LUND UNIVERSITY

Evolution of Yeast Respiro-Fermentative Lifestyle and the Underlying Mechanisms Behind Aerobic Fermentation

Hagman, Arne

2013

Evolution of Yeast Respiro-Fermentative Lifestyle and the Underlying Mechanisms Behind Aerobic Fermentation

A Comparative Physiology and Genomics Study of Saccharomycotina Yeast

Arne Hagman

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Science at Lund University, Sweden.

To be defended in public in the Biology House Lecture Hall, Sölvegatan 35, Lund.
Date 28th of October and time 9.15 am.

Faculty opponent

Karin Breunig

Organization	Document name
LUND UNIVERSITY	DOCTORAL DISSERTATION
	Date of issue: 2013-09-17
Author: Arne Hagman	Sponsoring organization

Title and subtitle

Evolution of Yeast Respiro-Fermentative lifestyle and the Underlying Mechanism Behind the Crabtree Effect / A Comparative Physiology and Genomics Study of *Saccharomycotina* Yeast

Abstract

Under aerobic conditions, most yeasts such as *Kluyveromyces lactis,* prefer the respiratory pathway and some, such as *Saccharomyces cerevisiae* prefer less energy efficient fermentative pathway for their energy metabolism. These two metabolic strategies are also known as Crabtree negative and Crabtree positive respectively, and the evolution of the latter has lately been explained by the "make-accumulate-consume" life strategy. Scientists have for more than a century tried to elucidate the mechanism behind the physiology and the evolution of the peculiar respiro-fermentation trait.

During the last decades, comparative genomics approaches have enabled the reconstruction of the evolutionary history of yeast, and several evolutionary events have been identified and postulated to have contributed to the development of the respiro-fermentative lifestyle in the *Saccharomyces* lineage. However, many of these inspiring studies have been verified with reference species only. Therefore, as parts of my thesis I conducted large-scale physiology studies of more than 40 yeast species and their central carbon metabolism under controlled conditions, in bioreactors. This was done in order to map the evolution of aerobic fermentation in yeast belonging to the *Saccharomyces* lineage that span over 200 million years of yeast evolution.

This evolutionary blueprint, which most likely will be an invaluable information source of primary data for future *in silico* studies on the evolution of Crabtree effect, has already verified the importance of evolutionary events, such as promoter rewiring, chromatin relaxation, whole genome duplication, gene duplication and lateral gene-transfers. I further propose a mechanism that provides an explanation for the origin of the respiro-fermentative lifestyle in yeast, and how this was subsequently, through a multistep process developed into the Crabtree effect as we know it in the modern yeasts today, such as *S. cerevisiae* and its sister species.

Key words

Crabtree effect, aerobic fermentation, respiro-fermentation, ethanol production, yeast, central carbon metabolism, evolution of alternative lifestyles

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language: English
ISSN and key title	ISBN: 978-91-7473-709-7; 978-91-7473-710-3 (pdf)
Recipient's notes	Number of pages 235; Price; Security classification

Signature _____ Date 2013-09-19

Evolution of Yeast Respiro-Fermentative Lifestyle and the Underlying Mechanisms Behind Aerobic Fermentation

A Comparative Physiology and Genomics Study of Saccharomycotina Yeast

Arne Hagman

LUND
UNIVERSITY

Faculty of Science, Department of Biology
Lund University
Sölvegatan 35
SE-223 62 LUND
SWEDEN

ISBN 978-91-7473-709-7 (printed version)
ISBN 978-91-7473-710-3 (electronic version)

Printed in Sweden by Media-Tryck, Lund University
Lund 2013

It's about the purpose to serve,
not the lifestyle to live...

Table of contents

Abstract

Under aerobic conditions, most yeasts such as *Kluyveromyces lactis*, prefer the respiratory pathway and some, such as *Saccharomyces cerevisiae* prefer less energy efficient fermentative pathway for their energy metabolism. These two metabolic strategies are also known as Crabtree negative and Crabtree positive respectively, and the evolution of the latter has lately been explained by the "make-accumulate-consume" life strategy. Scientists have for more than a century tried to elucidate the mechanism behind the physiology and the evolution of the peculiar respiro-fermentation trait.

During the last decades, comparative genomics approaches have enabled the reconstruction of the evolutionary history of yeast, and several evolutionary events have been identified and postulated to have contributed to the development of the respiro-fermentative lifestyle in the *Saccharomyces* lineage. However, many of these inspiring studies have been verified with reference species only. Therefore, as parts of my thesis I conducted large-scale physiology studies of more than 40 yeast species and their central carbon metabolism under controlled conditions, in bioreactors. This was done in order to map the evolution of aerobic fermentation in yeast belonging to the *Saccharomyces* lineage that span over 200 million years of yeast evolution.

This evolutionary blueprint, which most likely will be an invaluable information source of primary data for future *in silico* studies on the evolution of Crabtree effect, has already verified the importance of evolutionary events, such as promoter rewiring, chromatin relaxation, whole genome duplication, gene duplication and lateral gene-transfers. I further propose a mechanism that provides an explanation for the origin of the respiro-fermentative lifestyle in yeast, and how this was subsequently, through a multistep process developed into the Crabtree effect as we know it in the modern yeasts today, such as *S. cerevisiae* and its sister species.

Populärvetenskaplig sammanfattning

Jäst kan antingen bryta ner socker direkt till koldioxid och vatten genom andning, eller indirekt via koldioxid och etanol, genom jäsning. Andning är en process som förbrukar syre, och sker därför endast under syrerika (aeroba) förhållanden. Jäsning kräver till skillnad från andning inte tillgång till syre, och kan därför ske under syrefattiga (anaeroba) förhållanden.

Eftersom andning är en mer effektiv strategi för utvinning av energi från socker för tillväxt så föredrar många jästarter, som t.ex. *Kluyveromyces lactis* att andas under aeroba förhållanden. Andra jästarter, som t.ex. bryggeri- och bagerijästen *Saccharomyces cerevisiae* föredrar dock jäsning, även under aeroba förhållanden. Jäst som uppvisar aerob jäsning tillhör kategorin Crabtree positiva, vilket är en karakteristisk fysiologi hos olika typer av tumörceller, som upptäcktes 1929 av biokemisten Herbert Crabtree.

Det är ett mysterium varför Crabtree effekten existerar i naturen, och detta fenomen som ger upphov till jäsning under aeroba förhållanden har förbryllat många vetenskapsmän sedan upptäckten av bryggerijästen och dess aktivitet, av Louis Pasteur för mer än 150 år sedan. Tack vare genomsekvenseringen av *S. cerevisiae* och andra närbesläktade arter inom familjen *Saccharomycetaceae* under de senaste årtionden så har man, med hjälp av komparativa studier av jästgenom, kunnat rekonstruera släktträd och identifiera evolutionära händelser som kan ha bidragit till uppkomsten av aerob jäsning och Crabtree effekten.

Även om de inspirerande komparativa studierna av olika jästgenom har underlättats av den ständigt ökande tillgängligheten av sekvenserade genom, så har verifieringen av dessa studier oftast varit begränsade till fysiologin hos ett fåtal referensarter. Därför har jag, som en del i mitt doktorandprojekt utfört en storskalig fysiologisk undersökning av mer än 40 olika jästarter, under likformiga och kontrollerade förhållanden i bioreaktorer. Detta gjordes för att ”kartlägga” evolutionen av aerob jäsning och Crabtree effekten hos jästarter med en gemensam evolutionär historia på över 250 miljoner år.

Min storskaliga undersökning av jästfysiologin hos många olika jästarter har genererat en stor mängd rådata som troligen kommer att utgöra en fortsatt viktig informationskälla för framtida studier, främst inom systembiologin, och i utvecklingen av liknande fysiologi i parallella biologiska system. Resultatet från mina studier har redan bekräftat vikten av olika evolutionära mekanismer som kan ha bidragit till aerob jäsning och Crabtree effekten, som exempelvis

förändringar i promotor sekvenser, packning och upplindning av kromatinet, gen och genom duplicering, och horisontell genöverföring.

Slutligen så föreslår jag en fundamental mekanism som kan förklara uppkomsten av aerob jäsning i familjen *Saccharomycetaceae,* och hur detta sedermera har utvecklats till Crabtree effekten som vi känner till idag i de moderna jästarterna, *S. cerevisiae* och dess närbesläktade arter.

List of publications

Publications included in the thesis:

Paper I
Elzbieta Rozpedowska, Silvia Galafassi, Louise Johansson, Arne Hagman, Jure Piškur, Concetta Compagno
Candida albicans - a pre-whole genome duplication yeast - is predominantly aerobic and a poor ethanol producer.
FEMS Yeast Research 11, 285 (2011).

Paper II
Arne Hagman, Torbjörn Säll, Concetta Compagno, Jure Piškur
Yeast "Make-Accumulate-Consume" life strategy evolved as a multi-step process that predates the whole genome duplication.
PLoS One 8, e68734 (2013).

Paper III
Arne Hagman, Jure Piškur
The origin of the short-term Crabtree effect coincides with the long-term Crabtree effect.
Manuscript (2013).

Paper IV
Arne Hagman, Jure Piškur
Overflow metabolism is the underlying mechanism behind short- and long-term Crabtree effect in yeast.
Manuscript (2013).

All authors in the list of publications have given their consent for the use of their work in the thesis.

Additional publications that are not included in the thesis:

Paper I
Paul G. Becher, Gerhard Flick, Elzbieta Rozpedowska, Alexandra Schmidt, Arne Hagman, Sebastien Lebreton, Mattias C. Larsson, Bill S. Hansson, Jure Piškur, Peter Witzgall, Marie Bengtsson
Yeast, not fruit volatiles mediate Drosophila melanogaster attraction, oviposition and development.
Functional Ecology 26, 822–828 (2012).

Paper II
Peter Witzgall, Magali Proffit, Elzbieta Rozpedowska, Paul G. Becher, Stefanos Andreadis, Miryan Coracini, Tobias U. T. Lindblom, Lee J. Ream, Arne Hagman, Marie Bengtsson, Cletus P. Kurtzman, Jure Piškur, Alan Knight
"This is not an Apple"–Yeast Mutualism in Codling Moth.
Journal of Chemical Ecology, (2012).

List of contributions

Paper I
Candida albicans – a pre-whole genome duplication yeast – is predominantly aerobic and a poor ethanol producer

Established parts of the experimental setup.

Paper II
Yeast ”Make-Accumulate-Consume” life strategy evolved as a multi-step process that predates the whole genome duplication

Designed the experiments together with Jure Piškur and Concetta Compagno. Performed all experiments. Performed the data analysis. Analyzed the result in cooperation with Jure Piškur and Concetta Compagno. Performed the statistical analysis in cooperation with Torbjörn Säll.

Paper III
The origin of the short-term Crabtree effect coincides with the long-term Crabtree effect.

Designed the experiments together with Jure Piškur and Concetta Compagno. Performed all experiments. Performed the data analysis. Analyzed the result in cooperation with Jure Piškur. Wrote the paper in cooperation with Jure Piškur.

Paper IV
Overflow is the fundamental mechanism behind short- and long-term Crabtree effect in yeast

Performed the data analysis. Analyzed the result in cooperation with Jure Piškur. Performed the statistical analysis. Wrote the paper together with Jure Piškur.

List of abbreviations

ADH	Alcohol dehydrogenase
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
Ba	Billion years ago
By	Billion years
CIT1	Gene encoding a citrate synthase
DHODase	Dihydroorotate dehydrogenase (Ura1p and Ura9p)
FADH2	Flavin adenine dinucleotide
GAL2	Gene encoding a hexose (galactose/glucose) transporter
GLK1	Gene encoding a hexokinase
HGT	Horizontal gene transfer/Lateral gene transfer
HGT1	Gene encoding a low affinity hexose transporter (not a member of the *HXT*-gene family) in *K. lactis*
HXT	Gene family encoding hexose (glucose) transporters
HXK	Gene encoding a hexokinase
Ma	Million years ago
MIG1	Gene encoding a transcription factor involved in glucose repression (homologous to *MIG2* and *MIG3* in *S. cerevisiae*)
MRX	A protein complex that functions along with Sae2p endonuclease
My	Million years
NADH	Nicotinamide adenine dinucleotide
RAG1	Gene encoding a high affinity (member of the *HXT*-gene family) transporter in *K. lactis*
REV3	Gene encoding a catalytic subunit of polymerase ζ
RGE	*Cis*-acting element in yeast, upstream of ribosomal genes and genes belonging to central carbon metabolism
RGT2	Gene encoding a high glucose concentration sensor in *S. cerevisiae*

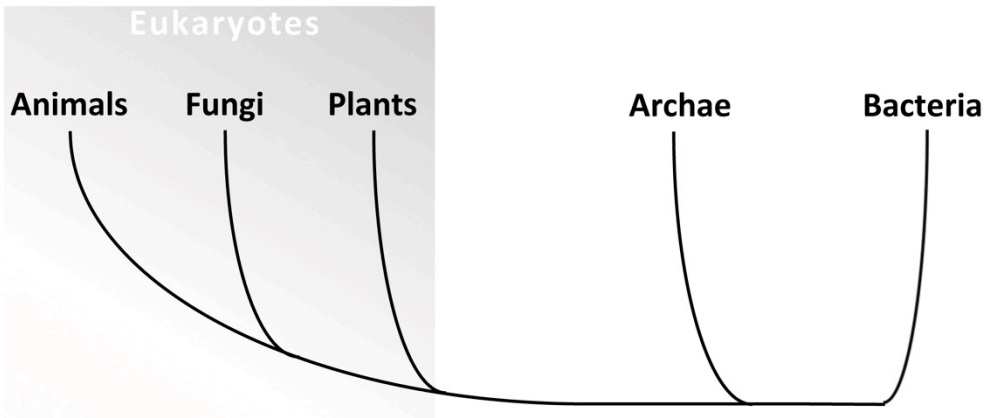
SAE2	Gene encoding an endonuclease, involved in double-strand break repair
SNF1	Gene encoding a Mig1p kinase, involved in the regulation of glucose repression in *S. cerevisiae*
SNF3	Gene encoding a sensor for glucose at low concentrations in *S. cerevisiae*
TCA cycle	Tricarboxylic acid cycle/Citric acid cycle/Krebs cycle
TF	Transcription factor
URA1	Cytoplasmic dihydroorotate dehydrogenase encoding gene
URA9	Mitochondrial dihydroorotate dehydrogenase encoding gene
WGD	Whole genome duplication

1. Fungal taxonomy – the traditional approach

1.1. Yeast and fungi, what are they?

The fungal kingdom forms next to plants and animals the eukaryote domain in the tree of life (Figure 1). Traditional taxonomy of fungi has relied on biological-species concepts, the observation of physiology such as metabolism and morphological characteristics such as sexual structures, sexual states, ultrastructure, and cell wall composition. A unique trait of fungi that distinguishes them from plants and animals is the possession of a chitinous cell wall. All fungi are chemoheterotrophic organisms that grow either as single cells or as multicellular mycelium, which consists of hyphae.

Figure 1. The tree of life can be divided into three domains, eukaryotes, archae and bacteria. Three kingdoms, animals, fungi and plants comprise eukaryotes. It is not certain whether plants or fungi diverged first from the common progenitor, but animals most likely diverged late in the history of eukaryotes. This figure is adapted from an unrooted multigene-concatenation phylogeny of 31 orthologous genes that occur in 191 sequenced genomes, spanning the tree of life (*1*).



Fungi are free-living, parasitic or mutualistic such as lichens and mychorizza, with both green algae and cyanobacteria or with plants respectively. Depending on the phenotypical characterization of sexual states, fungi can be grouped as either teleomorphic (sexual) or anamorphic (asexual). When fungi sporulate, they can give rise to a variety of morphologies. Based on these morphologies, fungi have traditionally been divided into three major phyla, (I) Zygomycota (II) Basidiomycota and (III) Ascomycota. Yeast is a unicellular form of fungi and does not produce fruiting bodies as compared to other fungi, such as mushrooms, bracket fungi, rusts and smuts. Multicellular fungi that produce

fruiting bodies exist in all fungal phyla, however yeast predominantly occupies Ascomycota and Basidiomycota (Figure 2).

