Engineering and characterization of alternative configurations of glycolysis in Saccharomyces cerevisiae

Knibbe, E.

DOI
10.4233/uuid:e05ce216-a710-459a-9530-d22c1e70ed00

Publication date
2022

Document Version
Final published version

Citation (APA)
Knibbe, E. (2022). *Engineering and characterization of alternative configurations of glycolysis in Saccharomyces cerevisiae*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:e05ce216-a710-459a-9530-d22c1e70ed00

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Engineering and characterization of alternative configurations of glycolysis in *Saccharomyces cerevisiae*

Ewout Knibbe

Engineering and characterization of alternative configurations of glycolysis in *Saccharomyces cerevisiae*

Dissertation

For the purpose of obtaining the degree of doctor

at Delft University of Technology

by the authority of the Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen

Chair of the Board for Doctorates

to be defended publicly on:

Wednesday 29 June 2022 at 15:00 o'clock

by

Ewout KNIBBE

Master of Science in Life Science & Technology, Delft University of Technology, The Netherlands

Born in Amsterdam, The Netherlands

This dissertation has been approved by the promotors
Prof. dr. P.A.S. Daran-Lapujade and prof. dr. J.T. Pronk

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof. dr. P.A.S. Daran-Lapujade	Delft University of Technology, Promotor
Prof. dr. J.T. Pronk	Delft University of Technology, Promotor

Independent members:

Dr. M. Krogh Jensen	DTU Biosustain, Denmark
Prof. dr. D. Claessen	Leiden University
Prof. dr. Ir. M. de Mey	Ghent University, Belgium
Prof. dr. Ir. M.C.M. van Loosdrecht	Delft University of Technology
Prof. dr. F. Hollmann	Delft University of Technology, reserve member

Other members:

Dr. C.L. Flores Mauriz	IIBM, Spain

The research presented in this thesis was performed at the Industrial Microbiology Group, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, The Netherlands. The project was funded by a consolidator grant; AdLibYeast; from the European Research Council (ERC).

Cover: Ewout Knibbe

Layout: Ewout Knibbe

Printed by: Ridderprint, www.ridderprint.nl

ISBN: 978-94-6384-339-3

Contents

Summary

One of the most conserved and central parts of metabolism is the pathway of glycolysis, that breaks down hexose sugars. While multiple variants of glycolytic pathway exist, the Emden-Meyerhof-Parnas pathway which converts glucose to two molecules of pyruvate with the net gain of two ATP and two reduced NADH moieties is the one predominantly found in eukaryotes. This well-studied metabolic pathway is important for industrial biotechnology, because sugars are often used as substrate for microbial production of fuels and chemicals, as well as human health, because of the involvement of glycolytic dysfunction and deregulation in diseases such as cancer. The analysis of glycolytic enzyme function is complicated by the presence in many organisms of a large range of secondary 'moonlighting' functions besides the main glycolytic function. Because of its robustness, genetic accessibility, ease of culture and highly expressed glycolytic pathway, baker's yeast, *Saccharomyces cerevisiae* has been key to research on glycolysis. Besides its role in food and beverage production and as industrial host for e.g. bioethanol and insulin production, this yeast has been intensively used as a model organism to study fundamental processes and pathways.

Despite *S. cerevisiae* glycolysis being the first metabolic pathway to be discovered, understanding of its regulation, which is largely achieved by metabolic inhibition and activation, as well of its complex secondary functions, is far from complete. Study and engineering of eukaryotic metabolism is complicated by genetic redundancy, as illustrated by the 26 isoenzymes encoded by 26 genes scattered across the genome in yeast, catalysing the 10 glycolytic steps and the 2 fermentation steps (converting pyruvate to ethanol). To enable simpler study and modification of yeast glycolysis, previously the essentiality of the various isoenzymes has been studied and the minor isoenzymes were deleted without noticeable impact on yeast physiology. In a further step, the genes encoding the 13 major isoenzymes were relocated to a single chromosomal locus, enabling the modular replacement of the entire pathway in two transformation steps. These Minimal Glycolysis (MG) and Switchable Yeast Glycolysis (SwYG) strains are the foundation for this thesis. The goals of this thesis are to investigate the potential of pathway swapping to study and remodel this central pathway and explore the metabolic and physiological impact of glycolytic transplantation in *S. cerevisiae*.

While the minimization of the glycolytic pathway had no visible effect on yeast physiology, the relocation of the genes encoding the major isoenzymes in the SwYG strain resulted in a 8-30% decrease in growth rate depending on the growth conditions. In **Chapter 2** the effects of genetic reorganization on yeast physiology,

and more particularly on the specific growth rate decrease was studied. Using the SwYG strain, several aspects of the pathway relocation and the concomitant genetic alterations were investigated, including presence of ARS sequences and gene direction. Additionally the changes resulting from pathway reorganization on glycolytic enzyme activity and expression from neighbouring loci were analysed, however none of these showed a distinct connection to strain physiology. Wild-type growth rate could be restored in the SwYG strain through adaptive evolution and reverse engineering, unexpectedly revealing the involvement of genes related to autophagy. While this study could not identify the exact cause of the slow growth phenotype of the SwyG strain, untargeted mutations caused by strain construction as well as an as-yet uncharacterised moonlighting function of Pgk1 were implicated in the growth defect. This chapter illustrates the challenges encountered by extensive strain construction programs, additionally the analysis of the impact of pathway reorganization provided design guidelines for modular pathways.

The humanization of genes in yeast has long been used to study protein and pathway function, allowing simpler modification and characterization compared to mammalian cells. **Chapter 3** reports the humanization of the glycolytic pathway in yeast. The combination of single gene complementation, full pathway swapping and adaptive laboratory evolution proved a powerful approach to evaluate the conservation of enzyme function and generate new humanized model strains. The activity of the 25 tested human glycolytic enzymes was largely conserved in the humanized yeast strains, with the exception of the muscle hexokinases, which were only functional in yeast after the acquisition of mutations that modified hexokinase sensitivity to its allosteric inhibitor glucose-6-phosphate. This requirement for reduced allosteric sensitivity was observed with single hexokinase complementation, as well as in yeast strains with fully humanized muscle glycolysis, showing fundamental differences in operation between human (muscle) and yeast cells. A surprising further finding was the apparent complementation of moonlighting functions in yeast by the human hexokinase, aldolase and enolase enzymes. The function of the human pathway was further optimized in *S. cerevisiae* by targeted gene overexpression and evolutionary engineering, which identified hexokinase and phosphoglycerate mutase as likely limiting steps in the pathway and pointed to the possibility of novel interactions between human aldolase and the yeast cytoskeleton. Finally, comparison of the glycolytic enzymes from humanized strains to those from cultured muscle cells indicated similar catalytic turnover numbers, showing the potential of humanized yeast strains as novel models to study metazoan glycolysis.

Integrating the lessons learned in the previous chapters, **Chapter 4** describes the use of pathway swapping to explore metabolic regulation of *S. cerevisiae* glycolysis. While such fast-acting regulation of glycolytic enzymes by metabolites is wide-

spread across organisms and thought to be important for dynamic flux regulation in *S. cerevisiae*, single enzymes studies have failed to identify a clear role for (allosteric) metabolic regulation in glycolysis. Using transplantation of non-allosteric glycolytic enzymes from the oleaginous yeast *Yarrowia lipolytica*, devoid of the typical metabolic regulations found in *S. cerevisiae*, these long-standing questions were put to the test. Full and partial pathway replacement, combined with single complementation revealed that while single enzymes can be replaced with non-regulated variants with marginal phenotypic effect, combination of expression of the *Y. lipolytica* glucokinase, phosphofructokinase and pyruvate kinase led to a dysfunctional pathway. Swapped glycolysis strains showed high sensitivity to glucose, but not to galactose, suggesting that the lack of metabolic regulation caused a carbon overflow in upper glycolysis and an imbalance between upper and lower glycolysis. Adaptive evolution revealed that reduction of the glucokinase activity suppressed this effect. Kinetic modelling combined with culturing under lower glucose conditions confirmed that the reduction of the flux in upper glycolysis restored glycolytic balance in the absence of *S. cerevisiae*-like metabolic regulations. These observations demonstrate the synergetic role played by metabolic regulation of hexokinase, phosphofructokinase and pyruvate kinase in *S. cerevisiae* during dynamic conditions, particularly during transition between glucose poor and rich environments. This chapter therefore demonstrates the potential of pathway swapping to study fundamental aspects of glycolysis and the importance of pathway level, systems approaches to study complex regulatory mechanisms.

Finally **Chapter 5** looks beyond yeast Synthetic Biology to the broader field of Synthetic Genomics, which focuses on the construction of synthetic chromosomes and genomes. Lessons about synthetic locus design learned from the swapping of a single pathway can eventually be used for larger scale engineering of metabolism, including modification of central carbon metabolism and implementation of product pathways. Such metabolic engineering can be made easier by the application of synthetic chromosomes, which would enable great flexibility for pathway rearrangement, integration and swapping. However, current DNA synthesis techniques are limited to short DNA fragments, necessitating assembly of multiple fragments even for a single synthetic gene. With current synthetic chromosomes ranging to thousands of base pairs and dozens of genes and potential synthetic genomes increasing in size beyond that, methods for DNA assembly are critical to the success of synthetic biology. In this chapter the use of *S. cerevisiae* and its powerful intrinsic homologous recombination machinery for assembly of synthetic DNA is reviewed. The use of this versatile microorganism has been critical to the development of the field so far, enabling the first assembly of a complete synthetic genome. The lack of efficient alternative DNA assembly methods and the intrinsic

limits of bacterial hosts for synthetic biology mean *S. cerevisiae* will likely remain a key genome foundry for the foreseeable future.

Samenvatting

De glycolyse, de metabole route waarin suikers worden afgebroken, is een van de meest behouden en centrale metabole routes. Hoewel meerdere varianten van deze route bestaan, komt in eukaryoten voornamelijk de Emden-Meyerhof-Parnas glycolyse voor, waarin glucose wordt omgezet tot twee moleculen pyruvaat (pyrodruivenzuur) met de productie van twee ATP en twee gereduceerde NADH. Dit uitgebreid bestudeerde deel van het metabolisme is van belang voor de industriële biotechnologie, waarin vaak gebruik wordt gemaakt van suikers als grondstof voor de productie van brandstoffen en chemicaliën, maar ook voor de menselijke gezondheid, omdat dysfunctie en deregulatie van de glycolyse verband houden met ziektes zoals kanker. Onderzoek naar de functie van glycolyse enzymen wordt gecompliceerd doordat in veel organismen deze eiwitten een reeks secundaire cellulaire functies hebben die niet gerelateerd zijn aan hun katalytische rol in de glycolyse. Bakkersgist, *Saccharomyces cerevisiae*, is een belangrijk organisme in het onderzoek naar deze metabole route vanwege zijn robuustheid, genetische toegankelijkheid, gemak van cultivatie en hoge expressie van glycolyse enzymen. Deze gist wordt toegepast in de productie van voedsel en dranken, en als industrieel gastheerorganisme voor onder andere bio-ethanol en insuline productie, maar is ook van groot belang als modelorganisme voor fundamenteel onderzoek. Hoewel de glycolyse van *S. cerevisiae* de eerst beschreven metabole route is, zijn niet alle aspecten volledig opgehelderd, waaronder de regulatie, die grotendeels via remming en activatie door metabolieten plaatsvindt, en de complexe secundaire functies. Het bestuderen en aanpassen van het metabolisme in eukaryoten zoals gist wordt verder bemoeilijkt doordat er voor veel reacties meerdere iso-enzymen aanwezig zijn. In *S. cerevisiae* worden de tien reacties van de glycolyse samen met de twee reacties van pyruvaat naar ethanol uitgevoerd door 26 enzymen en de coderende 26 genen liggen verspreid over het gistgenoom. Om de studie van de gist glycolyse eenvoudiger te maken is in eerder werk de noodzakelijkheid van de verschillende iso-enzymen in kaart gebracht en zijn de overtollige enzymen verwijderd, wat geen invloed bleek te hebben op de fysiologie van deze ‘geminimaliseerde glycolyse’ giststam. Vervolgens zijn de noodzakelijke 13 iso-enzymen verplaatst naar een enkele locus, in de zogenaamde ‘SwYG’ stam (Switchable Yeast Glycolysis), waardoor het mogelijk werd om deze route in enkele stappen te verwisselen voor andere varianten. Deze ‘minimale glycolyse’ en ‘verwisselbare glycolyse’ platformstammen vormen de basis voor deze thesis. De doelen van dit werk waren het onderzoeken van transplantatie van metabole routes als methode om de glycolyse te modificeren en te bestuderen en het in kaart brengen van de metabole en fysiologische invloed van zulke verwisselingen in *S. cerevisiae*.

Hoewel het minimaliseren van de glycolyse geen meetbare effecten had op de fysiologie van gist, resulteerde het verplaatsten van de essentiële glycolyse genen naar één locus in een 8-30% lagere groeisnelheid van de SwYG stam, afhankelijk van de condities. In **Hoofdstuk 2** worden de effecten van genetische reorganisatie op de gistfysiologie, en specifiek de lagere groeisnelheid, bestudeerd. Verschillende aspecten van de verplaatsing en reorganisatie van deze metabole route zijn onderzocht, waaronder de aanwezigheid van replicatie-sequenties en de transcriptie-richting van genen. Daarnaast zijn de veranderingen in glycolyse activiteit en transcriptie van nabijgelegen genen onderzocht, maar geen van deze factoren leek verband te houden met de verminderde groeisnelheid van deze stam. De groeisnelheid kon wel hersteld worden door gerichte evolutie toe te passen en de gevonden genetische wijzigingen terug te plaatsen. Enkele van deze genetische wijzigingen waren onverwacht gerelateerd aan autofagie processen. Hoewel deze studie niet eenduidig de oorzaken van het groeisnelheidsverschil kon identificeren, konden onbedoelde mutaties en een niet eerder beschreven secundaire functie van Pgk1 aangewezen worden als mogelijke oorzaken. Dit hoofdstuk laat de uitdagingen zien van uitgebreide genetische modificatie, daarnaast leidde de analyse van de invloed van genetische reorganisatie tot ontwerprichtlijnen voor modulaire metabole routes.

Het tot expressie brengen van menselijke genen in gist (vermenselijking) wordt al lang toegepast voor de studie van eiwitten en routes omdat het aanpassen en karakteriseren van gist eenvoudiger is in vergelijking tot zoogdiercellen. **Hoofdstuk 3** beschrijft de vermenselijking van de glycolyse in gist. De combinatie van de complementatie (vervanging) van enkelvoudige genen, het verwisselen van de complete glycolyse, en gerichte laboratoriumevolutie bleek een krachtige methode om nieuwe modelorganismen te verkrijgen en de conservatie van enzymfunctie te bestuderen. De 25 geteste menselijke glycolyse enzymen waren grotendeels functioneel in de vermenselijkte giststammen, met uitzondering van de menselijke spier-hexokinases. Deze bleken alleen functioneel te zijn in gist na de verkrijging van mutaties die de inhibitie door glucose-6-fosfaat verminderden. Deze vereiste voor verminderde inhibitie was zowel aanwezig in de enkele hexokinase complementatie stammen als in de stammen met een volledige menselijke glycolyse, en liet de fundamentele verschillen in werking zien tussen menselijke (spier-) cellen en gistcellen. Een verrassende vinding was verder de complementatie door menselijke hexokinase, aldolase en enolase enzymen van de secundaire functies van de gistenzymen. De werking van de humane glycolyse in *S. cerevisiae* werd vervolgens geoptimaliseerd door de overexpressie van enkele glycolyse genen en laboratoriumevolutie, waardoor hexokinase en fosfoglyceromutase als voornaamste limiterende stappen konden worden aangewezen. Ook werden aanwijzingen

gevonden voor de interactie van menselijke aldolase en het gist cytoskelet. Vergelijking van de in gist geproduceerde menselijke enzymen met de enzymen uit een cultuur van menselijke spiercellen liet verder zien dat de katalytische efficiëntie grotendeels vergelijkbaar was, wat de potentie aantoont van vermenselijkte gistcellen als nieuwe modellen voor onderzoek naar de menselijke glycolyse.

Met gebruik van de lessen uit de vorige hoofdstukken, gaat **Hoofdstuk 4** over het gebruik van route-verwisseling om de metabole regulatie van de gistglycolyse te onderzoeken. Zulke regulatie is wijdverspreid in verschillende organismen en wordt gezien als belangrijk voor de dynamische regulatie van de glycolyse flux in *S. cerevisiae*. Echter, studies naar enkele enzymen hebben tot dusver de rol van metabole regulatie niet eenduidig geïdentificeerd. Door transplantatie van niet-gereguleerde glycolyse enzymen uit de vetzuur-ophopende gist *Yarrowia lipolytica* naar *S. cerevisiae* konden vragen over metabole regulatie worden getest. De combinatie van volledige en gedeeltelijke vervanging samen met complementatie van enkele enzymen toonde aan dat enkele enzymen konden worden vervangen door niet-gereguleerde enzymen zonder grote effecten, maar dat de combinatie van expressie van de *Y. lipolytica* glucokinase, fosfofructokinase en pyruvaatkinase tot een dysfunctionele glycolyse leidde. Deze gedereguleerde glycolyse stammen waren overgevoelig voor glucose, maar niet voor galactose, wat suggereerde dat het gebrek aan metabole regulatie een disbalans tussen de bovenste en onderste helft van de glycolyse veroorzaakte. Laboratorium evolutie onthulde dat een vermindering van de glucokinase activiteit dit effect kon onderdrukken. Een kinetisch model samen met groei bij lagere glucose concentraties bevestigden dat een reductie van flux in het bovenste deel van de glycolyse de balans kon herstellen in afwezigheid van de *S. cerevisiae* metabole regulatie. Deze waarnemingen demonstreren de synergetische rol van metabole regulatie van hexokinase, fosfofructokinase en pyruvaatkinase in *S. cerevisiae* onder dynamische omstandigheden, specifiek bij de overgang tussen lage en hoge glucose concentraties. Dit hoofdstuk laat de mogelijkheden van route-verwisseling zien om fundamentele aspecten van de glycolyse te bestuderen en toont het belang van systemische methoden op route-niveau om regulatiemechanismen te onderzoeken.

Als laatste kijkt **Hoofdstuk 5** verder dan de synthetische biologie in gist naar het bredere veld van synthetische genomica (synthetic genomics), waarin de constructie van synthetische chromosomen en genomen centraal staat. Lessen uit het verwisselen van een enkele route zouden kunnen worden toegepast voor het modificeren van het metabolisme op grotere schaal, waaronder het veranderen van het centrale koolstofmetabolisme en het implementeren van specifieke product routes. Dit soort ‘metabolic engineering’ zou eenvoudiger kunnen worden door het toepassen van synthetische chromosomen, die een hoge mate van flexibiliteit voor

de integratie, verwisseling en herschikking van genen voor metabole routes mogelijk maken. De huidige methoden voor de synthese van DNA zijn echter beperkt tot korte fragmenten, waardoor zelfs voor een enkel gen meerdere fragmenten aan elkaar moeten worden gezet. De huidige synthetische chromosomen bevatten al duizenden baseparen en tientallen genen en toekomstige synthetische genomen zullen nog groter moeten zijn. Methoden voor assembleren van DNA fragmenten zijn dus kritiek voor het succes van de synthetische genomica en de synthetische biologie. In dit hoofdstuk wordt een overzicht gegeven van het gebruik van *S. cerevisiae* en zijn krachtige, inherente homologe recombinatie voor de assemblage van synthetische DNA fragmenten. Het gebruik van dit veelzijdige micro-organisme is essentieel geweest voor de ontwikkeling van het veld tot dusver, en heeft de constructie van het eerste synthetische genoom mogelijk gemaakt. Het gebrek aan efficiënte alternatieve DNA assemblage methoden en de intrinsieke limieten van bacteriële organismen voor synthetische biologie betekenen dat *S. cerevisiae* waarschijnlijk een belangrijke rol zal behouden in de constructie van synthetische chromosomen en genomen.

Chapter 1

General Introduction

Structural diversity of glycolytic pathways

The Emden-Meyerhof-Parnas pathway

All life forms primarily consist of organic molecules, characterized by carbon-carbon bonds. To take up molecules containing carbon from the environment and convert them into the lipids, proteins and nucleic acids that make up biomass, and extract energy from conversion of carbon molecules, life has evolved a network of enzyme-catalysed reactions, known as central carbon metabolism. Important cellular building blocks and carbon sources in many environments are hexose sugars such as glucose, galactose and fructose. These primary products of photosynthesis can be degraded for use as source of carbon, electrons and energy by many forms of life using a set of reactions known as glycolysis. Glycolysis is usually defined from glucose to the three-carbon organic acid pyruvate, which can subsequently be subject to different fates, including further oxidation, incorporation in biomass, or functioning as electron acceptor to maintain redox balance, depending on the needs and environment of the cell. A variety of glycolytic pathways has evolved, shaped by different cellular requirements and environments across the domains of life, and while most variants conserve energy in the form of ATP via substrate-level phosphorylation, these pathways differ in their component enzymes and ATP yield. In addition, glycolysis is an amphibolic pathway, where the same reaction steps can operate in the opposite direction (gluconeogenesis), from three-carbon intermediates to sugar phosphates, with input of free energy. These sugar-phosphates fulfil essential roles as cellular building blocks or can be converted to storage metabolites such as glycogen and starch even in organisms that don't use extracellular hexoses as primary carbon and energy source.

The most intensively studied glycolytic pathway is the Emden-Meyerhof-Parnas (EMP) pathway, shown in detail in Fig. 1. The abundance of knowledge on this pathway is largely explained by its high level of conservation among eukaryotes, including humans, but also by the fact that glycolysis was initially discovered and characterized in the model eukaryote *Saccharomyces cerevisiae* [1], popularly known as bakers' yeast. Besides in eukaryotes, the EMP pathway is also found in many bacteria and some archaea (though usually with some variations, discussed below). The EMP pathway consists of 10 biochemical reactions that convert one mole of glucose into two moles of pyruvate and generate two moles of reduced NADH and two moles of ATP. In the first three steps of the pathway, glucose is phosphorylated by the investment of two ATP molecules producing fructose-1,6-bisphosphate, subsequently this phosphorylated sugar is cleaved in the centre, generating two phosphorylated three-carbon moieties. These two compounds, glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP) are interconverted by the triose phosphate isomerase (TPI). In the lower part of the

pathway, which is largely reversible, glyceraldehyde-3-phosphate is oxidized, with the incorporation of inorganic phosphate, generating 1,3-bisphosphoglycerate and a reduced NAD(P)H moiety. The two phosphate groups are then used to conserve free energy in the form of ATP, a mechanism known as substrate level phosphorylation. The energy generated from substrate level phosphorylation enables this pathway to support growth in the absence of respiration, provided the two produced NADH moieties are re-oxidized. For example, yeasts convert pyruvate to acetaldehyde, which is reduced to form ethanol, while lactic acid bacteria and human muscle cells directly reduce pyruvate to form lactic acid. Besides these examples, many more fermentative reactions from pyruvate are described in diverse organisms, some contributing to substrate-level phosphorylation, (see [2] for a concise overview of possible reactions).

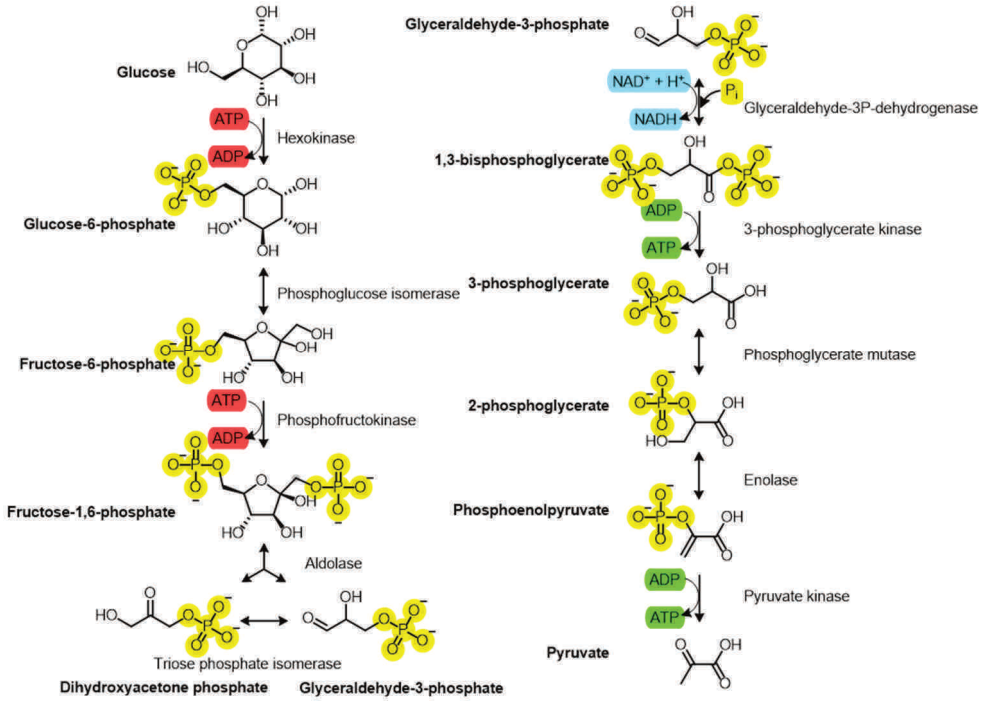


Figure 1 - The Emden-Meyerhof-Parnas pathway of glycolysis.
The EMP glycolytic pathway is shown, on the left the upper glycolysis, on the right the lower 'trunk' pathway. For every glucose molecule, two molecules of glyceraldehyde-3-phosphate are produced. Phosphate groups are indicated in yellow, ATP consuming reactions in red, ATP producing reactions in green, redox reaction in blue.

Alternative reactions and pathways

The classical EMP pathways is the textbook standard for glycolysis, but not the only biochemical pathway to convert hexoses to pyruvate. Biological constraints on the chemical properties and toxicity of intermediates as well as the feasibility of enzyme-catalysed reaction mechanisms do limit the possibilities for chemical pathways between glucose and pyruvate [3, 4]. Especially the lower part of the EMP pathway between glyceraldehyde-3-phosphate and pyruvate, the so-called trunk pathway, seems to have few biochemically viable alternatives [5]. Indeed most enzymes in this part of the pathway (triose-phosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and enolase) are highly conserved among kingdoms and are suspected to have been present in the last universal common ancestor (LUCA). This ancient pathway possibly evolved originally as part of a gluconeogenic pathway [6]. Corresponding to their ancient origins, evolutionary rates are slow for the enzymes of lower glycolysis, while more variation and faster evolutionary rates are found for upper glycolytic enzymes, especially the irreversible steps involved in regulation of the glycolytic flux, hexokinase and phosphofructokinase [7].

Variations on the EMP pathway as described above appear in various organisms (overview shown in Fig. 2), one common variation is the nature of the phosphate donor for the phosphorylation of glucose and fructose in upper glycolysis, which can affect the glycolytic ATP yield. In many bacteria, including *E. coli*, a membrane-bound phosphotransferase system (PTS) which uses PEP instead of ATP as phosphate group donor, couples glucose entry into the cell to its phosphorylation. This does not change the overall ATP yield of glucose assimilation, but has implications for the regulation of the glycolytic flux, since PEP is required for the start-up of glycolysis, and cells have to adjust their metabolite levels accordingly [8]. PTS-containing bacteria often also contain an ATP-dependent glucokinase, possibly to offer metabolic flexibility [9]. More strikingly different variants of the EMP pathway are found in archaea and thermophilic bacteria, which often have the typical EMP pathway structure but use ADP or pyrophosphate instead of ATP to phosphorylate glucose and fructose-6-phosphate, thereby potentially increasing the glycolytic ATP yield [10, 11]. Similarly, pyrophosphate-dependent phosphofructokinases are commonly found in plant cells cytosol, although ATP-dependent variants are also present [12]. Variations also exist in the more conserved lower part of glycolysis, for instance non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) or ferredoxin-dependent glyceraldehyde-3-phosphate oxidoreductase (GAPOR) can replace two steps of the standard pathway (glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase), leading to 3-phosphoglycerate formation from glyceraldehyde-3-phosphate without ATP generation from ADP.

Additionally pyruvate phosphate dikinase (PPDK) and pyruvate water dikinase, which use AMP and pyrophosphate/inorganic phosphate for pyruvate formation are sometimes found instead of pyruvate kinase [10, 13]. These pathway variations can change the ATP yield and in the case of GAPN and GAPOR, the redox cofactors generated (NADPH or Fd^{red} instead of NADH), changing the EMP pathway from generating two ATP to one without net ATP yield (with GAPN or GAPOR) or even a higher ATP yield (with ADP or PP_i dependent enzymes). These changes in pathway stoichiometry also alter the thermodynamic driving force of the pathway and are thought to result from an optimization specific growth conditions and requirements of the organisms in which these variants are found [10].

The Entner-Doudoroff (ED) pathway is a widely-spread variation of the glycolytic pathway among bacteria and archaea. The ED pathway is composed of two enzymes that catalyse the conversion of one molecule of 6P-gluconate, which can be derived from glucose-6-phoshate or gluconate, into one molecule of pyruvate and one molecule of glyceraldehyde-3-phosphate. Glyceraldehyde-3-phosphate is further metabolized via the same lower glycolysis that is part of the EMP pathway, leading to one ATP per glyceraldehyde-3-phosphate molecule. The direct production of a pyruvate molecule from glucose-6-phosphate shortcuts lower glycolysis and thereby reduces the ATP yield from two to one mole of ATP per mole of glucose as compared to the EMP pathway. Besides functioning as the main pathway for hexose utilization in many bacteria and archaea, the ED pathway also acts as specialized pathway for catabolism of gluconate and some other carbohydrates next to the EMP pathway, for example in *Escherichia coli* [9, 14]. The ED pathway is sometimes referred to as an example of a lesser evolved ‘paleo-metabolism’ since it has a lower yield compared to the EMP pathway [4]. However, despite its lower ATP yield, the ED pathway, less demanding in terms of protein cost than the EMP pathway, might be more optimal under certain conditions. This theory is supported by the preferential catabolism of glucose via the ED pathway in some organisms that harbour both the EMP and ED pathway [15]. Several variations on the ED pathway have been identified, including non-phosphorylative and semiphosphorylative variants mainly found in archaea [10, 16]. A particular, cyclic pathway found in *Pseudomonads*, combines the ED, pentose phosphate (PP) and EMP pathways [14, 17]. In *Pseudomonas putida*, glucose is first converted to gluconate and then metabolized through the ED pathway. The synthesis of intermediates of upper glycolysis, required for biosynthetic purposes, is catalysed via parts of the PP and reversed EMP pathways. This cyclic pathway is thought to endow *Pseudomonads* with metabolic flexibility and oxidative stress resistance, important features for survival and growth in their environmental niche [17, 18].

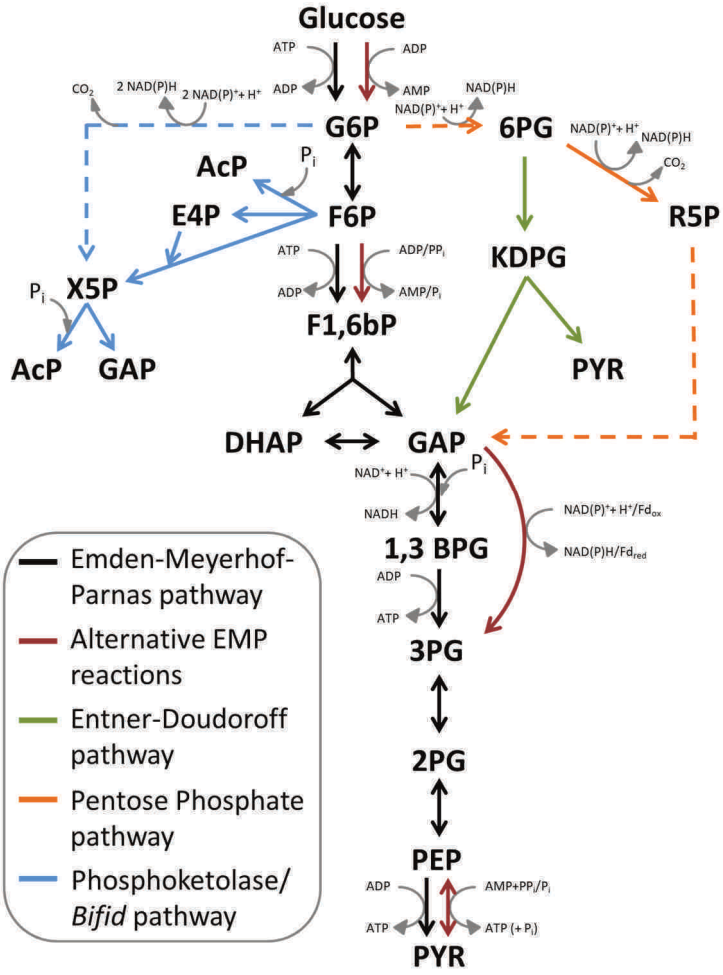


Figure 2 - Alternative pathways of glycolysis.
Schematic representation of variations in the glycolytic pathway. Reactions catalysed by multiple enzymes are indicated as dashed lines. In black the EMP pathway as present in *e.g.* humans and *S. cerevisiae*, in red the alternative reactions found in the EMP pathway in archaea and some plants and bacteria. In green the Entner-Doudoroff pathway, present in many bacteria. In orange the pentose phosphate pathway, present in most organisms but rarely used as main glycolytic pathway. In blue the phosphoketolase and *bifidobacterium* shunt pathways found in some prokaryotes. Stoichiometry of the sugar rearrangement reactions in the pentose-phosphate and phosphoketolase pathways is not indicated. Some pathways are not shown including the semi-phosphorylative and non-phosphorylative ED pathways (mainly found in archaea and some prokaryotes) and pathways via glucose dehydrogenase (gluconate) found in e.g. *Pseudomonas* species. Abbreviations: G6P: glucose-6-phosphate; F6P: fructose-6-phosphate; F1,6bP: fructose-1,6-bisphosphate; DHAP: dihydroxy-acetone phosphate; GAP: glyceraldehyde-3-phosphate; 1,3BPG: 1,3-bisphosphoglycerate; 3PG: 3-phosphoglycerate; 2PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate; 6PG: 6-phospho-gluconate; KDPG: 2-keto-3-deoxy-6-phosphogluconate; R5P: Ribulose-5-phosphate; AcP: Acetyl-phosphate; E4P: Erythrose-4-phosphate; X5P: Xylulose-5-phosphate.

Besides these two most widely spread glycolysis variants (EMP and ED), some other variations, mainly built on components of the PP, EMP and ED pathways, have been described in bacteria and archaea (see [2] and [10] for comprehensive reviews). The pentose phosphate pathway can convert glucose-6-phosphate to glyceraldehyde-3-phosphate with the release of CO_2. While in most organisms this route is used mainly as a way to produce NADPH equivalents for biosynthesis, it can replace glycolysis in some cases [2, 19]. The phosphoketolase (also called heterolactic) pathway splits the PP intermediate xylulose-5-phosphate into acetyl-phosphate and glyceraldehyde-3-phosphate, which can then be used in lower glycolysis. A related pathway known as the *Bifidobacterium* shunt (or bifid pathway) uses a broader spectrum phosphoketolase that is also active on fructose-6-phosphate, which enables degradation of glucose to acetyl-phosphate and pyruvate without CO_2 production [20].

In general, the structure of the glycolytic pathway is dependent on the metabolic needs and ecological niche of each organism. However, due to constraints, there is strong conservation of glycolysis across kingdoms, especially in the bottom section of the pathway, which is considered the most ancient part. The Emden-Meyerhof-Parnas pathway, topic of this thesis, is the main route for hexose utilisation in many organisms across the tree of life, including humans and many (micro)organisms relevant for food production and biotechnology.

Regulation of the glycolytic pathway

The importance of key-points

The glycolytic pathway does not operate in isolation, but is part of the larger cellular metabolic network and is connected to the extracellular environment. Since both the extracellular environment and the requirements for energy and building blocks can fluctuate in time, the glycolytic flux needs to be flexible to allow cells to survive and grow [21]. A variety of mechanisms have evolved to allow changes of enzyme activities on different time scales. A relatively slow way of adjusting the flux is through transcriptional regulation of the glycolytic genes. This type of regulation is often present on irreversible glycolytic and gluconeogenic enzymes, simultaneous operation of which would lead to futile cycles and should be prevented [7]. Regulation of the glycolytic enzymes abundance can also be regulated by mRNA stability and possibly even mRNA localization [9, 22] as well as protein translation and degradation [23]. For fast responses to environmental changes, other mechanisms have evolved that can act faster, by direct modification of the enzyme activity. Oligomerization can play an important role in regulation of glycolytic activities as most glycolytic enzymes operate as oligomers, which can change stability upon binding of other proteins or metabolites [7]. Binding to cellular structures, such as

the cytoskeleton or mitochondria, and re-localization are also thought to regulate activity of some glycolytic enzymes [24]. Another mechanism leading to direct changes in enzyme activities *in vitro* is covalent post-translational modification of enzymes, such as phosphorylation, which is wide-spread in eukaryotic central carbon metabolism [25]. Such post-translational modifications can be both activating or inhibiting and can also cross-talk with other regulations, for example, phosphorylation can cause re-localization of an enzyme, or change its degradation [26]. These types of regulation are difficult to measure in controlled *in vitro* systems, and knowledge on their importance in glycolytic regulation is therefore far from complete. The most direct and, in many systems, most important mechanism controlling glycolytic activities on short timescales is regulation by metabolites. Metabolite concentrations can affect the activity and kinetic properties of enzymes either through mass-action effects (substrates and products) or through binding to allosteric sites.

In each organism and cell type, regulation of the glycolytic flux occurs through some combination of all of these mechanisms, and this regulation can be subject to change over time, for instance by expression of different isoenzymes with different allosteric properties. Despite this large variety, some general trends in regulation have been observed that translate across different organisms. Not all reactions of the glycolytic pathway are equally targeted by regulations, instead, key-points identified across different systems are the irreversible ATP- and ADP-linked steps of glucose phosphorylation, fructose-6-phosphate phosphorylation and phosphoenolpyruvate dephosphorylation, and the enzymes catalysing those reactions (hexokinase/glucokinase/PTS, phosphofructokinase and pyruvate kinase respectively) are often sensitive to fast-acting metabolic regulation. Interestingly, in modified EMP routes in archaea, where ADP and PPi dependent hexokinases and phosphofructokinases catalyse reversible reactions and are not subject to regulation, instead such regulation is transferred to an irreversible step such as GAPN [10, 16]. Apart from these classical regulation points, glucose transport and product export can also be control points in some cases, for instance in immortalized human cell lines [27]. Control is therefore mainly centred on the irreversible reactions. The amphibolic nature of the glycolytic pathway means that the irreversible enzymes of glycolysis and gluconeogenesis need to be strictly regulated to avoid futile cycling and ATP dissipation which would occur if both were active simultaneously [28]. Another reason regulation of some steps is of particular importance is the possibility of metabolic imbalances resulting from the EMP pathway architecture, where ATP is invested in the first part of the pathway, but only recovered later. If a situation occurs where the first half of the pathway runs faster than the second half, so the pathway temporarily consumes more ATP than it produces, imbalances can occur that can

lead to a lethal depletion of cellular ATP [29]. The balancing of upper and lower glycolytic fluxes can therefore be considered another target of evolution of glycolytic regulation, as suggested by the frequently observed tight regulation of hexokinase and phosphofructokinase. In line with the central importance of ATP balancing, ATP itself is often involved in regulation of glycolytic enzyme activities, and in many systems, ATP demand can be seen as the main factor controlling glycolytic flux [21, 30]. This does however not hold for all organisms, in *Lactococcus lactis* for instance, the redox status seems to be a major regulator of the glycolytic flux instead of ATP, but how this control is exerted is not completely elucidated [8, 31, 32].

Despite decades of research into the control of glycolysis in various organisms, the understanding of how the flux is regulated and which regulations are important is not complete, even in model organisms. This incomplete understanding is underlined with the difficulty of generating predictive *in silico* kinetic models of glycolysis from biochemical data [33-35]. New mechanisms that modulate activities of glycolytic enzymes continue to be discovered, such as new metabolic regulations [36], a polymerization mechanism controlling activity [37] and indications of metabolon formation, linking multiple glycolytic enzymes together [38]. For such newly discovered mechanisms, as well as many previously discovered but not quantitatively understood types of regulation, their importance on *in vivo* flux regulation remains difficult to estimate.

Regulation of glycolysis in various eukaryotes

Below the regulation in three thoroughly investigated eukaryotic systems is described in detail to give a more complete picture of how the regulation is connected to the function of glycolysis in each organism. The chosen systems are yeast, mainly *Saccharomyces cerevisiae*, human muscle and the parasitic trypanosomes, which each have different regulatory mechanisms but share some marked similarities.

Yeast

The regulation of the glycolytic pathway has been most thoroughly studied in baker's yeast. *Saccharomyces cerevisiae* has a long history of use by humankind for products such as beer, wine and bread, for which the fast uptake of sugars and the production of fermentation products ethanol and CO_2 are critical. As a Crabtree-positive yeast, *S. cerevisiae* preferentially uses glycolysis coupled to ethanol fermentation as main catabolic pathway, when glucose is supplied in excess, even in presence of oxygen. Because glycolysis coupled to fermentation has a low ATP yield as compared to respiration, *S. cerevisiae* requires a high glycolytic flux to maintain fast growth rates. The evolutionary reasons for this metabolic strategy have been debated as there are multiple potential benefits [39]. Ethanol production could offer a competitive advantage in natural environments, because *S. cerevisiae* has a

high ethanol tolerance [40]. Additionally, the produced ethanol and organic acids can subsequently be consumed after glucose depletion, so carbon is not lost, but 'stored' in another form. Alternatively, glycolytic ATP production coupled to fermentation might simply be more efficient in terms of protein costs compared to respiration [41, 42].

On a pathway level, the metabolic strategy of *S. cerevisiae* means that the glycolytic proteins are highly expressed, making up approximately 7-13% of the proteome [41]. The high glycolytic capacity makes *S. cerevisiae* potentially sensitive to imbalances in the pathway, and various regulations have evolved on the glycolytic enzymes. Although transcriptional regulators have been described, the abundance of glycolytic proteins remains high under varying conditions and transcriptional regulation is therefore not thought to be a major regulatory mechanism to adjust the glycolytic flux on short timescales [23, 43]. Underlining this, overexpression of glycolytic enzymes does not increase the flux significantly [44, 45]. Instead, the major mechanism controlling glycolytic activities and glycolytic flux on short timescales is metabolic regulation on a few key enzymes.

S. cerevisiae hexokinase 2, the major isoenzyme in high glucose conditions, is strongly inhibited by trehalose-6-phosphate [46]. This metabolite is an intermediate of trehalose biosynthesis, produced from glucose-6-phosphate by trehalose-6-phosphate synthase (Tps1). The next regulated step, phosphofructokinase, is sensitive to multiple effectors, which sense energy status and substrate supply. The major regulatory interactions are ATP inhibition, AMP and ADP activation and activation by fructose-2,6-bisphosphate [35, 47]. Additionally the substrate fructose-6-phosphate binds cooperatively, and the product fructose-1,6-bisphosphate decreases the activation by fructose-2,6-bisphosphate [48]. Fructose-2,6-bisphosphate is a regulatory metabolite, generated by 6-phosphofructo-2-kinases and broken down by specific phosphatases, which are themselves subject to allosteric and transcriptional regulation [49-51]. Fructose-2,6-bisphosphate, besides activating phosphofructokinase, also inhibits the opposite reaction catalysed by fructose-1,6-bisphosphatase, and might thereby prevent simultaneous gluconeogenesis and glycolysis [21, 51]. The other major allosteric regulation is feed-forward activation of pyruvate kinase 1, the major isoform, by fructose-1,6-bisphosphate [52]. These regulations are thought to balance the upper and lower halves of the pathway by inhibition of hexokinase and phosphofructokinase and activation of pyruvate kinase, while sensing ATP demand and energy state, mainly through effectors of phosphofructokinase. See Fig. 3 for a schematic overview.

Even though the regulation of yeast glycolysis has been under scrutiny since its discovery, the picture is not yet complete. Proving the *in vivo* function of any of the

above described regulations has been very difficult, and removal of any single regulatory mechanism though mutations and deletions does not quantifiably alter physiology. For example, mutations on the enzymes modifying the fructose-2,6-bisphosphate concentration did not lead to drastic changes in glycolytic or gluconeogenic fluxes [49]. Similarly replacement of each of the key-points enzymes with heterologous or mutated variants lacking allosteric regulations has revealed limited effects on the flux and robustness of yeast cells [53-56]. One exception where a single deletion causes a dramatic deregulation of the glycolytic flux is the deletion of the gene encoding trehalose-6-phosphate synthase, *TPS1*. Mutants deficient for this enzyme fail to grow on glucose and accumulate hexose phosphates, leading to the hypothesis that the inhibition of hexokinase by its product, thehalose-6-phosphate, is critical for a balanced glycolytic pathway [57]. However, more detailed investigation revealed that alleviation of hexokinase inhibition was not the main mechanism of deregulated glycolysis in *tps1* deletion mutants, since trehalose-6-phosphate insensitive heterologous hexokinases can complement native yeast hexokinases. Instead the effect of Tps1 and trehalose-6-phosphate seems to be exerted via a combination of hexokinase inhibition and freeing up phosphate through the trehalose cycle [58]. Additionally, Tps1 and its product trehalose-6-phosphate have been implicated in transcriptional regulation, cellular signalling, apoptosis, and effects on intracellular pH suggesting regulatory roles outside glycolysis, which have not been fully elucidated to date [59-61]. The metabolic regulation of yeast glycolysis has been summarized in kinetic models [35, 58, 62]. These models are useful collections of biochemical data and have been used to successfully generate hypotheses, but are incomplete and when matched against experimental data sometimes lack predictive power, illustrating not all aspects of glycolytic regulation are currently understood [34].

In other yeasts glycolysis has not been studied in the same level of detail as in *S. cerevisiae*, but some marked similarities and differences have been found. In other Crabtree-positive yeasts, such as *Schizosaccharomyces pombe*, at least some of the metabolic regulations are similar. The *S. pombe* phosphofructokinase is similarly activated and inhibited by multiple metabolites, such as ATP and fructose-2,6-bisphosphate [63]. However, the *S. pombe* hexokinase is not regulated by trehalose-6-phosphate and deletion of the trehalose-6-phosphate synthase does not have a major effect on glycolysis, suggesting a more limited role of the trehalose cycle in flux regulation in this yeast [64]. Additionally the *S. pombe* pyruvate kinase is not activated to the same extent as the *S. cerevisiae* enzyme by fructose-1,6-bisphosphate. Even more differences are found in more distantly related yeasts, such as the oleaginous yeast *Yarrowia lipolytica.* This yeast relies mostly on a glucokinase insensitive to trehalose-6-phosphate, but additionally has a very

insensitive phosphofructokinase, which is only inhibited by phosphoenolpyruvate [65, 66]. In addition, its pyruvate kinase is seemingly not metabolically regulated [67]. These differences illustrate the optimal metabolic regulations of glycolysis will likely depend strongly on environment and substrate preference.

Human muscle

In metazoans, glycolytic regulation differs between cell types, since some cell types will require high glycolytic fluxes, such as muscle cells, while others have low glycolytic rates or switch between gluconeogenesis and glycolysis, such as liver cells. Expression of different genes and splicing isoforms generates a diverse set of glycolytic genes with different regulatory properties to cater to the regulatory requirements of different cell types. In human muscle cells, the glycolytic flux needs to change quickly to respond to the energy demands of exercise and this is achieved by metabolic regulation at the level of glucose phosphorylation (hexokinase/glucokinase) and phosphofructokinase [68, 69]. Glycogenolysis, the breakdown of glycogen, is the main source of glucose-6-phosphate during exercise and the glycogen phosphorylase is tightly regulated by a combination of endocrine signals, muscle activation (via pH and Ca^{2+}) and metabolites such as AMP and glucose-6-phosphate. Similarly the entry of glucose into the cells is regulated via localization of glucose transporters, and muscle hexokinases (isoforms 1 and 2) are sensitive to product inhibition by glucose-6-phosphate [69, 70]. Phosphofructokinase is a major regulatory hub, being sensitive to many effectors that sense cellular energy state (inhibition by ATP and activation by AMP and ADP), lactate accumulation (inhibition at low pH), TCA cycle flux (inhibition by citrate) and substrate supply and endocrine signals (activation by fructose-2,6-bisphosphate) [24, 69]. Additionally the enzyme appears to be regulated by complex phosphorylation, oligomerization and association with the cytoskeleton [24]. While the impact of all individual regulations *in vivo* is not fully understood, maintaining ATP supply under varying demand while balancing upper and lower glycolysis seems to be the major reason for the presence of these various (metabolic) regulations [21, 69, 71].

Trypanosomes

The requirement for some type of fast acting regulation of glycolytic enzyme activities seems to be conserved across different organisms, especially in cell types with high glycolytic capacities, although the specific regulatory mechanisms themselves are not [7]. An exception to this requirement is found in the glycolysis of trypanosomes, unicellular parasitic eukaryotes. *Trypanosoma brucei*, which causes African sleeping disease, can metabolize glucose at high rates in its bloodstream form, relying on glycolysis for most of its ATP production. Yet unusually, most typical eukaryotic glycolytic regulations are absent, including metabolic regulation on hexokinase and

phosphofructokinase [72]. Pyruvate kinase is sensitive to allosteric activation by fructose-2,6-bisphosphate, which acts on phosphofructokinase in most eukaryotes. However, all glycolytic enzymes from hexokinase until phosphoglycerate kinase are compartmentalized in a separate organelle, named the glycosome, while phosphoglycerate mutase, enolase and pyruvate kinase are present in the cytosol [73]. The glycosome is thought to function by separating the glycolytic ATP consuming steps from the producing steps, preventing imbalance between the two phases of the pathway, because there is only a limited amount of ATP that can be consumed inside the glycosome [74]. This unique configuration allows the absence of allosteric regulation on the ATP consuming steps, while still balancing ATP consumption and production.

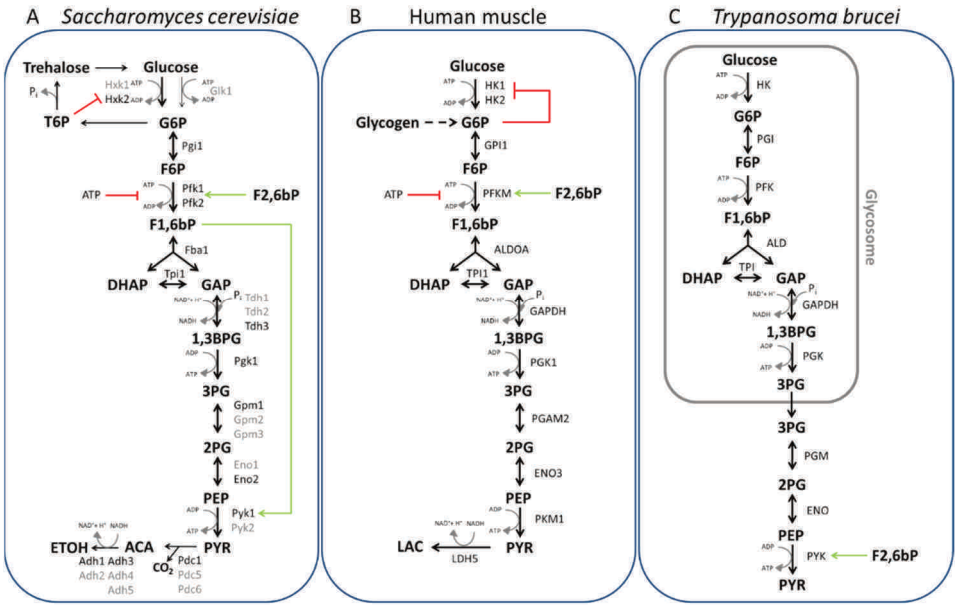


Figure 3 - Examples of different metabolic regulation strategies in glycolysis in eukaryotes. Schematic representations of the main metabolic regulations in glycolysis in three different eukaryotes. Inhibitions are shown with red lines, activations with green arrows. **A)** *S. cerevisiae* major isoenzymes are shown in black, minor isoenzymes shown in grey. For the hexokinase and pyruvate kinase regulation is different between the isoenzymes as indicated. **B)** Human muscle glycolysis, major isoenzymes are shown. **C)** *T. brucei* glycolysis is unique in its lack of metabolic regulation but instead is partially contained in a semi-permeable organelle called the glycosome. Abbreviations: G6P: glucose-6-phosphate; F6P: fructose-6-phosphate; F1,6bP: fructose-1,6-bisphosphate; DHAP: dihydroxy-acetone phosphate; GAP: glyceraldehyde-3-phosphate; 1,3BPG: 1,3-bisphosphoglycerate; 3PG: 3-phosphoglycerate; 2PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate; T6P: trehalose-6-phosphate; F2,6bP: fructose-2,6-bisphosphate; ACA: acetaldehyde; ETOH: Ethanol; LAC: lactate.

Moonlighting functions of glycolytic enzymes

In many (eukaryotic) organisms, glycolytic enzymes have additional cellular roles next to their catalytic function in glycolysis. Such secondary functions, known as moonlighting activities, are often completely unrelated to the main function of the enzymes, and are thought to have evolved after the primary metabolic function. Because of this more recent emergence, there is large variation in these functions between organisms. While glycolytic enzymes are not unique in having such secondary functions, highly conserved proteins, such as other enzymes of central carbon metabolism and ribosomal proteins seem to be implicated more often in moonlighting [75]. It has been hypothesized that their cytosolic localization and high expression levels, enabling efficient binding to other proteins, might make glycolytic enzymes specifically susceptible to evolving moonlighting functions [76]. Moonlighting functions have been described for most glycolytic enzymes, at least in some organisms, and especially for GAPDH and enolase a multitude of different functions have been described. These range from mitochondrial tRNA import (enolase, [77]), to mediating apoptosis (GAPDH, [78]), to acting as the structural lens protein in various vertebrates (described for both enzymes in different organisms, [79], [80]). Most of these functions are cytosolic, but some require relocation of the enzymes to other cellular compartments or even to the extracellular space, in case of multicellular and parasitic organisms [76, 81]. While glycolytic moonlighting functions are often not dependent on the catalytic function of enzymes, they are sometimes related to glucose metabolism, suggesting they might act as signal transducers for information on metabolic status. A well-studied example in *S. cerevisiae* concerns the role of Hxk2 in glucose repression, the transcriptional repression of genes involved in utilization of alternative carbon sources in the presence of glucose. When glucose is present, hexokinase 2 is partially relocated to the nucleus where it is part of a repressor complex that binds to DNA [82, 83]. This function is not directly related to the glucose phosphorylation activity of the enzyme, but the position of hexokinase as the first committed step of glucose utilization could be related to a function that controls transcription of other carbon-utilization pathways. Additionally the yeast aldolase enzyme, Fba1, has a secondary function in the assembly of the vacuolar ATPase, which is dependent on the presence of glucose [84, 85]. Interestingly the V-ATPases are highly conserved, and the interaction with aldolase was found in human cells and in yeast, as well as in yeast expressing human aldolase, even though the aldolases themselves share no significant homology [86].

Moonlighting functions of glycolytic enzymes can be important for normal cellular function and have potential roles in human health and disease. However most known moonlighting functions have been found ‘accidentally’ through association and

mutation studies, and then proven through detailed molecular investigations [75]. Detailed studies have not been undertaken in every suspected case of moonlighting, and there are undoubtedly many more to uncover. The fact that gene deletion or even changes in expression of glycolytic enzymes most often have deleterious effects on cellular metabolism, not only hampers the discovery of subtle phenotypic effects, but generally complicates study of the underlying mechanisms [82, 86]. Mutants deficient for either the moonlighting or the catalytic function are often challenging to construct and their construction requires knowledge on the mechanism and domains involved. Conversely, the presence of moonlighting activities can lead to unexpected effects upon metabolic engineering of the glycolytic pathway. Study of conservation of moonlighting functions between different organisms could help shed light on their mechanism of action, as well as their importance and evolution.

Engineering of the glycolytic pathway and its surrounding reactions

As the central pathway of sugar utilization, the glycolytic pathway in many organisms is highly relevant for microbial biotechnology. Not only is the production rate of fermentation products such as ethanol and lactate directly dependent on the glycolytic flux, but many other products are derived from phosphorylated sugar intermediates or from acetyl-CoA which is derived from pyruvate [87]. This has led to many attempts to engineer the glycolytic pathway and its surrounding reactions, ranging from overexpression of heterologous enzymes to complete rearrangements of the stoichiometry of the pathway. For ethanol production, attempts have been made to increase product formation rates by increasing the glycolytic flux in *Saccharomyces cerevisiae*. However, overexpression of single or multiple glycolytic enzymes in *S. cerevisiae* did not lead to increases in the glycolytic rate [44, 45]. Similarly, in other organisms attempts to directly increase the glycolytic flux have not always met with success [88], likely resulting from the complex multi-level regulation of this central pathway. Besides changes in expression of the pathway, engineering efforts have targeted expression of alternative reactions and pathways to change the stoichiometry of carbon metabolism. For example, a non-phosphorylating $NADP^+$-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPN) was expressed in yeast to reduce glycerol formation and improve ethanol production [89]. Another addition that allowed a change in the stoichiometry was the expression of the Calvin cycle enzymes phosphoribulokinase (PRK) and ribulose-1,5-bisphosphate carboxylase (Rubisco) which provide the ability to incorporate carbon from CO_2 using NADH, also reducing by-product formation [90]. Such rewiring of carbon metabolism

can be challenging, because it requires high activity of the expressed heterologous enzymes. Correspondingly, expression of the Entner-Doudoroff pathway enzymes in *S. cerevisiae*, which would enable modifying the ATP yield and redox cofactor production, has been attempted on several occasions but has not been successful to date [91, 92].

More complete rewiring and bypassing of glycolysis was obtained in bacteria, for example with the design and integration in *E. coli* of a novel pathway termed non-oxidative glycolysis (NOG) [93, 94]. This synthetic pathway uses various heterologous enzymes, including a phosphoketolase and was designed to avoid carbon loss as CO_2 and increase production of acetyl-CoA and products derived from it. Major rewiring of the *E. coli* glycolytic pathway was required, including deletion of GAPDH and PGK to allow efficient carbon flux through this pathway. Expression of heterologous phosphoketolases was similarly applied in yeast for improved production of acetyl-CoA derived products in *S. cerevisiae* [95] and *Yarrowia lipolytica* [96], although the extent of pathway rewiring was more limited. Another example of rewiring of the glycolytic pathway in *E. coli* showed that deletions in the EMP trunk pathway could be bypassed by two alternative routes, the methylglyoxal pathway and the serine shunt [97]. Both of these pathways contain toxic intermediates however, and the engineered bypasses did not lead to ATP production. Engineering of the many genes of glycolysis scattered around the genome can be challenging. To facilitate engineering, the *E. coli* EMP pathway was refactored in simplified expression cassettes that could be used in various Gram-negative bacteria, allowing the expression of the complete *E. coli* EMP pathway in two *Pseudomonas* species [98].

In most of the cases described, rewiring of glycolysis required extensive engineering and led to suboptimal growth, even if product formation was increased in some cases, reflecting the optimized nature of the pathway.

Yeast as a model organism

Besides its uses in food and beverage production and microbial biotechnology, the yeast *S. cerevisiae* is of interest as a model organism. Yeast is one of the most intensively studied eukaryotes and has been essential in the discovery and elucidation of many pathways and processes relevant for human health, including much of the glycolytic pathway [1]. Underlying this is its long history of use a fermenting agent in food and beverage production, but also ease of cultivation, GRAS safety status and ease of genetic manipulation. The intense study on yeast is reflected in the fact that it was the first eukaryote to be whole-genome sequenced, six-and-a-half years before the first human genome was reported [99]. Later

developments, such as the construction of the yeast gene deletion library [100, 101], and the improvement of genetic editing by the use of CRISPR-Cas9 endonuclease [102] have further accelerated the use of yeast as a model organism. Due to its eukaryotic nature, yeast cellular architecture is more similar to human cells than some other easily accessible microbial models (e.g. *E. coli*), leading to first uses as a model organism to study processes that are conserved between yeast and human [103]. Additionally, expression of human proteins in yeast has improved our understanding of many human proteins and their variants. Such humanization efforts vary from expressing single genes to complete humanization of pathways [104]. Single gene complementation in yeast has been used for functional annotation of human genes, for example, the function of the human mitochondrial voltage-dependant anion channel (VDAC) proteins was verified by expression in a yeast strain deleted for its native VDAC [105]. Such humanization of yeast genes allows the study of multiple gene variants in a simplified context, for example to examine the effect of mutations found in patients [106]. Humanized yeast strains can also be used to test the effect of drugs on specific human proteins *in vivo* [107].

A systematic study of complementation with 414 human-yeast gene ortholog pairs showed successful humanization for 47% of the genes tested. Successful humanization was found to be strongly process-dependent however, with genes from a single pathway or process often showing similar complementation [108]. This pathway dependency of complementation was also found in related studies systematically investigating the heme biosynthesis pathway, which was found to be highly replaceable, and the cytoskeleton, which proved challenging to humanize [109, 110]. Complete humanization of a metabolic pathway was recently shown for the adenine biosynthesis pathway [111]. Such modifications can be challenging however, and this pathway required extensive ‘debugging’ and analysis of suppressor mutants to identify the problems with functional expression. A major challenge of humanization lies in overcoming the differences between the yeast cellular environment from human cells to allow functional gene expression. An example of how such challenges can be overcome was shown recently, when the sterol synthesis of yeast cells was humanized by engineering of cholesterol synthesis to functionally express G-protein coupled receptors (GPCRs). Subsequently, this allowed the study of the dependency of GPCR activity on the presence of specific sterols and the generation of novel opioid biosensors [112].

Challenges in engineering the glycolytic pathway in yeast

For both metabolic engineering as well as fundamental studies on the operation of the glycolytic pathway and its enzymes, the yeast *S. cerevisiae* has played a key

role. Even in this model microorganism however, modifying complete metabolic pathways such as glycolysis is challenging. Engineering of the glycolytic reactions in *S. cerevisiae* is complicated due to its complex genetic organization. For most of the glycolytic steps, *S. cerevisiae* contains multiple isoenzymes (also see Fig. 3), encompassing 19 genes encoding the 10 steps from glucose to pyruvate spread across the genome. This redundancy is thought to result from an ancestral whole-genome duplication which led to the emergence of *Saccharomyces* yeasts [113]. For most pathways and processes, the duplicate genes resulting from this event have subsequently been lost in evolution, and the presence of multiple paralogs for glycolysis and other pathways of central-carbon metabolism is thought to be related to *S. cerevisiae's* fermentative lifestyle [114]. The multiple isoenzymes in several cases have different kinetic properties and metabolic regulations, probably allowing flexibility under different external conditions. Expression levels differ per isoenzyme, with a so-called 'major' isoenzyme constitutively expressed at high level, and condition-dependent 'minor' isoenzymes. Only the two phosphofructokinase isoenzymes are equally expressed and required for fast growth. Deletion of the minor isoenzymes of glycolysis and alcoholic fermentation surprisingly does not affect the physiology of *S. cerevisiae* under laboratory conditions [115]. Subsequently, to enable more efficient engineering of the complete pathway in this minimized glycolysis strain, the genes encoding the major isoenzymes of glycolysis and alcoholic fermentation were co-localized to a single locus and deleted from their native locus [116]. This simplification of the genetics of yeast glycolysis opens the way to modular pathway design which for metabolic rewiring and fundamental studies into the operation and conservation of the pathway.

Scope of this thesis

This PhD project was funded by an ERC consolidator grant with as overarching goal to facilitate engineering of yeast central carbon metabolism. Large scale rewiring of yeast central carbon metabolism could enable higher product yields, rates and titres in the many yeast-catalysed processes in industrial biotechnology as well as improve yeast as a model organism for fundamental studies. Such engineering is challenging however, limited by the available molecular tools, the genetic redundancy for many of the central-carbon metabolism reactions and the lack of fundamental knowledge on its regulation. Therefore, these limitations were addressed by the development of new molecular tools for *S. cerevisiae* engineering [117-119], minimization and relocalization of the central carbon metabolism genes [120] and the construction of synthetic chromosomes to serve as expression platforms [121]. Yeast glycolysis and its engineering formed a central topic in this project and was the subject of the three experimental chapters in this thesis.

The Minimal Glycolysis (MG) and Switchable Yeast Glycolysis (SwYG) yeast strains greatly facilitate engineering of glycolysis [115, 116]. These strains, intensively used in **Chapters 3 and 4**, enable easy and fast single complementation by heterologous glycolytic genes as well as swapping of the entire pathway by remodelled (heterologous) glycolytic configurations. However, while the phenotype of the MG strain was identical to the phenotype of the parental yeast strain with a full set of glycolytic isoenzymes, relocalization of the glycolytic genes to a single locus mildly but reproducibly affected the phenotype of the SwYG strain with a ca. 5-10% reduction in specific growth rate [116]. Since the effect of pathway swapping with heterologous variants was expected to be much more severe, this mild effect does not present a hurdle for the implementation of the SwYG strain to study the remodelling of yeast glycolysis. However, it might hold relevant information for the design of large synthetic gene constructs. The pathway swapping concept is based on the co-localization of genes on a chromosomal locus. Demonstrated in this thesis for glycolysis, this concept can be extended to other pathways or functions, and opens the way to the construction of modular genomes to facilitate genome engineering. Also, integration of co-localized (heterologous) genes in a host genome is a standard approach for the (over)expression of pathways in metabolic engineering. The occurrence of a negative effect of gene co-localization on yeast physiology might therefore bring important insight for the construction of synthetic pathways and chromosomes and thereby help define design criteria for synthetic constructs. The mechanisms underlying the slower growth phenotype of the SwYG strain carrying a single locus glycolysis were therefore studied in **Chapter 2**. This study explores the impact of intensive genetic engineering on the yeast genome and physiology, and illustrates the technical challenges encountered in the identification the causal relationship between genotype and phenotype.

In **Chapter 3** the MG and SwYG strains were used to their full extent to tackle the humanization of the glycolytic pathway in yeast. Humanization of single genes in yeast has been applied for decades to generate improved model organisms. Engineering of complete pathways, especially those that are central to metabolism remains challenging however, due to their high genetic complexity and essentiality. Despite extensive use of yeast as a model organism, and the importance of glycolysis for human health and disease, complementation in yeast has only been described for some of the human glycolytic genes and no systematic study on its complementation has been undertaken. The analysis of both single complementation strains as well as completely transplanted human glycolysis strains allows the study of the conservation of this essential pathway, including its regulation and moonlighting functions. Combining quantitative physiology, proteomics and laboratory evolution a detailed picture of the function of the human muscle glycolysis

in yeast was generated and the potential of humanized glycolysis strains as new model organisms was evaluated.

The physiological role of allosteric regulations in yeast glycolysis remains poorly defined. So far, the impact of these regulations has mostly been investigated by single gene complementation studies of the regulated glycolytic reactions. However, to date the replacement of individual glycolytic genes by insensitive variants failed to identify the physiological role of allosteric regulations in *S. cerevisiae*. Harnessing full pathway swapping, **Chapter 4** explores *S. cerevisiae* robustness to drastic remodelling of the allosteric regulation of glycolysis. The full set of *S. cerevisiae* glycolytic enzymes was swapped by the EMP pathway of the dimorphic oleaginous yeast *Yarrowia lipolytica*. Contrary to *S. cerevisiae*, hexoses are not a preferred carbon source for *Y. lipolytica*, and its glycolytic enzymes are devoid of the allosteric regulations found in *S. cerevisiae*. A combination of strains with full pathway swapping and smaller scale gene complementation (one to three genes), was investigated for balanced growth on glucose as well as transition between different sugar substrates. In-depth physiological and genetic characterization of these strains revealed the remarkable synergetic role played by the allosteric regulations of the glycolytic enzymes in the stability of the glycolytic flux.

The yeast genome is composed of a mosaic of genes that are not physically clustered according to their cellular function. With a focus on glycolysis, Chapters 3 and 4 demonstrate the power of a modular genetic organisation, in which relocalizing genes involved in the same cellular process enables its fast and easy remodelling. Glycolysis remodelling is but the first step in a larger effort to modularize the yeast genome. The developing field of Synthetic Genomics offers great possibilities for genome engineering using designer synthetic chromosomes. In the ERC project of which this thesis is a part, genetic modularization was expanded to central carbon metabolism, and the yeast genome was remodelled using specialized, modular synthetic chromosomes. With a focus on *S. cerevisiae*, **Chapter 5** reviews the recent advances in the field of Synthetic Genomics and evaluates the technical challenges to overcome for the construction of synthetic chromosomes to become a valuable approach for modular genome remodelling.

References

1. Barnett, J.A., A history of research on yeasts 5: the fermentation pathway. Yeast, 2003. **20**(6): p. 509-43.DOI: 10.1002/yea.986.
2. Folch, P.L., M.M.M. Bisschops, and R.A. Weusthuis, Metabolic energy conservation for fermentative product formation. Microb Biotechnol, 2021. **14**(3): p. 829-858.DOI: 10.1111/1751-7915.13746.
3. Bar-Even, A., et al., Rethinking glycolysis: on the biochemical logic of metabolic pathways. Nature Chemical Biology, 2012. **8**(6): p. 509-517.DOI: 10.1038/Nchembio.971.
4. Meléndez-Hevia, E., et al., Theoretical approaches to the evolutionary optimization of glycolysis: chemical analysis. European journal of biochemistry, 1997. **244**(2): p. 527-543.
5. Court, S.J., B. Waclaw, and R.J. Allen, Lower glycolysis carries a higher flux than any biochemically possible alternative. Nat Commun, 2015. **6**(1): p. 8427.DOI: 10.1038/ncomms9427.
6. Ronimus, R.S. and H.W. Morgan, Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. Archaea, 2003. **1**(3): p. 199-221.
7. Fothergill-Gilmore, L.A. and P.A. Michels, Evolution of glycolysis. Prog Biophys Mol Biol, 1993. **59**(2): p. 105-235.DOI: 10.1016/0079-6107(93)90001-z.
8. Dolatshahi, S., L.L. Fonseca, and E.O. Voit, New insights into the complex regulation of the glycolytic pathway in *Lactococcus lactis*. II. Inference of the precisely timed control system regulating glycolysis. Molecular Biosystems, 2016. **12**(1): p. 37-47.DOI: 10.1039/c5mb00726g.
9. Romeo, T. and J.L. Snoep, Glycolysis and flux control. EcoSal Plus, 2005. **1**(2).
10. Bräsen, C., et al., Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol Mol Biol Rev, 2014. **78**(1): p. 89-175.DOI: 10.1128/MMBR.00041-13.
11. Ronimus, R.S. and H.W. Morgan, The biochemical properties and phylogenies of phosphofructokinases from extremophiles. Extremophiles, 2001. **5**(6): p. 357-373.DOI: 10.1007/s007920100215.
12. Plaxton, W.C., The organization and regulation of plant glycolysis. Annual Review of Plant Physiology and Plant Molecular Biology, 1996. **47**(1): p. 185-214.DOI: 10.1146/annurev.arplant.47.1.185.
13. Taillefer, M. and R. Sparling, Glycolysis as the Central Core of Fermentation. Anaerobes in Biotechnology, 2016. **156**: p. 55-77.DOI: 10.1007/10_2015_5003.
14. Conway, T., The Entner-Doudoroff pathway: history, physiology and molecular biology. FEMS Microbiol Rev, 1992. **9**(1): p. 1-27.DOI: 10.1111/j.1574-6968.1992.tb05822.x.
15. Flamholz, A., et al., Glycolytic strategy as a tradeoff between energy yield and protein cost. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(24): p. 10039-10044.DOI: 10.1073/pnas.1215283110.
16. Verhees, C.H., et al., The unique features of glycolytic pathways in Archaea. Biochem J, 2003. **375**(Pt 2): p. 231-46.DOI: 10.1042/BJ20021472.
17. Nikel, P.I., et al., *Pseudomonas putida* KT2440 Strain Metabolizes Glucose through a Cycle Formed by Enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and Pentose Phosphate Pathways. Journal of Biological Chemistry, 2015. **290**(43): p. 25920-25932.DOI: 10.1074/jbc.M115.687749.
18. Chavarría, M., et al., The Entner–Doudoroff pathway empowers Pseudomonas putida KT 2440 with a high tolerance to oxidative stress. Environmental microbiology, 2013. **15**(6): p. 1772-1785.

19. Jacoby, J., C.P. Hollenberg, and J.J. Heinisch, Transaldolase Mutants in the Yeast Kluyveromyces-Lactis Provide Evidence That Glucose Can Be Metabolized through the Pentose-Phosphate Pathway. Molecular Microbiology, 1993. **10**(4): p. 867-876.DOI: 10.1111/j.1365-2958.1993.tb00957.x.
20. de Vries, W., S.J. Gerbrandy, and A.H. Stouthamer, Carbohydrate metabolism in Bifidobacterium bifidum. Biochim Biophys Acta, 1967. **136**(3): p. 415-25.DOI: 10.1016/0304-4165(67)90001-3.
21. van Heerden, J.H., F.J. Bruggeman, and B. Teusink, Multi-tasking of biosynthetic and energetic functions of glycolysis explained by supply and demand logic. BioEssays, 2015. **37**(1): p. 34-45.
22. Morales-Polanco, F., et al., Core Fermentation (CoFe) granules focus coordinated glycolytic mRNA localization and translation to fuel glucose fermentation. iScience, 2021. **24**(2): p. 102069.DOI: 10.1016/j.isci.2021.102069.
23. Daran-Lapujade, P., et al., The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. Proceedings of the National Academy of Sciences, 2007. **104**(40): p. 15753-15758.
24. Sola-Penna, M., et al., Regulation of mammalian muscle type 6-phosphofructo-1-kinase and its implication for the control of the metabolism. IUBMB life, 2010. **62**(11): p. 791-796.DOI: 10.1002/iub.393.
25. Oliveira, A.P. and U. Sauer, The importance of post-translational modifications in regulating *Saccharomyces cerevisiae* metabolism. FEMS yeast research, 2012. **12**(2): p. 104-117.
26. Chen, Y. and J. Nielsen, Flux control through protein phosphorylation in yeast. Fems Yeast Research, 2016. **16**(8).DOI: 10.1093/femsyr/fow096.
27. Tanner, L.B., et al., Four Key Steps Control Glycolytic Flux in Mammalian Cells. Cell Syst, 2018. **7**(1): p. 49-62 e8.DOI: 10.1016/j.cels.2018.06.003.
28. Boiteux, A. and B. Hess, Design of Glycolysis. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 1981. **293**(1063): p. 5-22.DOI: 10.1098/rstb.1981.0056.
29. Teusink, B., et al., The danger of metabolic pathways with turbo design. Trends in Biochemical Sciences, 1998. **23**(5): p. 162-169.
30. Koebmann, B.J., et al., The glycolytic flux in *Escherichia coli* is controlled by the demand for ATP. Journal of Bacteriology, 2002. **184**(14): p. 3909-3916.DOI: 10.1128/Jb.184.14.3909-3916.2002.
31. Koebmann, B.J., et al., The extent to which ATP demand controls the glycolytic flux depends strongly on the organism and conditions for growth. Molecular Biology Reports, 2002. **29**(1-2): p. 41-45.DOI: 10.1023/A:1020398117281.
32. Neves, A.R., et al., Is the Glycolytic Flux in Lactococcus lactis Primarily Controlled by the Redox Charge?: KINETICS OF NAD+ AND NADH POOLS DETERMINED *IN VIVO* BY 13C NMR. Journal of Biological Chemistry, 2002. **277**(31): p. 28088-28098.
33. Voit, E.O., et al., Regulation of glycolysis in *Lactococcus lactis*: an unfinished systems biological case study. Syst Biol (Stevenage), 2006. **153**(4): p. 286-98.DOI: 10.1049/ip-syb:20050087.
34. Van den Brink, J., et al., Dynamics of glycolytic regulation during adaptation of *Saccharomyces cerevisiae* to fermentative metabolism. Applied and environmental microbiology, 2008. **74**(18): p. 5710-5723.
35. Teusink, B., et al., Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem, 2000. **267**(17): p. 5313-29.DOI: 10.1046/j.1432-1327.2000.01527.x.
36. Gruning, N.M., et al., Inhibition of triosephosphate isomerase by phosphoenolpyruvate in the feedback-regulation of glycolysis. Open Biology, 2014. **4**(3): p. 130232.DOI: 10.1098/rsob.130232.

