

Form of paper: Regular paper

Field: Molecular Biology (Molecular Evolution / Bioinformatics)

Title: Molecular Mechanisms Involved in Robustness of Yeast Central Metabolism against Null Mutations

By line: Natalia Maltsev,<sup>1\*</sup> Elizabeth M. Glass,<sup>1</sup> Galina Ovchinnikova,<sup>1</sup> Zhenglong Gu<sup>2</sup>

From line: <sup>1</sup>Argonne National Laboratory, Mathematics and Computer Science Division, Argonne, IL 60439, USA; <sup>2</sup>University of Chicago, Department of Ecology and Evolution, Chicago, IL 60637, USA

Running title: "Tactics to Ensure Robustness in Yeast Central Metabolism"

Corresponding author:

Natalia Maltsev

Argonne National Laboratory

9700 S. Cass Avenue

Building 221/ B-211

Argonne, IL 60439

630-252-5195 (office)

630-252-5986 (fax)

[maltsev@mcs.anl.gov](mailto:maltsev@mcs.anl.gov)

Abbreviations:

EC, enzyme commission number; EMP, Embden Meyerhoff Parnas pathway, PPS, pentose phosphate shunt; TCA, tricarboxylic acid cycle

"The submitted manuscript has been created in part by the University of Chicago as Operator of Argonne National Laboratory ("Argonne") under Contract No. W-31-109-ENG-38 with the U.S. Department of Energy. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government."

## Summary

Adaptive strategies employed by the yeast *Saccharomyces cerevisiae* provide robustness and adaptability of its central metabolism. Since central metabolism in yeast has been well studied at the enzymatic and genetic levels, it represents an excellent system for evaluating the relative roles of duplicate genes and alternative metabolic pathways as possible mechanisms for the stability of central metabolism against null mutations. Yeast appears to employ a variety of mechanisms to ensure functional robustness of its central metabolism. Uninterrupted flow of energy and precursor metabolites through the pathways of central metabolism via glycolysis (EMP), pentose phosphate shunt (PPS), and the tricarboxylic acid (TCA) cycle are ensured by a variety of adaptive mechanisms. One of the most significant mechanisms appears to be gene duplication events that have produced a number of isozymes functioning under variable environmental and physiological conditions. Alternative pathways represent another important mechanism for increasing the robustness of the system. The robustness of the pathways of central metabolism is apparently higher than that of the other parts of metabolism, because of its exceptional importance to the organism's vitality. The proportion of duplicated viable genes also is substantially larger in central metabolism than that in a pool of other metabolic genes.

Keywords: enzymes, gene duplication, isozymes, phenotype, subunits

Recent advances in experimental and computational biology allow one to address complex questions regarding the evolution of biological adaptation. To survive and thrive under variable environmental conditions, most organisms have developed complex mechanisms to maintain their cellular homeostasis. During the long course of evolution, a number of adaptive strategies, such as efficient sensory systems, functional redundancy of cellular components, and optimization of metabolic and regulatory processes, have been developed to ensure an organism's robustness against perturbations. However, such robustness has a cost for a biological system in increased complexity and energy requirements (1). Therefore, a biological system can afford to increase robustness only against a limited number of possible environmental challenges, since the cost of such an adaptation would be too high. Featherstone and Broadie (2002) have suggested that a topology of adaptive changes is favored where disruption of one component may affect just a few "peripheral" components of the biological system, but not any central component; otherwise the system may suffer severe damage or death. Jeong et al. (2001) have demonstrated the importance of such robustness for protein networks.

In this paper, we study adaptive strategies employed by the yeast *S. cerevisiae* to provide robustness and adaptability of its central metabolism. Since central metabolism in yeast has been well studied at the enzymatic and genetic levels, it represents an excellent system for evaluating the relative roles of duplicate genes and alternative metabolic pathways as possible mechanisms for the stability of central metabolism against null mutations. Moreover, since the yeast genome has been completely sequenced and annotated, genes can be classified into single copy genes and duplicate genes (4). Gene duplication and emergence of alternative pathways have been considered the two principal mechanisms for genetic robustness against null mutations, but their relative importance has recently become a controversial issue (5). In this manuscript we examine the fitness effect of single-gene deletion for every gene in the pathways of central metabolism

(e.g., glycolysis, pentose-phosphate shunt, and the TCA cycle) in yeast and study adaptive strategies employed by evolution to ensure robustness of metabolic performance.

*S. cerevisiae* is one of the best-studied model organisms. This eukaryotic, facultative anaerobe shows substantial metabolic and physiological flexibility and is capable of utilizing a wide range of mono- and oligosaccharides as well as pyruvate, lactate, acetate ethanol, and glycerol. However, the preferred mode of operation in *S. cerevisiae* is alcohol fermentation of glucose (and other carbohydrates) via glycolysis. In the presence of glucose, expression of genes encoding enzymes involved in utilization of disaccharides, other than glucose, or nonfermentable carbon sources (e.g., ethanol, acetate, and pyruvate) is a subject of catabolite (glucose) repression under both anaerobic and aerobic conditions. When the concentration of glucose is low, these genes are “derepressed,” and cells are capable of growth on substrates other than glucose. If oxygen is available, cells consume the ethanol produced earlier and completely oxidize it to carbon dioxide and water via the TCA and glyoxylate cycles. Growth on nonfermentable carbohydrates requires the gluconeogenic production of glucose-6-phosphate. This compound is used for biosynthesis of the storage carbohydrate trehalose. Trehalose is required for “shuttling” around the pentose phosphate shunt that provides precursors for nucleotide biosynthesis and for the provision of sulfur-containing organic compounds (6).

The present study concentrates on the analysis of central metabolism that represents an energy and metabolic “axis” of yeast livelihood. Various low-molecular weight carbohydrates are channeled into central degradation pathways via a network of pathways of peripheral metabolism (e.g., catabolism of amino acids, fats, sugars, aromatic compounds). The pathways of central metabolism are the major source of ATP and other high-energy compounds, as well as universal precursor metabolites for cellular biosynthesis. These compounds are later extracted at various steps of central metabolism for the needs of biosynthesis (Figure 1). Central metabolism

traditionally includes glycolysis, which converts glucose to pyruvate; the TCA cycle, which oxidizes acetyl-CoA to CO<sub>2</sub>; and the pentose-phosphate shunt, which oxidizes glucose 6-phosphate to CO<sub>2</sub>. Linker reactions that join these pathways and enable the reversal of flow of metabolites through the network, and several reactions that allow bypassing portions of the pathway (e.g., the glyoxylate shunt) are also considered in this paper.

Central metabolic pathways operate under very tight genetic and metabolic control that ensures rapid and precise response to changes in metabolite concentrations. Analysis of the evolutionary strategies employed for stabilization of central metabolism in yeast is especially interesting because of its importance for the organism's survival.

## 2. Materials and Methods

### 2.1 Identification of duplicate and singleton genes

As described by Gu et al. (2003), the whole sets of *S. cerevisiae* protein sequences were downloaded from SGD (<http://genome-www.stanford.edu/Saccharomyces/>). An all-against-all FASTA search was conducted on each protein dataset independently. A singleton was defined as a protein that did not hit any other proteins in the FASTA search with  $E = 0.1$ . Duplicate genes were identified as described by Gu et al. (2003) ( $E < 10^{-10}$ ). We also used less stringent criteria to detect duplicate genes and obtained essentially the same results.