1.2. Yeast biodiversity

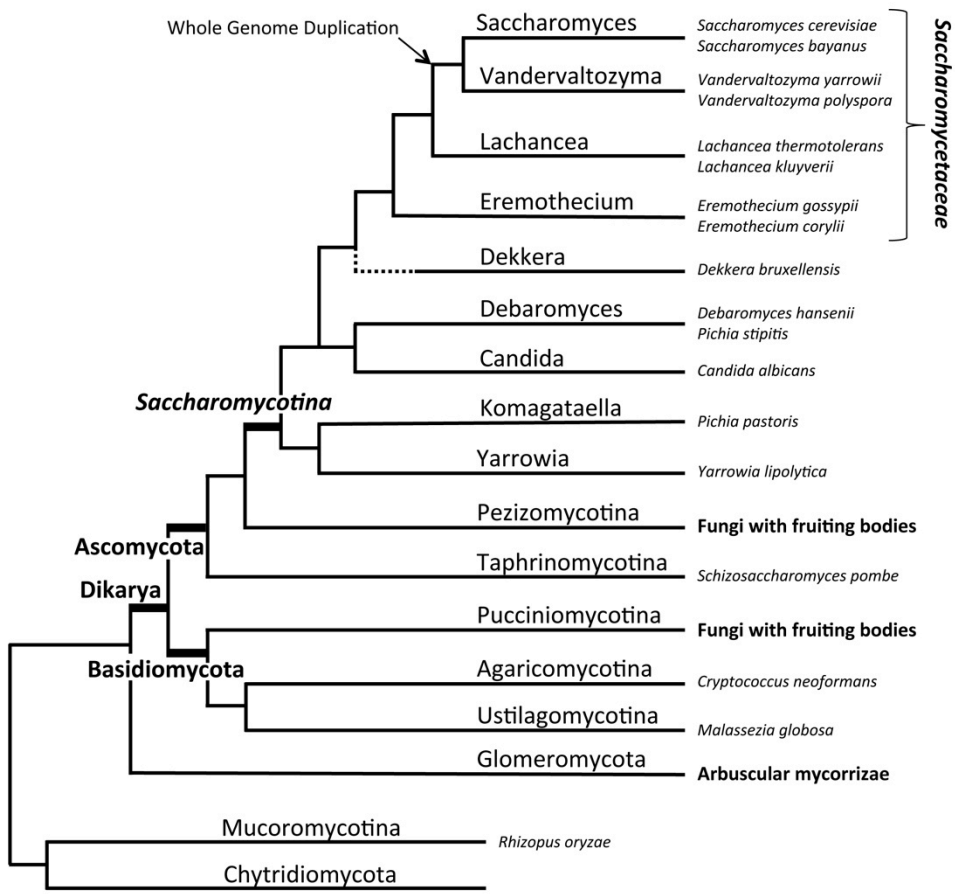
Fungi comprise one of the species-richest kingdoms of Eukaryotes. Ascomycota and Basidiomycota are by far the two largest fungal phyla with the highest biodiversity. Up to now, more than 64 000 species belonging to Ascomycota, and more than 31 000 Basidiomycota species have been identified (Figure 3). If these numbers reflect the true interrelationship in sizes between different phyla, the Basidiomycota would comprise almost one third, Ascomycota almost two third, and both together almost 98% of the total fungal biodiversity. According to the classical biological-species concept, only actual or potential interbreeding population that is isolated by intrinsic reproductive barriers can be considered as species. The biological-species concept works well for teleomorphic fungi, but has been problematic for the anamorphic ones. Development of modern molecular techniques such as DNA sequencing, and the incorporation of sequence divergence analysis of certain genomic regions into modern systematics as the basis for species identification have helped overcome these problems. One such region is the ca. 600-nucleotide D1/D2 domain that resides at the 5' end of 26S rDNA. In a large-scale study by Kurtzman and Robnett in 1998 (*2*), it was shown that all recognized ascomycetes at that time (nearly 500 species) could be resolved from their unique D1/D2 sequence domain. The approach of using molecular data from type strains for sequence divergence analysis form the basis of modern molecular species concept.

2. Fungal taxonomy – the modern systematics approach

2.1. The emergence of phylogenetic and phylogenomic

The transition from classical taxonomy to modern taxonomy involved the introduction of molecular techniques that enabled sequence analysis. The development of PCR and universal primers for the characterization of fungi is the background for the modern phylogenetic-species concept. DNA sequence analysis in classification of fungdi has not only contributed to increased number of estimated and identified species, but also to increased understanding of their relationship and evolutionary history.

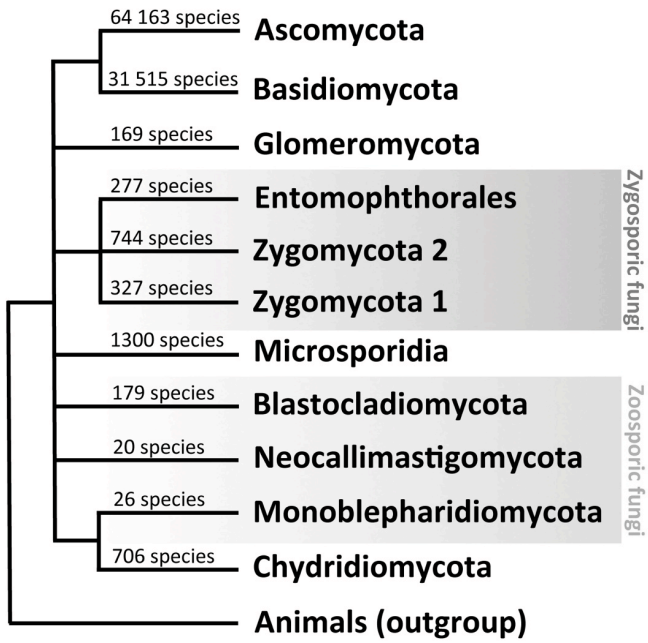
Figure 2. The fungal tree of life that span 600 to 1400 million years of fungal evolutionary history. Branch lengths are arbitrary, and some characteristic species are named [adapted from (*3*-*6*)].



2.2. Fungal biodiversity

At year 2004, roughly 80 000 species of fungi were described (*7*) from an estimated total number of 1.5 million (*8*). In 2011, almost 99 000 were described (*9*), and more recent data from high-throughput sequencing methodology estimated the total fungal biodiversity to 3.5 - 5.1 million species (*10*). This indicates that only 1.9 to 2.8 percent of the total fungal biodiversity has been discovered (Figure 3). Hibbett and colleagues estimated that it would take 1170 – 4170 years to discover all fungal species based on the mean yearly discovery rate of 1196 species per year, and the lower and upper limit of total fungal biodiversity (*8, 10*).

Figure 3. Fungal biodiversity – true yeasts occupy the two largest fungal phyla, Basidiomycota and Ascomycota phylum. Basidiomycota comprise almost 1/3rd, and Ascomycota almost 2/3rd of the total discovered biodiversity (*9*).



2.3. Phylogenetic of the basal fungal lineages

One of the earlier phylogenetic studies managed to divide the fungal kingdom into four major phyla, Chytridiomycota, Zygomycota, Ascomycota and Basidiomycota (*7*). However, a more recent attempt to resolve the basal groups

of fungi have revealed 5 new distinct sub-phyla within the previous Zygomycota phylum, which are Glomeromycota, Mucoromycotina, Entomophthoromycotina, Kickzellomycotina and Zoopagomycotina (*11*). With the sequencing of new species from the early diverging branches, more taxa can be added in future analysis, and improved resolution of the basal lineages are yet to come.

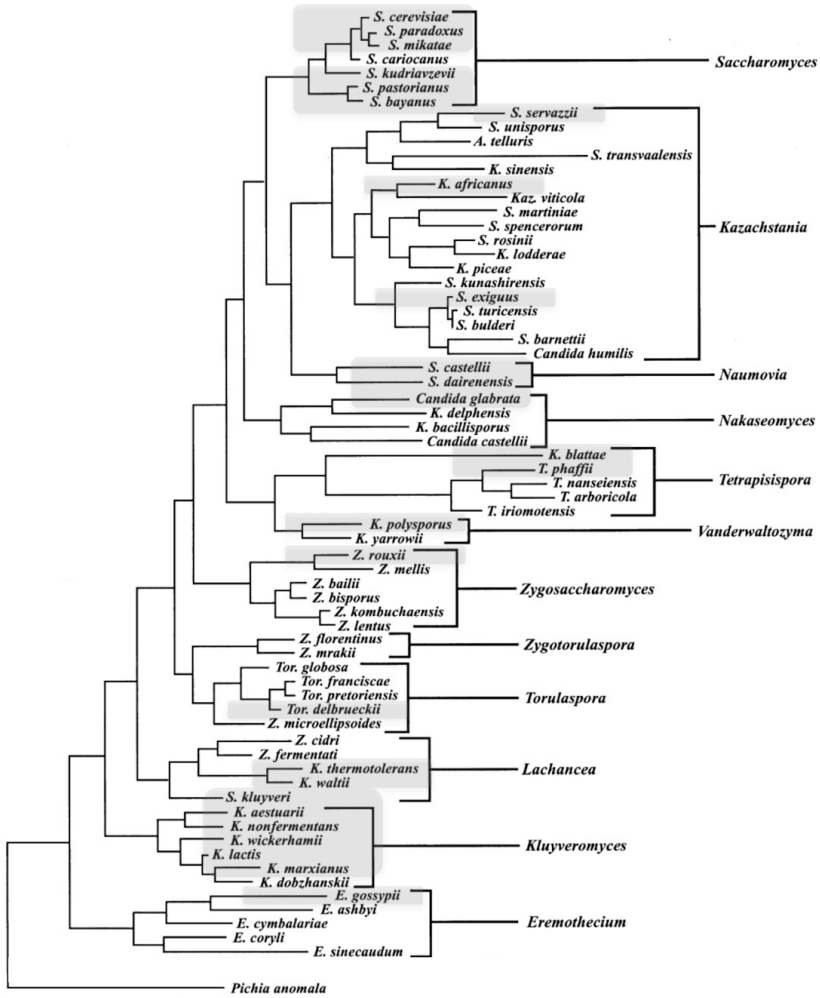
2.4. Fungal evolution and timing of the diverging branches

Taxa traditionally placed in the Chytridiomycota/Zygomycota-phyla represents the earliest diverging lineages of fungi, which proliferate in unicellular or filamentous form and produce flagellated cells that occur in terrestrial and aquatic habitats. Members of the Chytridiomycota phyla is believed to have changed little since the early history of eukaryotic evolution, and are thus considered to be extant representatives of the first fungi to appear. Most of the earliest-diverging lineages, such as members of the Chytridiomycota phyla are zoosporic true fungi, which suggest that the ancestor to the first fungi was evolved in the aquatic habitat (*12*).
Fossile data of fungi are rare, but there are evidence of lichen-like fossils that date more than 600 million years ago (Ma) (*13*) and arbuscular mycorrizae fossils that date more than 400 Ma. This suggests that fungi belonging to the Glomeromycota phylum, which have aerial spore dispersal and diverged after Chytridiomycota were among the first eukaryotes to colonize land (*5*). Unfortunately, fungal structures are poorly preserved, which open up the possibility of much later deposits of fungal fossils in the early history of each phylum (*14*). Molecular clock has i.e. estimated the divergence time of basal fungal lineages from other fungi up to 1400 Ma (*15*).

A majority of all described fungal species belong to the dikarya clade, which consists of the Basidiomycota and Ascomycota phylum. The oldest fossile record of ascomycetes dates back 400 Ma (*16*) whereas molecular clock dates the divergence of dikarya from the basal fungal lineages up to 1200 Ma (*15*). Whether the molecular clock predictions are accurate or not, it is clear from recent multigene concatenation and supertree phylogenetic analysis of whole genomes that the Glomeromycota, Basidiomycota and Ascomycota form a monophyletic tree (*12*) where dikarya has diverged from Glomeromycota.
Basidiomycota is the second largest fungal phyla (Figures 2, 3) and includes the Ustilagomycotina, Agariomycotina, and Pucciniomycotina subphylum (*6*). Ascomycota is by far the largest phylum amongst fungi (Figures 2, 3) and includes the Taphrinomycotina, Pezizomycotina and Saccharomycotina subphylum (*4, 6*). Among the different sub-phyla of Ascomycota yeast, special

focus has been on yeast species belonging to the Saccharomycetaceae family (Figure 4) such as *Saccharomyces cerevisae*, which has emerged as a model organism for several historical, academic as well as applied reasons.

Figure 4. The Saccharomycetacea family includes 12 phylogenetic clades, *Eremothecium, Kluyveromyces, Lachancea, Torulaspora, Zygotorulaspora, Zygosaccharomyces, Vandervaltozyma, Tetrapisispora, Nakaseomyces, Naumovozyma* (*Naumovia*), *Kazachstania* and *Saccharomyces*, which span an evolutionary history of more than 150 million years. Highlighted species have been sequenced and are publically available through databases such as NCBI (*17*), SGD (*18*), Génolevures (*19*) and YGOB (*20*). The phylogenetic tree was constructed from the alignment of concatenated sequences of three ribosomal RNA regions (18S, ITS, 26S) *EF-1α*, mitochondrial small-subunit rDNA and *COX II* [adapted from (*3*)].



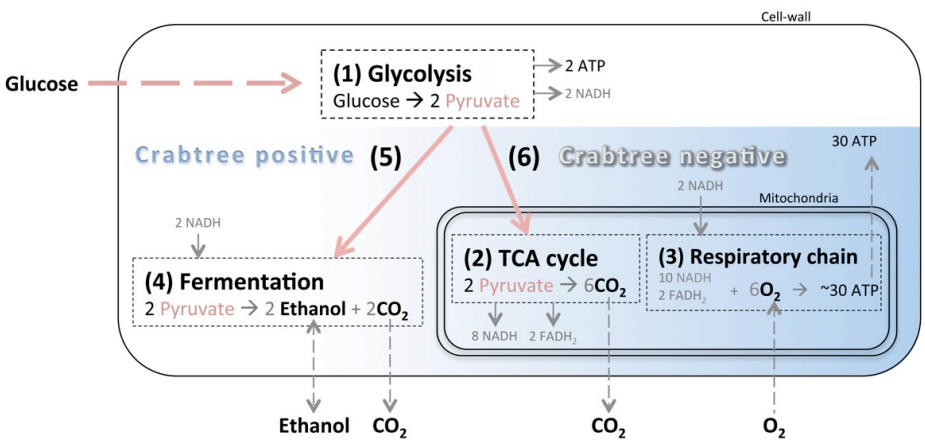
3. Emergence of yeast as a model organism

Yeast was probably domesticated during the Neolithic times and has for thousands of years been used for its fermenting properties for food preservation and alcohol production. Before Louis Pasteur discovered the role of brewers yeast in alcoholic fermentation, it was probably only known from its activity, and most manmade fermentations were spontaneous. The peculiar trait of the budding yeast *S. cerevisiae* to ferment even under aerobic conditions has made it the preferred organism for bread, alcoholic beverage and also lately for industrial bioethanol production (Figure 5).

3.1. Yeast physiology

Scientists have ever since its discovery tried to elucidate the underlying mechanism behind yeast physiology. In the beginning of the 20th century, observations that some yeast share physiology with certain tumour cells were made, which further increased its attention.

Figure 5. Yeast central carbon metabolism can be divided into several pathways; (1) glycolysis (2) TCA cycle (3) respiratory chain and (4) fermentation (see also box 1). Some yeasts such as *S. cerevisiae* predominantly utilises fermentation for the degradation of hexose sugars rather than respiration (TCA cycle and respiratory chain) even under aerobic conditions. This ability was originally attributed to glucose repression of respiration and named after its discoverer Herbert Crabtree. Thus, (5) Crabtree positive yeasts utilises less energy efficient pathway for ATP production, and "waste" energy by releasing ethanol into the surroundings. On the other hand, (6) Crabtree negative yeasts rather maximise their energy yield by respiration and the complete breakdown of hexoses into CO_2 and H_2O under aerobic conditions.



Box 1| Glossary of (yeast) central carbon metabolism

1) Glycolysis oxidises glucose 6-phosphate (C_6 compound originating from glucose) in a series of reaction to produce pyruvate (C_3 compound) energy (ATP) and reducing agents (NADH).

2) TCA cycle (Krebs cycle or the citric acid cycle) is a series of reactions that start with the condensation of acetyl CoA (originating from pyruvate) and oxaloacetate into citric acid. TCA cycle generates CO_2, precursors for anabolic reactions, and more reducing agents (NADH and $FADH_2$).

3) The respiratory chain utilises O_2 as the final electron acceptor in the oxidation of reducing agents, like NADH and $FADH_2$ to generate energy for cell-proliferation.

4) Fermentation is a pathway that does not require O_2 as a final electron acceptor for the re-oxidation of glycolytic NADH. Instead acetaldehyde is reduced into ethanol and NAD^+ can be recycled for glycolysis.

5) Crabtree positive yeasts predominantly ferment sugars (such as glucose) to ethanol, even under aerobic conditions.

6) Crabtree negative yeasts rather respire under aerobic conditions and can ferment ethanol only under O_2-limited conditions.

The importance of budding yeast in both industrial and academic research has accumulated an extensive knowledge about this eukaryote, which together with inherent properties such as, easy to culture and well-developed molecular tools has made it a good model organism.

3.2. The role of S. cerevisiae in the emergence of the genomics-era

In 1996, *S. cerevisiae* became the first eukaryote to have its genome fully sequenced (*21*). This was the beginning of the genomics era that enabled new *in silico* based research fields such as comparative genomics. The sequencing of *S. cerevisiae* and the extensive annotation of its genome have paved the way for subsequent genome sequencing projects. In fact, up to now there are more than 265 fungal genomes listed only on NCBI, which are publically available for comparative studies. Out of these 265 available fungal genomes, 191 belong to Ascomycota yeast, and 54 to the Saccharomycotina sub-phylum.

