A Minimal Set of Glycolytic Genes Reveals Strong Redundancies in \textit{Saccharomyces cerevisiae} Central Metabolism

Daniel Solis-Escalante, Niels G. A. Kuijpers, Nuria Barrajon-Simancas, Marcel van den Broek, Jack T. Pronk, Pascale Daran-Lapujade

Department of Biotechnology, Delft University of Technology, Delft, The Netherlands

As a result of ancestral whole-genome and small-scale duplication events, the genomes of \textit{Saccharomyces cerevisiae} and many eukaryotes still contain a substantial fraction of duplicated genes. In all investigated organisms, metabolic pathways, and more particularly glycolysis, are specifically enriched for functionally redundant paralogs. In ancestors of the \textit{Saccharomyces} lineage, the duplication of glycolytic genes is purported to have played an important role leading to \textit{S. cerevisiae}'s current lifestyle favoring fermentative metabolism even in the presence of oxygen and characterized by a high glycolytic capacity. In modern \textit{S. cerevisiae} strains, the 12 glycolytic reactions leading to the biochemical conversion from glucose to ethanol are encoded by 27 paralogs. In order to experimentally explore the physiological role of this genetic redundancy, a yeast strain with a minimal set of 14 paralogs was constructed (the "minimal glycolysis" [MG] strain). Remarkably, a combination of a quantitative systems approach and semiquantitative analysis in a wide array of growth environments revealed the absence of a phenotypic response to the cumulative deletion of 13 glycolytic paralogs. This observation indicates that duplication of glycolytic genes is not a prerequisite for achieving the high glycolytic fluxes and fermentative capacities that are characteristic of \textit{S. cerevisiae} and essential for many of its industrial applications and argues against gene dosage effects as a means of fixing minor glycolytic paralogs in the yeast genome. The MG strain was carefully designed and constructed to provide a robust prototrophic platform for quantitative studies and has been made available to the scientific community.

Gene duplication plays a key role in evolution by providing DNA templates for evolutionary innovation, while preventing interference with the cellular function of the original genes (1–3). Immediately after gene duplication, the resulting paralog pairs are usually identical and therefore functionally redundant. Unless duplication confers a selective advantage, either via gene dosage effects or via mutational acquisition of modified or new functions, duplicated genes will eventually be pseudogenized and/or lost from the genome (2, 3).

Gene duplication played a key role in the evolutionary history of the yeast \textit{Saccharomyces cerevisiae}, an intensively investigated model in eukaryotic evolutionary biology. A whole-genome duplication event (WGD) in an ancestor of \textit{S. cerevisiae}, ca. 100 million years ago, was followed by loss of ca. 90% of the resulting gene duplications (4, 5). Despite the long time interval that separates current \textit{S. cerevisiae} strains from the WGD, many surviving paralog pairs still exhibit a substantial degree of functional redundancy, as indicated by neutral or weak phenotypes of single-paralog deletion mutants (6–9). However, the selective pressures that caused the long-term retention of these functionally redundant paralogs remain elusive (1, 10).

The Embden-Meyerhof-Parnas (EMP) pathway, the main route for oxidation of glucose to pyruvate in all eukaryotes and in many other organisms, is among the most slowly evolving metabolic pathways (11, 12). In all taxa, paralog families are found for structural genes that encode the EMP enzymes (12–14). Under conditions of oxygen limitation and/or sugar excess, \textit{S. cerevisiae} couples the EMP pathway to the fermentative production of ethanol via pyruvate decarboxylase and NAD⁺-dependent alcohol dehydrogenase (15, 16). In this article, we use the term "glycolysis" to indicate the set of 12 enzyme-catalyzed reactions in yeast that convert intracellular glucose to ethanol.

In \textit{S. cerevisiae}, no fewer than 8 of the 12 enzyme reactions in glycolysis are represented by multiple paralogous genes (Fig. 1). This incidence represents a significant overrepresentation \((P = 1.9 \times 10^{-10},\) based on hypergeometric distribution analysis) relative to the ca. 26% of the yeast genome that consists of paralogous combinations (17). The WGD event and resulting duplication of genes involved in central metabolism have been implicated in the appearance of several key physiological characteristics of \textit{S. cerevisiae}. In particular, the duplication of glycolytic genes has been proposed to have contributed to the strong tendency of \textit{S. cerevisiae} to produce ethanol under aerobic conditions (Crabtree effect) and its high glycolytic capacity (18–20). However, the impact of reducing the number of glycolytic paralogs on these and other physiological characteristics of \textit{S. cerevisiae} has not been systematically explored. In all paralogous gene sets in
yeast glycolysis, with the notable exception of the phosphofructokinase gene, gene expression and gene deletion studies support the definition of a single major paralog and one to four minor paralogs (Fig. 1). Except for the pseudogenes GPM2 and GPM3, all paralogs have retained their original catalytic function, although their context-dependent expression profiles differ (Fig. 1). Deletion of minor paralogs for individual glycolytic enzymes has minor effects on enzyme activities in cell extracts and on the specific growth rate under standard laboratory conditions (usually shake flask cultivation on yeast extract, peptone, and glucose) (21–25) (Fig. 1). Several hypotheses have been forwarded to explain the neutral effect of paralog deletion (6, 26–31). These include experimental limitations, such as the poor sensitivity of fitness screens (32) and the narrow range of cultivation condition tested (generally complex medium) (33). Additionally, analysis of deletion mutants in which paralogs that encode a single glycolytic enzyme are inactivated cannot reveal synergistic effects of the minor paralogs of different glycolytic enzymes. Such synergistic effects might, for example, arise from the well-documented phenomenon of distribution of metabolic control over multiple enzymes in metabolic pathways (34, 35) or from other regulatory or catalytic interactions.