### 2.2 Metabolic pathways

Genes in *S. cerevisiae* metabolic pathways are defined according to the WIT (<http://wit.mcs.anl.gov/WIT2/>) database (8). The *S. cerevisiae* ORFs are categorized into central metabolic (CM) and noncentral metabolic (Non-CM) genes. Central metabolic genes are those that are involved in glycolysis, the pentose-phosphate shunt, and the TCA cycle (including the

glyoxylate shunt). The numbers of metabolic steps within CM and Non-CM with singletons and duplicates are counted. A metabolic step represents a biochemical reaction catalyzed by an enzyme. Whether a step has singleton or duplicate enzymes, we count it as one for singleton or one for duplicate.

### 2.3 Proportion of unduplicated genes and the number of duplications per gene

For each category (i.e., CM or All ORFs) or CM pathway (EMP, PPS, or TCA) under study, the number of unique types of genes is defined as the number of singletons plus the number of duplicated gene types in that category.

## 3. Results

### 3.1 General statistics

A gene is said to be essential if its deletion from the genome or loss of function as a result of mutation causes death of the organism; otherwise, it is called nonessential. Table I summarizes the distribution of essential genes in the pathways of central metabolism and the role of gene duplication events for stabilization of its functionality. With only one exception, all essential genes in central metabolism are involved in glycolysis; and only one essential gene is found in the pentose phosphate shunt. No essential genes are found in the TCA cycle.

Gene duplication appears to be the most important factor that contributes to the robustness against gene deletion (null mutation). None of the duplicate genes in central metabolism is lethal in the single-gene deletion experiment. All essential genes in central metabolism are singletons, that is, genes that do not have any recognizable homologous genes (paralogs) in the genome. Deletion of a nonessential gene, while it may cause a number of phenotypic changes (e.g., inability to utilize certain substrates), has not resulted in cell death under the conditions of the

experiment (9). As Table I shows, 61.1% of nonessential enzymes have one or more duplicate copies (paralogs) in the genome, while only 27.8% of them exist as a single copy.

The proportion of singletons that are essential genes in central metabolism is substantially lower than that of the general pool of metabolic genes (4.21% and 34.04%, respectively). The proportion of duplicated nonessential genes in central metabolism is considerably higher than that in the other pathways. The proportion of singleton nonessential genes in central metabolism is substantially lower than that in a general pool of metabolic genes (10.53% versus 29.46%, respectively).

These observations suggest the importance of duplicate genes in the genetic robustness of central metabolism in yeast.

### 3.2 Glycolysis and ethanol fermentation

Glycolysis via the EMP pathway and alcohol fermentation is of major importance to the physiology of *S. cerevisiae*. This significance is reflected by the fact that glycolytic enzymes represent 30–65% of its soluble protein in the yeast cell, depending on the physiological stage (10). Yeast glycolytic pathway functions under very tight metabolic control as suggested by metabolic control analysis that all enzymes in a biochemical pathway contribute to some extent to metabolic flux through that pathway (11). Such regulation of a metabolic process on the level of individual enzymes allows cells to achieve fast and precise response to changes in the concentrations of glycolytic metabolites.

Analysis of the distribution of essential (lethal) genes in *S. cerevisiae* glycolytic pathway has revealed a number of interesting observations. Three out of the four essential genes found in central metabolism participate in glycolysis, or more precisely, in its upper part (Figure 1). The

enzymes encoded by the three essential genes are phosphoglucose isomerase (*pgi*, EC 5.3.1.9); fructose-bisphosphate aldolase (*fba*, EC 4.1.2.13), and triosephosphate isomerase (*tpi*, EC 5.3.1.1). None of these enzymes have isozymes or close homologs in the yeast genome. The suggestion of the existence of *pgi* isozyme(s) appears to be an artifact from processing of the *pgi1* gene product (12). *Fba* appears to function under a stringent metabolic control: *fba* mutants fail to grow without supplying a carbon source on each side of the *Fba1p* reaction (13). *Fba* does not show recognizable homology to any other gene product in the yeast genome. The *Tpi* product of the third glycolytic essential gene in *S. cerevisiae*, triosephosphate isomerase, makes up about 2% of the soluble cellular protein (14); however it is also encoded by a singleton gene.

All glycolytic enzymes encoded by the three essential genes catalyze reactions leading to production of essential biosynthetic precursors: EC 5.3.1.9, glucose-6-phosphate; EC 4.1.2.13 and EC 5.3.1.1, glyceraldehydes-3-phosphate. These metabolites are common to glycolysis and the pentose-phosphate shunt. Apparently, because of the importance of these compounds for the overall homeostasis of the cell, notable restrictions were imposed by evolution against duplication and modification of the genes leading to their production. However, the other essential metabolites produced in glycolysis are the products of enzymes that do form isozymes in yeast.

Gene duplication appears to be a significant adaptation mechanism in yeast glycolysis. All nonessential enzymes in glycolysis appear to a subject of gene duplication. For example, *pgm1* and *pgm2* genes, encoding two phosphoglucomutases, (EC 5.4.2.2) in yeast, are closely related to one another and apparently arose by gene duplication. Boles et al. (1994) suggest that yeast possesses a complex network of mutases that can partially substitute for each other. Double mutants of *pgm1* and *pgm2*, could not grow on galactose (13).

Another important glycolytic enzyme in yeast, phosphofructokinase, is an octameric enzyme (EC 2.7.1.11) composed of four  $\alpha$ -subunits and four  $\beta$ -subunits, encoded by the genes

*pfk1* and *pfk2*, respectively. As shown by Heinisch et al. (1989) the amino acid sequences of these two genes showed 20% homology between the N- and the C-terminal halves of each subunit, and more than 50% homology between the two subunits. Apparently, two gene duplication events occurred in the evolution of the yeast *pfk* genes. The first duplication event took place soon after the separation of prokaryotic and eukaryotic lineage and the second in *Saccharomyces* later in the phylogeny (17). The three glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) genes also likely emerged as a consequence of two nontandem duplication events. All three gene products show substantial sequence homology to each other (18).

Pyruvate kinase, (EC 2.7.1.40) the last enzyme in the glycolytic pathway, catalyzes the irreversible conversion of phosphoenolpyruvate (PEP) into pyruvate. For a long time *S. cerevisiae* was thought to have one pyruvate kinase encoded by the *pyk1* gene (19). Mutants defective in the *pyk1* gene fail to grow on fermentable carbon sources and are even inhibited by them (20). However, they grow normally on ethanol or other gluconeogenic carbon sources. After completion of the yeast genome, unexpectedly a gene encoding the second pyruvate kinase (*pyk2*) was discovered. Pyk1 and Pyk2 protein sequences are 70% identical. Pyk2 was shown to operate under conditions of very low glycolytic flux. Overexpression of the *pyk2* gene restored growth on glucose in a *pyk1* mutant strain and could completely substitute for the *pyk1*-encoded enzymatic activity; *pyk2* gene expression is subject to glucose repression (21).

### 3.3 The pentose phosphate shunt

The pentose phosphate shunt is primarily an anabolic pathway (under aerobic conditions) (22) that utilize the six carbons of glucose to generate five carbon sugars and reducing equivalents. It is closely related to glycolysis and shares two common intermediates with it (fructose-6-phosphate and glyceraldehydes 3-phosphate). It has a number of functions important for an organism's livelihood, such as production of NADPH for reductive biosynthesis, provision of

erythrose 4-phosphate for biosynthesis of aromatic amino acids and vitamins, and ribose 5-phosphate for biosynthesis of nucleotides and some amino acids.