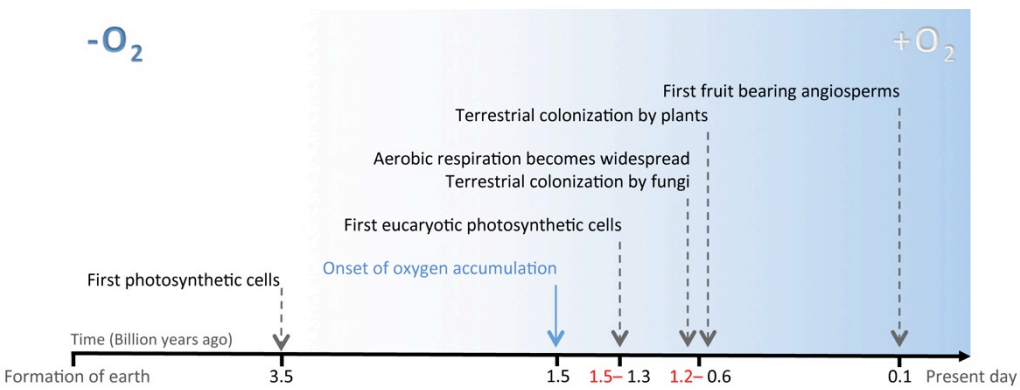
With the continuous development and improvement of new sequencing techniques, such as high-throughput sequencing, many more sequenced genomes are expected to become available in the next coming years.

4. Oxygen – its role in early eukaryote evolution of aerobes

O_2 play an essential role in energy metabolism of a majority of extant life forms. Different organisms, such as yeast can be classified as obligate anaerobes, facultative anaerobes or obligate aerobes depending on their need of O_2 to survive. Although a majority of modern organisms cannot proliferate without O_2, this has not always been the case.

It is believed that one of the earlier life forms to be evolved approximately 3.5 billion years ago (Ba), the progenitor to extant cyanobacteria and plant chloroplasts (*22*), had the ability to harness energy from sunlight for the production of organic chemicals, by incorporating inorganic CO_2. This water-splitting process, known today as photosynthesis, releases molecular O_2 and was the main reason for the reduction of CO_2 and accumulation of O_2 in the modern atmosphere today. The introduction of a more oxidative environment by ancient anaerobic eubacteria was most likely the major key event for the evolution of the modern life forms that possesses an alternative to the ancient anaerobic lifestyle – explaining the predominance of aerobic organisms today.

Figure 6. Dating of important evolutionary events that affected the accumulation of O_2 in the atmosphere – determined from the study of fossil records (in black) and from molecular clock estimations (in red). The first photosynthetic cells are believed to have evolved at least 3.5 Ba (*22*). The onset of O_2 accumulation in the atmosphere has been detected in deposits of iron oxide in rocks dating 1.5 Ba (*23*). The earliest fossile records of algae date more than 1.2 Ba (*24*) and the origin of the first plants and the colonization of land by one of the first terrestrial aerobes could have been as early as 1.2 Ba by the ancestor to extant fungus, which was followed by the first vascular plants, as early as 0.6 Ba (*5, 15, 25*). The first fruit bearing angiosperms dates to more than 0.1 Ba (*26*).



4.1. Proof of early photosynthetic bacteria and the consequence of their activities

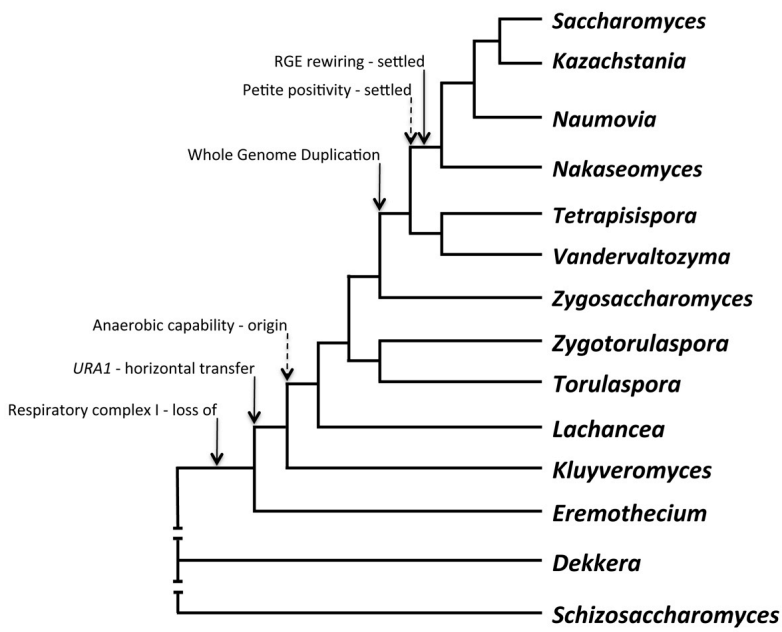
Due to the immense evolutionary time-span (billions of years), there is no direct neontology approach to infer the exact nature of the first life forms responsible for the above-mentioned events, but rather indirect paleontology approaches managed to determine their activity and therefore also the timing of the ancient evolutionary innovations that must have occurred. The origin of the first photosynthetic bacteria is for instance based on comparative examination of ca. 3.5 billion years (By) old stromatolite fossils and extant stromatolite consisting of large colonies of O_2-producing photosynthetic cyanobacteria (*22*). Moreover, the onset for O_2 accumulation in the atmosphere is believed to have caused the enormous deposits of iron oxide that can be found in 1.5 By-old rocks (*23*). The first photosynthetic eukaryotes (the progenitors to extant algae) are believed to have evolved more than 1.2 Ba (*24, 27*), and land would be colonized and transformed by probably the first terrestrial aerobes to evolve up to 1.2 Ba - the ancestor to extant arbuscular mycorrhizal fungus, which is thought to have paved the way for the first vascular plants by providing them with water and nutrients (phosphate and nitrogen) in exchange for carbohydrates (*5, 15, 20*). Not until ca. 125 Ba would the first angiosperms that dominate many environments today have been evolved (*26*).

5. Comparative genomics of fungi

5.1. A tool to reconstruct phylogenetic trees and identify evolutionary events important for the Crabtree effect

With the development of new sequencing techniques a readily increasing number of fully sequenced fungal genomes have become available, which has resulted in a new field of *in silico*-based approaches such as phylogenomics and comparative genomics. The combination of computational and paleontology approaches has contributed much to current knowledge about the structure of phylogenetic trees, such as the relationship between different species and clades, and also to the identification and timing of evolutionary events. Moreover, several comparative genomics studies of extant and experimentally evolved yeast strains have hypothesized the importance of some evolutionary events for the development of the fermentative lifestyle in yeasts that belong to the Ascomycota phyla (Box 2, Figures 4, 7).

Figure 7. The timing of evolutionary events, which are believed to be important for the evolution of aerobic fermentation and Crabtree effect in the Saccharomycetaceae family are illustrated (see also box 2). Evolutionary events, presented with broken arrows, have not been determined by *in silico* comparative genomics approaches. More detailed explanation for some evolutionary events, and others not yet timed is found in the following chapters (see also box 2).



Box 2| Summary of evolutionary events behind the respiro-fermentative lifestyle

1) The respiratory complex I is a part of the respiratory chain, and important for it's activity in the first step of NADH-oxidation for ATP production. All investigated extant yeast species belonging to the Saccharomycetacea and *Schizosaccharomyces* lineage lack the respiratory comlex I (*4, 28*).

2) A horizontal transfer of *URA1* (coding for a cytoplasmic dihydroorotate dehydrogenase) has uncoupled *de novo* pyrimidine synthesis from the respiratory chain, which is believed to have contributed to the anaerobic growth-capability in the Saccharomycetaceae lineage (*29, 30*).

3) A whole genome duplication event that occurred approximately 100 Ma is believed to have contributed to gross duplication of glycolytic genes and *trans*-acting factors, providing a genomic background necessary for the evolution of an efficient anaerobic fermentation capacity and the respiro-fermentative life-style in *Saccharomyces* yeasts (*31, 32*).

4) The rewiring of rapid growth elements (RGE) has been shown to be responsible for the uncoupled regulation of cytoplasmic and respiratory genes in yeast (*33*). The rewiring of RGE (AATTTT), which is reported as a *cis*-acting elements that has been lost in the respiratory genes are believed to have evolved independently in *Saccharomyces* yeasts (*32, 34*) and the *Dekkera* lineage (*35*).

6. Gene-dosage imbalance – an evolutionary driving force

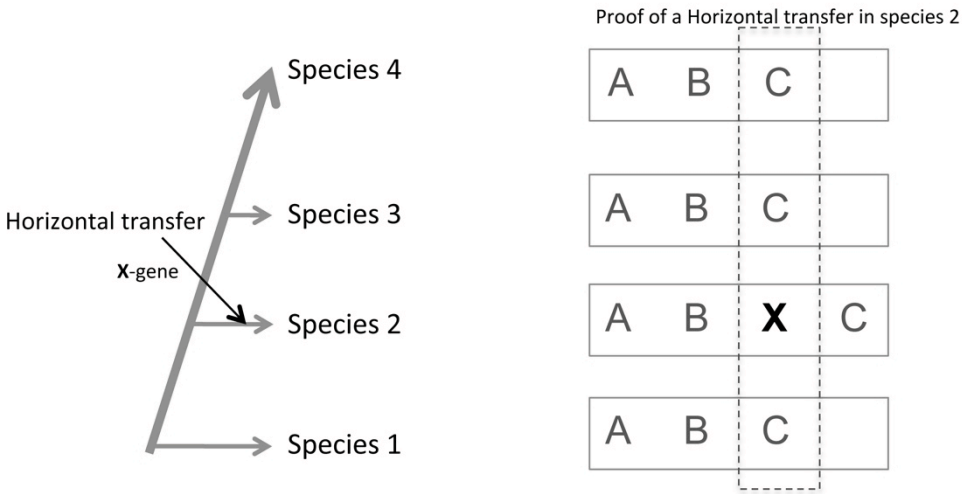
Small-scale gene duplications alter gene-dosage balances, which could be a mechanism that promotes rapid adaptation to changing niches. An organism that explores niches with scarce resources could i.e. increase its fitness through duplication (creating paralogs) or upregulation of endogenous genes important for rapid uptake to efficiently assimilate these resources. When exploring exotic niches, an organism might require the incorporation of exogenic material (either orthologs or exotic genes with new functions) for rapid adaptation. Likewise, gene duplication that causes redundancy could in theory result in subsequent loss of the less compatible gene-copy to reduce gene-dosage imbalances (*36, 37*).

Gene duplication could also confer to novel gene-functions by providing an "excess" of genetic material that can undergo evolution, and thereby avoid redundancy or gene-dosage imbalance issues. While small-scale duplications such as tandem gene-duplication and in some cases horizontal gene transfer (single gene duplications), segmental gene duplication (regional duplication of chromosomes) or aneuploidy (whole chromosome duplications) most likely result in negative gene-dosage imbalance, large-scale duplications (such as whole genome duplications) could in theory help overcome these problems by duplicating whole protein-protein complexes or pathways (*38*).

6.1. Horizontal gene transfer

Lateral gene transfer or horizontal gene transfer (HGT) is an event that involves the transfer of genetic information across mating barriers between different species, as compared to normal vertical transfer from parent to offspring (*39*). In bacteria, HGT is known to be an important mechanism for adaptation by providing the recipient cell with exogenic content that might increase its fitness when exploring new ecological niches. There are generally three possible outcomes of HGT, observed by comparative genome analysis of extant organisms, which are (I) acquisition of genes with new functions different from the endogenous genes (II) acquisition of paralogs to endogenous genes (III) acquisition and replacement of orthologous genes with better functions.

Figure 8. This figure illustrates an approach of comparative genomics for the detection of horizontal gene transfers. This stringent approach is however limited to inter-kingdom transfers only, such as from bacteria to yeast.



6.1.1. Importance of HGT in the development of alternative lifestyles

HGT has played an important role in early eukaryote evolution. Examples of some of the earliest known HGT were the endosymbiosis of the progenitors to all mitochondria, and the subsequent integration of their genomes (*40, 41*). The transition from a CO_2 rich atmosphere to a more oxidative one, most likely favoured the evolution of new types of more energy efficient organisms that could harness the oxidative potential of O_2 as a terminal electron acceptor. Thus, the incorporation and invention of a mitochondrial organelle in the obligate anaerobic progenitor of eukaryotes most likely provided the genetic background for the evolution of the predominant modern aerobes that exist today.

A more recent event, which involves a horizontal transfer of a gene between bacteria and yeast, *URA1*, was first postulated by Nagy and colleagues (*42*) and later confirmed by others (*29, 43*). This gene encodes the enzyme dihydroorotate dehydrogenase (DHODase), which is responsible for the only redox reaction involved in the *de novo* pyrimidine biosynthesis. DHODase has been shown to occupy different subcellular compartmentations in different kingdoms in the tree of life. Prokaryotic DHODases are either associated with membranes, i.e. as in *Escherichia coli* (*44*) or soluble as in *Lactococcus lactis*

(*45*). Moreover, cytoplasmic DHODase versions are more common amongst intracellular anaerobic parasites as compared to free-living unicellular organisms [(*46*) and references therein]. All investigated mammalians up to now contain only mitochondrial-associated versions (*47-49*). In yeast, both forms have been reported, where the cytoplasmic DHODase (Ura1p) seem to be overrepresented amongst facultative anaerobes such as *S. cerevisiae* and it's sister-species (*30, 43*).

6.1.2. Ura1p promotes growth in pyrimidine and oxygen depleted environments

The ability of *S. cerevisiae* to grow anaerobically on synthetic complete media supplemented with anaerobic growth factors but without added uracil is due to the possession of a cytoplasmic version of DHODase (*42*). This ability was most likely acquired from a HGT of *URA1* from a *Lactococcus lactis* after the divergence of the Saccharomycetaceae yeast from *C. albicans* (*43*). Yeast that diverged from the *Saccharomyces* lineage before *E. gossypii* (Figure 4), such as *S. pombe*, *C. albicans* and *K. lactis* have been shown to lack the ability to synthesize pyrimidines under strict anaerobic conditions (*35, 42*) and possess only the mitochondrial associated DHODase (Ura9p). Ura9p requires quinone in the respiratory chain as an electron acceptor for its function, while Ura1p can utilize cytoplasmic fumarate as an electron acceptor (*29*). Thus, yeast that possesses Ura1p does not require the respiratory chain for *de novo* pyrimidine biosynthesis and can proliferate under O_2 and pyrimidine limited conditions. Species belonging to the *Lachancea* clade such as *L. kluyveri* and *L. waltii* can be considered as an intermediate form in the transition from an aerobic to facultative anaerobic yeast, and they possess both *URA1* and *URA9* homologs (*29, 43*). Most species that diverged after the *Kluyveromyces/Lachancea* lineage have been reported as facultative anaerobes (*30*) and a majority of investigated species that diverged after the *Zygosaccharomyces* clade have lost their *URA9* gene (Figure 7) (*50*).

6.1.3. Identification of HGT

HGT has up to now been regarded as rare events in yeast, but this might be proved differently. The identification of HGT has often been restricted to genes with bacterial origin and usually starts with a blast for genes with best hit to bacteria, followed by comparative analysis between different yeast-genomes. If a putative HGT is found in more than one genome it has been discarded (*43, 51*). A slightly modified strategy has taken the advantage of conserved synteni amongst different species for the identification of species-specific genes within

these regions (*39*). Up to now, only 16 HGT candidates have been identified by these approaches. These stringent strategies most likely create a bias towards the identification of more recent HGT in yeast, and could therefore dismiss many transfer events that are more ancient or between different yeast species. Moreover, HGT among fungi is not easy to study, which might explain the low numbers identified up to now. Due to the fact that it took almost a decade since the sequencing of the *S. cerevisiae* before a large-scale search for HGT in its genome (*43*) was undertaken, many more will probably be identified in the near future. Whether the low numbers of identified HGT in yeast reflects the reality or if it is just a bias resulting from difficulties and the limitations that comes with the identification strategies, HGT has clearly played an important role in the early evolution of eukaryotes.

6.2. High-affinity glucose transporters promotes certain niche adaptation

Yeast such as *S. cerevisiae* and it's sister species are often associated with vineyards and vineries (*52*). Although they occur at extremely low populations on healthy, undamaged grapes and therefore in the beginning of spontaneous grape-juice fermentation, they often dominate the later stages of fermentation (*53*). There are three ways that sugars can enter a cell: (I) by net diffusion (II) by facilitated diffusion and (III) by active and energy-dependent transport. Freshly crushed grape juice contains large excess of sugars (>20% w/v), which is initially enough for net diffusion to occur down the concentration gradient into the cell. These sugars can be metabolized in the glycolytic pathway to pyruvate and further into acetaldehyde (Section 8.4, Figure 14). The reduction of acetaldehyde results in ethanol that can exit the cytoplasm where all reactions occur, through the plasma membrane. However, since yeast plasma membrane is not freely permeable by large polar molecules, other ways for an efficient glucose uptake are necessary when the concentration drops.