37. Stoddard, P.R., et al., Polymerization in the actin ATPase clan regulates hexokinase activity in yeast. Science, 2020. **367**(6481): p. 1039-1042.DOI: 10.1126/science.aay5359.
38. Araiza-Olivera, D., et al., A glycolytic metabolon in *Saccharomyces cerevisiae* is stabilized by F-actin. FEBS J, 2013. **280**(16): p. 3887-905.DOI: 10.1111/febs.12387.
39. Pfeiffer, T. and A. Morley, An evolutionary perspective on the Crabtree effect. Frontiers in Molecular Biosciences, 2014. **1**: p. 17.DOI: 10.3389/fmolb.2014.00017.
40. Dashko, S., et al., Why, when, and how did yeast evolve alcoholic fermentation? FEMS yeast research, 2014. **14**(6): p. 826-832.
41. Elsemman, I.E., et al., Whole-cell modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies. bioRxiv, 2021.DOI: 10.1101/2021.06.11.448029.
42. Nilsson, A. and J. Nielsen, Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. Sci Rep, 2016. **6**(1): p. 22264.DOI: 10.1038/srep22264.
43. Chambers, A., E.A. Packham, and I.R. Graham, Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*). Current Genetics, 1995. **29**(1): p. 1-9.DOI: 10.1007/Bf00313187.
44. Schaaff, I., J. Heinisch, and F.K. Zimmermann, Overproduction of Glycolytic-Enzymes in Yeast. Yeast, 1989. **5**(4): p. 285-290.DOI: 10.1002/yea.320050408.
45. Hauf, J., F.K. Zimmermann, and S. Muller, Simultaneous genomic overexpression of seven glycolytic enzymes in the yeast *Saccharomyces cerevisiae.* Enzyme Microb Technol, 2000. **26**(9-10): p. 688-698.DOI: 10.1016/s0141-0229(00)00160-5.
46. Blazquez, M.A., et al., Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. FEBS Letters, 1993. **329**(1-2): p. 51-54.
47. Bartrons, R., et al., The stimulation of yeast phosphofructokinase by fructose 2,6-bisphosphate. FEBS Lett, 1982. **143**(1): p. 137-40.DOI: 10.1016/0014-5793(82)80290-1.
48. Otto, A., et al., Kinetic effects of fructose-1,6-bisphosphate on yeast phosphofructokinase. Biomed Biochim Acta, 1986. **45**(7): p. 865-75.
49. Müller, S., F.K. Zimmermann, and E. Boles, Mutant studies of phosphofructo-2-kinases do not reveal an essential role of fructose-2, 6-bisphosphate in the regulation of carbon fluxes in yeast cells. Microbiology, 1997. **143**(9): p. 3055-3061.
50. François, J., E. Van Schaftingen, and H.G. HERS, Characterization of phosphofructokinase 2 and of enzymes involved in the degradation of fructose 2, 6-bisphosphate in yeast. European journal of Biochemistry, 1988. **171**(3): p. 599-608.
51. Navas, M.A. and J.M. Gancedo, The regulatory characteristics of yeast fructose-1,6-bisphosphatase confer only a small selective advantage. J Bacteriol, 1996. **178**(7): p. 1809-12.DOI: 10.1128/jb.178.7.1809-1812.1996.
52. Hess, B., R. Haeckel, and K. Brand, FDP-activation of yeast pyruvate kinase. Biochemical and biophysical research communications, 1966. **24**(6): p. 824-831.
53. Mayordomo, I. and P. Sanz, Human pancreatic glucokinase (GlkB) complements the glucose signalling defect of *Saccharomyces cerevisiae hxk2* mutants. Yeast, 2001. **18**(14): p. 1309-16.DOI: 10.1002/yea.780.
54. Bonini, B.M., P. Van Dijck, and J.M. Thevelein, Uncoupling of the glucose growth defect and the deregulation of glycolysis in *Saccharomyces cerevisiae* Tps1 mutants expressing trehalose-6-phosphate-insensitive hexokinase from *Schizosaccharomyces pombe*. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2003. **1606**(1-3): p. 83-93.
55. Boles, E., et al., Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in *Saccharomyces cerevisiae* that is catalytically insensitive to fructose-1,6-bisphosphate. J. Bacteriol, 1997. **179**(9): p. 2987-2993.

56. Estévez, A.M., J.J. Heinisch, and J.J. Aragón, Functional complementation of yeast phosphofructokinase mutants by the non-allosteric enzyme from *Dictyostelium discoideum.* FEBS letters, 1995. **374**(1): p. 100-104.
57. Thevelein, J.M. and S. Hohmann, Trehalose synthase: guard to the gate of glycolysis in yeast? Trends Biochem Sci, 1995. **20**(1): p. 3-10.DOI: 10.1016/s0968-0004(00)88938-0.
58. van Heerden, J.H., et al., Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. Science, 2014. **343**(6174): p. 1245114.DOI: 10.1126/science.1245114.
59. Deroover, S., et al., Trehalose-6-phosphate synthesis controls yeast gluconeogenesis downstream and independent of SNF1. FEMS Yeast Res, 2016. **16**(4): p. fow036.DOI: 10.1093/femsyr/fow036.
60. Van Leemputte, F., et al., Aberrant Intracellular pH Regulation Limiting Glyceraldehyde-3-Phosphate Dehydrogenase Activity in the Glucose-Sensitive Yeast *tps1Δ* Mutant. Mbio, 2020. **11**(5): p. e02199-20.
61. Peeters, K., et al., Fructose-1,6-bisphosphate couples glycolytic flux to activation of Ras. Nature Communications, 2017. **8**(1): p. 1-15.DOI: 10.1038/s41467-017-01019-z.
62. van Eunen, K., et al., Testing biochemistry revisited: how *in vivo* metabolism can be understood from *in vitro* enzyme kinetics. PLoS Comput Biol, 2012. **8**(4): p. e1002483.
63. Reuter, R., et al., Purification, molecular and kinetic characterization of phosphofructokinase-1 from the yeast *Schizosaccharomyces pombe*: evidence for an unusual subunit composition. Yeast, 2000. **16**(14): p. 1273-85.DOI: 10.1002/1097-0061(200010)16:14<1273::AID-YEA621>3.0.CO;2-#.
64. Blazquez, M.A., et al., Trehalose-6-P synthase is dispensable for growth on glucose but not for spore germination in *Schizosaccharomyces pombe.* Journal of Bacteriology, 1994. **176**(13): p. 3895-3902.
65. Flores, C.-L., et al., The dimorphic yeast *Yarrowia lipolytica* possesses an atypical phosphofructokinase: characterization of the enzyme and its encoding gene. Microbiology, 2005. **151**(5): p. 1465-1474.
66. Flores, C.-L., C. Gancedo, and T. Petit, Disruption of *Yarrowia lipolytica TPS1* gene encoding trehalose-6-P synthase does not affect growth in glucose but impairs growth at high temperature. PloS one, 2011. **6**(9): p. e23695.
67. Hirai, M., A. Tanaka, and S. Fukui, Difference in pyruvate kinase regulation among three groups of yeasts. Biochimica et Biophysica Acta (BBA)-Enzymology, 1975. **391**(2): p. 282-291.
68. Crowther, G.J., et al., Control of glycolysis in contracting skeletal muscle. I. Turning it on. American Journal of Physiology-Endocrinology And Metabolism, 2002. **282**(1): p. E67-E73.
69. Stanley, W.C. and R.J. Connett, Regulation of muscle carbohydrate metabolism during exercise 1. The FASEB journal, 1991. **5**(8): p. 2155-2159.
70. Wilson, J.E., Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. J Exp Biol, 2003. **206**(Pt 12): p. 2049-57.DOI: 10.1242/jeb.00241.
71. Lambeth, M.J. and M.J. Kushmerick, A computational model for glycogenolysis in skeletal muscle. Ann Biomed Eng, 2002. **30**(6): p. 808-27.DOI: 10.1114/1.1492813.
72. Bakker, B.M., et al., Glycolysis in bloodstream form *Trypanosoma brucei* can be understood in terms of the kinetics of the glycolytic enzymes. Journal of Biological Chemistry, 1997. **272**(6): p. 3207-3215.DOI: 10.1074/jbc.272.6.3207.
73. Opperdoes, F.R. and P. Borst, Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome. FEBS Lett, 1977. **80**(2): p. 360-4.DOI: 10.1016/0014-5793(77)80476-6.

74. Bakker, B.M., et al., Compartmentation protects trypanosomes from the dangerous design of glycolysis. Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2087-92.DOI: 10.1073/pnas.030539197.
75. Copley, S.D., Moonlighting is mainstream: Paradigm adjustment required. Bioessays, 2012. **34**(7): p. 578-588.DOI: 10.1002/bies.201100191.
76. Gomez-Arreaza, A., et al., Extracellular functions of glycolytic enzymes of parasites: Unpredicted use of ancient proteins. Molecular and Biochemical Parasitology, 2014. **193**(2): p. 75-81.DOI: 10.1016/j.molbiopara.2014.02.005.
77. Entelis, N., et al., A glycolytic enzyme, enolase, is recruited as a cofactor of tRNA targeting toward mitochondria in *Saccharomyces cerevisiae.* Genes & development, 2006. **20**(12): p. 1609-1620.
78. Chuang, D.-M., C. Hough, and V.V. Senatorov, Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. Annu. Rev. Pharmacol. Toxicol., 2005. **45**: p. 269-290.
79. Wistow, G.J., et al., Tau-Crystallin Alpha-Enolase - One Gene Encodes Both an Enzyme and a Lens Structural Protein. Journal of Cell Biology, 1988. **107**(6): p. 2729-2736.DOI: 10.1083/jcb.107.6.2729.
80. Jimenezasensio, J., et al., Glyceraldehyde 3-phosphate dehydrogenase is an enzyme-crystallin in diurnal geckos of the genus Phelsuma. Biochemical and biophysical research communications, 1995. **209**(3): p. 796-802.
81. Didiasova, M., L. Schaefer, and M. Wygrecka, When Place Matters: Shuttling of Enolase-1 Across Cellular Compartments. Front Cell Dev Biol, 2019. **7**: p. 61.DOI: 10.3389/fcell.2019.00061.
82. Herrero, P., C. Martínez-Campa, and F. Moreno, The hexokinase 2 protein participates in regulatory DNA-protein complexes necessary for glucose repression of the *SUC2* gene in *Saccharomyces cerevisiae.* FEBS letters, 1998. **434**(1-2): p. 71-76.
83. Ahuatzi, D., et al., The glucose-regulated nuclear localization of hexokinase 2 in *Saccharomyces cerevisiae* is Mig1-dependent. J Biol Chem, 2004. **279**(14): p. 14440-6.DOI: 10.1074/jbc.M313431200.
84. Lu, M., et al., Interaction between aldolase and vacuolar H+-ATPase: evidence for direct coupling of glycolysis to the ATP-hydrolyzing proton pump. Journal of Biological Chemistry, 2001. **276**(32): p. 30407-30413.
85. Lu, M., et al., The glycolytic enzyme aldolase mediates assembly, expression, and activity of vacuolar H+-ATPase. J Biol Chem, 2004. **279**(10): p. 8732-9.DOI: 10.1074/jbc.M303871200.
86. Lu, M., et al., Physical interaction between aldolase and vacuolar H+-ATPase is essential for the assembly and activity of the proton pump. J Biol Chem, 2007. **282**(34): p. 24495-503.DOI: 10.1074/jbc.M702598200.
87. Nielsen, J. and J.D. Keasling, Engineering Cellular Metabolism. Cell, 2016. **164**(6): p. 1185-1197.DOI: 10.1016/j.cell.2016.02.004.
88. Jojima, T. and M. Inui, Engineering the glycolytic pathway: a potential approach for improvement of biocatalyst performance. Bioengineered, 2015. **6**(6): p. 328-334.
89. Bro, C., et al., In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. Metabolic engineering, 2006. **8**(2): p. 102-111.
90. Guadalupe-Medina, V., et al., Carbon dioxide fixation by Calvin-Cycle enzymes improves ethanol yield in yeast. Biotechnol Biofuels, 2013. **6**(1): p. 125.DOI: 10.1186/1754-6834-6-125.
91. Benisch, F. and E. Boles, The bacterial Entner–Doudoroff pathway does not replace glycolysis in *Saccharomyces cerevisiae* due to the lack of activity of iron–sulfur cluster enzyme 6-phosphogluconate dehydratase. Journal of biotechnology, 2014. **171**: p. 45-55.

92. Morita, K., et al., Heterologous expression of bacterial phosphoenol pyruvate carboxylase and Entner–Doudoroff pathway in *Saccharomyces cerevisiae* for improvement of isobutanol production. Journal of bioscience and bioengineering, 2017. **124**(3): p. 263-270.
93. Bogorad, I.W., T.-S. Lin, and J.C. Liao, Synthetic non-oxidative glycolysis enables complete carbon conservation. Nature, 2013. **502**(7473): p. 693-697.
94. Lin, P.P., et al., Construction and evolution of an *Escherichia coli* strain relying on nonoxidative glycolysis for sugar catabolism. Proceedings of the National Academy of Sciences, 2018. **115**(14): p. 3538-3546.
95. Hellgren, J., et al., Promiscuous phosphoketolase and metabolic rewiring enables novel non-oxidative glycolysis in yeast for high-yield production of acetyl-CoA derived products. Metabolic Engineering, 2020. **62**: p. 150-160.
96. Kamineni, A., et al., Increasing lipid yield in *Yarrowia lipolytica* through phosphoketolase and phosphotransacetylase expression in a phosphofructokinase deletion strain. Biotechnology for Biofuels, 2021. **14**(1): p. 1-14.DOI: 10.1186/s13068-021-01962-6.
97. Iacometti, C., et al., Activating silent glycolysis bypasses in *Escherichia coli.* BioRxiv, 2021.DOI: 10.1101/2021.11.18.468982.
98. Sanchez-Pascuala, A., et al., Functional implementation of a linear glycolysis for sugar catabolism in *Pseudomonas putida.* Metabolic Engineering, 2019. **54**: p. 200-211.DOI: 10.1016/j.ymben.2019.04.005.
99. Goffeau, A., et al., Life with 6000 genes. Science, 1996. **274**(5287): p. 546-567.
100. Giaever, G. and C. Nislow, The yeast deletion collection: a decade of functional genomics. Genetics, 2014. **197**(2): p. 451-465.
101. Giaever, G., et al., Functional profiling of the *Saccharomyces cerevisiae* genome. nature, 2002. **418**(6896): p. 387-391.
102. DiCarlo, J.E., et al., Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res, 2013. **41**(7): p. 4336-43.DOI: 10.1093/nar/gkt135.
103. Mager, W.H. and J. Winderickx, Yeast as a model for medical and medicinal research. Trends in pharmacological sciences, 2005. **26**(5): p. 265-273.
104. Laurent, J.M., et al., Humanization of yeast genes with multiple human orthologs reveals functional divergence between paralogs. PLoS Biology, 2020. **18**(5): p. e3000627.
105. Blachly-Dyson, E., et al., Cloning and functional expression in yeast of two human isoforms of the outer mitochondrial membrane channel, the voltage-dependent anion channel. Journal of Biological Chemistry, 1993. **268**(3): p. 1835-1841.
106. Hamza, A., et al., Complementation of yeast genes with human genes as an experimental platform for functional testing of human genetic variants. Genetics, 2015. **201**(3): p. 1263-74.DOI: 10.1534/genetics.115.181099.
107. Hamza, A., et al., Cross-species complementation of nonessential yeast genes establishes platforms for testing inhibitors of human proteins. Genetics, 2020.DOI: 10.1534/genetics.119.302971.
108. Kachroo, A.H., et al., Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science, 2015. **348**(6237): p. 921-925.DOI: 10.1126/science.aaa0769.
109. Kachroo, A.H., et al., Systematic bacterialization of yeast genes identifies a near-universally swappable pathway. Elife, 2017. **6**.DOI: 10.7554/eLife.25093.
110. Garge, R.K., et al., Systematic humanization of the yeast cytoskeleton discerns functionally replaceable from divergent human genes. Genetics, 2020.
111. Agmon, N., et al., Phylogenetic debugging of a complete human biosynthetic pathway transplanted into yeast. Nucleic Acids Res, 2020. **48**(1): p. 486-499.DOI: 10.1093/nar/gkz1098.

112. Bean, B.D., et al., Functional expression of opioid receptors and other human GPCRs in yeast engineered to produce human sterols. bioRxiv, 2021.DOI: 10.1101/2021.05.12.443385.
113. Wolfe, K.H. and D.C. Shields, Molecular evidence for an ancient duplication of the entire yeast genome. Nature, 1997. **387**(6634): p. 708-713.
114. Escalera-Fanjul, X., et al., Whole-genome duplication and yeast's fruitful way of life. Trends in Genetics, 2019. **35**(1): p. 42-54.
115. Solis-Escalante, D., et al., A minimal set of glycolytic genes reveals strong redundancies in *Saccharomyces cerevisiae* central metabolism. Eukaryotic Cell, 2015. **14**(8): p. 804-816.DOI: 10.1128/EC.00064-15.
116. Kuijpers, N.G., et al., Pathway swapping: Toward modular engineering of essential cellular processes. Proceedings of the National Academy of Sciences, USA, 2016. **113**(52): p. 15060-15065.DOI: 10.1073/pnas.1606701113.
117. Świat, M.A., et al., *Fn*Cpf1: a novel and efficient genome editing tool for *Saccharomyces cerevisiae.* Nucleic acids research, 2017. **45**(21): p. 12585-12598.
118. Wijsman, M., et al., A toolkit for rapid CRISPR-SpCas9 assisted construction of hexose-transport-deficient *Saccharomyces cerevisiae* strains. FEMS Yeast Res, 2019. **19**(1).DOI: 10.1093/femsyr/foy107.
119. Boonekamp, F.J., et al., Design and Experimental Evaluation of a Minimal, Innocuous Watermarking Strategy to Distinguish Near-Identical DNA and RNA Sequences. ACS Synth Biol, 2020. **9**(6): p. 1361-1375.DOI: 10.1021/acssynbio.0c00045.
120. Postma, E., et al., Top-down, knowledge-based genetic reduction of yeast central carbon metabolism. BioRxiv, 2021.DOI: 10.1101/2021.08.24.457526.
121. Postma, E.D., et al., Modular, synthetic chromosomes as new tools for large scale engineering of metabolism. Metabolic Engineering, 2022. **72**: p. 1-13.

Chapter 2

What's wrong with SwYG? Reaching the limits of the *Saccharomyces cerevisiae* molecular and analytical toolbox

Ewout Knibbe[#], Eline D. Postma[#], Francine J. Boonekamp, Sofiia Dashko, Jordi Geelhoed, Anne-Marijn Maat, Marcel van den Broek, Marijke A. H. Luttik, Pascale Daran-Lapujade

[#] These authors contributed equally to this work and should be considered co-first authors.

Manuscript in preparation for submission

Abstract

The construction of powerful cell factories requires extensive remodelling of microbial genomes, entailing many rounds of transformations to perform the large number of desired gene modifications. However, increasing the number of genetic interventions inevitably increases the occurrence of unwanted mutations and effects. Using glycolysis as paradigm, a previous study developed a *Saccharomyces cerevisiae* strain in which the glycolytic genes, relocated to a single locus, can be easily swapped by any new design, thereby enabling fast and easy remodelling of the entire pathway. After 27 genetic modifications performed in 43 transformation rounds, the Switchable Yeast Glycolysis (SwYG) strain grew ca. 20% slower than its ancestor with the same glycolytic genes with native glycolysis design. Exploring the cause of this slower growth rate, the present study reflects on the genetic and analytical challenges encountered by extensive strain construction programs and provides design guidelines for integration of large constructs in the yeast genome. This study also suggests a potential involvement of the yeast glycolytic enzyme phosphoglycerate kinase (Pgk1) in PI(3)P synthesis and autophagy, as found in mammalian cells.

Introduction

Converting microbial cells, optimized to operate in defined natural environments, into powerful factories with high product yields and productivity in harsh industrial environments requires extensive remodelling of their genome. Due to the limited predictability of biological systems, the construction of such cell factories requires the iteration of design-build-test-learn cycles [1]. In the build phase of the cycle, microbial cells undergo many rounds of genetic manipulations meant to add new metabolic routes, but also to rewire native pathways and cellular machineries, indispensable to optimally supply energy-rich moieties, redox equivalents, precursors and co-factors to the newly added routes [2-4]. Despite the recent developments in DNA editing with CRISPR-Cas technologies, extensive rewiring of native pathways and processes still presents a technical challenge. Taking as paradigm the popular model organism and synthetic biology platform *Saccharomyces cerevisiae*, rewiring entire native pathways is complicated by a high degree of genetic redundancy and the scattering of genes over the 12 Mb and 16 chromosome-containing genome. For instance, glycolysis and alcoholic fermentation, central pathways for carbon conversion, cover a set of twelve biochemical reactions important for industrial applications of this yeast. Most of these twelve reactions are catalysed by multiple iso-enzymes, resulting in a set of 26 genes scattered over the yeast chromosomes. To facilitate remodelling of glycolysis and alcoholic fermentation in yeast, previous studies reported the construction of a strain in which genes encoding the 13 minor paralogs of glycolysis and fermentation were deleted and the remaining 13 major paralogs were relocalized to a single locus [5, 6]. In a process named pathway swapping, the resulting Switchable Yeast Glycolysis strain (SwYG) enables the two-step replacement of the entire glycolytic and fermentation pathways by any new design, and their translocation to any locus on native or synthetic chromosomes [6, 7]. The SwYG strain, harbouring a minimized set of glycolytic and fermentation genes clustered on a single chromosomal locus consistently displayed a 8-30% reduction in growth rate in synthetic chemically defined medium with glucose as carbon source, as compared to a strain with a complete set of glycolytic enzymes scattered over the yeast chromosomes [6]. This decrease in growth rate could not be attributed to the removal of the minor glycolytic paralogs, as a similar strain in which the same minimized set of genes was still present in their native loci grew just as fast as the non-minimized control strain, and therefore faster than SwYG [5]. While the SwYG strain is a valuable addition to *S. cerevisiae* molecular toolbox, this decrease in growth rate was puzzling and could hold key information regarding the genetic design of large DNA constructs. If the co-localization of thirteen highly expressed genes had deleterious effects on the host, it

should be taken into consideration for the design and engineering of cells factories, as well as the construction of synthetic chromosomes and genomes.

Many mechanisms could potentially cause the growth defect of the SwYG strain, and this study attempts to identify which ones are at play (Fig. 1). The growth defect could be directly caused by the integration of the set of 13 clustered genes (called SinLoG for Single Locus Glycolysis) in the yeast chromosomes, by altering the expression either of the genes in the SinLoG itself or of genes neighbouring the SinLoG integration site. Alternatively, the SinLoG might affect DNA replication either by causing collision between the replication machinery and the polymerase processing the highly transcribed colocalized glycolytic genes, or by excessively increasing the distance between adjacent autonomously replicating sequences (ARS) [8, 9]. The growth defect of the SwYG strain could also be indirectly caused by the mutations voluntarily (targeted deletion of genes, presence of selection markers and plasmids) or involuntarily (random mutations) induced during strain construction. Using a combination of targeted genetic engineering, and explorative evolutionary and reverse engineering, this study makes full use of the rich *S. cerevisiae* molecular and analytical toolbox to directly test these hypotheses and to discover yet unpredicted mechanisms.

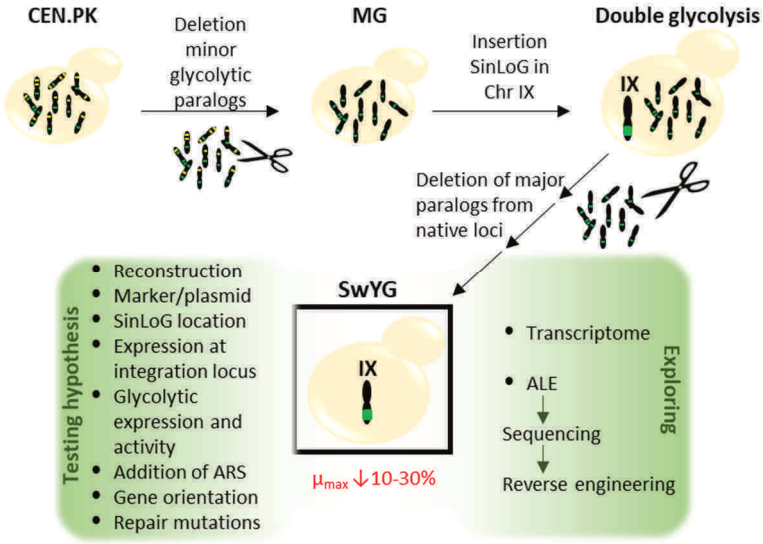


Figure 1 - Schematic overview of the experimental approach.
Strain construction strategy leading to the Switchable Yeast Glycolysis (SwYG) strain with a Single Locus Glycolysis (SinLoG). First the Minimal Glycolysis (MG) strain was constructed by deletion of the minor glycolytic paralogs without measurable phenotypic effects [5]. Subsequently the major glycolytic paralogs were integrated as SinLoG in chromosome IX and deleted from their native loci, resulting in the SwYG strain [6]. The SwYG strain displayed a decreased growth rate as compared to the MG strain. Several hypotheses for the mechanism underlying this growth rate decrease were tested, as well as exploratory transcriptomics and adaptive laboratory evolution (ALE).

Results

Confirmation that SinLoG integration in the MG strain causes a decrease in yeast specific growth rate

First, the specific growth rate of the SwYG strain (IMX589) and reference strain CEN.PK113-5D was measured again, confirming that the SwYG strain grew consistently 7.6±0.3% slower in shake flask as compared to the control strain harbouring a native set of glycolytic genes scattered over the genome (Fig. 2A). The SwYG strain is auxotrophic for uracil and harbours the amdS marker in the SinLoG. Repairing the uracil auxotrophy by integration of the *URA3* gene at the *TDH1* locus and removing the amdS marker did not affect growth rate, indicating that selection markers were not accountable for the SwYG strain slower growth phenotype (strain IMX1822, Fig. 2A). Re-localization of glycolytic genes to a single locus was originally performed in 18 transformation rounds, starting by the integration of the SinLoG and followed by the sequential deletion of the 13 glycolytic genes from their native loci (Fig. 1). To test which steps were involved in the SwYG strain slower growth, the specific growth rate of all relevant intermediate strains in the construction steps between the MG and the SwYG strains were measured (Fig 2B). It revealed that,

after integration of the SinLoG in the MG strain the following deletion steps did not significantly affect growth. It was further confirmed that the growth rate decrease caused by SinLoG integration was not an artefact specific to a single transformation event, by repeating the transformation of the MG strain with the SinLoG DNA parts (strain IMX382, Supplementary Fig. S1). Altogether these results confirmed that the growth rate reduction occurred at the time of integration of the SinLoG cassette and suggested that this integration was the most probable cause for the SwYG strain's slower growth rate phenotype.

SwYG decreased growth rate is not caused by secondary effects at the integration site

While the SinLoG seemed to be responsible for the SwYG strain's reduced growth rate, this effect might not be a direct consequence of the genetic clustering of the glycolytic genes. It could result from unexpected genetic effects at the integration site. To explore this hypothesis, prototrophic strains with the SinLoG integrated in two different loci, *CAN1* (chromosome V, IMX1821) and *SGA1* (chromosome IX, IMX1822) were constructed. However, this resulted in a similar growth rate as measured before for prototrophic SwYG strains IMX605 and IMX606 and auxotrophic strain IMX589 ([6]and Fig. 2A). The SinLoG is a 35 kb long DNA stretch, considering that autonomously replicating sequences (ARS) are distributed on average every 30 kb across the yeast genome, integration of the SinLoG might hinder replication [6, 9]. However, flanking the SinLoG by additional ARS sequences (strain IMX2109 Fig. 3, strain lineage in Supplementary Fig. S2) [7], had no positive effect on the growth rate of the SwYG strain. The presence of the SinLoG might also impact the expression of genes surrounding its integration site, however their transcript levels were largely unaffected by the SinLoG presence (Supplementary Fig. S3). Yeast appeared therefore extremely robust to the disruption of its chromosomes by integration of a 35 kb highly transcribed DNA sequence.

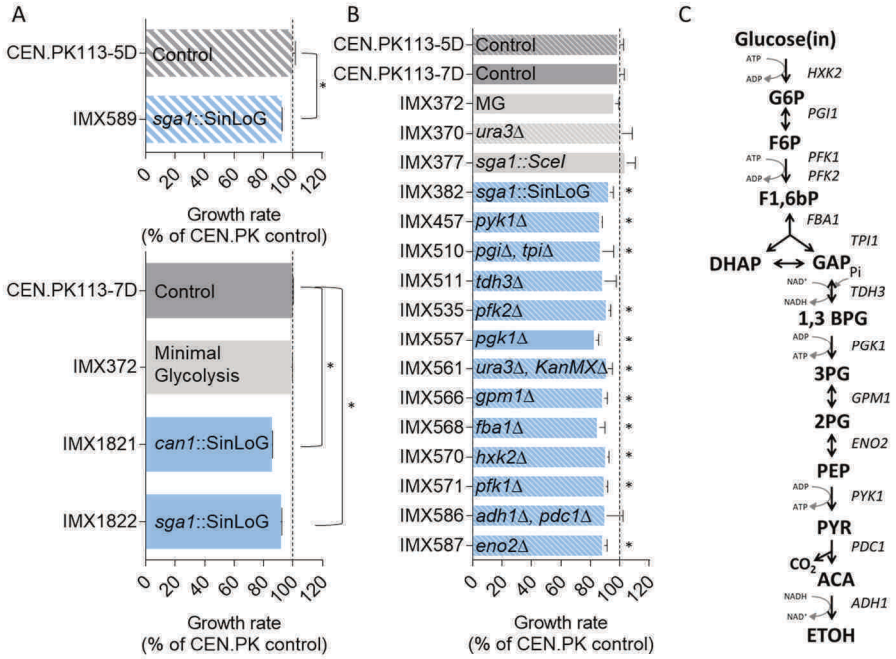


Figure 2 - Specific growth rates of single locus glycolysis strains.
The specific growth rates of various strains with a SinLoG integration are shown in blue and reference strains are in grey. Prototrophic strains are shown as filled column while uracil auxotrophic strains are depicted with dashed columns. All growth rates are normalized to the CEN.PK113-7D (prototrophic) or CEN.PK113-5D (uracil auxotrophic) control strains from the same experiment. **A)** Specific growth rates determined in duplicate shake flask cultures for several reference strains and single locus glycolysis strains. **B)** Growth rates of intermediate strains with glycolytic gene deletions during strain construction of the single locus glycolysis strains. Each modification follows below the previous one from top (MG strain) to bottom (SwYG strain). Growth rates are measured in 96-wells plates in at least culture triplicate. Statistically significant differences relative to the auxotrophic or prototrophic CEN.PK reference strain are indicated (* $P<0.05$, t-test 2-tailed, homoscedastic). **C)** Schematic overview of the glycolytic pathway of *S. cerevisiae*, reactions are indicated together with the glycolytic genes encoding the major isoenzymes which were relocated. G6P: glucose-6-phosphate; F6P: fructose-6-phosphate; F1,6bP: fructose-1,6-bisphosphate; DHAP: dihydroxy-acetone phosphate; GAP: glyceraldehyde-3-phosphate; 1,3BPG: 1,3-bisphosphoglycerate; 3PG: 3-phosphoglycerate; 2PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate; ACA: Acetaldehyde; ETOH: Ethanol.

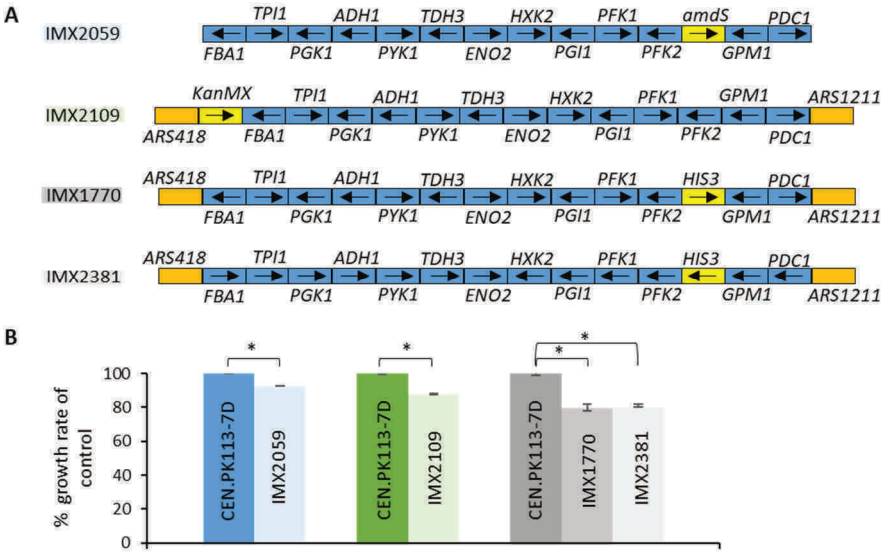


Figure 3 - Effects of ARS sequences and gene orientation.
A) Configurations of single locus glycolysis gene clusters with ARS sequences and various gene orientations. **B)** Specific growth rate of strains with SinLoG represented in panel A, as well as the prototrophic control strain with native glycolytic design CEN.PK113-7D. Bars represent the average and standard deviation of duplicate shake flask cultures expressed as a percentage of the control strain CEN.PK113-7D. Statistically significant differences are indicated (* $P<0.05$, t-test 2-tailed, homoscedastic).

SwYG decreased growth rate is not caused by altered expression of genes from the SinLoG

As the indirect effects considered so far were not responsible for the SwYG strain's slower growth rate, possible direct effects of the SinLoG itself were explored. Glycolysis is an essential pathway for growth on glucose as sole carbon source, altering its expression might therefore affect yeast growth. Transcript levels of the glycolytic genes and *in vitro* activity of the glycolytic enzymes were measured in the SwYG and control strains in batch conditions. Despite the re-localization and clustering of the glycolytic genes, their expression was very similar in the SwYG strain compared to the minimal glycolysis strain IMX372 (MG), with only *PGK1* expression changing more than 2-fold (Fig. 4). Accordingly, activity of the glycolytic enzymes was largely comparable to the MG control strain. Only Pgk1 activity was significantly decreased by 43±4% in SwYG as compared to MG (Fig. 4B). However, this reduced catalytic activity of Pgk1 is most likely not responsible for the lower growth rate of the SwYG strain as, like most glycolytic enzymes, Pgk1 is present with a large overcapacity and even with reduced activity it operates far from saturation [10]. Indeed, for Pgk1, the degree of saturation estimated from the glucose uptake rate and enzyme activity only increased from 17±3% in the MG strain to 24±3% in

the SwYG strain. Next to their key role in central carbon metabolism, several glycolytic enzymes (hexokinase, aldolase and enolase) are also involved in moonlighting functions outside glycolysis in yeast [11], however, since glycolytic transcript levels and *in vitro* activity are similar whether the glycolytic genes are scattered over the chromosomes or co-localized, there is no evidence suggesting that moonlighting functions play a role in the SwYG strain reduced growth rate.

There is yet another aspect to consider for the genetic co-localization of genes such as the glycolytic genes. Glycolytic genes are intensively and continuously transcribed in growing yeast and their transcripts are amongst the most abundant. While the strong and constitutive glycolytic promoters are an asset for yeast molecular biology toolbox, their co-localization might be deleterious for the host. During cell division, the replication machinery needs to access DNA, which might lead to head-on collisions with RNA polymerases in the SinLoG transcriptional hotspot [8]. In such a scenario, DNA replication might be stalled, and even lead to DNA damage, which could result in a slower growth rate. Since co-directional collisions are less detrimental than head-on collisions, aligning the orientation of transcription units with the nearest ARS sequences could attenuate conflicts between replication and transcription. This possibility was assessed by the design, construction and growth rate measurement of two strains with modified transcription-unit's orientation, one without (IMX1770) and one with co-directionality of transcription and replication (IMX2381). These strains showed no significant difference in growth rate (Fig. 3). The SwYG strain (auxothrophic IMX589 and prototrophic IMX1821) did not display alterations in the duration of the S phase of the cell cycle, which could be expected if DNA replication was competing with transcription (Supplementary Fig. S4). Although these results do not entirely exclude potential effects of the SinLoG on DNA replication, they strongly suggest that it is not the main cause of the SwYG strain's growth defect.

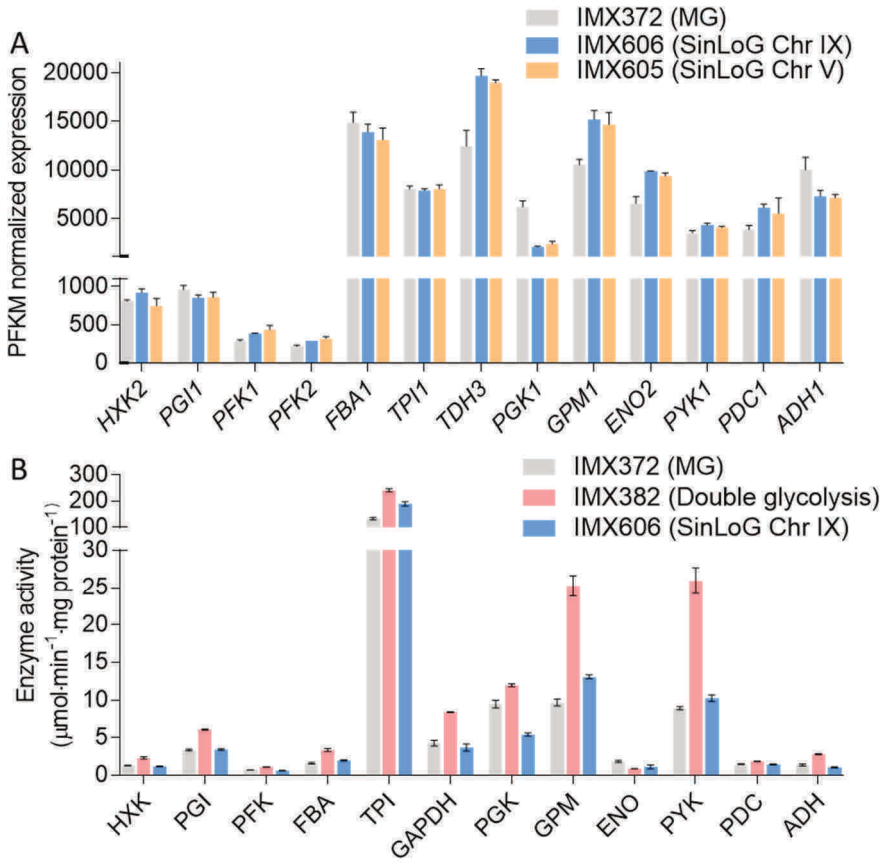


Figure 4 - Transcription and enzyme activity of the glycolytic genes in the single locus glycolysis strains.
A) PFKM-normalized transcript levels of the glycolytic genes in strains IMX372, IMX606 and IMX605 are compared. Error bars represent the standard error of the mean of independent culture duplicates. See Fig. 2C for reactions catalysed by each gene product. * indicates significant difference compared to IMX372 ($P_{adj} < 0.05$). **B)** Enzyme activities measured in cell-free extracts of the minimal glycolysis strain IMX372, the double glycolysis strain IMX382 and single locus glycolysis strain IMX606 (SwYG, SinLoG-IX). The data for IMX372 and IMX606 shown here is reproduced from [6] and is the average of at least four independent culture replicates. IMX382 was measured in biological duplicates. Error bars represent the standard error of the mean. * indicates significant difference to IMX372 ($P < 0.05$, student t-test, 2 tailed, homoscedastic).

Discovery-driven approaches to elucidate the molecular basis of SwYG slow growth rate

While the targeted engineering strategy enabled the elimination of factors that were not causal for the SwYG strain's slower growth rate, it failed to directly identify the responsible factors. In an attempt to identify these mechanisms, gene expression

profiles in the SwYG strains with a SinLoG in chromosome V or IX and their ancestor the MG strain were mined. There again, while genes were up- and down-regulated in the strains with a SinLoG as compared to the MG strain, transcriptome analysis did not identify specific genes, sets of genes or mechanisms potentially involved in the SwYG strain slow growth (Supplementary Fig. S5).

As transcriptome analysis proved inconclusive, an adaptive laboratory evolution approach was undertaken. To increase the growth rate of SwYG, repeated transfers of IMX1821 and the control strain CEN.PK113-7D were performed on chemically defined glucose medium for approximately 525 generations (Fig. 1). Three independent evolution lines successfully led to SwYG populations with growth rate comparable to that of the reference strain CEN.PK113-7D (Fig. 5 and Supplementary Fig. S6). Three fast growing single colony isolates were selected from independent evolution lines of SwYG and the control strain, and their genome was sequenced. All three independent single locus glycolysis evolution lines carried mutations in *ATG41, CNB1* and *SUR2* (Fig. 5B and for strain lineage Supplementary Fig. S5). *CNB1*, encoding a regulatory subunit of calcineurin, was also mutated in one of the evolution lines of the CEN.PK113-7D control strain, and was most likely not the causal mutation for the SwYG strain phenotype (Supplementary Fig. S7). Introducing the *CNB1* mutation of the evolved strains in the non-evolved SwYG strain confirmed this hypothesis (Fig. 5A). Conversely, *ATG41* and *SUR2* were not mutated in the control strain evolution lines and could therefore be involved in alleviating the slower growth phenotype of the SwYG strain. *SUR2* is involved in sphingolipids biosynthesis and encodes a sphinganine C4-hydroxylase [12-14]. Sphingolipids are important membrane components and act as messengers and bioactive molecules in cell growth, senescence, autophagy and apoptosis [15]. In all three evolution lines the Sur2 fatty acid hydroxylase domain (amino acids 162 to 297) was mutated. In one of the evolved strains the mutation caused the occurrence of an early stop codon, suggesting that the mutations most likely caused a loss of function of this non-essential gene [16]. Both *SUR2* deletion and reverse-engineering of the *SUR2* mutations were performed in the SwYG parental strain IMX1821, but did not result in a measurable phenotypic effect (Fig. 5A). The exact function of *ATG41* has not been identified, but it is associated with the formation of the autophagosome [17]. Similar to Sur2, several mutations found in Atg41 truncated the protein, removing its extreme C-terminal domain (amino acids 127-136). This region is essential for the interaction with protein partners (more particularly Atg9) and thus correct subcellular distribution and autophagy activity [17]. As *SUR2*, *ATG41* was therefore deleted both in the SwYG parental strain IMX1821 and modified with the mutations found in the evolved strain. While individual mutations hardly affected the growth rate of the SwYG strain, the combined deletion or mutation of both *SUR2* and *ATG41* successfully restored the growth rate of the evolved strains (Fig. 5A). These results

demonstrated that mutations in both genes were required to fix the slower growth rate phenotype of the SwYG strain.

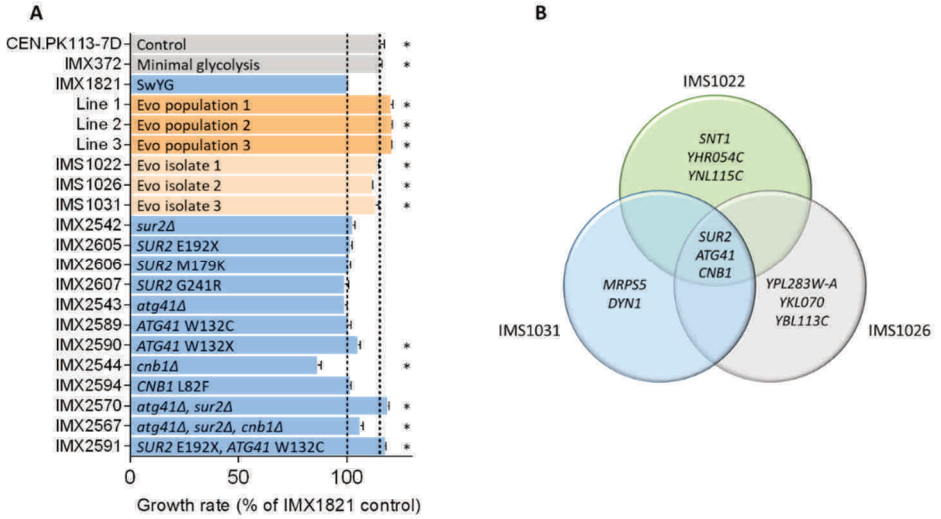


Figure 5 - Physiological and genetic characterization of the evolved and reverse engineered strains derived from the SwYG strain.
A) Specific growth rates of the evolved population of the SwYG strain, single colony isolates of the evolved populations and reverse engineered strains, shown as percentage of reference SwYG strain IMX1821. Control strains without a single locus glycolysis are shown in grey, evolved strains in orange and engineered SwYG strains are shown in blue, dashed lines indicate growth rates of the reference strains IMX1821 and CEN.PK113-7D. Significant differences compared to strain IMX1821 are indicated (*, $P<0.05$, T test, homoscedastic). **B)** Venn diagram showing genes in which mutations were identified in the evolved strains derived from IMX1821.

Discussion

With the ambition to consolidate the SwYG strain as platform for glycolysis swapping and to identify design principles for integration of large, modular pathways in *S. cerevisiae*, we aimed to characterize and identify underlying factors of the slower growth phenotype observed upon genetic clustering of the glycolytic pathway in the SwYG strain.

This in-depth study revealed that genetic factors directly or indirectly related to the integration of the SinLoG were not responsible for the SwYG strain slow growth rate, a promising outcome for integration of large DNA constructs in yeast. ALE rather identified the involvement of autophagy and sphingolipids biosynthesis in the slower growth rate phenotype, a response difficult to reconcile with the genetic interventions performed in the SwYG strain. These results might yet be explained if we consider an alternative theory, initially discarded due to its unlikelihood, but that might, in the light of all the data obtained in this study, explain the SwYG strain phenotype. The

SinLoG integration in the MG strain causes a duplication of the glycolysis and fermentation genes. Glycolysis and fermentation proteins are amongst the most abundant proteins in *S. cerevisiae* cells, representing around 15% of total cell proteins under fermentative conditions [18]. Integration of a second copy of these genes might increase the abundance of these proteins and, considering that proteins are extremely costly to synthetize, cause an energetic burden to the cells, reflected in a decreased growth rate. Enzyme assays confirmed the increased specific activity of most of these enzymes by 1.5- to 2-fold in the double glycolysis strain (IMX382, Fig. 4B). This increase in activity, with the reasonable assumption that it represents an equivalent increase in protein abundance, should result in a decrease in growth rate. Indeed individual overexpression of glycolytic enzymes has been shown to lead to decreased growth rates when around 15% of the total cellular protein is made up of gratuitous glycolytic protein [19]. Similarly overexpression of fluorescent proteins to high percentages of the proteome (estimated around 10-15%) has been shown to lead to a growth rate decrease of up to 15% from the cost of protein production [20]. This overexpression alone therefore could explain the decrease in growth rate measured for the double glycolysis strain (IMX382, Fig. 2B).

Building on the realistic working hypothesis that protein burden is the major cause of the decreased growth rate in the double glycolysis strain, subsequent removal of the gene duplicates in this strain is expected to alleviate this burden and to restore growth rate to the levels of the MG strain. However, deletion of the 13 duplicated genes from the double glycolysis strain did not alter growth rate (Fig. 2). This stability in growth rate upon deletion of the gene duplicates could reflect the compensation of the expected positive effect of sequential deletion by the occurrence of deleterious events during strain construction. While analysis of the genome sequence and transcriptome of the SwYG strain indicated that the targeted genetic modifications did not have measurable direct or indirect deleterious effects on yeast physiology, the 18 transformation rounds necessary to engineer IMX382 into the SwYG strain could have caused unwanted mutations in the yeast genome. Accordingly, the sequence of the SwYG strain revealed six mutations in coding regions of the native genome (Supplementary Table S1), three of which led to amino acid changes in Opt1 (plasma membrane oligopeptide transporter), Cwc25 (essential protein, component of spliceosomal complex implicated in the catalytic step of pre-mRNA splicing) and more interestingly Vps15 a serine/threonine protein kinase, key component in the synthesis of phosphatidylinositol 3-phosphate (PI(3)P) required for autophagosome formation [21-25]. The deletion of *VPS15* as well as the other main component of the PI3-kinase complex, *VPS34*, has been reported to lead to growth defects and cause aberrant endosome morphology [26, 27]. During ALE, mutations in *ATG41*, an important protein for autophagosome formation [15, 17], would

counterbalance the deleterious effect of the mutated *VPS15*. The role of *SUR2* mutations in restoring the growth rate of the SwYG strain might appear more puzzling, however, evidence for the role of ceramides, important membrane lipids, in autophagy is growing [12-15, 28]. However, repair of the *VPS15* in the IMX589 SwYG strain did not lead to any effect on growth rate, suggesting this mutation in itself is not responsible for the observed growth rate defect (Supplementary Fig. S8).

An additional, particularly attractive scenario in the context of engineering glycolysis, involves Pgk1. In mammals the Pgk1 homologue, also called Pgk1, plays a key role in the formation and activation of the complex responsible for PI(3)P synthesis and autophagosome formation [29]. The PI3-kinase complex is largely conserved between mammalian cells and yeast and the human Pgk1 interaction partner, Beclin1, complements the function of its yeast homolog Vps30 [30]. The potential involvement of the yeast Pgk1 in PI(3)P synthesis and autophagy has not been explored yet, but it is not unlikely considering the conservation of moonlighting functions of several glycolytic proteins between yeast and mammals [31-33]. Remarkably, the activity of Pgk1 is two-fold lower in SwYG strains as compared to the minimal glycolysis control strain (Fig. 4). While this lower activity of an enzyme operating far from saturation is most likely not hindering the glycolytic flux, it might be deleterious for a potential moonlighting function. If such a moonlighting exists, deletion of the second copy of Pgk1 in the double glycolysis strain (strain IMX557, Fig. 2) and ensuing lower Pgk1 abundance might have a negative effect on autophagosome formation and therefore growth. Further research will be required to test these hypotheses, including the integration of a second copy of Pgk1 in the SwYG strain.

While *S. cerevisiae* is a fantastic microbial cell factory and model organism, with a remarkable genetic accessibility, the present study illustrates the challenges faced by intensive strain construction. It also illustrates the limits of analytical tools and highlights the need to combine multiple techniques to identify the molecular basis of mild phenotypic responses. Still, this study enabled a first, in-depth investigation of the impact of integration of large DNA constructs in *S. cerevisiae* native chromosomes, thereby identifying design guidelines for pathway transplantation, but also pathway design. The *SGA1* and *CAN1* loci proved to be robust integration sites for constructs of 30-40 kb, and can most likely harbour longer constructs. Addition of ARS was not required, although the spacing between adjacent ARS was increased from 38 Kb to 73 Kb, an observation in line with previous reports [34]. Transcriptional hotspots, different from the native genetic configuration with a mosaic of genes with different transcription levels, did not harm yeast physiology, and the orientation of transcription did not require careful design. As in most organisms, in *S. cerevisiae* transformation is mutagenic, it is therefore recommended to keep the number of

transformations to a minimum, an achievement made possible using CRISPR-Cas mediated multiplex editing. Finally, efforts to relocalize genes for genome modularization should consider watermarking genes, such that only the desired (native) copy can be exclusively targeted and edited [35]. This approach minimizes the risk of excessive removal of DNA sequences during gene deletion. Applicable to native and synthetic chromosomes, these design guidelines should facilitate future large scale strain engineering endeavours.

Materials and Methods

Media, strain cultivation and stocking

The yeast strains used in this study are all derived from the CEN.PK lineage (Supplementary Table S2) [36]. For non-selective growth, *S. cerevisiae* strains were cultivated on Yeast extract Peptone Dextrose (YPD) medium, consisting of 10 g L^{-1} Bacto yeast extract, 20 g L^{-1} Bacto peptone and 20 g L^{-1} glucose. For selective growth, Synthetic Medium (SM) was used, containing: 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 5 g L^{-1} $(NH_4)_2SO_4$, 1 mL L^{-1} of a trace element solution [37]. YPD and SM medium were set to pH 6 by addition of KOH and 20 g L^{-1} Bacto agar was added in case of solid medium. YPD and SM media were autoclaved for 20 min at 110°C and 121°C, respectively. Finally, for Synthetic Medium with glucose (SMD), 1 ml L^{-1} of a filter sterilized vitamin solution and 20 g L^{-1} of glucose separately autoclaved for 20 min at 110°C was added. When needed SMD was supplemented with 125 mg L^{-1} histidine and/or 150 mg L^{-1} uracil. When the dominant markers *amdS*, *KanMX* or *hphNT1* were used with SM, $(NH_4)_2SO_4$ was omitted from the medium and replaced by 6.6 g L^{-1} K_2SO_4. For amdS selection, 1.8 g L^{-1} filter sterilized acetamide was used as nitrogen source. For selection of *KanMX* or *hphNT1*, 2.3 g L^{-1} urea was used as nitrogen source instead and 200 mg L^{-1} G418 and 200 mg L^{-1} hygromycin (Hyg) were added to the medium, respectively. For YPD medium, the G418 and hygromycin were directly added to the medium. Yeast cultures were grown at 30°C at 200 rpm in 50-/100-/500-mL shake flasks containing respectively 10-/20-/100 mL of medium, in an Innova 44 incubator shaker (New Brunswick Scientific, Edison, NJ). Yeast on solid media were incubated until single colonies were visible (approximately 3 days).

Escherichia coli XL1-Blue strains with ampR containing plasmids were cultivated on Lysogeny Broth (LB) medium, consisting of: 10 g L^{-1} tryptone, 5.0 g L^{-1} yeast extract, and 5 g L^{-1} NaCl, supplemented with 100 mg mL^{-1} ampicillin. *E. coli* cultivations were performed at 37°C and 200 rpm in an Innova 4000 incubator (New Brunswick Scientific) shaker, in 15 mL Greiner tubes containing 5 mL medium. Cultures on solid LB medium, containing 20 g L^{-1} Bacto agar, were incubated overnight at 37°C.

S. cerevisiae and *E. coli* strains were stored at -80°C in 1 mL aliquots of appropriate medium containing 30% (v/v) glycerol.

Molecular biology techniques

DNA used for strain construction purposes was amplified by PCR using Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA), according to the manufacturer's instructions. All diagnostic PCRs were performed with DreamTaq PCR Master Mix (Thermo Fisher Scientific), following the supplier's instruction. Primers were ordered desalted unless used for amplification of open reading frames (ORFs) in which case PAGE purified primers were used (Sigma-Aldrich, St. Louis, MO).

DNA was purified using the GenElute PCR Clean-Up kit (Sigma-Aldrich) or the GeneJET PCR Purification Kit (Thermo Fisher Scientific) when no unspecific bands were present. In case of unspecific bands, the DNA was separated on a 1% or 2% (w/v) agarose (TopVision Agarose, Thermo Fisher Scientific) gel in 1x Tris-acetate-EDTA (TAE) buffer (Thermo Fisher Scientific) or 1x Tris-Borate-EDTA (TBE) (Thermo Fisher Scientific) buffer. 10 µl L^{-1} SERVA (SERVA Electrophoresis GmbH, Heidelberg, Germany) was added to the gel for DNA staining. As size standards, the GeneRuler DNA Ladder mix (Sigma-Aldrich) or GeneRuler DNA Ladder 50bp (Sigma-Aldrich) were used. DNA was isolated from gel using the GenElute PCR Clean-Up kit (Sigma-Aldrich). The purity and quantity of DNA was assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). For more precise DNA quantification the Qubit dsDNA BR Assay kit (Thermo Fisher Scientific) in combination with the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) was used. Digestion of DNA with DpnI was performed with the FastDigest restriction enzyme (Thermo Fisher Scientific), following the manufacturer's protocol. Gibson assembly for construction of gRNA plasmids was performed with Gibson assembly master mix 2x (New England Biolabs, Ipswich, MA) according to the manufacturer's instructions but scaled down to a final volume of 5 µl.

DNA for plasmid assembly was transformed into *E.coli* XL1-Blue by chemical transformation as described by Inoue *et al.* [38]. Plasmids were isolated using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich, St. Louis, MO) or the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) and verified by diagnostic PCR or restriction analysis. *S. cerevisiae* was transformed with the lithium acetate/polyethylene glycol method [39]. Single colony isolates were obtained by three consecutive re-streaks on selective solid medium. Yeast DNA was isolated by either boiling in 0.02 N NaOH, the protocol described by Looke *et al.* [40], the YeaStar genomic DNA kit (Zymo Research, Irvine, CA) or the QIAGEN Blood & Cell

Culture Kit with 100/G Genomic-tips (Qiagen, Hilden, Germany) depending on the desired DNA purity.

Adaptive laboratory evolution

CEN.PK113-7D and IMX1821 were evolved for faster growth on aerobic SMD medium at 30°C and 150 rpm in 100 mL shake flasks with 20 mL medium. The strains were inoculated from a freezer stock and subsequently transferred every 24 hours, by a 1:100 dilution in fresh medium. For both strains, three separate evolution lines were maintained for 79 transfers, when all IMX1821 evolution cultures had reached the approximate growth rate of wild type CEN.PK113-7D. The first 23 transfers of the CEN.PK 113-7D evolution line were reported previously [41]. From each evolution line, single colonies were isolated. The single cell lines that showed a growth rate representative of the evolved population were stored at -80°C and whole-genome sequenced (Supplementary Table S2).

Plasmid Construction

To target the Cas9 protein to a desired locus and induce a double strand break, gRNA plasmids were constructed according to Mans *et al.* [42] with minor alterations. To summarize, plasmids containing two gRNAs were Gibson assembled from a backbone amplified with primer 6005 from a pROS plasmid, and two insert fragments amplified with primers 5974 and 5975 and a respective gRNA specific primer. In Supplementary Table S3A all constructed plasmids are listed. Supplementary Table S4A-B contains all primers to construct and verify the gRNA plasmids.

Strain construction

All strains constructed in this study originate from IMX589, the uracil auxothrophic SwYG strain, constructed by Kuijpers *et al.* [6] (Supplementary Table S2). This strain is characterized by the deletion of the minor paralogs of glycolysis and the relocalization of the major paralogs of glycolysis to the *SGA1* locus on chromosome IX. Some of the strains used in this study to test hypotheses were previously constructed and described [6, 7, 35, 43]. All strains constructed by CRISPR-Cas9 have been made as described by Mans *et al.* [42] and all strains were verified by diagnostic PCR (primers in Supplementary Table S4).

To construct a strain in which the transcription of the glycolytic genes was aligned with replication an *in silico* design was made (Fig. 3A). The watermarked glycolytic genes and auxiliary parts were amplified from template DNA with SHR containing primers to allow for homologous recombination (Supplementary Tables S2B, S3C and S4). Strain IMX1338 was transformed with 1.5 µg of plasmid p426-SNR52p-gRNA.CAN1.Y-SUP4t (Supplementary Table S3) and 200 fmol of each fragment (13 glycolytic genes*: FBA1, TPI1, PGK1, ADH1, PYK1, TDH3, ENO2, HXK2, PGI1, PFK1, PFK2, GPM1, PDC1*, the *HIS3* marker, *ARS418* and *ARS1211*) to integrate

an altered orientation SinLoG in the *CAN1* locus (Chr. V). The gRNA plasmid was removed by non-selective growth and the uracil auxotrophic strain was stocked as IMX2359. To construct IMX2381, the glycolytic cassette in the *sga1* locus on Chr. IX was deleted from IMX2359 using gRNA plasmid pUDR413 and repaired with a *KlURA3* transcriptional unit. The *KlURA3* fragment was amplified from the pMEL10 plasmid with primers containing *sga1* flanks (Supplementary Table S3 and S4E). The gRNA plasmid was removed before stocking.

For reverse engineering of the mutations in *ATG41, SUR2* and *CNB1* found in the IMX1821 evolution lines, the three genes were first deleted by insertion of a synthetic gRNA and PAM sequence. To this end, IMX1821 was separately transformed with pUDR748, pUDR749 or pUDR750, to delete *ATG41*, *CNB1* and *SUR2*, respectively. 120 bp of corresponding repair fragments made by annealing of complementary primers containing the synthetic gRNA and PAM sequence were also supplied (Supplementary Table S4F). gRNA plasmids were removed, and strains were stocked as IMX2542, IMX2543 and IMX2544 for the deletion of *SUR2*, *ATG41* and *CNB1*, respectively. The mutated *SUR2* gene was amplified from evolution strains IMS1022, IMS1026 and IMS1031 with primer pair 17559/17560 (Supplementary Table S4G). Strain IMX2542 was transformed with plasmid pUDR114 to cut in the synthetic gRNA and separate repair fragments, resulting after plasmid removal in strains IMX2605 ($SUR2^{E192stop}$), IMX2606 ($SUR2^{M179K}$) and IMX2607 ($SUR2^{G251R}$). Similarly, the *ATG41* and *CNB1* genes were amplified from the evolution lines (primers in Supplementary Table S4G) and transformed together with plasmid pUDR114 in strain IMX2543 and IMX2544, respectively. These transformations resulted in strains IMX2589 ($ATG41^{W132C}$), IMX2590 ($ATG41^{W132stop}$) and IMX2594 ($CNB1^{L82F}$). For double deletion and replacement with a synthetic gRNA and PAM of *ATG41* and *SUR2*, strain IMX1821 was transformed with gRNA plasmid pUDR751 and corresponding 120 bp repair fragments (Supplementary Table S4F). After plasmid removal the resulting strain was stocked as IMX2570. Transformation of this strain with pUDR114 to cut the synthetic gRNAs and the amplified repair fragment (primers in Supplementary Table S4G) of *ATG41* and *SUR2* from evolution strain IMS1022, resulted in strain IMX2591 ($ATG41^{W132C}$ $SUR2^{E192stop}$). Finally the triple deletion strain IMX2567 was made by transforming IMX2544 (*CNB1::(synthethic gRNA)*) with gRNA plasmid pUDR751 targeting *ATG41* and *SUR2* and the corresponding repair fragments with synthetic gRNA and PAM (Supplementary Table S4F).

The repair of the *VPS15* gene in strain IMX589 was performed in a similar manner to the reverse engineering described above. Strain IMX589 was transformed with pUDR754 to delete *VPS15* and integrate a synthetic gRNA sequence, resulting in strain IMX2769 after plasmid loss and genotype confirmation by PCR (primers in

Supplementary Table S4F). Subsequently in this strain the non-mutated allele for *VPS15* was re-introduced by transformation with plasmid pUDR114 and a repair fragment amplified from CEN.PK113-7D (primers in Supplementary Table S4G). The verification of the *VPS15* locus was done by PCR and Sanger sequencing and the strain was stocked as IMX2813.

Growth rate determination

Shake flasks

The growth of strains was monitored from 100 mL SMD cultures in 500 mL shake flasks by following the optical density at 660 nm (OD_{660}) with a JENWAY 7200 spectrophotometer (Cole-Parmer, Stone, UK). Maximum specific growth rates (μ_{max}) were determined in at least biological duplicates from a minimal of six time point in exponential phase. To ensure that the cultures were in exponentially phase for growth rate measurements, two successive pre-cultures were performed prior to inoculation in the measuring flasks.

96-well format

Pre-cultures inoculated from glycerol-stocks were performed in 1.5 mL of YPD medium in 12-wells plate, and grown at 30°C and 800 rpm in a thermoshaker (Grant-bio PHMP-4, United Kingdom) until stationary phase. Of each strain 20 µL were transferred to fresh SMD (with uracil) medium in a new 12-well plate and grown until mid-exponential phase. Each strain was then transferred with six biological replicates to a 96-well plate (CELLSTAR, Greiner Bio-One). Both the 12-well and 96-well plates were covered with sterile polyester acrylate sealing tape (Thermo Scientific). The OD_{660} from the cultures in the 96-well plates were determined every 20 min using a plate reader (TECAN infinite M200 Pro. Tecan, Männedorf, Switzerland). The growth rate was determined from at least 6 points in exponential phase.

Determination of glycolytic enzyme activities and degree of saturation calculation

Glycolytic enzyme activities were determined in cell extracts obtained from mid-exponential shake flask cultures of strain IMX382 as previously described [44]. Spectrophotometric assays were done at 30 °C and 340 nm in 1 mL reactions in 2 mL cuvettes using a Hitachi model 100-60 spectrophotometer. Assays were performed as previously described [45], for phosphofructokinase the assay was performed as described elsewhere [46]. Activities are expressed as µmol substrate converted per mg protein per minute (U/min). Independent biological duplicates were used with at least two analytic replicates per assay.

Degree of saturation of phosphoglycerate kinase was calculated using the glucose uptake rates and Pgk activities reported in [6]. The soluble protein content was

assumed to be 33% and enzyme activities were expressed in mmol per gram dry weight per hour to obtain the maximal flux capacity. Dividing the observed glucose uptake rate by this flux capacity gives an indication of the enzyme saturation under these conditions.

Ploidy staining and cell cycle phase estimation

Circa 10^7 cells were sampled from mid-exponential shake-flasks cultures on YPD and centrifuged (5 min 4700 g). The pellet was washed with demineralized water, centrifuged again and resupended in 800 µl 70% ethanol while vortexing. Another 800 µl 70% ethanol was added and fixed cells were stored at 4 °C until analysis. Staining of cells with SYTOX© Green Nucleic Acid Stain (Invitrogen S7020) was performed as described earlier [47]. Samples were analyzed on a BD Accuri C6 flow cytometer with a 488 nm laser (BD Biosciences, Breda, The Netherlands). Fluorescence intensity was represented using FlowJo (V. 10.6.1, FlowJo, LLC, Ashland, OR, USA).

For cell cycle analysis, stained cells were selected from all events based on fluorescence at 533/30 nm. Subsequently the G_1 and G_2 populations corresponding to 1n and 2n were selected with ellipsoid gates based on forward scatter and fluorescence. All cells not corresponding to G_1 or G_2 were designated to be in the S phase.

Sequencing

Whole genome sequencing and alignment

High quality genomic DNA was isolated with the QIAGEN Blood & Cell Culture Kit with 100/G Genomic-tips (Qiagen) from IMX1821 and evolution strains IMS1022, IMS1026, IMS1031, IMS1043, IMS1048 and IMS1053. The strains were sequenced in-house using an Illumina MiSeq Sequencer (Illumina, San Diego, CA) as described previously [7, 35]. For long-read DNA sequencing of IMX589, genomic DNA was loaded on a R9 flowcell and sequenced on a MinION (Oxford Nanopore, sequencing kit SQK-LSK109). Guppy (Oxford Nanopore, version 3.9.5_GPU) was used for basecalling. The resulting fastq files were filtered on length (>1 kb) and *de novo* assembled with Canu (version 1.8) [48]. Resulting contigs were error-corrected by first mapping the Illumina paired-end read library of IMX589 with Burrows-Wheeler Algorithm (BWA mem, version 0.7.15) [49], converting the alignments to binary alignment format (BAM, samtools version 1.3.1) [50] and further processed with Pilon [51].

A *de novo* assembled reference genome was constructed for IMX589 (auxothrophic SwYG) using MinION and Miseq data [6]. Using the Burrows-Wheeler Alignment (BWA) Tool [52] (version 0.7.15), sequencing data of IMX1821, IMS1022, IMS1026 and IMS1031 was aligned to the IMX589 reference genome and sequencing data of

IMS1043, IMS1048 and IMS1053 was aligned to a CEN.PK113-7D reference [53]. The data was further processed using SAMTools [52] (version 1.3.1) and single nucleotide polymorphisms (SNPs) were determined using Pilon (with -vcf setting; version 1.18)) [51]. The BWA.bam output file was visualized using the Integrative Genomics Viewer (version 2.4.0) [54], and copy numbers were estimated using Magnolya (version 0.15) [55]. SNPs were compared between CENPK113-7D and the three CEN.PK113-7D evolutions lines (IMS1043, IMS1048 and IMS1053) and IMX1821 and the three SwYG evolution lines (IMS1022, IMS1026 and IMS1031).
All whole genome sequencing data are available at NCBI (https://www.ncbi.nlm.nih.gov/) under bioproject PRJNA757374.

Sanger sequencing
Sanger sequencing was performed at Baseclear (Baseclear, Leiden, The Netherlands) or Macrogen (Macrogen Europe B.V., Amsterdam, The Netherlands) to check the reverse engineering of SNPs in *ATG41, SUR2, CNB1* and *VPS15.*

RNA sequencing and analysis
RNA sequencing results from the MG strain (IMX372) grown in aerobic SMD bioreactor cultures were obtained from Solis-Escalante *et al.* [5]. For the SwYG strains with the glycolytic cassette in Chr. IX (IMX606) and in Chr. V (IMX605), RNA was sampled from the biological duplicate aerobic SMD bioreactor cultures run by Kuijpers *et al.* [6]. The sampling and processing of the transcriptome samples from these bioreactors was executed as described by Solis-Escalante *et al.* [5]. Library preparation and RNA sequencing were performed by Novogen. Sequencing of samples IMX605 and IMX606 was done with Illumina paired end 150 bp sequencing read system (PE150) using a 250 bp insert strand specific library which was prepared by Novogen.

Sequencing reads of samples IMX606, IMX605 and IMX372 were aligned to the CEN.PK113-7D genome using STAR [56]. Expression was quantified by featureCounts (version 1.6.0) [57]. Normalized FPKM counts were obtained by applying the rpkm function from the edgeR package [58]. Differential expression analysis was done by applying DESeq2 [59].

Transcriptomics datasets were uploaded to the Gene Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE190122.

Supplementary data

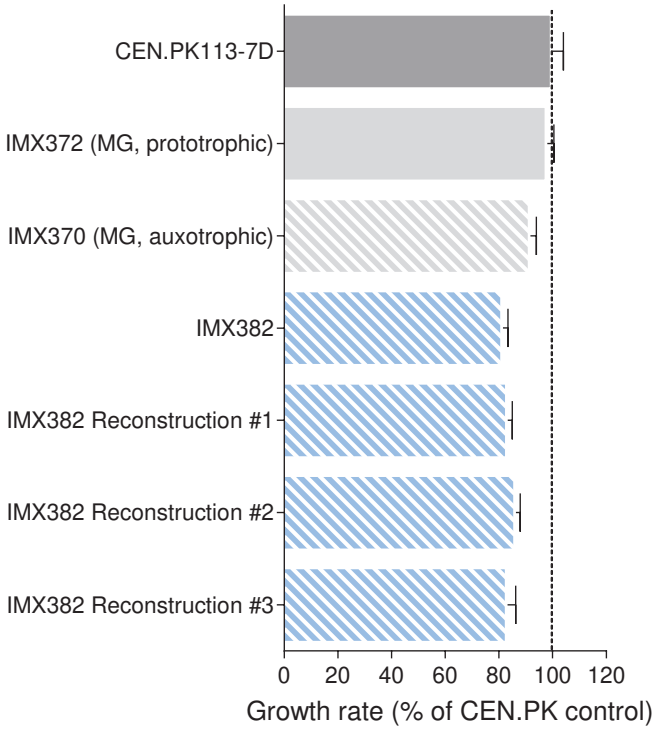


Figure S1 - Specific growth rates of yeast strains with Minimal Glycolysis background after integration of the single locus glycolysis.
Reconstruction of strain IMX382 (SinLoG integration in *sga1* in IMX370) led to a similar growth rate defect. All growth rates are normalized to the CEN.PK113-7D control strain from the same experiment, prototrophic strains are shown as filled bars, uracil auxotrophic strains with dashed fill. Statistically significant differences relative to the auxotrophic MG strain reference strain are indicated (* $P<0.05$, t-test 2-tailed, homoscedastic).

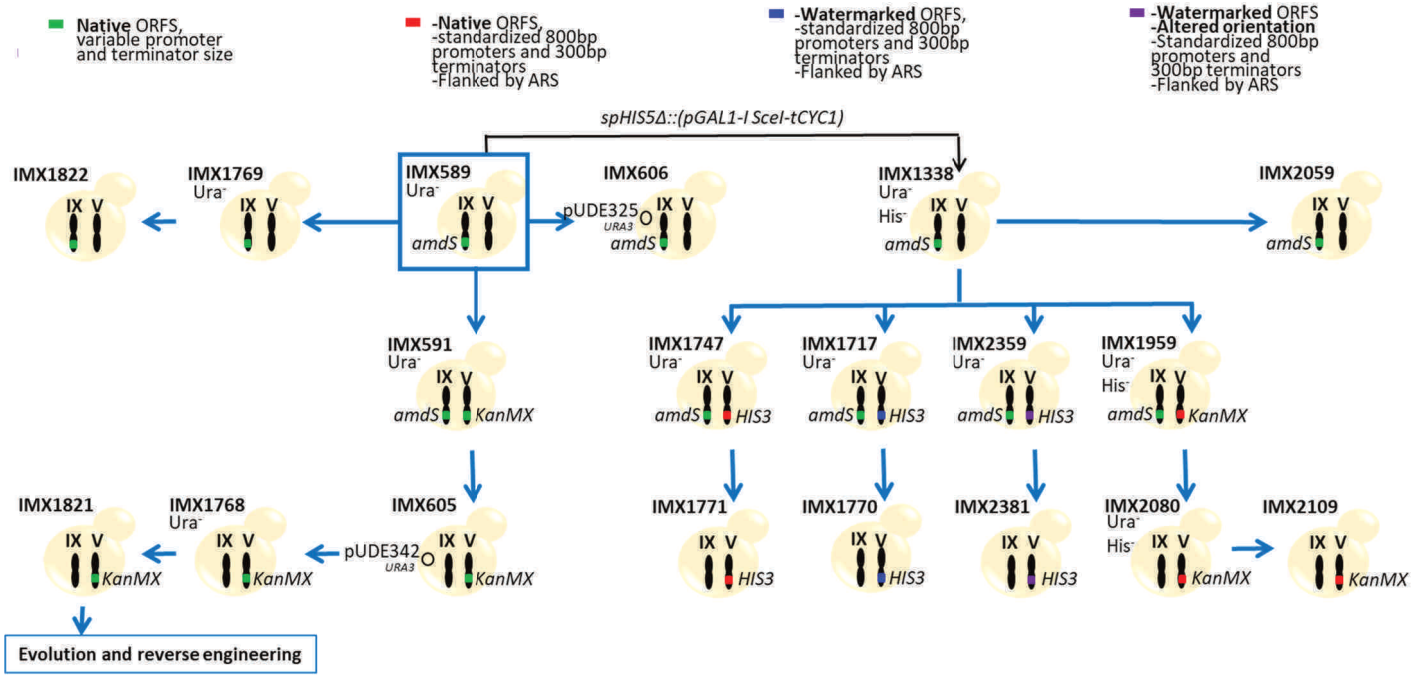
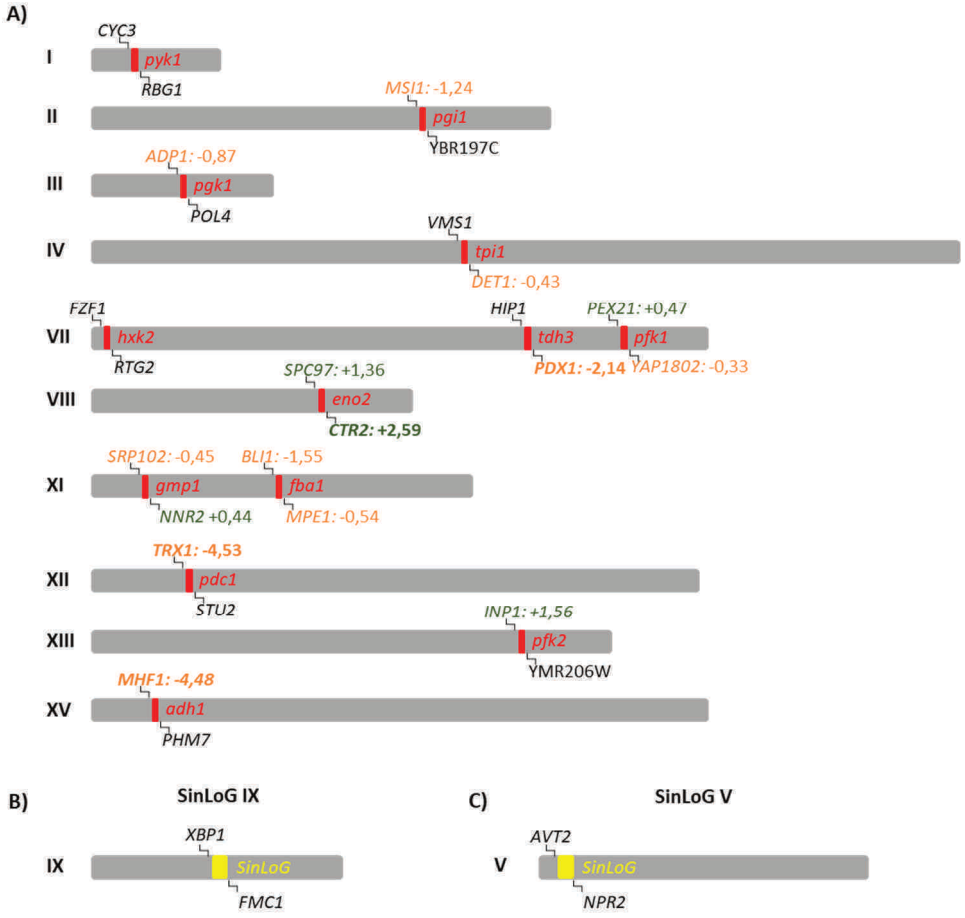


Figure S2 - Strain lineage of the single locus glycolysis strains.
The lineage of strain construction is indicated for the main single locus glycolysis strains used in this study. Various designs of SinLoG with changes in promotors and terminators, watermarking, orientation of genes, and presence of ARS sequences are indicated by color.


