Transcriptome analysis of the evolved mutants in galactose (gal) and glucose (glu) through principal component analysis (PCA).....	49
Fig. 3-11.	Patterns of common molecular changes of the evolved mutants in galactose (gal) and glucose (glu).....	50
Fig. 3-12.	Comparison of fermentation physiology of the site-directed mutants in galactose (gal) and glucose (glu).....	51
Fig. 3-13.	Summary of possible molecular mechanism for the trade-off in galactose and glucose utilization.....	52
Table. 1-1.	Chemicals are recently produced from biomass (companies and host strains).....	13
Table. 1-2.	Examples of products and strains of <i>S.cerevisiae</i>	14
Table. 1-3.	The price of the next-generation sequencing.....	22
Table. 1-4.	The performance of overall genome sequencing results.....	22
Table. 1-5.	Composition of cheese whey.....	28
Table. 3-1.	Genetic changes in the evolved mutants.....	40
Table. 3-2.	Reconstructed strains and control strains.....	43

ABBREVIATIONS AND SYMBOLS

β Gal: β -D-galactose

α Gal: α -D-galactose

α Gal-1P (Gal 1P, Galactose 1P, Galactose-1-P): α -D-galactose-1-phosphate

α Glu-1P (Glu 1P, Glucose 1P, Glucose-1-P): α -D-glucose-1-phosphate

α Glu-6P (Glu 6P, Glucose 6P, Glucose-6-P): α -D-glucose-6-phosphate

UDP-Glu (UDP-Glucose): Uridine diphosphate glucose

UDP-Gal (UDP-Galactose): Uridine diphosphate galactose

gal: galactose

glu: glucose

Trehalose 6P: Trehalose-6-phosphate

REF: Reference strain

KEGG: Kyoto Encyclopedia of Genes and Genomes

GO term: Gene Ontology term

PCA: Principal component analysis

PC: Principal component

SNPs: Single-nucleotide polymorphism

gTME: Global transcription machinery engineering

MAGE: Multiplex automated genome engineering

TRMR: Trackable multiplex recombineering

Yeast nomenclature

Gene name consists of three letters and up to three numbers, ex. *GAL4*, *MIG1*, *ura3*

Wild-type gene name is written with capital letters in italic, ex. *PGM2*, *RAS2*, *UGP1*

Recessive mutant gene name is written with small letters in italics, ex. *mig1*, *gal80*, *gal6*

Mutant alleles are named with a dash and a number, ex. *ura3-52*, *cdc28-2*

Deleted gene with the genetic marker is used for deletion, ex. *tps1Δ::HIS3*

The gene product, a protein, is written with a capital letter at the first letter and not in italics; often a "p" is added at the end, ex. Pgm2p, Ugp1p

Some genes, which are only found by systematic sequencing and their functions are not determined, get a landmark name, ex. YNL200C, YHL042W, YLR278C

(Y, yeast; the second letter, the chromosome (D=IV, M=XIII....); L or R, left or right chromosome arm; the three-digit number; the ORF counted from the centromere; C or W, Crick or Watson, i.e. direction of the ORF)

Exceptional case, ex. *HO*, *MAT α* , *MAT α*

Amino acid sequence change is described by gene name and changed amino acid with its position, ex. *RAS2^{Lys 77}*, *ERG5^{Pro 370}*

1. INTRODUCTION

1.1. Yeast *Saccharomyces cerevisiae* for future biorefineries

Even before recognizing the presence of microorganisms, mankind has used microbial fermentation to produce beverages and foods. Since 1920 industrial microbial fermentation has been used to manufacture organic acids, amino acids and vitamins (Kinoshita, *et al.*, 1957, Nakayama, *et al.*, 1961, Demain, 2000). The advent of genetic engineering in the 1970s led to the use of microbial fermentation for the production of pharmaceutical proteins such as human insulin and human growth hormone (Goeddel, *et al.*, 1979, Johnson, 1983). Currently, the world is confronting serious challenges such as climate changes due to greenhouse gas emission and the depletion of petroleum oil causing limitation of energy and chemical resources. Microbial fermentation is considered as one of the possible solutions to these grand challenges, because it uses renewable biomass that can also absorb carbon dioxide during growth, and produce fuels and chemicals in eco-friendly processes (Lipinsky, 1981, Werpy & Petersen, 2004, Vennestrom, *et al.*, 2011). There are already several successful industrial trials to produce chemicals from biomass by microbial fermentations (Table 1-1).

Table 1-1. Chemicals are recently produced from biomass, including major players and host strains

Chemicals	Products/Uses	Major players	Host strains
succinic acid	plastics, chemical intermediates, solvents, polyurethanes, plasticizers	BASF/Purac(CSM)	<i>Basfia succiniciproducens</i> (from Bovin rumen, Gram-negative)
3-hydroxypropionic acid	acrylic acid: plastics, fiber, coatings, paints, super-absorbent diapers	Novozymes/Cargill	<i>Escherichia coli</i> <i>Saccharomyces cerevisiae</i>
isoprene	synthetic rubber	Genencor(Danisco) /Goodyear	<i>Bacillus subtilis</i> , <i>Escherichia coli</i> , <i>Pantoea citrea</i> , <i>Trichoderma reesei</i> , <i>Yarrowia lipolytica</i>
lactic acid	plastics, synthetic fibers	Cargill	<i>Kluyveromyces marxianus</i>

lactic acid	plastics, synthetic fibers	Purac(CSM)/Arkema	thermophilic <i>Bacillus</i> , thermophilic <i>Geobacillus</i>
1,3-propanediol	engine coolant, cosmetics, surfactants, emulsifiers, preservatives, polymers	Dupont/Tate&Lyle	<i>Escherichia coli</i>
propylene	Thermoplastic	Braskem/Novozymes	<i>Propionibacterium acidipropionici</i>

Strain development is a pre-requisite to materialize bio-based chemical production, as it is directly related to not only improving yield, titer, and productivity of products, but also utilizing cheap raw materials efficiently (Tyo, *et al.*, 2007, Patnaik, 2008, Elkins, *et al.*, 2010). The yeast *Saccharomyces cerevisiae* has been used for the production of a wide range of industrial products due to its tolerance to industrial conditions and the vast amount of knowledge about its physiology, biochemistry, genetics, and long history of fermentation (Pronk, 2002, van Maris, *et al.*, 2006, Nevoigt, 2008, Nielsen & Jewett, 2008, Krivoruchko, *et al.*, 2011). Thus, its products range and available current technologies are quite broad (Table 1-2).

Table 1-2. Examples of products and strains of *S.cerevisiae*

(More detailed explanation is in **Paper IV.**)

Categories	Products	Strains	References
Biofuels	Ethanol	<i>CEN.PK102-3A (MATα ura3 leu2)</i>	(Guadalupe Medina, <i>et al.</i> , 2010)
	Biobutanol	<i>CEN.PK 2-1C (MATα leu2-3, 112 his3-Δ1 ura3-52 trp1-289 MAL2-8(Con) MAL3 SUC3)</i>	(Chen, <i>et al.</i> , 2011)
	Biodiesels	<i>YPH499 (MATα ura3-52 lys2-801_amber ade2-101_ochre trp1-D63 his3-D200 leu2-D1)</i>	(Yu, <i>et al.</i> , 2012)
	Bisabolene	<i>BY4742 (MATα his3D1 leu2D0 lys2D0 ura3D0)</i>	(Peralta-Yahya, <i>et al.</i> , 2011)
Bulk chemicals	1,2-propanediol	<i>NOY386αA (MATα ura3-52 lys2-801 trp1-Δ63 his3-Δ200 leu2-Δ1), BWG1-7a (MATα ade1-100 his4-519 leu2-3,112 ura3-52 GAL$^+$)</i>	(Lee & Dasilva, 2006)
	L-Lactic acid	<i>CEN. PK2-1C (MATα ura3-52 trp1-289 leu2-3,112 his3Δ1 MAL2-8C SUC2)</i>	(Zhao, <i>et al.</i> , 2011)
	Polyhydroxy-alkanoates	<i>BY4743 (MATα/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 ura3Δ0/ura3Δ0)</i>	(Zhang, <i>et al.</i> , 2006)

	Pyruvic acid	<i>CEN.PK113-7D (MATα MAL2-8C, SUC2)</i>	(van Maris, et al., 2004)
	Succinic acid	<i>AH22ura3 (MATα ura3Δ leu2-3 leu2-112 his4-519 can1)</i>	(Raab, et al., 2010)
	β -amyrin	<i>CEN.PK113-7D (MATα MAL2-8C SUC2)</i>	(Madsen, et al., 2011)
	β -carotene	<i>CEN.PK113-7D (MATα MAL2-8C SUC2)</i>	(Verwaal, et al., 2007)
	Amorpha-4, 11-diene	<i>CEN.PK2-1C (MATα ura3-52 trp1-289 leu2-3,112 his3Ä1 MAL2-8C SUC2), CEN.PK2-1D (MATα ura3-52 trp1-289 leu2-3,112 his3Ä1 MAL2-8C SUC2)</i>	(Westfall, et al., 2012)
Fine chemicals	Cinnamoyl anthranilates	<i>BY4742 (MATα his3D1 leu2D0 lys2D0 ura3D0)</i>	(Eudes, et al., 2011)
	Cubebol	<i>CEN.PK113-5D (MATα MAL2-8c SUC2 ura3-52)</i>	(Asadollahi, et al., 2010)
	Eicosapentaenoic acid (EPA)	<i>CEN.PK113-5D (MATα MAL2-8c SUC2 ura3-52)</i>	(Tavares, et al., 2011)
	Linalool	<i>BQS252 (MATα ura3-52 (derivative of FY1679))</i>	(Rico, et al., 2010)
	Methylmalonyl-coenzyme A	<i>InvSC1 (MATα, his3delta1, leu2, trp1-289, ura3-52 (Invitrogen, Carlsbad, CA, USA)) BJ5464 (MATα, ura3-52, trp1, leu2-delta1, his3-delta200, pep4::HIS3, prb1-delta1.6R, can1, GAL).</i>	(Mutka, et al., 2006)
	Patchoulol	<i>CEN.PK113-13D and CEN.PK113-5D</i>	(Albertsen, et al., 2011)
	Resveratrol	<i>FY23 (MATα ura3-52 trp1A63 leu2A1)</i>	(Becker, et al., 2003)
	Vanillin	<i>X2180-1A (MATα his3D1 leu2D0 met15D0 ura3D0 adh6::LEU2 bgl1::KanMX4 PTPI1::3DSD [AurC]::HsOMT [NatMX]::ACAR [HphMX])</i>	(Brochado, et al., 2010)
	Se-methylselenocysteine	<i>CEN.PK113-7D (MATα MAL2-8C SUC2)</i>	(Mapelli, et al., 2011)
Protein drugs	Non-ribosomal peptides	<i>CEN.PK113-11C (MATα MAL2-8c SUC2 ura3-52 his3-D1)</i>	(Siewers, et al., 2010)
	Insulin-like growth factor 1 (fIGF-1)	<i>GcP3 (MATα pep4-3 prb1-1122 ura3-52 leu2 gal2 cir°)</i>	(Vai, et al., 2000)
	Glucagon	<i>SY107 (MATα YPS1 Δtpi::LEU2 pep4-3 leu2 Δura3 cir⁺)</i>	(Egel-Mitani, et al., 2000)

single-chain antibodies (scFv)	<i>BJ5464</i> (<i>a ura3-52 trp1 leu2D1 his3D200 pep40HIS3 prb1D1.6R can1 GAL</i>)	(Hackel, <i>et al.</i> , 2006)
Hepatitis surface antigen (HBsAg)	<i>INVSc1</i> (<i>MATα his3D1 leu2 trp1-289 ura3-52</i>)	(Vellanki, <i>et al.</i> , 2007)
Parvovirus B19 VP2	<i>HT393</i> (<i>MATα leu2-3 leu2-112 ura3Δ5 prb1-1 prc1-1 pra1-1 pre1-1</i>)	(Lowin, <i>et al.</i> , 2005)
Epidermal Growth factor (EGF)	<i>W303-1A</i> (<i>MATα leu2-3,112 his3-11,15 ade2-1 ura3-1 trp1-1 can1-100</i>), <i>W303-1B</i> (<i>MATα leu2-3,112 his3-11,15 ade2-1 ura3-1 trp1-1 can1-100</i>)	(Chigira, <i>et al.</i> , 2008)
Immunoglobulin G	<i>BJ5464a</i> (<i>MATα ura3-52 leu2~1 his3~200 pep4::HIS3 prb1~1.6Rcan1 GAL</i>)	(Rakestraw, <i>et al.</i> , 2009)
Hepatitis B virus surface antigen (HBsAg)	<i>S.cerevisiae 2805</i> (<i>MATα pep4::HIS3 prb-Δ1.6 his3 ura3-52 gal2 can1</i>)	(Kim, <i>et al.</i> , 2009)
L1 protein of human papillomavirus (HPV) type16	<i>S.cerevisiae 2805</i> (<i>MATα pep4::HIS3 prb-Δ1.6 his3 ura3-52 gal2 can1</i>)	(Kim, <i>et al.</i> , 2010)

There is also extensive research on extending substrate range of this yeast. Resources for traditional fermentations have been derived from food crops like corn, wheat and sugar cane, but to replace the large amounts of fuels and chemicals currently derived from mineral oil, the use of abundant and renewable non-food resources such as switchgrass, corn-cob, bagasse, cheese whey and algae is necessary. These biomass resources are composed of diverse kinds of carbon structure: polymers (cellulose, starch, xylan), dimers (cellobiose, melibiose, lactose) and monomers (glucose, fructose, galactose, arabinose, xylose). Except the hexoses (glucose, fructose, galactose) and a few dimers (sucrose, maltose), most of these carbon compounds are not endogenously metabolized by *S. cerevisiae*. Even among the hexoses there are broad differences in uptake rate, for example the uptake rate of galactose is much lower than for the other hexoses. Therefore, the extension of substrate range of *S. cerevisiae* provides an excellent opportunity to enhance its suitability for biofuels and biochemicals production (van Maris, *et al.*, 2006, Hahn-Hagerdal, *et al.*, 2007, Nevoigt, 2008, Marie, *et al.*, 2009) (Fig. 1-1).

(For details, refer to **Paper IV**.)

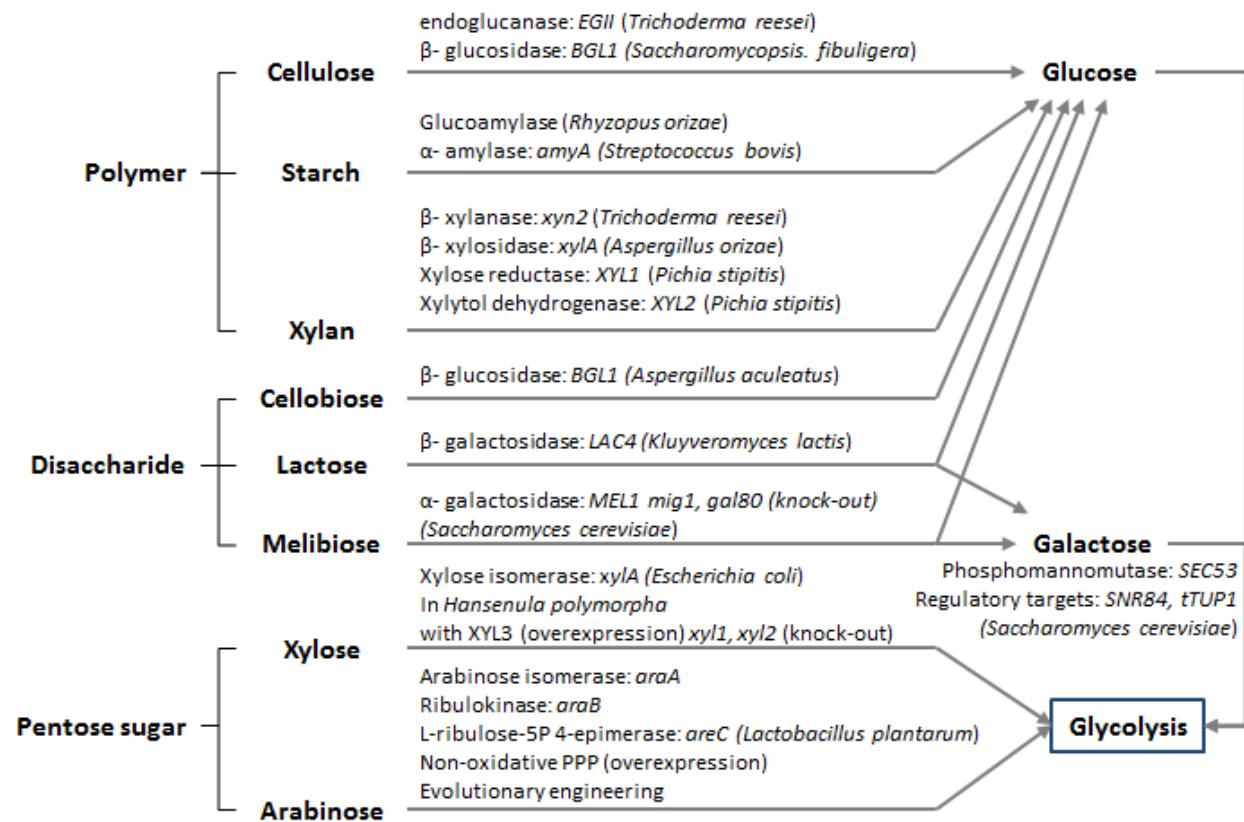


Fig. 1-1. Overview of relevant carbon sources for yeast fermentation. Heterologous enzymes that are currently introduced are summarized for non-utilizable carbon sources (polymer, disaccharide and pentose sugar) and non-preferred one (galactose) in *S. cerevisiae*.

1.2. Evolutionary approaches in strain development

1.2.1. Evolutionary engineering

Evolutionary engineering has been traditionally employed for strain development in industry, since it can generate specific traits relatively quickly at some level, even though governing biological principle may not be evident (Sauer, 2001, Zhang, *et al.*, 2002). The term of evolutionary engineering is composed of evolution and engineering. The evolution is a strategy of life to adapt to changed environments by natural selection. It is operated through iterative process of creating variation in population and selecting proper individuals, consequently a specific trait in the population is enriched. This nature's algorithms can be engineered to make biotechnological relevant traits by adjusting the rate of variant generation or defining new selection pressures. Therefore, evolutionary engineering is the application of suitable mutagenesis and artificially designed selection procedures based on evolutionary mechanisms for strain development. Success of evolutionary engineering is, hence, dependent on the ability to design mutagenesis and selection conditions (Sauer, 2001, Sonderegger & Sauer, 2003). On the one hand random mutagenesis by treatment of mutagens for adjusting mutation rate can be used; it generates a broader distribution of mutations in the genome, whereas it makes it more difficult to identify beneficial mutations (Sauer, 2001, Ikeda, *et al.*, 2006). On the other hand, the methods that can generate traceable mutations in specific regions have been developed, which is called genome engineering such as gTME, MAGE and TRMR (Santos & Stephanopoulos, 2008, Boyle & Gill, 2012). Another important consideration is to understand the underlying evolution mechanisms. Basically, evolutionary engineering relies on evolutionary mechanisms such as natural selection or natural preservation (Darwin regretted using selection more frequently than preservation). Since natural preservation is the fundamental evolutionary mechanism, this concept is used in the design stage of evolutionary engineering. It is also important to be aware of other relevant evolutionary mechanisms for strain development such as clonal interference, trade-offs in traits, negative epistasis (Elena & Lenski, 2003).

1.2.2. Inverse metabolic engineering

Developed strains based on evolutionary engineering can be directly used for industrial application. In most cases, only a few specific traits from the mutant strains are needed; however

industrial strains retain the combination of several non-necessary traits simultaneously (Ohnishi, *et al.*, 2002, Ikeda, *et al.*, 2006, Ikeda, *et al.*, 2009, Warner, *et al.*, 2009). Therefore, additionally the concept of inverse metabolic engineering has been used to make evolutionary engineering more useful. Inverse metabolic engineering starts from the identification of genetic basis of obtained phenotypes, and completed by transfer of that specific genotype(s) to an industrial strain (Bailey, *et al.*, 2002, Ikeda, *et al.*, 2006). The important part in this engineering is the identification of the genetic basis; not only in order to enable the transfer of genetic changes related to the gained trait, but also the specific trait(s) may not easily be reached to its optimum stage because of evolutionary constraints such as negative epistasis, clonal interference; therefore, additional engineering based on the identified genetic changes is sometimes required (Warner, *et al.*, 2009). For these reasons, the identification of the genetic bases of selected traits is crucial. Recently, analytical capabilities that can scan molecular or genetic alteration at genome scale have been developed such as omics tools and next generation whole genome sequencing. The integration of data generated from those tools is expected to facilitate identification of molecular and genetic changes in a more comprehensive fashion (Bro & Nielsen, 2004, Heinemann & Sauer, 2010, Oud, *et al.*, 2012).

1.2.3. Adaptive evolution

Adaptive evolution is often confused with several similar terms such as adaptive laboratory evolution, experimental evolution, and even evolutionary engineering (Sauer, 2001, Elena & Lenski, 2003, Conrad, *et al.*, 2011, Portnoy, *et al.*, 2011, Dettman, *et al.*, 2012). Adaptive evolution has been used to explain adaptation process of life in biology. When this process can be imitated in a laboratory to understand evolution mechanisms or applied to strain development, derivative words have been generated. Therefore, adaptive evolution includes both natural processes in basic science and a tool in biotechnology. It generates mutations spontaneously based on the cell's endogenous system, and finds a phenotype that have improved fitness to a given environment than an ancestor strain, simply by continuous exposure of a population to the given environment over a period of time.

1.3. Characterization of evolved mutants by genome-scale analysis

Evolutionary engineering has made commercially successful stories in strain development, while the identification of molecular or genetic basis that is involved in phenotypic changes has remained an enduring challenge (Ohnishi, et al., 2002, Warner, et al., 2009). There have been several efforts to find genetic changes. If biological information about an obtained phenotype is present at the pathway level or its regulation, one can check molecular changes in that specific pathway. For example, the strain producing high concentration of lysine was characterized based on analysis of specific amino acid production pathways (Ikeda, et al., 2006). Key mutations that were likely related to release of allosteric regulation were detected, and partial contribution of the mutations on the overall phenotype was confirmed. Technological advance has led to accumulation of huge amount of knowledge about the biological reactions and regulations, facilitating better predictability of relative molecular changes. However, since the changes are happened at the whole genome level, and the complexity of biological reactions and regulations are still beyond full comprehension, advanced analytical tools that can scan overall molecular changes in a system level of a cell are required. During the last decade, omics techniques have been developed for genome-wide analysis (Bro & Nielsen, 2004, Herrgard, et al., 2008, Petranovic & Vemuri, 2009, Snyder & Gallagher, 2009), and omics approaches established a new field in life science, so called Systems Biology that aims to understand a cell in an holistic view by using high-throughput omics data and mathematical models. Systems Biology has been implemented by quantifying each level of molecules through whole-genome sequencing, transcriptome, proteome, metabolome. Their usefulness and limitation especially for the characterization of evolved stains has been recently reviewed (Oud, et al., 2012).

1.3.1. Transcriptome analysis for the characterization of evolved strains

Transcriptome analysis has been routinely used in the last decade because of standardization of techniques and data with the support of bioinformatics and models (Bro, et al., 2005, Patil & Nielsen, 2005, Bengtsson, et al., 2008, Reimand, et al., 2011). Not only technical maturation, but also it has the best coverage among other omics tools (Herrgard, et al., 2008, Reimand, et al., 2011). The effect of environmental or genetic perturbation can be checked easily by counting the number of significantly changed genes. Identified differentially expressed genes between evolved mutants and a reference strain are routinely analyzed to find altered pathways, metabolisms, and

regulation circuits based on several gene enrichment methods. Therefore, the transcriptome analysis can enumerate all possible transcriptional changes that are related to obtained phenotypes. Although there are several restrictions such as mixing of transcriptional changes between cause and consequence or the desired phenotype related and the experimental condition related and so on, transcriptome analysis is essentially useful as a first scan of molecular changes in evolved mutants. Additionally, comparison of multi strains or combination with other omics data has identified key molecular changes in different mutants (Ideker, et al., 2001, Bro, et al., 2005, Bengtsson, et al., 2008, Vijayendran, et al., 2008, Hazelwood, et al., 2009).

1.3.2. Metabolome analysis for the characterization of evolved strains

Metabolites play important roles as intermediates of biochemical reactions, which means their concentration is a key factor for controlling the reaction rate and they further are involved in regulation of the metabolic network through allosteric regulation. Thus, the level of metabolites represents integrative information of the cellular function; they can give critical clues to define the phenotype in evolved mutants (Zaldivar, et al., 2002, Kummel, et al., 2010). However, since metabolites have very diverse molecular kinds, it is almost impossible to analyze and quantify all metabolites in a cell simultaneously unlike the transcriptome. Practically targeted metabolome that analyze and quantify selected metabolites therefore has been more frequently used than metabolite profiling that tries to increase the number of covering metabolites. Targeted metabolomics can get clues from transcriptome data in selected metabolites of interest; and these metabolites data can be used to provide additional proof about the link between a desired phenotype and molecular changes.

1.3.3. Whole-genome sequencing for the characterization of evolved strains

A genetic change is the first and direct origin of a phenotypic change. Other molecular alterations are reflections of the genetic change. Therefore identification of driving genetic changes is crucial in inverse metabolic engineering. The importance of the identification of genetic changes was mentioned by Bailey et al. in 1996, The power of the technology for deciphering the genetic basis for a given phenotype is a critical determinant of the feasibility of inverse metabolic engineering (Bailey, et al., 2002). At that time the main limitation was in the technical part, since whole genome sequencing was time consuming and had a high cost.

However, with next-generation sequencing there has been a revolution in genome sequencing technologies and this has reduced the costs many folds (Herring, et al., 2006, Mardis, 2008, Shendure & Ji, 2008, Le Crom, et al., 2009, MacLean, et al., 2009, Metzker, 2010, Oud, et al., 2012). These techniques show the possibility of substantial reduction of the time and cost of genome sequencing such that it can be used for routine application similar to transcriptome analysis. As an example, the results of sequencing three yeast evolved mutants that were used in this thesis are explained in Table 3 and 4.

Table 1-3. The price of the next-generation sequencing (Illumina/Solexa, at 13th January 2010)

Description	Quantity	Unit Price (€)	Total Price (€)
Sample preparation for Genome Analysis, Genomic Shotgun	3 (strains)	544	1632
Sample preparation with bar-coded adapters	3	34	102
Sequencing on the Genome Analyzer GAIx, 1 paired-ends channel 2x38 bp	1	3,808	3,808
Additional bar-coded sample in the same channel, paired-ends	2	136	272
Bioinformatics analyses	0	340	0
		Total	5,814

Table 1-4. The performance of overall genome sequencing results

Sequencing Parameters	Mutant A	Mutant B	Mutant C
No. of Reads	5,605,504	18,203,846	5,239,106
Total Bases (bp) *	213,009,152	691,746,148	199,086,028
Coverage Fold	17	55	16
Undetermined Base	158,723	86,791	171,362
Genome percent reference coverage (%) [†]	98.7	99.3	98.6
No. of supercontigs	17	17	17
Chromosomes	16	16	16
Mitochondria	1	1	1

* 38 bases per sequence read for 2 cycles

[†] Based on genome consensus sequence length of CEN.PK113-7D of 12,155,742 base pairs

In addition, whole genome sequencing of evolved mutants can give genetic proofs for evolutionary theories or related questions as Dettman et al. mentioned in 2012, *How many mutations underlie adaptive evolution, and how are they distributed across the genome and through time? Are there general rules or principles governing which genes contribute to adaptation, and are certain kinds of genes (e.g. regulatory vs. structural) more likely to be targets than others? How common is epistasis among adaptive mutations, and what, if anything, does this reveal about the variety of genetic routes to adaptation? How common is parallel evolution, where the same mutations evolve repeatedly and independently in response to similar selective pressures?* (Dettman, et al., 2012) Phenotypic results of mutations are constrained by evolutionary genetic context such as epistasis, pleiotropy, hitch-hiking of negative mutation with beneficial ones, and so on. Therefore, the whole genome sequencing can give vast amount of possibility for increasing our understanding about evolution itself, and prediction our ability to use evolutionary strategies in setting the further design in strain development.

1.4. Improving galactose utilization in *Saccharomyces cerevisiae*

Galactose metabolism in *S. cerevisiae* was selected to generate evolutionary strategies and explore them through genome-scale analyses in this study. The galactose regulon of *S. cerevisiae* has been extensively investigated, since it has very strict gene expression control properties and it is a model system for human disease, galactosemia (Lai, *et al.*, 2009). In addition, yeast strains retaining higher galactose utilization ability have been developed for industrial application, because galactose is one of the abundant renewable carbon sources (Panesar, *et al.*, 2007, Wi, *et al.*, 2009, Kim, *et al.*, 2012). In previous studies, direct genetic engineering approach in galactose metabolism showed successful results (Ostergaard, *et al.*, 2000, Bro, *et al.*, 2005, Garcia Sanchez, *et al.*, 2010, Lee, *et al.*, 2011). For the next turn in the metabolic engineering cycle, new strategies are required.

1.4.1. Galactose metabolism in *Saccharomyces cerevisiae*

1.4.1.1. Leloir Pathway

Even though the molecular structure of galactose is very similar with glucose, more enzymatic reactions for galactose utilization are needed to reach glucose-6-phosphate, a precursor of glycolysis. And the number of transporter specialized for galactose is just one, while there are at least 20 transporters for glucose (Boles & Hollenberg, 1997, Ozcan & Johnston, 1999, Wieczorke, *et al.*, 1999). Galactose is metabolized through the Leloir pathway, after the Nobel Prize laureate, biochemist Louis Leloir (Cabib, 1970). This pathway is composed of 5 enzymes: galactose mutarotase (*GAL10*), galactokinase (*GAL1*), galactose-1-phosphate uridylyltransferase (*GAL7*), UDP-galactose 4-epimerase (*GAL10*) and phosphoglucomutase (*PGM1/PGM2*), and expression of those enzymes is controlled by very tight regulatory system consisting of 3 regulators, Gal3p, Gal4p and Gal80p (Timson, 2007). Further regulation is mediated by Mig1p, i.e. glucose is present, Mig1p is de-phosphorylated resulting in its transfer into the nucleus where it inhibits expression of *GAL1* and *GAL4* by binding to upstream repression site (URS) of those genes (Timson, 2007) (Fig. 2). Galactose enters the cells mainly through Gal2p, a specific galactose transporter. Intracellular galactose is structurally changed to alpha-D-galactose from beta-D-galactose by Gal10p, and phosphorylated to galactose-1-phosphate with ATP by galactose kinase, Gal1p (Holden, *et al.*, 2003). Galactose-1-phosphate reacts with UDP-glucose resulting in

glucose-1-phosphate and UDP-galactose in a reaction catalyzed by galactose-1-phosphate uridylyltransferase, Gal7p. This reaction has been studied more intensively because of its relation with the human disease galactosemia. Failure of this reaction accumulates galactose-1-phosphate, which is a marker for diagnostic of the disease (Lai, *et al.*, 2009). The toxicity of high concentration of galactose-1-phosphate is not clear, while there have been several proposes about the reasons such as inhibition of enzymes and ATP drain (Lai, *et al.*, 2009). UDP-galactose is converted into UDP-glucose by a reaction of UDP-galactose 4-epimerase, Gal10p, which is also galactose mutarotase. This enzyme has dual activity, which is a unique feature of *S. cerevisiae* and *Kluyveromyces fragilis* (Thoden & Holden, 2005). Prokaryotes and higher eukaryotes have different enzymes to provide these two enzyme activities (Holden, *et al.*, 2003). Since the Leloir pathway is a highly conserved system in most organisms, and yeast supposedly occupies a position between prokaryotes and higher eukaryotes, this distinctive evolutionary history is an open question. In the last step, glucose-1-phosphate is converted into glucose-6-phosphate in a reaction of isomerization by Pgm1p and Pgm2p. Pgm2p is responsible for about 80% of the total activity (Timson, 2007). Further detailed knowledge about the kinetic properties and structures of the enzymes in the Leloir pathway are well explained in biochemistry references (Daugherty, *et al.*, 1975, Schell & Wilson, 1977, Segawa & Fukasawa, 1979, Fukasawa, *et al.*, 1980, Reifenberger, *et al.*, 1997, Holden, *et al.*, 2003).

1.4.1.2. Regulation of GAL genes

Regulation of *GAL* genes is an excellent model for studying a regulated eukaryal gene expression system (Acar, *et al.*, 2005, Ramsey, *et al.*, 2006, Pannala, *et al.*, 2010). The promoter of the *GAL* genes has been used as a strong expression system with galactose induction (Li, *et al.*, 2008). Each of the galactose catabolism enzymes Gal1p, Gal7p and Gal10p exist at about 0.3 to 1.5% of total soluble cytoplasmic protein during growth on galactose (St John & Davis, 1981). There are three regulation mechanisms. First, the presence of glucose represses expression of the *GAL* genes through the transcription factor Mig1p (Timson, 2007) (Fig. 2). The Mig1p interacts with the transcriptional co-repressor complex Cyc8p (Ssn6p)-Tup1p (Treitel & Carlson, 1995). The complex of these three proteins activates the histone deacetylases Hda1p, Hos1p, Hos2p and Rpd3p (Davie, *et al.*, 2003), which ensures keeping the chromatin deacetylated, compact, and hereby in a transcriptionally inactive state. Second, at high glucose concentrations, the Mig1p is

dephosphorylated and stays in the nucleus where it together with the co-repressors block expression of the *GAL* genes, especially *GAL4* that is an activator of the *GAL* genes. Therefore the presence of glucose completely blocks expression of the *GAL* genes.

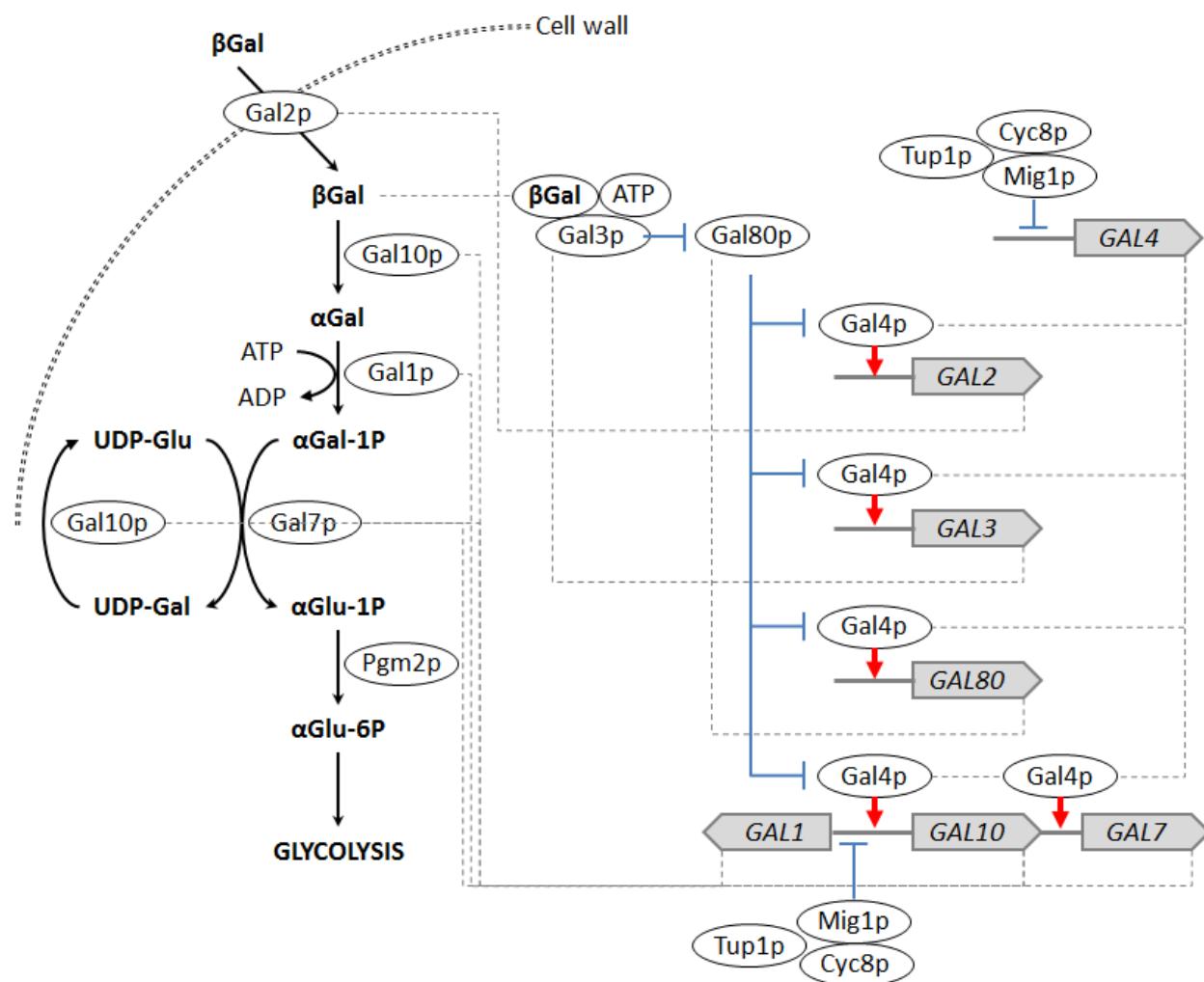


Fig. 1-2. Galactose pathway (Leloir pathway) and regulation in *S. cerevisiae*. Pointed arrows mean conversion of intracellular metabolites by enzymatic reactions, red arrows indicate transcriptional activation. Blue blunt arrows mean inhibition. Dotted lines indicate the direct connection between genes and proteins.

The absence of glucose, however, is not adequate to induce the galactose metabolizing enzymes. The existence of galactose is also necessary (Fig. 2). When external galactose is present as a sole carbon source, it can transfer at low rate to the cell through hexose transporters (Hxts), which are not specific to galactose as they have very high Km values for galactose transport. Galactose transporter gene, *GAL2* can be expressed after galactose enters to the cells. Intracellular galactose is combined with Gal3p that is a sensor of galactose. Gal3p binds galactose and ATP, and then traps the repressor Gal80p. Gal80p is present in the cytoplasm and nucleus, which interferes proper binding of Gal4p to the upstream activating sequences (UAS_{GAL}) of the *GAL* genes. Thus, in the absence of galactose, the Gal80p blocks Gal4p and hereby prevents induction of the *GAL* gene expression. When only the complex of Gal3p with galactose and ATP is present, the blocking of the Gal80p binding to Gal4p is released since this complex catches Gal80p (Yano & Fukasawa, 1997). Dual feedback loops have been well elucidated in the galactose control system (Ramsey, *et al.*, 2006). The Gal4p induces not only the Gal2p and Gal3p, but also Gal80p. Induction of Gal2p and Gal3p is positive feed-back loop because the increased expression of Gal2p and Gal3p result in further activation of Gal4p, while induction of Gal80p provides a negative feed-back loop since higher expression of Gal80p blocks Gal4p activation (Fig. 2). Simultaneous operation of these dual opposite controls has provided an excellent model for studies of the dynamics of gene expression regulation in eukaryotes. Third, Lap3p/Gal6p is supposed as a possible regulator, because deletion of this gene increases expression of the *GAL* genes (2.5 fold) (Zheng, *et al.*, 1997). The Lap3p is a cysteine protease and the *S. cerevisiae* homologue of this enzyme is Gal6p. The exact mechanism of how Gal6p carries negative regulation remains unclear (Zheng, *et al.*, 1997).

1.4.2. Galactose as a feedstock in industrial biotechnology

1.4.2.1. Galactose content of biomass

In terms of its use as a carbon and energy source for production of fuels and chemicals galactose is mostly found in cheese whey, but with the prospect of using algae as source of biomass it is interesting to note that red seaweed has a high content of galactose (*Gelidium amansii*) (Wi, *et al.*, 2009, Kim, *et al.*, 2012). Cheese whey is an eluent from the dairy industry; and it contains about 85-95% of the milk volume and 55% of milk nutrients. Two types of the cheese whey, sweet (pH 6~7) and acid (pH < 5) are produced dependent on the procedure of

casein precipitation. The main components are lactose, whey protein and minerals (Table 5) (Jelen, 1979, Siso, 1996, Panesar, *et al.*, 2007). Lactose is a disaccharide sugar composed of galactose and glucose through beta-1, 6-linkage. Galactose therefore contains around 22~26 g/l in cheese whey.

Table 1-5. Typical composition of sweet and acid whey (Source: Jelen, 1979, Panesar, *et al.*, 2007)

Components	Sweet whey (g/l)	Acid whey (g/l)
Total solids	63-70	63-70
Lactose	45-52	44-46
Protein	6-10	6-8
Calcium	0.4-0.6	1.2-1.6
Phosphate	1-3	2-4.5
Lactate	2	6.4
Chloride	1.1	1.1

Recently algae have been considered as an attractive biomass source for bio-based products, due to several advantages compared to terrestrial plant biomass such as high production yield, non-food and land usage, little recalcitrant lignin and crystalline cellulose, a higher growth rate and others (Kim, *et al.*, 2011, Wargacki, *et al.*, 2012). One algae, the red seaweed (*Gelidium amansii*) has high galactose content even comparable to the amount of glucose (Wi, *et al.*, 2009, Kim, *et al.*, 2012). Carbohydrate compositions of different biomass sources are given in Fig. 3.

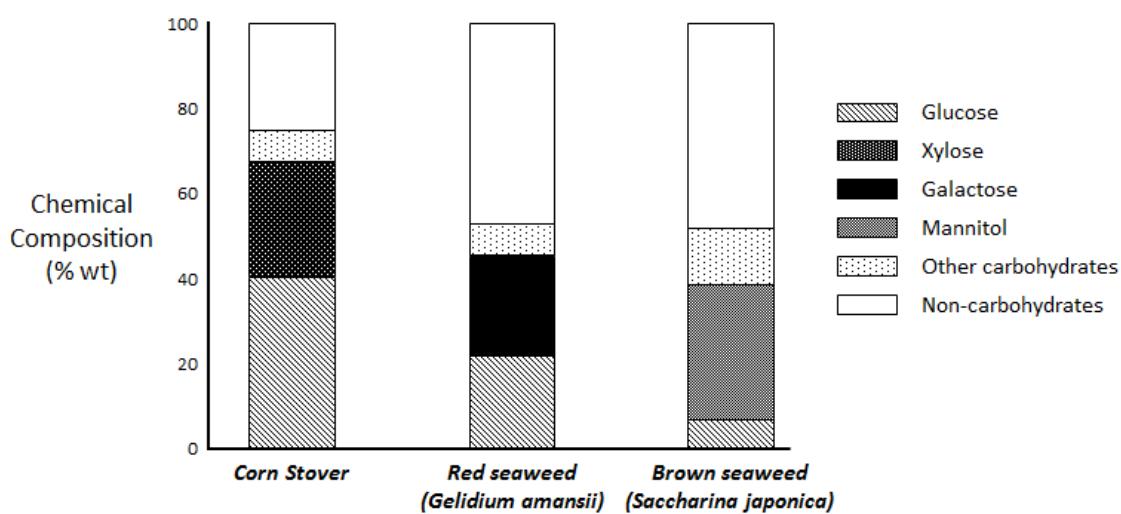


Fig. 1-3. Composition of non-food biomass (Source: Kim, *et al.*, 2012)

1.4.2.2. Metabolic engineering for improved galactose utilization

Intensive research on galactose metabolism has generated vast amount of information about its metabolic structure and regulation. By exploiting this abundant resource, many elegant metabolic engineering approaches have been executed to improve galactose utilization in terms of a specific galactose uptake rate. Direct genetic modification of *GAL* genes and regulatory genes were done. Over-expression of *GAL* catabolic genes was implemented in a high-copy number plasmid with different combinations (de Jongh, *et al.*, 2008). Over-expression of the *GAL* catabolic genes was expected to increase flux from galactose to glycolysis; however, the result showed reduction of galactose uptake and growth rate. The reason was that changed expression level of the *GAL* genes triggered the fluctuation of concentration of intermediate metabolites in the Leloir pathway. One of them, galactose-1-phosphate, was known as a toxic intracellular metabolite that interfere with galactose metabolism. Therefore genetic modification was focused on regulatory genes. By over-expression of the transcriptional activator, *GAL4* and deletion of negative regulators, *GAL80*, *MIG1* and *GAL6* showed improved galactose uptake rate without growth retardation (Ostergaard, *et al.*, 2000). Especially, the triple knock-out mutant (SO16, which was used as a control strain in this thesis study) showed the highest specific galactose uptake rate. In a follow up study, transcriptome data of these strains was used to find target genes that were related to improvement of galactose availability (Bro, *et al.*, 2005). Commonly changed genes were screened in galactose related pathways and based on this *PGM2*, encoding phosphoglucomutase, was the only gene that showed significant up-regulation. Application of this gene in a high-copy number plasmid clearly showed improvement of galactose utilization (this strain was called PGM2, which was used as another control strain in this study). Since this gene was supposed to be quite highly expressed even at non-galactose growth condition, the result that the rate limiting step enzyme of the Leloir pathway was *PGM2* was surprising. When galactose-1-phosphate was measured, this strain showed no reduction of this metabolite. Higher activity of phosphoglucomutase was checked by checking higher concentration of sugar-6-phosphates that were considered as products of this enzyme such as galactose-6-phosphate, glucose-6-phosphate, mannose-6-phosphate and fructose-6-phosphate. Therefore, even though basal expression level of *PGM2* was relatively higher than other *GAL* genes in the wild-type, over-expression of this gene was still needed to improve galactose utilization. Another study also supported the importance of higher activity of *PGM2* for improved galactose utilization (Lee, *et*

al., 2011). In this study, a genomic library was used to find target genes that were related to galactose utilization. The constructed library was transformed into the wild-type strain, and improved strains were screened. Three beneficial over-expression targets, *SEC3*, *tTUP1*, and *SNR84* were identified. All three targets displayed higher phosphoglucomutase activity. Two of them, Sec3p (phosphomannomutase having activity as phosphoglucomutase) and truncated Tup1p (complex of Mig1p repressor) were confirmative with the previous work due to the function of those genes; while the last target was a new discovery. *SNR84* codes for H/ACA box small nucleolar RNA, and there is no report on the effect of this gene to galactose metabolism. However, higher activity of phosphoglucomutase in the transformant over-expressing *SNR84* proposed a relationship between this gene and galactose metabolism.

2. OVERVIEW OF THE THESIS

The motivation of this study is to apply genome-scale analyses for unraveling the molecular and genetic basis of evolutionary strategy of microorganism.

The galactose metabolism of yeast *S. cerevisiae* was chosen as a target for evolution, because of the following reasons.

- 1) It has relevance for developing yeast strains that can use galactose more efficiently like other hexose carbon sources; since galactose is an abundant sugar in some renewable resources, and *S. cerevisiae* is a vastly useful strain in industrial applications.
- 2) The galactose metabolism in yeast has been extensively studied, which has led to many trials for the construction of yeast mutant strains by direct genetic engineering. Consequently, several genetic targets related to the improvement of galactose utilization have already been identified, which means it is difficult to find new targets. However, rather less attention was paid on evolutionary engineering, thus if different targets are generated from an evolutionary approach, they are useful for a next round of strain development.
- 3) Moreover, which was the main purpose of this thesis; can genome-scale analyses be used for the characterization of evolved mutants with the objective to find driving mutations? If this answered positively it could open up for wider use of evolutionary strategies in metabolic engineering.

For these reasons, yeast *S. cerevisiae* CEN.PK113-7D was evolved on galactose minimal media through adaptive evolution for 62 days. Three evolved mutants were generated from independent populations grown in identical serial transfers. Improved galactose utilization ability was confirmed in precisely controlled bioreactors, and genome-scale analyses through transcriptome, metabolome and whole-genome analyses were applied to understand evolutionary strategies of the galactose-evolved yeast mutants. Furthermore, inverse metabolic engineering was applied using identified mutations and new combinations of the genetic changes. The comparison of reconstructed strains with the evolved mutants provided a good example how evolution and engineering work synergistically in strain development. Further characterization of the evolved mutants was done in glucose minimal media to explore the pleiotropy of obtained

traits. Molecular and genetic bases of that pleiotropy were elucidated by genome-scale analyses. The result increased the understanding of evolutionary strategies of the evolved mutants. Consequently, three research studies were designed.

Paper I: *Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis*

Adaptive evolution generated improved galactose availability with different physiology. The molecular and genetic bases that were supposed to be related to improved galactose utilization were analyzed. The significant molecular changes in transcripts and metabolites were detected in around galactose metabolic pathways, but no mutations were found in those regions. Instead the Ras/PKA signaling pathway was detected as a common pathway that had mutations in all the evolved mutants. Introduction of one of those mutations in a reference strain partially provided the genetic bases of the galactose evolved physiology. It was confirmed that adaptive evolution can generate key mutations in unpredictable regions or non-canonical pathways. And the genetic basis (mutations) and resulting molecular basis (transcriptome and metabolome) for evolutionary changes were found to happen in different regions.

Paper II: *Recovery of phenotypes obtained by adaptive evolution through inverse metabolic engineering*

Through adaptive evolution and genome-scale analyses, new genetic targets for improving galactose utilization were identified. As only a few mutations were selected from many mutations, it was necessary to evaluate whether the adaptive phenotype can be recovered by a few mutations. Furthermore, it was speculated how inverse metabolic engineering could give more chances beyond evolutionary engineering itself for strain development. Two groups of engineered mutants were constructed; site-directed mutants that had the identified mutations in the reference strain genetic background, and combined mutants that had new combinations by transforming the *PGM2* over-expression plasmid into the site-directed mutants. Surprisingly, some of the constructed strains showed complete recovery of the galactose adaptive phenotype with just one

or two genetic modifications. Even one of the reconstructed mutants exhibited further improved galactose utilization. These results indicated that far fewer genetic changes were enough to reach the same phenotype as the evolved mutants. Therefore inverse metabolic engineering is an essential step in the application of evolutionary approaches for strain development, i.e. it can enable more strategies for further improvement of desired phenotypes by sieving out beneficial mutations from negative ones and generated new artificial combinations of mutations. Detailed molecular changes by the mutations were also analyzed using transcriptome analysis and the level of a few metabolites. The introduction of key mutations that recovered the adaptive phenotype triggered fewer molecular changes compared to the evolved mutants. This result indicated again that all molecular changes were not necessary for reaching the same phenotype.

Paper III: Adaptively evolved yeast mutants on galactose show trade-offs in carbon utilization on glucose

The evolved mutants obtained the trait that was an ability to utilize galactose more efficiently than the ancestor strain. It was, however, interesting to evaluate whether this trait was associated with other effects, i.e. pleiotropy. The galactose-evolved mutants were therefore grown in glucose minimal media. Interestingly, these cultivations showed reduced glucose utilization in the evolved strains. This means that there is trade-off in galactose utilization and glucose utilization. In other words, the evolved mutants likely obtained the increased galactose availability by partly losing their ability to very efficiently utilize glucose. The underlying mechanisms of this trade-off were studied at the molecular and genetic level by integrated genome-scale analyses. Antagonistic pleiotropy was found to be the dominant evolutionary trade-off mechanism. The tight regulation system of glucose catabolic repression was loosened by the mutations in Ras/PKA signaling pathway and unidentified mutations that may be involved in hexokinase regulation and reserve carbohydrates metabolism. Therefore, the glucose utilization ability is likely collateral cost for having improved galactose availability in the evolved mutants. This finding indicates that genetic context such as pleiotropy causing trade-off in traits should be considered, when evolutionary approaches is applied in strain development.

3. RESULTS AND DISCUSSION

This section provides a summary of the results, whereas the attached papers in the end of this thesis include detailed materials, methods and experimental design with expanded explanations.

3.1. Molecular and genetic basis of evolutionary strategies of the galactose-evolved mutants (Paper I)

Three evolved populations were generated from *S. cerevisiae* CEN.PK113-7D, an ancestor strain, by three independent serial transfers in a galactose (20g/l) minimal medium for 62 days (Fig. 3-1). Single clone isolates were obtained from the last shake flasks, and designated 62A, 62B and 62C. Two engineered strains, SO16 (*gal6Δ gal80Δ mig1Δ*) and PGM2 (over-expression of *PGM2* gene), showed improved galactose uptake rates in previous studies were used as control strains to elucidate unique strategies of adaptive evolution (Ostergaard, *et al.*, 2000, Bro, *et al.*, 2005).

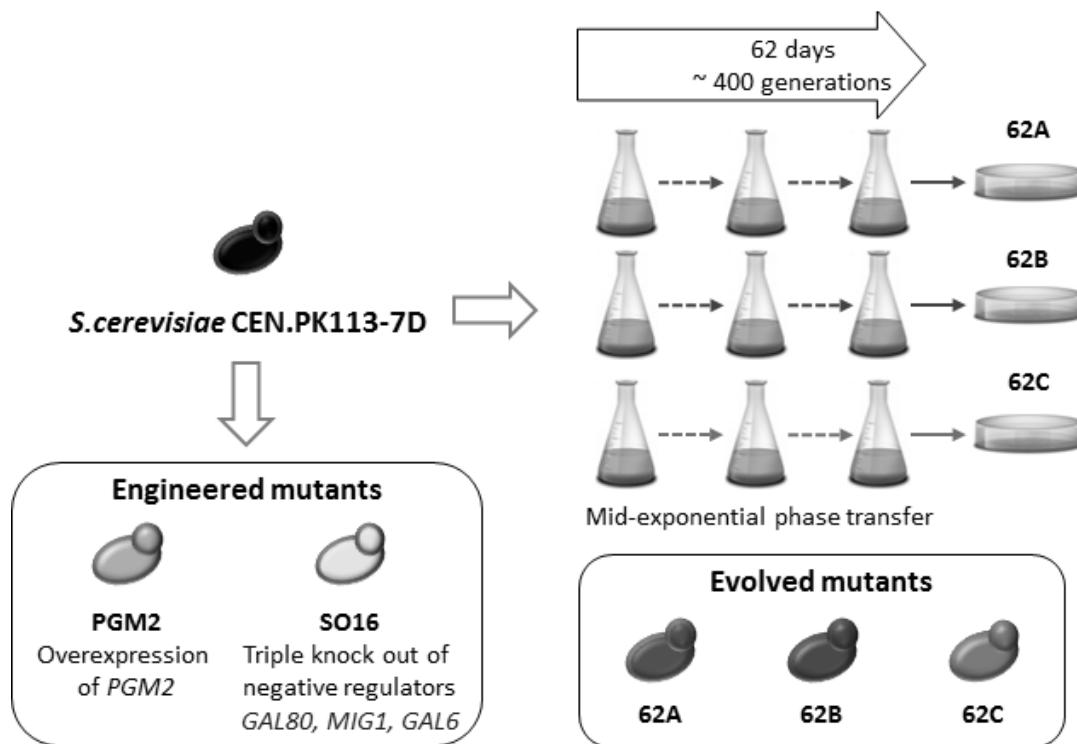


Fig. 3-1. Yeast strains used in this study. Engineered mutants were constructed in previous studies (Ostergaard, *et al.*, 2000, Bro, *et al.*, 2005) whereas the three evolved mutants were generated in this study.