Insight into the importance, under laboratory conditions, of the minor glycolytic paralogs is essential for understanding an apparent genetic redundancy in a key ubiquitous metabolic pathway in an important model organism and industrial platform. In addition, analysis of the extent of genetic redundancy in central metabolism is highly relevant for the complete redesign and construction of entirely synthetic yeast genomes, as pursued in the Synthetic Yeast 2.0 initiative (36). Moreover, if complexity in yeast glycolysis can be significantly reduced by elimination of “redundant” isoenzymes, this could eliminate uncertainties and thereby facilitate the formulation and validation of mathematical models that describe the kinetics of this key metabolic pathway (37, 38).

The goals of the present study are to experimentally explore genetic redundancy in yeast glycolysis by cumulative deletion of minor paralogs and to provide a new experimental platform for fundamental yeast research by constructing a yeast strain with a functional “minimal glycolysis” (MG). To this end, we deleted 13 minor paralogs, leaving only the 14 major paralogs for the S. cerevisiae glycolytic pathway. The cumulative impact of deletion of all minor paralogs was investigated by two complementary approaches. A first, quantitative analysis focused on the impact on glycolytic flux under a number of controlled cultivation conditions that in wild-type strains result in different glycolytic fluxes. These quantitative growth studies were combined with transcriptome, enzyme activity, and intracellular metabolite analysis to capture potential small phenotypic effects. A second, semi-quantitative characterization explored the phenotype of the “minimal glycolysis” (MG) strain under a wide array of experimental conditions to identify potential context-dependent phenotypes.

**MATERIALS AND METHODS**

**Strains and strain construction.** Plasmid propagation and isolation were performed with chemically competent Escherichia coli DH5α (Z-compotent transformation kit; Zymo Research, Orange, CA) cultivated in lysogeny broth (LB) medium (39, 40) supplemented with 100 mg liter⁻¹ ampicillin (LBamp) when required. All yeast strains are derived from the CEN.PK family (41–43) and are listed in Table S3 in the supplemental material. All strains were stored at −80°C in 1-ml aliquots of 30% glycerol and the appropriate agar (CEN.PK102-12A was selected as the parental strain for the minimal glycolysis (MG) strain. The order of gene deletions that led to the final MG strain IMX372 was GLK1, HXK1, TDH1, TDH2, GPM2, GPM3, ENO1, PYK2, PDC5, PDC6, ADH2, ADH5, and ADH4. The genes GLK1, HXK1, TDH1, and TDH2 were deleted using the auxotrophic and dominant markers Sphi5 (44, 45), KILEU2 (45, 46), KanMX (47, 48), and hphNT1 (49, 50), respectively. These markers remained in the genome during the whole strain construction process. GPM2 to ADH4 deletions were performed using a strategy of selection and counterselection with the KlURA3/S-5-fluoroorotic acid (5-FOA) system (51) for the recovery of the marker module. KlURA3 was removed seamlessly as previously described (52, 53). The dominant marker modules KanMX and hphNT1 were removed using deletion cassettes containing KlURA3 and amdsSYM, respectively. These markers were sequentially removed as reported previously (54). To obtain a prototrophic strain, a cassette containing the marker module KlURA3 was integrated in the TDH1 locus; this generated the MG strain. All genetic modifications were confirmed by PCR and later by whole-genome sequencing.

**Molecular biology techniques.** All integrative cassettes were constructed using Phusion Hot Start II high-fidelity polymerase (Thermo Scientific, Landsmeer, The Netherlands) following the manufacturer’s recommendations and with the primer pairs and plasmids listed in Tables S4 and S5 in the supplemental material, respectively, as the templates. Correct integrations, deletions, and marker excision were confirmed by PCR using Dream Taq polymerase (Thermo Scientific) and following the manufacturer’s recommendations. The primers used for confirmation are listed in Table S4. Genomic DNA that served as the template for PCR was obtained by extraction with 0.05 N NaOH directly from single colonies or by purification using the Yeastar genomic DNA kit (Zymo Research, Orange, CA) following the manufacturer’s recommendations. All PCR products were loaded on gels containing 1% (wt/vol) agarose (Thermo Scientific) and stained with 0.5 μg ml⁻¹ ethidium bromide. Molecular weight markers were added to the gel. DNA bands were visualized using a long-wave ultraviolet transilluminator and photographed with a gel documentation system (Alpha Innotech, San Leandro, CA).

**Strains and plasmids.** All strains were derived from the S. cerevisiae genome sequence (accession nos. NC_001021 and NC_001022). The following eight enzymes in yeast glycolysis are encoded by paralogous genes: hexokinase/glucokinase (HXK), phosphofructokinase (PFK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate mutase (GPM), enolase (ENO), pyruvate kinase (PYK), pyruvate decarboxylase (PDC), and alcohol dehydrogenase (ADH). Percentages of gene and protein similarity (GeSy and PrSy, respectively) between paralogs were calculated by using Blast2Go (http://blast2go.org/) (119). Scatter plots show the comparisons between expression levels of the different glycolytic paralogs in S. cerevisiae measured under 170 different conditions (75) (see Table S6 in the supplemental material). Except for PFK, where both paralogs PFK1 and PFK2 were considered paralogs with equivalent contributions, the expression levels of the major glycolytic paralogs (labeled in red) are indicated by a red line. Bar graphs display in vitro enzyme activities of glycolytic enzymes in mutants carrying individual and multiple glycolytic gene knockouts. Values are presented as percentages of in vitro enzyme activities measured in reference strains (RS [denoted by an orange dotted line]) as reported in the literature (21–25).