A)
I
CYC3
pyk1
RBG1
II
MSI1: -1,24
pgi1
YBR197C
III
ADP1: -0,87
pgk1
POL4
IV
VMS1
tpi1
DET1: -0,43
VII
FZF1
hxk2
RTG2
HIP1
tdh3
PDX1: -2,14
PEX21: +0,47
pfk1
YAP1802: -0,33
VIII
SPC97: +1,36
eno2
CTR2: +2,59
XI
SRP102: -0,45
gmp1
NNR2 +0,44
BLI1: -1,55
fba1
MPE1: -0,54
XII
TRX1: -4,53
pdc1
STU2
XIII
INP1: +1,56
pfk2
YMR206W
XV
MHF1: -4,48
adh1
PHM7
B)
SinLoG IX
IX
XBP1
SinLoG
FMC1
C)
SinLoG V
V
AVT2
SinLoG
NPR2

D)

Gene	Function	PFKM normalized counts
MSI1	Subunit of chromatin assembly factor I (CAF-1); chromatin assembly by CAF-1 affects multiple processes	53
ADP1	Putative ATP-dependent permease of the ABC transporter family	32
DET1	Acid phosphatase; involved in the non-vesicular transport of sterols in both directions between the endoplasmic reticulum and plasma membrane; deletion confers sensitivity to nickel	89
PDX1	E3-binding protein of the mitochondrial pyruvate dehydrogenase complex	105
PEX21	Peroxin required for peroxisomal matrix protein targeting	125
YAP1802	Protein of the AP180 family, involved in clathrin cage assembly	53
SPC97	Component of the microtubule-nucleating Tub4p (gamma-tubulin) complex	18
CTR2	Low-affinity copper transporter of the vacuolar membrane	148
SRP102	Signal recognition particle (SRP) receptor beta subunit; involved in SRP-dependent protein targeting	125
NNR2	Widely-conserved NADHX dehydratase; converts (S)-NADHX to NADH in ATP-dependent manner	45
BLI1	Subunit of the BLOC-1 complex involved in endosomal maturation	61
MPE1	Subunit of CPF cleavage and polyadenylation factor and E3 Ub-ligase; involved in 3' end formation of mRNA via cleavage and polyadenylation of pre-mRNA	51
TRX1	Cytoplasmic thioredoxin isoenzyme; part of the thioredoxin system that protects cells from oxidative and reductive stress	1580
INP1	Peripheral membrane protein of peroxisomes; involved in peroxisomal inheritance	16
MHF1	Component of the MFH histone-fold complex	115

Figure S3 - Relative transcription of genes neighboring the deletion loci of native glycolytic genes.

A) Relative expression in strain IMX606 (SinLoG-Chr IX) relative to the MG strain IMX372 of the genes neighbouring the deleted glycolytic loci is indicated. Genes which are decreased in expression are marked in orange, genes which are increased in expression are marked in green and genes which are not expressed or did not change in expression are marked black (Mean PFKM_MG>10 or Mean PFKM IMX606>10, padj<0.05). Genes with $\log_2$FoldChange>2 are marked in bold. **B,C)** Expression of genes neighbouring the SinLoG loci on Chr. IX (IMX606) and Chr. V (IMX605). **D)** Function and PFKM count of IMX372 (MG) of genes that are changed in expression. Genes were considered not expressed if PFKM<10 in all strains [7]; high expression was considered for PFKM>300.

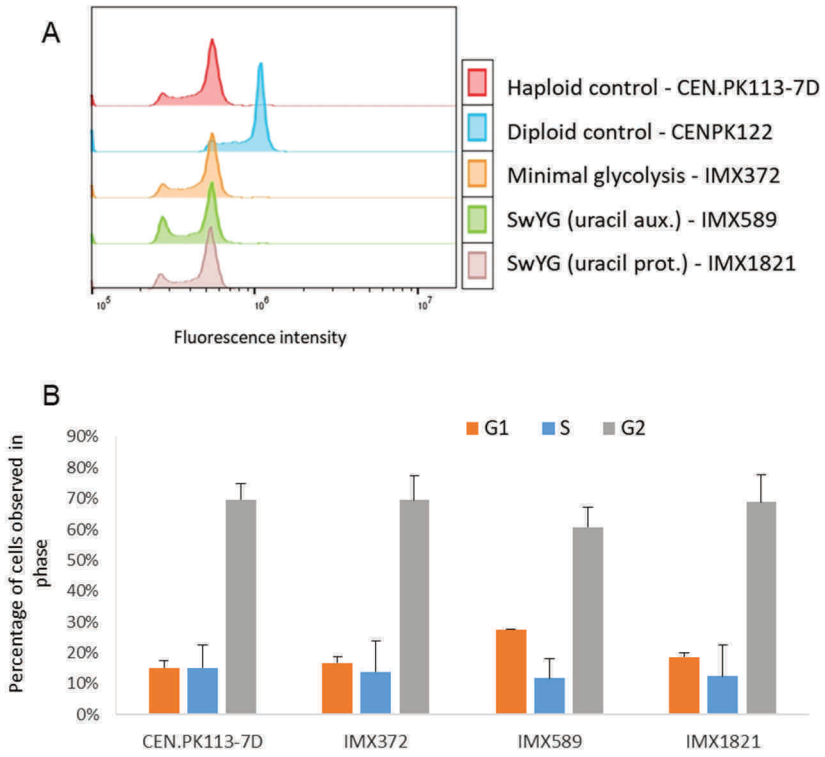


Figure S4 - Ploidy and cell cycle distribution estimation by flow cytometric analysis.
A) Ploidy measurement based on DNA staining and flow cytometry. All analysed single locus glycolysis strains match the haploid reference. **B)** Cell cycle distribution based on the DNA content shown in panel A. The estimated percentage of cells in each cell cycle phase is shown. Error bars represent standard deviation between duplicate samples. The single locus glycolysis strain IMX589 (Chr IX) shows a slight increase in cells in G1 compared to the CEN.PK113-7D and IMX372 (minimal glycolysis) control strains.

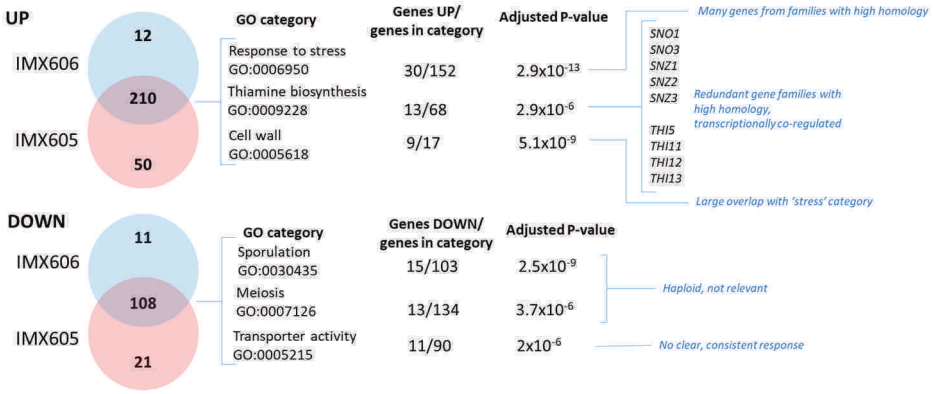


Figure S5 - Global transcription effects from the single locus glycolysis.
Global transcriptional changes in strains IMX605 (SinLoG on Chr V) and IMX606 (SinLoG on Chr IX) compared to the minimal glycolysis strain IMX372. Since both strains carrying a SinLoG share a similar growth rate enrichment analysis was performed on significantly up and down regulated genes in both strains to identify possible causal factors for this growth defect. Up- and down regulated genes with an adjusted p-value below 0.05 and a fold change greater than 2 were selected and doublons were removed. Enrichment analysis was performed with FunSpec (http://funspec.med.utoronto.ca/cgi-bin/funspec), GO terms with a Bonferroni-corrected p-value < 10^{-5} are reported. Although several categories are significantly enriched, no link to processes related to the integration of the SinLoG were identified.

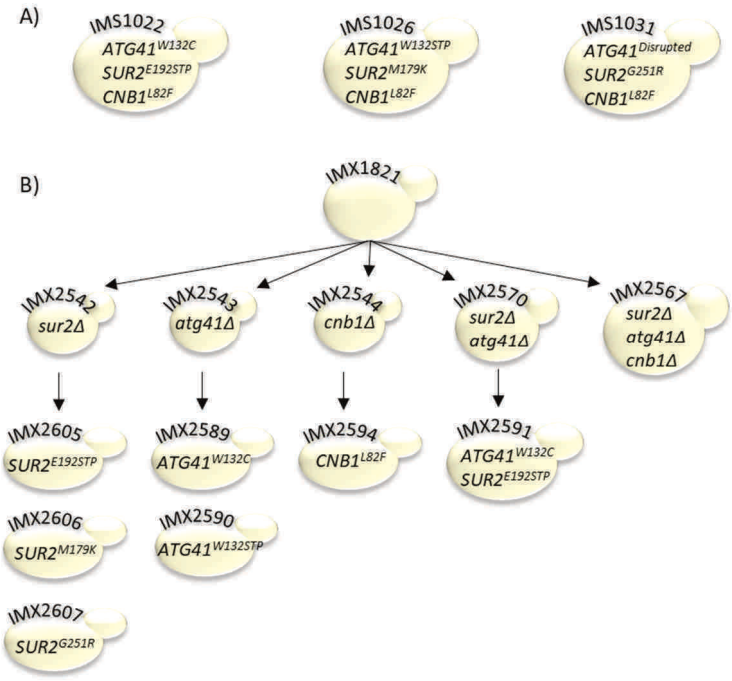


Figure S6 - Strain lineage of evolved isolates of the SwYG strain and reverse engineered strains.
A) Mutations in *ATG41*, *SUR2* and *CNB1* in evolved strains IMS1022, IMS1026 and IMS1031. **B)** Reverse engineering strategy of the *SUR2*, *ATG41* and *CNB1* mutations in the SwYG strain IMX1821.

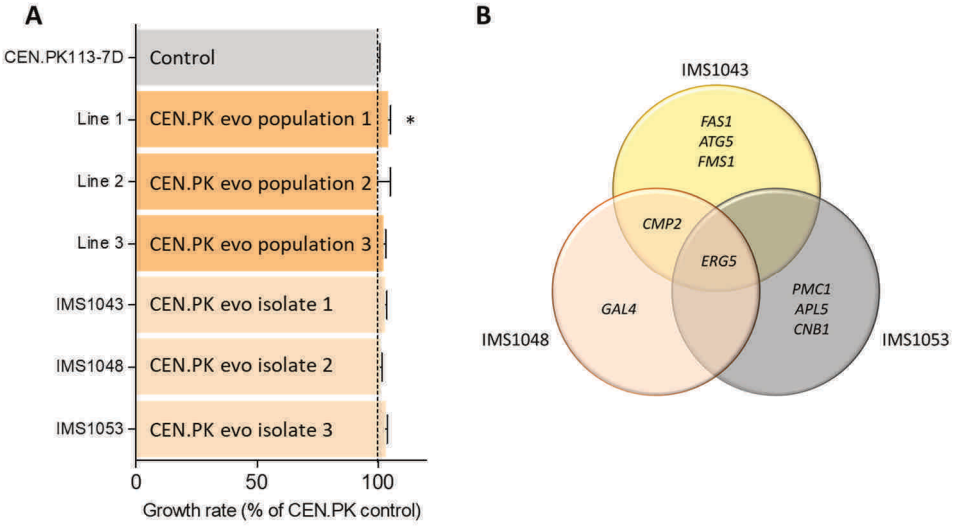


Figure S7 - Physiological and genetic characterization of populations and single colony isolates after evolution of the CEN.PK113-7D reference strain.
A) Specific growth rate CEN.PK113-7D and derived strains after evolution, represented as evolved population and single colony isolates. Growth rates are shown as percentage of the unevolved reference strain. No significant increase in growth rate was found except a slight increase for the population of line 1. Bars represent the average growth rate and standard deviation of duplicate shake flask cultures. Significant differences as compared to unevolved control are indicated (*, $P<0.05$, T test, homoscedastic). **B)** Venn diagram showing genes in which mutations were identified in the evolved strains derived from CEN.PK113-7D.

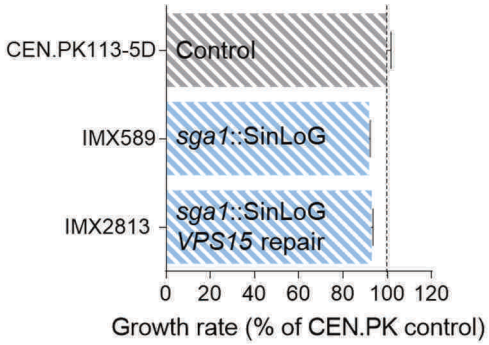


Figure S8 - Physiological characterization of a VPS15 repaired SwYG strain.
Specific growth rates measured in shake flask of the control strain CEN.PK113-7D, the uracil auxotrophic SwYG strain IMX589 and strain IMX2813, in which the IMX589 *VPS15* allele was restored to the native CEN.PK sequence.

Table S1 - Single nucleotide polymorphisms in the auxotrophic SwYG strain (IMX589) compared to MG strain (IMX372)
Information from Kuijpers *et al.* [6]

Mutations (SNP)	Genome	SinloG
All	15	9
In ORF	6[#]	1*
MisSense	3 (*VPS15, OPT1, CWC25*)	0

 # Information on the 6 in ORF SNPs in the genome

Systematic name	Name	Type	Amino Acid change
YBR079W	*VPS15*	Non-synonymous	E474K
YJL212C	*OPT1*	Non-synonymous	I463T
YNL245C	*CWC25*	Non-synonymous	P62L
YDL079C	*MRK1*	Synonymous	I190I
YLR180W	*SAM1*	Synonymous	V217V
YNL262W	*POL2*	Synonymous	F1536F

*** Information on the single SNP in the SinLoG**

Systematic name	Name	Type	Amino Acid change
YOL086C	*ADH1*	Synonymous	A180A

Tables S2-S5
Supplementary tables S2-S5 can be found in the supplementary materials file at the data.4TU repository via: https://doi.org/10.4121/16539840

References

1. Nielsen, J. and J.D. Keasling, Engineering Cellular Metabolism. Cell, 2016. **164**(6): p. 1185-1197.DOI: 10.1016/j.cell.2016.02.004.
2. Galanie, S., et al., Complete biosynthesis of opioids in yeast. Science, 2015. **349**(6252): p. 1095-100.DOI: 10.1126/science.aac9373.
3. Paddon, C.J., et al., High-level semi-synthetic production of the potent antimalarial artemisinin. Nature, 2013. **496**(7446): p. 528-32.DOI: 10.1038/nature12051.
4. Kondo, A., et al., Development of microbial cell factories for bio-refinery through synthetic bioengineering. Journal of Biotechnology, 2013. **163**(2): p. 204-216.DOI: 10.1016/j.jbiotec.2012.05.021.
5. Solis-Escalante, D., et al., A Minimal Set of Glycolytic Genes Reveals Strong Redundancies in *Saccharomyces cerevisiae* Central Metabolism. Eukaryot Cell, 2015. **14**(8): p. 804-16.DOI: 10.1128/EC.00064-15.
6. Kuijpers, N.G., et al., Pathway swapping: Toward modular engineering of essential cellular processes. Proceedings of the National Academy of Sciences, USA, 2016. **113**(52): p. 15060-15065.DOI: 10.1073/pnas.1606701113.
7. Postma, E.D., et al., A supernumerary designer chromosome for modular *in vivo* pathway assembly in *Saccharomyces cerevisiae.* Nucleic Acids Res, 2021. **49**(3): p. 1769-1783.DOI: 10.1093/nar/gkaa1167.
8. Hamperl, S. and K.A. Cimprich, Conflict Resolution in the Genome: How Transcription and Replication Make It Work. Cell, 2016. **167**(6): p. 1455-1467.DOI: 10.1016/j.cell.2016.09.053.
9. Dhar, M.K., S. Sehgal, and S. Kaul, Structure, replication efficiency and fragility of yeast ARS elements. Res Microbiol, 2012. **163**(4): p. 243-53.DOI: 10.1016/j.resmic.2012.03.003.
10. Tai, S.L., et al., Control of the glycolytic flux in *Saccharomyces cerevisiae* grown at low temperature - A multi-level analysis in anaerobic chemostat cultures. Journal of Biological Chemistry, 2007. **282**(14): p. 10243-10251.DOI: 10.1074/jbc.M610845200.
11. Gancedo, C. and C.L. Flores, Moonlighting proteins in yeasts. Microbiol Mol Biol Rev, 2008. **72**(1): p. 197-210, table of contents.DOI: 10.1128/MMBR.00036-07.
12. Haak, D., et al., Hydroxylation of *Saccharomyces cerevisiae* ceramides requires Sur2p and Scs7p. J Biol Chem, 1997. **272**(47): p. 29704-10.DOI: 10.1074/jbc.272.47.29704.
13. Grilley, M.M., et al., Syringomycin action gene SYR2 is essential for sphingolipid 4-hydroxylation in *Saccharomyces cerevisiae.* J Biol Chem, 1998. **273**(18): p. 11062-8.DOI: 10.1074/jbc.273.18.11062.
14. Megyeri, M., et al., Yeast ceramide synthases, Lag1 and Lac1, have distinct substrate specificity. J Cell Sci, 2019. **132**(12).DOI: 10.1242/jcs.228411.
15. Li, Y., et al., The pleiotropic roles of sphingolipid signaling in autophagy. Cell Death Dis, 2014. **5**(5): p. e1245.DOI: 10.1038/cddis.2014.215.
16. Desfarges, L., et al., Yeast mutants affected in viability upon starvation have a modified phospholipid composition. Yeast, 1993. **9**(3): p. 267-77.DOI: 10.1002/yea.320090306.
17. Yao, Z., et al., Atg41/Icy2 regulates autophagosome formation. Autophagy, 2015. **11**(12): p. 2288-99.DOI: 10.1080/15548627.2015.1107692.
18. Metzl-Raz, E., et al., Principles of cellular resource allocation revealed by condition-dependent proteome profiling. Elife, 2017. **6**: p. e28034.DOI: 10.7554/eLife.28034.
19. Eguchi, Y., et al., Estimating the protein burden limit of yeast cells by measuring the expression limits of glycolytic proteins. Elife, 2018. **7**: p. e34595.DOI: 10.7554/eLife.34595.

20. Kafri, M., et al., The Cost of Protein Production. Cell Rep, 2016. **14**(1): p. 22-31.DOI: 10.1016/j.celrep.2015.12.015.
21. Chiu, Y.F., et al., Cwc25 is a novel splicing factor required after Prp2 and Yju2 to facilitate the first catalytic reaction. Mol Cell Biol, 2009. **29**(21): p. 5671-8.DOI: 10.1128/MCB.00773-09.
22. Tseng, C.K., et al., A central role of Cwc25 in spliceosome dynamics during the catalytic phase of pre-mRNA splicing. RNA, 2017. **23**(4): p. 546-556.DOI: 10.1261/rna.059204.116.
23. Osawa, H., G. Stacey, and W. Gassmann, ScOPT1 and AtOPT4 function as proton-coupled oligopeptide transporters with broad but distinct substrate specificities. Biochem J, 2006. **393**(Pt 1): p. 267-75.DOI: 10.1042/BJ20050920.
24. Stack, J.H., et al., Vesicle-mediated protein transport: regulatory interactions between the Vps15 protein kinase and the Vps34 PtdIns 3-kinase essential for protein sorting to the vacuole in yeast. J Cell Biol, 1995. **129**(2): p. 321-34.DOI: 10.1083/jcb.129.2.321.
25. Obara, K., et al., Transport of phosphatidylinositol 3-phosphate into the vacuole via autophagic membranes in Saccharomyces cerevisiae. Genes Cells, 2008. **13**(6): p. 537-47.DOI: 10.1111/j.1365-2443.2008.01188.x.
26. Grunau, S., et al., The phosphoinositide 3-kinase Vps34p is required for pexophagy in *Saccharomyces cerevisiae.* Biochemical Journal, 2011. **434**(1): p. 161-170.DOI: 10.1042/Bj20101115.
27. Laidlaw, K.M., et al., Endosomal recycling to the surface mediated by Gpa1 and PI3-Kinase is inhibited by glucose starvation. bioRxiv, 2021.DOI: 10.1101/2021.04.02.438183.
28. Yamagata, M., K. Obara, and A. Kihara, Sphingolipid synthesis is involved in autophagy in *Saccharomyces cerevisiae.* Biochem Biophys Res Commun, 2011. **410**(4): p. 786-91.DOI: 10.1016/j.bbrc.2011.06.061.
29. Qian, X., et al., Phosphoglycerate Kinase 1 Phosphorylates Beclin1 to Induce Autophagy. Mol Cell, 2017. **65**(5): p. 917-931 e6.DOI: 10.1016/j.molcel.2017.01.027.
30. Liang, X.H., et al., Induction of autophagy and inhibition of tumorigenesis by beclin 1. Nature, 1999. **402**(6762): p. 672-6.DOI: 10.1038/45257.
31. Lu, M., et al., Physical interaction between aldolase and vacuolar H+-ATPase is essential for the assembly and activity of the proton pump. J Biol Chem, 2007. **282**(34): p. 24495-503.DOI: 10.1074/jbc.M702598200.
32. Lu, M., et al., The glycolytic enzyme aldolase mediates assembly, expression, and activity of vacuolar H+-ATPase. J Biol Chem, 2004. **279**(10): p. 8732-9.DOI: 10.1074/jbc.M303871200.
33. Mayordomo, I. and P. Sanz, Human pancreatic glucokinase (GlkB) complements the glucose signalling defect of *Saccharomyces cerevisiae hxk2* mutants. Yeast, 2001. **18**(14): p. 1309-16.DOI: 10.1002/yea.780.
34. Karas, B.J., Y. Suzuki, and P.D. Weyman, Strategies for cloning and manipulating natural and synthetic chromosomes. Chromosome Research, 2015. **23**(1): p. 57-68.DOI: 10.1007/s10577-014-9455-3.
35. Boonekamp, F.J., et al., Design and Experimental Evaluation of a Minimal, Innocuous Watermarking Strategy to Distinguish Near-Identical DNA and RNA Sequences. ACS Synth Biol, 2020. **9**(6): p. 1361-1375.DOI: 10.1021/acssynbio.0c00045.
36. Entian, K.-D. and P. Kötter, 25 Yeast Genetic Strain and Plasmid Collections, in *Yeast Gene Analysis - Second Edition*, I. Stansfield and M.J.R. Stark, Editors. 2007, Academic Press. p. 629-666.
37. Verduyn, C., et al., Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast, 1992. **8**(7): p. 501-17.DOI: 10.1002/yea.320080703.

38. Inoue, H., H. Nojima, and H. Okayama, High efficiency transformation of *Escherichia coli* with plasmids. Gene, 1990. **96**(1): p. 23-8.DOI: 10.1016/0378-1119(90)90336-p.
39. Gietz, R.D. and R.A. Woods, Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods in enzymology, 2002. **350**: p. 87-96.
40. Looke, M., K. Kristjuhan, and A. Kristjuhan, Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques, 2011. **50**(5): p. 325-8.DOI: 10.2144/000113672.
41. Perli, T., et al., Adaptive Laboratory Evolution and Reverse Engineering of Single-Vitamin Prototrophies in *Saccharomyces cerevisiae.* Appl Environ Microbiol, 2020. **86**(12).DOI: 10.1128/AEM.00388-20.
42. Mans, R., et al., CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae.* FEMS Yeast Research, 2015. **15**(2).DOI: 10.1093/femsyr/fov004.
43. Boonekamp, F.J., et al., A yeast with muscle does not run faster: full humanization of the glycolytic pathway in *Saccharomyces cerevisiae.* bioRxiv, 2021.DOI: 10.1101/2021.09.28.462164.
44. Postma, E., et al., Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae.* Appl Environ Microbiol, 1989. **55**(2): p. 468-77.DOI: 10.1128/aem.55.2.468-477.1989.
45. Jansen, M.L.A., et al., Prolonged selection in aerobic, glucose-limited chemostat cultures of *Saccharomyces cerevisiae* causes a partial loss of glycolytic capacity. Microbiology (Reading), 2005. **151**(Pt 5): p. 1657-1669.DOI: 10.1099/mic.0.27577-0.
46. Cruz, L.A., et al., Similar temperature dependencies of glycolytic enzymes: an evolutionary adaptation to temperature dynamics? BMC Syst Biol, 2012. **6**: p. 151.DOI: 10.1186/1752-0509-6-151.
47. Haase, S.B. and S.I. Reed, Improved flow cytometric analysis of the budding yeast cell cycle. Cell Cycle, 2002. **1**(2): p. 132-6.
48. Koren, S., et al., Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res, 2017. **27**(5): p. 722-736.DOI: 10.1101/gr.215087.116.
49. Li, H., Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997, 2013.
50. Li, H., et al., The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009. **25**(16): p. 2078-9.DOI: 10.1093/bioinformatics/btp352.
51. Walker, B.J., et al., Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One, 2014. **9**(11): p. e112963.DOI: 10.1371/journal.pone.0112963.
52. Li, H. and R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 2009. **25**(14): p. 1754-60.DOI: 10.1093/bioinformatics/btp324.
53. Salazar, A.N., et al., Nanopore sequencing enables near-complete de novo assembly of *Saccharomyces cerevisiae* reference strain CEN.PK113-7D. FEMS Yeast Res, 2017. **17**(7).DOI: 10.1093/femsyr/fox074.
54. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform, 2013. **14**(2): p. 178-92.DOI: 10.1093/bib/bbs017.
55. Nijkamp, J.F., et al., *De novo* detection of copy number variation by co-assembly. Bioinformatics, 2012. **28**(24): p. 3195-202.DOI: 10.1093/bioinformatics/bts601.
56. Dobin, A., et al., STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 2013. **29**(1): p. 15-21.DOI: 10.1093/bioinformatics/bts635.

57. Liao, Y., G.K. Smyth, and W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 2014. **30**(7): p. 923-30.DOI: 10.1093/bioinformatics/btt656.
58. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 2010. **26**(1): p. 139-40.DOI: 10.1093/bioinformatics/btp616.
59. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 2014. **15**(12): p. 550.DOI: 10.1186/s13059-014-0550-8.

Chapter 3

A yeast with muscle doesn't run faster: full humanization of the glycolytic pathway in *Saccharomyces cerevisiae*

Francine J. Boonekamp[#], Ewout Knibbe[#], Marcel A. Vieira-Lara, Melanie Wijsman, Marijke A.H. Luttik, Karen van Eunen, Maxime den Ridder, Reinier Bron, Ana Maria Almonacid Suarez, Patrick van Rijn, Justina C. Wolters, Martin Pabst, Jean-Marc Daran, Barbara Bakker, Pascale Daran-Lapujade

[#] These authors contributed equally to this work and should be considered co-first authors.

Manuscript under review at Cell Reports

Abstract

While transplantation of single genes in yeast plays a key role in elucidating gene functionality in metazoans, technical challenges hamper the humanization of full pathways and processes. Empowered by advances in synthetic biology, this study demonstrates the feasibility and implementation of full humanization of glycolysis in yeast. Single gene and full pathway transplantation revealed the remarkable conservation of both glycolytic and moonlighting functions and, combined with evolutionary strategies, brought to light novel, context-dependent responses. Remarkably, human hexokinase 1 and 2, but not 4, required mutations in their catalytic or allosteric sites for functionality in yeast, while hexokinase 3 was unable to complement its yeast ortholog. Comparison with human tissues cultures showed the preservation of turnover numbers of human glycolytic enzymes in yeast and human cell cultures. This demonstration of transplantation of an entire, essential pathway paves the way to the establishment of species, tissue and disease-specific metazoan models.

Introduction

Due to its tractability and genetic accessibility, *S. cerevisiae* has played and still plays a key role as simplified model organism for higher eukaryotes. Many discoveries in yeast native processes such as the cell cycle and ribosome biogenesis were pivotal for understanding their mammalian equivalents [1, 2]. In addition, yeast is used to study a wide range of diseases such as cancer, diabetes and neurodegenerative diseases [3]. In a large part of these studies, the heterologous expression of human genes in yeast enables the detailed investigation of human biology and disease-specific variations of human genes [4]. As the yeast and human genome share over 2000 groups of orthologs [4], several large initiatives have explored the complementarity of human genes in yeast and shown a high degree of functional conservation [4-11]. These studies are however complicated by the genetic redundancy of eukaryotic genomes [12], which is even more prominent for genes encoding proteins with metabolic functions [13, 14]. While individual gene complementation in yeast is an interesting approach to characterize single human proteins, humanization of entire pathways or processes would greatly increase their usefulness. Such 'next level' yeast models hold the potential to capture the native functional context of the humanized proteins and to enable the study of more complex, multigene phenotypes, and epistatic interactions between genes. The feasibility of such extensive humanization projects depend largely on the replaceability of yeast genes by their human orthologs. Recent large scale humanization and bacterialization efforts of the yeast genome suggested that replaceability was better predicted on pathway- or process-basis than by sequence conservation [8, 15]. To date, reports of full humanization of pathways or protein complexes are scarce [5, 16-19]. However, rapid developments in synthetic biology have tremendously increased the ability to extensively remodel microbial genomes, and promise to bring more examples of large scale humanization in the future.

The Embden-Meyerhof-Parnas (EMP) pathway of glycolysis, which is near ubiquitous to eukaryotes, has a central role in carbon metabolism and is involved in a wide range of diseases in mammals, including cancer with the well-known Warburg effect [20]. So far, few single human glycolytic enzymes have been transplanted into yeast, mostly in large-scale complementation studies [6, 8, 9, 22-24]. Whether all human glycolytic enzymes can complement their yeast orthologs is however unknown. It is a particularly fascinating question as glycolytic enzymes, both in yeast and human are characterized by their versatility in moonlighting capabilities [25, 26]. The degree of conservation of these moonlighting functions between these two distant organisms has hardly been explored to date, with the exception of the human aldolase B (*Hs*ALDOB) and the glucokinase (*Hs*HK4) [23, 24]. To overcome the difficulty caused by genetic redundancy, a yeast strain in which the set of genes

encoding glycolytic enzymes has been minimized from 19 to 11 was previously constructed [27]. This minimal glycolysis (MG) strain is a perfect platform for single glycolytic gene complementation. Furthermore, a strain in which this minimized set of yeast glycolytic genes has been fully relocated to a single chromosomal locus (SwYG strain) enables the swapping of the entire yeast glycolytic pathway by any designer glycolysis with minimal genetic engineering [28]. In the present study, glycolysis swapping with the SwYG strain was used to evaluate the functionality of an entire human muscle glycolytic pathway in yeast to enable the construction of new model yeast strains. A combination of single gene complementation, full pathway humanization and adaptive laboratory evolution was used to explore the functionality of all human glycolytic genes in yeast. This led to the identification of mutations in human hexokinase 1 and 2 related to allosteric inhibition by glucose-6-phosphate, which appear to be required for functional expression in yeast. Finally the validity of yeast strains with humanized glycolysis as a model was evaluated by comparing the protein turnover number (k_{cat}) of the human glycolytic enzymes expressed in yeast with enzymes in their native environment from human skeletal muscle myotube cell cultures.

Results

All human glycolytic genes directly complement their yeast ortholog except for hexokinases 1-3

With the exception of hexokinases and F1,6bP aldolases, human and yeast glycolytic enzymes are highly conserved with 43% to 65% identity at protein level, as compared to the 32% identity at whole proteome level [8] (Fig. 1A). The human and yeast F1,6bP-aldolases belong to two different classes of enzymes and do not share homology at all at protein level [29] (Fig. 1A and Table S1). Among the four human hexokinases (*Hs*HK1 to *Hs*HK4), *Hs*HK4 is closest in size and sequence to *Sc*Hxk2 (ca. 30% protein identity) while *Hs*HK1, *Hs*HK2 and *Hs*HK3 are roughly twice the size of their yeast orthologs, with each subunit sharing ca. 30% identity with *Sc*Hxk2 ([30], Table 1). Due to the genetic redundancy of metabolic pathways in eukaryotes [13, 14], and associated difficulty of complementation studies, so far complementation in *S. cerevisiae* was only tested for eight human glycolytic genes [7-9, 22, 24], of which only *HsPGAM2* was unsuccessful (Table S1, [9]). Implementation of the MG yeast strain, which carries a single isoenzyme for each glycolytic step, with the notable exception of the *Sc*Pfk1 and *Sc*Pfk2 subunits of the hetero-octameric phosphofructokinase [27], considerably facilitates complementation studies (Fig. 1A). The ability of 25 human glycolytic genes to complement their yeast ortholog(s) was systematically explored by individual gene complementation in the MG strain. For enzymes with multiple splicing variants, the

canonical version was used (Fig. 1A and Table S1). However, as the two pyruvate kinase genes *HsPKLR* and *HsPKM* have tissue-specific splicing variants (*HsPKL* and *HsPKR for HsPKLR*, and *HsPKM1* and *HsPKM2* for *HsPKM*), all four variants were tested (Fig. 1A, Table S1 [31]). The 25 genes were codon-optimized, cloned downstream strong, constitutive promoters (Table S2) and individually cloned in the MG strain, after which the yeast ortholog(s) were removed (Fig. S1). Remarkably, 22 out of these 25 genes demonstrated direct complementation of their yeast orthologs for growth on glucose in their native form (Fig. 1B, Fig. S2). Additionally, *HsHK1* and *HsHK2* but not *HsHK3* also complemented their yeast orthologs, but only after a period of adaptation of several days. While most strains were only marginally affected by single humanization of the glycolytic genes, strains harbouring a human hexokinase 2, the aldolases, phosphoglycerate mutases and the glyceraldehyde-3P dehydrogenase GAPDH variant S had a strongly reduced growth rate, the strongest decrease (30%) occurring with *Hs*ALDOB (Fig. 1B). No clear correlation could be found between growth rate and conservation between human and yeast gene sequences or promoter strength (Fig S3). All human genes were Sanger-sequenced in the complementation strains, revealing that all besides *HsHK1* and *HsHK2* had the expected sequence (see following section). This study therefore demonstrated the absence of complementation of the native human *HsHK3* and the remarkable complementation by 22 out of 25 human genes of their yeast orthologs.

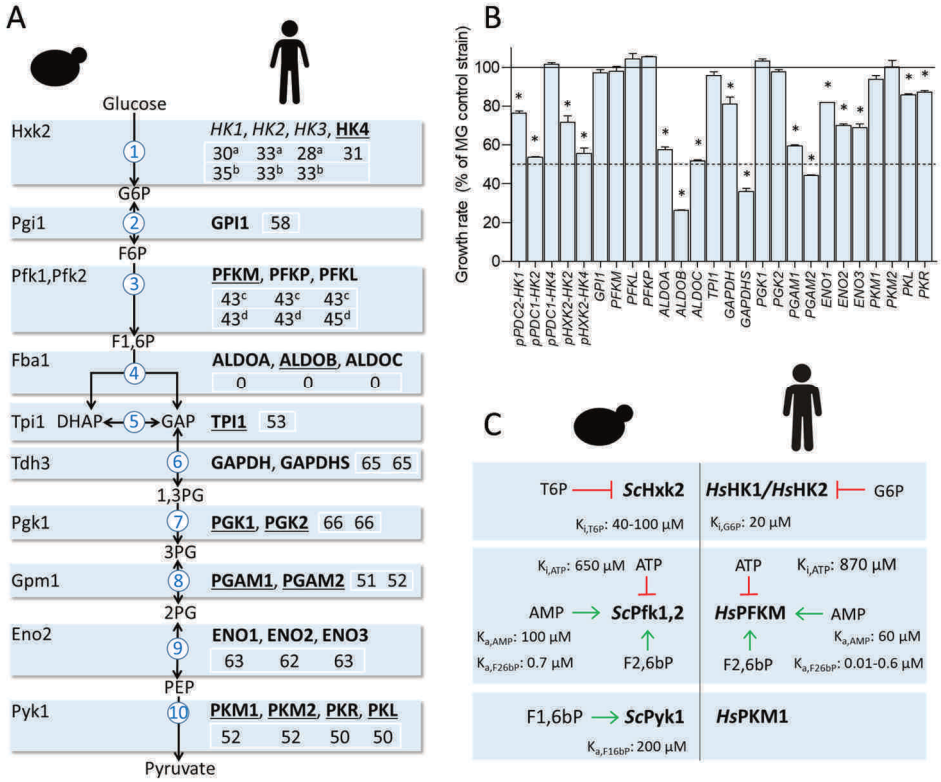


Figure 1 - Glycolytic human and yeast enzymes relevant for this study and single complementation assays.
A) Major glycolytic isoenzymes in yeast (left) and human enzymes used in this study (right). Underlined human enzymes were previously shown to complement their yeast ortholog. Bold enzymes complemented their yeast counterpart in this study. Percentage identity at protein level is shown of the human enzymes as compared to their yeast orthologs: [a, b], 1st subunit and 2nd of human hexokinases vs Hxk2, respectively, [c,d] human phosphofructokinases vs ScPfk1 and ScPfk2, respectively. Also see Table S1. **B)** Specific growth rates of the single gene complementation strains grown in SM glucose shown as percentage of the MG control strain IMX372, see also Fig S2. *HsHK2* and *HsHK4* were expressed with the *PDC1* promoter or the *HXK2* promotor, leading to different growth rates. Average and standard deviation of at least three independent replicates. * p-values between complementation and control strain below 0.01 (Student t-test, two-tailed, homoscedastic). **C)** Main allosteric regulators of the glycolytic yeast and human kinases and their regulation constants [32-37].

Human *Hs*HK1 and *Hs*HK2 can only complement the yeast hexokinases upon mutation

Upon transformation, strains expressing the human *HsHK1* or *HsHK2* as sole hexokinase grew well on galactose, a carbon source phosphorylated by galactokinase that does not require hexokinase activity, while exposure to glucose led to 1-2 days lag phase. Strains solely cultured on galactose displayed native *HsHK1* and *HsHK2* sequences, whereas exposure to glucose led to the systematic occurrence of single mutations in these genes, leading to an amino acid substitution

or deletion (Fig. 2A-C). *Hs*HK1 and *Hs*HK2 of strains solely exposed to galactose were active *in vitro* (IMX1689 and IMX2419 for *Hs*HK1 and *Hs*HK2 respectively), revealing that impaired growth on glucose was most likely not caused by lack of functionality of the human hexokinases in yeast (Fig. 2D). Considering that native and mutated alleles of human hexokinases (strains IMS1137 and IMX1690) had similar catalytic activities *in vitro* (Fig. 2D), growth defects upon exposure to glucose might result from inhibition of native human hexokinases in the yeast context. The observed mutations could then alleviate this inhibition to enable hexokinase activity *in vivo*. Mutations were observed in different regions of the *HsHK2* sequence in different strains, while in three separate *HsHK1* mutants the mutations were reproducibly localized at the glucose-6-phosphate binding site (Fig. 2B and C). The activity of both human hexokinases is sensitive to substrate concentration [38, 39], but is also allosterically inhibited by the product of the reaction, glucose-6-phosphate (G6P) [39, 40]. The elevated intracellular G6P concentrations reported for yeast (0.5-2 mM) are well above the $K_{i,G6P}$ of *Hs*HK1 and *Hs*HK2 (0.02 mM) and might inhibit these enzymes when expressed in yeast [33, 41].

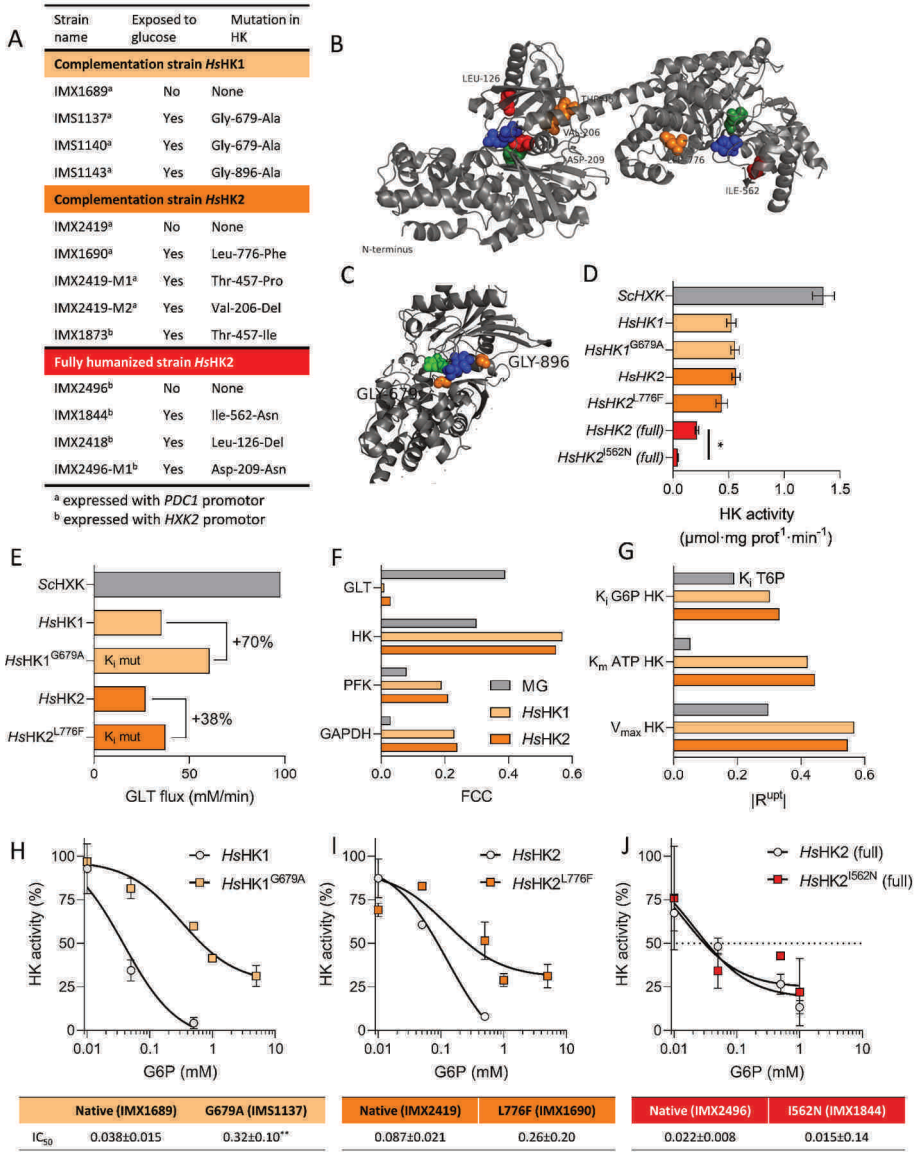
A computational 'hexokinase complementation model' built to address this phenomenon (Appendix 1) predicted a functional glycolytic pathway, able to reach a stable flux for both *Hs*HK1 and *Hs*HK2 with glucose as carbon source (Fig. 2E), but with a remarkable shift in the control of the glycolytic flux from glucose import (as predicted with yeast hexokinases) to hexokinase, as demonstrated by the increased flux control coefficient (FCC) (Fig. 2F). The glycolytic flux was more specifically predicted to be sensitive to the magnitude of the $K_{i,G6P}$, $K_{m,ATP}$ and V_{max} of both hexokinases (Fig 2G). In agreement with this prediction, the tested hexokinase variants of both *Hs*HK1 and *Hs*HK2 (*Hs*HK1^{G679A} from IMS1137 and *Hs*HK2^{L776F}) were less sensitive to G6P inhibition than the native alleles (Fig. 2H-I). For the *Hs*HK1^{G679A} mutation, our results are in direct contradiction to a previous study, where this mutation was found to have no impact on inhibition by the glucose-6P analog 1,5-anhydroglucitol-6P in purified *Hs*HK1 expressed in *E.coli* [38], which could be due to the different host organism. The $K_{i,G6P}$ of the humanized computational glycolytic model was modified using our experimental data (see Fig. 2H-I) to mimic the response of the mutated *Hs*HK1^{G679A} and *Hs*HK2^{L776F} variants, resulting in a predicted increase of the glycolytic flux of 70% for *Hs*HK1^{G679A} and 38% for *Hs*HK2^{L776F}, as compared to their native variants (Fig. 2E). Conversely, the sensitivity of the native and L776F variants to ATP, ADP and glucose measured *in vitro* were similar (Fig. S4) and trehalose-6-phosphate, a major inhibitor of yeast hexokinase not present in human cells, only mildly affected human hexokinase activity *in vitro* and the predicted *in silico* glycolytic flux (Fig S4, Appendix 1) [32]. Glucose-6-phosphate inhibition is therefore most likely the main mechanism

underlying the inability of native *Hs*HK1 and *Hs*HK2 to complement their yeast ortholog.

A

Strain name	Exposed to glucose	Mutation in HK
Complementation strain *Hs*HK1		
IMX1689[a]	No	None
IMS1137[a]	Yes	Gly-679-Ala
IMS1140[a]	Yes	Gly-679-Ala
IMS1143[a]	Yes	Gly-896-Ala
Complementation strain *Hs*HK2		
IMX2419[a]	No	None
IMX1690[a]	Yes	Leu-776-Phe
IMX2419-M1[a]	Yes	Thr-457-Pro
IMX2419-M2[a]	Yes	Val-206-Del
IMX1873[b]	Yes	Thr-457-Ile
Fully humanized strain *Hs*HK2		
IMX2496[b]	No	None
IMX1844[b]	Yes	Ile-562-Asn
IMX2418[b]	Yes	Leu-126-Del
IMX2496-M1[b]	Yes	Asp-209-Asn

[a] expressed with *PDC1* promotor
[b] expressed with *HXK2* promotor



	Native (IMX1689)	G679A (IMS1137)
IC_{50}	0.038±0.015	0.32±0.10**

Native (IMX2419)	L776F (IMX1690)
0.087±0.021	0.26±0.20

Native (IMX2496)	I562N (IMX1844)
0.022±0.008	0.015±0.14

Figure 2 - Characterization of human hexokinase mutants.
A) Mutations in human hexokinases after growth on glucose. **B) and C)** Localization of the amino acid substitutions in *Hs*HK2 and *Hs*HK1 variants, respectively. Colour coding as in panel A. Green, glucose binding site in the catalytic domain and blue, glucose-6P allosteric binding site. *Hs*HK2 crystal structure from [38] and *Hs*HK1 from [42]. **D)** Hexokinase *in vitro* activity assay from *S. cerevisiae* strains grown on galactose. *Sc*Hxk2 control activity was measured using strain IMX2015. Non-mutated *Hs*HK1 and *Hs*HK2 were assayed using strains IMX1689, IMX2419, and IMX2496 (*Hs*Gly-HK2), that were never exposed to glucose. Activities of the mutated variants were measured in extracts obtained from IMS1137, IMX1690 and IMX1844. In the complementation strains and the control *Sc*Hxk strain, hexokinase was expressed with the strong *ScPDC1* promoter while in the fully humanized strains the *ScHXK2* promotor was used. * significant change in activity as compared to unmutated enzymes ($p<0.01$, n=2, unpaired t-test). **E)** Simulated glucose uptake rate (GLT, glucose transporter) for the MG control, native *Hs*HK1 and *Hs*HK2 complementation strains. The effect of the observed change in K_i,G6P on the flux is modelled for each enzyme. **F)** Flux Control Coefficients (FCC's) of the four enzymes with the highest control over the flux, GLT, hexokinase (HK), phosphofructokinase (PFK) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH). **G)** Absolute values of the response coefficient $|R^{upt}|$ of the three parameters with highest control over the steady-state glucose uptake flux. **H)-I)** *In vitro* activity of the native and mutated hexokinase variants at various concentrations of the competitive inhibitor glucose-6-phosphate, expressed as percentage of activity without inhibitor. IC_{50}: half maximal inhibitory concentration of G6P (mM). ** significant difference of the mutated as compared to the native variant ($p<0.01$, extra-sum-of-squares F test, n=2).

Successful humanization of the entire glycolytic pathway in yeast

The successful complementation of individual glycolytic enzymes suggested that transplantation of a complete human glycolytic pathway might be possible. The success of humanization at full pathway level would depend on a combination of expression, kinetic and moonlighting properties of the whole set of enzymes (Fig. 1C) [43, 44], which together may have a severe impact on the growth of the yeast host. The muscle glycolytic pathway, characterized by fast *in vivo* rates, was chosen for transplantation in yeast [45-47]. Based on public databases, the most expressed isoenzymes in muscle tissue were selected. While *Hs*HK1 and *Hs*HK2 are the most abundant enzymes in muscle tissue [48, 49], *Hs*HK4 was initially chosen as hexokinase due to its ability to readily complement *Sc*Hxk2 (Fig. 1B). Despite the systematic mutations found in single complementation strains, *Hs*HK2 was also transplanted in a second glycolysis version, to test whether mutations were also required for its functionality in a human glycolytic context. *Hs*HK2 was chosen over *Hs*HK1 because it is usually annotated as the main muscle isoenzyme, while *Hs*HK1 is usually annotated as the brain hexokinase. Using both *Hs*HK4 and *Hs*HK2 was also interesting considering their difference in sequence and in kinetic and regulatory properties, *Hs*HK4 having a substantially lower affinity for glucose and being insensitive to glucose-6P [30, 50, 51]. Transplantation in a SwYG strain [28] of the entire set of human glycolytic genes resulted in the *Hs*Gly-HK2 strain with *Hs*HK2 as hexokinase and the *Hs*Gly-HK4 strain with *Hs*HK4 (Fig. 3A). Expression of the human genes was driven by strong, constitutive yeast promoters (Table S2). Note

that expression of *HsHK2* and *HsHK4* using the yeast *ScHXK2* promoter in these fully humanized strains, led to a slower growth rate than complementation using the stronger *ScPDC1* promoter (Fig. 1B and Fig S2). Pathway transplantation was successful as both the *Hs*Gly-HK2 and *Hs*Gly-HK4 strains displayed remarkably fast growth (ca. 0.15 h^{-1}, around 40% of the control SwYG strain with native glycolysis *Sc*Gly (IMX1821)) in minimal medium with glucose as sole carbon source (Fig. 3B). Exposure to glucose of the *Hs*Gly-HK2 strain led to long lag phase, and sequencing of *HsHK2* from several culture isolates revealed the systematic presence of single mutations in the vicinity of the catalytic and glucose-6P binding sites (Fig. 2A, Table S3). Remarkably a *ca.* 5-fold decrease in hexokinase V_{max} was observed in an isolate carrying the *Hs*Gly-HK2^{I562N} variant, while no evidence was found for other changes in kinetic parameters (Fig 2D and J and Fig S5). In another isolate, the *Hs*HK2^{D209N} mutation affected an amino acid key to the activity of the C-terminal active site [39]. The decrease in V_{max} but maintenance of glucose-6P sensitivity of *Hs*HK2^{I562N} was in stark contrast with the stable V_{max} but decreased glucose-6P sensitivity observed for *Hs*HK2 variants in single complementation strains (Fig. 2H-J). This suggested that the humanized glycolytic context might result in a different intracellular environment (particularly metabolite concentrations) and thereby lead to different requirements for hexokinase functionality and different evolutionary solutions.

Next to hexokinase, human and yeast pyruvate kinases also differ in allosteric regulations, *Hs*PKM1 being insensitive to the feed-forward activation by F1,6bP characteristic of *Sc*Pyk1 (Fig. 1C, Fig. S6). Despite the proposed role of this allosteric regulation for yeast cellular adaption to transitions [44, 52], the ability of the humanized yeast strains was not visibly impaired during transition between alternative (galactose) and favourite (glucose) carbon source (Fig. S7).

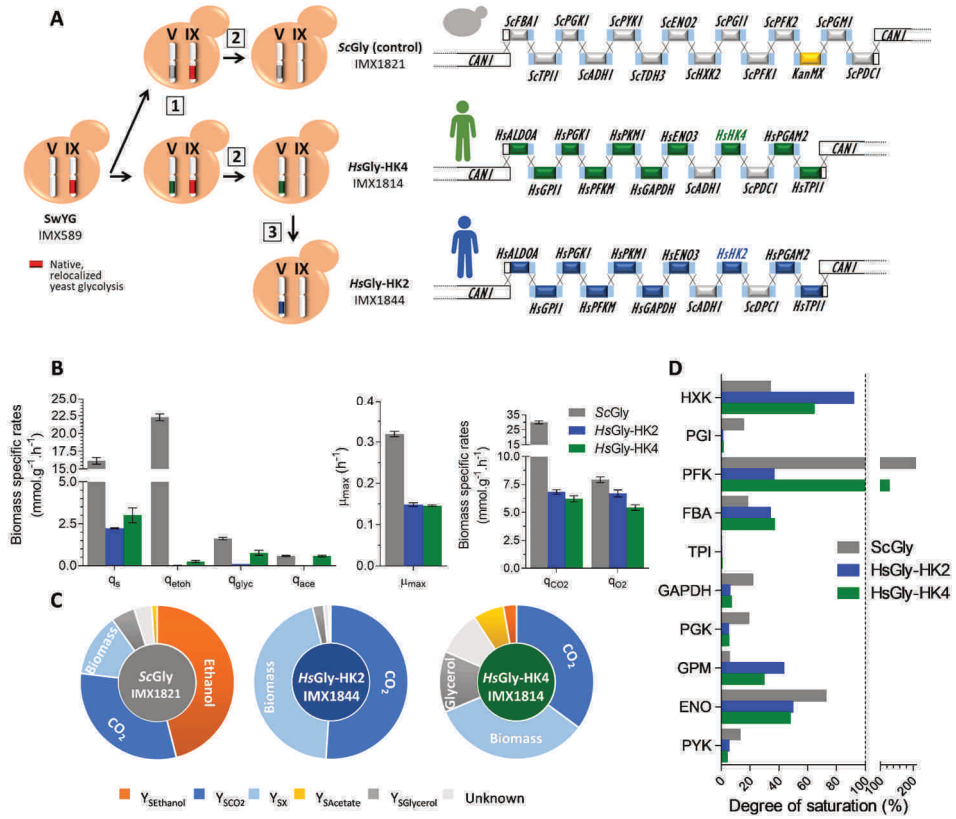


Figure 3 - Construction and physiological characterization of strains with fully humanized glycolysis.

A) Strain construction strategy and glycolytic pathway strains with native, and humanized co-localized glycolysis. **B)**, **C)** and **D)** Physiological characterization of strains shown in panel A in duplicate bioreactors on SM glucose. **B)** Specific rates for glucose and oxygen uptake (q_{glu} and q_{O2}), and for ethanol (q_{etoh}), glycerol (q_{glyc}), acetate (q_{ace}), CO_2 (q_{CO2}) and biomass (μ_{max}) production. Average and SEM of biological duplicates. **C)** Yields on glucose (CMol/CMol) of ethanol, CO2, biomass, acetate and glycerol are indicated ($Y_{SEthanol}$, Y_{SCO2}, Y_{SX}, $Y_{SAcetate}$, $Y_{SGlycerol}$ respectively). **D)** Estimation of the degree of saturation of glycolytic enzymes based on *in vitro* assays from cell extracts (activities reported in Fig. S8). *in vivo* fluxes were approximated from the q_{glu}. The dashed line indicates 100% saturation.

S. cerevisiae favors a mixed respiro-fermentative metabolism when glucose is present in excess (see IMX1821 in Fig. 3B and C), a phenomenon known as the Crabtree effect, analogous to the Warburg effect in mammalian cells [53]. *Hs*Gly-HK2 mostly respired glucose, with only traces of ethanol and glycerol being produced, while *Hs*Gly-HK4 displayed a respiro-fermentative metabolism, more similar to that of the *Sc*Gly control strain, although with far lower substrate uptake and ethanol production rates (Fig. 3B and C and Table S4). The fact that these physiological responses were similar to those observed for the respective *Hs*HK2 and *Hs*HK4 single complementation strains suggested that the human hexokinases strongly contributed to this switch between fermentative and respiratory metabolism, but they might not be the only players (Fig. 4A). Excepted *Hs*GPI1, the activity of the human enzymes was two to fifty times lower than the activity of their yeast ortholog (Fig. S8). With the notable exception of phosphofructokinase, sensitive *in vivo* to many effectors, the yeast glycolytic enzymes generally operate at overcapacity ([54-56] and Fig. 3D). In the humanized yeast strains hexokinase, aldolase and phosphoglycerate mutase showed higher degrees of saturation compared to the control strain with the native yeast glycolysis, suggesting that the activity of these enzymes could exert higher control on the glycolytic flux in the humanized strains (Fig. 3D). In line with this hypothesis, these three enzymes also led to low growth rates in single complementation strains (Fig. 1B). Remarkably, the activity of *Hs*PFKM was 2.6-fold lower in *Hs*Gly-HK4 than in *Hs*Gly-HK2, while the same protein abundance was found (Fig. S8). Consequently, while *Hs*PFKM *in vivo* operated above its *in vitro* capacity in *Hs*Gly-HK4, similarly to what is typically observed in *S. cerevisiae* and in IMX1821, the flux through *Hs*PFKM in *Hs*Gly-HK2 was only at ca. 30% of its *in vitro* capacity (Fig. 3D).

Overall the transplantation of a complete human glycolytic pathway to yeast was successful despite the structural, kinetic and regulatory differences between yeast and human enzymes. Global proteomics revealed increases in protein abundance of mainly metabolic enzymes, corresponding to the altered physiology (Fig. S9). The fully humanized glycolysis strains grew remarkably well, where the reduced glycolytic flux and growth rate as compared to yeast strains with a native glycolysis (μ_{max} ca. 60% slower), were in agreement with the lower *in vitro* enzymatic capacity of the human glycolytic enzymes.

Complementation of moonlighting functions

Many eukaryotic glycolytic enzymes, next to their glycolytic functions, have other cellular activities. These moonlighting functions might not be conserved across species [26] and failure of the human orthologs to complement the yeast moonlighting activities could strongly affect the humanized yeast strains. Three

glycolytic enzymes in *S. cerevisiae* have documented moonlighting functions: hexokinase, aldolase and enolase.

*Sc*Hxk2 is involved in glucose repression and the Crabtree effect by partially localizing to the nucleus in the presence of excess glucose where it represses the expression of genes involved in respiration and the utilization of alternative carbon sources such as the sucrose hydrolysing invertase *SUC2* [57, 58]. Accordingly, invertase activity is not detected in *S. cerevisiae* cultures with excess glucose, while it is expressed and active when glucose repression is alleviated (Fig. 4A and B). Conversely, double deletion of *ScHXK1* and *ScHXK2* alleviates glucose repression and enables invertase expression and activity in the presence of excess glucose condition. Invertase assays suggested that *Hs*HK4 but not *Hs*HK2^{L776F} was able to complement the role in glucose repression of *Sc*Hxk2 (Fig. 4A and B). However, since invertase repression is known to be sensitive to growth rate and the *Hs*HK2-Gly strains grew slowly, our findings do not completely rule out the possibility that *Hs*HK2 plays a role in glucose repression, although based on the difference in sequence, size and structure with yeast hexokinase, this seems unlikely. The role of *Hs*HK4 in glucose repression, suggested in an earlier report [24], is in line with the Crabtree effect we observed for the complementation and fully humanized strains carrying *Hs*HK4 despite their slow growth rate (Fig 3B and C, [59-61]).

Yeast aldolase is involved in assembly of vacuolar proton-translocating ATPases (V-ATPases), leading to the inability of aldolase deficient strains to grow at alkaline pH [62]. This function has been reported to be highly conserved between the yeast Fba1 and the human *Hs*ALDOB despite the absence of any sequence homology between the two proteins [23, 62]. The present study shows that the other human aldolases (*Hs*ALDOA and *Hs*ALDOC), which share ca. 70% identity with *Hs*ALDOB, can complement the moonlighting functions of *Sc*Fba1. All three human aldolase complementation strains as well as the fully humanized strains *Hs*Gly-HK2 and *Hs*Gly-HK4 showed no growth defects at pH 7.5 (Fig. 4C and Fig. S10), indicating that the vacuolar function was complemented by all three human aldolases.

Furthermore, the yeast enolase plays a role in mitochondrial import of a tRNA^{Lys}, a mechanism important at growth temperature above 37˚C, particularly on non-fermentable carbon sources [63, 64]. This mechanism seems well conserved in mammals, as yeast tRK1 is imported *in vitro* and *in vivo* in human mitochondria, in an enolase-dependent manner [65-67]. All three human enolase complementation strains show only minor growth defects at 37˚C on both glucose and non-fermentable carbon sources compared to the MG control strain (Fig. 4D). The fully humanized strains similarly show no growth defect at 37˚C (Fig. S10). This suggests that, in addition to the vacuolar function of *Sc*Eno2, the human enolase enzymes are also

able to fully take over its role in mitochondrial import of tRK1. Additionally the yeast enolases *Sc*Eno1 and *Sc*Eno2, are involved in vacuolar fusion and protein transport to the vacuole [68]. Whether the human enolases can take over this function in yeast is unknown. While enolase-deficient yeast strains display a fragmented vacuole phenotype and growth defects, this phenotype was not observed for the MG strain expressing *Sc*Eno2 only ([27], (Fig. 4E) and was also not observed for complementation strains expressing any of the three human enolases (Fig. 4E). This vacuolar moonlighting function seems therefore to be conserved between yeast and human enolases.

With the exception of *Hs*HK2, no phenotypic defects previously associated with an absence of glycolytic moonlighting activities could be observed in the humanized strains, suggesting that, in the tested conditions, the moonlighting functions of all glycolytic enzymes could be sufficiently complemented by their human orthologs.

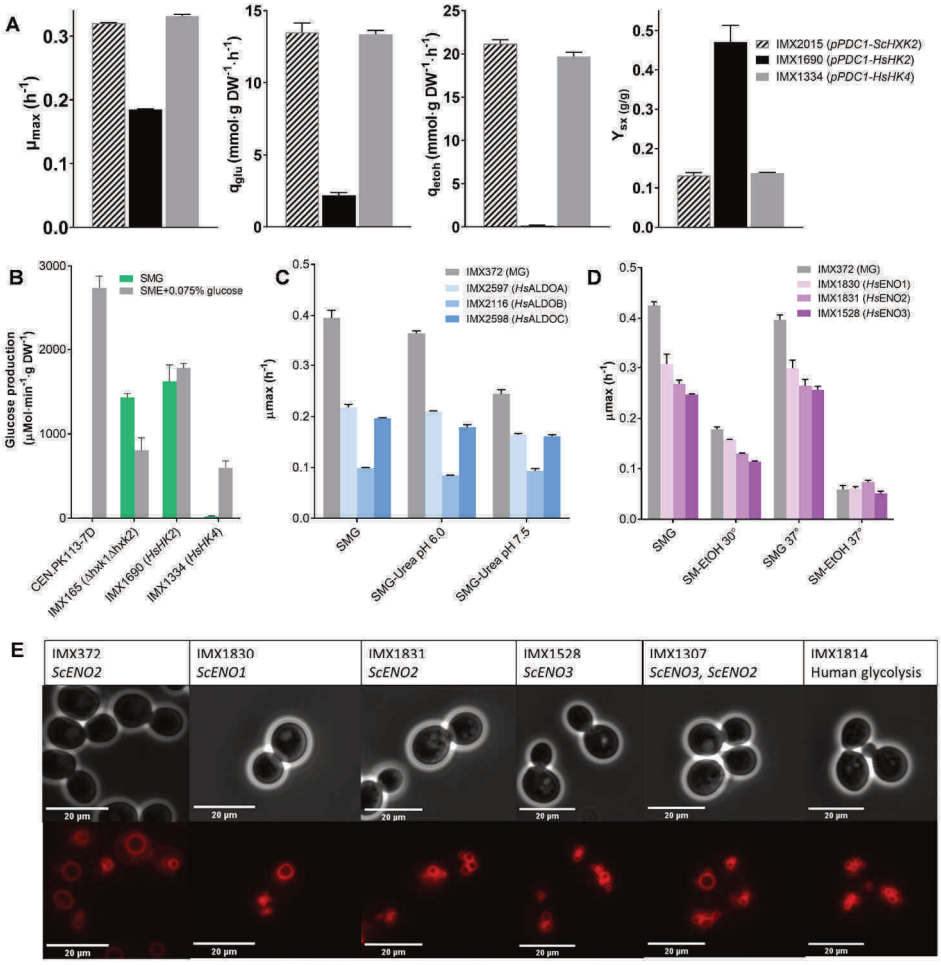


Figure 4 - Complementation of moonlighting functions.
A) Specific growth (μ_{max}), glucose consumption (q_{glu}) and ethanol production (q_{Eth}) rates, and biomass yield (Y_{SX}) of single hexokinase complementation strains grown in shake flask in SM glucose. Average and SEM of two biological replicates. **B)** Extracellular invertase activity of cultures with SM glucose (repressing condition) or SM ethanol + 0.075% glucose (inducing condition). Average and SEM of two biological replicates. **C)** and **D)** Specific growth rate of strains with single complementation of the three human aldolases and the three human enolases and of the MG control strain at different pH or at 30°C and 37 °C, with glucose or ethanol as carbon source as indicated. Average and standard deviation of biological triplicates. SM medium was used, but ammonium was replaced by urea to maintain pH in panel C). **E)** Staining of membranes with FM4-64 in *S. cerevisiae* strains expressing the yeast *Sc*Eno2 (IMX372 control) or the human *Hs*ENO3, *Hs*ENO2 and *Hs*ENO3 (IMX1830, IMX1831, IMX1528) as single complementation or as fully humanized glycolysis (*Hs*ENO3, IMX1814). IMX1307 carries one copy of the human *HsENO3* and the yeast *ScENO2* gene.

Engineering and evolutionary approaches to accelerate the slow growth of humanized glycolysis strains

Several enzymes (*Hs*HK2, *Hs*HK4, *Hs*ALDOA and *Hs*PGAM2) showed a significantly higher degree of saturation in the humanized strains (two to six-fold higher as compared to *Sc*Gly, Fig. 3D). The corresponding complementation strains also grew slower than the control strain, suggesting that the capacity of these enzymes might be limiting the glycolytic flux. Indeed simultaneous overexpression of *HsHK2*, *HsALDOA* and *HsPGAM2* in *Hs*Gly-HK2, and of *HsHK4*, *HsALDOA* and *HsPGAM2* in *Hs*Gly-HK4 successfully increased their specific growth rate by 63% and 48% respectively (Fig. 5A). These optimized, humanized yeast strains still grew 30% to 40% slower than the control strain with native, minimized yeast glycolysis (Fig. 5A). Growth at 37˚C, optimal temperature for human enzymes, instead of 30˚C did not improve the growth rate of the humanized yeast strains (Fig. 5A, also Fig. S10).

Many other mechanisms could explain the slow growth phenotype of the humanized strains, such as (allosteric) inhibition of the enzymes *in vivo*, incompatibility of substrate and co-factors concentrations with enzyme kinetic requirements, or deleterious effects of moonlighting activities of the human orthologs, more than could reasonably be tested by design-build-test-learn approaches. An adaptive laboratory evolution (ALE) strategy, particularly powerful to elucidate complex phenotypes [69], was therefore used to improve the fitness of the humanized strains. After approximately 630 generations in glucose medium, evolved populations of humanized yeast strains grew ca. twofold faster than their *Hs*Gly-HK2 and *Hs*Gly-HK4 ancestors (Fig. S11). Single colony isolates from six independent evolution lines, three per humanized yeast strain, confirmed the increased growth rate of the evolved humanized yeast strains (strains IMS0987 to IMS0993, Fig. 5A, Table S7). These strains evolved towards a higher glycolytic flux and a more fermentative metabolism, producing ethanol, albeit with a lower yield compared to the *Sc*Gly control (Fig. S12). This increase in fermentation was in line with the increased specific growth rate and glucose uptake rate [59]. These evolved strains were further characterized in an attempt to elucidate the molecular basis of the slow growth phenotype of the humanized yeast strains.

Exploring the causes of the slow growth phenotype of humanized glycolysis strains

The activity of several human glycolytic enzymes was affected by evolution (Fig. S13). Across the six evolution lines, the activity of both hexokinases (*Hs*HK2 and *Hs*HK4) and *Hs*PGAM2, for which activity was much lower than their yeast variants, was increased two to three-fold during evolution (Fig. 5B and C). These enzymes also led to the strongest decrease in growth rate upon complementation (Fig. 1B and

Fig. S14). Protein abundance of *Hs*HK4 and *Hs*PGAM2 was accordingly increased, albeit not with the same magnitude, but the change in *in vitro* activity of *Hs*HK2 was not reflected in protein abundance (Fig. 5C, Fig S13). The activity of *Hs*ALDOA, which also led to a large decrease in growth rate upon complementation and for which the activity was strongly reduced in the fully humanized strains, was not markedly altered by evolution. The response of *Hs*PFKM was particularly interesting. While its activity was already lower in the humanized yeast strains than in the *Sc*Gly strain, *Hs*PFKM was the only enzyme for which the activity was substantially decreased during evolution, by a factor of 2 to 8 as compared to their non-evolved humanized ancestors. For *Hs*HK2 and *Hs*PFKM, the changes in *in vitro* activity were not reflected in protein abundance. Global proteomics showed few proteins changed significantly in abundance after evolution (Fig. S15).

With the exception of phosphofructokinase, the genome sequence of the evolved strains offered little insight into the mechanisms leading directly to the above-mentioned alterations in *in vitro* specific activities or abundance of the glycolytic enzymes. However, interesting mutations were identified. The promoter, coding or terminator regions of the human glycolytic genes were exempt of mutations in the evolved strains. Only *Hs*PFKM carried a single mutation in its coding region in all three evolution lines of *Hs*Gly-HK4 (Fig. 5D, Table S5), one located in the N-terminal catalytic domain of the protein and the two others in the C-term regulatory domain where several allosteric effectors can bind (F2,6biP, ATP, ADP, citrate, etc. [70, 71]). The impact of these mutations cannot be inferred directly from the location, but they are most likely involved in the strong decrease in *in vitro* activity of PFKM in the strains evolved from *Hs*Gly-HK4. All three evolved strains from the *Hs*Gly-HK4 strain were also mutated in *TUP1* (general repressor of transcription with a role in glucose repression), with a conserved non-synonymous mutation resulting in an amino acid substitution (Fig. 5D, Table S5). Other transcription factors involved in the regulation of the activity of the yeast glycolytic promoters (Rap1, Abf1, Gcr1, Gcr2) did not harbour mutations in any strain. Overall few mutations were conserved between the evolution lines of the two humanized strains, but mutations in a single gene, *STT4*, were found in all six evolution lines (Fig 5D). Remarkably the six identified mutations were located within 164 amino acids, in the C-terminus of the protein harbouring its catalytic domain (Table S5). *STT4* encodes a phosphatidylinositol-4P (PI4P) kinase that catalyses the phosphorylation of PI4P into $PI4,5P_2$. As Stt4 is essential in yeast [72], the mutations present in the evolved strains could not cause a loss of function. Phosphoinositides are important signalling molecules in eukaryotes, involved in vacuole morphology and cytoskeleton organisation via actin remodelling [73]. PI4P kinases are conserved eukaryotic proteins [74], and Stt4 shares similarities with human PI3 kinases [75]. In mammals, activation of PI3K remodels actin, thereby

releasing aldolase A trapped in the actin cytoskeleton in an inactive state and increasing cellular aldolase activity [76, 77]. As yeast and human forms of actin are highly conserved (89% identity at protein level), a similar mechanism could be active in yeast and enable the evolved, humanized yeast strains to increase aldolase activity *in vivo* without increasing its concentration. Reverse engineering of two of the mutations found in the evolved strains IMS0990 and IMS0992 was performed in the non-evolved strain backgrounds with native yeast glycolysis and humanized glycolysis, by mutating the native *STT4* gene (Fig. 5A). The increases in specific growth rate did not match the growth rates of the evolved strains, suggesting other parallel mechanisms. Interestingly, in the reverse engineered strains *STT4* mutations resulted in a fragmented vacuole phenotype (Fig. S16), confirming that the mutations interfered with Stt4 activity and PI4P signalling. Such a phenotype was not observed in the evolved strains, however in these strains vacuoles also displayed abnormal morphologies with collapsed structures, indicating that specific mechanisms might have evolved in parallel to mitigate the effect of *STT4* mutations on vacuolar morphology (Fig. S16).

These findings suggest that evolution led to optimization of the human glycolytic pathway function in yeast through several mechanisms. Hexokinase 4 and phosphoglycerate mutase abundance and activity increased, allowing a higher glycolytic flux. For hexokinase 2 and phosphofructokinase, posttranslational mechanisms to modify enzyme activity must be present, counter-intuitively decreasing *Hs*PFKM activity *in vitro*. In all evolution lines, mutations in Stt4 occurred, which could potentially benefit *in-vivo* aldolase activity through modulation of actin structures. These adaptations reveal that the enzymes with the largest impact on growth rate in single complementation models (*Hs*HK2, *Hs*HK4, *Hs*PGAM2 and *Hs*ALDOA), and not those with the lowest activity, are the main targets for evolution. Changes in enzyme abundance, cellular environment and posttranslational modifications, and not direct mutations of the glycolytic genes, appear to be the most effective evolutionary strategy to improve flux of this heterologous pathway.