Table II and Figure 1 summarize the distribution and phenotypic properties of homologs for genes involved in PPS. A sole essential gene, *rki*, found in the pentose phosphate shunt encodes ribose 5-phosphate epimerase (EC 5.3.1.6), which does not have any paralogs or nonhomologous isozymes in yeast. A product of this enzyme, ribose-5-phosphate, is essential for producing nucleotides, tryptophan, and histidine.

Deleterious mutations of genes encoding nonessential enzymes in yeast PPS result in disruption of some its collateral metabolic roots and therefore inability to grow on various other substrates. For example, deletion of the gene for glucose 6-phosphate isomerase (*pgi*) (EC 5.3.1.9) increases the requirement for organic sulphur source without causing noticeable effect on carbohydrate metabolism (6). Mutation in the *gnd1* gene, which encodes 6-phosphogluconate dehydrogenase (EC 1.1.1.44), compromises ability of yeast to grow on glucose; however, mutations in *gnd1* are suppressed by *zwf1* mutations (23). Deletion of either the *tkl1* gene (24) or the *rpe1* gene (D-ribulose-5-phosphate 3-epimerase, EC 5.1.3.1) reduced the ability of *S. cerevisiae* to grow on xylulose; however, mutation of *tkl2* was not associated with any phenotypic changes (13). Some of the enzymes participating in PPS belong to multigene families (e.g., Sol1,2,4).

As follows from the results of analysis of gene duplication events in *S. cerevisiae* glycolysis and PPS described above, although this organism uses gene duplication and alternative pathways to ensure robustness of the central metabolic pathways, four enzymes are not protected in this manner. Glucose-6-phosphate isomerase (EC 5.3.1.9), fructose-bisphosphate

aldolase (EC 4.1.2.13), triosephosphate isomerase (EC 5.3.1.1) of glycolysis and ribose phosphate isomerase (EC 5.3.1.6) of PPS are directly involved in production of common precursor metabolites for biosynthesis. Our analysis shows that these enzymes are singletons in *S. cerevisiae* and, when mutated or knocked out, prove lethal to this organism. An interesting scientific question is whether or not the restrictions on duplication of these genes in nature are ubiquitous. To answer this question, we have compiled information regarding gene duplications, and occurrence of non-homologous isozymes for these enzymes from a diverse set of taxa (eubacteria, archaea, deep-branching, and higher eukaryotes). Our observations show that similar restrictions are observed in other yeastlike *S. pombe*, the gram-negative bacteria, *E. coli* and *C. jejuni*, and the archaea, *A. fulgidus* and *P. furiosus*. Below we examine in more detail duplication events for enzymes outlined in Table IV.

The glycolytic enzyme, glucose-6-phosphate isomerase *gpi* (EC 5.3.1.9) catalyzes the interconversion of glucose-6 phosphate and fructose-6 phosphate. As in *S. cerevisiae*, no duplicates of the *gpi* gene were found in the gram-positive eubacteria, *B. subtilis*, *L. monocytogenes*, gram-negative *E. coli* K12 and *C. jejuni*, the archaea *P. furiosus*, and the low eukaryotes *S. pombe* and *P. falciparum*. The genome of the thermophilic archaea *A. fulgidus*, contains two genes encoding multifunctional enzymes performing glucose-6-phosphate isomerase activity. These enzymes are homologous and may have arisen from gene duplication events. The multifunctional features of these enzymes may help compartmentalize the reaction(s) through metabolic channeling in order to protect chemically labile intermediates in thermophilic environments (25). The higher eukaryote *Mus musculus* has a duplicate gene for *gpi*; but taking into consideration that mutation of the *gpi* gene yields a lethal phenotype, we can conclude that the duplicate

gene is not functionally competent. The plant *A. thaliana* has two homologous isozymes, which appear to function in different compartments (26); however, it is unclear whether these two can compensate for function.

Fructose 1,6-bisphosphate aldolase (EC 4.1.2.13) is a key enzyme in cellular metabolism, catalyzing the reversible aldol condensation/cleavage reaction between glyceraldehyde 3-phosphate and dihydroxyacetone phosphate to yield fructose 1,6-bisphosphate (FBP). Two mechanistically and structurally distinct forms of FBP are known. Both class I and II of the FBP aldolases are found both in prokaryotes and eukaryotes, a fact that suggests an ancient evolutionary origin of both versions of this enzyme. Most organisms contain only one class of FBP aldolase, although there are few exceptions. *E. coli* is one example: it expresses both classes of FBP aldolases, demonstrating a catalytic redundancy largely eliminated by most organisms. *M. musculus* appears to have homologous tissue-specific isozymes of Class I aldolases; *A. thaliana* also has duplicate copies of Class I gene; however, their subcellular location is not documented, and no experimental data is available about whether these copies can compensate for loss of function. *L. monocytogenes* has duplicate copies of Class II aldolase that have redundant functions, which apparently protects its central metabolic core, and its ability to survive in a wide range of environmental conditions. Duplication of *fba* gene in *Bacillus subtilis* does not provide protection for mutations, although the genes for the two copies of Class II aldolases are similar, the mutation of *fbaB* is lethal for this organism (27).

Triosphosphate isomerase (EC 5.3.1.1) performs a reversible interconversion of dihydroxyacetone phosphate and glyceraldehydes 3-phosphate. This enzyme is represented by a singleton gene in most of organisms under consideration (see Table IV), with the exception of *A. thaliana*, *P.falciparum*, and *L. monocytogenes*, where the viability of this reaction may be ensured by duplicate isozymes. However, no experimental data is available that will allow one to evaluate this hypothesis. The vulnerability of the singleton gene, *tpi* in *M. musculus*, is unclear, as mutation causes only a reduction in output.

Ribose-5-phosphate isomerase (EC 5.3.1.6) catalyzes the interconversion of ribose-5-phosphate and ribulose-5-phosphate. This activity is essential to the pentose phosphate shunt (nonoxidative branch) and to the plant Calvin cycle. Two nonhomologous enzymes are known to catalyze this reaction. Both versions are sometimes found in the same organism. For example, in the case of *E. coli*, analysis of strains containing *rpiA* and *rpiB* mutants revealed that both isozymes are equally efficient in catalyzing isomerization in either direction (28). *RpiA* genes are distributed broadly in nature and have been identified in more than 100 species. *RpiB* genes are found almost exclusively in bacteria. A number of bacteria have only *rpiB*, while another group, including *E. coli* and *L. monocytogenes*, have both *rpiA* and *rpiB*.

#### 3.4 The TCA and glyoxylate cycles

The eight enzymes of the TCA cycle are encoded by 15 nuclear genes (29). Four of these enzymes are encoded by single genes, while the others are composed of subunits or have isoforms.

The TCA cycle carries out many roles in cell metabolism. The experiments on *S. cerevisiae* mutants having defects in each of 15 genes that encode 15 subunits of the eight enzymes show that the expression of >400 genes changed at least threefold in response to the TCA cycle dysfunction (30).

*S. cerevisiae* can grow under so-called repressing and nonrepressing conditions. Repressing means anaerobic growth with glucose. Under aerobic nonrepressing conditions the TCA cycle in yeast has a dual function: first, it provides the majority of cellular ATPs via aerobic oxidization of pyruvate, and second, it plays an important role in the metabolism of universal biosynthetic precursors (e.g., citrate,  $\alpha$ -ketoglutarate, oxaloacetate). Under repressing conditions the TCA cycle operates in a branched fashion to fulfill biosynthetic demands only (22). Under anaerobic, non-repressing conditions, the glyoxylate cycle is essential for growth on the two-carbon substrates (e.g. acetate, ethanol) and fulfills an anaplerotic role in the provision of precursors for biosynthesis. However, the glyoxylate cycle is inactive under repressing conditions only (22).