**FIG 1** Separation of major and minor glycolytic paralogs in Saccharomyces cerevisiae. The following eight enzymes in yeast glycolysis are encoded by paralogous genes: hexokinase/glucokinase (HXK), phosphofructokinase (PFK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate mutase (GPM), enolase (ENO), pyruvate kinase (PYK), pyruvate decarboxylase (PDC), and alcohol dehydrogenase (ADH). Percentages of gene and protein similarity (GeSy and PrSy, respectively) between paralogs were calculated by using Blast2Go (http://blast2go.org/) (119). Scatter plots show the comparisons between expression levels of the different glycolytic paralogs in S. cerevisiae measured under 170 different conditions (75) (see Table S6 in the supplemental material). Except for PFK, where both paralogs PFK1 and PFK2 were considered paralogs with equivalent contributions, the expression levels of the major glycolytic paralogs (labeled in red) are indicated by a red line. Bar graphs display in vitro enzyme activities of glycolytic enzymes in mutants carrying individual and multiple glycolytic gene knockouts. Values are presented as percentages of in vitro enzyme activities measured in reference strains (RS [denoted by an orange dotted line]) as reported in the literature (21–25).
Media. Complex and nonselective media for growth rate determina-
tion and propagation contained 10 g liter\(^{-1}\) yeast extract, 20 g liter\(^{-1}\)
peptone, and 20 g liter\(^{-1}\) glucose (YPD). When selection was required,
YPD was supplemented with 200 g liter\(^{-1}\) G418 (YPD+G418) or 200
mg liter\(^{-1}\) hygromycin (YPD+Hyg). Synthetic medium (SM) containing
3 g liter\(^{-1}\) KH\(_2\)PO\(_4\), 0.5 g liter\(^{-1}\) MgSO\(_4\), 7H\(_2\)O, 5 g liter\(^{-1}\) (NH\(_4\))\(_2\)SO\(_4\), 1
ml liter\(^{-1}\) of a trace element solution, and 1 ml liter\(^{-1}\) of a vitamin solu-
tion as previously described (56). SM was supplemented with 20 g liter\(^{-1}\)
glucose (SMG) for propagation, growth rate determination, and batch
cultures. When auxotrophic strains were cultivated, SMG was supple-
mented with 150 mg liter\(^{-1}\) uracil, 125 mg liter\(^{-1}\) histidine, and/or 500
mg liter\(^{-1}\) leucine, according to strain auxotrophy (57). Selection of
strains lacking the KIURA3 marker was done in SMG supplemented
with 150 mg liter\(^{-1}\) uracil and 1 g liter\(^{-1}\) 5-fluoroorotic acid (SMG–5-FOA).
Solid media were obtained by addition of 20 g liter\(^{-1}\) agar or agarose.
Unless otherwise stated, liquid cultures were performed in 500-ml shake
flasks with a working volume of 100 ml and incubated at 30°C and 250
rpm, the targeted starting optical density at 660 nm (OD\(_{660}\), measured
using a Libra 811 spectrophotometer (Biocrom, Cambridge, United King-
dom), being 0.5. Unless otherwise stated, cultures on solid media were
incubated for 3 days at 30°C.

Semiquantitative phenotypic characterization on plates. Growth of
the MG strain and of the prototrophic control strain CEN.PK113-7D was
performed on agar plates containing YPD or SM plus 2% glucose (SMG),
SM plus 2% galactose (SMGal), SM plus 2% maltose (SMMal), SM plus
2% sucrose (SMSuc), SM plus 3% (vol/vol) ethanol (SMEthOH), SM plus
6% glycerol (SMGlc), SM plus 5% glucose (SMHG), SM plus 0.5% glu-
cose (SMG), SMG plus 2.5, 5, or 10 mM LiCl (SMG-Lit), SMG plus 200
or 500 mM NaCl (SMGNaS), SMG plus 50, 100, or 200 μM CdCl\(_2\) (SMG-
Cad), SMG plus 1, 1.5, or 2 M sorbitol (SMG-Sor), SMG at pH 4, 6, or 7.5
(SMGph), or SMG plus 1 mM H\(_2\)SO\(_4\) (SMGox). Spot plates were inocu-
lated with 10\(^{-3}\) to 10\(^{-6}\) cells obtained by serial dilutions of liquid cultures
grown to the exponential phase on SMG. All plates were incubated at 30°C
for 3 days, except for plates containing SMGlc, which were incubated for
6 days.

Quantitative characterization of the MG strain’s physiology in
shake flasks. Growth rate determination was performed by OD\(_{660}\) measure-
cment of cultures grown in 500-ml shake flasks containing 100
ml of YPD, SMG, or SMEthOH. Cultures for growth rate measurements
were inoculated with precultures grown until the late exponential
phase under identical conditions (i.e., the same medium and temper-
ature). Glucose/ethanol/glucose switches were started by inoculation of SMG
shake flasks with a preculture grown overnight on the same medium.
At the mid-exponential phase (OD\(_{660}\) of ca. 4) samples were taken, washed
twice with sterile demineralized water, and used to inoculate shake flasks
containing 100 ml of SMEthOH. When reaching an OD\(_{660}\) of 3, a sample
from the culture was washed twice with sterile demineralized water and
used to inoculate a 500-ml shake flask containing 100 ml SMG. Growth
was then monitored until the late exponential phase.