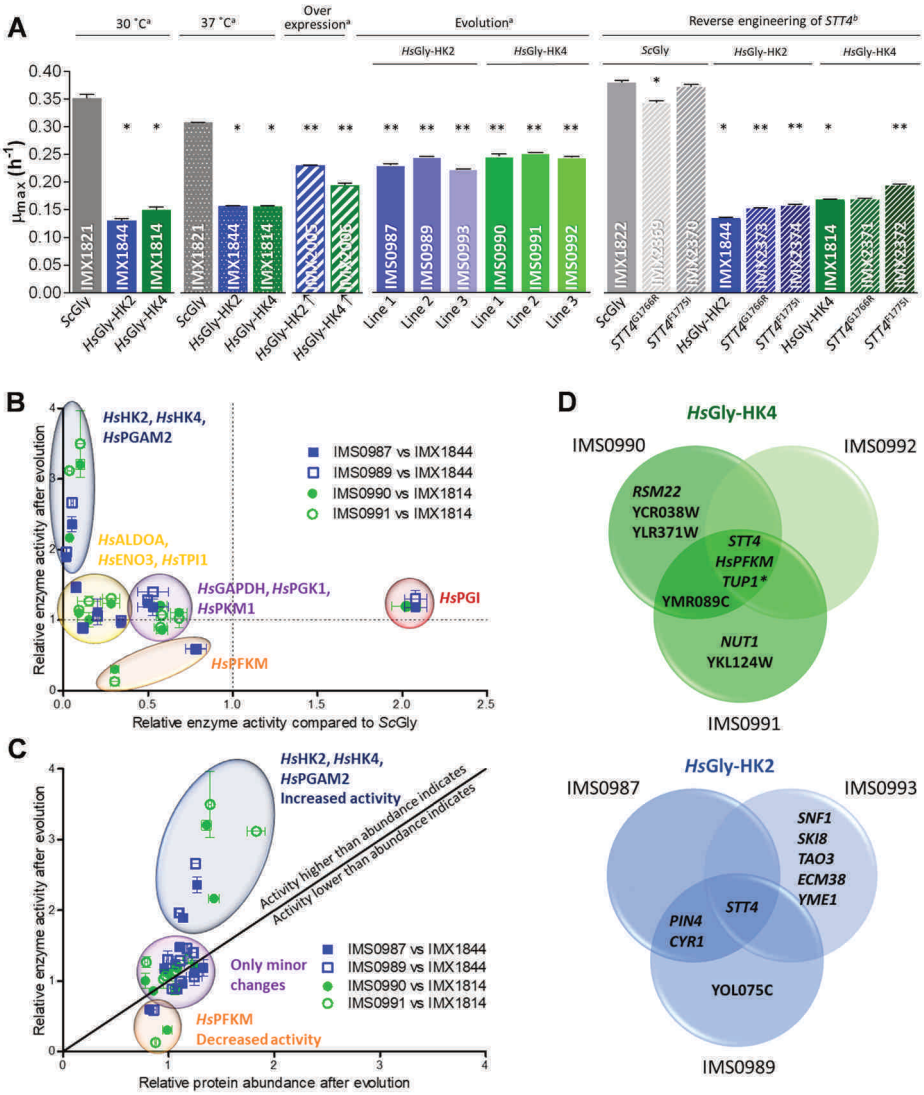


Figure 5 - Strategies to improve the growth rates of fully humanized yeast strains.
A) Specific growth rate of humanized, evolved and reverse-engineered strains in SM glucose. * indicates significant differences between *Hs*Gly-HK2 or *Hs*Gly-HK4 and the control strain IMX1821 and ** between the evolved and reverse-engineered strains and their respective parental strain *Hs*Gly-HK2 or *Hs*Gly-HK4 (n=2, student t-test, two-tailed, homoscedastic, p-value<0.05). [a] aerobic shake-flasks, [b] Growth Profiler. **B)** Comparison between the changes in glycolytic enzyme activity caused by humanization of yeast glycolysis and by evolution of the humanized strains. Activities available in Fig. S8 and S13. Error bars represent SEM of biological duplicates. Enzymes with a similar response are grouped. **C)** Comparison of changes in enzyme activity and in protein abundance caused by evolution of the humanized yeast strains. Error bars represent SEM of duplicate enzyme activity measurements and triplicate proteomics measurements. **D)** Mutations found in single colony isolates from independent evolution lines of humanized yeast strains.

Relevance of yeast as model for human glycolysis

The yeast intracellular environment could interfere with folding or posttranslational modifications of human enzymes and thereby alter their catalytic turnover number (k_{cat}). To explore this possibility, the k_{cat} of the human glycolytic proteins in yeast and in their native, human environment in myotube was experimentally determined. Enzyme activities (V_{max}) were measured in cell extracts using *in vivo*-like assay conditions. In these assays phosphofructokinase and hexokinase activities were too low for detection, although both could be detected at the protein level (Fig. S17). Overall the V_{max} of glycolytic enzymes were of the same order of magnitude in humanized yeast and muscle cells (Fig. 6A). *Hs*GPI1, *Hs*ALDOA, *Hs*GAPDH and *Hs*PGK1 activity was higher in yeast cells than in muscle cells, particularly for *Hs*GPI1 (seven-fold), while the activity of *Hs*PGAM2 was 5.5 fold lower in yeast compared to muscle cells (Fig. 6A). The differences in *in vitro* activity between yeast and human isoenzymes were mirrored in the peptide abundance for these proteins (Fig. S17), suggesting that the turnover rates of the human proteins expressed in human and yeast cells were not substantially different.

For *Hs*GPI, *Hs*ALDOA and *Hs*PGK1, the k_{cat} values were calculated by dividing the V_{max} values by the respective protein concentrations (Fig. S17C). This revealed no differences in the turnover rate between yeast and myotubes, irrespective of which of the standard peptides was used for protein quantification (Fig. 6B). For the remaining enzymes, calculation of the turnover rate was complicated by the presence of isoenzymes other than the canonical muscle glycolytic enzymes in myotube cultures. In addition to the canonical muscle isoenzymes, the isoenzymes *Hs*PFKL, *Hs*PGAM1, *Hs*ENO1 were present at equivalent or higher concentrations and *Hs*HK1, *Hs*PFKP and *Hs*PKM2 were present in low concentrations (Fig S17). This difference in isoenzyme abundance between tissue and isolated cell lines has been reported before for *in vitro* muscle cultures and, to a lower extent, for muscle biopsies [78]. Therefore, for enolase, phosphoglycerate mutase and pyruvate kinase, the apparent k_{cat} was assumed to be the V_{max} divided by the sum of all detected isoforms catalysing the specific reaction. The k_{cat} values of *Hs*ENO1 and *Hs*ENO3 are reported to be similar [79], but the *Hs*PGAM2 and *Hs*PKM1 have a higher k_{cat} than their respective isozymes [80, 81]. Taking these proportions into account, we found that apparent k_{cat} values for enolase and pyruvate kinase were similar between humanized yeast and myotubes while the k_{cat} of *Hs*PGAM was lower in humanized yeast (2.5 fold lower than the value in myotubes) (Fig. 6C). This may suggest that the yeast intracellular environment has a negative impact on posttranslational processing of the enzyme or the influence of the *Hs*PGAM2 isozyme in muscle cells. Taken together, these results demonstrate that out of the

six enzymes for which a turnover rate could be determined, five were not catalytically altered by the yeast environment, with *Hs*PGAM2 as potential exception.

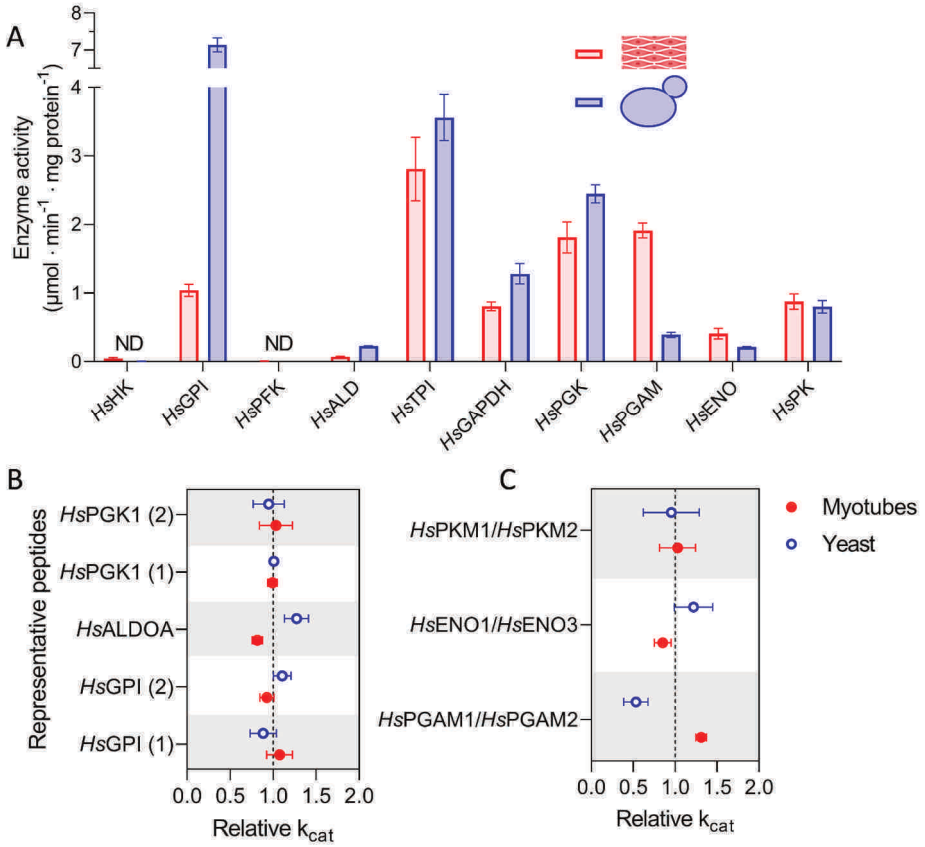


Figure 6 - Enzyme activity and k_{cat} of human glycolytic enzymes in human myotubes and humanized yeast.
A) Specific activity of human glycolytic enzymes measured *in vitro* with *in-vivo* like assaying conditions. Blue, yeast strain *Hs*Gly-HK2, red, muscle myotube cultures. **B)** and **C)** k_{cat} values for enzymes present as single isoform **B)** or multiple isoforms **C)** in the myotubes proteome data (Fig. S17). Data represent the mean values and standard deviation of three and two independent culture replicates for the myotubes and yeast cultures respectively. ND: not detected.

Discussion

The human glycolytic genes showed remarkable levels of complementation in yeast, both individually and as a complete pathway, with conservation of their secondary functions and turnover numbers similar to human muscle glycolysis. The combination of strains presented here can thus serve as new models to study fundamental aspects of human glycolysis in a simplified experimental setting, including moonlighting functions, the effects of PTMs and allosteric regulators, and cross-talk between enzymes. The extensive genetic accessibility and tractability of yeast enables the fast construction and testing of libraries of humanized yeast strains that carry different glycolytic designs.

The 100 kDa human hexokinases 1 to 3 did not show immediate complementation, however for *Hs*HK1 and *Hs*HK2, single amino acid substitutions were sufficient to restore their functionality in the yeast cellular environment. The requirement for mutations illustrates that these hexokinases have evolved to function in a particular metabolic niche. The ease of complementation with the human glycolytic genes is remarkable since glycolytic enzymes are known to be involved in numerous different moonlighting functions in yeast and human cells. In line with earlier work [24], we found that *Hs*HK4 but not *Hs*HK2 is able to complement the yeast Hxk2 function in invertase repression. The ability of *Hs*HK4 to transduce glucose signalling in yeast is surprising since this enzyme is not reported to have a transcriptional regulation function in human cells. *Hs*HK4 also lacks the decapeptide required for the translocation of *Sc*Hxk2 to the nucleus and its binding to the Mig1 transcription factor, although it has previously been shown to localize to the yeast nucleus [58, 82, 83]. The secondary functions of yeast aldolase and enolase in vacuolar ATPase assembly, vacuolar fusion and transport, and mitochondrial tRNA import were complemented by all human aldolase and enolase isozymes. This extraordinary conservation of glycolytic moonlighting functions observed between human and yeast glycolytic enzymes challenges our understanding of the underlying molecular mechanisms and reveals evolutionarily conserved functions.

The successful humanization of the entire glycolytic pathway in yeast and the availability of a library of strains with single complementation offer a unique opportunity to study potential synergetic effects between glycolytic enzymes and the impact of a full pathway on individual enzymes. A good example is the different evolutionary strategies found by fully humanized and single complementation strains to restore functionality of human hexokinases in a yeast context. All tested single complementation strains alleviated G6P inhibition on *Hs*Hk1 and *Hs*Hk2, while a fully humanized strain reduced *Hs*Hk2 activity without altering G6P sensitivity. G6P is a key metabolite at the branchpoint of several pathways in both yeast and human.

While the capacity in glycogen synthesis and hexokinase activity, as measured by their V_{max}'s, does not largely differ between yeast and skeletal muscle [37, 84-87], several yeast to muscle differences could account for the higher cellular G6P levels found in yeast [33, 41]. Critically, glucose uptake in yeast and muscle cells is very differently regulated. In the skeletal muscle, glycolytic fluxes are very dynamic and respond to the physiological status (such as meal status and exercise), a response largely controlled by glucose transport [88, 89] and ATP demand [90]. Even at their maximum capacity upon stimulation by insulin, glucose uptake is *ca*. two orders of magnitude lower than in yeast cells, and the muscle environment offers a much higher phosphorylation/uptake ratio than yeast. Other differences influencing the G6P concentration could be the presence in yeast of the trehalose cycle [91], and the greater capacity in yeast of glucose 6-phosphate dehydrogenase, first step of the pentose phosphate pathway, compared to human muscle [84, 92-94]. G6P has also been implicated in transcriptional regulation via the ChREBP and MondoA-MlX transcription factor complexes, which in turn modulate glycolytic gene expression in human cells [95, 96]. Altogether these factors contribute to yeast to skeletal cells differences in cellular G6P levels, resulting in supra-inhibitory levels for *Hs*HK1 and *Hs*HK2 in the yeast context. The different evolutionary strategy found in strains with fully humanized glycolysis might originate from different factors. The isomerization of G6P into F6P does most likely not account for large differences in G6P levels in fully humanized and single complementation strains as both human *Hs*PGI and native *Sc*Pgi1 operate near-equilibrium and have similar *in vitro* activity [97]. Conversely, the substantially lower activity of several human glycolytic enzymes as compared to their yeast equivalent (e.g. ALDOA and PGAM activities are ca. 10-fold) and resulting low glycolytic flux might alter the yeast cellular context (i.e. metabolite concentrations), and thereby the selection pressure exerted on hexokinase. Measuring intracellular metabolites in the different humanized strains should shed some light on the impact of partial and full humanization on the yeast cellular context. Another difference between human and yeast hexokinase 2 is the VDAC-dependent mitochondrial binding of the human variant, a binding not likely to be conserved in the humanized strains, unless the human VDAC protein is heterologously expressed [98, 99]. Humanization of glucose transport or mitochondrial VDAC proteins in yeast could be extremely useful to elucidate specific aspects of human hexokinases regulation and function in a human-like context.

A potential crosstalk between *Hs*PFKM and hexokinase was also revealed by comparing single gene and full pathway transplantation. *Hs*PFKM displayed a 2.5-fold higher *in vitro* activity in a strain expressing *Hs*HK2 as sole hexokinase as compared to a strain expressing *Hs*HK4, while protein abundance was identical. In the *Hs*Gly-HK4 strain the low *in vitro* activity of *Hs*PFKM activity does not match the

predicted *in vivo* activity based on the observed fluxes. This discrepancy between *in vitro* and *in vivo* was even stronger in evolved isolates of *Hs*Gly-HK4, in which *Hs*PFKM was systematically mutated. Conversely, no mutations were found in *Hs*PFKM in the evolved *Hs*Gly-HK2 strains. *Hs*PFKM is regulated at multiple levels (e.g. post-translational modification, binding to various cytoskeleton components, etc.) and most likely does not operate optimally in yeast [70, 100]. Notably, stabilization of *Hs*PFKM oligomerization promoted by calmodulin in human cells might be impaired in yeast considering the difference between yeast and human calmodulin (ca. 60% identity at the protein level) [101, 102]. The present results suggest the existence of a yet unknown, hexokinase-dependent mechanisms controlling *Hs*PFKM.

Both ALE and overexpression identified hexokinase and *Hs*PGAM2 as critical enzymes for glycolytic flux and growth rate improvement of the fully humanized strains. *Hs*PGAM2 lower activity and k_{cat} as compared to human myotube cells suggested that the yeast cellular environment is not favourable for this enzyme, a problem that both humanized strains solved by increasing *Hs*PGAM2 activity. The suboptimal activity of the hexokinases was similarly solved during evolution. However, for *Hs*PGAM2 and *Hs*HK2, as well as *Hs*PFKM, which decreased in activity during evolution, the changes in *in vitro* activity in evolved strains could not be fully explained by the changes in protein abundance and did not result from mutations in non-coding or coding regions of the corresponding genes. Modulation of enzyme activity through interactions with the cellular environment or direct posttranslational covalent modifications are most likely responsible for this discrepancy between protein level and *in vitro* enzyme activity. The regulation of several human glycolytic proteins occurs via interaction with the cytoskeleton, as mentioned above for *Hs*PFKM and calmodulin. In mammalian cells, phosphoinositide signalling via PI3-kinase regulates aldolase activity by actin remodelling. The systematic mutation in the evolved humanized strains of *STT4*, encoding a PI4-kinase involved in cellular signalling for many cellular processes, including actin organization in yeast, suggests that ALDOA activity might also be modulated in yeast by binding to actin and altered by phosphoinositide-mediated signalling. Overall optimization of human glycolysis in yeast seems to be largely exerted by posttranslational mechanisms, and ALE is a powerful strategy to identify the mechanisms causing suboptimal functionality in yeast.

This first successful humanization of the skeletal muscle glycolysis in yeast offers new possibilities to explore human glycolysis. Since many complex interactions with various organelles and signalling pathways that are present in human cells will be absent in yeast, such model strains can be applied to study the pathway in a ‘clean’ background. Transplantation to the yeast context enables to dissect metabolic from

signalling-related mechanisms in the control and regulation of glucose metabolism, mechanisms often debated in the field of diabetes and muscle insulin resistance. As an example, whether glycolytic enzymes themselves could be inhibited by intermediates of lipid metabolism in the muscle and consequently impact enzyme activity and glycolytic fluxes remains an open question to be tested [103, 104]. Beyond muscle tissue, the glycolysis swapping concept can be extended to any glycolytic configuration. Complete pathway transplantation can in the future be used to generate translational microbial models to study fundamental aspects of evolutionary conservation between species and tissues, and to unravel mechanisms of related diseases.

Limitations of this study

While this work demonstrates that pathway swapping enables fast and facile transplantation of an entire pathway, a general limitation to the proposed approach is the characterization of the physiological and cellular impact of the transplantation. Comprehensively capturing all the effects of such drastic genetic changes is very time- and resource-consuming, which has guided the decision made in the present manuscript to focus on a subset of human glycolytic isoenzymes for humanization, and on in-depth characterization of a few important findings such as the hexokinases. Of the many intriguing results uncovered by this study, many could not (yet) lead to mechanistic understanding and require further research. While adaptive laboratory evolution did bring new insight in the interaction between the human glycolytic isoenzymes and the yeast environment, many questions remain unanswered such as the mechanisms that cause the reduction in growth rate in humanized strains or that enable yeast cells to tune the abundance of human glycolytic proteins. Finally, an unfortunate technical limitation prevented the measurement of human hexokinase and phosphofructokinase activity in yeast and human cells using *in-vivo* like assay conditions. The reasons underlying this lack of *in vitro* activity of these two key enzymes have to be further explored to complete the comparison of glycolytic human enzymes functionality in yeast and human context.

Materials and Methods

Data and code availability

The whole genome DNA sequencing data generated in this project are accessible at NCBI under bioproject PRJNA717746 and are publicly available as of the date of publication.

Mass spectrometric raw data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE repository

and are publicly available as of the date of publication with the dataset identifier PXD025349.

Yeast and *E. coli* strains details

All yeast strains used and generated in this study are derived from the CEN.PK background [105] and are listed in table S7. Yeast strains were propagated at 30 ºC except otherwise indicated on YP medium containing 10g L^{-1} Bacto Yeast extract, 20 g L^{-1} Bacto Peptone or synthetic medium containing 5 g L^{-1} $(NH_4)_2SO_4$, 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4 \cdot 7 \cdot H_2O$, and 1 mL L^{-1} of a trace elements and vitamin solution [106]. Media were supplemented with 20 g L^{-1} glucose or galactose or 2% (v/v) ethanol. For the physiological characterization of the individual hexokinase complementation strains (Fig. 4A) and to test the aldolase moonlighting function (Fig. 4C and S10B), $(NH_4)_2SO_4$ was replaced with 6.6 g L^{-1} K_2SO_4 and 2.3 g L^{-1} urea to reduce acidification of the medium. Urea was filter sterilized and added after heat sterilization of the medium at 121°C. When indicated, 125 mg L^{-1} histidine was added. For solid media 2% (w/v) agar was added to the medium prior to heat sterilization. The pH of SM was adjusted to pH 6 by addition of 2 M KOH. For selection, YP medium was supplemented with 200 mg L^{-1} G418 (KanMX) or 100 mg L^{-1} nourseothricin (Clonat). For removal of the native yeast glycolysis cassette from the *sga1* locus the SM glucose (SMG) medium was supplemented with 2.3 g L^{-1} fluoracetamide to counter select for the *AmdS* marker present in the cassette [107]. Yeast strains were stored at -80 °C after addition of 30% (v/v) glycerol to an overnight grown culture.

For plasmid propagation chemically competent *Escherichia coli* XL1-Blue (Agilent Technologies, Santa Clara, CA) cells were used which were grown in lysogeny broth (LB) supplemented with 100 mg L^{-1} ampicillin, 25 mg L^{-1} chloramphenicol or 50 mg L^{-1} kanamycin when required [108, 109]. *E. coli* strains were stored at -80 °C after addition of 30% (v/v) glycerol to an overnight grown culture.

Human myotube culture details

Human myoblasts were obtained from orbicularis oculi muscle biopsies, which has previously been described [110]. Briefly, a muscle biopsy was obtained from a healthy 60-year-old female donor, undergoing blepharoplasty, from which myoblasts were isolated as previously described. After 8 passages in culture, a subclone V49 that expressed Pax7, MyoD and Myogenin was used for the assays here described. Cells were maintained in high glucose Dulbecco's Modified Eagle's Medium (DMEM, Sigma-Aldrich/Merck) in the presence of L-glutamine, 20% fetal bovine serum (FBS, Life Technologies Gibco/Merck) and 1% penicillin/streptomycin (p/s, Sigma-Aldrich/Merck). For differentiation, cells were seeded on 10 cm dishes covered with polydimethylsiloxane (PDMS) gradients at 5,000 cells/cm^2. After reaching

confluence, the medium was changed to DMEM in the presence of 2% FBS, 1% p/s, 1% Insulin-Transferrin-Selenium (Life Technologies Gibco/Merck) and 1% dexamethasone (Sigma-Aldrich/Merck). The presence of PDMS gradients allows cells to grow aligned, which in turn improves myotube maturity and functionality [111]. Cells were harvested after 5 days in differentiation medium. In short, cells were washed twice with ice-cold Dulbecco's Phosphate Buffered Saline (DPBS, Gibco) and scraped in DPBS in the presence of Complete Protease Inhibitor Cocktail (Merck, 11836145001, 1:25 v/v after resuspension according to manufacturer's guidelines). Cell extracts were frozen at -80 °C.

Growth rate determinations and laboratory evolution

For all growth experiments in shake flasks, 100 mL medium in a 500 mL shake flask was used except for the shake flask growth study with individual complementation strain for which 20 mL in a 100 mL volume shake flaks was used. Strains were incubated with constant shaking at 200 rpm and at 30°C unless stated otherwise. Strains were inoculated from glycerol stocks in YPD and grown overnight. This culture was used to inoculate the pre-culture (SMG) from which the exponentially growing cells were transferred to new shake flasks to start a growth study. Growth rates were determined from OD_{660} data by linear regression of the log-linear plots of OD_{660} versus time over at least six consecutive time-points in a range where the OD_{660} doubled twice, choosing the range to maximize the squared Pearson correlation coefficient (R^2).

Growth studies in microtiter plate were performed at 30 °C and 250 rpm using a Growth Profiler 960 (EnzyScreen BV, Heemstede, The Netherlands). Strains from glycerol freezer stocks were inoculated and grown overnight in 10 mL YPD or YPGal medium in a 50 mL volume shake flask. This culture was used to inoculate a preculture in a 24-wells plate with a 1 mL working volume (EnzyScreen, type CR1424f) or a shakeflask with 15 mL of the medium of interest, which was grown until mid/late-exponential growth. From this culture the growth study was started in a 96-wells microtiter plate (EnzyScreen, type CR1496dl), with final working volumes of 250 µL and starting OD_{660} of 0.1-0.2. Microtiter plates were closed with a sandwich cover (EnzyScreen, type CR1296). Images of cultures were made at 30 min intervals. Green-values for each well were corrected for position in the plate using measurements of a culture of OD_{660} 5.05 of CEN.PK113-7D. Corrected green values were converted to OD-values based on 15-point calibrations, fitted with the following equation: OD-equivalent = $a \times GV(t) + b \times GV(t)^c - d$ in which GV(t) is the corrected green-value measured in a well at time point 't'. This resulted in curves with a = 0.0843, b = 5.35×10^{-8}, c = 4.40 and d = 0.42 for the data in Fig. 1, 5 and S2 and a = 0.077, b = 1.66×10^{-7}, c = 3.62 and d = 1.61 for the data in Fig. 4 and S10. Growth

rates were calculated in a time frame where the calculated OD was between 2 and 10 in which OD doubled at least twice.

Adaptive laboratory evolution of IMX1814 and IMX1844 was performed in SMG at 30°C in 100 mL volume shake flasks with a working volume of 20 mL. Initially every 48 hours 200 µL of the culture was transferred to a new shake flask with fresh medium, after 22 transfers (approximately 170 generations) this was done every 24h. For both strains three evolution lines were run in parallel. At the end of the experiment single colony isolates were obtained by restreaking three times on YPD plates (Table S7D).

Molecular techniques, gene synthesis and Golden Gate plasmid construction
PCR amplification for cloning purposes was performed with Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, MA) according to the manufacturers recommendations except that the primer concentration was lowered to 0.2 µM. PCR products for cloning and Sanger sequencing were purified using the Zymoclean Gel DNA Recovery kit (Zymo Research, Irvine, CA) or the GeneJET PCR Purification kit (Thermo Fisher Scientific). Sanger sequencing was performed at Baseclear BV (Baseclear, Leiden, The Netherlands) and Macrogen (Macrogen Europe, Amsterdam, The Netherlands). Diagnostic PCR to confirm correct assembly, integration of the constructs and sequence verification by Sanger sequencing was done with DreamTaq mastermix (Thermo Fisher Scientific) according to the manufacturers recommendations. To obtain template DNA, cells of single colonies were suspended in 0.02 M NaOH, boiled for 5 min and spun down to use the supernatant. All primers used in this study are listed in Table S8. Primers for cloning purposes were ordered PAGE purified, the others desalted. To obtain gRNA and repair fragments the designed forward and reverse primers were incubated at 95°C for 5 min to obtain a double stranded piece of DNA. PCR products were separated in gels containing 1% agarose (Sigma) in Tris-acetate buffer (TAE). Genomic DNA from CEN.PK113-7D was extracted using the YeaStar™ Genomic DNA kit (Zymo Research Corporation, Irvine, CA, USA). Cloning of promoters, genes and terminators was done using Golden Gate assembly. Per reaction volume of 10 µl, 1 µl T4 buffer (Thermo Fisher Scientific), 0.5 µl T7 DNA ligase (NEB New England Biolabs, Ipswich, MA) and 0.5 µl BsaI (Eco31I) (Thermo Fisher Scientific) or BsmBI (NEB) was used and DNA parts were added in equimolar amounts of 20 fmol as previously described [112]. First a plasmid backbone was constructed from parts of the yeast toolkit [112] using a kanamycin marker, *URA3* marker, bacterial origin of replication, 3'and 5' *ura3* integration flanks and a *GFP* marker resulting in pGGKd002 (Table S9C). In a second assembly, the *GFP* gene in this plasmid was replaced by a transcriptional unit containing a *S. cerevisiae* promoter and terminator and a human glycolytic gene. The sequences of the human glycolytic genes were obtained

from the Uniprot database, codon optimized for *S. cerevisiae* and ordered from GeneArt Gene Synthesis (Thermo Fisher Scientific). Genes were synthetized flanked with BsaI restriction sites to use them directly in Golden Gate assembly (Table S9A). The *PKL* gene which is a shorter splicing variant of *PKR* was obtained by amplifying it from the *PKR* plasmid pGGKp024 using primers containing BsaI restriction site flanks (Table S8G). *S. cerevisiae* promoters and terminators were PCR amplified from genomic DNA using primers flanked with BsaI and BsmBI restriction sites (Table S8A) [113]. The resulting PCR product was directly used for Golden Gate assembly. For long term storage of the fragments, the promoters and terminators were cloned into the pUD565 entry vector using BsmBI Golden Gate cloning resulting in the plasmids pGGKp025-048 listed in Table S9B. For the *HXK2* and *TEF2* promoters and *HXK2* and *ENO2* terminators already existing plasmids were used (Table S9B). For the construction of pUDE750 which was used as PCR template for the amplification of the *HsHK4* fragment used in IMX1814, first a dropout vector (pGGKd003) was constructed from the yeast toolkit parts pYTK002, 47, 67, 74, 82 and 84 (Table S9C). In this backbone, *ScHXK2p, HK4* and *ScHXK2t* were assembled as described above (Table S9E). Plasmid isolation was done with the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, St. Louis, MO). Yeast transformations were performed according to the lithium acetate method [114].

Construction of individual gene complementation strains

To enable CRISPR/Cas9 mediated gene editing, *Cas9* and the *NatNT1* marker were integrated in the *SGA1* locus of the minimal glycolysis strain IMX370 by homologous recombination, resulting in strain IMX1076 [27]. *Cas9* was PCR amplified from p414-*TEF1p-Cas9-CYC1t* and *NatNT1* from pUG-*natNT1* (Tables S8G and S9E) and 750 ng of both fragments were, after gel purification, used for transformation.

For the individual gene complementation study, 400 ng of the constructed plasmids containing the human gene transcriptional units (Table S9D) were linearized by digestion with NotI (FastDigest, Thermo Fisher Scientific) according to the manufacturer's protocol for 30 min and subsequently the digestion mix was directly transformed to IMX1076. The linearized plasmids were integrated by homologous recombination in the disrupted *ura3-52* locus of strain IMX1076 and the transformants were plated on SMG. After confirmation of correct integration by PCR (Table S8B), in a second transformation the orthologous yeast gene (or genes, in case of *PFK1* and *PFK2*) was removed using CRISPR/Cas9 according to the protocol of Mans *et al.* [115]. Since only the yeast gene and not the human ortholog should be targeted, the gRNAs were designed manually (Table S8D). For deletion of *FBA1*, *GPM1*, and *PFK1* and *PFK2,* the plasmids containing the gRNA were preassembled as previously described [115] using Gibson assembly and a PCR amplified pROS13 backbone containing the KanMX marker (Tables S8D and S9F).

For *HXK2* deletion, the double stranded gRNA and a PCR amplified pMEL13 backbone were assembled using Gibson assembly (Table S8D). The constructed plasmids were verified by PCR. The rest of the gRNA plasmids for yeast gene deletion were assembled *in vivo* in yeast and were not stored as individual plasmid afterwards. For the *in vivo* assembly approach the strains were co-transformed with 100 ng of the PCR amplified backbone of pMEL13 (Table S8D and S9F), 300 ng of the double stranded gRNA of interest (Table S8D) and 1 µg repair fragment to repair the double stranded break (Table S8E). For pre-assembled plasmids, strains were co-transformed with 0.6-1 µg of plasmid (Table S9F) and 1 µg repair fragment (Table S8E). Transformants were plated on YPD + G418 and for the *HsHK1-HK3* strains on YPGal + G418. Successful gene deletion was confirmed with diagnostic PCR (Table S8B, Fig. S18). gRNA plasmids were afterwards removed by several restreaks on non-selective medium. To test if the complementation was successful, the strains were tested for growth in SMG. For *HsHK2,* three complementation strains were made. IMX1690 (*pScPDC1-HsHK2)* and IMX1873 (*pScHXK2-HsHK2*) which were grown on glucose medium and contain a mutation in *HsHK2* and IMX2419 (*pScPDC1-HsHK2)* which was never exposed to glucose and does not contain mutations. Similarly for *HsHK1,* complementation strain IMX1689 was not grown on glucose, after growth on glucose mutations occurred and strains IMS1137, IMS1140 and IMS1143 were stocked. *HsHK4* was also expressed both with the *ScHXK2* and *ScPDC1* promoter, resulting in IMX1874 and IMX1334 respectively (Table S7A,B). An overview of the workflow is provided in Fig. S1. To test for the occurrence of mutations, the human gene transcriptional units were PCR amplified using the primers listed in Table S8 and sent for Sanger sequencing.

Full human glycolysis strain construction

For the construction of the strains containing a full human glycolysis, the transcriptional units of the *HsHK2, HsHK4, HsGPI, HsPFKM, HsALDOA, HsTPI1, HsGAPDH, HsPGAM2, HsENO3,* and *HsPKM1* gene were PCR amplified from the same plasmids as were used for the individual gene complementation using primers with flanks containing synthetic homologous recombination (SHR) sequences (Table S8C and S9D). For the *HsHK2* and *HsHK4* gene for which pUDE750 and pUDI207 were used as template, which contain the Sc*HXK2* promoter and terminator. An overview of the promoters used for the human gene expression is provided in Table S2. The yeast *PDC1* and *ADH1* genes were amplified with their corresponding promoter and terminator regions from genomic DNA from CEN.PK113-7D (Table S7E). The fragments were gel purified and the fragments were assembled in the *CAN1* locus of strain IMX589 by *in vivo* assembly. 160 fmol per fragment and 1 µg of the pMEL13 plasmid targeting *CAN1* was used. Transformation mix was plated on YPD + G418 and correct assembly was checked by PCR and resulted in strain

IMX1658. In a second transformation, the cassette in the *SGA1* locus containing the native *S. cerevisiae* glycolytic genes and the *AmdS* marker was removed. To this end IMX1658 was transformed with 1 µg of the gRNA plasmid pUDE342 (Table S9F) and 2 µg repair fragment (counter select oligo) (Table S8E) and plated on SMG medium with fluoracetamide to counter select for the *AmdS* marker. From the resulting strain the pUDE342 plasmid was removed and it was stored as IMX1668. To replace the *HsHK4* gene with *HsHK2,* the *HsHK2* gene was PCR amplified from pGGKp002 using primers flanked with sequences homologous to the *ScHXK2* promoter and terminator to allow for recombination (Table S8G). IMX1668 was co-transformed with this fragment and pUDR387 containing the gRNA targeting the *HsHK4* gene and the cells were plated on YPD + G418 (Table S9F). After confirmation of correct integration by PCR and plasmid removal, the strain was stored as IMX1785. The pUDR387 gRNA plasmid was constructed with Gibson Assembly from a pMEL13 backbone and double stranded *HsHK4* gRNA fragment (Table S8D). To make the constructed yeast strains prototrophic, the *ScURA3* marker was PCR amplified from CEN.PK113-7D genomic DNA using primers with flanks homologous to the *TDH1* region (Table S8G) and integrated in the *tdh1* locus of IMX1785 and IMX1668, by transforming the strains with 500 ng of the fragment and plating on SMG. This resulted in IMX1844 and IMX1814 respectively. These strains were verified by whole genome sequencing (Table S3) and the ploidy was verified (Fig. S19). IMX2418 (*Hs*Gly *HK2* strain without mutation in Hs*HK2*) was constructed by transforming IMX1814 with pUDR387 and the *HsHK2* fragment amplified as described above. The cells were plated on YPGal + G418 and later restreaked on YPGal plates to remove the plasmid. For the overexpression of *HsALDOA, HsPGAM2* and *HsHK2/HsHK4* resulting in IMX2005 and IMX2006*,* the expression cassettes were PCR amplified from pUDI141, pUDI150, pUDI134 and pUDI136 respectively using primer sets 12446/12650, 12467/14542 and 14540/14541 (Table S8C, S9D). IMX1844 and IMX1814 were transformed with 160 fmol per fragment and 1 µg of the plasmid pUDR376 containing a gRNA targeting the X2 locus [116] and plated on SMG-acetamide plates. To obtain the reference strain IMX1821 which contains a yeast glycolysis cassette integrated in *CAN1,* the pUDE342 plasmid was removed from the previously described strain IMX605 [28] and the *URA3* fragment was integrated in *tdh1* in the manner as described above*.* An overview of strain construction is provided in Fig. S20.

STT4 reverse engineering

The single nucleotide polymorphisms (SNPs) which were found in the *STT4* gene of evolved strains IMS0990 and IMS0992 resulting in amino acid changes G1766R and F1775I respectively, were introduced in the *STT4* genes of the non-evolved strains IMX1814, IMX1844 and IMX1822 using CRISPR/Cas9 editing [115] (Table S7D).

Two gRNA plasmids pUDR666 and pUDR667 were constructed using Gibson Assembly of a backbone amplified from pMEL13 (Table S8D, S9F) and a gRNA fragment consisting of oligo 16748+16749 and 16755+16756 respectively (Table S8D). For introduction of the G1766R mutation, strains were transformed with 500 ng of pUDR666 and 1 µg of repair fragment (oligo 16750+16751) and for introduction of F1775I with 500 ng of pUDR667 and 1 µg of repair fragment (oligo 16757+16758) (Table S8E, S9F). Strains were plated on YPD + G418 and introduction of the mutation was verified by Sanger sequencing. The control strain IMX1822 containing the native yeast minimal glycolysis in the *SGA1* locus originates from strain IMX589 [28]. From this strain the *AmdS* marker was removed by transforming the strain with 1 µg repair fragment (oligo 11590+11591) and 300 ng of a gRNA fragment (oligo 11588+11589, Table S8D) targeting *AmdS* and 100 ng of backbone amplified from pMEL10 resulting in a *in vivo* assembled gRNA plasmid. After removal of the plasmid by restreaking on non-selective medium, this strain, IMX1769, was made prototrophic by integrating *ScURA3* in *tdh1* (Table S8G, S7E), resulting in IMX1822.

Visualization of hexokinase mutants and mathematical modelling

Sequencing of the *HsHK1* and *HsHK2* carrying strains showed the presence of mutations in all strains after growth on glucose. All found mutations were mapped unto the protein sequence and visualized on the structural model with PDB code 1HKB [42] for *Hs*HK1 and 2NZT [38] for *Hs*HK2 using the PyMOL Molecular Graphics System, version 1.8.6 (Schrödinger LLC).

Native human hexokinase complementation strains were simulated with the use of a previously published computational model of yeast glycolysis [44]. The SBML version of the model was downloaded from *jjj.bio.vu.nl/models/vanheerden1* and imported into COPASI (software version 4.23) [117]. All concentrations in the model are expressed as mM and time in minutes. Equilibrium constants (K_{eq}'s) were obtained from Equilibrator (V 2.2) [118] at pH 6.8 and ionic strength 360 mM [119]. The V_{max}'s from the MG strain from Kuijpers et al., 2016 [28] were initially incorporated to create a control strain. The forward V_{max}'s from GAPDH and PGI were calculated from the measured reverse V_{max}'s and model K_{eq}'s and K_m's according to the Haldane relationship. For the complementation strains, the kinetic equation of hexokinase was first adapted to include competitive terms from G6P and ADP inhibition and trehalose 6-phosphate inhibition of the human enzymes was disregarded based on our kinetic results. Mammal kinetic parameters were obtained from [48]. *Hs*HK1 and *Hs*HK2 glycolysis models were subsequently obtained by incorporating the hexokinase V_{max} measured from strains IMX1689 and IMX2419, respectively. Steady-state fluxes were calculated with the integration of ordinary differential equations. Flux control coefficients (FCC) and response coefficients (R) were calculated in COPASI under Metabolic Control Analysis and Sensitivities

according to equations 1 and 2 below, respectively. J_{ss} represents a steady-state flux, for which the glucose uptake flux was used. $V_{max\ i}$ represents the V_{max} while P_i stands for a kinetic parameter of an enzyme 'i' in the pathway. Summation theorem was verified for FCC calculations [120]. A complete overview of model construction and assumptions can be found in Appendix 1.

$$FCC^i = \frac{\partial \ln J_{ss}}{\partial \ln V_{max\,i}} \quad (1)$$

$$R^i = \frac{\partial \ln J_{ss}}{\partial \ln P_i} \quad (2)$$

Construction of *Δhxk1Δhxk2* strain IMX165 and control strain IMX2015

The *Δhxk1Δhxk2* strain IMX165 which was used as control in the invertase assay was constructed in three steps. The *HXK1* and *HXK2* deletion cassettes were PCR amplified from pUG73 and pUG6 respectively using the primers listed in Table S8F. First, *HXK1* was removed from CEN.PK102-12A by transformation with the *HXK1* deletion cassette containing the *Kluyveromyces lactis LEU2* marker flanked with loxP sites and *HXK1* recombination flanks resulting in strain IMX075. To remove the *LEU2* marker from this strain, it was transformed with the plasmid pSH47 containing the galactose inducible Cre recombinase [121]. Transformants were plated on SMG with histidine and were transferred to YPGal for Cre recombinase induction to remove *LEU2,* resulting in strain IMS0336. Subsequently, this strain was transformed with the *HXK2* deletion cassette containing the *KanMX* marker flanked with LoxP sites and *HXK2* recombination flanks, resulting in IMX165. IMX2015 was constructed as control strain for the characterization of the human hexokinase complementation strains. In this strain Sc*HXK2* is expressed with the p*PDC1* promoter instead of the native *HXK2* promoter. *pPDC1* was PCR amplified from genomic DNA from CEN.PK113-7D with primer 14670 and 14671 containing *HXK2* recombination flanks. 500 ng of this fragment was transformed to IMX1076 together with 800 ng of pUDE327 containing a gRNA targeting the *HXK2* promoter (table S9F).

Illumina whole genome sequencing

Genomic DNA for sequencing was isolated with the the Qiagen 100/G kit according to the manufacturer's description (Qiagen, Hilden, Germany) and library preparation and sequencing was done as described previously using Illumina Miseq sequencing (Illumina, San Diego, CA) [113]. Data was aligned to a CEN.PK113-7D reference [122] and data was further processed using SAMTools [123], and SNPs were determined using Pilon [124]. The output was visualized using the Integrative Genomics Viewer and chromosome copy numbers were determined using Magnolya

[125]. A list of mutations is provided in Table S3. For the mutation found in the *SBE2* gene which is involved in bud growth, it is unlikely to have an effect since it has a functionally redundant paralog *SBE22* [126]. No abnormalities were observed under the microscope. The sequencing data generated in this project are accessible at NCBI under bioproject PRJNA717746.

Quantitative aerobic batch cultivations

Quantitative characterization of strains IMX1821, IMX1814 and IMX1844 was done in 2 L bioreactors with a working volume of 1.4 L (Applikon, Schiedam, The Netherlands). The cultivation was done in synthetic medium supplemented with 20 g L^{-1} glucose, 1.4 mL of a vitamin solution [106] and 1.4 mL of 20% (v/v) Antifoam emulsion C (Sigma, St. Louise, USA). During the fermentation 0.5 mL extra antifoam was added when necessary. The salt and antifoam solution were autoclaved separately at 121°C and the glucose solution at 110°C for 20 min. During the fermentation the temperature was kept constant at 30°C and the pH at 5 by automatic addition of 2 M KOH. The stirring speed was set at 800 rpm. The medium was flushed with 700 mL min^{-1} of air (Linde, Gas Benelux, The Netherlands). For preparation of the inoculum, freezer stocks were inoculated in 100 mL YPD and grown overnight. From this culture the pre-culture was inoculated in 100 mL SMG which was incubated till mid-exponential growth phase. This culture was used to inoculate the inoculum flasks which were incubated till OD 4.5. The cells were centrifuged for 10 min at 3000g and the pellet was suspended in 100 mL demineralized water and added to the fermenter to start the fermentation with an OD of 0.25-0.4.

Biomass dry weight determination was done as previously described [106] by filtering 10 mL of culture on a filter with pore-size 0.45 mm (Whatman/GE Healthcare Life Sciences, Little Chalfont, United Kingdom) in technical duplicate. For extracellular metabolite analysis 1 mL of culture was centrifuged for 3 min at 20000g and the supernatant was analysed using high performance liquid chromatography (HPLC) using an Aminex HPX-87H ion-exchange column operated at 60°C with 5 mM H_2SO_4 as the mobile phase with a flow rate of 0.6 mL min^{-1} (Agilent, Santa Clara). Off-gas was cooled in a condenser and dried with a Perma Pure Dryer (Perma Pure, Lakewood, NJ), O_2 and CO_2 concentrations were measured using a NGA 2000 Rosemount gas analyzer (Emerson, St. Louis, MO). The OD_{660} was measured with a Jenway 7200 spectrophotometer (Jenway, Staffordshire, UK) at 660 nm. Per strain at least two independent fermentations were performed. The carbon balances for all reactors closed within 5%.

Growth rates (μ_{max}) were determined from OD_{660} data by linear regression of the log-linear plots of OD_{660} versus time over at least six consecutive time-points in a range

where the OD_{660} doubled twice, choosing the range to maximize the squared Pearson correlation coefficient (R^2). Dry weights for all time points were estimated through a linear correlation of OD_{660} and dry weight based on the dry-weights from the same experiment. Molar yields for the products biomass, CO_2, ethanol, glycerol and acetate on glucose were determined as the slope of the correlation of the product concentration against glucose concentration during the exponential phase (the same time period used for growth rate determination). The specific glucose uptake rate (q_{glu}) was determined by dividing the growth rate through the biomass yield on glucose. Specific product production rates and the oxygen consumption rate was determined by multiplying the molar yields with the specific glucose uptake rate.

Sample preparation and enzymatic assays for comparison of yeast and humanized yeast samples

Yeast samples were prepared as previously described [127], from exponentially growing cultures (62 mg dry weight per sample) from bioreactor main activity determinations and for testing of allosteric effectors and for comparison of the evolved strains from shake flask. Sonication was used for cell-free extract preparation except for the hexokinase measurements (Fig. 2, S4 and S5) where fast-prep was used. All determinations were performed at 30°C and 340 nm (εNAD(P)H at 340 nm/6.33 mM^{-1}). In most cases glycolytic V_{max} enzyme activities were determined in 1 mL reaction volume (in 2 mL cuvettes), using a Hitachi model 100-60 spectrophotometer, using previously described assays [56], except for phosphofructokinase activity which was determined according to Cruz *et al.*[128]. When required extracts were diluted in water.

Briefly for each enzyme assays contained:

Hexokinase (HXK; EC2.7.1.1) – 50 mM Imidazole-HCl (pH 7.6), 1 mM $NADP^+$, 10 mM $MgCl_2$, 10 mM Glucose, 1.8 U/mL glucose-6-phosphate dehydrogenase (EC1.1.1.49), and 1 mM ATP as start reagent.

Phosphoglucose isomerase (GPI; EC5.3.1.9) – 50 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 0.4 mM $NADP^+$, 1.8 U/mL glucose-6-phosphate dehydrogenase (EC1.1.1.49), and 2 mM fructose 6-phosphate as start reagent.

Phosphofructokinase (PFK; EC2.7.1.11) – 50 mM Imidiazole-HCl (pH 7.0), 1 mM AMP, 5 mM $MgCl_2$, 0.15 mM NADH, 0.45 U/mL aldolase (EC4.1.2.13), 0.6 U/mL glycerol-3P-dehydrogenase (EC1.1.1.8), 1.8 U/mL triosephosphate isomerase (EC5.3.1.1), 0.5 mM Fructose-6-phosphate, and 1 mM ATP as start reagent.

Aldolase (ALDO; EC4.1.2.13) – 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 0.15 mM NADH, 0.6 U/mL glycerol-3P-dehydrogenase (EC1.1.1.8), 1.8 U/mL

triosephosphate isomerase (EC5.3.1.1), and 2 mM fructose 1,6-bisphosphate as start reagent.

Triosephosphate isomerase (TPI, EC5.3.1.1) – 200 mM Triethanolamine (pH 8.2), 0.15 mM NADH, 5 U/mL Glycerol 3P-dehydrogenase (EC1.1.1.8), 17.4 mM Glyceraldehyde-3-phosphate as start reagent.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH; EC1.2.1.12) – 100 mM Triethanolamine-HCl (pH 7.6), 1 mM ATP, 1 mM EDTA, 1.5 mM $MgSO_4$, 0.15 mM NADH, 16.5 U/mL 3-phosphoglycerate kinase (EC2.7.2.3), and 5 mM 3-phosphoglyceric acid as start reagent.

3-phosphoglycerate kinase (PGK; EC2.7.2.3) – 100 mM Triethanolamine-HCl (pH 7.6), 1 mM ATP, 1 mM EDTA, 1.5 mM $MgSO_4$, 0.15 mM NADH, 8 U/mL glyceraldehyde-3-phosphate dehydrogenase (EC1.2.1.12), and 5 mM 3-phosphoglyceric acid as start reagent.

Phosphoglycerate mutase (PGAM; EC5.4.2.1) – 100 mM Triethanolamine-HCl (pH 7.6), 1.5 mM $MgSO_4$, 0.15 mM NADH, 10 mM ADP, 1.25 mM 2,3-diphospho-glyceric acid, 6.3 U/mL enolase (EC4.2.1.11), 13 U/mL pyruvate kinase (EC2.7.1.40), 11.3 U/mL L-lactate dehydrogenase (EC1.1.1.27), and 5 mM 3-phosphoglyceric acid as start reagent.

Enolase (ENO; EC4.2.1.11) – 100 mM Triethanolamine-HCl (pH 7.6), 1.5 mM $MgSO_4$, 0.15 mM NADH, 10 mM ADP, 9 U/ml pyruvate kinase (EC2.7.1.40), 11.3 U/mL L-lactate dehydrogenase (EC1.1.1.27), and 1 mM 2-phosphoglyceric acid as start reagent.

Pyruvate kinase (PK; EC2.7.1.40) – 100 mM Cacodylic Acid (pH 6.2), 100 mM KCl, 10 mM ADP, 1 mM fructose 1,6-bisphosphate, 25 mM $MgCl_2$, 0.15 mM NADH, 8 U/mL L-lactate dehydrogenase (EC1.1.1.27) and 2 mM phosphoenolpyruvate as start reagent.

To increase throughput, the specific activities of the evolved strains and the glucose, ATP and ADP dependency of the hexokinase complementation strains (Fig S4 C-E) were assayed using a TECAN infinite M200 Pro. (Tecan, Männedorf, Switzerland) microtiter plate reader. Samples were prepared manually in microtiter plates (transparent flat-bottom Costar plates; 96 wells) using a reaction volume of 300 µl per well. The assays were the same as for the cuvette-based assays. For determination of glucose-6-phosphate inhibition of hexokinase, cell extracts were prepared in Tris-HCl buffer (50 mM, pH 7.5) to limit phosphate concentrations, which could influence glucose-6-phosphate inhibition. For these measurements an alternative enzyme assay coupled by pyruvate kinase and lactate dehydrogenase

was used based on [129], buffer and metabolite concentrations were kept the same as the yeast hexokinase assay.

Alternative assay hexokinase (HXK; EC2.7.1.1) – 50 mM Imidazole-HCl (pH 7.6), 10 mM $MgCl_2$, 0.15 mM NADH, 0.5 mM Phosphoenolpyruvate, 1 mM ATP, 2 U/mL Pyruvate Kinase (EC2.7.1.40), 4 U/ml L-lactate dehydrogenase (EC1.1.1.27) and 10 mM Glucose as start reagent, supplemented with various glucose-6-phosphate concentrations.

The reported data are based on at least two independent biological replicate samples, with at least two analytic replicates per sample per assay, including two different cell free extract concentrations. The protein concentration was determined using the Lowry method with bovine serum albumin as a standard [130]. Enzyme activities are expressed as µmol substrate converted (mg protein)$^{-1}$ min^{-1}.

To calculate the degree of saturation of glycolytic enzymes, the specific activity in µmol.mg$_{protein}^{-1}$.h^{-1} was converted into mmol.g$_{DW}^{-1}$.h^{-1} considering that soluble proteins represent 30% of cell dry weight. This value represents the maximal enzyme flux capacity. The *in vivo* flux in the glycolytic reactions were approximated from the glucose specific uptake rate (q_{glu}). Reactions in the top of glycolysis (hexokinase to triosephosphate isomerase) were assumed to equal the q_{glu}, while reactions in the bottom of glycolysis (glyceraldhyde-3P dehydrogenase to pyruvate kinase) were calculated as the q_{glu} times two. The degree of saturation was calculated as follows:

$$degree\ of\ saturation = \frac{approximated\ in\ vivo\ flux}{maximal\ flux\ capacity} \times 100$$

Invertase enzyme assay

The invertase assay was performed on whole cells previously described [131]. Exponentially growing cells in SMG were washed with sterile dH_2O, transferred to shake flasks (at OD 3) with 100 mL fresh SMG or SME+0.075% glucose and incubated for 2h at 30°C and shaking at 200 rpm. Afterwards the dry weight of the cultures was determined and the cells were washed in 50mM sodium acetate buffer wit 50mM NaF to block the metabolism and were then suspended till a concentration of 2.5-7.5 mg dry weight per mL. 4 mL of this cell suspension were added to a dedicated vessel thermostated at 30°C, and kept under constant aeration by flushing with air (Linde, Gas Benelux, The Netherlands) and stirring with a magnetic stirrer. The reaction was started by addition of 1 mL 1M sucrose and 1 mL reaction mix was taken at 0, 1, 2, 3 and 5 minutes, directly filtered using 13 mm diameter 0.22 µm pore size nylon syringe filters to remove cells and put on ice. Afterwards the glucose concentration resulting from sucrose hydrolysis by invertase was determined using a D-Glucose assay kit (Megazyme, Bray, Ireland). The glucose production rate was calculated in µMol·min^{-1}·g dry weight^{-1}.

Staining of vacuoles

Yeast strains were stained with the red fluorescent dye FM 4-64 (excitation/emission, 515/640 nm) (Thermo Fisher Scientific). Exponentially growing cells were incubated at an OD of 0.5-1 in YPD with 2 µM FM4-64 in the dark for 30 minutes at 30˚C. Afterwards cells were spun down, washed and incubated for 2-3 h in 5 mL YPD. For analysis, cells were spun down and suspended in SMG medium. Yeast cells and vacuoles were visualized with an Imager-Z1 microscope equipped with an AxioCam MR camera, an EC Plan-Neofluar 100x/1.3 oil Ph3 M27 objective, and the filter set BP 535/25, FT 580, and LP 590 (Carl-Zeiss, Oberkochen, Germany). Images were analysed in ImageJ (NIH).

Ploidy determination by flow cytometry

Samples of culture broth (equivalent to circa 10^7 cells) were taken from mid-exponential shake-flask cultures on YPD and centrifuged (5 min, 4700g). The pellet was washed once with demineralized water, and centrifuged again (5 min, 4700g) and suspended in 800 µL 70% ethanol while vortexing. After addition of another 800 µL 70% ethanol, fixed cells were stored at 4˚C until further staining and analysis. Staining of cells with SYTOX® Green Nucleic Acid Stain (Invitrogen S7020) was performed as described [132]. Samples were analysed on a BD Accuri C6 flow cytometer equipped with a 488 nm laser (BD Biosciences, Breda, The Netherlands). The fluorescence intensity (DNA content) was represented using FlowJo (v. 10.6.1, FlowJo, LLC, Ashland, OR, USA), (Fig. S19).

Transition experiment

For testing transitioning between carbon sources, strains were grown overnight in SMGal medium till mid-exponential phase. These cultures were used to plate single cells on SMG and SMGal plates (96 cells per plate) using a BD FACSAriaII (Franklin Lakes, NJ). After 5 days the percentage of cells which was growing was determined.

Whole cell lysate proteomics of humanized yeast strains

Sampling, cell lysis, protein extraction, in-solution proteolytic digestion and TMT labelling

For proteomics analysis, yeast strains grown to exponential phase in SMG shake flasks were inoculated to fresh shake flasks in biological triplicates. During mid-exponential phase samples of 8 ml were taken, centrifuged for 10 min. at 5000 g at 4°C and the pellet was stored at -80°C. Approx. 50mg of cell pellets (wet weight) were lysed using beads beating in 1% SDS, 100mM TEAB, including protease inhibitor and phosphate inhibitor. The lysed cells were centrifuged and the supernatant was transferred to a new tube. The proteins were reduced with dithiothreitol and alkylated using iodoacetamide (IAA), where the protein content was precipitated and washed using ice cold acetone. The protein pellets were dissolved

in 100mM ammonium bicarbonate buffer, and subjected to overnight digestion using proteomics grade Trypsin at 37°C and under gentle shaking using an Eppendorf incubator. The peptides were desalted using solid-phase extraction on a Waters Oasis HLB 96-well µElution plate according to the manufacturers protocol, SpeedVac dried and stored at -20°C until analysed. One aliquot of each sample was dissolved in 3% acetonitrile in H_2O, containing 0.1% formic acid and subjected to nLC-Orbitrap-MS analysis for digestion quality control purpose. One aliquot of each sample was further dissolved in 100mM TEAB for further labelling using a TMT 10plex labelling kit (Thermo scientific, catalog number: 90110). The peptide content of the samples were estimated using a NanoDrop photospectrometer, and the samples were diluted to achieve an approx. equal concentration. The TMT labelling agents were dissolved by adding 40µL anhydrous acetonitrile, and 5µL of each label was added to the individual samples. The samples were incubated at 25°C under gentle shaking, for 75 minutes. Then the reaction was stopped by adding diluted hydroxylamine solution, and incubated at 25°C under gentle shaking. Equal amounts of sample were combined in a LoBind Eppendorf tube. After dilution with aqueous buffer, solid phase extraction was once more performed according to the manufacturers protocol. The samples were SpeedVac dried, and stored at -20°C until analysed. Before analysis, the samples were solubilised in 3% acetonitrile and 0.01% TFA and subjected to nLC-Orbitrap-MS analysis.

Whole cell lysate shotgun proteomics

An aliquot corresponding to approx. 500ng of every TMT 10plex peptide mixture (SET1, SET2 and SET3, where 3 strains were mixed and compared within one SET, and each strain within one SET was present as triplicate) was analysed by duplicate analysis employing one-dimensional shotgun proteomics. The sets were chosen to be able to directly compare the evolved strains with their parental strain and to compare the humanized strains with the *Sc*Gly strain, SET1 contained strains IMX1814, IMS0990 and IMS0991, SET2 IMX1844, IMS0987 and IMS0989, and SET3 IMX1814, IMX1844 and IMX1821. Briefly, TMT labelled peptides were analysed using a nano-liquid-chromatography system consisting of an EASY nano LC 1200, equipped with an Acclaim PepMap RSLC RP C18 separation column (50 µm x 150 mm, 2µm and 100 Å), online coupled to a QE plus Orbitrap mass spectrometer (Thermo Scientific, Germany). The flow rate was maintained at 350nL/min over a linear gradient from 5% to 25% solvent B over 178 minutes, and finally to 55% B over 60 minutes. Solvent A consisted of H2O containing 0.1% formic acid, and solvent B consisted of 80% acetonitrile in H2O, plus 0.1% formic acid. The Orbitrap was operated in data-dependent acquisition mode acquiring peptide signals form 385-1450 m/z at 70K resolution, 75ms max IT, and an AGC target of 3e6, where the top 10 signals were isolated at a window of 1.6m/z and fragmented at a NCE of

32. The peptide fragments were measured at 35K resolution, using an AGC target of 1e5 and allowing a max IT of 100ms.

MS raw data processing **and determination of protein expression levels**
For protein identification, raw data were processed using PEAKS Studio 10.0 (Bioinformatics Solutions Inc.) allowing 20ppm parent and 0.01Da fragment mass error tolerance, TMT10plex and Carbamidomethylation as fixed, and methionine oxidation and N/Q deamidation as variable modifications. Data were matched against an in-house established yeast protein sequence database, including the GPM crap contaminant database (https://www.thegpm.org/crap/) and a decoy fusion for determining false discovery rates. Peptide spectrum matches were filtered against 1% false discovery rate (FDR) and protein identifications were accepted as being significant when having 2 unique peptides matches minimum. Quantitative analysis of the global proteome changes between the individual yeast strains was performed using the PEAKS-Q software package (Bioinformatics Solutions Inc.), considering a quantification mass tolerance of 10ppm, a FDR threshold of 1%, using auto normalisation and ANOVA as the significance method. Significance (-10log(p)) vs fold change volcano plots were created using the scatter function in Matlab2019b (Fig. S9 and S15). TIC normalized signal intensity was calculated by dividing the signal intensity by the total intensity of each sample (Fig. S8 and S13). Mass spectrometric raw data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE repository with the dataset identifier PXD025349.

Comparison human and humanized yeast glycolytic enzymes

Cell-free extract preparation and V_{max} enzyme assays
Human cells stored at -80°C were thawed, centrifuged at 20000 g for 10 minutes at 4 °C and the pellet was discarded to obtain cell-free extracts.

Yeast samples (IMX1844) were harvested as previously described [127] from exponentially growing cultures (62 mg dry weight per sample) from bioreactor. Cell-free extract preparation for yeast cells was done using YeastBuster™ Protein Extraction Reagent supplemented with 1% of 100x THP solution according to the description (Novagen, San Diego, CA, USA). To a pellet with a wet weight of 0.3 g, 3.5 mL YeastBuster and 35 µl THP solution was added. The pellet was suspended and incubated for 20 min at room temperature. Afterwards the cell debris was removed by centrifugation at 20000 g for 15 min at 4 °C and the supernatant was used for the assays.

Prior to experimentation, YeastBuster™ Protein Extraction Reagent with 1% THP (Novagen) was added to the human cell samples and DPBS supplemented with protease inhibitor was added to the yeast samples (both as 50% of final volume).

This strategy was taken in order to equalize the buffer composition of yeast and human culture samples to perform enzyme kinetics assays and proteomics.

V_{max} assays for comparison of yeast and human cell extracts were carried out with freshly prepared extracts via NAD(P)H-linked assays at 37 °C in a Synergy H4 plate reader (BioTek™). The reported V_{max} values represent total capacity of all isoenzymes in the cell at saturating concentrations of all substrates and expressed per extracted cell protein. Four different dilutions of extract were used to check for linearity. Unless otherwise stated, at least 2 dilutions were proportional to each other and these were used for further calculation. All enzymes were expressed as µmoles of substrate converted per minute per mg of extracted protein. Protein determination was carried out with the Bicinchoninic Acid kit (BCA™ Protein Assay kit, Pierce) with BSA (2 mg/ml stock solution of bovine serum albumin, Pierce) as standard.

Based on the cytosolic concentrations described in literature, we have designed an assay medium that was as close as possible to the *in vivo* situation, whilst at the same time experimentally feasible. The standardized *in vivo*-like assay medium contained 150 mM potassium[133-136], 5 mM phosphate [133, 137], 15 mM sodium [133, 138], 155 mM chloride [139, 140], 0.5 mM calcium, 0.5 mM free magnesium [133, 141, 142] and 0.5-10.5 mM sulfate. For the addition of magnesium, it was taken into account that ATP and ADP bind magnesium with a high affinity. The amount of magnesium added equalled the concentration of either ATP or ADP plus 0.5 mM, such that the free magnesium concentration was 0.5 mM. Since the sulfate salt of magnesium was used the sulfate concentration in the final assay medium varied in a range between 0.5 and 10.5 mM. The assay medium was buffered at a pH of 7.0 [143-148] by using a final concentration of 100 mM Tris-HCl (pH 7.0). To end up with the above concentrations, an assay mixture containing 100 mM Tris-HCl (pH 7.0), 15 mM NaCl, 0.5 mM $CaCl_2$, 140 mM KCl, and 0.5-10.5 mM $MgSO_4$ was prepared.

In addition to the assay medium, the concentrations of the coupling enzymes, allosteric activators and substrates for each enzyme were as follows:

Hexokinase (HK; EC2.7.1.1) – 1.2 mM $NADP^+$, 10 mM Glucose, 1.8 U/mL glucose-6-phosphate dehydrogenase (EC1.1.1.49), and 10 mM ATP as start reagent.

Phosphoglucose isomerase (GPI; EC5.3.1.9) – 0.4 mM $NADP^+$, 1.8 U/mL glucose-6-phosphate dehydrogenase (EC1.1.1.49), and 2 mM fructose 6-phosphate as start reagent.

Phosphofructokinase (PFK; EC2.7.1.11) – 0.15 mM NADH, 1 mM ATP, 0.5 U/mL aldolase (EC4.1.2.13), 0.6 U/mL glycerol-3P-dehydrogenase (EC1.1.1.8), 1.8 U/mL triosephosphate isomerase (EC5.3.1.1), 65 µM fructose 2,6-bisphosphate as

activator (synthesized as previously described [149]), and 10 mM fructose 6-phosphate as start reagent.

Aldolase (ALDO; EC4.1.2.13) – 0.15 mM NADH, 0.6 U/mL glycerol-3P-dehydrogenase (EC1.1.1.8), 1.8 U/mL triosephosphate isomerase (EC5.3.1.1), and 2 mM fructose 1,6-bisphosphate as start reagent.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH; EC1.2.1.12) – 0.15 mM NADH, 1 mM ATP, 24 U/mL 3-phosphoglycerate kinase (EC2.7.2.3), and 5 mM 3-phosphoglyceric acid as start reagent.

3-phosphoglycerate kinase (PGK; EC2.7.2.3) – 0.15 mM NADH, 1 mM ATP, 8 U/mL glyceraldehyde-3-phosphate dehydrogenase (EC1.2.1.12), and 5 mM 3-phosphoglyceric acid as start reagent.

Phosphoglycerate mutase (PGAM; EC5.4.2.1) – 0.15 mM NADH, 1 mM ADP, 2.5 mM 2,3-diphospho-glyceric acid, 5 U/mL enolase (EC4.2.1.11), 50 U/mL pyruvate kinase (EC2.7.1.40), 60 U/mL L-lactate dehydrogenase (EC1.1.1.27), and 5 mM 3-phosphoglyceric acid as start reagent.

Enolase (ENO; EC4.2.1.11) – 0.15 mM NADH, 1 mM ADP, 50 U/ml pyruvate kinase (EC2.7.1.40), 15 U/mL L-lactate dehydrogenase (EC1.1.1.27), and 1 mM 2-phosphoglyceric acid as start reagent.

Pyruvate kinase (PK; 2.7.1.40) – 0.15 mM NADH, 1 mM ADP, 1 mM fructose 1,6-bisphosphate, 60 U/mL L-lactate dehydrogenase (EC1.1.1.27) and 2 mM phosphoenolpyruvate as start reagent.

Determination of absolute enzyme concentrations [E]
Absolute concentrations of glycolytic targets was performed by targeted proteomics [150]. Isotopically labelled peptides with ^{13}C lysines and arginines were designed for human glucose metabolism and a list of peptides of interest detected in our samples can be found in Table S6.

Turnover number (k_{cat}) calculations
Turnover numbers were estimated based on the equation $V_{max} = k_{cat} \cdot [E]$, the maximal enzyme activity was divided by the concentration of each individual peptide detected by proteomics when no other isoform was detected. In the human skeletal muscle samples, more than one isoform was detected for certain proteins. In these cases (phosphoglycerate mutase, enolase and pyruvate kinase) the sum of the concentrations of all isoforms was used to estimate the turnover number. In Fig. S17 k_{cat} values were calculated for the specific isoforms based on the ratio between turnover numbers found in the literature for the human isoforms. Protein concentrations [E] were measured as $pmol \cdot mg\ protein^{-1}$ and V_{max}'s as $\mu mol \cdot min^{-1}$

· *mg protein*$^{-1}$. In order to obtain k_{cat} values in *min*$^{-1}$, the following equation was used for each enzymatic reaction in the dataset:

$$k_{cat} = \frac{V_{max}}{[E]} \cdot 10^6$$

Quantification and statistical analysis

Statistical analysis was performed in Graphpad Prism and Microsoft Excel. K_m and IC_{50} estimations were performed in Graphpad Prism. Statistical details (number of replicates and used tests) of experiments can be found in the figure legends. Significance was usually defined by p-values < 0.05.

Acknowledgements

We thank Jordi Geelhoed for his work on reverse engineering and growth characterization, Agnes Hol and Ingeborg van Lakwijk for their valuable work on the human hexokinase strains and Lycka Kamoen, Eveline Vreeburg and Daniel Solis-Escalante for their contribution to construction of reference strains. We thank Carol de Ram for processing proteomics samples. We thank Pilar de la Torre and Marcel van den Broek for whole genome sequencing and analysis, and Philip de Groot for help with bioreactor sampling.

Supplementary data

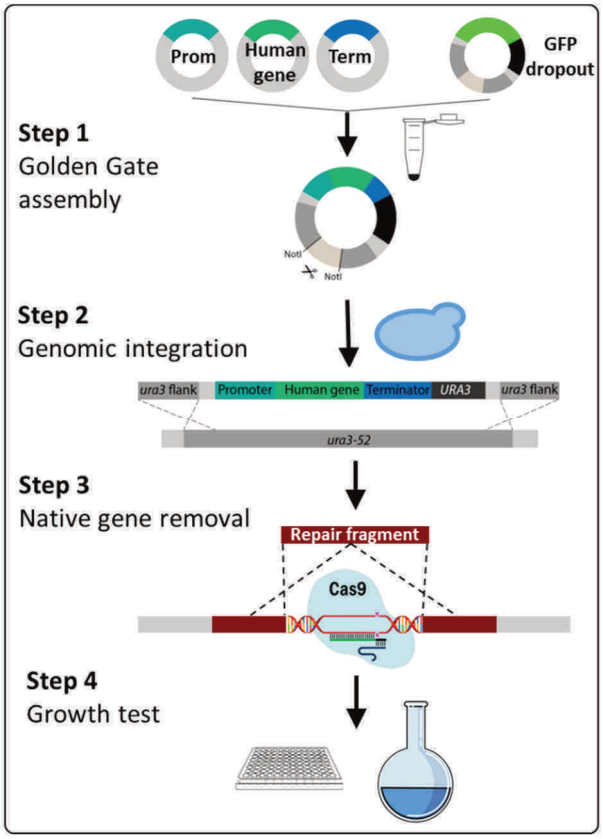


Figure S1 - Complementation strategy
Codon-optimized human glycolytic genes were stitched to yeast promoters and terminators using Golden Gate assembly (step 1). The resulting plasmids were linearized by restriction with NotI and integrated in the *ura3-52* locus of the minimal glycolysis strain IMX1076 which was plated on SMG (Step 2). In a second transformation round, the yeast ortholog was selectively removed using CRISPR/Cas9-mediated DNA editing. Transformed cells were plated on YPD + G418, except for the *HsHK1-3* strains which were plated on YPGal + G418 (step 3). Gene integration and deletion was checked for all strains by PCR and Sanger sequencing. All strains were tested for growth on chemically defined medium with glucose as sole carbon source (SMG).

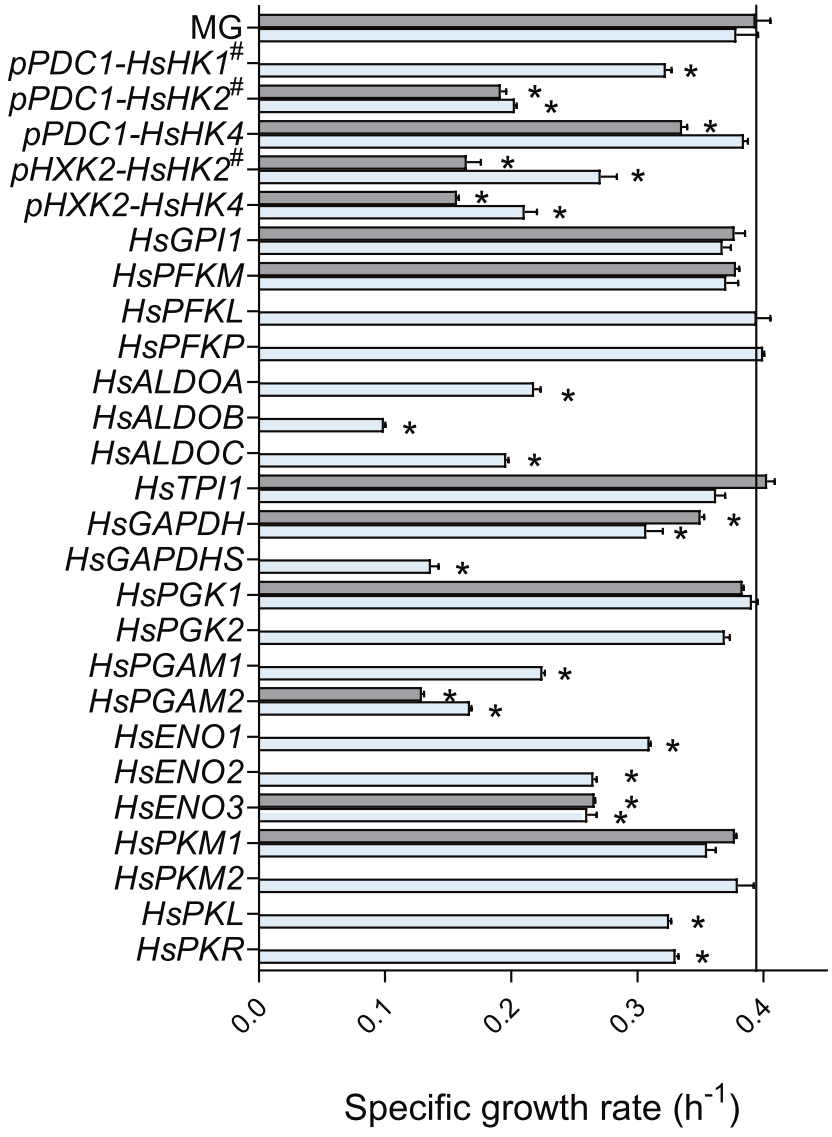


Figure S2 - Growth rate of complementation strains on SMG
Specific growth rates of complementation strains and minimal glycolysis (MG) control strain determined in shake flask (grey) and growth profiler (light blue) using chemically defined medium with glucose as sole carbon source (SMG). # Indicates strains with mutations after growth on glucose, *pPDC1-HsHK1*: IMS1137, *pPDC1-HsHK2*: IMX1689, *pHXK2-HsHK2*: IMX1873. Error bars represent SEM of duplicates (shake flask experiments) or triplicates (growth profiler). * indicates significant difference from control strain MG (IMX372) ($P < 0.01$, Student t-test, two-tailed, homoscedastic).