Figure 2 shows the metabolic steps and major intermediates involved in the TCA and glyoxylate pathways. Here we describe a couple of anaplerotic reactions that are tightly connected with the TCA cycle. The malic enzyme (EC 1.1.1.38) plays a key role for mitochondrial pyruvate metabolism (in both the aerobic and anaerobic regime) only (22). Another important reaction is provided by the pyruvate carboxylase (EC 6.4.1.1), which has two isoforms *pyc1* and *pyc2*. In the aerobic regime 75% of mitochondrial oxaloacetate comes from carboxylation of pyruvate only (22). The distribution of duplicated genes in these pathways is shown in Table III. All metabolic steps common to the TCA and glyoxylate cycles are catalyzed by enzymes that have formed isozymes or have recognizable homologs in the yeast genome. Citrate synthase (EC 2.3.3.1), which catalyses the first rate limiting step of TCA, has three isozymes: Cit1, a

mitochondrial enzyme active in TCA; Cit2, a peroxisomal enzyme participating in the glyoxylate cycle; and Cit3, a second active mitochondrial copy (31). The nuclear gene *aco1*, which encodes aconitase (4.2.1.3), a mitochondrial enzyme participating in both pathways, has a close homolog (YJL200c) (32). (However, the authors were not able to express this gene under the chosen experimental conditions.) *S. cerevisiae* possesses two isozymes for both NAD and NADP-dependent isocitrate dehydrogenase. The two subunits of the NAD-dependent mitochondrial version of the enzyme (EC 1.1.1.41) are encoded by *idh1* and *idh2*. The mitochondrial and cytosolic versions of NADP-dependent isocitrate dehydrogenase (EC 1.1.1.42) are encoded by *idp1* and *idp2*. It was suggested that these two NADP-dependent isocitrate dehydrogenases do not play any significant role in the TCA cycle, but rather contribute to formation of  $\alpha$ -ketoglutarate necessary for the biosynthesis of glutamate even under the conditions when the full TCA cycle is not functioning (33).

A number of enzymes involved exclusively in the glyoxylate cycle also have homologs: the gene *icl1* encoding isocitrate lyase (EC 4.1.3.1) has a homolog *icl2* for which the function is not established. The *icl2* gene appears to be inactive, both in vitro and in vivo. No other function has been elucidated for this gene (34). Presumably, two highly homologous isozymes of malate synthase exist in yeast: one of them, Mls1, is involved in the glyoxylate cycle, while the other, Dal7, participates in allantoin degradation. Expression of the genes that encode these two proteins is specific for their physiological roles in carbon and nitrogen metabolism. Based on observations with null mutations in these genes, Hartig et al. (1992) suggested that *S. cerevisiae* contains at least one and perhaps two additional malate synthase (EC 2.3.3.9) genes; their suggestion is not supported by our analysis, however, unless the proposed isozymes emerged as a result of nonhomologous replacement. Of the three isozymes of malate dehydrogenase (EC 1.1.1.37) found in yeast, the cytoplasmic version encoded by *mdh2*, is believed to participate in the glyoxylate cycle, while the products of *mdh1* and *mdh3* are mostly involved in energy production

The importance of the TCA and glyoxylate cycles dictates a need for strict coordination of functionality of the pathways and individual enzymes. Several factors should be taken into consideration. First, during the diauxic shift the expression of many nuclear-encoded mitochondrial proteins, such as components of the electron transport chain and enzymes of the tricarboxylic acid (TCA) cycle, become derepressed. A process of derepression is regulated by *hap* (CCAAT-binding factor) that represents a mechanism for global control of expression of key components of respiratory metabolism. Moreover, the compartmentalization of the yeast cell, which leads to a dissection of TCA and glyoxylate metabolism into subnetworks localized in different compartments—in cytosol, *Idp2*; or in organelles, mitochondria (*Cit1*, *Cit3*, *Aco*, *Idp1*, *Idh1*) or peroxisomes (*Cit2*, *Mdh3*), requires strict control and coordination of gene expression. The pathway of three-way communication between mitochondria, the nucleus, and peroxisomes involving the *rtg* (retrograde regulation) genes achieves such control. As was shown by Liu and Butow (1999), when the cell's respiratory function is reduced or eliminated, the expression of four TCA cycle genes (*cit1*, *aco1*, *idh1*, and *idh2*) switches from the *hap* control to control by three genes, *rtg1*, *rtg2*, and *rtg3*. The expression of four additional TCA cycle genes downstream of *idh1* and *idh2* is independent of the *rtg* genes. The same authors previously reported that the *rtg* genes control the retrograde pathway responsible for a change in the expression of a subset of nuclear genes (e.g., the glyoxylate cycle *cit2* gene) in response to changes in the functional state of mitochondria. Apparently, in cells with compromised mitochondrial function, the *rtg* genes take control of the expression of genes leading to the synthesis of  $\alpha$ -ketoglutarate. It ensures sufficient supply of glutamate for biosynthetic processes and increased flux in the glyoxylate cycle, via elevated *CIT2* expression, that provides a supply of metabolites entering the TCA cycle sufficient to support anabolic pathways.

#### 4. Discussion

Yeast appears to employ a variety of mechanisms to ensure functional robustness of its central metabolism. Uninterrupted flow of energy and precursor metabolites through the pathways of glycolysis via EMP, PPS, and the TCA cycle are ensured by variety of adaptive mechanisms. One of the most significant mechanisms appears to be gene duplication events that have produced a number of isozymes functioning in organism under variable environmental and physiological conditions. The results of gene disruption experiments strongly suggest that the presence of isozymes provides various degrees of compensation for missing function (e.g., *Eno1* and *Eno2*, *Tdh1*, *Tdh2*, and *Tdh3*).

One can also suggest the importance of multisubunit composition of enzymes in stabilization of function. For example, yeast phosphofructokinase is an octameric enzyme composed of four  $\alpha$ -subunits and four  $\beta$ -subunits, encoded by the genes *pfk1* and *pfk2*, respectively. Subunits of this enzyme apparently have emerged via two gene duplication events. Possibly, such redundant subunit composition allows increasing functional stability of this important glycolytic enzyme. However, further detailed investigation is needed to make this hypothesis more conclusive.