In *S. cerevisiae*, members of the hexose transporter family are responsible for the transport of sugars such as glucose and fructose through the plasma membrane. The hexose transporter family of *S. cerevisiae* is large and consists of 20 genes that encode high, intermediate and low affinity transporters. These transporters are specific for different sugars and concentrations, and it can be assumed that the process of glucose uptake could be the first rate-limiting step for the carbon flux via glycolysis. Indeed this has been confirmed by experimental evolution approaches. Several genome rearrangements have been observed in experimentally evolved yeast strains under glucose-limited conditions (*54*) that resulted in the amplification of high-affinity hexose transporters, *HXT6* and *HXT7*. One case of local amplification of a chimeric

gene consisting of the coding region of *HXT6* and the promoter region of *HXT7* was identified on the right arm of chromosome 4 (*55*). Two independent cases of amplification of the same right arm of chromosome 4 were also confirmed (*56*).

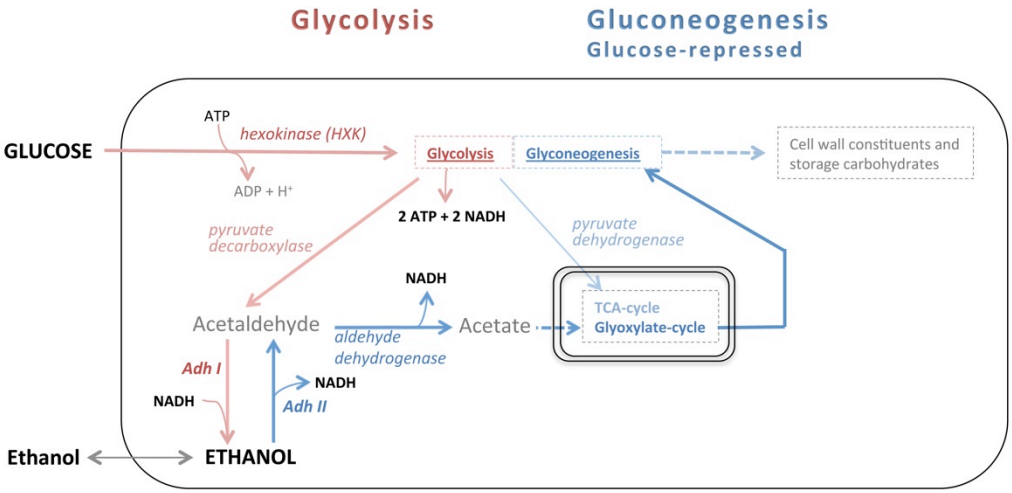
Since hexose-transporters are directly involved in the uptake of glucose, the amplification of high-affinity hexose transporters could be a relatively fast mechanism to increase the glucose uptake in the evolved strains. An increased gene-dosage of hexose transporters could in other words have contributed to an increased fitness of evolved strains in glucose limited conditions, which had improved their affinity for glucose.

6.3. ADH2 is important for a well-regulated fermentative lifestyle

The Crabtree effect is responsible for the fundamental physiology of *S. cerevisiae*, which is to make and accumulate ethanol when grown on glucose, even under aerobic conditions. Later, when glucose is depleted and if O_2 is still present, ethanol will be consumed. This observation has led to the postulation of a "make-accumulate-consume" strategy (*57*), which is a modern interpretation of the Crabtree effect that incorporates it in an ecological context. It was hypothesized that before the duplication of an ancient alcohol dehydrogenase (ADH_A) that resulted in two paralogs (ADH_1 and ADH_2), yeast did not accumulate ethanol for later consumption (*32, 57*). Instead ADH_A would have been used to recycle NADH, produced from anaerobic glycolysis (Figure 9).

The *S. cerevisiae* ADH_1 is constitutively expressed and involved in the production of ethanol (*58*). The other paralog, ADH_2 is under glucose repression (*59*) and has a Michaelis constant (K_M) 26 times lower than ADH_1, consistent with its role in ethanol consumption after glucose depletion (Figure 9). A "resurrection" of ADH_A with the help of site-directed mutagenesis, cloning, expression and subsequent kinetic analysis confirmed that ADH_A most likely had elevated K_M values for ethanol, similar to ADH_1 in *S. cerevisiae*, *K. lactis* and *S. pombe* (*57, 60*). Thus, Thompson and colleagues concluded that ADH_2 had diverged from its original role and evolved to consume ethanol, while its paralog ADH_1 has retained the ancient function since the duplication event approximately 80 Mya (*57*).

Figure 9. When yeast proliferates on glucose, it can recycle NADH that was produced from glycolysis by the production of ethanol in the fermentative pathway. Crabtree negative yeasts such as *K. lactis* would normally only ferment ethanol when there is no O_2 available for the oxidation of NADH in the mitochondria. However, Crabtree positive yeasts such as *S. cerevisiae* has evolved the ability to ferment even in presence of O_2, thus it prefer ethanol fermentation for the recycling of its glycolytic NADH.



6.4. The underlying mechanism behind small-scale duplications

During the last decades, experimental evolution of *S. cerevisiae* has begun to shed light on the underlying mechanisms behind adaptive responses, due to selective pressure from environments. Early laboratory evolution experiments demonstrated that asexual aerobic propagation of isogenic haploid and diploid clones under glucose-limited conditions (for 250-450 generations) resulted in more rapidly evolved diploid strains as compared to haploids (*54*). Moreover, the evolved strains would usually have an increased (two to eight times higher) glucose assimilation rate as compared to the non-evolved clones from which they were derived (*55*).

The use of Southern blot and comparative genomics hybridization with microarrays has proved genome-rearrangements as the preferred mechanism employed by yeast to rapidly increase its capacity to assimilate a limiting substrate (*56*). Small-scale alterations in gene-dosage are results from local gene duplications such as tandem duplication (*55*), segmental aneuploidy due to intra- and inter-chromosomal translocations, and whole chromosome

aneuploidy (*56*). It was further shown that all rearrangements were flanked with Ty-elements, transposons, transposon fragments or tRNAs.

Not much is known about the underlying mechanisms and factors behind the observed rearrangements that have resulted in gene amplification in yeast. However, short inverted repeats appears to play an important role in hairpin formations upon fusion of double stranded DNA, either through 5'-3' resection by exonucleases after a double strand break or through cleavage by Holliday junction resolvase (*61*). These hairpin-capped double strand breaks can be resolved by the Mre11-Rad50-Xrs2 complex along with Sae2p in yeast (*62*). This creates open ends, which are prone to recombination through break induced replication and strand invasion that could lead to altered gene-dosage and aneuploidy.

6.5. Small-scale duplications as a possible tool for rewiring of cis-elements

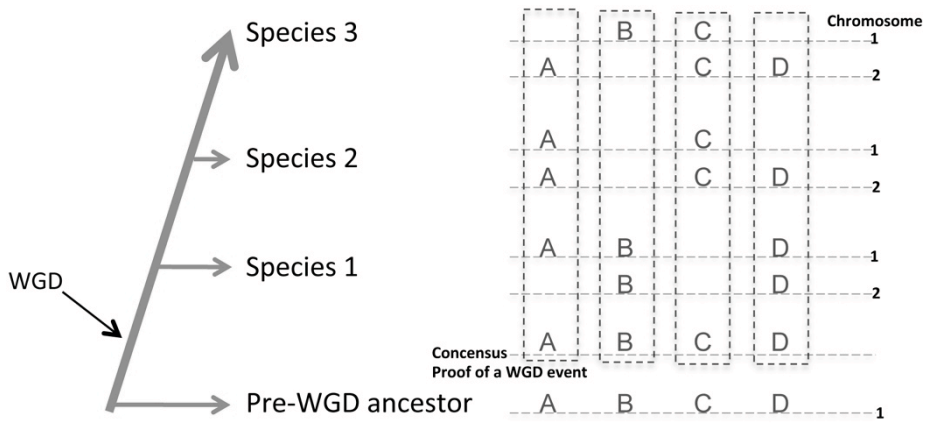
Transcriptome analysis of three independently evolved *S. cerevisiae* strains (for 250 and 500 generations) in continuous glucose limited aerobic condition, compared against a common parental *S. cerevisiae* strain provided a plausible explanation for the shift from fermentative to a more oxidative metabolism (*63*). All three evolved strains exhibited increased expression of respiratory genes such as, mitochondrial associated gene products involved in the Krebs cycle and the electron transport chain as compared to the parental strain. All evolved strains also exhibited a decreased expression of a majority of genes involved in glycolysis and gluconeogenesis as compared to the parental strain.

Up to now, it can only be speculated of what might have caused the altered gene-regulation, responsible for the observed transition from respiro-fermentative to a more oxidative metabolism in the evolved strains. One hypothesis is that a change in the *cis*-element of citrate synthase (*CIT1*), which encodes a product important for the key enzymatic step in Krebs cycle, might have been responsible for the up-regulation of this pathway in the evolved strains (Figures 5 and 15). Many of the breakpoints where recombination occur, are in repetitive sequences. One example of an inverse repeat is a transposon, found in the promoter region of *CIT1*. It was postulated that the observed independent breakpoint in that promoter region could have recruited an active cis-element of a transposon, resulting in derepression of this key regulated point in the Krebs cycle, even in the presence of glucose (*56*).

7. Whole genome duplication

With the increasing number of available genomes, research is not anymore limited to the study of evolution and function of single genes only. During the last decades, more focus has been on the evolution and function of genomes as whole, which resulted in the discovery of large-scale duplications. It is becoming more apparent that not only has small-scale gene duplications played an important role in the shaping of genomes during the course of evolution, but so has whole genome duplications (WGD).

Figure 10. This figure illustrates how comparative genomics can detect WGD-events by comparing several lineages and based on synteny, reconstruct an ancestral consensus gene-order. This gene-order should share synteny with pre-WGD ancestral species.



A WGD event was first identified in yeast. It was postulated to have occurred approximately 100 Ma in the *Saccharomyces* lineage (*31*). The WGD event divides the *Saccharomyces* lineage into two major groups: (I) species that diverged before the WGD event (pre-WGD species) and (II) species that diverged after the WGD event (WGD species). By searching for homologous genes based on amino-acid sequence on different chromosomes, and with the help of synteny between duplicated gene pairs in the WGD specie, a set of duplicated regions could be identified in *S. cerevisiae*. Wolfe and colleagues found that the amount of identified segmental duplications were significantly higher than what would be expected if the homologous genes were randomly distributed, and when they compared gene sequence and gene order against orthologs in *K. lactis*, a WGD event was concluded.

The WGD hypothesis was later confirmed by direct comparison of the *S. cerevisiae* genome to fully sequenced genomes of pre-WGD yeast species that diverged from a common progenitor approximately 150 Mya, such as *L. waltii* (*64*) and *E. gossyppii* (*65*). Not much is known about the mechanism behind the WGD events, and so far no laboratory evolution experiment has been able to capture this event. It is however believed that WGD could be the result from different mechanisms, such as hybridization events that causes allopolyploidy or spontaneous genome doubling events that causes autopolyploidy.

Up to now, several ancient WGD events have been described in most kingdom in the tree of life such as; plants (*66-70*), fungi and yeast (*31, 64, 65, 71*), in the early ancestor of vertebrates (*72*), teleost fishes (*73, 74*) and the frog *Xenopus laevis* (*75*).

7.1. Possible outcome of WGD-events

WGD could have contributed to increased biodiversity in certain lineages, since reciprocal gene-loss adversely affects positive outcomes from between population mating events. As an example, the identified whole genome duplications that were followed by an extensive gene-loss have provided an explanation for the radiation of species observed in yeast (*37*) and teleost fishes (*75*).

Moreover, WGD event is believed to provide an important source of evolutionary novelty. It was shown that shortly after the WGD event in yeast, most duplicated genes would eventually be deleted, either as large segmental deletions or single gene deletions (*64*). Some duplicates might be retained due to (I) subfunctionalization or (II) neofunctionalization, where one paralog maintain its essential function and the other copy evolves a novel function (*37*). Gene-fixation in the WGD-yeasts is estimated to almost 10 percent, which means that more than 90 percent the duplicated genes were lost since the WGD event in all investigated WGD-species (*37, 64, 65*).

The outcome of WGD events involves gross alteration in gene-dosage that could, similar to small-scale duplication, directly affect the amount of gene products. The abundance of gene products could also be altered through changes in gene-regulation when protein domains of i.e. trans-factors are modified in the process of subfunctionalization or neofunctionalization. There are limited amounts of WGD genes characterized up to now, and most studies have been conducted in yeast. Most likely, all outcomes of small-scale gene duplication events could occur after a WGD followed by rapid gene-loss, but

different outcomes could have played more important roles for the evolution of certain lineages and lifestyles.

7.2. Gene-dosage changes as a tool for development of alternative lifestyles

The concept of genetic balance was set in the early days of genetics when it was shown that aneuploidy in plants had greater deleterious impact on the phenotype than whole-genome changes [reviewed in (*76*)]. According to the gene balance hypothesis, imbalances between components of a protein-protein complex are most likely to be deleterious if not compensated for, by i.e. altered gene-regulation to uphold the balance between gene-products (*36*).

Changed ratios of components in a protein complex could be due to altered gene-copy numbers through duplications, and thereby changing the net gene-product amount. Other changes could involve altered *cis*-elements or enhancer elements of trans-factors, or of the gene encoding the protein-component itself. When analyzing the fitness of diploid heterozygotes, carrying knockouts of single essential genes in *S. cerevisiae*, it was concluded that fitness was negatively correlated with the number of protein-interactions of a deleted gene product (*36*). It was for instance revealed that imbalance between co-expressed interacting proteins decreased fitness in *S. cerevisiae*, and that members of large gene-families rarely constitute complexes (*36*).

Interestingly, duplication of single chromosomes in haploid yeast due to missegregation that results in aneuploidy, favours the accumulation of mutations that increases whole chromosome-instability in *S. cerevisiae*. It was shown that the majority of aneuploidy strains (haploid or triploid) relied on translesion synthesis activity of REV3p (the catalytic subunit of Pol ζ) to compensate for defects in recombinatorial and DNA repair (*77*). Furthermore, non-deleterious changes in gene-dosage such as duplication of whole pathways could increase the fitness of an organism when exploiting new niches and resources. Compared to small-scale duplications, WGD would thus provide the opportunity to duplicate complete protein-protein complexes and metabolic pathways instantaneously, without encountering protein-protein-imbalance issues.

An example is the outcome of a WGD in the ancestor to *Rhizopus oryzae*, which is a multicellular fungus that belongs to the basal fungal lineage (Figure 1) that diverged up to 1200 Ma (*15*). The WGD in the ancestor to *R. oryzae* resulted in the retention of all duplicated protein complexes in the respiratory chain, which is hypothesized to have contributed to the rapid growth of this

organism (*71*). Another example is the WGD in the *Saccharomyces*-lineage, which resulted in duplication of trans-factors such as TOR and an increased copy number of glycolytic genes (*31*). These duplications were hypothesized to have been important for the necessary genomic background required for the major change in gene-regulation and metabolism for the development of efficient anaerobic sugar fermentation (*31*) and even under aerobic conditions (*32, 38*).

7.3. Rewiring of cis-elements: a mechanism for development of new phenotypes

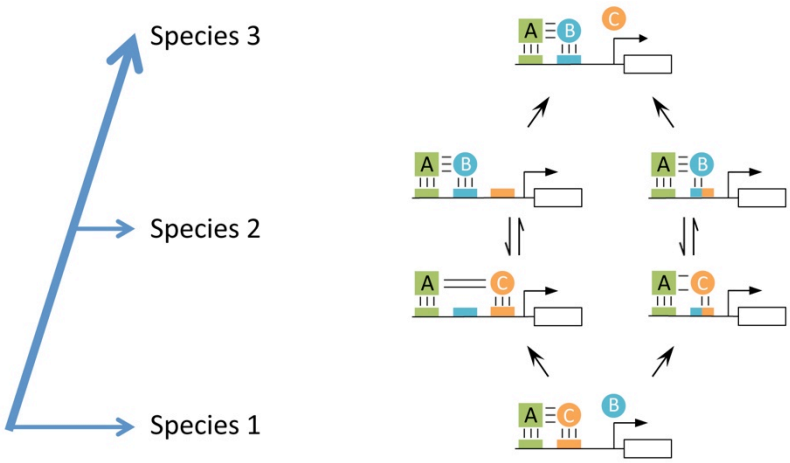
The consequence of altered gene-dosage can be changed levels of gene-products, which could result in the development of new phenotypes in organisms with increased fitness in certain habitats. However, changes in phenotypes can also be a consequence of altered gene expression-levels. It is known that a cell's activity depends on a network of interacting modules that consists of co-regulated genes whose expression are controlled by transcription factors (TFs) (*78*). TFs often bind cooperatively to cis-acting elements (generally 5 to 10 nucleotides long) with varying positions, relative to the gene whose transcription they control. This type of control can also be seen as a combinatorial control, where small alterations in TF composition can result in large changes in the network of interacting modules of co-regulated genes.