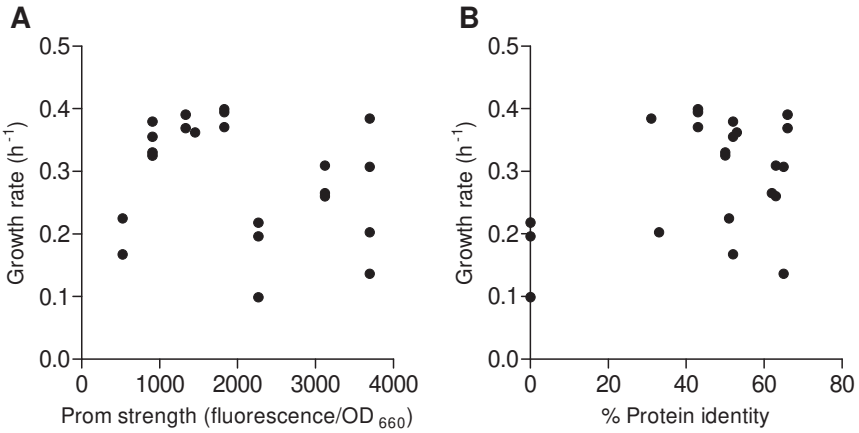
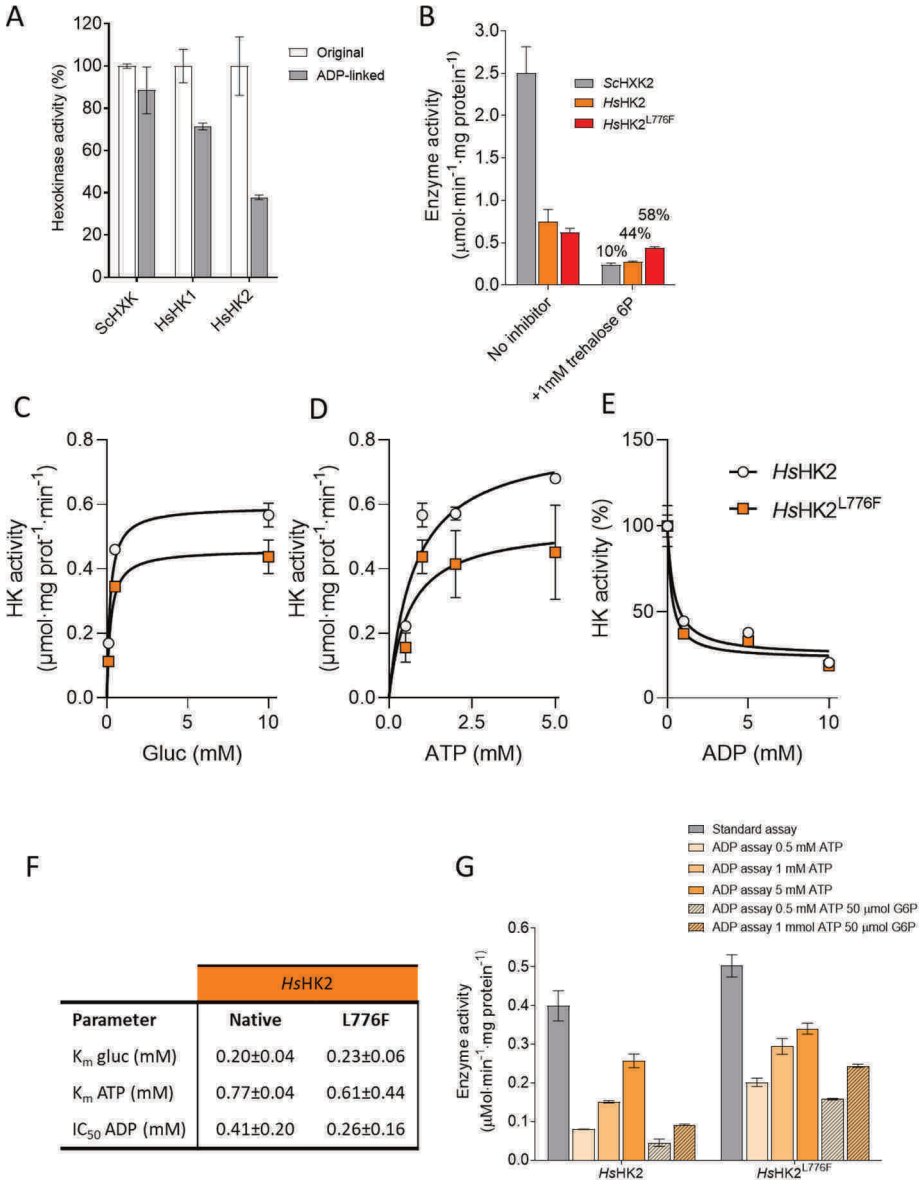


Figure S3 - Comparison of promoter strength, protein identity and growth rate in the complementation strains

The growth rate of the complementation strains measured in the growth profiler was plotted against: **A)** the strength of the promoter used to express the human gene. Promoter strengths were obtained from Boonekamp *et al.* 2018 [1]. **B)** Percentage protein identity between the human gene and the corresponding yeast ortholog.



	*Hs*HK2	
Parameter	**Native**	**L776F**
K_m gluc (mM)	0.20±0.04	0.23±0.06
K_m ATP (mM)	0.77±0.04	0.61±0.44
IC_{50} ADP (mM)	0.41±0.20	0.26±0.16

Figure S4 - Human hexokinase 2 complementation strain characterization
In vitro biochemical characterization was performed for the non-mutated *Hs*HK2 and *Hs*HK2^{L776F} variant (strains IMX2419 and IMX1690 respectively). Strains were grown in shake flask on SM-Galactose medium to prevent any selection for mutations. **A)** As the standard hexokinase assay was based on G6P detection (original assay), a different assay based on ADP detection (ADP-linked assay) was implemented to quantify hexokinase sensitivity to G6P. The ADP-linked assay systematically detected a lower hexokinase activity, possibly due to G6P accumulation during the assay. Hexokinase activity expressed as % of the activity measured with the original assay for the same strain. For all other measurements (panel B-F) the original G6PDH-linked assay was used. Error bars represent SEM of duplicates. **B)** Inhibition by trehalose-6-phosphate of the yeast *Sc*Hxk2 (strain IMX2015) and the native and mutated variants of *Hs*HK2 enzymes. 1 mM trehalose-6P resulted in 90% inhibition of the yeast *Sc*Hxk2, while it only marginally affected the human *Hs*HK2. As concentrations of Trehalose-6P above 1 mM are not often encountered in yeast cells, [2-4], trehalose-6P inhibition was probably not responsible for the lack of activity of the native *Hs*HK2 in yeast cells. Inhibition of *Hs*HK2 and *Hs*HK2^{L776F} was not significantly different (t-test, two-tailed, homoscedastic, p>0.05). Error bars represent SEM of biological duplicates, for T6P measurements individual technical replicates were measured. **C-E)** Response to varying concentrations of glucose, ATP and ADP of the two *Hs*HK2 variants. These measurements were performed in a 96-wells reader in biological duplicate with technical triplicates (different extract concentrations). **F)** Calculated kinetic parameters of *Hs*HK2 and *Hs*HK2^{L776F} based on the data shown in previous panels. **G)** Hexokinase activity at various concentrations of ATP and G6P concentrations in the ADP-linked assay. Increasing ATP concentrations led to an increase in measured activity in the ADP-linked assay consistent with an effect of the production of competitive inhibitor G6P which is not removed in this assay, SEM of duplicates is shown.

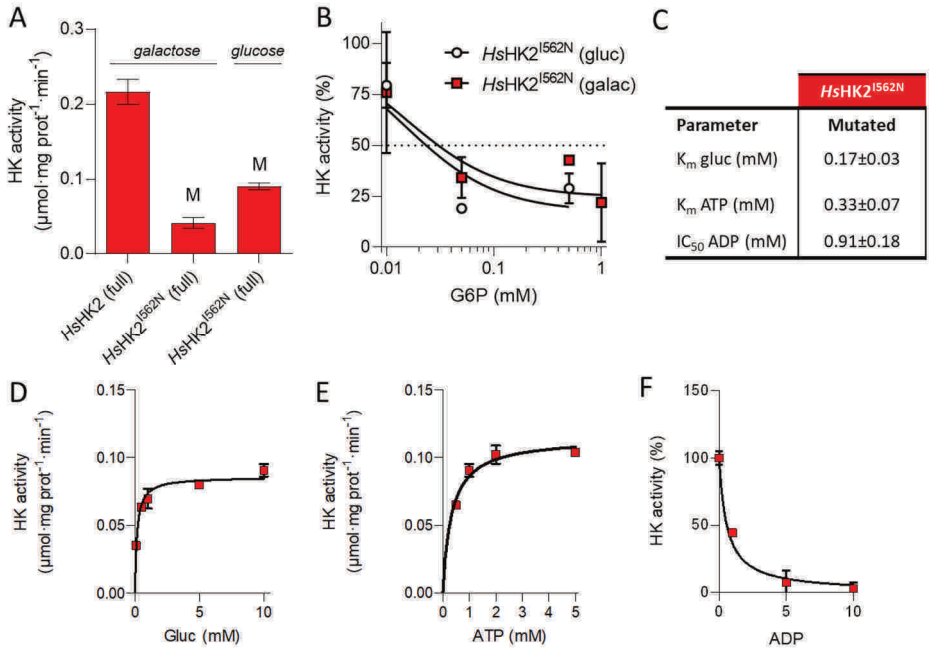


Figure S5 - Human hexokinase 2 characterization in the fully humanized strains
In vitro characterization of hexokinase in cell free extracts from the fully humanized *Hs*Gly-HK2 strain IMX1844 (*Hs*HK2^{I562N}) and fully humanized control strain IMX2496 (*Hs*Gly-HK2 without mutation in HK2). In both strains, *Hs*HK2 is expressed from the *ScHXK2* promotor, which is repressed in galactose cultures, hence both glucose and galactose grown cultures were compared for strain IMX1844. **A)** comparison of hexokinase activity measured in the IMX2496 and IMX1844 strains shows a strong decrease in activity in the mutant strain, SEM of duplicates is shown. **B)** The sensitivity of the mutated hexokinase to glucose-6-phosphate is independent of the carbon source used to grow the fully humanized strain. **C)** Calculated kinetic parameters of hexokinase *Hs*HK2^{I562N} in strain IMX1844 from the data shown in panels D-F. For these measurements strain IMX1844 was grown on SM glucose (non-repressive condition). All values are in the same order of magnitude as the values measured for the *Hs*HK2 complementation strains shown in Fig. S4 and hence do not explain the lower measured activity. **D-F)** Hexokinase activity of glucose-grown IMX1844 (*Hs*Gly-HK2) at various concentrations of glucose, ATP and ADP, error bars represent SEM of duplicates.

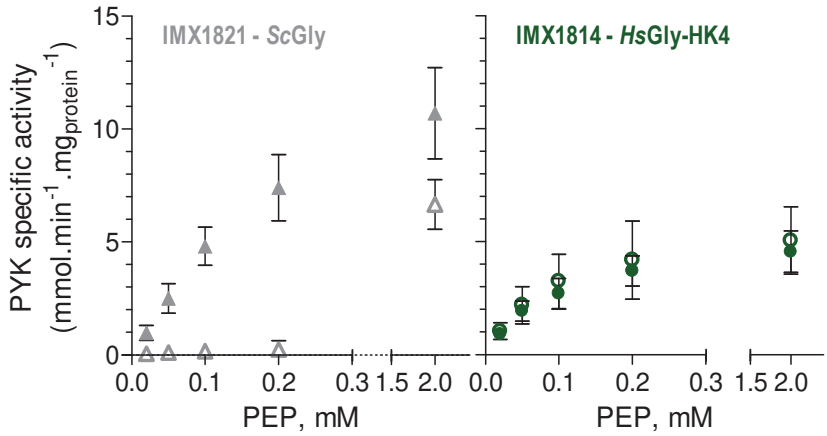


Figure S6 - Sensitivity of human and yeast pyruvate kinase to fructose-1,6bisP
Yeast (left) and human (right) specific pyruvate kinase activity *in vitro* assayed with fructose-1,6bisP (closed symbols) and without fructose-1,6bisP (open symbols) in IMX1821 (*Sc*gly) and IMX1814 (*Hs*gly) with different concentrations of the substrate phosphoenolpyruvate (PEP). Symbols and error bars represent the average and SEM of three biological replicates.

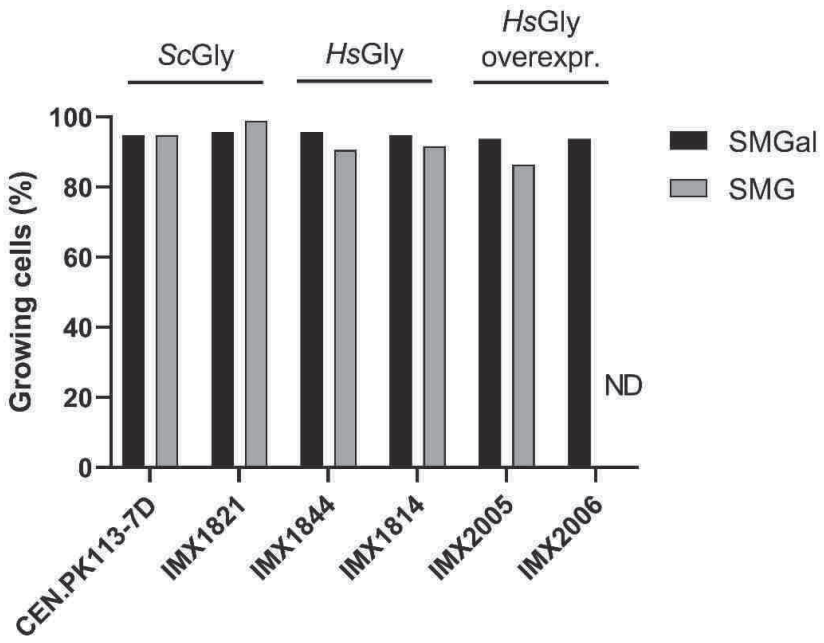


Figure S7 - Robustness to transitioning between carbon sources
The control *Sc*Gly and humanized glycolysis strains were cultivated in liquid synthetic medium with galactose as sole carbon source (SMGal), from these cultures, single cells were plated on plates with fresh SM medium with glucose (SMG) or galactose as sole carbon source. Per condition 96 single cells were plated. Five days after plating, the colonies growing on SMGal and SMG were counted. These data originate from a single experiment. ND; not determined. Strains with humanized glycolysis displayed no impairment in transitioning between respiratory and fermentative carbon sources.

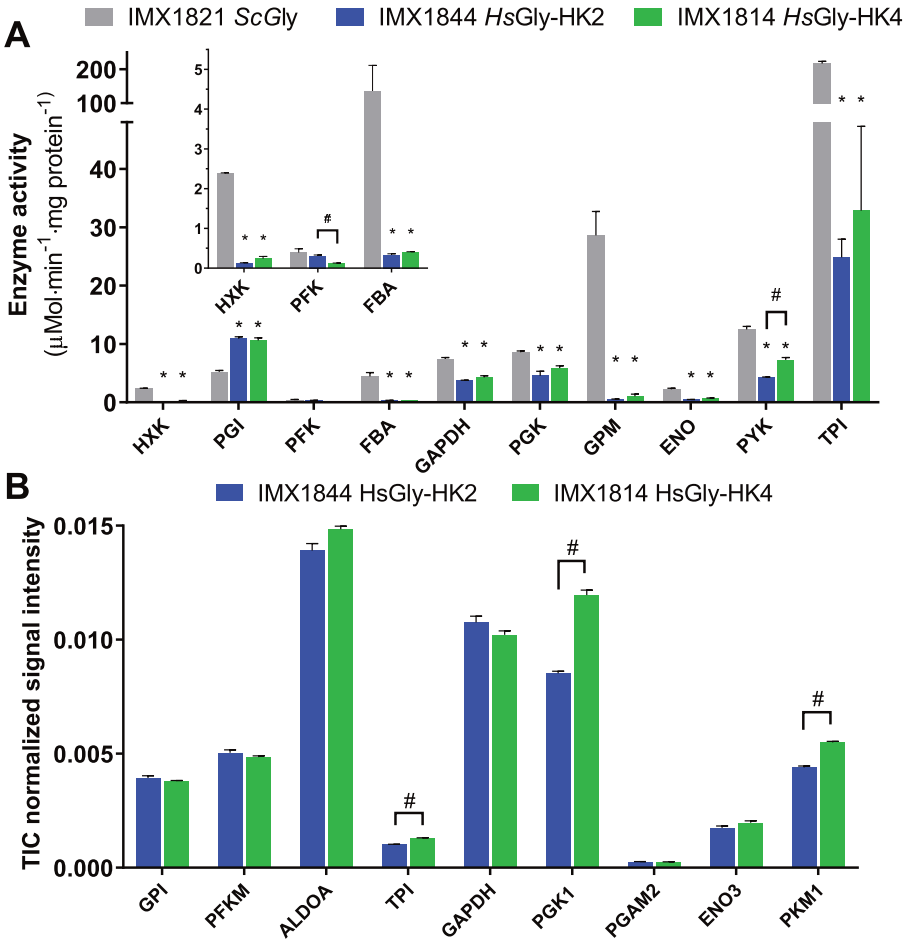
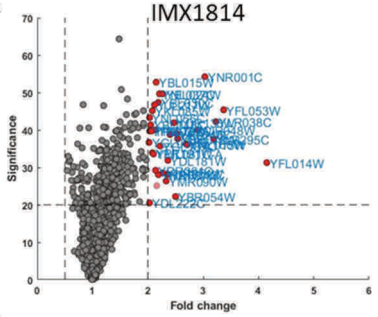


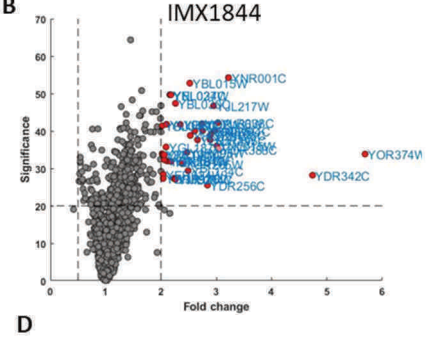
Figure S8 - Specific activity and relative levels of glycolytic enzymes in fully humanized yeast strains

A) Specific enzyme activities measured *in vitro* with cell free extracts from batch cultures in bioreactors. Error bars represent the SEM from two biological replicates. * indicate values that are significantly different from the control strain IMX1821 with *S. cerevisiae* glycolysis, # indicate significant difference between the two humanized strains (Student t-test, two-tailed, homoscedastic, P<0.05). **B)** Relative enzyme abundance of the human glycolytic enzymes (normalized to the total signal intensity) in the *Hs*Gly-HK2 (IMX1844) and *Hs*Gly-HK4 (IMX1814) strains. Means with SEM are shown for biological triplicates from one of the duplicate injections. # indicates significant differences between IMX1814 and IMX1844 (p-value <0.05 student t-test, two tailed, homoscedastic).

A



B



C

Systematic name	Common name	Fold change IMX1814 vs IMX1821	P value	TIC normalized abundance IMX1821
YFL014W	Hsp12	4.15	7.4E-04	1.5E-04
YFL053W	Dak2	3.37	2.9E-05	1.5E-05
YLR038C	Cox12	3.23	6.0E-05	2.8E-04
YLR295C	Atp14	3.19	1.7E-04	1.8E-04
YNR001C	Cit2/Cit1	3.03	3.8E-06	2.4E-03
YJR048W	Cyc7/Cyc1	2.9	1.0E-04	1.4E-03
YNL015W	Pbi2	2.71	2.5E-04	2.7E-04
YNL160W	Ygp1/Sps100	2.69	2.4E-04	1.2E-04
YBR072W	Hsp26	2.54	1.7E-04	2.6E-04
YKL148C	Sdh1	2.52	1.2E-04	3.6E-04
YBR054W	Mrh1/Yro2	2.5	6.0E-03	4.5E-05
YOR136W	Idh2	2.47	6.4E-05	8.9E-04
YHL021C	Aim17	2.4	1.3E-04	2.0E-05
YDR342C	hypothetical protein	2.36	1.5E-03	2.2E-05
YDL181W	Inh1/Stf1	2.36	6.5E-04	1.5E-04
YMR090W	hypothetical protein	2.33	2.3E-03	3.6E-05
YEL024W	Rip1	2.27	1.1E-05	5.0E-04
YHR051W	Cox6	2.27	1.4E-03	7.4E-04
YGL187C	Cox4	2.22	2.7E-04	5.1E-04
YNL037C	Idh1	2.21	1.1E-05	1.2E-03
YBL030C	Aac3/Pet9	2.19	1.8E-05	2.4E-03
YHR087W	Rtc3	2.19	1.6E-03	3.8E-05
	Ddr48	2.16	3.1E-03	1.0E-05
YBL015W	Ach1	2.15	5.3E-06	1.6E-03
YNL104C	Leu4/Leu9	2.14	9.9E-05	9.2E-04
YDR361C	Bcp1	2.14	1.2E-03	1.0E-05
YJL217W	hypothetical protein	2.12	2.1E-05	4.2E-04
YML081C-A	Atp18	2.1	4.4E-04	8.6E-05
YPR191W	Qcr2	2.08	4.1E-04	1.7E-03
YKL085W	Mdh1	2.08	3.1E-05	5.2E-04
YPR020W	Atp20	2.07	1.1E-04	1.8E-04
YBL045C	Cor1	2.05	7.3E-05	1.4E-03
YNL055C	Por2/Por1	2.03	4.6E-05	2.0E-03
YPL078C	Atp4	2.03	1.1E-04	6.8E-04
YDL222C	Fmp45	2.03	8.9E-03	6.0E-06
YGL037C	Pnc1	2.02	2.2E-04	1.2E-04

D

Systematic name	Common name	Fold change IMX1844 vs IMX1821	P value	TIC normalized abundance IMX1821
YOR374W	Ald4	5.69	4.1E-04	1.5E-03
YDR342C	hypothetical protein	4.74	1.5E-03	2.2E-05
YNR001C	Cit2/Cit1	3.22	3.8E-06	2.4E-03
YOR388C	Fdh1	3.04	3.1E-04	2.0E-04
	hypothetical protein	3.04	3.1E-04	2.0E-04
YLR038C	Cox12	3.03	6.0E-05	2.8E-04
YNL015W	Pbi2	3.01	2.5E-04	2.7E-04
YNL104C	Leu4/Leu9	2.97	9.9E-05	9.2E-04
YJL217W	hypothetical protein	2.95	2.1E-05	4.2E-04
YKL148C	Sdh1	2.9	1.2E-04	3.6E-04
YLR295C	Atp14	2.89	1.7E-04	1.8E-04
YDR256C	Cta1	2.84	2.8E-03	4.3E-05
YER065C	Icl1	2.75	9.9E-05	1.3E-04
YOR136W	Idh2	2.74	6.4E-05	8.9E-04
YLR174W	Idp3/Idp2	2.66	1.7E-04	4.2E-04
YJR048W	Cyc7/Cyc1	2.61	1.0E-04	1.4E-03
YHL021C	Aim17	2.53	1.3E-04	2.0E-05
YBL015W	Ach1	2.52	5.3E-06	1.6E-03
YPL134C	Odc2/Odc1	2.49	1.2E-03	5.7E-06
YJR095W	Sfc1	2.47	3.7E-04	1.2E-04
YOR285W	Rdl1	2.38	7.4E-04	3.5E-04
YPR002W	Pdh1	2.35	6.7E-05	8.2E-04
YBL030C	Aac3/Pet9	2.26	1.8E-05	2.4E-03
YML054C	Cyb2	2.26	2.0E-03	5.9E-05
YJR148W	Bat2/Bat1	2.23	1.8E-03	7.3E-05
YEL024W	Rip1	2.18	1.1E-05	5.0E-04
YDL181W	Inh1/Stf1	2.18	6.5E-04	1.5E-04
YNL037C	Idh1	2.16	1.1E-05	1.2E-03
YGR234W	Yhb1	2.11	6.4E-04	4.5E-03
YLL041C	Sdh2	2.1	6.6E-05	2.2E-04
YGL187C	Cox4	2.09	2.7E-04	5.1E-04
YER057C	Mmf1/Hmf1	2.06	5.8E-04	9.2E-04
YML081C-A	Atp18	2.06	4.4E-04	8.6E-05
YER024W	Yat2	2.04	1.4E-03	1.9E-05
YBR092C	Pho3	2.04	1.8E-03	1.4E-04
YPR191W	Qcr2	2.02	4.1E-04	1.7E-03
YBL045C	Cor1	2.02	7.3E-05	1.4E-03

Figure S9 - Global proteome response to humanization of the glycolytic pathway
A-B) Volcano plots showing fold change and significance of the complete set of identified proteins of the humanized strains relative to the *Sc*Gly strain IMX1821, data based on biological triplicates, one of the duplicate injections is shown. Proteins that changed more than 2-fold with a significance above 20 (corresponding to a p value below 0.01) are highlighted red. **C-D)** Proteins that changed >2-fold are listed with common name, fold change, p-value and normalized abundance in reference strain (*Sc*Gly, IMX1821).

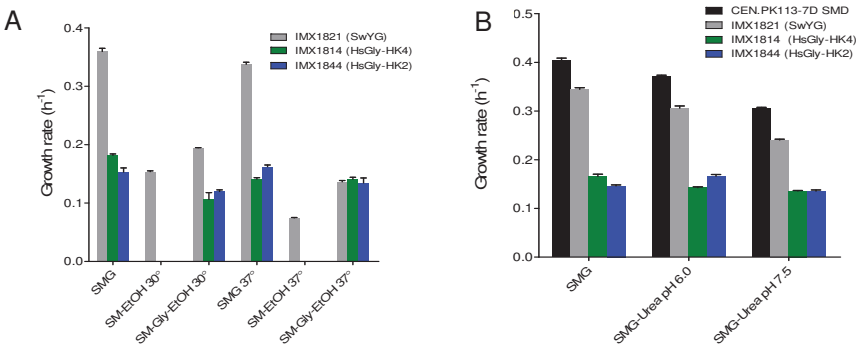


Figure S10 - Moonlighting functions in fully humanized strains.
Growth rates of the fully humanized strains IMX1844 and IMX1814 were determined in conditions where moonlighting functions of enolase and aldolase are known to play a role. Specific growth rates were determined in 96-well plates using the growth profiler, data represents average and standard deviation of independent culture triplicates. **A)** Growth rates at 30 and 37 °C. Strain IMX1814 and IMX1844 (*Hs*Gly-HK4 and *Hs*Gly-HK2) did not show growth in synthetic medium with only ethanol as carbon source. **B)** Growth rates at different media pH, acidification was prevented by use of urea instead of ammonium as nitrogen source. In both high temperature and high pH conditions no decrease in growth rate specific to the *Hs*Gly strains was observed.

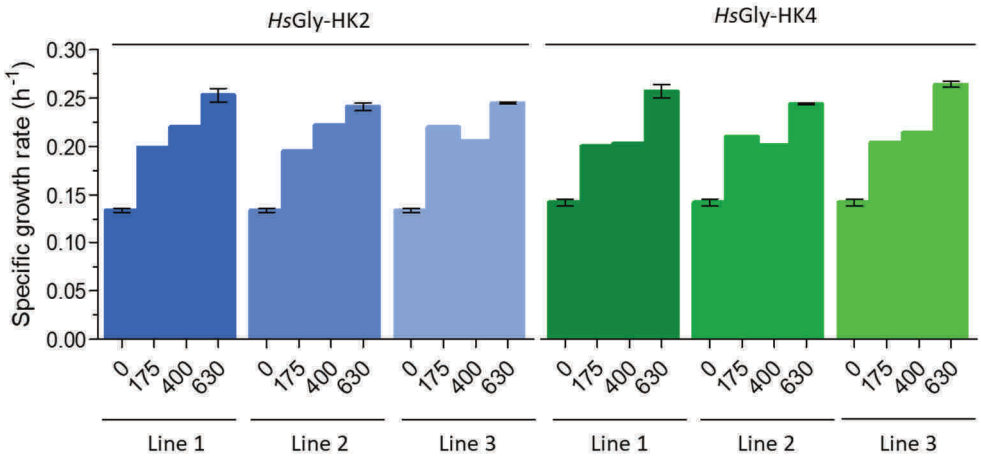


Figure S11 - Specific growth rate during evolution of IMX1844 and IMX1814
The specific growth rates were measured in shake-flask with SMG. Independent duplicates were only done for 0 and 630 generations. Measurements at 400 and 630 generations are from single colony isolates, while at 175 generations they were done with the whole evolved population.

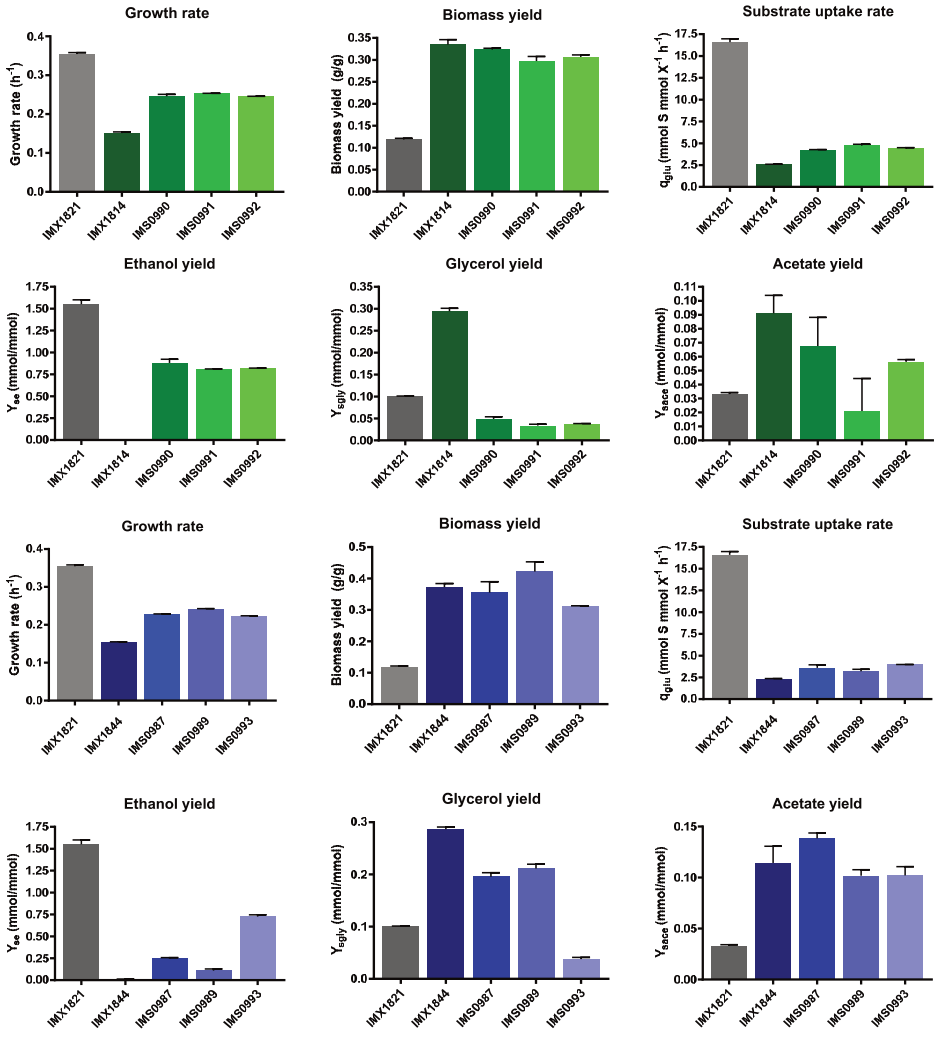


Figure S12 - Physiology of evolved, humanized glycolysis strains in shake flask
Strains were grown in SMG at 30 °C in shake flask cultures. Data represent the average and SEM of two independent culture replicates.

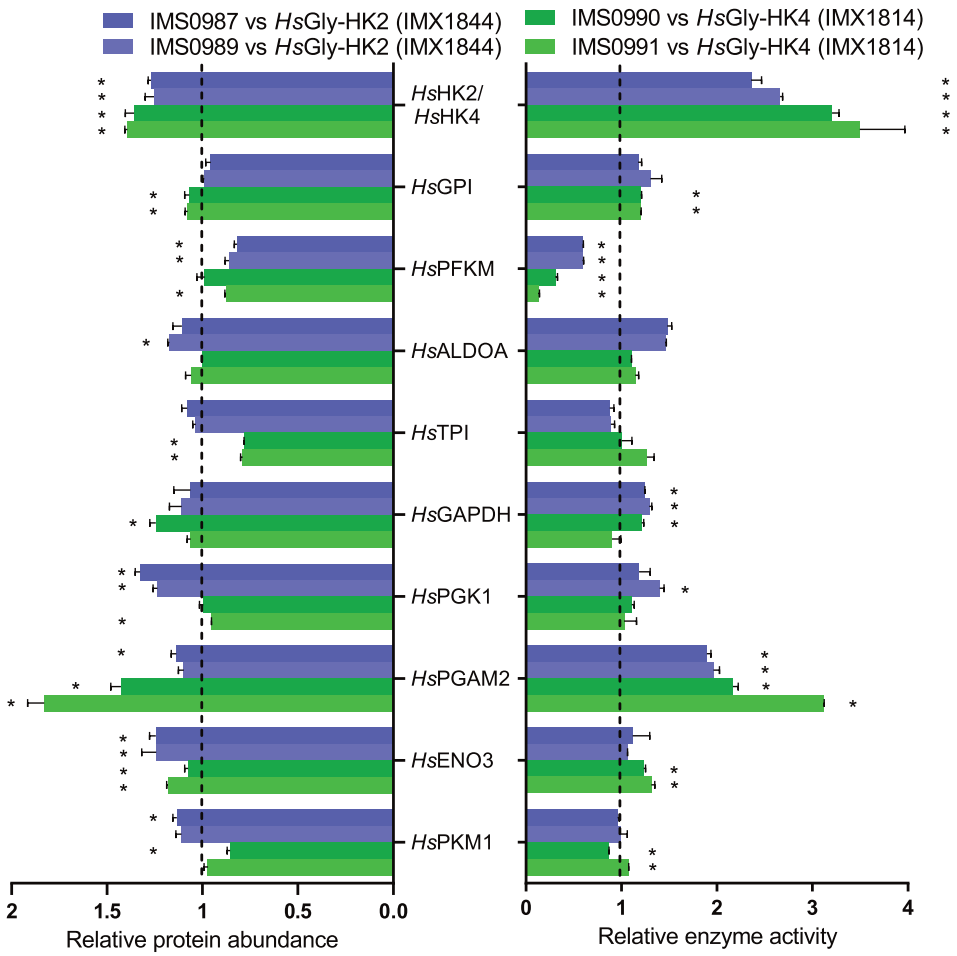


Figure S13 - Relative levels and activity of the glycolytic enzymes in evolved humanized yeast strains

Left side shows relative abundances of the human glycolytic enzymes in the evolved strains relative to unevolved strains IMX1814 and IMX1844. Means with SEM are shown for biological triplicates from one of the duplicate injections. * indicates significant difference with IMX1814/IMX1844 (p-value <0.05 student t-test, two tailed, homoscedastic). Right side shows relative enzyme activities compared to IMX1814/IMX1844. Glycolytic enzymes were assayed from shake flask cultures on SMG at 30˚C. Activities were assayed in 96-well plates. The data represent the average and mean deviation of biological culture duplicates. The asterisks indicate statistical significance (Student t-test, two-tailed, homoscedastic, P<0.05).

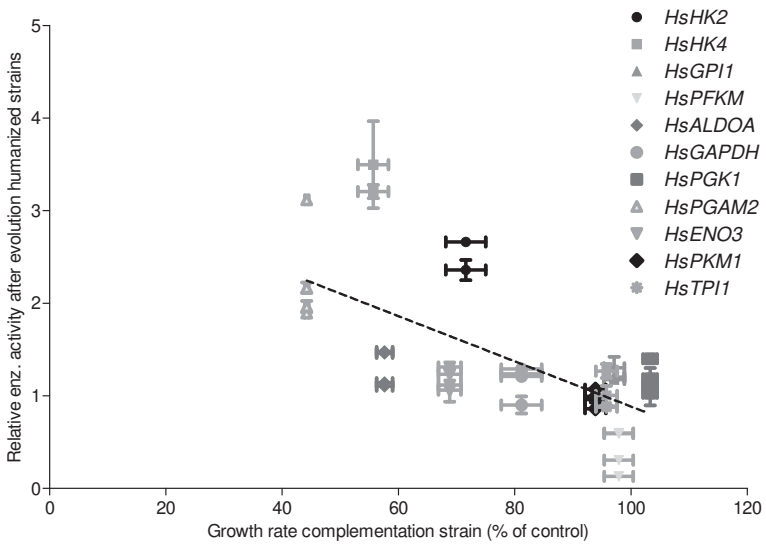
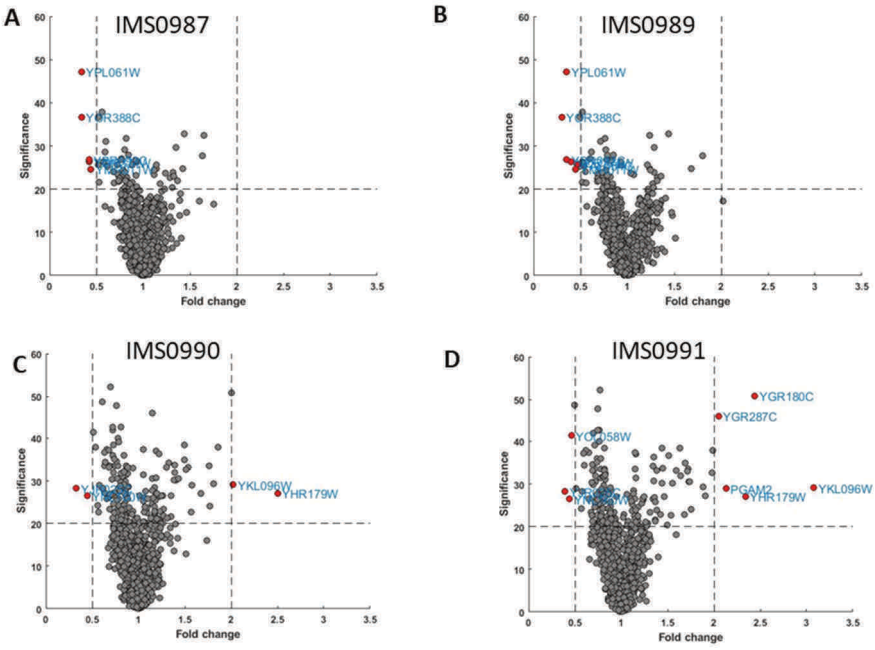


Figure S14 - Correlation between the change in enzyme activity after evolution of the fully humanized strains and growth rate of single complementation strains

The change in enzyme activity after evolution is plotted against the growth rate of the corresponding complementation strains. Since a low growth rate in the single complementation strains suggests a suboptimal expression and activity, this is expected to correlate with increased activity after evolution. This is the case for several enzymes, especially *PGAM2* and the hexokinases. A linear decreasing trendline is shown (R^2 = 0.44). For the hexokinases the values of the complementation strains with *ScHXK2* promoters were used since the same promotor was used in the fully humanized strains



E

Systematic name	Common name	Fold change vs IMX1844	P value	TIC normalized abundance IMX1844
IMS0987				
YMR011W	hypothetical protein	0.44	3.5E-03	5.5E-04
YOR374W	Ald4	0.42	2.4E-03	1.8E-02
YBR092C	Pho3	0.42	2.1E-03	4.3E-04
YPL061W	Ald6	0.34	1.9E-05	1.9E-02
YOR388C	Fdh1	0.34	2.2E-04	1.2E-03
IMS0989				
YAL054C	Acs1	0.46	2.8E-03	1.3E-03
YMR011W	hypothetical protein	0.45	3.5E-03	5.5E-04
YOR374W	Ald4	0.4	2.4E-03	1.8E-02
YPL061W	Ald6	0.35	1.9E-05	1.9E-02
YBR092C	Pho3	0.35	2.1E-03	4.3E-04
YOR388C	Fdh1	0.3	2.2E-04	1.2E-03

F

Systematic name	Common name	Fold change vs IMX1814	P value	TIC normalized abundance IMX1814
IMS0990				
YHR179W	Oye2	2.5	2.0E-03	2.6E-03
YKL096W	Cwp2	2.02	1.2E-03	2.7E-05
YNL160W	Ygp1/Sps100	0.45	2.2E-03	4.8E-04
YJR025C	Bna1	0.33	1.5E-03	5.8E-05
IMS0991				
YKL096W	Cwp2	3.08	1.2E-03	2.7E-05
YGR180C	Rnr4/Rnr2	2.44	8.3E-06	1.0E-03
YHR179W	Oye2	2.34	2.0E-03	2.6E-03
	PGAM2	2.13	1.3E-03	4.0E-04
YGR287C	Ima1	2.05	2.4E-05	4.0E-04
YOL058W	Arg1	0.46	7.4E-05	2.1E-03
YNL160W	Ygp1/Sps100	0.44	2.2E-03	4.8E-04
YJR025C	Bna1	0.39	1.5E-03	5.8E-05

Figure S15 - Global proteome response to evolution of humanized strains

A) - D) Volcano plots showing fold change and significance of the complete set of identified proteins of the evolved strains relative to their parental strain. Proteins that changed more than 2-fold with a significance above 20 (corresponding to a p value of 0.01) are indicated. E-F) Proteins changed > 2-fold in abundance after evolution are indicated with common name, fold change, p-value and normalized abundance in parental strain

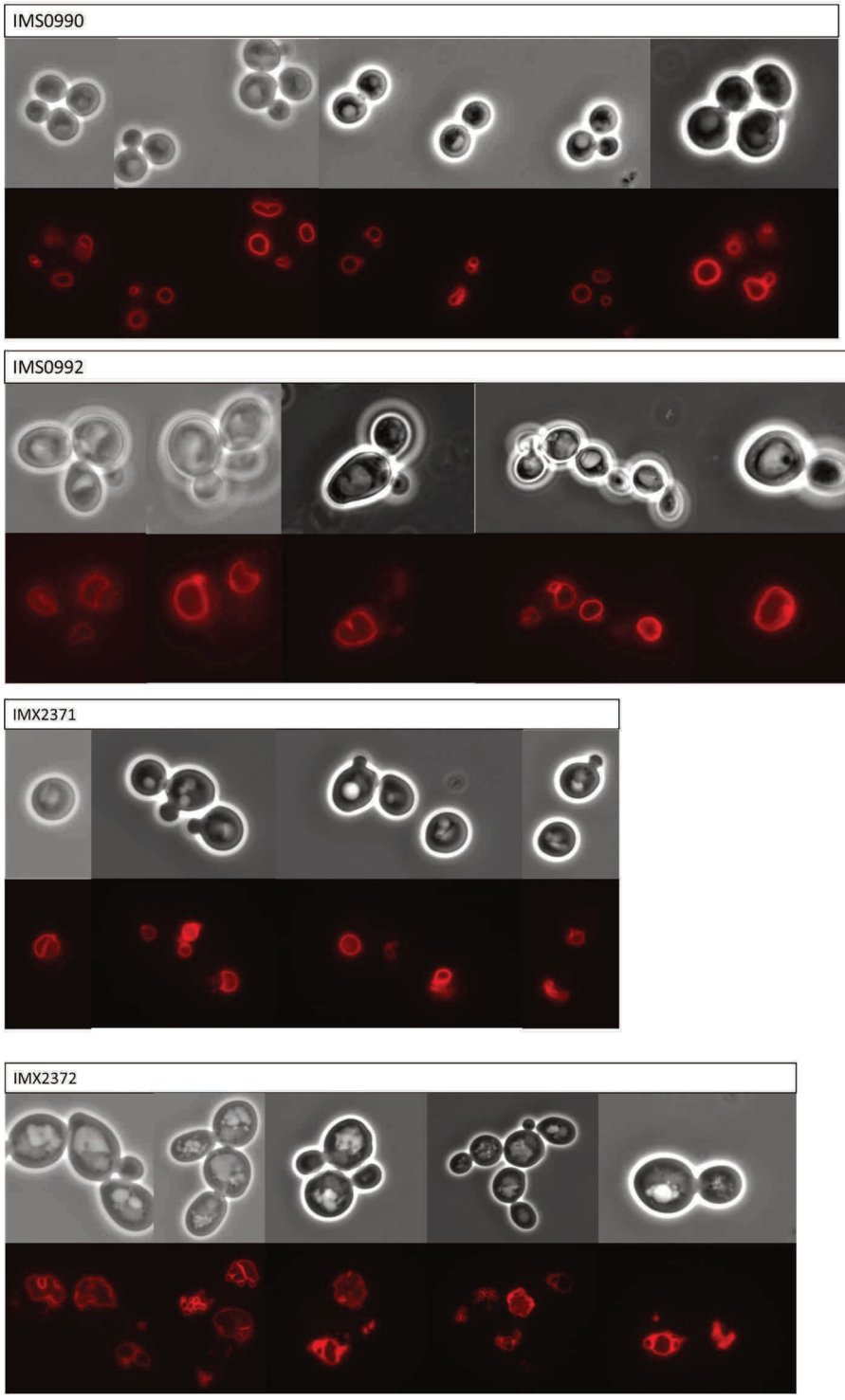

IMS0990
IMS0992
IMX2371
IMX2372

Figure S16 - Vacuole morphology in evolved humanized yeast strain
Cells were stained with the red fluorescent dye FM4-64 and visualized with an Imager-Z1 microscope (Carl-Zeiss). Strain IMS0990 and IMS0992 are *Hs*Gly evolved strains and IMX2371 and IMX2372 are the non-evolved, genetically engineered variants in which the *STT4* gene was mutated.

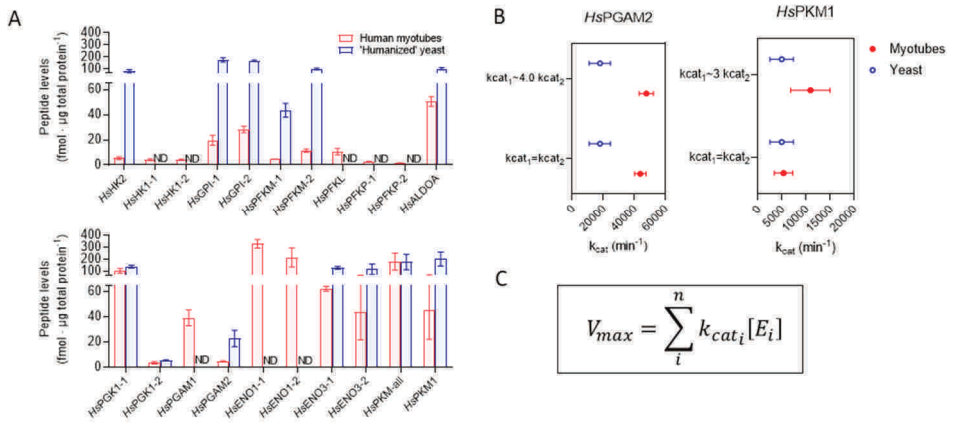
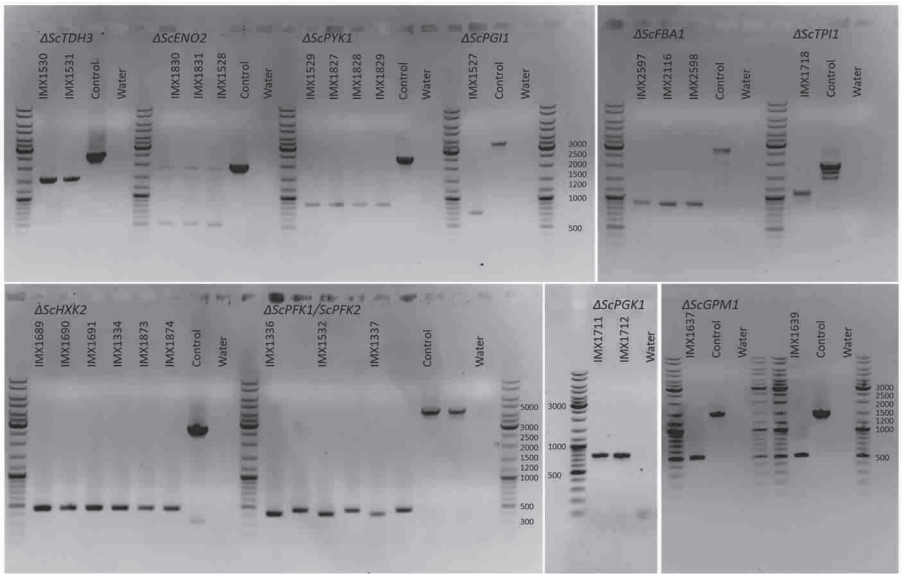


Figure S17 - Peptide abundance
A) Peptide abundance in fmol·µg prot^{-1} for the human glycolytic enzymes in yeast (*Hs*Gly-HK2 strain IMX1844, blue) and myotube (red) cell extracts. Peptides identified in the targeted proteomics analysis are represented on the x-axis. When more than one peptide was quantified per protein, they are indicated by -1 or -2 after the protein name (see Table S6 for the peptide sequences). For the myotubes, three independent cultures were averaged, for yeast, two independent cultures were averaged. No absolute quantification could be made for TPI and GAPDH due to the lack of standard peptides. **B)** Estimated k_{cat}'s for *Hs*PKM1 and *Hs*PGAM2 with different assumptions, k_{cat} ratios for isozymes derived from the literature. Data show means ± SD. **C)** Equation used to estimate k_{cat}, $[E_i]$ represents the concentration of each isoform, number of subunits per enzyme complex not included.



Yeast gene	Deletion (bp)	No deletion (bp)
ENO2	520	1839
PYK1	860	2373
TDH3	1530	2534
PGI1	710	3453
PGK1	755	2362
GPM1	512	1481
FBA1	865	2714
HXK2	460	2653
TPI1	1090	1840
PFK1	415	4418
PFK2	375	4302

Figure S18 - Diagnostic PCR of single complementation strains
PCR confirmation of yeast gene deletions resulting in the individual human gene complementation strains which are indicated per well. The forward and reverse primer were chosen upstream and downstream of the region of gene deletion. The control shows the PCR product resulting from a strain (IMX1076) in which the gene is not deleted. In case of the *ENO2* deletion the primers were binding in the promoter and terminator which resulted in amplification of the human gene as well.

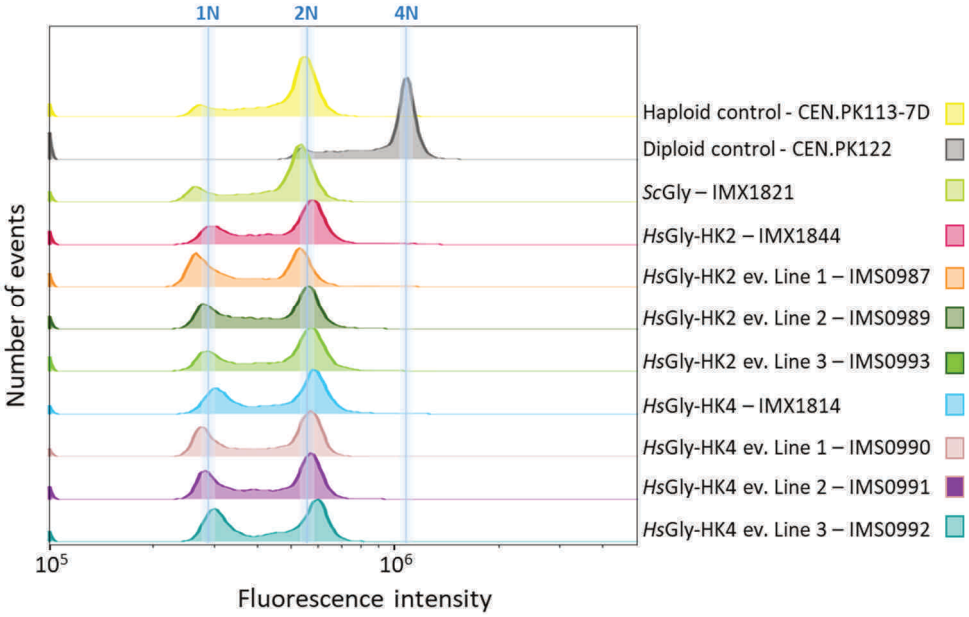


Figure S19 - Ploidy analysis of the strains used in this study
Ploidy measurement based on DNA content determination by flow cytometry. The two top plots represent the haploid and diploid controls. All humanized strains are haploids as expected.

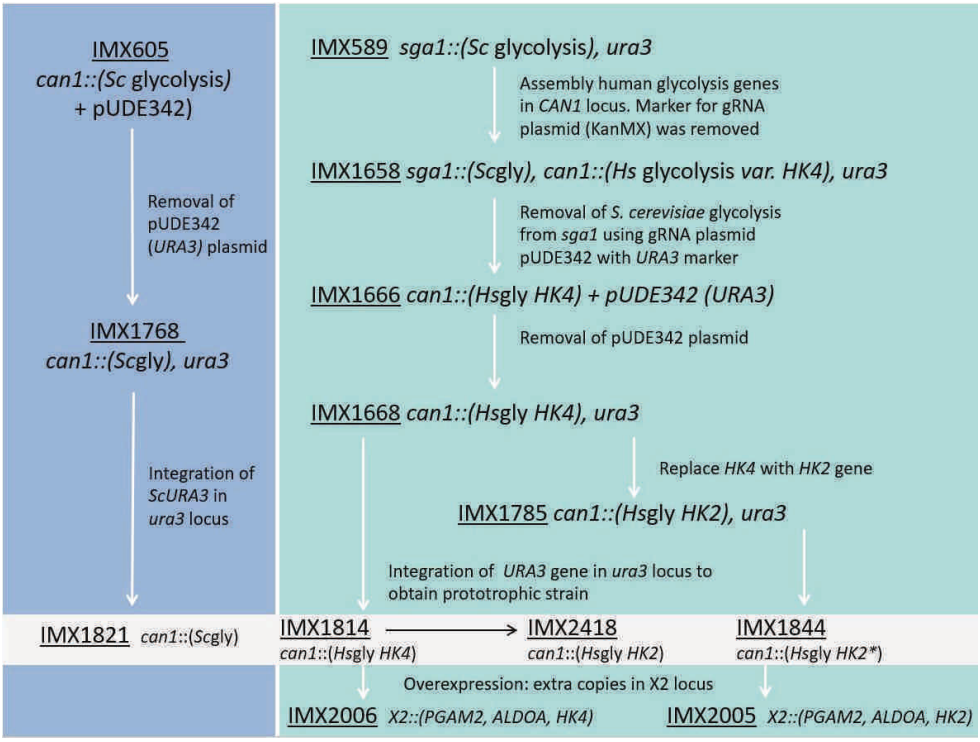


Figure S20 - Overview of single locus glycolysis strain construction.
Construction of IMX605 and IMX589 is described in Kuijpers *et al.* 2016.

Table S1 - Comparison of the human glycolytic proteins to their yeast orthologs.

Human enzyme	Size (aa)	Closest yeast enzyme	Size (aa)	% Identity at protein level	Complementation	Shown previously?
HK1	917	Hxk2	486	30% to subunit 1, 35% to subunit 2	Yes, with mutation	No
HK2	917	Hxk2	486	33% to both subunits	Yes, with mutation	No
HK3	923	Hxk2	486	28% to subunit 1 33% to subunit 2	No	No
HK4	465	Hxk2	486	31%	Yes	Yes, [5]
GPI1	558	Pgi1	554	58%	Yes	No
PFKM	780	Pfk1, Pfk2	987/959	43% to PFK1, 43% to PFK2	Yes	Yes, [6]
PFKP	784	Pfk1, Pfk2	987/959	43% to PFK1, 43% to PFK2	Yes	No
PFKL	780	PFK1, PFK2	987/959	43% to PFK1, 45% to PFK2	Yes	No
ALDOA	364	Fba1	359	-	Yes	No
ALDOB	364	Fba1	359	-	Yes	Yes, [7]
ALDOC	364	Fba1	359	-	Yes	No
TPI	286	Tpi1	248	53%	Yes	Yes
GAPDH	335	Tdh3	332	65%	Yes	No
GAPDHS	408	Tdh3	332	65%	Yes	No

PGK1	417	Pgk1	416	66%	Yes	Yes, [8]
PGK2	417	Pgk1	416	66%	Yes	Yes, [8]
PGAM1	253	Gpm1	247	51%	Yes	Yes, [8]
PGAM2	253	Gpm1	247	52%	Yes	Tested, but negative, [8]
ENO1	434	Eno2	437	63%	Yes	No
ENO2	434	Eno2	437	62%	Yes	No
ENO3	434	Eno2	437	63%	Yes	No
PKM1*	531	Pyk1	500	52%	Yes	Yes, [8]
PKM2*	531	Pyk1	500	52%	Yes	No
PKR#	574	Pyk1	500	50%	Yes	Yes, [8]
PKL#	553	Pyk1	500	50%	Yes	Yes, [8]

*,# Splicing variants

Table S2 - Genetic composition of the glycolytic transcriptional units used for the single complementation strains and strains with fully humanized glycolysis.

Human gene	Yeast prom	Yeast term	Yeast gene	Yeast prom	Yeast term
HsHK1	*PDC1*	*PDC1*	*ScHXK2*	*HXK2*	*HXK2*
HsHK2#	*PDC1*	*PDC1*			
HsHK2#	*HXK2*	*HXK2*			
HsHK3	*PDC1*	*PDC1*			
HsHK4#	*PDC1*	*PDC1*			
HsHK4#	*HXK2*	*HXK2*			
HsGPI	*TEF2*	*TEF2*	*ScPGI1*	*PGI1*	*PGI1*
HsPFKM	*TEF1*	*TEF1*	*ScPFK1*	*ScPFK1*	*ScPFK1*
HsPFKP	*TEF1*	*TEF1*	*ScPFK2*	*ScPFK2*	*ScPFK2*
HsPFKL	*TEF1*	*TEF1*			
HsALDOA	*FBA1*	*FBA1*	*ScFBA1*	*FBA1*	*FBA1*
HsALDOB	*FBA1*	*FBA1*			
HsALDOC	*FBA1*	*FBA1*			
HsTPI	*TPI1*	*TPI1*	*ScTPI1*	*TPI1*	*TPI1*
HsGAPDH	*TDH3*	*TDH3*	*ScTDH3*	*TDH3*	*TDH3*
HsGAPDHS	*TDH3*	*TDH3*			
HsPGK1	*PGK1*	*PGK1*	*ScPGK1*	*PGK1*	*PGK1*
HsPGK2	*PGK1*	*PGK1*			
HsPGAM1	*GPM1*	*GPM1*	*ScGPM1*	*GPM1*	*GPM1*
HsPGAM2	*GPM1*	*GPM1*			
HsENO1	*ENO2*	*ENO2*	*ScENO2*	*ENO2*	*ENO2*
HsENO2	*ENO2*	*ENO2*			
HsENO3	*ENO2*	*ENO2*			
HsPKM1	*PYK1*	*PYK1*	*ScPYK1*	*PYK1*	*PYK1*
HsPKM2	*PYK1*	*PYK1*			
HsPKR	*PYK1*	*PYK1*			
HsPKL	*PYK1*	*PYK1*			

For hexokinase expression in full human glycolysis strains *HXK2* promoter and terminator was used and corresponding complementation strains with the same promotor and terminator were constructed.

Table S3 - Whole genome sequence analysis of the strains with fully humanized glycolysis.

Systematic name	Name	Type	Amino acid change
Mutations in IMX1666 (human glycolysis var. *HK4*) compared to IMX589 (SwYG)			
YDR351W	SBE2	NSY	Ser-412-Thr
YFL064C	uncharacterized protein	NSY	Asn-66-Asp
YHR219W	uncharacterized protein	NSY	Val-383-Leu
YJL223C	PAU1	SYN	Ala-13-Ala
	*Hs*PFKM	SYN	Val-529-Val
-	tPDC1/SHR BC	4 bp gap	Last 2 bp missing and 2 bp from SHR BC
Mutations in IMX1844 (human glycolysis var. HK2) relative to IMX1666			
YJL225C	uncharacterized protein	NSY	Arg-267-Ser
-	*Hs*HK2	NSY	Ile-562-Asn
Mutations in IMX2005 relative to IMX1844			
YEL074W	Putative protein	NSY	His-66-Pro
No mutations in IMX2006 relative to IMX1666			

Table S4 - Physiology of humanized glycolysis strains in bioreactors.
Yields and biomass specific conversion rates of IMX1844 (*Hs*Gly-*HK2*), IMX1814 (*Hs*Gly-*HK4*) and the control strain IMX1821 (*Sc*Gly). Strains were grown in bioreactor at 30°C under aerobic conditions in SMG. Maximum specific growth rate (μ_{max}), biomass specific rates of glucose and oxygen consumption (q_{glu} and q_{O2}, respectively), biomass specific rates of production of ethanol (q_{etoh}), glycerol (q_{gly}), acetate (q_{ace}) and carbon dioxide (q_{CO2}) and RQ: respiration quotient (q_{CO2}/q_{O2}). Yields on glucose for biomass (Y_{sx}), ethanol (Y_{setoh}), glycerol (Y_{sgly}), and acetate (Y_{sace}). Values show the average and standard error of the mean from at least two independent bioreactors per strain. DW: biomass dry weight. BDL: below detection level.

	*Sc*Gly IMX1821	HsGly-HK2 IMX1844	HsGly-HK4 IMX1814
Biomass specific rates			
μ_{max} (h^{-1})	0.32 ± 0.01	0.15 ± 0.00	0.15 ± 0.00
$-q_{glu}$ ($mmol.g_{DW}^{-1}.h^{-1}$)	16.1 ± 0.8	2.2 ± 0.03	3.0 ± 0.44
q_{etoh} ($mmol.g_{DW}^{-1}.h^{-1}$)	22.3 ± 0.8	0.03 ± 0.00	0.25 ± 0.07
q_{gly} ($mmol.g_{DW}^{-1}.h^{-1}$)	1.6 ± 0.1	0.11 ± 0.00	0.77 ± 0.15
q_{ace} ($mmol.g_{DW}^{-1}.h^{-1}$)	0.59 ± 0.06	BDL	0.58 ± 0.05
Respiration			
q_{CO2} ($mmol.g_{DW}^{-1}.h^{-1}$)	29.9 ± 1.6	6.8 ± 0.21	6.2 ± 0.27
q_{O2} ($mmol.g_{DW}^{-1}.h^{-1}$)	7.9 ± 0.4	6.7 ± 0.31	5.4 ± 0.25
RQ	3.8 ± 0.01	1.0 ± 0.01	0.87 ± 0.00
Yields			
Y_{sx} (g_{DW}/g_{glu})	0.11 ± 0.01	0.37 ± 0.01	0.28 ± 0.04
Y_{setoh} ($mol_{ethanol}/mol_{glu}$)	1.4 ± 0.0	0.01 ± 0.00	0.08 ± 0.01
Y_{sgly} ($mol_{glycerol}/mol_{glu}$)	0.10 ± 0.00	0.05 ± 0.00	0.25 ± 0.01
Y_{sace} ($mol_{acetate}/mol_{glu}$)	0.04 ± 0.00	BDL	0.19 ± 0.01

Table S5 - Mutations in the coding regions of evolved strains.
Text in grey italics indicates synonymous mutation. Dark grey background indicates the mutations common to all six evolved strains and light grey background the mutations common to all evolution lines from a single humanized strain.

		IMS0990		IMS0991		IMS0992	
Evolved strains with IMX1814 (HsGly-HK4) background	TUP1	G to A	His-489-Tyr	G to A	His-489-Tyr	G to A	His-489-Tyr
	STT4	C to T	Gly-1766-Arg	C to G	Arg-1707-Pro	A to T	Phe-1775-Ile
	PFKM	A to G	Thr-81-Ala	G to C	Arg-623-Ser	G to T	Arg-673-Ile
	NUT1			T to G	Tyr-432-stop		
	RSM22	G to A	Arg-120-Cys				
	YMR089C	C to A	Pro-217-Thr	G to C	Asp-430-His		
	YCR038W	*G to A*	*Leu-123-Leu*				
	YLR371W	*A to T*	*Ala-74-Ala*				
	YKL124W			*G to A*	*Thr-346-Thr*		
	NMa111	*G to A*	*Glu-40-Glu*	*G to A*	*Glu-40-Glu*	*G to A*	*Glu-40-Glu*

		IMS0987		IMS0989		IMS0993	
Evolved strains with IMX1844 (Hx-Gly-HK2) background	STT4	C to A	Asp-1650-Tyr	A to G	Ile-1771-Thr	G to C	Ser-1611-Cys
	PIN4	A to T	Phe-130-Leu	G to T	Gln-482-stop		
	SNF1					T to G	Phe-261-Cys
	SKI8					G to A	Pro-219-Ser
	TAO3					G to T	His-167-Asn
	CYR1	G to T	Gly-1612-Val	G to A	Gly-1768-Ser		
	ECM38					T to G	Tyr-650-Asp
	YOL075C			G to C	Pro-246-Ala		
	YME1					G to T	STOP-748-Leu
	YOR114W	*C to A*	*Pro-146-Pro*				

Tables S6 – S9 and Appendix 1.
Supplementary tables S6-S9 as well as appendix 1 on mathematical modelling can be found online at:

https://www.biorxiv.org/content/10.1101/2021.09.28.462164v1.supplementary-material

References

1. Nurse, P.M., Nobel lecture: cyclin dependent kinases and cell cycle control. Bioscience reports, 2002. **22**(5-6): p. 487-499.
2. Woolford, J.L. and S.J. Baserga, Ribosome biogenesis in the yeast *Saccharomyces cerevisiae.* Genetics, 2013. **195**(3): p. 643-681.
3. Petranovic, D. and J. Nielsen, Can yeast systems biology contribute to the understanding of human disease? Trends in biotechnology, 2008. **26**(11): p. 584-590.
4. Laurent, J.M., et al., Efforts to make and apply humanized yeast. Briefings in Functional Genomics, 2016. **15**(2): p. 155-163.DOI: 10.1093/bfgp/elv041.
5. Garge, R.K., et al., Systematic humanization of the yeast cytoskeleton discerns functionally replaceable from divergent human genes. Genetics, 2020.
6. Hamza, A., et al., Cross-species complementation of nonessential yeast genes establishes platforms for testing inhibitors of human proteins. Genetics, 2020.DOI: 10.1534/genetics.119.302971.
7. Hamza, A., et al., Complementation of yeast genes with human genes as an experimental platform for functional testing of human genetic variants. Genetics, 2015. **201**(3): p. 1263-74.DOI: 10.1534/genetics.115.181099.
8. Kachroo, A.H., et al., Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science, 2015. **348**(6237): p. 921-925.DOI: 10.1126/science.aaa0769.
9. Laurent, J.M., et al., Humanization of yeast genes with multiple human orthologs reveals functional divergence between paralogs. PLoS Biology, 2020. **18**(5): p. e3000627.
10. Sun, S., et al., An extended set of yeast-based functional assays accurately identifies human disease mutations. Genome Res, 2016. **26**(5): p. 670-80.DOI: 10.1101/gr.192526.115.
11. Zhang, N., et al., Using yeast to place human genes in functional categories. Gene, 2003. **303**: p. 121-129.
12. Prince, V.E. and F.B. Pickett, Splitting pairs: the diverging fates of duplicated genes. Nature Reviews Genetics, 2002. **3**(11): p. 827-837.
13. Steinke, D., et al., Three rounds (1R/2R/3R) of genome duplications and the evolution of the glycolytic pathway in vertebrates. BMC biology, 2006. **4**(1): p. 16.
14. Gordon, J.L., K.P. Byrne, and K.H. Wolfe, Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. PLoS genetics, 2009. **5**(5).
15. Kachroo, A.H., et al., Systematic bacterialization of yeast genes identifies a near-universally swappable pathway. Elife, 2017. **6**.DOI: 10.7554/eLife.25093.
16. Agmon, N., et al., Phylogenetic debugging of a complete human biosynthetic pathway transplanted into yeast. Nucleic Acids Res, 2020. **48**(1): p. 486-499.DOI: 10.1093/nar/gkz1098.
17. Hamilton, S.R. and D. Zha, Progress in yeast glycosylation engineering. Methods Mol. Biol, 2015. **1321**: p. 73-90.DOI: 10.1007/978-1-4939-2760-9_6.
18. Labunskyy, V.M., et al., The insertion Green Monster (iGM) method for expression of multiple exogenous genes in yeast. G3 (Bethesda), 2014. **4**(7): p. 1183-91.DOI: 10.1534/g3.114.010868.
19. Truong, D.M. and J.D. Boeke, Resetting the yeast epigenome with human nucleosomes. Cell, 2017. **171**(7): p. 1508-1519 e13.DOI: 10.1016/j.cell.2017.10.043.
20. Warburg, O., The metabolism of carcinoma cells. The Journal of Cancer Research, 1925. **9**(1): p. 148-163.

21. Mans, R., J.-M.G. Daran, and J.T. Pronk, Under pressure: evolutionary engineering of yeast strains for improved performance in fuels and chemicals production. Current opinion in biotechnology, 2018. **50**: p. 47-56.
22. Heinisch, J.J., Expression of heterologous phosphofructokinase genes in yeast. FEBS letters, 1993. **328**(1-2): p. 35-40.
23. Lu, M., et al., Physical interaction between aldolase and vacuolar H+-ATPase is essential for the assembly and activity of the proton pump. J Biol Chem, 2007. **282**(34): p. 24495-503.DOI: 10.1074/jbc.M702598200.
24. Mayordomo, I. and P. Sanz, Human pancreatic glucokinase (GlkB) complements the glucose signalling defect of *Saccharomyces cerevisiae hxk2* mutants. Yeast, 2001. **18**(14): p. 1309-16.DOI: 10.1002/yea.780.
25. Sriram, G., et al., Single-gene disorders: what role could moonlighting enzymes play? American Journal of Human Genetics, 2005. **76**(6): p. 911-24.DOI: 10.1086/430799.
26. Gancedo, C. and C.L. Flores, Moonlighting proteins in yeasts. Microbiol Mol Biol Rev, 2008. **72**(1): p. 197-210, table of contents.DOI: 10.1128/MMBR.00036-07.
27. Solis-Escalante, D., et al., A minimal set of glycolytic genes reveals strong redundancies in *Saccharomyces cerevisiae* central metabolism. Eukaryotic Cell, 2015. **14**(8): p. 804-816.DOI: 10.1128/EC.00064-15.
28. Kuijpers, N.G., et al., Pathway swapping: Toward modular engineering of essential cellular processes. Proceedings of the National Academy of Sciences, USA, 2016. **113**(52): p. 15060-15065.DOI: 10.1073/pnas.1606701113.
29. Marsh, J.J. and H.G. Lebherz, Fructose-bisphosphate aldolases: an evolutionary history. Trends Biochem Sci, 1992. **17**(3): p. 110-3.DOI: 10.1016/0968-0004(92)90247-7.
30. Cárdenas, M.a.L., A. Cornish-Bowden, and T. Ureta, Evolution and regulatory role of the hexokinases. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1998. **1401**(3): p. 242-264.
31. Israelsen, W.J. and M.G. Vander Heiden, Pyruvate kinase: Function, regulation and role in cancer. Semin Cell Dev Biol, 2015. **43**: p. 43-51.DOI: 10.1016/j.semcdb.2015.08.004.
32. Blazquez, M.A., et al., Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. FEBS Letters, 1993. **329**(1-2): p. 51-54.
33. Wilson, J.E., Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. J Exp Biol, 2003. **206**(Pt 12): p. 2049-57.DOI: 10.1242/jeb.00241.
34. van Eunen, K., et al., Testing biochemistry revisited: how *in vivo* metabolism can be understood from *in vitro* enzyme kinetics. PLoS Comput Biol, 2012. **8**(4): p. e1002483.
35. Lambeth, M.J. and M.J. Kushmerick, A computational model for glycogenolysis in skeletal muscle. Ann Biomed Eng, 2002. **30**(6): p. 808-27.DOI: 10.1114/1.1492813.
36. Hohmann, S., et al., Evidence for trehalose-6-phosphate-dependent and-independent mechanisms in the control of sugar influx into yeast glycolysis. Molecular microbiology, 1996. **20**(5): p. 981-991.
37. Teusink, B., et al., Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem, 2000. **267**(17): p. 5313-29.DOI: 10.1046/j.1432-1327.2000.01527.x.
38. Nawaz, M.H., et al., The catalytic inactivation of the N-half of human hexokinase 2 and structural and biochemical characterization of its mitochondrial conformation. Biosci Rep, 2018. **38**(1).DOI: 10.1042/bsr20171666.
39. Ardehali, H., et al., Functional organization of mammalian hexokinase II. Retention of catalytic and regulatory functions in both the NH2- and COOH-terminal halves. J Biol Chem, 1996. **271**(4): p. 1849-52.DOI: 10.1074/jbc.271.4.1849.

40. Tsai, H.J. and J.E. Wilson, Functional organization of mammalian hexokinases: both N- and C-terminal halves of the rat type II isozyme possess catalytic sites. Arch Biochem Biophys, 1996. **329**(1): p. 17-23.DOI: 10.1006/abbi.1996.0186.
41. Lagunas, R. and C. Gancedo, Role of phosphate in the regulation of the Pasteur effect in *Saccharomyces cerevisiae*. European journal of biochemistry, 1983. **137**(3): p. 479-483.
42. Aleshin, A.E., et al., The mechanism of regulation of hexokinase: new insights from the crystal structure of recombinant human brain hexokinase complexed with glucose and glucose-6-phosphate. Structure, 1998. **6**(1): p. 39-50.
43. Teusink, B., et al., The danger of metabolic pathways with turbo design. Trends in Biochemical Sciences, 1998. **23**(5): p. 162-169.
44. van Heerden, J.H., et al., Lost in transition: startup of glycolysis yields subpopulations of nongrowing cells. Science, 2014.DOI: science.1245114 [pii];10.1126/science.1245114 [doi].
45. Crowther, G.J., et al., Control of glycolysis in contracting skeletal muscle. II. Turning it off. American Journal of Physiology-Endocrinology And Metabolism, 2002. **282**(1): p. E74-E79.
46. Crowther, G.J., et al., Control of glycolysis in contracting skeletal muscle. I. Turning it on. American Journal of Physiology-Endocrinology And Metabolism, 2002. **282**(1): p. E67-E73.
47. van Hall, G., Lactate kinetics in human tissues at rest and during exercise. Acta physiologica, 2010. **199**(4): p. 499-508.
48. Grossbard, L. and R.T. Schimke, Multiple Hexokinases of Rat Tissues PURIFICATION AND COMPARISON OF SOLUBLE FORMS. Journal of Biological Chemistry, 1966. **241**(15): p. 3546-3560.
49. Ritov, V.B. and D.E. Kelley, Hexokinase isozyme distribution in human skeletal muscle. Diabetes, 2001. **50**(6): p. 1253-1262.
50. Bell, G.I., et al., Glucokinase mutations, insulin secretion, and diabetes mellitus. Annu Rev Physiol, 1996. **58**: p. 171-86.DOI: 10.1146/annurev.ph.58.030196.001131.
51. Matschinsky, F.M., Glucokinase as glucose sensor and metabolic signal generator in pancreatic beta-cells and hepatocytes. Diabetes, 1990. **39**(6): p. 647-52.DOI: 10.2337/diab.39.6.647.
52. Boles, E., et al., Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in *Saccharomyces cerevisiae* that is catalytically insensitive to fructose-1,6-bisphosphate. J. Bacteriol, 1997. **179**(9): p. 2987-2993.
53. Diaz-Ruiz, R., et al., Tumor cell energy metabolism and its common features with yeast metabolism. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2009. **1796**(2): p. 252-265.
54. Van den Brink, J., et al., Dynamics of glycolytic regulation during adaptation of *Saccharomyces cerevisiae* to fermentative metabolism. Applied and environmental microbiology, 2008. **74**(18): p. 5710-5723.
55. Tai, S.L., et al., Control of the glycolytic flux in *Saccharomyces cerevisiae* grown at low temperature - A multi-level analysis in anaerobic chemostat cultures. Journal of Biological Chemistry, 2007. **282**(14): p. 10243-10251.DOI: 10.1074/jbc.M610845200.
56. Jansen, M.L.A., et al., Prolonged selection in aerobic, glucose-limited chemostat cultures of *Saccharomyces cerevisiae* causes a partial loss of glycolytic capacity. Microbiology (Reading), 2005. **151**(Pt 5): p. 1657-1669.DOI: 10.1099/mic.0.27577-0.
57. Fernandez-Garcia, P., et al., Phosphorylation of yeast hexokinase 2 regulates its nucleocytoplasmic shuttling. J Biol Chem, 2012. **287**(50): p. 42151-64.DOI: 10.1074/jbc.M112.401679.