Parallel fermentations at aerobic batch mode in precisely controlled bioreactors was performed to estimate physiological parameters and to take samples for omics analyses; and transcriptome and targeted metabolome analysis were applied to all strains including the two engineered strains; and whole-genome analysis were performed on the evolved mutants (Fig. 3-2).

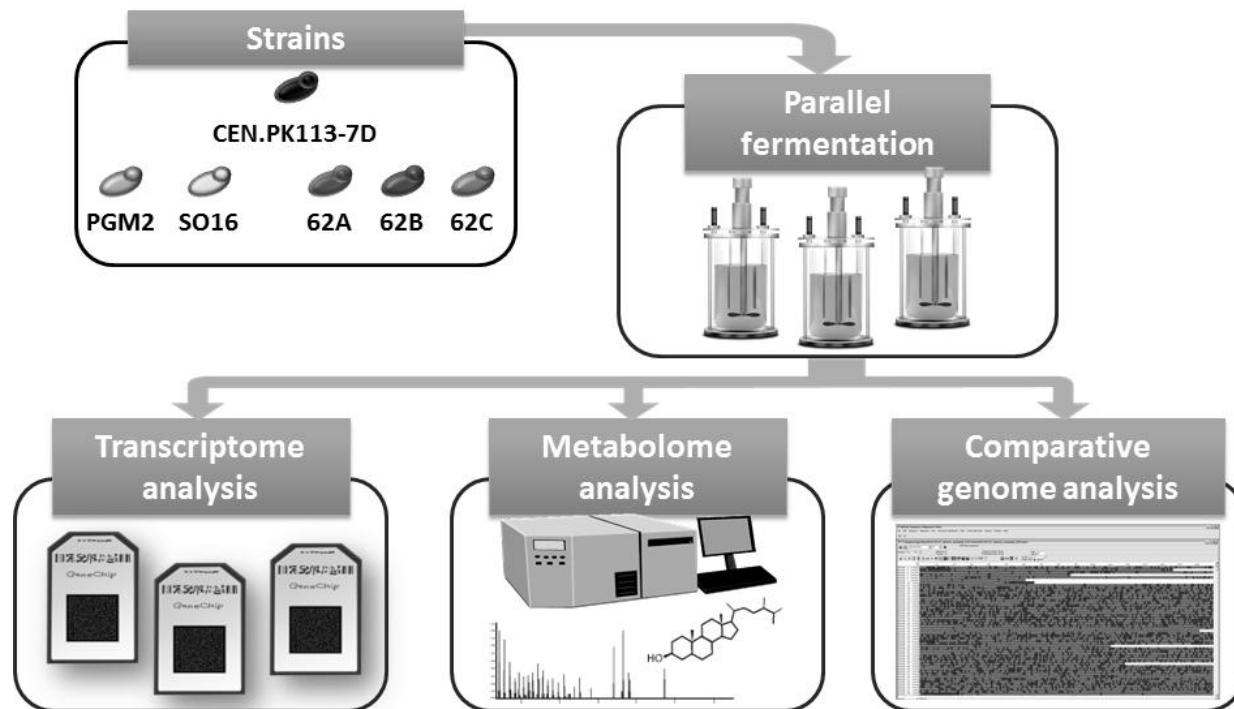


Fig. 3-2. Overall flow of the experiments in this study. 6 strains were cultivated in bioreactors, and at mid-exponential phase, samples for omics analyses were collected.

Exact identification of improved galactose utilization in the evolved mutants was compulsory, since only one colony from each of the populations were selected. Population was composed of many diverse individuals, thus it was not sure if the selected colony was really evolved in terms of improved galactose utilization. Of course, based on Darwin's theory, natural preservation, the variants that had higher fitness would take more portions in the population; therefore, there was high chance to select evolved clones with improved fitness. The purpose of this study was to detect evolutionary strategies; hence the confirmation of improved phenotype in the evolved mutants was a prerequisite. The evolved mutants achieved improved galactose

utilization in terms of a maximum specific growth rate and a specific galactose uptake rate, which were a different phenotype compared with the engineered strains (Fig. 3-3). The galactose-evolved mutants showed a 24% increase in the maximum specific growth rate and 18 ~ 36% increase in the specific galactose uptake rate compared to the reference strain. Interestingly, clear grouping was observed between all three evolved mutants and the reference strains in the plot of the specific galactose uptake rate versus a specific ethanol production rate (Fig. 3-3B). These two groups were separated by different regression curves, which observation means that the adaptive evolution has resulted in different phenotypes compared with the engineered strains.

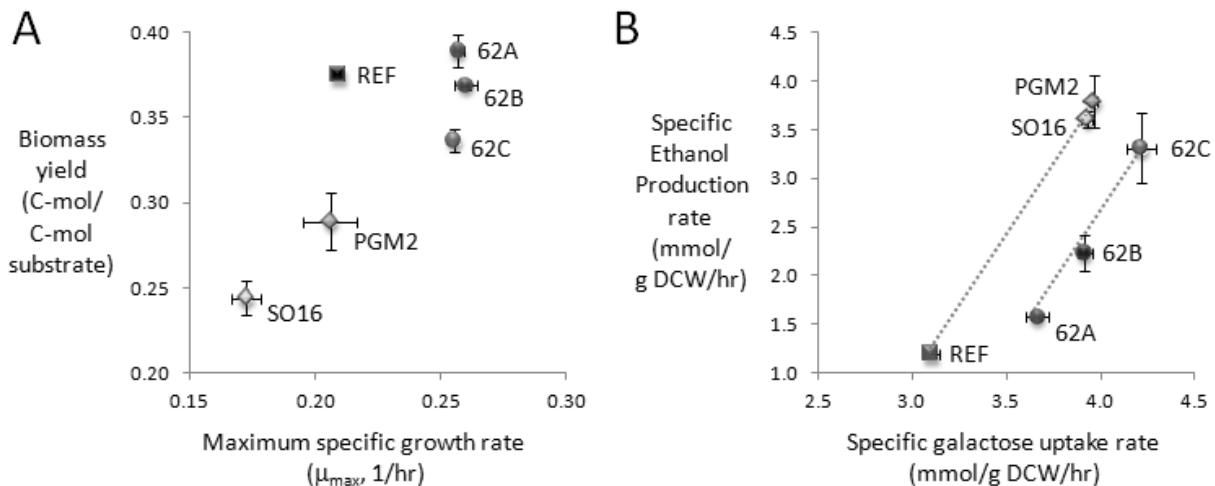


Fig. 3-3. Phenotypic changes of evolved mutant strains 62A, 62B and 62C compared with the reference strain CEN.PK113-7D and the two engineered strains SO16 and PGM2. (A) Correlation between a maximum specific a growth rate and biomass yield. (B) Correlation between a specific galactose uptake rate and a specific ethanol production rate. The regression curves of the two lines (from right to left) had a slope of 2.95 and 3.19 and intercept of minus 7.95 ($R^2 = 0.99$) and minus 10.157 ($R^2 = 0.98$), respectively. Both slope values were around 3, which indicated that catabolic repression induced flux re-direction from respiratory metabolism to fermentation one, because if there was not that repression, slope should be around 2.

To investigate the molecular basis, firstly transcriptome analysis was used to check overall changes, and select significantly altered pathways (Paper I, Fig. 2). Secondly, target metabolome was implemented; around 40 metabolites were measured based on the results from the transcriptome data and quantified by diverse analytical instruments (Paper I, Fig. 3). Both data

sets were used to find the molecular basis for the evolutionary strategies by selecting commonly changed metabolisms in all evolved mutants (Fig. 3-4).

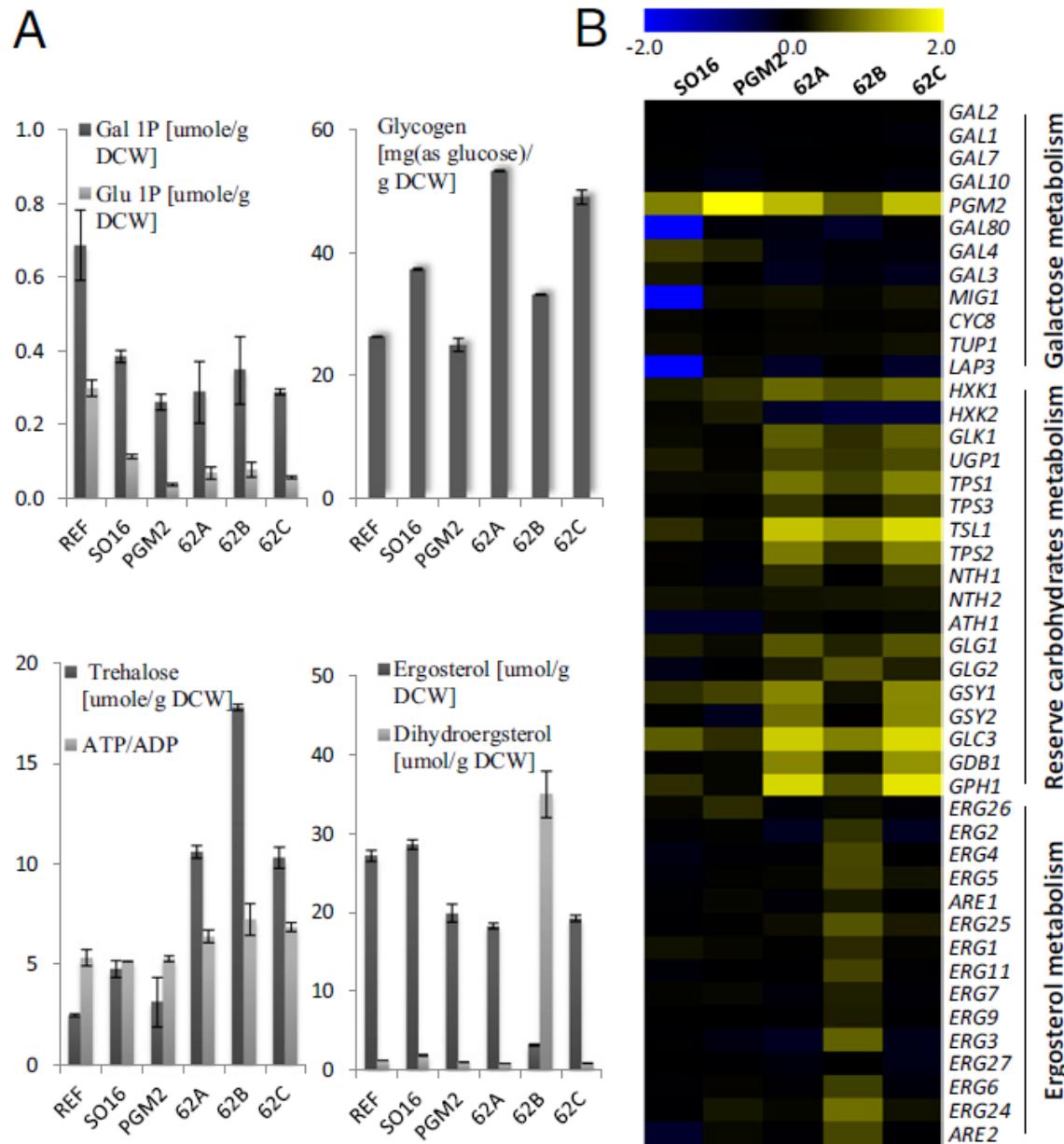


Fig. 3-4. Changes in the galactose, reserve carbohydrates, and ergosterol metabolism in the evolved mutants are illustrated by changes in the concentration of metabolites and fold changes of transcriptome compared with the other strains. (A) The concentrations of sugar phosphates, storage carbohydrates, and sterols and the ratio of ATP to ADP. (B) Fold changes of all genes

involved in galactose, reserve carbohydrates, and ergosterol metabolism are compared with the reference strain.

Up-regulation of the *PGM2* gene and lower concentration of galactose-1-phosphate and glucose-1-phosphate in the galactose pathway were common in all evolved mutants and engineered mutants compared to the reference strain; while up-regulation of genes in reserve carbohydrates metabolism and down-regulation of *HXK2* that was one of the main glucose catabolic repression controllers (Gancedo, 1998), were unique for the evolved mutants. A unique change among evolved mutants was found in ergosterol metabolism. The 62B strain only showed up-regulation of *ERG* genes with different ratio of the concentration of ergosterol and dihydroergosterol. In terms of fermentation physiology, there was big difference between the 62A and 62C strains; however, in terms of transcriptome and metabolome data, they looked almost identical. 62B was positioned between them in terms of gross physiology, whereas this evolved mutant showed vast differences in the transcriptome and the metabolome. The reason of differences among the evolved mutants was not clear, while the common changes of the evolved mutants from the reference strain likely explained the molecular bases of evolutionary strategies for improving galactose utilization.

To identify the genetic basis of evolutionary strategies, whole-genome sequencing was applied to the evolved mutants (explanation of overall process and raw data are in the supplementary data of Paper I). Surprisingly, there were no mutations or duplications in the *GAL* genes and the regulatory genes involved in galactose metabolism including their promoter and terminator regions. The reaction step by *PGM2* was earlier found as a rate-controlling step in galactose metabolism; hence *PGM2* over-expression was already proven as a beneficial target for metabolic engineering, and several genetic modifications that induced higher expression of this gene were also identified. However, mutations from previous studies were not detected in the evolved mutants. This result indicated that new mutations that induced up-regulation of *PGM2* were generated. Furthermore, genes of the reserve carbohydrates metabolism and hexokinases had no mutations, even though they showed significant alteration in their transcription. Exceptionally, the 62B evolved mutant had mutation in the *EGR5* gene that seemed to be related to changes in ergosterol metabolism (Kelly, *et al.*, 1995) (Table 3-1). No mutations in the metabolisms that showed common molecular changes in all the evolved mutants implied that the key mutations may be involved in regulatory regions. Common genes, pathway or cellular

metabolism that had mutation in all three evolved mutants was searched, and it was found that there were common mutations in the regulatory, Ras/PKA signaling pathways (Table 3-1). The Ras/PKA signaling pathway has been known to take key role in global regulation of glucose sensing and stress response (Estruch, 2000). And *PGM2* and *UGP1* had STER element in their promoter region. Therefore, the mutations in Ras/PKA signaling pathway were suggested as a driven mutation that increased galactose utilization by triggering the activation of galactose and reserve carbohydrates metabolism (Fig. 3-5). Introduction of mutations in *RAS2* genes into a reference strain clearly showed the increase of galactose utilization (supplementary data from Paper I and Paper II). The 62B unique mutation in the *EGR5* gene could explain the changes of transcripts and metabolites in the ergosterol pathway; it may also be explicable why this mutant showed large differences from other evolved mutants by this mutation. However it was not clear how there is a relationship between galactose metabolism and ergosterol pathway.

Table 3-1. Genetic changes in the evolved mutants

Strains	Mutations	Functions	Specific features
62A	<i>RAS2</i> [Gln ⁷⁷ → Lys]		
62B	<i>RAS2</i> [Asp ¹¹² → Tyr]	Ras/PKA signaling pathway	Commonly mutated pathway
62C	<i>CYR1</i> [Asp ⁸²² → Asn]		
62B	<i>ERG5</i> [Arg ³⁷⁰ → Pro]	Ergosterol metabolism	Uniquely mutated gene

In conclusion, key genetic changes were identified in non-canonical metabolism, but in the Ras/PKA signaling pathway; which meant no mutation detected in galactose metabolism not like other direct genetic engineering studies (Ostergaard, *et al.*, 2000, Bro, *et al.*, 2005, Lee, *et al.*, 2011). And, molecular changes were well related to canonical metabolisms; up-regulation of *PGM2* in galactose metabolism, and up-regulation of genes in reserve carbohydrates metabolism that shared the intermediate of galactose metabolic pathway, i.e. glucose-1-phosphate. Hypothetical evolutionary changes were plotted in Fig. 3-5. Therefore, insight about evolutionary strategy that results in non-canonical genetic changes with canonical molecular changes could be applied as an evolutionary approach in strain development.

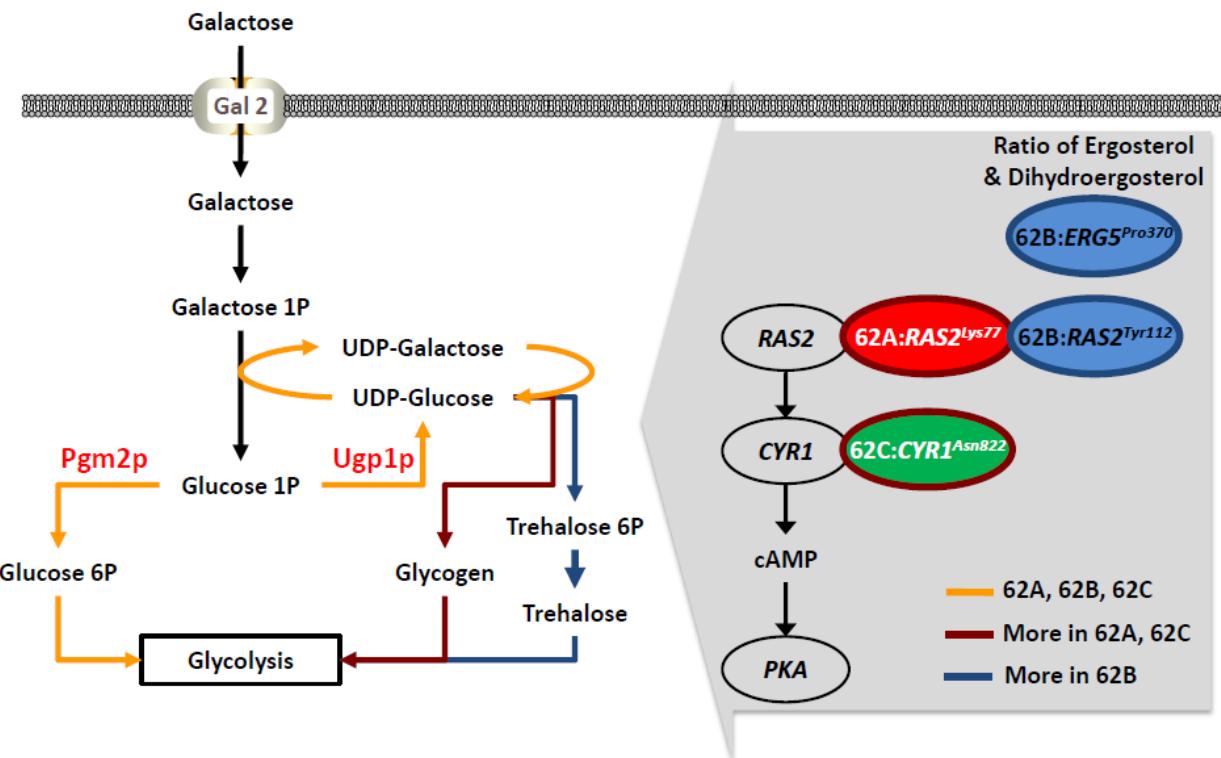


Fig. 3-5. Summary of evolution changes in the three evolved mutants; 62A, 62B, and 62C. Color circular boxes indicate genes having genetic mutations. Color lines indicate activated fluxes inferred from transcriptome and metabolome analysis.

3.2. Complete recovery of adaptive phenotype through inverse metabolic engineering (Paper II)

The initial purpose of this step was to evaluate the identified mutations. Furthermore, the objective was to explore how inverse metabolic engineering was useful in connection with evolutionary approaches for strain development. Site-directed mutants and combined mutants were constructed for the application of inverse metabolic engineering.

The genome-scale analyses suggested that the genetic basis for the improved galactose utilization could be present in the Ras/PKA signaling pathway, which is not directly involved in galactose metabolism, since all the evolved mutants commonly had mutations in this signaling pathway. One of the evolved mutants, 62B showed significant changes in ergosterol metabolism both at the level of transcripts and metabolites and it carried a mutation in the *ERG5* gene. To clearly confirm the effect of identified mutations on galactose availability, and to examine how much those mutations recovered the adaptive phenotype of the evolved mutants, site-directed mutants carrying each of the mutations independently were constructed. In addition, combined mutants were constructed by introduction of the known beneficial change (*PGM2* overexpression) into the site-directed mutants. These combined mutants were designed to generate new combination of the genetic basis for improving galactose utilization that was not present in the evolved mutants (Table 3-2).

The gross phenotype of the reconstructed strains was compared to the evolved strains (Fig. 3-6). The results of the site-directed mutants clearly confirmed the effects of the identified mutations on galactose utilization. Two site-directed mutants (RAU and RBU) that had mutations in the *RAS2* gene showed a significant increase in the maximum specific growth rate and the specific galactose uptake rate compared with their reference strain (5DU). Especially, the RAU strain that carried the mutation *RAS2*^{Lys 77} exhibited the highest specific galactose uptake rate among all the strains including the evolved mutants. Additionally, when its improvement of maximum specific growth rate was compared to the evolved mutants in terms of increased extent from each of their reference strains, i.e. RAU from 5DU, and the evolved mutants from 7D, even the RAU has a higher relative increase in the specific growth rate than the evolved mutants. Interestingly, even though two mutations were positioned in the same gene, their effect on galactose utilization was quite different. These results highlighted why the concept of inverse

metabolic engineering was important for strain development, because new targets from evolutionary engineering should be re-evaluated and there was space for more improvement of the desired traits by surveying more mutations in that target gene. The results of the combined mutants also showed improvement of the galactose availability. They almost fully recovered the adaptive phenotype of the evolved mutants, since the maximum specific growth rate and the specific galactose uptake rate were in the same level as the evolved mutants. This result again confirmed the importance of inverse metabolic engineering in connection with evolutionary approaches for strain development, because the same phenotype was realized with much fewer traceable genetic modifications providing more space for new engineering strategies.

Table 3-2. Reconstructed strains and control strains. *Saccharomyces cerevisiae* CEN.PK113-5D was used to construct site-directed mutants and combined mutants due to its availability of *URA3* marker gene. Prototrophic site-directed mutants (RAU, RBU and EBU) were constructed by transformation with the plasmid, *pSP-GM2* containing the *URA3* gene. The combined mutants RAP, RBP and EBP were constructed by transformation of the plasmid *pPGM2* into the site-directed mutants.

Strains	Ancestor strains and Genotype	Groups	References
7D	<i>MATa SUC2 MAL2-8^c</i> (CEN.PK113-7D)	Reference strain	SR&D*
62A	7D, total no. SNPs: 21 including <i>RAS2</i> ^{Lys 77}		
62B	7D, total no. SNPs: 104 including <i>RAS2</i> ^{Tyr112} , <i>ERG5</i> ^{Pro 370}	Evolved mutants	This study
62C	7D, total no. SNPs: 29 including <i>CYR1</i> ^{Asn822}		
5D	<i>MATa SUC2 MAL2-8^c ura3-52</i> (CEN.PK113-5D)		SR&D*
5DU	5D, <i>pSP-GM2(URA3)</i>	Reference strain	This study
RAU	5D, <i>pSP-GM2(URA3)</i> ; <i>RAS2</i> ^{Lys 77} (from 62A)		
RBU	5D, <i>pSP-GM2(URA3)</i> ; <i>RAS2</i> ^{Tyr112} (from 62B)	Site-directed mutants	This study
EBU	5D, <i>pSP-GM2(URA3)</i> ; <i>ERG5</i> ^{Pro 370} (from 62B)		
PGM2	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i>	Engineered mutant	Bro et al 2005
RAP	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i> ; <i>RAS2</i> ^{Lys 77}		
RBP	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i> ; <i>RAS2</i> ^{Tyr112}	Combined mutants	This study
EBP	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i> ; <i>ERG5</i> ^{Pro 370}		

*Scientific Research & Development GmbH, Oberursel, Germany.

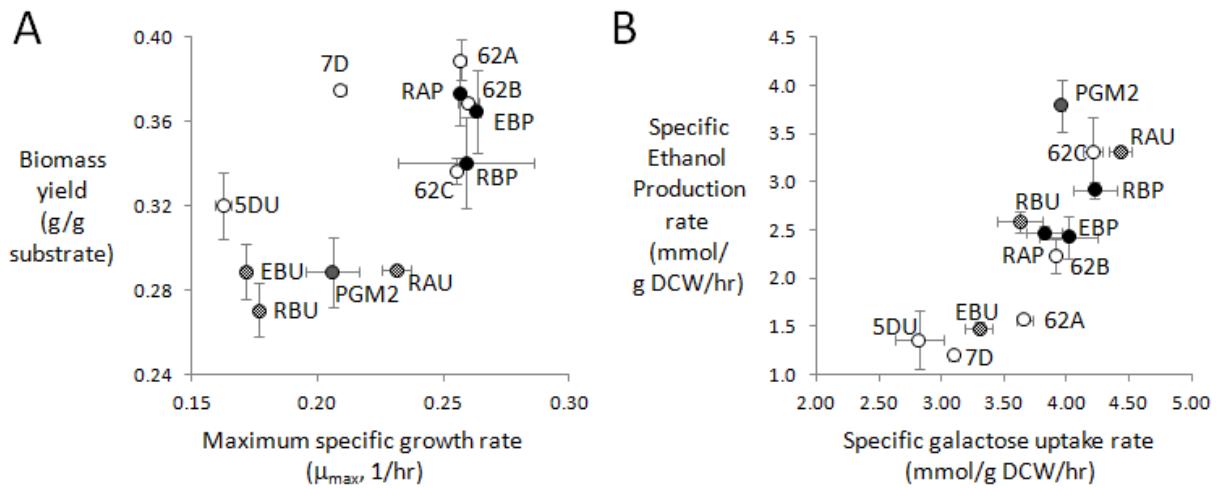


Fig. 3-6. Data on overall fermentation physiology of the site-directed mutants (RAU, RBU and EBU) and the combined mutants (RAP, RBP and EBP) are compared to the reference strains 5DU, 7D and the engineered strain PGM2, and the corresponding evolved mutants 62A and 62B. A: Correlation between the maximum specific growth rate and biomass yield B: Correlation between the specific galactose uptake rate and the specific ethanol production rate. Error bars represent standard deviation from biological duplicates.

The critical points were, 1) the evolved mutants accumulated many genetic changes that seemed to be not necessary for improving galactose utilization, because the reconstructed strains showed full recovery of the galactose adaptive phenotype with much fewer genetic changes; 2) the combination effect of genetic changes was different from the sum of each of the changes, for example, the RAP strain that contained a combination of RAU ($RAS2^{Lys~77}$) and *PGM2* overexpression, showed an increase of the maximum specific growth rate (58% from 5DU); however, the sum of each of genetic changes was bigger (69% = 42% (RAU) + 27% (PGM2 strain)). This phenomenon was much clearer in the specific galactose uptake rate; the combination (RAP) showed even reduced value compared with the RAU and PGM2 strains. Another combination case, the RBP strain with a combination of RBU and PGM2, also showed a negative synergy of beneficial genetic changes even though the extent of the physiological changes was different. On the contrast, the combination of *ERG5* mutation and *PGM2* over-expression looked like synergistic epistasis. The mutation in the *ERG5* gene showed only a small effect on galactose utilization. It looked almost neutral when it was solely present, while the

combination of this mutation with the over-expression of *PGM2* presented the same phenotype like the other combined mutants. In evolutionary biology, the accumulation of negative or neutral mutations and epistasis among mutations is a well-known event during adaptive evolution (Ikeda, *et al.*, 2006, Warner, *et al.*, 2009). That was one of the reasons why cells may not always reach to the optimum point of a specific trait by adaptive evolution, especially in asexual reproduction. Even for a versatile biological system, natural selection or preservation that could enrich only beneficial mutations would possibly require infinite generation time. Thus, to reach the optimum point by laboratorial adaptive evolution could be almost impossible (Sauer, 2001). Therefore, there is likely space for further improvement of desired traits by removing negative mutations and reconstruction of new combinations that may generate synergetic epistasis. Because of these reasons, inverse metabolic engineering is an essential step in evolutionary approaches for strain development.

The molecular basis of the reconstructed strains was investigated to clarify the relationship between the identified mutations and the molecular changes of transcripts and metabolites in specific pathways. First, the overall number of differentially expressed genes was compared (Fig. 3-7). Like the case of genetic changes, the reconstructed mutants showed a much smaller number of differentially expressed genes than the evolved mutants. It confirmed again that many changes in the evolved mutants were not necessary to reach the same phenotype. Second, the detailed molecular changes indicated that the mutations in the *RAS2* gene induced *PGM2*, but not reserve carbohydrates metabolism (Fig. 3-8). This result indicated that there were unidentified mutations triggering up-regulation of reserve carbohydrates metabolism. Maybe the up-regulation of this metabolism was not closely related to improving galactose utilization, or there would be negative epistasis between the mutations in Ras/PKA signaling pathway and the unidentified mutations that activated the reserve carbohydrates metabolism. Both cases could explain the recovery of the galactose adaptive phenotype by the mutations in the *RAS2* gene. Another finding is that the two mutations in the *RAS2* gene showed substantial difference in terms of molecular changes. The RAU strain (*RAS2^{Lys 77}*) showed much fewer numbers of transcriptional changes than RBU (*RAS2^{Tyr112}*), while the RAU exhibited higher improvement of galactose utilization than RBU (Fig. 3-7). This finding again emphasized the space for further improvement of galactose utilization by inverse metabolic engineering. The *ERG5* mutation was confirmed as a reason of the changes in the ergosterol pathway (Paper II).

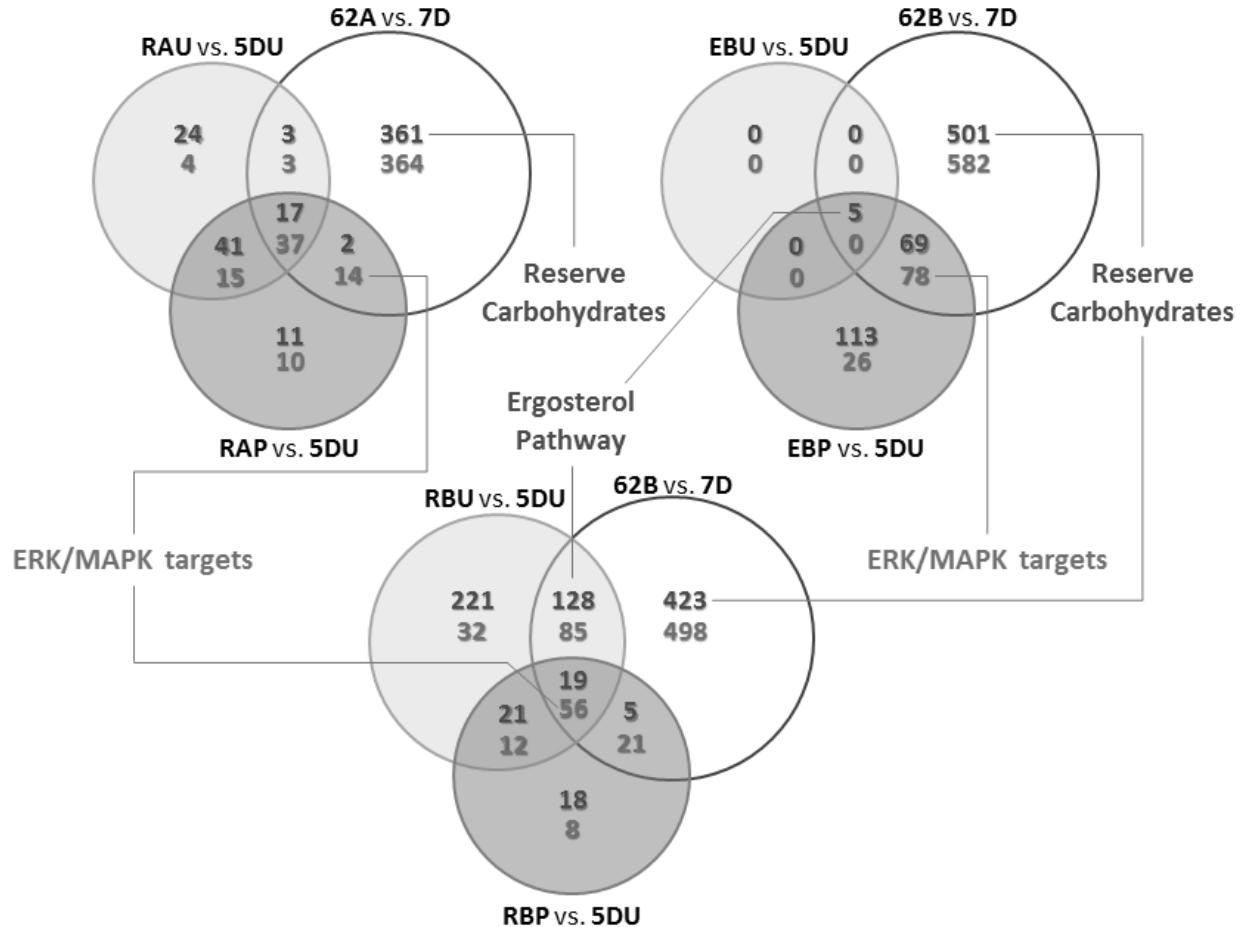


Fig. 3-7. Effect of reconstructed strains compared to the evolved strains 62A and 62B by differentially expressed genes. Differentially expressed genes (adjust p < 0.01) are categorized as Venn diagrams. The functional enrichment of genes in each part was analyzed by hypergeometric distribution based on the KEGG, Reactome and GO term databases. Upper numbers in a pair of two numbers mean up-regulation and lower number mean down-regulation.

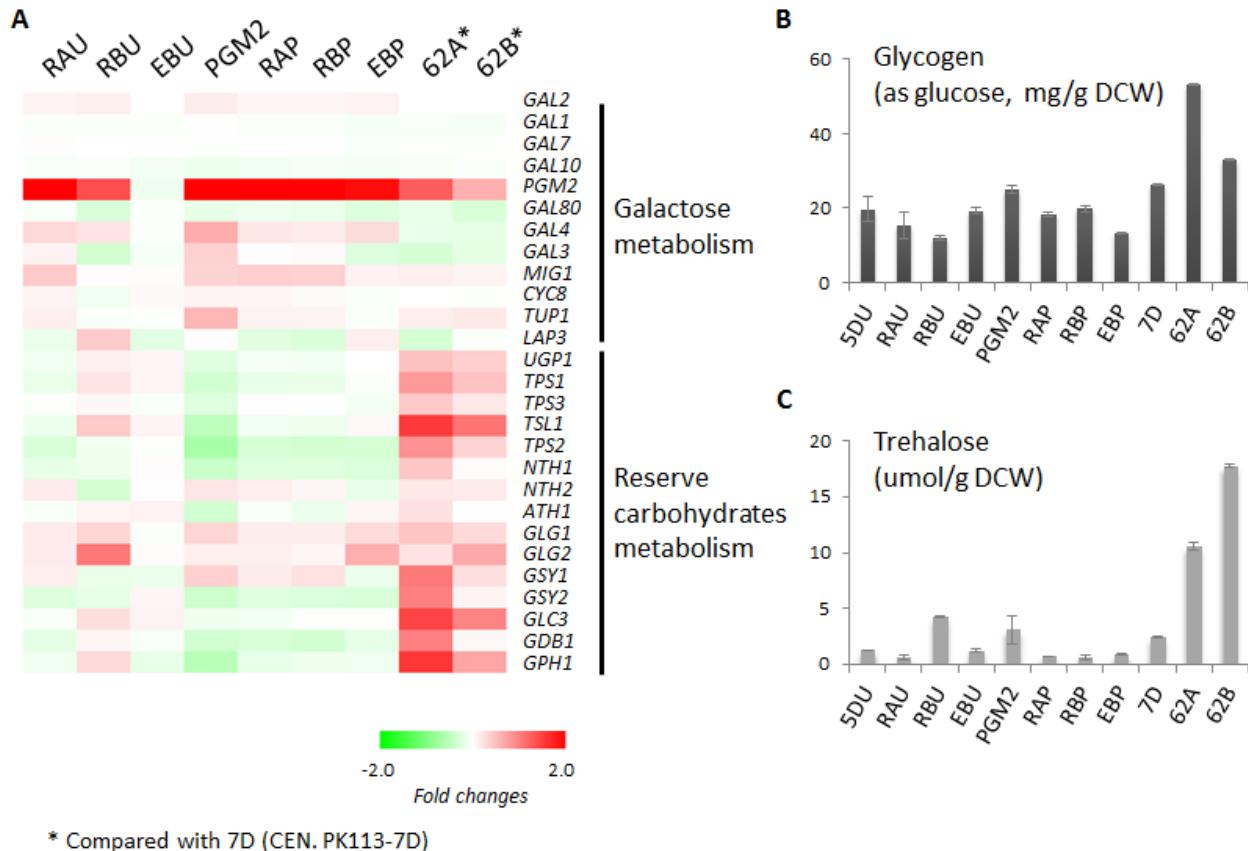


Fig. 3-8. Changes in the galactose and reserve carbohydrates metabolisms in the reconstructed strains are shown by changes in fold changes of the transcriptome and the concentration of carbohydrates. A: Fold changes of all genes involved in galactose and reserve carbohydrates metabolism are compared to the reference strains. B: The concentrations of glycogen. C: The concentration of trehalose. Error bars represent standard deviation from biological duplicates.

3.3. Characterization of molecular mechanism of trade-offs in carbon utilization (Paper III)

When galactose is used as a carbon source in industry, glucose would almost always also be present. Therefore, further characterization of the galactose-evolved mutants for growth on glucose was carried out. In addition, it was wondered if there was another effect from the obtained phenotype or traits, or there was collateral cost to get new traits. Considering the short adaptive evolutionary history of the evolved mutants to grow faster on galactose compared to the millions of year of evolution to maximize growth on glucose, a decline in glucose utilization to compensate for the cost of improving galactose utilization was not expected. However, interestingly all galactose-evolved mutants showed reduced glucose utilization (Fig. 3-9). In other word, the trade-off in carbon utilization between galactose and glucose was clearly detected in the evolved mutants. Two engineered mutants, PGM2 and SO16 strains also showed the trade-offs in the specific carbon uptake rate and the specific ethanol production rate (Fig. 3-9B). Since the genetic changes of these engineered strains were known, the genetic bases of this trade-off were easily identified. However, in case of evolved mutants, they showed different pattern of trade-offs, for example the maximum specific growth rate (Fig. 3-9A). Characterization of molecular and genetic bases of this trade-off was the main purpose of this study.

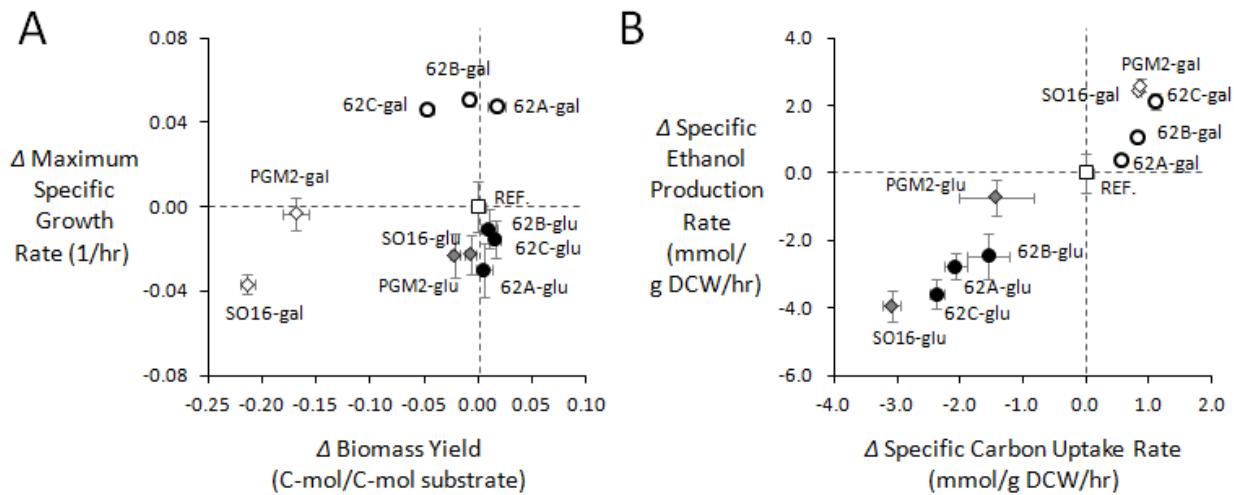


Fig. 3-9. Fermentation physiology of the evolved mutants and the engineered mutants compared to the reference strain in galactose (gal) and glucose (glu) through correlation between different values (Δ) of maximum specific growth rate and biomass yield (A), specific carbon uptake rate and specific ethanol production rate (B). Error bar represents standard error from biological duplicates in bioreactors.

Trade-offs among traits is one of the fundamental concepts in evolutionary biology. Two mechanisms for evolutionary trade-off have been suggested (Cooper & Lenski, 2000, Elena & Lenski, 2003, Wenger, *et al.*, 2011); antagonistic pleiotropy (AP) in which the same mutation is related to gain and loss of adaptation in different environment, and mutations accumulation (MA) where different mutations are responsible for the gain and loss of adaptation. Characterization of the trade-off mechanisms is important in the evolutionary approach for strain development, since the strategy for inverse metabolic engineering will dependent on the reason for evolutionary trade-off, AP or MA.

In this study, integrated genome-scale analyses were again applied to elucidate molecular and genetic evolutionary mechanism of the trade-off in the evolved mutants. Firstly, overall transcriptome profile was compared by principal component analysis; the distance between the evolved mutants and the reference strains looked almost identical during growth on both carbon sources (Fig. 3-10). This means that the evolved mutants responded to both carbon sources by similar transcriptional changes. More detailed molecular changes were analyzed by comparison of the differentially expressed genes and functional enrichment of them.

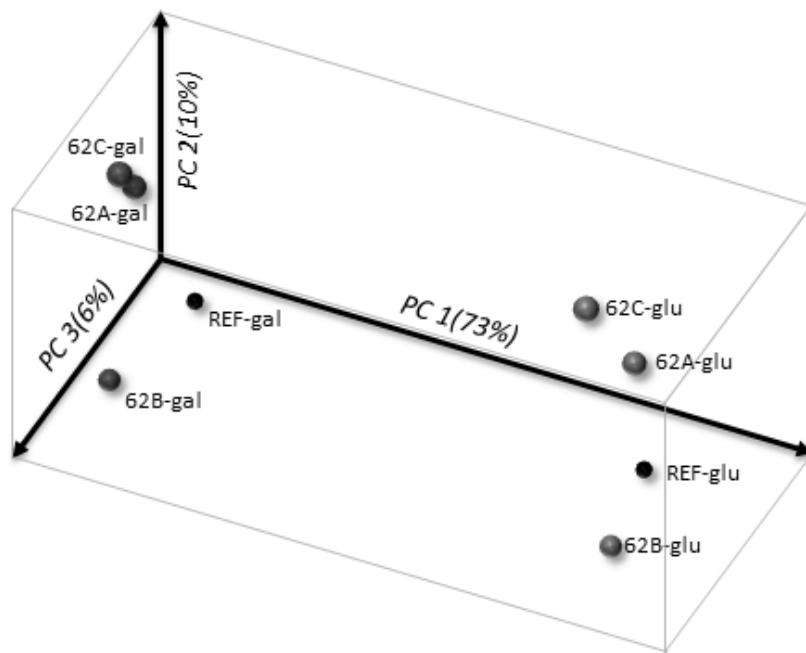


Fig. 3-10. Transcriptome analysis of the evolved mutants and the reference strain in galactose (gal) and glucose (glu) through principal component analysis (PCA). The results are projected by the first three PCs, which covered 89% of the variance.

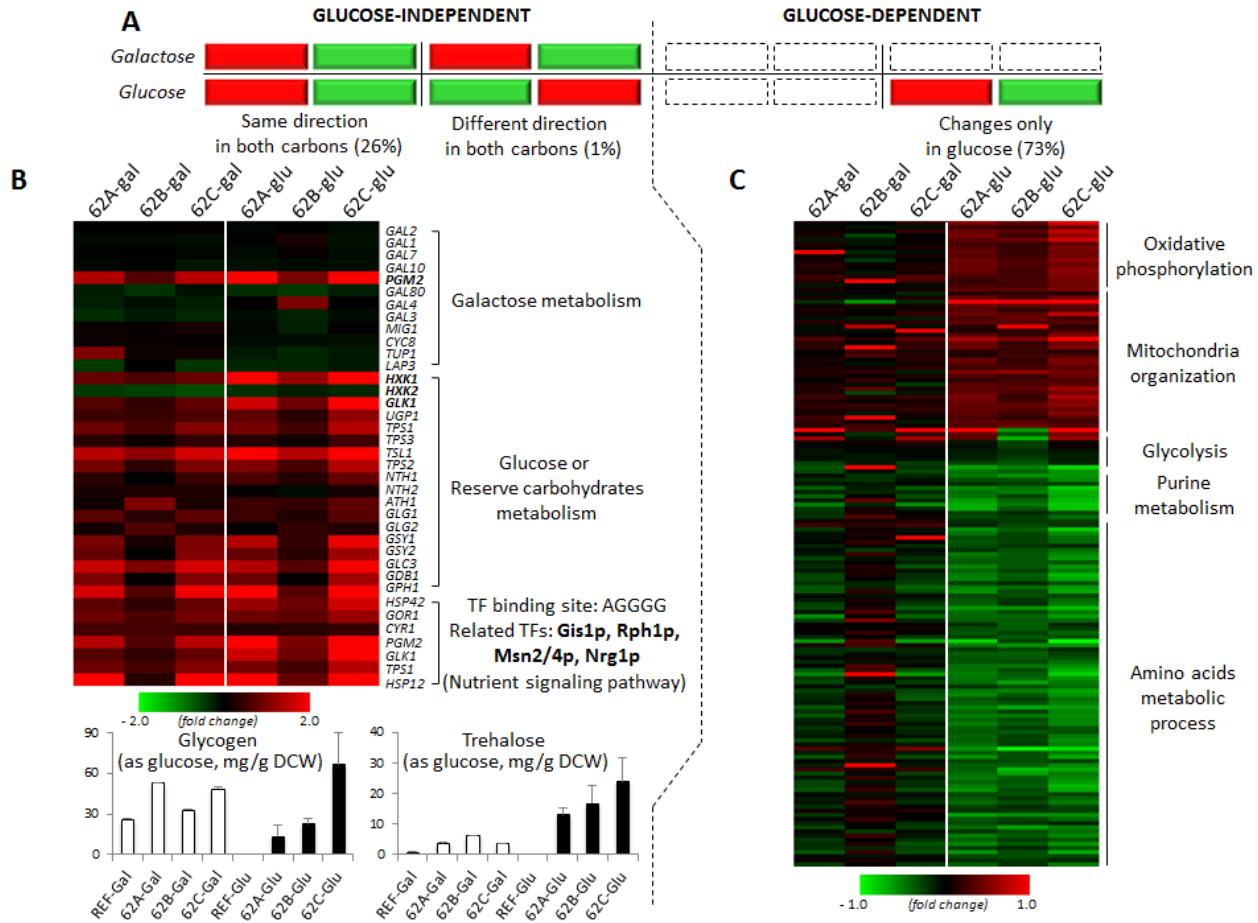


Fig. 3-11. Patterns of common molecular changes of the evolved mutants compared to the reference strain based on separation of glucose-independent and glucose-dependent. Glucose-independent, differentially expressed genes in both carbons; glucose-dependent, differentially expressed genes only in glucose (A), Specific pathways and targeted metabolites in glucose-independent (B), Specific pathways in glucose-dependent (C). Error bars in the concentration of glycogen and trehalose represent standard error from biological duplicates in bioreactors.

Conserved pattern of transcripts in specific parts of the metabolisms and reserve carbohydrates were detected (Fig. 3-11); and specific molecules that were likely involved in the trade-off mechanism were identified such as up-regulation of 1) *PGM2*, 2) two non-glucose inducible hexokinase *HXK1*, *GLK1* and 3) genes in reserve carbohydrates metabolism, and down regulation of 4) *HXK2* that is one of the key enzymes of glucose metabolism and is also a regulator of glucose catabolic repression (Gancedo, 1998). Additionally, commonly up-regulated genes on growing both carbon sources had the same transcription factor (TF) binding site (AGGGG). And TFs related to this site were *Gis1p, Rph1p, Msn2/4p* and *Nrg1p*, which are

involved in nutrient signaling pathway (Orzechowski Westholm, *et al.*, 2012). These results at the molecular level implied that antagonistic pleiotropy was the dominant mechanism for the trade-off, and the result was loosening the tight glucose control of metabolism.

The genetic bases of the trade-off in carbon utilization were explored. There were three identified mutations; two of them were confirmed as a beneficial mutation for galactose utilization such as mutations in the *RAS2* genes, and one of them was neutral for galactose availability. The site-directed mutants that had each of those mutations supported antagonistic pleiotropy as the mechanism for trade-off (Fig. 3-12).

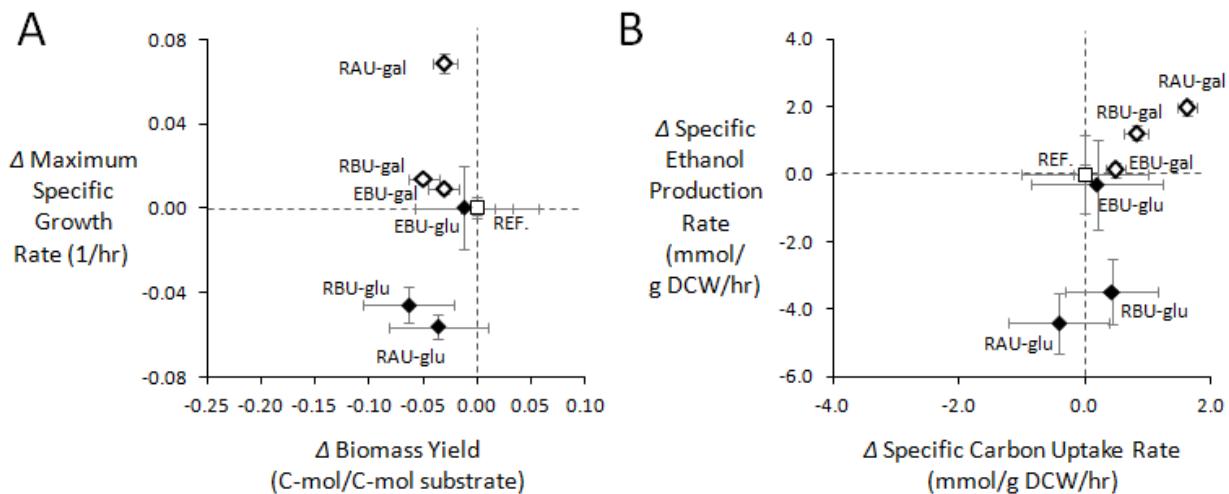


Fig. 3-12. Fermentation physiology of the site-directed mutants compared to the reference strain in galactose (gal) and glucose (glu) through correlation between different values (Δ) of maximum specific growth rate and biomass yield (A), specific carbon uptake rate and specific ethanol production rate (B). A reference strain for site-directed mutants is CEN.PK 113-5D having URA3 marker in plasmid. Error bar represents standard error from biological duplicate on galactose in bioreactors and biological triplicate on glucose in baffled flasks. Longer error bars in the reference were from glucose culture, shorter ones came from galactose culture.

The Ras/PKA signaling pathway is involved in the control of transcription factors, Gis1p, Rph1p, Msn2/4p and Nrg1p (Orzechowski Westholm, *et al.*, 2012). The identified mutations in the *RAS2* gene showed up-regulation of *PGM2* in the previous study (Paper II). The mutations

that triggered the molecular changes in reserve carbohydrates metabolism and hexokinases are not still clear. There would be unidentified mutations, which mutations may adjust the change between maximum specific growth rate and specific glucose uptake rate, because that change was the main difference between the evolved mutants and the site-directed mutants containing mutations in the *RAS2* gene. Hypothetical interpretation of the trade-off mechanism in the galactose evolved mutants is illustrated in Fig. 3-13.

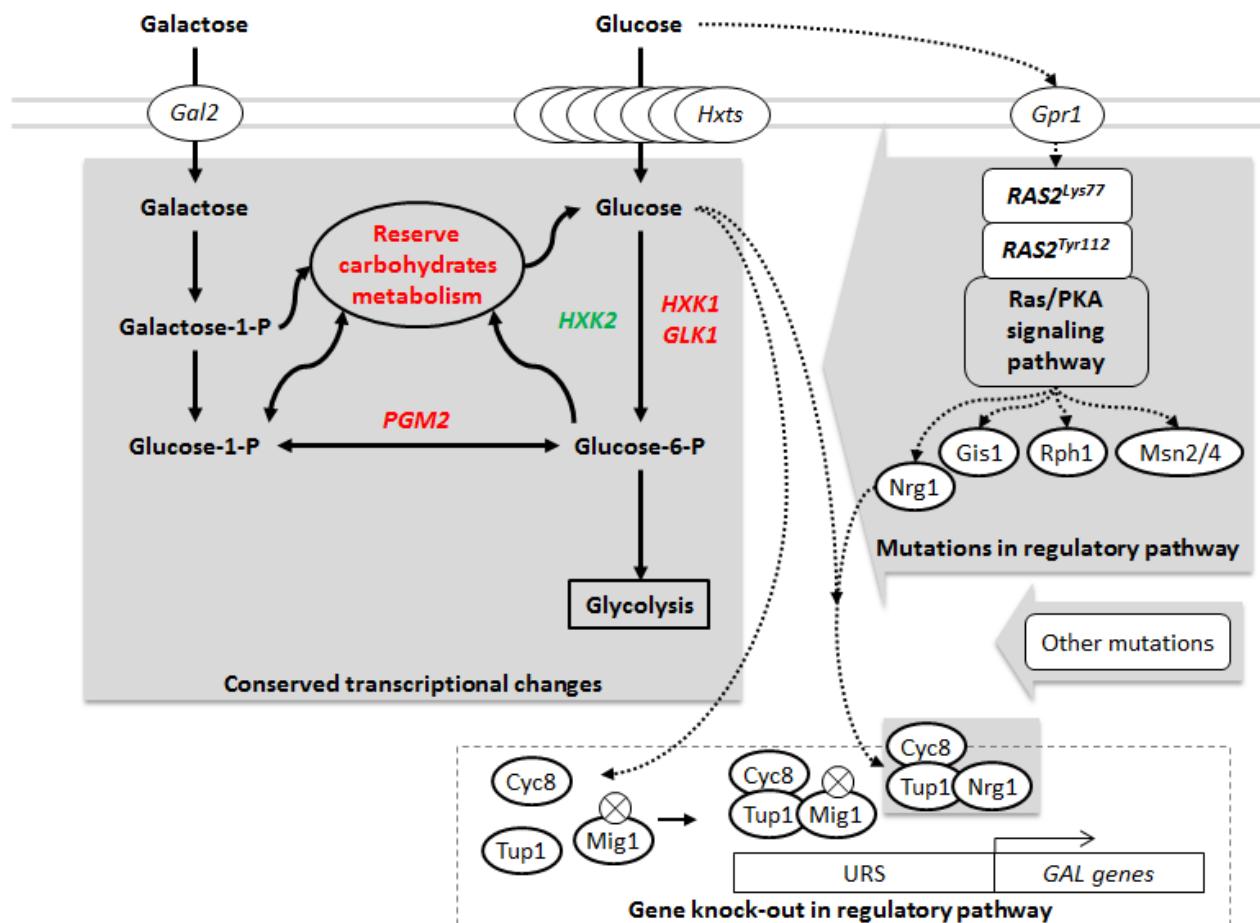


Fig. 3-13. Summary of possible molecular mechanism for the trade-off in galactose and glucose utilization. Colored letters (red and green) mean transcriptional change, up-regulation and down-regulation, respectively. Gray boxes (square and arrow shape) exhibit the changes in the evolved mutants. A dotted box means the change only in SO16 strain (knock-out of *MIG1*). Dot arrows represent signaling flow, solid arrows represent metabolic flow.