Analysis of the survival evolutionary strategies used for increasing the robustness of the yeast TCA cycle in order to ensure supply of vitally important precursors for cellular biosynthesis has revealed the following:

1. The fact that all of the enzymes participating in reactions leading to the production of essential metabolites have isozymes confirms the importance of gene duplication in increasing the robustness of the system. TCA enzymes that are not directly involved in

- their production, namely those encoded by, *kgd1*, *kdg2*, *lpd*, *lsc1*, and *fum*, do not have isozymes or close homologs in the yeast genome.
2. Alternative pathways represent another important mechanism for increasing of the robustness of the system. A good example of this mechanism is the use of the glyoxylate cycle and a pathway regulated by *rtg* genes to ensure sufficient production of biosynthetic precursors. Such an adaptation mechanism is widely used by evolution for securing vital function. Alternative pathways are used by the yeast under the conditions when the full TCA cycle is not functional:
    - a. The glyoxylate cycle, which leads to the production of essential metabolite oxaloacetate
    - b. A pathway regulated by *rtg* genes that involves the gene products from *cit1*, *acol*, *idh1*, *idh2*, and *idp2*, which leads to the production of  $\alpha$ -ketoglutarate under anaerobic conditions
    - c. A pool of essential metabolites, produced by the functioning TCA cycle and supplemented by various catabolic processes in the cell (e.g., degradation of amino acids, benzoate, propanoate catabolism)

The robustness of the pathways of central metabolism is apparently greater than that of the other parts of metabolism, because of its exceptional importance to organism's vitality. The proportion of duplicated viable enzymes in central metabolism is substantially larger in central metabolism than that in a pool of other metabolic enzymes (61.1% vs 24.7%, respectively). However, it is not clear why glycolytic enzymes encoded by essential genes that catalyze reactions leading to production of essential biosynthetic precursors, namely, EC 5.3.1.9 (glucose-6-phosphate), EC 4.1.2.13 and EC 5.3.1.1 (glyceraldehydes-3-phosphate), do not form isozymes and contain no recognizable homologs in the yeast genome. The same trend was observed with some modifications in another yeast, *S. pombe*;; archaea, *P. furiousus* and *A. fulgidus*; some eubacteria, *B.*

*subtilis* and *C. jejuni*; and deep branching eukaryotes. However, *E. coli* and *L. monocytogenes* as well as high eukaryotes had more prevalent modifications in regard to duplications of these enzymes. A number of these duplications were tissue-specific isozymes, nonhomologous enzyme replacements and poorly annotated homologous genes, which may indeed not be functionally competent, so it is difficult to interpret the impact of such duplications on robustness. One possible explanation for the occurrence of singleton essential genes catalyzing production of common metabolites may be the importance of these compounds for the overall homeostasis of the cell. Apparently, restrictions were imposed by evolution against duplication and modification of these genes.

Understanding of the evolutionary mechanisms involved in securing the functionality of metabolic processes in the cell is essential for progress in medical and biotechnological applications.

## Acknowledgments

This work was supported in part by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-ENG-38.

## REFERENCES

1. Carlson, J.M. and Doyle, J. (1999) Highly optimized tolerance: a mechanism for power laws in designed systems. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*. 60, 1412–1427.
2. Featherstone, D.E. and Broadie, K. (2002) Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network. *Bioessays*. 24, 267–274.
3. Jeong, H., Mason, S.P., Barabasi, A.L., Oltvai, Z.N. (2001) Lethality and centrality in protein networks. *Nature*. 411, 41–42.
4. Marland, E., Prachumwat, A., Maltsev, N., Gu, Z., Li, W-H. (2004) Higher Gene Duplicabilities for Metabolic Protein than for Non-Metabolic Proteins in Yeast and E.coli. (in press, *Mol Evol*).
5. Kitami, T., Nadeau, J.H. (2002) Biochemical networking contributes more to genetic buffering in human and mouse metabolic pathways than does gene duplication. *Nat Genet*. 32, 191–194.
6. Thomas, D., Cherest, H., Surdin-Kerjan, Y. (1991) Identification of the structural gene for glucose-6-phosphate dehydrogenase in yeast. Inactivation leads to a nutritional requirement for organic sulfur. *EMBO J*. 10, 547–553.
7. Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., & Li, W.-H. (2003) Role of Duplicate Genes in Genetic Robustness against Null Mutations. *Nature*. 421, 63–66.
8. Overbeek, R., Larsen, N., Pusch, G.D., D'Souza, M., Selkov, E. Jr., Kyrpides, N., Fonstein, M., Maltsev, N., Selkov, E. (2000) WIT: integrated system for high-throughput genome sequence analysis and metabolic reconstruction. *Nucleic Acids Res*. 28, 123–125.
9. Mewes, H. W., Albermann, K. Heumann, K., Liebl, S., Pfeiffer, F. (1997) MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acids Res*. 25, 28.
10. Banuelos, M., Fraenkel, D.G. (1982) *Saccharomyces carlsbergensis* fdp mutant and futile

- cycling of fructose 6-phosphate. *Mol Cell Biol.* 2, 921–929.
11. Kacser, H., Burns, J.A. (1973) The control of flux. *Symp Soc Exp Biol.* 27, 65–104.
  12. Green, J.B., Wright, A.P., Cheung, W.Y., Lancashire, W.E., Hartley, B.S. (1988) The structure and regulation of phosphoglucose isomerase in *Saccharomyces cerevisiae*. *Mol Gen Genet.* 215, 100–106.
  13. Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A.P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.D., Flaherty, P., Foury, F., Garfinkel, D.J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J.H., Hempel, S., Herman, Z., Jaramillo, D.F., Kelly, D.E., Kelly, S.L., Kotter, P., LaBonte, D., Lam, D.C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S.L., Revuelta, J.L., Roberts, C.J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D.D., Sookhai-Mahadeo, S., Storms, R.K., Strathern, J.N., Valle, G., Voet, M., Volckaert, G., Wang, C.Y., Ward, T.R., Wilhelmy, J., Winzeler, E.A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J.D., Snyder, M., Philippsen, P., Davis, R.W., Johnston, M. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature.* 2002 418, 387-391.
  14. Scott, E.W., Baker, H.V. (1993) Concerted action of the transcriptional activators REB1, RAP1, and GCR1 in the high-level expression of the glycolytic gene TPI. *Mol Cell Biol.* 13, 543–550.
  15. Boles, E., Schulte, F., Miosga T., Freidel, K., Schluter, E., Zimmermann, F.K., Hollenberg, C.P., Heinisch, J.J. (1997) Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in *Saccharomyces cerevisiae* that is catalytically insensitive to fructose-1,6-bisphosphate. *J Bacteriol.* 179, 2987–2993.
  16. Heinisch, J., Ritzel, R.G., von Borstel, R.C, Aguilera, A., Rodicio, R., Zimmermann, F.K.

- (1989) The phosphofructokinase genes of yeast evolved from two duplication events. *Gene*. 78, 309–321.
17. Koonin, E.V., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Krylov, D.M., Makarova, K.S., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Rogozin, I.B., Smirnov, S., Sorokin, A.V., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A. (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5, R7.
18. Holland, J.P., Labieniec, L., Swimmer, C., Holland, M.J. (1983) Homologous nucleotide sequences at the 5' termini of messenger RNAs synthesized from the yeast enolase and glyceraldehyde-3-phosphate dehydrogenase gene families. The primary structure of a third yeast glyceraldehyde-3-phosphate dehydrogenase gene. *J Biol Chem.* 258, 5291–5299.
19. Burke, R.L., Tekamp-Olson, P., Najarian, R. (1983) The isolation, characterization, and sequence of the pyruvate kinase gene of *Saccharomyces cerevisiae*. *J Biol Chem.* 258, 2193–2201.
20. Pearce, A.K., Crimmins, K., Groussac, E., Hewlins, M.J., Dickinson, J.R., Francois, J., Booth, I.R., Brown, A.J. (2001) Pyruvate kinase (Pyk1) levels influence both the rate and direction of carbon flux in yeast under fermentative conditions. *Microbiology.* 147, 391–401.
21. Boles, E., Liebetrau, W., Hofmann, M., Zimmermann, F.K. (1994) A family of hexosephosphate mutases in *Saccharomyces cerevisiae*. *Eur J Biochem.* 220, 83–96.
22. Maaheimo H., Flaux J., Cakar Z.P., Bailey J.E., Souer U., Szyperski T. (2001) Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional C(13) labeling of common amino acids. *Eur. J. Biochem.* 268, 2464–2479.
23. Lobo, Z., Maitra, P.K. (1982) Pentose phosphate pathway mutants of yeast. *Mol Gen Genet.* 185, 367–368.
24. Schaaff-Gerstenschlager, I., Zimmermann, F.K. (1993) Pentose-phosphate pathway in