To evaluate tolerance of MG to oxidative stress, cells grown to the
exponential phase in shake flasks containing SMG were transferred to a
shaking bioreactor. Aerobic and anaerobic batch cultures were performed in 2-
liter laboratory fermenters (Applikon, Schiedam, The Netherlands) with a
1-liter working volume. SM was used and set at pH 5 before autoclaving at
121°C. Antifoam emulsion C (Sigma) was added to a concentration of 0.2
g liter\(^{-1}\) from a 20% (vol/vol) solution autoclaved separately (121°C).
Prior to inoculation, glucose was added to a final concentration of 20 g
liter\(^{-1}\) from a sterile (110°C) 50% glucose solution and 1 ml of a vitamin
solution (56). Anaerobic cultures were supplemented with 0.01 g liter\(^{-1}\)
ergosterol and 0.42 g liter\(^{-1}\) Tween 80 dissolved in ethanol as previously
described (56). The pH was maintained at 5 by the automatic addition of
2 M KOH. Compressed air or gaseous nitrogen (quality, 4.5; >99.995%
vol N\(_2\), <10 volumes per million [vpm] O\(_2\) pollution) (Linde Gas,
Schiedam, The Netherlands) for aerobic and anaerobic cultures, respec-
tively, was sparged to the bioreactor at a rate of 500 ml min\(^{-1}\) via an Ion
Science Saga digital flow meter (Cambridge, United Kingdom). The tem-
perature in the fermenters was kept at 30°C. Complete mixing of the
medium was ensured by stirring at 800 rpm. Four independent batch
cultures were performed for each condition and each strain. Precultures
were started by inoculation of SMG shake flasks with 1 stock vial of the
appropriate strain. After ca. 16 h of incubation, these cultures were used to
inoculate new SMG shake flasks. The starting OD\(_{660}\) was tightly controlled
to reach an OD\(_{660}\) of 1 after 8 h of incubation. These exponentially grow-
ing cultures were washed twice with demineralized water (spinning at
5,000 rpm for 4 min) and resuspended in 100 ml demineralized water.
This was the cell suspension that was used to inoculate fermenters with a
starting OD\(_{660}\) of 0.1.

The OD\(_{660}\) of diluted cell suspensions was measured as described
above. Biomass dry weight was determined by filtration as previously
described (56). The concentration of extracellular metabolite in culture
supernatants was measured by high-performance liquid chromatography
(HPLC) analysis using a Aminex HPX-87H ion-exchange column oper-
ated at 60°C with 5 mM H\(_2\)SO\(_4\) as the mobile phase at a flow rate of 0.6 ml
min\(^{-1}\).

Chromosome separation using CHEF electrophoresis. Agarose
plugs for the different strains were prepared using the contour-clamped
homogenous electric field (CHEF) yeast genomic DNA plugs kit (Bio-
Rad, Richmond, CA), following the manufacturer’s recommendations,
and used for CHEF electrophoresis. CHEF electrophoresis was performed
as previously described (54).

Vacuole staining. MG and CEN.PK113-7D vacuoles were stained with
a red fluorescent dye FM4-64 (excitation/emission, ~515/640 nm) (Thermo Fisher Scientific)
following the manufacturer’s recommendations. Yeast cells and vacuoles were visualized
with an Imager-Z1 microscope equipped with an AxioCam MR camera, an EC Plan-NeoFluar
100×/1.3 oil Ph3 M27 objective, and the filter set BP 535/25, FT 580, and
LP 590 (Carl-Zeiss, Oberkochen, Germany).

Enzyme activities. Samples equivalent to 62.5 mg of dry weight bio-
mass taken at the mid-exponential phase (ca. 10 h after inoculation) of
aerobic batch cultures were used to obtain cell extracts as previously
described (58). Measurement of the activity of glycolytic enzymes was car-
rried out as previously described (59), except for phosphofructokinase,
the activity of which was determined as described in reference 60. Enzyme
activities are expressed as micromoles of substrate per minute per milli-
gram of protein or units per milligram of protein. Protein concentrations
in cell extracts were determined according to reference 61 with bovine
serum albumin as a standard.