58. Herrero, P., C. Martínez-Campa, and F. Moreno, The hexokinase 2 protein participates in regulatory DNA-protein complexes necessary for glucose repression of the *SUC2* gene in *Saccharomyces cerevisiae.* FEBS letters, 1998. **434**(1-2): p. 71-76.
59. Elbing, K., et al., Role of hexose transport in control of glycolytic flux in *Saccharomyces cerevisiae.* Applied and Environmental Microbiology, 2004. **70**(9): p. 5323-5330.
60. Van Dijken, J., et al., An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. Enzyme and Microbial Technology, 2000. **26**(9): p. 706-714.
61. Van Hoek, P., J.P. Van Dijken, and J.T. Pronk, Effect of specific growth rate on fermentative capacity of baker's yeast. Appl. Environ. Microbiol., 1998. **64**(11): p. 4226-4233.
62. Lu, M., et al., The glycolytic enzyme aldolase mediates assembly, expression, and activity of vacuolar H+-ATPase. J Biol Chem, 2004. **279**(10): p. 8732-9.DOI: 10.1074/jbc.M303871200.
63. Entelis, N., et al., A glycolytic enzyme, enolase, is recruited as a cofactor of tRNA targeting toward mitochondria in *Saccharomyces cerevisiae.* Genes & development, 2006. **20**(12): p. 1609-1620.
64. Schneider, A., Mitochondrial tRNA import and its consequences for mitochondrial translation. Annu Rev Biochem, 2011. **80**: p. 1033-53.DOI: 10.1146/annurev-biochem-060109-092838.
65. Entelis, N.S., et al., RNA delivery into mitochondria. Adv Drug Deliv Rev, 2001. **49**(1-2): p. 199-215.DOI: 10.1016/s0169-409x(01)00135-1.
66. Kolesnikova, O.A., et al., Nuclear DNA-encoded tRNAs targeted into mitochondria can rescue a mitochondrial DNA mutation associated with the MERRF syndrome in cultured human cells. Hum Mol Genet, 2004. **13**(20): p. 2519-34.DOI: 10.1093/hmg/ddh267.
67. Baleva, M., et al., A Moonlighting Human Protein Is Involved in Mitochondrial Import of tRNA. International Journal of Molecular Sciences, 2015. **16**(5): p. 9354-9367.DOI: 10.3390/ijms16059354.
68. Decker, B.L. and W.T. Wickner, Enolase activates homotypic vacuole fusion and protein transport to the vacuole in yeast. Journal of Biological Chemistry, 2006. **281**(20): p. 14523-14528.
69. Sandberg, T.E., et al., The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. Metab Eng, 2019. **56**: p. 1-16.DOI: 10.1016/j.ymben.2019.08.004.
70. Sola-Penna, M., et al., Regulation of mammalian muscle type 6-phosphofructo-1-kinase and its implication for the control of the metabolism. IUBMB life, 2010. **62**(11): p. 791-796.DOI: 10.1002/iub.393.
71. Banaszak, K., et al., The crystal structures of eukaryotic phosphofructokinases from baker's yeast and rabbit skeletal muscle. J Mol Biol, 2011. **407**(2): p. 284-97.DOI: 10.1016/j.jmb.2011.01.019.
72. Trotter, P.J., et al., A genetic screen for aminophospholipid transport mutants identifies the phosphatidylinositol 4-kinase, STT4p, as an essential component in phosphatidylserine metabolism. J Biol Chem, 1998. **273**(21): p. 13189-96.DOI: 10.1074/jbc.273.21.13189.
73. Audhya, A., M. Foti, and S.D. Emr, Distinct roles for the yeast phosphatidylinositol 4-kinases, Stt4p and Pik1p, in secretion, cell growth, and organelle membrane dynamics. Mol Biol Cell, 2000. **11**(8): p. 2673-89.DOI: 10.1091/mbc.11.8.2673.
74. Tolias, K.F. and L.C. Cantley, Pathways for phosphoinositide synthesis. Chem Phys Lipids, 1999. **98**(1-2): p. 69-77.DOI: 10.1016/s0009-3084(99)00019-5.

75. Yoshida, S., et al., A novel gene, *STT4*, encodes a phosphatidylinositol 4-kinase in the PKC1 protein kinase pathway of *Saccharomyces cerevisiae.* Journal of Biological Chemistry, 1994. **269**(2): p. 1166-1172.
76. Hu, H., et al., Phosphoinositide 3-kinase regulates glycolysis through mobilization of aldolase from the actin cytoskeleton. Cell, 2016. **164**(3): p. 433-46.DOI: 10.1016/j.cell.2015.12.042.
77. Kusakabe, T., K. Motoki, and K. Hori, Mode of interactions of human aldolase isozymes with cytoskeletons. Arch Biochem Biophys, 1997. **344**(1): p. 184-93.DOI: 10.1006/abbi.1997.0204.
78. Deshmukh, A.S., et al., Deep proteomics of mouse skeletal muscle enables quantitation of protein isoforms, metabolic pathways, and transcription factors. Mol Cell Proteomics, 2015. **14**(4): p. 841-53.DOI: 10.1074/mcp.M114.044222.
79. Shimizu, A., F. Suzuki, and K. Kato, Characterization of alpha alpha, beta beta, gamma gamma and alpha gamma human enolase isozymes, and preparation of hybrid enolases (alpha gamma, beta gamma and alpha beta) from homodimeric forms. Biochim Biophys Acta, 1983. **748**(2): p. 278-84.DOI: 10.1016/0167-4838(83)90305-9.
80. Durany, N. and J. Carreras, Distribution of phosphoglycerate mutase isozymes in rat, rabbit and human tissues. Comp Biochem Physiol B Biochem Mol Biol, 1996. **114**(2): p. 217-23.DOI: 10.1016/0305-0491(95)02135-3.
81. Harkins, R.N., J.A. Black, and M.B. Rittenberg, M2 isozyme of pyruvate kinase from human kidney as the product of a separate gene: its purification and characterization. Biochemistry, 1977. **16**(17): p. 3831-3837.
82. Ahuatzi, D., et al., The glucose-regulated nuclear localization of hexokinase 2 in *Saccharomyces cerevisiae* is Mig1-dependent. J Biol Chem, 2004. **279**(14): p. 14440-6.DOI: 10.1074/jbc.M313431200.
83. Riera, A., et al., Human pancreatic β-cell glucokinase: subcellular localization and glucose repression signalling function in the yeast cell. Biochemical Journal, 2008. **415**(2): p. 233-239.
84. Albe, K.R., M.H. Butler, and B.E. Wright, Cellular Concentrations of Enzymes and Their Substrates. Journal of Theoretical Biology, 1990. **143**(2): p. 163-195.DOI: 10.1016/S0022-5193(05)80266-8.
85. Bogardus, C., et al., Correlation between muscle glycogen synthase activity and *in vivo* insulin action in man. The Journal of Clinical Investigation, 1984. **73**(4): p. 1185-1190.DOI: 10.1172/JCI111304.
86. Gollnick, P.D., et al., Human soleus muscle: A comparison of fiber composition and enzyme activities with other leg muscles. Pflügers Archiv, 1974. **348**(3): p. 247-255.DOI: 10.1007/BF00587415.
87. Simoneau, J.A. and C. Bouchard, Human variation in skeletal muscle fiber-type proportion and enzyme activities. Am J Physiol, 1989. **257**(4 Pt 1): p. E567-72.DOI: 10.1152/ajpendo.1989.257.4.E567.
88. Hansen, P.A., et al., Skeletal muscle glucose transport and metabolism are enhanced in transgenic mice overexpressing the Glut4 glucose transporter. J Biol Chem, 1995. **270**(4): p. 1679-84.DOI: 10.1074/jbc.270.5.1679.
89. Ren, J.M., et al., Evidence from transgenic mice that glucose transport is rate-limiting for glycogen deposition and glycolysis in skeletal muscle. J Biol Chem, 1993. **268**(22): p. 16113-5.
90. Schmitz, J.P., et al., Silencing of glycolysis in muscle: experimental observation and numerical analysis. Exp Physiol, 2010. **95**(2): p. 380-97.DOI: 10.1113/expphysiol.2009.049841.
91. Argüelles, J.-C., Why Can't Vertebrates Synthesize Trehalose? Journal of Molecular Evolution, 2014. **79**(3): p. 111-116.DOI: 10.1007/s00239-014-9645-9.

92. Heinisch, J.J., J. Knuesting, and R. Scheibe, Investigation of Heterologously Expressed Glucose-6-Phosphate Dehydrogenase Genes in a Yeast *zwf1* Deletion. Microorganisms, 2020. **8**(4): p. 546.
93. Bresolin, N., et al., Muscle glucose-6-phosphate dehydrogenase deficiency. Journal of Neurology, 1989. **236**(4): p. 193-198.DOI: 10.1007/BF00314498.
94. Battistuzzi, G., et al., Tissue-specific levels of human glucose-6-phosphate dehydrogenase correlate with methylation of specific sites at the 3' end of the gene. Proceedings of the National Academy of Sciences, 1985. **82**(5): p. 1465-1469.DOI: 10.1073/pnas.82.5.1465.
95. Li, M.V., et al., Glucose-6-phosphate mediates activation of the carbohydrate responsive binding protein (ChREBP). Biochem Biophys Res Commun, 2010. **395**(3): p. 395-400.DOI: 10.1016/j.bbrc.2010.04.028.
96. Peterson, C.W., et al., Glucose controls nuclear accumulation, promoter binding, and transcriptional activity of the MondoA-Mlx heterodimer. Mol Cell Biol, 2010. **30**(12): p. 2887-95.DOI: 10.1128/MCB.01613-09.
97. Canelas, A.B., et al., An *in vivo* data-driven framework for classification and quantification of enzyme kinetics and determination of apparent thermodynamic data. Metab Eng, 2011. **13**(3): p. 294-306.DOI: 10.1016/j.ymben.2011.02.005.
98. Blachly-Dyson, E., et al., Cloning and functional expression in yeast of two human isoforms of the outer mitochondrial membrane channel, the voltage-dependent anion channel. Journal of Biological Chemistry, 1993. **268**(3): p. 1835-1841.
99. Wilson, J.E., Homologous and heterologous interactions between hexokinase and mitochondrial porin: Evolutionary implications. Journal of Bioenergetics and Biomembranes, 1997. **29**(1): p. 97-102.DOI: Doi 10.1023/A:1022472124746.
100. Rais, B., et al., Quantitative characterization of homo- and heteroassociations of muscle phosphofructokinase with aldolase. Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology, 2000. **1479**(1-2): p. 303-314.DOI: Doi 10.1016/S0167-4838(00)00047-9.
101. Marinho-Carvalho, M.M., et al., Calmodulin upregulates skeletal muscle 6-phosphofructo-1-kinase reversing the inhibitory effects of allosteric modulators. Biochim Biophys Acta, 2009. **1794**(8): p. 1175-80.DOI: 10.1016/j.bbapap.2009.02.006.
102. Schmitz, J.P., et al., Combined *in vivo* and *in silico* investigations of activation of glycolysis in contracting skeletal muscle. Am J Physiol Cell Physiol, 2013. **304**(2): p. C180-93.DOI: 10.1152/ajpcell.00101.2012.
103. Jenkins, C.M., et al., Reversible high affinity inhibition of phosphofructokinase-1 by acyl-CoA: a mechanism integrating glycolytic flux with lipid metabolism. J Biol Chem, 2011. **286**(14): p. 11937-50.DOI: 10.1074/jbc.M110.203661.
104. Thompson, A.L. and G.J. Cooney, Acyl-CoA inhibition of hexokinase in rat and human skeletal muscle is a potential mechanism of lipid-induced insulin resistance. Diabetes, 2000. **49**(11): p. 1761-1765.DOI: 10.2337/diabetes.49.11.1761.
105. Entian, K.-D. and P. Kötter, 25 Yeast genetic strain and plasmid collections. Methods in Microbiology, 2007. **36**: p. 629-666.
106. Verduyn, C., et al., Effect of benzoic acid on metabolic fluxes in yeasts: a continuous culture study on the regulation of respiration and alcoholic fermentation. Yeast, 1992. **8**(7): p. 501-517.DOI: 10.1002/yea.320080703.
107. Solis-Escalante, D., et al., *amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae.* FEMS Yeast Research, 2013. **13**(1): p. 126-39.DOI: 10.1111/1567-1364.12024.
108. Bertani, G., Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. Journal of Bacteriology, 2004. **186**(3): p. 595-600.
109. Bertani, G., Studies on lysogenis I.: The mode of phage liberation by lysogenic *Escherichia coli.* Journal of Bacteriology, 1951. **62**(3): p. 293-300.

110. Koning, M., et al., A global downregulation of microRNAs occurs in human quiescent satellite cells during myogenesis. Differentiation, 2012. **84**(4): p. 314-21.DOI: 10.1016/j.diff.2012.08.002.
111. Almonacid Suarez, A.M., et al., Directional topography gradients drive optimum alignment and differentiation of human myoblasts. Journal of tissue engineering and regenerative medicine, 2019. **13**(12): p. 2234-2245.
112. Lee, M.E., et al., A highly characterized yeast toolkit for modular, multipart assembly. ACS Synthetic Biology, 2015. **4**(9): p. 975-986.DOI: 10.1021/sb500366v.
113. Boonekamp, F.J., et al., The genetic makeup and expression of the glycolytic and fermentative pathways are highly conserved within the *Saccharomyces* genus. Frontiers in genetics 2018. **9**: p. 504.DOI: 10.3389/fgene.2018.00504.
114. Gietz, R.D. and R.A. Woods, Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods in enzymology, 2002. **350**: p. 87-96.
115. Mans, R., et al., CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae.* FEMS Yeast Research, 2015. **15**(2).DOI: 10.1093/femsyr/fov004.
116. Mikkelsen, M.D., et al., Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metabolic engineering, 2012. **14**(2): p. 104-111.
117. Hoops, S., et al., COPASI—a complex pathway simulator. Bioinformatics, 2006. **22**(24): p. 3067-3074.
118. Flamholz, A., et al., eQuilibrator—the biochemical thermodynamics calculator. Nucleic acids research, 2012. **40**(D1): p. D770-D775.
119. van Eunen, K., et al., Measuring enzyme activities under standardized in vivo-like conditions for systems biology. FEBS J, 2010. **277**(3): p. 749-60.DOI: 10.1111/j.1742-4658.2009.07524.x.
120. Fell, D.A., Metabolic control analysis: a survey of its theoretical and experimental development. Biochem J, 1992. **286 (Pt 2)**: p. 313-30.DOI: 10.1042/bj2860313.
121. Güldener, U., et al., A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic acids research, 1996. **24**(13): p. 2519-2524.
122. Salazar, A.N., et al., Nanopore sequencing enables near-complete de novo assembly of *Saccharomyces cerevisiae* reference strain CEN.PK113-7D. FEMS Yeast Res, 2017. **17**(7).DOI: 10.1093/femsyr/fox074.
123. Li, H., et al., The sequence alignment/map format and SAMtools. Bioinformatics, 2009. **25**(16): p. 2078-2079.
124. Walker, B.J., et al., Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One, 2014. **9**(11): p. e112963.DOI: 10.1371/journal.pone.0112963.
125. Nijkamp, J.F., et al., *De novo* detection of copy number variation by co-assembly. Bioinformatics, 2012. **28**(24): p. 3195-202.DOI: 10.1093/bioinformatics/bts601.
126. Santos, B. and M. Snyder, Sbe2p and sbe22p, two homologous Golgi proteins involved in yeast cell wall formation. Mol Biol Cell, 2000. **11**(2): p. 435-52.DOI: 10.1091/mbc.11.2.435.
127. Postma, E., et al., Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae.* Appl Environ Microbiol, 1989. **55**(2): p. 468-77.DOI: 10.1128/aem.55.2.468-477.1989.
128. Cruz, L.A., et al., Similar temperature dependencies of glycolytic enzymes: an evolutionary adaptation to temperature dynamics? BMC Syst Biol, 2012. **6**: p. 151.DOI: 10.1186/1752-0509-6-151.
129. Rijksen, G. and G.E. Staal, Regulation of human erythrocyte hexokinase. The influence of glycolytic intermediates and inorganic phosphate. Biochim Biophys Acta, 1977. **485**(1): p. 75-86.DOI: 10.1016/0005-2744(77)90194-2.

130. Lowry, O.H., et al., Protein measurement with the Folin phenol reagent. J Biol Chem, 1951. **193**(1): p. 265-75.
131. Silveira, M.C., E. Carvajal, and E.P. Bon, Assay for *in vivo* yeast invertase activity using NaF. Analytical Biochemistry, 1996. **238**(1): p. 26-8.DOI: 10.1006/abio.1996.0244.
132. Haase, S.B. and S.I. Reed, Improved flow cytometric analysis of the budding yeast cell cycle. Cell Cycle, 2002. **1**(2): p. 132-6.
133. Guynn, R.W., et al., The concentration and control of cytoplasmic free inorganic pyrophosphate in rat liver *in vivo.* Biochemical Journal, 1974. **140**(3): p. 369-375.
134. Rorsman, P. and G. Trube, Glucose dependent K+-channels in pancreaticβ-cells are regulated by intracellular ATP. Pflügers Archiv, 1985. **405**(4): p. 305-309.
135. Tschopp, J. and K. Schroder, NLRP3 inflammasome activation: The convergence of multiple signalling pathways on ROS production? Nature reviews immunology, 2010. **10**(3): p. 210-215.
136. Kristensen, L.O., Associations between transports of alanine and cations across cell membrane in rat hepatocytes. American Journal of Physiology-Gastrointestinal and Liver Physiology, 1986. **251**(5): p. G575-G584.
137. Conley, K., et al., Activation of glycolysis in human muscle *in vivo.* American Journal of Physiology-Cell Physiology, 1997. **273**(1): p. C306-C315.
138. Lidofsky, S.D., et al., Vasopressin increases cytosolic sodium concentration in hepatocytes and activates calcium influx through cation-selective channels. Journal of Biological Chemistry, 1993. **268**(20): p. 14632-14636.
139. Breitwieser, G.E., A.A. Altamirano, and J.M. Russell, Osmotic stimulation of Na (+)-K (+)-Cl-cotransport in squid giant axon is [Cl-] i dependent. American Journal of Physiology-Cell Physiology, 1990. **258**(4): p. C749-C753.
140. Janssen, L. and S. Sims, Acetylcholine activates non-selective cation and chloride conductances in canine and guinea-pig tracheal myocytes. The Journal of Physiology, 1992. **453**(1): p. 197-218.
141. Ingwall, J.S., Phosphorus nuclear magnetic resonance spectroscopy of cardiac and skeletal muscles. American Journal of Physiology-Heart and Circulatory Physiology, 1982. **242**(5): p. H729-H744.
142. Murphy, E., et al., Cytosolic free magnesium levels in ischemic rat heart. Journal of Biological Chemistry, 1989. **264**(10): p. 5622-5627.
143. Bárány, M., *Biochemistry of smooth muscle contraction*. 1996: Elsevier.
144. Bruch, P., K.D. Schnackerz, and R.W. Gracy, Matrix-Bound Phosphoglucose Isomerase: Formation and Properties of Monomers and Hybrids. European journal of biochemistry, 1976. **68**(1): p. 153-158.
145. Groen, A.K., et al., Control of gluconeogenesis in rat liver cells. I. Kinetics of the individual enzymes and the effect of glucagon. Journal of Biological Chemistry, 1983. **258**(23): p. 14346-14353.
146. Ishibashi, H. and G.L. Cottam, Glucagon-stimulated phosphorylation of pyruvate kinase in hepatocytes. Journal of Biological Chemistry, 1978. **253**(24): p. 8767-8771.
147. Oudard, S., et al., High glycolysis in gliomas despite low hexokinase transcription and activity correlated to chromosome 10 loss. British journal of cancer, 1996. **74**(6): p. 839-845.
148. Civelek, V.N., et al., Intracellular pH in adipocytes: effects of free fatty acid diffusion across the plasma membrane, lipolytic agonists, and insulin. Proceedings of the National Academy of Sciences, 1996. **93**(19): p. 10139-10144.
149. Van Schaftingen, E. and H.G. Hers, Formation of fructose 2,6-bisphosphate from fructose 1,6-bisphosphate by intramolecular cyclisation followed by alkaline hydrolysis. European Journal of Biochemistry, 1981. **117**(2): p. 319-23.

150. Wolters, J.C., et al., Translational targeted proteomics profiling of mitochondrial energy metabolic pathways in mouse and human samples. Journal of proteome research, 2016. **15**(9): p. 3204-3213.

Chapter 4

Simpler is not always better: transplanting the *Yarrowia lipolytica* glycolytic pathway into *Saccharomyces cerevisiae* reveals essential synergetic regulatory mechanisms

Ewout Knibbe, Francine J. Boonekamp, Rachel Stuij, Koen A. J. Pelsma, Liset Jansen, Carmen-Lisset Flores, Pascale Daran-Lapujade

Manuscript in preparation for submission

Abstract

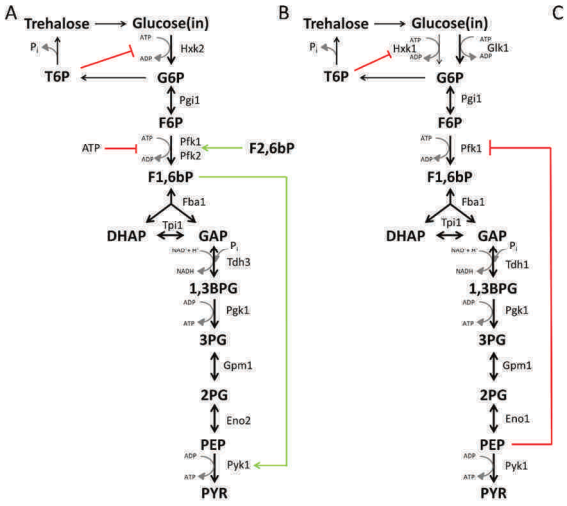
The Embden-Meyerof-Parnas pathway of glycolysis is a widely distributed and intensively investigated metabolic route. While allosteric regulation is thought to be essential for glycolytic flux dynamics in many organisms including yeast, to date single enzyme complementation studies with non-allosteric glycolytic enzymes have failed to experimentally demonstrate this essentiality and quantify the overall contribution of allosteric regulation in tuning the glycolytic flux. This study brings new insight in the synergetic metabolic role of allosteric regulation by implementing pathway swapping, a strategy enabling to remodel, in two simple genetic interventions, the entire glycolytic pathway of *Saccharomyces cerevisiae*. *S. cerevisiae* equipped with the full set of non-allosteric glycolytic enzymes from the oleaginous yeast *Y. lipolytica* lost the ability to grow on media containing 2% glucose and displayed dynamic responses suggesting metabolic imbalance between upper and lower glycolysis. Single and combined gene complementation demonstrated that this phenotype was caused by the simultaneous deregulation of the three key kinases: hexokinase, phosphofructokinase and pyruvate kinase. 'Deregulated glycolysis' *S. cerevisiae* strains could naturally restore glycolytic stability and growth on glucose by evolving mutations in the *Y. lipolytica* glucokinase, causing a strong decrease in glucokinase activity and glycolytic flux. This solution could be recapitulated in non-evolved deregulated glycolysis *S. cerevisiae* strains by experimentally tuning glucose import. Supported by kinetic modelling, the present work demonstrates the major synergetic role played by allosteric regulations in preventing metabolic imbalance in glycolysis and highlights the power of synthetic biology in addressing long-standing questions in systems biology.

Introduction

Glycolysis, the main route for sugar utilization across all kingdoms of life, is an important biochemical pathway in industrial biotechnology [1] and is involved in several mammalian diseases, such as cancer (via the Warburg effect) and diabetes. Eukaryotes and several prokaryotes favour the Embden-Meyerhof-Parnas (EMP) pathway of glycolysis (Fig. 1A), a set of ten biochemical reactions largely elucidated in the model and industrial yeast *Saccharomyces cerevisiae*. Catalysing the conversion of hexoses to pyruvate, glycolysis plays a key role in energy conservation and the supply of precursors for biomass formation. To meet cellular demands and adjust to varying carbon and energy supplies, cells constantly need to tune the glycolytic flux in response to their environment dynamics. While the glycolytic biochemical conversions are highly conserved between different organisms, evolution in dynamic environments has equipped cells with a myriad of multi-layered regulatory mechanisms to control the flux, ranging from gene expression regulation to modulation of *in vivo* activity of the glycolytic enzymes by post-translational protein modifications and binding of metabolites (also known as metabolic regulation) [2, 3]. Metabolic regulation of glycolytic enzymes is exerted via simple mass action (concentrations of substrates and products) or through allosteric activation and inhibition by specific effectors belonging to or closely related to glycolysis. The nature of these metabolic regulations is organism-dependent and even tissue-dependent in metazoans. While functions have been linked to some regulations, e.g. phosphofructokinase inhibition by ATP to sense cellular energy demand [4] and feedforward activation of pyruvate kinase to limit accumulation of lower glycolytic intermediates [5], these are difficult to confirm experimentally and for many regulatory interactions a functional understanding is lacking. While the EMP pathway of *S. cerevisiae* is undoubtedly one of the most studied and best understood examples of glycolysis, the precise physiological role of several well-described metabolic regulations remains uncertain. Hexokinase (Hxk), first step in glycolysis, is inhibited by trehalose-6-phophate (T6P), phosphofructokinase (Pfk) is activated by fructose-2,6-bisphosphate (F2,6bP) and inhibited by ATP and pyruvate kinase (Pyk) is sensitive to feedforward activation by fructose-1,6-bisphosphate (F1,6bP) (Fig. 1A). While some other metabolic effectors have been described, these main metabolic regulations are thought to help balance production and consumption of ATP in a pathway in which an initial energy investment is required (ATP hydrolysed by Hxk and Pfk) before the pay-off phase in the lower part of glycolysis [5-7]. However, to date individual complementation of these three glycolytic steps (HXK, PFK and PYK) with insensitive variants have failed to demonstrate the purported role of these metabolic regulation in maintaining a balanced glycolysis [8-12].

While individual gene complementation is a powerful approach for the functional characterization of proteins, it might fail to capture synergetic mechanisms where multiple proteins are involved in a specific pathway or function. This shortcoming can be tackled by multigene complementation approaches, targeting multiple enzymes at once. However, multigene complementation of essential pathways such as glycolysis is technically very challenging and is hindered in eukaryotes by a high degree of genetic redundancy in central metabolism. To address these issues and enable facile modular complementation of the entire glycolytic pathway, we engineered *S. cerevisiae* for glycolysis swapping (the Minimal Glycolysis and Switchable Yeast Glycolysis strains [13, 14]) and recently demonstrated the two-step full humanization of *S. cerevisiae* glycolysis using these platforms [15]. Glycolysis swapping can be used to challenge *S. cerevisiae* native metabolic regulation, for instance by transplanting glycolytic variants from organisms that have evolved in radically different environments. The oleaginous yeast *Yarrowia lipolytica* is a pre-whole genome duplication, Crabtree-negative yeast and is phylogenetically very distant from the *Saccharomyces* genus [16]. *Y. lipolytica* thrives in environments where glucose, *S. cerevisiae's* favourite carbon source, is scarce, instead favouring glycerol and lipids as carbon source. Although *Y. lipolytica* can grow on hexoses, it does not ferment these to ethanol, favouring their respiratory dissimilation. These nutritional preferences have understandably shaped the regulation of *Y. lipolytica*'s glycolysis differently from *S. cerevisiae*. All steps in *Y. lipolytica* glycolysis are catalysed by a single isoenzyme, with the exception of the glucose phosphorylation step (Fig. 1B and C). *Y. lipolytica* relies mostly on a glucokinase (Glk), insensitive to any known allosteric regulators, while hexokinase (Hxk), sensitive to T6P inhibition and known to complement in *S. cerevisiae*, plays a minor role [17, 18]. Furthermore, while phosphofructokinase (Pfk) is a hub of allosteric regulation in many organisms, including *S. cerevisiae*, the *Y. lipolytica* Pfk is only strongly inhibited by phosphoenolpyruvate (PEP). This enzyme has been shown to complement individually in *S. cerevisiae* [11]. Finally *Y. lipolytica* pyruvate kinase is not affected by F1,6bP, activator of *S. cerevisiae* Pyk1 [19]. The *Y. lipolytica* glycolytic pathway therefore lacks most of the allosteric regulations found in the model yeast *S. cerevisiae*.

Using *S. cerevisiae* glycolysis as paradigm, the present study explores the potential of partial and full pathway swapping to identify the metabolic role of complex allosteric regulations.



C

Reaction Step	*S. c.* protein	*Y.l.* protein	Amino acid identity
HXK	*Sc*Hxk2	*Yl*Hxk1	55.7 %
	*Sc*Glk1	*Yl*Glk1	39.9 %
PGI	*Sc*Pgi1	*Yl*Pgi1	74.9 %
PFK	*Sc*Pfk1	*Yl*Pfk1	47.8%
	*Sc*Pfk2	*Yl*Pfk1	48.5 %
FBA	*Sc*Fba1	*Yl*Fba1	67.9 %
TPI	*Sc*Tpi1	*Yl*Tpi1	66.7 %
GAPDH	*Sc*Tdh3	*Yl*Tdh1	67.2 %
PGK	*Sc*Pgk1	*Yl*Pgk1	68.4 %
GPM	*Sc*Gpm1	*Yl*Gpm1	72.6 %
ENO	*Sc*Eno2	*Yl*Eno1	73.0 %
PYK	*Sc*Pyk1	*Yl*Pyk1	66.1 %

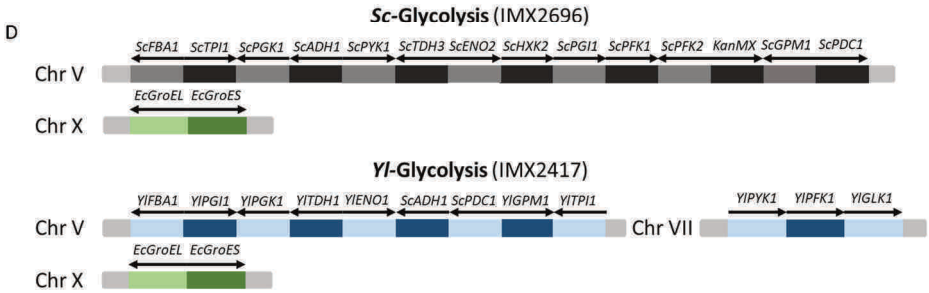


Figure 1 - Schematic overview of the allosteric regulation of glycolysis in *S. cerevisiae* and *Y. lipolytica*.
A) Glycolytic pathway in *S. cerevisiae*, only the major isoenzymes are shown. **B)** Glycolysis in *Y. lipolytica* with all isoenzymes shown. In A and B red lines indicate inhibition, green arrows indicate activation. G6P: glucose-6-phosphate; T6P: trehalose-6-phosphate; F6P: fructose-6-phosphate; F1,6bP: fructose-1,6-bisphosphate; DHAP: dihydroxy-acetone phosphate; GAP: glyceraldehyde-3-phosphate; 1,3BPG: 1,3-bisphosphoglycerate; 3PG: 3-phosphoglycerate; 2PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate. **C)** Percentage identity between *S. cerevisiae* major glycolytic isoenzymes and their *Y. lipolytica* counterparts. HXK: hexokinase/glucokinase; PGI: Phosphoglucose isomerase; PFK: Phosphofructokinase; FBA: Fructose bisphosphate aldolase; TPI: Triosephosphate isomerase; GAPDH: Glyceraldehyde-3P dehydrogenase; PGK: Phosphoglycerate kinase; GPM: Phosphoglycerate mutase; ENO: Enolase; PYK: Pyruvate kinase; PDC: Pyruvate decarboxylase; ADH: Alcohol dehydrogenase. **D)** Schematic overview of the genetic loci containing glycolytic and chaperone genes in the control *Sc*-Glycolysis strain and the fully swapped *Yl*-Glycolysis strain.

Results

General experimental approach for pathway transplantation

While *S. cerevisiae* has a set of 26 glycolytic enzymes, the presence of a single enzyme for nine out of the ten glycolytic steps in *Y. lipolytica* simplified the choice of genes to transplant. The complete set of ten *Y. lipolytica* enzymes leading to the conversion of glucose to pyruvate was transplanted in *S. cerevisiae* using the SwYG strain, harbouring a minimal set of glycolytic genes relocated to a single locus [13]. The T6P-insensitive glucokinase (*Yl*Glk1) was chosen for expression in *S. cerevisiae* as part of the complete pathway. The *E. coli* folding chaperones GroEL and GroES were co-expressed to facilitate functional expression of *Y. lipolytica* enzymes [20]. The constructed IMX2417 strain, called *Yl*-Glycolysis, was devoid of *S. cerevisiae* glycolytic genes and entirely relied on the *Y. lipolytica* glycolysis for glucose dissimilation (Fig. 1D). A reference strain, IMX2696 (named *Sc*-Glycolysis), expressing the native, minimized *S. cerevisiae* glycolytic pathway from the same chromosomal locus and expressing the GroEL/GroES chaperones was also constructed. It is important to note that, to minimize the requirement for fast glycolytic flux and the risk of glycolytic imbalance, during all construction steps, strains expressing *Yarrowia* glycolytic genes were purposely grown on galactose or on a mixture of ethanol and glycerol, but not on glucose. Galactose utilization occurs via the Leloir pathway in yeast, and the lower capacity of this pathway compared to hexokinase is hypothesized to facilitate transitions, for example in *tps1 S. cerevisiae* strains [21, 22].

S. cerevisiae with a *Y. lipolytica* glycolysis displays growth defects on glucose medium

Upon full pathway transplantation, growth on galactose in minimal, chemically defined medium was immediately observed for the *Yl*-Glycolysis strain IMX2417 (Fig. 2A and B), proving functional expression in *S. cerevisiae* of all *Yarrowia lipolytica* glycolytic enzymes besides glucokinase (*Yl*Glk1). Growth on ethanol was similarly observed immediately, although at a reduced rate. However, the same strain could not readily grow upon transfer to glucose medium, and displayed slow but exponential growth only after 3 to 4 days of incubation (Fig. 2B). Intracellular pH (pH_i) is a good indicator of the cellular metabolic status, considering that cells unable to conserve energy are incapable of maintaining pH homeostasis [21, 23, 24]. Using pHluorin as proxy for pH_i showed that addition of either glucose or galactose to a galactose-grown *Sc*-Glycolysis strain did not cause a decrease in pH_i, but instead a slight increase, consistent with previous data of control strains (Fig. 2C and 2D and Supplementary Fig. S1) [21, 23]. A similar response was observed immediately after galactose was added to the *Yl*-Glycolysis strain (Fig. 2E), while in contrast addition of glucose to the medium caused a substantial and rapid drop in pH_i, which lasted

for a prolonged period of time (Fig. 2E and F and Supplementary Fig. S2), matching the initial lack of growth observed on glucose. This response to glucose might have several causes such as a non-functional *Y. lipolytica* glucokinase, an insufficient capacity of the *Y. lipolytica* glycolytic enzymes during growth on glucose, or a regulatory deficiency preventing glucose catabolism by otherwise functional glycolytic enzymes. To determine the functionality of *Y. lipolytica* glycolytic enzymes expressed in *S. cerevisiae*, *in vitro* assays were performed on the *Yl*-Glycolysis strain grown on galactose. All *Y. lipolytica* glycolytic enzymes, including glucokinase, showed activity in *in vitro* assays from cell extracts of galactose-grown cultures (Fig. 2G). Compared to the *Sc*-glycolysis reference strain IMX2696, the *Y. lipolytica* phosphoglucose isomerase (PGI), triosephosphate isomerase (TPI) and phosphoglycerate mutase (PGM) showed substantially lower specific activities (19%, 23% and 7% of the activity in the *Sc*-glycolysis strain, respectively), while the kinases glucokinase (GLK), phosphofructokinase (PFK) and pyruvate kinase (PYK), which were expressed as codon-optimized genes, were highly active *in vitro* (89%, 259% and 67% of *Sc*-Glycolysis strain activity).

The *Y. lipolytica* glycolytic enzymes were all expressed in *S. cerevisiae* and functional *in vitro*. Fast growth on galactose (0.24±0.01 h^{-1} as compared to 0.25±0.00 h^{-1} for the control *Sc*-Glycolysis strain, Fig. 2A) demonstrated that enzymes from phosphoglucose isomerase to pyruvate kinase were active *in vivo* and able to carry the glycolytic flux, although the lower activities of some enzymes might be responsible for the lower growth rates measured in ethanol and glucose. Growth on galactose but inability to grow upon transition to glucose of the *Yl*-Glycolysis strain despite substantial *in vitro* activity of *Yl*Glk, suggested a defect in *Yl*Glk activity *in vivo* or a regulatory imbalance in the pathway. The appearance of growth after an extended lag phase upon galactose/glucose transition could be explained by the resolution of *Yl*Glk *in vivo* activity defects either by evolution (i.e. mutations) or metabolic adaptation.

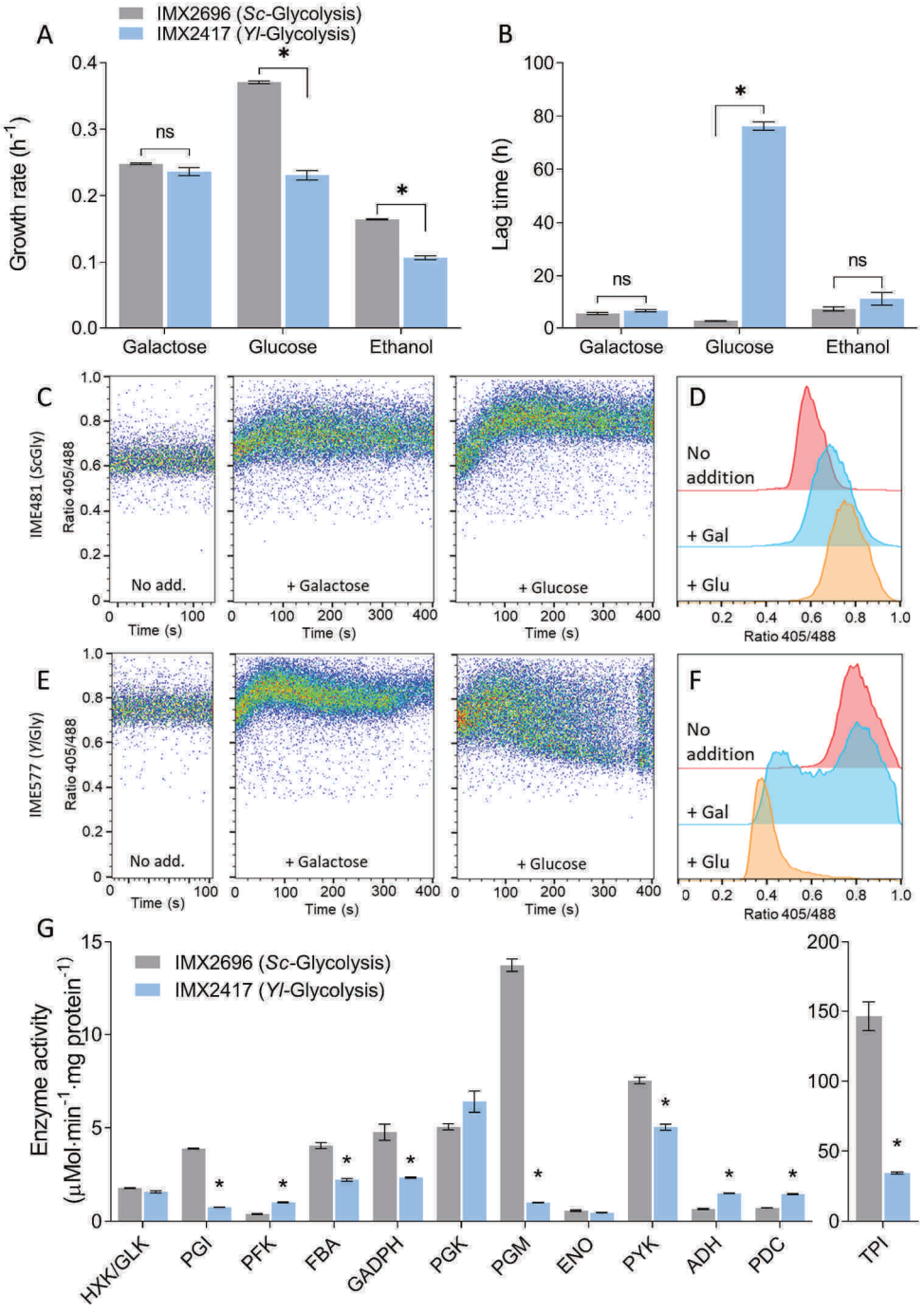

A
IMX2696 (Sc-Glycolysis)
IMX2417 (Yl-Glycolysis)
Growth rate (h-1)
ns
*
*
Galactose
Glucose
Ethanol
B
Lag time (h)
ns
*
ns
Galactose
Glucose
Ethanol
C
IME481 (ScGly)
Ratio 405/488
No add.
+ Galactose
+ Glucose
Time (s)
D
No addition
+ Gal
+ Glu
Ratio 405/488
E
IME577 (YlGly)
Ratio 405/488
No add.
+ Galactose
+ Glucose
Time (s)
F
No addition
+ Gal
+ Glu
Ratio 405/488
G
IMX2696 (Sc-Glycolysis)
IMX2417 (Yl-Glycolysis)
Enzyme activity (µMol·min-1·mg protein-1)
HXK/GLK
PGI
PFK
FBA
GADPH
PGK
PGM
ENO
PYK
ADH
PDC
TPI

Figure 2 - Characterization of a fully swapped glycolysis strain.
A) Growth rates measured in the growth profiler on galactose, glucose and ethanol for the full *Yarrowia lipolytica* glycolysis strain (*Yl*-Glycolysis, IMX2417) and a control strain expressing the native *S. cerevisiae* glycolytic pathway from a single locus (*Sc*-Glycolysis, IMX2696). **B)** Lag phase duration upon transfer from SM-galactose to SM with either galactose, glucose or ethanol. For **A)** and **B)** the mean and SEM of triplicates are shown, * indicates significant difference (t-test, homoscedastic, unpaired with P-value<0.05). **C)** and **E)** Dynamics of intracellular pH, measured as pHluorin signal ratio, in a strain with the native glycolytic pathway (**C**) and the *Yl*-Glycolysis strain (**E**) without, and immediately after galactose or glucose addition. **D)** and **F)** pHluorin signal ratio in the population of native glycolysis strain IME481 (**D**) and *Yl*-Glycolysis strain (**F**) after incubation without carbon-source addition (red, after 56 and 62 min for **D)** and **F)** respectively) or with galactose (blue, after 66 and 72 min) or glucose (orange, after 76 and 82 min). Duplicate pHluorin experiments are shown in Supplementary Fig. S1 and S2. **G)** *In vitro* enzyme activities of the glycolytic enzymes in the *Yl*-Glycolysis and *Sc*-Glycolysis strains grown on galactose. Mean and SEM of triplicates are shown, * indicates significant differences (t-test, homoscedastic, unpaired with P-value<0.05).

Full pathway swapping reveals that glucose phosphorylation is a key regulatory node for glycolysis transplantation in *S. cerevisiae*

To identify whether the delayed growth on glucose of the *S. cerevisiae* strain with the *Yarrowia* glycolysis resulted from evolution or adaptation, serial transfers of the *Yl*-Glycolysis strain in media alternating galactose and glucose as carbon source were performed. The long lag phase disappeared in the second and following galactose/glucose transitions, revealing that the ability to grow on glucose of the *Yl*-Glycoysis strain was most probably of genetic origin (Supplementary Fig. S3). Systematic mutations in *YlGLK1* were accordingly found in single colonies isolated from three independent glucose-grown cultures (Fig. 3D and S1B). All three isolates were capable of fast growth on glucose without lag phase (strains IMS1202 to 1204, Fig. 3A and B). Whole genome sequencing identified mutations leading to single, distinct amino-acid substitutions in all isolates, scattered over the *Yl*Glk protein (Supplementary Fig. S4). *In vitro* assays revealed a marked (10- to 146-fold) decrease in glucokinase activity for the three mutated *YlGLK* variants as compared to the native enzyme (Fig. 3E and Supplementary Fig. S5), and an apparent increase in the $K_{m,glucose}$ for two of them (Supplementary Fig. S5A). More detailed physiological characterization of the evolved strains in shake flasks on glucose urea medium showed reduced growth rates as compared to the *Sc*-Glycolysis strain (ca. two-fold decreased, Fig. 3F). This decrease in growth rate corresponded to a fully respiratory metabolism, with a higher biomass yield, a strong decrease in glucose uptake rate, and an absence of ethanol production (Fig. 3 G-I). The lower flux through glycolysis appears to match the lower glucokinase activity in these strains, although for IMS1203 and IMS1204 the *in vivo* glucokinase activity appeared too low to sustain the *in vivo* glucose uptake rate (Supplementary Fig. S6). This discrepancy suggested an underestimation of the *in vivo* glucokinase activity based on *in vitro* enzyme assay for the evolved strains. Nonetheless, all together, the ability to grow readily on

galactose but not on glucose of the *Yl*-Glycolysis strain, the systematic requirement for mutations in glucokinase for growth on glucose, and the decreased *in vitro* glucokinase activity and low glucose uptake rate in independent isolates suggested that the activity of glucokinase, the first step in glycolysis during glucose utilization, might be too high for growth on glucose in the *Yl*-Glycolysis strain.

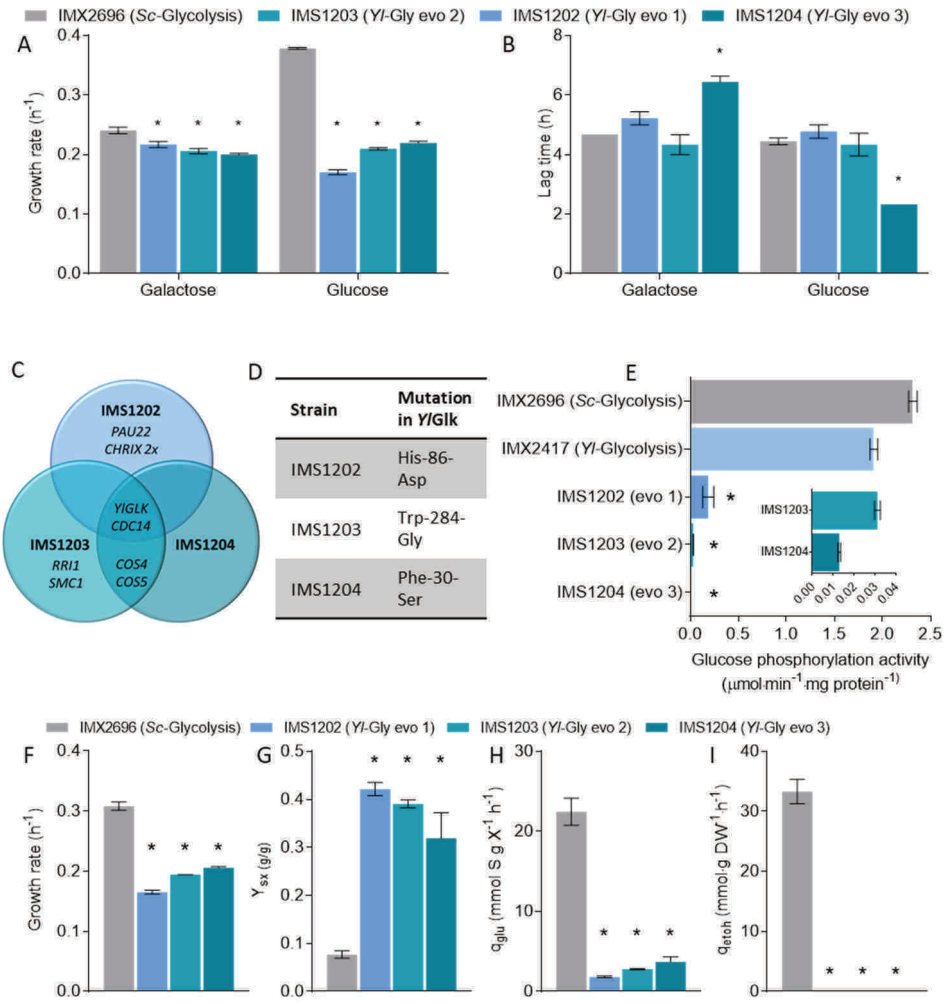


Figure 3 - Characterization of glucose-grown isolates of the *Yl*-Glycolysis strain.
A) Growth rates on galactose and glucose of the *Sc*-Glycolysis reference strain and of three single colony isolates obtained after serial transfers of the *Yl*-Glycolysis strain on glucose, data from biological triplicates. **B)** Duration of lag phase upon galactose/galactose and galactose/glucose transition for the same strains as panel A. **C)** Genes with mutations in the three independent isolates of the glucose-grown *Yl*-Glycolysis strain, *CHRIX* 2x indicates duplication of chromosome 9. **D)** Mutations in the *Y. lipolytica* glucokinase observed in each of the three isolates. **E)** *In vitro* glucose phosphorylation activity in the *Yl*-Glycolysis strain and in the evolved isolates measured in duplicate. To enable comparison with the non-mutated *Yl*-glycolysis strain IMX2417, all strains were grown in medium with galactose as sole carbon source. **F-I)** Physiological characterization of the evolved *Yl*-Glycolysis strains and control *Sc*-Glycolysis strain during growth on SM-glucose in duplicate shake-flasks. Growth rate, estimated biomass yield (Y_{SX}), specific glucose uptake rates (q_{glu}) and specific ethanol production rates (q_{etoh}) are shown. For all panels * indicate significant differences to control (t-test, homoscedastic, unpaired with P-value<0.05).

The glycolytic context determines the physiological impact of glucokinase deregulation

Previous studies have shown that *S. cerevisiae* hexokinase can be successfully complemented by heterologous, 'unregulated' variants insensitive to known effectors [8, 9, 15]. The requirement for mutations in *YlGLK1* for growth on glucose of *S. cerevisiae* and their impact on the kinetic properties of glucokinase was therefore unexpected and required further investigation. The major difference with these earlier studies was the presence of the complete set of glycolytic enzymes from *Yarrowia lipolytica* in the present study. The ability of T6P-insensitive variants to complement *Sc*Hxk2 might therefore be context-dependent. To test this hypothesis a single complementation strain was constructed using the Minimal Glycolysis (MG) strain [14], a strain with a minimized set of glycolytic genes, specifically engineered to facilitate complementation studies. The *YlHXK* gene, encoding the *Y. lipolytica* hexokinase, which is strongly inhibited by T6P, was similarly tested for complementation. During galactose/glucose transitions the *YlGLK* single complementation strain (IMX2062) largely responded as its parent strain MG with the complete set of *S. cerevisiae* enzymes, with an identical growth rate on glucose (Fig. 4A). The IMX2062 strain lag phase on glucose was slightly longer than that of MG (6.0±0.6 hours vs 2.7±0.3 hours, Fig. 4B) but was substantially shorter than the 3 to 4 days required for the *Yl*-glycolysis strain. The *YlHXK* strain did not show any measurable lag phase and only a very slight effect on growth rate. For both strains, characterization in shake flask showed similar growth rates, both strains produced ethanol when grown on glucose, only at a slightly lower rate (Supplementary Fig. S7). The intracellular pH however showed a decrease upon glucose addition in the *YlGLK* complementation strain, matching the slightly longer adaptation time required for galactose/glucose transition (Fig. 4F), but the effect was again not as severe as that of the *Yl*-glycolysis strain. Expression of *YlGLK1* therefore visibly but mildly affected *S. cerevisiae* transition from galactose to glucose, and did not result in

mutation in *YlGLK1* upon repeated transfers between galactose and glucose (Supplementary Fig. S3). Single complementation of the *S. cerevisiae* hexokinase with the T6P-insensitive *Yl*Glk1 was therefore not recapitulating the phenotype of the strain with a full *Y. lipolytica* glycolysis.

Next to hexokinase, phosphofructokinase and pyruvate kinase are considered as 'pacemakers' in *S. cerevisiae* glycolysis and are subject to allosteric regulations that are absent in their *Y. lipolytica* orthologs. To check if these differential regulations affected *S. cerevisiae* physiology, more particularly during carbon source switches, single complementation strains were also constructed for *YlPFK* and *YlPYK*, resulting in strains IMX2236 and IMX2235, respectively. These complementation strains grew with rates close to the MG control strain on glucose (0.43±0.007 and 0.42±0.001 h^{-1} compared to 0.44±0.008 h^{-1}), and displayed neither lag phase nor pH_i decrease upon transition from galactose to glucose (Fig. 4A and B, Supplementary Fig. S8). This absence of phenotype for individual complementation with insensitive isoenzymes was in line with earlier reports for phosphofructokinase [11, 12], and pyruvate kinase [15, 25]. In line with these observations, a strain expressing the entire *Y. lipolytica* glycolysis gene set except the phosphofructokinase (strain IMX2734) encoded by the native *ScPFK1* and *ScPFK2* genes, responded very similarly to the *Yl*-Glycolysis strain with full *Y. lipolytica* glycolysis during growth on glucose, further supporting the notion that single enzymes were not responsible for the growth defect on glucose of the *Yl*-Glycolysis strain (Supplementary Fig. S2 and S9).

Individual deregulation of hexokinase, phosphofructokinase and pyruvate kinase was therefore not affecting *S. cerevisiae* ability to switch between galactose and glucose, revealing that functional expression of the *Y. lipolytica* enzymes was dependent on the glycolytic context.

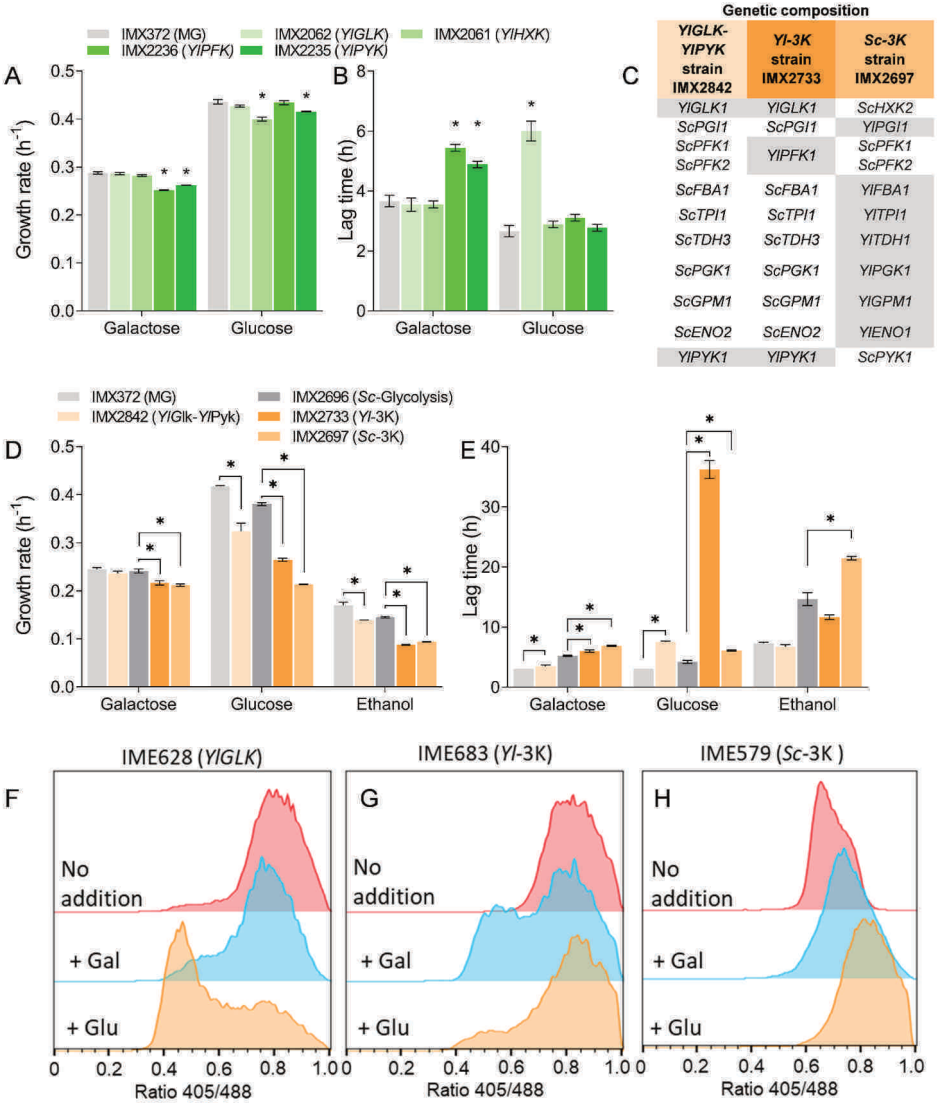


Figure 4 - Response of complementation and 'mosaic' glycolysis strains to glucose and galactose.

A) and **B)** Growth rate and lag phase duration of single *YlGLK*, *YlHXK, YlPFK* and *YlPYK* complementation strains. Measured in biological triplicates, * indicates significant differences to reference strain (t-test, homoscedastic, unpaired with P-value<0.05). **C)** Composition of the glycolysis of the *YlGLK-YlPYK, Yl*-3K and *Sc*-3K strains, *Y. lipolytica* genes are indicated in grey. **D)** and **E)** Growth rate and lag phase duration of the 'mosaic' glycolysis strains *YlGLK-YlPYK, Yl*-3K, *Sc*-3K and the control strains with native, minimized glycolysis IMX372 (MG) and IMX2696 (*Sc*-Glycolysis). Measured in biological triplicates, * indicates significant differences to reference strain (t-test, homoscedastic, unpaired with P-value<0.05). **F-H)** pHluorin-based pH_i profiles of strains grown on galactose transferred to glucose and Galactose. **F)** *YlGLK* complementation strain, **G)** *Yl*-3K strain and **H)** *Sc*-3K strain. Duplicate pHluorin experiments and time-course data are shown in Supplementary Fig. S2 and S8.

Simultaneous expression of deregulated glycolytic kinases reproduces the glycolytic imbalance phenotype

While individual expression of the *Y. lipolytica* key kinases did not affect *S. cerevisiae* phenotype, the coordinated regulation of these three kinases or a combination of two of them might still play an important role for metabolic adaptation to sugar transition. To test this hypothesis, several strains with 'mosaic glycolytic configurations' were constructed. Strain IMX2842 (*Yl*Glk-*Yl*Pyk) had the native *S. cerevisiae* glycolysis except for the *Y. lipolytica* glucokinase and pyruvate kinase. Strain IMX2733 (*Yl*-3K) harboured the native *S. cerevisiae* glycolysis but with the three *Y. lipolytica* kinases (glucokinase, phosphofructokinase and pyruvate kinase), while strain IMX2697 (*Sc*-3K strain) carried the *Y. lipolytica* glycolysis with the three *S. cerevisiae* kinases (Fig. 4C). The response to glucose exposure of these strains was radically different. The *Yl*Glk-*Yl*Pyk and *Sc*-3K strains showed phenotypes similar to the *S. cerevisiae* control strains with native glycolysis, albeit with a somewhat slower growth rate (78±4% and 56±0.2% of the control strain growth rate on glucose, Fig. 4D), possibly caused by lower activities of some of the *Y. lipolytica* enzymes in *S. cerevisiae*, as determined in the *Yl*-Glycolysis strain (Fig. 2G). Conversely, the *Yl*-3K strain showed a long lag phase reminiscent of the full *Yl*-Glycolysis strain, although the duration of this lag phase was clearly shorter (36±3 hours compared to 76±3 hours on average for the *Yl*-3K and *Yl*-Glycolysis strains, Fig. 4E). Also, the pH_i response upon exposure to glucose was not as marked in the *Yl*-3K as in the *Yl*-Glycolysis strain (Fig. 4 and Supplementary Fig. S2).

Despite this somewhat milder response of the *Yl*-3K strain as compared to the strain with full *Y. lipolytica* glycolysis to glucose medium, repeated glucose/galactose transfers again resulted in systematic mutations in *YlGLK1* (A394T, A471S and G270S) and a decrease in *in vitro* glucokinase activity (Fig. 5 A-D and Supplementary Fig. S3, single cell isolates IMS1207 to IMS1209). Characterization in shake flasks of these evolved strains showed low specific growth and glucose uptake rates, as compared to the *Sc*-Glycolysis control strain, but some ethanol production was observed, unlike for the evolved *Yl*-Glycolysis strain isolates (Fig. 5 E-H, Supplementary Fig. S6). The lower specific growth and glucose uptake rates of the evolved strains as compared to the *Sc*-Glycolysis control strain are most likely explained by a limited capacity of the mutated glucokinases. Indeed, in these strains the maximum glycolytic flux estimated from *in vitro* glucokinase activity was similar to the glycolytic flux estimated from the glucose uptake rate (Supplementary Fig. S6). These results brought new insight in the genotype to phenotype relationship of the *Yl*-glycolysis strain. Firstly, the absence of any phenotype during transition to glucose of the *Sc*-3K strain revealed that the relatively low activity of several of the glycolytic enzymes in the *Yl*-glycolysis strain was not causing the deficiency in

transitioning between carbon sources. However, these lower activities are probably responsible for the slower growth rate of the *Yl*-glycolysis strain as compared to the control strain with native *S. cerevisiae* glycolysis. Secondly, the combined expression of the three deregulated *Y. lipolytica* kinases was identified as the major cause of the growth and transition defects of the *Yl*-Glycolysis strain.

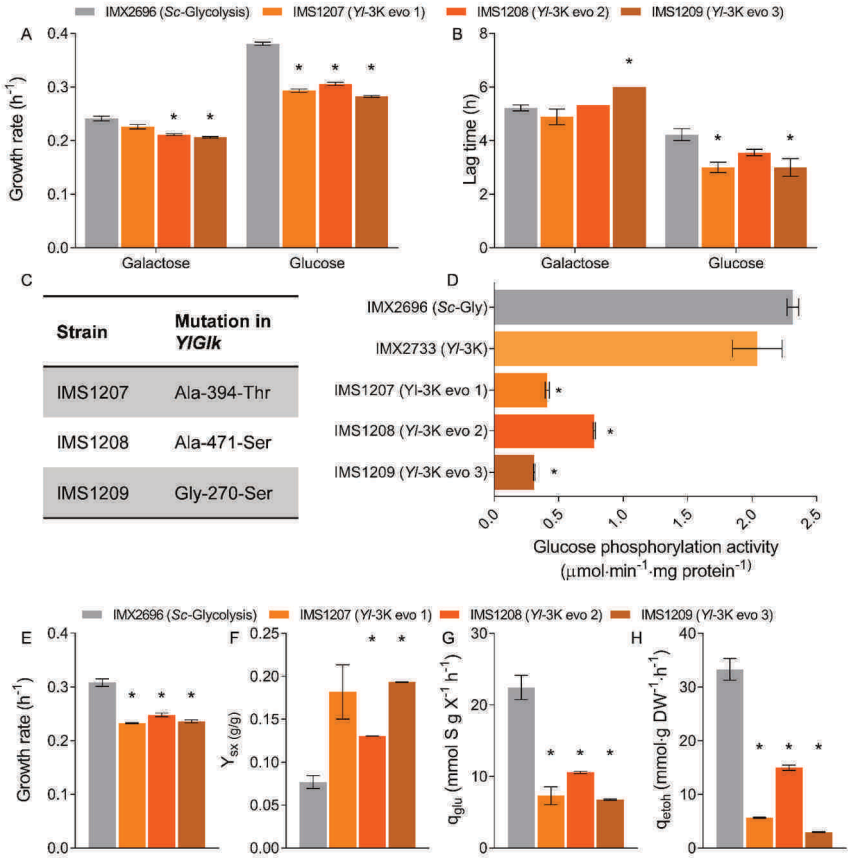


Figure 5 - Glucose grown isolates of a mosaic glycolysis strain and lag phase dependency on glucose concentration.

A) and **B)** Growth rate and lag phase duration on glucose and galactose medium of three independent isolates of the *Yl*-3K strain (IMX2733) that acquired the ability to grow on glucose, biological triplicates were measured. **C)** Mutations in the *Yl* glucokinase found in each of the three isolates. **D)** *In vitro* glucokinase activity of the *Yl*-3K strain and its evolved isolates determined from biological duplicate SM-Galactose cultures. **E-H)** Physiological characterization of the three evolved isolates of the *Yl-3*K strain and the *Sc*-Glycolysis reference strain on glucose urea medium in shake flasks in biological duplicates. The same data for *Sc*-Glycolysis strain IMX2696 is shown as in Fig. 3 F-I as control. Growth rates, biomass yield (Y_{SX}), specific glucose uptake rates (q_{glu}), and specific ethanol production rates (q_{etoh}) are shown. In all panels asterisks indicate significant differences compared to the relevant reference strain (t-test, homoscedastic, unpaired, with P-value<0.05).

Reducing the flux through top glycolysis compensates for the lack of metabolic regulation of the Y. lipolytica kinases in S. cerevisiae

The results suggested that lowering the glycolytic flux by limiting the rate of glucose phosphorylation was the optimal cellular strategy of the *Yl*-Glycolysis and *Yl*-3K strains to balance glycolysis under glucose excess. Kinetic modelling was used to explore whether reducing hexokinase activity could indeed compensate for the absence of *S. cerevisiae*-like regulation of the three kinases and to guide further experimental design. Studying a *tps1Δ* mutant, Van Heerden *et al*. demonstrated that the Emden-Meyerhof-Parnas glycolysis can switch between balanced and imbalanced states depending on the genetic and metabolic context [21]. Mutants with a *tps1* deletion cannot synthetize T6P, regulator of *S. cerevisiae* hexokinases, and cannot maintain inorganic phosphate (P_i) homeostasis, resulting in the inability to grow under high glucose conditions. Cellular levels of F1,6bP and inorganic phosphate (P_i) are key determinants of glycolysis stability in *tps1* mutants. Considering cell-to-cell variation in F1,6bP and P_i concentrations, upon exposure to glucose only the fraction of cells with the appropriate concentrations can reach a balanced state and resume growth. Van Heerden and colleagues used a kinetic model of glycolysis to describe and predict the start-up of glycolysis as a function of F1,6bP and P_i levels. The reference *S. cerevisiae* strain with native regulation displayed a balanced glycolysis for a wide range of physiologically-relevant F1,6bP and P_i levels (Fig. 6A, plot 1). Only at very low initial F1,6bP and P_i concentrations glycolytic imbalance occurred, resulting in intracellular accumulation of F1,6bP and ATP depletion (Supplementary Fig. S10). Such an imbalance would consequently lead to the inability to grow on glucose of the small fraction of the cellular population that occupied this concentration range at the moment of glucose exposure. Starting from this model, the regulations of the three kinases were removed, individually and simultaneously, to mimic *in silico* the replacement of *S. cerevisiae* kinases by their *Y. lipolytica* variants. Removing the regulations, individually (Fig. 6A plots 2 to 6) or simultaneously (plot 7) resulted in most cases in an increase in conditions leading to the in imbalanced state, with the strongest effect obtained by removing ATP inhibition of PFK or combining removal of multiple regulations. Combinatorial removal of multiple regulations also resulted in a decrease of balanced model outcomes (Supplementary Fig. S11), showing their overlapping effects. The strong impact of PFK regulation in the model did not match our experimental observations (Fig. 4A). PFK is an enzyme with complex regulatory pattern, and its activity has been proven to be difficult to mathematically describe in *S. cerevisiae* [7, 26]. Using an imperfect *S. cerevisiae* PFK model to reflect *Yarrowia*'s variant most likely exacerbates this problem, resulting in an *in silico* oversensitivity to phosphofructokinase regulation and activity. However, the overall modelling response agreed well with experimental data, as factors increasing the flux through HXK and PFK destabilized glycolysis,

while decreasing the flux through upper glycolysis (e.g. by decreasing activation on PFK) stabilized it. Accordingly, reducing the hexokinase V_{max} in the model resulted in higher glycolytic stability, with all model configurations reaching a balanced state at 10% of the original hexokinase activity (Fig. 6A plots 8-21).

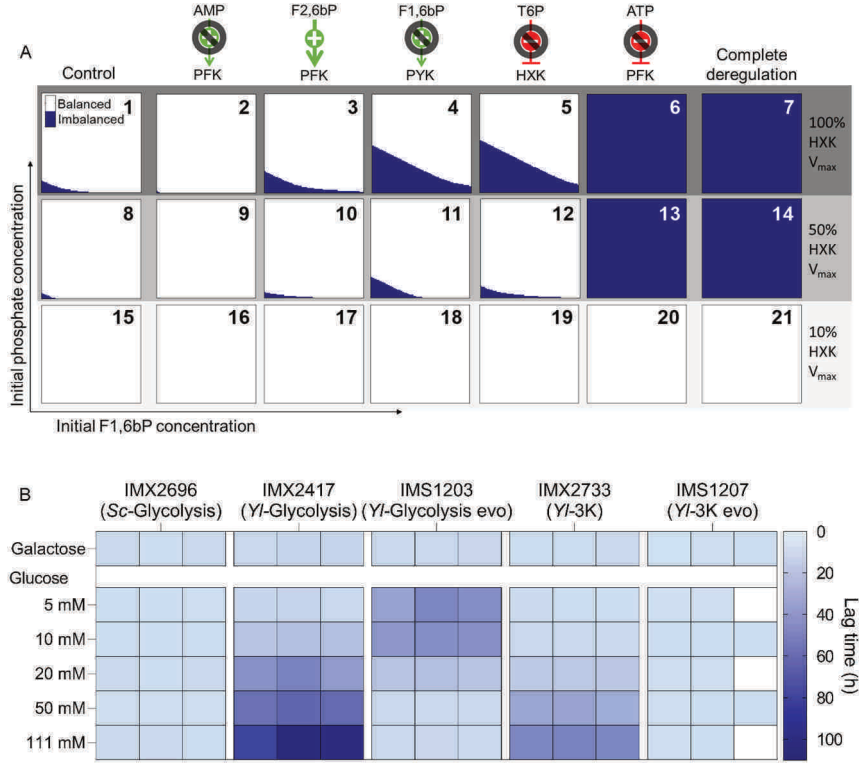


Figure 6 - Kinetic modelling to study the impact of deregulation of glycolysis.
A) Results from kinetic modelling analysis of balanced and imbalanced states. For each graph, the x-axis indicates the initial fructose-1,6-bisphosphate concentration (between 0 and 5 mM) and the y-axis the initial phosphate concentration (between 0 and 20 mM), with identical conditions used for each plot. White colouring indicates the model reaches a balanced steady state with those conditions within 100 minutes, dark blue colouring shows imbalance. In the top left (plot 1) the situation in the unaltered control model is shown, which has a small imbalanced region at low phosphate and F1,6bP concentrations. For each of five allosteric regulations in yeast glycolysis, the impact of their removal on the stability of the system is shown, as well as the impact of combinations of removal of all regulations simultaneously. Tested regulations were removal of AMP activation on PFK (2, 9, 16), constant F2,6bP activation of PFK (3, 10, 17) removal of F1,6bP activation on PYK (4, 11, 18), removal of T6P inhibition on HXK (5, 12, 19), removal of ATP inhibition of PFK (6, 13, 20) and combination of removal of all together (7, 14, 21). The lower rows show the impact of hexokinase activity on the stability of each system, with the second row (8-14) showing 50% hexokinase V_{max} and the bottom row (15-21) 10% hexokinase V_{max}. **B)** Dependency of the lag-phase duration of the *Yl*-Glycolysis and *Yl*-3K strains and two of their evolved isolates on the glucose concentration. 111 mM galactose is shown as control condition, growth rates and lag phase duration shown in Supplementary Fig. S12. Individual replicates are shown for each strain, those marked in white were not measured.

To experimentally test whether a too high flux in upper glycolysis was the main causal factor of the growth defect on glucose of the Yl-Glycolysis and Yl-3K strains, glucose uptake rate was tuned by exposing the strains to media with varying glucose concentrations (5, 10, 20, 50 and 111 mM, Fig. 6B and Supplementary Fig. S12). As expected, the lag phase of evolved strains with YlGLK mutations (IMS1203 and IMS1207) was mostly insensitive to glucose concentration. Conversely, for both Yl-Glycolysis and Yl-3K strains, while specific growth rate was only marginally affected by glucose concentrations, the lag phase duration was strongly positively correlated with glucose concentration. At the lowest tested glucose concentration of 5 mM, growth resumed upon transfer to glucose medium without lag phase, demonstrating that all cells can cope with slow glucose influx. These results suggest that slow glucose import rates mimicked the reduced glucokinase activity of the evolved strains, leading to a stable system. At the other extreme, high glucose concentrations (111 mM) prevent growth and, as demonstrated above in the Yl-Glycolysis and Yl-3K strains, require reduction of the glucokinase activity. At intermediate concentrations of 10 - 50 mM, shorter lag phases of 20 to 60 hours were observed. Considering the measured growth rates, a single mutant cell present at the start of the culture would reach measurable biomass concentration after ca. 70 hours of growth for Yl-Glycolysis strain IMX2417, a value corresponding to the lag phase of cultures with 111 mM glucose in which hexokinase mutations were systematically observed. The shorter lag phases measured at low glucose concentrations can therefore not be reasonably explained by the occurrence of mutations and outgrowth of the population by a mutant. Such a graded response to glucose concentration, reminiscent of tps1 mutants [21], suggested the result of deregulation of the glycolytic pathway was population heterogeneity with an increasing fraction of the population unable to grow due to glycolytic imbalance with increasing glucose concentration.

Discussion

The limits of single complementation to study metabolic regulation

While metabolic control theory has long suggested that modification of the activity of a single enzyme (by expression or regulation) is usually not sufficient to greatly change pathway flux [27, 28], single enzyme studies are still routinely performed. The present work illustrates that, despite technical challenges, the complexity of biology requires experimental investigation at the larger scale of pathway or function. The EMP pathway of glycolysis, widely distributed across kingdoms, is characterized by a structure in which ATP is first invested in the top of the pathway before it can be recouped further downstream. Such a configuration is prone to imbalance when cells have to transition between conditions poor and rich in hexoses [6], a problem

for which various solutions have emerged. Organisms such as the parasite *Trypanosoma* control ATP use by glycolysis by sequestering glycolytic enzymes in specialized compartments called glycosomes [29, 30]. Conversely, in many prokaryotes hexose uptake and phosphorylation are coupled to lower glycolysis via the phosphotransferase system [31, 32], while many eukaryotes are equipped with hexose phosphorylating enzymes inhibited by T6P or G6P [33, 34] that apparently keep upper glycolytic flux in check during fluctuating substrate supply. However, all studies reporting the complementation of *S. cerevisiae* Hxk2 by insensitive variants suggest that Hxk2 inhibition by T6P is not by itself essential for the regulation of the glycolytic flux [9, 35]. The same holds for pyruvate kinase and its activation by F1,6bP [10] and phosphofructokinase and its myriad of metabolic regulations [11, 12, 36]. This insensitivity of yeast physiology to alterations in glycolytic activity and allosteric regulation has been subject of debate for decades [37-40]. Similarly in prokaryotic model organisms, while overexpression and deregulation of phosphofructokinase and pyruvate kinase had some minor effects on flux distribution and intracellular metabolites, tight regulation of these reactions appears far from essential [41-43]. The widespread nature of allosteric regulation of several key glycolytic enzymes raises the question on the function and essentiality of these mechanisms. By combining the expression of multiple unregulated enzymes in *S. cerevisiae* glycolysis, the present study explored the essentiality of allosteric regulation using novel synthetic biology methods. Remarkable differences in phenotype were observed between *S. cerevisiae* strains expressing single, triple or the complete set of *Yarrowia* glycolytic enzymes. While it was confirmed that individual expression of insensitive enzyme variants could lead to functional complementation in *S. cerevisiae*, combination of three insensitive kinases did not. This study therefore identifies for the first time the essential but redundant functions of these metabolic regulations for allowing transitioning of *S. cerevisiae* between low sugar and high sugar environments. The results strongly suggest allosteric regulations in glycolysis, while not required for pathway operation per se, have been subject to strong evolutionary selection as they allow efficient and balanced pathway operation under dynamic conditions.