4. CONCLUSIONS

Adaptive evolution generated new strategies for improving galactose utilization in yeast *S.cerevisiae*. Those evolutionary strategies were characterized by integrated genome-scale analyses. Significantly, through this approach, evolutionary strategies of galactose-evolved mutants were elucidated at the molecular and genetic level. This characterization allowed inverse metabolic engineering to be more useful in evolutionary approaches for strain development. In addition, more characterization of the evolved mutants elucidated pleiotropy of the obtained traits. Through examples of this study, one could know what can be expectable or predictable in the application of evolutionary approaches for strain development. The most important findings in this Ph.D. study could be summarized as follow,

Evolutionary changes of galactose-evolved yeast mutants can be characterized by integrated genome-scale analyses

- Integration of genome scale analyses such as transcriptome, metabolome and whole-genome sequencing is crucial to identify the molecular and genetic basis of evolutionary changes. Each of these techniques does not allow for drawing comprehensive interpretation, but the combination of them provide a picture that enables understanding of the evolutionary strategies.
- It is important to use several evolved mutants with a reference strain, because this comparison allow identification of conserved mutations that result in the same phenotype. Each of the three evolved mutants has several mutations that probably do not contribute to the evolved phenotype, but by identifying conserved mutations, a clear picture emerged.

Non-canonical genetic changes results in canonical molecular changes of the evolved mutants

Transcriptome and metabolome analyses lists up significantly changed metabolism; and among them, molecular changes likely related to galactose metabolism were found. Whole-genome sequencing identified several mutations, while there were no mutations in the genes or promoter regions that show the molecular changes. Key mutations for improving galactose utilization were found in non-canonical pathways.

- Common molecular changes in all the galactose-evolved mutants are 1) Up-regulation of *PGM2* with reduced concentration of galactose-1-phosphate and glucose-1-phosphate, 2) Up-regulation of reserve carbohydrates metabolism with increased concentration of trehalose and glycogen (Fig. 6-4).
- Adaptive evolution of yeast on galactose generates no mutations in the galactose pathway and its regulatory region, which had been considered as modification targets for metabolic engineering; also no mutations in reserve carbohydrate metabolism.
- A common pathway that contains mutations in all the evolved mutants was the Ras/PKA signaling pathway (Table 6-1).
- Two identified mutations in the *RAS2* gene result in improved galactose utilization with up-regulation of *PGM2* but not reserve carbohydrate metabolism (Fig. 6-6, Fig. 6-8).

Importance of inverse metabolic engineering in connection with use of evolutionary approaches for strain development

Few genetic and transcriptional changes are required to reach adaptive phenotypes. Accumulation of deleterious mutations or negative epistasis among beneficial mutations seems to be quite high during adaptive evolution. Therefore, inverse metabolic engineering can give a lot of new strategies for further engineering, such as sieving out beneficial mutations from negative ones and generation of new artificial combination of mutations.

- Site-directed mutants containing only one mutation in the *RAS2* gene, [Gln⁷⁷→Lys] or [Asp¹¹²→Tyr] show similar improvement in the specific galactose uptake rate with the evolved mutants; also those strains display much smaller transcription changes compared to the evolved mutants (Fig. 6-6, Fig. 6-7).
- The site-directed mutant having *RAS2* [Gln⁷⁷→Lys] mutation even presents the highest specific galactose uptake rate among all the evolved and engineered strains (Fig. 6-6), and also relatively the highest maximum specific growth rate.
- Two mutations in the *RAS* gene have different effects on galactose utilization.
- New combinations of beneficial genetic changes almost completely recovers adaptive phenotypes in terms of galactose utilization, such as constitutive *PGM2* over-expression on a plasmid combined with mutation in *RAS2* [Gln⁷⁷→Lys] or [Asp¹¹²→Tyr], and in *ERG5*

[Arg³⁷⁰→Pro], respectively. These results indicate that new combinations of beneficial targets are one of the strategies for inverse metabolic engineering.

Molecular and genetic bases of evolutionary pleiotropy: trade-offs in carbon utilization

Galactose-evolved yeast mutants show trade-offs in carbon utilization between galactose and glucose. Adaptation on galactose seems to be realized by losing capacity for glucose utilization. The characterization results at the molecular and genetic level of this trade-off mechanism reveals that antagonistic pleiotropy is the dominant mechanism in the evolved mutations and this is likely realized by loosening the tight glucose catabolic repression system.

- The cost for improving galactose utilization may come from diminishing glucose utilization.
- Transcriptional changes with key metabolites of the three evolved mutants reveal antagonistic pleiotropy between glucose and galactose.
- Conserved molecular changes on both carbon sources are considered underlying the molecular mechanism by loosening tight glucose catabolic repression such as up-regulation of 1) *PGM2*, 2) non-glucose metabolism related hexokinase *HXK1*, *GLK1* and 3) reserve carbohydrate metabolism; down-regulation of 4) glucose catabolic repression regulator *HXK2*; and 4) involvement of transcription factors in nutrient sensing, *GIS1*, *RPH1*, *MSN2/4*, and *NRG1*.
- The mutations in the *RAS2* gene indicate antagonistic pleiotropy mechanisms for trade-off in carbon utilization by covering the phenotypic changes of the evolved mutants on both carbon sources.
- As mutations in the *RAS2* gene triggered up-regulation of *PGM2* and involved the transcription factors in nutrient sensing (Paper II), there are other unidentified mutations that induce transcriptional changes in the reserve carbohydrate metabolism and hexose kinases.
- Antagonistic pleiotropy between galactose and glucose utilization by attenuation of glucose regulation

5. PERSPECTIVE

Engineers have established significant development in the massive production of fuels and chemicals from petroleum and our generation is taking benefits from these technical advances. However, since the petroleum based production is using limited resources and generating serious environmental problems, our generation should prepare new technologies for the next generation, which uses renewable resources and alleviates environmental issues. Microbial fermentation processes could be one of the possible solutions, because this process utilizes biomass that is continuously produced with absorbing carbon dioxide in connection with its growth.

Engineering or reconstructing of microorganisms is the requisite step for the development of fermentation process. The engineering of biological systems is certainly different from mechanical or chemical engineering, since the biology is not only vastly complicated in their reaction networks and regulations, but also has emergent properties. Endy suggested four challenges in the engineering of biological system; 1) biological complexity, 2) the tedious and unreliable construction and characterization of synthetic biological systems, 3) the apparent spontaneous physical variation of biological system behavior, 4) evolution (Endy, 2005). One of the strategies for engineering the microorganism is to learn and apply nature's algorithm (Rothschild, 2010). Nature has produced relevant traits in specific environment; one also has used this valuable mechanism for making domesticated species from wild ones. Currently there are tools available for analysis genome-wide molecular and genetic changes. This means one can trace nature's strategies for obtaining new traits.

In this thesis, mutations in the *RAS2* gene were identified as the genetic bases for improving galactose utilization in yeast *S. cerevisiae*. Firstly, these mutations were only designable by nature' random mutagenesis; because not only the relationship between these mutations and galactose utilization was not predictable, but also even though they were located in the same gene, the effects from each of them were vastly different. Therefore, there are still vast amounts of opportunity to find new strategies for strain development by evolutionary approaches. Recently, artificial mutagenesis methods have been developed, which are called genome engineering techniques that make it possible to generate random mutations on specific regions such as promoters, regulators and limited pathway genes (Santos & Stephanopoulos, 2008, Boyle & Gill, 2012). However, these techniques still cannot cover the mutations that were found by random

mutagenesis. Secondly, these mutations were only detectable by genome-scale analyses, since these analyses can only scan whole genome level changes. Of course there were still unidentified mutations that could be related to the changes in reserve carbohydrates metabolism and hexokinases. Whole-genome sequencing in this study had limitation such as incomplete coverage of whole DNA, insufficient coverage folds, missing copy number changes and rearrangement and so on. In spite of these limitations, whole genome sequencing detected the key mutations. It was also important point to focus on common changes by employing several parallel evolved mutants.

In industry, a lot of mutations are normally accumulated in producing strains, because of long history of evolution and high mutation rate by treating mutagen. The limitation in the number of evolved mutants could make it difficult to find common mutations generated in the same gene or pathways. Therefore, identification of beneficial genetic changes is practically very difficult. As shown in this study, there is a possible solution, namely to do more characterization of the mutants at different conditions with other omics tools such as transcriptome and metabolome analysis. Perturbation of culture conditions could separate conserved changes from others. And those conserved related mutations could be the main reason for the obtained phenotype. For example, the galactose-evolved mutants kept the changes of transcripts and metabolites in specific metabolism when growing on two different carbon sources. So those changes could be interpretive as induced by the same mutations. This process could reduce the number of mutations that is involved in desired traits.

Another point to consider is that engineering of the Ras/PKA signaling pathway might be an efficient way to achieve multiple phenotypes of industrial interest. Two mutations in the same gene showed different phenotypes, and just one mutation was enough to reach the entire adaptive phenotype. These results indicate that the effect of mutations in the *RAS2* gene could be beyond the change of activity of Ras/PKA signaling pathway. In addition, when some mutations in the *RAS2* gene were combined with *PGM2* over-expression, the more diverse phenotype could be expectable. Therefore, constructing a mutation library of the *RAS2* gene or another mutation library of the whole Ras/PKA signaling pathway with adjusting *PGM2* expression could be very useful for the next step in strain development.

It is important to keep in mind that there are evolutionary mechanisms related to strain development, such as negative epistasis and trade-offs in traits. These mechanisms indicate that

there are many chances to do further engineering, which means that adaptive evolution could overcome the lack of detecting beneficial mutations and hereby to design new combinations of them. It is therefore crucial to accumulate examples about evolutionary mechanisms at detailed molecular levels supported by genome-scale analyses for further advancement the use of evolutionary engineering in industrial biotechnology.

6. ACKNOWLEDGMENTS

It is a lucky chance to take my doctoral period in Sweden, because of having experience of mature society. It is more luck to live in Göteborg, one of the best cities for taking care of family and lovely atmosphere. It is even greater to study in Systems and Synthetic Biology group at Chalmers. During this time, I have learned and disciplined not only to step in an independent researcher in science and engineering but also to build my view of world and personality.

First of all, I am sincere grateful to my supervisor Prof. Jens Nielsen, for giving an opportunity to be his PhD student, making profound freedom in my idea development with encouragement, trust and patience. He also has supported a lot of resources in research and education with guiding how my work fit into the big picture of biotechnology. His excellent ability and professional attitude in science are my next goal that I would like to achieve in my life.

I would like to thank Prof. Christer Larsson, Assoc. Prof. Joakim Norbeck, Assist. Prof. Dina Petranovic for many productive discussion and good feedback; also, Assoc. Prof. Carl Johan Franzén, and Assist. Prof. Maurizio Bettiga for good teaching and discussion.

I would like to thank co-authors, Assist. Prof. Goutham Vemuri for his supervision and inspiration in research, Assist Prof. Wanwipa Vongsangnak, Assist. Prof. Jin Hou, Assist. Prof. Sergio Bordel Velasco and Saeed Shoae for productive collaboration.

In particular, I would like to thank seniors at CJ CheilJedang Dr. Young-Lyeol Yang, Dr. Chang-Gyeom Kim, Sung-Sik Park, Jin-Su Chung for helping to start my PhD; and Seong-Ah Kim, Jaeyong Park, Dr. Jaeyeong Ju, Kyung-Suk Lee, Director of R&D Sangjo Lim, and CEO Chul-Ha Kim for their practical supporting and helping; my leaders, Vice Director Jinman Cho, Dr. Jong-Kwon Han, Dr. Kwang-Myung Cho, Dr. Kwang-Ho Lee, Assist Prof. Hyun-Soo Kim for their advises and valuable guidance; and Byoung-chun Lee, Jun-Ok Moon for their encouragement.

A special acknowledgement for Dr. Rahul Kumar, he hasn't hesitated to support me with his vast amount of time, effort and attention to my research and writing. He always gives very valuable input and discussion in diverse field, which allow me to have broad knowledge.

I would like to thank my colleagues, Dr. Jie Zhang, Dr. Pramote Chumnanpuen for their helping to initiate experimental systems biology works and cheering; Dr. Luis Caspetá-Guadarrama, Dr. Yun Chen, Klaas Buijs for revising the thesis and good discussion; Kanokarn Kocharin, Zihe Liu, Lifang Liu, Marta Papini, Siavash Partow, Gionata Scalcinati, Fredrik Karlsson, Tobias Österlund, Christoph Knuf, Juan Valle, Dr. Roberto Olivares-Hernández, Dr. Shuobo Shi, Dr. José L. Martínez, Dr. Adil Mardinoglu, Dr. António Roldão and other lab mates for making a lab life fun and checking manuscripts. Thank Dr. Verena Siewers, Dr. Intawat

Nookaew, Dr. Subir Kumar Nandy, Assist. Prof. Keith Tyo for their valuable comments. Thank Govindprasad Bhutada and Sinisa Bratulic for their questioning and cheers. I would like to deeply thank Erica Dahlin, Martina Butorac, Malin Nordvall, Marie Nordqvist, Ximena Sevilla, Suwanee Jansa-Ard for their supporting.

There are also very nice Korean friends and seniors who I met in Sweden, Dr. Hee Chul Park, Sang-Mi Na, Hyeongcheol Yoo, Dr. Byung-Soo Kang, Eunmi Choi, Dr. Keunjae Kim, Kwang-Sup Lee, Myung-Ja Kim, Youngsoon Um and her husband Ingemar Johansson. Without their helps, my family would face many difficulties; thank so much for their support.

I would like to thank Prof. Jin Byung Park for guiding to PhD and close friends Taehee Park, Hyungun Kim for their encouragement and keeping good discussion in many issues.

My sincere gratitude goes to my parents, Sung-Hwan Hong, Kyung-Ja Lee, my younger brother Duk-Ki Hong, parents-in-law Jae-Suk-Kang, Sung-Won Hong, and sister-in-law Min-Sun Kang. Their devotion, pray for me and enthusiasm in their life are great resource for sustaining and keeping values of my life.

And most of all, I would like to thank my wife Min-Jin Kang, for exploring unpredictable life together, with being beautiful, lively, wise and encouraging me to overcome any barriers in the doctoral works. It is great to see and learn how to adapt new environment from my daughter Jin-Seo Hong; she is very well along with her international friends, and I am busy to drop her on and off her friends' home. Jin-Ha Hong, my son doesn't care about his dad's situation, just always try to hug me and sit on my knees. Thank my angels for your presence.

Kuk-Ki Hong

Aug 2012

at Chalmers

in Göteborg, Sweden

7. REFERENCES

- Acar M, Becskei A & van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435: 228-232.
- Albertsen L, Chen Y, Bach LS, et al. (2011) Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Appl Environ Microbiol* 77: 1033-1040.
- Asadollahi MA, Maury J, Schalk M, Clark A & Nielsen J (2010) Enhancement of farnesyl diphosphate pool as direct precursor of sesquiterpenes through metabolic engineering of the mevalonate pathway in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 106: 86-96.
- Bailey JE, Sburlati A, Hatzimanikatis V, Lee K, Renner WA & Tsai PS (2002) Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnol Bioeng* 79: 568-579.
- Becker JV, Armstrong GO, van der Merwe MJ, Lambrechts MG, Vivier MA & Pretorius IS (2003) Metabolic engineering of *Saccharomyces cerevisiae* for the synthesis of the wine-related antioxidant resveratrol. *FEMS Yeast Res* 4: 79-85.
- Bengtsson O, Jeppsson M, Sondergaard M, Parachin NS, Sauer U, Hahn-Hagerdal B & Gorwa-Grauslund MF (2008) Identification of common traits in improved xylose-growing *Saccharomyces cerevisiae* for inverse metabolic engineering. *Yeast* 25: 835-847.
- Boles E & Hollenberg CP (1997) The molecular genetics of hexose transport in yeasts. *FEMS Microbiol Rev* 21: 85-111.
- Boyle NR & Gill RT (2012) Tools for genome-wide strain design and construction. *Curr Opin Biotechnol.*
- Bro C & Nielsen J (2004) Impact of 'ome' analyses on inverse metabolic engineering. *Metab Eng* 6: 204-211.
- Bro C, Knudsen S, Regenberg B, Olsson L & Nielsen J (2005) Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* 71: 6465-6472.
- Brochado AR, Matos C, Moller BL, Hansen J, Mortensen UH & Patil KR (2010) Improved vanillin production in baker's yeast through in silico design. *Microb Cell Fact* 9: 84.
- Cabib E (1970) Research on sugar nucleotides brings honor to Argentinian biochemist (Luis Leloir). *Science* 170: 608-609.
- Chen X, Nielsen KF, Borodina I, Kielland-Brandt MC & Karhumaa K (2011) Increased isobutanol production in *Saccharomyces cerevisiae* by overexpression of genes in valine metabolism. *Biotechnol Biofuels* 4: 21.
- Chigira Y, Oka T, Okajima T & Jigami Y (2008) Engineering of a mammalian O-glycosylation pathway in the yeast *Saccharomyces cerevisiae*: production of O-fucosylated epidermal growth factor domains. *Glycobiology* 18: 303-314.
- Conrad TM, Lewis NE & Palsson BO (2011) Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol* 7: 509.
- Cooper VS & Lenski RE (2000) The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407: 736-739.
- Daugherty JP, Kraemer WF & Joshi JG (1975) Purification and properties of phosphoglucomutase from Fleischmann's yeast. *Eur J Biochem* 57: 115-126.
- Davie JK, Edmondson DG, Coco CB & Dent SY (2003) Tup1-Ssn6 interacts with multiple class I histone deacetylases in vivo. *J Biol Chem* 278: 50158-50162.
- de Jongh WA, Bro C, Ostergaard S, Regenberg B, Olsson L & Nielsen J (2008) The roles of galactitol, galactose-1-phosphate, and phosphoglucomutase in galactose-induced toxicity in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 101: 317-326.
- Demain AL (2000) Microbial biotechnology. *Trends Biotechnol* 18: 26-31.
- Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF & Kassen R (2012) Evolutionary insight

- from whole-genome sequencing of experimentally evolved microbes. *Mol Ecol* 21: 2058-2077.
- Egel-Mitani M, Andersen AS, Diers II, Hach M, Thim L, Hastrup S & Vad K (2000) Yield improvement of heterologous peptides expressed in *yps1*-disrupted *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 26: 671-677.
- Elena SF & Lenski RE (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* 4: 457-469.
- Elkins JG, Raman B & Keller M (2010) Engineered microbial systems for enhanced conversion of lignocellulosic biomass. *Curr Opin Biotechnol* 21: 657-662.
- Endy D (2005) Foundations for engineering biology. *Nature* 438: 449-453.
- Estruch F (2000) Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiol Rev* 24: 469-486.
- Eudes A, Baidoo E, Yang F, et al. (2011) Production of tranilast [*N*-(3',4'-dimethoxycinnamoyl)-anthranilic acid] and its analogs in yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biot* 89: 989-1000.
- Fukasawa T, Obonai K, Segawa T & Nogi Y (1980) The enzymes of the galactose cluster in *Saccharomyces cerevisiae*. II. Purification and characterization of uridine diphosphoglucose 4-epimerase. *J Biol Chem* 255: 2705-2707.
- Gancedo JM (1998) Yeast carbon catabolite repression. *Microbiology and Molecular Biology Reviews* 62: 334-+.
- Garcia Sanchez R, Hahn-Hagerdal B & Gorwa-Grauslund MF (2010) PGM2 overexpression improves anaerobic galactose fermentation in *Saccharomyces cerevisiae*. *Microb Cell Fact* 9: 40.
- Goeddel DV, Heyneker HL, Hozumi T, et al. (1979) Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. *Nature* 281: 544-548.
- Guadalupe Medina V, Almering MJ, van Maris AJ & Pronk JT (2010) Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Appl Environ Microbiol* 76: 190-195.
- Hackel BJ, Huang D, Bubolz JC, Wang XX & Shusta EV (2006) Production of soluble and active transferrin receptor-targeting single-chain antibody using *Saccharomyces cerevisiae*. *Pharm Res* 23: 790-797.
- Hahn-Hagerdal B, Karhumaa K, Fonseca C, Spencer-Martins I & Gorwa-Grauslund MF (2007) Towards industrial pentose-fermenting yeast strains. *Appl Microbiol Biotechnol* 74: 937-953.
- Hazelwood LA, Walsh MC, Luttik MA, Daran-Lapujade P, Pronk JT & Daran JM (2009) Identity of the growth-limiting nutrient strongly affects storage carbohydrate accumulation in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 75: 6876-6885.
- Heinemann M & Sauer U (2010) Systems biology of microbial metabolism. *Curr Opin Microbiol* 13: 337-343.
- Herrgard MJ, Swainston N, Dobson P, et al. (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat Biotechnol* 26: 1155-1160.
- Herring CD, Raghunathan A, Honisch C, et al. (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 38: 1406-1412.
- Holden HM, Rayment I & Thoden JB (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *Journal of Biological Chemistry* 278: 43885-43888.
- Ideker T, Thorsson V, Ranish JA, et al. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292: 929-934.
- Ikeda M, Ohnishi J, Hayashi M & Mitsuhashi S (2006) A genome-based approach to create a minimally mutated *Corynebacterium glutamicum* strain for efficient L-lysine production. *J Ind Microbiol Biotechnol* 33: 610-615.

- Ikeda M, Mitsuhashi S, Tanaka K & Hayashi M (2009) Reengineering of a *Corynebacterium glutamicum* L-arginine and L-citrulline producer. *Appl Environ Microbiol* 75: 1635-1641.
- Jelen P (1979) Industrial Whey Processing Technology - Overview. *Journal of Agricultural and Food Chemistry* 27: 658-661.
- Johnson IS (1983) Human insulin from recombinant DNA technology. *Science* 219: 632-637.
- Kelly SL, Lamb DC, Corran AJ, Baldwin BC, Parks LW & Kelly DE (1995) Purification and reconstitution of activity of *Saccharomyces cerevisiae* P450 61, a sterol Delta(22)-desaturase. *FEBS Lett* 377: 217-220.
- Kim E-J, Park Y-K, Lim H-K, Park Y-C & Seo J-H (2009) Expression of hepatitis B surface antigen S domain in recombinant *Saccharomyces cerevisiae* using GAL1 promoter. *J Biotech* 141: 155-159.
- Kim HJ, Lee SJ & Kim H-J (2010) Optimizing the secondary structure of human papillomavirus type 16 L1 mRNA enhances L1 protein expression in *Saccharomyces cerevisiae*. *J Biotech* 150: 31-36.
- Kim NJ, Li H, Jung K, Chang HN & Lee PC (2011) Ethanol production from marine algal hydrolysates using *Escherichia coli* KO11. *Bioresour Technol* 102: 7466-7469.
- Kim SR, Ha SJ, Wei N, Oh EJ & Jin YS (2012) Simultaneous co-fermentation of mixed sugars: a promising strategy for producing cellulosic ethanol. *Trends Biotechnol* 30: 274-282.
- Kinoshita S, Ueda S & Shimono M (1957) Studies on the amino acid fermentation. Part I. Production of L-glutamic acid by various microorganisms. *J. Gen. Appl. Microbiol.* 3: 193-205.
- Krivoruchko A, Siewers V & Nielsen J (2011) Opportunities for yeast metabolic engineering: Lessons from synthetic biology. *Biotechnol J* 6: 262-276.
- Kummel A, Ewald JC, Fendt SM, et al. (2010) Differential glucose repression in common yeast strains in response to HXK2 deletion. *FEMS Yeast Res* 10: 322-332.
- Lai K, Elsa LJ & Wierenga KJ (2009) Galactose toxicity in animals. *IUBMB Life* 61: 1063-1074.
- Le Crom S, Schackwitz W, Pennacchio L, et al. (2009) Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* 106: 16151-16156.
- Lee KS, Hong ME, Jung SC, et al. (2011) Improved galactose fermentation of *Saccharomyces cerevisiae* through inverse metabolic engineering. *Biotechnol Bioeng* 108: 621-631.
- Lee W & Dasilva NA (2006) Application of sequential integration for metabolic engineering of 1,2-propanediol production in yeast. *Metab Eng* 8: 58-65.
- Li A, Liu Z, Li Q, Yu L, Wang D & Deng X (2008) Construction and characterization of bidirectional expression vectors in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 8: 6-9.
- Lipinsky ES (1981) Chemicals from biomass: petrochemical substitution options. *Science* 212: 1465-1471.
- Lowin T, Raab U, Schroeder J, Franssila R & Modrow S (2005) Parvovirus B19 VP2-proteins produced in *Saccharomyces cerevisiae*: comparison with VP2-particles produced by baculovirus-derived vectors. *J Vet Med B Infect Dis Vet Public Health* 52: 348-352.
- MacLean D, Jones JD & Studholme DJ (2009) Application of 'next-generation' sequencing technologies to microbial genetics. *Nat Rev Microbiol* 7: 287-296.
- Madsen KM, Udatha GD, Semba S, et al. (2011) Linking genotype and phenotype of *Saccharomyces cerevisiae* strains reveals metabolic engineering targets and leads to triterpene hyper-producers. *PLoS ONE* 6: e14763.
- Mapelli V, Hillestrøm PR, Kápolna E, Larsen EH & Olsson L (2011) Metabolic and bioprocess engineering for production of selenized yeast with increased content of seleno-methylselenocysteine. *Metab Eng* 13: 282-293.
- Mardis ER (2008) Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 9: 387-402.

- Marie G-G, Bärbel H-H & Maurizio B (2009) Metabolic Engineering in Yeast. *The Metabolic Pathway Engineering Handbook*, ed.^eds.), p.^pp. 22-21-22-48. CRC Press.
- Metzker ML (2010) Sequencing technologies - the next generation. *Nat Rev Genet* 11: 31-46.
- Mutka SC, Bondi SM, Carney JR, Da Silva NA & Kealey JT (2006) Metabolic pathway engineering for complex polyketide biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 6: 40-47.
- Nakayama K, Kitada S & Kinoshita S (1961) Studies on lysine fermentation I. The control mechanism on lysine accumulation by homoserine and threonine. *J. Gen. Appl. Microbiol.* 7: 145-154.
- Nevoigt E (2008) Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 72: 379-412.
- Nielsen J & Jewett MC (2008) Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res* 8: 122-131.
- Ohnishi J, Mitsuhashi S, Hayashi M, Ando S, Yokoi H, Ochiai K & Ikeda M (2002) A novel methodology employing *Corynebacterium glutamicum* genome information to generate a new L-lysine-producing mutant. *Appl Microbiol Biotechnol* 58: 217-223.
- Orzechowski Westholm J, Tronnersjo S, Nordberg N, Olsson I, Komorowski J & Ronne H (2012) Gis1 and Rph1 regulate glycerol and acetate metabolism in glucose depleted yeast cells. *PLoS One* 7: e31577.
- Ostergaard S, Olsson L, Johnston M & Nielsen J (2000) Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat Biotechnol* 18: 1283-1286.
- Oud B, van Maris AJ, Daran JM & Pronk JT (2012) Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. *FEMS Yeast Res* 12: 183-196.
- Ozcan S & Johnston M (1999) Function and regulation of yeast hexose transporters. *Microbiol Mol Biol Rev* 63: 554-569.
- Panesar PS, Kennedy JF, Gandhi DN & Bunko K (2007) Bioutilisation of whey for lactic acid production. *Food Chemistry* 105: 1-14.
- Pannala VR, Bhat PJ, Bhartiya S & Venkatesh KV (2010) Systems biology of GAL regulon in *Saccharomyces cerevisiae*. *Wiley Interdiscip Rev Syst Biol Med* 2: 98-106.
- Patil KR & Nielsen J (2005) Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc Natl Acad Sci U S A* 102: 2685-2689.
- Patnaik R (2008) Engineering complex phenotypes in industrial strains. *Biotechnol Prog* 24: 38-47.
- Peralta-Yahya PP, Ouellet M, Chan R, Mukhopadhyay A, Keasling JD & Lee TS (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat Commun* 2: 483.
- Petranovic D & Vemuri GN (2009) Impact of yeast systems biology on industrial biotechnology. *J Biotechnol* 144: 204-211.
- Portnoy VA, Bezdan D & Zengler K (2011) Adaptive laboratory evolution--harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol* 22: 590-594.
- Pronk JT (2002) Auxotrophic yeast strains in fundamental and applied research. *Appl Environ Microbiol* 68: 2095-2100.
- Raab AM, Gebhardt G, Bolotina N, Weuster-Botz D & Lang C (2010) Metabolic engineering of *Saccharomyces cerevisiae* for the biotechnological production of succinic acid. *Metab Eng* 12: 518-525.
- Rakestraw JA, Sazinsky SL, Piatesi A, Antipov E & Wittrup KD (2009) Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 103: 1192-1201.
- Ramsey SA, Smith JJ, Orrell D, et al. (2006) Dual feedback loops in the GAL regulon suppress cellular heterogeneity in yeast. *Nat Genet* 38: 1082-1087.
- Reifenberger E, Boles E & Ciriacy M (1997) Kinetic characterization of individual hexose

- transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanisms of glucose repression. *Eur J Biochem* 245: 324-333.
- Reimand J, Arak T & Vilo J (2011) g:Profiler--a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res* 39: W307-315.
- Rico J, Pardo E & Orejas M (2010) Enhanced production of a plant monoterpene by overexpression of the 3-hydroxy-3-methylglutaryl coenzyme A reductase catalytic domain in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 76: 6449-6454.
- Rothschild LJ (2010) A powerful toolkit for synthetic biology: Over 3.8 billion years of evolution. *Bioessays* 32: 304-313.
- Santos CN & Stephanopoulos G (2008) Combinatorial engineering of microbes for optimizing cellular phenotype. *Curr Opin Chem Biol* 12: 168-176.
- Sauer U (2001) Evolutionary engineering of industrially important microbial phenotypes. *Adv Biochem Eng Biotechnol* 73: 129-169.
- Schell MA & Wilson DB (1977) Purification and properties of galactokinase from *Saccharomyces cerevisiae*. *J Biol Chem* 252: 1162-1166.
- Segawa T & Fukasawa T (1979) The enzymes of the galactose cluster in *Saccharomyces cerevisiae*. Purification and characterization of galactose-1-phosphate uridylyltransferase. *J Biol Chem* 254: 10707-10709.
- Shendure J & Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol* 26: 1135-1145.
- Siewers V, San-Bento R & Nielsen J (2010) Implementation of communication-mediated domains for non-ribosomal peptide production in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 106: 841-844.
- Siso MIG (1996) The biotechnological utilization of cheese whey: A review. *Bioresour Technol* 57: 1-11.
- Snyder M & Gallagher JE (2009) Systems biology from a yeast omics perspective. *FEBS Lett* 583: 3895-3899.
- Sonderegger M & Sauer U (2003) Evolutionary engineering of *Saccharomyces cerevisiae* for anaerobic growth on xylose. *Appl Environ Microbiol* 69: 1990-1998.
- St John TP & Davis RW (1981) The organization and transcription of the galactose gene cluster of *Saccharomyces*. *J Mol Biol* 152: 285-315.
- Tavares S, Grotkjaer T, Oksen T, Haslam RP, Napier JA & Gunnarsson N (2011) Metabolic engineering of *Saccharomyces cerevisiae* for production of Eicosapentaenoic Acid, using a novel {Delta}5-Desaturase from *Paramecium tetraurelia*. *Appl Environ Microbiol* 77: 1854-1861.
- Thoden JB & Holden HM (2005) The molecular architecture of galactose mutarotase/UDP-galactose 4-epimerase from *Saccharomyces cerevisiae*. *J Biol Chem* 280: 21900-21907.
- Timson DJ (2007) Galactose Metabolism in *Saccharomyces cerevisiae*. *Dynamic Biochemistry, Process Biotechnology and Molecular Biology* 1: 63-73.
- Treitel MA & Carlson M (1995) Repression by SSN6-TUP1 is directed by MIG1, a repressor/activator protein. *Proc Natl Acad Sci U S A* 92: 3132-3136.
- Tyo KE, Alper HS & Stephanopoulos GN (2007) Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol* 25: 132-137.
- Vai M, Brambilla L, Orlandi I, Rota N, Ranzi BM, Alberghina L & Porro D (2000) Improved secretion of native human insulin-like growth factor 1 from gas1 mutant *Saccharomyces cerevisiae* cells. *Appl Environ Microbiol* 66: 5477-5479.
- van Maris AJ, Geertman JM, Vermeulen A, et al. (2004) Directed evolution of pyruvate decarboxylase-negative *Saccharomyces cerevisiae*, yielding a C2-independent, glucose-tolerant, and pyruvate-hyperproducing yeast. *Appl Environ Microbiol* 70: 159-166.
- van Maris AJ, Abbott DA, Bellissimi E, et al. (2006) Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie Van Leeuwenhoek* 90: 391-418.

- Vellanki RN, Komaravelli N, Tatineni R & Mangamoori LN (2007) Expression of hepatitis B surface antigen in *Saccharomyces cerevisiae* utilizing glyceraldehydehyde-3-phosphate dehydrogenase promoter of *Pichia pastoris*. *Biotechnol Lett* 29: 313-318.
- Vennestrom PN, Osmundsen CM, Christensen CH & Taarning E (2011) Beyond petrochemicals: the renewable chemicals industry. *Angew Chem Int Ed Engl* 50: 10502-10509.
- Verwaal R, Wang J, Meijnen JP, Visser H, Sandmann G, van den Berg JA & van Ooyen AJ (2007) High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol* 73: 4342-4350.
- Vijayendran C, Barsch A, Friebs K, Niehaus K, Becker A & Flaschel E (2008) Perceiving molecular evolution processes in *Escherichia coli* by comprehensive metabolite and gene expression profiling. *Genome Biol* 9: R72.
- Wargacki AJ, Leonard E, Win MN, et al. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* 335: 308-313.
- Warner JR, Patnaik R & Gill RT (2009) Genomics enabled approaches in strain engineering. *Curr Opin Microbiol* 12: 223-230.
- Wenger JW, Piotrowski J, Nagarajan S, Chiotti K, Sherlock G & Rosenzweig F (2011) Hunger Artists: Yeast Adapted to Carbon Limitation Show Trade-Offs under Carbon Sufficiency. *Plos Genetics* 7.
- Werpy T & Petersen G (2004) *Top value added chemicals from biomass. Volume 1, Results of screening for potential candidates from sugars and synthesis gas*. U.S. Dept. of Energy, [[Washington, D.C.].
- Westfall PJ, Pitera DJ, Lenihan JR, et al. (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A* 109: E111-118.
- Wi SG, Kim HJ, Mahadevan SA, Yang DJ & Bae HJ (2009) The potential value of the seaweed Ceylon moss (*Gelidium amansii*) as an alternative bioenergy resource. *Bioresour Technol* 100: 6658-6660.
- Wieczorke R, Krampe S, Weierstall T, Freidel K, Hollenberg CP & Boles E (1999) Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett* 464: 123-128.
- Yano K & Fukasawa T (1997) Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 94: 1721-1726.
- Yu KO, Jung J, Kim SW, Park CH & Han SO (2012) Synthesis of FAEEs from glycerol in engineered *Saccharomyces cerevisiae* using endogenously produced ethanol by heterologous expression of an unspecific bacterial acyltransferase. *Biotechnol Bioeng* 109: 110-115.
- Zaldivar J, Borges A, Johansson B, Smits HP, Villas-Boas SG, Nielsen J & Olsson L (2002) Fermentation performance and intracellular metabolite patterns in laboratory and industrial xylose-fermenting *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 59: 436-442.
- Zhang B, Carlson R & Srienc F (2006) Engineering the monomer composition of polyhydroxylalkanoates synthesized in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 72: 536-543.
- Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WP & del Cardayre SB (2002) Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* 415: 644-646.
- Zhao L, Wang J, Zhou J, Liu L, Du G & Chen J (2011) Modification of carbon flux in *Saccharomyces cerevisiae* to improve L-lactic acid production. *Wei Sheng Wu Xue Bao* 51: 50-58.
- Zheng WJ, Xu HE & Johnston SA (1997) The cysteine-peptidase bleomycin hydrolase is a member of the galactose regulon in yeast. *Journal of Biological Chemistry* 272: 30350-30355.

PAPER I

Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis

Kuk-Ki Hong^{a,b}, Wanwipa Vongsangnak^{a,1}, Goutham N. Vemuri^a, and Jens Nielsen^{a,c,2}

^aDepartment of Chemical and Biological Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; ^bCJ Research Institute of Biotechnology, CJ CheilJedang, Seoul 157-724, Korea; and ^cNovo Nordisk Centre for Biosustainability, Technical University of Denmark, DK2800 Lyngby, Denmark

Edited by Arnold L. Demain, Drew University, Madison, NJ, and approved June 8, 2011 (received for review February 28, 2011)

Identification of the underlying molecular mechanisms for a derived phenotype by adaptive evolution is difficult. Here, we performed a systems-level inquiry into the metabolic changes occurring in the yeast *Saccharomyces cerevisiae* as a result of its adaptive evolution to increase its specific growth rate on galactose and related these changes to the acquired phenotypic properties. Three evolved mutants (62A, 62B, and 62C) with higher specific growth rates and faster specific galactose uptake were isolated. The evolved mutants were compared with a reference strain and two engineered strains, SO16 and PGM2, which also showed higher galactose uptake rate in previous studies. The profile of intermediates in galactose metabolism was similar in evolved and engineered mutants, whereas reserve carbohydrates metabolism was specifically elevated in the evolved mutants and one evolved strain showed changes in ergosterol biosynthesis. Mutations were identified in proteins involved in the global carbon sensing Ras/PKA pathway, which is known to regulate the reserve carbohydrates metabolism. We evaluated one of the identified mutations, *RAS2*^{Tyr112}, and this mutation resulted in an increased specific growth rate on galactose. These results show that adaptive evolution results in the utilization of unpredicted routes to accommodate increased galactose flux in contrast to rationally engineered strains. Our study demonstrates that adaptive evolution represents a valuable alternative to rational design in bioengineering of improved strains and, that through systems biology, it is possible to identify mutations in evolved strain that can serve as unforeseen metabolic engineering targets for improving microbial strains for production of biofuels and chemicals.

In the field of industrial biotechnology, there is a need to develop efficient cell factories for the production of fuels and chemicals. Even though the concept of metabolic engineering (1) is frequently used in both academia and industry for the development of unique cell factories, evolutionary engineering methods are still widely performed (2). The power of adaptive evolution, sometimes in combination with metabolic engineering, is well illustrated in several recent examples (3, 4). Despite its advantages, conventional random mutagenesis and screening are hampered by the difficulties associated with finding the underlying molecular mechanisms for a derived phenotype and, hence, the combination of adaptive evolution with more rational approaches like metabolic engineering is attractive. Tools from systems biology and the ability to perform deep sequencing of several strains have offered new opportunities for establishing links between genotype and phenotype and, hereby, allow for combinations of random and rational approaches to strain improvement (5, 6).

Understanding the evolutionary strategies of a cell to metabolize nonfavored carbon sources is an integral part of strain development in cost efficient bioprocesses. Galactose is an abundant sugar in nonfood crops (7), and it is therefore interesting to generate strains that can efficiently use galactose as a carbon source. However, the yeast *Saccharomyces cerevisiae*, which is a frequently used cell factory in industrial biotechnology, grows at half the rate on galactose compared with glucose, despite the structural similarity between galactose and glucose (8).

There is extensive knowledge on the regulation of the Leloir pathway in *S. cerevisiae*, the catabolic route for galactose metabolism, because this regulon has served as a paradigm for understanding eukaryotic transcription principles (9). Consequently, much information on the regulation and structure of the components involved in galactose metabolism has accumulated. There is also a vast amount of high throughput data available that elucidates galactose metabolism (10). Exploiting this abundant information, many elegant metabolic engineering approaches have been implemented to increase the galactose uptake rate by modification of transporters, regulators, metabolic genes, or a combination of them (11, 12). Based on analysis of these and other strains, it has been found that accumulation of metabolic intermediates in the galactose metabolism, such as galactose-1-phosphate and glucose-1-phosphate, may inhibit the flux through the Leloir pathway and, hence, lead to a lower galactose uptake rate (10). Therefore, successful strain development was performed by balanced expression of structural genes through modification of the regulatory system (12) or through overexpression of the final enzyme of galactose metabolism, *PGM2*, which converts glucose-1-phosphate to glucose-6-phosphate (11). Both these engineered cells had lower concentrations of the intermediates and a higher galactose uptake rate (8). Despite these successes on improving galactose uptake in yeast through metabolic engineering, there has so far not been any description of using an evolutionary approach for improving galactose utilization.

Adaptive evolution of bacteria has enabled increasing the specific growth rate due to mutations that were not predicted by rational engineering (5). We therefore decided to apply the concept of adaptive evolution for the improvement of galactose utilization by yeast, with the objective to evaluate if unique strategies for improving galactose uptake could be identified. Furthermore, through detailed characterization of evolved mutants, we expected to expand our understanding of the galactose metabolism in yeast. We therefore characterized adaptively evolved mutants at the systems level for gaining better understanding of the molecular mechanisms that are responsible for acquired phenotypes. We evolved a laboratory strain of yeast for ≈ 400 generations in three different serial transfer lines and analyzed the changes in transcriptome, metabolome, and genome sequence that contribute to the phenotypic changes. Here, we present results of integrated analysis of the data from the three evolved mutants compared with wild-type yeast and two engineered

Author contributions: K.-K.H., G.N.V., and J.N. designed research; K.-K.H. performed research; K.-K.H. contributed new reagents/analytic tools; K.-K.H., W.V., G.N.V., and J.N. analyzed data; and K.-K.H., G.N.V., and J.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. [GSE27185](#)).

¹Present address: Center for Systems Biology, Soochow University, Suzhou 215006, China.

²To whom correspondence should be addressed. E-mail: nielsenj@chalmers.se.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103219108/-DCSupplemental.

strains that were developed by the rational approach in previous studies (11, 12). Based on our analysis of the evolved strains, unique strategies for improving galactose uptake were identified, and one of these strategies was proven to result in an improved galactose uptake. Furthermore, our analysis highlights that it is only by an integrated systems biology approach that it is possible to map out the mechanisms underlying evolved phenotypes.

Results

Physiological Changes in Evolved Mutants. Three adaptively evolved strains that had 24% faster specific growth rates on galactose were obtained through 62-d serial transfers of cultures to fresh medium with galactose as the sole carbon source. The strains were isolated from the last culture and were designated 62A, 62B, and 62C. The gross phenotype of the three evolved strains was compared with those of the reference strain (CEN.PK113-7D) and two engineered strains characterized in earlier studies: SO16 with deletion of *MIG1*, *GAL80*, and *GAL6* (12) and PGM2 with overexpression of *PGM2* (11). All mutants showed an improved specific galactose uptake rate compared with the reference strain (Fig. 1 and Table S1). The difference between the evolved mutants and the engineered strains was found in the biomass yield, the specific ethanol production rate, and the specific growth rate (Fig. 1). The engineered strains exhibited the highest ethanol yield at the expense of the biomass yield, whereas the evolved mutants showed a similar biomass yield with the reference strain. The evolved mutants from the three lineages commonly exhibited a 24% increase in the maximum specific growth rate compared with the reference strain, whereas they differed in their specific galactose uptake rates and their specific ethanol production rates. The specific galactose uptake rates varied from an 18% increase in 62A strain to a 36% increase in 62C. The specific ethanol production rate in the evolved mutants was increased from 31 to 170%. In an earlier study, we found that

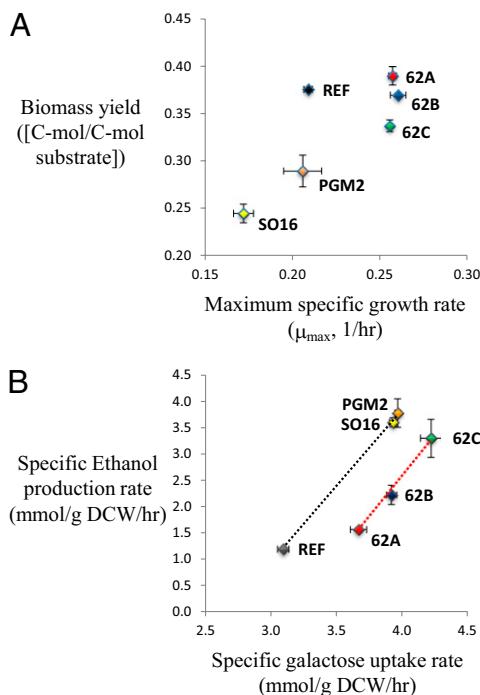


Fig. 1. Phenotypic changes of evolved mutant strains 62A, 62B, and 62C compare with the reference strain CEN.PK113-7D and the two engineered strains SO16 and PGM2. (A) Correlation between a maximum specific growth rate and biomass yield. (B) Correlation between a specific galactose uptake rate and a specific ethanol production rate. The regression curves of the two lines (from right to left) have a slope of 2.953 and 3.187 and intercept of -7.9472 ($R^2 = 0.991$) and -10.157 ($R^2 = 0.98$), respectively.

several engineered strains were lying on a linear regression curve when their specific ethanol production rate was plotted against the specific galactose uptake rate (12). This experience led us to plot data from all of the strains in a similar kind of plot, but we found there to be a grouping pattern that clearly separated all of the strains into two groups (Fig. 1B). The reference and the two engineered strains were on the same regression curve ($R^2 = 0.99$), whereas all of the evolved mutants were on a different regression curve ($R^2 = 0.98$), indicating a common phenomenon underlying the increased galactose uptake rates. If data from all of the strains are included in the same linear regression, a rather poor correlation coefficient is obtained ($R^2 = 0.64$).

Changes in the Transcriptome and the Metabolome. The evolved mutants showed clear separation from the reference and the two engineered strains in their transcriptome profile, using principal component analysis (PCA) (Fig. 2A). The first principal component (PC) (35%) separated the evolved mutants from the other strains. Among the differentially expressed genes between these two groups, the KEGG pathways involving trehalose and glycogen metabolism were overrepresented ($P < 1e-4$).

Transcriptional differences ($P < 0.01$) between the evolved mutants and the reference strain were categorized into those that commonly changed in the three lineages and those that were mutant-specific (Fig. 2B). Genes involved in trehalose and glycogen metabolism were commonly up-regulated in all of the evolved mutants, whereas genes encoding proteins involved in the MAPK signaling pathway (also based on KEGG pathway)

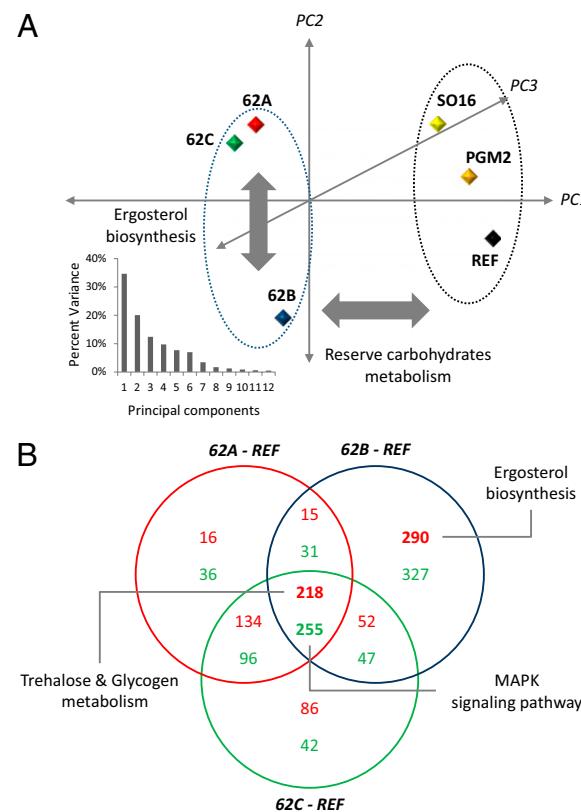
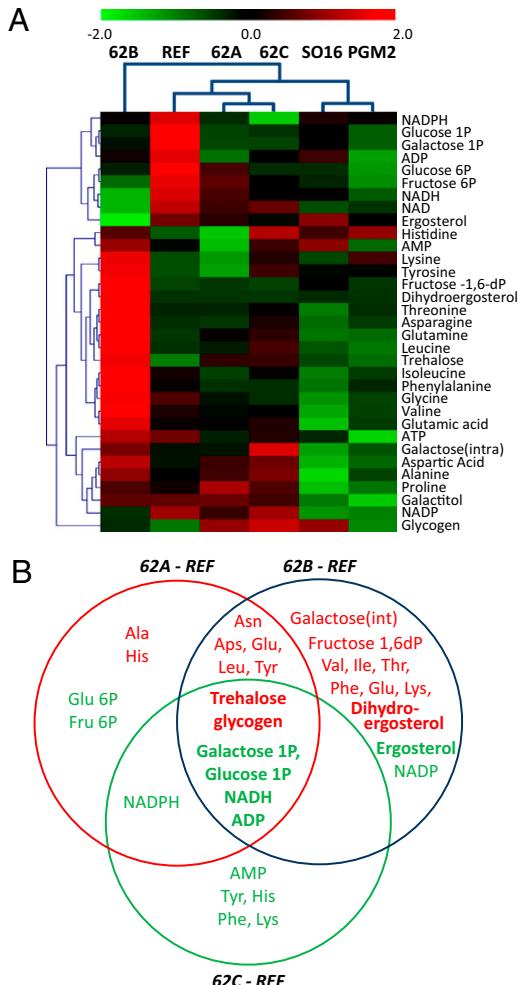


Fig. 2. Results from transcriptome analysis. (A) Principle component analysis of all of the strains is used to highlight the differences among strains. The box shows how much variance is observed by the different principle components. The results are projected by the first three PCs, which covered 85% of the variance. (B) Differentially expressed genes ($P < 0.01$) are categorized in a Venn diagram. The functions of genes in each part are analyzed by hypergeometric model using KEGG, Reactome, and GO term database ($P < 0.001$). Red color letter means up-regulation; green one means down-regulation.



were down-regulated, relative to the reference strain (cumulative hypergeometric probability, $P < 1e-4$). Interestingly, few genes appeared to be differentially expressed between 62A and 62C, even though they showed very different physiology. However, several genes (≈ 600 genes) were differentially expressed in 62B, and the genes involved in ergosterol biosynthesis were overrepresented among these genes.

We also measured the concentration of intracellular metabolites of the Leloir pathway, redox cofactors, and amino acids. Prompted by the transcriptional changes in genes involved in glycogen, trehalose, and ergosterol metabolism (Table S2), we also measured the levels of these compounds in the cell. Hierarchical clustering of all of the metabolites separated 62B from the other two evolved mutants, similar to the transcriptome data (Fig. 3A). The evolved and engineered strains commonly showed lower intracellular concentrations of sugar phosphates than the reference strain. In general, the level of free amino acids was higher only in 62B, relative to the other strains. The two engineered ones exhibited the lowest levels of amino acids.

Significantly changed metabolites were classified as those whose intracellular concentration changed commonly in all of the evolved mutants or in one specific strain compared with the reference strain ($P < 0.05$) (Fig. 3B). Trehalose and glycogen

were at a higher concentration only in all of the evolved mutants, whereas the concentrations of galactose-1-phosphate and glucose-1-phosphate, NADH, and ADP were lower in both the engineered and evolved strains. Interestingly, despite having increased transcriptional activity of the ergosterol pathway, the intracellular concentration of ergosterol was much lower in 62B. However, 62B was found to produce a higher amount of dihydroergosterol compared with any of the other strains, indicating a redirection of carbon flux in the ergosterol biosynthesis pathway to this metabolite.

Changes in Genotype. The genomes of the evolved mutants were sequenced and compared with that of the parent strain, CEN.PK113-7D (13) (Table S3). Only a small fraction of the genome of the reference strain (0.7–1.4%) was not covered by the reads for the three mutant strains, allowing for high quality mapping of the genome sequence of the three evolved mutants. Based on the raw sequence data, we identified in the order of 6,000 putative mutations in all of the three evolved strains, but after filtering the number of mutations, was reduced to 44, 334, and 40 in the 62A, 62B, and 62C, respectively (Table S4, Fig. S1, and Dataset S1). We aimed for 20 \times coverage in the sequencing, but we received a substantially higher coverage for strain 62B ($\approx 55\times$) (Table S3). We applied a filtering process, and this resulted in maintaining a larger number of mutations in the 62B strain, which is likely due to the high sequencing coverage for this strain. Of the identified mutations, 21, 104, and 29 were single nucleotide polymorphism (SNPs), respectively, whereas the remainder were insertions and deletions (INDELS). Only about one-third of the SNPs were in coding regions, and only very few of the INDELS were in coding regions (none in 62A, 11 of 230 in 62B, and 3 of 11 in 62C). The number of mutations is much larger than observed in an earlier study on evolution of *E. coli* (14), but we believe that this difference could be explained by the rather poor growth of the reference strain on galactose and possible also by the larger genome size of *S. cerevisiae*. Surprisingly, no mutations were detected in galactose regulatory and structural genes, even in *PGM2*, which was considered the most beneficial target for increasing the galactose uptake rate in previous studies based on metabolic engineering (Dataset S2) (11). Furthermore, no mutations in the trehalose and glycogen pathway genes were found, which showed significant changes in the transcriptome and metabolome level in all of the evolved mutants. Only genes encoding proteins of the Ras/PKA signaling pathway were found to carry mutations in all three evolved mutants (Table 1). 62A and 62B had a mutation in *RAS2* at different positions, whereas 62C had one mutation in *CYR1*, encoding adenylate cyclase. Both genes are related to the cAMP-dependent stress response signaling pathway. A mutation in *ERG5* in 62B was identified, which is a gene encoding one of the enzymes of the ergosterol pathway. To evaluate whether these mutations are causing the increased specific growth rate on galactose, we reconstructed the mutation in position 112 of Ras2p resulting in an amino acid substitution of aspartate with tyrosine. The resulting strain was evaluated for its growth on galactose in shake flasks, and it was found to have a 10% higher specific growth rate than the reference strain ($P = 0.05$) (Table S5).