Intracellular metabolite determination. Samples (1.2 ml each) were
taken from aerobic batch cultures using a rapid sampling setup (62)
and placed directly into vials containing 5 ml 80% methanol precooled
at –40°C. The samples were washed with precooled 80% methanol,
and the extraction was performed with boiling ethanol as previously
described (63). 13C-labeled cell extract was used as an internal stan-
dard for metabolite quantification (64). Intracellular glucose (Gluc),
glucose-6-phosphate (Gluc-6-P), fructose-6-phosphate (Fruc-6-P),
dihydroxyacetone phosphate (DHAP), glyceraldehyde-3-phosphate
(GAP), 3-phosphoglycerate (3PG), 2-phosphoglycerate (2PG),
and pyruvate (Pyr) were measured by gas chromatography-mass spec-
trometry (GC-MS) according to reference 65. Intracellular fructose-
1,6-bisphosphate (Fruc-1,6-bP) and phosphoenolpyruvate (PEP)
were measured by liquid chromatography-mass spectrometry (LC-
MS) according to reference 66. Acetaldehyde determination was
performed as previously described (67).
RNA-seq. Samples for transcriptome sequencing (RNA-seq) were obtained from aerobic batch cultures at the mid-exponential phase of growth on glucose (ca. 10 h after inoculation). Sampling, rapid quenching in liquid nitrogen, and RNA extraction were performed as previously described (68). Sequencing was performed with the Illumina HiSeq 2500 and carried out by BaseClear (Leiden, The Netherlands). A data set of 51-bp single reads of at least 1 Gb was generated. The genome sequence of CEN.PK113-7D (42) was used for all analyses. The data were aligned to the reference using the Burrows-Wheeler alignment tool BWA (69, 70). Gene expression levels were estimated using fragments per kilobase per million (FPKM) values by the Cufflinks software (71). To identify differential gene expression between strains CEN.PK113-7D and IMX372, RNA-seq data comparison was performed and statistically assessed using Cuffdiff (71).

Whole-genome sequencing. High-quality genomic DNA of the strains CEN.PK102-12A and IMX372 (MG) was obtained using the Qiagen 100/G kit (Qiagen, Hilden, Germany) following the manufacturer's recommendations. Libraries of 300-bp inserts were constructed and paired end sequenced (100-bp reads) using an illumina HiSeq 2500 sequencer (Baseclear BV, Leiden, The Netherlands). A minimum data quantity of 950 Mb was generated, representing a minimum 80-fold coverage. The sequence reads were mapped onto CEN.PK113-7D genome (42) using the Burrows-Wheeler Alignment tool BWA and further processed using SAMtools (69, 70). Single-nucleotide variations were extracted from the mapping using SAMtools' varFilter. Default settings were used, except that the maximum read depth was set to 400 (−D 400). To minimize false-positive mutation calls, custom Perl scripts were used for further mutation filtering as follows: (i) mutation calls containing ambiguous bases in mapping consensus were filtered out, (ii) only the single-nucleotide variations with a quality of at least 20 were kept (with variant quality defined as the Phred-scaled probability that the mutation call is incorrect [72, 73]), (iii) mutations with a depth of coverage below 10× were discarded, and (iv) the mutations found in CEN.PK102-12A were subtracted from the list sequence. Eventually, the single-nucleotide variations were physically positioned and functionally annotated according to the CEN.PK113-7D sequence annotation.

Microarray data accession number. The RNA-seq data generated in this study have been submitted to the Genome Expression Omnibus database and assigned accession no. GSE63884. The sequence data generated in this study are searchable at NCBI-Entrez (http://www.ncbi.nlm.nih.gov/) under Bioproject PRJNA269221.

RESULTS

From 27 to 14 glycolytic genes: minimal yeast glycolysis. To enable construction of a “minimal glycolysis” (MG) yeast strain, the first goal was to identify the major paralogs that should be retained in the final strain design. This assessment was based on information from the literature on phenotypes of relevant deletion mutants and on nonglycolytic roles (“moonlighting functions”) (74) of glycolytic isoenzymes. Furthermore, a compendium of S. cerevisiae transcriptome data obtained under a wide range of controlled cultivation conditions (75) was used to compare expression profiles of glycolytic paralogs. Major paralogs were elected based on the following, nonexclusive criteria (Fig. 1; see Fig. S1 in the supplemental material): (i) the highest transcript level over a range of growth conditions, (ii) the most extensive loss of enzyme activity in cell extracts upon deletion, (iii) moonlighting functions with a strong impact on specific growth rate or robustness (HXK2, ENO1, and ENO2) (76, 77), and (iv) the strongest decrease in specific growth rate upon deletion.

Based on these criteria, 11 of the 27 glycolytic genes in S. cerevisiae were assessed to be major paralogs: HXK2, PGi, FBA1, TPI1, TDH3, GPM1, ENO1, ENO2, PGi1, PYK1 (CDC19), PDC1, and ADH1, while 13 minor paralogs (i.e., HXK1, GLK1, TDH1, TDH2, GPM2, GPM3, ENO1, PYK2, PDC5, PDC6, ADH2, ADH4, and ADH5) were selected for gene deletion. In S. cerevisiae, the PFK1 and PFK2 paralogs have evolved into subunits of a hetero-octameric phosphofructokinase (78). Since deletion of either gene substantially decreases fitness (79–81) (see Fig. S1 in the supplemental material), both were retained in the minimal glycolysis design. ADH3 encodes a mitochondrial alcohol dehydrogenase involved in an anaerobic redox shuttle across the mitochondrial inner membrane (82). To prevent reduced growth rates under anaerobic conditions, ADH3 was therefore also retained in the minimal glycolysis design.

Sequential deletion of the 13 minor paralogs in a haploid S. cerevisiae strain belonging to the CEN.PK family (41–43) yielded the prototrophic strain IMX372, which we will refer to as the “MG” (minimal glycolysis) strain. Resequencing of the genome of the MG strain confirmed the correct deletion of all 13 minor glycolytic genes. Although transformation of S. cerevisiae can be mutagenic (83, 84), only 20 single-nucleotide differences were detected in the MG strain relative to its ancestor strain, CEN.PK102-12A. Eleven of these differences occurred within open reading frames (ORFs), of which 10 resulted in amino acid changes (see Table S1 in the supplemental material). None of these mutations affected glycolytic genes or genes related to central carbon metabolism. Whole-genome sequencing and karyotyping indicated the duplication of two short sections of chromosome III (16.5 kbp, from YCR019W to YCR027C) and chromosome V (19.1 kbp, from YER093C-A to YER104W) (see Fig. S2 and Table S1 in the supplemental material). These regions do not carry genes involved in central carbon metabolism, and no interchromosomal rearrangements were observed.