Mutations in *Y. lipolytica* glucokinase reduce glycolytic flux and act as a suppressor for a deregulated glycolysis

Replacement of *S. cerevisiae* by *Y. lipolytica* glycolysis resulted in a phenotype strongly suggesting glycolytic bistability. Single, double and triple kinase mutants and kinetic modelling demonstrated a strong synergetic role of hexokinase, phosphofructokinase and pyruvate kinase in transitioning between glucose-poor and glucose-rich medium. The graded response of *Yl*-Glycolysis and *Yl*-3K strains to sugar concentration strongly suggested the presence of two states in the yeast

population, a fraction able to cope with glucose inflow with a balanced glycolysis, and a fraction trapped in an imbalanced state. Such a stochastic effect, with a fraction of the population behaving differently from the majority, has previously been observed in yeast and other organisms during transition between carbon sources [21, 44]. Single-cell analysis studies using Phluorin or fluorescent ATP-sensors [45] as proxy for glycolytic (im)balance would enable more fine characterization of the population dynamics during transition at various glucose concentrations. Remarkably, the loss of these complex, synergetic regulations could be easily fixed by a single mutation of glucokinase. The six characterized mutants carried different mutations of the *YlGLK1* gene, with as common feature a strong reduction in the *in vitro* glucose phosphorylating activity, and concomitant low glucose uptake rate. A previous study demonstrated that functional complementation with the G6P-sensitive human HK1 and HK2 in *S. cerevisiae* also required single mutations in these hexokinases for growth on glucose [15]. These mutations were however clustered in a specific region of the protein, leading to the alleviation of G6P inhibition. *Yarrowia* glucokinase has no known effectors, and the random localization of the amino acid substitutions suggest that protein domains may not be relevant, as long as the substitutions enable the reduction of the glucose phosphorylating activity. This 'simple solution' comes however at a price, shown by the more than two-fold reduced glucose uptake rate of the *Yl*-3K strain as compared to the reference *S. cerevisiae* strain (Fig. 5G), while the specific activity of PFK and PYK was not expected to be different between the two strains (Fig. 2G). Thus while constantly imposing a slow flux in the top part of glycolysis allowed growth without defect under dynamic conditions, the maximal growth rate was limited. Both evolved and unevolved *Yl*-3K strains would likely be rapidly outcompeted by the reference *S. cerevisiae* strain with native regulations, which is capable of high glucose uptake rates and efficient transition to high glucose conditions. The evolved *Yl*-3K strains are however fantastic testbeds to explore the evolution of regulatory mechanisms, for instance by long term exposure to sugar transitions in adaptive laboratory evolution experiments. The strains constructed in this study also enable to explore the potential role of the regulatory synergy between the three kinases in dynamic situations beyond sugar transitions. This could be achieved by growing the evolved and non-evolved *Yl*-3K strains and strains with single kinase complementation in a broad range of dynamic growth conditions.

How does *Yarrowia lipolytica* cope with high hexose concentrations?

Replacement of the key regulatory enzymes in glycolysis with *Y. lipolytica* variants led to a dysfunctional pathway in *S. cerevisiae* during growth on glucose, unless the flux was reduced. The EMP pathway in *Y. lipolytica* has the same set of reactions, and therefore the same intrinsic risk of metabolic imbalance as its distant relative *S.*

cerevisiae, yet *Y. lipolytica* is capable of fast growth on media with high glucose concentrations (around 0.37 h^{-1} with over 55 mM of glucose [46]). The mechanisms enabling *Y. lipolytica* glycolysis to operate in its native context are therefore intriguing and most probably find their roots in the strong difference in ecological niches and life style of these two yeasts. While *S. cerevisiae* thrives in hexose rich environments and has a very limited carbon substrate range, *Y. lipolytica* is much more versatile and can use a broad range of carbon sources, such as glycerol, fatty acids and hydrocarbons, but is more limited in its utilization of sugars [47-49]. These nutritional preferences are likely reflected in the mechanisms controlling glycolysis. Two likely strategies that would enable *Yarrowia* to control the glycolytic flux without the stringent regulations on the regulatory key-point enzymes glucokinase, phosphofructokinase and pyruvate kinase that seem to be essential in *S. cerevisiae* are (i) imposing a low flux in the first, ATP-consuming steps of glycolysis or (ii) strictly controlling extracellular glucose influx in the cells. In line with the first idea, *in vitro* glycolytic enzyme activities in *Yarrowia lipolytica* are far lower than those found in *S. cerevisiae*. This is especially the case for upper glycolysis (Supplementary Fig. S13) with glucose phosphorylation activity 4-6 fold lower in *Y. lipolytica* as compared to *S. cerevisiae* (in line with previous studies [46]) and phosphofructokinase activity 3-4 fold lower (Supplementary Fig. S13). This lower rate of glucose phosphorylation and glycolysis as a whole might be sufficient for *Y. lipolytica* to cope with glucose fluctuations in its environment without imbalance in the pathway, as observed in the glucokinase mutant *Yl*-Glycolysis strains. Additionally, transport might further contribute to maintaining a slow glucose influx. Several hexose transporters have been identified in *Y. lipolytica*, however little is known about their sugar preference and kinetic properties. Whether import plays a role in controlling the glycolytic flux in *Y. lipolytica* can be explored by hexose transporters engineering, as performed earlier with *S. cerevisiae* [50-52]. Finally, considering the knowledge gap on *Y. lipolytica* lifestyle and glycolytic enzymes kinetic properties and regulation, we cannot rule out the possibility that some yet unidentified metabolites outside glycolysis are also involved in the regulation of the glycolytic flux in *Y. lipolytica*.

Materials and Methods

Strains and cultivation conditions

All *S. cerevisiae* strains used are derived from the CEN.PK lineage [53] and are listed in Supplementary Table S1. *Yarrowia lipolytica* strain W29 and CJM246 (PO1a) were used as controls for enzyme activity measurements. Yeast strains were grown on either YP medium containing 10 g L^{-1} Bacto Yeast extract and 20 g L^{-1} Bacto Peptone or synthetic medium (SM) containing 5 g L^{-1} $(NH_4)_2SO_4$, 3 g L^{-1} KH_2PO_4, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 1 mL L^{-1} of a trace element solution and 1 mL of a vitamin solution

added after autoclaving [54]. For physiological characterization of the *Sc*-Glycolysis and evolved isolates of the *Yl*-Glycolyis and *Yl*-3K strains in shake flasks (Fig. 3 F-I and Fig. 5 E-H), SMD-urea was used, where the $(NH_4)_2SO_4$ was replaced with 6.6 g L^{-1} K_2SO_4 as source of sulfate and 2.3 g L^{-1} urea. Similarly, when the dominant marker *amdS* was used, $(NH_4)_2SO_4$ was replaced by 6.6 g L^{-1} K_2SO_4 and 1.8 g L^{-1} acetamide. The pH of SM was set to 6 prior to autoclaving by addition of KOH. Media were supplemented with relevant carbon sources after autoclaving, final concentrations were 20 g L^{-1} glucose (YPD/SMD), 20 g L^{-1} galactose (YP-Gal/SM-Gal), 2% (v/v) ethanol (YPE/SME) or 1% (v/v) ethanol and 1% (v/v) glycerol (YPEG/SMEG). For solid media 2% (w/v) agar was added before heat sterilization. For counter selection of the *amdS* marker, 2.3 g L^{-1} fluoroacetamide was added to SM [55] and for selection on the *KanMX* and *hphNT1*, 200 mg L^{-1} G418 or 200 mg L^{-1} hygromycin (Hyg) were added to YP medium respectively. Yeast cultures were grown at 30 °C at 200 rpm in an Innova 44 incubator shaker (New Brunswick Scientific, Edison, NJ, USA) in 50-/100-/500-mL shake flasks containing respectively 10-/20-/100 mL of medium. For plasmid propagation and maintenance *Escherichia coli* XL1-Blue cells (Agilent Technologies, Santa Clara, CA, USA) were used grown in Lysogeny Broth (LB) medium containing: 10 g L^{-1} tryptone, 5.0 g L^{-1} yeast extract and 4 g L^{-1} NaCl supplemented with 100 mg mL^{-1} ampicillin, 25 mg L^{-1} chloramphenicol or 50 mg L^{-1} kanamycin when required. Yeast and *E. coli* strains were stored at -80 °C in 1 mL aliquots of appropriate medium after addition of 30% (v/v) glycerol.

Molecular biology techniques

PCR amplification for strain construction purposes was performed using Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA), according to the manufacturers recommendations except that the primer concentration was lowered to 0.2 µM. Diagnostic PCR amplification was performed using the DreamTaq PCR Master Mix (Thermo Fisher Scientific) following the supplier's recommendations. Primers for cloning purposes were ordered PAGE purified, all other primers were ordered desalted (Sigma-Aldrich, St. Louis, MO, USA). To obtain double stranded gRNA and repair fragments the designed forward and reverse oligo's, ordered as PAGE purified primers, were incubated at 95°C for 5 min and allowed to cool to room temperature. PCR products were separated by gel electrophoresis with gels containing 1% agarose (TopVision Agarose, Thermo Fisher Scientific) in 1x Tris-acetate-EDTA (TAE) buffer (Thermo Fisher Scientific), 10 µl L^{-1} SERVA (SERVA Electrophoresis GmbH, Heidelberg, Germany) was added to the gel for DNA staining. As size standard the GeneRuler DNA Ladder Mix (Sigma-Aldrich) was used. PCR fragments used for cloning obtained from plasmids were treated with addition of 1 µL DpnI FastDigest restriction enzyme (Thermo Fisher

Scientific) for 1 hour to remove remaining template DNA. DNA was purified using the GenElute PCR Clean-Up kit (Sigma-Aldrich) or the GeneJET PCR Purification Kit (Thermo Fisher Scientific) when no unspecific bands were present, otherwise products were purified from gel using the Zymoclean Gel DNA Recovery kit (Zymo Research, Irvine, CA). Purity and quantity of DNA was assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), for more precise DNA quantification the Qubit dsDNA BR Assay kit (Thermo Fisher Scientific) in combination with the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) was used.

Cloning of promoters, genes and terminators was done using Golden Gate assembly or *in vivo* assembly in yeast. For Golden Gate assembly per reaction volume of 10 µl, 1 µl T4 buffer (Thermo Fisher Scientific), 0.5 µl T7 DNA ligase (New England Biolabs, Ipswich, MA) and 0.5 µl BsaI (Eco31I) (Thermo Fisher Scientific) or BsmBI (NEB) was used and DNA parts were added in equimolar amounts of 20 fmol as previously described [56]. *In vivo* assembly of plasmids in yeast was performed according to [57] using 60 bp homologous flanks added by PCR and transformation of all fragments to *S. cerevisiae*. After transformation and colony PCR verification of correct assembly, plasmids were purified using the Zymoprep Yeast Plasmid Miniprep (II) kit (Zymo Research). Gibson assembly for construction of gRNA plasmids was performed with Gibson assembly master mix 2x (New England Biolabs, Ipswich, MA) according to the manufacturer's instructions but scaled down to a final volume of 5 µl.

Plasmids were transformed into *E.coli* XL1-Blue by chemical transformation for amplification [58]. Plasmids were isolated using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich) or the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) and verified by diagnostic PCR or restriction analysis using FastDigest restriction enzymes with FastDigest Green Buffer (Thermo Fisher Scientific) according to the manufacturer's instructions. Transformation of *S. cerevisiae* was done with the lithium acetate/PEG/ssDNA method [59], and colonies were verified by diagnostic PCR. Single colony isolates were obtained by three consecutive re-streaks on selective solid medium. Yeast DNA was isolated by either boiling in 0.02 N NaOH, the protocol described by Lõoke *et al.* [60], or the QIAGEN Blood & Cell Culture Kit with 100/G Genomic-tips (Qiagen, Hilden, Germany) depending on the desired DNA purity.

Plasmid and strain construction

Isolation of the Yarrowia lipolytica glycolytic genes and plasmid construction
Gene, cDNA and protein sequences of the *Y. lipolytica* glycolysis were obtained from the Genome Resources for Yeast Chromosomes database (https://gryc.inra.fr). The

genes *YlPGI1, YlFBA1, YlTPI1, YlTDH1, YlGPM1*, and *YlENO1* were identified based on sequence similarity to the *S. cerevisiae* glycolytic genes (Table 1). A *Yarrowia lipolytica* cDNA library from strain W29 [61] kindly provided by C.-L. Flores, was used for the cloning of glycolytic gene sequences except the genes for *HXK1*, *GLK1*, *PFK1* and *PYK1*, which were obtained codon-optimized for *S. cerevisiae* by GeneArt Gene Synthesis (Thermo Fisher Scientific). *Y. lipolytica* genes were amplified from the cDNA library using the primers listed in Supplementary Table S4A and subsequently assembled into part plasmids using Golden Gate assembly with the pUD565 entry vector. Subsequently the genes were assembled into transcriptional cassettes with *S. cerevisiae* promotors and terminators using Golden Gate assembly (*YlFBA1, YlPGK1*) or *in vivo* yeast assembly (*YlENO1, YlPGI1, YlTPI1, YlTDH1* and *YlGPM1*), used promotors and terminators are shown in Supplementary Table S2. For both methods the backbone was obtained from plasmid pGGKd017 [62] and *S. cerevisiae* promotors previously assembled onto part plasmids (Supplementary Table S3, [15]). For *in vivo* assembly the pGGKd017 backbone and required *S. cerevisiae* promotor and terminator parts were amplified, as well as the *Y. lipolytica* genes, with 60 bp homologous regions. The genes that were synthesized codon-optimized (*YlHXK1, YlGLK1, YlPFK1* and *YlPYK1*) were assembled using Golden Gate Assembly directly with plasmid pGGKd002 and promotor and terminator part plasmids to obtain plasmids that could be integrated in the *S. cerevisiae* genome directly after linearization (Supplementary Table S3). CRISPR-Cas9 yeast genetic engineering and construction of gRNA plasmids was performed as described [63].

Table 1 - *Yarrowia lipolytica* glycolytic gene accession numbers

Gene	Previously functionally annotated	Accession number GRYC
YlGLK1	[18]	YALI0E15488g
YlHXK1	[17]	YALI0B22308g
YlPGI1	No	YALI0F07711g
YlPFK1	[11]	YALI0D16357g
YlFBA1	No	YALI0E26004g
YlTPI1	No	YALI0F05214g
YlTDH1	[64]	YALI0C06369g
YlPGK1	[65]	YALI0D12400g
YlGPM1	No	YALI0B02728g
YlENO1	No	YALI0F16819g
YlPYK1	[66]	YALI0F09185g

Similarly, to facilitate amplification of the native *S. cerevisiae* glycolytic gene cassettes, these were amplified from CEN.PK113-7D with their native promotor and terminator sequences and assembled into glycolytic expression cassettes using golden gate assembly with dropout vector pGGKd017 (Supplementary Table S3).

Construction of a full Y. lipolytica *glycolysis expression strain and* Sc*-glycolysis reference strain*

For strain construction Cas9 expressing background strains were used to facilitate CRISPR/Cas9 genome editing according to [63]. An overview of the most important genetic modifications and strain pedigree is shown in Fig. S14. To facilitate the construction of strains expressing various combinations of *Y. lipolytica* and *S. cerevisiae* glycolytic genes, first the seven glycolytic genes excluding the key kinases (*YlPGI1, YlFBA1, YlTPI1, YlTDH1, YlPGK1, YlGPM1* and *YlENO2*) were integrated in the *CAN1* locus of platform strain IMX1822 together with the *S. cerevisiae ADH1* and *PDC1* genes for alcoholic fermentation. These genes were amplified from their respective expression cassette plasmids flanked by *S. cerevisiae* promotors and terminators and 60 base pair homology flanks were added during PCR to enable *in vivo* assembly (primers in Supplementary Table S4B). 200-350 ng of each repair fragment and 1 µg of gRNA plasmid pMEL13 targeting *CAN1* was used to transform SwYG strain IMX1822. The resulting strain was verified by diagnostic PCR to contain all integrated genes (diagnostic primers in Supplementary Table S4E) and named IMX2065. Subsequently the *E. coli* GroEL and GroES chaperone genes were integrated in the X2 locus and the *URA3* marker gene was deleted to enable selection on uracil prototrophy. To this end gRNA plasmid pUDR591 was constructed by Gibson assembly of the pROS13 backbone and a 2µ fragment amplified by primers containing the gRNA sequences targeting X2 and *URA3* (primers in Supplementary Table S4F). pUDR591 was co-transformed with the chaperone gene cassettes amplified from pUDE232 and pUDE233 and a repair fragment for the *URA3* locus. After selection on YPD-G418, this resulted in strain IMX2151. In this strain the *Y. lipolytica* glycolytic kinases *YlGLK1, YlPFK1* and *YlPYK1* were integrated in the *SPR3* locus by transformation of the three expression cassettes amplified from pUDI226-pUDI228 and the gRNA plasmid pUDR596 targeting the *SPR3* locus. pUDR596 was constructed by Gibson assembly of the pROS10 plasmid backbone with a 2µ fragment flanked by the *SPR3* gRNA sequence (Supplementary Table S4F). The gRNA plasmid was removed by non-selective growth and the uracil auxotrophic strain was stocked as IMX2333.

To remove the native *S. cerevisiae* glycolytic genes and the *AmdS* marker from the *sga1* locus, IMX2333 was transformed with recycle plasmid pUDE342 and a repair fragment (counter select oligo) and selected on SM-EtOH supplemented with fluoroacetamide to select against the presence of the *AmdS* marker in the *sga1*

locus. After PCR verification and removal of the gRNA plasmid, the strain, which only carries *Y. lipolytica* glycolytic genes, was stocked as IMX2363. This strain was verified by whole genome sequencing. In this strain the *ura3-52* locus was repaired to generate a prototrophic strain for physiological characterization by transformation of a *URA3* gene cassette amplified from plasmid pYTK074, the correct integration of *URA3* was verified by Sanger sequencing of the amplified *URA3* locus and the resulting strain was named IMX2417.

To generate a comparable control strain with the native *S. cerevisiae* glycolytic genes the *E. coli* GroEL/ES genes were integrated in strain IMX1821, which carries the native glycolytic genes in the *can1* locus. The chaperone genes were integrated in the X2 locus by co-transformation of gRNA plasmid pUDR547 and the GroEL/ES expression cassettes and selection on hygromycin, leading to strain IMX2696, which was similarly verified by diagnostic PCR.

Construction of mosaic glycolysis and single complementation strains

To verify the function of the *Y. lipolytica* glycolytic genes several strains with combinations of the *Y. lipolytica* and *S. cervevisiae* glycolysis genes were constructed. Single complementation strains expressing the *YlHXK, YlGLK, YlPFK* and *YlPYK* were constructed by transformation of 750 ng of the plasmids pUDI225-pUDI228, linearized by NotI digestion (FastDigest, Thermo Fisher Scientific) according to the manufacturers protocol, to uracil auxotrophic MG strain IMX1076 [15]. The linear plasmids were integrated in the disrupted *ura3-52* locus and strains were selected on SMD medium. Integration was verified by diagnostic PCR. In the resulting strains (IMX2047 - IMX2050) the native yeast glycolytic genes were subsequently removed in a second round of transformation using CRISPR/Cas9. *HXK2* was deleted by transformation with pUDR371, *PFK1* and *PFK2* were deleted by transformation with pUDR265 and *PYK1* was deleted by transformation with the pMEL13 backbone and the *PYK1* gRNA for *in vivo* assembly. Cells were plated and restreaked on YP-Ethanol supplemented with G418 to avoid selection for mutations allowing growth on glucose. This resulted in strains IMX2061 (*YlHXK* complementation), IMX2062 (*YlGLK* complementation), IMX2236 (*YlPFK* complementation) and IMX2235 (*YlPYK* complementation). Strains IMX2236 and IMX2235 were verified by whole genome sequencing.

To construct a double *YlGLK-YlPYK* complementation strain, *YlPYK* complementation strain IMX2235 was transformed with *HXK2* gRNA plasmid pUDR371 and *HXK2* deletion repair fragments to delete the native *HXK2* gene, generating strain IMX2812 after selection on YP-Gal supplemented with G418. Subsequently, in this strain the *YlGLK1* gene was integrated in the X2 locus by transformation with gRNA plasmid pUDR547, targeting the X2 locus and a repair

fragment containing the *YlGLK1* gene flanked with homology flanks to the X2 locus. After selection on YP-Gal with hygromycin, verification by PCR and restreaking the strain was stocked as IMX2842.

To make the *Sc*-3K strain, the *S. cerevisiae* kinases *ScHXK2, ScPFK1, ScPFK2* and *ScPYK1* and the seven *Y. lipolytica* genes *YlPGI1, YlPGK1, YlTDH1, YlENO1, YlGPM1, YlGPM1,YlGPM1* and *YlTPI1*, were integrated in one transformation step in the *CAN1* locus of SwYG strain IMX589. To this end the *Yl* glycolytic gene cassettes were amplified from the expression cassette plasmids with homology flanks as before but with some modifications to the flanks to incorporate the *S. cerevisiae* gene cassettes (primers in Supplementary Tables S4B and C). The fragments were co-transformed with the *URA3* containing plasmid pMEL10, the resulting strain was selected on SM-Ethanol and named IMX1751. In this strain the native glycolytic genes were removed from the *sga1* locus by transformation with the recycle plasmid pUDE342 and selection on SM-Ethanol-fluoroacetamide, leading to strain IMX1803, which was verified by whole genome sequencing. This strain was subsequently transformed with a *URA3* gene fragment to repair the *URA3* locus and generate a prototrophic strain as described above, leading to strain IMX2465. To make this strain directly comparable to the fully swapped glycolysis strain the *E. coli* chaperones GroEL and GroES were integrated in the *X2* locus by co-transformations of the GroEL and GroES expression cassettes and X-2 gRNA plasmid pUDR547, leading to strain IMX2697.

Similarly a pathway with the *Yarrowia lipolytica* kinases *YlGLK*, *YlPFK* and *YlPYK* and seven *S. cerevisiae* genes was designed. As background strain the *E. coli* chaperones GroEL and GroES were integrated first in SwYG strain IMX589, again by co-transformation of pUDR547 and the chaperone expression cassettes, leading to strain IMX2694. Integration of the mosaic glycolytic pathway in this strain was done by co-transformation of the expression cassettes with pMEL13 targeting the *CAN1* locus, selection was performed on YP-Gal supplemented with G418, this led to strain IMX2703. Deletion of the native glycolytic genes from *sga1* in the same manner described above with selection on SM-Gal-fluoroacetamide led to strain IMX2718. In this strain the *ura3* locus was repaired in the same manner described above to obtain a uracil prototrophic strain, named IMX2733 which was also described as the *Yl*-3K strain.

Finally a strain with the full *Yarrowia* glycolysis except the phosphofructokinase was constructed, starting from intermediate strain IMX2151, which carries the seven *Yarrowia* glycolytic genes excluding the three key kinases, and the GroES and GroEL chaperone genes. Similar to the construction of the full *Yarrowia* glycolysis strain, expression cassettes for the keypoints, *YlGLK1, ScPFK1, ScPFK2,* and

YlPYK1, were integrated in the *SPR3* locus. Cassettes were amplified from pUDI226 (*YlGLK1*), pUDE769 (*ScPFK1*), pUDE770 (*ScPFK2*) and pUDI228 (*YlPYK1*), and flanks were adapted to allow *in vivo* assembly of these four genes in the *SPR3* locus. Transformation with gRNA plasmid pUDR596 and the expression cassettes resulted in strain IMX2164 after PCR verification and single colony isolation. From this strain the native glycolytic genes were removed similar to described above, by co-transformation of gRNA plasmid pUDE342 and repair fragments with homology to the *SGA1* locus, resulting in strain IMX2182 after selection on SM-EtOH and counterselection of the AmdS marker on fluoroacetamide. This strain was verified by whole genome sequencing. To generate a prototrophic strain the *URA3* marker was repaired by transformation with the *URA3* fragment amplified from pYTK074 and selection on SM-Galactose.

Construction tps1 *strain*

To verify the function of pHluorin we constructed a *tps1* deletion strain in the CEN.PK113-7D background. To enable deletion of *TPS1* we constructed gRNA plasmid pUDR626 using Gibson assembly with the pMEL13 backbone and a double stranded *TPS1* gRNA fragment (Supplementary Table S4F). This plasmid was transformed to Cas9 expressing strain IMX581 together with a double stranded repair fragment and selected on YP-Gal-G418 and the deletion was verified by PCR. The resulting strain was stocked as IMX2243.

Construction of pHluorin expressing strains

To enable estimation of the intracellular pH the plasmid pYES2-P_{ACT1}-pHluorin [23], which was kindly shared by Bas Teusink, was transformed into several uracil auxotrophic strains (Table 2). Selection was performed in each case on SM-Galactose or SM-Ethanol medium and presence of the plasmid was verified by observation of fluorescence. For the single complementation strains the *URA3* marker was deleted to allow transformation of this plasmid by transformation with *URA3* gRNA plasmid pUDR107 and selection on YP-galactose-hygromycin, generating strains IMX2549-IMX2551.

Table 2 - pHluorin expressing strains
Uracil auxotrophic strains with various genotypes were transformed with plasmid pYES2-P_{ACT1}-pHluorin, host and resulting strain are indicated.

Strain characteristic genotype	Uracil auxotrophic host strain	pHluorin expressing strain
Control strain	CEN.PK113-5D	IME480
Sc-Glycolysis	IMX589	IME481
tps1 control strain	IMX2243	IME576
Full *Yl*-glycolysis	IMX2363	IME577
Sc-3K strain	IMX1803	IME579
Yl-3K strain	IMX2718	IME683
YlHXK complementation strain	IMX2549	IME627
YlGLK complementation strain	IMX2550	IME628
YlPFK complementation strain	IMX2552	IME631
YlPYK complementation strain	IMX2551	IME632
Yl-glycolysis *ScPFK*	IMX2182	IME609

Growth rate and lag phase determination and adaptive evolution

Growth profiler

For growth rate and lag phase measurements growth cultures were grown at 30 °C and 250 rpm using a Growth Profiler 960 (EnzyScreen BV, Heemstede, The Netherlands). Strains were inoculated from glycerol freezer stocks and grown overnight in 10 mL YP-Gal medium in shake flask. These cultures were transferred to 20 mL SM-Gal medium which was grown until mid-exponential growth. From this culture cells were re-suspended in SM without added carbon source and inoculated in 96-well square-well microtiter plates (EnzyScreen, type CR1496dl or CR1496dg), pre-filled with appropriate media, with final working volumes of 250 µL to a starting OD_{660} of 0.2. Microtiter plates were closed with a sandwich cover (EnzyScreen, type CR1296). Images of cultures were made at 20 minute intervals. Green values for each well were corrected for the position in the plate using measurements of a culture of OD_{660} of 5 of control strain CEN.PK113-7D. Corrected green values were converted to OD-values based on calibration measurements with the control strain CEN.PK113-7D, fitted with the following equation: OD-equivalent = $a \times GV(t) + b \times GV(t)^c - d$ in which GV(t) is the corrected green-value measured in a well at time point 't'. This resulted in curves with the following values for a, b, c and d: 0.07742; $1.662*10^{-7}$; 3.624; -1.615 for plates of the CR1496dl type, and 0.09622; $5.968*10^{-6}$;

3.254; -0.7939 for plates of the CR1496dg type. Growth rates were calculated in a time frame where the calculated OD was between 1 and 10 in which OD doubled at least twice except for the low glucose experiments (Fig. 6B), where cell-densities remained low. Linear regression of the log-transformed OD data versus time was used to determine the growth rate.

Shake flask growth characterization

Growth rates and extracellular metabolite consumption and production were estimated from duplicate 100 mL shake flask cultures on SMD-urea. OD_{660} was measured on a JENWAY 7200 spectrophotometer (Cole-Parmer, Stone, UK). Wake-up cultures were inoculated in 10 mL YP-Gal medium and grown overnight, from there pre-cultures were inoculated in 20 mL SMD and grown until exponential phase and transferred to SMD-urea. Samples were taken and OD_{660} was measured and 1 mL samples were centrifuged for 5 min at 20000g for extracellular metabolite determination. The supernatants were analysed using an Aminex HPX-87 ion-exchange column operated at a 60°C and a flow rate of 0.6 mL/min with 5 mM H_2SO_4 as mobile phase (Agilent). Biomass dry weights were estimated from a correlation with dry weights measured on filters with pore-size 0.45 µm as described previously [54]. Growth rates were determined by linear regression on log-linear OD_{660} data over at least six consecutive points, over which the optical density doubled twice. The optimal range was chosen by maximization of the R^2. Molar yields were estimated as the slope of the product concentration versus glucose concentration during the exponential phase. The specific substrate uptake rate was estimated by dividing the growth rate with the biomass yield on glucose. Specific ethanol production was estimated by multiplying the molar yield with the specific glucose uptake rate.

Shake flask adaptive evolution

Tests for adaptation on glucose medium of strains IMX2417, IMX2733 and IMX2062 were performed in shake flasks. Pre-cultures were inoculated in 20 mL non-selective YP-Gal medium and grown overnight. From there cultures were transferred to 20 mL SM-Gal and grown until exponential phase. Exponential SM-Gal cultures were inoculated in triplicate 100 mL SMD cultures to a starting OD_{660} of 0.2. After growth on glucose containing medium cultures were transferred to 100 mL SM-Gal cultures, again to a starting OD_{660} of 0.2. After growth on SM-Gal cultures were re-inoculated at OD_{660} 0.2 on SMD medium to verify whether a lag phase was still present. From each of these SMD cultures single colonies were isolated by triplicate restreaking on solid SMD medium. A single isolate of each of the three shake flasks was stocked for each experiment, resulting in strains IMS1203, IMS1204 and IMS1205 from IMX2417; IMS1207, IMS1208 and IMS1209 from IMX2733 and IMS1218, IMS1219 and IMS1220 from IMX2062.

Sequencing

High quality genomic DNA was isolated with the QIAGEN Blood & Cell Culture Kit with 100/G Genomic-tips (Qiagen) from strains IMX2363, IMX2182, IMX1803, IMX2063 and IMX2064 and sequenced in-house using an Illumina MiSeq Sequencer (Illumina, San Diego, CA) as described previously [67, 68]. For IMS1202, IMS1203 and IMS1204 DNA was obtained in the same manner but sequenced at NovoGene (NovoGene, Leiden, The Netherlands).

A *de novo* assembled reference genome was previously constructed for IMX589 (auxothrophic SwYG) using MinION and Miseq data. Using the Burrows-Wheeler Alignment (BWA) Tool [69] (version 0.7.15), sequencing data of SwYG-derived strains (IMX2363, IMX2182, IMX1803, IMS1202, IMS1203 and IMS1204) was aligned to the IMX589 reference genome and sequencing data of all strains was additionally aligned to a CEN.PK113-7D reference [70]. The data was further processed using SAMTools [69] (version 1.3.1) and single nucleotide polymorphisms (SNPs) were determined using Pilon (with -vcf setting; version 1.18)) [71]. The BWA.bam output file was visualized using the Integrative Genomics Viewer (version 2.4.0) [72], and copy numbers were estimated using Magnolya (version 0.15) [73]. SNPs were compared between previously obtained sequencing data of parental SwYG strains IMX589 and IMX605 [13] and the SwYG-derived strains to verify the absence of mutations during strain construction and between the unevolved *Yl*-Glycolysis strain IMX2363 and evolved strains IMS1202, IMS1203 and IMS1204 to find mutations after growth on glucose (Fig. 3C and D). Sanger sequencing was performed at Baseclear BV (Baseclear, Leiden, The Netherlands) and Macrogen (Macrogen Europe, Amsterdam, The Netherlands). PCR amplified fragments of the *YlGLK1* gene were sequenced from *Yl*-3K strains IMS1207, IMS1208 and IMS1209 and *YlGLK* complementation strains IMS1218, IMS1219 and IMS1220 with the primers listed in Supplementary Table S4E.

pHluorin pH_i determinations

pHluorin pH_i response was verified by measurement of the fluorescence signal in control strain IME480 after permeabilization by incubation with digitonin in Citrate-Na_2PO_4 buffers with known pH, as in [23] in a TECAN infinite M200 Pro microtiter plate reader (TECAN, Männedorf, Switzerland) (Supplementary Fig. S1). Flow cytometry for pHluorin fluorescence ratios was performed on a BD FACSCelesta (Becton Dickinson Biosciences, Breda, The Netherlands). Excitation was by a 405-nm laser (Violet) and a 488-nm laser (Blue) and emission was detected through BD Horizon Brilliant Violet 510 (525/50 nm) filter and a BD Horizon Brilliant Blue 515, FITC (530/30 nm) filter, FlowJo™ v10 (BD Biosciences) was used to analyze and visualize FACS data. pHluorin expressing strains and the CEN.PK113-7D non-fluorescent control strain were inoculated from freezer stocks in SM-Gal medium.

These cultures were transferred to 15 mL SM-Gal which was grown overnight to mid-exponential phase. Exponential cultures were harvested by centrifugation at 5000 g and washed in 10 mL SM without C-source. Cultures were diluted to an OD_{660} of 0.5 and 260 µL aliquots were placed in round bottom 96-well microtiter plates. Glucose or galactose was added to a final concentration of 20 g L^{-1}. For the time-course measurements there aws approximately 1 minute between addition of sugar and start of the measurement. The control strains IME480 and IME481 and the *tps1Δ* strain IME576 were first tested for their pH_i response (Supplementary Fig. S1). At least 20000 events were measured for each condition, fluorescent cells were gated based on fluorescence in both channels by comparing with the non-fluorescent control strain. Events on the edges (maximum detectable fluorescence) were removed to avoid skewing the ratio. Settings and voltages were kept the same for all experiments to verify reproducible ratio's.

Cell free extract preparation and enzyme assays

S. cerevisiae samples for enzyme activity determinations were prepared as previously described [74], from exponentially growing cultures (approx. 62 mg dry weight per sample) from shake flask. For *Yarrowia lipolytica* (Supplementary Fig. S13) a similar procedure was followed but approximately double the amount of biomass was sampled (based on OD_{660}). All determinations were performed at 30°C and 340 nm (εNAD(P)H at 340 nm/6.33 mM^{-1}). Glycolytic V_{max} enzyme activities were determined in 1 mL reaction volume in 2 mL cuvettes, using a Hitachi model 100-60 spectrophotometer, using previously described assays [75], except for phosphofructokinase activity which was determined according to Cruz *et al.*[76]. *Y. lipolytica* glucokinase activity was assayed with increased glucose and ATP concentrations in strains IMX2417, IMX2733, IMS1202, IMS1203, IMS1204, IMS1207, IMS1208 and IMS1209 (Supplementary Fig. S5). The reported data are based on at least two independent biological replicate samples, with at least two analytic replicates per sample per assay with different cell free extract concentrations except the measurements at higher ATP concentrations (where no effect was seen). The protein concentration was determined using the Lowry method with bovine serum albumin as a standard [77]. Enzyme activities are expressed as µmol substrate converted (mg protein)$^{-1}$ min^{-1}.

Kinetic modelling of glycolysis

The kinetic model of yeast glycolysis of Van Heerden et al. [21] was obtained from *jjj.bio.vu.nl/models/vanheerden1* and imported as a system of ordinary differential equations in Python using PySceS [78]. The system of ODE's was solved using the solve_ivp function in Python 3.6. Adaptations to the rate equations were made as follows. Removal of trehalose-6-phosphate inhibition of hexokinase was performed

by removing the G6P inhibition term $\frac{G6P}{K_{i,G6P}}$ in the hexokinase rate equation, since T6P inhibition is modelled as G6P inhibition. ATP inhibition of phosphofructokinase was removed by changing parameter C_i,ATP from 100 to 1. AMP activation of PFK was removed by changing parameter C_i,AMP from 0.0845 to 1. Fructose-2,6-bisphosphate activation of PFK was 0.02 to 0.1. Fructose-1,6-bisphosphate activation on PYK was removed by removing the F1,6bP activation term $\frac{F16bP}{K_{m,F16bP}}$ from the pyruvate kinase rate equation.

The various model configurations were solved with initial phosphate and FBP concentrations ranging between 0 and 20 and 0 and 5 respectively. 100 different concentrations were run for each metabolite, resulting in 10,000 initial conditions for each model configuration. Steady state was evaluated after 100 minutes by checking if the FBP concentration changed more than 1% of its original initial concentration over the last 5 simulated time points.

Acknowledgements

We thank Marijke Luttik for technical assistance and enzyme activity measurements of *Yarrowia lipolytica* strains, Erik de Hulster for assistance with growth profiler measurements and Nigell de Ronde for strain and plasmid construction. We thank Diederik Laman-Trip and Diego Gomez-Alvarez for assistance with flow cytometry. We thank Carlos Gancedo and Bas Teusink for insightful commentary and discussion.

Supplementary data

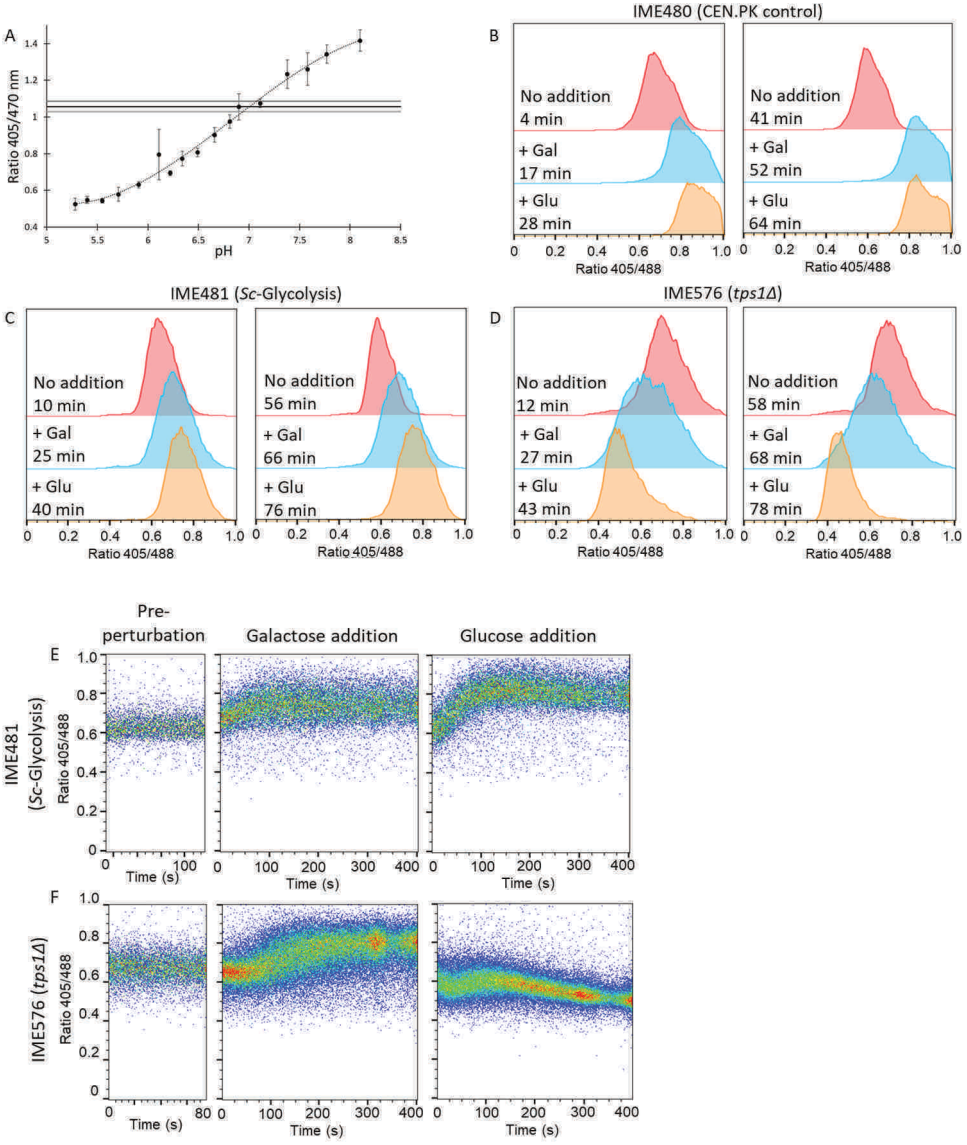


Figure S1 - Verification of pHluorin function.
Expression of pHluorin and its response to pH changes were verified. **A)** Strain IME480, a reference strain expressing pHluorin was incubated in presence of digitonin to permeate the cell membrane in Citrate-Na_2PO_4 buffers of known pH. The ratio of fluorescence intensity at 512 nm after excitation at 405 and 470 nm was determined. Triplicate wells were measured for each pH, mean and standard deviation are shown and a cubic trendline is plotted through the points. The horizontal line represents the signal measured in non-permeabilized cells in SM, grey lines indicate the standard deviation. **B)-D)** pHluorin signal in the fluorescent population of control strain IME480 (CEN.PK113-5D background), IME481 (SwYG, *Sc*-Glycolysis background) and IME576 (CEN.PK113-5D *tps1Δ* background) as measured by flow cytometry after incubation without C-source addition (red) or with galactose (blue) or glucose (orange) in duplicate experiments. The ratio between the fluorescence excitation at 510 and 515 nm after excitation at 405 and 488 respectively is shown. Time of incubation (with or without C-source) is indicated. The control and SwYG (*Sc*-Glycolysis) strains behaved as expected, with an increase in the pH_i signal after sugar addition. The *tps1Δ* strain shows as strong decrease in the pHluorin signal upon glucose but not galactose addition, corresponding to previously published data. **E)** and **F)** Response of the pHluorin signal directly after addition of glucose or galactose in the IME481 (*Sc*-Glycolysis control) and IME576 (*tps1Δ*) strains. Again an immediate and sharp decrease in signal is seen for the *tps1Δ* strain upon glucose addition, while the addition of galactose leads to an increase in the signal.

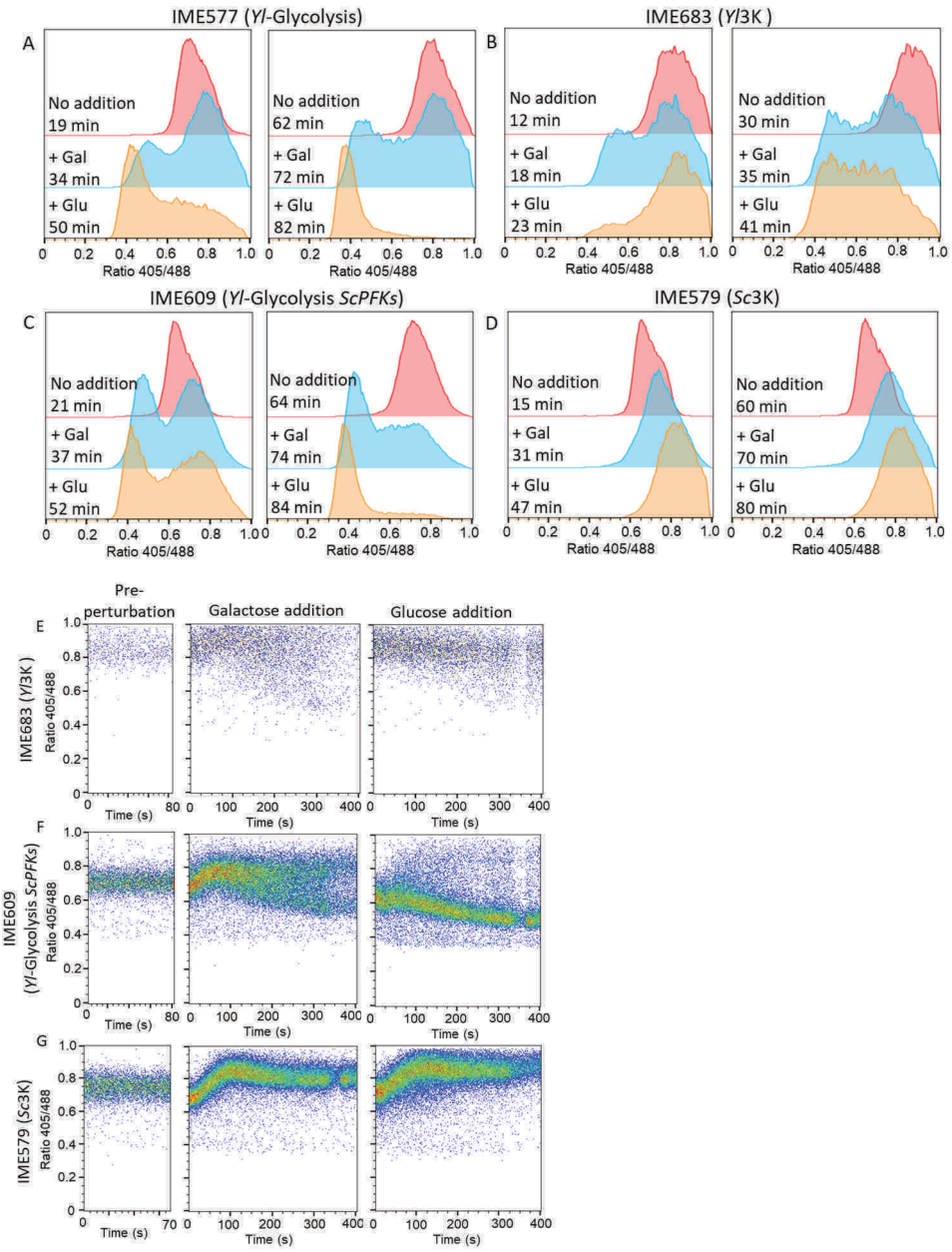
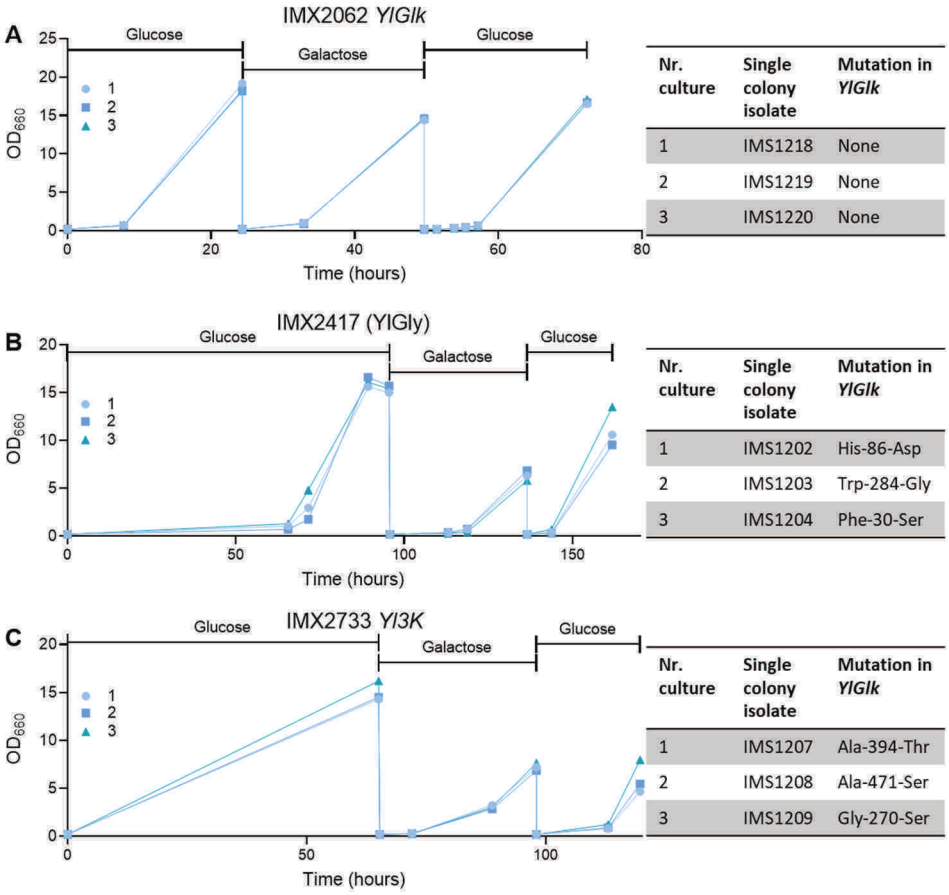


Figure S2 - pHluorin responses of the *Yl*-Glycolysis and mosaic glycolysis strains.
A)-D) pHluorin signal in the fluorescent population of strains IME577, IME683, IME609 and IME579 which express different combinations of *Yarrowia* glycolytic enzymes after incubation without C-source addition (red) or with galactose (blue) or glucose (orange) in duplicate experiments. Time of incubation (with or without C-source) is indicated. **E)-G)** Response of the pHluorin signal directly after addition of glucose or galactose in the IME683, IME609 and IME579 strains.



Nr. culture	Single colony isolate	Mutation in *YIGlk*
1	IMS1218	None
2	IMS1219	None
3	IMS1220	None

Nr. culture	Single colony isolate	Mutation in *YIGlk*
1	IMS1202	His-86-Asp
2	IMS1203	Trp-284-Gly
3	IMS1204	Phe-30-Ser

Nr. culture	Single colony isolate	Mutation in *YIGlk*
1	IMS1207	Ala-394-Thr
2	IMS1208	Ala-471-Ser
3	IMS1209	Gly-270-Ser

Figure S3 - Repeated transfer in glucose medium.
To determine whether adaptation to glucose medium was genetic or not strains were pre-grown on SM-Galactose liquid medium, then transferred to glucose in three independent cultures (indicated with 1,2,3,) at a starting OD_{660} of 0.2, after growth was observed these cultures were re-inoculated to non-selective SM-Galactose medium and after growth re-inoculated in glucose medium. After the final glucose cultures single colony isolates were checked for mutations. **A)** IMX2062 (*YI*Glk complementation strain) showed a similar short lag phase when inoculated to a glucose culture for the second time, isolates IMS1218, IMS1219 and IMS1220 did not show mutations in the *YIGLK* gene. **B)** IMX2417 (*YI*-Glycolysis strain) showed a lag phase on the first glucose culture of approximately 70 hours, consistent with that observed in Growth Profiler cultures. The second glucose culture appeared to start growth immediately and the *Y. lipolytica* glucokinase gene was mutated in all three resulting single colony isolates (see also Fig. 3C) **C)** IMX2733 (*YI*-3K strain) similarly showed immediate growth in the second glucose culture and mutations the *YIGLK* gene were observed in each resulting single colony isolate.

A

	Percentage identity/similarity	
Protein	*Sc*Glk	*Kl*Glk
*Yl*Glk	39.9%/57.4%	40.7%/57.7%
*Sc*Glk		60.8%/74.5%

B

Strain	Mutated residue	Corresponding residue	
	*Yl*Glk	*Kl*Glk	*Sc*Glk
IMS1204 (YlGly)	F30	F31	F39
IMS1202 (YlGly)	H86	H93	H102
IMS1209 (Yl3K)	G270	G275	G291
IMS1203 (YlGly)	W284	Y289	Y305
IMS1207 (Yl3K)	A394	A392	A413
IMS1208 (Yl3K)	A471	A473	A494

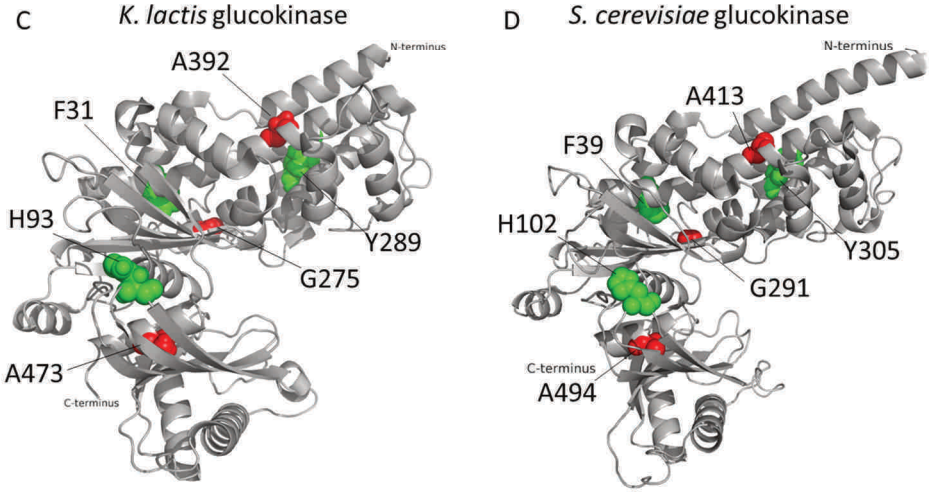


Figure S4 - Location of glucokinase mutations.
No crystal structure is currently available for the *Yarrowia lipolytica* glucokinase *(Yl*Glk*)*, this enzyme does however share similarity to the *Kluyveromyces lactis* and *S. cerevisiae* glucokinases (*Kl*Glk and *Sc*Glk) for which crystal structures are available. **A)** Table showing percentages identity and similarity as determined by global pairwise alignment of the protein sequences (EMBOSS Needle). B) Comparison of the protein sequences shows the mutations found in this study occurred mostly in conserved amino acid residues which have a corresponding residue in each of the glucokinases. **C)** and **D)** Residues corresponding to those mutated in the *Yl*Glk shown in the *Kl*Glk and *Sc*Glk crystal structures. In green those found in the mutants of the *Yl*-glycolysis strain and in red those found in the mutants of the *Yl*-3K strain. Mutations are spread over the different domains and are not directly in the active site. PDB identifiers and source of structure *Kl*Glk: 6R2N from [79], *Sc*Glk: 6p4x [80].

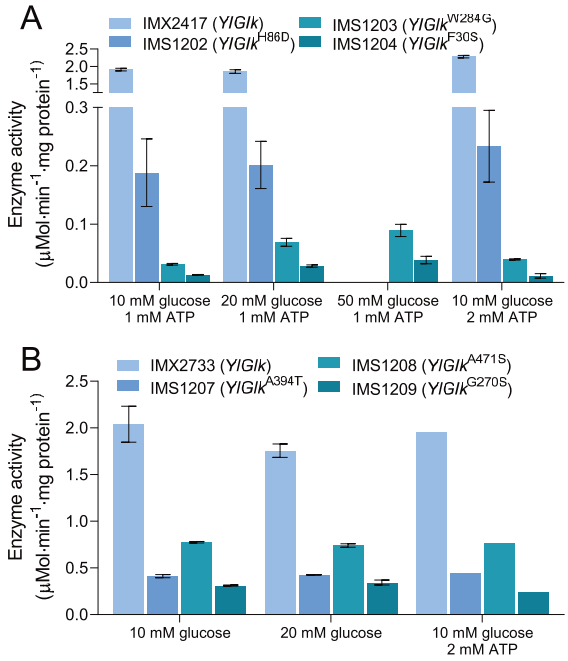


Figure S5 - Characterization of glucokinase mutants.
A) The glucokinase activity of the evolved isolates derived from *Yl*-Glycolysis strain IMX2417 (IMS1202-IMS1204) was determined at different glucose and ATP concentrations. The standard concentrations used were 10 mM glucose and 1 mM ATP. Increasing the glucose concentration to 20 and 50 mM increased activity of mutants *Yl*Glk^{W284G} and *Yl*Glk^{F30S}, but not the native enzyme or mutant *Yl*Glk^{H86D} suggesting an increased $K_{m,glucose}$ for those two mutants. **B)** Increasing the glucose or ATP concentration did not increase the activity of the *Yl*Glk mutants derived from strain IMX2733 (*Yl*-3K strain). Mean and SEM is shown for duplicate measurements except the increased ATP measurements of the IMX2733 strains which were measured in only one replicate.

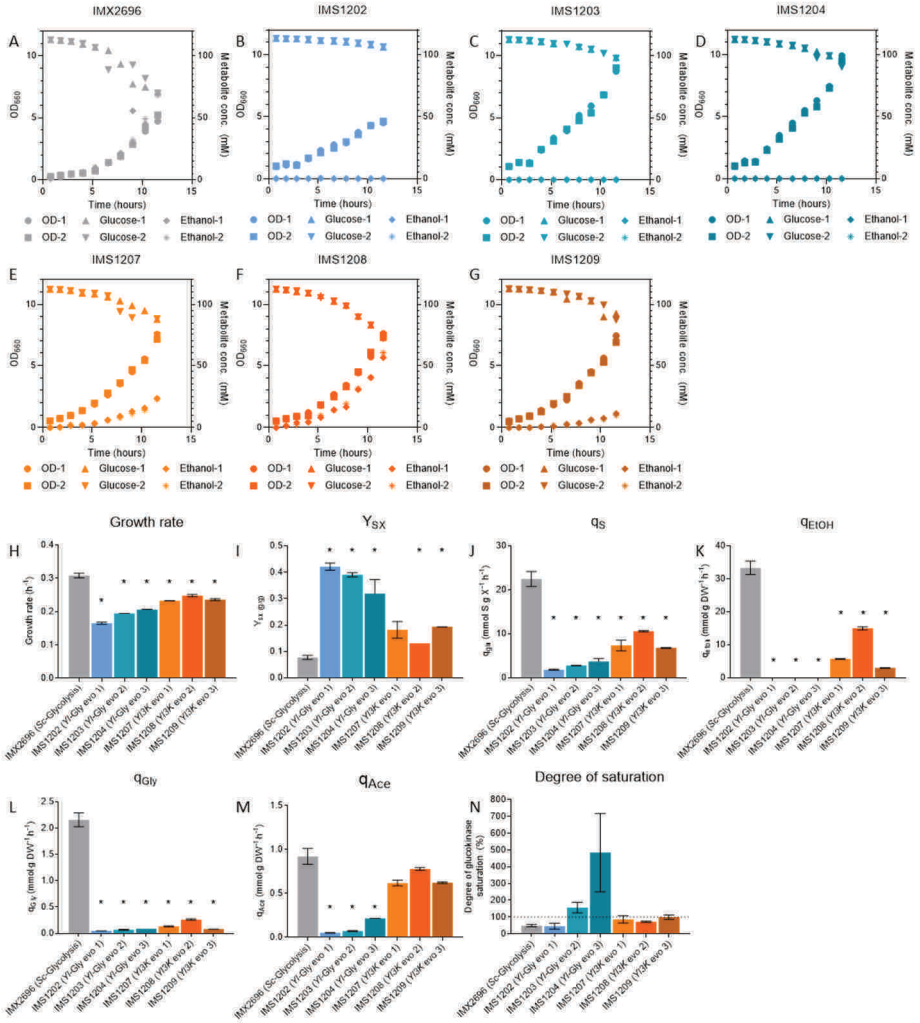


Figure S6 - Characterization of evolved isolates of the *Yl*-Glycolysis and *Yl*-3K strains.
The evolved isolates of the *Yl*-Glycolysis and *Yl*-3K strains and the control *Sc*-Glycolysis strain were grown on glucose minimal medium with urea to prevent acidification. **A)-G)** OD_{660} and metabolite profiles over time of duplicate cultures of each strain. **H)-M)** Estimations of growth rate, biomass yield, glucose uptake rate and ethanol, glycerol and acetate production rates based on the measured metabolite profiles, mean and SEM are shown significant differences to control strain IMX2696 indicated by * (T.Test, homoscedastic, unpaired P<0.05). **N)** Estimated degree of saturation of the hexokinase/glucokinase reaction based on the highest measured activities (Supplementary Fig. S5) and the glucose uptake rate (q_{glu}) estimated for each strain. Error bars indicate summed relative standard deviation for both measurements.

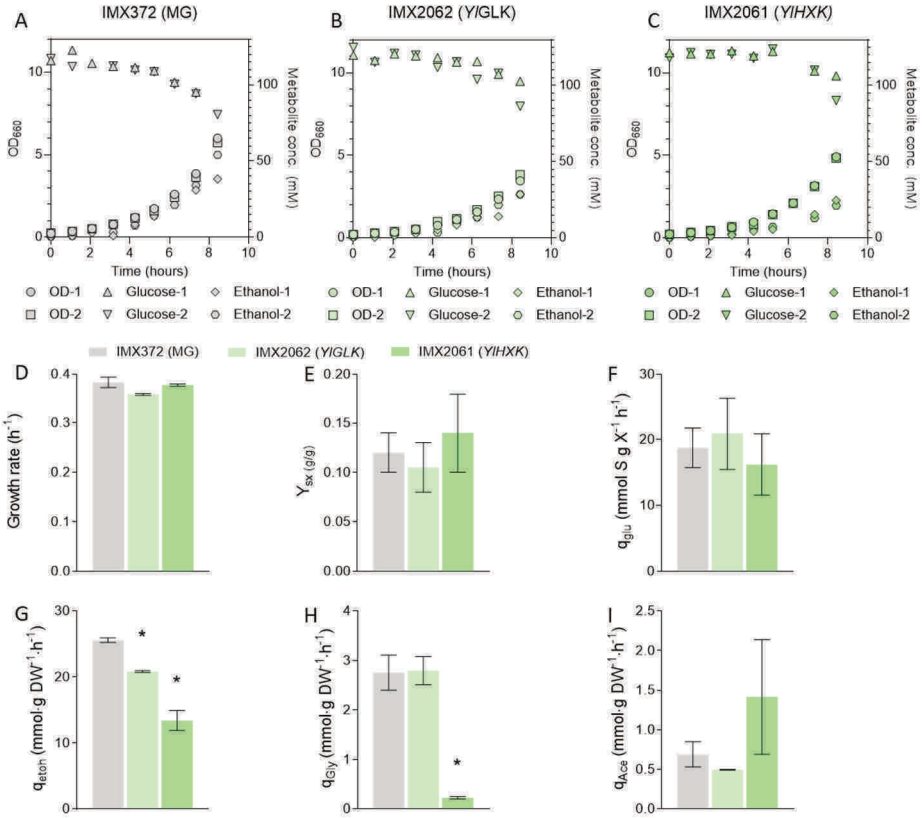


Figure S7 - Characterization of glucokinase and hexokinase complementation strains.
The complementation strains expressing the *YlGLK* and *YlGLK* genes and the control Minimal Glycolysis strain were grown on glucose minimal medium to measure growth rate, glucose uptake and ethanol production. **A)-C)** OD_{660} and metabolite profiles over time of duplicate cultures of each strain. **D)-I)** Estimations of growth rate, biomass yield, glucose uptake rate and ethanol, glycerol and acetate production rates based on the measured metabolite profiles, mean and SEM are shown, significant differences to control strain IMX372 indicated by * (T.Test, homoscedastic, unpaired $P<0.05$).

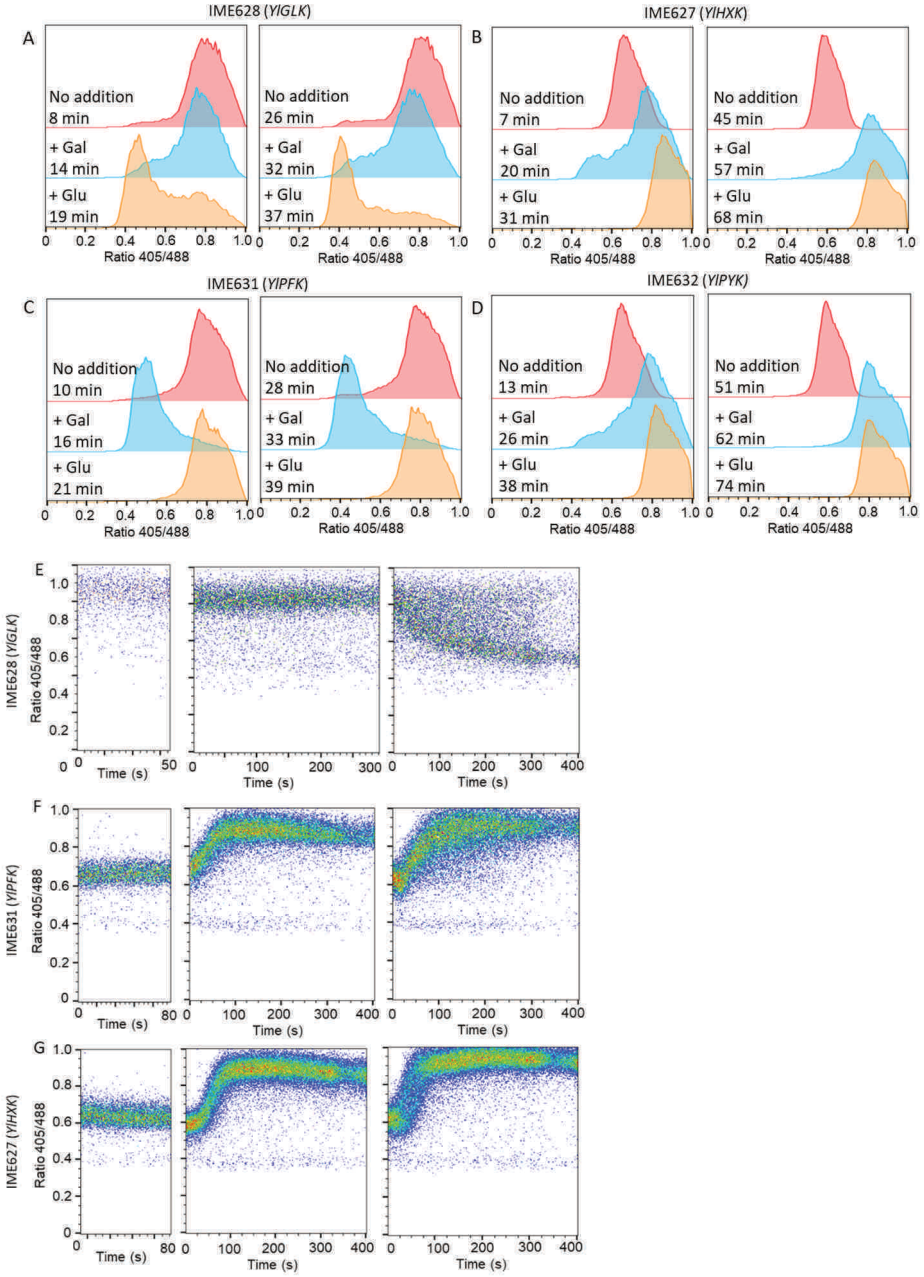


Figure S8 - pHluorin response of single complementation strains.
A)-D) pHluorin signal in the fluorescent population of strains IME628, IME627, IME631 and IME632 which express different single *Yarrowia* glycolytic enzymes after incubation without C-source addition (red) or with galactose (blue) or glucose (orange) in duplicate experiments. Time of incubation (with or without C-source) is indicated. **E)-G)** Response of the pHluorin signal directly after addition of glucose or galactose in the IME628, IME631 and IME627 strains.

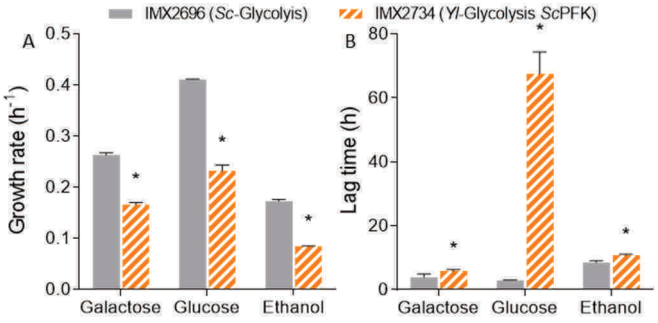


Figure S9 - Growth rate and lag phase of *Yl*-Glycolysis strain with *Sc*Pfk.
A) Growth rates measured in the growth profiler on galactose, glucose and ethanol for strain IMX2734, expressing the *Yarrowia lipolytica* glycolysis except phosphofructokinase, for which it has the *ScPFK* genes and control strain IMX2696 (*Sc*-Glycolysis, IMX2696). **B)** Growth on glucose was only observed after a lag phase of up to 75 hours similar to the *Yl*-Glycolysis strain. Mean and SEM of triplicates are shown, * indicates significant difference (T.Test, homoscedastic, unpaired, P<0.05).

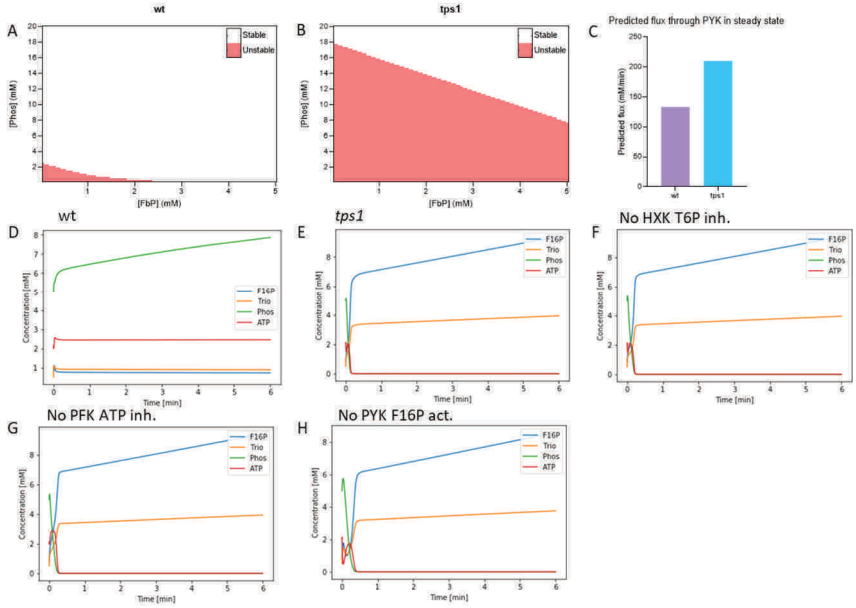


Figure S10 - Predicted metabolite time courses.
The model results were reproduced for a 'wt' and a *tps1* deletion mutant. **A)** and **B)** Division between balanced and imbalanced states in the wildtype and *tps1* models. With red showing the imbalanced state and white steady state. **C)** Predicted flux through the pyruvate kinase in the balanced state for both model types. **D)-H)** Predicted metabolite time courses for the first six minutes with various model configurations. Imbalanced starting concentrations of FBP and Phosphate were chosen (FBP_i: 0.836 mM, $Phos_i$: 5.0 mM). Time courses show similar behaviour for all imbalanced systems, with accumulation of FBP and depletion of phosphate. For the wildtype a steady state is reached with these initial conditions after ~50 minutes.

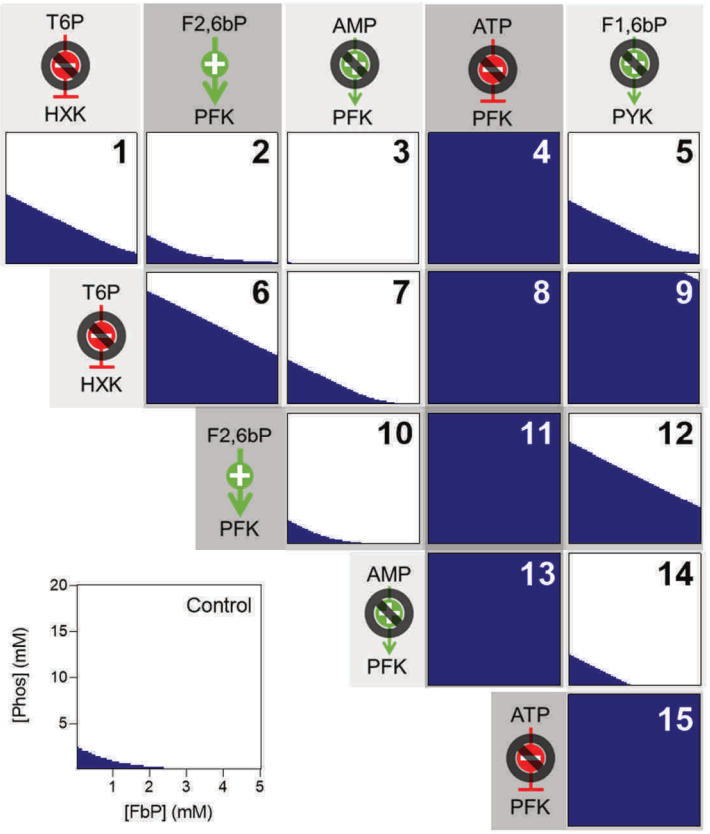


Figure S11 - Combinatorial effects of removing regulation in a mathematical model of glycolysis. The outcome of the glycolytic model is shown as a function of the initial concentrations of F1,6bP and phosphate, with dark blue indicating an imbalanced outcome and white a balanced steady state. In the bottom left the situation in the unmodified control model is shown, the same initial concentrations were tested for all model configurations. In plots 1-5 the effect on the model outcome of removal of single allosteric regulations is shown similar to Figure 6A. In the plots below combinatorial removal of two regulations is shown, with one removed in each row. Overall combinatorial removal was detrimental to stability, with the exception of removing AMP activation of PFK, which is also stabilizing on its own.

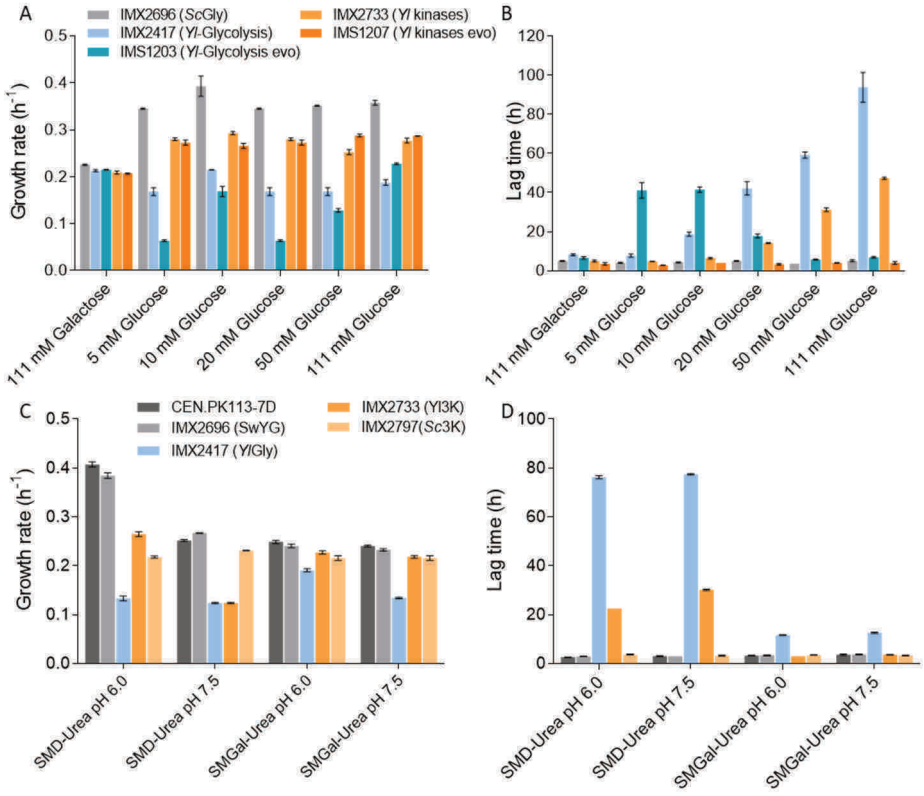


Figure S12 - Characterization of strains on different glucose concentrations and at high pH.
A) and **B)** Growth rates and lag-times determined by growth in the growth profiler of the *Yl*-glycolysis and *Yl*-3K strains and the control *Sc*-Glycolysis strain on galactose and at various glucose concentrations. **C)** and **D)** Growth rates of the *Yl*-Glycolysis and *Yl*-3K and *Sc*-3K strains in normal pH (6.0) and high pH (7.5) media to check the presence of a growth defect from dysfunction of the moonlighting function of yeast aldolase. Growth rates and lag-times were largely unaffected by the increased pH for the strains expressing *Y. lipolytica* glycolytic genes.