Discussion

In this study, we characterized adaptive evolution of yeast to acquire faster growth and galactose utilization. First, we analyzed

Table 1. Genetic changes

Strains	Mutations	Functions
Commonly mutated pathway		
62A	<i>RAS2</i> [Gln ⁷⁷ →Lys]	Ras/PKA signaling pathway
62B	<i>RAS2</i> [Asp ¹¹² →Tyr]	
62C	<i>CYR1</i> [Asp ⁸²² →Asn]	
Uniquely mutated genes		
62B	<i>ERG5</i> [Arg ³⁷⁰ →Pro]	Ergosterol metabolism

Fig. 3. Analysis of the metabolite data from the evolved strains compared with the reference and the two recombinant strains. (A) Hierarchical clustering of all strains is computed after standardizing of metabolites concentration to z score. (B) Differentially produced metabolites are categorized in a Venn diagram by calculation of significance ($P < 0.05$) based on a Student t test. Red color letter means higher levels; green one means lower levels.

the phenotypic differences by gross kinetic parameters such as maximum specific growth rate, specific galactose uptake rate, and yield coefficients. These values indicated that the three evolved mutants had obtained improved growth on galactose by obtaining different mutations and also by different means than the strategy applied in metabolic engineering studies for improving the galactose uptake. To understand the underlying metabolic changes that conferred the phenotypic differences of the evolved mutants, we measured the variation of the transcriptome and the metabolome of the evolved strains and compared these with the reference strain and the earlier constructed engineered strains. Changed transcription of genes involved in trehalose and glycogen metabolism were detected in all three evolved strains, and genes involved the ergosterol pathway were noticed to have uniquely changed expression in the 62B strain. From comparative genome analysis of the evolved mutants, we could find no mutations in galactose and reserve carbohydrates metabolisms, whereas proteins of the Ras/PKA signaling pathway was discovered to contain mutations in all of the evolved mutants. Furthermore, 62B had a unique missense mutation in *ERG5*.

Adaptive Evolution Achieves Improved Galactose Availability with Different Physiology. All mutants, including the evolved mutants and the two engineered strains, showed lower CO₂ yield than the reference strain, even though they had a higher ethanol yield (Table S1). The lower CO₂ yield along with the higher ethanol yield indicates a reduced TCA cycle activity. The significant decrease in the NADH concentration partially supports this phenomenon because NADH is the main cofactor of the TCA cycle (Fig. 3B). All mutants showed lower levels of sugar phosphates, such as galactose-1-phosphate and glucose-1-phosphate, than the reference strain (Fig. 4A). Decrease in their concentration is likely due to an increased flux downstream of these

metabolites. In the case of the two engineered strains, this reduction is realized solely by amplification of *PGM2*, which is a final enzyme in galactose metabolism converting glucose-1-phosphate to glucose-6-phosphate (11). For the evolved mutants there is, however, besides increased expression of *PGM2*, likely to be an increased flux through trehalose and glycogen.

Another common feature of the evolved mutants and the engineered strains is that in a plot of the specific ethanol production rate versus the specific galactose uptake rate, the three evolved mutants lie on one regression line with a slope of ≈ 3 and the engineered strains lie on another regression line together with the reference strain also with a slope of ≈ 3 (Fig. 1B). The results for the engineered strains are consistent with our earlier findings that engineering of the GAL-regulon results in a slope >2 in this kind of plot (12). If an increase in galactose uptake resulted solely in ethanol production, this slope should be 2, but because both sets of strains are ≈ 3 , it shows that an increase in the galactose uptake results in a redirection of flux from respiratory metabolism, i.e., TCA cycle, to fermentative metabolism. Thus, when the galactose uptake is increasing, then carbon catabolite repression of the respiratory systems sets in, but the transcriptome analysis did not provide any indication of increased carbon catabolite repression and/or decreased respiratory metabolism in the metabolically engineered strains compared with the reference strain and in 62C and 62B compared with 62A. Despite the similarity in slope for the two sets of strains, it is interesting that the evolved mutants seems to find a different metabolic operation that allows a higher galactose uptake without increasing the ethanol production, probably due to the redirection of flux through the storage carbohydrates glycogen and trehalose.

Up-Regulation of *PGM2* and Activation of Reserve Carbohydrates Metabolism Are Detected as Common Changes in All Evolved Mutants. The significant changes in the levels of metabolic intermediates of the Leloir pathway did not arise because of transcriptional differences between the strains, except for *PGM2* (Fig. 4B). The concentration of galactose-1-phosphate and glucose-1-phosphate were much lower in all of the evolved mutants compared with the reference strain (Fig. 4A). Both the changes in the levels of sugar phosphates and expression of *PGM2* were common with the two engineered strains. These results reveal that adaptive evolution used partially the same strategy as designed using rational engineering. However, for the evolved strains, trehalose and glycogen metabolism showed significant change in the transcriptome and metabolome data in all three evolved mutants compared with the other strains (Figs. 2 and 3). The concentration of glycogen and trehalose were clearly increased (Fig. 4A), and the transcription level of structural genes in the metabolism of these storage carbohydrates was remarkably increased (Fig. 4B). The induction of this pathway for increasing galactose utilization could be explained simply by the fact that glucose-1-phosphate is used as a precursor for the production of these reserve carbohydrates. Ugp1p converts glucose-1-phosphate to UDP-glucose, which is a branch point metabolite in the trehalose and glycogen metabolism.

Trehalose and glycogen are reserve carbohydrates for maintaining the energy charge of yeast cells (15). Their metabolism is very closely linked to the concentration of glucose, which is the most favored carbon source to yeast (16). Upon sensing glucose limitation, yeast mobilizes energy reserve to balance the rate of glycolytic activity and retain the energy charge. Trehalose production is induced especially under stress condition. However, galactose seems not to trigger activation of glycogen and trehalose metabolism. Wild-type yeast showed no change of the concentration of trehalose and the glycogen during growth on galactose (17). The evolved mutants probably have higher energy charge and, therefore, are capable of having an increased specific growth rate. This interpretation is supported by higher ratio of ATP/ADP in these cells (Fig. 4A; ref. 18).

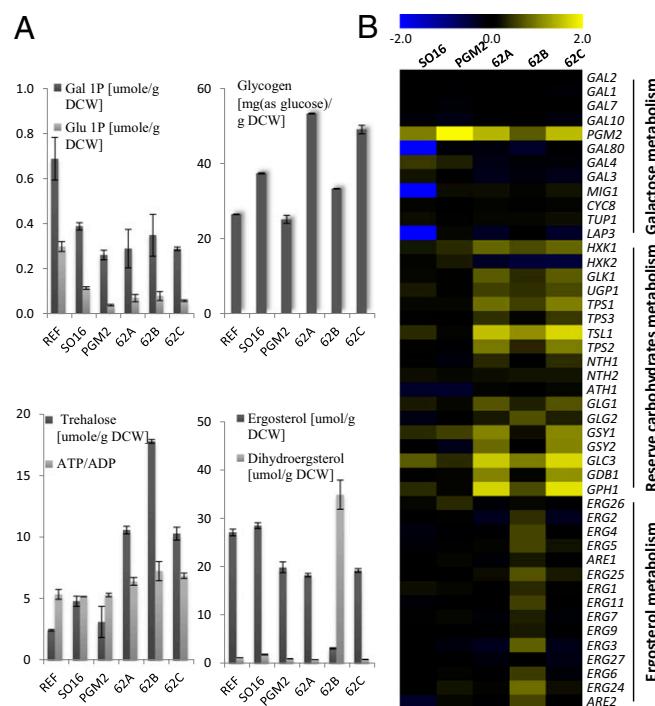


Fig. 4. Changes in the galactose, reserve carbohydrates, and ergosterol metabolism in the evolved mutants are illustrated by changes in the concentration of metabolites and fold changes of transcriptome compared with the other strains. (A) The concentrations of sugar phosphates, storage carbohydrates, and sterols and the ratio of ATP to ADP. (B) Fold changes of all genes involved in galactose, reserve carbohydrates, and ergosterol metabolism are compared with the reference strain.

Adaptive Evolution of Yeast on Galactose Generates No Mutations in Galactose and Reserve Carbohydrates Metabolism. In general, when microorganisms face a new carbon source, they evolve the structural or regulatory genes that are related to the metabolism of it. For example, all adaptively evolved mutants of *E. coli* on glycerol had mutations commonly in *glpK*, glycerol kinase (5). However, in our case, no mutations were identified in the galactose metabolism, including promoter regions of the *GAL* genes (there were also no considerable changes in transcription levels). Furthermore, no mutations were detected in genes involved in the reserve carbohydrates metabolism, but there were significant changes in both the transcripts and the levels of trehalose and glycogen. These findings indicate that the phenotypic changes are consequences of mutations in regulatory systems and, indeed, all evolved mutants had mutations in proteins involved in the Ras/PKA signaling pathway, which is related to *PGM2* overexpression and activation of reserve carbohydrates metabolism (19). Reduced activity of Ras2 or Cyr1 lead to a decreased concentration of cAMP, and lower levels of this metabolite can release the blocking effect of PKA on the transcription factors Msn2/4. Release of the Msn2/4 transcription factors from PKA control results in up-regulation of genes having STRE elements in their promoter region, such as *PGM2* and *UGP1*. 62A and 62B have mutations commonly in *RAS2* (*RAS2^{Lys77}* and *RAS2^{Tyr112}*, respectively), whereas 62C has a mutation in *CYR1* (*CYR1^{Asn822}*). *RAS2^{Lys77}* in 62A and *CYR1^{Asn822}* in 62C appear to exert similar control, as reflected by the close patterns in their transcriptome and metabolome profiles. In contrast, 62B showed a very different pattern in its omics data, even though it had mutations in *RAS2* like 62A. The trehalose concentration in this strain was the highest among the evolved mutants, whereas the concentration of glycogen was the lowest (Fig. 4A). Also, the transcript level of this strain in reserve carbohydrates metabolism was much lower than for the other two evolved mutants, the 62A and 62C strains (Fig. 4B). These unique features of 62B indicate that the changes of the reserve carbohydrates metabolism caused a mechanism that is different from that in 62A and 62C. It is interesting to note that the ergosterol biosynthesis and the concentration of ergosterol showed marked changes in this strain (Fig. 4A). A unique mutation was identified in *ERG5^{Pro370}*. It has been found that a knockout mutant of *ERG5* (desaturase) produces dihydroergosterol instead of ergosterol (20). Because 62B had low levels of ergosterol and high levels of dihydroergosterol, it seems plausible that the mutation in *ERG5* results in reduced activity of Erg5p, and we speculate that this declined activity may also be closely linked to the increased trehalose production. Both ergosterol and trehalose are regarded as protectants for stress response (21), and a changed sterol composition in the cell membrane can induce trehalose accumulation. The 62B strain will have changed cell membrane rigidity, because dihydroergosterol has more loosened structure than ergosterol because of the loss of one double bond (22), and this looseness may trigger the accumulation of trehalose and, consequently, affect the galactose metabolism by consuming glucose-1-phosphate as a precursor of trehalose. The effect of one of the identified mutations in the Ras/PKA signaling pathway was evaluated, and site-directed mutagenesis of amino acid 112 in Ras2p resulted in an increased specific growth rate on galactose.

The different strategies of adaptive evolution of yeast for improving galactose metabolism are summarized in Fig. 5. As mentioned above, increasing the flux through trehalose and glycogen results in a drain of glucose-1-phosphate, which may have a positive effect on the galactose uptake. Furthermore, the increased flux toward glycogen and trehalose may lead to increased levels of UDP-glucose, which is a cosubstrate in the conversion of galactose-1-phosphate to glucose-1-phosphate, and this reaction may lead to improved conversion of galactose-1-phosphate, which also acts as a feed-forward inhibitor of Pgm2p (8). That increased concentration of UDP-glucose may have a positive effect on the flux through the Leloir pathway is partly supported by earlier findings that the Gal7/Gal10 enzyme system is not controlling the flux through the pathway (23).

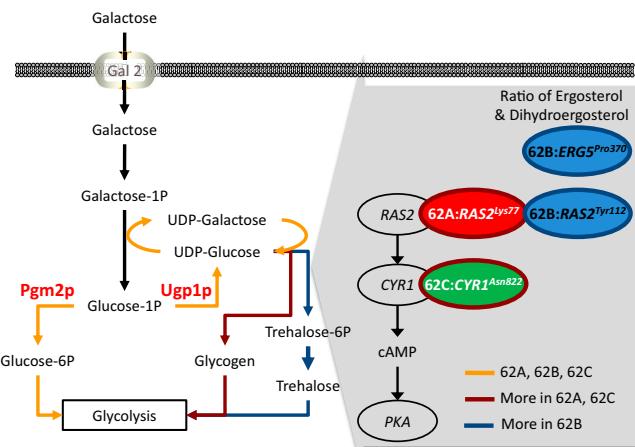


Fig. 5. Summary of evolution changes in the three evolved mutants; 62A, 62B, and 62C. Color circular boxes indicate genes having genetic mutations. Color lines indicate activated fluxes inferred from transcriptome and metabolome analysis.

In conclusion, our study showed that through adaptive evolution the cells may find different ways to ensure an increased flux through the Leloir pathway. This pathway is, despite its few steps, quite complex because it involves the conversion of galactose-1-phosphate to glucose-1-phosphate with the concurrent conversion of UDP-glucose to UDP-galactose by galactose-1-phosphate uridylyltransferase (encoded by *GAL7*), and further regeneration of UDP-glucose from UDP-galactose by UDP-glucose 4-epimerase (encoded by *GAL10*). Because there is further feed-forward inhibition of galactose-1-phosphate on phosphoglucomutase (encoded by *PGM2*), it is clear that activity of the Gal7/Gal10 enzyme system is very important for proper function of the pathway. We, however, earlier found that over-expression of Gal7/Gal10, either alone or together with the other structural *GAL* genes, does not result in an improved flux through the pathway (on the contrary the flux was decreased), indicating a high level of metabolic regulation. Through adaptive evolution, the cells find a way to circumvent this regulation. It up-regulates expression of *PGM2*, which partly takes care of the problem of feed-forward inhibition by galactose-1-phosphate, as in the engineered strains, but it further up-regulates the flux from glucose-1-phosphate to glycogen and trehalose. This alteration is likely to result in increased levels of UDP-galactose that may allow an increased flux through the Gal7/Gal10 enzyme system and, hence, an increased flux through the pathway. Thus, as in other studies on adaptive evolution, it is clear that this strategy allows the generation of new strategies that cannot be found from a rational approach. The most crucial advantage of adaptive laboratory evolution is the finding of unpredictable and unexpected beneficial mutations. In *E. coli*, this phenomenon is well known (5). Thus, when this bacterium was evolved for improved growth on glycerol, it resulted in mutations in *glk* (glycerol kinase) that is a key enzyme in the glycerol pathway. However, more effective mutations were detected in unpredicted genes such as RNA polymerase β and β' subunits (24). In the case of yeast, adaptive laboratory evolution on galactose generated no mutations in the galactose pathway but generated unforeseen ones in other pathways such as carbon regulatory pathway and ergosterol biosynthesis. The main hurdle of galactose metabolism in yeast may therefore not be in Leloir pathway, and increased expression of the *GAL* genes results in reduced galactose metabolism (12). The beneficial changes are detected in the consuming reaction of glucose-1-phosphate that is directly linked to glycolysis and reserve carbohydrates metabolism, and these changes comes around through mutations in the regulatory PKA pathway.

Besides our findings on unique strategies for improving the galactose uptake, our study clearly provides two key lessons for success in terms of identifying the underlying genotypes of mutants with improved phenotypes:

It is essential to combine detailed phenotypic analysis, e.g., involving transcriptome and metabolome analysis, with genome sequencing. Each of these techniques do not allow for drawing solid conclusions, but combined they provide a clear picture of the consequences of identified mutations.

It is important to analyze several evolved mutants with different control strains because this comparison allows for identification of conserved mutations that result in the same phenotype. Each of the three evolved mutants has several mutations that probably do not contribute to the evolved phenotype, but by identifying conserved mutations, a clear picture emerged.

From these two lessons, we are confident that there is opened up for wider use of systems biology and genome-sequencing for identifying the underlying genotypes for evolved phenotypes in eukaryotic cells.

Materials and Methods

Yeast Strains and Adaptive Evolution. *S. cerevisiae* CEN.PK113-7D was used as a reference and starting strain for the adaptive evolution. Two engineered strains, SO16 ($\Delta gal4 \Delta gal80 \Delta mig1$) and PGM2 (overexpression of *PGM2*) were constructed in previous study (11, 12). Three adaptively evolved strains, 62A, 62B, and 62C, were generated from CEN.PK 113-7D after daily serial dilution for 62 d (≈ 400 generations) on galactose (20 g/L) minimal media. Three culture lines were independently carried out at 30 °C and $\times 0.76$ g in cotton-covered 500-mL Erlenmeyer flasks with baffles with 100 mL of media.

1. Nielsen J (2001) Metabolic engineering. *Appl Microbiol Biotechnol* 55:263–283.
2. Patnaik R (2008) Engineering complex phenotypes in industrial strains. *Biotechnol Prog* 24:38–47.
3. Wisselink HW, Toirkens MJ, Wu Q, Pronk JT, van Maris AJ (2009) Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered *Saccharomyces cerevisiae* strains. *Appl Environ Microbiol* 75: 907–914.
4. José Manuel Otero DC, Patil KR, Poulsen SG, Olsson L, Nielsen J (2011) Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory. *PLoS ONE*, in press.
5. Herring CD, et al. (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 38: 1406–1412.
6. Ikeda M, Ohnishi J, Hayashi M, Mitsuhashi S (2006) A genome-based approach to create a minimally mutated *Corynebacterium glutamicum* strain for efficient L-lysine production. *J Ind Microbiol Biotechnol* 33:610–615.
7. Albers E, Larsson C (2009) A comparison of stress tolerance in YPD and industrial lignocellulose-based medium among industrial and laboratory yeast strains. *J Ind Microbiol Biotechnol* 36:1085–1091.
8. de Jongh WA, et al. (2008) The roles of galactitol, galactose-1-phosphate, and phosphoglucomutase in galactose-induced toxicity in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 101:317–326.
9. Ramsey SA, et al. (2006) Dual feedback loops in the GAL regulon suppress cellular heterogeneity in yeast. *Nat Genet* 38:1082–1087.
10. Ideker T, et al. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292:929–934.
11. Bro C, Knudsen S, Regenberg B, Olsson L, Nielsen J (2005) Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* 71:6465–6472.
12. Ostergaard S, Olsson L, Johnston M, Nielsen J (2000) Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat Biotechnol* 18:1283–1286.
13. Otero JM, et al. (2010) Whole genome sequencing of *Saccharomyces cerevisiae*: From genotype to phenotype for improved metabolic engineering applications. *BMC Genomics* 11:723–740.
14. Barrick JE, et al. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.
15. François J, Parrou JL (2001) Reserve carbohydrates metabolism in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* 25:125–145.
16. Hazelwood LA, et al. (2009) Identity of the growth-limiting nutrient strongly affects storage carbohydrate accumulation in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 75:6876–6885.
17. van den Brink J, et al. (2009) Energetic limits to metabolic flexibility: responses of *Saccharomyces cerevisiae* to glucose-galactose transitions. *Microbiology* 155: 1340–1350.
18. Boer VM, Crutchfield CA, Bradley PH, Botstein D, Rabinowitz JD (2010) Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. *Mol Biol Cell* 21:198–211.
19. Estruch F (2000) Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiol Rev* 24:469–486.
20. Pierce AM, Unrau AM, Oehlschlager AC, Woods RA (1979) Azasterol inhibitors in yeast. Inhibition of the delta 24-sterol methyltransferase and the 24-methylene sterol delta 24 (28)-reductase in sterol mutants of *Saccharomyces cerevisiae*. *Can J Biochem* 57:201–208.
21. Swan TM, Watson K (1999) Stress tolerance in a yeast lipid mutant: membrane lipids influence tolerance to heat and ethanol independently of heat shock proteins and trehalose. *Can J Microbiol* 45:472–479.
22. Kelly SL, et al. (1995) Purification and reconstitution of activity of *Saccharomyces cerevisiae* P450 61, a sterol delta 22-desaturase. *FEBS Lett* 377:217–220.
23. Ostergaard S, Olsson L, Nielsen J (2001) In vivo dynamics of galactose metabolism in *Saccharomyces cerevisiae*: Metabolic fluxes and metabolite levels. *Biotechnol Bioeng* 73:412–425.
24. Applebee MK, Herrgård MJ, Palsson BO (2008) Impact of individual mutations on increased fitness in adaptively evolved strains of *Escherichia coli*. *J Bacteriol* 190: 5087–5094.

The cells were cultivated until they reached midexponential phase, before they were transferred to new fresh media. Single clone isolates were obtained from the last shake flasks. The construction of site-directed mutant (*RAS2*^{Tyr112}) and growth rate measurements in a flask is explained in *SI Materials and Methods*.

Batch Fermentation and Measurement of Cell Mass and Extracellular Metabolites. Biological duplication of all strains was performed from seed culture to sample preparation for omics data analyses. Additional details are provided in *SI Materials and Methods*. The dry cell weight and extracellular metabolites were determined as described (12).

Transcriptome Analysis. Affymetrix Yeast Genome 2.0 Array was used for transcriptome analysis. Gene expression data were deposited to the Gene Expression Omnibus (GEO) database with accession number GSE27185. Detailed methods are included in *SI Materials and Methods*.

Metabolome Analysis. Quenching and extraction of intracellular metabolites were done with biological duplicates of all strains as described (8). Venn diagram was used to represent common and specific features of all of the evolved mutants.

Illumina/Solexa Genome Sequencing. Genome sequencing was performed by Fasteris SA who used Illumina/Solexa technology. Details are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Aldo Jesorka for assistance with dihydroergosterol analysis and Rahul Kumar for valuable discussion. We thank CJ CheiJedang Doctoral Fellowship for financial support (to K.-K.H.). We acknowledge the Chalmers Foundation, The Knut and Alice Wallenberg Foundation, the European Union funded projects UNICELLSYS (Contract 201142) and SYSINBIO (Contract 212766), and European Research Council Grant 247013 for financial contributions to this project.

Supporting Information

Hong et al. 10.1073/pnas.1103219108

SI Materials and Methods

Site-Directed Mutant Generation and Measurements of a Maximum Growth Rate.

i) Steps for site-directed mutagenesis

Step 1) Two cassettes preparation by fusion PCR of Partial *RAS2^{Tyr112}* and *URA3*

Partial *RAS2^{Tyr112}* (805 bp) Partial 5' *URA3* (716 bp)

Partial 3' *URA3* (1,028 bp)

Partial *RAS2^{Tyr112}* (805 bp)

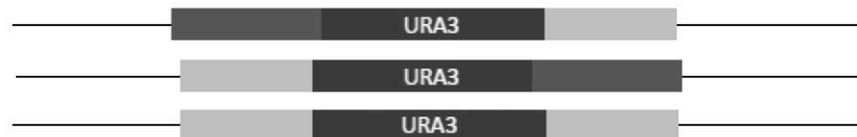
Transformation into CEN.PK113-5D (*ura3*)

Step 2) Homologous recombination for introducing *RAS2^{Tyr112}* with *URA3* marker



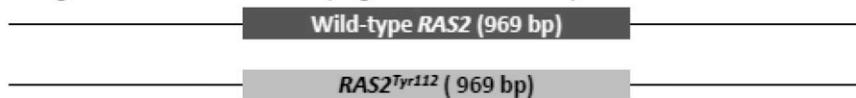
Spreading on uracil media

Step 3) Selection strains that have *URA3* marker



Streaking on 5-FOA media
(5-Fluoroorotic Acid Monohydrate)

Step 4) Homologous recombination for looping out *URA3* marker by 5-FOA media based selection



Colony PCR with mutant specific primers
Whole RAS2 gene sequencing

Step 5) Selection mutant (*RAS2^{Tyr112}*) by PCR with mutants specific primers and confirming by sequencing

RAS2^{Tyr112} (969 bp)

ii) Primer lists for construction of site-directed mutants

For fusion of <i>RAS2</i> and <i>URA3</i>	
RAS2-5P1	CCACTCTTATCTGACTCTCTGC
RAS2-3P1	CCAATTGCCATGAAGCCGAA
	TTCACATTTACCGTTGGCAGC
For fusion of <i>RAS2</i> and <i>URA3</i>	
RAS2-5P2	GGAATTTCGATATCAAGCTTATCGAT
RAS2-3P2	CCACTCTTATCTGACTCTCTGC
	TTCACATTTACCGTTGGCAGC
For selection of mutants	
RAS2-4BW (wild-type)	AATTGGAACATAGTCGGTATC
RAS2-4BM (mutant)	AATTGGAACATAGTCGGTATA
For sequencing of whole <i>RAS2</i> genes	
RAS2-5P	GTTCCTAGCCGGTGTCTCTT
RAS2-3P	GTTCCTTCGCTTAGCGTTT

iii) Characterization of maximum growth rate of site-directed mutants in a flask culture

Biological triplicates of all strains were performed in flask scale. Uracil (150 mg/L) was supplied to galactose (20 g/L) minimal media. *P* value was calculated by assumption of one-tailed distribution and homoscedasticity.

Batch Fermentation and Sampling. Aerobic batch cultures were carried out in 1.0-L DasGip stirrer-pro vessels (DasGip) containing 0.8-L working volume. The same media with preculture was used, and the temperature was kept at 30 °C. The airflow rate was set to 1 L per liquid volume per minute, and the pH of the media was maintained at 5.0 by automatic addition of 2 M KOH. All conditions of temperature, agitation, air supply, pH, and off-gas analysis were controlled and monitored by the DasGip system. The concentration of dissolved oxygen was measured with an autoclavable polarographic oxygen electrode (Mettler Toledo). The off-gas from the bioreactors was analyzed for real-time determination of oxygen and carbon dioxide concentration by DasGip fedbatch pro gas analysis systems with the off gas analyzer 1 GA4 based on zirconium dioxide and two-beam infrared sensor (DasGip). Samples for measuring biomass and extracellular metabolites were obtained hourly after starting the exponential growth phase. When the concentration of carbon dioxide reached ≈1.7%, which was almost midexponential phase in the aerobic batch culture with 20 g/L galactose, samples for DNA microarray and intracellular metabolite analysis were taken.

Transcriptome Analysis. PCA was applied to normalized data from all strains by using the TM4 software (1). To find common changes and specific changes of evolved mutants, differentially expressed genes (Student's *t* test; *P* < 0.01) were identified by comparison among all evolved mutants and the other strains. Differentially expressed genes were used as input to g:Profiler to find significantly changed metabolic pathways, reactions, and GO terms by using a cumulative hyper-geometric test (2). The results

were visualized by using software of MultiExperiment Viewer (Dana-Farber Cancer Institute).

Metabolome Analysis. The concentration of sugar phosphates was measured by ion chromatography (3). Quantification of redox cofactors was done by assay kits of BioVision. Free amino acids were quantified by EZ:faast amino acid analysis kit (Phenomenex) by using gas chromatography-mass spectrometry. To analyze carbohydrates and sterols, harvested cells were directly used for extraction of those metabolites; 10 mg and 100 mg of dry cell weight were used, respectively. Trehalose and glycogen were quantified by measuring the amount of glucose released after treatment with enzymes that degrade these carbohydrates (4). Ergosterol and dihydroergosterol were identified by gas chromatography-mass spectrometry (5) and quantified by high performance liquid chromatography. It was difficult to find a standard for dihydroergosterol, and we therefore quantified this by using a standard curve with ergosterol. For hierarchical clustering, all quantified metabolites were standardized to *z* score to remove different weights for each metabolite. The result of the clustering was visualized by using the software of Multi-Experiment Viewer. To analyze significant changes in metabolite concentrations, a Student *t* test (*P* < 0.05) was applied between the three evolved mutants and the reference strain, respectively. Venn diagram was used to represent common and specific features of all of the evolved mutants.

Illumina/Solexa Genome Sequencing. Genome Analyzer (GA-IIx) was used with sequencing cycles of 2 × 38 + 7 (index) by Chrysalis 36 cycles version 4.0 as a sequencing kit. The Mapping and Assembly with Quality (MAQ) software version 0.7.1 (<http://maq.sourceforge.net>) was used with a maximum set at 2 mismatches in the first 24 bases on the reference sequence that was original CEN.PK113-7D genome sequence (6) available at the CENPK genome database (www.sysbio.se/cenpk). The MAQ software was then used to validate the mapping obtained by comparing the paired end of each read orientation and position. The maximal insert size was set to 400 base pairs for this process. The results from the mapping were further used for calling of single nucleotide polymorphisms (SNPs) and insertions and deletions (INDELS) in the reads compared with the reference sequence. The raw candidates of SNPs and InDels having identical mutations in all evolved mutants were removed for further comparative analysis, because those mutations came from the ancestor strain. It was confirmed by resequencing of some of identical mutations with an ancestor strain. Filtering for SNPs was done by selecting of SNPs with Phred-like consensus quality >30 and a coverage depth >3. Some of INDELS were filtered out when the number of reads with INDELS was lower than the number of reads without INDELS. To perform annotation, GlimmerHMM was used for prediction the gene structures by training the gene models from strain CEN.PK113-7D and S288c (6). Especially, to calculate the coverage folds of gal genes in each evolved mutant, the count function of integrative genome viewer (IGV software) was used. A 200-bp window was set to estimate the average coverage folds. The results are summarized in Dataset S2.

- Saeed AI, et al. (2003) TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* 34:374–378.
- Reimand J, Kull M, Peterson H, Hansen J, Vilo J (2007) g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res* 35 (Web Server issue):W193–200.
- Smits HP, Cohen A, Buttler T, Nielsen J, Olsson L (1998) Cleanup and analysis of sugar phosphates in biological extracts by using solid-phase extraction and anion-exchange chromatography with pulsed amperometric detection. *Anal Biochem* 261:36–42.
- Parrou JL, François J (1997) A simplified procedure for a rapid and reliable assay of both glycogen and trehalose in whole yeast cells. *Anal Biochem* 248:186–188.
- Kelly SL, et al. (1995) Purification and reconstitution of activity of *Saccharomyces cerevisiae* P450 61, a sterol delta 22-desaturase. *FEBS Lett* 377:217–220.
- Otero JM, et al. (2010) Whole genome sequencing of *Saccharomyces cerevisiae*: From genotype to phenotype for improved metabolic engineering applications. *BMC Genomics* 11:723–740.

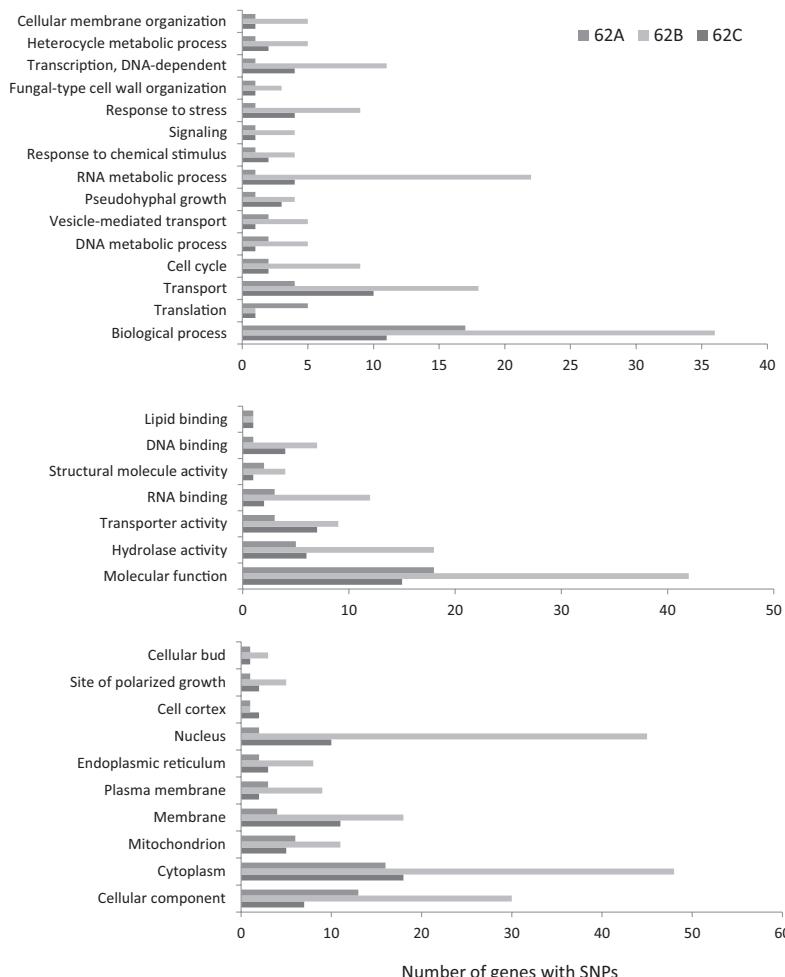


Fig. S1. Gene ontology (GO) enrichment analysis by GO slim mapper based on *S. cerevisiae* genome background.

Table S1. Characterization of phenotype in aerobic batch culture on galactose

Strains	Maximum specific growth rate (h ⁻¹)	Biomass yield (C-mol/C-mol galactose)	Ethanol yield (C-mol/C-mol galactose)	CO ₂ yield (C-mol/C-mol galactose)
REF	0.21 ± 0.00	0.46 ± 0.00	0.13 ± 0.01	0.39 ± 0.00
SO16	0.17 ± 0.01	0.30 ± 0.01	0.31 ± 0.01	0.31 ± 0.00
PGM2	0.21 ± 0.01	0.35 ± 0.02	0.32 ± 0.02	0.27 ± 0.03
62A	0.26 ± 0.00	0.48 ± 0.01	0.14 ± 0.00	0.31 ± 0.05
62B	0.26 ± 0.00	0.45 ± 0.00	0.19 ± 0.01	0.29 ± 0.02
62C	0.26 ± 0.00	0.41 ± 0.01	0.26 ± 0.02	0.30 ± 0.01

Table S2. The concentration of measured metabolites

$\mu\text{mol/g DCW}$	REF-1	REF-2	SO16-1	SO16-2	PGM2-1	PGM2-2	62A-1	62A-2	62B-1	62B-2	62C-1	62C-2
Galactose (intracellular)	149.0	169.1	72.2	114.7	117.5	139.4	215.5	309.0	198.5	240.7	136.5	179.3
Galactitol	29.37	40.23	19.42	20.88	12.20	16.71	29.93	35.98	35.42	33.74	35.53	34.79
Galactose 1P	0.618	0.751	0.373	0.396	0.242	0.272	0.225	0.346	0.279	0.411	0.278	0.290
Glucose 1P	0.279	0.310	0.106	0.115	0.037	0.032	0.054	0.078	0.061	0.090	0.051	0.057
Glucose 6P	0.246	0.305	0.209	0.200	0.155	0.199	0.181	0.229	0.207	0.208	0.227	0.242
Fructose 6P	0.051	0.068	0.052	0.044	0.037	0.050	0.041	0.060	0.048	0.042	0.051	0.056
Fructose 1,6 dP	0.540	0.620	0.954	0.817	0.570	0.711	0.630	0.539	1.828	2.246	0.650	0.627
ATP	3.380	3.857	3.073	2.746	2.003	2.122	3.381	3.221	3.422	4.359	2.950	2.934
ADP	0.668	0.683	0.594	0.532	0.372	0.409	0.512	0.522	0.512	0.559	0.440	0.418
AMP	0.041	0.045	0.059	0.043	0.033	0.043	0.047	0.048	0.056	0.048	0.033	0.033
NADPH	0.167	0.256	0.126	0.202	0.139	0.177	0.090	0.097	0.183	0.133	0.099	0.180
NADP	0.277	0.306	0.197	0.231	0.173	0.266	0.287	0.299	0.235	0.248	0.220	0.318
NADH	0.275	0.301	0.212	0.201	0.115	0.207	0.186	0.223	0.149	0.094	0.210	0.253
NAD	3.989	5.057	2.814	2.752	2.360	3.566	3.956	4.032	2.585	1.517	4.076	3.396
Alanine	0.683	0.768	0.478	0.390	0.475	0.797	0.813	0.949	1.010	0.798	0.793	0.832
Glycine	0.571	0.516	0.361	0.339	0.315	0.524	0.411	0.468	0.648	0.699	0.450	0.468
Valine	0.323	0.293	0.152	0.159	0.179	0.284	0.305	0.261	0.479	0.431	0.260	0.282
Leucine	0.106	0.095	0.096	0.091	0.071	0.104	0.136	0.110	0.159	0.149	0.093	0.115
Isoleucine	0.096	0.074	0.047	0.052	0.049	0.080	0.063	0.093	0.155	0.116	0.045	0.081
Threonine	0.844	0.736	0.608	0.600	0.666	0.885	0.918	0.891	1.494	1.342	0.676	0.902
Proline	0.226	0.178	0.000	0.000	0.036	0.106	0.233	0.255	0.240	0.247	0.293	0.345
Asparagine	0.298	0.266	0.243	0.239	0.220	0.323	0.344	0.321	0.494	0.459	0.253	0.297
Aspartic Acid	2.247	1.890	1.377	1.428	1.618	1.835	2.846	2.288	2.910	2.788	2.296	2.499
Phenylalanine	0.040	0.053	0.018	0.021	0.037	0.033	0.055	0.012	0.094	0.097	0.023	0.043
Glutamic acid	7.528	7.467	4.897	4.720	5.511	7.474	8.076	7.361	10.288	9.354	7.057	7.623
Glutamine	2.432	1.685	1.075	1.553	1.008	1.823	3.765	2.921	5.588	5.611	2.949	2.912
Lysine	0.676	0.696	0.701	0.684	0.672	0.996	0.842	0.768	1.026	0.938	0.617	0.601
Histidine	0.586	0.601	0.722	0.712	0.533	0.805	0.821	0.775	0.782	0.681	0.492	0.519
Tyrosine	0.269	0.272	0.302	0.292	0.265	0.331	0.340	0.311	0.408	0.365	0.230	0.246
Trehalose	2.37	2.47	5.09	4.49	2.22	3.99	10.80	10.36	17.90	17.70	10.65	9.91
Glycogen [mg (as glucose)/g DCW]	26.33	26.22	37.04	37.34	24.22	25.67	53.20	53.04	33.18	33.00	49.68	48.05
Ergosterol	27.60	26.61	28.95	28.11	19.06	20.66	18.52	18.00	3.02	3.23	19.51	18.94
Dihydroergosterol*	1.16	1.19	1.88	1.75	0.90	0.96	0.75	0.79	32.79	37.06	0.85	0.80

*Dihydroergosterol was estimate from ergosterol standard curve.

Table S3. Overall Illumina/Solexa genome sequencing results

Sequencing parameters	62A	62B	62C
No. of reads	5,605,504	18,203,846	5,239,106
Total bases,* bp	213,009,152	691,746,148	199,086,028
Coverage fold	17	55	16
Undetermined base	158,723	86,791	171,362
Genome percent reference coverage, † %	98.7	99.3	98.6
No. of supercontigs	17	17	17
Chromosomes	16	16	16
Mitochondria	1	1	1

*38 bases per sequence read for two cycles.

†Based on genome consensus sequence length of CEN.PK113-7D of 12,155,742 base pairs.

Table S4. Detection of mutations across different genomes

Sequencing parameters	62A/others*	62B/others	62C/others
Total number of mutations	44	334	40
Total number of SNPs	21	104	29
Coding region	6	29	11
Noncoding region	15	75	18
Total number of insertions and deletions	23	230	11
Coding region	0	11	3
Noncoding region	23	219	8

*All other strains including the reference strain, CEN.PK113-7D and the other evolved strains.

Table S5. Characterization of maximum growth rate of site-directed mutants in flask culture

Strain	Genotype	U_{\max}	STD ($n = 3$)	Increase, %	t test (P value)
CEN.PK117-5D	<i>MATa MAL2-8c SUC2 ura3</i>	0.199	0.013	—	—
RB	<i>MATa MAL2-8c SUC2 ura3 RAS2^{Tyr-112}</i>	0.219	0.010	10	0.05

Other Supporting Information Files

[Dataset S1 \(XLS\)](#)

[Dataset S2 \(XLS\)](#)

PAPER II

Recovery of phenotypes obtained by adaptive evolution through inverse metabolic engineering

Kuk-Ki Hong^{1,2}, Jens Nielsen^{1,3#}

¹Novo Nordisk Foundation Center for Biosustainability, Department of Chemical and Biological Engineering, Chalmers University of Technology, SE412 96 Gothenburg, Sweden

²Research Institute of Biotechnology, CJ CheilJedang, Seoul, Korea

³Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK2970 Hørsholm, Denmark

#Corresponding author

E-mail: nielsenj@chalmers.se

Tel: +46 (0)31 772 3804

Fax: +46(0)31 772 3801

Running title: inverse metabolic engineering in systems biology

Abstract

In a previous study system level analysis of adaptively evolved yeast mutants showing improved galactose utilization revealed relevant mutations. The governing mutations were suggested to be in the Ras/PKA signaling pathway and ergosterol metabolism. Here site-directed mutants having one of the mutations, *RAS2*^{Lys 77}, *RAS2*^{Tyr112} and *ERG5*^{Pro 370} were constructed and evaluated. The mutants were also combined with over-expression of *PGM2*, earlier proved as a beneficial target for galactose utilization. The constructed strains were analyzed for their gross phenotype, transcriptome and targeted metabolites; and the results were compared to those obtained from reference strains and the evolved strains. The *RAS2*^{Lys 77} mutation resulted in the highest specific galactose uptake rate among all the strains with an increased maximum specific growth rate on galactose. The *RAS2*^{Tyr112} mutation also improved the specific galactose uptake rate and also resulted in many transcriptional changes, including ergosterol metabolism. The *ERG5*^{Pro 370} mutation only showed a small improvement, but when it was combined with *PGM2* over-expression, the phenotype was almost the same as that of the evolved mutants. Combination of the *RAS2* mutations with *PGM2* over-expression also showed a complete recovery of the adaptive phenotype in galactose utilization. Recovery of the gross phenotype by the reconstructed mutants was achieved with much fewer changes in the genome and transcriptome than for the evolved mutants. Our study demonstrates how the identification of specific mutations by systems biology can direct new metabolic engineering strategies for improving galactose utilization by yeast.

Introduction

Microbial based production of fuels and chemicals has been extensively investigated as it may contribute to establishment of a more sustainable society. In this context the development of microorganisms having efficient substrate utilization and product formation is a requirement. Evolutionary engineering has traditionally been used for this kind of improvement in industry, since it can result in strategies that are unpredicted and hence cannot be obtained through rational design (23). Evolutionary strategies have gained renewed interest with the progress of tools in systems biology as this has allowed for identification of governing mutations that can subsequently be implemented through site-directed mutagenesis using the concept of inverse metabolic engineering (1, 15). Furthermore, understanding the evolution process is useful for identification of novel metabolic engineering strategies (2, 5). One of the typical patterns during adaptive evolution is the saturation of fitness with a proportional increase of mutations with the number of generations (2, 10). This phenomenon seems to be partially explained by accumulation of deleterious mutations. Therefore introduction of beneficial mutations into a parental strain to remove negative mutations is normally the last step, when evolutionary engineering is applied for strain development (25). Another finding from adaptive evolution studies is that the decline of fitness is due to negative epistasis among beneficial mutations, which means that the impact of combining beneficial mutations is less than the sum of individual mutations (4, 20). This finding implicates that different combinations of beneficial genetic changes may result in more advanced phenotype and this opens for multiple inverse metabolic engineering strategies.

In a previous study we characterized evolved mutants of yeast that showed improved galactose utilization compared with a reference strain (17). Three isolated mutants from different populations derived from identical adaptive evolution processes were analyzed by using systems biology tools; and the evolved mutants were also compared with metabolically engineered strains. No mutations were detected in known galactose pathway and regulatory pathways, whereas all three evolved mutants had mutations in the Ras/PKA signaling pathway, and application of one of these mutations showed an improved specific galactose uptake rate. One unique mutation that may link with specific changes of transcripts and metabolites in the ergosterol pathway was also identified, namely a mutation in *ERG5*. In addition to the above mentioned mutations, *PGM2* encoding phosphoglucomutase was validated as a metabolic engineering target as it was found to be over-expressed in all the evolved mutants, which was consistent with earlier finding that over-

expression of this enzyme resulted in improved galactose uptake (3, 13, 17), whereas over-expression of *GAL*-genes either individually or in combination has shown not to result in increased galactose uptake (7). Here we therefore constructed site-directed mutants that have each of the identified mutations and combined them with *PGM2* overexpression to evaluate how much the phenotypes of the adaptively evolved mutants could be recovered by these defined genetic changes. The site-directed mutants were constructed by introducing three point mutations, *RAS2*^{Lys 77} (from evolved mutant 62A), *RAS2*^{Tyr112} (from evolved mutant 62B) and *ERG5*^{Pro 370} (from evolved mutant 62B) to a reference strain; and the mutants were further engineered by transformation of a high-copy number plasmid with constitutive promoter for over-expression of *PGM2*. We present result of a comparative analysis of the mutants with only site-directed mutations and the mutants having a combination of the site-directed mutations and *PGM2* over-expression. All the engineered strains were compared to a reference strain, a previously constructed strain and the two evolved strains. The comparative analysis involved quantitative fermentation physiology and measurement of the transcriptome and targeted intracellular metabolites. Based on our analysis of the constructed mutants, additional strategies for improving galactose utilization are discussed together with the use of adaptive evolution for improving strain performance.

Materials and Methods

Yeast strains and plasmids. *Saccharomyces cerevisiae* strains used in this study are summarized in Table 1. *Saccharomyces cerevisiae* CEN.PK113-5D was used to construct site-directed mutants and combined mutants due to its availability of *URA3* marker gene with the same genotype background of 7D except that gene. Reference strains, evolved mutants and site-directed mutants were preserved on a synthetic minimal (SM) media containing yeast nitrogen base without amino acids, 6.9 g/l (Formedium, Hun-stanton, UK); complete supplement mixture 0.77 g/l (MP Biomedicals, Solon, OH, USA); glucose, 20 g/l and agar, 20 g/l. Other strains carrying plasmids were maintained on the same SM media except the complete supplement mixture without uracil, 0.77 g/l. To avoid uracil supplement related effects, an empty plasmid having the *URA3* gene as a marker, *pSP-GM2*, was used for transformation into the CEN.PK113-5D strain and the site-directed mutants (22). To over-express *PGM2* gene, *pPGM2* was used, which is a high copy number plasmid containing the constitutive promoter of the *PMA1* gene and

URA3 as a marker (3).

Strain construction and cultivation. Site-directed mutants were constructed by previously described method (17). Three point mutations, *RAS2* ^{Lys 77} (identified in evolved mutant 62A), *RAS2* ^{Tyr 112} (identified in evolved mutant 62B) and *ERG5* ^{Pro 370} (identified in evolved mutant 62B) were introduced to the strain CEN.PK113-5D strain separately. Oligomers used in this study were summarized in Table 2. Around 1kb sequence including the mutation in the center region and part of *URA3* gene (20~22 mer) in the end was amplified by PCR. The marker gene, *URA3* was amplified into two parts. 5'-end part (716 base pairs) and 3'-end part (1,028 base pairs) of *URA3* gene were amplified separately. Fusion PCR between the 1kb sequence and each partial *URA3* was implemented; and then two cassettes were prepared. The cassettes were transformed into *S.cerevisiae* CEN.PK113-5D and the sequence that had the mutation in the target gene together with the *URA3* marker was integrated into the chromosome by homologous recombination. The strains that had *URA3* could be screen out in selection media, which were SD media without uracil. These strains were streaked on the same media one more time to purify single colonies. The single isolate clones were streaked out on 5-FOA media, which was SD media supplemented with 30mg/l uracil and 750mg/l 5-Fluoroortic acid monohydrate (Formedium), to loop out the *URA3* marker. In this process two types of clones can be obtained, one containing the point mutation in the target gene and the other having the wild-type sequence of the target gene. The site-directed mutants were selected through colony PCR by comparing different concentration of PCR product when the primer that had the mutation in the 3 'end was used. To check the selected clones, the genes were sequenced and presence of the site-directed mutation was confirmed. Finally prototrophic site-directed mutants (RAU, RBU and EBU) were constructed by transformation with the plasmid, *pSP-GM2* containing the *URA3* gene. The combined mutants RAP, RBP and EBP were constructed by transformation of the plasmid *pPGM2* into the site-directed mutants. The culture media and condition used in this study were as described earlier (17). The fermentation physiology data of *Saccharomyces cerevisiae* CEN.PK113-7D, PGM2, 62A and 62B strains was generated in a previous study (17).

Transcriptome analysis. The transcriptome was measured by using Affymetrix Yeast Genome 2.0 Arrays. The set of differentially expressed genes (Limma test from Bioconductor by using the R Language version 2.13.2, adjust *p-value* < 0.01) were analyzed by comparing the site-directed mutants, the combined mutants and the PGM2 strain after background subtraction of

the transcriptome of the 5DU strain. A second set of differentially expressed genes (adjust *p*-value < 0.01) was identified by contrast among the combined mutants and the evolved mutants. For the evolved mutants, their parental strain is 7D, and background subtraction was therefore based on the 7D strain. The transcriptome data of *Saccharomyces cerevisiae* CEN.PK113-7D, PGM2, 62A and 62B strains was used from previous study (17). Functional gene enrichment was done by updated g:Profiler (24). The heat maps of specific pathways were visualized by using the software of MultiExperiment Viewer (Dana-Farber Cancer Institute, 44 Binney St, Boston, MA, USA). Gene expression data were deposited to the Gene Expression Omnibus (GEO) database with accession number GSE36118.

Analysis of carbohydrates and sterols. Harvested cells were directly used for extraction of carbohydrates and sterols, 10mg and 100mg dry cell weight, respectively. The quantification of carbohydrates (trehalose and glycogen) and sterols (ergosterol and dihydroergosterol) used in this study was described earlier (17). The metabolite concentration data of the CEN.PK113-7D, PGM2, 62A and 62B strains were measured in a previous study (17) and used for combined analysis here.

Results

Fermentation Physiology of reconstructed mutants. The *S cerevisiae* strains used in this study are summarized in Table 1. Three point mutations, *RAS2*^{Lys 77} (identified in evolved mutant 62A), *RAS2*^{Tyr 112} (identified in evolved mutant 62B) and *ERG5*^{Pro 370} (identified in evolved mutant 62B) were introduced into the CEN.PK113-5D strain separately. To avoid uracil supplement related effects, an empty plasmid having the *URA3* gene as a marker, *pSP-GM2*, was transformed into the CEN.PK113-5D strain and the strains carrying the site-directed mutations. The resulting prototrophic strains were designated 5DU (reference strain), RAU (carrying *RAS2*^{Lys 77}), RBU (carrying *RAS2*^{Tyr 112}) and EBU (carrying *ERG5*^{Pro 370}). From comparison of these four strains it was found that each of the single point mutations resulted in improvement of both the maximum specific growth rate and the specific galactose uptake rate (Fig. 1). The largest effect was observed for RAU and this single mutation resulted in a 42% increase in the maximum specific growth rate and a 57% increase in the specific galactose uptake rate compared to the reference strain 5DU. This improvement was much higher than what was obtained in the evolved mutants 62A and 62B derived from the prototrophic CEN.PK113-7D strain, namely a 23 ~ 24%

increase in the maximum specific growth rate and 18 ~ 26 % increase in the maximum specific galactose uptake rate. Though that for the RBU strain that carries another point mutation in the same gene (*RAS2*), the improvement of galactose utilization was much less. The EBU strain showed the smallest increase of galactose utilization among the three site-directed mutants.

Combination of the beneficial point mutations with over-expression of *PGM2* was evaluated by transformation of *pPGM2* including the *URA3* marker instead of the empty plasmid *pSP-GM2* into the site-directed mutants and the reference strain CEN.PK113-5D, and the resulting strains were named RAP, RBP, EBP and PGM2 (Table 1). The PGM2 strain exhibited a 27 % increase of the maximum specific growth rate and a 40 % increase of the specific galactose uptake rate compared to 5DU. These improvements were better than observed for the RBU and EBU strains (Fig. 1), but consistent with what we have reported earlier (3). Combination of the single point mutations from the evolved mutants and *PGM2* overexpression resulted in strains that have almost completely identical gross phenotypes with the evolved mutants, i.e. in terms of the maximum specific growth rate on galactose, the specific galactose uptake rate and the specific ethanol production rate, they are very similar (Fig. 1 and Table S1). Not only did the two mutations in *RAS2*, but also the *ERG5* ^{Pro 370} mutation, combined with *PGM2* overexpression result in an about 60% increase in the maximum specific growth rate and a 35-50 % increase in the specific galactose uptake rate. This is remarkable as the two reference strains 5DU and 7D showed some differences in terms of the maximum specific growth rate, but still RAP, RBP and EBP reached the same level of galactose utilization as the evolved mutants. We performed an evaluation of the overall carbon fluxes, and this shows that the addition of *PGM2* over-expression in the single point mutations resulted in increased respiration and reduced fermentation even with the increase in the maximum specific growth rate (Fig. 2 and Table S2). The biomass yield for the combined mutants (RAP, RBP and EBP) showed a similar range as those for the evolved mutants.

Molecular differences among the reconstructed mutants. The site-directed mutants (RAU, RBU and EBU), the PGM2 strain, and the combined mutants (RAP, RBP and EBP) were compared to find common and unique changes in their transcriptome (Fig. 3). The transcriptome in each of the strains were compared with the reference strain (CEN.PK113-5D containing an empty plasmid) and genes with significant changes in its transcription were identified (correcting for multiple testing). Conserved transcriptional changes in the different mutants were evaluated using Venn diagrams as illustrated in Fig. 3. The PGM2 strain showed the largest number of

transcriptional alterations among all the reconstructed mutants (762 significantly changed genes), whereas the combined mutants, even though they also have *PGM2* over-expression, had much lower number of genes with significant changed transcription, i.e. around 150-300 genes. Functional enrichments by GO term for the transcriptional changes in the PGM2 strain were mostly in organelle related genes (Fig. 3 and Data set S1). A functional category of genes with significant changes in transcription in the combined mutants RAP and EBP was ERK/MAPK target related genes based on the REACTOME database. The RBU strain also showed many more transcriptional changes (389 up-regulated and 185 down-regulated) than RAU (85 up-regulated and 59 down-regulated). Especially, one of the gene enrichment categories in the uniquely up-regulated genes of RBU (254 genes) was ergosterol metabolism. Transcripts and metabolites of the ergosterol pathway were therefore analyzed in more detail (Fig. 4). All strains that have mutation of *ERG5^{Pro 370}* showed overall up-regulation of *ERG* genes and a higher ratio of dihydroergosterol to ergosterol. The mutation of *RAS2^{Tyr112}* also showed this pattern, while the combination of this mutation with *PGM2* overexpression results in a loss of this pattern (RBP strain). The mutation of *ERG5^{Pro 370}* had no significant changes in its transcriptome except for 5 genes in the ergosterol pathway, which are significantly up-regulated. The combination with *PGM2* overexpression (EBP strain), however, induced more transcriptional changes; but when this strain is compared to the PGM2 strain, the change in the transcriptome is much smaller.