Elimination of all minor glycolytic paralogs has minimal impacts on growth kinetics, intracellular metabolite concentrations, and gene expression. To explore the physiological impact of the simultaneous deletion of all 13 minor glycolytic paralogs, we quantitatively compared specific growth rates, product formation, and gene expression in the MG strain and in a congenic, prototrophic reference strain with a full complement of glycolytic genes. In these studies, a synthetic, chemically defined medium was used, since complex media do not enable cells to express their full genetic potential (33) and complicate quantitative physiological analysis.

In aerobic, glucose-grown bioreactor batch cultures, specific rates of growth, substrate consumption, and product formation were not significantly affected by deletion of all 13 minor glycolytic paralogs (Fig. 2A). The only exception to this observation concerned acetate production, which was slightly higher in the MG strain. Also after the diauxic shift, where the glycolytic flux changes direction as the aerobic yeast cultures consumed the ethanol and acetate produced during the initial growth phase on glucose, the biomass formation and substrate consumption rates of the two strains were virtually identical. In anaerobic yeast cultures, the absence of oxidative phosphorylation makes glycolysis the only pathway for energy conservation. Also under these more demanding conditions, the specific growth rate and metabolic fluxes of the MG strain did not differ significantly from those of the congenic reference strain (Fig. 2B).

To further compare growth of the MG strain in aerobic bioreactor batch cultures, its glycolytic enzyme activities in cell extracts, intracellular metabolite concentrations, and transcriptome were...
compared with those of a congenic reference strain. Remarkably, no significant differences were observed in the transcript levels of any of the major glycolytic paralogs or in vitro glycolytic enzyme activities (Fig. 3). With the exception of slightly higher intracellular concentrations of acetaldehyde, concentrations of glycolytic intermediates in the MG strain did not significantly differ from those in the reference strain. To explore potential impacts of the deletion of 13 minor glycolytic paralogs outside glycolysis, genome-wide transcript levels of the MG strain and the reference strain were compared. As few as 17 genes showed a significantly different transcript level (see Table S2 in the supplemental material), 12 of which were located on the duplicated regions on chromosomes III and V. An in-depth, comprehensive analysis of the MG strain in glucose-grown bioreactor batch cultures therefore failed to identify substantial impacts of the minor glycolytic paralogs on fluxes, intracellular metabolite concentrations, or gene expression.

**Minimal impact of minor glycolytic paralogs under a wide range of conditions.** If during evolution of *S. cerevisiae*, its glycolytic paralogs have evolved different roles through sub- or neo-functionalization, their deletion may only cause an observable phenotype under specific growth conditions. The growth rates of the MG and reference strains were therefore compared under a wide range of selected growth conditions. During fast growth in shake flasks on complex (yeast extract-peptone-glucose [YPD]) medium (Fig. 4) and during growth at low temperature (12°C) (Fig. 4), the glycolytic pathway operates at rates closer to its maximum capacity than during growth at 30°C on synthetic medium (60, 85). However, no difference in specific growth rates between the MG and reference strains was observed under these conditions. The MG strain also showed the same growth rate as the reference strain at high temperature (37°C) (Fig. 4).

Absence of the minor glycolytic paralogs did not affect growth of the MG strain on ethanol (Fig. 2A and 4A). These results supported the conclusion from the aerobic bioreactor batch cultures that efficient gluconeogenesis does not require any of the minor paralogs (as has previously been proposed for Adh2 [86]). While minor glycolytic paralogs apparently do not contribute to maximum specific growth rates on individual carbon sources, they might be involved in substrate transitions. For instance, Pyk2, the glucose-repressed and fructose-1,6-bisphosphate-insensitive paralog of Pyk1 (21), has been proposed to prevent futile cycling resulting from simultaneous operation of glycolysis and gluconeogenesis during transitions between fermentable and nonfermentable carbon sources (87, 88). However, no significant impact of the deletion of minor glycolytic paralogs was observed during transitions from glucose to ethanol and back to glucose (Fig. 4B).

To investigate possible phenotypes of the MG strain under a wider range of environmental conditions, growth on solid medium was investigated under 24 experimental conditions. Some of
these were chosen without a specific focus on individual glycolytic paralogs. For example, growth was analyzed at high osmotic pressure (Fig. 5a to c), on different carbon sources (Fig. 5p to w), and at high concentrations of lithium and sodium ions, to which strains from the CEN.PK lineage are hypersensitive (89) (Fig. 5b to l). Only growth with glycerol as a carbon source indicated a slightly reduced growth rate of the MG strain. However, this phenotype was not observed during growth on liquid medium with glycerol (see Fig. S3 in the supplemental material) and may therefore have been caused by a contaminating substrate in the agar used for the plate experiments.