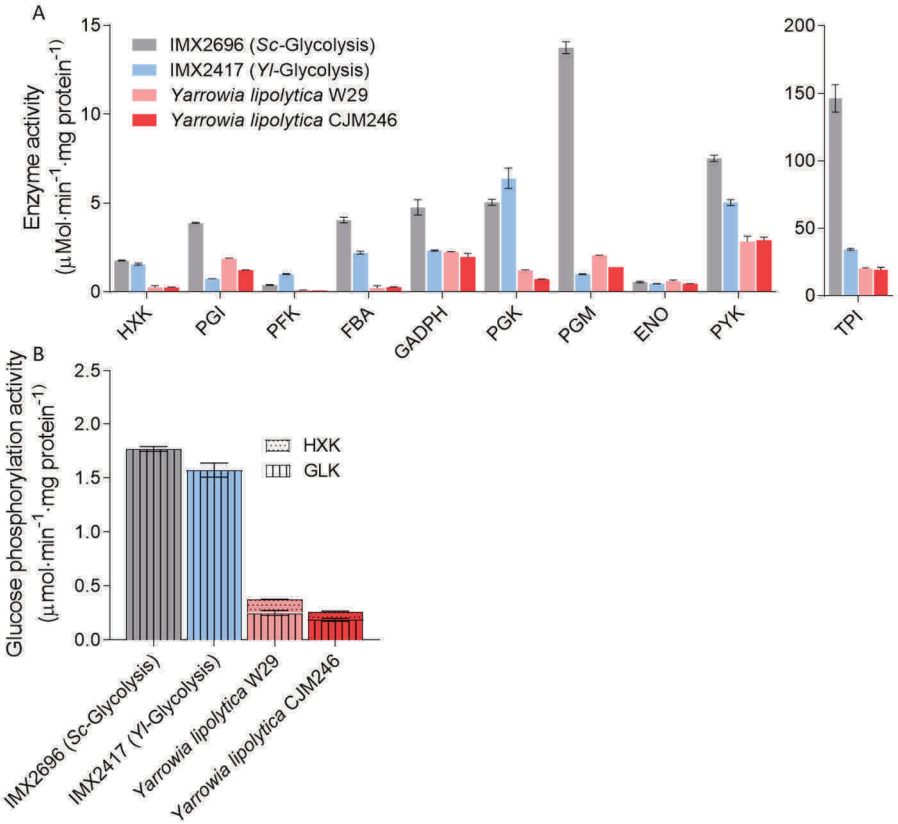


Figure S13 - Glycolytic activities in *Yarrowia lipolytica.*
Activity of glycolytic enzymes in *Yarrowia lipolytica* extracts. **A)** *In vitro* measured activities for two *Yarrowia lipolytica* strains, W29 a wildtype strain and laboratory strain CJM246 (PO1a). Activities of the *S. cerevisiae* strains IMX2696 and IMX2417 expressing the *Yarrowia* enzymes are shown for comparison. All glycolytic enzyme activities are significantly lower in *Y. lipolytica.* **B)** Separation of the glucokinase and hexokinase activities, based on measurements of the glucose and fructose phosphorylation activity, assuming a fructose/glucose phosphorylation ratio of 1.4 for hexokinase and an absence of glucokinase activity on fructose [18]. Glucokinase was the major isoenzyme in *Y. lipolytica* as expected, accounting for 66 and 71% of glucose phosphorylation activity in the W29 and CJM246 respectively.

Figure S14 - Overview of strain construction.
A) Overview of strain construction leading to the most important single locus glycolysis strains. Each step signifies a round of transformation and selection and a genetic modification. The construction of IMX589, IMX1821 and IMX1822 is described elsewhere [13, 15]. **B)** Overview of the main genetic loci in the key strains, *S. cerevisiae* genes are indicated in black and grey, *Yarrowia lipolytica* genes in blue, bacterial genes in green.

Table S1-S4
Supplementary tables S1-S4 can be found in the supplementary materials file at the data.4TU repository via: https://figshare.com/s/29c7f3dac3e16c4f4d97.

References

1. Folch, P.L., M.M.M. Bisschops, and R.A. Weusthuis, Metabolic energy conservation for fermentative product formation. Microb Biotechnol, 2021. **14**(3): p. 829-858.DOI: 10.1111/1751-7915.13746.
2. Oliveira, A.P. and U. Sauer, The importance of post-translational modifications in regulating *Saccharomyces cerevisiae* metabolism. FEMS yeast research, 2012. **12**(2): p. 104-117.
3. Daran-Lapujade, P., et al., The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. Proceedings of the National Academy of Sciences, 2007. **104**(40): p. 15753-15758.
4. van Heerden, J.H., F.J. Bruggeman, and B. Teusink, Multi-tasking of biosynthetic and energetic functions of glycolysis explained by supply and demand logic. BioEssays, 2015. **37**(1): p. 34-45.
5. van Eunen, K., et al., Testing biochemistry revisited: how *in vivo* metabolism can be understood from *in vitro* enzyme kinetics. PLoS Comput Biol, 2012. **8**(4): p. e1002483.
6. Teusink, B., et al., The danger of metabolic pathways with turbo design. Trends in Biochemical Sciences, 1998. **23**(5): p. 162-169.
7. Teusink, B., et al., Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem, 2000. **267**(17): p. 5313-29.DOI: 10.1046/j.1432-1327.2000.01527.x.
8. Mayordomo, I. and P. Sanz, Human pancreatic glucokinase (GlkB) complements the glucose signalling defect of *Saccharomyces cerevisiae hxk2* mutants. Yeast, 2001. **18**(14): p. 1309-16.DOI: 10.1002/yea.780.
9. Bonini, B.M., P. Van Dijck, and J.M. Thevelein, Uncoupling of the glucose growth defect and the deregulation of glycolysis in *Saccharomyces cerevisiae* Tps1 mutants expressing trehalose-6-phosphate-insensitive hexokinase from *Schizosaccharomyces pombe.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2003. **1606**(1-3): p. 83-93.
10. Boles, E., et al., Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in *Saccharomyces cerevisiae* that is catalytically insensitive to fructose-1,6-bisphosphate. J. Bacteriol, 1997. **179**(9): p. 2987-2993.
11. Flores, C.-L., et al., The dimorphic yeast *Yarrowia lipolytica* possesses an atypical phosphofructokinase: characterization of the enzyme and its encoding gene. Microbiology, 2005. **151**(5): p. 1465-1474.
12. Estévez, A.M., J.J. Heinisch, and J.J. Aragón, Functional complementation of yeast phosphofructokinase mutants by the non-allosteric enzyme from *Dictyostelium discoideum.* FEBS letters, 1995. **374**(1): p. 100-104.
13. Kuijpers, N.G., et al., Pathway swapping: Toward modular engineering of essential cellular processes. Proceedings of the National Academy of Sciences, USA, 2016. **113**(52): p. 15060-15065.DOI: 10.1073/pnas.1606701113.
14. Solis-Escalante, D., et al., A minimal set of glycolytic genes reveals strong redundancies in *Saccharomyces cerevisiae* central metabolism. Eukaryotic Cell, 2015. **14**(8): p. 804-816.DOI: 10.1128/EC.00064-15.
15. Boonekamp, F.J., et al., A yeast with muscle does not run faster: full humanization of the glycolytic pathway in *Saccharomyces cerevisiae.* bioRxiv, 2021.DOI: 10.1101/2021.09.28.462164.
16. Nicaud, J.M., *Yarrowia lipolytica.* Yeast, 2012. **29**(10): p. 409-418.DOI: 10.1002/yea.2921.
17. Petit, T. and C. Gancedo, Molecular cloning and characterization of the gene *HXK1* encoding the hexokinase from *Yarrowia lipolytica.* Yeast, 1999. **15**(15): p. 1573-84.DOI: 10.1002/(SICI)1097-0061(199911)15:15<1573::AID-YEA478>3.0.CO;2-3.

18. Flores, C.-L., C. Gancedo, and T. Petit, Disruption of *Yarrowia lipolytica TPS1* gene encoding trehalose-6-P synthase does not affect growth in glucose but impairs growth at high temperature. PloS one, 2011. **6**(9): p. e23695.
19. Hirai, M., A. Tanaka, and S. Fukui, Difference in pyruvate kinase regulation among three groups of yeasts. Biochimica et Biophysica Acta (BBA)-Enzymology, 1975. **391**(2): p. 282-291.
20. Guadalupe-Medina, V., et al., Carbon dioxide fixation by Calvin-Cycle enzymes improves ethanol yield in yeast. Biotechnology for biofuels, 2013. **6**(1): p. 125.
21. van Heerden, J.H., et al., Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. Science, 2014. **343**(6174): p. 1245114.DOI: 10.1126/science.1245114.
22. Neves, M.J., et al., Control of Glucose Influx into Glycolysis and Pleiotropic Effects Studied in Different Isogenic Sets of *Saccharomyces cerevisiae* Mutants in Trehalose Biosynthesis. Current Genetics, 1995. **27**(2): p. 110-122.DOI: 10.1007/Bf00313424.
23. Orij, R., et al., *In vivo* measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in *Saccharomyces cerevisiae* reveals a relation between intracellular pH and growth. Microbiology, 2009. **155**(1): p. 268-278.
24. Miesenböck, G., D.A. De Angelis, and J.E. Rothman, Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. Nature, 1998. **394**(6689): p. 192-195.
25. Maitra, P.K. and Z. Lobo, Yeast Pyruvate Kinase: A Mutant Form Catalytically Insensitive to Fructose 1, 6-Bisphosphate. European journal of biochemistry, 1977. **78**(2): p. 353-360.
26. Van den Brink, J., et al., Dynamics of glycolytic regulation during adaptation of *Saccharomyces cerevisiae* to fermentative metabolism. Applied and environmental microbiology, 2008. **74**(18): p. 5710-5723.
27. Fell, D.A. and S. Thomas, Physiological control of metabolic flux: the requirement for multisite modulation. Biochem J, 1995. **311 (Pt 1)**(Pt 1): p. 35-9.DOI: 10.1042/bj3110035.
28. Veech, R.L. and D.A. Fell, Distribution control of metabolic flux. Cell Biochem Funct, 1996. **14**(4): p. 229-36.DOI: 10.1002/cbf.697.
29. Opperdoes, F.R. and P. Borst, Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome. FEBS Lett, 1977. **80**(2): p. 360-4.DOI: 10.1016/0014-5793(77)80476-6.
30. Haanstra, J.R., B.M. Bakker, and P.A. Michels, In or out? On the tightness of glycosomal compartmentalization of metabolites and enzymes in *Trypanosoma brucei.* Mol Biochem Parasitol, 2014. **198**(1): p. 18-28.DOI: 10.1016/j.molbiopara.2014.11.004.
31. Romeo, T. and J.L. Snoep, Glycolysis and flux control. EcoSal Plus, 2005. **1**(2).
32. Dolatshahi, S., L.L. Fonseca, and E.O. Voit, New insights into the complex regulation of the glycolytic pathway in *Lactococcus lactis*. II. Inference of the precisely timed control system regulating glycolysis. Molecular Biosystems, 2016. **12**(1): p. 37-47.DOI: 10.1039/c5mb00726g.
33. Wilson, J.E., Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. J Exp Biol, 2003. **206**(Pt 12): p. 2049-57.DOI: 10.1242/jeb.00241.
34. Blazquez, M.A., et al., Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. FEBS Letters, 1993. **329**(1-2): p. 51-54.
35. Ernandes, J.R., et al., During the initiation of fermentation overexpression of hexokinase PII in yeast transiently causes a similar deregulation of glycolysis as deletion of Tps1. Yeast, 1998. **14**(3): p. 255-269.

36. Boles, E., H.W. Gohlmann, and F.K. Zimmermann, Cloning of a second gene encoding 6-phosphofructo-2-kinase in yeast, and characterization of mutant strains without fructose-2,6-bisphosphate. Mol Microbiol, 1996. **20**(1): p. 65-76.DOI: 10.1111/j.1365-2958.1996.tb02489.x.
37. Fraenkel, D.G., Mutants in glucose metabolism. Annu Rev Biochem, 1986. **55**(1): p. 317-37.DOI: 10.1146/annurev.bi.55.070186.001533.
38. Manjrekar, J., Allosteric Regulation in Search of a Role. Current Science, 1993. **65**(6): p. 443-447.
39. Thevelein, J.M. and S. Hohmann, Trehalose synthase: guard to the gate of glycolysis in yeast? Trends Biochem Sci, 1995. **20**(1): p. 3-10.DOI: 10.1016/s0968-0004(00)88938-0.
40. Hauf, J., F.K. Zimmermann, and S. Muller, Simultaneous genomic overexpression of seven glycolytic enzymes in the yeast *Saccharomyces cerevisiae.* Enzyme Microb Technol, 2000. **26**(9-10): p. 688-698.DOI: 10.1016/s0141-0229(00)00160-5.
41. Emmerling, M., J.E. Bailey, and U. Sauer, Altered regulation of pyruvate kinase or co-overexpression of phosphofructokinase increases glycolytic fluxes in resting *Escherichia coli.* Biotechnol Bioeng, 2000. **67**(5): p. 623-7.DOI: 10.1002/(sici)1097-0290(20000305)67:5<623::aid-bit13>3.0.co;2-w.
42. Emmerling, M., J.E. Bailey, and U. Sauer, Glucose catabolism of *Escherichia coli* strains with increased activity and altered regulation of key glycolytic enzymes. Metab Eng, 1999. **1**(2): p. 117-27.DOI: 10.1006/mben.1998.0109.
43. Ramos, A., et al., Effect of pyruvate kinase overproduction on glucose metabolism of *Lactococcus lactis.* Microbiology (Reading), 2004. **150**(Pt 4): p. 1103-1111.DOI: 10.1099/mic.0.26695-0.
44. Kotte, O., et al., Phenotypic bistability in *Escherichia coli*'s central carbon metabolism. Mol Syst Biol, 2014. **10**(7): p. 736.DOI: 10.15252/msb.20135022.
45. Botman, D., J.H. van Heerden, and B. Teusink, An Improved ATP FRET Sensor For Yeast Shows Heterogeneity During Nutrient Transitions. ACS Sensors, 2020. **5**(3): p. 814-822.DOI: 10.1021/acssensors.9b02475.
46. Lazar, Z., et al., Hexokinase - a limiting factor in lipid production from fructose in *Yarrowia lipolytica.* Metabolic engineering, 2014. **26**: p. 89-99.
47. Workman, M., P. Holt, and J. Thykaer, Comparing cellular performance of *Yarrowia lipolytica* during growth on glucose and glycerol in submerged cultivations. AMB Express, 2013. **3**(1): p. 1-9.
48. Ledesma-Amaro, R. and J.-M. Nicaud, Metabolic engineering for expanding the substrate range of *Yarrowia lipolytica.* Trends in Biotechnology, 2016. **34**(10): p. 798-809.
49. Lubuta, P., et al., Investigating the Influence of Glycerol on the Utilization of Glucose in *Yarrowia lipolytica* Using RNA-Seq-Based Transcriptomics. G3-Genes Genomes Genetics, 2019. **9**(12): p. 4059-4071.DOI: 10.1534/g3.119.400469.
50. Elbing, K., et al., Role of hexose transport in control of glycolytic flux in *Saccharomyces cerevisiae.* Applied and Environmental Microbiology, 2004. **70**(9): p. 5323-5330.
51. Otterstedt, K., et al., Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae.* EMBO Rep, 2004. **5**(5): p. 532-7.DOI: 10.1038/sj.embor.7400132.
52. Wieczorke, R., et al., Characterisation of mammalian GLUT glucose transporters in a heterologous yeast expression system. Cell Physiol Biochem, 2003. **13**(3): p. 123-34.DOI: 10.1159/000071863.
53. Entian, K.-D. and P. Kötter, 25 Yeast Genetic Strain and Plasmid Collections, in *Yeast Gene Analysis - Second Edition*, I. Stansfield and M.J.R. Stark, Editors. 2007, Academic Press. p. 629-666.

54. Verduyn, C., et al., Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast, 1992. **8**(7): p. 501-17.DOI: 10.1002/yea.320080703.
55. Solis-Escalante, D., et al., *amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae.* FEMS Yeast Research, 2013. **13**(1): p. 126-39.DOI: 10.1111/1567-1364.12024.
56. Lee, M.E., et al., A highly characterized yeast toolkit for modular, multipart assembly. ACS Synthetic Biology, 2015. **4**(9): p. 975-986.DOI: 10.1021/sb500366v.
57. Kuijpers, N.G., et al., A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. Microbial cell factories, 2013. **12**(1): p. 47.
58. Inoue, H., H. Nojima, and H. Okayama, High efficiency transformation of *Escherichia coli* with plasmids. Gene, 1990. **96**(1): p. 23-8.DOI: 10.1016/0378-1119(90)90336-p.
59. Gietz, R.D. and R.A. Woods, Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods in enzymology, 2002. **350**: p. 87-96.
60. Looke, M., K. Kristjuhan, and A. Kristjuhan, Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques, 2011. **50**(5): p. 325-8.DOI: 10.2144/000113672.
61. Lopez, M.C., et al., A Phosphatidylinositol Phosphatidylcholine Transfer Protein Is Required for Differentiation of the Dimorphic Yeast *Yarrowia Lipolytica* from the Yeast to the Mycelial Form. Journal of Cell Biology, 1994. **125**(1): p. 113-127.DOI: 10.1083/jcb.125.1.113.
62. Wronska, A.K., et al., Exploiting the diversity of *Saccharomycotina* yeasts to engineer biotin-independent growth of *Saccharomyces cerevisiae.* Applied & environmental microbiology, 2020. **86**(12): p. e00270-20.
63. Mans, R., et al., CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae.* FEMS Yeast Research, 2015. **15**(2).DOI: 10.1093/femsyr/fov004.
64. Shi, S., et al., Improved production of fatty acids by *Saccharomyces cerevisiae* through screening a cDNA library from the oleaginous yeast *Yarrowia lipolytica.* FEMS yeast research, 2016. **16**(1): p. fov108.
65. Le Dall, M.-T., et al., The 3-phosphoglycerate kinase gene of the yeast *Yarrowia lipolytica* de-represses on gluconeogenic substrates. Current genetics, 1996. **29**(5): p. 446-456.
66. Strick, C., et al., The isolation and characterization of the pyruvate kinase-encoding gene from the yeast *Yarrowia lipolytica*. Gene, 1992. **118**(1): p. 65-72.
67. Boonekamp, F.J., et al., Design and Experimental Evaluation of a Minimal, Innocuous Watermarking Strategy to Distinguish Near-Identical DNA and RNA Sequences. ACS Synth Biol, 2020. **9**(6): p. 1361-1375.DOI: 10.1021/acssynbio.0c00045.
68. Postma, E.D., et al., A supernumerary designer chromosome for modular *in vivo* pathway assembly in *Saccharomyces cerevisiae.* Nucleic Acids Res, 2021. **49**(3): p. 1769-1783.DOI: 10.1093/nar/gkaa1167.
69. Li, H. and R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 2009. **25**(14): p. 1754-60.DOI: 10.1093/bioinformatics/btp324.
70. Salazar, A.N., et al., Nanopore sequencing enables near-complete de novo assembly of *Saccharomyces cerevisiae* reference strain CEN.PK113-7D. FEMS Yeast Res, 2017. **17**(7).DOI: 10.1093/femsyr/fox074.
71. Walker, B.J., et al., Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One, 2014. **9**(11): p. e112963.DOI: 10.1371/journal.pone.0112963.

72. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform, 2013. **14**(2): p. 178-92.DOI: 10.1093/bib/bbs017.
73. Nijkamp, J.F., et al., *De novo* detection of copy number variation by co-assembly. Bioinformatics, 2012. **28**(24): p. 3195-202.DOI: 10.1093/bioinformatics/bts601.
74. Postma, E., et al., Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae.* Appl Environ Microbiol, 1989. **55**(2): p. 468-77.DOI: 10.1128/aem.55.2.468-477.1989.
75. Jansen, M.L.A., et al., Prolonged selection in aerobic, glucose-limited chemostat cultures of *Saccharomyces cerevisiae* causes a partial loss of glycolytic capacity. Microbiology (Reading), 2005. **151**(Pt 5): p. 1657-1669.DOI: 10.1099/mic.0.27577-0.
76. Cruz, L.A., et al., Similar temperature dependencies of glycolytic enzymes: an evolutionary adaptation to temperature dynamics? BMC Syst Biol, 2012. **6**: p. 151.DOI: 10.1186/1752-0509-6-151.
77. Lowry, O.H., et al., Protein measurement with the Folin phenol reagent. J Biol Chem, 1951. **193**(1): p. 265-75.
78. Olivier, B.G., J.M. Rohwer, and J.-H.S. Hofmeyr, Modelling cellular systems with PySCeS. Bioinformatics, 2005. **21**(4): p. 560-561.
79. Zak, K.M., et al., Crystal Structure of *Kluyveromyces lactis* Glucokinase (*Kl*Glk1). Int J Mol Sci, 2019. **20**(19): p. 4821.DOI: 10.3390/ijms20194821.
80. Stoddard, P.R., et al., Polymerization in the actin ATPase clan regulates hexokinase activity in yeast. Science, 2020. **367**(6481): p. 1039-1042.DOI: 10.1126/science.aay5359.

Chapter 5

Synthetic Genomics from a yeast perspective

Charlotte C. Koster[#], Eline D. Postma[#], Ewout Knibbe[#], Céline Cleij[#], Pascale Daran-Lapujade

[#] These authors contributed equally to this work and should be considered co-first authors.

Essentially as published in Frontiers in Bioengineering and Biotechnology 2022, (https://doi.org/10.3389/fbioe.2022.869486)

Abstract

Synthetic Genomics focuses on the construction of rationally designed chromosomes and genomes and offers novel approaches to study biology and to construct synthetic cell factories. Currently, progress in Synthetic Genomics is hindered by the inability to synthesize DNA molecules longer than a few hundred base pairs, while the size of the smallest genome of a self-replicating cell is several hundred thousand base pairs. Methods to assemble small fragments of DNA into large molecules are therefore required. Remarkably powerful at assembling DNA molecules, the unicellular eukaryote *Saccharomyces cerevisiae* has been pivotal in the establishment of Synthetic Genomics. Instrumental in the assembly of entire genomes of various organisms in the past decade, the *S. cerevisiae* genome foundry has a key role to play in future Synthetic Genomics developments.

Introduction

Synthetic Genomics (SG) is a recent Synthetic Biology discipline that focuses on the construction of rationally designed chromosomes and genomes. SG offers a novel approach to address fundamental biological questions by restructuring, recoding, and minimizing (parts of) genomes (as recently reviewed by [1]). SG is now spurring technological developments in academia and has a strong future potential in industry [2, 3]). Humankind's best microbial friend, the baker's yeast *Saccharomyces cerevisiae*, has played, and continues to play a key role in SG advances, both by enabling the construction of chromosomes for other hosts, and in the refactoring of its own genome. This mini review explores the reasons for this strategic positioning of *S. cerevisiae* in SG, surveys the main achievements enabled by this yeast and reflects on future developments.

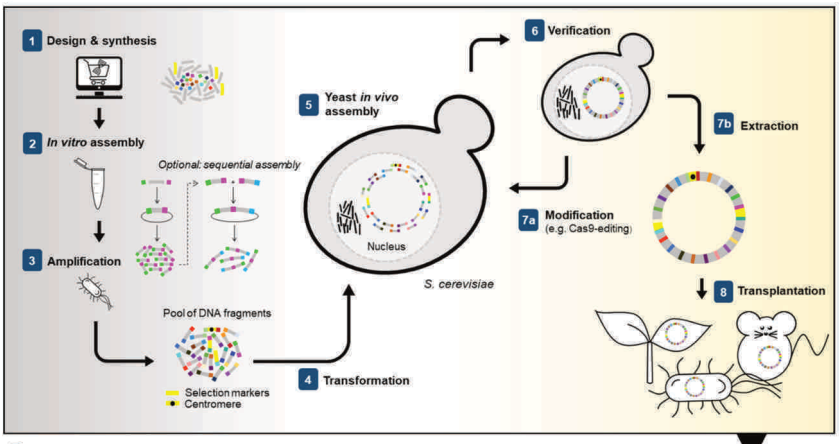
Current limitations of genome assembly

While small-sized viral chromosomes were the first to be chemically synthetized, the breakthrough in the field of SG came with the synthesis and assembly of the 592 kilobase (kb) chromosome of *Mycoplasma genitalium* [4, 5]. The unicellular eukaryote *Saccharomyces cerevisiae* has made a key contribution to this famous milestone. To understand how this microbe, commonly used in food and beverages, contributes to the assembly of synthetic genomes, let us recapitulate how synthetic chromosomes can be constructed (Fig. 1).

It starts with the customized synthesis of short DNA molecules called oligonucleotides. Oligonucleotides are mostly synthetized via phosphoramidite chemistry, a 40 year-old method [6] that, despite decades of technological developments, struggles to deliver error-free oligonucleotides longer than 200 base pairs (bp). While the implementation of microarrays has substantially decreased the synthesis cost, it has not increased oligo length, an achievement that requires new synthesis methods [7]. Enzymatic alternatives for DNA synthesis are under development [8, 9], but still have considerable shortcomings regarding automation and scalability that must be overcome before commercial scale can be considered (reviewed in [10-13]). Considering that a theoretical minimal genome would be around 113 kb long [14] and that the first fully synthesized genome of *M. genitalium* contains 583 kb [4], thousands of oligos must be stitched together to construct a complete synthetic genome. These DNA oligos can be assembled into longer DNA fragments owing to a plethora of *in vitro* methods (reviewed in [11, 15, 16]). A method that has gained tremendous popularity since its development is the homology-based Gibson isothermal assembly [17], devised to assemble the *M. genitalium* genome. As all *in vitro* methods, Gibson assembly is limited by the number of fragments that can reliably be stitched together in one reaction, usually around a dozen, requiring a stepwise assembly procedure of increasingly large genomic DNA constructs [18]. DNA must be recovered from the reaction, amplified and verified in each round, to

allow further processing. Selection and amplification of correctly cloned DNA is routinely performed in *Escherichia coli,* however, maintenance of large constructs of exogenous DNA, especially from prokaryotic origins, in this bacterium is often limited by expression and toxicity of gene products [19]. *In vitro* alternatives for efficient and faithful selection and amplification of correctly assembled DNA are under development, but these are currently limited in length of amplified DNA and scalability [20-23]. While in principle stepwise *in vitro* assembly can lead to a DNA molecule of any size, and selection and amplification in *E. coli* worked well for DNA constructs up to 72 kb, *E. coli* had great difficulties maintaining quarter *M. genitalium* genomes, causing Gibson and colleagues to turn to baker's yeast [4, 5].

A



B

	In vitro	*E. coli*	*B. subtilis*[1]	*S. cerevisiae*
One-pot assembly				
Number of fragments	Up to 20 (Chandran, S., 2017)	Up to 6 (Kostylev, M. *et al.*, 2015)	2 (Itaya, M. *et al.*, 2008)	Up to 44 (Postma, E.D. *et al.*, 2021a)
Construct size	Up to 144 kb (Gibson, D.G. *et al.*, 2008a)	Up to 180 kb (Kostylev, M. *et al.*, 2015)	Up to 200 kb (Itaya, M. *et al.*, 2018)	Up to 1.66 Mb (Tagwerker, C. *et al.*, 2012)
Sequential assembly				
Number of assembled fragments	600 (Gibson, D.G. *et al.*, 2010b)[2]	-	31 (Itaya, M. *et al.*, 2008)	1078 (Gibson, D.G. *et al.*, 2010a)
Number of sequential HR events	4	-	18	3
Final construct size	16.3 kb	-	134.5 kb	1.08 Mb
Other factors				
Compatibility with heterologous genes	***	*	***	***
DNA yield	*	***	Unknown	**
Overall proficiency to assemble synthetic genomes	-[3]	*	**	***

Figure 1- *In vivo* and *in vitro* approaches for DNA assembly in synthetic genomics.
A) Simplified overview of chromosome construction using *Saccharomyces cerevisiae* for genome assembly and production. **B)** Strengths and weaknesses of *in vitro* and *in vivo* assembly methods. (1) Assembly of fragments in *B. subtilis* is performed by integration into the host genome, (2) Between rounds of sequential assembly, transformation into *E. coli* is conventional for selection and amplification of constructs. (3) Requires *in vivo* amplification and selection in a microbial host.

Saccharomyces cerevisiae as a genome foundry

S. cerevisiae seems a logical host for SG as it naturally maintains a 12 Mb genome consisting of 16 chromosomes ranging from 230 to 1500 kb in its haploid version, lives as polyploid in natural environments, and is extremely robust to changes in genome content and architecture [24]. The extreme robustness of *S. cerevisiae* to supernumerary, chimeric chromosomes, a key feature for SG, was already demonstrated in the late '80s [25, 26]. A second key feature of *S. cerevisiae* is its preference for homologous recombination (HR) to repair double-strand DNA breaks[27], a rare trait among eukaryotes. *S. cerevisiae* ability to efficiently and with high fidelity stitch together linear DNA molecules that present homologous regions as short as 40 bp [28] at their ends, was rapidly valorized for genetic manipulations and assembly of heterologous DNA. Recently renamed *in vivo* (or *in yeasto*) assembly, this cloning technique (Fig. 1) contributes to the remarkable genetic tractability and popularity of *S. cerevisiae* as model and industrial microbe [17, 29]. The combination of *S. cerevisiae*'s HR efficiency and fidelity, chromosome maintenance and propagation enabled the construction of the full *Mycoplasma* genome. Reflecting that "*in the future, it may be advantageous to make greater use of yeast recombination to assemble chromosomes*", this study propelled *S. cerevisiae* as powerful 'genome foundry' [4]. In the challenge to synthesize genomes, Ostrov and colleagues rightfully identified assembly of these long DNA constructs as '*the most critical hurdle*' [10]. To date, *S. cerevisiae* has been key to assembling entire or partial genomes in most synthetic genome projects (Table 1). For instance, the entire 785 kb refactored *Caulobacter crescentus* (renamed *C. ethensis*) genome was assembled *in vivo* from 16 fragments [30], while the recoded *E. coli* genome was split over 10 fragments of 91 to 136 kb individually assembled in yeast, and then sequentially integrated in the *E. coli* chromosome to replace native segments [31] (Table 1). *In vivo* assembly also proved to be powerful in assembling and modifying genomes of organisms that are poorly amenable to genome editing; the rapid and faithful HR-based assembly of *S. cerevisiae* recently enabled the reconstruction of a synthetic SARS-CoV-2 genome in a single week [32], and has been shown to be a promising host for *in vivo* assembly and modification of other viral genomes [33] as well as the genomes of various pathogens [34] and even a 101 kb human gene, which was transplanted into mouse embryonic cells [35] (Table 1). Moreover, *S. cerevisiae* was selected for the construction of the first synthetic eukaryotic genome. The international Sc2.0 consortium, spearheaded by Jef Boeke, undertook less than ten years ago the daunting task of synthesizing recoded versions of the 16 yeast chromosomes. Via stepwise, systematic replacement of 30 to 40 kb (using ca. 12 DNA fragments of 2 to 4 kb) of the native yeast sequence, the consortium is close to the completion of the largest synthetic genome to date [36, 37], with the ambition to reshape and minimize the *S. cerevisiae* genome [38].

While *S. cerevisiae* is not the only microbial host available for the construction of (neo)chromosomes (Fig. 1), several key features make it superior to its bacterial alternatives *Bacillus subtilis* and *E. coli* as genome foundry: i) *S. cerevisiae* has the natural ability to carry large amounts of DNA and therefore to host multiple exogenous bacterial genomes [34]; ii) *E. coli* frequently struggles with toxicity caused by the expression of exogenous bacterial sequences [5, 19, 39], while S. *cerevisiae* is very robust to the presence of heterologous DNA from prokaryotic or eukaryotic origin [40]; iii) *S. cerevisiae* can, in a single transformation, assemble many DNA oligonucleotides into (partial)genomes. *B. subtilis* can also maintain large exogenous DNA constructs, but requires a stepwise method for DNA assembly, in which each DNA part is integrated sequentially into *B. subtilis* genome [41]. This approach is intrinsically more labor-intensive and time-consuming than *S. cerevisiae* single transformation assembly.

Surprised by *S. cerevisiae* genetic tractability, Gibson and colleagues wondered *"how many pieces can be assembled in yeast in a single step?"* [4]. Pioneering a SG approach for metabolic engineering based on modular, specialized synthetic chromosomes, Postma *et al.* probed this limit recently in our lab by constructing 100 kb artificial linear and circular neochromosomes from 44 DNA parts in a single transformation[42, 43]. The remarkable efficiency of *in vivo* assembly (36% of assemblies faithful to design) revealed that its limit has clearly not been reached yet, and that future systematic studies are required to evaluate the true potential of *S. cerevisiae* as a genome foundry. The supernumerary chromosomes were shown to stably maintain complete heterologous pathways as well as the yeast's central carbon metabolism, underlining the potential of yeast synthetic genomics in the development of optimized cell-factories. Once assembled, synthetic chromosomes could be easily edited in *S. cerevisiae* thanks to its efficient HR and rich molecular toolbox.

Challenges in genome assembly using yeast

While *S. cerevisiae* is natively proficient for SG, several aspects of '*in yeasto*' assembly are still far from optimal. Firstly, compared to bacterial alternatives, *S. cerevisiae* cells grow slowly with a maximum specific growth rate around 0.4 - 0.5 h^{-1} and are hard to disrupt due to their sturdy cell wall. Considering that large DNA constructs above a few hundred kilobases are sensitive to shear stress, chromosome extraction and purification from *S. cerevisiae* is possible, but remains tenuous and inefficient, leading to low DNA yields and potentially damaged chromosomes [44]. Secondly, the strength of *S. cerevisiae* can become its weakness, as the HR machinery can be overzealous and recombine any (short) DNA sequence with homology within or between the (neo)chromosomes, which may lead to misassemblies. Lastly, non-homologous end joining and microhomology-

mediated end joining, DNA repair mechanisms that assemble pieces of DNA with no or minimal homology, are present in *S. cerevisiae* with low activity [45, 46], and can also cause misassemblies. Similar to how *E. coli* was engineered to become a lab tool for DNA amplification, these shortcomings could be alleviated by engineering *S. cerevisiae* into a more powerful genome foundry.

Are there future alternatives to *S. cerevisiae*? Naturally, *B. subtilis* and *E. coli* could also be engineered. However, considering the minute fraction of the vast microbial biodiversity that has been tested for genetic accessibility and DNA assembly, it is likely that microbes yet to be discovered are even better genome foundries. Environments causing extreme DNA damage (high radiation, toxic chemicals, etc.) might be a source of HR-proficient organisms (e.g. [47, 48]) better suited for SG.

In a more distant future, *in vitro* alternatives might replace the need for live DNA foundries altogether, thereby accelerating and simplifying genome construction. However, this will require major technological advances in *in vitro* DNA assembly and amplification. Already substantial efforts have led to the development of methods for DNA amplification, such as rolling circle amplification by the phage φ29 DNA polymerase [49, 50], recently implemented for the amplification of a 116 kb multipartite genome [20] and the *in vitro* amplification of synthetic genomes using the *E. coli* replisome, which already demonstrated to be capable of amplification of 1Mb synthetic genomes [22]. Targets for improvement of these methods are the maximal length of amplified DNA fragments, the yield of amplification, the need for restriction of the amplified, concatenated molecules or the formation of non-specifically amplified products. The development of an *in vitro* approach that can parallel *S. cerevisiae in vivo* assembly capability seems even more challenging. While an interesting avenue might be to transplant *S. cerevisiae* HR DNA repair *in vitro*, it presents a daunting task considering that all players and their respective role have not been fully elucidated yet [46, 51]. Still, considering that highly complex systems such as the transcription and translation machineries have been successfully implemented *in vitro* and are commercially available [52], cell-free *S. cerevisiae* HR might become a reality in the coming years.

Outlook

Since the first genome synthesis in 2008, relatively few genomes have been synthetized. Low-cost, customizable construction of designer genomes, currently accessible for small viral, organellar or bacterial constructs, is still out of reach for large (eukaryotic) genomes. There are still numerous technical, financial, and computational hurdles that must be overcome on the road to microbial designer genomes, tailored to applications in bio-based industry. Here we reviewed why the yeast *S. cerevisiae* is a key organism in the field of SG, however, the spectrum of

available hosts is expected to increase as research in SG advances. For example, a recent study shows improving the HR capacity of the industrially relevant yeast *Yarrowia lipolytica* could greatly expand the potential applications of SG in bio-based processes [53].

In the near future, SG is anticipated to contribute to various fields, such as a platform technology for industrial biotechnological processes [3, 43], as a new means for data storage [54] and for the development of new cell therapies and other medical applications, which is the ambition of the Genome Project-Write [55]. In parallel, worldwide bottom-up approaches endeavor to construct synthetic cells from scratch, such as the European consortia BaSyC (http://www.basyc.nl), MaxSynBio (https://www.maxsynbio.mpg.de) and the Synthetic cell initiative (http://www.syntheticcell.eu) and the US-based Build-a-cell initiative (http://buildacell.io) (reviewed in [56]). Looking further ahead, SG may even assist in understanding and engineering entire ecosystems by assembly of a metagenomes in a single cell [57]. SG, albeit still in its infancy and mostly limited to academic research, has bright days ahead, and *S. cerevisiae* is foreseen to remain a valuable, if not indispensable, SG tool for the coming decade.

Table 1 - Overview of the contribution of *S. cerevisiae* in synthetic genomics by the assembly of large (>100 kb) DNA constructs.

	Donor DNA	Nr. of transformed fragments[1]	Size of transformed fragments[1,2]	Size of final construct	Aim of yeast assembly	Reference
Viruses	Herpes simplex type 1	11	14 kb	152 kb	Assembly and modification of viral genome, transfection and reconstitution in mammalian cells	[58]
	Autographa californica nucleopolyhedrovirus	4	45 kb	145 kb	Assembly and modification of viral genome, transfection and reconstitution in insect cells.	[59]
	Cytomegalovirus isolate Toledo	3	116 kb	230 kb	Assembly and modification of viral genome, transfection and reconstitution in mammalian cells.	[60]
Prokaryotes	*Mycoplasma genitalium*	6	Up to 144 kb	592 kb	Assembly of synthetic *M. genitalium* genome which could not be stably maintained in *E. coli*.	[4]
	Mycoplasma genitalium	25	17-35 kb	592 kb	Assembly of synthetic *M. genitalium* genome from short fragments, exploring assembly capacity in yeast.	[5]
	Mycoplasma mycoides	11	100 kb	1 Mb	Assembly of synthetic *M. mycoides* genome, transplantation to a recipient cell to create the first bacterial cell controlled by a synthesized genome.	[61]
	Mycoplasma pneumonia	2	10 kb - 816 kb	826 kb	Insertion of yeast regulatory elements in the full *M. pneumonia* genome to allow for cloning and engineering of the genome.	[34, 62]
	Mycoplasma hominis	2	5 kb – 665 kb	670 kb	Insertion of yeast regulatory elements in the full *M. hominis* genome to allow for cloning and engineering of the genome.	[63]
	Acholeplasma laidlawii	3[3]	121-897 kb	1.38 Mb	Exploring potential toxicity when assembling bacterial genomes in yeast.	[64]

	Escherichia coli	3	185 – 660 kb	1.03 Mb	Assembly of a minimal *E. coli* genome by Cas9-induced recombination of partial genomes.	[65]
	Escherichia coli	7-14	6-13 kb	100 kb	Assembly of recoded *E. coli* partial genomes, used to replace the *E. coli* genome by a recoded synthetic genome.	[31]
	Caulobacter crescentus	16	38-65 kb	785 kb	Assembly of a minimized and synthetic *C. crescentus* genome, recoded to be compatible with chemical DNA synthesis and transplanted in a recipient cell.	[30]
	Prechlorococcus marinus	2	580-675 kb	1.66 Mb	Exploring assembly capacity and DNA stability of exogenous genomes in yeast.	[40]
	Synechococcus elongatus	4	100 - 200 kb	454 kb	Exploring the ability to clone genomes with high G/C-content in yeast.	[66]
Algae	*Phaeodactylum tricornutum*	5	106-128 kb	497 kb	Assembly of DNA with a moderate G + C content as a case study for assembly and modification of eukaryotic chromosomes in yeast.	[67]
	Chlamydomonas reinhardtii chloroplast genome	6	34-129 kB	230 kb	Assembly of a partial *C. reinhardtii* chloroplast genome to create genetic diversity at multiple loci at once	[68]
Yeasts	Yeast chromosome XII	33[4]	26-39 kb	976 kb	Assembly of a megabase synthetic yeast chromosome harboring the highly repetitive ribosomal DNA locus.	[69]
	Single-chromosome yeast	15[4]	230-1500 kb	11 Mb	Assembly of all sixteen *S. cerevisiae* chromosomes into a single chromosome .	[24]
	Yeast neochromosome	44	2.5 kb	100 kb	Assembly of a circular supernumerary *S. cerevisiae* neochromosome that can act as a platform for modular genome engineering.	[42]
	Yeast neochromosome for pathway engineering	43	2.5 – 5 kb	100 kb	Assembly of a circular and linear supernumerary *S. cerevisiae* neochromosomes for expression of heterologous and essential metabolic pathways.	[43]

Other	Human *HPRT1* gene	13	3-83 kb	125 kb	Assembly of a synthetic human *HRPT1* gene and transplantation and expression in mammalian cells.	[35]
	Artificial data storage chromosome	5	40 kb	254 kb	Assembly of a *S. cerevisiae* artificial chromosome containing data-encoded DNA for digital data storage.	[54]

[1] In case of a sequential assembly, the fragment number and size of the last assembly is used
[2] Short backbones containing regulatory elements such as CEN/ARS and markers not included
[3] Initial assembly of the entire genome failed due to gene toxicity
[4] Assembly was performed by stepwise integration in multiple rounds

References

1. Coradini, A.L.V., C.B. Hull, and I.M. Ehrenreich, Building genomes to understand biology. Nature Communications, 2020. **11**(1).DOI: 10.1038/s41467-020-19753-2.
2. Zhang, W., et al., Synthetic genomes. Annual Review of Biochemistry, 2020. **89**: p. 77-101.
3. Schindler, D., Genetic Engineering and Synthetic Genomics in Yeast to Understand Life and Boost Biotechnology. Bioengineering, 2020. **7**(4): p. 137.
4. Gibson, D.G., et al., Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. science, 2008. **319**(5867): p. 1215-1220.
5. Gibson, D.G., et al., One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. Proceedings of the National Academy of Sciences, 2008. **105**(51): p. 20404-20409.
6. Beaucage, S. and M. Caruthers, Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron letters, 1981. **22**(20): p. 1859-1862.
7. Hughes, R.A. and A.D. Ellington, Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. Cold Spring Harbor Perspectives in Biology, 2017. **9**(1): p. a023812.DOI: 10.1101/cshperspect.a023812.
8. Lee, H., et al., Photon-directed multiplexed enzymatic DNA synthesis for molecular digital data storage. Nat Commun, 2020. **11**(1): p. 5246.DOI: 10.1038/s41467-020-18681-5.
9. Lee, H.H., et al., Terminator-free template-independent enzymatic DNA synthesis for digital information storage. Nat Commun, 2019. **10**(1): p. 2383.DOI: 10.1038/s41467-019-10258-1.
10. Ostrov, N., et al., Technological challenges and milestones for writing genomes. Science, 2019. **366**(6463): p. 310-312.
11. Paul, S.S., et al., Advances in long DNA synthesis, in *Microbial Cell Factories Engineering for Production of Biomolecules*. 2021, Elsevier. p. 21-36.
12. Hao, M., J. Qiao, and H. Qi, Current and Emerging Methods for the Synthesis of Single-Stranded DNA. Genes, 2020. **11**(2): p. 116.DOI: 10.3390/genes11020116.
13. Eisenstein, M., Enzymatic DNA synthesis enters new phase. Nature Biotechnology, 2020. **38**(10): p. 1113-1115.DOI: 10.1038/s41587-020-0695-9.
14. Forster, A.C. and G.M. Church, Towards synthesis of a minimal cell. Mol Syst Biol, 2006. **2**: p. 45.DOI: 10.1038/msb4100090.
15. Chao, R., Y. Yuan, and H. Zhao, Recent advances in DNA assembly technologies. FEMS Yeast Research, 2014: p. n/a-n/a.DOI: 10.1111/1567-1364.12171.
16. Casini, A., et al., Bricks and blueprints: methods and standards for DNA assembly. Nature Reviews Molecular Cell Biology, 2015. **16**(9): p. 568-576.DOI: 10.1038/nrm4014.
17. Gibson, D.G., et al., Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods, 2009. **6**(5): p. 343-345.DOI: 10.1038/nmeth.1318.
18. Gibson, D.G., et al., Chemical synthesis of the mouse mitochondrial genome. Nature methods, 2010. **7**(11): p. 901-903.
19. Karas, B.J., Y. Suzuki, and P.D. Weyman, Strategies for cloning and manipulating natural and synthetic chromosomes. Chromosome Research, 2015. **23**(1): p. 57-68.DOI: 10.1007/s10577-014-9455-3.
20. Libicher, K., et al., In vitro self-replication and multicistronic expression of large synthetic genomes. Nature communications, 2020. **11**(1): p. 1-8.
21. van Nies, P., et al., Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. Nat Commun, 2018. **9**(1): p. 1583.DOI: 10.1038/s41467-018-03926-1.

22. Mukai, T., et al., Overcoming the Challenges of Megabase-Sized Plasmid Construction in *Escherichia coli.* ACS Synthetic Biology, 2020. **9**(6): p. 1315-1327.
23. Su'Etsugu, M., et al., Exponential propagation of large circular DNA by reconstitution of a chromosome-replication cycle. Nucleic Acids Research, 2017. **45**(20): p. 11525-11534.DOI: 10.1093/nar/gkx822.
24. Shao, Y., et al., Creating a functional single-chromosome yeast. Nature, 2018. **560**(7718): p. 331-335.
25. Burke, D.T., G.F. Carle, and M.V. Olson, Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science, 1987. **236**(4803): p. 806-812.
26. Larionov, V., et al., Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. Proceedings of the National Academy of Sciences, 1996. **93**(1): p. 491-496.
27. Kunes, S., D. Botstein, and M.S. Fox, Transformation of yeast with linearized plasmid DNA: formation of inverted dimers and recombinant plasmid products. Journal of molecular biology, 1985. **184**(3): p. 375-387.
28. Noskov, V., et al., Defining the minimal length of sequence homology required for selective gene isolation by TAR cloning. Nucleic acids research, 2001. **29**(6): p. e32-e32.
29. Larionov, V., et al., Transformation-associated recombination between diverged and homologous DNA repeats is induced by strand breaks. Yeast, 1994. **10**(1): p. 93-104.DOI: 10.1002/yea.320100109.
30. Venetz, J.E., et al., Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. Proceedings of the National Academy of Sciences, 2019. **116**(16): p. 8070-8079.
31. Fredens, J., et al., Total synthesis of *Escherichia coli* with a recoded genome. Nature, 2019. **569**(7757): p. 514-518.
32. Thao, T.T.N., et al., Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. Nature, 2020. **582**(7813): p. 561-565.
33. Vashee, S., Y. Arfi, and C. Lartigue, Budding yeast as a factory to engineer partial and complete microbial genomes. Current Opinion in Systems Biology, 2020.
34. Benders, G.A., et al., Cloning whole bacterial genomes in yeast. Nucleic acids research, 2010. **38**(8): p. 2558-2569.
35. Mitchell, L.A., et al., De novo assembly and delivery to mouse cells of a 101 kb functional human gene. Genetics, 2021. **218**(1): p. iyab038.
36. Pretorius, I. and J. Boeke, Yeast 2.0—connecting the dots in the construction of the world's first functional synthetic eukaryotic genome. FEMS yeast research, 2018. **18**(4): p. foy032.
37. Eisenstein, M., How to build a genome. Nature, 2020. **578**(7796): p. 633-635.DOI: 10.1038/d41586-020-00511-9.
38. Dai, J., et al., Sc3.0: revamping and minimizing the yeast genome. Genome Biology, 2020. **21**(1).DOI: 10.1186/s13059-020-02130-z.
39. Sorek, R., et al., Genome-Wide Experimental Determination of Barriers to Horizontal Gene Transfer. Science, 2007. **318**(5855): p. 1449-1452.DOI: 10.1126/science.1147112.
40. Tagwerker, C., et al., Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. Nucleic acids research, 2012. **40**(20): p. 10375-10383.
41. Itaya, M., et al., Far rapid synthesis of giant DNA in the *Bacillus subtilis* genome by a conjugation transfer system. Scientific reports, 2018. **8**(1): p. 8792.
42. Postma, E.D., et al., A supernumerary designer chromosome for modular in vivo pathway assembly in *Saccharomyces cerevisiae.* Nucleic Acids Research, 2021. **49**(3): p. 1769-1783.DOI: 10.1093/nar/gkaa1167.

43. Postma, E.D., et al., Modular, synthetic chromosomes as new tools for large scale engineering of metabolism. Metabolic Engineering, 2022. **72**: p. 1-13.
44. Blount, B.A., M.R.M. Driessen, and T. Ellis, GC Preps: Fast and Easy Extraction of Stable Yeast Genomic DNA. Scientific Reports, 2016. **6**(1): p. 26863.DOI: 10.1038/srep26863.
45. Lee, K., et al., Microhomology Selection for Microhomology Mediated End Joining in *Saccharomyces cerevisiae.* Genes, 2019. **10**(4): p. 284.DOI: 10.3390/genes10040284.
46. Ranjha, L., S.M. Howard, and P. Cejka, Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. Chromosoma, 2018. **127**(2): p. 187-214.DOI: 10.1007/s00412-017-0658-1.
47. Albarracín, V.H., et al., Extremophilic *Acinetobacter* Strains from High-Altitude Lakes in Argentinean Puna: Remarkable UV-B Resistance and Efficient DNA Damage Repair. Origins of Life and Evolution of Biospheres, 2012. **42**(2-3): p. 201-221.DOI: 10.1007/s11084-012-9276-3.
48. Sato, T., et al., Integration of large heterologous DNA fragments into the genome of *Thermococcus kodakarensis.* Extremophiles, 2020. **24**(3): p. 339-353.DOI: 10.1007/s00792-020-01159-z.
49. Dean, F.B., Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primed Rolling Circle Amplification. Genome Research, 2001. **11**(6): p. 1095-1099.DOI: 10.1101/gr.180501.
50. Lau, Y.H., et al., Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA. Nucleic acids research, 2017. **45**(11): p. 6971-6980.
51. Kwon, Y., J.M. Daley, and P. Sung, Reconstituted system for the examination of repair DNA synthesis in homologous recombination. Methods in enzymology, 2017. **591**: p. 307-325.
52. Shimizu, Y., et al., Cell-free translation reconstituted with purified components. Nature biotechnology, 2001. **19**(8): p. 751.
53. Guo, Z.-p., et al., An artificial chromosome ylAC enables efficient assembly of multiple genes in *Yarrowia lipolytica* for biomanufacturing. Communications biology, 2020. **3**(1): p. 1-10.
54. Chen, W., et al., An artificial chromosome for data storage. National Science Review, 2021. **8**(5): p. nwab028.
55. Boeke, J.D., et al., The Genome Project-Write. Science, 2016. **353**(6295): p. 126-127.DOI: 10.1126/science.aaf6850.
56. Mutschler, H., et al., Special Issue on Bottom-Up Synthetic Biology. ChemBioChem, 2019. **20**(20): p. 2533-2534.DOI: 10.1002/cbic.201900507.
57. Belda, I., et al., Seeding the idea of encapsulating a representative synthetic metagenome in a single yeast cell. Nature Communications, 2021. **12**(1): p. 1-8.
58. Oldfield, L.M., et al., Genome-wide engineering of an infectious clone of herpes simplex virus type 1 using synthetic genomics assembly methods. Proceedings of the National Academy of Sciences, 2017. **114**(42): p. E8885-E8894.DOI: 10.1073/pnas.1700534114.
59. Shang, Y., et al., Construction and rescue of a functional synthetic baculovirus. ACS synthetic biology, 2017. **6**(7): p. 1393-1402.
60. Vashee, S., et al., Cloning, assembly, and modification of the primary human cytomegalovirus isolate Toledo by yeast-based transformation-associated recombination. Msphere, 2017. **2**(5): p. e00331-17.
61. Gibson, D.G., et al., Creation of a bacterial cell controlled by a chemically synthesized genome. Science, 2010. **329**(5987): p. 52-56.

62. Ruiz, E., et al., CReasPy-cloning: a method for simultaneous cloning and engineering of megabase-sized genomes in yeast using the CRISPR-Cas9 system. ACS synthetic biology, 2019. **8**(11): p. 2547-2557.
63. Rideau, F., et al., Cloning, stability, and modification of *Mycoplasma hominis* genome in yeast. ACS synthetic biology, 2017. **6**(5): p. 891-901.
64. Karas, B.J., et al., Cloning the *Acholeplasma laidlawii* PG-8A genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. ACS synthetic biology, 2012. **1**(1): p. 22-28.
65. Zhou, J., et al., CasHRA (Cas 9-facilitated Homologous Recombination Assembly) method of constructing megabase-sized DNA. Nucleic acids research, 2016. **44**(14): p. e124-e124.
66. Noskov, V.N., et al., Assembly of large, high G+C bacterial DNA fragments in yeast. ACS synthetic biology, 2012. **1**(7): p. 267-273.
67. Karas, B.J., et al., Assembly of eukaryotic algal chromosomes in yeast. Journal of biological engineering, 2013. **7**(1): p. 30.
68. O'Neill, B.M., et al., An exogenous chloroplast genome for complex sequence manipulation in algae. Nucleic acids research, 2012. **40**(6): p. 2782-2792.
69. Zhang, W., et al., Engineering the ribosomal DNA in a megabase synthetic chromosome. Science, 2017. **355**(6329).

Outlook

In this thesis, swapping of the glycolytic pathway in *Saccharomyces cerevisiae* was used to study different aspects of this essential and central pathway. In **Chapter 2** the phenotypic and transcriptional impact of a relocated glycolytic pathway was characterized in the original platform strain for pathway swapping, which served as a basis for the other chapters. Co-localization of genes belonging to a single pathway enables simpler metabolic engineering of microbial cell factories and can, in future work, lead to a modular genome design, where the desired genes and metabolic pathways can be selected, synthetized and integrated together. However, this would be greatly complicated by the presence of physiological defects resulting only from genomic reorganization through secondary effects on DNA transcription and translation. While pathway integration and even genome engineering in yeast is applied relatively frequently, the physiological effects of genome reorganization are usually not studied in detail [1, 2]. Using the modular glycolysis strains as paradigm, however, no direct effects from pathway relocation and genetic factors such as integration locus, presence of ARS sequences and gene direction on physiology were observed, as previously suspected [3]. These findings, combined with the power of evolutionary engineering to resolve growth defects resulting from rational engineering suggest modular pathway design can be freely expanded to other essential pathways and larger scale genomic reorganization for increased genetic accessibility. Indeed genome reorganization combined with randomized rearrangement and selection is already being applied to complete synthetic *S. cerevisiae* chromosomes [4, 5]. Rational strain improvement however will require the proper assessment of the effects of genetic rearrangement, which is complicated by the unknown impact of untargeted mutations resulting from strain construction, mutagenesis or adaptive evolution. Since each genetic transformation round has the potential for accumulation of such mutations, these cannot easily be repaired by further genetic engineering. This work showed the challenge of associating a simple phenotype with the causal genetic intervention and this is only likely to be multiplied if the scope of genetic modification increases. To limit the impact of untargeted mutations, genome engineering should therefore be carried out in as few steps as possible, using the latest developments in genetic engineering [6-8] and assembly of modular pathways in a limited number of genomic loci. Additionally, whole-genome sequencing and physiological characterization of intermediate strains will be of paramount importance in complex strain construction programs.

In **Chapter 3** the minimal glycolysis and pathway swapping strains were used to humanize the glycolytic pathway in yeast and the use of humanized strains as model

organisms was investigated. Humanization of the glycolytic pathway in yeast proved possible, with a complete transplantation of the human muscle glycolytic pathway leading to a functional pathway. Properties of the glycolytic enzymes were shown to be largely comparable to those obtained from human muscle cells, opening the way to the use of humanized glycolysis strains as models to study this essential pathway. Humanization of yeast has previously been used to generate experimental models to study specific diseases such as Parkinson's and Alzheimer [9, 10] and the humanization of complete pathways and processes can expand the use of yeast as model organism [11]. With the hypothesized role of glycolytic flux and specific glycolytic isoenzymes in diseases such as cancer [12, 13], implementation in yeast of disease-specific glycolytic pathways can be pursued to help unravel the role of the pathway and its component enzymes. The absence in yeast of human regulatory elements and cell components can limit the use of such models on one hand, since interactions between glycolytic enzymes and other cell components undoubtedly impacts their function and role in various diseases [14, 15]. However, exactly the absence of such interaction partners, allowing the study of enzyme variants in a 'clean' yeast background, could enable the untangling of complex functions and the testing of hypotheses. Challenges of the pathway swapping approach were associated with the complexity of the regulation and function of glycolytic enzymes. The requirement of mutations relieving product inhibition of the muscle hexokinases for functional expression in yeast show the importance of allosteric regulation in glycolysis and the possibility of quick adaptation of an enzyme to a new cellular environment. Additionally, moonlighting functions of the glycolytic enzymes proved to be largely complemented after pathway transplantation. This poorly studied aspect of heterologous enzyme function is critical to the interpretation of pathway swapping studies, since disentangling the effects of moonlighting and catalytic functions can be challenging. Information on moonlighting functions is far from complete, and glycolytic, as well as other metabolic enzymes, likely have more unknown functions and interaction partners [16]. Unravelling the effects of the glycolytic pathway and its diverse functions in human health and disease therefore proves to be a major future challenge in which humanized yeast strains can play an important role.

In **Chapter 4** the concept of glycolysis swapping was applied to explore the role of allosteric regulation in glycolysis by expressing glycolytic enzymes from the oleaginous yeast *Yarrowia lipolytica* in *S. cerevisiae*. Functional replacement of the key regulatory steps of hexokinase, phosphofructokinase and pyruvate kinase showed the requirement for metabolic regulation of these enzymes and exposed large differences in functional expression in different pathway contexts. Similar to **Chapter 3**, regulation and activity of the glucose phosphorylation step proved critical

for functional pathway transplantation, as the heterologous glycolytic pathway could only be functional with a reduced rate of glucose phosphorylation. These results confirmed previous hypotheses on the roles and importance of metabolic regulation for glycolytic stability [17-19], and showed new redundancies and interconnections of regulation. Pathway swapping, combined with metabolic modelling and detailed biochemical characterization of enzymes proved therefore to be a powerful approach to study metabolic regulations on a pathway level. Especially the ability to construct and test various mosaic pathways was essential to study metabolic regulation, although the combinations tested were not exhaustive. Testing of additional combinations of regulated and deregulated enzymes, as well as the confirmation of observations with *S. cerevisiae* enzyme mutants devoid of specific interactions would strengthen the conclusions of the current work. This method can be further applied in future work to unravel the role of specific regulations, modifications, and interactions in glycolysis and other metabolic pathways.

In this thesis pathway swapping was used mainly to study fundamental questions on the function, regulation and conservation of glycolysis, however the same principle could be applied in metabolic engineering. Engineering of metabolism remains challenging due to the number and complexity of metabolic genes and pathways and could be simplified by modular organization of the genome according to function, a concept that has already been explored for minimal bacterial genomes [20]. Integration of heterologous, highly expressed, multi-gene product pathways is already commonly employed in yeast metabolic engineering (e.g. [1, 21, 22]), and the large-scale synthesis and reorganization of the yeast genome is currently being finalized [4, 23]. Additionally, it was recently shown that synthetic chromosomes can be used as expression platforms for glycolysis as well as new product pathways [24, 25]. The combination of these technologies with the modular rewiring of central carbon metabolism could lead to powerful and flexible cell factories, starting from a modular genome platform strain. Challenges remain however, besides limited fundamental understanding of enzyme and pathway function, rational design and assembly of large DNA constructs will require efficient methods of DNA assembly. In **Chapter 5** the role of yeast in the construction of rationally designed chromosomes and genomes is reviewed, showing the extensive contribution of this microorganism to this emerging field. Eventually, progression in synthetic biology could enable the design of custom pathways, chromosomes and even microorganisms ideal for production of specific chemical compounds.

Overall swapping of the glycolytic pathway of *S. cerevisiae* proved to be a powerful addition to the genetic engineering toolbox, with potential benefits for fundamental understanding as well as metabolic engineering. However, the wide range of possible effects resulting from transplanting essential metabolic pathways that are

entwined with many cellular functions through regulation, moonlighting functions and physical protein interactions can make interpretation of results difficult. Analysis of multiple pathway variants with different enzymes can help bridge this knowledge gap by pinpointing the essential steps, but requires relatively high-throughput methods to be efficient. For metabolic engineering, automation of strain design and construction and improvements in genetic engineering show great promise, and designer pathways and chromosomes appear to be within reach [21, 25, 26]. However, if strains with different pathway variants can be made faster and easier, this will shift the bottle-neck from strain construction to strain characterization. To move beyond screening for growth rate or product formation and generate new understanding of fundamental processes, broad physiological characterization under multiple different conditions, including dynamic environments, in combination with classical biochemical methods will be required. Such methods are however difficult to scale-up to many strains, and leveraging the new developments in automation, high throughput screening and *in silico* strain design for application to fundamental questions will be a great challenge for synthetic biology. Ultimately however, this will be necessary to build a better understanding of how metabolic pathways operate and interact and how they can be manipulated most effectively for human benefit.

References

1. Galanie, S., et al., Complete biosynthesis of opioids in yeast. Science, 2015. **349**(6252): p. 1095-100.DOI: 10.1126/science.aac9373.
2. Zhang, W., et al., Engineering the ribosomal DNA in a megabase synthetic chromosome. Science, 2017. **355**(6329).
3. Kuijpers, N.G., et al., Pathway swapping: Toward modular engineering of essential cellular processes. Proceedings of the National Academy of Sciences, USA, 2016. **113**(52): p. 15060-15065.DOI: 10.1073/pnas.1606701113.
4. Richardson, S.M., et al., Design of a synthetic yeast genome. Science, 2017. **355**(6329): p. 1040-1044.DOI: 10.1126/science.aaf4557.
5. Dymond, J. and J. Boeke, The *Saccharomyces cerevisiae* SCRaMbLE system and genome minimization. Bioengineered, 2012. **3**(3): p. 168-71.DOI: 10.4161/bbug.19543.
6. Mans, R., et al., CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae.* FEMS Yeast Research, 2015. **15**(2).DOI: 10.1093/femsyr/fov004.
7. Randazzo, P., et al., gEL DNA: A Cloning-and Polymerase Chain Reaction–Free Method for CRISPR-Based Multiplexed Genome Editing. The CRISPR Journal, 2021.
8. Malci, K., L.E. Walls, and L. Rios-Solis, Multiplex Genome Engineering Methods for Yeast Cell Factory Development. Front Bioeng Biotechnol, 2020. **8**: p. 589468.DOI: 10.3389/fbioe.2020.589468.
9. Pereira, C., et al., A yeast model of the Parkinson′ s disease-associated protein Parkin. Experimental cell research, 2015. **333**(1): p. 73-79.
10. Seynnaeve, D., et al., Recent insights on Alzheimer's disease originating from yeast models. International journal of molecular sciences, 2018. **19**(7): p. 1947.
11. Laurent, J.M., et al., Efforts to make and apply humanized yeast. Briefings in Functional Genomics, 2016. **15**(2): p. 155-163.DOI: 10.1093/bfgp/elv041.
12. Mathupala, S., Y.a. Ko, and P. Pedersen, Hexokinase II: cancer's double-edged sword acting as both facilitator and gatekeeper of malignancy when bound to mitochondria. Oncogene, 2006. **25**(34): p. 4777-4786.
13. Altenberg, B. and K.O. Greulich, Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. Genomics, 2004. **84**(6): p. 1014-20.DOI: 10.1016/j.ygeno.2004.08.010.
14. Blaha, C., et al., A novel non-catalytic scaffolding activity of Hexokinase 2 contributes to EMT and metastasis. bioRxiv, 2021.DOI: 10.1101/2021.04.08.439049.
15. Sola-Penna, M., et al., Regulation of mammalian muscle type 6-phosphofructo-1-kinase and its implication for the control of the metabolism. IUBMB life, 2010. **62**(11): p. 791-796.DOI: 10.1002/iub.393.
16. Copley, S.D., Moonlighting is mainstream: Paradigm adjustment required. Bioessays, 2012. **34**(7): p. 578-588.DOI: 10.1002/bies.201100191.
17. Teusink, B., et al., The danger of metabolic pathways with turbo design. Trends in Biochemical Sciences, 1998. **23**(5): p. 162-169.
18. van Heerden, J.H., F.J. Bruggeman, and B. Teusink, Multi-tasking of biosynthetic and energetic functions of glycolysis explained by supply and demand logic. BioEssays, 2015. **37**(1): p. 34-45.
19. van Heerden, J.H., et al., Lost in transition: startup of glycolysis yields subpopulations of nongrowing cells. Science, 2014.DOI: science.1245114 [pii];10.1126/science.1245114 [doi].
20. Hutchison, C.A., et al., Design and synthesis of a minimal bacterial genome. Science, 2016. **351**(6280).

21. Young, E.M., et al., Iterative algorithm-guided design of massive strain libraries, applied to itaconic acid production in yeast. Metab Eng, 2018. **48**: p. 33-43.DOI: 10.1016/j.ymben.2018.05.002.
22. Koopman, F., et al., De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae.* Microb Cell Fact, 2012. **11**(1): p. 155.DOI: 10.1186/1475-2859-11-155.
23. Annaluru, N., et al., Total synthesis of a functional designer eukaryotic chromosome. science, 2014. **344**(6179): p. 55-58.
24. Postma, E.D., et al., A supernumerary designer chromosome for modular *in vivo* pathway assembly in *Saccharomyces cerevisiae.* Nucleic Acids Res, 2021. **49**(3): p. 1769-1783.DOI: 10.1093/nar/gkaa1167.
25. Postma, E.D., et al., Modular, synthetic chromosomes as new tools for large scale engineering of metabolism. Metabolic Engineering, 2022. **72**: p. 1-13.
26. Si, T., et al., Automated multiplex genome-scale engineering in yeast. Nat Commun, 2017. **8**(1): p. 15187.DOI: 10.1038/ncomms15187.

Acknowledgements

Finally it is time to write what most likely will be the most read part of this book, the part where I thank everyone who made this work possible.

First of all I would like to thank Pascale, who as my promotor and primary supervisor was responsible for designing the project and continuously guiding me in the right direction. Your focus on people and making your PhD students independent researchers as well as your hard work, curiosity, great ambition and critical attitude made a project under your supervision a challenging but fun and very educational experience. I think it took some time for us to really understand each other, but once we did, our meetings became very fun and useful.

Also thanks to my other promotor, Jack, despite not being involved in the scientific part of my project as much as you would no doubt have liked due to administrative responsibilities, your infectious enthusiasm and great project management advice were definitely appreciated. Your management of first IMB and then the department have created a great environment for science. Jean-Marc, your extensive knowledge of all things yeast and molecular biology were always useful, thank you for your ideas, input and guidance, both during my PhD and earlier during my MSc project. Also your taking over of the group leadership during difficult times and no-nonsense attitude helped the group continue to do great work. I also want to thank Robert and Rinke for discussions on students, yeast physiology, mutations and generally always being willing to help and think along.

Almost no part of this work was done alone, and the most intensive collaborations were with my two main co-authors, Francine and Eline. Francine, your hard work on both the *Yarrowia* and human glycolysis projects, which turned out more difficult than expected and started before I even joined the group made this work possible. Your thesis is definitely the one I most frequently opened for inspiration (and to look things up). Our collaboration and office sharing including mutual venting of frustrations made my project much more fun. Eline, thanks for your endless positivity. Your drive and attention to detail really pushed our collaboration forward and helped make the chapter on the SwYG growth rate, as well as the review, happen. I also really enjoyed being Francine's paranymph together with you as well as being yours!

I also want to thank the other ERC/synthetic biology group members, Sofia, Melanie, Charlotte, Paola, Jordi and Céline, for their input during our biweekly meetings. Sitting through stories on little growth rate changes and failed PCRs is not everyone's idea of fun, but I think we got a lot of help out of each other.

I want to thank the staff technicians for keeping IMB running and generally being great colleagues. Erik, thanks for the help with fermentations, the growth profiler and the fun times giving practicals. Marijke, thank you for all the help with enzyme assays, the FACS, more enzyme assays, always thinking along and even more enzyme assays. Pilar and Clara, thanks for your help with sequencing, strain registration and keeping the molbio lab operational. Marcel, thanks for all the help with sequencing and transcriptomic data, uploading datasets and showing me the way around your sequencing analysis methods. The MSD ladies, Astrid, Jannie and Apilena are some of the most important people to keeping the lab running, thanks a lot for all your - mostly behind the scenes - efforts in autoclaving and media prep. Also thanks to all other (past and present) members of IMB for all their enthusiasm, great comments, questions, help, advice, lab retreats and fun times over the years. IMB has been and still is a place where everyone wants their colleagues to succeed and it was great to be a part of it for almost five years.

During these years, I had the honour of supervising ten students either partially or completely. Every student project was a new learning moment for me and we had to overcome some hurdles together, but I hope you all found your projects worth-while. Thank you all for your motivation, bright ideas and hard work Jelle, Linda, Jimmy, Liset, Nigell, Giulia, Anne-Marijn, Luca, Brian and Rachel.

Also thanks to my own MSc project supervisor, Arthur. Your drive, constant belief in other people, (especially your students) and always being ready for social activities made it great to have you as MSc project supervisor and I learned a lot from you (including tips on how to deal with French supervisors).

I also had the opportunity to work together with several people outside the comfort of IMB. Thanks to Marcel Vieira-Lara and Barbara Bakker for the fun collaboration on the human glycolysis work. Marcel, thank you for your enthusiasm and all the help with modelling. I'm glad I didn't scare you away from IMB and wish you lots of luck with your postdoc. Thanks to Bas Teusink and Philipp Savakis for their thinking along with the *Yarrowia* work and the exciting experiments (which didn't make it to the thesis) we managed to do in Amsterdam. I also want to thank all the other people who thought along and were willing to help, Peter-Leon Hagedoorn and Greg Bokinsky, thank you for your assistance with experiments and knowledge on Fe-S clusters (which also didn't make it into the thesis, but I learned a lot from it). Thanks to Carmen-Lisset Flores and Carlos Gancedo on their help with the *Yarrowia* work.

Finally I would like to thank Teun, Diederik, Robert and Sophie, it was fun on our holidays and trips together, complaining to each other about our various PhD frustrations and (sometimes) sharing successes. Sophie, also thanks for being a great senior PhD office buddy.

Lastly I want to thank my family, who always supported me throughout my various educational stages and Emma, who was always supportive and ready to distract me from thinking too much about work.

Curriculum Vitae

Ewout Knibbe was born on the 23rd of November 1993 in Amsterdam, The Netherlands. He grew up in Den Haag, Middelharnis and Middelburg before a move to Nieuwegein, where he finished his Gymnasium education at the Oosterlicht College, including an extracurricular program at the University of Utrecht. Subsequently, he moved to Leiden in 2012 and enrolled in the Bachelor Life Science & Technology at the University of Leiden and Delft University of Technology. After a minor following courses in Basel, Switzerland at the department of Biosystems Science and Engineering of the ETH Zürich he returned to Leiden for a Bachelor thesis project with Lizah van der Aart and prof. Gilles van Wezel, studying vancomycin resistance in *Streptomyces coelicolor*. Completing the Bachelor *Summa Cum Laude*, he moved on to the Master's program Life Science & Technology at Delft University of Technology in 2015 with specialization Cell Factory. In the second year of the Master he joined the Industrial Microbiology group for his Master thesis, supervised by Arthur Gorter de Vries and dr. Jean-Marc Daran. Working on chromosomal copy number variation in the beer-brewing yeast *Saccharomyces pastorianus*, this project further sparked his scientific interest. After an industrial internship on computational biology supervised by Jurgen Nijkamp and Joep Schmitz at DSM in Delft he completed his Master *Cum Laude* in 2017. Returning to the TU Delft in the same year he started a PhD under supervision of prof. dr. Pascale Daran-Lapujade and prof. dr. Jack T. Pronk in the Industrial Microbiology group. The results of this PhD project on yeast synthetic biology are described in this thesis.

List of publications

- Gorter de Vries[#], A. R., **Knibbe, E.** [#], van Roosmalen, R., van den Broek, M., de la Torre Cortes, P., O'Herne, S. F., Vijverberg, P. A., El Masoudi, A., Brouwers, N., Pronk, J. T., & Daran, J. G. (2020). Improving Industrially Relevant Phenotypic Traits by Engineering Chromosome Copy Number in *Saccharomyces pastorianus*. *Front Genet, 11*, 518. doi:10.3389/fgene.2020.00518

- Koster, C. C.[#], Postma, E. D.[#], **Knibbe, E.**[#], Cleij, C.[#], & Daran-Lapujade, P. (2022) Synthetic Genomics from a yeast perspective. *Frontiers in Bioengineering and Biotechnology,* 10, doi:10.3389/fbioe.2022.869486

- den Ridder, M., **Knibbe, E.**, van den Brandeler, W., Daran-Lapujade, P., & Pabst, M. (2022). A systematic evaluation of yeast sample preparation protocols for spectral identifications, proteome coverage and post-isolation modifications. *Journal of Proteomics,* 104576

- Boonekamp, F. J.[#], **Knibbe, E.**[#], Vieira-Lara, M. A., Wijsman, M., Luttik, M. A. H., van Eunen, K., den Ridder, M., Bron, R., Almonacid Suarez, A. M., van Rijn, P., Wolters, J. C., Pabst, M., Daran, J.-M., Bakker, B., & Daran-Lapujade, P. (2022). A yeast with muscle does not run faster: full humanization of the glycolytic pathway in *Saccharomyces cerevisiae*. *Submitted for publication in Cell Reports*

- **Knibbe E.**[#], Postma E. D.[#], Boonekamp, F. J., Dashko, S., Geelhoed, J., Maat, A.-M., van den Broek, M., Luttik, A. H., & Daran-Lapujade, P. (2022) What's wrong with SwYG? Reaching the limits of the *Saccharomyces cerevisiae* molecular and analytical toolbox. *Manuscript in preparation for submission*

- **Knibbe E.**, Boonekamp, F. J., Stuij, R., Pelsma, K. A. J., Jansen, L., Flores, C.-L., & Daran-Lapujade, P. (2022) Simpler is not always better: transplanting the *Yarrowia lipolytica* glycolytic pathway into *Saccharomyces cerevisiae* reveals essential synergetic regulatory mechanisms. *Manuscript in preparation for submission*

[#]These authors contributed equally