Comparison with evolved mutants. We next compared the reconstructed strains with the evolved mutants at the transcriptome level to analyze mutual and exclusive changes in their transcriptome (Fig. 5). The transcriptome of the evolved strains, 62A and 62B, was analyzed in our earlier study (17), and was found to display a much larger number of transcriptional changes than all the reconstructed mutants, i.e. around 700-1,100 genes had significant changes in their transcription. The result of functional enrichments for uniquely changed genes in the evolved mutants included reserve carbohydrate metabolism (Fig. 5 and Data set S2). The reconstructed strains, RAU and RAP covered only 7.5 % and 8.7 % of the transcriptional changes of the 62A evolved mutant, respectively. The RBU strain showed 23.3 % of coverage of the transcriptional changes of the 62B evolved mutant, whereas the combined mutant, RBP covered only 8.2 % of the transcriptional changes. In case of the *ERG5^{Pro 370}* mutation derived from the evolved mutant 62B (EBU strain), only 0.4% of the transcriptional changes of the 62B strain were covered. However, when *PGM2* over-expression was introduced together with the *ERG5^{Pro 370}* mutation

(EBP strain) almost 12.3% of the transcriptional changes in 62B were covered. The common change in all the strains that showed improved galactose availability was up-regulation of *PGM2*. The other common change in the combined mutants and the evolved mutants was the functional category of ERK/MPPK targets. Genes involved in the ergosterol pathway were detected to have significantly changed expression in both RBU and 62B, while this is not the case for the RBP strain. Also, the common change among the EBU, EBP and 62B strains was genes of the ergosterol pathway based on the KEGG database.

Changes in galactose and reserve carbohydrates metabolism. The effects of the mutations in *RAS2* for regulation of *PGM2* and reserve carbohydrates metabolism were analyzed in terms of transcripts and metabolites. The mutations, *RAS2*^{Lys⁷⁷} and *RAS2*^{Tyr¹¹²} induced up-regulation of *PGM2* but not reserve carbohydrate metabolism (Fig. 6). The level of *PGM2* up-regulation by the mutations in *RAS2* was almost similar with what is found in the *PGM2* strain that have over-expression of *PGM2* through constitutive expression using a high-copy number plasmid. Transcripts and metabolites in reserve carbohydrates metabolism showed no comparable changes for any of the reconstructed mutants as found for the evolved mutants 62A and 62B.

Discussion

Few genetic and transcriptional changes are required to reach adaptive phenotypes. During adaptive evolution, there is the possibility to have negative/neutral mutations and these may be inherited for many generations, since their effect could be compensated for by other beneficial mutations (10). Identification of beneficial mutations and removing negative mutations is clearly one of the important points in strain development (25). Also negative epistasis among beneficial mutations is possible (20). In *E. coli* undergoing adaptive evolution it has been found that spontaneous mutations appear proportionally with the number of cell generations; however, the increase of fitness to a specific environment showed a saturation pattern (2). There are therefore clear discordant between mutation generation rate and adaptation rate. Recently, this observation has been explained with the concept of diminishing returns epistasis among beneficial mutations (4, 20). *E. coli* generated and accumulated beneficial mutations; however, the improvement of fitness by their combination was lower than by the sum of individual mutations. Therefore, interaction among beneficial mutations could be negative to phenotypic progress. In other words, expressed phenotypes may be changeable dependent on the

combination of beneficial changes. This phenomenon should be very carefully considered in metabolic engineering as reconstruction of strains based on a combination of beneficial targets identified in evolved mutants may result in improved strains as it not only results in removing negative or neutral mutations but also removes negative epistasis effects. Evolved mutant 62A achieved the improved galactose availability with 21 point mutations including 6 mutations being in coding regions (17), whereas the site-directed mutant RAU showed even higher specific galactose uptake rate with only one point mutation *RAS2*^{Lys 77} identified in 62A with much reduced overall transcriptional alteration. The 62A strain showed a slightly higher maximum specific growth rate than RAU; however, based on comparison with the corresponding reference strains 5DU and 7D, respectively, the improvement in specific growth rate is much higher for RAU. It is unclear why the two reference strains showed differences in terms of the maximum specific growth rate and biomass yield. Their transcriptional differences indicate the changes in cell wall and membrane composition (Data Set S3), and this may be generated by different Ura3p activity because of different copy numbers. The combined mutants RAP, RBP and EBP, even though they had much fewer genetic and molecular changes than the evolved mutants, also achieved almost the same galactose utilization like the evolved ones. These results indicate that fewer genetic changes leading to less transcriptional alteration resulted in better performance of the strain compared with the cumulative mutations obtained from adaptive evolution. The changes of reserve carbohydrates metabolism were considered critical for galactose utilization, because of not only common changes in the three evolved mutants but also close linkage with galactose metabolism by sharing metabolites (17). However, the reconstructed mutants obtained the same gross phenotype as the evolved mutants without any changes in reserve carbohydrate metabolism. This finding indicates that all changes occurring in connection with adaptive evolution, even present in different evolved mutants, may be at least non-essential for improving the phenotype. There could, however, be a beneficial effect of the change in reserve carbohydrate metabolism observed in all the evolved mutants, and it is possible that the effect was not detected in this study. However, it seems that many changes that are not related to galactose utilization appeared during the adaptive evolution.

***RAS2*^{Lys77} (RAU) and *RAS2*^{Tyr112} (RBU) induced up-regulation of *PGM2* but not reserve carbohydrate metabolism; *RAS2*^{Tyr112} (RBU) and *ERG5*^{Pro370} (EBU) up-regulate ergosterol pathway.** In our previous study we inferred that the identified mutations in the Ras/PKA

signaling pathway may have triggered up-regulation of *PGM2* and reserve carbohydrates metabolism, because not only were the mutations identified in all the evolved mutants but on the promoter region of *PGM2* and genes involved in reserve carbohydrate metabolism contained STER elements involved in Ras activation (27). Up-regulation of *PGM2* in the evolved strains could increase galactose utilization as proven in earlier studies (3, 13, 17, 21). In this study, mutations *RAS2*^{Lys77} (RAU) and *RAS2*^{Tyr112} (RBU) showed certain relation to the up-regulation of *PGM2* but not any changes of reserve carbohydrate metabolism. The Ras/PKA signaling pathway is involved in glucose sensing, stress response and many other cellular functions (27). The Ras2 protein is one of the regulatory components that control the activity of protein kinase A (PKA). Ras2 interacts with several molecules such as GTP, guanine exchanging factor (Cdc25), GTPase and adenylate cyclase (Cyr1), and protein structure studies to elucidate binding site with those molecules have been implemented using site-directed mutations (14, 26). One of the mutations *RAS2*^{Val119} in yeast has been well investigated by evaluating its effect on PKA activity and cellular metabolism (16, 18). Incorporation of that mutation resulted in constitutive activation of PKA by strong GTP binding through inhibition of GTPase (26). The effects of the mutation were transcriptional down-regulation of genes in galactose metabolism and reduction of glycogen accumulation (18). The mutations *RAS2*^{Lys77} (RAU) and *RAS2*^{Tyr112} (RBU) in this study seem to have more complex roles than the change of PKA activity. Their positions in *RAS2* are not directly involved in known binding sites of molecules (26). No changes in galactose metabolic genes (*GAL1*, *GAL7* and *GAL10*) were detected (Fig. 6A). In addition to these different effects compared to mutations in other studies such as *RAS2*^{Val119} mutation, the *RAS2*^{Lys77} (RAU) and *RAS2*^{Tyr112} (RBU) showed their unique features in terms of differences in the extent of galactose utilization and the induction of transcriptional changes. One of the unique changes in *RAS2*^{Tyr112} (RBU) was a change in expression of genes involved in the ergosterol pathway (Fig. 4). The transcriptional level of *ERG* genes and the ratio of ergosterol and dihydroergosterol in this mutant were very similar with the strain containing the *ERG5*^{Pro370} (EBU) mutation. There have been no reports about direct relationship between Ras2 and ergosterol metabolism in yeast. This phenomenon seems to be a distinctive feature of the *RAS2*^{Tyr112} (RBU) mutation.

Concerning the effect of the *ERG5* mutation on galactose metabolism we speculate that the higher amount of dihydroergosterol instead of ergosterol may loosen the cell membrane rigidity, which may result in induction of trehalose biosynthesis for stress protection. Increased production

of trehalose may increase galactose uptake by consuming glucose-1-phosphate, which is a feed-forward inhibitor of Pgm2 (17). However, in this study, only negligible effect of mutation in the ergosterol pathway on galactose metabolism was detected. This result means that the *ERG5*^{Pro 370} (EBU) mutation has no direct effect on improving galactose utilization, and this may be the reason why EBU only showed a very small improvement in galactose utilization (Fig. 1). It is, however, interesting that the unique changes in the ergosterol pathway observed in the evolved mutant 62B may be derived from not only the *ERG5*^{Pro 370} (EBU) mutation but also the *RAS2*^{Tyr112} (RBU) mutation.

Combination of *PGM2* overexpression and mutations in *RAS2* and *ERG5* recovers adaptive phenotypes in galactose utilization. The combination of *ERG5*^{Pro 370} mutation with *PGM2* over-expression (EBP strain) showed a similar pattern as the other combined strains, which indicate a role of *ERG5*^{Pro 370} not only in ergosterol metabolism but also as having a positive interaction with *PGM2* over-expression for improving galactose utilization. *PGM2* over-expression by constitutive promoter in a high-copy number plasmid (PGM2 strain) mainly increased fermentation as seen by the higher specific ethanol production rate and up-regulation of organelle related genes (Fig. 1 and Fig. 3) (8, 19). When the three point mutations were combined with *PGM2* over-expression, the metabolism changed to have increased respiration resulting in a higher maximum specific growth rate and hereby also a higher biomass yield on galactose with down-regulation of stress related signaling ERK/MAPK targets, which are related to stress response of osmolarity and starvation. *PGM2* over-expression in RAU and RBU strains, which already had up-regulation of this gene, may somehow result in different Pgm2p functions that could not only be dedicated to galactose metabolism but also to the cellular state. The main function of *PGM2* (phosphoglucomutase) is the conversion of glucose-1-phosphate to glucose-6-phosphate, but other functions such as involvement in the synthesis of cell wall and cellular calcium ion homeostasis have been reported (6, 12). Moreover, conditional posttranscriptional modification of *PGM2* has been reported (9, 11). This modification is controlled by culture condition such as heat shock and carbon sources.

In conclusion, we find that only a few mutations identified in adaptively evolved yeast strains are necessary to confer the same gross phenotype. Furthermore, these mutations may be additive in the sense that over-expression of a single gene may result in a large genome-wide

transcriptional response, but when this over-expression is combined with point mutations many of these transcriptional changes disappear. This is most likely due to an effect of regulation of cellular stress response in the mutations of the Ras/PKA signaling pathway. Thus, it is clear that the use of adaptive evolution with identification of beneficial changes is a powerful strategy in metabolic engineering as it allows for identification of strategies that involves modulation of the cells regulatory system.

Acknowledgements

We thank CJ CheilJedang Doctoral Fellowship for financial support (to K.-K.H.). We acknowledge the Chalmers Foundation, The Knut and Alice Wallenberg Foundation, the European Union funded projects UNICELLSYS (Contract 201142), European Research Council (Grant 247013), and the Novo Nordisk Foundation for financial contributions to this project.

References

1. Atsumi S, Wu TY, Machado IM, Huang WC, Chen PY, Pellegrini M, Liao JC. 2010. Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Mol Syst Biol* 6:449.
2. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243-1247.
3. Bro C, Knudsen S, Regenberg B, Olsson L, Nielsen J. 2005. Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* 71:6465-6472.
4. Chou HH, Chiu HC, Delaney NF, Segre D, Marx CJ. 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332:1190-1192.
5. Conrad TM, Lewis NE, Palsson BO. 2011. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol* 7:509.
6. Daran JM, Bell W, Francois J. 1997. Physiological and morphological effects of genetic alterations leading to a reduced synthesis of UDP-glucose in *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* 153:89-96.
7. de Jongh WA, Bro C, Ostergaard S, Regenberg B, Olsson L, Nielsen J. 2008. The roles of galactitol, galactose-1-phosphate, and phosphoglucomutase in galactose-induced toxicity in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering* 101:317-326.
8. DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
9. Dey NB, Bounelis P, Fritz TA, Bedwell DM, Marchase RB. 1994. The glycosylation of phosphoglucomutase is modulated by carbon source and heat shock in *Saccharomyces cerevisiae*. *J Biol Chem* 269:27143-27148.
10. Elena SF, Lenski RE. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* 4:457-469.
11. Fu L, Bounelis P, Dey N, Browne BL, Marchase RB, Bedwell DM. 1995. The posttranslational modification of phosphoglucomutase is regulated by galactose induction and glucose repression in *Saccharomyces cerevisiae*. *J Bacteriol* 177:3087-3094.

12. Fu L, Miseta A, Hunton D, Marchase RB, Bedwell DM. 2000. Loss of the major isoform of phosphoglucomutase results in altered calcium homeostasis in *Saccharomyces cerevisiae*. *J Biol Chem* 275:5431-5440.
13. Garcia Sanchez R, Hahn-Hagerdal B, Gorwa-Grauslund MF. 2010. PGM2 overexpression improves anaerobic galactose fermentation in *Saccharomyces cerevisiae*. *Microb Cell Fact* 9:40.
14. Hermann-Le Denmat S, Jacquet M. 1997. Yeast RAS2 mutations modulating the ras-guanine exchange factor interaction. *FEBS Lett* 403:95-99.
15. Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR, Albert TJ, Blattner FR, van den Boom D, Cantor CR, Palsson BO. 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 38:1406-1412.
16. Hlavata L, Aguilaniu H, Pichova A, Nystrom T. 2003. The oncogenic RAS2(val19) mutation locks respiration, independently of PKA, in a mode prone to generate ROS. *EMBO J* 22:3337-3345.
17. Hong KK, Vongsangnak W, Vemuri GN, Nielsen J. 2011. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc Natl Acad Sci U S A* 108:12179-12184.
18. Howard SC, Deminoff SJ, Herman PK. 2006. Increased phosphoglucomutase activity suppresses the galactose growth defect associated with elevated levels of Ras signaling in *S. cerevisiae*. *Curr Genet* 49:1-6.
19. James TC, Campbell S, Donnelly D, Bond U. 2003. Transcription profile of brewery yeast under fermentation conditions. *J Appl Microbiol* 94:432-448.
20. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332:1193-1196.
21. Lee KS, Hong ME, Jung SC, Ha SJ, Yu BJ, Koo HM, Park SM, Seo JH, Kweon DH, Park JC, Jin YS. 2011. Improved galactose fermentation of *Saccharomyces cerevisiae* through inverse metabolic engineering. *Biotechnol Bioeng* 108:621-631.
22. Partow S, Siewers V, Bjorn S, Nielsen J, Maury J. 2010. Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*. *Yeast* 27:955-964.

23. Patnaik R. 2008. Engineering complex phenotypes in industrial strains. *Biotechnol Prog* 24:38-47.
24. Reimand J, Arak T, Vilo J. 2011. g:Profiler--a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res* 39:W307-315.
25. Warner JR, Patnaik R, Gill RT. 2009. Genomics enabled approaches in strain engineering. *Curr Opin Microbiol* 12:223-230.
26. Wood DR, Poulet P, Wilson BA, Khalil M, Tanaka K, Cannon JF, Tamanoi F. 1994. Biochemical characterization of yeast RAS2 mutants reveals a new region of ras protein involved in the interaction with GTPase-activating proteins. *J Biol Chem* 269:5322-5327.
27. Zaman S, Lippman SI, Zhao X, Broach JR. 2008. How *Saccharomyces* responds to nutrients. *Annu Rev Genet* 42:27-81.

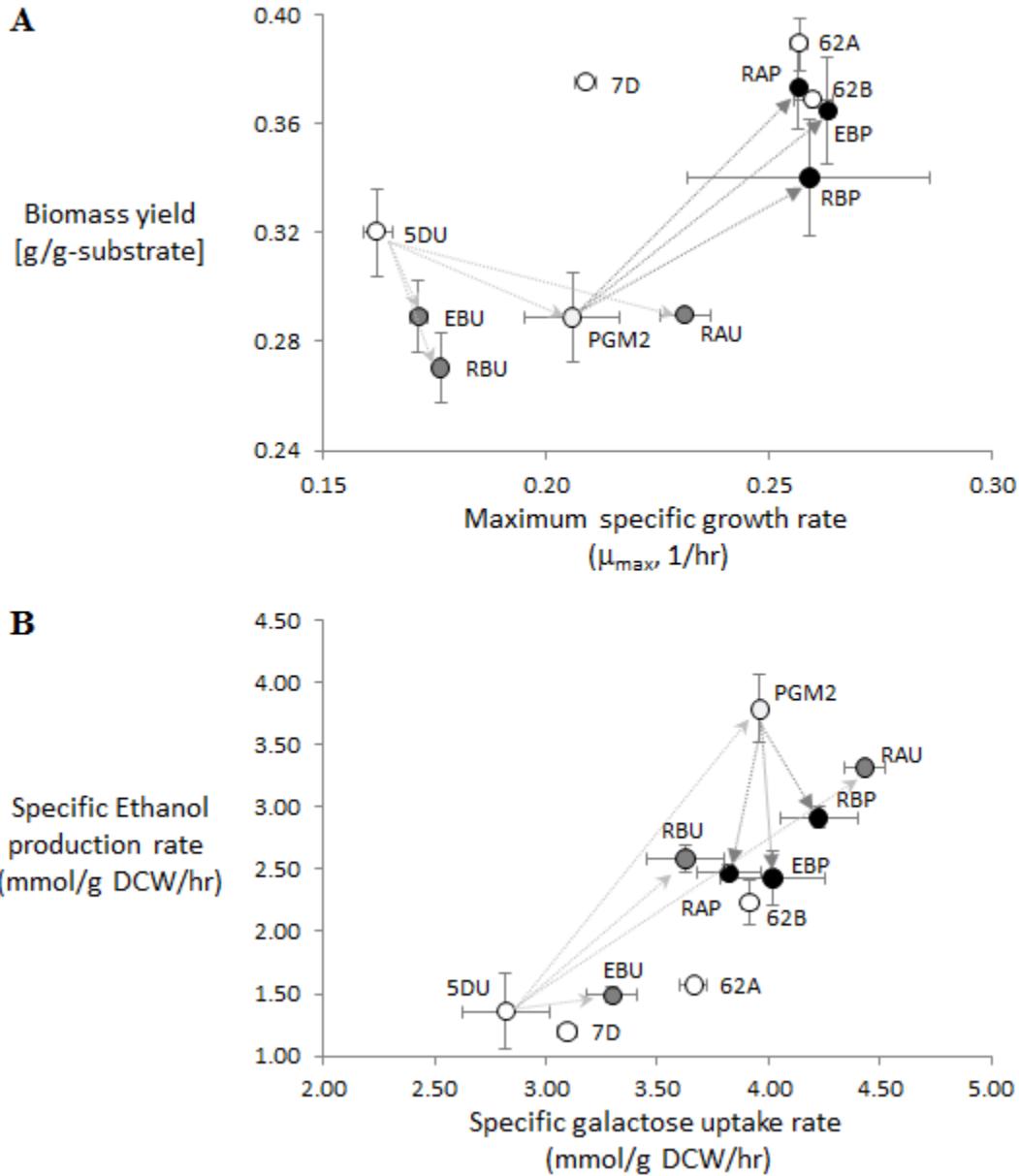


Fig. 1. Data on overall fermentation physiology of the site-directed mutants (RAU, RBU and EBU) and the combined mutants (RAP, RBP and EBP) compare to the reference strains 5DU, 7D and the engineered strain PGM2, and the corresponding evolved mutants 62A and 62B. **A:** Correlation between the maximum specific growth rate and biomass yield **B:** Correlation between the specific galactose uptake rate and the specific ethanol production rate. The arrows indicate relation between reference strains and mutants. Error bars represent standard deviation from biological duplicates.

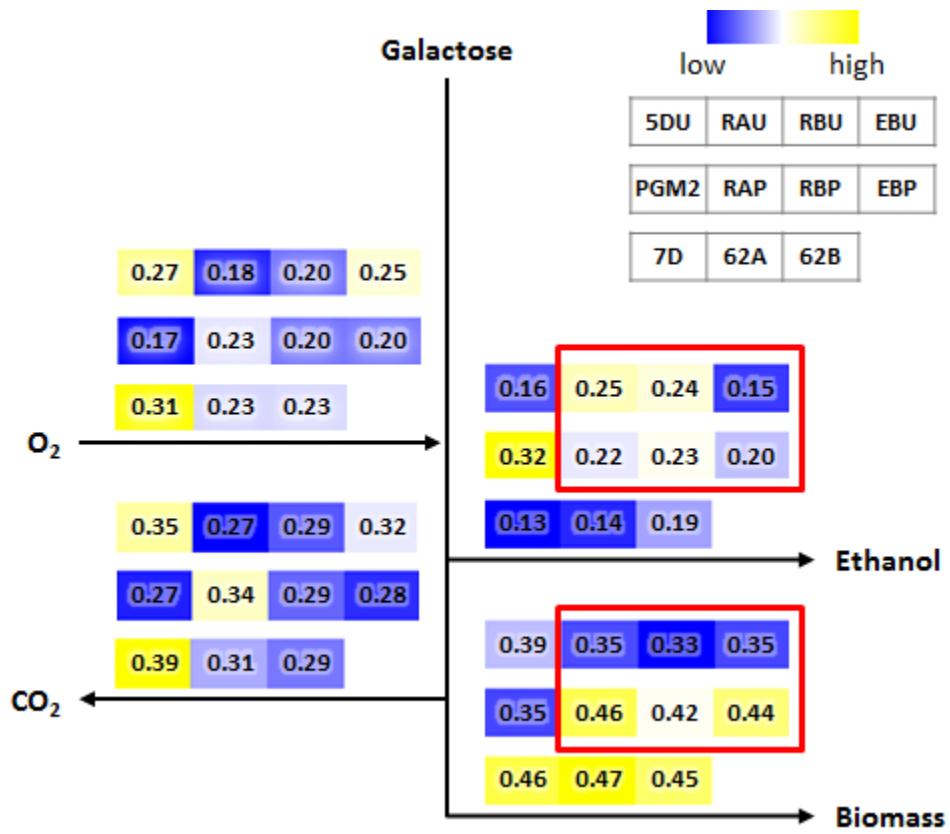


Fig. 2. Comparison of overall carbon fluxes for all the strains used in this study. The boxes represent yield values in unit C-mol/C-mol galactose (mol/C-mol galactose for oxygen). Colors indicate relative size of values for each line of strains. Red boxes indicate additional effect of *PGM2* overexpression in the site-directed mutants.

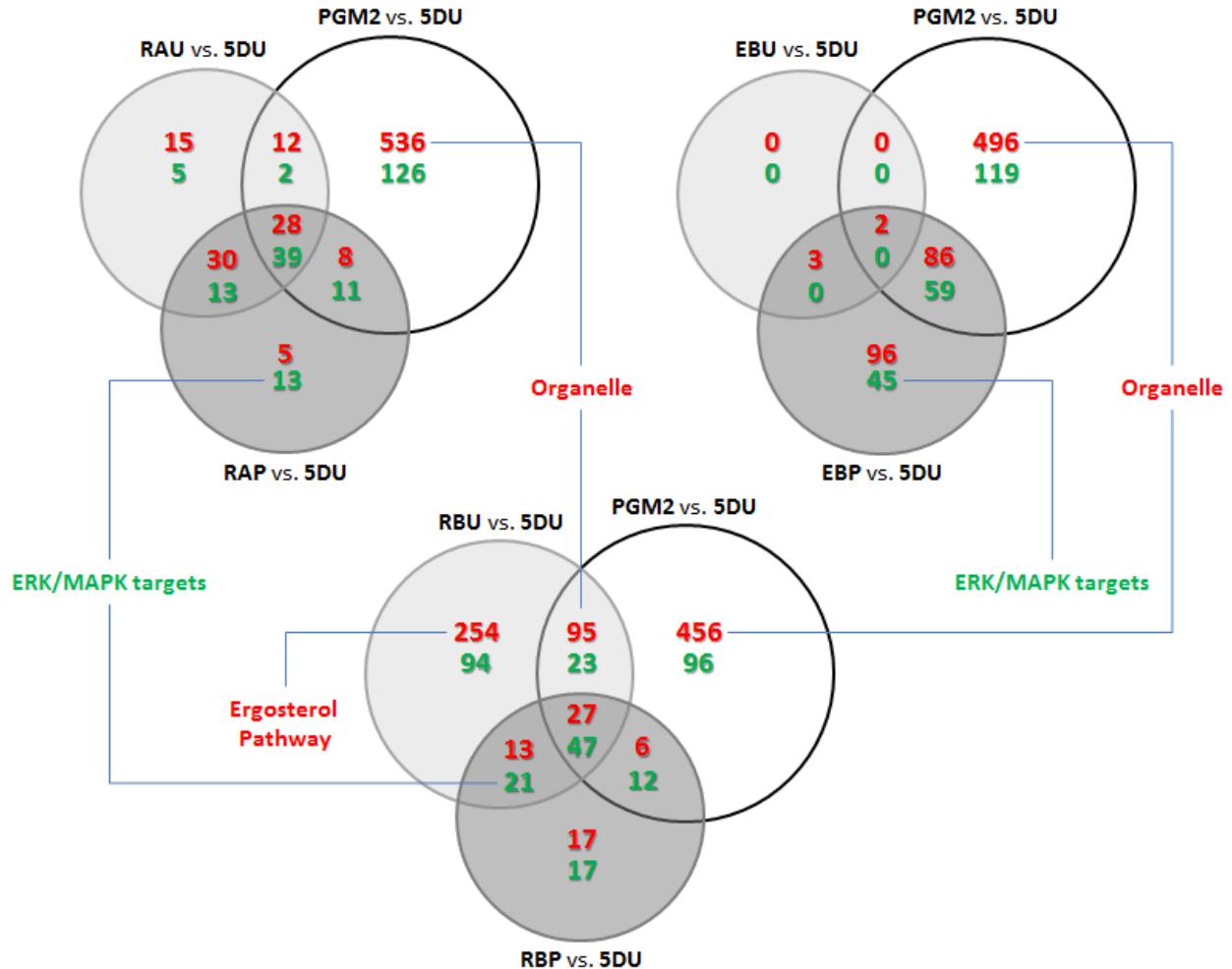


Fig. 3. Effect of mutations ($RAS2^{Lys77}$, $RAS2^{Tyr112}$ and $ERG5^{Pro\ 370}$) and the combination with $PGM2$ over-expression compared with the engineered strain $PGM2$ by differentially expressed genes. Differentially expressed genes ($p<0.01$) are categorized as Venn diagrams. The functional enrichment of genes in each part was analyzed by hyper-geometric distribution based on the KEGG, Reactome and GO term databases. Red color letter means up-regulation and green means down-regulation.

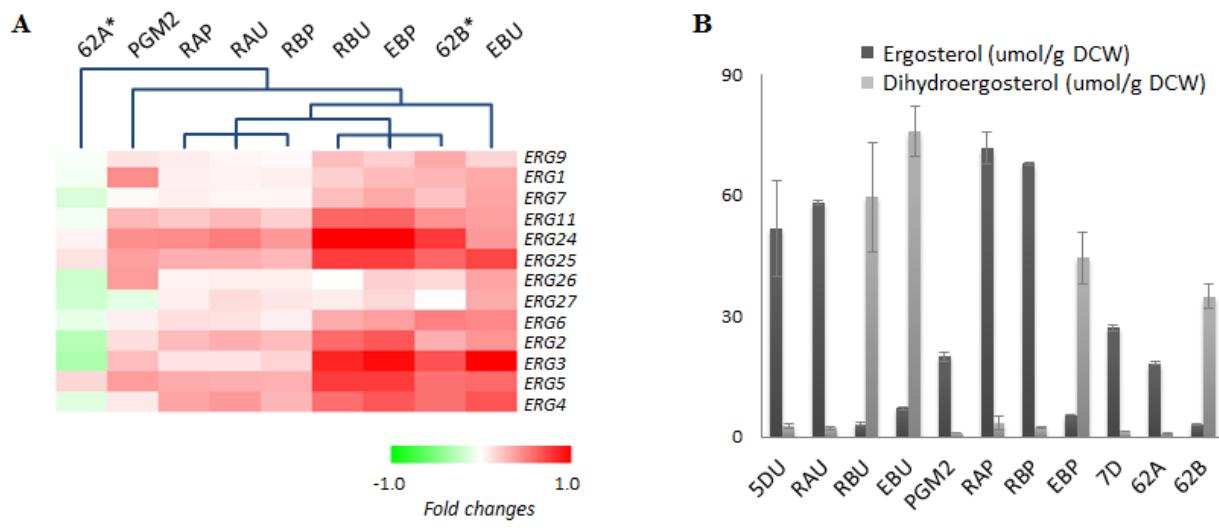


Fig. 4. Changes in ergosterol metabolism in the reconstructed strains illustrated by changes in fold changes of gene transcription and the concentration of ergosterol and dihydroergosterol compared to the other strains. **A:** Hierarchical clustering of all strains is computed using fold changes of genes in ergosterol metabolism. **B:** The concentrations of ergosterol and dihydroergosterol. Error bars represent standard deviation from biological duplicates.

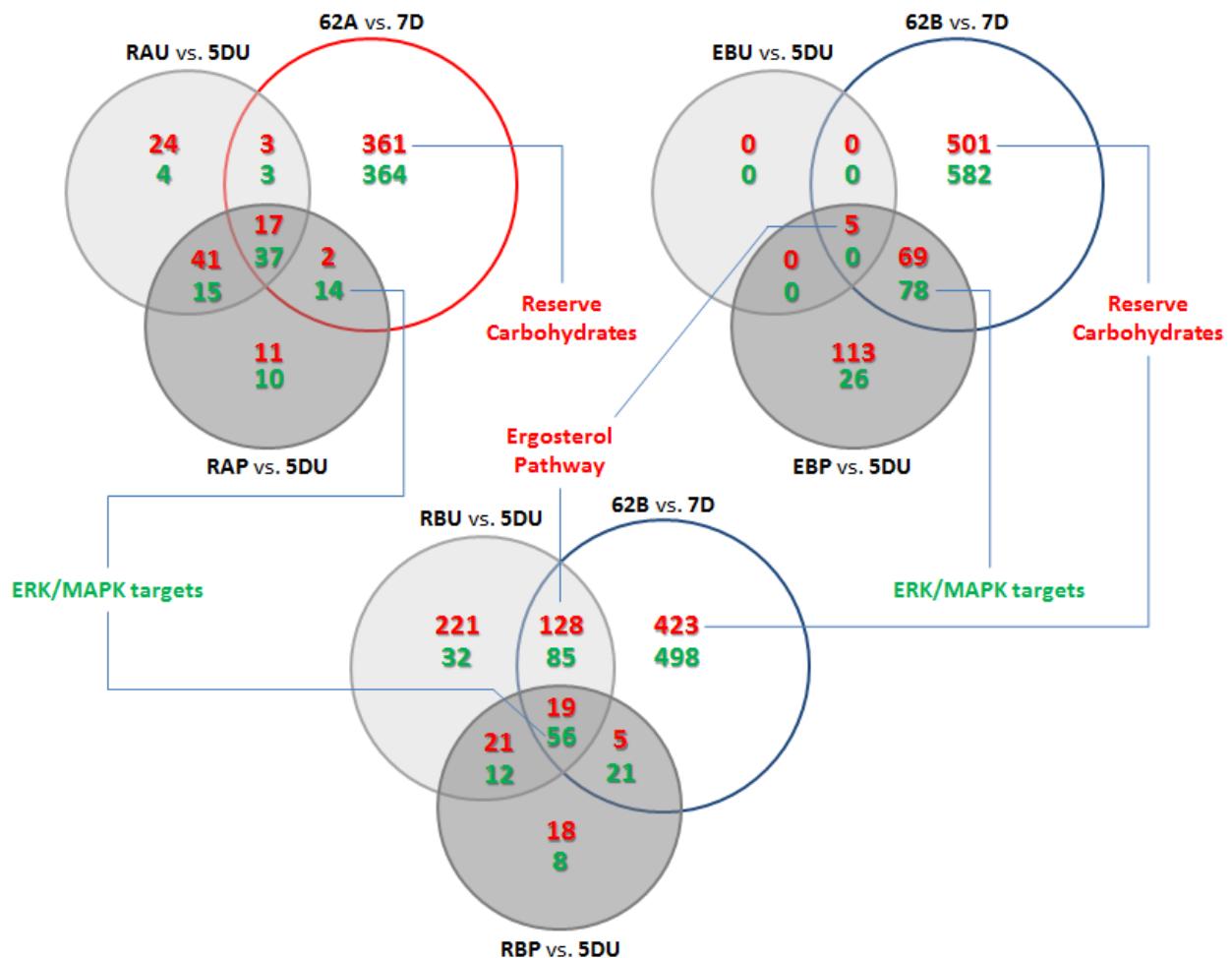


Fig. 5. Effect of reconstructed strains compared to the evolved strains 62A and 62B by differentially expressed genes. Differentially expressed genes ($p<0.01$) are categorized as Venn diagrams. The functional enrichment of genes in each part was analyzed by hyper-geometric distribution based on the KEGG, Reactome and GO term databases. Red color letter means up-regulation and green means down-regulation.

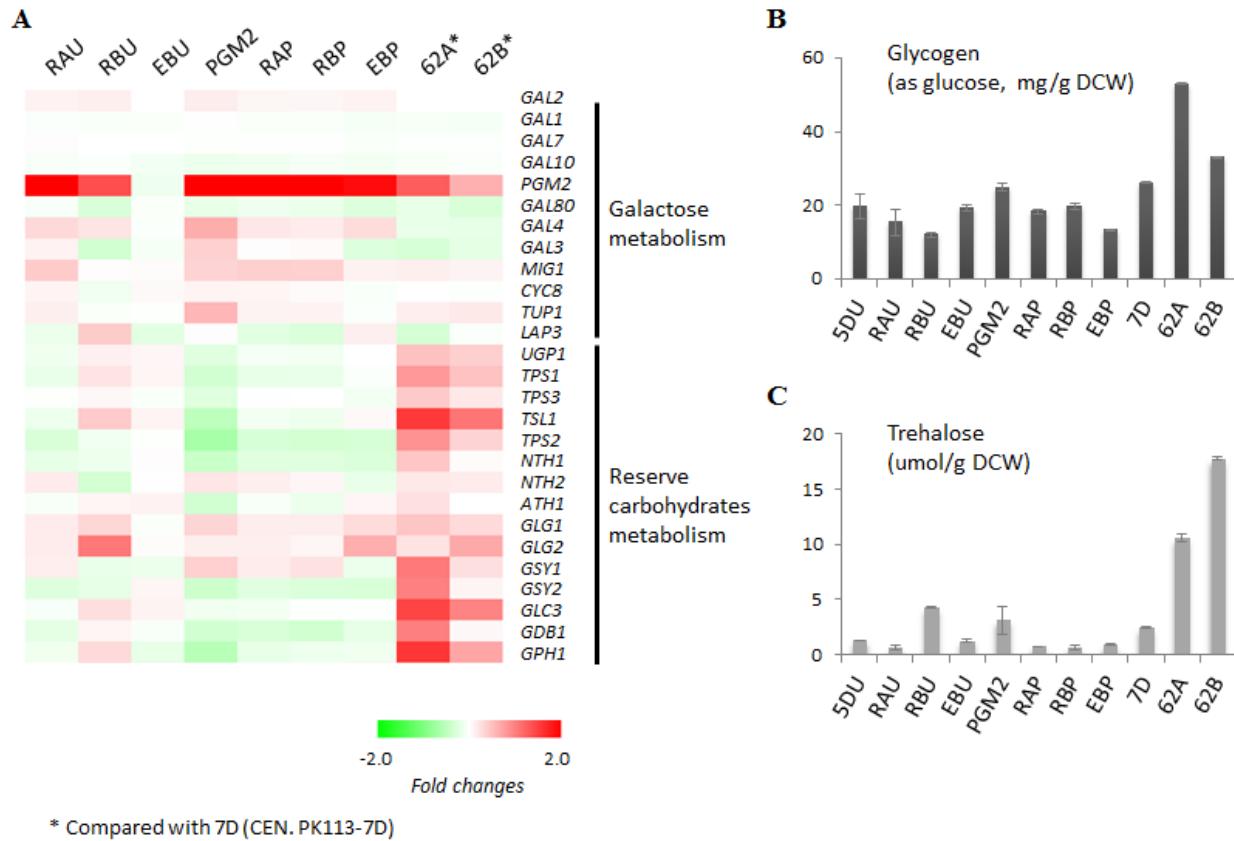


Fig. 6. Changes in the galactose and reserve carbohydrates metabolisms in the reconstructed strains are shown by changes in fold changes of the transcriptome and the concentration of carbohydrates. **A:** Fold changes of all genes involved in galactose and reserve carbohydrates metabolisms are compared to the reference strains. **B:** The concentrations of glycogen. **C:** The concentration of trehalose. Error bars represent standard deviation from biological duplicates.

Table 1. Yeast strains used in this study

Name of Strains	Ancestor strains and Genotype	Name of Groups	References
7D	<i>MATa SUC2 MAL2-8^c</i> (CEN.PK113-7D)	Reference strain	SR&D*
62A	7D, total no. SNPs: 21 including <i>RAS2</i> ^{Lys 77}	Evolved mutants	Hong et al 2011
62B	7D, total no. SNPs: 104 including <i>RAS2</i> ^{Tyr112} , <i>ERG5</i> ^{Pro 370}		Hong et al 2011
5D	<i>MATa SUC2 MAL2-8^c ura3-52</i> (CEN.PK113-5D)		SR&D*
5DU	5D, <i>pSP-GM2(URA3)</i>	Reference strain	This study
RAU	5D, <i>pSP-GM2(URA3)</i> ; <i>RAS2</i> ^{Lys 77} (from 62A)		This study
RBU	5D, <i>pSP-GM2(URA3)</i> ; <i>RAS2</i> ^{Tyr112} (from 62B)	Site-directed Mutants	Hong et al 2011 (plasmid added in this study)
EBU	5D, <i>pSP-GM2(URA3)</i> ; <i>ERG5</i> ^{Pro 370} (from 62B)		This study
PGM2	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i>	Engineered mutant	Bro et al 2005
RAP	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i> ; <i>RAS2</i> ^{Lys 77}		This study
RBP	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i> ; <i>RAS2</i> ^{Tyr112}	Combined mutants	This study
EBP	5D, <i>pPGM2(URA3, P_{PMA1}-PGM2)</i> ; <i>ERG5</i> ^{Pro 370}		This study

*Scientific Research & Development GmbH, Oberursel, Germany.

Table 2. Oligomers used in this study*

Targets	Names	Sequence (5'-3')	Use
<i>URA3</i>	URA3-5P1	TTCGGCTTCATGGCAATTCC	Half of <i>URA3</i> (5' <i>URA3</i>)
	URA3-3P1	GAGCAATGAACCCAATAACGAAATC	
	URA3-5P2	CTTGACGTTCGTTCGACTGATGAGC	Half of <i>URA3</i> (3' <i>URA3</i>)
	URA3-3P2	ATCGATAAGCTTGATATCGAATTCC	
<i>RAS2</i>	RAS2-5P1	CCACTCTTATCTGACTCTTCTGC	
	RAS2-3P1	GGAATTGCCATGAAGCCGAATTCACATT TTTA CCGTTGGCAGC	Fusion with 5' <i>URA3</i>
	RAS2-5P2	GGAATTCGATATCAAGCTTATCGATCCACTCT TTATCTGACTCTTCTGC	Fusion with 3' <i>URA3</i>
	RAS2-3P2	TTCACATTTTACCGTTGGCAGC	
<i>RAS2</i>	RAS2-4AW	GCCGTTGCGCATGTATT <u>G</u>	Selection of wild-type, <i>RAS2</i>
	RAS2-4AM	GCCGTTGCGCATGTATT <u>T</u>	Selection of mutants, <i>RAS2</i> ^{Lys 77}
	RAS2-4BW	AATTGGAACATAGTCGGTAT <u>C</u>	Selection of wild-type, <i>RAS2</i>
	RAS2-4BM	AATTGGAACATAGTCGGTAT <u>A</u>	Selection of mutants, <i>RAS2</i> ^{Tyr112}
<i>RAS2</i>	RAS2-5P	GTTTAGCCGTTGTCTTCTCTT	
	RAS2-3P	GTTCTTCGCTTAGCGTTTC	Sequencing of <i>RAS2</i>
<i>ERG5</i>	ERG5-5P1	CGCCTTATCATTGAACCTTT	
	ERG5-3P1	GGAATTGCCATGAAGCCGAAGGGAAAATTGTA GCGAAAAC	Fusion with 5' <i>URA3</i>
	ERG5-5P2	GGAATTCGATATCAAGCTTATCGATCGCCTTA TCATTGAACCTTT	Fusion with 3' <i>URA3</i>
	ERG5-3P2	GGGAAAATTGTAGCGAAAAC	
<i>ERG5</i>	ERG5-4BW	GTAGACATGTCATTGTTAC	Selection of wild-type, <i>ERG5</i>
	ERG5-4BM	GTAGACATGTCATTGTT <u>A</u>	Selection of mutants, <i>ERG5</i> ^{Pro 370}
<i>ERG5</i>	ERG5-5P	GAGAATAACTACGAGCCCCAG	
	ERG5-3P	GGTCTCTCTTTGAAAGTCAG	Sequencing of <i>ERG5</i>

* The sequence for making fusion region with *URA3* are marked by bold font. Different sequences to separate wild-type and mutant are underlined.

PAPER III

Adaptively evolved yeast mutants on galactose show trade-offs in carbon utilization on glucose

Kuk-Ki Hong^{1,2}, Jens Nielsen^{1,3#}

¹Novo Nordisk Foundation Center for Biosustainability, Department of Chemical and Biological Engineering, Chalmers University of Technology, SE412 96 Gothenburg, Sweden

²Research Institute of Biotechnology, CJ CheilJedang, Seoul, Korea

³Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK2970 Hørsholm, Denmark

#Corresponding author

E-mail: nielsenj@chalmers.se

Tel: +46 (0)31 772 3804

Fax: +46(0)31 772 3801

ABSTRACT

Adaptive evolution has provided new strategies in metabolic engineering; however, several constraints still exist as evolutionary trade-offs impose collateral cost to obtain new traits. Its application for strains development could be further improved by elucidating the molecular mechanism. In this study, adaptively evolved yeast mutants with increased galactose utilization ability showed decreased glucose utilization. The molecular genetic basis of this trade-off was investigated by systems biology approach. Molecular changes incorporated the improvement of galactose utilization were maintained even under glucose cultivation. Moreover, glucose repression related genes showed conserved expression patterns towards both sugars. Mutations in the *RAS2* gene that were identified as beneficial ones for galactose utilization in the evolved mutants exhibited significant correlation to attenuation of glucose utilization. These results indicate that antagonistic pleiotropy is the dominant mechanism in this trade-off, and it is likely realized by the changes in glucose signaling.

1. Introduction

Integrative analysis of gross phenotypes using diverse high-throughput molecular data has allowed for improved understanding of the underlying molecular mechanisms (Daran-Lapujade et al., 2007; Hong et al., 2011). This salient achievement is enabled by development of genome scale tools such as DNA microarray, whole genome sequencing and metabolomics (Bro and Nielsen, 2004; Herring et al., 2006), and identification of molecular changes in evolutionary processes has started to take advantages of this development (Conrad et al., 2011; Wenger et al., 2011). Fundamental theories in evolutionary biology were traditionally proved mostly at the phenotypic level; however, application of high-throughput techniques is enabling the elucidation of molecular mechanisms driving specific phenotypes (Dettman et al., 2012). Furthermore, in recent studies evolved mutants having improved specific traits for industrial applications could be precisely characterized resulting in identification of beneficial mutations (Conrad et al., 2010; Hong and Nielsen, 2012; Smith and Liao, 2011).

A trade-off in relative fitness between different environments is one of the prevailing concepts in evolutionary biology (Cooper and Lenski, 2000; Elena and Lenski, 2003). The presence of this specific feature has been suggested by theoretical studies; and the comparison of changed fitness of descendants to an ancestor in altered environments proved it experimentally (Bennett and Lenski, 2007; Chang and Leu, 2011; Cooper and Lenski, 2000). Mainly two mechanisms for the evolutionary trade-off has been suggested, antagonistic pleiotropy (AP) in which the same mutation is involved in gain and loss of adaptation in different conditions, and mutation accumulation (MA) where the different mutations are related to that advantage and disadvantage (Cooper and Lenski, 2000; Elena and Lenski, 2003). Physiological tests were used to find the trade-offs in carbon sources, temperature, pH; and applied to discriminate the plausible mechanism (Bennett and Hughes, 2009; Bennett and Lenski, 2007; Cooper and Lenski, 2000). Recently, the advent of tools in systems biology has accelerated the study of the trade-off mechanisms resulting from molecular genetic changes (Wenger et al., 2011). Besides gaining new fundamental insight the characterization of trade-off mechanisms is important for strain development in industrial biotechnology. For example, if an evolved mutant has improved traits on one carbon source and the mutant shows reduced feature on another carbon, then it is relevant to know if this trade-off is AP or MA. Thus, if the trade-off has happened by the same mutation, i.e. by AP, adjustment of that mutation by saturation mutagenesis or making it conditional

operational should be considered. On the other hand, if MA is the dominant mechanism in the trade-off, removing the mutation that is neutral in one condition but deleterious in other condition can eliminate the trade-off effect. Understanding trade-off mechanism in evolutionary changes can therefore provide new strategies in metabolic engineering.

In our previous study, we characterized three galactose-evolved *Saccharomyces cerevisiae* mutants that were independently generated from different populations on galactose (Hong et al., 2011). Their improvement of galactose utilization were confirmed in precisely controlled bioreactors and correlated with molecular changes; we further identified beneficial mutations through site-directed mutagenesis (Hong and Nielsen, 2012). Here, the response of the galactose-evolved yeast mutants on glucose was characterized at the systems level. Glucose is the preferred carbon source for yeast, and yeast therefore has a very tightly regulated glucose metabolism, and it is the most widely used carbon source in industrial fermentations. The galactose-evolved mutants were generated by serial cultivation on galactose minimal media over 62 days. Considering the short evolutionary history of adapting the strains to grow faster on galactose compared to the millions of years of evolution to maximize growth on glucose, a decline in glucose utilization to compensate for the cost of improving galactose utilization was not expected. However, interestingly all galactose-evolved mutants show reduced maximum specific growth rate and specific carbon uptake rate on glucose compared to their ancestor strain. The purpose of this study was to characterize these trade-offs through the integration of transcriptome, target metabolites profile and mutation analysis based on whole-genome sequencing data. Besides studying the mechanisms of trade-off, our study is the first example of how systems biology can be used to distinguish the trade-off mechanisms as being AP or MA at the molecular level.

2. Materials and methods

2.1. Yeast strains

All yeast strains used in this study were from previous studies (Bro et al., 2005; Hong and Nielsen, 2012; Hong et al., 2011; Ostergaard et al., 2000). *Saccharomyces cerevisiae* CEN.PK113-7D was an ancestor strain for the evolved mutants and was used as reference strain. The adaptively evolved clones, 62A, 62B and 62C strains were selected from three independent

populations that were evolved on galactose (20 g/L) minimal media for 62 days through daily serial transfer (Hong et al., 2011). Engineered strains, SO16 ($\Delta gal6 \Delta gal80 \Delta mig1$) and PGM2 (over-expression of *PGM2*) were constructed by direct genetic modification (Bro et al., 2005; Ostergaard et al., 2000). Site-directed mutants, RAU (mutation, *RAS2*^{Lys⁷⁷} identified in evolved mutant 62A), RBU (mutation *RAS2*^{Tyr¹¹²} identified in evolved mutant 62B) and EBU (mutation, *ERG5*^{Pro³⁷⁰} identified in evolved mutant 62B) were constructed by site-directed mutagenesis (Hong and Nielsen, 2012).

2.2. Measurement of gross phenotype

Fermentation physiology data of all strains on galactose were used from previous studies (Hong and Nielsen, 2012; Hong et al., 2011). Biological duplicate of all strains was implemented in bioreactor scale (1L). In this study, the same culture condition and media composition except carbon source to glucose instead of galactose was used in the same bioreactor or baffled shake flasks. Evolved mutants (62A, 62B and 62C), engineered strains (PGM2, SO16) and a reference strain (CEN.PK113-7D) were cultivated in the bioreactors with biological duplicate. Site-directed mutants (RAU, RBU and EBU) and their reference strain (CEN.PK113-5D having *URA3* marker in plasmid) were cultivated in baffled shake flasks with biological triplicate. In the bioreactor scale, firstly inoculation was prepared from seed culture broth in shake flask. Media for the seed was prepared as following, salt solution $(\text{NH}_4)_2\text{SO}_4$, 7.5 g/L; KH_2PO_4 , 14.4 g/L; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.50 g/L; Antifoam 289 (A-5551, Sigma-Aldrich, St. Louis, MO, USA), 0.05 ml/L; and trace metal solution, 1 ml/L, pH 6.5 with 2M NaOH. These solutions were sterilized and mixed with filtrated vitamin mixture solution, 1 ml/L; and glucose was separately autoclaved, 10g/L. The trace metal solution was composed of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.45 g/L; $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0.3 g/L; $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.3 g/L; $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.3 g/L; H_3BO_3 , 0.1g/L; KI, 0.1 g/L; $\text{MnCl}_2 \cdot 2\text{H}_2\text{O}$, 1 g/L; $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.4 g/L; $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 0.45 g/L; and EDTA (sodium salt), 15.0 g/L. The pH 4.0 of this solution was adjusted with 2 M NaOH. The vitamin solution was composed of biotin, 0.05 g/L; Ca-pantothenate, 1 g/L; myo-inositol, 25 g/L; nicotinic acid, 1 g/L; p-amino benzoic acid, 0.2 g/L; pyridoxine-HCl, 1 g/L; and thiamine-HCl, 1 g/L. The pH 6.5 of this solution was adjusted with 2 M NaOH. Inoculation was implemented with 1 mg dry cell weight/L from the seed culture. Aerobic batch cultures were carried out with 1.0 L DasGip stirrer-pro® vessels (DasGip, Jülich, Germany) containing 0.8 L working volume. Temperature was

maintained at 30°C; an air flow rate was kept to 1 liter per liquid volume per minute (1vvm); the pH of the culture media was maintained at 5.0 by automatic addition of 2N KOH; an agitation rate was set from 400 to 1,200 rpm for keeping over 30% of dissolved oxygen in the media. The concentration of dissolved oxygen was measured with an autoclavable polarographic oxygen electrode (Mettler Toledo, Columbus, OH, USA). The off-gas from the bioreactors was monitored for real-time determination of oxygen and carbon dioxide concentration by DasGip fedbatch pro® gas analysis systems with the off gas analyzer 1 GA4 based on zirconium dioxide and two-beam infrared sensor (DasGip, Jülich, Germany). For the cultivation of the site-directed mutants and their reference strain in baffled shake flask scale, medium composition was the same as the shake flask for seed culture except the concentration of the $(\text{NH}_4)_2\text{SO}_4$, 5g/L; the KH_2PO_4 ; 3 g/L; and glucose 20g/L. Cultivation was done by cotton-covered 500-mL Erlenmeyer baffled flasks with 100 mL of the media at 30 °C and $\times 0.76$ g. In both scales, samples were taken hourly during exponential growth phase; and they were used for measuring biomass and extracellular metabolites. These data was used to calculate yields and rates.

2.3. Transcriptome analysis

Samples for analysis of transcriptome were obtained at mid-exponential phase in bioreactors. Transcriptome chip-file of the evolved mutants (62A, 62B and 62C) and the reference strain (CEN.PK113-7D) in galactose were hired from previous study (Hong et al., 2011). In this study, their transcriptome data in glucose were prepared. Affymetrix Yeast Genome 2.0 Array was used to analyze transcriptome of strains. In previous study, since the transcriptional difference between 62A and 62C strains were insignificant on galactose, initially one array chip for each of them on glucose was applied. In case of other mutant, 62B and the reference strain on glucose, two array chips for biological duplicate were used. After normalization of all chips, principal component analysis (PCA) was operated by TM4 software (Saeed et al., 2003). Even in glucose, 62A and 62C were also quite closer than other strains, hence they were considered as a replicate of one strain for analysis of differentially expressed genes (Supplementary Figure 1A and 1B). The analysis of differentially expressed genes were performed through Limma test from Bioconductor by using the R Language version 2.13.2 (adjust *p-value* < 0.05) among the evolved mutants and the reference strain in one carbon. Common transcriptional changes among the evolved mutants compared to the reference strain in glucose ware compared to the changes in galactose

(Supplementary Figure 1C). Also the specific changes of each mutant from reference strain in glucose were compared to the change in galactose. Functional gene enrichment was performed by updated g:Profiler (Reimand et al., 2011). This web-software showed significant changes based on GO terms, REACTOME, KEGG and TRANSFAC database by using a cumulative hypergeometric test. The heat maps of specific pathways were generated by using the TM4 software (Saeed et al., 2003). Gene expression data were deposited to the Gene Expression Omnibus (GEO) database with accession number GSE39065.

2.4. Analysis of carbohydrates and sterols

Harvested cells from mid-exponential phase in bioreactors were used for measurement of carbohydrates and sterols. The quantification of these metabolites was followed in previous study (Hong et al., 2011). To extract trehalose and glycogen, alkaline digestion to 10 mg cell (dry cell weight) at 95 °C for 4 hours with 0.25ml of 0.25 M Na₂CO₃ was implemented; then 0.15 ml of 1M acetic acid and 0.6 ml of 0.2 M sodium acetate was added to adjust pH to pH 5.8. Quantification of trehalose and glycogen were performed by measuring the amount of glucose released after enzymatic reaction of trehalase (0.05 U/ml) (Sigma Cat. No. T-8778) at 37 °C and amyloglucosidase from *Aspergillus niger* (1.2 U/ml) (Sigma Cat. No. A7420) at 57 °C. The released glucose was quantified by glucose assay kit (Sigma Cat. No. GAGO20). Extraction of ergosterol and dihydroergosterol was done by saponification to 100 mg cell (dry cell weight) following treatment of 0.4 ml of 0.2 N HCl at 85 °C for 1 hour; cell pellet was centrifuged and responded in 0.2 ml of pyrogallol in methanol 0.2% (w/v) and 0.1 ml of 4N KOH at 85 °C for 2 hours; after cooling down to room temperature, 0.5ml of heptane was added and mixed vigorously for 2min; then left the mixture for 2 hours on the table. Consequently the heptane layer was transferred to a new clean glass vial for injection to high-performance liquid chromatograph (HPLC). Quantification of ergosterol and dihydroergosterol was implemented by HPLC (UltiMates 3000 Nano, Dionex) with a reverse phase Develosil column (C30-UG-5; Nomura Chemicals, Aichi, Japan) at 40 °C with 100% methanol as a mobile phase at 1 ml/min of a flow rate; and UV-visible light absorbance detector set at 280 nm (Photodiode Array Detector, Dionex) was used to identify the peaks of two sterols. The metabolite concentration data of the evolved mutants and the reference strain on galactose was used from previous study (Hong et al., 2011).