Additional growth conditions tested on solid media were chosen based on information from the literature (74, 90) on sub- or neofunctionalization of specific deleted glycolytic paralogs. In addition to their enolase activity, Enol1 and Enol2 have both been implicated in vacuolar assembly, but in the absence of Enol1, Enol2 is able to support both functions (76). In the vacuole and the ATPase associated with it, disruption causes a deleterious phenotype in S. cerevisiae when cultivated under alkaline conditions (91). As expected, the MG strain grew normally at pH 7.5, thereby indicating the absence of major vacuolar mal-function (Fig. 5f; see Fig. S4 in the supplemental material). Tdh3 and Tdh2, the two minor isoenzymes of glyceraldehyde-3-phosphate dehydrogenase, have been proposed to protect S. cerevisiae against oxidative stress because of their different levels of thio-lation (92). However, exposure to oxidative stress by growth on hydrogen peroxide-containing plates (Fig. 5g) showed that the absence of Tdh2 did not affect the oxidative stress resistance of the MG strain. Sub- and neofunctionalization have resulted in a different glucose-dependent regulation of GLK1/HXX1 and PYK2 and in different regulatory properties of the proteins that they encode from those encoded by HXX2 and PYK1, respectively (21, 93). However, growth of the MG strain at low, intermediate, and high glucose concentrations was not impaired relative to that of the reference strain (Fig. 5p to r). Finally, PDC6, which encodes a pyruvate decarboxylase isoenzyme with a low cysteine and methionine content, is specifically induced under sulfur limitation (94). Cultivation in the presence of cadmium increases abundance of Pdc6 (94–96); since cells require a high level of glutathione production for detoxification, sulfur amino acids are then used for this process and are less available for protein synthesis. Despite the absence of PDC6, the MG strain grew as well as the reference strain in the presence of cadmium.

**DISCUSSION**

Absence of phenotypes upon deletion of 13 glycolytic genes. Deletion of 13 of the 27 glycolytic paralogs in S. cerevisiae yielded a “minimal glycolysis” (MG) yeast strain whose most spectacular characteristic was the absence of any pronounced phenotype under a wide range of laboratory growth conditions. One of the hypotheses that has been proposed to explain retention of functionally redundant paralogs during evolution is a contribution to gene dosage and, thereby, to the capacity of the pathway or process in which they operate (97, 98). The high glycolytic rates in anaerobic cultures of the MG strain (18 mmol glucose per g of dry biomass per h) did not significantly differ from those of wild-type strains (99) (Fig. 2). This result argues against gene dosage effects as a means for fixing minor glycolytic paralogs in the yeast genome. Instead, our observations indicate that duplication of glycolytic genes is not a prerequisite for achieving the high glycolytic fluxes and fermentative capacities that are characteristic of S. cerevisiae and essential for many of its industrial applications (100, 101). It might be argued that the gene dosage hypothesis does apply to phosphofructokinase, as deletion of either PFK1 or PFK2 substantially reduces enzyme activity and fitness (79–81) (see Fig. S1 in the supplemental material). However, the presence of Pfk1 and Pfk2 in a hetero-octameric complex can also be seen as a case of neofunctionalization, in which two paralogs have been fixed by an acquired mutual dependency. The near-wild-type growth kinetics of the MG strain in aerobic and anaerobic conditions is difficult to reconcile with the hypothesis that duplication of glycolytic genes during the WGD event played a major role in increasing its glycolytic capacity or in causing the phenomenon of aerobic fermentation (the Crabtree effect) (20).

Our inability to identify a phenotype after deletion of all minor paralogs of glycolytic genes in S. cerevisiae does not imply that such a phenotype does not exist. The range of conditions tested represents an infinitesimal fraction of the environmental conditions to which S. cerevisiae may have been exposed to in its evolutionary history. The absence of a clear phenotype under standard laboratory conditions makes the MG strain an even more interesting platform for future high-throughput studies to investigate its phenotype under more conditions. Other characteristic features of S. cerevisiae like tolerance to high ethanol concentrations or growth at high gravity are interesting conditions to test, as well as the robustness of the MG strain under nonstandard conditions, such as sporulation, starvation, severe calorie restriction, and dynamic nutrient supply regimens. However, guessing the environmental factors that have conferred an evolutionary advantage to strains carrying gene duplications presents a formidable challenge, more particularly because the real niche for S. cerevisiae is still a matter of debate, to the extent that S. cerevisiae may be considered a generalist that does not necessarily favor a specific type of environment (102). For further functional analysis studies of the minor glycolytic paralogs, the set of intermediate strains used for construction of the MG strains (see Table S3 in the supplemental material) can be used to rapidly identify which minor paralog or paralogs contribute to any newly identified phenotypes. Moreover, the sensitivity of fitness analyses can be improved—for example, by competitive cultivation of wild-type and MG strains in mixed cultures.

In S. cerevisiae, overrepresentation of paralogous gene sets is not unique for glycolysis (23 of 27 glycolytic genes are paralogs, corresponding to 85%) but also occurs for metabolic genes in general. Of all metabolic genes, 44% are paralogs (98), while this percentage is only 5% for the entire yeast genome (5, 17). Currently, a large research effort is under way to enable design and assembly of entirely synthetic yeast genomes (36). For the design of compact, functional synthetic genomes, it will be highly interesting to investigate whether our results on yeast glycolysis can be extrapolated to minor paralogs of genes encoding key enzymes in other central metabolic pathways. At present, predictions of the outcome of such experiments can only be highly speculative. The clear impact of “minor paralogs” of transaldolase and transketolase on pentose fermentation kinetics by engineered S. cerevisiae (103) is a clear example that, at least under some conditions, minor paralogs can make a clear contribution to metabolic flux.