What is common for all cells is that their growth is highly coupled to the cellular energy status and the abundance of ribosomes, which act as the cellular protein-producing machinery that drives biomass production. In fact, the amount of ribosomal RNA in rapidly growing yeast cells accounts for nearly 80% of total cellular nucleic acid and almost 60% of RNA polymerase I, and 50% of RNA polymerase II transcription is devoted to ribosomal biogenesis (*79*). It is therefore not strange that the ribosomal module is strictly regulated. Although orthologous transcriptional modules of co-regulated genes are usually conserved across different domains of life, from bacteria to humans (*80, 81*), their transcriptional regulators are not. This has for example been observed in the shift of TFs that control the ribosomal module between different Ascomycota fungi (*82, 83*).

Not much is known about the mechanism behind the observed recruitment of new TFs and the loss of ancient TFs from modules of co-regulated genes (such as the ribosomal genes) but rewiring of cis-acting elements has emerged as a potential underlying mechanism (Figure 11) (*79*). Moreover, the rewiring of RGE (*33*) (box 2) is believed to have occurred through the split of an ancient ribosomal module into two functional sub-modules, comprised by cytoplasmic

and mitochondrial ribosomal-genes (*84*). This however, appears controversially since no TF has been identified that bind to the RGEs up to now (*85*). Instead the loss of A/T-rich sequences (such as RGE) in the mitochondrial ribosomal, and respiratory gene-promoters could have enabled the recruitment of nucleosomes to these genomic regions, which would provide an alternative explanation to the down regulation of respiratory genes observed in the WGD yeasts (*85, 86*).

Figure 11. Different pathways for rewiring modules of co-regulated genes are suggested. Small black lines represent protein-protein (between different TFs) and protein-DNA (between TFs and cis-elements) interactions. Both pathways contain an intermediate stage where factor B and C may act redundantly in an intermediate species. The combination A-B has evolved from an ancient combination A-C, but the modules of co-regulated genes may remain unchanged [modified from (*87*)].



8. Yeast central carbon metabolism

8.1. The Crabtree effect share a common origin with glucose repression

In 1861 Louis Pasteur discovered a quantitative interdependence between fermentative and oxidative pathways in the brewer's yeast. He could show that yeast had many times higher growth yield, per gram of consumed sugar, under aerobic as compared to anaerobic conditions. This discovery has kept many biochemists occupied since then and has proven significant for understanding the biochemistry of various types of cells that can proliferate both under aerobic and anaerobic conditions. Of all the studies made during the first half of the twentieth century, some observations had great impact on the understanding of the interdependence between fermentation and respiration. In the early 1920, when Otto Warburg investigated the Pasteur effect in certain rat tumour cells he could conclude that (I) cancer cells had the same rate of O_2 consumption as compared to normal cells and (II) cancer cells had higher rate of lactate production, even in the presence of O_2.

Warburg also observed an increase of glycolysis upon addition of ethyl isocyanide that inhibits cellular respiration, and he would later hypothesize the origin of cancer with the irreversible damage of respiration in certain "cellular granula", known today as mitochondria (*88*). As a follow up to Warburgs finding that certain tumour cells had an increased glycolysis as compared to normal cells, Herbert Crabtree could in 1929 observe that certain tumour cells had a repressed respiration as a response to glucose, even under aerobic conditions. Up to now, the underlying mechanism behind this physiology is still not known in all details, but it was originally believed to occur mainly due to imbalanced ATP/ADP ratio between mitochondria and cytoplasm [reviewed in (*89*)].

When De Deken in 1966 observed similar physiology in several yeast species (*90*), the term "Crabtree effect" was introduced in yeast research after it's discoverer, and has ever since often been used as a synonym to "glucose effect" or "glucose/catabolite repression". Glucose repression was later shown to mainly act on the transcriptional level by downregulating genes in pathways associated with respiration (*91*), gluconeogenesis (*92*) or metabolism of alternative carbon substrates (*93*).

8.1.1. Crabtree effect – Long-term

The fundamental characteristic of *S. cerevisiae* is its utilization of the less energy efficient glycolytic pathway as compared to respiration for rapid growth, even in the presence of O_2 (*90*). Glucose repression of respiration has, since its discovery in tumour cells (*94*) provided a plausible mechanism for this peculiar trait in yeast (*90*). The repression of respiratory associated genes, attributed to glucose repression in *S. cerevisiae*, is also known as "long-term regulation", "long-term Crabtree effect" or just Crabtree effect. Since glucose repression involves the relatively slow regulation on transcriptional level as compared to posttranslational regulation, it can be investigated in bioreactors during batch growth (Box 3, Figure 12).

Box 3 | Glossary for experimental setup of bioreactors

Batch culture: Yeast can be grown under controlled conditions in bioreactors. If no fresh growth-media is added to the culture, substrate will deplete, and product formation and growth cease.

Chemostat: Upon continuous inflow of fresh growth-media to a well-mixed culture, and an outflow equal to the inflow rate (dilution rate) of biomass and other formed products, a steady-state condition can normally be obtained within 5 residence times.

Dilution rate: Upon addition of fresh growth-media to a chemostat, dilution occurs at a rate that corresponds to the inflow rate (volume/hour) divided by the culture volume and has the unit (1/h).

Residence time: Also known as holding time, which is the reciprocal of the dilution rate.

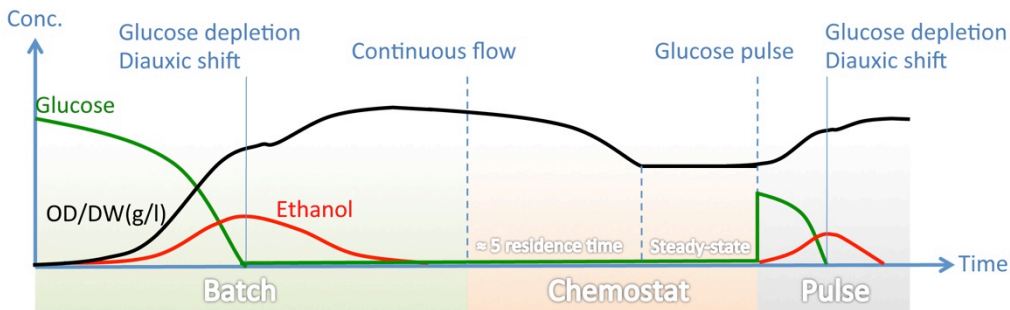
Steady-state: At unchanged culture conditions - when the culture volume, substrates, products and biomass concentration remain constant, steady-state is achieved.

8.1.2. Crabtree effect – Short-term

With the development of continuous culture methodology, a contradicting "short-term regulation" was discovered, which is believed to be the result from overflow metabolism of the respiratory pathway in yeast. It was shown that fully respiring continuously growing *S. cerevisiae* cells under derepressed

conditions would have an immediate production and accumulation of ethanol upon sudden addition of glucose to the media. This response is faster than any transcriptional repression of respiratory associated genes and is hypothesized to depend on posttranslational regulation of pyruvate dehydrogenase and pyruvate decarboxylase activities (*95*). Since overflow metabolism involves fast response time, attributed to current enzymatic properties (on posttranslational level) upon a glucose pulse to fully respiring steady-state cell cultures, a chemostat approach is required to capture these types of mechanisms (Box 3 and Figure 12). Despite many years of study, there is still hitherto no consensus to be found in the literature on what causes the onset of the Crabtree effect, a long-term glucose associated repression of respiratory genes or a short-term limiting capacity of the respiratory pathway.

Figure 12. Different experimental setup is required for the study of long- and short-term Crabtree effect in yeast (see also box 3). The aerobic batch culture approach is sufficient to quantify the long-term Crabtree effect. Crabtree positive yeasts make and accumulate ethanol during growth on glucose, and ethanol is later consumed upon depletion of glucose and diauxic shift. On the other hand, Crabtree negative yeasts only produce biomass, and do not make ethanol under aerobic conditions. For the quantification of short-term Crabtree effect, a glucose pulse approach of a steady-state culture is required. Yeasts that exhibit overflow of respiratory metabolism start to produce ethanol immediately (detectable within 20 minutes) upon glucose addition.



8.2. Different mechanism of sugar uptake in yeast

Before a cell can assimilate hexoses such as glucose and fructose, it needs to be transported across the plasma membrane. Hexose uptake in yeast is mainly mediated by hexose transporters (*HXT*), which depend on two different mechanisms, carrier-mediated facilitated diffusion systems and active proton sugar symport systems. The facilitated diffusion is energy independent and transport hexoses such as glucose and fructose down a concentration gradient, while the active symport system is energy dependent (*96*) and is probably

under the control of glucose repression in several yeast (*97*). It has been shown that Crabtree positive yeasts such as *S. cerevisiae* possesses only hexose transporters involved in facilitated uptake while Crabtree negative yeasts such as *Kluyveromyces* can possess both types of transporters (*98-101*).

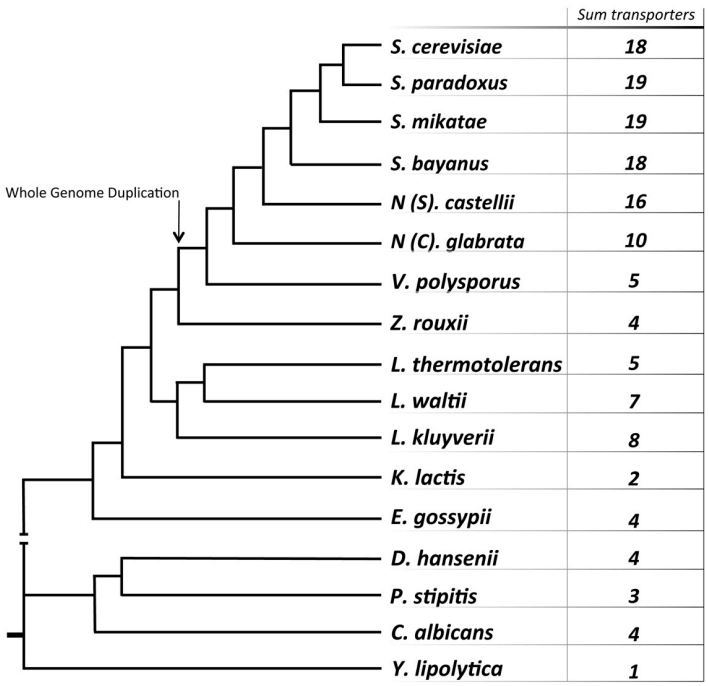
8.2.1. Hexose transporters and their role in sensing glucose

There are in total 20 putative hexose transporters identified *S. cerevisiae HXT1-17*, *GAL2, SNF3* and *RGT2* of which 17 are involved in glucose transport with high, intermediate and low glucose affinity (*98, 102*). Snf3p is glucose repressed and also act as a glucose sensor that facilitate glucose repression of high affinity Hxt6p and Hxt7p or induce expression of high/intermediate affinity Hxt2p and Hxt4p. Rgt2p is constitutively expressed and also act as a glucose sensor that together with Rgt1p induce expression of the low affinity Hxt1p. It is apparent that low affinity transporters in *S. cerevisiae* are responsible for glucose uptake at high concentrations, while high affinity transporters are derepressed under low glucose concentrations and are therefore responsible for the uptake under these conditions.

8.2.2. Evolution of hexose transporters - their role in the respiro-fermentative lifestyle

Two distinct families of sugar transporters have been identified in yeast. The Crabtree negative yeast *K. lactis* possess *RAG1* that belong to the *HXT* family and encodes a low affinity glucose transporter, which is induced by high concentration of glucose (*100*). The other glucose transporter in *K. lactis*, *HGT1* only share 27% sequence identity with *RAG1*, and belongs to a family of glucose transporter that separated from the *HXT* gene family early in hemiascomycetes evolution (*98*). The *HGT1* gene encodes a high-affinity glucose transporter, which is responsible for glucose uptake at low glucose concentrations by *K. lactis* (*99*). It has been postulated that high affinity (*HGT*) and low affinity (*HXT*) hexose transporters were present in the common ancestor of budding yeasts, and that low affinity transporters were retained and rapidly expanded in the adaptation of respiratory yeast to glucose-limited environment. However, respiro-fermentative yeasts seem to have lost *HGT* genes in the adaptation to glucose-rich environment, and instead rapid duplications of *HXT* genes have occurred (Figure 13) (*98*).

Figure 13. This figure illustrates the expansion of glucose transporters in Saccharomycotina yeast. It is appears as if the expansion of hexose transporters could have played an important role in the evolution of the *Saccharomyces sensu stricto* group [adapted from (*3, 4, 98*)].



8.3. Hexokinases – the first step of glycolysis

In *S. cerevisiae* there are 3 hexokinases, *HXK1*, *HXK2* and *GLK1* that can phosphorylate glucose at C6 after its uptake. The phosphorylation of glucose adds a charge on the sugar-molecule, which prevents its diffusion out of the cell and is the first step that dedicates it to central carbon metabolism such as glycolysis (Figure 14). The idea that glucose phosphorylation activity could trigger glucose repression is appealing, but was early rejected when *HXK1/HXK2* double knockout strains with overexpressed *GLK1* was shown to be insensitive to glucose, even if the phosphorylating activity increased three-fold (*103*). Amongst the three kinases in *S. cerevisiae*, Hxk2p has been considered to be the important enzyme for glucose repression for many years (*104, 105*). The discovery that Hxk2p is transported into the nucleus (*106*) supported the idea of a dual kinase and regulatory activity of this enzyme, both important for triggering glucose repression (*107*).

More recently, it has been shown that Hxk2p directly interacts with Mig1p in the nucleus and that this binding is required for its retention in there (*108*). This also suggests that the main role of Hxk2p in glucose repression is to maintain the repressing capacity of Mig1p by binding and thereby blocking Mig1p-phosphorylation by Snf1p (*108*). Although *HXK2* play an important role in glucose repression, it has lately been shown not to be required for glucose to affect the transcription of many genes [reviewed in (*104*)].

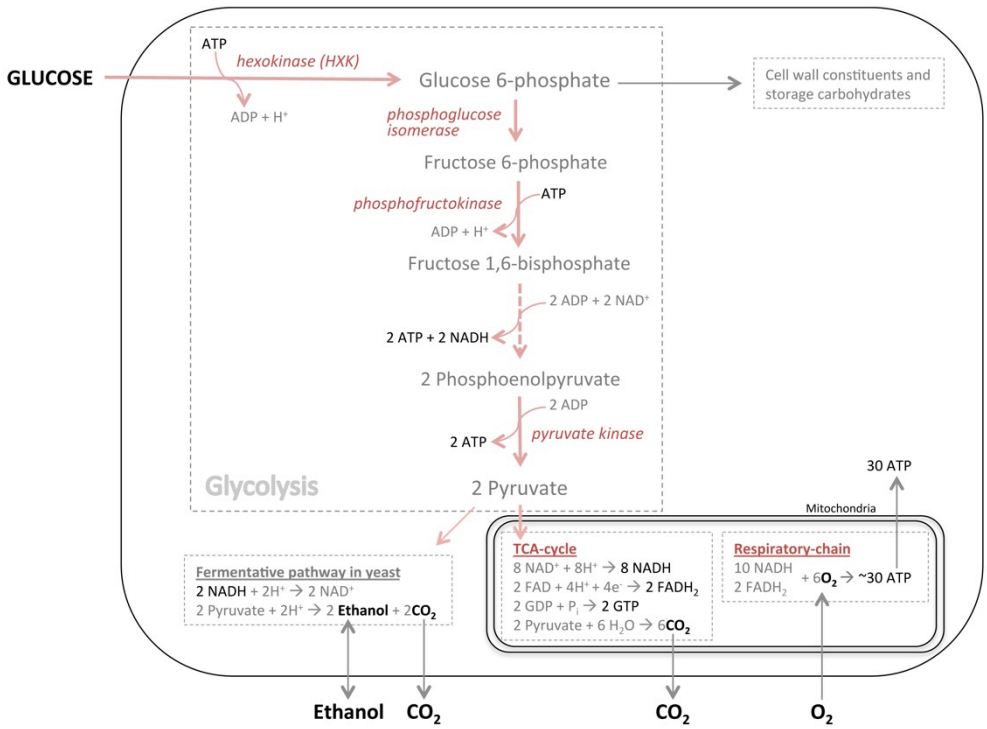
Up to now, it can only be speculated how glucose repression might be triggered but it appears as if glycolytic flux could accomplish this (*109*). This would however require a system for sensing and monitoring the velocity of reactions, such as the glucose uptake mediated by hexose transporters or the phosphorylation of glucose mediated by hexokinases that could then trigger glucose repression (*92*).