- Saccharomyces cerevisiae*: analysis of deletion mutants for transketolase, transaldolase, and glucose 6-phosphate dehydrogenase. *Curr Genet.* 24, 373–376.
25. Massant, J., Verstreken, P., Durbecq, V., Kholti, A., Legrain, C., Beckmans, S., Cornelis, P., Glansdorff, N. (2002) Metabolic channeling of carbamoyl phosphate, a thermolabile intermediate: evidence for physical interaction between carbamate kinase-like carbamoyl-phosphate synthetase and ornithine carbamoyltransferase from the hyperthermophile *Pyrococcus furiosus*. *J Biol Chem.* 277, 18517-22.
26. Yu, T.S., Lue, W.L., Wang, S.M., Chen, J. (2000) Mutation of *Arabidopsis* plastid phosphoglucose isomerase affects leaf starch synthesis and floral initiation. *Plant Physiol.* 123, 319-326.
27. [http://bacillus.genome.jp/bsorf\\_mutant\\_list/Page14.htm](http://bacillus.genome.jp/bsorf_mutant_list/Page14.htm)
28. Sorensen, K.I., Hove-Jensen, B. (1996) Ribose catabolism of *Escherichia coli*: characterization of the *rpiB* gene encoding ribose phosphate isomerase B and of the *rpiR* gene, which is involved in regulation of *rpiB* expression. *J Bacteriol.* 178, 1003-1011.
29. McAlister-Henn L., Small W.C. (1997) Molecular genetics of yeast TCA cycle isozymes. *Prog. Nucleic Acids Res.. Mol. Biol.* 57, 317–339.
30. McCammon M.T., Epstein C.B., Przybyla-Zawislak B., McAlister-Henn L., Butow R. (2003) Global transcription analysis of Krebs Tricarboxylic Acid cycle mutants reveals an alternating pattern of gene expression and effects on hypoxic and oxidative genes. *Molecular Biology of the Cell.* 14, 958–972.
31. Jia Y.K., Becam A.M., Herbert C.J. (1997) The *CIT3* gene of *Saccharomyces cerevisiae* encodes a second mitochondrial isoform of citrate synthase. *Mol Microbiol.* 24, 53–59.
32. Purnelle, B., Coster, F., Goffeau, A. (1994) The sequence of a 36 kb segment on the left arm of yeast chromosome X identifies 24 open reading frames including *NUC1*, *PRP21* (*SPP91*), *CDC6*, *CRY2*, the gene for *S24*, a homologue to the aconitase gene *ACO1* and two homologues to chromosome III genes. *Yeast.* 10, 1235–1249.

33. Haselbeck, R.J., McAlister-Henn, L. (1993) Function and expression of yeast mitochondrial NAD- and NADP-specific isocitrate dehydrogenases. *J Biol Chem.* 268, 12116–12122.
34. Heinisch, J.J., Valdes, E., Alvarez, J., Rodicio, R.(1996) Molecular genetics of ICL2, encoding a non-functional isocitrate lyase in *Saccharomyces cerevisiae*. *Yeast.* 12, 1285–1295.
35. Hartig, A., Simon, M.M., Schuster, T., Daugherty, J.R., Yoo, H.S., Cooper, T.G. (1992) Differentially regulated malate synthase genes participate in carbon and nitrogen metabolism of *S. cerevisiae*. *Nucleic Acids Res.* 20, 5677–5686.
36. Liu, Z., Butow, R.A. (1999) A transcriptional switch in the expression of yeast tricarboxylic acid cycle genes in response to a reduction or loss of respiratory function. *Mol Cell Biol.* 19, 6720–6728.

Table I. Phenotypic distribution of central metabolic enzymes in *S. cerevisiae*. Enzymes are categorized as singleton or duplicate and as viable or lethal phenotype. Distribution of duplicate genes in central metabolic pathways is more than double in all yeast ORFs. Duplicate lethal genes are defined as genes that are homologous, but are unable to compensate for the others central metabolic function, and therefore produce a lethal phenotype.

<u>Central Metabolic Pathway</u>	<u>Singleton Viable</u>	<u>Duplicate Viable</u>	<u>Singleton Lethal</u>	<u>Duplicate Lethal</u>
Glycolysis		5.4.2.2		
			5.3.1.9	
		2.7.1.11		
			4.1.2.13	
			5.3.1.1	
		6.4.1.1		
		1.2.1.12		
	2.7.2.3			
		5.4.2.1		
		4.2.1.11		
		2.7.1.40		
		6.4.1.1		
<i>Totals</i>	9.1%	63.6%	27.3%	0.0%
Pentose Phosphate Shunt	1.1.1.49			
		3.1.1.31		
		1.1.1.44		
	5.1.3.1			
			5.3.1.6	
		2.2.1.1		
		2.2.1.2		
<i>Totals</i>	28.6%	57.1%	14.3%	0.0%
Intermediate Steps	1.2.4.1			
	2.3.1.12			
<i>Totals</i>	100.0%	0.0%	0.0%	0.0%
TCA Cycle & Glyoxylate Cycle		2.3.3.1		
		4.2.1.3		
		1.1.1.41		
		1.1.1.42		
	1.2.4.2			
	2.3.1.61			
		1.8.1.4		
	6.2.1.4			
		1.3.5.1		
	4.2.1.2			
		1.1.1.37		
		4.1.3.1		
		2.3.3.9		
	1.1.1.38			
<i>Totals</i>	35.7%	64.3%	0.0%	0.0%
Fermentation		4.1.1.1		
		1.1.1.1		
<i>Totals</i>	0.0%	100.0%	0.0%	0.0%
<i>Central Metabolic Pathway Totals</i>	27.8%	61.1%	11.1%	0.0%
<i>All ORFs Totals</i>	57.5%	24.7%	15.2%	2.6%

Gene Category	Singleton Viable	Duplicate Viable	Singleton Lethal	Duplicate Lethal
All Genes	1819	1197	2102	1058
Central Metabolic	10	81	4	0
	% Singleton Viable	% Singleton Lethal	% Duplicate Viable	% Duplicate Lethal
All Genes	29.46%	34.04%	19.38%	17.13%
Central Metabolic	10.53%	4.21%	85.26%	0.00%
	Duplicate Viable Gene Types	Duplicate Lethal Gene Types	% of Unduplicated Genes	# of Duplicates per Gene
All Genes	282	431	84.63%	1.33
Central Metabolic	24	0	36.84%	2.50

Table II. Distribution and phenotype properties of homologs for genes involved in the pentose phosphate shunt.