The deletion of all minor glycolytic paralogs had a surpris-
FIG 3 In vitro enzyme activities and intracellular metabolite concentrations in aerobic batch cultures in bioreactors of the minimal glycolysis (MG) strain and a congenic reference strain. Thirteen minor glycolytic paralogs were deleted in the MG strain. (A) Average values of four independent culture replicates of the in vitro activities for the glycolytic enzymes in the MG strain (white bars) and the prototrophic reference strain CEN.PK113-7D (black bars). HXK, hexokinase; glucokinase; PGI, phosphoglucone isomerase; PFK, phosphofructokinase; FbPA, fructose-bisphosphate aldolase; TPI, triose-phosphate isomerase; GAPdh, glyceraldehyde-3-phosphate dehydrogenase; PGK, phosphoglycerate kinase; GPM, phosphoglycerate mutase; ENO, enolase; PYK, pyruvate kinase; PDC, pyruvate decarboxylase; ADH, alcohol dehydrogenase. The denoted error bars represent standard deviations. (B) Intracellular glycolytic metabolite profiles of the MG strain (open circles) and of CEN.PK113-7D (close circles) from aerobic batch cultures. Average values from two independent culture replicates are shown, and the average deviations of the mean are indicated by error bars. The vertical orange dotted line indicates the time at which glucose was depleted.

FIG 4 Growth of the minimal glycolysis (MG) strain and a congenic reference strain under different growth conditions in shake flasks. (A) Specific growth rates of S. cerevisiae strains MG (white bars) and CEN.PK113-7D (reference strain, black bars). MG, synthetic medium with glucose; SMEOH, synthetic medium with ethanol as the carbon source; YPD, complex medium with glucose as the carbon source. The labels “37°C” and “12°C” indicate growth on SMG at 37°C and 12°C, respectively. Average specific growth rates (per hour) are denoted above each bar. (B) Growth, measured as change in the culture’s optical density at 660 nm (OD660), of the MG strain (open circles) and reference strain (closed circles) during carbon source switches. Strains were successively grown in glucose, ethanol, and glucose. All data represent the average and average deviation of the mean from two independent culture replicates.

Minimal glycolysis yeast: a versatile new platform for quantitative research. As the MG strain was constructed before the advent of CRISPR-Cas9 technology (106), its construction involved 16 consecutive rounds of transformation and 11 marker recycling steps. A previous study (107) in which all 20 hexose transporter (HXT) genes in S. cerevisiae were deleted by repeated rounds of transformation and marker recycling resulted in massive genomic rearrangements, which were recently explained from repeated use of the LoxP system for marker recovery (120). Repeated use of the LoxP system enables recombination across LoxP sites that are left in the genome after marker removal (108). The “HXT-null” strain is mainly used as an excellent platform for functional analysis of individual transporter strains (107, 109–117), in which role its chromosomal rearrangements are hardly relevant. While the MG strain similarly offers an interesting platform for functional analysis of (heterologous) glycolytic genes by one-step gene replacement, we designed and constructed it with the specific aim of building a robust platform for quantitative studies in systems biology. By using URA3 and amdSYM (53) as counterselectable, recyclable marker genes, major chromosomal rearrangements could be avoided. Furthermore, interference of auxotrophies in the interpretation of quantitative growth studies (57) was avoided by making the MG strain prototrophic.

The availability of a well-defined yeast platform with a minimal complement of glycolytic enzymes should provide clear advantages for quantitative modeling of the kinetics and regulation of glycolysis, as it eliminates the intrinsic uncertainties caused by the simultaneous, context-dependent expression of different isoenzymes. In view of the important role of glucose transport in the kinetics of yeast glycolysis, we are currently constructing MG variants with a single hexose transporter. In addition to providing a relevant test bed for mathematical modeling of glycolysis, the MG strain provides an interesting, simplified starting point for laboratory studies on yeast glycolysis. Previous studies have indicated that long-term laboratory evolution can have a major impact on glycolytic genes and their expression (59, 118). Comparison of evolution of the MG strain with that of strains that carry a full complement of glycolytic genes provides an interesting starting point for studies on the impact of gene duplication on evolutionary flexibility of a ubiquitous central metabolic pathway.

with the backup theory for fixation of duplicated genes (17), which postulates that (minor) paralogs provide a buffer against deleterious mutations. Although upregulation of minor paralogs and even repair of a null mutation in a major glycolytic paralog by recombination with a minor paralog have been demonstrated (pdc1/PDC5) (104), the general evolutionary significance of the backup theory is still a matter of debate (98, 105). The selective advantages responsible for the overrepresentation of paralogous genes among the structural genes of yeast glycolysis therefore remain an intriguing conundrum.

ingularly small effect on the yeast transcriptome in aerobic, glucose-grown batch cultures. This result is in marked contrast with several studies in which deletion of individual major glycolytic paralogs led to transcriptional upregulation of its minor paralogs. An example of such a “compensatory” upregulation of minor paralogs is the upregulation of PDC5 upon deletion of PDC1 (104). Such an asymmetric cross-regulation is consistent
ACKNOWLEDGMENTS

This work was supported by the Technology Foundation STW (Vidi grant 10776).

We thank Walter van Gulik for expert support for metabolome analysis and Lizanne Bosman for experimental contribution to the construction of the MG strain.

REFERENCES


6. Dean EJ, Davis JC, Davis RW, Petrov DA. 2008. Pervasive and persis-


Downloaded from ec.asm.org on August 18, 2015 by BIBLIOTHEEK TU DELFT