3. Results

3.1. Physiological response of galactose-evolved yeast mutants to glucose

In a previous study we generated three independent, adaptively evolved mutants 62A, 62B and 62C of *S. cerevisiae* to improve galactose utilization (Hong et al., 2011). They were evolved for faster growth in galactose minimal media over 62 days, and they showed improved galactose utilization manifested by a higher maximum specific growth rate and specific galactose uptake rate. In addition, two engineered strains SO16 and PGM2 that have increased specific galactose uptake rates were evaluated. SO16 and PGM2 carry deletion in *MIG1*, *GAL80*, *GAL6* and over-expression of *PGM2*, respectively (Bro et al., 2005; Ostergaard et al., 2000). In this study, the gross phenotype of these strains on glucose was quantified and compared with data for these strains grown on galactose (taken from our previous study (Hong et al., 2011)). Differences in the value of the gross phenotype parameters among the strains on the two carbon sources are shown in Fig. 1. The maximum specific growth rate of the galactose-evolved mutants compared to the reference strain growing on glucose was reduced (Δ -value: around -0.02 h^{-1}), while they showed an increased value on galactose (Δ -value: around $+0.04\text{ h}^{-1}$) (Fig. 1A). This trade-off effect was also observed for the specific carbon uptake rates and the specific ethanol production rate (Fig. 1B). The degree of decline in the specific glucose uptake rate (Δ -value: around -2 mmoles/g DW/h) and the specific ethanol production rate (Δ -value: around -3 mmoles/g DW/h) in the evolved mutants was larger than the extent of improvement of galactose metabolism, (Δ -value: around 1 mmoles/g DW/h) and (Δ -value: around 1 mmoles/g DW/h), respectively. To sum up, the trade-off effect for the utilization of the two different carbon sources was significant for the evolved mutants, but the magnitude varied among the strains. In comparison, the two engineered strains showed no trade-off effect in the maximum specific growth rate (Fig. 1A), while there was a trade-off effect on the specific carbon uptake rate and the specific ethanol production rate (Fig. 1B). Both engineered strains, PGM2 (overexpression of *PGM2* on plasmid) and SO16 (triple gene knock-out, *mig1*, *gal80* and *gal6*) showed a reduction of the specific glucose uptake rates; especially the SO16 strain exhibited the largest decrease in the specific glucose uptake rate among all the strains.

3.2. Molecular responses of galactose-evolved mutants on glucose.

To explore molecular changes among the different strains, transcriptome analysis was carried out. Transcriptional differences of the galactose-evolved mutants to the reference strain on the two carbon sources were analyzed by principal component analysis (PCA) (Fig. 2). The distances between the evolved mutants and the reference strain in each of the carbon sources were very similar. Principal component 1 (PC1), representing 73 % of the variance, showed the largest alteration between galactose and glucose for all the strains, whereas PC2 and PC3, corresponding to 10 % and 6 % of the variance, respectively, displayed that the degree of distance between the reference strain and each of the evolved mutants on galactose was almost same as on glucose. To investigate the detailed transcriptional responses on glucose, differentially expressed genes common for all three evolved mutants on glucose (adjusted p-value < 0.05) were analyzed and compared with the differentially expressed genes on galactose (Supplementary Figure 1). Differentially expressed genes during growth on glucose were separated into two groups, the first one contained the genes that also showed changes during growth on galactose and the second had the genes that changed their expression level significantly only during growth on glucose (Fig. 3). The first group was designated as glucose-independent, as its expression level was changed during growth on both carbon sources; while the second one was categorized as glucose-dependent, because it was only significantly changed during growth on glucose (Fig. 3A). The fraction of genes having transcriptional change in a glucose-dependent fashion (73%) was much higher than that being glucose-independent (27%; same direction in both carbons, 26%; different direction in both carbons, 1%). Functional gene enrichment was done to inspect which pathways or genes were associated with the two groups (Fig. 3B and 3C). In the group of genes being glucose-independent, genes involved in galactose, glucose and reserve carbohydrates metabolisms were detected in all three evolved mutants. Furthermore, genes having a transcription factor binding site (AGGGG) were also identified in this group. Based on the YEASTRACT database this binding site is for binding of Gis1p, Rph1p, Msn2/4p and Nrg1p. Especially, *PGM2*, *HXK1* and *GLK1* were highly up-regulated, while *HXK2* was down-regulated in all evolved mutants on both carbon sources. At the metabolite level, the concentration of trehalose and glycogen were significantly increased in all three evolved strains compared to the reference strain during growth on both carbon sources. The genes that showed significant change only during growth on glucose are also shown in the heat map in Fig. 3C. Most of these genes

were related to the central carbon metabolism; up-regulation of genes involved in oxidative phosphorylation, mitochondria organization and down-regulation of genes involved in glycolysis, purine metabolism and amino acid metabolic processes. One of the evolved mutants, 62B strain was quite different from the other two evolved mutants based on the PCA analysis (Fig. 2). This separation can mainly be explained by differences in expression of genes involved in the ergosterol pathway (Supplementary Figure 1) (Hong et al., 2011). The genes in this pathway were up-regulated during growth on both carbon sources; and unique for this strain is the high level of dihydroergosterol and low of ergosterol (Supplementary Figure 3).

3.3. Fermentation physiology of site-directed mutants

In order to evaluate whether the reduced glucose utilization of the galactose-evolved mutants was caused by beneficial mutations or neutral mutations for galactose utilization we used whole genome sequencing data of the evolved mutants from our previous study (Hong et al., 2011). Based on analysis of the whole genome sequencing data we identified several site-directed mutations, and we found that by introducing some of these point mutations in the reference strain the gross phenotype can be partly restored in terms of galactose metabolism. Two mutations in the *RAS2* gene, *RAS2*^{Lys 77} (from evolved mutant 62A) and *RAS2*^{Tyr112} (from evolved mutant 62B), were proven to be beneficial mutations for galactose utilization (Hong and Nielsen, 2012). A mutation identified in *ERG5*, *ERG5*^{Pro 370} in evolved mutant 62B, could explain the changes of transcripts and metabolites in the ergosterol pathway, but it showed almost no effect in terms of galactose utilization, and this mutation can therefore be categorized as a neutral mutation during growth on galactose. The site-directed mutants that carried these three mutations individually were designated as RAU (mutation *RAS2*^{Lys 77}), RBU (mutation *RAS2*^{Tyr112}) and EBU (mutation *ERG5*^{Pro 370}). Here we characterized the growth of these mutants on glucose and compared the growth kinetics with data on galactose metabolism obtained in our earlier study (Hong and Nielsen, 2012) (Fig. 4). Even though the single point mutations could not explain all the physiological changes of the evolved mutants for growth on both carbon sources, the site-directed mutants that had mutations in the *RAS2* gene presented trade-off physiology in terms of carbon source utilization. As found earlier (Hong and Nielsen, 2012), the site-directed mutants RAU and RBU showed improved maximum specific growth rate on galactose, whereas the strains clearly had a reduction in this kinetic parameter during growth on glucose (Fig. 4A). The RAU strain

exhibited even more extended changes than the evolved mutants on both carbon sources (Fig. 1A and Fig. 4A). Even though the specific glucose uptake rate of the RAU and RBU strains was insignificant compared to the increase of their specific galactose uptake rate, the specific ethanol production rate was significantly decreased on glucose like the evolved mutants (Fig. 4B). Overall, the mutations in the *RAS2* gene were positive for galactose utilization but negative for glucose. However, in case of the EBU strain carrying a mutation in *ERG5*, the physiological differences compared to the reference strain on both carbon sources looked smaller not only in terms of the maximum specific growth rate but also in terms of the specific galactose uptake rate (Fig. 4).

4. Discussion

Three adaptively evolved mutants of *S. cerevisiae* selected for improved galactose metabolism showed attenuated glucose utilization (Fig. 1). The strains were evolved on galactose for around 400 generations resulting in improved galactose utilization (Hong et al., 2011). This finding indicated that physiological trade-off in the availability of carbon sources between galactose and glucose happened in the evolved mutants. To investigate how this phenotypic change was related to the intracellular molecular state, analyses of the transcriptome and targeted metabolites were implemented. Common transcriptional changes in the three evolved mutants compared to the reference strain were almost similar when the strains were grown in galactose and glucose (Fig. 2 and Fig. 3). Thus, even though the carbon source was different, the responses of the evolved mutants at the transcriptome level were the same. The results from targeted metabolites showed a similar common change. The genetic changes that could be involved in this trade-off were further explored and as mutations in *RAS2* were identified as key mutations for improving galactose utilization (Hong and Nielsen, 2012) it was interesting to evaluate whether these mutations results in trade-off in terms of glucose metabolism. The site-directed mutants carrying these mutations showed attenuated glucose utilization like the evolved mutants (Fig. 4). In comparison, another mutation in *ERG5*, that was considered a neutral mutation for galactose utilization, did not show any change in glucose utilization (Fig. 4 and Supplementary Figure 3).

4.1. The cost for improving galactose utilization may come from diminishing glucose utilization.

The glucose utilization ability seems to be collateral cost for having improved galactose availability in the evolved mutants (Fig.1). Two engineered mutants were used as control strains. These engineered mutants having an improved specific galactose uptake rate also showed a reduction in the specific glucose uptake rate; hence in this case, the physiological trade-off in the carbon consumption rate was able to be linked to known genetic modifications. Among those modifications, overexpression of *PGM2* gene was commonly detected in all galactose-evolved mutants on both carbons. The enzyme phosphoglucomutase (*PGM2*) involves the bidirectional conversion between glucose-6-phosphate and glucose-1-phosphate. When glucose is used in yeast, glucose-6-phosphate is directly generated from glucose by hexokinase (*HXK2*). In contrast, when galactose is consumed, glucose-1-phosphate is an intermediate in the formation of glucose-6-phosphate. Since the conversion of glucose-1-phosphate into glucose-6-phosphate is an important step in the galactose pathway, up-regulation of *PGM2* has been shown to be a beneficial target for improving galactose utilization (Bro et al., 2005; Hong et al., 2011; Lee et al., 2011). However, as Pgm2p also can convert glucose-6-phosphate to glucose-1-phosphate, it could delay the supply of glucose-6-phosphate to glycolysis by converting it to glucose-1-phosphate. This reversibility of Pgm2p may likely contribute to the trade-off physiology between glucose and galactose utilization. The results of growing the *PGM2* strain on both carbon sources in terms of the specific carbon uptake rate support this explanation. Not only the *PGM2* strain but also *SO16* and all the evolved mutants showed significant up-regulation of *PGM2* during growth on galactose (Hong et al., 2011), and also, during growth on glucose all galactose-evolved mutants showed significant upregulation of *PGM2* (Fig. 3B). Therefore over-expression of *PGM2* can be one of the reasons for the observed trade-off. The *SO16* strain contains a deletion of *MIG1*, which is key regulator for glucose repression (Ostergaard et al., 2000). Even though the deletion of *MIG1* was related to the improvement of the specific galactose uptake rate, a severe effect was detected on the specific glucose uptake rate (Fig. 1B). This result implies that the mutation in regulatory genes seems to be related to the trade-off effect.

In addition to the observation of trade-off in terms of utilizing galactose and glucose, the magnitude in improvement and attenuation of carbon source utilization was not symmetrical (Fig. 1). The galactose-evolved mutants showed more improvement of the maximum specific growth rate on galactose than there was attenuation of the specific growth rate on glucose (Fig. 1A); whereas the relative increase of the specific galactose uptake rate was smaller the relative

decrease in the specific glucose uptake rate (Fig. 1B). This could indicate that there could be other types of trade-off in utilization of other carbon sources metabolism or at conditions beyond the interrelationship between galactose and glucose.

4.2. Transcriptional changes with key metabolites of three evolved mutants reveal antagonistic pleiotropy between glucose and galactose.

An evolutionary trade-off provides the theoretical perspective about the gain and loss of specific traits during the evolutionary process (Elena and Lenski, 2003). Two molecular mechanisms inducing the evolutionary trade-offs have been suggested; 1) antagonistic pleiotropy (AP), where the same mutations are related to beneficial traits in one environment and detrimental in another, and 2) mutation accumulation (MA), where different mutations are associated with the trade-off consequence. The comparison of fitness changes in different environments between mutants and an ancestral strain has been studied to prove the presence and plausible mechanism of the evolutionarily trade-offs (Cooper and Lenski, 2000). Recently, the underlying molecular genetic basis for adaptation to glucose limitation has been investigated by transcriptome analysis and whole genome sequencing (Wenger et al., 2011). Similarly, in this study, transcriptome data was used to interpret the trade-off mechanism between growth on glucose and galactose.

Through comparison of the overall transcriptional profile of all the strains by PCA, the transcriptional changes in the evolved mutants during growth on glucose was found to be similar to the changes observed during growth on galactose (Fig. 2). Since the distance of PC1 between the reference strain grown on galactose and on glucose was almost the same as for each of the evolved mutants, the change could most likely be explained by differences in carbon source and not by the mutations; while the distances of PC2 and PC3 between the evolved mutants and the reference strain for each carbon source may be triggered by the mutations. The length of these distances was also similar for growth on galactose and glucose. These results indicate that transcriptional changes observed during growth on glucose could be induced by the mutations that also caused transcriptional alteration during growth on galactose. If there are beneficial transcriptional changes for galactose utilization and these changes are kept during growth on glucose they could be related to the decreased glucose utilization. Consequently, antagonistic pleiotropy (AP) could be suggested as the main mechanism for the trade-off in carbon source

utilization in the evolved mutants.

Assuming that the trade-off mechanism is AP, specific pathways or genes involved in these transcriptional changes were investigated. Differentially expressed genes common in all evolved mutants grown on glucose were paired with the ones identified for growth on galactose (Supplementary Figure 1C). The genes that changed on both carbon sources were designated to be glucose-independent, and the genes that had altered expression only during growth on glucose were categorized as glucose-dependent (Fig. 3A). Analysis of both groups could give clue about which trade-off mechanism is dominant. For example, if the genes in the glucose-independent group are the main contributor to the differentially expressed genes and the functions of those genes are related to utilization of both carbon sources, AP can be considered the prevailing mechanism. Otherwise, if the glucose-dependent group is the major contributor and the functions of differentially expressed genes in this group are associated to glucose utilization but not galactose utilization, MA could be the governing mechanism for the trade-off (Supplementary Figure 2). First, since the glucose-dependent group is composed of a much higher portion (73 %) of the differentially expressed genes than the glucose-independent group (27 %), this may point to MA as the dominant mechanism. In this case, unique changes on glucose could be triggered by the mutations that are neutral on galactose. However, the results of functional enrichment in both groups revealed that the dominant trade-off mechanism seems to be AP (Fig. 3B and 3C). Thus, in the glucose-independent group *PGM2* was up-regulated during growth of both carbons. As explained earlier, up-regulation of this gene was beneficial for galactose utilization and negative for glucose utilization (Fig. 1B). Secondly, the galactose-evolved mutants showed down-regulation of *HXK2* not only on galactose but also on glucose compared to the reference strain; and other hexokinases (*HXK1* and *GLK1*) were up-regulated (Fig. 3B). Hexokinase isoenzyme 2 (*HXK2*) is predominant during growth on glucose, while other hexokinase *HXK1* and *GLK1* are induced during growth on non-glucose sugars. When hexokinase isoenzyme 2 is involved in glucose metabolism the two other hexokinases are repressed (Rodriguez et al., 2001). Therefore the change in expression level of these hexokinases is likely related to the trade-offs in galactose and glucose utilization, which again suggests the AP mechanism. Thirdly, reserve carbohydrate metabolism commonly displays significant changes in transcripts and metabolites on both carbon sources (Fig. 3B). Activation of reserve carbohydrates metabolism is possibly helpful to increase galactose utilization, because this metabolism consumes one of the accumulated intermediates in

galactose pathway, glucose-1-phosphate. In contrast, this activation may be negative for glucose utilization by delaying supply of glucose-6-phosphate to glycolysis. Lastly, the genes that have a transcription factor binding site for Gis1p, Rph1p, Msn2/4p and Nrg1p were commonly up-regulated. These transcription factors are associated to nutrient signaling. Gis1p, Rph1p and Msn2/4p are involved in glucose limitation and controlled by the Ras/PKA, TOR and Sch9 pathways (Orzechowski Westholm et al., 2012). Nrg1p are related to glucose repression by forming a complex with Cyc8p and Tup1p like Mig1 (Zhou and Winston, 2001). When one of these genes is inactivated, the sensitivity to glucose is reduced. Therefore, it could be possible that the evolved mutants released the tight control of glucose metabolism resulted in improved galactose metabolism, but resulting in a trade-off in terms of glucose utilization. This result also points to the AP mechanism as being the dominant. The conserved transcriptional changes are summarized in Fig. 5.

In case of the glucose-dependent group, most of the significantly changed pathways were in the central carbon metabolism representing increased respiratory metabolism (Fig. 3B). The change of these pathways is not uniquely related to glucose metabolism; and the expression pattern on galactose was also quite similar even though at a lower significance level. This result implies that the expression profile of the evolved mutants is conserved on galactose and glucose. Consequently, this finding also support that MA is not the dominant mechanism for the trade-offs on the utilization of these two carbon source.

In addition to common changes in all three evolved mutants, unique change in one evolved mutant was also significantly conserved on both carbons the 62B strain shows a similar pattern of transcripts and metabolites in the ergosterol pathway during growth on both carbon sources (Supplementary Figure 3). Although the influence of changes in the ergosterol pathway on galactose and glucose metabolism is not clearly known, this common pattern could be due to the same mutation. This result also partially supports the AP mechanism.

4.3. Mutations in the RAS2 gene indicate antagonistic pleiotropy mechanism for trade-off in carbon source utilization.

The elucidation of the trade-off mechanism at the genetic level was challenging, since a large number of mutations were identified in the evolved mutants (Hong et al., 2011), and this is a typical scenario in adaptive evolution. Therefore, it is necessary to have efficient screening

methods for confining the number of governing mutations. So, other omics data were used to find significantly changed metabolic functions or genes. This selection gave the hint to find related mutations; and the analysis of common pathway that had mutations in all three evolved mutants also provided additional clue. Consequently, mutations were identified in the Ras/PKA signaling pathway, where all evolved mutants had mutation; and in the ergosterol pathway, where only one evolved mutant had mutation, but it was clearly linked to transcripts and metabolite changes. Characterization of the site-directed mutants containing these point mutations individually proved that they were mutations being beneficial or neutral to galactose utilization (Hong and Nielsen, 2012). These mutations are therefore conceivable targets for testing trade-off mechanism of the galactose-evolved mutants for growth on glucose. If a site-directed mutant that shows positive result in terms of galactose utilization displays negative outcome in terms of glucose utilization, it may be suggestive of the AP mechanism; otherwise it could prove the MA mechanism indirectly. Also, if another site-directed mutant that has neutral mutation on galactose shows deleterious glucose utilization, this result may indicate MA mechanism. The site-directed mutants, RAU and RBU that had beneficial mutation in *RAS2* gene for galactose utilization presented significant reduction in the maximum specific growth rate on glucose (Fig. 4A). Moreover, even though the specific glucose uptake rate was not changed, there was a substantial decrease in the specific ethanol production rate on glucose (Fig. 4B), which means that there is a reduction of the fermentation metabolism in the evolved mutations when they are grown on glucose. These results demonstrate AP mechanism effects by mutations in *RAS2*. *PGM2* up-regulation by these *RAS2* mutations was proven in our previous study (Hong and Nielsen, 2012), and as explained, *PGM2* over-expression could trigger the trade-off between galactose and glucose utilization. The Ras/PKA signaling pathway has broad roles for nutrient and stress related metabolism in yeast (Tamaki, 2007). The activity of this pathway is closely linked to the concentration of extracellular glucose. High concentration of glucose stimulates this activity resulted in repression of other carbon utilization pathway. The mutations in *RAS2* may interfere with this signaling; and these changes could therefore reduce glucose utilization. As detected in the *MIG1* knock-out strain, SO16, controlling the activity in regulatory pathway seems to be related to the trade-off in carbon utilization (Fig. 5). The site-directed mutant, EBU that had neutral mutations in terms of galactose utilization also showed negligible changes in glucose utilization (Fig. 4). The *ERG5* mutation did, however, affect the transcripts of genes and metabolites in the ergosterol pathway

during growth on galactose, so this mutation was expected to be a possible target for a MA mechanism. However, the result of this mutation during growth on glucose was also neutral and the changes in transcripts of genes and metabolites were the same as during growth on galactose (Fig. 4). Therefore any evidence for the MA mechanism could not be proved using this target. However, it is curious why the evolved mutant kept molecular changes that were not related to the utilization of both carbon sources. This mutation was identified only in one evolved mutant, so it is possible that this mutation is just randomly generated during the evolutionary process. On the other hand, this mutation has influence in combination with other mutations to generate epistatic effect (Hong and Nielsen, 2012), i.e. standalone the *EGR5* mutation displays only molecular changes without alteration in carbon utilization. In conclusion, AP mainly explains the trade-off mechanism in carbon source utilization of the galactose-evolved mutants through the mutations in the Ras/PKA signaling pathway (Fig. 5).

5. Conclusion

Evolutionary engineering has become a useful tool in the development of novel strains to be used for industrial production of fuels, chemicals and materials, and here we demonstrate how systems biology tools can be used to elucidate the link between phenotype and genotype. In addition, this kind of system level analysis opens the possibility to explore evolution mechanism in many other cases, e.g. how an evolved mutant gained its improved traits based on natural constraints of evolution. This can further feedback to metabolic engineering where new knowledge can be used to design evolutionary engineering strategies. In this study, the trade-off mechanisms in carbon source utilization for galactose-evolved mutants were investigated at the systems level. We found that there is a trade-off between galactose and glucose utilization mostly due to conserved transcriptional changes in galactose and glucose metabolic pathway by *RAS2* mutations (Fig. 5). One of the possible strategies to further improve galactose utilization or to make strains that could utilize both carbon sources efficiently, which would be required in industrial fermentations, may be to generate a mutation library of the *RAS2* gene and then screen for improved strains. Dependent on mutations one may move from strains that are specialist for one carbon source to strains that are generalist for more than one source.

It is, however, important to keep in mind that there are also other evolutionary mechanisms

related to strain development, such as negative epistasis and clonal interference among beneficial mutations (Elena and Lenski, 2003; Kao and Sherlock, 2008; Khan et al., 2011; Kvitek and Sherlock, 2011). These mechanisms indicate that there are many chances to lose beneficial mutations, which means that adaptive evolution could lead to lack of detecting beneficial mutations and hereby to design new combinations of them. It is therefore important to accumulate examples about evolutionary mechanisms at detailed molecular levels supported by systems biology approaches for further advancing the use of adaptive evolution in industrial biotechnology.

Acknowledgements

We thank CJ CheilJedang Doctoral Fellowship for financial support to KKH. We acknowledge the Chalmers Foundation, The Knut and Alice Wallenberg Foundation, the EU funded project UNICELLSYS (contract no. 201142) and the European Research Council (contract no. 247013) for financial contributions to this project.

References

- Bennett, A. F., Hughes, B. S., 2009. Microbial experimental evolution. Am. J. Physiol. Regul. Integr. Comp. Physiol. 297, R17-25.
- Bennett, A. F., Lenski, R. E., 2007. An experimental test of evolutionary trade-offs during temperature adaptation. Proc. Natl. Acad. Sci. USA 104 Suppl 1, 8649-54.
- Bro, C., Knudsen, S., Regenberg, B., Olsson, L., Nielsen, J., 2005. Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. Appl. Environ. Microbiol. 71, 6465-72.
- Bro, C., Nielsen, J., 2004. Impact of 'ome' analyses on inverse metabolic engineering. Metab. Eng. 6, 204-11.
- Chang, S. L., Leu, J. Y., 2011. A tradeoff drives the evolution of reduced metal resistance in natural populations of yeast. Plos Genetics 7, e1002034.
- Conrad, T. M., Frazier, M., Joyce, A. R., Cho, B. K., Knight, E. M., Lewis, N. E., Landick, R., Palsson, B. O., 2010. RNA polymerase mutants found through adaptive evolution reprogram Escherichia coli for optimal growth in minimal media. Proc. Natl. Acad. Sci. USA 107, 20500-5.
- Conrad, T. M., Lewis, N. E., Palsson, B. O., 2011. Microbial laboratory evolution in the era of genome-scale science. Mol. Syst. Biol. 7, 509.
- Cooper, V. S., Lenski, R. E., 2000. The population genetics of ecological specialization in evolving Escherichia coli populations. Nature 407, 736-9.
- Daran-Lapujade, P., Rossell, S., van Gulik, W. M., Luttik, M. A., de Groot, M. J., Slijper, M., Heck, A. J., Daran, J. M., de Winde, J. H., Westerhoff, H. V., Pronk, J. T., Bakker, B. M., 2007. The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. Proc. Natl. Acad. Sci. USA 104, 15753-8.
- Dettman, J. R., Rodrigue, N., Melnyk, A. H., Wong, A., Bailey, S. F., Kassen, R., 2012. Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. Mol. Ecol. 21, 2058-77.
- Elena, S. F., Lenski, R. E., 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat. Rev. Genet. 4, 457-69.
- Herring, C. D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M. K., Joyce, A. R., Albert, T. J., Blattner, F. R., van den Boom, D., Cantor, C. R., Palsson, B. O., 2006. Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale. Nat. Genet. 38, 1406-12.
- Hong, K. K., Nielsen, J., 2012. Recovery of phenotypes obtained by adaptive evolution through inverse metabolic engineering. Appl. Environ. Microbiol. [Under Review: 1st revision].
- Hong, K. K., Vongsangnak, W., Vemuri, G. N., Nielsen, J., 2011. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. Proc. Natl. Acad. Sci. USA 108, 12179-84.
- Kao, K. C., Sherlock, G., 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. Nat. Genet. 40, 1499-504.
- Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., Cooper, T. F., 2011. Negative epistasis between beneficial

- mutations in an evolving bacterial population. *Science* 332, 1193-6.
- Kvitek, D. J., Sherlock, G., 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *Plos Genetics* 7, e1002056.
- Lee, K. S., Hong, M. E., Jung, S. C., Ha, S. J., Yu, B. J., Koo, H. M., Park, S. M., Seo, J. H., Kweon, D. H., Park, J. C., Jin, Y. S., 2011. Improved galactose fermentation of *Saccharomyces cerevisiae* through inverse metabolic engineering. *Biotechnol. Bioeng.* 108, 621-31.
- Orzechowski Westholm, J., Tronnersjo, S., Nordberg, N., Olsson, I., Komorowski, J., Ronne, H., 2012. Gis1 and Rph1 regulate glycerol and acetate metabolism in glucose depleted yeast cells. *PLoS One*. 7, e31577.
- Ostergaard, S., Olsson, L., Johnston, M., Nielsen, J., 2000. Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat. Biotechnol.* 18, 1283-6.
- Reimand, J., Arak, T., Vilo, J., 2011. g:Profiler--a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res.* 39, W307-15.
- Rodriguez, A., De La Cera, T., Herrero, P., Moreno, F., 2001. The hexokinase 2 protein regulates the expression of the GLK1, HXK1 and HXK2 genes of *Saccharomyces cerevisiae*. *Biochem. J.* 355, 625-31.
- Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., Quackenbush, J., 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34, 374-8.
- Smith, K. M., Liao, J. C., 2011. An evolutionary strategy for isobutanol production strain development in *Escherichia coli*. *Metab. Eng.* 13, 674-81.
- Tamaki, H., 2007. Glucose-stimulated cAMP-protein kinase A pathway in yeast *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* 104, 245-50.
- Wenger, J. W., Piotrowski, J., Nagarajan, S., Chiotti, K., Sherlock, G., Rosenzweig, F., 2011. Hunger Artists: Yeast Adapted to Carbon Limitation Show Trade-Offs under Carbon Sufficiency. *Plos Genetics* 7.
- Zhou, H., Winston, F., 2001. NRG1 is required for glucose repression of the SUC2 and GAL genes of *Saccharomyces cerevisiae*. *BMC Genet.* 2, 5.

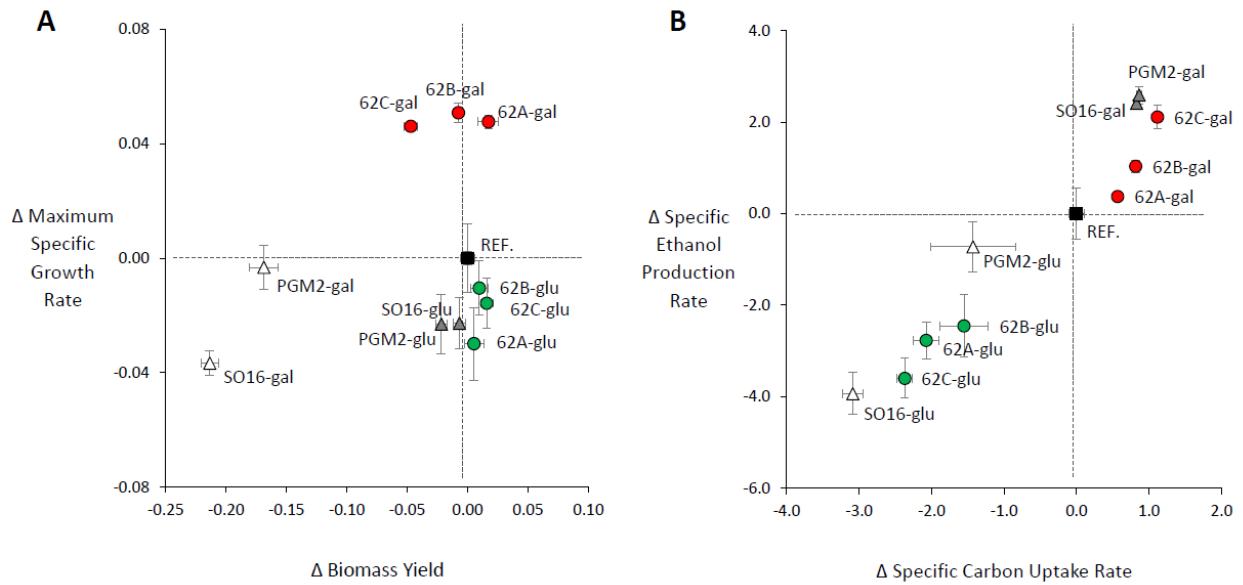


Fig. 1: Fermentation physiology of the evolved mutants and the engineered mutants compared to the reference strain in galactose (gal) and glucose (glu). The data are shown as: A. Absolute changes (Δ) in the maximum specific growth rate (h^{-1}) and a biomass yield (C-mol/C-mol substrate). B. Absolute changes in the specific carbon uptake rate (mmoles/g DW/h) and the specific ethanol production rate (mmoles/g DW/h). Error bar represents standard error from biological duplicates in bioreactors.

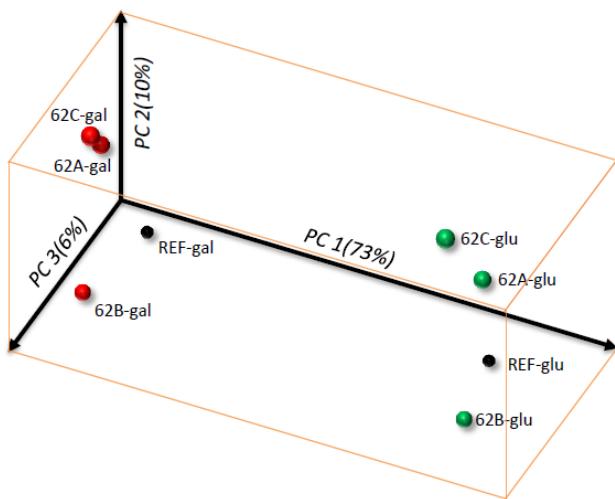


Fig. 2: Transcriptome analysis of the evolved mutants and the reference strain in galactose (gal) and glucose (glu) through principal component analysis (PCA). The results are projected by the first three PCs, which covered 89% of the variance.

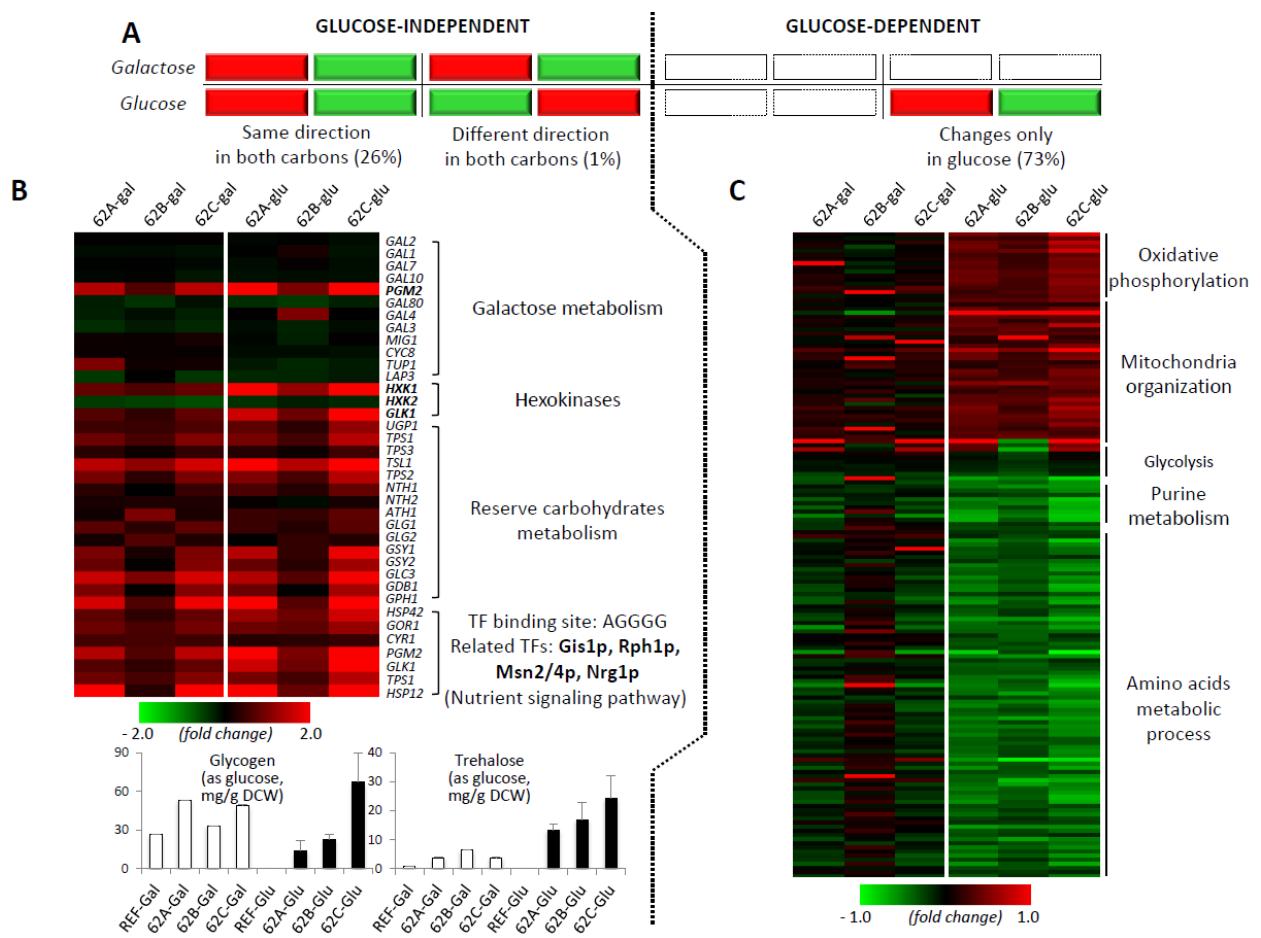


Fig. 3: Patterns of common molecular changes of the evolved mutants compared to the reference strain based on separation of glucose-independent and glucose-dependent genes. Glucose-independent genes are differentially expressed genes during growth on both carbon sources; glucose-dependent genes are differentially expressed genes only during growth on glucose (A), Specific pathways and targeted metabolites being glucose-independent (B), Specific pathways being glucose-dependent (C). Error bars in the concentration of glycogen and trehalose represent standard error from biological duplicates in bioreactor experiments.

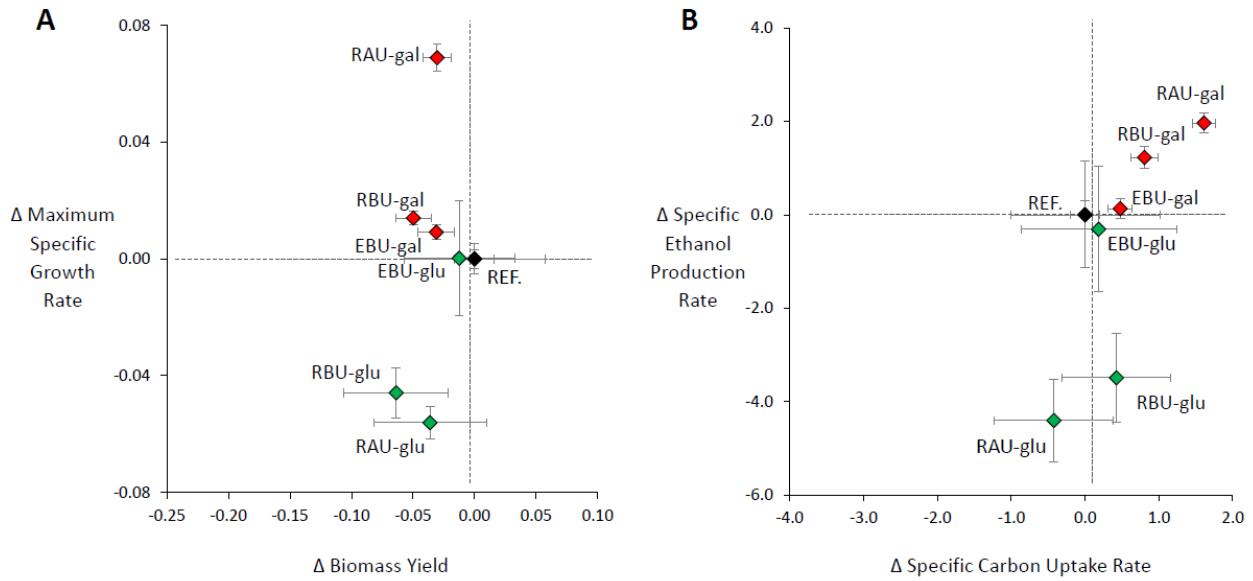


Fig. 4: Fermentation physiology of the site-directed mutants compared to the reference strain in galactose (gal) and glucose (glu) through correlation between different values (Δ) of a maximum specific growth rate and a biomass yield (A), a specific carbon uptake rate and a specific ethanol production rate (B). The reference strain for the site-directed mutants is CEN.PK 113-5D having *URA3* marker on an empty plasmid. Error bar represents standard error from biological duplicates on galactose in bioreactors and biological triplicate on glucose in baffled flasks.

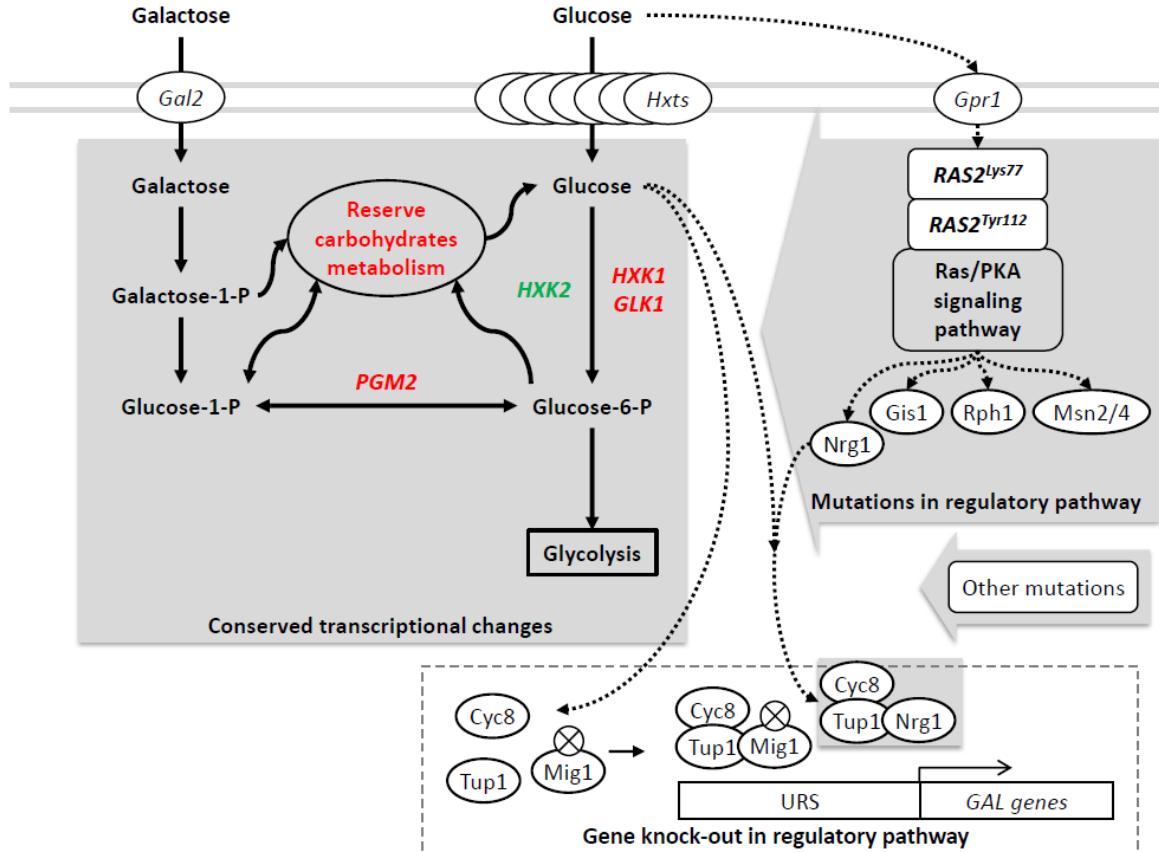
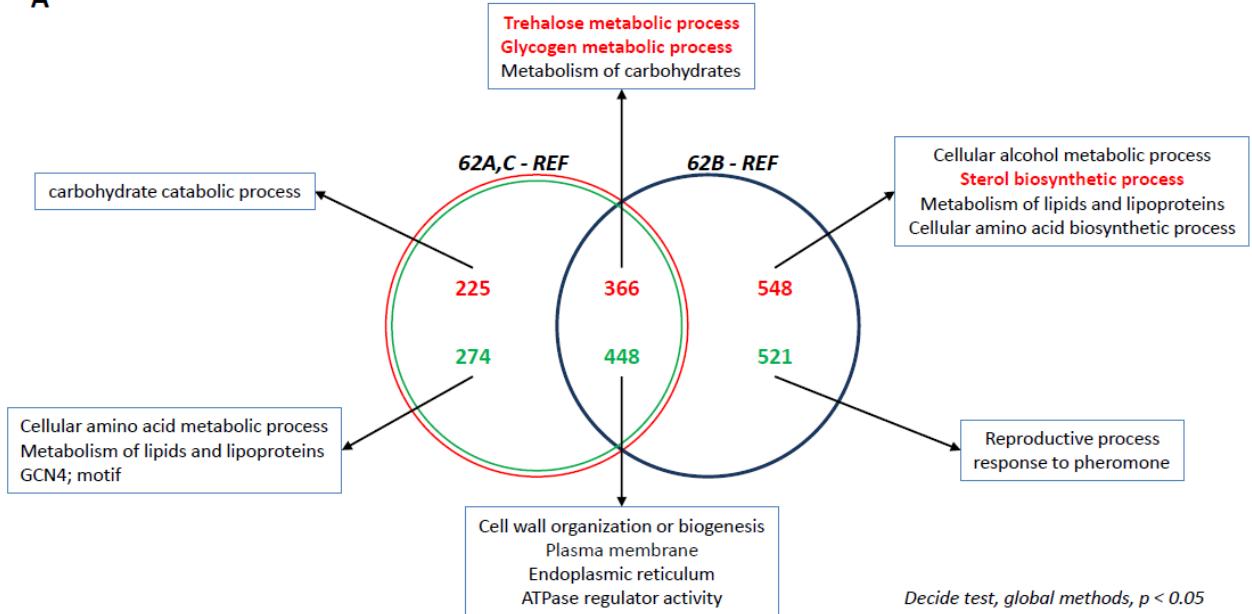
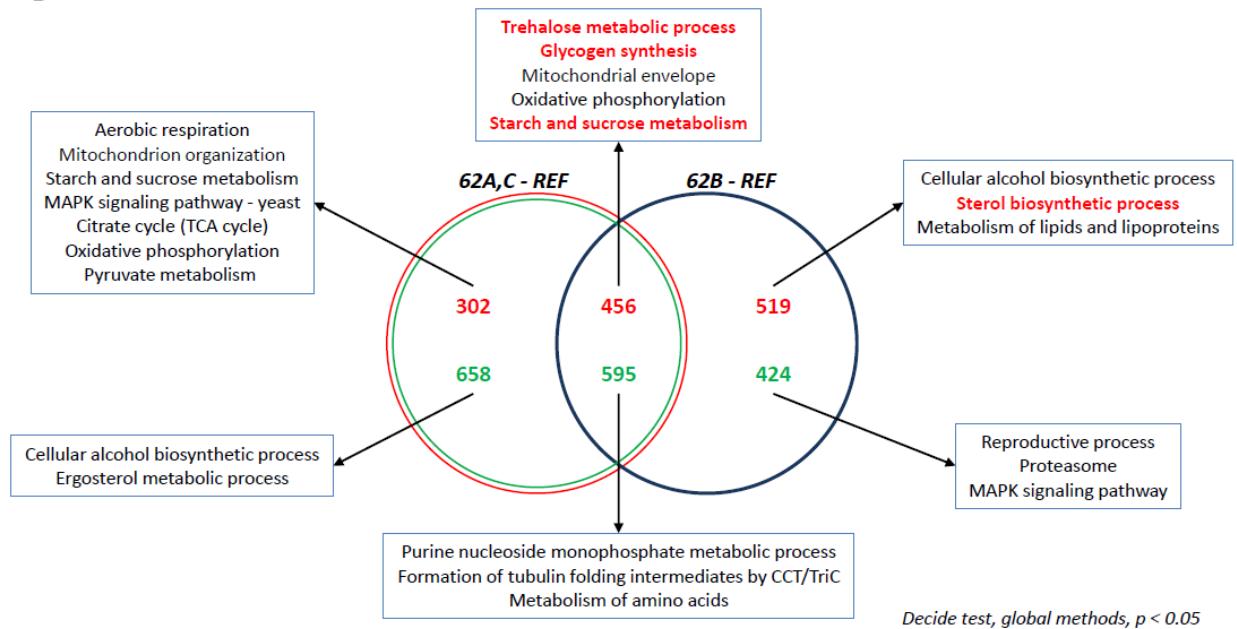


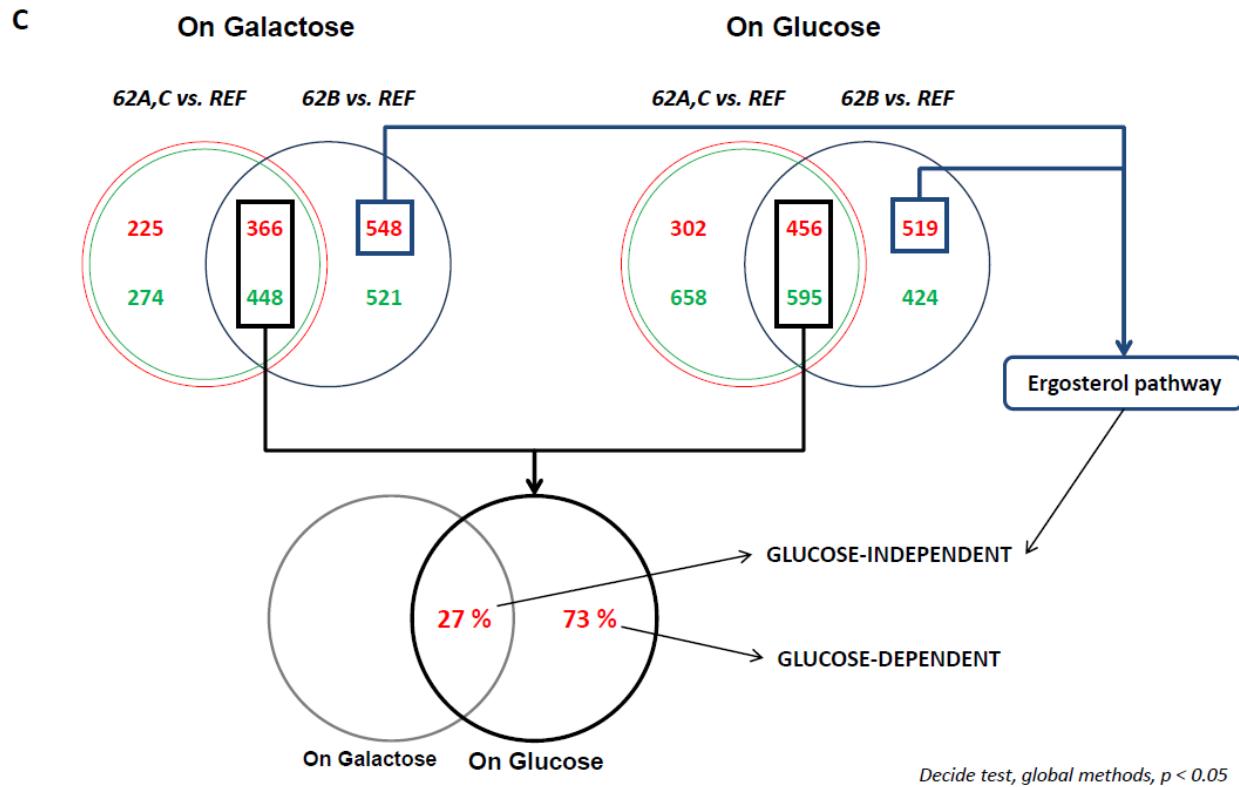
Fig. 5: Summary of possible molecular mechanism for the trade-off in galactose and glucose utilization by yeast. Colored letters (red and green) mean transcriptional changes with up- and down-regulation, respectively. Gray boxes (square and arrow shape) exhibit the changes in the evolved mutants. A dotted box means the change only occurred only in SO16 (knock-out of *MIG1*). Dot arrows represent signaling flow and solid arrows represent carbon flow.

A

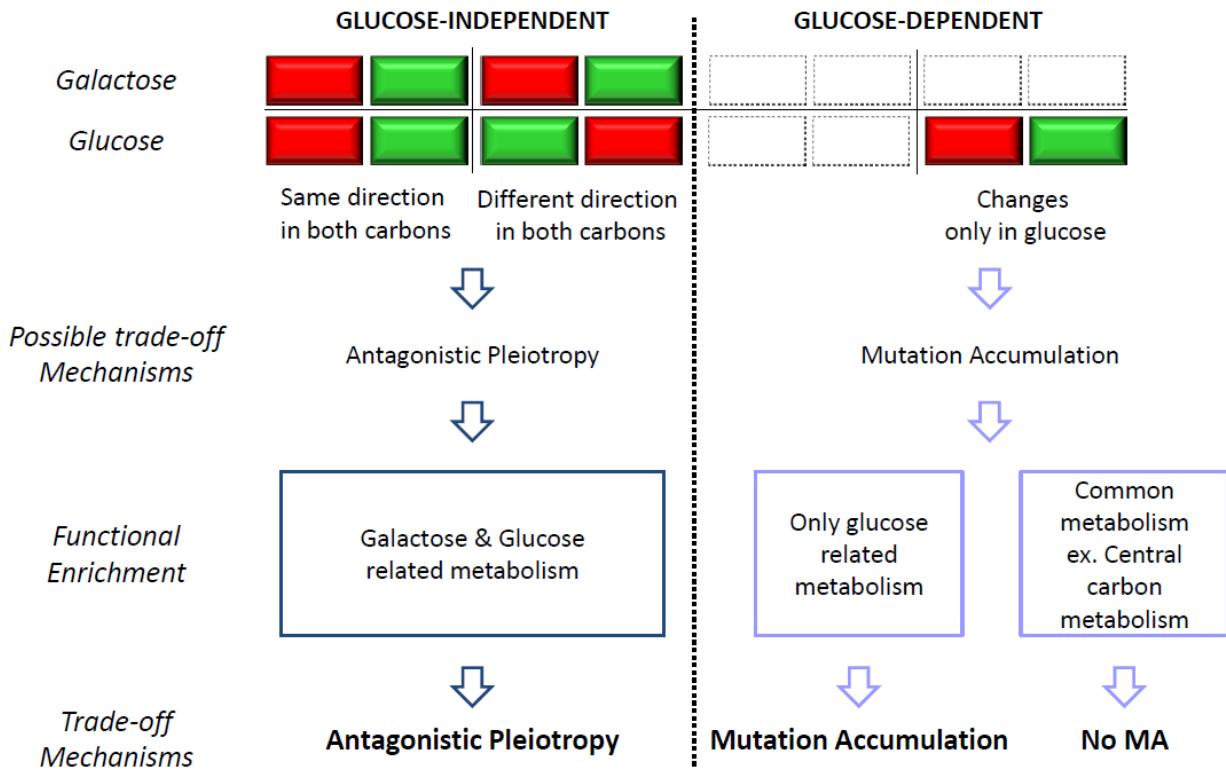


B



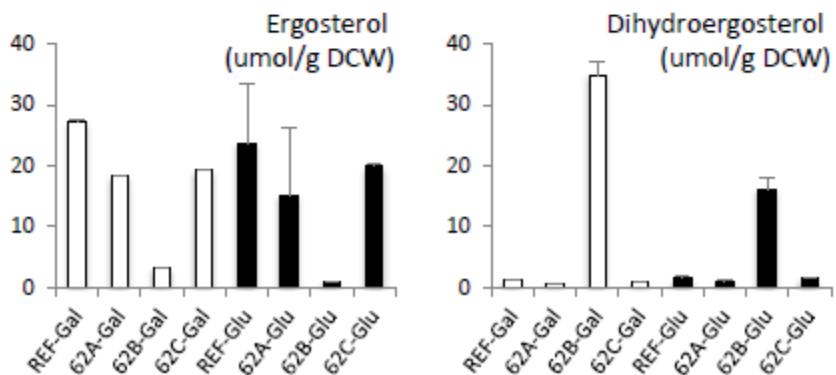
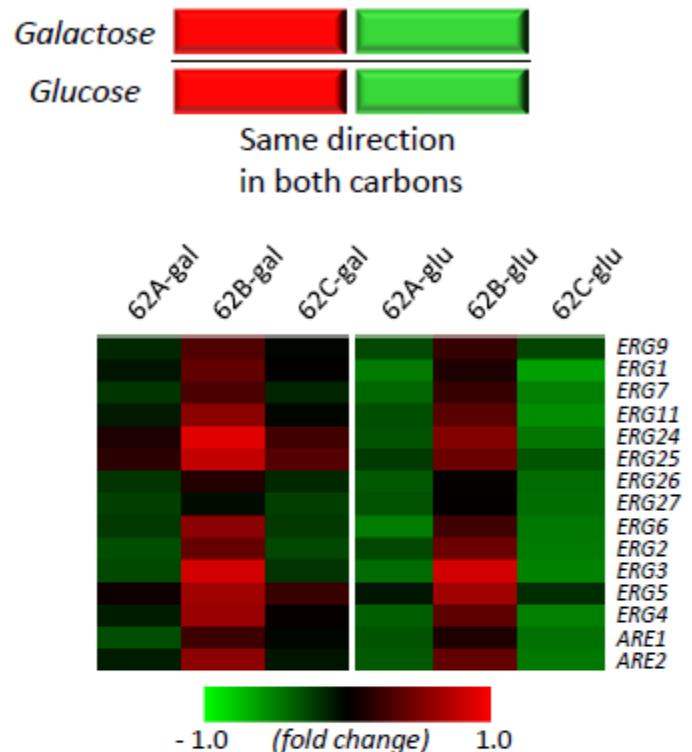


Supplementary Figure 1: Analysis of differentially expressed genes among the evolved mutants compared to the reference strain. Differentially expressed genes (adjust p-value < 0.05) on galactose (A) and on glucose (B) are categorized in a Venn diagram. The functions of genes in each part are analyzed by using a cumulative hyper-geometric model based on GO terms, REACTOME, KEGG and TRANSFAC database ($P < 0.001$). The process that separates common transcriptional changes on both carbon sources into glucose-independent and glucose-dependent are depicted (C). Red color letter without % mark means the number of genes being up-regulated; green one means down- regulation.