<u>Pentose Phosphate Genes</u>	<u>E.C. Number</u>	<u>Phenotype</u>	<u>Duplicates</u>	<u>Adaptive Strategies</u>
HXK1   HKA	2.7.1.1	Viable	yes	Isozymes
HXK2   HKB   HEX1	2.7.1.1	Viable	yes	Isozymes
GLK1	2.7.1.2	Viable	yes	Glycolytic
ZWF1   MET19	1.1.1.49	Viable	no	Alternative Pathway?
Sol1	3.1.1.31	Viable	yes	Homologs
Sol2	3.1.1.31	Viable	yes	Homologs
Sol3	3.1.1.31	Viable	yes	Homologs
Sol4	3.1.1.31	Viable	yes	Homologs
Gnd1	1.1.1.44	Viable	yes	Isozymes
Gnd2	1.1.1.44	Viable	yes	Isozymes
RPE1   POS18	5.1.3.1	Viable	no	Alternative Pathway
RKT1	5.3.1.6	Lethal	no	Lethal - Ribose 5-Phosphate
TKL1	2.2.1.1	Viable	yes	Isozymes
TKL2	2.2.1.1	Viable	yes	Isozymes
TAL1   YGR043C	2.2.1.2	Viable	yes	2 enzymes appear to be encoded by same gene

Table III. The distribution of duplicated genes in the TCA and glyoxylate cycles and their subcellular location. All metabolic steps common to the TCA and glyoxylate cycles are catalyzed by enzymes that have formed isozymes or have recognizable homologs in the yeast genome.

<u>E.C. Number</u>	<u>TCA</u>	<u>Glyoxylate</u>	<u>Unknown</u>	<u>Phenotype</u>	<u>Duplicates</u>	<u>Localization</u>
2.3.3.1	Cit1			Viable	yes	mitochondria
		Cit2		Viable	yes	peroxisome
			Cit3	Viable	yes	mitochondria
4.2.1.3	Aco1	Aco1		Viable	yes	mitochondria
4.1.3.1		Icl1		Viable	yes	peroxisome
2.3.3.9		Mls1		Viable	yes	glyoxosome
		Dal7		Viable	yes	glyoxosome
1.1.1.41	Idh1			Viable	yes	mitochondria
	Idh2			Viable	yes	mitochondria
1.1.1.42		Idp1		Viable	yes	mitochondria
		Idp2		Viable	yes	cytoplasm
1.1.1.37	Mdh1			Viable	yes	mitochondria
		Mdh2		Viable	yes	cytoplasm
	Mdh3			Viable	yes	peroxisome
1.2.4.2 / 2.3.1.61 / 1.8.1.4	Kgd1			Viable	no	mitochondria
	Kgd2			Viable	no	mitochondria
	Lpd1			Viable	yes	mitochondria
6.2.1.4	Lsc1			Viable	no	mitochondria
	Lsc2			Viable	no	mitochondria
1.3.5.1 / 1.3.99.1	Sdh1			Viable	yes	mitochondria
	Sdh2			Viable	yes	mitochondria
	Sdh3			Viable	yes	mitochondria
	Sdh4			Viable	yes	mitochondria
4.2.1.2	Fum1			Viable	no	mitochondria / cytoplasm
1.1.1.38	Mae1			Viable	no	mitochondria

Table IV. The distribution of duplicates for EC 5.3.1.9, 4.1.2.13, 5.3.1.1 of glycolysis and EC 5.3.1.6 of the pentose phosphate shunt. GN=gene name; ORF=open reading frame; EF=enzyme function; SL=subcellular location; COM=comments. All known essential genes are in bold type, duplicate genes are underlined, while all known viable genes are italicized.

		EC 5.3.1.9					
Tax. Group	Organism	GN	ORF	EF	SL	COM	
High Eukaryote	<i>M. musculus</i>	<i>gpi</i>		isozyme	cytosol	Frag- ment	
	<i>M. musculus</i>	<u><i>gpi1</i></u>		isozyme	cytosol		
	<i>A. thaliana</i>	<u>F22K18.180</u>	<u>At4g24620</u>	isozyme	cytosol		
	<i>A. thaliana</i>	<u>MJB21.12</u>	<u>At4g42740</u>	isozyme	cytosol		
Low Eukaryote	<i>S. pombe</i>	<i>pgi1</i>	SPBC1604.05		cytosol		
	<i>P. falciparum</i>		PF14_0341		cytosol		
Gram +	<i>B. subtilis</i>	<i>pgi</i>	BSU31350		cytosol		
	<i>L. monocytogenes</i>	<i>pgi</i>	lmo2367		cytosol		
Gram -	<i>E. coli</i> K12	<i>pgi</i>	b4025		cytosol		
	<i>C. jejuni</i>	<i>pgi</i>	Cj1535c		cytosol		
Archaea	<i>A. fulgidis</i>	<i>manC</i>	AF1097	multifunctional	cytosol		
	<i>A. fulgidis</i>		AF0035	bifunctional	cytosol		
	<i>P. furiosus</i>	<i>pgiA</i>	PF0196		cytosol		
<b>EC 4.1.2.13</b>							
High Eukaryote	<i>M. musculus</i>	<i>aldoa</i>	-	isozyme	muscul	class I	
	<i>M. musculus</i>	<i>aldob</i>	-	isozyme	liver	class I	
	<i>M. musculus</i>	<i>aldoc</i>	-	isozyme	brain	class I	
	<i>A. thaliana</i>	<u>F10A8.2</u>	<u>At2g01140</u>	isozyme	cytosol	class I	
	<i>A. thaliana</i>	<u>F3K23.9</u>	<u>At2g21330</u>	isozyme		class I	
	<i>A. thaliana</i>	<u>F1011.9</u>	<u>At2g36460</u>	isozyme		class I	
	<i>A. thaliana</i>	<u>F8J2.100</u>	<u>At3g52930</u>	isozyme		class I	
	<i>A. thaliana</i>	<u>M3E9.50</u>	<u>At4g26520</u>	isozyme		class I	
	<i>A. thaliana</i>	<u>M3E9.40</u>	<u>At4g26530</u>	isozyme		class I	
	<i>A. thaliana</i>	<u>F19H22.70</u>	<u>At4g38970</u>	isozyme		class I	
	<i>A. thaliana</i>	<u>F17c15.110</u>	<u>At5g03690</u>	isozyme		class I	
Low Eukaryote	<i>S. pombe</i>	<i>fba1</i>	SPBC19C2.07			cytosol	class II
	<i>P. falciparum</i>		PF14_0425			cytosol	class I
Gram +	<i>B. subtilis</i>	<i>fbaA</i>	BSU39670	isozyme		cytosol	class II
	<i>B. subtilis</i>	<b><i>fbaB</i></b>	<b>BSU37120</b>	isozyme	cytosol	class II	
	<i>L. monocytogenes</i>	-	<u>lmo0359</u>	isozyme	cytosol	class II	
	<i>L. monocytogenes</i>	-	<u>lmo2133</u>	isozyme	cytosol	class II	
Gram -	<i>L. monocytogenes</i>	<i>fbaA</i>	lmo2556	isozyme	cytosol	class II	
	<i>E. coli</i> K12	<i>fbaA</i>	b2925	isozyme	cytosol	class II	
	<i>E. coli</i> K12	<i>fbaB</i>	b2097	isozyme	cytosol	class I	