8.4 Glycolysis

Adenosine triphosphate (ATP) acts as a short-term carrier of chemical energy that fuel energy-dependent biosynthetic reactions in all known life forms. When cells are growing, a constant supply of ATP from catabolic pathways is required to drive anabolic pathways. Glycolysis is an ancient anaerobic ATP-producing pathway in the central carbon metabolism and is believed to have existed in the last common ancestor of bacteria, eukaryotes and archae (*110*), before O_2 from the first photosynthetic organisms became abundant in the atmosphere. Glycolysis is responsible for a stepwise oxidation of glucose, which allows release of small packages of energy for reduction of carrier molecules such as NADH and ATP. When hexoses are catabolized to pyruvate, ATP and reducing agents (NADH) are produced, which is needed for cell-proliferation. Pyruvate can be further oxidized into CO_2 and H_2O in the TCA-cycle and produce more reducing agents ($NADH/FADH_2$). Under aerobic conditions, most yeasts utilizes O_2 for the oxidation of reducing agents in the respiratory chain to produce more ATP, which enable the recycling of reducing agents to maintain a flux through the TCA cycle. Under O_2 limited conditions yeasts are dependent on glycolysis for the energy requirements, and fermentative pathways for the recycling of NADH to maintain a high glycolytic flux for growth. The net result of glycolysis is:

(1) Glucose ($C_6H_{12}O_6$) + 2 ADP + 2 NAD^+ → 2 Pyruvate ($C_3H_3O_3^-$) + 2 ATP + 2 NADH

Figure 14. This figure illustrates the central carbon metabolism and energy production in yeast. Glycolysis is an anaerobic energy-producing pathway in the central carbon metabolism, and the oxidation and breakdown of many different sugars converts here. The respiratory pathway produces approximately 15 times more energy in terms of ATP yield, which can be used to drive energy-dependent anabolic pathways in a cell. Crabtree negative yeasts predominantly depend on respiration when O_2 is available, while the Crabtree positive yeast, *S. cerevisiae* have been shown to repress respiration and ferment ethanol, even under aerobic conditions.



Crabtree negative yeasts such as *K. lactis* and *C. albicans* are most likely adapted to glucose-limited aerobic environments (*35*). As a result, Crabtree negative yeasts have an economical carbon and energy-source usage, which is reflected in their maximization of energy output by only utilizing respiration for ATP production. When all reducing agents produced from the oxidation of a glucose molecule in glycolysis and TCA cycle are processed in the respiratory chain, the net energy yield is, including the energy requirement for transporting 2 NADH in, and ATP out from the mitochondria:

(2) 10 NADH + 2 $FADH_2$ → 30 ATP

From (1) and (2) it can be seen that complete oxidation of a glucose molecule yields roughly 30 ATP, which corresponds to approximately 15 times higher energy yield in a Crabtree negative as compared to a Crabtree positive yeast.

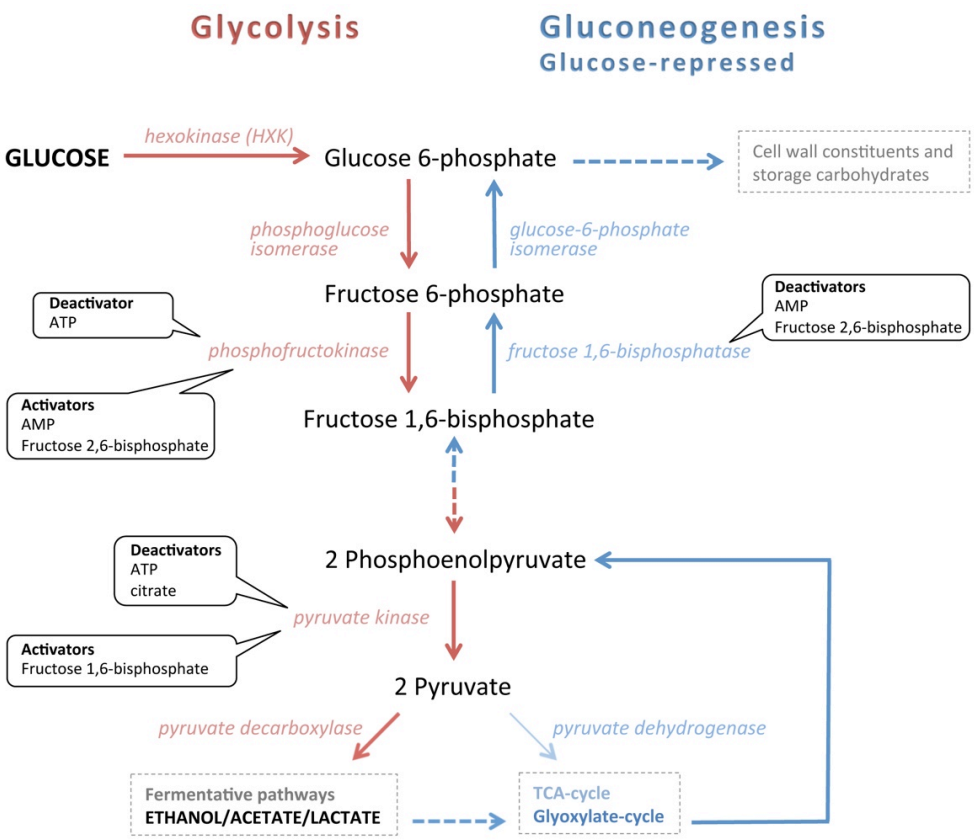
Crabtree positive yeast have on the other hand, most likely evolved in glucose-rich and O_2-limited environments (*111*). This is reflected in its use of the less energy-efficient anaerobic glycolysis for energy-production, and an apparent waste of energy-rich carbon sources by releasing acids and alcohols during growth.

In terms of energy efficiency, it is strange why the Crabtree effect has evolved, and this has puzzled many researchers since its discovery almost a century ago. However, recent hypothesis have tried to incorporate the Crabtree effect into a broader ecological context, which could have identified the key biological force for the evolution of a respiro-fermentative lifestyle in glucose-rich environments (*32, 43*).

8.4.1 Gluconeogenesis

When yeast proliferate on non-fermentable carbon sources such as glycerol, acetate and ethanol, a production of hexoses for storage carbohydrates and cell wall constituents is necessary for growth. This is accomplished by gluconeogenesis, which contain reverse steps of glycolysis that also occur in the cytoplasm. Since many of these reactions involve intermediates from both the glycolysis and gluconeogenesis, these two pathways need to be tightly regulated. This regulation is accomplished at two levels, which are (I) through allosteric deactivation or activation of key-enzymatic reactions in glycolysis and gluconeogenesis, and (II) through glucose repression of all enzymes in gluconeogenesis when cells proliferate on hexoses such as glucose (Figure 15).

Figure 15. Allosteric activation/deactivation of key enzymes in glycolysis and gluconeogenesis, where metabolites acts as sensors of the energy state and substrate levels in the cell. Moreover, all enzymes in gluconeogenesis are also under glucose repression during growth on glucose, and derepressed upon glucose depletion [adapted from (*92, 112*)].



When *S. cerevisiae* proliferates on non-fermentable carbon-sources (i.e. upon a diauxic shift), high activity of enzymes involved in gluconeogenesis to provide sugar precursors is needed. This is partly accomplished by derepression of genes involved in gluconeogenesis, such as alcohol dehydrogenase II, phosphoenolpyruvate carboxykinase and fructose bisphosphatase (Figures 9, 15).

8.5. Respiration

The modern mitochondria harbours many complex biochemical processes, such as (I) the citric acid cycle, mainly responsible for the anabolic and catabolic reactions of fatty-acids and aminoacids, and (II) the chemiosmotic coupled reactions, which is responsible for energy production in forms of ATP from the oxidation of NADH and $FADH_2$ in aerobic organisms (Figure 16). In addition to hosting these processes, the organelle also transcribes and synthesizes proteins encoded in it's own genome (*113*).

Since mitochondria harbour important respiratory pathways such as the TCA cycle and chemiosmotic coupling, mitochondrial defects are often deleterious. The budding yeast and its sister species are however facultative anaerobes, which are believed to have evolved in O_2-limited conditions. In fact, all WGD and some pre-WGD yeasts can proliferate under anaerobic conditions without external supply of pyrimidine (*30*). It has been demonstrated that a lateral gene transfer of *URA1* has enabled uncoupled *de novo* synthesis of pyrimidine from mitochondrial activity, which has contributed to the anaerobic capabilities of yeast in the *Saccharomyces* lineage. Moreover, spontaneous respiratory deficient mutants that have mutations or have lost a significant part of their mitochondrial DNA, lack mitochondrial biogenesis and can be readily isolated in some of these yeasts (*30, 114*). Respiratory deficient mutants are also known as *petites* (meaning small in French), and appear as small colonies on plates with non-fermentable carbon sources and limited amounts of glucose.

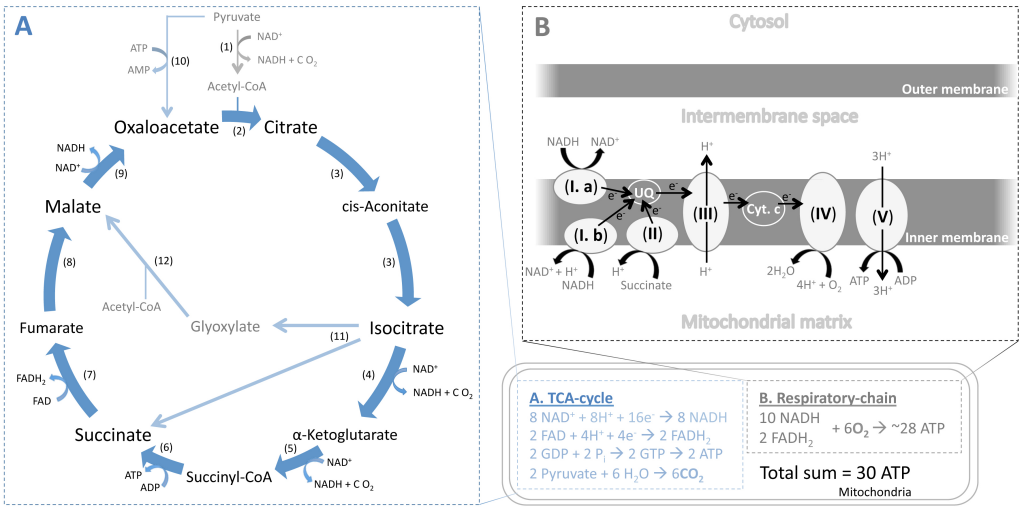
It has been assumed for many decades that respiration is under glucose repression in Crabtree positive yeast, such as *S. cerevisiae*. However, physiological studies on glucose effect on respiratory capacity are limited and appear controversial.

A study on the kinetics of glucose repression on cytochrome c in *S. cerevisiae* has revealed that the regulation was on the transcript level and not on a post-translational level (*91*). When high concentrations of glucose (5 – 10%) were added to a derepressed culture, the rate of cytochrome c synthesis was reduced to a repressed level and the half-life of cytochrome c mRNA decreased to 2 minutes. It was further shown that intermediate level of glucose (0.5 – 1%) would result in a less reduced half-life (3 minutes). Thus, this study on expression and mRNA stability of cytochrome c demonstrated that glucose represses the rate of cytochrome c synthesis by reducing the level of translatable cytochrome c mRNA, without any posttranslational regulation.
However, comparative *in vitro* studies that quantified the respiratory capacity of mitochondria from *S. cerevisiae* and the respiratory capacity of

mitochondria from a Crabtree negative yeast, *C. utilis* could not detect any significant differences (*115*).

More recent transcriptome studies have also revealed a negative effect of glucose on expression of respiratory genes that encodes the TCA cycle (*116*) (Figure 16A), but it is still not clear what causes the Crabtree effect and more physiological studies are needed.

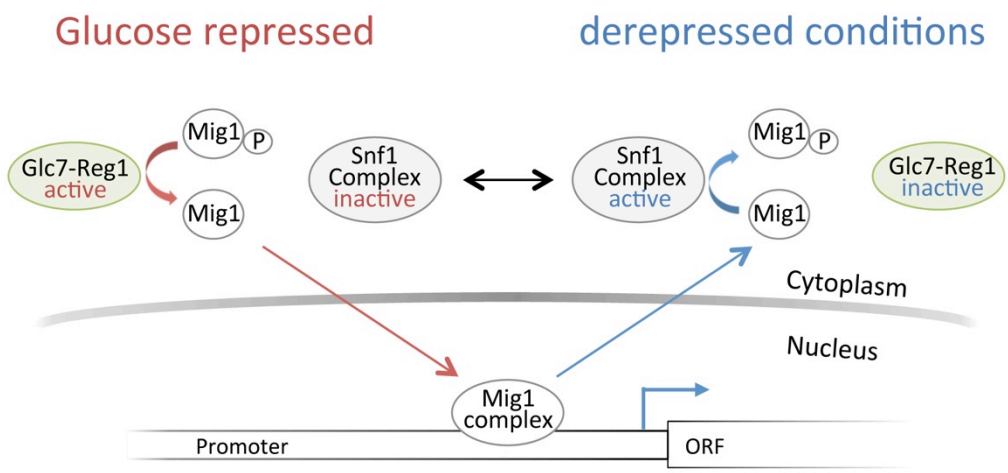
Figure 16. (A) The TCA cycle degrades pyruvate in a stepwise manner producing biomolecules, ATP and energy carriers (NADH, $FADH_2$) for cell-proliferation. The expression of all genes that comprises the TCA- and glyoxylate-cycle in *S. cerevisiae* are under glucose repression, except for (1) pyruvate dehydrogenase, which remain unchanged before and after diauxic shift (*116*). The other enzymes are: (2) citrate synthase; (3) aconitase; (4) isocitrate dehydrogenase; (5) α-ketoglutarate dehydrogenase; (6) succinate thiokinase; (7) succinate dehydrogenase; (8) fumarase; (9) malate dehydrogenase; (10) pyruvate carboxylase; (11) isocitrate lyase; (12) malate synthase. (B) The respiratory chain produces more ATP from the oxidation of energy carriers from glycolysis and the TCA cycle. High energy electrons (e^-) are taken up by external and internal NADH dehydrogenases (1. a,b), and complex (II), which are subsequently released to ubiquinone (UQ). UQ diffuses freely in the lipid membrane and transfer electrons to complex (III), which are further transferred to cytochrome c and finally to O_2 at complex (IV). The redox-reactions that occur at complex (III) and (IV) releases energy that drives the pumping of protons (H^+) against its concentration gradient, out from the mitochondrial matrix. Moreover the diffusion of H^+ down its gradient, back into the mitochondrial matrix, is coupled to ATP synthesis by the proton-conducting complex (V).



8.6. Glucose repression

Glucose can affect cell physiology both on posttranslational level (catabolite inactivation) and on transcriptional level (glucose repression). It is clear that the protein kinase Snf1p plays a central role in mediating the signal from external glucose to the cell nucleus. Snf1p can exist in two forms in a cell, which are (I) an inactive phosphorylated form or (II) an active dephosphorylated form. The equilibrium between these two forms of Snf1p can be shifted by glucose, and in its active form Snf1p can phosphorylate and deactivate the key repressor, Mig1p [reviewed in (*104*)]. Nearly all glucose-repressed genes have a Mig1p binding site that recruits the Mig1p complex, which subsequently blocks transcription of downstream genes in glucose rich conditions (Figure 17).

Figure 17. The Snf1p kinase complex is active and phosphorylates Mig1p in glucose depleted conditions. Mig1p is in its phosphorylated form located in the cytoplasm, where it cannot exert glucose repression (shown in blue). The Glc-Reg1 phosphatase complex is active and dephosphorylates Mig1p under glucose rich conditions. Mig1p is in its dephosphorylated form located in the nucleus where it binds to the promoter regions of glucose repressed genes, and form complex with other proteins such as Hxk2p, Cyc8p and Tup1p to block transcription of the downstream ORF (shown in red) [adapted from (*104, 112*)].



It is clear that *MIG1* and its homologs play an important role in glucose repression and the physiology of *S. cerevisiae*. Mig1p and its homologs exist in many yeasts, but it is not clear how conserved its functions are through yeast evolution. Functional studies of the *K. lactis* Mig1p homolog revealed that it

could complement and rescue a *S. cerevisiae mig1* mutant (*117*). On the other hand, functional studies of *C. albicans* Mig1p homolog, which lacks a putative Snf1p phosphorylation site, could only partly complement a *S. cerevisiae mig1* mutant (*118*). Glucose repression of respiratory genes could indeed be the underlying mechanism behind the Crabtree effect in the *Saccharomyces* lineage. Laboratory evolution experiments have hypothesized that changes in cis-elements of a glucose-repressed gene could increase its expression to a derepressed level, even during growth on glucose (*56*). It could therefore be assumed that the opposite would also be true, where genes important for respiration could "recruit" and accumulate Mig1 *cis*-elements, resulting in glucose repression of respiration.