Supplementary Figure 2: The process for the analysis of trade-off mechanism using transcriptome data. The categorization of differentially expressed genes like in Supplementary Figure 1 makes one deduce possible trade-off mechanism. Then, through functional analysis of genes in those categories, one can deduce which trade-off mechanism is dominant.

GLUCOSE-INDEPENDENT only in 62B strains



Supplementary Figure 3: Patterns of unique molecular changes in the 62B strain. Glucose-independent changes in the ergosterol pathways are expressed by transcripts and targeted metabolites. Error bars in the concentration of ergosterol and dihydroergosterol represent standard error from biological duplicates in bioreactors.

PAPER IV

REVIEW

Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries

Kuk-Ki Hong · Jens Nielsen

Received: 27 October 2011 / Revised: 7 February 2012 / Accepted: 15 February 2012 / Published online: 3 March 2012
 © Springer Basel AG 2012

Abstract Metabolic engineering is the enabling science of development of efficient cell factories for the production of fuels, chemicals, pharmaceuticals, and food ingredients through microbial fermentations. The yeast *Saccharomyces cerevisiae* is a key cell factory already used for the production of a wide range of industrial products, and here we review ongoing work, particularly in industry, on using this organism for the production of butanol, which can be used as biofuel, and isoprenoids, which can find a wide range of applications including as pharmaceuticals and as biodiesel. We also look into how engineering of yeast can lead to improved uptake of sugars that are present in biomass hydrolyzates, and hereby allow for utilization of biomass as feedstock in the production of fuels and chemicals employing *S. cerevisiae*. Finally, we discuss the perspectives of how technologies from systems biology and synthetic biology can be used to advance metabolic engineering of yeast.

Keywords Metabolic engineering · Yeast · Substrate range · Biobutanol · Isoprenoids · Industrial biotechnology

K.-K. Hong · J. Nielsen (✉)
 Novo Nordisk Centre for Biosustainability,
 Department of Chemical and Biological Engineering,
 Chalmers University of Technology,
 412 96 Gothenburg, Sweden
 e-mail: nielsenj@chalmers.se

K.-K. Hong
 Research Institute of Biotechnology,
 CJ CheilJedang, Seoul 157-724, Korea

Introduction

Microbial fermentations have been used for the production of fermented food and beverages since ancient times. Already around 1920 microbial fermentation was introduced for the production of citric acid, and this was the first large-scale industrial production process of a chemical compound based on microbial fermentation. With the development of genetic engineering in the 1970s, it became possible to produce compounds that are not native to microbes, such as pharmaceutical proteins like human insulin and human growth hormone using fermentation technology. Genetic engineering also allowed the transformation of microbes into “cell factories” for the production of chemicals through so-called metabolic engineering [1–3], a field dedicated to design of microbial metabolism to efficiently convert cheap raw materials like glucose, sucrose, and biomass-derived sugars into fuels and chemicals. With the further development of genomics and omics analysis and advanced modeling tools in the field of systems biology [4–8], it has become possible to perform very detailed phenotypic characterization of microorganisms that can serve as efficient cell factories for the production of fuels and chemicals. Thus, the last 10 years have witnessed a substantial technology push in terms of cell factory design, and with the recent desire to develop more sustainable processes for the production of fuels, chemicals, and materials, the chemical industry is trying to exploit these technological developments. The net result is the development of what is generally referred to as industrial biotechnology. There are already several examples of how the fuel and chemical industry is trying to develop novel bioprocesses that can change the primary feedstock from oil to agricultural-based products:

- Dupont, one of the largest chemical companies in the world, has launched a process for production of 1,3-propanediol using a recombinant *Escherichia coli*. 1,3-propanediol is used as one of the key chemicals in the production of the polymer Sorona®, which is used for the manufacturing of fabrics, carpets, and a wide range of plastic-based materials. The process was developed through close collaboration of Dupont, Genencor, and Tate & Lyle.
- DSM has launched a completely biotech route for production of the antibiotic cephalexin, which was earlier produced based on chemical conversion of penicillin.
- BASF has launched a completely biotech route for production of the vitamin B2, riboflavin. The previous process relied on several chemical synthesis steps and the biotech route resulted in both a reduction in raw materials and energy usage [9].
- Dupont has entered into a joint venture with British Petroleum (BP), Butamax, on developing a bio-based production of butanol as a sustainable biofuel.
- Dupont has also entered into another joint venture with Danisco for the development of a second-generation bioethanol production plant that will rely on the use of lignocellulosics as raw materials for ethanol production. BP recently acquired Verenium also with the objective of developing second-generation bioethanol production, but using a different technology.
- ExxonMobil has entered into a joint venture with Synthetic Genomics to develop a novel microalgal-based process for the production of biodiesel.
- Novozymes has entered a joint venture with Cargill with the objective of developing a bio-based process for the production of 3-hydroxypropionic acid, which is to be used for the production of acrylates for the production of a range of personal care products, e.g., diapers and other hygienic products.
- Gevo has launched a process for production of isobutanol that can find application as a biofuel and a commodity chemical.
- Amyris has developed a yeast-based fermentation process for production of farnese that can be used as biodiesel as well as be converted into squalene, which is used in cosmetics.

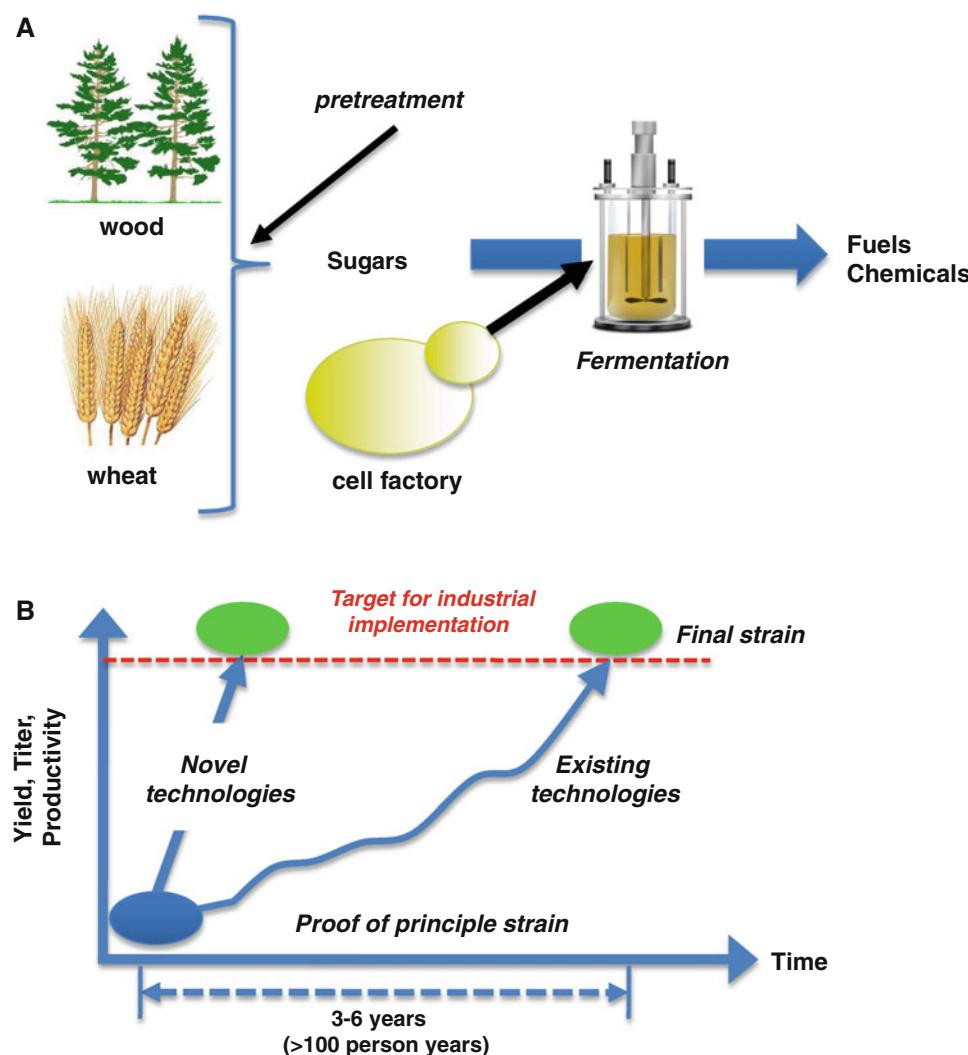
These and many other examples clearly demonstrate two key points: (1) the large chemical and fuel companies are turning to biotech as the solution to develop sustainable processes for the production of fuels and chemicals, and (2) most novel processes are developed through close collaboration/joint ventures involving two or more companies, and often also involve academia or small technology-based companies as a provider of novel technologies. The reason

for the latter is that the development of a novel bioprocess requires a wide range of competences. Traditional chemical companies hold the necessary engineering competence required for scale-up and plant construction, but they may lack competence on the biotech part.

The overall idea in the above-mentioned and many other ventures is to develop bio-based processes that use wheat, corn, sugarcane, or biomass as a raw material for production of fuels and chemicals using so-called biorefineries (Fig. 1a). In a biorefinery, there is pre-treatment of the plant-based raw materials, and this has been well implemented for sugar and starch-based raw materials, i.e., sugarcane, sugar beet, wheat, and corn, and today there is very large scale production of bioethanol and other chemicals using these raw materials. Very efficient enzymes have been developed for the degradation of starch and today many processes are based on simultaneous saccharification and fermentation, where the enzymes are added directly to the fermentation process. Despite the success of sugar- and starch-based bioprocesses, these raw materials will not be able to meet the increasing demand for bio-based products, and there is therefore a need to move to biomass-based raw materials. Here, the pre-treatment is far more complicated and it depends on the raw material to be used, but recently there has been progress on several different fronts, including the development of efficient enzymes that can hydrolyze celluloses and hemicelluloses [10–12]. A key part of developing novel bioprocesses for production of fuels and chemicals is the construction of the cell factory. This cell factory has to meet commercial requirements for yield, productivity, and titer. Often it is possible to quite rapidly develop a proof-of-principle strain that produces the product of interest, whereas it is generally far more time consuming to develop a strain that meets the commercial targets for yield, productivity, and titer (Fig. 1b). There is therefore much interest to develop novel technologies that may speed up this development process, and here it is expected that tools from systems biology may assist. Advancement in the field of synthetic biology may also allow faster development of efficient cell factories, as synthetic biology may provide novel tools for controlled expression of genes, assembly of complete pathways on scaffold proteins, and completely novel enzymatic functions [13]. Nielsen and Keasling [2] recently discussed the synergies between synthetic biology and metabolic engineering.

Microbial fermentation is already today used for the production of a whole range of products (Fig. 2a), but not all of these will fit naturally into the biorefinery concept, which is primarily geared towards fuels and bulk chemicals where there is a requirement for cheap raw materials that are available in large quantities. However, also for products

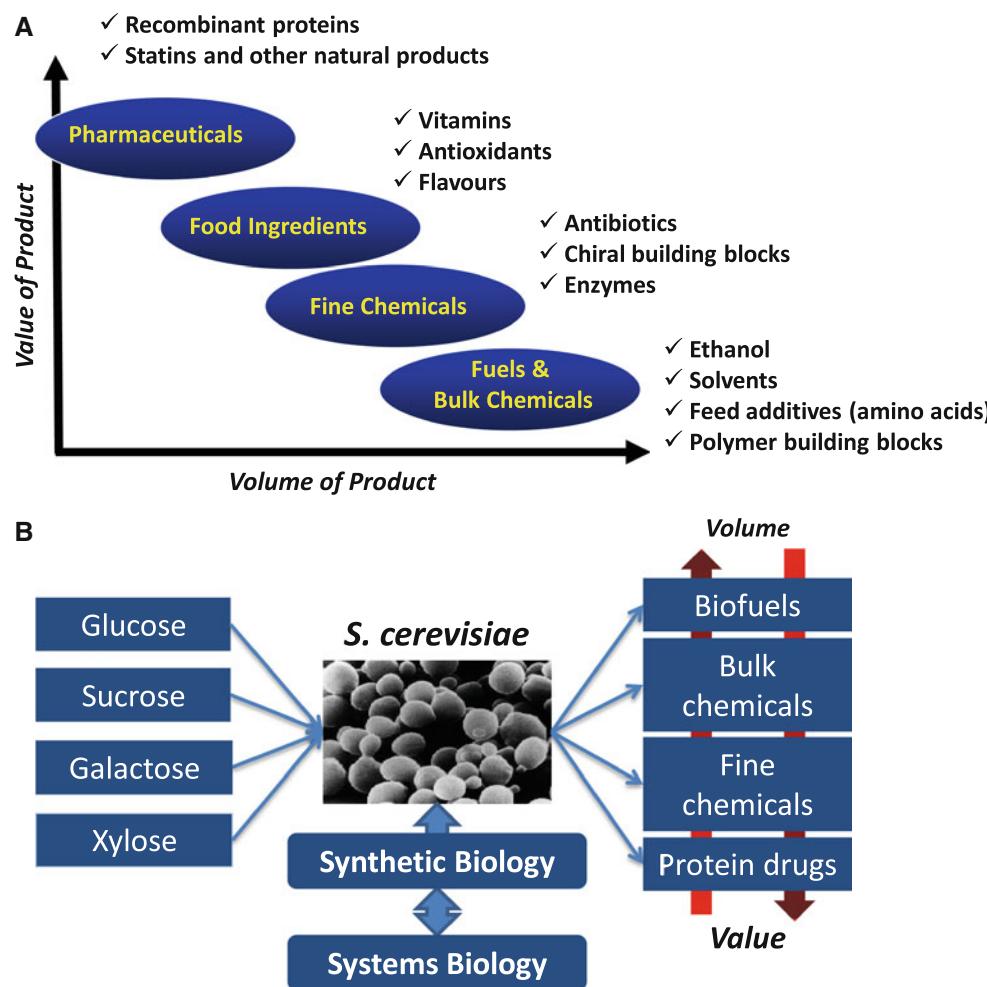
Fig. 1 Illustration of the biorefinery concept and the development time of novel bioprocesses. **a** In a biorefinery, plant-based feedstocks such as sugarcane, corn, wheat, or biomass are converted into sugars that are subsequently used for microbial fermentations. In the fermentation process, cell factories convert the sugars into fuels and chemicals. **b** The development of cell factories is the central research and development process in connection with the development of a novel bioprocess. Construction of an efficient cell factory requires large investment, in particular in connection with bringing the cell factory from proof-of-principle stage where it is producing small amounts of the desired product to a final strain that produces the product at yields, titers, and productivities that make the process financially competitive with fossil fuel-based processes



that do not fit into the biorefinery concept, there is a need for decreasing the development time such that novel products to be used as food and pharmaceutical ingredients can be brought to the market faster. In this context, the type of cell factory plays a very important role, and in recent years there has been some consolidation towards the use of a few industrial platform cell factories that include (not an exhaustive list) the yeast *Saccharomyces cerevisiae*, the bacteria *Escherichia coli*, *Corynebacterium glutamicum* and *Bacillus subtilis*, and the filamentous fungi *Aspergillus niger* and *Aspergillus oryzae*. *S. cerevisiae* is a very attractive cell factory, as it has been demonstrated to be very well suited for industrial production of a range of products due to its robustness and tolerance towards industrial conditions. Thus, it is used for the production of bioethanol, the largest-volume fermentation product by far, and it is also used for the production of several pharmaceuticals, e.g., human insulin, hepatitis vaccines, and human papillomavirus vaccines, and the production of

nutraceuticals, e.g., resveratrol, has been announced to be produced by this cell factory (Fig. 2b). Furthermore, a number of academic studies have illustrated the suitability of this cell factory for the production of a range of chemicals [5, 14], e.g., lactic acid, glycerol, and malic acid. Several recent reviews provide an overview of the many different metabolic engineering examples using yeast as a cell factory [15, 16], and Table 1 provides a summary of some of these key developments. In addition, the wide use of this organism is illustrated by the very large number of patents filed on the use of yeast and/or *S. cerevisiae* for production of fuels and chemicals (Table 2). Yeast also serves as an important model eukaryote, and many fundamental studies have therefore been performed on this organism [54]. It was also the first eukaryotic organism to have its genome sequenced and a number of high-throughput studies have been pioneered using this organism as a model [55–57]. Thus, there is an extensive technology platform in terms of systems biology and

Fig. 2 Range of products and illustration of the key research problems associated with cell factory design and development. **a** Biotech products range from high-value-added to low-value-added products, with the latter being produced in large quantities and the former in small quantities. Examples of the different types of products are indicated. The yeast *S. cerevisiae* is used for the production of products in the whole spectrum. **b** In connection with the development of yeast for the production of different types of products using different sugars as feedstock, there is a need for an extensive platform of technologies from synthetic and systems biology



synthetic biology available for this organism (Fig. 2b), and this makes it a promising host for rapid development as a cell factory for production of novel fuels and chemicals. *S. cerevisiae*, however, has one major limitation in its use, and that is its lack of ability to efficiently grow and metabolize pentoses that are present in hemicelluloses, and therefore result from biomass hydrolysis. However, there has been much interest in metabolic engineering of *S. cerevisiae* for improving its ability to use pentoses, in particular xylose.

Here we will review recent advances in the use of *S. cerevisiae* for production of novel fuels and chemicals with a focus on bio-butanol and isoprenoids. Considering the importance of using pentoses in future biorefineries we will, however, also review work on expanding the substrate range of yeast. Much of the work we will be reviewing has been carried out within companies, and there is therefore limited information in the published literature, and we will therefore to a large extent rely on patents and patent applications. We chose to focus on butanol as an example of an advanced biofuel and isoprenoids as examples of a class of valuable biochemical, as there have been much

development within several companies in the production of these compounds recently.

Extended substrate range

Ethanol is a first-generation biofuel that is being used either in pure form or blended with gasoline, and it is primarily being produced using *S. cerevisiae*. Most of the currently produced fuel ethanol is produced in Brazil and the US. In 2009, about 55% of the world's ethanol production was in Brazil, where the main carbon source was sugarcane, while about 35% is produced in the USA through corn fermentation [58]. To avoid the resource competition between food and fuel, there are demands for the use of the abundant and sustainable non-food resource such as switchgrass and waste cheese whey as well as agricultural by-products like corn-cob and bagasse. Currently available non-food resources are not sufficient to fully replace the fuel produced from oil; however, several efforts to increase non-food biomass have been implemented [59, 60]. These biomass resources comprise diverse types of carbon

Table 1 Example of products and strains of *S. cerevisiae*

Categories	Products	Specific applications	Strains	References
Biofuels	Ethanol	Redox balance problem by inhibiting glycerol formation in anaerobic culture was solved by combining gene deletion (<i>GPD1</i> and <i>GPD2</i>) and integration (<i>mhpF</i> from <i>E. coli</i>) with acetic acid supplementation, which was presented at substantial quantities in lignocellulosic hydrolysates of agricultural residues	CEN.PK102-3A (<i>MATa ura3 leu2</i>)	[17]
	Biobutanol	Overexpression of genes in valine metabolism, <i>ILV2</i> , <i>ILV3</i> , <i>ILV5</i> , and <i>BAT2</i> showed an increased production of isobutanol in <i>S. cerevisiae</i> , which strain was decided as a host because of relative tolerance to alcohols, and robustness in industrial fermentation	CEN.PK 2-1C (<i>MATα leu2-3, 112 his3-Δ1 ura3-52 trp1-289 MAL2-8(Con) MAL3 SUC3</i>)	[18]
	Biodiesels	Glycerol utilization for production of fatty acid ethyl esters (FAEEs) was done by amplification of ethanol production pathway, which is used for the transesterification in FAEEs synthesis, with overexpression of an unspecific acyltransferase from <i>Acinetobacter baylyi</i>	YPH499 (<i>MATa ura3-52 lys2-801_amber ade2-101_ochre trp1-D63 his3-D200 leu2-D1</i>)	[19]
Bulk chemicals	Bisabolene (D2 diesel fuel, bisabolane)	Bisabolene, the immediate precursor to bisabolane, was produced by (1) using the strategy for increasing pool of farnesyl diphosphate (FPP) in artemisinic acid production [20] and (2) screening and codon-optimizing bisabolene synthases (sesquiterpene synthases). The final titers were over 900 mg/l in shake flasks	BY4742 (<i>MATα his3D1 leu2D0 lys2D0 ura3D0</i>)	[21]
	1,2-propanediol	The combination effects of different copy number (from 0 to 3) of two <i>E. coli</i> genes (<i>mgs</i> and <i>gldA</i>) were studied. Although the three copy numbers of two genes showed the highest level of 1,2-propanediol, specific activity of Mgs and inhibitory relationship by GldA was considered more importantly for the production of 1,2-propanediol	NOY386 α A (<i>MATα ura3-52 lys2-801 trp1-Δ63 his3-Δ200 leu2-Δ1</i>)	[22]
	D-ribose and ribitol	The flux from glucose to pentose phosphate pathway was amplified by inactivation of both phosphogluucose isomerase and transketolase with overexpression of sugar phosphate phosphatase (DOG1). Fructose was supplied and redox balance was controlled by overexpression of NAD ⁺ -specific glutamate dehydrogenase (<i>GDH2</i>) of <i>S. cerevisiae</i> or NADPH-utilizing glyceraldehyde-3-phosphate dehydrogenase (<i>gapB</i>) of <i>Bacillus subtilis</i>	BWG1-7a (<i>MATa ade1-100 his4-519 leu2-3,112 ura3-52 GAL⁺</i>)	
Metabolic engineering	L-lactic acid	Improved production of L-lactic acid was achieved by overexpression of <i>LDH</i> gene coding L-lactic acid dehydrogenase from bovine and knocked out a <i>PDC1</i> gene coding pyruvate decarboxylase to redirect the fluxes to L-lactic acid; and overexpression of an NADH oxidase (<i>nox</i>) from <i>Streptococcus pneumoniae</i> into the cytoplasm to reduce the ratio of NADH/NAD ⁺	CEN.PK2-1D (<i>VW-1B; MATα, leu2-3/112 ura3-52 trp1-289 his3Δ1 MAL2-8c SUC2</i>)	[23]
	Polyhydroxy-alkanoates	The synthesis of diverse size of PHA polymer (C4 to C14) was investigated by cytosolic expression of mcl-PHA synthase from <i>Pseudomonas oleovorans</i> or peroxisomal expression of scl-PHA synthase from <i>Ralstonia eutropha</i>	BY4743 (<i>MATα/his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 ura3Δ0/ura3Δ0</i>)	[24]
	Pyruvic acid	Pyruvate decarboxylase-negative [Pdc(-)] strains were evolved in glucose-limited chemostat cultivation by progressively lowering the acetate content in the feed to obtain an acetate-independent Pdc (-) mutant. Maximum yield was 0.54 g of pyruvate/g glucose	CEN.PK113-7D (<i>MATa MAL2-8C, SUC2</i>)	[25]
Metabolic engineering	Succinic acid	The deletion of the genes <i>SDH1</i> , <i>SDH2</i> , <i>IDH1</i> , and <i>IDP1</i> made higher flux to succinic acid production. Maximum yield was 0.11 mol of succinic acid/mol of glucose	AH22ura3 (<i>MATa ura3Δ leu2-3 leu2-112 his4-519 can1</i>)	[26]
				[27]

Table 1 continued

Categories	Products	Specific applications	Strains	References
Fine chemicals	β -amyrin	The differences of phenotype and genotype in two yeast strains, CEN.PK113-7D and S288C, were compared. CEN.PK113-7D had more contents of ergosterol and fatty acids with non-silent SNPs in relative metabolism, <i>ERG8</i> , <i>ERG9</i> , and <i>HFA1</i> . Amplification of those genes exhibited a fivefold increase of β -amyrin	CEN.PK113-7D (<i>MATA MAL2-8C SUC2</i>)	[28]
	β -carotene	Genomic integration and overexpression of carotenogenic genes from <i>X. dendrorhous</i> (<i>crtYB</i> , <i>crtE</i> , and <i>crtI</i>) and <i>S. cerevisiae</i> (<i>BTS1</i> and truncated <i>HMG1</i>) with change of copy number achieved high levels of β -carotene, up to 5.9 mg/g dry cell weight	CEN.PK113-7D (<i>MATA MAL2-8C SUC2</i>)	[29]
	Amorpha-4, 11-diene	Amplification of mevalonate pathway in CEN.PK2 was engineered and compared to previously constructed strain S288C [20]. Artemisinic acid production was doubled, while amorpha-4, 11-diene was tenfold higher, over 40 g/l	CEN.PK2-1C (<i>MATA ura3-52 trp1-289 leu2-3,112 his3Δ1 MAL2-8C SUC2</i>) CEN.PK2-1D (<i>MATα ura3-52 trp1-289 leu2-3,112 his3Δ1 MAL2-8C SUC2</i>)	[30]
	Valencene and amorphadiene	Co-expression of heterologous enzymes, farnesyl diphosphate synthases (FDPSs), and sesquiterpene synthase (ex. <i>Citrus sinensis</i> valencene synthase <i>CsTPS1</i> , <i>Artemisia annua</i> terpene synthase, amorpha-4,11-diene synthase ADS) in mitochondria and cytosol improved the production of valencene and amorphadiene	W303-1A (<i>MATA ade2-1 trp1-1 leu2-3, 112 his3-11, 15 ura3-1</i>) mBDXe (a uracilauxotroph derivative of strain BDX, Lallemand, Rexdale, Ontario, Canada)	[31]
	Casbene (an anti-fungal diterpene)	Genes of putative Casbene synthases from different Euphorbiaceae species were isolated and applied for production of diterpenes. Maximum concentration of Casbene was 31 mg/l	BY4742 (<i>MATα his3D1 leu2D0 lys2D0 ura3D0</i>)	[32]
	Cinnamoyl anthranilates	Twenty-six different cinnamoyl anthranilates molecules were produced by co-expressing a 4-coumarate/CoA ligase (4CL, EC 6.2.1.12) from <i>Arabidopsis thaliana</i> and a hydroxycinnamoyl/benzoyl-CoA/anthranilate N-hydroxycinnamoyl/benzoyltransferase (HCBT, EC 2.3.1.144) from <i>Dianthus caryophyllus</i>	BY4742 (<i>MATα his3D1 leu2D0 lys2D0 ura3D0</i>)	[33]
	Cubebol	Overexpression of <i>GFTpsC</i> (a sesquiterpene synthase isolated from <i>Citrus paradisi</i> and encoding for a cubebol synthase) with integration of <i>tHMG1</i> into genome and reduction of <i>ERG9</i> gene expression produced cubebol up to 10 mg/l	CEN.PK113-5D (<i>MATA MAL2-8c SUC2 ura3-52</i>)	[34]
	Eicosapentaenoic acid (EPA)	Five heterologous fatty acid desaturases and an elongase were identified by a BLAST search and assayed their substrate preferences activity. Without supplement of fatty acids, EPA/ARA were produced	CEN.PK113-5D (<i>MATA MAL2-8c SUC2 ura3-52</i>)	[35]
	Farnese and geranyl geraniol	<i>ERG9</i> deletion and overexpression of two isoforms of HMGCoA reductases (<i>HMG1</i> and <i>HMG2</i>) was implemented in a host strain with overexpression of diverse FPP synthases and GGPP synthases	FL100 (<i>MATA</i> , ATCC: 28383)	[36]
	L-ascorbic acid	About 100 mg of L-ascorbic acid per liter was produced by overexpression of D-arabinono-1,4-lactose oxidase from <i>S. cerevisiae</i> and L-galactose dehydrogenase from <i>Arabidopsis thaliana</i>	GRF18U (<i>MATα his3 leu2 ura3; NRRL Y-30320</i>) W303 1B (<i>MATα ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100</i>)	[37]
	Linalool	Overexpression of <i>Clarkia breweri</i> linalool synthase gene (LIS) in wine strain T ₇₃ showed higher levels of linalool than conventional laboratory strains. Combining with deregulation of HMG-CoA reductase improved linalool yield	BQS252 (<i>MATA ura3-52</i> (derivative of FY1679))	[38]
	Methylmalonyl-coenzyme A	Polyketide precursor (Methylmalonyl-CoA) pathway was constructed by introducing propionyl-CoA carboxylase and malonyl/methylmalonyl-CoA ligase from <i>Streptomyces coelicolor</i>	InvSC1 (<i>MATA, his3delta1, leu2, trp1-289, ura3-52</i> (Invitrogen, Carlsbad, CA, USA)) BJ5464 (<i>MATα, ura3-52, trp1, leu2-delta1, his3-delta200, pep4::HIS3, prb1-delta1.6R, can1, GAL</i> .)	[39]

Table 1 continued

Categories	Products	Specific applications	Strains	References
	Patchoulol	A physical fusion between native (farnesyl diphosphate synthase) and heterologous enzymes (patchoulol synthase of plant origin, <i>Pogostemon cablin</i>) was successfully applied to produce patchoulol, 25 mg/l	CEN.LA100 (<i>MATA/MATα</i> <i>ERG20/erg20::hph</i> <i>MAL2-8c/</i> <i>MAL2-8c SUC2/SUC2</i> <i>ura3-52/</i> <i>ura3-52</i>)	[40]
	Resveratrol	Co-expression of the coenzyme-A ligase-encoding gene (<i>4CL216</i>) from a hybrid poplar and the grapevine resveratrol synthase gene (<i>vst1</i>) from <i>Vitis vinifera</i> with supplement of <i>p</i> -coumaric acid produced resveratrol, 1.45 mg/L	FY23 (<i>MATA</i> <i>ura3-52 trp1A63</i> <i>leu2A1</i>)	[41]
	Vanillin	Knock-out targets, <i>PDC1</i> and <i>GDH1</i> , suggested by in silico metabolic model was applied and production of vanillin was improved up to fivefold	X2180-1A (<i>MATA his3D1 leu2D0</i> <i>met15D0 ura3D0 adh6::LEU2</i> <i>bg11::KanMX4 PTP11::3DSD</i> <i>[AurC]::HsOMT</i> <i>[NatMX]::ACAR [HphMX]</i>)	[42]
	Se-methylselenocysteine	Combination of metabolic (codon optimization of heterologous selenocysteine methyltransferase) and bioprocess (tuning carbon-and sulfate-limited fed-batch) engineering achieved 24-fold increase in Se-methylselenocysteine production	CEN.PK113-7D (<i>MATA MAL2-8C</i> <i>SUC2</i>)	[43]
	Non-ribosomal peptides	Separated non-ribosomal peptide synthetase modules with compatible communication-mediated domains showed functional interaction, which meant that new module combinations could produce novel non-ribosomal peptides	CEN.PK113-11C (<i>MATA MAL2-8C</i> <i>SUC2 ura3-52 his3-D1</i>)	[44]
Protein drugs	Insulin-like growth factor 1 (fhlGF-1)	Inactivation of <i>GAS1</i> increased the yield of human insulin-like growth factor1, from 8 to 55 mg/l	GcP3 (<i>MAT a pep4-3 prb1-1122</i> <i>ura3-52 leu2 gal2 cir^o</i>)	[45]
	Glucagon	Disruption of <i>YPS1</i> encoded aspartic protease increased glucagon, 17.5 mg/l	SY107 (<i>MATα YPS1 Δtpi::LEU2</i> <i>pep4-3 leu2 Ara3 cir⁺</i>)	[46]
	Single-chain antibodies (scFv)	Production of an anti-transferrin receptor single-chain antibody (OX26 scFv) was optimized by adjusting expression temperature and gene dosage and final yield was 0.5 mg/l	BJ5464 (<i>MATA ura3-52 trp1</i> <i>leu2D1 his3D200 pep40HIS3</i> <i>prb1D1.6R can1 GAL</i>)	[47]
	Hepatitis surface antigen (HBsAg)	Glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter of <i>Pichia pastoris</i> was used for HBsAg production and final yield was 19.4 mg/l	INVSc1 (<i>MATA his3D1 leu2</i> <i>trp1-289 ura3-52</i>)	[48]
	Parvovirus B19 VP2	The major-capsid protein VP2 of Parvovirus B19 produced in <i>S. cerevisiae</i> showed similar properties to native virus or produced by baculovirus system in size, molecular weight, and antigenicity. The yield was 400 mg/l	HT393 (<i>MATA leu2-3 leu2-112</i> <i>ura3Δ5 prb1-1 prc1-1</i> <i>pra1-1 pre1-1</i>)	[49]
	Epidermal growth factor (EGF)	<i>O</i> -glycosylation pathway was constructed by introduction of GFR (GDP-fucose transporter), POFUT1 (<i>O</i> -fucosyltransferase 1), <i>manic fringe</i> gene (β 1,3-N-acetylglucosaminyltransferase) from human and <i>MUR1</i> (GDP-mannose-4,6-dehydratase), <i>AtFX/GER1</i> (GDP-4-keto-6-deoxy-mannose-3,5-epimerase/4-reductase) from <i>Arabidopsis thaliana</i> producing <i>O</i> -glycosylated EGF protein	W303-1A (<i>MATA leu2-3,112 his3-11,15 ade2-1 ura3-1 trp1-1</i> <i>can1-100</i>)	[50]
	Immunoglobulin G	Leader peptides for the enhanced secretion of proteins constructed by directed evolution allowed for a 180-fold increase in secretion of full-length, functional, glycosylated human IgG	W303-1B (<i>MATα leu2-3,112</i> <i>his3-11,15 ade2-1 ura3-1 trp1-1</i> <i>can1-100</i>)	[50]
	Hepatitis B virus surface antigen (HBsAg)	The yield of S domain of hepatitis B virus surface antigen (sfHBsAg) was increased by co-expression of disulfide isomerase (<i>PDII</i>) with adjusting fermentation mode	W303-1A (<i>MATα ura3-52 leu2~1</i> <i>his3~200 pep4::HIS3</i> <i>prb1~1.6Rcan1 GAL</i>)	[51]
	L1 protein of human papillomavirus (HPV) type16	Optimization of the secondary structure of HPV16 L1 mRNA increased the expression level of that protein up to fourfold than of wild-type	<i>S. cerevisiae</i> 2805 (<i>MATα</i> <i>pep4::HIS3 prb-Δ1.6 his3</i> <i>ura3-52 gal2 can1</i>)	[52]
			<i>S. cerevisiae</i> 2805 (<i>MATα</i> <i>pep4::HIS3 prb-Δ1.6 his3</i> <i>ura3-52 gal2 can1</i>)	[53]

structures such as polymers (cellulose, starch, xylan), dimers (cellobiose, melibiose, lactose) and monomers (glucose, fructose, galactose, arabinose, xylose). Except for

the hexoses (glucose, fructose, galactose) and a few dimers (sucrose and maltose), most of these compounds are not naturally metabolized by *S. cerevisiae*. Even among the

Table 2 Overall analysis of patents

Search words or phrases in title and abstract	<i>S. cerevisiae</i>	Yeast	<i>E. coli</i>	Fermentation
Strains or fermentation	4,630	13,769	14,914	9,065
Production ^a	3,080	7,812	9,306	5,211
Pharmaceuticals ^a	985	2,686	3,343	956
Food ingredients ^a	24	41	21	60
Chemical ^a	1,702	4,830	5,683	2,651
Production and chemical ^a	1,342	3,367	4,003	1,899
Production and fine chemical ^a	58	64	80	19
Fuel ^a	66	145	110	369
Production and fuel ^a	59	126	94	321

US Granted 5,266,192 patents searched (July 31, 2011)

^a Search term: Strains or fermentation and a keyword

hexoses there are wide differences in terms of uptake, e.g., the uptake rate of galactose is much lower than for the other hexoses. Therefore, the extension of substrate range in *S. cerevisiae* has been a major priority in connection with the use of yeast for biofuel and biochemical production, and this has recently been covered by several reviews [14, 61–63]. An overview of alternative substrates is provided in Fig. 3. Recently, since tools from systems biology have been available, they have been applied to identify new target genes and understand metabolism at the whole-cell level. Comparative genomics among xylose-fermenting fungi were used to identify new pathways or genes for increasing xylose utilization [73]. The capacity of xylose utilization by those genes was demonstrated by engineering *S. cerevisiae*, and genome-scale modeling was implemented to assess that global flux analysis could predict the effect of co-factor balancing for pentose utilization [74]. In the following section, we review recent work improving galactose utilization and co-fermentation of cellobiose/xylose.

The modification targets to improve galactose utilization have been well elucidated in *S. cerevisiae*, which include engineering of the regulatory network (inactivation of repressors and up-regulation of activator) and over-expression of the final enzyme, phosphoglucomutase (*PGM2*) in the Leloir pathway responsible for galactose catabolism [75–77]. All these targets were directly associated with the galactose metabolic pathway, but using a cDNA library, another target that is not part of galactose metabolism was also found [71]. In this study, three beneficial over-expression targets, *SEC3*, *tTUP1*, and *SNR84* were identified. Although two of them were confirmative with previous works due to the function of those genes, *SEC3* (phosphomannomutase having activity as phosphoglucomutase) and truncated *TUP1* (repressor); the last target was unpredicted. *SNR84* codes for H/ACA box small nucleolar RNA, and higher activity of phosphoglucomutase

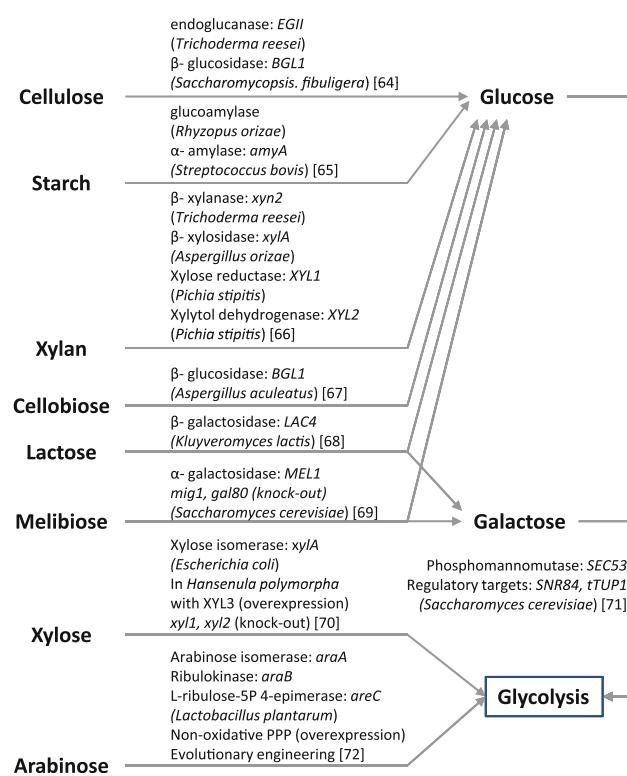


Fig. 3 Illustration of relevant substrates that have been considered for yeast fermentation. Heterologous enzymes that are currently applied are summarized for non-utilizable carbon sources in *S. cerevisiae* such as polymers (cellulose, starch, xylan), disaccharide (cellobiose, lactose, melibiose), pentose sugar (xylose, arabinose). In case of galactose, which is utilized slowly compared to glucose, over-expression targets of innate enzymes for improving galactose availability were screened

in the transformant over-expressing *SNR84* proposed a relationship between this gene and galactose metabolism. Recently, evolutionary engineering was also implemented to find unforeseen targets using systems biology [78]. The combination of different systems biology techniques

enabled linking phenotype and genotype. Also, to identify true-positive targets, three different evolved clones on galactose were compared, and all evolved mutants showed higher transcripts and metabolite level in storage carbohydrate metabolism together with up-regulation of phosphoglucomutase, whereas there were no mutations in any of the GAL-genes and *PGM2* including 1kb up- and downstream. However, based on analysis of all three mutants, the Ras2/PKA signaling pathway was strongly suggested to induce the observed phenotypic changes because this signaling pathway has mutations commonly in all three evolved mutants. In another study, co-fermentation of cellobiose and xylose was implemented by co-expression of cellobextrin transporter (*cdt-1*), β -glucosidase (*gh1-1*), and xylose enzymes (*XYL1*, *mXYL1*, *XYL2*, and *XKS1*) [79]. *mXYL1* is a mutant enzyme that exhibits much higher preference for NADH as co-factor unlike the wild-type *XYL1* that prefers the use of NADPH as co-factor. Introduction of these genes minimized glucose repression of xylose fermentation, since glucose was generated after transporting of cellobiose by the cellobextrin transporter and degraded by β -glucosidase inside the cell. These modifications allowed *S. cerevisiae* to co-ferment cellobiose and xylose with improved ethanol yield.

There are also indications of industrial application of pentose metabolism by yeast, in particular the utilization of the pentose sugars xylose and arabinose, as several patents cover the metabolism of these sugars [80–87]. There are two Dutch companies that have patents on the use of these two carbon sources, namely DSM and Royal Nedalco. Both focus on isomerase-based pathways to avoid the co-factor balancing problem in xylose and arabinose utilization. DSM over-expressed xylose isomerase from *Thermotoga maritime* MSB8 and arabinose genes (*araA*, *araB*, and *araC*) from *Lactobacillus plantarum* with codon optimization and constitutive expression of the genes in the non-oxidative pentose phosphate pathway [82, 83, 86]. Royal Nedalco employed xylose isomerase of *Piromyces* sp. E2 (ATCC 76762) and arabinose genes of *Arthrobacter aurescens*, *Clavibacter michiganensis*, or *Gramella forsetii* [80, 81, 85]. The main focus of both approaches was application of higher activity of isomerases from heterologous sources (Fig. 4).

Bio-butanol production

Butanol has gained much attention as a potential biofuel to replace ethanol, currently by far the dominating biofuel. Butanol has a number of advantages as a biofuel compared to ethanol. It has a higher energy density than ethanol and only around 4% less than that of gasoline [88–92]. Furthermore, butanol blends better with gasoline

than with ethanol and it is non-corrosive, which allows it to be used in the existing petrochemical infrastructure. There are four different types of butanol (Table 3): *n*-butanol, *sec*-butanol, isobutanol, and *tert*-butanol. *n*-butanol and *sec*-butanol have a stretched carbon chain and a hydroxyl group at position 1 or 2. Isobutanol and *tert*-butanol have two or three branched carbon chains, respectively. The branched structure results in a higher octane number [93] (Table 3), which means a higher anti-knock property. Although *tert*-butanol has a higher octane number than isobutanol, its much higher melting temperature (25°C) than that of isobutanol (-101.9°C) prohibits its use as a pure fuel source. Isobutanol is therefore preferred over the other butanols.

Butanol is naturally produced by different species of *Clostridia*, but most of these processes are limited by relatively low yields and titers, and therefore much interest has been expressed in the development of novel cell factories that can be used for bio-based production of butanol [89, 94, 95]. Mainly, two strategies for producing bio-butanol have been used, (1) use of a host that has an innate butanol pathway and improving its yield and productivity, and (2) re-construction of an efficient butanol pathway in strains that are already widely used for industrial production of other fuels and chemicals. Yeast is one of the hosts that produce butanol naturally through the so-called Ehrlich pathway for fusel alcohol production [96, 97]. Also, yeast is widely used for industrial ethanol production due to its high ethanol tolerance and its robustness towards harsh industrial conditions, e.g., high osmotic stress and low pH [14, 98]. In academic papers, however, most results presented have been based on work using *Clostridia* or *E. coli* [99, 100]. So far, only two papers have been published on the use of yeast for bio-butanol production, one describing *n*-butanol production based on reconstruction of a pathway from *Clostridia* and the other describing isobutanol production based on engineering of yeast's natural pathway, and both presented very low yields and productivities [18, 101]. There are several recent reviews that summarized these academic studies [88, 94], and from comparisons of results on different organisms yeast do not seem to be an attractive host for bio-butanol production because of its low yield and productivities, especially compared to metabolically engineered *E. coli*. However, in contrast to academic research, most of the companies announcing work towards commercial production of bio-butanol are using yeast as a production organism [90, 91, 102–104]. We will therefore here review the main strategies employed in industry based on analysis of information provided in patents and patent applications. Based on patent applications and issued patents, the three dominating companies for producing bio-butanol by yeast are presently Gevo, Butamax Advanced Biofuels, and Butalco.

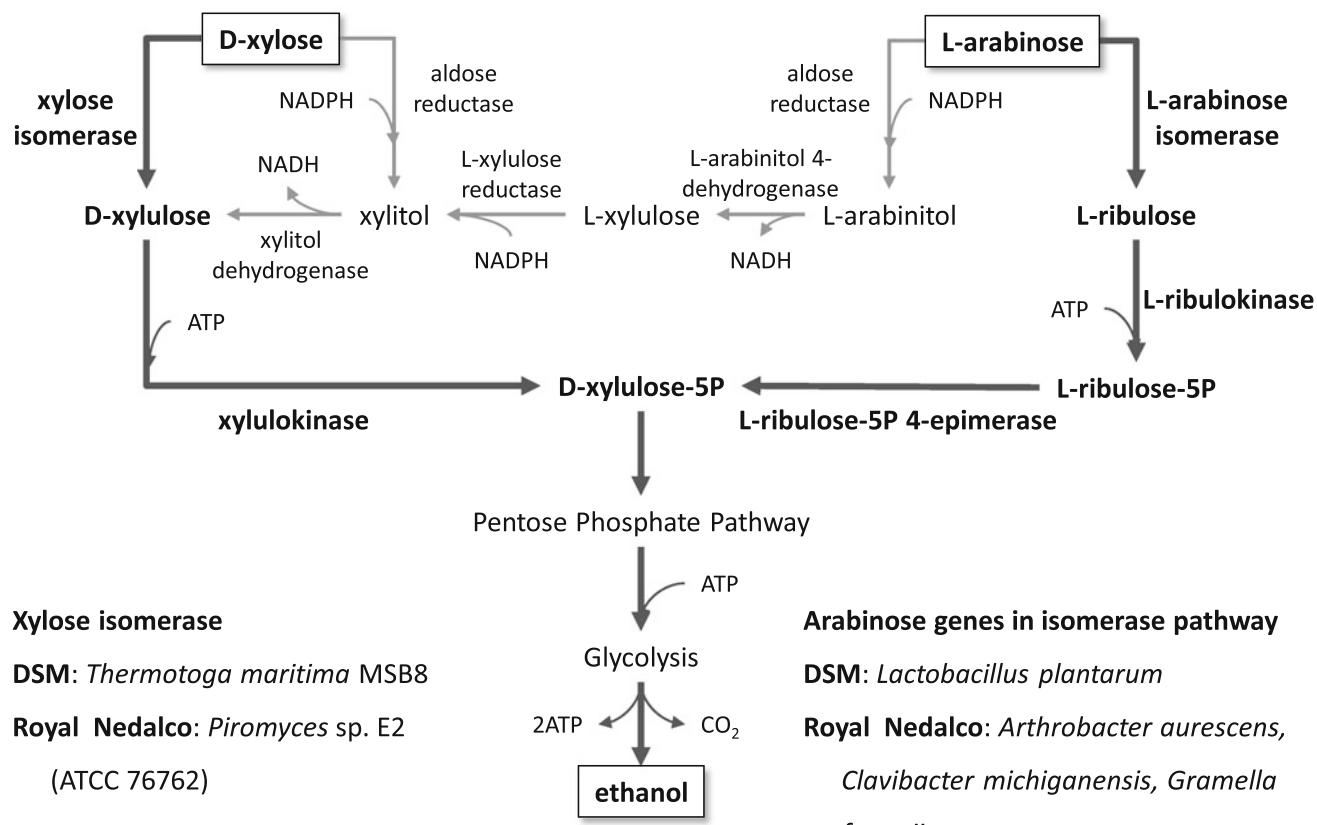


Fig. 4 Overview of pathways for pentose utilization covered by patent applications of DSM and Royal Nedalco. D-xylose and L-arabinose can be utilized by two pathways: (1) aldose reductase NADPH-dependent and (2) isomerase cofactors-independent. In case of the latter, requirement of cofactor balance is eliminated and enhancing activity of isomerase remains main issue. Two Dutch companies, DSM and Royal Nedalco, claim over-expression of heterologous xylose isomerases from *Thermotoga maritime* MSB8

and *Piromyces* sp. E2 (ATCC 76762), respectively. Xylulokinase and enzymes in pentose phosphate pathway were also amplified simultaneously. Arabinose genes in isomerase pathway such as L-arabinose isomerase, L-ribulokinase, and L-ribulose-5P 4-epimerase originated from *Lactobacillus plantarum* in DSM and *Arthrobacter aurescens*, *Clavibacter michiganensis*, *Gramella forsetii* in Royal Nedalco were amplified with enzymes in pentose phosphate pathways

Table 3 Comparison of butanol isomers [49]

	<i>n</i> -butanol	<i>sec</i> -butanol	Isobutanol	<i>tert</i> -butanol
Research octane number (RON)	96	101	113	105
Motor octane number (MON)	78	32	94	89
Melting temperature (°C)	-89.5	-114.7	-108	25.7
Boiling temperature (°C)	117.7	99.5	108	82.4
Enthalpy of vaporization at T _{boil} (kJ/kg)	582	551	566	527

Gevo is one of the pioneers in bio-butanol production, and they received a combined grant from the US Department of Agriculture and the Department of Energy for developing a yeast fermentation system to produce iso-

butanol from cellulosic-derived sugars. Gevo's patents describe different strategies to produce *n*-butanol and iso-butanol. First, to produce *n*-butanol, they propose to increase cytosolic acetyl-CoA pool and incorporate the

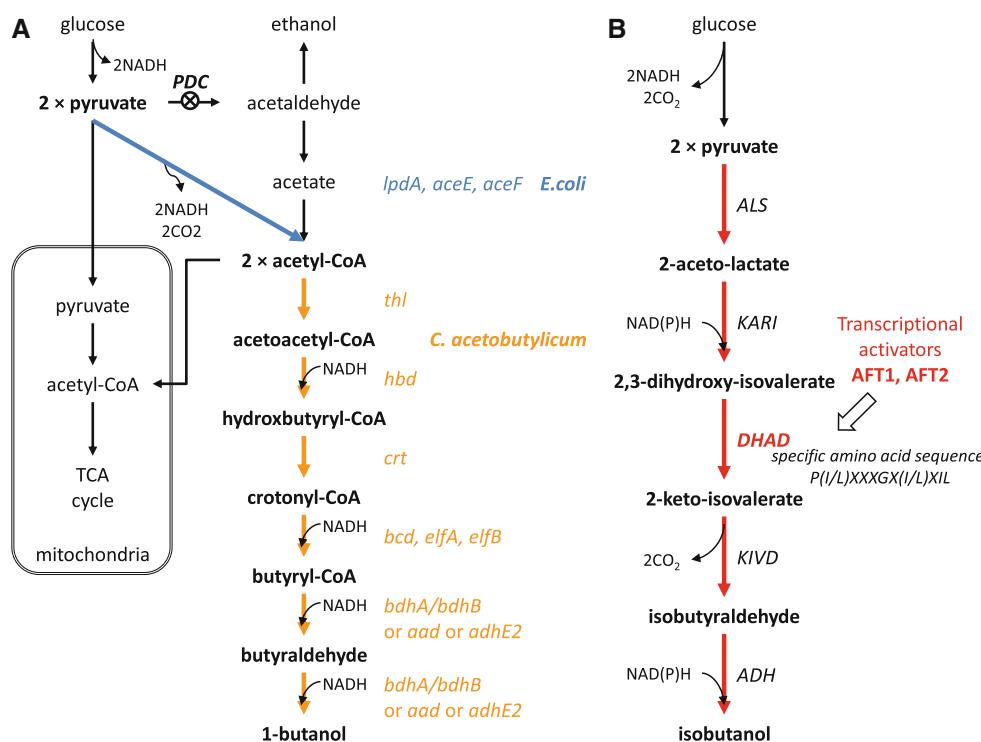


Fig. 5 Illustration of Gevo's strategies for *n*-butanol and isobutanol production in the cytosol [90, 91, 105]. **a** *n*-butanol production was attempted by amplification of heterologous genes such as the pyruvate dehydrogenase multienzyme complex (*lpdA, aceE, aceF*) from *E. coli* for increasing the cytosolic acetyl-CoA pool, and the genes in butanol synthetic pathway from *Clostridia* species. Moreover, the activity of pyruvate decarboxylase (PDC) was reduced. **b** Isobutanol was

produced in the cytosol to avoid cofactor balancing in the mitochondria; all the genes in isobutanol pathway were over-expressed in cytosol. Especially, dihydroxyacid dehydratases (DHAD) from *Lactococcus lactis* and *Neurospora crassa* were used, which had specific amino sequence, P(I/L)XXXGX(I/L)XIL. Also, the transcriptional activators *AFT1/AFT2* were over-expressed to increase DHAD activity

butanol synthetic pathway from *Clostridia* species into yeast [91]. To increase the cytosolic acetyl-CoA pool, the pyruvate dehydrogenase multienzyme complex (*lpdA, aceE, aceF*) from *E. coli* was expressed to establish a direct pathway from pyruvate to acetyl-CoA not passing via acetaldehyde that can be converted to ethanol. In addition, the activity of pyruvate decarboxylase (PDC), which converts pyruvate to acetaldehyde, is reduced (Fig. 5a). Second, isobutanol production has been tried intensively using different strategies [105, 106]. As mentioned previously, yeast can naturally produce isobutanol, and this pathway shares valine synthesis from pyruvate to 2-keto-isovalerate in the mitochondria. 2-keto-isovalerate can be exported from the mitochondria to the cytosol where decarboxylation by pyruvate decarboxylase (PDC) and dehydrogenation by alcohol dehydrogenases (ADH) can lead to production of isobutanol [18]. Gevo worked on reconstructing the isobutanol pathway in the mitochondria and the cytosol separately using innate yeast or heterologous genes. Using a mitochondrial targeting sequence, all enzymes involved in isobutanol synthesis were localized in the mitochondria and isobutanol was produced there (Fig. 6a). In this strategy, cofactor balancing was

considered important, since NADH is produced in the glycolysis and NADPH is needed for isobutanol production. First, NADPH-dependent enzymes were engineered to an NADH-dependent form, and then NADH was supplied by using a NADH shuttle concept [107]. Acetaldehyde and ethanol produced by fermentation can freely transport across membranes, and alcohol dehydrogenase in the mitochondria (encoded by *ADH3*) can provide one NADH by conversion of ethanol to acetaldehyde. Moreover, isobutyraldehyde, an intermediate of the isobutanol pathway, can be transported to the cytosol where it can be converted to isobutanol under consumption of NADH generated by glycolysis in the cytosol (Fig. 6b). To avoid cofactor balancing in the mitochondria, the whole isobutanol pathway was in another strategy expressed in the cytosol [90, 105, 108]. In this case, modification or amplification of dihydroxyacid dehydratase (DHAD) was emphasized, and obtained from *Lactococcus lactis* and *Neurospora crassa*. Additionally, the transcriptional activators *AFT1/AFT2* are over-expressed to increase DHAD activity (Fig. 5b). Gevo is currently developing 18 million gallons per year (MGPY) plant in the USA and have plans to develop 350 MPGY of new capacity by 2015.