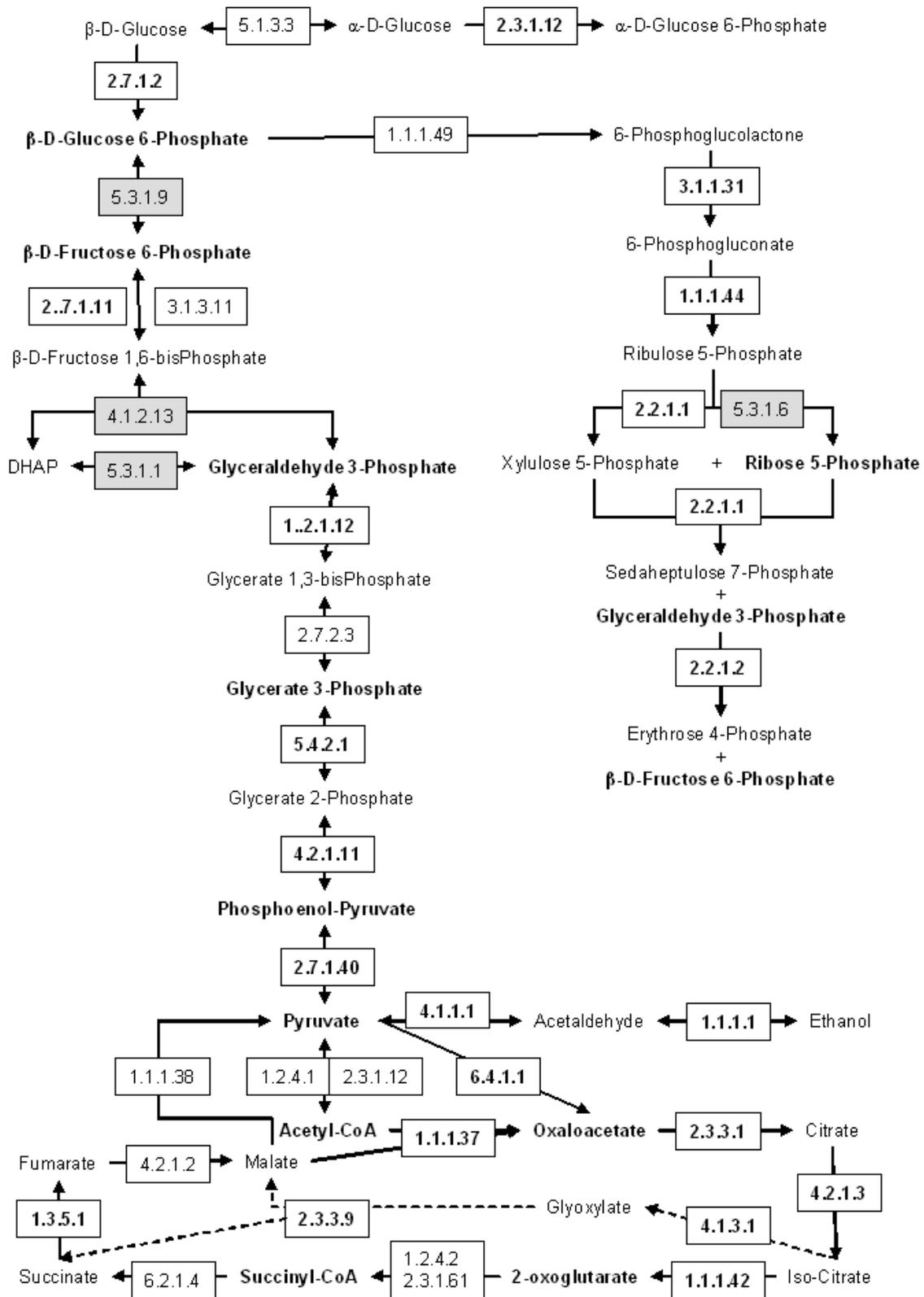
	<b>C. jejuni</b>	<i>fba</i>	Cj0597		cytosol	class II
<b>Archaea</b>	<b>A. fulgidis</b>		AF0230		cytosol	class I
	<b>P. furiosus</b>	<i>fba</i>	PF1956		cytosol	class I
<b>EC 5.3.1.1</b>						
<b>High Eukaryote</b>	<b>M. musculus</b>	<i>tpi1</i>				
	<b>A. thaliana</b> <b>A. thaliana</b>	F26H11.7 T22E16.100	At2g21170 At3g55440	isozyme isozyme	chloroplast cytosol	
<b>Low Eukaryote</b>	<b>S. pombe</b>	<i>tpi</i>	SPCC24B10.21		cytosol	
	<b>P. falciparum</b> <b>P. falciparum</b>		PF14_0378 PFC0831w	isozyme isozyme	cytosol cytosol	
<b>Gram +</b>	<b>B. subtilis</b>	<i>tpiA</i>	<i>BSU33920</i>	isozyme	cytosol	
	<b>L. monocytogenes</b> <b>L. monocytogenes</b>	<i>tpiA1</i> <i>tpiA2</i>	<i>lmo2457</i> <i>lmo0346</i>	isozyme isozyme	cytosol cytosol	
	<b>E. coli K12</b>	<i>tpiA</i>	b3919		cytosol	
<b>Gram -</b>	<b>C. jejuni</b>	<i>tpiA</i>	Cj1401c		cytosol	
	<b>A. fulgidis</b>	<i>tpiA</i>	AF1304		cytosol	
<b>Archaea</b>	<b>P. furiosus</b>	<i>tpiA</i>	PF1920		cytosol	
<b>EC 5.3.1.6</b>						
<b>High Eukaryote</b>	<b>M. musculus</b>	<i>rpiA</i>				class A
	<b>A. thaliana</b>	F23N20.9	At1g71100	isozyme	cytosol	class A
	<b>A. thaliana</b> <b>A. thaliana</b>	F10A8.17 T9J14.26	At2g01290 At3g04790	isozyme isozyme	chloroplast? chloroplast	expres sed class A
<b>Low Eukaryote</b>	<b>S. pombe</b>		SPAC144.12		cytosol	class A
	<b>P. falciparum</b>		PFE0730c		cytosol	class A
<b>Gram +</b>	<b>B. subtilis</b>	<i>ywIF</i>	<i>BSU36920</i>		cytosol	class B
	<b>L. monocytogenes</b> <b>L. monocytogenes</b> <b>L. monocytogenes</b> <b>L. monocytogenes</b> <b>L. monocytogenes</b> <b>L. monocytogenes</b> <b>L. monocytogenes</b>	<i>rpiA</i>	<i>lmo0975</i> <i>lmo0345</i> <i>lmo0498</i> <i>lmo0736</i> <i>lmo2662</i> <i>lmo2674</i>	isozyme isozyme isozyme isozyme isozyme isozyme	cytosol cytosol cytosol cytosol cytosol cytosol	class A class B class B class B class B class B
	<b>E. coli K12</b>	<i>rpiA</i>	b2914	isozyme	cytosol	class A
	<b>E. coli K12</b>	<i>rpiB</i>	b4090	isozyme	cytosol	class B
	<b>C. jejuni</b>	<i>rpiB</i>	Cj0925			cytosol class B
	<b>Archaea</b>	<b>A. fulgidis</b>	<i>rpi</i>	AF0943		cytosol class A
		<b>P. furiosus</b>	<i>rpiA</i>	PF1258		cytosol class A

## Legends to Figures

Figure 1. Central metabolic pathways in *S. cerevisiae*. Distribution of duplicate genes and those that have lethal phenotype are noted by their EC numbers in bold type or grey boxes, respectively. Precursor metabolites are labeled in bold type.

Figure 2. Metabolic steps and major intermediates involved in the TCA and glyoxylate pathways. Under aerobic conditions the TCA cycle in yeast has a dual function, to provide energy via oxidation of pyruvate and to produce metabolic precursors (in bold type). Under anaerobic conditions, the glyoxylate shunt is essential for growth on two carbon substrates and fulfills an anaplerotic role in the provision of precursors for biosynthesis. Duplicate genes are noted in grey boxes; gene products located in the cytoplasm are in hexagons; gene products located in the peroxisome are in ovals; and gene products located in the mitochondria are in rectangles. RTG regulated genes are noted with four-point stars. Precursor metabolites are in bold type.

TOP, Figure 1, Maltsev N.



TOP, Figure 2, Maltsev, N.