9. Summary of papers

9.1. Paper I

Candida albicans – a pre-whole genome duplication yeast – is predominantly aerobic and a poor ethanol producer

Yeast are a very divergent group that consist of obligate aerobes, obligate fermenters and yeast that possess a respiro-fermentative metabolism. Yeast species that belong to the lineage that underwent the whole genome duplication (WGD), including *Saccharomyces cerevisiae*, are not dependent on O_2 and can accumulate ethanol under aerobic and glucose rich conditions. On the other hand, several species that diverged from *Saccharomyces* lineage before the WGD event (approx. 100 mya), including *Kluyveromyces lactis* are more dependent on O_2 and do not accumulate ethanol in the presence of O_2. In this study we investigated the physiology and carbon metabolism of yeast species such as *Debaromyces hansenii* and *Candida albicans*, which belong to the lower branches that diverged more than 200 million years ago from the *Saccharomyces* lineage. These studies, which were conducted under fully controlled conditions in bioreactors, revealed that *C. albicans* and *D. hansenii* are poor ethanol producers under aerobic conditions as well as incapable of proliferation under anaerobic conditions. These results were interpreted as that the last common ancestor to yeast possessed an aerobic lifestyle, which was dependent on O_2 for proliferation.

9.2. Paper II

Yeast ”Make-Accumulate-Consume” life strategy evolved as a multi-step process that predates the whole genome duplication

In the age of genomics, fully sequenced yeast genomes have become more readily available for *in silico* based research approaches such as phylogenetics and comparative genomics. This has enabled the reconstruction of the evolutionary history of yeasts, and the identification and timing of several evolutionary events that are believed to explain the appearance of the respiro-fermentative lifestyle in the *Saccharomyces* lineage. Some of these events are (I) the loss of respiratory complex I, which is the first step in chemiosmotic coupled ATP production in mitochondria (II) the lateral transfer of a cytoplasmic version of dihydroorotate dehydrogenase encoding gene (*URA1*), which have enabled uncoupling of *de novo* pyrimidine synthesis from

respiration (III) a whole genome duplication event, which among others increased the gene-dosage of glycolytic genes (IV) a duplication of a alcohol dehydrogenase, which enabled a well regulated fermentative lifestyle where ethanol is accumulated by yeast for later use, and (V) rewiring of rapid growth elements, which are important for co-regulation of cytoplasmic and mitochondrial ribosomal genes.
In this study we characterized over forty yeast species for their carbon metabolism under fully aerobic conditions when grown on glucose. With the help of comparative physiology and comparative genomics approach, we show that the origin of respiro-fermentative lifestyle occurred after the lateral transfer of *URA1* and the loss of respiratory complex I, but before the divergence of *Lachancea* - *Saccharomyces* lineage from the rest of the Saccharomycetaceae yeasts. We also demonstrated that fermentative activity such as ethanol yield is positively correlated to glucose consumption rate and negatively correlated to biomass yield among different yeast species. In overall, closely related yeast species, as determined from phylogeny, also exhibit similar physiology. It appears as if the origin of the fermentative lifestyle in yeast coincided with the origin of anaerobic growth capability more than 130 mya, when the first fruit-bearing angiosperm appeared.

9.3. Paper III

The origin of the short-term Crabtree effect coincides with the long-term Crabtree effect

S. cerevisiae possesses a fermentative life-style and a peculiar trait, which is shared by certain tumour cells and known as the Crabtree effect. There are several definitions of the Crabtree effect and it appears as a complex trait that can be divided into a short- and long-term response to glucose. We have in a previous study (paper II) managed to determine the origin of the long-term Crabtree effect in the *Saccharomyces* lineage. In this study we analysed the response of ten yeast species to glucose pulse. We demonstrated that the origin of the short- and long-term Crabtree effects appear to coincide.

9.4. Paper IV

Overflow is the fundamental mechanism behind long- and short-term Crabtree effect in yeast

Although there are many studies on the short- and long-term Crabtree effect, no consensus for the mechanism behind these traits have been reached. In this paper we use the literature data combined with our own recent wet laboratory data to develop a model that explains the Crabtree effect. We show an interrelationship between anabolic and catabolic pathways, which is conserved among all investigated Ascomycota yeasts. Our model reveal overflow metabolism as the common mechanism behind short- and long-term Crabtree effect. We also attempt to explain the benefits of overflow metabolism from an energy point of view, and why this trait further evolved into a "make-accumulate-consume" life strategy, which appears confined only to certain strains of *S. cerevisiae* and its sister-species.

10. Concluding remarks

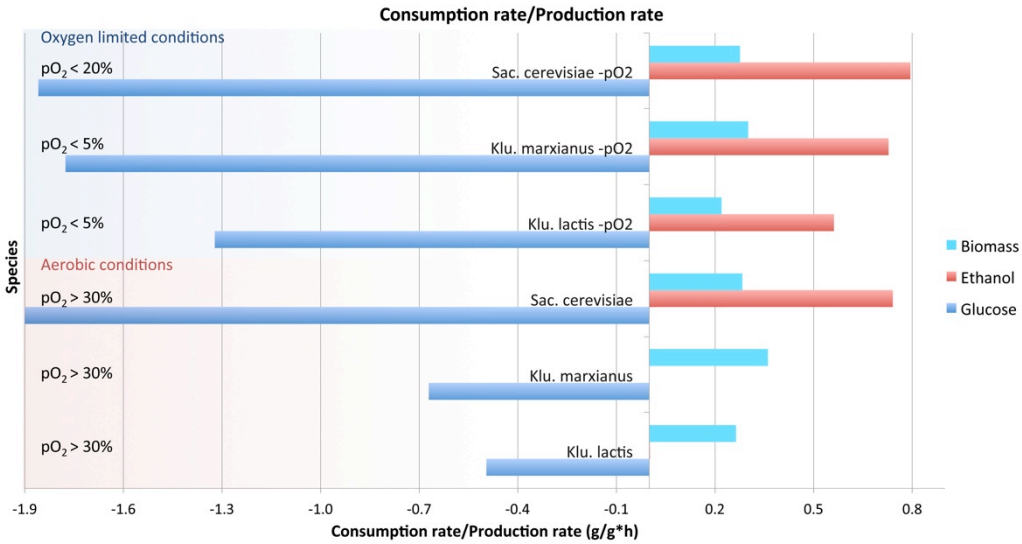
Laboratory evolution experiments have given insight into the importance of aneuploidy in the adaptation of yeast to aerobic glucose limited conditions (section 6.2). Thus, one would assume that similar mechanism could have played an important role in the adaptation to new glucose-rich environments that coincided with the origin of the first angiosperms with fruits more than 125 million years ago (*26*). Moreover, *in silico* studies have clearly indicated that rapid multiplication of low affinity hexose-transporters could be a result from the adaptation of certain yeast species to sugar-rich environment for increased sugar uptake-rate (section 8.2.2), while experimental evolution studies have revealed the opposite to occur – the expansion of high affinity transporters under aerobic glucose-limited conditions.
From these results one could assume that sugar uptake-rate and the aerobic fermentation capacity would increase with increasing number of hexose transporters. However, this assumption might not always be true for other conditions, which can be illustrated by the fact that glucose uptake rate of *K. lactis* and *K. marxianus* is similar to that of *S. cerevisiae* under semi-aerobic conditions, and even the ethanol yield and production-rate are similar for all three species under this condition. Thus, the Pasteur effect, observed in *Kluyveromyces* and the reversed Pasteur effect, observed in *S. cerevisiae* is therefore neither due to an intrinsic limited capacity of glucose uptake-rate nor of glycolytic flux, in *K. lactis* and *K. marxianus*.

My unpublished results (Figure 18) clearly indicate an aberration in the regulation between the respiratory and glycolytic pathways as the most possible mechanism for the transition from a Pasteur effect exhibiting and Crabtree negative yeast, to a reversed Pasteur effect exhibiting and Crabtree positive yeast.

For long time catabolite repression was believed to be the original cause of Crabtree effect in yeast by repressing mitochondrial cytochromes and respiration associated pathways (section 8.5). The earliest studies from the first half of the 20th century had limitations, and relied solely on the quantification of physiological traits, mainly conducted on reference species and often not under controlled and uniform conditions. Subsequent studies from the second half of the 20th century would however incorporate mRNA expression and stability measurements of single respiratory genes, as a response to glucose. However, more recent comparative investigations do not confirm any significant difference in the respiratory capacity of mitochondria from Crabtree positive and Crabtree negative yeast (*115*). It is clear that these limited studies

are insufficient, and more yeast need to be investigated in this matter before any conclusions can be drawn.

Figure 18. Glucose consumption rates and ethanol production rates are similar in *S. cerevisiae, K. marxianus* and *K. lactis* under semi-aerobic conditions (pO_2 < 5% of air saturation). No distinct Pasteur effect can be observed in *S. cerevisiae* under semi-aerobic condition (pO_2 < 20% of air saturation) as compared to aerobic condition (pO_2 > 30% of air saturation).



When the sequencing of *S. cerevisiae* was completed in 1996, new powerful *in silico* approaches enabled global transcriptomics and comparative genomics analysis. Shortly after the discovery of a WGD-event in yeast in 1997, focus was shifted from glucose repression and the peculiar trait of Crabtree effect has been explained in the view of ecology as the "make-accumulate-consume" strategy (section 6.3). Recent studies have claimed to provide the key evolutionary events behind the aerobic fermentation ability in *S. cerevisiae*, such as: (I) a whole genome duplication event (section 7), (II) a global rewiring of the transcriptional network (section 7.3) and (III) the duplication of ancient alcohol dehydrogenases (section 6.3). On the other hand, the Crabtree effect could have also occurred as a net result from combinations of events (Box 2). Piskur and colleagues were the first to incorporate several of these events into the existing "make-accumulate-consume" strategy, as a way for yeast to intoxicate and outcompete less ethanol tolerant bacteria during glucose growth-phase (*32*).

Whether the above-mentioned events have independently contributed to the onset of Crabtree effect in *S. cerevisiae* or not, a consensus in that it most likely occurred after the WGD-event has emerged during the last decades, after the introduction of phylogenetics (*30-35, 38, 57, 86, 119*). Furthermore, there have been reports on yeast with aerobic fermentation abilities that diverged before the *Saccharomyces* lineage, but this was explained as results from parallel evolution in separate lineages, of either same evolutionary events (*35*) or similar regulatory reprogramming events (*98*).

It should be mentioned that many of the conclusions drawn from these inspiring studies, were based on *in silico* approaches, which have been verified by limited experimental data from reference species only, or not at all. Although comparative genomics is a powerful approach for the reconstruction of evolutionary relationship between species and for the identification of evolutionary events, these data should always be backed up with results from comparative physiology experiments, preferably under controlled conditions.

Therefore, as parts of my thesis, large-scale characterization of Saccharomycotina yeasts were conducted to investigate their respiro-fermentative properties and anaerobic growth-capability (Paper I – section 12), and to confirm (or reject) the importance of individual evolutionary events behind the evolution of ethanol production and Crabtree effect in yeast (Paper II – section 12). This large-scale study was followed up by a subsequent investigation for the origin of short-term Crabtree effect (Paper III – section 12) that led to the identification of a putative underlying mechanism behind both long- and short-term Crabtree effect among all investigated Ascomycota yeast species, and a model describing this trait (Paper IV – section 12).

It appears from my studies that the evolution of Crabtree effect, as we know it from *S. cerevisiae* was preceded by overflow metabolism and involved several evolutionary events that have fine-tuned the balance between anabolic and catabolic pathways in yeast. Crabtree negative yeasts appear to possess highly co-regulated glycolysis with respiration, which are tightly coupled to growth. In other words, Crabtree negative yeasts have a coordinated energy metabolism with cell proliferation under aerobic conditions. Crabtree negative yeasts can under O_2-limited conditions upregulate their glucose uptake rates and carbon flux through anaerobic glycolysis to cover the requirements for biomolecules and energy to drive anabolic reactions. Respiro-fermenting and Crabtree positive yeasts on the other hand, appear to have a relaxed co-regulation between glycolysis and respiration, which are less coupled to growth. In other words, yeasts that possess aerobic fermentation capability have a less coordinated energy metabolism with growth that lead to overconsumption of

glucose, increased flux through anaerobic glycolytic pathway and ethanol formation, even under aerobic conditions.

Rewiring of promoter regions (section 7.3) to relax co-regulation of respiration from fermentation (*33*) could indeed have contributed to the observed Crabtree effect in some WGD species, whether this is mediated by any trans-factors of RGE (*35*) or by nucleosome depletion (*86*). From what is known in the literature, it can also be concluded that the Crabtree effect could be a result from rewiring of the *MIG* motif that resulted in repression of respiration. However, our data can best be explained by overflow metabolism, which appears to have evolved before the WGD event, glucose repression of respiration and a settled rewiring of the RGE elements in the Saccharomycetaceae yeast. Therefore, the rewiring of a yet unidentified motif, or the identification of an unknown activator could provide the explanation for the observed upregulation of anaerobic glycolytic pathway, which causes what appears to be an imbalance between catabolism and anabolism due to "overconsumption" of glucose, and results in overflow metabolism in pre-WGD respiro-fermentative yeasts.

Acknowledgements

I thank my supervisor Jure Piškur for accepting me to work in his lab, for all his patience, for much appreciated input into my projects, and support in the realization of project ideas as well as personal development plan and choice of career path – your efficiency and skills as a coordinator is truly inspiring.

Thanks to Khadija, Olena and Zeeshan for being great coworkers – the office is a better place with you there. Also, thanks to all current and former lab-members for creating a great working environment (you would be an "asset" to any laboratories!).

Special thanks to Karin Regnefors, and GENECO, for arranging interesting meetings, scientific activities and for making my visit to Český Krumlov and Milan possible. Thanks to Rickard Jonsson, for being my mentor while I was enrolled in GENECO – our conversations have been much appreciated.

My grateful thanks to Concetta Compagno for accepting me to work in her laboratory during my stay in Milan – I have learnt much during my visit. Thank you also for reading and commenting on my work. Louise Johansson, I cannot thank you enough for your help with the initial setup of our bioreactor platform in Lund, and for introducing me to the different aspects of fermentation, both in terms of practical as well as theoretical part.

I would also like to thank Torbjörn, for his scientific expertise and guidance, and for reading and commenting on my work - I am glad to be able to discuss ideas and statistical problems with you. You make statistics look simple and fun to learn! My grateful thanks are also extended to Bengt Olle Bengtsson, for reading and commenting on my work, and for making my visit to Cornell possible.

Thanks to Zhenglong Gu for accepting me to work in his laboratory in Cornell. I am also grateful to Huifeng Jiang, Kaixiong Ye and Brandon Barker for introducing me to bioinformatics – I have learnt much and am looking forward to see where our ongoing collaboration might lead. Also thanks to Daniel Jarosz for taking interest in my work, and for inviting me to his laboratory in Stanford.

I would also like to extend my thanks to all current and past members of the Biology building A – for contributing to a nice atmosphere, and last but not least, for always providing fresh coffee on time.

Finally, thanks to my family, Boonnak, Erik, Göran and Kay for your never-ending support, and for being my source of inspiration - you are the most determined and hardworking persons I know of, and I am proud of you. I hope one day I can inspire others, like you have inspired me. This work is dedicated to you.

11. References

1. F. D. Ciccarelli *et al.*, Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**, 1283 (2006).
2. C. P. Kurtzman, C. J. Robnett, Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Leeuwenhoek* **73**, 331 (1998).
3. C. Kurtzman, Phylogenetic circumscription of *Saccharomyces*, *Kluyveromyces* and other members of the Saccharomycetaceae, and the proposal of the new genera *Lachancea*, *Nakaseomyces*, *Naumovia*, *Vanderwaltozyma* and *Zygotorulaspora*. *FEMS Yeast Res.* **4**, 233 (2003).
4. B. Dujon, Yeast evolutionary genomics. *Nature reviews. Genetics* **11**, 512 (2010).
5. M. Parniske, Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature reviews. Microbiology* **6**, 763 (2008).
6. E. M. Medina, G. W. Jones, D. A. Fitzpatrick, Reconstructing the fungal tree of life using phylogenomics and a preliminary investigation of the distribution of yeast prion-like proteins in the fungal kingdom. *J. Mol. Evol.* **73**, 116 (2011).
7. F. Lutzoni *et al.*, Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits. *Am. J. Bot.* **91**, 1446 (2004).
8. D. L. Hawksworth, The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* **105**, 1422 (2001).
9. M. Blackwell, The Fungi: 1, 2, 3 ... 5.1 Million Species? *Am. J. Bot.* **98**, 426 (2011).
10. H. E. O'Brien, J. L. Parrent, J. A. Jackson, J. M. Moncalvo, R. Vilgalys, Fungal community analysis by large-scale sequencing of environmental samples. *Appl. Environ. Microbiol.* **71**, 5544 (2005).
11. D. S. Hibbett *et al.*, A higher-level phylogenetic classification of the Fungi. *Mycol. Res.* **111**, 509 (2007).
12. T. Y. James *et al.*, Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443**, 818 (2006).
13. X. Yuan, S. Xiao, T. N. Taylor, Lichen-like symbiosis 600 million years ago. *Science* **308**, 1017 (2005).
14. D. Redecker, J. B. Morton, T. D. Bruns, Ancestral lineages of arbuscular mycorrhizal fungi (*Glomales*). *Mol. Phylogenet. Evol.* **14**, 276 (2000).
15. D. S. Heckman *et al.*, Molecular evidence for the early colonization of land by fungi and plants. *Science* **293**, 1129 (2001).
16. T. N. Taylor, H. Hass, H. Kerp, The oldest fossil ascomycetes. *Nature* **399**, 648 (1999).

17. National Center for Biotechnology Information. http://www.ncbi.nlm.nih.gov/.
18. *Saccharomyces* Genome database. http://www.yeastgenome.org/.
19. Génolevures. http://www.genolevures.org/.
20. Yeast Gene Order Browser. http://wolfe.gen.tcd.ie/ygob/.
21. A. Goffeau *et al.*, Life with 6000 genes. *Science* **274**, 546 (1996).
22. J. W. Schopf, Fossil evidence of Archaean life. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **361**, 869 (2006).
23. H. D. Holland, The oxygenation of the atmosphere and oceans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **361**, 903 (2006).
24. H. S. Yoon, J. D. Hackett, C. Ciniglia, G. Pinto, D. Bhattacharya, A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21**, 809 (2004).
25. D. Redecker, R. Kodner, L. E. Graham, Glomalean fungi from the Ordovician. *Science* **289**, 1920 (2000).
26. G. Sun, D. L. Dilcher, H. Wang, Z. Chen, A eudicot from the Early Cretaceous of China. *Nature* **471**, 625 (2011).
27. A. Zimmer *et al.*, Dating the early evolution of plants: detection and molecular clock analyses of orthologs. *Molecular Genetics and Genomics : MGG* **278**, 393 (2007).
28. M. Marcet-Houben, G. Marceddu, T. Gabaldon, Phylogenomics of the oxidative phosphorylation in fungi reveals extensive gene duplication followed by functional divergence. *BMC Evol. Biol.* **9**, 295 (2009).
29. Z. Gojkovic *et al.*, Horizontal gene transfer promoted evolution of the ability to propagate under anaerobic conditions in yeasts. *Molecular Genetics and Genomics : MGG* **271**, 387 (2004).
30. A. Merico, P. Sulo, J. Piskur, C. Compagno, Fermentative lifestyle in yeasts belonging to the *Saccharomyces* complex. *The FEBS journal* **274**, 976 (2007).
31. K. H. Wolfe, D. C. Shields, Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708 (1997).
32. J. Piskur, E. Rozpedowska, S. Polakova, A. Merico, C. Compagno, How did *Saccharomyces* evolve to become a good brewer? *Trends Genet.* **22**, 183 (2006).
33. J. Ihmels *et al.*, Rewiring of the yeast transcriptional network through the evolution of motif usage. *Science* **309**, 938 (2005).
34. H. Jiang, W. Guan, D. Pinney, W. Wang, Z. Gu, Relaxation of yeast mitochondrial functions after whole-genome duplication. *Genome Res.* **18**, 1466 (2008).
35. E. Rozpedowska *et al.*, Parallel evolution of the make-accumulate-consume strategy in *Saccharomyces* and *Dekkera* yeasts. *Nature communications* **2**, 302 (2011).

36. B. Papp, C. Pal, L. D. Hurst, Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194 (2003).
37. D. R. Scannell *et al.*, Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8397 (2007).
38. G. C. Conant, K. H. Wolfe, Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Molecular Systems Biology* **3**, 129 (2007).
39. T. Rolland, C. Neuveglise, C. Sacerdot, B. Dujon, Insertion of horizontally transferred genes within conserved syntenic regions of yeast genomes. *PLoS One* **4**, e6515 (2009).
40. C. G. Kurland, S. G. Andersson, Origin and evolution of the mitochondrial proteome. *Microbiology and molecular biology reviews : MMBR* **64**, 786 (2000).
41. D. Bhattacharya, J. M. Archibald, A. P. Weber, A. Reyes-Prieto, How do endosymbionts become organelles? Understanding early events in plastid evolution. *BioEssays : news and reviews in molecular, cellular and developmental biology* **29**, 1239 (2007).
42. M. Nagy, F. Lacroute, D. Thomas, Divergent evolution of pyrimidine biosynthesis between anaerobic and aerobic yeasts. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8966 (1992).
43. C. Hall, S. Brachat, F. S. Dietrich, Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. *Eukaryot. Cell* **4**, 1102 (2005).
44. D. Karibian, P. Couchoud, Dihydro-orotate oxidase of *Escherichia coli* K12: purification, properties, and relation to the cytoplasmic membrane. *Biochim. Biophys. Acta* **364**, 218 (1974).
45. P. S. Andersen, P. J. Jansen, K. Hammer, Two different dihydroorotate dehydrogenases in *Lactococcus lactis*. *J. Bacteriol.* **176**, 3975 (1994).
46. T. Annoura, T. Nara, T. Makiuchi, T. Hashimoto, T. Aoki, The origin of dihydroorotate dehydrogenase genes of kinetoplastids, with special reference to their biological significance and adaptation to anaerobic, parasitic conditions. *J. Mol. Evol.* **60**, 113 (2005).
47. B. Bader, W. Knecht, M. Fries, M. Loffler, Expression, purification, and characterization of histidine-tagged rat and human flavoenzyme dihydroorotate dehydrogenase. *Protein Expr. Purif.* **13**, 414 (1998).
48. R. Bisson, G. Schiavo, C. Montecucco, ATP induces conformational changes in mitochondrial cytochrome c oxidase. Effect on the cytochrome c binding site. *The Journal of biological chemistry* **262**, 5992 (1987).
49. M. Loffler, J. Jockel, G. Schuster, C. Becker, Dihydroorotat-ubiquinone oxidoreductase links mitochondria in the biosynthesis of pyrimidine nucleotides. *Mol. Cell. Biochem.* **174**, 125 (1997).

50. K. H. Wolfe, Comparative genomics and genome evolution in yeasts. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **361**, 403 (2006).
51. B. Dujon *et al.*, Genome evolution in yeasts. *Nature* **430**, 35 (2004).
52. I. S. Pretorius, Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. *Yeast* **16**, 675 (2000).
53. G. H. Fleet, Wine yeasts for the future. *FEMS Yeast Res.* **8**, 979 (2008).
54. C. Paquin, J. Adams, Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature* **302**, 495 (1983).
55. C. J. Brown, K. M. Todd, R. F. Rosenzweig, Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol. Biol. Evol.* **15**, 931 (1998).
56. M. J. Dunham *et al.*, Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16144 (2002).
57. J. M. Thomson *et al.*, Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat. Genet.* **37**, 630 (2005).
58. C. Wills, Production of yeast alcohol dehydrogenase isoenzymes by selection. *Nature* **261**, 26 (1976).
59. M. Ciriacy, Cis-dominant regulatory mutations affecting the formation of glucose-repressible alcohol dehydrogenase (ADHII) in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **145**, 327 (1976).
60. A. Bozzi, M. Saliola, C. Falcone, F. Bossa, F. Martini, Structural and biochemical studies of alcohol dehydrogenase isozymes from *Kluyveromyces lactis*. *Biochim. Biophys. Acta* **1339**, 133 (1997).
61. D. K. Butler, L. E. Yasuda, M. C. Yao, Induction of large DNA palindrome formation in yeast: implications for gene amplification and genome stability in eukaryotes. *Cell* **87**, 1115 (1996).
62. K. S. Lobachev, D. A. Gordenin, M. A. Resnick, The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **108**, 183 (2002).
63. T. L. Ferea, D. Botstein, P. O. Brown, R. F. Rosenzweig, Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9721 (1999).
64. M. Kellis, B. W. Birren, E. S. Lander, Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617 (2004).
65. F. S. Dietrich *et al.*, The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science* **304**, 304 (2004).
66. T. Thiel *et al.*, Evidence and evolutionary analysis of ancient whole-genome duplication in barley predating the divergence from rice. *BMC Evol. Biol.* **9**, 209 (2009).

67. O. Jaillon *et al.*, The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463 (2007).
68. The map-based sequence of the rice genome. *Nature* **436**, 793 (2005).
69. C. Simillion, K. Vandepoele, M. C. Van Montagu, M. Zabeau, Y. Van de Peer, The hidden duplication past of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13627 (2002).
70. T. J. Vision, D. G. Brown, S. D. Tanksley, The origins of genomic duplications in *Arabidopsis*. *Science* **290**, 2114 (2000).
71. L. J. Ma *et al.*, Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. *PLoS Genet* **5**, e1000549 (2009).
72. P. Dehal, J. L. Boore, Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* **3**, e314 (2005).
73. S. C. Le Comber, C. Smith, Polyploidy in fishes: patterns and processes. *Biol. J. Linn. Soc.* **82**, 431 (2004).
74. O. Jaillon *et al.*, Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946 (2004).
75. M. Semon, K. H. Wolfe, Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8333 (2008).
76. J. A. Birchler, R. A. Veitia, The gene balance hypothesis: from classical genetics to modern genomics. *The Plant cell* **19**, 395 (2007).
77. J. M. Sheltzer *et al.*, Aneuploidy drives genomic instability in yeast. *Science* **333**, 1026 (2011).
78. E. Segal *et al.*, Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat. Genet.* **34**, 166 (2003).
79. J. R. Warner, The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24**, 437 (1999).
80. S. Bergmann, J. Ihmels, N. Barkai, Similarities and differences in genome-wide expression data of six organisms. *PLoS Biol.* **2**, E9 (2004).
81. J. M. Stuart, E. Segal, D. Koller, S. K. Kim, A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302**, 249 (2003).
82. I. Wapinski *et al.*, Gene duplication and the evolution of ribosomal protein gene regulation in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5505 (2010).
83. A. Tanay, A. Regev, R. Shamir, Conservation and evolvability in regulatory networks: the evolution of ribosomal regulation in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7203 (2005).
84. D. J. Wohlbach, D. A. Thompson, A. P. Gasch, A. Regev, From elements to modules: regulatory evolution in Ascomycota fungi. *Curr. Opin. Genet. Dev.* **19**, 571 (2009).

85. M. C. Bosio, R. Negri, G. Dieci, Promoter architectures in the yeast ribosomal expression program. *Transcription* **2**, 71 (2011).
86. Y. Field *et al.*, Gene expression divergence in yeast is coupled to evolution of DNA-encoded nucleosome organization. *Nat. Genet.* **41**, 438 (2009).
87. B. B. Tuch, H. Li, A. D. Johnson, Evolution of eukaryotic transcription circuits. *Science* **319**, 1797 (2008).
88. O. Warburg, On the origin of cancer cells. *Science* **123**, 309 (1956).
89. K. H. Ibsen, The Crabtree effect: a review. *Cancer Res.* **21**, 829 (1961).
90. R. H. De Deken, The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol* **44**, 149 (1966).
91. R. S. Zitomer, D. L. Nichols, Kinetics of glucose repression of yeast cytochrome c. *J. Bacteriol.* **135**, 39 (1978).
92. H. Ronne, Glucose repression in fungi. *Trends Genet.* **11**, 12 (1995).
93. H. J. Federoff, T. R. Eccleshall, J. Marmur, Carbon catabolite repression of maltase synthesis in *Saccharomyces carlsbergensis*. *J. Bacteriol.* **156**, 301 (1983).
94. H. G. Crabtree, Observations on the carbohydrate metabolism of tumours. *The Biochemical journal* **23**, 536 (1929).
95. M. Petrik, O. Kappeli, A. Fiechter, An Expanded Concept for the Glucose Effect in the Yeast *Saccharomyces-Uvarum* - Involvement of Short-Term and Long-Term Regulation. *J Gen Microbiol* **129**, 43 (1983).
96. R. A. Weusthuis, J. T. Pronk, P. J. van den Broek, J. P. van Dijken, Chemostat cultivation as a tool for studies on sugar transport in yeasts. *Microbiol Rev* **58**, 616 (1994).
97. M. C. Loureiro-Dias, Movements of protons coupled to glucose transport in yeasts. A comparative study among 248 yeast strains. *Antonie Leeuwenhoek* **54**, 331 (1988).
98. Z. Lin, W. H. Li, Expansion of hexose transporter genes was associated with the evolution of aerobic fermentation in yeasts. *Mol. Biol. Evol.* **28**, 131 (2011).
99. P. Billard *et al.*, Glucose uptake in *Kluyveromyces lactis*: Role of the *HGT1* gene in glucose transport. *J. Bacteriol.* **178**, 5860 (1996).
100. X. J. Chen, M. Wesolowski-Louvel, H. Fukuhara, Glucose transport in the yeast *Kluyveromyces lactis*. II. Transcriptional regulation of the glucose transporter gene *RAG1*. *Mol. Gen. Genet.* **233**, 97 (1992).
101. H. van Urk, E. Postma, W. A. Scheffers, J. P. van Dijken, Glucose transport in crabtree-positive and crabtree-negative yeasts. *J Gen Microbiol* **135**, 2399 (1989).
102. E. Boles, C. P. Hollenberg, The molecular genetics of hexose transport in yeasts. *FEMS Microbiol. Rev.* **21**, 85 (1997).
103. M. Rose, W. Albig, K. D. Entian, Glucose repression in *Saccharomyces cerevisiae* is directly associated with hexose phosphorylation by

hexokinases PI and PII. *European journal of biochemistry / FEBS* **199**, 511 (1991).
104. J. M. Gancedo, The early steps of glucose signalling in yeast. *FEMS Microbiol. Rev.* **32**, 673 (2008).
105. J. M. Gancedo, Carbon catabolite repression in yeast. *European journal of biochemistry / FEBS* **206**, 297 (1992).
106. F. Randez-Gil, P. Herrero, P. Sanz, J. A. Prieto, F. Moreno, Hexokinase PII has a double cytosolic-nuclear localisation in *Saccharomyces cerevisiae*. *FEBS Lett.* **425**, 475 (1998).
107. K. D. Entian, K. U. Frohlich, D. Mecke, Regulation of enzymes and isoenzymes of carbohydrate metabolism in the yeast *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **799**, 181 (1984).
108. D. Ahuatzi, P. Herrero, T. de la Cera, F. Moreno, The glucose-regulated nuclear localization of hexokinase 2 in *Saccharomyces cerevisiae* is *Mig1*-dependent. *The Journal of biological chemistry* **279**, 14440 (2004).
109. L. F. Bisson, V. Kunathigan, On the trail of an elusive flux sensor. *Res. Microbiol.* **154**, 603 (2003).
110. A. H. Romano, T. Conway, Evolution of carbohydrate metabolic pathways. *Res. Microbiol.* **147**, 448 (1996).
111. A. Hagman, T. Sall, C. Compagno, J. Piskur, Yeast "make-accumulate-consume" life strategy evolved as a multi-step process that predates the whole genome duplication. *PLoS One* **8**, e68734 (2013).
112. J. A. Barnett, K. D. Entian, A history of research on yeasts 9: regulation of sugar metabolism. *Yeast* **22**, 835 (2005).
113. T. L. Ulery, S. H. Jang, J. A. Jaehning, Glucose repression of yeast mitochondrial transcription: kinetics of derepression and role of nuclear genes. *Mol. Cell. Biol.* **14**, 1160 (1994).
114. V. Fekete, M. Cierna, S. Polakova, J. Piskur, P. Sulo, Transition of the ability to generate petites in the Saccharomyces/Kluyveromyces complex. *FEMS Yeast Res.* **7**, 1237 (2007).
115. H. Van Urk, P. M. Bruinenberg, M. Veenhuis, W. A. Scheffers, J. P. Van Dijken, Respiratory capacities of mitochondria of *Saccharomyces cerevisiae* CBS 8066 and *Candida utilis* CBS 621 grown under glucose limitation. *Antonie Leeuwenhoek* **56**, 211 (1989).
116. J. L. DeRisi, V. R. Iyer, P. O. Brown, Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680 (1997).
117. J. P. Cassart, I. Georis, J. Ostling, H. Ronne, J. Vandenhaute, The *MIG1* repressor from *Kluyveromyces lactis*: cloning, sequencing and functional analysis in *Saccharomyces cerevisiae*. *FEBS Lett.* **371**, 191 (1995).
118. O. Zaragoza, C. Rodriguez, C. Gancedo, Isolation of the *MIG1* gene from *Candida albicans* and effects of its disruption on catabolite repression. *J. Bacteriol.* **182**, 320 (2000).

119. H. Chen, L. Xu, Z. Gu, Regulation dynamics of WGD genes during yeast metabolic oscillation. *Mol. Biol. Evol.* **25**, 2513 (2008).