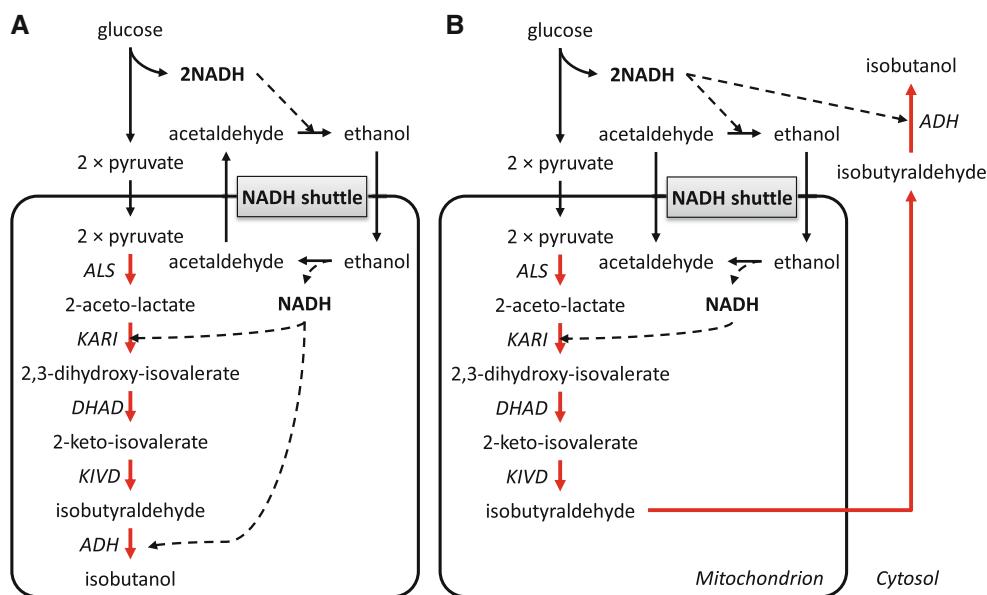


Fig. 6 Gevo's isobutanol production strategy in the mitochondria [106]. **a** All enzymes involved in isobutanol synthesis were localized in the mitochondria; KIVD and ADH being in cytosol were expressed in mitochondria with signal sequence. NADPH-dependent enzymes were engineered to an NADH-dependent form, and then NADH was supplied by using an NADH shuttle concept. Acetaldehyde and ethanol produced by fermentation were transported across

membranes, and alcohol dehydrogenase in the mitochondria (encoded by *ADH3*) provided NADH by conversion of ethanol to acetaldehyde. **b** Isobutyraldehyde was transferred to the cytosol from mitochondria where it is converted to isobutanol under consumption of NADH generated by glycolysis in the cytosol to make more precise cofactor balance

Butamax Advanced Biofuels was formed in 2009 as a joint venture between Dupont and BP, two large companies in the chemical and energy industry, respectively. Before establishing Butamax, Dupont investigated many different approaches in strain development for *n*-butanol, *sec*-butanol, and isobutanol production [109–111]. Consequently, one of their foundational patents for isobutanol production was officially granted in August 2011 by the US Patent and Trademark Office [102]. Dupont's (and now Butamax's) strategy is to find and introduce many different heterologous genes related to butanol biosynthesis and introduce these into yeast. For isobutanol production they considered four different pathways, which include conversion of valine to isobutyrylamine and butyryl-CoA to isobutyryl-CoA (Fig. 7). A total of 11 enzyme reactions were considered and each reaction could be catalyzed by at least three to four heterologous enzymes. For example, acetolactate synthase (ALS), which catalyzes the first step, could be *Klebsiella pneumoniae budB*, *Bacillus subtilis alsS*, or *Lactococcus lactis als*. Key findings claimed were the use of *Pseudomonas ilvC* (ketol acid reductoisomerase: KARI) and a method to amplify *ilvD* (DHAD), which contains a Fe–S cluster. Isobutanol production in mitochondria was also considered [109, 112, 113]. Here, the genes involved in substrate competing reactions, *BAT1*, *ILV1*, and *LEU4* were deleted and the activity of the E1 alpha subunit of the pyruvate dehydrogenase (PDH) complex (*PDA1*), which converts

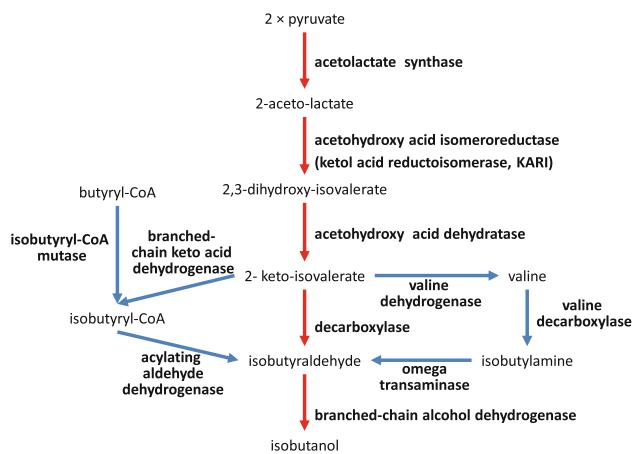


Fig. 7 Butamax's strategies for isobutanol production in cytosol [102]. Four different pathways for isobutanol production were suggested: (1) pyruvate to isobutanol directly (red arrows), (2) pyruvate through valine bypass (blue arrows), (3) pyruvate through isobutyryl-CoA bypass (blue arrows), and (4) butyryl-CoA to isobutanol (blue arrows). A total of 11 enzyme reactions were considered and at least three to four heterologous enzymes in each step were claimed in patents of Butamax

pyruvate to acetyl-CoA, was reduced by promoter exchange. Furthermore, NADH kinase (*POS5*) was over-expressed to ensure sufficient supply of NADPH required by the KARI enzyme (Fig. 8). Butamax also investigated the butanol tolerance by modification of the regulatory network

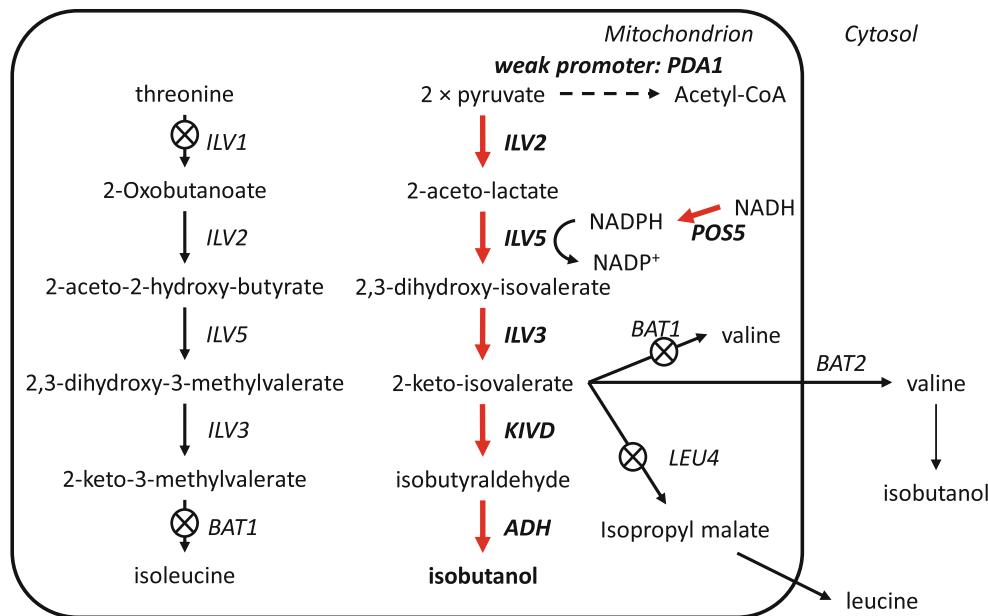


Fig. 8 Butamax's isobutanol production strategies in the mitochondria [103]. To block substrate-competing reactions *BAT1*, *ILV1*, and *LEU4* were deleted and the activity of the E1 alpha subunit of the pyruvate dehydrogenase (*PDH*) complex (*PDA1*) was reduced by

promoter exchange to a weak one. NADH kinase (*POSS*) was over-expressed to ensure sufficient supply of NADPH required by the KARI enzyme. Red arrows mean over-expression of genes

Table 4 Targets for increasing butanol tolerance in yeast (Butamax)

Targeting	Modified genes	Butanol tolerance [growth yield improvement in butanol % (w/v)]	References
Multidrug resistance ATP-binding cassette transporter	Pdr5p, CDR1, BFR1	~ 1.8-fold in 0.75%	[114]
Cell wall integrity pathway	SLT2p	~ 25% in 1%	[115]
Osmolality/glycerol response pathway	PBS2p	~ 40% in 1%	[116]
Filamentous growth response pathway	MSS11p	~ 2-fold in 1.5%	[117]
Amino acid starvation	Gcn1p, Gcn2p, Gcn3p, Gcn4p, Gcn5p, Gcn20p	~ 1.8-fold in 2.0%	[118]

[114–118]. Yeast, especially *S. cerevisiae*, is known to have higher butanol tolerance than other microorganisms and it can grow in butanol concentrations higher than 20 g/l [119], but still several targets were identified for improving butanol tolerance and these are summarized in Table 4. The first operational plant for commercial production of isobutanol by Butamax is scheduled to be operational by 2013.

The third company working actively on bio-butanol production by yeast is Butalco, which is a biofuel company that also develops ethanol producing yeast that can use xylose and arabinose as carbon sources. Isobutanol production strategies by this company are not to use heterologous genes; but rely solely on endogenous yeast genes [104]. As mentioned above, yeast has all enzymes necessary for isobutanol production; and Butalco is employing its core technology for genetic optimization of

yeast. Four points are mainly considered in their strategy: (1) different intracellular location of the isobutanol synthetic enzyme, (2) the weak activity of the enzymes, (3) cofactor imbalance, and (4) the formation of secondary products. All isobutanol synthetic genes were suggested to be express in the cytosol using strong promoters and the activity of pyruvate decarboxylase (PDC), which catalyzes pyruvate to ethanol reactions, were removed or reduced. To ensure cofactor balance, the acetoxyhydroxy acid reducto-isomerase (KARI) and the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were proposed to ensure balancing of the co-factors. Either NADH-dependent yeast KARI enzyme (*ILV5^{NADH}*) could use NADH from the glycolysis or NADP+-dependent yeast GAPDH (*GLD5*) could produce NADPH required by NADPH-dependent KARI enzyme (*ILV5*).

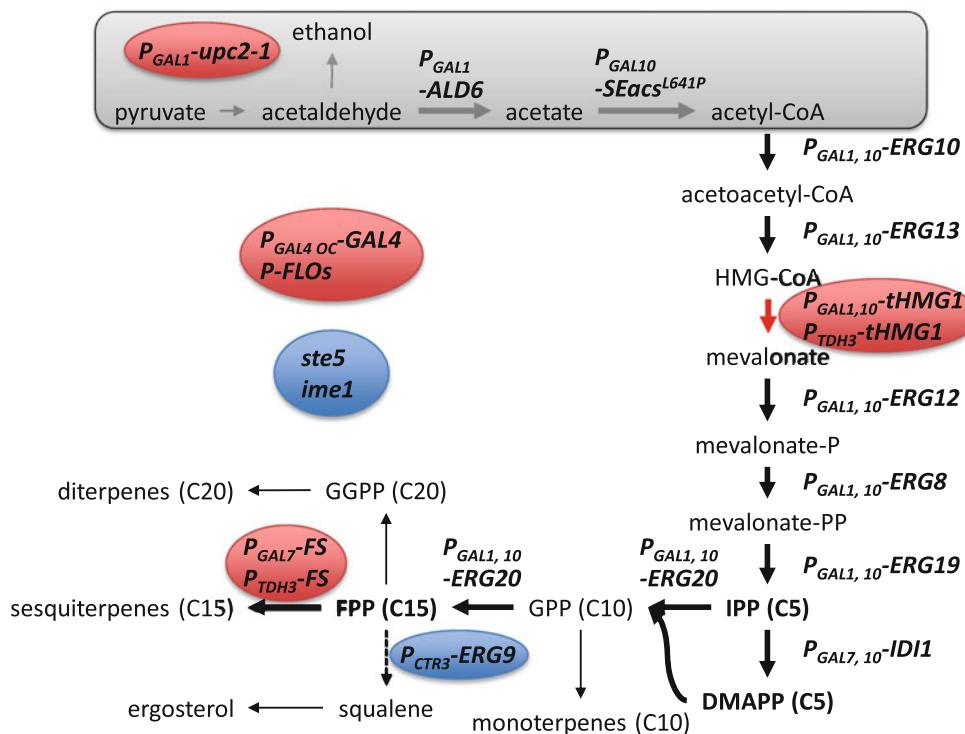
Isoprenoids production

Isoprenoids are a chemically diverse group of natural compounds that have many different biological functions, and they have found applications as medicines, perfumes, food additives, and fine chemical intermediates [120]. Recently, the possibility that they can be applied as fuels has been discussed [121]. Although the standard definition of isoprenoids has not been clearly set, the compounds that have isopentenyl diphosphate (IPP) or dimethylallyl diphosphate (DMAPP) as a unit molecule are typically categorized into isoprenoids. Those unit molecules have five carbons and the combination of them can generate bigger unit molecules such as geranyl diphosphate (GPP, ten carbons), farnesyl diphosphate (FPP, 15 carbons), and geranylgeranyl diphosphate (GGPP, 20 carbons), which are precursors for monoterpenes, sesquiterpenes and diterpenes, or carotenoids, respectively [122].

The production of isoprenoids has been limited by insignificant quantity in their natural sources, mostly plants and low yield of extraction. Moreover, the diversity of isoprenoids is enormous, which means that new bioactive compounds can be produced by finding new enzymes and they can even be used in combination to make chimeric pathways resulting in novel compounds [123]. For these reasons, there has been much interest in developing microbial-based production of isoprenoids by re-creating a plant-like pathway in yeast with genetic modifications of leader sequence and codon-optimization for re-locating and overexpression of relative genes. Two biosynthetic pathways for IPP and DMAPP are known—the mevalonate-dependent (MVA) pathway that converts acetyl-CoA to IPP and the deoxyxylulose-5-phosphate (DXP) pathway that converts glyceraldehyde-3-phosphate and pyruvate to IPP/DMAPP. The MVA pathway is present in the cytosol and mitochondria of plants and yeast, while the DXP pathway is found in bacteria. In higher plants, both pathways are present; with the DXP pathway being present in chloroplast, probably of bacterial origin [120]. Extensive endeavors have focused on (1) optimization of these pathways for increasing the metabolic flux of isoprenoids and (2) discovery of heterologous enzymes for production of different isoprenoid-based products. As a result, there have been remarkable achievements; especially, elucidation of key targets for metabolic engineering of the MVA pathway in yeast such as over-expression of truncated *HMG1* (*tHMG1*), *upc2-1*, and *ERG20* combined with repression of *ERG9* [20] (Fig. 9). In addition to optimization of the MVA pathway, a strategy for the increase of the acetyl-CoA pool has also been implemented by over-expression of *ALD6* of *S. cerevisiae* and a constantly active acetyl-CoA synthase mutant from *Salmonella enterica* [124]. In order to overcome regulation of the MVA

pathway, attempts to reconstruct the DXP pathway in yeast have also been evaluated [125, 126]. Much of the work on metabolic engineering of yeast for isoprenoid production has been reviewed recently [122, 123, 127, 128], and we will therefore here focus on industrial applications using analysis of the patent literature.

Isoprenoid and yeast were used as keywords in the Delphion™ Web site that analyzes and clusters relative patents. A list of companies and their patents were generated. Most of patents that are related to the flux increase of isoprenoid precursors (FPP) are assigned to University of California and Amyris, and is mainly derived from work of Jay Keasling's research group. Many of the strategies described in the patents are very similar to published scientific articles, but there are a few additional important strategies described for improving isoprenoid production at an industrial scale [129, 130] (Fig. 9). Firstly, the production host was selected among industrial strains, resulting in selection of *Saccharomyces cerevisiae* PE-2, which has been used in the Brazilian fuel ethanol industry since 1994. This strain was selected because of its higher tolerance to industrial fermentation conditions and its ability to tolerate yeast recycling typically used in Brazilian ethanol production plants. The range of tolerance includes high ethanol concentration, high cell density, high temperature, osmotic stress, low pH and sulfite, and bacterial contamination [131]. Secondly, all genetic modifications were done at the chromosomal DNA level and not using plasmids as is often done in academic research groups [20, 132]. In order to ensure strong expression of the genes, the promoters of all the genes in the MVA pathway were changed. In the case of *tHMG1*, which should be higher expressed than HMG-CoA synthase to avoid accumulation of HMG-CoA [133], one more copy with another strong promoter (*P_{TDH3}*) was integrated into the chromosome. These changes are comparable to previously published papers, which used over-expression of *upc2-1* to increase the expression level of genes in the MVA pathway. Over-expression of *GAL4* and knock-out of *GAL80* were constructed to induce expression of the genes under *GAL1*, 7, and 10 promoters. Especially, the modified *GAL4* promoter (*P_{GAL4 OC}*) that has no *MIG1* binding site was used for induction of *GAL4* expression in fermentation media containing glucose [134]. Repression of *ERG9* (squalene synthase) was done by replacing the native promoter to the *P_{CTR3}*, which is controlled by copper. Formerly, this repression was performed by using the methionine repressible promoter *P_{MET3}*, a strategy that has been used by different research groups in academia [20, 34, 135]. Ergosterol is the end-product from IPP/DMAPP condensation in yeast (Fig. 9). Reducing the activity of Erg5 theoretically enables an increase in flux to other terpenes while cellular growth can be maintained. Thirdly,



Saccharomyces cerevisiae PE-2 (Brazilian fuel ethanol industry since 1994)

Fig. 9 Overview of Amyris metabolic engineering strategies. Industrial strain *Saccharomyces cerevisiae* PE-2 was used as a production host because of its higher tolerance to the industrial environment [129]. All promoters of mevalonate genes were exchanged to strong one in chromosome. Gray box means the strategies that were used in a

scientific article [124] but not in the patent. Red color circles mean even higher expression than other overexpressed genes. Blue color circles mean knock-out of genes or reduction of expression level. Thick arrows mean amplified steps based on plasmids in a scientific article [20]. The dotted arrow indicates reduction of flux

additional modifications were performed in other pathways besides the MVA pathway. To facilitate the purification process, flocculation proteins (*FLOs*) were over-expressed. Furthermore, sporulation (*IME1*) and endogenous mating were impaired by disruption of responsible genes. The pheromone response genes (*STEs*) were also functionally inactivated in order to prevent mating. These three strategies were simultaneously implemented in one patent application [129]. Other strategies are also described in the patent literature, e.g., to increase C1 metabolism serine, which is a precursor of this metabolism, were supplied to the media [136] and GTR reductase and ALA synthase were over-expressed to increase the heme pool [137]. Besides genetic modification, Amyris also developed a monitoring method that can concurrently quantify cofactors, energy molecules, and intermediates in the MVA pathway by using LC/MS/MS [138]. This technique allows observation of the metabolite concentration change in the overall pathway during the fermentation process, and this was found very useful in finding steps that needed further modification.

Other companies that have patents on increasing isoprenoid production are Arkion Life Science and Allylix,

and these patents mostly focus on the development of highly active HMGR (HMG-CoA reductase, *HMG1*, *HMG2*) and reduction of *ERG9*, which have already been recognized as critical targets [139, 140]. HMGR is the main enzyme controlling the flux through the MVA pathway. *ERG9* encodes squalene synthase that converts FPP to squalene, and by attenuating the *ERG9* expression, the FPP pool can be increased and flux can be directed towards sesquiterpenes. However, yeast cannot grow without *ERG9* at aerobic condition, even if ergosterol is fed to the medium, and generation of *erg9* mutants that can grow aerobically with ergosterol in the medium were developed by these companies. Dupont also holds many patents related to isoprenoids production, but most of these are related to the identification of novel enzymes for conversion of FPP to different isoprenoids [141–144].

Conclusions

Based on our review of ongoing metabolic engineering projects in industry and academy, it is clear that there are several high-profile projects ongoing on using *S. cerevisiae*

for the production of novel fuels and chemicals. The vast knowledge on this organism combined with the robustness of this organism to harsh industrial conditions makes it a preferred organism for production of many fermentation-derived products. Compared to *E. coli*, which is widely used in academia, yeast cannot be contaminated by phages. It is very osmo-tolerant and can hence tolerate very high sugar concentrations, and it tolerates a lower pH than most bacteria. These advantages typically outweigh the fact that compared to *E. coli* it is more difficult to engineer yeast and that the rates of conversion typically are lower in yeast. Despite the extensive synthetic and systems biological toolbox available for yeast, it is necessary to further advance technologies that can be used in metabolic engineering. Still strain development is time-consuming and often serial in nature, meaning that development time cannot be reduced through increased resources for a short time. Among the novel technologies that would allow for faster development are tools for deletion of several genes in one round of transformation, a larger promoter library, in particular better characterization of promoters and their activity at different industrially relevant fermentation conditions, inducible promoters, and tools for rapid identification of how flux through different metabolic pathways are controlled [4, 13]. With the development of such tools combined with experience from development of yeast-based fermentation processes for novel biofuels like isobutanol and farnesene we are, however, confident that the way will be paved for faster development of other processes for sustainable production of fuels and chemicals to the benefit of society. This may lead to the establishment of yeast platform strains that can use a range of relevant sugars efficiently, i.e., sucrose, glucose, galactose, xylose, and arabinose, and efficiently convert each of these sugars to relevant building blocks that serve as precursors for different types of industrially relevant products. Such yeast platforms will have great prospects for application in biorefineries, where the same yeast platform can then be combined with different product formation pathways, and hereby the biorefinery can use a plug-and-play solution for production of a range of different products, and hence allow for easy adjustment to the market developments. Key factors for major advancement towards this scenario will be the ability to construct new strains faster, the improved ability to establish stable strains that can be used in industrial conditions, and focus on using yeast strains that have industrial relevance. Concerning the latter, there has fortunately been development towards a wider use of the CEN.PK yeast strain background, which, besides serving as a good laboratory strain, has excellent performance for industrial use [145]. Even though in the last 10 years there have been increased interactions between academia and industry, it will be necessary to strengthen these ties, as this

will ensure faster implementation of novel developments as well as gaining fundamental understanding of cellular processes that are relevant for industrial processes.

References

- Nielsen J (2001) Metabolic engineering. *Appl Microbiol Biotechnol* 55(3):263–283
- Nielsen J, Keasling JD (2011) Synergies between synthetic biology and metabolic engineering. *Nat Biotechnol* 29(8):693–695
- Tyo KE, Alper HS, Stephanopoulos GN (2007) Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol* 25(3):132–137
- Tyo KE, Kochan K, Nielsen J (2010) Toward design-based engineering of industrial microbes. *Curr Opin Microbiol* 13(3):255–262
- Nielsen J, Jewett MC (2008) Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res* 8(1):122–131
- Conrad TM, Lewis NE, Palsson BO (2011) Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol* 7:509
- Petranovic D, Vemuri GN (2009) Impact of yeast systems biology on industrial biotechnology. *J Biotechnol* 144(3):204–211
- Lee JW, Kim TY, Jang YS, Choi S, Lee SY (2011) Systems metabolic engineering for chemicals and materials. *Trends Biotechnol* 29(8):370–378
- Saling P (2005) Eco-efficiency analysis of biotechnological processes. *Appl Microbiol Biotechnol* 68(1):1–8
- Chandel AK, Singh OV (2011) Weedy lignocellulosic feedstock and microbial metabolic engineering: advancing the generation of 'Biofuel'. *Appl Microbiol Biotechnol* 89(5):1289–1303
- Elkins JG, Raman B, Keller M (2010) Engineered microbial systems for enhanced conversion of lignocellulosic biomass. *Curr Opin Biotechnol* 21(5):657–662
- Wilson DB (2011) Microbial diversity of cellulose hydrolysis. *Curr Opin Microbiol* 14(3):259–263
- Krivoruchko A, Siewers V, Nielsen J (2011) Opportunities for yeast metabolic engineering: lessons from synthetic biology. *Biotechnol J* 6(3):262–276
- Nevoigt E (2008) Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 72(3):379–412
- de Jong B, Siewers V, Nielsen J (2011) Systems biology of yeast: enabling technology for development of cell factories for production of advanced biofuels. *Curr Opin Biotechnol* (in press)
- Kim IK, Roldao A, Siewers V, Nielsen J (2012) A systems-level approach for metabolic engineering of yeast cell factories. *FEMS Yeast Res* 12(2):228–248
- Guadalupe Medina V, Almering MJ, van Maris AJ, Pronk JT (2010) Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Appl Environ Microbiol* 76(1):190–195
- Chen X, Nielsen KF, Borodina I, Kielland-Brandt MC, Karhumaa K (2011) Increased isobutanol production in *Saccharomyces cerevisiae* by overexpression of genes in valine metabolism. *Bio-technol Biofuels* 4:21
- Yu KO, Jung J, Kim SW, Park CH, Han SO (2012) Synthesis of FAEs from glycerol in engineered *Saccharomyces cerevisiae*

- using endogenously produced ethanol by heterologous expression of an unspecific bacterial acyltransferase. *Biotechnol Bioeng* 109(1):110–115
20. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MC, Withers ST, Shiba Y, Sarpong R, Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086):940–943
 21. Peralta-Yahya PP, Ouellet M, Chan R, Mukhopadhyay A, Keasling JD, Lee TS (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat Commun* 2:483
 22. Lee W, Dasilva NA (2006) Application of sequential integration for metabolic engineering of 1,2-propanediol production in yeast. *Metab Eng* 8(1):58–65
 23. Toivari M, Maheimo H, Penttilä M, Ruohonen L (2010) Enhancing the flux of D-glucose to the pentose phosphate pathway in *Saccharomyces cerevisiae* for the production of D-ribose and ribitol. *Appl Microbiol Biot* 85(3):731–739
 24. Zhao L, Wang J, Zhou J, Liu L, Du G, Chen J (2011) Modification of carbon flux in *Saccharomyces cerevisiae* to improve L-lactic acid production. *Wei Sheng Wu Xue Bao* 51(1):50–58
 25. Zhang B, Carlson R, Srienc F (2006) Engineering the monomer composition of polyhydroxyalkanoates synthesized in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 72(1):536–543
 26. van Maris AJ, Geertman JM, Vermeulen A, Groothuisen MK, Winkler AA, Piper MD, van Dijken JP, Pronk JT (2004) Directed evolution of pyruvate decarboxylase-negative *Saccharomyces cerevisiae*, yielding a C2-independent, glucose-tolerant, and pyruvate-hyperproducing yeast. *Appl Environ Microbiol* 70(1):159–166
 27. Raab AM, Gebhardt G, Bolotina N, Weuster-Botz D, Lang C (2010) Metabolic engineering of *Saccharomyces cerevisiae* for the biotechnological production of succinic acid. *Metab Eng* 12(6):518–525
 28. Madsen KM, Udatha GD, Semba S, Otero JM, Koetter P, Nielsen J, Ebizuka Y, Kushiro T, Panagiotou G (2011) Linking genotype and phenotype of *Saccharomyces cerevisiae* strains reveals metabolic engineering targets and leads to triterpene hyper-producers. *PLoS One* 6(3):e14763
 29. Verwaal R, Wang J, Meijnen JP, Visser H, Sandmann G, van den Berg JA, van Ooyen AJ (2007) High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllum dendrorhous*. *Appl Environ Microbiol* 73(13):4342–4350
 30. Westfall PJ, Pitera DJ, Lenihan JR, Eng D, Woolard FX, Regentin R, Horning T, Tsuruta H, Melis DJ, Owens A, Fickes S, Diola D, Benjamin KR, Keasling JD, Leavell MD, McPhee DJ, Renninger NS, Newman JD, Paddon CJ (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci USA* 109(3):E1111–1118
 31. Farhi M, Marhevka E, Masci T, Marcos E, Eyal Y, Ovadis M, Abeliovich H, Vainstein A (2011) Harnessing yeast subcellular compartments for the production of plant terpenoids. *Metab Eng* 13(5):474–481
 32. Kirby J, Nishimoto M, Park JG, Withers ST, Nowroozi F, Behrendt D, Rutledge EJG, Fortman JL, Johnson HE, Anderson JV, Keasling JD (2010) Cloning of casbene and neocembrene synthases from Euphorbiaceae plants and expression in *Saccharomyces cerevisiae*. *Phytochemistry* 71(13):1466–1473
 33. Eudes A, Baidoo E, Yang F, Burd H, Hadi M, Collins F, Keasling J, Loqué D (2011) Production of tranolast [*N*-(3',4'-dimethoxycinnamoyl)-anthranilic acid] and its analogs in yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 89(4):989–1000
 34. Asadollahi MA, Maury J, Schalk M, Clark A, Nielsen J (2010) Enhancement of farnesyl diphosphate pool as direct precursor of sesquiterpenes through metabolic engineering of the mevalonate pathway in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 106(1):86–96
 35. Tavares S, Grotkjaer T, Olseni T, Haslam RP, Napier JA, Gunnarsson N (2011) Metabolic engineering of *Saccharomyces cerevisiae* for production of eicosapentaenoic acid, using a novel $\{\delta\}5$ -desaturase from *Paramecium tetraurelia*. *Appl Environ Microbiol* 77(5):1854–1861
 36. Millis JRKWI, Maurina-Bunker JAWI, McMullin TWMWI (2003) Production of farnesol and geranylgeraniol. US Patent US 2003/0092144 A1
 37. Sauer M, Branduardi P, Valli M, Porro D (2004) Production of L-ascorbic acid by metabolically engineered *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*. *Appl Environ Microbiol* 70(10):6086–6091
 38. Rico J, Pardo E, Orejas M (2010) Enhanced production of a plant monoterpene by overexpression of the 3-hydroxy-3-methylglutaryl coenzyme A reductase catalytic domain in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 76(19):6449–6454
 39. Mutka SC, Bondi SM, Carney JR, Da Silva NA, Kealey JT (2006) Metabolic pathway engineering for complex polyketide biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 6(1):40–47
 40. Albertsen L, Chen Y, Bach LS, Rattleff S, Maury J, Brix S, Nielsen J, Mortensen UH (2011) Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. *Appl Environ Microbiol* 77(3):1033–1040
 41. Becker JV, Armstrong GO, van der Merwe MJ, Lambrechts MG, Vivier MA, Pretorius IS (2003) Metabolic engineering of *Saccharomyces cerevisiae* for the synthesis of the wine-related antioxidant resveratrol. *FEMS Yeast Res* 4(1):79–85
 42. Brochado AR, Matos C, Moller BL, Hansen J, Mortensen UH, Patil KR (2010) Improved vanillin production in baker's yeast through *in silico* design. *Microb Cell Fact* 9:84
 43. Mapelli V, Hillestrøm PR, Kápolna E, Larsen EH, Olsson L (2011) Metabolic and bioprocess engineering for production of selenized yeast with increased content of seleno-methylselenocysteine. *Metab Eng* 13(3):282–293
 44. Siewers V, San-Bento R, Nielsen J (2010) Implementation of communication-mediated domains for non-ribosomal peptide production in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 106(5):841–844
 45. Vai M, Brambilla L, Orlandi I, Rota N, Ranzi BM, Alberghina L, Porro D (2000) Improved secretion of native human insulin-like growth factor 1 from gas1 mutant *Saccharomyces cerevisiae* cells. *Appl Environ Microbiol* 66(12):5477–5479
 46. Egel-Mitani M, Andersen AS, Diers II, Hach M, Thim L, Hastrup S, Vad K (2000) Yield improvement of heterologous peptides expressed in yps1-disrupted *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 26(9–10):671–677
 47. Hackel BJ, Huang D, Bubolz JC, Wang XX, Shusta EV (2006) Production of soluble and active transferrin receptor-targeting single-chain antibody using *Saccharomyces cerevisiae*. *Pharm Res* 23(4):790–797
 48. Vellanki RN, Komaravelli N, Tatineni R, Mangamoori LN (2007) Expression of hepatitis B surface antigen in *Saccharomyces cerevisiae* utilizing glyceraldehyde-3-phosphate dehydrogenase promoter of *Pichia pastoris*. *Biotechnol Lett* 29(2):313–318
 49. Lowin T, Raab U, Schroeder J, Fransila R, Modrow S (2005) Parvovirus B19 VP2-proteins produced in *Saccharomyces cerevisiae*: comparison with VP2-particles produced by baculovirus-derived vectors. *J Vet Med B Infect Dis Vet Public Health* 52(7–8):348–352

50. Chigira Y, Oka T, Okajima T, Jigami Y (2008) Engineering of a mammalian *O*-glycosylation pathway in the yeast *Saccharomyces cerevisiae*: production of *O*-fucosylated epidermal growth factor domains. *Glycobiology* 18(4):303–314
51. Rakestraw JA, Sazinsky SL, Piatesi A, Antipov E, Wittrup KD (2009) Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 103(6):1192–1201
52. Kim E-J, Park Y-K, Lim H-K, Park Y-C, Seo J-H (2009) Expression of hepatitis B surface antigen S domain in recombinant *Saccharomyces cerevisiae* using GAL4 promoter. *J Biotech* 141(3–4):155–159
53. Kim HJ, Lee SJ, Kim H-J (2010) Optimizing the secondary structure of human papillomavirus type 16 L1 mRNA enhances L1 protein expression in *Saccharomyces cerevisiae*. *J Biotechnol* 150(1):31–36
54. Petranovic D, Tyo K, Vemuri GN, Nielsen J (2010) Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism. *FEMS Yeast Res* 10(8):1046–1059
55. Snyder M, Gallagher JE (2009) Systems biology from a yeast omics perspective. *FEBS Lett* 583(24):3895–3899
56. Canelas AB, Harrison N, Fazio A, Zhang J, Pitkänen JP, van den Brink J, Bakker BM, Bogner L, Bouwman J, Castrillo JI, Cankorur A, Chumnanpuen P, Daran-Lapujade P, Dikiciglu D, van Eunen K, Ewald JC, Heijnen JJ, Kirdar B, Mattila I, Menonides FI, Niebel A, Penttila M, Pronk JT, Reuss M, Salusjarvi L, Sauer U, Sherman D, Siemann-Herzberg M, Westerhoff H, de Winde J, Petranovic D, Oliver SG, Workman CT, Zamboni N, Nielsen J (2010) Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains. *Nat Commun* 1:145
57. Alberghina L, Cirulli C (2010) Proteomics and systems biology to tackle biological complexity: yeast as a case study. *Proteomics* 10(24):4337–4341
58. Josling T, Blandford D, Earley J (2010) Biofuel and biomass subsidies in the US, EU and Brazil: towards a transparent system of notification. IPC position paper
59. Vennestrom PN, Osmundsen CM, Christensen CH, Taarning E (2011) Beyond petrochemicals: the renewable chemicals industry. *Angew Chem Int Ed Engl* 50(45):10502–10509
60. Perlack RD, Wright LL, Turhollow AF, Graham RL, Stokes BJ, Erbach DC (2005) Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. (Tech Rep ORNL/TM-2006/66, Oak Ridge National Laboratory, Oak Ridge, TN). Also available at http://feedstockreviewnlgov/pdf/billion_ton_visionpdf
61. Hahn-Hagerdal B, Karhumaa K, Fonseca C, Spencer-Martins I, Gorwa-Grauslund MF (2007) Towards industrial pentose-fermenting yeast strains. *Appl Microbiol Biotechnol* 74(5):937–953
62. Pronk JT, van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH, Wisselink HW, Scheffers WA, van Dijken JP (2006) Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Anton Leeuw Int J G* 90(4):391–418
63. Bettiga M, Gorwa-Grauslund MF, Hahn-Hägerdal B (2009) Metabolic engineering in yeast. In: The metabolic pathway engineering handbook. Metabolic pathway engineering handbook. CRC Press, Boca Raton pp 22–21–22–48
64. Den Haan R, Rose SH, Lynd LR, van Zyl WH (2007) Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. *Metab Eng* 9(1):87–94
65. Shigechi H, Koh J, Fujita Y, Matsumoto T, Bito Y, Ueda M, Satoh E, Fukuda H, Kondo A (2004) Direct production of ethanol from raw corn starch via fermentation by use of a novel surface-engineered yeast strain codisplaying glucoamylase and alpha-amylase. *Appl Environ Microbiol* 70(8):5037–5040
66. Katahira S, Fujita Y, Mizuike A, Fukuda H, Kondo A (2004) Construction of a xylan-fermenting yeast strain through codisplay of xylanolytic enzymes on the surface of xylose-utilizing *Saccharomyces cerevisiae* cells. *Appl Environ Microbiol* 70(9):5407–5414
67. Fujita Y, Ito J, Ueda M, Fukuda H, Kondo A (2004) Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. *Appl Environ Microbiol* 70(2):1207–1212
68. Siso MIG, Beccerra M, Prado SD, Rodriguez-Belmonte E, Cerdan ME (2002) Metabolic engineering for direct lactose utilization by *Saccharomyces cerevisiae*. *Biotechnol Lett* 24(17):1391–1396
69. Ostergaard S, Roca C, Ronnow B, Nielsen J, Olsson L (2000) Physiological studies in aerobic batch cultivations of *Saccharomyces cerevisiae* strains harboring the MEL1 gene. *Biotechnol Bioeng* 68(3):252–259
70. Dmytryuk OV, Voronovsky AY, Abbas CA, Dmytryuk KV, Ishchuk OP, Sibirny AA (2008) Overexpression of bacterial xylose isomerase and yeast host xylulokinase improves xylose alcoholic fermentation in the thermotolerant yeast *Hansenula polymorpha*. *FEMS Yeast Res* 8(1):165–173
71. Lee KS, Hong ME, Jung SC, Ha SJ, Yu BJ, Koo HM, Park SM, Seo JH, Kweon DH, Park JC, Jin YS (2011) Improved galactose fermentation of *Saccharomyces cerevisiae* through inverse metabolic engineering. *Biotechnol Bioeng* 108(3):621–631
72. Wisselink HW, Toirkens MJ, Del Rosario Franco Berriel M, Winkler AA, van Dijken JP, Pronk JT, van Maris AJ (2007) Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl Environ Microbiol* 73(15):4881–4891
73. Wohlbach DJ, Kuo A, Sato TK, Potts KM, Salamov AA, Labutti KM, Sun H, Clum A, Pangilinan JL, Lindquist EA, Lucas S, Lapidus A, Jin M, Gunawan C, Balan V, Dale BE, Jeffries TW, Zinkel R, Barry KW, Grigoriev IV, Gasch AP (2011) Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. *Proc Natl Acad Sci USA* 108(32):13212–13217
74. Ghosh A, Zhao H, Price ND (2011) Genome-scale consequences of cofactor balancing in engineered pentose utilization pathways in *Saccharomyces cerevisiae*. *PLoS One* 6(11):e27316
75. Garcia Sanchez R, Hahn-Hagerdal B, Gorwa-Grauslund MF (2010) PGM2 overexpression improves anaerobic galactose fermentation in *Saccharomyces cerevisiae*. *Microb Cell Fact* 9:40
76. Bro C, Knudsen S, Regenberg B, Olsson L, Nielsen J (2005) Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* 71(11):6465–6472
77. Ostergaard S, Olsson L, Johnston M, Nielsen J (2000) Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat Biotechnol* 18(12):1283–1286
78. Hong KK, Vongsangnak W, Vemuri GN, Nielsen J (2011) Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc Natl Acad Sci USA* 108(29):12179–12184
79. Ha SJ, Galazka JM, Kim SR, Choi JH, Yang X, Seo JH, Glass NL, Cate JH, Jin YS (2011) Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc Natl Acad Sci USA* 108(2):504–509
80. Teunissen AWRH, De Bont JAM (2010) Xylose isomerase genes and their use in fermentation of pentose sugars. WO Patent WO 2010/074577 A1
81. Op Den Camp HJMO, Harhangi HR, Van Der Drift C, Pronk JT (2009) Transformed eukaryotic cells that directly convert xylose to xylulose. US Patent US 7622284

82. Klaassen P, van der Laan JM, Gielesen BEM, van Suylen GP (2009) A pentose sugar fermenting cell. WO Patent WO 2009/109630 A1
83. Klaassen P, van der Laan JM, Gielesen BEM, van Suylen GP (2009) A pentose sugar fermenting cell. WO Patent WO 2009/109631 A1
84. Ho NWY, Chen Z-D (2009) Stable recombinant yeasts for fermenting xylose to ethanol. US Patent US 7527927
85. De Bont JAM (2009) Novel arabinose-fermenting eukaryotic cells. WO Patent WO 2009/011591 A2
86. Van Maris AJA, Pronk JT, Wisselink W, Van Dijken JP, Winkler AA, De Winde H (2008) Metabolic engineering of arabinose-fermenting yeast cells. WO Patent WO 2008/041840 A1
87. Hughes SR, Butt TR (2010) *Saccharomyces cerevisiae* engineered for xylose utilization. WO Patent WO 2010/039692 A2
88. Mainguet SE, Liao JC (2010) Bioengineering of microorganisms for C3 to C5 alcohols production. Biotechnol J 5(12):1297–1308
89. Yan Y, Liao JC (2009) Engineering metabolic systems for production of advanced fuels. J Ind Microbiol Biotechnol 36(4):471–479
90. Feldman RMR, Gunawardena U, Urano J, Meinhold P, Aristidou A, Dundon CA, Smith C (2011) Yeast organism producing isobutanol at a high yield. US Patent US 2011/0020889 A1
91. Gunawardena U, Meinhold P, Peters MW, Urano J, Feldman RMR (2010) Butanol production by metabolically engineered yeast. US Patent US 2010/0062505 A1
92. Ramey DE (2008) Butanol: the other alternative fuel. agricultural biofuels: technology, sustainability and profitability, 137–147. (http://nabc.cals.cornell.edu/pubs/nabc_19/NABC19_SPlenary2_Ramey.pdf)
93. Wallner T, Miers SA, McConnell S (2009) A comparison of ethanol and butanol as oxygenates using a direct-injection, spark-ignition engine. J Eng Gas Turb Power 131(3):032802. (<http://scitation.aip.org/getabservlet/GetabsServlet?prog=normal&id=JETPEZ000131000003032802000001&idtype=cvips&gifs=yes>)
94. Weber C, Farwick A, Benisch F, Brat D, Dietz H, Subtil T, Boles E (2010) Trends and challenges in the microbial production of lignocellulosic bioalcohol fuels. Appl Microbiol Biotechnol 87(4):1303–1315
95. Nielsen DR, Leonard E, Yoon SH, Tseng HC, Yuan C, Prather KL (2009) Engineering alternative butanol production platforms in heterologous bacteria. Metab Eng 11(4–5):262–273
96. Hazelwood LA, Daran JM, van Maris AJ, Pronk JT, Dickinson JR (2008) The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. Appl Environ Microbiol 74(8):2259–2266
97. Dickinson JR, Harrison SJ, Hewlins MJ (1998) An investigation of the metabolism of valine to isobutyl alcohol in *Saccharomyces cerevisiae*. J Biol Chem 273(40):25751–25756
98. Porro D, Gasser B, Fossati T, Maurer M, Branduardi P, Sauer M, Mattanovich D (2011) Production of recombinant proteins and metabolites in yeasts: when are these systems better than bacterial production systems? Appl Microbiol Biotechnol 89(4):939–948
99. Lee SY, Park JH, Jang SH, Nielsen LK, Kim J, Jung KS (2008) Fermentative butanol production by Clostridia. Biotechnol Bioeng 101(2):209–228
100. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature 451(7174):86–89
101. Steen EJ, Chan R, Prasad N, Myers S, Petzold CJ, Redding A, Ouellet M, Keasling JD (2008) Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. Microb Cell Fact 7:36
102. Donaldson GK, Eliot AC, Flint D, Maggio-Hall LA, Nagarajan V (2011) Fermentative production of four carbon alcohols. US Patent US 7993889 B1
103. Anthony LC, Huang LL, Ye RW (2010) Production of isobutanol in yeast mitochondria. US Patent US 2010/0129886 A1
104. Festel G, Boles E, Weber C, Brat D (2011) Fermentative production of isobutanol with yeast. US Patent US 2011/0053235 A1
105. Urano J, Dundon CA, Meinhold P, Feldman RMR, Aristidou A, Hawkins A, Buelter T, Peters M, Lies D, Porter-Scheinman S, Berry R, Kalra I (2011) Cytosolic isobutanol pathway localization for the production of isobutanol. US Patent US 2011/0076733 A1
106. Buelter T, Meinhold P, Smith C, Aristidou A, Dundon CA, Urano J (2010) Engineered microorganisms for the production of one or more target compounds. WO Patent WO 2010/075504 A2
107. Bakker BM, Overkamp KM, van Maris AJ, Kotter P, Luttik MA, van Dijken JP, Pronk JT (2001) Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. FEMS Microbiol Rev 25(1):15–37
108. Dundon CA, Aristidou A, Hawkins A, Lies D, Albert LH (2011) Methods of increasing dihydroxy acid dehydratase activity to improve production of fuels, chemicals, and amino acids. US Patent US 2011/0183393 A1
109. Liao D-I, Nelson MJ, Bramucci MG (2008) Fermentative production of isobutanol using highly active ketol-acid reductoisomerase enzymes. US Patent US 2008/0261230 A1
110. Donaldson GK, Huang LL, Maggio-Hall LA, Nagarajan V, Nakamura CE, Suh W (2008) Fermentative production of four carbon alcohols. US Patent US 2008/0182308 A1
111. Donaldson GK, Eliot AC, Nagarajan V, Nakamura CE, Huang LL (2007) Fermentative production of four carbon alcohols. US Patent US 2007/0259410 A1
112. Flint D, Rothman SC, Suh W, Tomb J-F, Ye RW (2010) Identification and use of bacterial [2Fe-2S] dihydroxy-acid dehydratases. US Patent US 2010/0081154 A1
113. Anthony LC, Maggio-Hall LA, Rothman SC, Tomb J-F (2010) Increased heterologous Fe-S enzyme activity in yeast. US Patent US 2010/0081179 A1
114. Van Dyk TK (2010) Yeast with increased butanol tolerance involving a multidrug efflux pump gene. US Patent US 2010/0221801 A1
115. Bramucci MG, Larossa RA, Smulski DR (2010) Yeast with increased butanol tolerance involving cell wall integrity pathway. US Patent US 2010/0167364 A1
116. Bramucci MG, Larossa RA, Smulski DR (2010) Yeast with increased butanol tolerance involving high osmolarity/glycerol response pathway. US Patent US 2010/0167365 A1
117. Bramucci MG, Larossa RA, Singh M (2010) Yeast with increased butanol tolerance involving filamentous growth response. US Patent US 2010/0167363 A1
118. Larossa RA (2009) Yeast strain for production of four carbon alcohols. US Patent US 2009/0280546 A1
119. Knoshaug EP, Zhang M (2009) Butanol tolerance in a selection of microorganisms. Appl Biochem Biotechnol 153(1–3):13–20
120. Maury J, Asadollahi MA, Moller K, Clark A, Nielsen J (2005) Microbial isoprenoid production: an example of green chemistry through metabolic engineering. Adv Biochem Eng Biotechnol 100:19–51
121. Zhang F, Rodriguez S, Keasling JD (2011) Metabolic engineering of microbial pathways for advanced biofuels production. Curr Opin Biotechnol 22(6):775–783
122. Withers ST, Keasling JD (2007) Biosynthesis and engineering of isoprenoid small molecules. Appl Microbiol Biotechnol 73(5):980–990

123. Kirby J, Keasling JD (2008) Metabolic engineering of microorganisms for isoprenoid production. *Nat Prod Rep* 25(4):656–661
124. Shiba Y, Paradise EM, Kirby J, Ro DK, Keasling JD (2007) Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* 9(2):160–168
125. Clark A, Maury J, Asadollahi MA, MØLLer K, Nielsen J, Schalk M (2009) Method for producing terpenes and MEP-transformed microorganisms therefore. US Patent US 2009/0155874 A1
126. Maury J, Asadollahi MA, Moller K, Schalk M, Clark A, Formenti LR, Nielsen J (2008) Reconstruction of a bacterial isoprenoid biosynthetic pathway in *Saccharomyces cerevisiae*. *FEBS Lett* 582(29):4032–4038
127. Muntendam R, Melillo E, Ryden A, Kayser O (2009) Perspectives and limits of engineering the isoprenoid metabolism in heterologous hosts. *Appl Microbiol Biotechnol* 84(6):1003–1019
128. Chang MC, Keasling JD (2006) Production of isoprenoid pharmaceuticals by engineered microbes. *Nat Chem Biol* 2(12):674–681
129. Ubersax JA, Platt DM (2010) Genetically modified microbes producing isoprenoids. WO Patent WO 2010/141452 A1
130. Tsuruta H, Lenihan JR, Regentin R (2009) Production of isoprenoids. WO Patent WO 2009/042070 A2
131. Basso LC, de Amorim HV, de Oliveira AJ, Lopes ML (2008) Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res* 8(7):1155–1163
132. Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* 21(7):796–802
133. Keasling JD, Newman JD, Pitera DJ (2010) Method for enhancing production of isoprenoid compounds. US Patent US 7670825
134. Johnston M, Flick JS, Pexton T (1994) Multiple mechanisms provide rapid and stringent glucose repression of GAL gene expression in *Saccharomyces cerevisiae*. *Mol Cell Biol* 14(6):3834–3841
135. Asadollahi MA, Maury J, Moller K, Nielsen KF, Schalk M, Clark A, Nielsen J (2008) Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of ERG9 repression on sesquiterpene biosynthesis. *Biotechnol Bioeng* 99(3):666–677
136. Kizer J (2009) Methods of increasing isoprenoid or isoprenoid precursor production. WO Patent WO 2009/005704 A1
137. Chang M, Krupa RA, Ro D-K, Yoshikuni Y, Keasling JD (2009) Nucleic acids encoding modified cytochrome P450 enzymes and methods of use thereof. US Patent US 2009/0098626 A1
138. Bajad S, Leavell M (2009) Methods of monitoring metabolic pathways. WO Patent WO 2009/097339 A1
139. Millis JR, Maurina-Brunker J, McMullin TW (2011) Production of isoprenoids. US Patent US 7927862
140. Julien B, Burlingame R (2010) Method for production of isoprenoids. US Patent US 2010/0151519 A1
141. Kinney AJ, Ni H, Rouviere PE, Suh W (2007) Method to increase hydrophobic compound titer in a recombinant microorganism. US Patent US 7256014
142. Brzostowicz PC, Rouviere PE, Pollak DM (2006) Biological production of tetrahydrodihydrocannabinol. US Patent US 7087403
143. Bramucci MG, Brzostowicz PC, Cheng Q, Kostichka KN, Rouviere PE, Nagarajan V, Tao L, Thomas SM (2006) Genes involved in isoprenoid compound production. US Patent US 7034140
144. Cheng Q, Koffas M, Norton KC, Odom JM, Picataggio SK, Rouviere PE, Schenzle A, Tomb J-F (2002) Genes involved in isoprenoid compound production. WO Patent WO 2002/020733 A2
145. van Dijken JP, Bauer J, Brambilla L, Duboc P, Francois JM, Gancedo C, Giuseppin ML, Heijnen JJ, Hoare M, Lange HC, Madden EA, Niederberger P, Nielsen J, Parrou JL, Petit T, Porro D, Reuss M, van Riel N, Rizzi M, Steensma HY, Verrips CT, Vindelov J, Pronk JT (2000) An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 26(9–10):706–714

CHALMERS UNIVERSITY OF TECHNOLOGY
SE-412 96 Göteborg, Sweden
Telephone: +46-(0)-31 772 1000
www.chalmers.se

Engineers have established significant development in the massive production of fuels and chemicals from petroleum and our generation is taking benefits from these technical advances. However, since the petroleum based production is using limited resources and generating serious environmental problems, our generation should prepare new technologies for the next generation, which uses renewable resources and alleviates environmental issues.

Microbial fermentation processes could be one of the possible solutions, because this process utilizes biomass that is continuously produced with absorbing carbon dioxide in connection with its growth. Engineering or reconstructing of microorganisms is the requisite step for the development of fermentation process. The engineering of biological systems is certainly different from mechanical or chemical engineering, since the biology is not only vastly complicated in their reaction networks and regulations, but also has evolution. One of the strategies for engineering the microorganism is to learn and apply nature's algorithm. Currently there are tools available for analysis genome-wide molecular and genetic changes. This means one can trace nature's strategies for obtaining new traits.



In this thesis, mutations in the *RAS2* gene were identified as the genetic bases for improving galactose utilization in yeast *S. cerevisiae*. Firstly, these mutations were only designable by nature's random mutagenesis; because not only the relationship between these mutations and galactose utilization was not predictable, but also even though they were located in the same gene, the effects from each of them were vastly different. Secondly, these mutations were only detectable by genome-scale analyses, since these analyses can only scan whole genome level changes. The implications of evolutionary strategies and the impact of genome-scale analyses on characterization of evolved mutants are discussed.

KUK-KI HONG

Systems and Synthetic Biology
Department of Chemical and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY