



Compositions and methods for modeling *Saccharomyces cerevisiae* metabolism

Palsson, Bernard; Famili, Imandokht; Fu, Pengcheng; Nielsen, Jens B.; Förster, Jochen

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Palsson, B., Famili, I., Fu, P., Nielsen, J. B., & Förster, J. (2012). Compositions and methods for modeling *Saccharomyces cerevisiae* metabolism. (Patent No. WO03036296).

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

(19)



(11)

EP 2 463 654 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
13.06.2012 Bulletin 2012/24

(51) Int Cl.:
G01N 33/48 (2006.01)

(21) Application number: 11182034.6

(22) Date of filing: 24.10.2002

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
IE IT LI LU MC NL PT SE SK TR**

(30) Priority: 02.10.2002 US 263901
26.10.2001 US 344447 P

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
02784306.9 / 1 438 580

(71) Applicant: **The Regents of The University of California
Oakland, CA 94607-5200 (US)**

(72) Inventors:

- Palsson, Bernard O.
La Jolla, CA 92037 (US)
- Famili, Imandokht
San Diego, CA 92126 (US)

- Fu, Pengcheng
Honolulu, HI 96822 (US)
- Nielsen, Jens B.
Lyngby (DK)
- Forster, Jochen
Lyngby (DK)

(74) Representative: **Hollywood, Jane Constance
Kilburn & Strode LLP
20 Red Lion Street
London WC1R 4PJ (GB)**

Remarks:

- This application was filed on 20-09-2011 as a divisional application to the application mentioned under INID code 62.
- Claims filed after the date of filing of the application/ after the date of receipt of the divisional application (Rule 68(4) EPC).

(54) Compositions and methods for modeling *saccharomyces cerevisiae* metabolism

(57) The invention provides an *in silico* model for determining a *S. cerevisiae* physiological function. The model includes a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, a constraint set for the plurality of *S. cerevisiae* reactions, and commands for determining a distribution of flux through the reactions that is predictive of a *S. cerevisiae* physiological function. A model of the invention can further include a gene database containing information characterizing the associated gene or genes. The invention further provides methods for making an *in silico* *S. cerevisiae* model and methods for determining a *S. cerevisiae* physiological function using a model of the invention. The invention provides an *in silico* model for determining a *S. cerevisiae* physiological function. The model includes a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, a constraint set for the plurality of *S. cerevisiae* reactions, and commands for determining a distribution of flux through the reactions that is predictive of a *S. cerevisiae* physiological function. A model of the invention can further include a gene database containing information characterizing the associated gene or genes. The invention further provides methods for making an *in silico*

S. cerevisiae model and methods for determining a *S. cerevisiae* physiological function using a model of the invention.

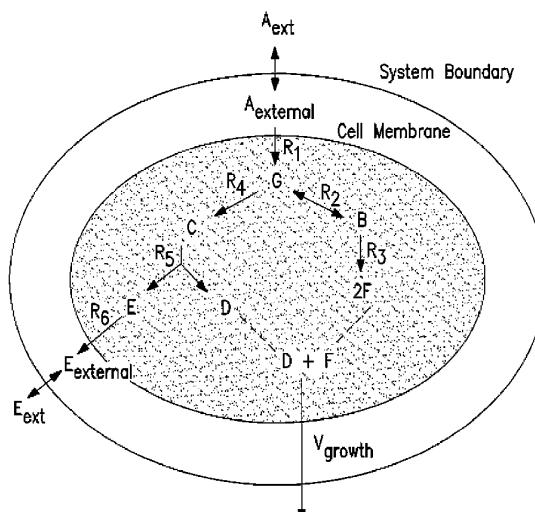


FIG. 1

Description

[0001] This invention was made with United States **Government support** under grant NIH RO1HL59234 awarded by the National Institutes of Health. The U.S. Government has certain rights in this invention.

5

BACKGROUND OF THE INVENTION

[0002] This invention relates generally to analysis of the activity of a chemical reaction network and, more specifically, to computational methods for simulating and predicting the activity of *Saccharomyces cerevisiae* (*S. cerevisiae*) reaction networks.

10

[0003] *Saccharomyces cerevisiae* is one of the best-studied microorganisms and in addition to its significant industrial importance it serves as a model organism for the study of eukaryotic cells (Winzeler et al. Science 285: 901-906 (1999)). Up to 30% of positionally cloned genes implicated in human disease have yeast homologs.

15

[0004] The first eukaryotic genome to be sequenced was that of *S. cerevisiae*, and about 6400 open reading frames (or genes) have been identified in the genome. *S. cerevisiae* was the subject of the first expression profiling experiments and a compendium of expression profiles for many different mutants and different growth conditions has been established. Furthermore, a protein-protein interaction network has been defined and used to study the interactions between a large number of yeast proteins.

20

[0005] *S. cerevisiae* is used industrially to produce fuel ethanol, technical ethanol, beer, wine, spirits and baker's yeast, and is used as a host for production of many pharmaceutical proteins (hormones and vaccines). Furthermore, *S. cerevisiae* is currently being exploited as a cell factory for many different bioproducts including insulin.

25

[0006] Genetic manipulations, as well as changes in various fermentation conditions, are being considered in an attempt to improve the yield of industrially important products made by *S. cerevisiae*. However, these approaches are currently not guided by a clear understanding of how a change in a particular parameter, or combination of parameters, is likely to affect cellular behavior, such as the growth of the organism, the production of the desired product or the production of unwanted by-products. It would be valuable to be able to predict how changes in fermentation conditions, such as an increase or decrease in the supply of oxygen or a media component, would affect cellular behavior and, therefore, fermentation performance. Likewise, before engineering the organism by addition or deletion of one or more genes, it would be useful to be able to predict how these changes would affect cellular behavior.

30

[0007] However, it is currently difficult to make these sorts of predictions for *S. cerevisiae* because of the complexity of the metabolic reaction network that is encoded by the *S. cerevisiae* genome. Even relatively minor changes in media composition can affect hundreds of components of this network such that potentially hundreds of variables are worthy of consideration in making a prediction of fermentation behavior. Similarly, due to the complexity of interactions in the network, mutation of even a single gene can have effects on multiple components of the network. Thus, there exists a need for a model that describes *S. cerevisiae* reaction networks, such as its metabolic network, which can be used to simulate many different aspects of the cellular behavior of *S. cerevisiae* under different conditions. The present invention satisfies this need, and provides related advantages as well.

35

SUMMARY OF THE INVENTION

40

[0008] The invention provides a computer readable medium or media, including: (a) a data structure relating a plurality of reactants in *S. cerevisiae* to a plurality of reactions in *S. cerevisiae*, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product, (b) a constraint set for the plurality of *S. cerevisiae* reactions, and (c) commands for determining at least one flux distribution that minimizes or maximizes an objective function when the constraint set is applied to the data representation, wherein at least one flux distribution is predictive of a physiological function of *S. cerevisiae*. In one embodiment, at least one of the cellular reactions in the data structure is annotated to indicate an associated gene and the computer readable medium or media further includes a gene database including information characterizing the associated gene. In another embodiment, at least one of the cellular reactions in the data structure is annotated with an assignment of function within a subsystem or a compartment within the cell.

45

[0009] The invention also provides a method for predicting physiological function of *S. cerevisiae*, including: (a) providing a data structure relating a plurality of *S. cerevisiae* to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product; (b) providing a constraint set for the plurality of *S. cerevisiae* reactions; (c) providing an objective function, and (d) determining at least one flux distribution that minimizes or maximizes the objective function when the constraint set is applied to the data structure, thereby predicting a *S. cerevisiae* physiological function. In one embodiment, at least one of the *S. cerevisiae* reactions in the

data structure is annotated to indicate an associated gene and the method predicts a *S. cerevisiae* physiological function related to the gene.

[0010] Also provided by the invention is a method for making a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions in a computer readable medium or media, including: (a) identifying a plurality of *S. cerevisiae* reactions and a plurality of reactants that are substrates and products of the reactions; (b) relating the plurality of reactants to the plurality of reactions in a data structure, wherein each of the reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product; (c) determining a constraint set for the plurality of *S. cerevisiae* reactions; (d) providing an objective function; (e) determining at least one flux distribution that minimizes or maximizes the objective function when the constraint set is applied to the data structure, and (f) if at least one flux distribution is not predictive of a physiological function of *S. cerevisiae*, then adding a reaction to or deleting a reaction from the data structure and repeating step (e), if at least one flux distribution is predictive of a physiological function of the eukaryotic cell, then storing the data structure in a computer readable medium or media. The invention further provides a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of reactions, wherein the data structure is produced by the method.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] Figure 1 shows a schematic representation of a hypothetical metabolic network.

[0012] Figure 2 shows the stoichiometric matrix (S) for the hypothetical metabolic network shown in Figure 1.

[0013] Figure 3 shows mass balance constraints and flux constraints (reversibility constraints) that can be placed on the hypothetical metabolic network shown in Figure 1. (∞ , infinity; Y_1 , uptake rate value)

[0014] Figure 4 shows an exemplary metabolic reaction network in *S. cerevisiae*.

[0015] Figure 5 shows a method for reconstruction of the metabolic network of *S. cerevisiae*. Based on the available information from the genome annotation, biochemical pathway databases, biochemistry textbooks and recent publications, a genome-scale metabolic network for *S. cerevisiae* was designed. Additional physiological constraints were considered and modeled, such as growth, non-growth dependent ATP requirements and biomass composition.

[0016] Figure 6 shows a Phenotypic Phase Plane (PhPP) diagram for *S. cerevisiae* revealing a finite number of qualitatively distinct patterns of metabolic pathway utilization divided into discrete phases. The characteristics of these distinct phases are interpreted using ratios of shadow prices in the form of isolines. The isolines can be used to classify these phases into futile, single and dual substrate limitation and to define the line of optimality. The upper part of the figure shows a 3-dimensional *S. cerevisiae* Phase Plane diagram. The bottom part shows a 2-dimensional Phase Plane diagram with the line of optimality (LO) indicated.

[0017] Figure 7 shows the respiratory quotient (RQ) versus oxygen uptake rate (mmole/g-DW/hr) (upper left) on the line of optimality. The phenotypic phase plane (PhPP) illustrates that the predicted RQ is a constant of value 1.06

[0018] Figure 8 shows phases of metabolic phenotype associated with varying oxygen availability, from completely anaerobic fermentation to aerobic growth in *S. cerevisiae*. The glucose uptake rate was fixed under all conditions, and the resulting optimal biomass yield, as well as respiratory quotient, RQ, are indicated along with the output fluxes associated with four metabolic by-products: acetate, succinate, pyruvate, and ethanol.

[0019] Figure 9 shows anaerobic glucose limited continuous culture of *S. cerevisiae*. Figure 9 shows the utilization of glucose at varying dilution rates in anaerobic chemostat culture. The data-point at the dilution rate of 0.0 is extrapolated from the experimental results. The shaded area or the infeasible region contains a set of stoichiometric constraints that cannot be balanced simultaneously with growth demands. The model produces the optimal glucose uptake rate for a given growth rate on the line of optimal solution (indicated by Model (optimal)). Imposition of additional constraints drives the solution towards a region where more glucose is needed (i.e. region of alternative sub-optimal solution). At the optimal solution, the *in silico* model does not secrete pyruvate and acetate. The maximum difference between the model and the experimental points is 8% at the highest dilution rate. When the model is forced to produce these by-products at the experimental level (Model (forced)), the glucose uptake rate is increased and becomes closer to the experimental values. Figure 9B and 9C show the secretion rate of anaerobic by-products in chemostat culture. (q, secretion rate; D, dilution rate).

[0020] Figure 10 shows aerobic glucose-limited continuous culture of *S. cerevisiae* *in vivo* and *in silico*. Figure 10A shows biomass yield (Y_X), and secretion rates of ethanol (Eth), and glycerol (Gly). Figure 10B shows CO_2 secretion rate (q_{CO_2}) and respiratory quotient (RQ; i.e. $q_{\text{CO}_2}/q_{\text{O}_2}$) of the aerobic glucose-limited continuous culture of *S. cerevisiae*. (exp, experimental).

DETAILED DESCRIPTION OF THE INVENTION

[0021] The present invention provides an *in silico* model of the baker's and brewer's yeast, *S. cerevisiae*, that describes

the interconnections between the metabolic genes in the *S. cerevisiae* genome and their associated reactions and reactants. The model can be used to simulate different aspects of the cellular behavior of *S. cerevisiae* under different environmental and genetic conditions, thereby providing valuable information for industrial and research applications. An advantage of the model of the invention is that it provides a holistic approach to simulating and predicting the metabolic activity of *S. cerevisiae*.

[0022] As an example, the *S. cerevisiae* metabolic model can be used to determine the optimal conditions for fermentation performance, such as for maximizing the yield of a specific industrially important enzyme. The model can also be used to calculate the range of cellular behaviors that *S. cerevisiae* can display as a function of variations in the activity of one gene or multiple genes. Thus, the model can be used to guide the organismal genetic makeup for a desired application. This ability to make predictions regarding cellular behavior as a consequence of altering specific parameters will increase the speed and efficiency of industrial development of *S. cerevisiae* strains and conditions for their use.

[0023] The *S. cerevisiae* metabolic model can also be used to predict or validate the assignment of particular biochemical reactions to the enzyme-encoding genes found in the genome, and to identify the presence of reactions or pathways not indicated by current genomic data. Thus, the model can be used to guide the research and discovery process, potentially leading to the identification of new enzymes, medicines or metabolites of commercial importance.

[0024] The models of the invention are based on a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product.

[0025] As used herein, the term "*S. cerevisiae* reaction" is intended to mean a conversion that consumes a substrate or forms a product that occurs in or by a viable strain of *S. cerevisiae*. The term can include a conversion that occurs due to the activity of one or more enzymes that are genetically encoded by a *S. cerevisiae* genome. The term can also include a conversion that occurs spontaneously in a *S. cerevisiae* cell. Conversions included in the term include, for example, changes in chemical composition such as those due to nucleophilic or electrophilic addition, nucleophilic or electrophilic substitution, elimination, isomerization, deamination, phosphorylation, methylation, glycolysation, reduction, oxidation or changes in location such as those that occur due to a transport reaction that moves a reactant within the same compartment or from one cellular compartment to another. In the case of a transport reaction, the substrate and product of the reaction can be chemically the same and the substrate and product can be differentiated according to location in a particular cellular compartment. Thus, a reaction that transports a chemically unchanged reactant from a first compartment to a second compartment has as its substrate the reactant in the first compartment and as its product the reactant in the second compartment. It will be understood that when used in reference to an *in silico* model or data structure, a reaction is intended to be a representation of a chemical conversion that consumes a substrate or produces a product.

[0026] As used herein, the term "*S. cerevisiae* reactant" is intended to mean a chemical that is a substrate or a product of a reaction that occurs in or by a viable strain of *S. cerevisiae*. The term can include substrates or products of reactions performed by one or more enzymes encoded by *S. cerevisiae* gene(s), reactions occurring in *S. cerevisiae* that are performed by one or more non-genetically encoded macromolecule, protein or enzyme, or reactions that occur spontaneously in a *S. cerevisiae* cell. Metabolites are understood to be reactants within the meaning of the term. It will be understood that when used in reference to an *in silico* model or data structure, a reactant is intended to be a representation of a chemical that is a substrate or a product of a reaction that occurs in or by a viable strain of *S. cerevisiae*.

[0027] As used herein the term "substrate" is intended to mean a reactant that can be converted to one or more products by a reaction. The term can include, for example, a reactant that is to be chemically changed due to nucleophilic or electrophilic addition, nucleophilic or electrophilic substitution, elimination, isomerization, deamination, phosphorylation, methylation, reduction, oxidation or that is to change location such as by being transported across a membrane or to a different compartment.

[0028] As used herein, the term "product" is intended to mean a reactant that results from a reaction with one or more substrates. The term can include, for example, a reactant that has been chemically changed due to nucleophilic or electrophilic addition, nucleophilic or electrophilic substitution, elimination, isomerization, deamination, phosphorylation, methylation, reduction or oxidation or that has changed location such as by being transported across a membrane or to a different compartment.

[0029] As used herein, the term "stoichiometric coefficient" is intended to mean a numerical constant correlating the number of one or more reactants and the number of one or more products in a chemical reaction. Typically, the numbers are integers as they denote the number of molecules of each reactant in an elementally balanced chemical equation that describes the corresponding conversion. However, in some cases the numbers can take on non-integer values, for example, when used in a lumped reaction or to reflect empirical data.

[0030] As used herein, the term "plurality," when used in reference to *S. cerevisiae* reactions or reactants is intended to mean at least 2 reactions or reactants. The term can include any number of *S. cerevisiae* reactions or reactants in the range from 2 to the number of naturally occurring reactants or reactions for a particular strain of *S. cerevisiae*. Thus,

the term can include, for example, at least 10, 20, 30, 50, 100, 150, 200, 300, 400, 500, 600 or more reactions or reactants. The number of reactions or reactants can be expressed as a portion of the total number of naturally occurring reactions for a particular strain of *S. cerevisiae* such as at least 20%, 30%, 50%, 60%, 75%, 90%, 95% or 98% of the total number of naturally occurring reactions that occur in a particular strain of *S. cerevisiae*.

[0031] As used herein, the term "data structure" is intended to mean a physical or logical relationship among data elements, designed to support specific data manipulation functions. The term can include, for example, a list of data elements that can be added combined or otherwise manipulated such as a list of representations for reactions from which reactants can be related in a matrix or network. The term can also include a matrix that correlates data elements from two or more lists of information such as a matrix that correlates reactants to reactions. Information included in the term can represent, for example, a substrate or product of a chemical reaction, a chemical reaction relating one or more substrates to one or more products, a constraint placed on a reaction, or a stoichiometric coefficient.

[0032] As used herein, the term "constraint" is intended to mean an upper or lower boundary for a reaction. A boundary can specify a minimum or maximum flow of mass, electrons or energy through a reaction. A boundary can further specify directionality of a reaction. A boundary can be a constant value such as zero, infinity, or a numerical value such as an integer and non-integer.

[0033] As used herein, the term "activity," when used in reference to a reaction, is intended to mean the rate at which a product is produced or a substrate is consumed. The rate at which a product is produced or a substrate is consumed can also be referred to as the flux for the reaction.

[0034] As used herein, the term "activity," when used in reference to *S. cerevisiae* is intended to mean the rate of a change from an initial state of *S. cerevisiae* to a final state of *S. cerevisiae*. The term can include, the rate at which a chemical is consumed or produced by *S. cerevisiae*, the rate of growth of *S. cerevisiae* or the rate at which energy or mass flow through a particular subset of reactions.

[0035] The invention provides a computer readable medium, having a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product

[0036] The plurality of *S. cerevisiae* reactions can include reactions of a peripheral metabolic pathway. As used herein, the term "peripheral," when used in reference to a metabolic pathway, is intended to mean a metabolic pathway that includes one or more reactions that are not a part of a central metabolic pathway. As used herein, the term "central," when used in reference to a metabolic pathway, is intended to mean a metabolic pathway selected from glycolysis, the pentose phosphate pathway (PPP), the tricarboxylic acid (TCA) cycle and the electron transfer system (ETS), associated anapleurotic reactions, and pyruvate metabolism.

[0037] A plurality of *S. cerevisiae* reactants can be related to a plurality of *S. cerevisiae* reactions in any data structure that represents, for each reactant, the reactions by which it is consumed or produced. Thus, the data structure, which is referred to herein as a "reaction network data structure," serves as a representation of a biological reaction network or system. An example of a reaction network that can be represented in a reaction network data structure of the invention is the collection of reactions that constitute the metabolic reactions of *S. cerevisiae*.

[0038] The methods and models of the invention can be applied to any strain of *S. cerevisiae* including, for example, strain CEN.PK113.7D or any laboratory or production strain. A strain of *S. cerevisiae* can be identified according to classification criteria known in the art. Classification criteria include, for example, classical microbiological characteristics, such as those upon which taxonomic classification is traditionally based, or evolutionary distance as determined for example by comparing sequences from within the genomes of organisms, such as ribosome sequences.

[0039] The reactants to be used in a reaction network data structure of the invention can be obtained from or stored in a compound database. As used herein, the term "compound database" is intended to mean a computer readable medium or media containing a plurality of molecules that includes substrates and products of biological reactions. The plurality of molecules can include molecules found in multiple organisms, thereby constituting a universal compound database. Alternatively, the plurality of molecules can be limited to those that occur in a particular organism, thereby constituting an organism-specific compound database. Each reactant in a compound database can be identified according to the chemical species and the cellular compartment in which it is present. Thus, for example, a distinction can be made between glucose in the extracellular compartment versus glucose in the cytosol. Additionally each of the reactants can be specified as a metabolite of a primary or secondary metabolic pathway. Although identification of a reactant as a metabolite of a primary or secondary metabolic pathway does not indicate any chemical distinction between the reactants in a reaction, such a designation can assist in visual representations of large networks of reactions.

[0040] As used herein, the term "compartment" is intended to mean a subdivided region containing at least one reactant, such that the reactant is separated from at least one other reactant in a second region. A subdivided region included in the term can be correlated with a subdivided region of a cell. Thus, a subdivided region included in the term can be, for example, the intracellular space of a cell; the extracellular space around a cell; the periplasmic space; the interior space of an organelle such as a mitochondrion, endoplasmic reticulum, Golgi apparatus, vacuole or nucleus;

or any subcellular space that is separated from another by a membrane or other physical barrier. Subdivided regions can also be made in order to create virtual boundaries in a reaction network that are not correlated with physical barriers. Virtual boundaries can be made for the purpose of segmenting the reactions in a network into different compartments or substructures.

5 [0041] As used herein, the term "substructure" is intended to mean a portion of the information in a data structure that is separated from other information in the data structure such that the portion of information can be separately manipulated or analyzed. The term can include portions subdivided according to a biological function including, for example, information relevant to a particular metabolic pathway such as an internal flux pathway, exchange flux pathway, central metabolic pathway, peripheral metabolic pathway, or secondary metabolic pathway. The term can include portions subdivided according to computational or mathematical principles that allow for a particular type of analysis or manipulation of the data structure.

10 [0042] The reactions included in a reaction network data structure can be obtained from a metabolic reaction database that includes the substrates, products, and stoichiometry of a plurality of metabolic reactions of *S. cerevisiae*. The reactants in a reaction network data structure can be designated as either substrates or products of a particular reaction, each with a stoichiometric coefficient assigned to it to describe the chemical conversion taking place in the reaction. Each reaction is also described as occurring in either a reversible or irreversible direction. Reversible reactions can either be represented as one reaction that operates in both the forward and reverse direction or be decomposed into two irreversible reactions, one corresponding to the forward reaction and the other corresponding to the backward reaction.

15 [0043] Reactions included in a reaction network data structure can include intra-system or exchange reactions. Intra-system reactions are the chemically and electrically balanced interconversions of chemical species and transport processes, which serve to replenish or drain the relative amounts of certain metabolites. These intra-system reactions can be classified as either being transformations or translocations. A transformation is a reaction that contains distinct sets of compounds as substrates and products, while a translocation contains reactants located in different compartments. Thus, a reaction that simply transports a metabolite from the extracellular environment to the cytosol, without changing its chemical composition is solely classified as a translocation, while a reaction such as the phosphotransferase system (PTS) which takes extracellular glucose and converts it into cytosolic glucose-6-phosphate is a translocation and a transformation.

20 [0044] Exchange reactions are those which constitute sources and sinks, allowing the passage of metabolites into and out of a compartment or across a hypothetical system boundary. These reactions are included in a model for simulation purposes and represent the metabolic demands placed on *S. cerevisiae*. While they may be chemically balanced in certain cases, they are typically not balanced and can often have only a single substrate or product. As a matter of convention the exchange reactions are further classified into demand exchange and input/output exchange reactions.

25 [0045] The metabolic demands placed on the *S. cerevisiae* metabolic reaction network can be readily determined from the dry weight composition of the cell which is available in the published literature or which can be determined experimentally. The uptake rates and maintenance requirements for *S. cerevisiae* can be determined by physiological experiments in which the uptake rate is determined by measuring the depletion of the substrate. The measurement of the biomass at each point can also be determined, in order to determine the uptake rate per unit biomass. The maintenance requirements can be determined from a chemostat experiment. The glucose uptake rate is plotted versus the growth rate, and the y-intercept is interpreted as the non-growth associated maintenance requirements. The growth associated maintenance requirements are determined by fitting the model results to the experimentally determined points in the growth rate versus glucose uptake rate plot.

30 [0046] Input/output exchange reactions are used to allow extracellular reactants to enter or exit the reaction network represented by a model of the invention. For each of the extracellular metabolites a corresponding input/output exchange reaction can be created. These reactions can either be irreversible or reversible with the metabolite indicated as a substrate with a stoichiometric coefficient of one and no products produced by the reaction. This particular convention is adopted to allow the reaction to take on a positive flux value (activity level) when the metabolite is being produced or removed from the reaction network and a negative flux value when the metabolite is being consumed or introduced into the reaction network. These reactions will be further constrained during the course of a simulation to specify exactly which metabolites are available to the cell and which can be excreted by the cell.

35 [0047] A demand exchange reaction is always specified as an irreversible reaction containing at least one substrate. These reactions are typically formulated to represent the production of an intracellular metabolite by the metabolic network or the aggregate production of many reactants in balanced ratios such as in the representation of a reaction that leads to biomass formation, also referred to as growth. As set forth in the Examples, the biomass components to be produced for growth include L-Alanine, L-Arginine, L-Asparagine, L-Aspartate, L-Cysteine, L-Glutamine, L-Glutamate, Glycine, L-Histidine, L-Isoleucine, L-Leucine, L-Lysine, L-Methionine, L-Phenylalanine, L-Proline, L-Serine, L-Threonine, L-Tryptophan, L-Tyrosine, L-Valine, AMP, GMP, CMP, UMP, dAMP, dCMP, dTMP, dGMP, Glycogen, alpha, alpha-

Trehalose, Mannan, beta-D-Glucan, Triacylglycerol, Ergosterol, Zymosterol, Phosphatidate, Phosphatidylcholine, Phosphatidylethanolamine, Phosphatidyl-D-myo-inositol, Phosphatidylserine, ATP, Sulfate, ADP and Orthophosphate, with exemplary values shown in Table 1.

5 **Table 1.** Cellular components of *S. cerevisiae* (mmol/gDW).

ALA	0.459	CMP	0.05
ARG	0.161	dAMP	0.0036
ASN	0.102	dCMP	0.0024
ASP	0.297	dGMP	0.0024
CYS	0.007	DTMP	0.0036
GLU	0.302	TAGLY	0.007
GLN	0.105	ERGOST	0.0007
GLY	0.290	ZYMST	0.015
HIS	0.066	PA	0.0006
ILE	0.193	PINS	0.005
LEU	0.296	PS	0.002
LYS	0.286	PE	0.005
MET	0.051	PC	0.006
PHE	0.134	GLYCOGEN	0.519
PRO	0.165	TRE	0.023
SER	0.185	Mannan	0.809
THR	0.191	13GLUCAN	1.136
TRP	0.028	SLF	0.02
TYR	0.102	ATP	23.9166
VAL	0.265	ADP	23.9166
AMP	0.051	PI	23.9456
GMP	0.051	Biomass	1
UMP	0.067		

40 [0048] A demand exchange reaction can be introduced for any metabolite in a model of the invention. Most commonly these reactions are introduced for metabolites that are required to be produced by the cell for the purposes of creating a new cell such as amino acids, nucleotides, phospholipids, and other biomass constituents, or metabolites that are to be produced for alternative purposes. Once these metabolites are identified, a demand exchange reaction that is irreversible and specifies the metabolite as a substrate with a stoichiometric coefficient of unity can be created. With these specifications, if the reaction is active it leads to the net production of the metabolite by the system meeting potential production demands. Examples of processes that can be represented as a demand exchange reaction in a reaction network data structure and analyzed by the methods of the invention include, for example, production or secretion of an individual protein; production or secretion of an individual metabolite such as an amino acid, vitamin, nucleoside, antibiotic or surfactant; production of ATP for extraneous energy requiring processes such as locomotion; or formation of biomass constituents.

45 [0049] In addition to these demand exchange reactions that are placed on individual metabolites, demand exchange reactions that utilize multiple metabolites in defined stoichiometric ratios can be introduced. These reactions are referred to as aggregate demand exchange reactions. An example of an aggregate demand reaction is a reaction used to simulate the concurrent growth demands or production requirements associated with cell growth that are placed on a cell, for example, by simulating the formation of multiple biomass constituents simultaneously at a particular cellular growth rate.

50 [0050] A hypothetical reaction network is provided in Figure 1 to exemplify the above-described reactions and their interactions. The reactions can be represented in the exemplary data structure shown in Figure 2 as set forth below.

The reaction network, shown in Figure 1, includes intrasystem reactions that occur entirely within the compartment indicated by the shaded oval such as reversible reaction R_2 which acts on reactants B and G and reaction R_3 which converts one equivalent of B to two equivalents of F. The reaction network shown in Figure 1 also contains exchange reactions such as input/output exchange reactions A_{xt} and E_{xt} , and the demand exchange reaction, V_{growth} , which represents growth in response to the one equivalent of D and one equivalent of F. Other intrasystem reactions include R_1 which is a translocation and transformation reaction that translocates reactant A into the compartment and transforms it to reactant G and reaction R_6 which is a transport reaction that translocates reactant E out of the compartment.

[0051] A reaction network can be represented as a set of linear algebraic equations which can be presented as a stoichiometric matrix S, with S being an m x n matrix where m corresponds to the number of reactants or metabolites and n corresponds to the number of reactions taking place in the network. An example of a stoichiometric matrix representing the reaction network of Figure 1 is shown in Figure 2. As shown in Figure 2, each column in the matrix corresponds to a particular reaction n, each row corresponds to a particular reactant m, and each S_{mn} element corresponds to the stoichiometric coefficient of the reactant m in the reaction denoted n. The stoichiometric matrix includes intra-system reactions such as R_2 and R_3 which are related to reactants that participate in the respective reactions according to a stoichiometric coefficient having a sign indicative of whether the reactant is a substrate or product of the reaction and a value correlated with the number of equivalents of the reactant consumed or produced by the reaction. Exchange reactions such as $-E_{xt}$ and $-A_{xt}$ are similarly correlated with a stoichiometric coefficient. As exemplified by reactant E, the same compound can be treated separately as an internal reactant (E) and an external reactant ($E_{external}$) such that an exchange reaction (R_6) exporting the compound is correlated by stoichiometric coefficients of -1 and 1, respectively.

However, because the compound is treated as a separate reactant by virtue of its compartmental location, a reaction, such as R_5 , which produces the internal reactant (E) but does not act on the external reactant ($E_{external}$) is correlated by stoichiometric coefficients of 1 and 0, respectively. Demand reactions such as V_{growth} can also be included in the stoichiometric matrix being correlated with substrates by an appropriate stoichiometric coefficient.

[0052] As set forth in further detail below, a stoichiometric matrix provides a convenient format for representing and analyzing a reaction network because it can be readily manipulated and used to compute network properties, for example, by using linear programming or general convex analysis. A reaction network data structure can take on a variety of formats so long as it is capable of relating reactants and reactions in the manner exemplified above for a stoichiometric matrix and in a manner that can be manipulated to determine an activity of one or more reactions using methods such as those exemplified below. Other examples of reaction network data structures that are useful in the invention include a connected graph, list of chemical reactions or a table of reaction equations.

[0053] A reaction network data structure can be constructed to include all reactions that are involved in *S. cerevisiae* metabolism or any portion thereof. A portion of *S. cerevisiae* metabolic reactions that can be included in a reaction network data structure of the invention includes, for example, a central metabolic pathway such as glycolysis, the TCA cycle, the PPP or ETS; or a peripheral metabolic pathway such as amino acid biosynthesis, amino acid degradation, purine biosynthesis, pyrimidine biosynthesis, lipid biosynthesis, fatty acid metabolism, vitamin or cofactor biosynthesis, transport processes and alternative carbon source catabolism. Examples of individual pathways within the peripheral pathways are set forth in Table 2, including, for example, the cofactor biosynthesis pathways for quinone biosynthesis, riboflavin biosynthesis, folate biosynthesis, coenzyme A biosynthesis, NAD biosynthesis, biotin biosynthesis and thiamin biosynthesis.

[0054] Depending upon a particular application, a reaction network data structure can include a plurality of *S. cerevisiae* reactions including any or all of the reactions listed in Table 2. Exemplary reactions that can be included are those that are identified as being required to achieve a desired *S. cerevisiae* specific growth rate or activity including, for example, reactions identified as ACO1, CDC19, CIT1, DAL7, ENO1, FBA1, FDP1, FUM1, GND1, GDPM1, HXK1, ICL1, EDH1, IDH2, IDP1, IDP2, IDP3, KGD1, KGD2, LPD1, LSC1, LSC2, MDH1, MDH2, MDH3, MLS1, PDC1, PFK1, PFK2, PGI1, PGK1, PGM1, PGM2, PYC1, PYC2, PYK2, RKG1, RPE1, SOL1, TAL1, TDH1, TDH2, TDH3, TKL1, TPI1, ZWF1 in Table 2. Other reactions that can be included are those that are not described in the literature or genome annotation but can be identified during the course of iteratively developing a *S. cerevisiae* model of the invention including, for example, reactions identified as MET6_2, MNADC, MNADD1, MNADE, MNADF_1, MNADPHPS, MNADG1, MNADG2, MNADH, MNPT1.

50

55

5
10
15
20
25
30
35
40
45
50
55

Table 2

Locus #	E.C. #	Gene	Gene Description	Reaction	Rxn Name
Carbohydrate Metabolism					
Glycolysis/Gluconeogenesis					
YCL040W	2.7.1.2	GLK1	Glucokinase	GLC + ATP -> G6P + ADP	glk1_1
YCL040W	2.7.1.2	GLK1	Glucokinase	MAN + ATP -> MAN6P + ADP	glk1_2
YCL040W	2.7.1.2	GLK1	Glucokinase	bDGLC + ATP -> bDG6P + ADP	glk1_3
YFR053C	2.7.1.1	HXK1	Hexokinase I (PI) (also called Hexokinase A)	bDGLC + ATP -> G6P + ADP	hxk1_1
YFR053C	2.7.1.1	HXK1	Hexokinase I (PI) (also called Hexokinase A)	GLC + ATP -> G6P + ADP	hxk1_2
YFR053C	2.7.1.1	HXK1	Hexokinase I (PI) (also called Hexokinase A)	MAN + ATP -> MAN6P + ADP	hxk1_3
YFR053C	2.7.1.1	HXK1	Hexokinase I (PI) (also called Hexokinase A)	ATP + FRU -> ADP + F6P	hxk1_4
YGL253W	2.7.1.1	HXK2	Hexokinase II (PII) (also called Hexokinase B)	bDGLC + ATP -> G6P + ADP	hxk2_1
YGL253W	2.7.1.1	HXK2	Hexokinase II (PII) (also called Hexokinase B)	GLC + ATP -> G6P + ADP	hxk2_2
YGL253W	2.7.1.1	HXK2	Hexokinase II (PII) (also called Hexokinase B)	MAN + ATP -> MAN6P + ADP	hxk2_3
9		HXK2	Hexokinase II (PII) (also called Hexokinase B)	ATP + FRU -> ADP + F6P	hxk2_4
YBR196C	5.3.1.9	PGI1	Glucose-6-phosphate isomerase	Glucose -> F6P	pgi1_1
YBR196C	5.3.1.9	PGI1	Glucose-6-phosphate isomerase	Glucose -> bDG6P	pgi1_2
YBR196C	5.3.1.9	PGI1	Glucose-6-phosphate isomerase	bDG6P -> F6P	pgi1_3
YMR205C	2.7.1.11	PFK2	phosphofructokinase beta subunit	F6P + ATP -> FDP + ADP	pfk2
YGR240C	2.7.1.11	PFK1	phosphofructokinase alpha subunit	F6P + ATP -> FDP + ADP	pfk1_1
YGR240C	2.7.1.11	PFK1	phosphofructokinase alpha subunit	ATP + TAG6P -> ADP + TAG16P	pfk1_2
YGR240C	2.7.1.11	PFK1	phosphofructokinase alpha subunit	ATP + S7P -> ADP + S17P	pfk1_3
YKL060C	4.1.2.13	FBA1	fructose-bisphosphate aldolase	FDP -> T3P2 + T3P1	fba1_1
YDR050C	5.3.1.1	TPI1	triosephosphate isomerase	T3P2 -> T3P1	tpi1
YJL052W	1.2.1.12	TDH1	Glyceraldehyde-3-phosphate dehydrogenase 1	T3P1 + Pi + NAD ->> NADH + 13PDG	tdh1

5
10
15
20
25
30
35
40
45
50
55

Locus #	E.C. #	Gene	Gene Description	Reaction	Rxn Name
(continued)					
Carbohydrate Metabolism					
Glycolysis/Gluconeogenesis					
YJR009C	1.2.1.12	TDH2	glyceraldehyde 3-phosphate dehydrogenase	T3P1 + Pi + NAD \leftrightarrow NADH + 13PDG	tdh2
YGR192C	1.2.1.12	TDH3	Glyceraldehyde-3-phosphate dehydrogenase 3	T3P1 + Pi + NAD \leftrightarrow NADH + 13PDG	tdh3
YCR012W	2.7.2.3	PGK1	phosphoglycerate kinase	13PDG + ADP \leftrightarrow 3PG + ATP	pgk1
YKL152C	5.4.2.1	GPM1	Phosphoglycerate mutase	13PDG \leftrightarrow 23PDG	gpm1_1
YKL152C	5.4.2.1	GPM1	Phosphoglycerate mutase	3PG \leftrightarrow 2PG	gpm1_2
YDL021 W	5.4.2.1	GPM2	Similar to GPM1 (phosphoglycerate mutase)	3PG \leftrightarrow 2PG	gpm2
YOL056W	5.4.2.1	GPM3	phosphoglycerate mutase	3PG \leftrightarrow 2PG	gpm3
YGR254W	4.2.1.11	ENO1	enolase I	2PG \leftrightarrow PEP	eno1
YHR174W	4.2.1.11	ENO2	enolase	2PG \leftrightarrow PEP	eno2
YMR323W	4.2.1.11	ERR1	Protein with similarity to enolases	2PG \leftrightarrow PEP	eno3
YPL281C	4.2.1.11	ERR2	enolase related protein	2PG \leftrightarrow PEP	eno4
YOR393W	4.2.1.11	ERR1	enolase related protein	2PG \leftrightarrow PEP	eno5
YAL038W	2.7.1.40	CDC19	Pyruvate kinase	PEP + ADP \rightarrow PYR + ATP	cdc19
YOA347C	2.7.1.40	PYK2	Pyruvate kinase, glucose-repressed isoform	PEP + ADP \rightarrow PYR + ATP	pyk2
YER178w	1.2.4.1	PDA1	pyruvate dehydrogenase (lipoamide) alpha chain precursor, E1 component, alpha unit	PYRm + COAm + NADm \rightarrow NADHm + CO2m + ACCOAm	pda1
YBR221c	1.2.4.1	PDB1	pyruvate dehydrogenase (lipoamide) beta chain precursor, E1 component, beta unit	dihydrolipoamide S-acetyltransferase, E2 component	
YNL071w	2.3.1.12	LAT1			
Citrate cycle (TCA cycle)					
YNR001C	4.1.3.7	CIT1	Citrate synthase, Nuclear encoded	ACCOAm + OAm \rightarrow COAm + CItm	cit1

(continued)

Citrate cycle (TCA cycle)			
YCR005C	4.1.3.7	CIT2	
YPR001W	4.1.3.7	cit3	Citrate synthase, non-mitochondrial citrate synthase
YLR304C	4.2.1.3	acol	Citrate synthase, Mitochondrial isoform of citrate synthase
YJL200C	4.2.1.3	YJL200C	Aconitase, mitochondrial
YNL037C	1.1.1.41	IDH1	aconitate hydratase homolog
YOR136W	1.1.1.41	IDH2	Isocitrate dehydrogenase (NAD+) mito, subunit1 Isocitrate dehydrogenase (NAD+) mito, subunit2
YDL066W	1.1.1.42	IDP1	Isocitrate dehydrogenase (NADP+) Isocitrate dehydrogenase (NADP+)
YLR174W	1.1.1.42	IDP2	Isocitrate dehydrogenase (NADP+) OSUC
YNL009W	1.1.1.42	IDP3	Isocitrate dehydrogenase (NADP+) OSUC
YDL066W	1.1.1.42	IDP1	Isocitrate dehydrogenase (NADP+) OSUC
YLR174W	1.1.1.42	IDP2	Isocitrate dehydrogenase (NADP+) OSUC
YNL009W	1.1.1.42	IDP3	Isocitrate dehydrogenase (NADP+) OSUC
YIL125W	1.2.4.2	kgd1	alpha-ketoglutarate dehydrogenase complex, E1 component CO2m + NADHm + SUCCOAm
YDR148C	2.3.1.61	KGD2	Dihydrolipoamide S- succinyltransferase, E2 component
YGR244C	6.2.1.4/6. 2.1.5	LSC2	Succinate-CoA ligase (GDP-forming)
YOR142W	6.2.1.4/6. 2.1.5	LSC1	succinate-CoA ligase alpha subunit
<i>Electron Transport System, Complex II</i>		SDH3	
YKL141W	1.3.5.1	SDH1	succinate dehydrogenase cytochrome b succinate dehydrogenase cytochrome b
YKL148C	1.3.5.1		

Electron Transport System, Complex II

YLL041c 1.3.5.1

SDH2

Succinate dehydrogenase
(ubiquinone) iron-sulfur protein

YDR178W 1.3.5.1

SDH4

succinate dehydrogenase membrane
subunit

YLR164W 1.3.5.1

YLR164W

strong similarity to SDH4P
YMR118C 1.3.5.1

YMR118C

strong similarity to succinate

YJL045w 1.3.5.1

YJL045w

dehydrogenase

YEL047c 1.3.99.1

YEL047c

strong similarity to succinate
dehydrogenase flavoprotein

YJR051W 1.3.99.1

osm1

soluble fumarate reductase,
cytoplasmic

YPL262W 4.2.1.2

YPL262W

Mitochondrial soluble fumarate
reductase involved in osmotic
regulation

YPL262W 4.2.1.2

YPL262W

Fumarate

YKL085W 1.1.1.37

MDH1

Fumarate

YDL078C 1.1.1.37

MDH3

mitochondrial malate dehydrogenase

YOL126C 1.1.1.37

MDH2

MALATE DEHYDROGENASE,
PEROXISOMAL

YER085C 4.1.3.1

ICL1

Malate synthase

YPR006C 4.1.3.1

ICL2

isocitrate lyase

YIR031C 4.1.3.2

dal7

isocitrate lyase, may be nonfunctional

YNL117W 4.1.3.2

MLS1

Malate synthase

YKR097W 4.1.1.49

pck1

phosphoenolpyruvate carboxylkinase

YLR377C 3.1.3.11

FBP1

fructose-1,6-bisphosphatase

FDP -> F6P + Pi

FADH2m + FUM -> SUCC +

FADm

FADH2m + FUMm ->

SUCCm + FAdm

fads1

osm1

fum1_1

fum1_2

mdh1

mdh3

mls1

icl1

icl2

dal7

ACC OA + GLX -> COA +

MAL

ACCOA + GLX -> COA +

MAL

ACC OA + GLX -> COA +

MAL

OA + ATP -> PEP + CO2 +

ADP

pck1

fbp1

(continued)

55 50 45 40 35 30 25 20 15 10 5

(continued)

55	50	45	40	35	30	25	20	15	10
YBR299W	3.2.1.20	MAL32	Maltase	MLT -> 2 GLC	mal32a				
YGR287C	3.2.1.20	YGR287 C	putative alpha glucosidase	MLT -> 2 GLC	mal32b				
YGR292W	3.2.1.20	MAL12	Maltase	MLT -> 2 GLC	mal12a				
YIL172C	3.2.1.20	YIL172C	putative alpha glucosidase	MLT -> 2 GLC	mal12b				
YJL216C	3.2.1.20	YJL216C	probable alpha glucosidase (MAL Tase)	MLT -> 2 GLC	mal12c				
YJL221C	3.2.1.20	FSP2	homology to maltase (alpha-D-glucosidase)	MLT -> 2 GLC	fsp2a				
YJL221C	3.2.1.20	FSP2	homology to maltase (alpha-D-glucosidase)	6DGlc -> GLAC + GLC	fsp2b				
YBR018C	2.7.7.12	GAL7	UDPG glucose--hexose-1-phosphate uridylyltransferase	UDPG + GAL1P <-> G1P + UDPGAL	unkrx10				
Trehalose		TPS1	trehalose-6-P synthetase, 56 kD synthase subunit of trehalose-6-phosphate synthaseVphosphatase complex	UDPG + G6P -> UDP + TRE6P tps1					
YBR126C	2.4.1.15	tsl1	trehalose-6-P synthetase, 123 kD regulatory subunit of trehalose-6-phosphate synthaseVphosphatase complex; homologous to TPS3 gene product	UDPG + G6P -> UDP + TRE6P ts1					
YML100W	2.4.1.15		trehalose-6-P synthetase, 115 kD regulatory subunit of trehalose-6-phosphate synthaseVphosphatase complex	UDPG + G6P -> UDP + TRE6P tps3					
YMR261C	2.4.1.15	TPS3	Trehalose-6-phosphate phosphatase Acid trehalase	TRE6P -> TRE + PI	tps2				
			Neutral trehalase, highly homologous to Nth1p	TRE -> 2 GLC	ath1				
YDR074W	3.1.3.12	TPS2	Neutral trehalase, highly homologous to Nth1p	TRE -> 2 GLC	nth2				
YPR026W	3.2.1.28	ATH1	neutral trehalase	TRE -> 2 GLC	nth1				
YBR001C	3.2.1.28	NTH2	Branching enzyme, 1,4-glucan-6-(1,4-glucano)-transferase	GLYCOGEN + PI -> G1P	glc3				
YDR001C	3.2.1.28	NTH1	Glycogen phosphorylase	GLYCOGEN + PI -> G1P	gph1				
YEL011W	2.4.1.18	gic3							
YPR160W	2.4.1.1	GPH1							

(continued)

(continued)

Glycogen Metabolism (sucrose and sugar metabolism)

YFR015C	2.4.1.11	GSY1		Glycogen synthase (UDP-glucosese-starch glucosyltransferase)	UDPG -> UDP + GLYCOGEN	gsy1
YLR258W	2.4.1.11	GSY2		Glycogen synthase (UDP-glucosese-starch glucosyltransferase)	UDPG -> UDP + GLYCOGEN	gsy2
Pyruvate Metabolism						
YAL054C	6.2.1.1	acs1				
YLR153C	6.2.1.1	ACS2				
YDL168W	1.2.1.1	SFA1				
YJL068C	3.1.2.12	-				
YGR087C	4.1.1.1	PDC6				
YLR134W	4.1.1.1	PDC5				
YLR044C	4.1.1.1	pdc1				
YBL015W	3.1.2.1	ACH1				
YBL015W	3.1.2.1	ACH1				
YDL131W	4.1.3.21	LYS21				
YDL182W	4.1.3.21	LYS20				
YDL182W	4.1.3.21	LYS20				
YGL256W	1.1.1.1	adh4				
YMR083W	1.1.1.1	adh3				
YMR303C	1.1.1.1	adh2				
YBR145W	1.1.1.1	ADH5				
YOL086C	1.1.1.1	adh1				
YDL168W	1.1.1.1	SFA1				

5

10

15

20

25

30

35

40

45

50

55

(continued)

Electron Transport System, Complex III

YBL045C 1.10.2.2

COR1

ubiquinol-cytochrome c reductase core

subunit 1

ubiquinol-cytochrome c reductase core

subunit 2

ubiquinol-cytochrome c reductase

ubiquinol-cytochrome c reductase

subunit 6

ubiquinol-cytochrome c reductase

subunit 7

ubiquinol-cytochrome c reductase

subunit 8

ubiquinol-cytochrome c reductase

subunit 9

ubiquinol-cytochrome c reductase

subunit 10

Electron Transport System, Complex IV

COX1

cytochrome c oxidase subunit I

QH2m + 2 FERl m + 1.5 + Hm cytr

-> Qm 2 FER Om

COX2

cytochrome c oxidase subunit I

COX3

cytochrome c oxidase subunit I

COX9

cytochrome c oxidase subunit I

COX4

cytochrome c oxidase subunit I

COX13

cytochrome c oxidase subunit I

COX6

cytochrome c oxidase subunit I

COX5B

cytochrome c oxidase subunit I

COX12

cytochrome c oxidase subunit I

COX8

cytochrome c oxidase subunit I

COX7

cytochrome c oxidase subunit I

COX5A

cytochrome c oxidase subunit I

F1F0-ATPase complex, F1 alpha

subunit

F1F0-ATPase complex, F1 epsilon

subunit

ADPm + PIm -> ATPm + 3 Hm atp1

ATP15

3.6.1.34

ATP1

3.6.1.34

5

10

15

30

35

40

50

ATP Synthase								
YDL004W	3.6.1.34							
Q0085	3.6.1.34	ATP16	F-type H+-transporting ATPase delta chain					
YBR039W	3.6.1.34	ATP6	F1FO-ATPase complex, FO A subunit					
YBR127C	3.6.1.34	ATP3	F1FO-ATPase complex, F1 gamma subunit					
YPL078C	3.6.1.34	VMA2	H+-ATPase V1 domain 60 KD subunit, vacuolar					
YDR298C	3.6.1.34	ATP4	F1FO-ATPase complex, F1 delta subunit					
YDR377W	3.6.1.34	ATP5	F1FO-ATPase complex, OSCP subunit					
YJR121W	3.6.1.34	ATP17	ATP synthase complex, subunit f					
YKL016C	3.6.1.34	ATP2	F1FO-ATPase complex, F1 beta subunit					
YLR295C	3.6.1.34	ATP7	F1FO-ATPase complex, FO D subunit					
Q0080	3.6.1.34	ATP14	ATP synthase subunit h					
Q0130	3.6.1.34	ATP8	F-type H+-transporting ATPase subunit 8					
YOL077W-A	3.6.1.34	ATP9	F-type H+-transporting ATPase subunit c					
YPR020W	3.6.1.34	ATP19	ATP synthase k chain, mitochondrial					
YLR447C	3.6.1.34	ATP20	subunit G of the dimeric form of mitochondrial F1FO-ATP synthase					
YGR020C	3.6.1.34	VMA6	V-type H+-transporting ATPase subunit AC39					
YKL080W	3.6.1.34	VMA7	V-type H+-transporting ATPase subunit F					
YDL185W	3.6.1.34	VMA5	V-type H+-transporting ATPase subunit C					
YBR127C	3.6.1.34	TFP1	V-type H+-transporting ATPase subunit A					
YOR332W	3.6.1.34	VMA2	V-type H+-transporting ATPase subunit B					
YEI027W	3.6.1.34	VMA4	V-type H+-transporting ATPase subunit E					
		CUP5	V-type H+-transporting ATPase proteolipid subunit					

(continued)

5

10

15

20

30

35

40

50

55

(continued)

ATP Synthase

YHR026W

3.6.1.34

PPA1

YPL234C

3.6.1.34

TFP3

YMR054W

3.6.1.34

STV1

YOR270C

3.6.1.34

VPH1

YEL051W

3.6.1.34

VMA8

YHR039C-A

3.6.1.34

VMA10

YPR036W

3.6.1.34

VMA13

Electron Transport System, Complex IV

Q0045

1.9.3.1

COX1

Q0275

1.9.3.1

COX3

V-type H+-transporting ATPase

proteolipid subunit

V-type H+-transporting ATPase

proteolipid subunit

V-type H+-transporting ATPase subunit

vacuolar ATP synthase subunit G

V-type H+-transporting ATPase 54 kD

subunit

cytochrome-c oxidase subunit I

cytochrome-c oxidase subunit II,

mitochondrially-coded

cytochrome-c oxidase subunit II

Cytochrome-c oxidase

cytochrome-c oxidase chain IV

cytochrome-c oxidase chain V_{ia}

cytochrome-c oxidase subunit VI

cytochrome-c oxidase chain V_b

cytochrome-c oxidase, subunit VII

cytochrome-c oxidase chain VII

cytochrome-c oxidase, subunit VII

cytochrome-c oxidase chain V.A

precursor

Lactic acid dehydrogenase

2 FERlOm + LLACm -> PYRm + cyb2

2 FERlOm

2 FERlOm + LACm -> PYRm + dld1

2 FERlOm

55	50	45	40	35	30	25	20	15	10	5
(continued)										
Methane metabolism										
YPL275W	1.2.1.2				YPL275 W	putative formate dehydrogenase/ putative pseudogene	FOR + NAD \rightarrow CO ₂ + NADH	tfo1a		
YPL276W	1.2.1.2				YPL276 W	putative formate dehydrogenase/putative pseudogene	FOR + NAD \rightarrow CO ₂ + NADH	tfo1b		
YOR388C	1.2.1.2				FDH1	Protein with similarity to formate dehydrogenases	FOR + NAD \leftrightarrow CO ₂ + NADH	fdh1		
Nitrogen metabolism										
YBR208C	6.3.4.6				DUR1	urea amidolyase containing urea carboxylase / allophanate hydrolase	ATP + UREA + CO ₂ \leftrightarrow ADP + PI + UREAC	dur1		
YBR208C	3.5.1.54				DUR1	Allophanate hydrolase	UREAC \rightarrow 2 NH ₃ + 2 CO ₂	dur2		
YJL126W	3.5.5.1				NIT2	nitrilase	ACNL \rightarrow INAC + NH ₃	nit2		
Sulfur metabolism (Cysteine biosynthesis maybe)										
YJR137C	1.8.7.1				ECM17	Sulfite reductase	H ₂ SO ₃ + 3 NADPH \leftrightarrow H ₂ S + 3 NADP	ecm17		
Lipid Metabolism										
Fatty acid biosynthesis										
YER015W	6.2.1.3				FAA2	Long-chain-fatty-acid-CoA ligase, Acyl-CoA synthetase	ATP + LCCA + COA \leftrightarrow AMP + PPI + ACOA	faa2		
YIL009W	6.2.1.3				FAA3	Long-chain-fatty-acid-CoA ligase, Acyl-CoA synthetase	ATP + LCCA + COA \leftrightarrow AMP + PPI + ACOA	faa3		
YOR317W	6.2.1.3				FAA1	Long-chain-fatty-acid-CoA ligase, Acyl-CoA synthetase	ATP + LCCA + COA \leftrightarrow AMP + PPI + ACOA	faa1		
YMR246W	6.2.1.3				FAA4	Acyl-CoA synthase (long-chain fatty acid CoA ligase); contributes to activation of imported myristate	ATP + LCCA + COA \leftrightarrow AMP + PPI + ACOA	faa4		
YKR009C	1.1.1.-				FOX2	3-Hydroxyacyl-CoA dehydrogenase	HACOA + NAD \leftrightarrow OACOA + NADH	fox2b		
YIL160C	2.3.1.16				pot1	3-Ketoacyl-CoA thiolase	OACOA + COA \rightarrow ACOA + ACCOA	pot1_1		
YPL028W	2.3.1.9				erg10	Acetyl-CoA C-acetyltransferase, ACETOACETYL-COA THIOLASE	2 ACCOA \leftrightarrow COA + ACCOA	erg10_1		
YPL028W	2.3.1.9				erg10	Acetyl-CoA C-acetyltransferase, ACETOACETYL-COA THIOLASE (mitoch)	2 ACCOAm \leftrightarrow COAm + ACCOAm	erg10_2		

Fatty Acids Metabolism**Mitochondrial type II fatty acid synthase**

ER061CYO	1/-	55	50	45	40	35	30	25	20	15	10	5
(continued)												
R221CYKL												
055C	YKL192CY	1	ACP1/C	EM1/MC	T1/OAR							
ER061CYO	1.6.5.3/-		EM1/MC									
R221CYKL			T1/OAR									
055C	YKL192CY	1	ACP1/C	EM1/MC	T1/OAR							
ER061CYO	1.6.5.3/-		EM1/MC									
R221CYKL			T1/OAR									
055C	YKL192CY	1	ACP1/C	EM1/MC	T1/OAR							
ER061CYO	1.6.5.3/-		EM1/MC									
R221CYKL			T1/OAR									
055C	YKL192CY	1	ACCPm + 8 MALACPm + 16 TypeII_7	NADPHm -> 13 NADPm + C161ACPm + 7 CO2m + 7 ACPm								
ER061CYO	1.6.5.3/-		NADPHm -> 16 NADPm + C180ACPm + 8 CO2m + 8 ACPm									
R221CYKL												
055C	YKL192CY	1	ACACPm + 8 MALACPm + 15 TypeII_8	NADPHm -> 15 NADPm + C181ACPm + 8 CO2m + 8 ACPm								
ER061CYO	1.6.5.3/-		NADPHm -> 14 NADPm + C182ACPm + 8 CO2m + 8 ACPm									
R221CYKL												
055C	YKL192CY	1	ACCOA + ATP + CO2 -> accl	ACCOA + ADP + PI								
ER061CYO	1.6.5.3/-		ACCOA + ADP + PI	MALCOA + ACP <-> MALACP + COA								
R221CYKL												
YNR016C	6.4.1.2 6.3.4.14	ACC1	fatty-acyl-CoA synthase, beta chain	fatty-acyl-CoA synthase, alpha chain								
YKL182w	4.2.1.61; 1.3.1.9; 2.3.1.38; 2.	fas1										
YPL231w	3.1.39; 3.1.2.14; 2.3.1.86	FAS2										
YKL182w	2.3.1.85; 1.1.1.100 ; 2.3.1.41	fas1										
YER061C	4.2.1.61; 1.3.1.9; 2.3.1.38; 2.	CEM1										
YGR037CY	3.1.39; 3.1.2.14; 2.3.1.86	ACB1/A										
NR016CYK	6.4.1.2;	CC1/fas1										

(continued)

Cytosolic fatty acid synthesis		(continued)	
L182W/YPL	2.3.1.85;	/FAS2/	ACP
231W	1.1.1.100 ;2.3.1.41; 4.2.1.61		
YGR037C/Y	6.4.1.2;	ACB1/A	b-Ketoacyl-ACP synthase (C12,0), fatty acyl CoA synthase
NR016C/YK	6.3.4.1; 4	CC1/fas1	ACACCP + 5 MALACP + 10 NADPH -> 10 NADP + C120ACP + 5 CO2 + 5 ACP
L182W/YPL	2.3.1.85;	/FAS2/	ACACCP + 6 MALACP + 12 NADPH
231W	1.1.1.100 ;2.3.1.41; 4.2.1.61	ACB1/A	-> 12 NADP + C140ACP + 6 CO2 + 6 ACP
YGR037C/Y	6.4.1.2;	CC1/fas1	ACACCP + 6 MALACP + 11 NADPH
NR016C/YK	6.3.4.1; 4	/FAS2/	-> 11 NADP + C141ACP + 6 CO2 + 6 ACP
L182W/YPL	2.3.1.85;		ACACCP + 7 MALACP + 14 NADPH -> 14 NADP + C160ACP + 7 CO2 + 7 ACP
231W	1.1.1.100 ;2.3.1.41; 4.2.1.61		c160s
YGR037C/Y	6.4.1.2;	ACB1/A	ACACCP + 7 MALACP + 13 NADPH -> 13 NADP + C161ACP + 7 CO2 + 7 ACP
NR016C/YK	6.3.4.1; 4	CC1/fas1	c161s
L182W/YPL	2.3.1.85;	/FAS2/	
231W	1.1.1.100 ;2.3.1.41; 4.2.1.61		
YGR037C/Y	6.4.1.2;	ACB1/A	
NR016C/YK	6.3.4.1;	4 CC1/fas1	
L182W/YPL	2.3.1.85;	/FAS2/	
231W	1.1.1.100 ;2.3.1.41; 4.2.1.61		

55	50	45	40	35	30	25	20	15	10	5
Cytosolic fatty acid synthesis										
YGR037C/Y	6.4.1.2;									
NR016C/YK	6.3.4.1; 4									
L182W/YPL231w	2.3.1.85; 1.1.1.100 ; 2.3.1.41; 4.2.1.61	ACB1/A CC1/fas1 /FAS2/	b-Ketoacyl-ACP synthase I (C18,0)	ACACAP + 8 MALACP + 16 NADPH -> 16 NADP + C180ACP + 8 CO2 + 8 ACP	c180sy					
YGR037C/Y	6.4.1.2;									
NR016C/YK	6.3.4.1; 4									
L182W/YPL231w	2.3.1.85; /FAS2/ 1.1.1.100 ; 2.3.1.41; 4.2.1.61	ACB1/A CC1/fas1	b-Ketoacyl-ACP synthase I (C18,1)	ACACAP + 8 MALACP + 15 NADPH c181sy -> 15 NADP + C181ACP + 8 CO2 + 8 ACP						
YGR037C/Y	6.4.1.2;									
NR016C/YK	6.3.4.1;									
L182W/YPL231w	2.3.1.85; 1.1.1.100 ; 2.3.1.41; 4.2.1.61	ACB1/A 4 CC1/fas1 /FAS2/	b-Ketoacyl-ACP synthase I (C18,2)	ACACAP + 8 MALACP + 14 NADPH -> 14 NADP + C182ACP + 8 CO2 + 8 ACP	c182sy					
YKL182W	4.2.1.61	fas1	3-hydroxypalmitoyl-[facyl]-carrier protein] dehydratase	3HPACP ->> 2HDACP	fas1_3					
YKL182W	1.3.1.9	fas1	Enoyl-ACP reductase	AACP + NAD ->> 23DAACP	fas1_4					
				+ NADH						
Fatty acid degradation										
YGL205W/	1.3.3.6/2.	POX1/F	Fatty acid degradation	C140 + ATP + 7 COA + 7	c140dg					
YKR009C/Y	3.1.18	OX2/PO		FADm + 7 NAD -> AMP + PPI						
IL160C		T3		+ 7 FADH2m + 7 NADH + 7						
YGL205W/	1.3.3.6/2.	POX1/F	Fatty acid degradation	ACCOA						
YKR009C/Y	3.1.18	OX2/PO		C160 + ATP + 8 COA + 8	c160dg					
IL160C		T3		FADm + 8 NAD -> AMP + PPI						
YGL205W/	1.3.3.6/2.	POX1/F	Fatty acid degradation	+ 8 FADH2m + 8 NADH + 8						
				ACCOA						
YKR009C/Y	3.1.18	OX2/PO		C180 + ATP + 9 COA + 9	c180dg					
IL160C		T3		FADm + 9 NAD -> AMP + PPI						
				+ 9 FADH2m + 9 NADH + 9						
				ACCOA						

5
10
15
20
25
30
35
40
45
50
55

Phospholipid Biosynthesis

(continued)

-	Glycerol-3-phosphate acyltransferase	GL3P + 0.017 C100ACP + 0.062 C120ACP + 0.1 C140ACP + 0.27 C160ACP + 0.169 C161ACP + 0.055 C180ACP + 0.235 C181ACP + 0.093 C182ACP -> AGL3P + ACP	Gat1_1
-	Glycerol-3-phosphate acyltransferase	GL3P + 0.017 C100ACP + 0.062 C120ACP + 0.1 C140ACP + 0.27 C160ACP + 0.169 C161ACP + 0.055 C180ACP + 0.235 C181ACP + 0.093 C182ACP -> AGL3P + ACP	Gat2_1
-	Glycerol-3-phosphate acyltransferase	T3P2 + 0.017 C100ACP + 0.062 C120ACP + 0.1 C140ACP + 0.27 C160ACP + 0.169 C161ACP + 0.055 C180ACP + 0.235 C181ACP + 0.093 C182ACP -> AT3P2 + ACP	Gat1_2
-	Glycerol-3-phosphate acyltransferase	T3P2 + 0.017 C100ACP + 0.062 C120ACP + 0.1 C140ACP + 0.27 C160ACP + 0.169 C161ACP + 0.055 C180ACP + 0.235 C181ACP + 0.093 C182ACP -> AT3P2 + ACP	Gat2_2
-	Acyldihydroxyacetonephosphate reductase	AT3P2 + NADPH -> AGL3P + NADP ADHAPR	

55	50	45	40	35	30	25	20	15	10
YDL052C	2.3.1.51	SLC1	-	1-Acylglycerol-3-phosphate acyltransferase	AGL3P + 0.017 C100ACP + 0.062 C120ACP + 0.100 C140ACP + 0.270 C160ACP + 0.169 C161ACP + 0.055 C180ACP + 0.235 C181ACP + 0.093 C182ACP -> PA + ACP	slc1			
	(continued)			1-Acylglycerol-3-phosphate acyltransferase	AGL3P + 0.017 C100ACP + 0.062 C120ACP + 0.100 C140ACP + 0.270 C160ACP + 0.169 C161ACP + 0.055 C180ACP + 0.235 C181ACP + 0.093 C182ACP -> PA + ACP	AGAT			
	2.3.1.51	-	-	CDP-Diacylglycerol synthetase	PAm + CTPm -> CDPDGm	cds1a			
				CDP-Diacylglycerol synthetase	+ PPIm				
				phosphatidylserine synthase	PA + CTP -> CDPDG + PPI	cds1b			
				cho1	CDPDG + SER -> CMP + PS	cho1a			
				Phosphatidylserine synthase	CMPGm + SERm -> CMPm + PSm	cho1b			
				phosphatidylserine decarboxylase located in vacuole or Golgi	PS -> PE + CO2	psd2			
				Phosphatidylserine Decarboxylase 1	PSm -> PEm + CO2m	psd1			
				Phosphatidylethanolamine N-methyltransferase	SAM + PE -> SAH + PMME	cho2			
				Methylene-fatty-acyl-phospholipid synthase.	SAM + PMME -> SAH + PDME	op13_1			
				N-methyltransferase	PDME + SAM -> PC + SAH	op13_2			
				Choline kinase	ATP + CHO -> ADP + PCHO	ck11			
				Cholinephosphate cytidylyltransferase	PCHO + CTP -> CDPCCHO + PPI	pct1			
				Diacylglycerol cholinephosphotransferase	CDPCHO + DAGLY -> PC + CMP	CPT1	cpt1		

55	50	45	40	35	30	25	20	15	10	5
(continued)										
Phospholipid Biosynthesis										
YDR147W	2.7.1.82	EKI1								
YGR007W	2.7.7.14	MUQ1								
YHR123W	2.7.8.1	EPT1								
YJL153C	5.5.1.4	ino1								
YHR046C	3.1.3.25	INM1								
YPR113W	2.7.8.11	PLS1								
YJR066W	2.7.1.137	tor1								
YKL203C	2.7.1.137	tor2								
YLR240W	2.7.1.137	vps34								
YNL267W	2.7.1.67	PIK1								
YLR305C	2.7.1.67	STT4								
YFR019W	2.7.1.68	FAB1								
YDR208W	2.7.1.68	MSS4								
YPL268W	3.1.4.11	plc1								
YCL004W -	2.7.8.8 3.1.3.27	PGS1								
YDL142C	2.7.8.5	CRD1								

5	10	15	20	25	30	35	40	45	50	55
(continued)										
Phospholipid Biosynthesis										
YDR284C	DPP1	diacylglycerol pyrophosphate phosphatase			PA -> DAGLY + PI					dpp1
YDR503C	LPP1	lipid phosphate phosphatase		DGPP -> PA + PI						lpp1
Sphingoglycolipid Metabolism										
YDR062W	LCB2	Serine C-palmitoyltransferase		PALCOA + SER -> COA + DHSPH + CO2						lcb2
YMR296C	LCB1	Serine C-palmitoyltransferase		PALCOA + SER -> COA + DHSPH + CO2						lcb1
YBR285W	TSC10	3-Dehydroosphinganine reductase		DHSPH + NADPH -> SPH + NADP						tsc10
YDR297W	SUR2	SYRINGOMYCIN RESPONSE PROTEIN 2		SPH + 02 + NADPH -> PSPH + NADP						sur2
-		Ceramide synthase		PSPH + C260COA -> CER2 + COA						csyna
-		Ceramide synthase		PSPH + C240COA -> CER2 + COA						c symb
YMR272C	SCS7	Ceramide hydroxylase that hydroxylates the C-26 fatty-acyl moiety of inositol-phosphorylceramide		CER2 + NADPH + 02 -> CER3 + NADP						scs7
YKL004W	AUR1	IPS synthase, AUREOBASIDIN A RESISTANCE PROTEIN		CER3 + PINs -> IPC						aur1
YBR036C	CSG2	Protein required for synthesis of the mannosylated sphingolipids		IPC + GDPMAN -> MIPI C						csg2
YPL057C	SUR1	Protein required for synthesis of the mannosylated sphingolipids		IPC + GDPMAN -> MIPI C						sur1
YDR072C	IPT1	MIPI2C synthase, MANNOSYL DIPHOSPHORYLINOSITOL CERAMIDE SYNTHASE		MIPI C + PINs -> MIPI2C						ipt1
YOR171C	LCB4	Long chain base kinase, involved in sphingolipid metabolism		SPH + ATP -> DHSP + ADP						lcb4_1
YLR260W	LCB5	Long chain base kinase, involved in sphingolipid metabolism		SPH + ATP -> DHSP + ADP						lcb5_1
YOR171C	LCB4	Long chain base kinase, involved in sphingolipid metabolism		PSPH + ATP -> PHSP + ADP						lcb4_2

5							
10							
15							
20							
25							
30							
35							
40							
45							
50							
55							
(continued)							
Sphingoglycolipid Metabolism							
YLR260W	LCB5	Long chain base kinase, involved in sphingolipid metabolism	PSPH + ATP -> PHSP + ADP	lcb5_2			
YJL134W	LCB3	Sphingoid base-phosphate phosphatase, putative regulator of sphingolipid metabolism and stress response	DHSP -> SPH + PI	lcb3			
YKR053C	YSR3	Sphingoid base-phosphate phosphatase, putative regulator of sphingolipid metabolism and stress response	DHSP -> SPH + PI	ysr3			
Sterol biosynthesis							
YDR294C	DPL1	Dihydrosphingosine-1-phosphate lyase	DHSP -> PETHM + C16A	dpl1			
YML126C	HMGS	3-hydroxy-3-methylglutaryl/coenzyme A synthase	H3MCOA + COA -> ACCOA + AACCOA	hmgs			
YLR450W	hmg2	3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase isozyme	MVL + COA + 2 NADP <-> H3MCOA + 2 NADPH	hmg2			
YML075C	hmg1	3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase isozyme	MVL + COA + 2 NADP <-> H3MCOA + 2 NADPH	hmg1			
YMR208W	erg12	A (HMG-CoA) reductase isozyme	MVL + COA + 2 NADP <-> H3MCOA + 2 NADPH	erg12_1			
YMR208W	erg12	mevalonate kinase	ATP + MVL -> ADP + PMVL	erg12_2			
YMR208W	erg12	mevalonate kinase	CTP + MVL -> CDP + PMVL	erg12_3			
YMR208W	erg12	mevalonate kinase	GTP + MVL -> GDP + PMVL	erg12_4			
YMR220W	ERG8	48 kDa Phosphomevalonate kinase	UTP + MVL -> UDP + PMVL	erg8			
YNR043W	MVD1	Diphosphomevalonate decarboxylase	ATP + PMVL -> ADP + PMVL	mvd1			
YPL117C	idi1	Isopentenyl diphosphate:dimethylallyl diphosphate isomerase (IPP isomerase)	ATP + PMVL -> ADP + PI + IPP + CO2	idi1			
		IPPP -> DMPP					
Prenylation							
YJL167W	ERG20	prenyltransferase	DMPP + IPPP -> GPP + PPI	erg20_1			
YJL167W	ERG20	Farnesyl diphosphate synthetase (FPP synthetase)	GPP + IPPP -> FPP + PPI	erg20_2			

55	50	45	40	35	30	25	20	15	10	5
(continued)										
YDR294C	DPL1									
YHR190W	ERG9									
YGR175C	ERG1									
YHR072W	ERG7									
YHR007c	erg1									
YNL280c	ERG24									
YGR060w	ERG25									
YGL001c	ERG26									
YLR100C	YLR100 C									
YGR060w	ERG25									
YGL001c	ERG26									
YLR100C	YLR100 C									
YML008c	erg6									
YMR202W	ERG2									
YLR056w	ERG3									
YMR015c	ERG5									
YGL012w	ERG4									

55	50	YDR294C	DPL1	(continued)	Dihydroxyinosine-1-phosphate lyase	DHSP -> PETHM + C16A MZYMST + 3'02 + 4 NADPH + NAD -> ZYMST + CO2 + 4 NADP + NADH	dpl1 unkrxn4
				5.3.3.5	Cholestenol delta-isomerase	ZYMST + SAM -> ERGOST + SAH	cdiso4
					Histidine Biosynthesis		
	45				PRSS5	R5P + ATP <-> PRPP + AMP	prs5
					PRS4	R5P + ATP <-> PRPP + AMP	prs4
					PRS2	R5P + ATP <-> PRPP + AMP	prs2
					PRS3	R5P + ATP <-> PRPP + AMP	prs3
					PRS1	R5P + ATP <-> PRPP + AMP	prs1
					dai1	ATN <-> ATT	dai1
					dai2	ATT <-> UGC + UREA	dai2
					dai3	UGC <-> GLX + 2 NH3 + CO2	dai3
					CYR1	ATP -> cAMP + PPI	cyr1
					GUK1	GMP + ATP <-> GDP + ADP	guk1_1
					GUK1	DGMP + ATP <-> DGDP + ADP	guk1_2
					GLU1	GMP + DATP <-> GDP + DADP	guk1_3
						PRPP + GLN -> PPI + GLU + PRAM	
					ade4		
					ade5,7	PRAM + ATP + GLY <-> ADP	ade5
						+ PI + GAR	
					YGL234W	glycinamide ribotide synthetase	
						aminooimidazole ribotide synthetase	
						glycinamide ribotide transformylase	
						5'-phosphoribosylformyl glycinamide	
						synthetase	
					YGR061C	FGAR + ATP + GLN -> GLU + ADP + PI + FGAM	ade6
					YGL234W	FGAM + ATP -> ADP + PI + AIR	ade7
						Phosphoribosylformyl glycinamide	
						cyclo-ligase	

55	50	45	40	35	30	25	20	15	10	5			
Purine metabolism													
YOR128C	4.1.1.21	ade2											
YAR015W	6.3.2.6	ade1											
YLR359W	4.3.2.2	ADE13											
YLR028C	2.1.2.3	ADE16											
YMR120C	2.1.2.3	ADE17											
YLR028C	3.5.4.10	ADE16											
YMR120C YNL220W	2.1.2.3 6.3.4.4	ADE17 ade12											
YLR359W	4.3.2.2	ADE13											
YAR073 W	1.1.1.205	fun63											
YHR216W YML056C	1.1.1.205 1.1.1.205	pur5 IMD4											
YLR432W	1.1.1.205	IMD3											
YAR075 W	1.1.1.205	YAR075 W											
YMR217W	6.3.5.2, 6.3.4.1	GUAI											
YML035C	3.5.4.6	amd1											
(continued)													
phosphoribosylamino-imidazole-carboxylase													
phosphoribosyl amino imidazolesuccinocarbozamide synthetase													
5'-Phosphoribosyl-4-(N-succinocarboxamide)-5'-aminoimidazole lyase													
5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase/VIMP cyclohydrolase													
5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase/VIMP cyclohydrolase													
5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) transformylase/VIMP cyclohydrolase													
IMP cyclohydrolase													
adenylosuccinate synthetase													
Adenylosuccinate Lyase putative inosine-5'-monophosphate dehydrogenase													
purine excretion													
probable inosine-5'-monophosphate dehydrogenase (IMP)													
probable inosine-5'-monophosphate dehydrogenase (IMP)													
Protein with strong similarity to inosine-5'-monophosphate dehydrogenase, pseudogene													
GMP synthase													
AMP deaminase													
XMP + ATP + GLN -> GLU + gual													
AMP + PPi + GMP													
AMP -> IMP + NH3													
amid1													

		(continued)	
5			
10			
15			
20			
25			
30			
35			
40			
45			
50			
55			
	Purine metabolism		
YGL248W	3.1.4.17	PDE1	cAMP -> AMP cAMP -> AMP
YOR360C	3.1.4.17	pde2	3',5'-Cyclic-nucleotide phosphodiesterase, low affinity 3',5'-Cyclic-nucleotide phosphodiesterase, high affinity
YOR360C	3.1.4.17	pde2	cdAMP -> DAMP
YOR360C	3.1.4.17	pde2	cIMP -> IMP
YOR360C	3.1.4.17	pde2	cGMP -> GMP
YOR360C	3.1.4.17	pde2	cCMP -> CMP
YDR530C	2.7.7.53	APA2	ADP + ATP -> PI + ATRP
YCL050C	2.7.7.53	apaI	ADP + GTP -> PI + ATRP
YCL050C	2.7.7.53	apaI	GDP + GTP -> PI + GTRP II
			apaI_-
			apa2_-
			ura4
			ura1_-
	Pyrimidine metabolism		
YJL130C	2.1.3.2	ura2	Aspartate-carbamoyltransferase
YL420W	3.5.2.3	ura4	dihydroorotate e
YKL216W	1.3.3.1	ura1	dihydroorotate dehydrogenase
			Dihydroorotate dehydrogenase
YKL216W	1.3.3.1	PYRD	Dihydroorotate dehydrogenase
			Orotate phosphoribosyltransferase 1
			Orotate phosphoribosyltransferase 2
			orotidine-5'-phosphate decarboxylase
			Nucleoside-phosphate kinase
			UPRTase, Uracil
			phosphoribosyltransferase
			cytosine deaminase
			Thymidine (deoxyuridine) kinase
			Thymidine (deoxyuridine) kinase
			Uridine kinase
			Cytidine kinase
YPR062W	3.5.4.1	FCY1	CYTS -> URA + NH3
	-	-	DU + ATP -> DUMP + ADP
	2.7.1.21	-	DT + ATP -> ADP + DTMP
	2.7.1.21	-	URI + GTP -> UMP + GDP
	2.7.1.48	URK1	CYTID + GTP -> GDP + CMP
	2.7.1.48	URK1	urk1_-

55	50	45	40	35	30	25	20	15	10	5
(continued)										
Pyrimidine metabolism										
YNR012W	2.7.1.48		URK1		Uridine kinase, converts ATP and uridine to ADP and	URI + ATP -> ADP + UMP UMP	urk1_3			
YLR209C	2.4.2.4		PNP1		Protein with similarity to human nucleoside phosphorylase, Thymidine (deoxyuridine) phosphorylase, Purine nucleotide phosphorylase	purine DU + Pi -> URA + DR1P	deap1			
YLR209C	2.4.2.4		PNP1		Protein with similarity to human nucleoside phosphorylase, Thymidine (deoxyuridine) phosphorylase	purine DT + Pi -> THY + DR	deap2			
YLR245C	3.5.4.5	CDD1			Cytidine deaminase	CYTID -> URI + NH3	cdd1_1			
YLR245C	3.5.4.5	CDD1			Cytidine deaminase	DC -> NH3 + DU	cdd1_2			
YJR057W	2.7.4.9	cdc8			dTMP kinase	DTMP + ATP -> ADP + DTDP	cd8			
YDR353W	1.6.4.5	TRR1			Thioredoxin reductase	OTHIO + NADPH -> NADP + RTHO	trr1			
YHR106W	1.6.4.5	TRR2			mitochondrial thioredoxin reductase	OTHOm + NADPHm -> NADPm + RTHOm	trr2			
YBR252W	3.6.1.23	DUT1			dUTP pyrophosphatase (dUTPase)	DUTP -> PPi + DUMP	dut1			
YOR074C -	2.1.1.45 2.7.4.14	cdc21			Thymidylate synthase Cytidine kinase	DUMP + METTHF -> DHF + DTMP DCMP + ATP ->> ADP + DCDP	cd21 cmka1			
-	2.7.4.14				Cytidylate kinase	CMP + ATP ->> ADP + CDP	cmka2			
YHR144C	3.5.4.12	DCD1			dCMP deaminase	DCMP ->> DUMP + NH3	dc01			
YBL039C	6.3.4.2	URA7			CTP synthase, highly homologus to URA8 CTP synthase	UTP + GLN + ATP -> GLU + CTP + ADP + Pi	ura7_1			
YJR103W	6.3.4.2	URA8			CTP synthase	UTP + GLN + ATP -> GLU + CTP + ADP + Pi	ura8_1			
YBL039C	6.3.4.2	URA7			CTP synthase, highly homologus to URA8 CTP synthase	ATP + UTP + NH3 -> ADP + Pi + CTP	ura7_2			
YJR103W	6.3.4.2	URA8			CTP synthase	ATP + UTP + NH3 -> ADP + Pi + CTP	ura8_2			
YNL292W	4.2.1.70	PUS4			Pseudouridine synthase	URA + R5P -> PUR15P	pus4			
YPL212C	4.2.1.70	PUS1			intranuclear protein which exhibits a nucleotide-specific intron-dependent tRNA pseudouridine synthase activity	URA + R5P -> PUR15P	pus1			

5	10	15	20	25	30	35	40	45	50	55
(continued)										
Pyrimidine metabolism										
YGL063W	4.2.1.70	PUS2	pseudouridine synthase 2		URA + R5P \leftrightarrow PUR15P					pus2
YFL001W	4.2.1.70	deg1	Similar to rRNA methyltransferase (<i>Caenorhabditis elegans</i>) and hypothetical 28K protein (alkaline endoglucanase gene 5' region) from <i>Bacillus</i> sp.		URA + R5P \leftrightarrow PUR15P					deg1
Salvage Pathways										
YML022W	2.4.2.7	APT1	Adenine phosphoribosyltransferase		AD + PRPP \rightarrow PPI + AMP					apt1
YDR441C	2.4.2.7	APT2	similar to adenine phosphoribosyltransferase		AD + PRPP \rightarrow PPI + AMP					apt2
YNL141W	3.5.4.4	AAH1	adenine aminohydrolase (adenine deaminase)		ADN \rightarrow INS + NH3					aah1a
YNL141W	3.5.4.4	AAH1	adenine aminohydrolase (adenine deaminase)		DA \rightarrow DIN + NH3					aah1b
YLR209C	2.4.2.1	PNP1	Purine nucleotide phosphorylase, Xanthosine phosphorylase		DIN + PI \leftrightarrow HYXN + DR1P					xapa1
YLR209C	2.4.2.1	PNP1	Xanthosine phosphorylase, Purine nucleotide phosphorylase		DA + PI \leftrightarrow AD + DR1P					xapa2
YLR209C	2.4.2.1	PNP1	Xanthosine phosphorylase		DG + PI \leftrightarrow GN + DR1P					xapa3
YLR209C	2.4.2.1	PNP1	Xanthosine phosphorylase, Purine nucleotide phosphorylase		HYXN + R1P \leftrightarrow INS + PI					xapa4
YLR209C	2.4.2.1	PNP1	Xanthosine phosphorylase		AD + R1P \leftrightarrow PI + ADN					xapa5
YLR209C	2.4.2.1	PNP1	Xanthosine phosphorylase, Purine nucleotide phosphorylase		GN + R1P \leftrightarrow PI + GSN					xapa6
YLR209C	2.4.2.1	PNP1	Xanthosine phosphorylase, Purine nucleotide phosphorylase		XAN + R1P \leftrightarrow PI + XTSINE					xapa7
YJR133W	2.4.2.22	XPT1	Xanthine-guanine phosphoribosyltransferase		XAN + PRPP \rightarrow XMP+PPI					gpt1
YDR400W	3.2.2.1	urhl	Purine nucleosidase		GSN \rightarrow GN + RIB					pur21
YDR400W	3.2.2.1	urhl	Purine nucleosidase		ADN \rightarrow AD + RIB					pur11
YJR105W	2.7.1.20	YJR105 W	Adenosine kinase		ADN + ATP \rightarrow AMP + ADP					pm2
YDR226W	2.7.4.3	adk1	cytosolic adenylylate kinase		ATP + AMP \leftrightarrow 2 ADP					adk1_1
YDR226W	2.7.4.3	adk1	cytosolic adenylylate kinase		GTP + AMP \leftrightarrow ADP + GDP					adk1_2

Salvage Pathways

YDR226W	2.7.4.3	adk1	cytosolic adenylate kinase	ITP + AMP <-> ADP + IDP	adk1_3
YER170W	2.7.4.3	ADK2	Adenylate kinase (mitochondrial GTP: AMP phosphotransferase)	ATPM + AMPm <-> 2 ADPm	adk2_1
YER170W	2.7.4.3	adk2	Adenylate kinase (mitochondrial GTP: AMP phosphotransferase)	GTPm + AMPm <-> ADPm + GDPm	adk2_2
YER170W	2.7.4.3	adk2	Adenylate kinase (mitochondrial GTP: AMP phosphotransferase)	ITPm + AMPm <-> ADPm + IDPm	adk2_3
YGR180C	1.17.4.1	RNR4	ribonucleotide reductase, small subunit (alt), beta chain	ADP + RTHIO -> DADP + OTHIO	rnr3
YIL066C	1.17.4.1	RNR3	Ribonucleotide reductase (ribonucleoside-diphosphate reductase) large subunit, alpha chain	ADP + RTHIO -> DADP + OTHIO	rnr3
YJL026W	1.17.4.1	mr2	small subunit of ribonucleotide reductase, beta chain	ADP + RTHIO -> DADP + OTHIO	rnr3
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	UDP + ATP <-> UTP + ADP	ynk1_1
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	CDP + ATP <-> CTP + ADP	ynk1_2
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	DGDP + ATP <-> DGTP + ADP	ynk1_3
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	DUDP + ATP <-> DUTP + ADP	ynk1_4
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	DCDP + ATP <-> DCUTP + ADP	ynk1_5
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	DTDP + ATP <-> DTTP + ADP	ynk1_6
YKL067W	2.7.4.6	YNK1	Nucleoside-diphosphate kinase	DADP + ATP <-> DATP + ADP	ynk1_7
YKL067W	2.7.4.6	YNK1	Nucleoside diphosphate kinase	GDP + ATP <-> GTP + ADP	ynk1_8
YKL067W	2.7.4.6	YNK1	Nucleoside diphosphate kinase	IDP + ATP <-> ITP + IDP	ynk1_9
-	2.7.4.11	AAH1	Adenylyl kinase, dAMP kinase	DAMP + ATP <-> DADP + ADP	dampk
YNL141W	3.5.4.2	-	Adenine deaminase	AD -> NH3 + HYXN	yicp
-	2.7.1.73	-	Inosine kinase	INS + ATP -> IMP + ADP	gsk1
-	2.7.1.73	-	Guanosine kinase	GSN + ATP -> GMP + ADP	gsk2
YDR399W	2.4.2.8	HPT1	Hypoxanthine phosphoribosyltransferase	HYXN + PRPP -> PPI + IMP	hpt1_1

(continued)

5 10 15 20 25 30 35 40 45 50 55

Salvage Pathways	YDR399W	2.4.2.8	HPT1	GMP + PRPP → PPI + GMP	hpt1_2
-	YKL024C	2.4.2.3	UR1	Uridine phosphoryltransferase	
YKL024C	2.1.4.-	URA6	UR1 + PI <-> URA + R1P	udp	
YKL024C	2.1.4.-	URA6	UMP + ATP <-> UDP + ADP	pyrh1	
-	3.2.2.10	3.2.2.10	DUMP + ATP <-> DUDP + ADP	pyrh2	
YHR144C	3.5.4.13	dCTP	CMP -> CYTS + R5P	cmpg	
-	3.1.3.5	dCTP deaminase	DCTP -> DUTP + NH3	dcd	
-	3.1.3.5	5'-Nucleotidase	DUMP -> DU + PI	usha1	
-	3.1.3.5	5'-Nucleotidase	DTMP -> DT + PI	usha2	
-	3.1.3.5	5'-Nucleotidase	DAMP -> DA + PI	usha3	
-	3.1.3.5	5'-Nucleotidase	DGMP -> DG + PI	usha4	
-	3.1.3.5	5'-Nucleotidase	DCMP -> DC + PI	usha5	
-	3.1.3.5	5'-Nucleotidase	CMP -> CYTD + PI	usha6	
-	3.1.3.5	5'-Nucleotidase	AMP -> PI + ADN	usha7	
-	3.1.3.5	5'-Nucleotidase	GMP -> PI + GSN	usha8	
-	3.1.3.5	5'-Nucleotidase	IMP -> PI + INS	usha9	
-	3.1.3.5	5'-Nucleotidase	XMP -> PI + XTSINE	usha12	
-	3.1.3.5	5'-Nucleotidase	UMP -> PI + URI	usha11	
YER070W	1.17.4.1	RNR1	ADP + RTHIO -> DADP + OTHIO	rnr1_1	
YER070W	1.17.4.1	RNR1	Ribonucleoside-diphosphate reductase	GDP + RTHIO -> DGDP + OTHIO	rnr1_2
YER070W	1.17.4.1	RNR1	Ribonucleoside-diphosphate reductase	CDP + RTHIO -> DCDP + OTHIO	rnr1_3
YER070W	1.17.4.1	RNR1	Ribonucleoside-diphosphate reductase	UDP + RTHIO -> OTHIO + OTHIO	rnr1_4
-	1.17.4.2	1.17.4.2	Ribonucleoside-triphosphate reductase	ATP + RTHIO -> DATP + OTHIO	nrd1
-	1.17.4.2	1.17.4.2	Ribonucleoside-triphosphate reductase	GTP + RTHIO -> DGTP + OTHIO	nrd2
-	1.17.4.2	1.17.4.2	Ribonucleoside-triphosphate reductase	CTP + RTHIO -> DCTP + OTHIO	nrd3

(continued)

Salvage Pathways	1.17.4.2	55	50	45	40	35	30	25	20	15	10	5
(continued)												
-									Ribonucleoside-triphosphate reductase	UTP + RTTHIO -> OTTHIO +		nrdd4
3.6.1.-									DUTP			
3.6.1.-									GTP -> GSN + 3 Pi			mutt1
3.2.2.4									DGTP -> DG + 3 Pi			mutt2
YML035C	AMD1								AMP -> AD + R5P			amn1
YBR284W	YBR284W								AMP -> AD + R5P			amn1
YJL070C	3.2.2.4								AMP -> AD + R5P			amn2
Amino Acid Metabolism												
Glutamate Metabolism (Aminosugars met)												
YMR250W	4.1.1.15								Glutamate decarboxylase B	GLU -> GABA + CO2		btn2
YGR019W	2.6.1.19								Aminobutyrate aminotransaminase 2	GABA + AKG -> SUCCSAL +		uga1
YBR006w	1.2.1.16								GLU			
YKL104C	2.6.1.16								Succinate semialdehyde dehydrogenase —NADP	+ NADPH		gabda
									Glutamine_fuctose-6-phosphate amidotransferase (glucosamine-6-phosphate synthase)	F6P + GLN -> GLU + GA6P		gfal1
YFL017C	2.3.1.4								Glucosamine-phosphate N-acetyltransferase	ACCOA + GA6P -> COA + NAGA6P		gnal
YEL053W	5.4.2.3								Phosphoacetylglucosamine Mutase	NAGA1P -> NAGA6P		pcm1a
YDL103C	2.7.7.23								N-Acetylglucosamine-1-phosphate-uridylyltransferase	UTP + NAGA1P -> UDPNAG + PPI		qrl
									chitin synthase 3	UDPNAG -> CHIT + UDP		chs3
YBR028C	2.4.1.16								chitin synthase 2	UDPNAG -> CHIT + UDP		chs2
YBR038W	2.4.1.16								chitin synthase 2	UDPNAG -> CHIT + UDP		chs1
YNL192W	2.4.1.16								delta-1-pyrrole-5-carboxylate dehydrogenase	GLUGSALm + NADPm -> NADPHm + GLUm		put2_1
YHR037W	1.5.1.12								P5Cm + NADm -> NADHm + GLUm			
									AKG + GLN + NADH -> NAD + 2 GLU			glt1
YDL171C	1.4.1.14								GLU + NAD -> AKG + NH3 + NADH			gdh2
YDL215C	1.4.1.4											

Amino Acid Metabolism Glutamate Metabolism (Aminosugars met) YAL062W 1.4.1.4 YOR375C 1.4.1.4 YPR035W 6.3.1.2 YEL058W 5.4.2.3 - 3.5.1.2 - 3.5.1.2 Glucosamine - Arabinose YBR149W 1.1.1.117 YBR149W 1.1.1.117 Xylose YGR194C Mannitol - Alanine and Aspartate Metabolism YKL106W 2.6.1.1 YLR027C 2.6.1.1 YAR035W 2.3.1.7 YML042W 2.3.1.7 YDR111C 2.6.1.2 YLR089C 2.6.1.2	GDH3 GDH1 gln1 PCM1 gln1 gln1 gln1 gln1 nagb ARA1 ARA1 ARA1 XKS1 XKS1 AAT1 AAT2 YAT1 CAT2 ACAR PYR YGL089C	NADP-linked glutamate dehydrogenase AKG + NH3 + NADPH <-> GLU + NADP AKG + NH3 + NADPH <-> GLU gdh3 NADP-specific glutamate dehydrogenase glutamine synthetase gln1 Phosphoglucosamine mutase Glutaminase A Glutaminase B Glucosamine-6-phosphate deaminase D-arabinose 1-dehydrogenase (NAD(P)) ARAB + NAD → ARABLAC + NADH D-arabinose 1-dehydrogenase (NAD(P)) ARAB + NADP → ARABLAC + ara1_1 ara1_2 D-arabinose 1-dehydrogenase (NAD(P)) ARAB + NADP → ARABLAC + NADPH XUL + ATP → X5P + ADP xks1 MNT6P + NAD ↔ F6P + mtld OAm + GLUm ↔ ASPm + AKGm OA + GLU ↔ ASP + AKG COAm + ACARm → ACCOAm + CARm ACCOA + CAR → COA + ACAR PYR + GLU ↔ AKG + ALA YGL089C ALAm putative alanine transaminase alanine aminotransferase, mitochondrial precursor (glutamic--
55	50	5
45	40	10
35	30	15
20	25	5
10	15	5
(continued)		
40		

55	50	45	40	35	30	25	20	15	10	5
(continued)										
YNL277W	2.3.1.31	met2			homoserine O-trans-acetylase	ACCOA + HSER ↔ COA +				
YLR303W	4.2.99.10	MET17			O-Acetylhomoserine (thiol)-lyase	OAHSER + METH → MET +	met17_1			
YLR303W	4.2.99.8	MET17			O-Acetylhomoserine (thiol)-lyase	OAHSER + H2S → AC +	met17_2			
YLR303W	4.2.99.8, 4.2.99.10	met17			O-acetylhomoserine sulfhydrylase	OAHSER + H2S → AC +	met17_3			
YML082W	4.2.99.9	YML082W			(OAH SHLase); converts O-acetylhomoserine into homocysteine	OAHSER + H2S → AC +				
YDR502C	2.5.1.6	sam2			putative cystathione gamma-synthase	OSLHSER ↔ SUCC + OBUT	met17h			
YLR180W	2.5.1.6	sam1				+ NH4				
YLR172C	2.1.1.98	DPH5			S-adenosylmethionine synthetase	MET + ATP → PPI + PI + SAM	sam2			
Cysteine Biosynthesis										
YJR010W	2.7.7.4	met3			S-adenosylmethionine synthetase	MET + ATP → PPI + PI + SAM	sam1			
YKL001C	2.7.1.25	met14			Diphthine synthase	MET + ATP → PPI + PI + SAM	sam1			
YFR030W	1.8.1.2	met10				SAM + CALH → SAH + DPTH	dph5			
-	2.3.1.30				ATP sulfurylase	SLF + ATP → PPI + APS	met3			
YGR012W	4.2.99.8	YGR012W			adenylylsulfate kinase	APS + ATP → ADP + PAPS	met14			
YOL064C	3.1.3.7	MET22			sulfite reductase	H2SO3 + 3 NADPH <-> H2S +	met10			
YPR167C	1.8.99.4	MET16				3 NADP				
YCL050C	2.7.7.5	apal			Serine transacetylase	SER + ACCOA → COA +	cys1			
Branched Chain Amino Acid Metabolism (Valine, Leucine and Isoleucine)										
YHR208W	2.6.1.42	BAT1			putative cysteine synthase (O-acetylisserine sulfhydrylase) (O-3'-5'Bisphosphate nucleotidase)	ASER + H2S → AC + CYS	sul11			
YHR208W	2.6.1.42	BAT1			PAPS Reductase	PAP → AMP + PI	met22			
						PAPS + RTHIO → OTTHIO +	met16			
						H2SO3 + PAP				
						ADP + SLF → PI + APS	apal_2			
Branched chain amino acid aminotransferase										
YHR208W	2.6.1.42	BAT1			Branched chain amino acid aminotransferase	OICAPm + GLUm ↔ AKGm +	bat1_1			
YHR208W	2.6.1.42	BAT1			Branched chain amino acid aminotransferase	LEUm				
						OMVALm + GLUm ↔ AKGm	bat1_2			
						+ IEm				

(continued)

Lysine biosynthesis/degradation

YGL154C	1.2.1.31	lys5	L-Aminoadipate-semialdehyde dehydrogenase, small subunit	AMA + NADH + ATP → AMASA + NAD + AMP + PPI	lys2_2
YBR115C	1.2.1.31	lys2	L-Aminoadipate-semialdehyde dehydrogenase, large subunit		
YGL154C	1.2.1.31	lys5	L-Aminoadipate-semialdehyde dehydrogenase, small subunit		
YNR050C	1.5.1.10	lys9	Saccharopine dehydrogenase (NADP+, L-glutamate forming)	GLU + AMASA + NADPH ↔ SACP + NADP	lys9
YIR034C	1.5.1.7	lys1	Saccharopine dehydrogenase (NAD+, L-lysine forming)	SACP + NAD ↔ LYS + AKG + NADH	lys1a
YDR037W	6.1.1.6	krs1	lysyl-tRNA synthetase, cytosolic	ATP + LYS + LTRNA → AMP + PPI + LLTRNA	krs1
YNL073W	6.1.1.6	msk1	lysyl-tRNA synthetase, mitochondrial	ATPm + LYSm + LTRNAm → msk1	
YDR368W	1.1.1.-	YPR1	similar to aldo-keto reductase	AMPm + PPIm + LLTRNAm	
Arginine metabolism					
YMR062C	2.3.1.1	ECM40	Amino-acid N-acetyltransferase	GLUm + ACCOAm → COAm + NAGLUm	ecm40_1
YER069W	2.7.2.8	arg5	Acetylglutamate kinase	NAGLUm + ATPm → ADPm + NAGLUpm	arg6
YER069W	1.2.1.38	arg5	N-acetyl-gamma-glutamyl-phosphate reductase and acetylglutamate kinase	NAGLUpm + NADPHm → NADPm + PIm + NAGLUsm	arg5
YOL140W	2.6.1.11	arg8	Acetylornithine aminotransferase	NAGLUsm + GLUm → AKGm + NAORNm	arg8
YMR062C	2.3.1.35	ECM40	Glutamate N-acetyltransferase	NAORNm + GLUm → ORNm + NAGLUm	ecm40_2
YJL130C	6.3.5.5	ura2	carbamoyl-phosphate synthetase, aspartate transcarbamylase, and glutamine amidotransferase	GLN + 2 ATP + CO2 → GLU + CAP + 2 ADP + PI	ura2_2
YJR109C	6.3.5.5	CPA2	carbamoyl phosphate synthetase, chain	GLN + 2 ATP + CO2 → GLU + CAP + 2 ADP + PI	cpa2
YOR303W	6.3.5.5	cpa1	Carbamoyl phosphate synthetase, small chain, arginine specific		
YJL088W	2.1.3.3	arg3	Ornithine carbamoyltransferase	ORN + CAP → CITR + PI	arg3
YLR438W	2.6.1.13	car2	Ornithine transaminase	ORN + AKG → GLUGSAL + GLU	car2

5 10 15 20 25 30 35 40 45 50 55

5	10	15	20	25	30	35	40	45	50	55	5
(continued)											
Arginine metabolism											
YOL058W	6.3.4.5	arg1									
YHR018C	4.3.2.1		arginosuccinate synthetase								
YKL184W	4.1.1.17	arg4									
YOL052C	4.1.1.50	spe1	argininosuccinate lyase								
YPR069C	2.5.1.16	spe2	Ornithine decarboxylase								
		SPE3	S-adenosylmethionine decarboxylase								
			putrescine aminopropyltransferase								
			(spermidine synthase)								
YLR146C	2.5.1.22	SPE4	Spermine synthase								
YDR242W	3.5.1.4	AMD2	SFRM								
YMR293C	3.5.1.4	YMR293 C	GBAD → GBAT + NH3								
YPL111W	3.5.3.1	car1	GBAD → GBAT + NH3								
YDR341C	6.1.1.19	YDR341 C	ARG → ORN + UREA								
			ATP + ARG + ATRNA → AMP								
YHR091C	6.1.1.19	MSR1	+ PPI + ALTRNA								
YHR068W	1.5.99.6	DYS1	ATP + ARG + ATRNA → AMP								
			+ PPI + ALTRNA								
			SPRMD + Qm → DAPRP +								
			QH2m								
Histidine metabolism											
YER055C	2.4.2.17	his1	ATP phosphoribosyltransferase								
YCL030C	3.6.1.31	his4	phosphoribosyl-AMP cyclohydrolase								
			phosphoribosyl-ATP								
			pyrophosphorylase / histidinol								
			dehydrogenase								
			histidinol dehydrogenase								
			phosphoribosyl-5-amino-1-								
			phosphoribosyl-4-								
			imidazolecarboxiamide isomerase								
YCL030C	3.5.4.19	his4	PRBAMP → PRFP								
YIL020C	5.3.1.16	his6	PRFP → PRLP								
YOR202W	4.2.1.19	his3	his4_2								
YIL116W	2.6.1.9	his5	his5								
YFR025C	3.1.3.15	his2	his6								
			his3								
			IMGP → IMACP								
			IMACP + GLU → AKG +								
			HISOLP								
			HISOLP → PI + HISOL								

55	50	45	40	35	30	25	20	15	10
Histidine metabolism					(continued)				
YCL030C	1.1.1.23		his4		phosphoribosyl-AMP cyclohydrolase phosphoribosyl-ATP pyrophosphohydrolase / histidinol dehydrogenase	/ HISOL + 2 NAD → HIS + 2 NADH	his4_3		
YBR248C	2.4.2.-		his7		glutamine amidotransferase:cyclase	PRLP + GLN → GLU + AICAR	his7		
YPR033C	6.1.1.21		his1		histidyl-tRNA synthetase	DIMGP ATP + HIS + HTRNA → AMP + PPI + HHTRNA		hts1	
YBR034C	2.1.1.-		hmt1		hnRNP arginine N-methyltransferase	SAM + HIS → SAH + MHIS		hmt1	
YCL054W	2.1.1.-		spb1		putative RNA methyltransferase				
YML110C	2.1.1.-		coq5		ubiquinone biosynthesis methyltransferase COQ5				
YOR201C	2.1.1.-		pet56		rRNA (guanosine-2'- O-)methyltransferase				
YPL266W	2.1.1.-		dim1		dimehyladenosine transferase				
YBR249C	4.1.2.15		ARO4		3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase	E4P + PEP → PI + 3DDAH7P	aro4		
YDR035W	4.1.2.15		ARO3		DAHP synthase; a.k.a. phospho-2-dehydro-3-deoxyheptonate aldolase, phenylalanine-inhibited; phospho-2-keto-3-deoxyheptonate aldolase; 2-dehydro-3-deoxyphosphoheptonate aldolase; 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase	E4P + PEP → PI + 3DDAH7P	aro3		
YDR127W	4.6.1.3		aro1		pentafunctional arom polypeptide (contains: 3-dehydroquinate synthase, 3-dehydroquinate dehydratase (3-dehydroquinase), shikimate 5-dehydrogenase, shikimate kinase, and epsp synthase)	3DDAH7P → DQT + PI	aro1_1		
YDR127W	4.2.1.10		aro1		3-Dehydroquinate dehydratase	DQT → DHSK	aro1_2		
YDR127W	1.1.1.25		aro1		Shikimate dehydrogenase	DHSK + NADPH → SME + NADP	aro1_3		

(continued)

Phenylalanine, tyrosine and tryptophan biosynthesis (Aromatic Amino Acids)

		SME + ATP → ADP + SME5P aro1_4
		SME5P + PEP → 3PSME + PI aro1_5
YDR127W	2.7.1.71	Shikimate kinase I, II
YDR127W	2.5.1.19	3-Phosphoshikimate-1-carboxyvinyltransferase
YGL148W	4.6.1.4	Chorismate synthase
YPR060C	5.4.99.5	Chorismate mutase
YNL316C	4.2.1.51	prephenate dehydratase
YHR137W	2.6.1.-	putative aromatic amino acid aminotransferase
YBR166C	1.3.1.13	Prephenate dehydrogenase (NADP+)
YGL202W	2.6.1.-	aromatic amino acid aminotransferase I
YHR137W	2.6.1.-	aromatic amino acid aminotransferase II
-	1.3.1.12	Prephenate dehydrogenase
YER090W	4.1.3.27	Anthranilate synthase
YKL211C	4.1.3.27	Anthranilate synthase
YDR354W	2.4.2.18	anthranilate phosphoribosyl transferase
YDR007W	5.3.1.24	n-(5'-phosphoribosyl)-anthranilate isomerase
YKL211C	4.1.1.48	Indoleglycerol phosphate synthase
YGL026C	4.2.1.20	tryptophan synthetase
YDR256C	1.11.1.6	catalase A
YGR088W	1.11.1.6	cytoplasmic catalase T
YKL106W	2.6.1.1	Aspartate aminotransferase
YLR027C	2.6.1.1	Aspartate aminotransferase
YMR170C	1.2.1.5	Cytosolic aldehyde dehydrogenase
YMR169C	1.2.1.5	strong similarity to aldehyde dehydrogenase
YOR374W	1.2.1.3	mitochondrial aldehyde dehydrogenase
YOR374W	1.2.1.3	mitochondrial aldehyde dehydrogenase
		ACALm + NADm → NADHm + ACm
		ACALm + NADPm → NADPHm + ACm

		(continued)
Phenylalanine, tyrosine and tryptophan biosynthesis (Aromatic Amino Acids)		
YER073W	1.2.1.3	ALD5 mitochondrial Aldehyde Dehydrogenase
YPL061W	1.2.1.3	ALD6 Cytosolic Aldehyde Dehydrogenase
YJR078W	1.13.11.11	YJR078 W Protein with similarity to indoleamine 2,3-dioxygenases, which catalyze conversion of tryptophan and other indole derivatives into kynurenes, Tryptophan 2,3-dioxygenase
	-	YLR231 C Kynurene formamidase
	3.5.1.9	YBL098 W probable kynureninase (L-kynurene hydrolyase)
	3.7.1.3	YBL098 W Kynurene 3-hydroxylase, NADPH-dependent flavin monooxygenase that catalyzes the hydroxylation of kynurene to 3-hydroxykynurene in tryptophan degradation and nicotinic acid synthesis, Kynurene 3-monooxygenase
	3.7.1.3	YLR231 C probable kynureninase (L-kynurene hydrolyase)
	1.14.13.9	BNA1 3-hydroxyanthranilate 3,4-dioxygenase (3-HAO) (3- hydroxyanthranilic acid dioxygenase) (3-hydroxyanthranilate hydroxylase oxygenase) (3-hydroxyanthranilate decarboxylase)
	3.7.1.3	YLR231 C 3-hydroxyanthranilatehydroxyanthranilic acid dioxygenase (3-acid dioxygenase) (3-hydroxyanthranilate oxygenase)
	1.13.11.6	YJR025C Picolinic acid decarboxylase
	4.1.1.45	-
	1.2.1.32	-
	1.5.1.-	-
	1.3.11.27	-
	1.13.11.5	-
	5.2.1.2	-
	3.7.1.2	-
		Homogenisate 1,2-dioxygenase
		Maleyl-acetoacetate isomerase
		Fumarylacetoacetate
		trydego
		tyrdego
		trydego
		CO2
		HOMOGEN + O2 → MACAC
		MACAC → FUACAC
		FUACAC → FUM + ACTAC

(continued)

Phenylalanine, tyrosine and tryptophan biosynthesis (Aromatic Amino Acids)

YDR268w	6.1.1.2	MSW1	tryptophanyl-tRNA synthetase, mitochondrial	ATPm + TRPm + TRNAm → ms w1
YDR242W	3.5.1.4	AMD2	putative amidase	AMPm + PPIm + TRPTRNAm
YDR242W	3.5.1.4	AMD2	putative amidase	PAD → PAC + NH3
-	2.6.1.29	-	Diamine transaminase	IAD → IAC + NH3
-	1.5.3.11	-	Polyamine oxidase	SPRMD + ACCOA → ASPERMD + COA
-	1.5.3.11	-	Polyamine oxidase	ASPERMD + O2 → APRUT + sprb
-	2.6.1.29	-	Diamine transaminase	APROA + H2O2
-	1.5.3.11	-	Polyamine oxidase	APRUT + O2 → GABAL + sprc
YDR242W	3.5.1.4	AMD2	Polyamine oxidase	APROA + H2O2
-	2.6.1.29	-	Diamine transaminase	SPRM + ACCOA → ASPRM + sprd
-	1.5.3.11	-	Polyamine oxidase	COA
YDR268w	6.1.1.2	MSW1	Polyamine oxidase	ASPRM + O2 → ASPERM D + spre
YDR268w	6.1.1.2	MSW1	Polyamine oxidase	APROA + H2O2

Proline biosynthesis

YDR300C	2.7.2.11	pro1	gamma-glutamyl kinase, glutamate kinase	GLU + ATP → ADP + GLUP pro1
YOR323C	1.2.1.41	PRO2	gamma-glutamyl phosphate reductase	GLUP + NADH → NAD + PI + pro2_1
YOR323C	1.2.1.41	pro2	gamma-glutamyl phosphate reductase	GLUGSAL
-	-	-	spontaneous conversion (Strathern)	GLUP + NADPH → NADP + PI pro2_2
YER023W	1.5.1.2	pro3	spontaneous conversion (Strathern)	+ GLUGSAL
YER023W	1.5.1.2	pro3	Pyrroline-5-carboxylate reductase	GLUGSAL ↔ P5C
YER023W	1.5.1.2	pro3	Pyrroline-5-carboxylate reductase	GLUGSALm ↔ P5Cm
YER023W	1.5.1.2	pro3	Pyrroline-5-carboxylate reductase	P5C + NADPH → PRO + NADP
YER023W	1.5.1.2	pro3	Pyrroline-5-carboxylate reductase	NADP
YER023W	1.5.1.2	pro3	Pyrroline-5-carboxylate reductase	PHC + NADPH → HPRO + NADP
YER023W	1.5.1.2	pro3	Pyrroline-5-carboxylate reductase	PHC + NADH → HPRO + NAD
YLR142W	1.5.3.-	PUT1	Proline oxidase	PROm + NADm → P5Cm + NADHm

Metabolism of Other Amino Acids
beta-Alanine metabolism

YER073W	1.2.1.3	ALD5	aldehyde dehydrogenase, mitochondrial 1 mitochondrial Aldehyde Dehydrogenase	GABALm + NADm → GABAam + NADHm LACALm + NADm ↔ LLACm + NADHm	ald1 aldo5_2
Cyanoamino acid metabolism					
YJL126W	3.5.5.1	NIT2		APROP → ALA + NH3	nit2_1
YJL126W	3.5.5.1	NIT2		ACYBUT → GLU + NH3	nit2_2
Proteins, Peptides and Aminoacids Metabolism					
YLR195C	2.3.1.97	nmt1		TCOA + GLP → COA + TGLP	nmt1
YDL040C	2.3.1.88	nat1		ACCOA + PEPD → COA + APEP	nat1
YGR147C	2.3.1.88	NAT2		ACCOA + PEPD → COA + APEP	nat2
Glutathione Biosynthesis					
YJL101C	6.3.2.2	GSH1	gamma-glutamylcysteine synthetase	CYS + GLU + ATP → GC + PI	gsh1
YOL049W	6.3.2.3	GSH2	Glutathione Synthetase	GLY + GC + ATP → RGTT + PI	gsh2
YBR244W	1.11.1.9	GPX2	Glutathione peroxidase	+ ADP	
YIR037W	1.11.1.9	HYR1	Glutathione peroxidase	2 RGTT + H2O2 ↔ OGT	gpx2
YKL026C	1.11.1.9	GPX1	Glutathione peroxidase	2 RGTT + H2O2 ↔ OGT	hyr1
YPL091W	1.6.4.2	GLR1	Glutathione oxidoreductase	2 RGTT + H2O2 ↔ OGT	gpx1
YLR299W	2.3.2.2	ECM38	gamma-glutamyltranspeptidase	NADPH + OGT → NADP + RGTT	glr1
Metabolism of Complex Carbohydrates					
Starch and sucrose metabolism					
YGR032W	2.4.1.34	GSC2	1,3-beta-Glucan synthase	UDPG → 13GLUCAN + UDP	gsc2
YLR342W	2.4.1.34	FKS1	1,3-beta-Glucan synthase	UDPG → 13GLUCAN + UDP	fks1
YGR306W	2.4.1.34	FKS3	Protein with similarity to Fks1p and Gsc2p	UDPG → 13GLUCAN + UDP	fks3
YDR261C	3.2.1.58	exg2	Exo-1,3-beta-glucanase	13GLUCAN → GLC	exg2

(continued)

Metabolism of Complex Carbohydrates**Starch and sucrose metabolism**

YGR282C	3.2.1.58	BGL2	Cell wall endo-beta-1,3-glucanase	13GLUCAN → GLC
YLR300W	3.2.1.58	exg1	Exo-1,3-beta-glucanase	13GLUCAN → GLC
YOR190W	3.2.1.58	spr1	sporulation-specific exo-1,3-beta-glucanase	13GLUCAN → GLC

Glycoprotein Biosynthesis / Degradation

YMR013C	2.7.1.108	sec59	Dolichol kinase	CTP + DOL → CDP + DOLP
YPR183W	2.4.1.83	DPM1	Dolichyl-phosphate beta-D-mannosyltransferase	GDPMAN + DOLP → GDP + DPM1
YAL023C	2.4.1.109	PMT2	Dolichyl-phosphate-mannose-protein mannosyltransferase	DOLMANP → DOLP + PMT2
YDL093W	2.4.1.109	PMT5	Dolichyl-phosphate-mannose-protein mannosyltransferase	DOLMANP → DOLP + PMT5
YDL095W	2.4.1.109	PMT1	Dolichyl-phosphate-mannose-protein mannosyltransferase	DOLMANP → DOLP + PMT1
YGR199W	2.4.1.109	PMT6	Dolichyl-phosphate-mannose-protein mannosyltransferase	DOLMANP → DOLP + PMT6
YJR143C	2.4.1.109	PMT4	Dolichyl-phosphate-mannose-protein mannosyltransferase	DOLMANP → DOLP + PMT4
YOR321W	2.4.1.109	PMT3	Dolichyl-phosphate-mannose-protein mannosyltransferase	DOLMANP → DOLP + PMT3
YBR199W	2.4.1.131	KTR4	Glycolipid 2-alpha-mannosyltransferase	MAN2PD + 2 GDPMAN → 2 KTR4
			mannosyltransferase	GDP + 2MANPD
YBR205W	2.4.1.131	KTR3	Glycolipid 2-alpha-mannosyltransferase	MAN2PD + 2 GDPMAN → 2 KTR3
YDR483W	2.4.1.131	kre2	Glycolipid 2-alpha-mannosyltransferase	GDP + 2MANPD
YJL139C	2.4.1.131	yur1	Glycolipid 2-alpha-mannosyltransferase	MAN2PD + 2 GDPMAN → 2 kre2
YKR061 W	2.4.1.131	KTR2	Glycolipid 2-alpha-mannosyltransferase	GDP + 2MANPD
YOR099W	2.4.1.131	KTR1	Glycolipid 2-alpha-mannosyltransferase	MAN2PD + 2 GDPMAN → 2 yur1
YPL053C	2.4.1.131	KTR6	Glycolipid 2 alpha-manno syltransferase	GDP + 2MANPD
				MAN2PD + 2 GDPMAN → 2 KTR1
				MAN2PD + 2 GDPMAN → 2 KTR6
				GDP + 2MANPD

(continued)

5	10	15	20	25	30	35	40	45	50	55
(continued)										
Aminosugars metabolism										
YER062C	3.1.3.21									
YIL053W	3.1.3.21									
YLR307W	3.5.1.41									
YLR308W	3.5.1.41									
Metabolism of Complex Lipids										
Glycerol (Glycerolipid metabolism)										
YFL053W	2.7.1.29									
YML070W	2.7.1.29									
YDL022W	1.1.1.8									
YOL059W	1.1.1.8									
YHL032C	2.7.1.30									
YIL155C	1.1.99.5									
Thiamine (Vitamin B1) metabolism										
YOR143C	2.7.6.2									
YOR143C	2.7.6.2									
-	-									
YOL055C	2.7.1.49									
Metabolism of Cofactors, Vitamins, and Other Substances										
Thiamine (Vitamin B1) metabolism										
THI80										
THI80										
THI10										
Thiamin pyrophosphokinase										
Thiamin pyrophosphokinase										
thiC protein										
Bipartite protein consisting of N-terminal										
hydroxymethylpyrimidine phosphate										
(HMP-P) kinase domain, needed for										
thiamine biosynthesis, fused to C-										
terminal Pet1 8p-like domain of										
indeterminant function										
ATP + THIAMIN → AMP + TPP thi80_1										
ATP + TPP → AMP + TPP thi80_2										
AIR → AHM thiC										
AHM + ATP → AHMP + ADP thi20										

Metabolism of Cofactors, Vitamins, and Other Substances**Thiamine (Vitamin B1) metabolism**

YPL258C	2.7.1.49	THI21	Bipartite protein consisting of N-terminal hydroxymethylpyrimidine phosphate (HMP-P) kinase domain, needed for thiamine biosynthesis, fused to C-terminal Pet18p-like domain of indeterminant function	AHM + ATP → AHMP + ADP	thi21
YPR121 W	2.7.1.49	THI22	Bipartite protein consisting of N-terminal hydroxymethylpyrimidine phosphate (HMP-P) kinase domain, needed for thiamine biosynthesis, fused to C-terminal Pet18p-like domain of indeterminant function	AHM + ATP → AHMP + ADP	thi22
YOL055C	2.7.4.7	THI20	HMP-phosphate kinase	AHM + ATP → AHMPPP + ADP	thid
-	-	-	Hypothetical thiG protein	T3P1 + PYR → DTP DTP + TYR + CYS → THZ + HBA + CO2	unkrxn1
-	-	-	thiE protein	DTP + TYR + CYS → THZ + HBA + CO2	thie
-	-	-	thiF protein	DTP + TYR + CYS → THZ + HBA + CO2	thif
-	-	-	thiH protein	DTP + TYR + CYS → THZ + HBA + CO2	thih
YPL214C	2.7.1.50	THI6	Hydroxyethylthiazole kinase	THZ + ATP → THZP + ADP	thim
YPL214C	2.5.1.3	THI6	TMF pyrophosphorylase, hydroxyethylthiazole kinase	THZP + AHMPPP → THMP + PPI	thi6
-	2.7.4.16	-	Thiamin phosphate kinase (DL)-glycerol-3-phosphatase 2	THMP + ATP ↔ TPP + ADP	thiI
-	3.1.3.-	-	THMP → THIAMIN + PI	THMP → THIAMIN + PI	unkrxn8
Riboflavin metabolism		rib1	GTP cyclohydrolase II	GTP → D6RP5P + FOR + PP1	rib1
YBL033C	3.5.4.25	RIB7	HTP reductase, second step in the riboflavin biosynthesis pathway	D6RP5P → A6RP5P + NH3	ribd1
YBR153W	3.5.4.26	-	Pyrimidine reductase	A6RP5P + NADPH → A6RP5P2 + NADP	rib7
YBR153W	1.1.1.193	rib7	-	-	-

(continued)

Riboflavin metabolism - - YBR256C YOL143C YAR071W YDR236C YDR236C YDL045C	55 50 45 40 35 30 25 20 15 10 5	(continued) Pyrimidine phosphatase 3,4 Dihydroxy-2-butanoate-4-phosphate synthase Riboflavin biosynthesis pathway enzyme, 6,7-dimethyl-8-ribityllumazine synthase, alpha chain Riboflavin biosynthesis pathway enzyme, 6,7-dimethyl-8-ribityllumazine synthase, beta chain Acid phosphatase Riboflavin kinase Riboflavin kinase FAD synthetase FAD synthetase	A6RP5P2 → A6RP + PI RL5P → DB4P + FOR DB4P + A6RP → D8RL + PI FMN → RIBFLAV + PI RIBFLAV + ATP → FMN + ADP RIBFLAVm + ATPm → FMNm + ADPm FMN + ATP → FAD + PPI FMNm + ATPm → FAdm + PPIm	prm ribb rib5 pho11 fmn1_1 fmn1_2 fad1 fad1b
Vitamin B6 (Pyridoxine) Biosynthesis metabolism - - - YBR035C YBR035C YBR035C YBR035C YBR035C YOR184W YCR053W	2.7.1.35 2.7.1.35 2.7.1.35 1.4.3.5 1.4.3.5 1.4.3.5 1.4.3.5 1.4.3.5 1.4.3.5 2.6.1.52 4.2.99.2 3.1.3.-	Pyridoxine kinase Pyridoxine kinase Pyridoxine kinase Pyridoxine 5'-phosphate oxidase Pyridoxine 5'-phosphate oxidase Pyridoxine 5'-phosphate oxidase Pyridoxine 5'-phosphate oxidase Pyridoxine 5'-phosphate oxidase Hypothetical transaminase/ phosphoserine transaminase Threonine synthase Hypothetical Enzyme	PYRDX + ATP → P5P + ADP PDLA + ATP → PDLA5P + ADP PL + ATP → PL5P + ADP PDLA5P + 02 → PL5P + H2O2 + NH3 P5P + 02 ↔ PL5P + H2O2 PYRDX + 02 ↔ PL + H2O2 PL + 02 + NH3 ↔ PDLA + H2O2 PDLA5P + 02 → PL5P + H2O2 + NH3 OHB + GLU ↔ PHT + AKG PHT → 4HILT + PI PDLA5P → PDLA + PI	pdxka pdxkb pdxkc pdx3_1 pdx3_2 pdx3_3 pdx3_4 pdx3_5 ser1 thr4 hor2b

55	50	45	40	35	30	25	20	15	10
YBR176W	2.1.2.11	ECM31	Ketopentoate hydroxymethyl transferase	OVAL + METTHF → AKP + THF	ecm31				5
YHR063C	1.1.1.169	PAN5	Putative ketopantoate reductase (2-dehydropantoate 2-reductase) involved in coenzyme A synthesis, has similarity to Cbs2p; Ketopantoate reductase	AKP + NADPH → NADP + PANT	pane				
YLR355C	1.1.1.86	ilv5	Ketol-acid reductoisomerase	AKPm + NADPHm → NADPm + PANNm	ilv5_3				
YIL145C	6.3.2.1	YIL145C	Pantoate-b-alanine ligase	PANT + bALA + ATP → AMP + PPTO + PNTO	panca				
YDR531W	2.7.1.33	YDR531 W	Putative pantothenate kinase involved in coenzyme A biosynthesis, Pantothenate kinase	PNTO + ATP → ADP + 4PNTO	coaa				
-	6.3.2.5		Phosphopantothenate-cysteine decarboxylase	4PPNTO + CTP + CYS → CMP + PPI + 4PPNCYS	pclig				
-	4.1.1.36		Phosphopantothenate-cysteine adenyllyltransferase	4PPNCYS → CO2 + 4PPNTE	pccl				
-	2.7.7.3		Phospho-pantethiene adenyllyltransferase	4PPNTE + ATP → PPI + DPCOA	patra				
-	2.7.7.3		Phospho-pantethiene adenyllyltransferase	4PPNTEM + ATPm → PPIm + DPCOAm	patranb				
-	2.7.1.24		DiphosphoCoA kinase	DPCOA + ATP → ADP + COA	dphcoaka				
-	2.7.1.24		DiphosphoCoA kinase	DPCOAm + ATPm → ADPm + COAm	dphcoakb				
-	4.1.1.11		ASPARTATE ALPHA-DECARBOXYLASE	ASP → CO2 + bALA	pancb				
YPL148C	2.7.8.7	PPT2	Acyl carrier-protein synthase, phosphopantetheine protein transferase for Acp1p	COA → PAP + ACP	acps				

(continued)

Coenzyme A Biosynthesis

55	50	45	40	35	30	25	20	15	10	5
(continued)										
NAD Biosynthesis										
YGL037C	3.5.1.19	PNC1								
YOR209C	2.4.2.11	NPT1								
	1.4.3.-	Aspartate oxidase								
YFR047C	1.4.3.16	QPT1								
	2.4.2.19	Quinolate synthase								
YLR328W	2.7.7.18	YLR328 W								
YHR074W	6.3.5.1	QNS1								
YJR049C	2.7.1.23	utrl								
YEL041W	2.7.1.23	YEL041 W								
YPL188W	2.7.1.23	POS5								
	3.1.2.-	NADP phosphatase								
	3.2.2.5	NAD kinase, POLYPHOSPHATE KINASE (EC 2.7.4.1) / NAD+ KINASE (EC 2.7.1.23)								
	2.4.2.1	strong similarity to purine-nucleoside phosphorylases								
	2.4.2.1	strong similarity to purine-nucleoside phosphorylases								
Nicotinic Acid synthesis from TRP										
YFR047C	2.4.2.19	QPT1								
YLR328W	2.7.7.18	YLR328 W								
YLR328W	2.7.7.18	YLR328 W								
YHR074W	6.3.5.1	QNS1								

55	50	45	40	35	30	25	20	15	10	5
(continued)										
Nicotinic Acid synthesis from TRP										
YJR049c	2.7.1.23	utrl								
YPL188W	2.7.1.23	POSS								
YEL041W	2.7.1.23	YEL041 w								
-	3.1.2.-									
YLR209C	2.4.2.1	PNP1								
YLR209C	2.4.2.1	PNP1								
YGL037C	3.5.1.19	PNC1								
YOR209C	2.4.2.11 3.2.2.5	NPT1								
Uptake Pathways										
Porphyrin and Chlorophyll Metabolism										
YDR232W	2.3.1.37	hem1								
YGL040C	4.2.1.24	HEM2								
YDL205C	4.3.1.8	HEM3								
YOR278W	4.2.1.75	HEM4								
YDR047W	4.1.1.37	HEM12								
YDR044W	1.3.3.3	HEM13								
YER014W	1.3.3.4	HEM14								
YOR176W	4.99.1.1	HEM15								
YGL245W	6.1.1.17	YGL245 W								
YOL033W	6.1.1.17	MSE1								
YKR069W	2.1.1.107	met1								

					5
					10
					15
					20
					25
					30
					35
				(continued)	40
					45
					50
					55
Quinone Biosynthesis	4.1.3.27	trp3	anthranilate synthase ComponentII and indole-3-phosphate (multifunctional enzyme)	CHOR → 4HBZ + PYR	trp3_3
YKL211C	4.1.3.27	trp2	BET2	I CHOR → 4HBZ + PYR 4HBZ + NPP → N4HBZ + PPI	trp2_2 beta2
YER090W	4.1.3.27	2.5.1.-	BET4	geranylgeranyltransferase type II beta subunit	geranylgeranyltransferase type II alpha
YPR176C	2.5.1.-	cdc43	geranylgeranyltransferase type I beta subunit	geranylgeranyltransferase type II alpha	beta1
YJL031C	2.5.1.-	YGL155W	2.5.1.-	CDC43	Hexasprenylpyrophosphate synthetase, catalyzes the first step in coenzyme Q (ubiquinone) biosynthesis pathway
YBR003W	2.5.1.-	COQ1	COQ2	para-hydroxybenzoate-- polypropenyltransferase	4HBZ + NPP → N4HBZ + PPI
YNR041C	2.5.1.-	YPL172C	2.5.1.-	COX10	protoheme IX farnesylyltransferase, mitochondrial precursor
YDL090C	2.5.1.-	YKL019W	2.5.1.-	ram1	4HBZ + NPP → N4HBZ + PPI
YBR002C	2.5.1.-	RER2	RAM2	protein farnesylyltransferase alpha	ram1
YMR101C	2.5.1.-	SRT1	RER2	putative dehydrodolichyl diposphosphate	4HBZ + NPP → N4HBZ + PPI
YDR538W	4.1.1.-	PAD1	SRT1	putative dehydrodolichyl diposphosphate	4HBZ + NPP → N4HBZ + PPI
-	1.13.14.-	-	COQ5	Octaprenyl-hydroxybenzoate decarboxylase	N4HBZ → CO2 + 2NPPP
YPL266W	2.1.1.-	DIM1	-	2-Octaprenylphenol hydroxylase	pad1_2
-	1.14.13.-	-	-	2-Octaprenyl-6-methoxyphenol hydroxylase	-
YML110C	2.1.1.-	-	-	2-Octaprenyl-6-methoxy-1,4- benzoquinone methylase	ubio
YGR255C	1.14.13.-	-	-	COQ6 monooxygenase	dim1
YOL096C	2.1.1.64	-	-	3-Dimethylubiquinone 3- methyltransferase	coq6b
-	-	-	-	2NPMBm + SAMm → 2NPMMBm + SAHm	coq6
-	-	-	-	2NPMBm + O2m → 2NMHM	2NMHM
-	-	-	-	2NMHM + SAMm → QH2m + SAHm	ubiq

5
10
15
20
25
30
35
40
45
50
55

(continued)

Membrane Transport**Mitochondrial Membrane Transport**

The following diffuse through the inner mitochondrial membrane in a non-carrier-mediated manner:

O2 ↔ O2m	mco2
CO2 ↔ CO2m	meth
ETH ↔ ETHm	mnh3
NH3 ↔ NH3m	mmthn
MTHN ↔ MTHNm	mtfh
THFm ↔ THF	mmthf
METTHFm ↔ METTHF	mser
SERm ↔ SER	mgly
GLYm ↔ GLY	mcoh
CBHCAPm ↔ CBHCAP	moicap
OICAPm ↔ OICAP	mpro
PROm ↔ PRO	mcmp
CMPm ↔ CMP	mac
ACm ↔ AC	macar_
ACAR → ACARm	mcar_
CARm → CAR	maclac
ACLAC ↔ ACLACm	mactc
ACTAC ↔ ACTACm	msif
SLF → SLFm + Hm	mtfr
THRm ↔ THR	makA
AKAm → AKA	aac1
ADP + ATPm + PI → Hm +	
ADPm + ATP + Plm	
pet9	pet9
ADP/ATP carrier protein (MCF)	
YBL030C	
ADP/ATP carrier protein (MCF)	
YBR085w	
ADP/ATP carrier protein (MCF)	
MIR1	aac3
YER053 C	
phosphate carrier	
similarity to <i>C.elegans</i> mitochondrial	
phosphate carrier	
dicarboxylate carrier	
DIC1	
YJR077C	
YER053C	
YLR348C	
YLR348C	
YAL + SUCCm ↔ MAlm +	
SUCC	
MAL + Plm ↔ MAlm + PI	
DIC1_1	
DIC1_2	

5 10 15 20 25 30 35 40 45 50 55

Memberane Transport

Mitochondrial Membrane Transport

The following diffuse through the inner mitochondrial membrane in a non-carrier-mediated manner:

YLR348C	DIC1	SUCC + Plm → SUCCm + Pi dic1_3 mmmt
YKL120W	OAC1	MALT + Plm ↔ MALTm + Pi OA ↔ OAm + Hm
YBR291C	CTP1	CIT + MALM ↔ CITm + MAL CT + PEPm ↔ CITm + PEP
YBR291 C	CTP1	CIT + ICITm ↔ CITm + ICIT
YBR291C	CTP1	IPPMAL ↔ IPPMALm LAC ↔ LACm + Hm
		pyrca
		gca
		gcb
		ortl
		crc1
		moival
		momval
		mfad
		RIBFLAV ↔ RIBFLAVm DTB ↔ DTBm H3MCOA ↔ H3MCOAm MVL ↔ MVLm
		PA ↔ PAm 4PPNTE ↔ 4PPNTEM AD ↔ ADm PRPP ↔ PRPPm DHF ↔ DHFm QA ↔ QAm OPP ↔ OPPm SAM ↔ SAMm SAH ↔ SAHm
		mrbo mrdb mmcoa mrml mpa mppnt mad mrpp mdhf mqa mopp msam msah

(continued)

5

10

15

20

25

30

35

40

45

50

55

Memberane Transport**Mitochondrial Membrane Transport***The following diffuse through the inner mitochondrial membrane in a non-carrier-mediated manner:*

YJR095W

SFC1

The following diffuse through the inner mitochondrial membrane in a non-carrier-mediated manner:

SUCC + FUMm → SUCCm + FUM

sfc1

(continued)

Mitochondrial membrane
succinate-fumarate
transporter, member of the
mitochondrial carrier family
(MCF) of membrane
transporters

YPL134C ODC1 AKGm + OXA ↔ AKG + OXAm
YOR222W ODC2 AKGm + OXA ↔ AKG + OXAm

*Malate Aspartate Shuttle**Included elsewhere**Glycerol phosphate shuttle*

T3P2m → T3P2
GL3P → GL3Pm

Plasma Membrane Transport**Carbohydrates**

YHR092c HXT4

YLR081w GAL2

YOL156w HXT11

YDR536w stl1

YHR094c hxt1

YOL156w HXT11

YEL069c HXT13

YDL245c HXT15

YJR158w HXT16

YFL011w HXT10

YNR072w HXT17

moderate- to low-affinity
glucose transporter
galactose (and glucose)
permease
low affinity glucose transport
protein
Protein member of the hexose
transporter family
High-affinity hexose (glucose)
transporter
Glucose permease
high-affinity hexose
transporter
Hexose transporter
hexose permease
high-affinity hexose
transporter
Putative hexose transporter

hxt4

gal2_3

hxt11

stl1_1

hxt1_1

hxt11_1

hxt13_1

hxt15_1

hxt16_1

hxt10_1

hxt17_1

5

10

15

20

25

30

35

45

50

55

Plasma Membrane Transport**Carbohydrates**

YMR011w	HXT2	high affinity hexose transporter-2	GLCxt → GLC	hxt2_1
YHR092c	hxt4	High-affinity glucose transporter	GLCxt → GLC	hxt4_1
YDR345c	hxt3	Low-affinity glucose transporter	GLCxt → GLC	hxt3_1
YHR096c	HXT5	hexose transporter	GLCxt → GLC	hxt5_1
YDR343c	HXT6	Hexose transporter	GLCxt → GLC	hxt6_1
YDR342c	HXT7	Hexose transporter	GLCxt → GLC	hxt7_1
YJL214w	HXT8	hexose permease	GLCxt → GLC	hxt8_4
YJL219w	HXT9	hexose permease	GLCxt → GLC	hxt9_1
YLR081w	gal2	galactose permease	GLACxt + HEXT → GLAC	gal2_1
YFL011w	HXT10	high-affinity hexose transporter	GLACxt + HEXT → GLAC	hxt10_4
YOL156w	HXT11	Glucose permease	GLACxt + HEXT → GLAC	hxt11_4
YNL318c	HXT14	Member of the hexose transporter family	GLACxt + HEXT → GLAC	hxt14
YJL219w	HXT9	hexose permease	GLACxt + HEXT → GLAC	hxt9_4
YDR536W	stl1	Protein member of the hexose transporter family	GLACxt + HEXT → GLAC	stl1_4
YFL055w	AGP3	Amino acid permease for serine, aspartate, and glutamate	GLUxt + HEXT ↔ GLU	agp3_3
YDR536W	stl1	Protein member of the hexose transporter family	GLUxt + HEXT ↔ GLU	stl1_2
YKR039W	gap1	General amino acid permease	GLUxt + HEXT ↔ GLU	gap8
YCL025C	AGP1	Amino acid permease for most neutral amino acids	GLUxt + HEXT ↔ GLU	gap24
YPL265W	DIP5	Dicarboxylic amino acid permease	GLUxt + HEXT ↔ GLU	dip10
YDR536W	stl1	Protein member of the hexose transporter family	GLUxt + HEXT ↔ GLU	stl1_3
YHR094c	hxt1	High-affinity hexose (glucose) transporter	FRUxt + HEXT → FRU	hxt1_2

(continued)

Plasma Membrane Transport**Carbohydrates**

YFL011w

HXT10

high-affinity hexose transporter

Glucose permease
high-affinity hexose transporter

Hexose transporter

hexose permease
Putative hexose transporter

high affinity hexose transporter-2

Low-affinity glucose transporter

High-affinity glucose transporter

hexose transporter

Hexose transporter

Hexose transporter

hexose permease
hexose permease

High-affinity hexose (glucose) transporter

high-affinity hexose transporter

Glucose permease
high-affinity hexose transporter

Hexose transporter

hexose permease
Putative hexose transporter

high affinity hexose transporter-2

Low-affinity glucose transporter

(continued)

FRUxt + HEXT → FRU

FRUxt + HEXT → MAN

MANxt + HEXT → MAN

hxt10_2

hxt11_2
hxt13_2

hxt15_2

hxt16_2
hxt17_2

hxt2_2

hxt3_2

hxt4_2

hxt5_2

hxt6_2

hxt7_2

hxt8_5

hxt9_2

hxt1_5

hxt10_3

hxt11_3
hxt13_3

hxt15_3

hxt16_3
hxt17_3

hxt2_3

hxt3_3

5

10

15

20

25

30

40

45

50

55

Plasma Membrane Transport Carbohydrates

YHR092c

hxt4

YHR096c

HXT5

YDR343c

HXT6

YDR342c

HXT7

YJL214w

HXT8

YJL219w

HXT9

YDR497c

ITRI

YOL103w

ITR2

YIL162W 3.2.1.26

SUC2

hxt4_3

High-affinity glucose transporter

hexose transporter

Hexose transporter

Hexose transporter

hexose permease

hexose permease

myo-inositol transporter

myo-inositol transporter

Maltase permease

invertase (sucrose hydrolyzing enzyme)

sucrose

Dicarboxylates

a-Ketoglutarate/malate translocator

a-methylglucoside

Sorbose

Arabinose (low affinity)

Fucose

Glucitol

Glucosamine

Glycerol

Lactate transport

Mannitol

Melibiose

N-Acetylglucosamine

Rhamnose

Ribose

Trehalose

XYLxt ↔ XYL

MANxt + HEXT → MAN

Mixt + HEXT → MI

Mixt + HEXT → MI

MLTxt + HEXT → MLT

SUCxt → GLCxt + FRUxt

SUCxt + HEXT → SUC

MALxt + HEXT ↔ MAL

MALxt + AKG ↔ MAL + AKGxt

AMGxt ↔ AMG

SORxt ↔ SOR

ARABxt ↔ ARAB

FUCxt + HEXT ↔ FUC

GLTLxt + HEXT → GLTL

GLTxt + HEXT → GLT

GLAMxt + HEXT ↔ GLAM

GLxt ↔ GL

LACxt + HEXT ↔ LAC

MNTxt + HEXT → MNT

MEIxt + HEXT → MEI

NAGxt + HEXT → NAG

RMNxt + HEXT → RMN

RIBxt + HEXT → RIB

TRExt + HEXT → TRE

TRExt → AA TRE6P

xyup

(continued)

Amino Acids

YKR039W	gap1	General amino acid permease	ALAxt + HEXT \leftrightarrow ALA	gap1_1
YPL265W	DIP5	Dicarboxylic amino acid permease	ALAxt + HEXT \leftrightarrow ALA	dip5
YCL025C	AGP1	Amino acid permease for most neutral amino acids	ALAxt + HEXT \leftrightarrow ALA	gap25
YOL020W	TAT2	Tryptophan permease	ALAxt + HEXT \leftrightarrow ALA	tat5
YOR348C	PUT4	Proline permease	ALAxt + HEXT \leftrightarrow ALA	put4
YKR039W	gap1	General amino acid permease	ARGxt + HEXT \leftrightarrow ARG	gap2
YEL063C	can1	Permease for basic amino acids	ARGxt + HEXT \leftrightarrow ARG	can1_1
YNL270C	ALP1	Protein with strong similarity to permeases	ARGxt + HEXT \leftrightarrow ARG	alp1
YKR039W	gap1	General amino acid permease	ASNxt + HEXT \leftrightarrow ASN	gap3
YCL025C	AGP1	Amino acid permease for most neutral amino acids	ASNxt + HEXT \leftrightarrow ASN	gap21
YDR508C	GNP1	Glutamine permease (high affinity)	ASNxt + HEXT \leftrightarrow ASN	gnp2
YPL265W	DIP5	Dicarboxylic amino acid permease	ASNxt + HEXT \leftrightarrow ASN	dip6
YFL055W	AGP3	Amino acid permease for serine, aspartate, and glutamate	ASPxt + HEXT \leftrightarrow ASP	agp3_2
YKR039W	gap1	General amino acid permease	ASPxt + HEXT \leftrightarrow ASP	gap4
YPL265W	DIP5	Dicarboxylic amino acid permease	ASPxt + HEXT \leftrightarrow ASP	dip7
YKR039W	gap1	General amino acid permease	CYSxt + HEXT \leftrightarrow CYS	gap5
YDR508C	GNP1	Glutamine permease (high affinity)	CYSxt + HEXT \leftrightarrow CYS	gnp3
YBR068C	BAP2	Branched chain amino acid permease	CYSxt + HEXT \leftrightarrow CYS	bap2_1

(continued)

5

10

15

20

30

35

40

45

50

55

Amino Acids

YDR046C	BAP3	Branched chain amino acid permease	CYSxt + HEXT ↔ CYS	bap3_1
YBR069C	VAP1	Amino acid permease	CYSxt + HEXT ↔ CYS	vap7
YOL020W	TAT2	Tryptophan permease	CYSxt + HEXT ↔ CYS	tat7
YKR039W	gap1	General amino acid permease	GLYxt + HEXT ↔ GLY	gap6
YOL020W	TAT2	Tryptophan permease	GLYxt + HEXT ↔ GLY	tat6
YPL265W	DIP5	Dicarboxylic amino acid permease	GLYxt + HEXT ↔ GLY	dip8
YOR348C	PUT4	Proline permease	GLYxt + HEXT ↔ GLY	put5
YKR039W	gap1	General amino acid permease	GLNxt + HEXT ↔ GLN	gap7
YCL025C	AGP1	Amino acid permease for most neutral amino acids	GLNxt + HEXT ↔ GLN	gap22
YDR508C	GNP1	Glutamine permease (high affinity)	GLNxt + HEXT ↔ GLN	gnp1
YPL265W	DIP5	Dicarboxylic amino acid permease	GLNxt + HEXT ↔ GLN	dip9
YGR191W	HIP1	Histidine permease	HISxt + HEXT ↔ HIS	hip1
YKR039W	gap1	General amino acid permease	HISxt + HEXT ↔ HIS	gap9
YCL025C	AGP1	Amino acid permease for most neutral amino acids	HISxt + HEXT ↔ HIS	gap23
YBR069C	VAP1	Amino acid permease	HISxt + HEXT ↔ HIS	vap6
YBR069C	TAT1	Amino acid permease that transports valine, leucine, isoleucine, tyrosine, tryptophan, and threonine	ILExt + HEXT ↔ ILE	tat1_2
YKR039W	gap1	General amino acid permease	ILExt+ HEXT ↔ ILE	gap10
YCL025C	AGP1	Amino acid permease for most neutral amino acids	ILExt + HEXT ↔ ILE	gap32
YBR068C	BAP2	Branched chain amino acid permease	ILExt + HEXT ↔ ILE	bap2_2
YDR046C	BAP3	Branched chain amino acid permease	ILExt + HEXT ↔ ILE	bap3_2

(continued)

5

10

15

20

30

35

40

45

50

55

Amino Acids

YBR069C

YBR069C

VAP1
TAT1Amino acid permease
Amino acid permease that transports valine, leucine, isleucine, tyrosine, tryptophan, and threonine

YKR039W

gap1

General amino acid permease
Amino acid permease for most neutral amino acids

YCL025C

AGP1

gap33

Branched chain amino acid permease
Branched chain amino acid permease

YBR068C

BAP2

bap2_3

General amino acid permease
Amino acid permease

YDR046C

BAP3

bap3_3

Glutamine permease (high affinity)
General amino acid permease

YBR069C

VAP1
GNP1vap4
gnp7Glutamine permease (high affinity)
General amino acid permease

YKR039W

gap1

gap13

General amino acid permease
Amino acid permease for most neutral amino acids

YCL025C

AGP1

gap26

Glutamine permease (high affinity)
Branched chain amino acid permease

YDR508C

GNP1 I

gnp4

Branched chain amino acid permease
Branched chain amino acid permease

YBR068C

BAP2

bap2_4

High-affinity methionine permease
Low-affinity methionine permease

YDR046C

BAP3

bap3_4

General amino acid permease
Amino acid permease for most neutral amino acids

YGR055W

MUP1

mup1

Tryptophan permease
General amino acid permease

YHL036W

MUP3

mup3

TAT2

gap14

Tryptophan permease

gap29

YCL025C

AGP1

tat4

TAT2

(continued)

5

10

15

20

25

30

35

40

45

50

55

(continued)

Amino Acids

YBR068C	BAP2	Branched chain amino acid permease	PHExt + HEXT ↔ PHEN	bap2_5
YDR046C	BAP3	Branched chain amino acid permease	PHExt + HEXT ↔ PHEN	bap3_5
YKR039W	gap1	General amino acid permease	PROxt + HEXT ↔ PRO	gap15
YOR348C	PUT4	Proline permease	PROxt + HEXT ↔ PRO	put6
YBR069C	TAT1	Amino acid permease that transports valine, leucine, isoleucine, tyrosine, tryptophan, and threonine	TRPxxt + HEXT ↔ TRP	tat1_6
YKR039W	gap1	General amino acid permease	TRPxxt + HEXT ↔ TRP	gap18
YBR069C	VAP1	Amino acid permease	TRPxxt + HEXT ↔ TRP	vap2
YOL020W	TAT2	Tryptophan permease	TRPxxt + HEXT ↔ TRP	tat3
YBR068C	BAP2	Branched chain amino acid permease	TRPxxt + HEXT ↔ TRP	bap2_6
YDR046C	BAP3	Branched chain amino acid permease	TRPxxt + HEXT ↔ TRP	bap3_6
YBR069C	TAT1	Amino acid permease that transports valine, leucine, isoleucine, tyrosine, tryptophan, and threonine	TYRxxt + HEXT ↔ TYR	tat1_7
YKR039W	gap 1	General amino acid permease	TYRxxt + HEXT ↔ TYR	gap19
YCL025C	AGP1	Amino acid permease for most neutral amino acids	TYRxxt + HEXT ↔ TYR	gap28
YBR068C	BAP2	Branched chain amino acid permease	TYRxxt + HEXT ↔ TYR	bap2_7
YBR069C	VAP1	Amino acid permease	TYRxxt + HEXT ↔ TYR	vap1
YOL020W	TAT2	Tryptophan permease	TYRxxt + HEXT ↔ TYR	tat2
YDR046C	BAP3	Branched chain amino acid permease	TYRxxt + HEXT ↔ TYR	bap3_7
YKR039W	gap1	General amino acid permease	VALxt + HEXT ↔ VAL	gap20

5

10

15

20

30

35

40

45

50

55

Amino Acids

YCL025C	AGP1	Amino acid permease for most neutral amino acids Branched chain amino acid permease	VALxt + HEXT \leftrightarrow VAL	gap31
YDR046C	BAP3	Amino acid permease	VALxt + HEXT \leftrightarrow VAL	bap3_8
YBR069C	VAP1	Branched chain amino acid permease	VALxt + HEXT \leftrightarrow VAL	vap5
YBR068C	BAP2	Branched chain amino acid permease	VALxt + HEXT \leftrightarrow VAL	bap2_8
YFL055W	AGP3	Amino acid permease for serine, aspartate, and glutamate	SERxt + HEXT \leftrightarrow SER	agp3_1
YCL025C	AGP1	Amino acid permease for most neutral amino acids	SERxt + HEXT \leftrightarrow SER	gap27
YDR508C	GNP1	Glutamine permease (high affinity)	SERxt + HEXT \leftrightarrow SER	gnp5
YKR039W	gap1	General amino acid permease	SERxt + HEXT \leftrightarrow SER	gap16
YPL265W	DIP5	Dicarboxylic amino acid permease	SERxt + HEXT \leftrightarrow SER	dip11
YBR069C	TAT1	Amino acid permease that transports valine, leucine, isleucine, tyrosine, tryptophan, and threonine	THRxt + HEXT \leftrightarrow THR	tat1_1
YCL025C	AGP1	Amino acid permease for most neutral amino acids	THRxt + HEXT \leftrightarrow THR	gap30
YKR039W	gap 1	General amino acid permease	THRxt+ HEXT \leftrightarrow THR	gap17
YDR508C	GNP1	Glutamine permease (high affinity)	THRxt+ HEXT \leftrightarrow THR	gnp6
YNL268W	LYP1	Lysine specific permease (high affinity)	LYSxt + HEXT \leftrightarrow LYS	lyp1
YKR039W	gap1	General amino acid permease	LYSxt + HEXT \leftrightarrow LYS	gap12
YLL061W	MMP1	High affinity S-methylmethionine permease	MMETxt + HEXT \rightarrow MMET	mmp1

(continued)

Amino Acids
YPL274W

(continued)

YOR348C	PUT4	SAM3	High affinity S-adenosylmethionine permease	SAMxt + HEXT → SAM	sam3
YDL210W	uga4		Proline permease	GABAxt + HEXT → GABA	put7
YBR132C	AGP2		Amino acid permease with high specificity for GABA	GABAxt + HEXT → GABA	uga4
YGL077C	HNM1		Plasma membrane carnitine transporter	CARxt ↔ CAR	agp2
YNR056C	BIO5		Choline permease	CHOxt + HEXT → MET	hnm1
YDL210W	uga4		transmembrane regulator of KAPA/DAPA transport	BIOxt + HEXT → BIO	bio5a
YKR039W	gap1		Amino acid permease with high specificity for GABA	ALAVxt + HEXT → ALAV	uga5
YEI063C	can1		General amino acid permease	ORNxt + HEXT ↔ ORN	gap1b
			Permease for basic amino acids	ORNxt + HEXT ↔ ORN	can1b
			Putrescine	PTRSCxt + HEXT → PTRSC	ptrup
			Spermidine & putrescine	SPRMDDxt + HEXT → SPRMD	sprup1
			Dipeptide	DIPExxt + HEXT → DIPEP	ptr2
			Oligopeptide	OPEPxxt + HEXT → OPEP	ptr3
			Peptide	PEPTxt + HEXT → PEPT	ptr4
			Uracil	URAxxt + HEXT → URA	urau1
			Nicotinamide mononucleotide transporter	NMNxt + HEXT → NMN	nmmup
			Cytosine purine permease	CYTSxt + HEXT → CYTS	fey2_1
			Adenine	ADxt + HEXT → AD	fey2_2
			Guanine	GNxt + HEXT ↔ GN	fey2_3
			Cytosine purine permease	CYTSxt + HEXT → CYTS	fey21_1
			Adenine	ADxt + HEXT → AD	fey21_2
			Guanine	GNxt + HEXT ↔ GN	fey21_3
			Cytosine purine permease	CYTSxt + HEXT → CYTS	fey22_1
			Adenine	ADxt + HEXT → AD	fey22_2
			FCY22		
YER056C	FCY2				
YER056C	FCY2				
YER056C	FCY2				
YER060W	FCY21				
YER060W	FCY21				
YER060W	FCY21				
YER060W-	FCY22				
A					
YER060W-					
A					

5

10

15

20

25

30

35

40

45

50

55

Amino Acids

YER060W-	FCY22	Guanine A	GNxt + HEXT ↔ GN
YGL186C	YGL186 C	Cytosine purine permease	CYT Sxt + HEXT → CYTS
YGL186C	YGL186 C	Adenine	ADxt + HEXT → AD
YGL186C	YGL186 C	Guanine	GNxt + HEXT ↔ GN
		G-system	ADNxt + HEXT → ADN
		G-system	GSNxt + HEXT → GSN
		Uridine permease, G system	URIxt + HEXT → URI
		G-system	CYT Dxt + HEXT → CYTD
		G-system (transports all nucleosides)	INSxt + HEXT → INS
		G-system	XTSINExt + HEXT → XTSINE
		G-system	DTxt + HEXT → DT
		G-system	DINxt + HEXT → DIN
		G-system	DGxt + HEXT → DG
		G-system	DAXt + HEXT → DA
		G-system	DCxt + HEXT → DC
		G-system	DUXt + HEXT → DU
		C-system	ADNxt + HEXT → ADN
		Uridine permease, C-system	URIxt + HEXT → URI
		C-system	CYT Dxt + HEXT → CYTD
		C-system	DTxt + HEXT → DT
		C-system	DAXt + HEXT → DA
		C-system	DCxt + HEXT → DC
		C-system	DUXt + HEXT → DU
		Nucleosides and deoxynucleoside	ADNxt + HEXT → ADN
		Nucleosides and deoxynucleoside	GSNxt + HEXT → GSN
		Uridine permease,	URIxt + HEXT → URI
		Nucleosides and deoxynucleoside	ncup3
		Nucleosides and deoxynucleoside	NCup4
		Nucleosides and deoxynucleoside	NCup5
YBL042C	FU11		

(continued)

5

10

15

20

30

35

45

50

55

Amino Acids

(continued)

Nucleosides and deoxynucleoside	DTxt + HEXT → DT	ncup7
Nucleosides and deoxynucleoside	DINxt + HEXT → DIN	ncup8
Nucleosides and deoxynucleoside	DGxt + HEXT → DG	ncup9
Nucleosides and deoxynucleoside	DAxt + HEXT → DA	ncup10
Nucleosides and deoxynucleoside	DCxt + HEXT → DC	cup11
Nucleosides and deoxynucleoside	DUxt + HEXT → DU	ncup12
Hypoxanthine	HYXNxt + HEXT ↔ HYXN	hyxup
Xanthine	XANxt ↔ XAN	xanup
Metabolic By-Products		
YCR032W	BPH1	
	Probable acetic acid export pump, Acetate transport	acup
	Formate transport	forup
	Ethanol transport	ethup
	Succinate transport	succup
	Pyruvate lactate proton symport	jem1_1
YKL217W	JEAN1	
	UREAxt + 2 HEXT ↔ UREA	dur3
	Ammonia transport	NH3xt ↔ NH3
	Ammonia transport, low capacity high affinity	NH3xt ↔ NH3
	Ammonia transport, high capacity low affinity	MEP3
Other Compounds		
YHL016C	dur3	
YGR121C	MEP1	
YNL142W	MEP2	
YPR138C	MEP3	

5

10

15

20

25

30

35

40

45

50

55

Other Compounds

YJL129C trk1

(continued)

Potassium transporter of the plasma membrane, high affinity, member of the potassium transporter (TRK) family of membrane transporters

YBR294W
YLR092W
YGR125W
YML123C

SUL1
SUL2
YGR125W
pho84

protein

Citrate

Dicarboxylates

Fatty acid transport

a-Ketoglutarate

Putative Na+/H+ antiporter

Pantothenate

ATP drain flux for constant

maintenance requirements

H+-ATPase subunit, plasma

membrane

H+-ATPase subunit, plasma

membrane

H+-transporting P-type

ATPase, major isoform,

plasma membrane

H+-transporting P-type

ATPase, minor isoform,

plasma membrane

Glyceraldehyde transport

YGL008C

PMP1

PMP2

YPL036W

PMA1

PMA2

glalx

sul1

sul2

sulup

pho84

citup

fumup

faup1

faup2

faup3

faup4

faup5

akgup

nha1

fen2

atpm1

pmp1

pmp2

GLALxt ↔ GLAL

Kxt + HEXT ↔ K

trk1

5

10

15

20

25

30

35

40

45

50

55

Other Compounds

(continued)

YLR237W	THI7	Acetaldehyde transport	ACALxt ↔ ACAL
YOR071C	YOR071 C	Thiamine transport protein	THMxt + HEXT → THIAMIN
		Probable low affinity thiamine transporter	THMxt + HEXT → THIAMIN
YOR192C	YOR192 C	Probable low affinity thiamine transporter	THMxt + HEXT → THIAMIN
YIR028W	dal4	ATNxt → ATN	dal4
YJR152W	da15	ATTxt → ATT	da15
		MTHINxt ↔ MTHIN	mthup
		PAPxt ↔ PAP	papx
		DTPPxxt ↔ DTPP	dtpx
		THYxt ↔ THY + HEXT	thyx
		GA6Pxt ↔ GA6P	ga6up
		H+/biotin symporter and BTxt + HEXT ↔ BT	btup
YGR065C	VHT1	member of the allantioate permease family of the major facilitator superfamily	
		AONAxt + HEXT ↔ AONA	kapaup
		DANNAxt + HEXT ↔ DANNA	dapaup
		OGTxt → OGT	ogtup
		SPRMxt → SPRM	sprmup
		PIMExt → PIME	pimeup
		O2xt ↔ O2	o2tx
		CO2xt ↔ CO2	co2tx
		ERGOSTxt ↔ ERGOST	ergup
		ZYMSTxt ↔ ZYMST	zymup
		RFLAVxt + HEXT → RIBFLAV	rflup
YOR011W	AUS1		
YOR011W	AUS1		

EP 2 463 654 A1

[0055] Standard chemical names for the acronyms used to identify the reactants in the reactions of Table 2 are provided in Table 3.

5

10

15

20

25

30

35

40

45

50

55

5

10

15

20

25

30

35

40

50

55

TABLE 3

Abbreviation	Metabolite
13GLUCAN	1,3-beta-D-Glucan
13PDG	3-Phospho-D-glyceroyl phosphate
23DAACP	2,3-Dehydroacyl-[acyl]-carrier-protein]
23PDG	2,3-Bisphospho-D-glycerate
2HDACP	Hexadecenoyl-[acyl]
2MANPD	(α -D-mannosyl)(2)-"beta"-D-mannosyl-diacylglycosyldiphosphodolichol
2N6H	2-Nonaprenyl-6-hydroxyphenol
2NMHM	3-Demethylubiquinone-9M
2NMMBm	3-Demethylubiquinone-9M
2NPMBm	2-Nonaprenyl-6-methoxy-1,4-benzoquinoneM
2NPMBm	2-Nonaprenyl-3-methyl-6-methoxy-1,4-benzoquinoneM
2NPPM	2-Nonaprenyl-6-methoxyphenol
2NPPM	2-Nonaprenyl-6-methoxyphenolM
2NPPP	2-Nonaprenylphenol
2PG	2-Phospho-D-glycerate
3DDAH7P	2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate
3HPACP	(3R)-3-Hydroxypalmitoyl-[acyl]-carrier protein]
3PG	3-Phospho-D-glycerate
3PSER	3-Phosphoserine
3PSME	5-O-(1-Carboxyvinyl)-3-phosphoshikimate
4HBZ	4-Hydroxybenzoate
4HLT	4-Hydroxy-L-threonine
4HPP	3-(4-Hydroxyphenyl)pyruvate
4PPNCYS	(R)-4'-Phosphopantethenoyl-L-cysteine
4PPNTE	Pantetheine 4'-phosphate
4PPNTEm	Pantetheine 4'-phosphateM
4PPNTO	D-4'-Phosphopantethenate
5MTA	5'-Methylthioadenosine
6DGLC	D-Gal alpha 1 → 6D-Glucose
A6RP	5-Amino-6-(5'-phosphoribosylamino)uracil
A6RP5P	5-Amino-6-(5'-phosphoribitylamino)uracil
A6RP5P2	5-Amino-6-(5'-phosphoribitylamino)uracil

5

10

15

20

25

35

40

50

55

Abbreviation	Metabolite
AACCOA	Acetoacetyl-CoA
AACP	Acyl-[acyl-carrier-protein]
AA TRE6P	alpha, alpha'-Trehalose 6-phosphate
ABUTm	2-Aceto-2-hydroxy butyrateM
AC	Acetate
ACACP	Acyl-[acyl-carrier protein]
ACACPm	Acyl-[acyl-carrier protein]M
ACAL	Acetaldehyde
ACALm	AcetaldehydeM
ACAR	4-Acetyl[carnitine]
ACARM	O-Acetyl[carnitine]M
ACCOA	Acetyl-CoA
ACCOAm	Acetyl-CoAM
ACLAC	2-Acetolactate
ACLACm	2-AcetolactateM
Acm	AcetateM
ACNL	3-Indoleacetonitrile
ACOA	Acyl-CoA
ACP	Acyl-carrier protein
ACPm	Acyl-carrier proteinM
ACTAC	Acetoacetate
ACTACm	AcetoacetateM
ACYBUT	gamma-Amino-gamma-cyanobutanoate
AD	Adenine
ADCHOR	4-amino-4-deoxychorismate
Adm	AdenineM
ADN	Adenosine
ADNm	AdenosineM
ADP	ADP
ADPm	ADPM
ADPRIB	ADPRibose
ADPRIBm	ADPRiboseM
AGL3P	Acyl-sn-glycerol 3-phosphate

5

10

15

20

25

30

35

40

45

50

55

(continued)

Abbreviation	Metabolite
AHHMD	2-Amino-7,8-dihydro-4-hydroxy-6-(diphosphooxymethyl)pteridine
AHHMP	2-Amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine
AHM	4-Amino-5-hydroxymethyl-2-methylpyrimidine
AHMP	4-Amino-2-methyl-5-phosphomethylpyrimidine
AHMPP	2-Methyl-4-amino-5-hydroxymethylpyrimidine diphosphate
AHTD	2-Amino-4-hydroxy-6-(erythro-1,2,3-trihydroxypropyl)-dihydropteridine triphosphate
AICAR	1-(5'-Phosphoribosyl)-5-amino-4-imidazolecarboxamide
AIR	Aminimidazole ribotide
AKA	2-Oxoadipate
AKAm	2-OxoadipateM
AKG	2-Oxoglutarate
AKGm	2-OxoglutarateM
AKP	2-Dehydropantoate
AKPm	2-DehydropantoateM
ALA	L-Alanine
ALAGLY	R-S-Alanyl glycine
ALAm	L-AlanineM
ALAV	5-Aminolevulinate
ALAVm	5-AminolevulinateM
ALTRNA	L-Arginyl-tRNA(Arg)
AM6SA	2-Aminomuconate 6-semialdehyde
AMA	L-2-Aminoadipate
AMASA	L-2-Aminoadipate 6-semialdehyde
AMG	Methyl-D-glucoside
AMP	AMP
AMPm	AMPM
AMUCO	2-Aminomuconate
AN	Anthranilate
AONA	8-Amino-7-oxononanoate
APEP	Nalpha-Acetylpeptide
APROA	3-Aminopropanal
APROP	alpha-Aminopropiononitrile
APRUT	N-Acetylputrescine

5

10

15

20

25

35

40

50

55

Abbreviation	Metabolite
APS	Adenylylsulfate
ARAB	D-Arabinose
ARABLAC	D-Arabinono-1,4-lactone
ARG	L-Arginine
ARGSUCC	N-(L-Arginino)succinate
ASER	O-Acetyl-L-serine
ASN	L-Asparagine
ASP	L-Aspartate
ASPERMD	N1-Acetylspermidine
ASPM	L-AspartateM
ASPRM	N1-Acetylsperrmine
ASPSA	L-Aspartate 4-semialdehyde
ASPTRNA	L-Asparaginyl-tRNA(Asn)
ASPTRNAm	L-Asparaginyl-tRNA(Asn)M
ASUC	N6-(1,2-Dicarboxyethyl)-AMP
AT3P2	Acylidihydroxyacetone phosphate
ATN	Allantoin
ATP	ATP
ATPm	ATPM
ATRNA	tRNA(Arg)
A TRP	P1,P4-Bis(5'-adenosyl) tetraphosphate
ATT	Allantoate
bALA	beta-Alanine
BASP	4-Phospho-L-aspartate
bDG6P	beta-D-Glucose 6-phosphate
bDGLC	beta-D-Glucose
BIO	Biotin
BT	Biotin
C100ACP	Decanoyl-[acyl]
C120ACP	Dodecanoyl-[acyl]-carrier protein]
C120ACPm	Dodecanoyl-[acyl]-carrier protein]M
C140	Myristic acid
C140ACP	Myristoyl-[acyl]-carrier protein]

(continued)

5

10

15

20

25

35

(continued)

Abbreviation	Metabolite
C140ACPm	Myristoyl-[acyl]-carrier protein M
C141ACP	Tetradecenoyl-[acyl]-carrier protein]
C141ACPm	Tetradecenoyl-[acyl]-carrier protein M
C160	Palmitate
C160ACP	Hexadecanoyl-[acp]
C160ACPm	Hexadecanoyl-[acp]M
C161	1-Hexadecene
C161ACP	Palmitoyl-[acyl]-carrier protein]
C161ACPm	Palmitoyl-[acyl]-carrier protein M
C16A	C16_aldehydes
C180	Stearoyl-[acyl]-carrier protein]
C180ACP	Stearoyl-[acyl]-carrier protein M
C180ACPm	1-Octadecene
C181	Oleoyl-[acyl]-carrier protein]
C181ACP	Oleoyl-[acyl]-carrier protein M
C181ACPm	Linolenoyl-[acyl]-carrier protein
C182ACP	Linolenoyl-[acyl]-carrier protein M
CAASP	N-Carbamoyl-L-aspartate
CAIR	1-(5-Phospho-D-ribosyl)-5-amino-4-imidazolecarboxylate
CALH	2-(3-Carboxy-3-aminopropyl)-L-histidine
cAMP	3',5'-Cyclic AMP
CAP	Carbamoyl phosphate
CAR	Carnitine
CARM	CarnitineM
CBHCAP	3-Isopropylmalate
CBHCAPm	3-IsopropylmalateM
cCMP	3',5'-Cyclic CMP
cdAMP	3',5'-Cyclic dAMP
CDP	CDP
CDPCHO	CDPcholine
CDPDG	CDPdiacylglycerol
CDPDGm	CDPdiacylglycerolM
CDPETN	CDPEthanolamine

5

10

15

20

25

30

35

40

45

50

55

(continued)

Abbreviation	Metabolite
CER2	Ceramide-2
CER3	Ceramide-3
CGL Y	Cys-Gly
cGMP	3',5'-Cyclic GMP
CHCoA	6-Carboxyhexanoyl-CoA
CHIT	Chitin
CHITO	Chitosan
CHO	Choline
CHOR	Chorismate
cIMP	3',5'-Cyclic IMP
CIT	Citrate
CITm	CitrateM
CITR	L-Citrulline
CLm	CardiolipinM
CMP	CMP
CMPm	CMPM
CMUSA	2-Amino-3-carboxymuconate semialdehyde
CO2	CO2
CO2m	CO2M
COA	CoA
COAm	CoAM
CPAD5P	1-(2-Carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate
CPP	Coproporphyrinogen
CTP	CTP
CTPm	CTPM
CYS	L-Cysteine
CYTD	Cytidine
CYTS	Cytosine
D45P	1-Phosphatidyl-D-myo-inositol 4,5-bisphosphate
D6PGC	6-Phospho-D-glucuronate
D6PGL	D-Glucono-1,5-lactone 6-phosphate
D6RP5P	2,5-Diamino-6-hydroxy-4-(5'-phosphoribosylamino)-pyrimidine
D8RL	6,7-Dimethyl-8-(1-D-ribityl)umazine
DA	Deoxyadenosine
dADP	

5

10

15

20

25

30

35

40

50

55

(continued)

Abbreviation	Metabolite
DAGLY	Diacylglycerol
DAMP	dAMP
dAMP	dAMP
DANNA	7,8-Diaminononanoate
DAPRP	1,3-Diaminopropane
DATP	dATP
DB4P	L-3,4-Dihydroxy-2-butanone 4-phosphate
DC	Deoxycytidine
DCDP	dCDP
DCMP	dCMP
DCTP	dCTP
DFUC	alpha-D-Fucoside
DG	Deoxyguanosine
DGDP	dGDP
DGMP	dGMP
DGPP	Diacylglycerol pyrophosphate
DGTP	dGTP
DHF	Dihydrofolate
DHFm	DihydrofolateM
DHMAm	(R)-2,3-dihydroxy-3-methylbutanoateM
DHP	2-Amino-4-hydroxy-6-(D-erythro-1,2-trihydroxypropyl)-7,8-dihydropteridine
DHPP	Dihydronopterin phosphate
DHPT	Dihydropteroate
DHSK	3-Dehydroshikimate
DHSP	Sphinganine 1-phosphate
DHSPh	3-Dehydrosphinganine
DHVAlm	(R)-3-Hydroxy-3-methyl-2-oxobutanoateM
DIMGP	D-erythro-1-(imidazo-4-yl)glycerol 3-phosphate
DIN	Deoxyinosine
DIPEP	Dipeptide
DISAC1P	2,3-bis(3-hydroxytetradecanoyl)-D-glucosaminyl-1,6-beta-D-2,3-bis(3-hydroxytetradecanoyl)-beta-D-glucosaminyl 1-phosphate
DLIPOM	DihydrolipoamideM
DMPP	Dimethylallyl diphosphate

5

10

15

20

25

(continued)

Abbreviation	Metabolite
DMZYMST	4,4-Dimethylzymosterol
DOL	Dolichol
DOLMANP	Dolichyl beta-D-mannosyl phosphate
DOLP	Dolichyl phosphate
DOLPP	Dehydrololichol diphosphate
DOROA	(S)-Dihydroorotate
DPCOA	Dephospho-CoA
DPCOAm	Dephospho-CoAM
DPTH	2-[3-Carboxy-3-(methylammonio)propyl]-L-histidine
DQT	3-Dehydroquinate
DR1P	Deoxy-ribose 1-phosphate
DR5P	2-Deoxy-D-ribose 5-phosphate
DRIB	Deoxyribose
DSAM	S-Adenosylmethioninamine
DT	Thymidine
DTB	Dethiobiotin
DTBm	DethiobiotinM
DTDP	dTDP
DTMP	dTMP
DTP	1-Deoxy-d-threo-2-pentulose
DTTP	dTTP
DU	Deoxyuridine
DUDP	dUDP
DUMP	dUMP
DUTP	dUTP
E4P	D-Erythrose 4-phosphate
EPM	Epimelibiose
EPST	Episterol
ER4P	4-Phospho-D-erythronate
ERGOST	Ergosterol
ERTEOL	Ergosta-5,7,22,24(28)-tetraenol
ERTROL	Ergosta-5,7,24(28)-triol
ETH	Ethanol
ETHm	EthanolM
ETHM	Ethanolamine

5

10

15

20

25

30

35

(continued)

Abbreviation	Metabolite
F1P	D-Fructose 1-phosphate
F26P	D-Fructose 2,6-bisphosphate
F6P	beta-D-Fructose 6-phosphate
FAD	FAD
FADH2m	FADH2M
FADM	FADM
FALD	Formaldehyde
FDP	beta-D-Fructose 1,6-bisphosphate
FERIm	Ferricytochrome cM
FEROm	Ferrocytocochrome cM
FEST	Fecosterol
FGAM	2-(Formamido)-N1-(5'-phosphoribosyl)acetamidine
FGAR	5'-Phosphoribosyl-N-formylglycaminide
FGT	S-Formylglutathione
FKYN	L-Formylkynurenine
FMN	FMN
FMNm	N-Formylmethionyl-tRNAM
FOR	Formate
FORm	FormateM
FPP	trans,trans-Farnesyl diphosphate
FRU	D-Fructose
FTHF	10-Formyltetrahydrofolate
FTHfM	10-FormyltetrahydrofolateM
FUACAC	4-Fumarylacetooacetate
FUC	beta-D-Fucose
FUM	Fumarate
FUMm	FumaramM
G1P	D-Glucose 1-phosphate
G6P	alpha-D-Glucose 6-phosphate
GA1P	D-Glucosamine 1-phosphate
GA6P	D-Glucosamine 6-phosphate
GABA	4-Aminobutanate
GABAL	4-Aminobutyraldehyde
GABAIm	4-AminobutyraldehydeM

5

10

15

20

25

35

40

50

55

(continued)

Abbreviation	Metabolite
GABA _m	4-AminobutanoateM
GAL1P	D-Galactose 1-phosphate
GAR	5'-Phosphoribosylglycinamide
GBAD	4-Guanidino-butanamide
GBAT	4-Guanidino-butanooate
GC	gamma-L-Glutamyl-L-cysteine
GDP	GDP
GDPM	GDPM
GDPMAN	GDPMannose
GGL	Galactosylglycerol
GL	Glycerol
GL3P	sn-Glycerol 3-phosphate
GL3Pm	sn-Glycerol 3-phosphateM
GLAC	D-Galactose
GLACL	1-alpha-D-Galactosyl-myo-inositol
GLAL	Glycolaldehyde
GLAM	Glucosamine
GLC	alpha-D-Glucose
GLCN	Gluconate
GLN	L-Glutamine
GLP	Glycylpeptide
GLT	L-Glucitol
GLU	L-Glutamate
GLUGSAL	L-Glutamate 5-semialdehyde
GLUM	L-Glutamate 5-semialdehydeM
GLUP	GlutamateM
GLX	alpha-D-Glutamyl phosphate
GLY	Glyoxylate
	Glycine

5

10

15

20

25

30

(continued)

35

40

45

50

55

Abbreviation	Metabolite
GLYCOGEN	Glycogen
GLYm	GlycineM
GLYN	Glycerone
GMP	GMP
GN	Guanine
GNm	GuanineM
GPP	Geranyl diphosphate
GSN	Guanosine
GSNm	GuanosineM
GTP	GTP
GTPm	GTPM
GtRNA	L-Glutamyl-tRNA(Glu)
GtRNAm	L-Glutamyl-tRNA(Glu)M
GTRP	P1,P4-Bis(5'-guanosyl) tetraphosphate
H2O2	H2O2
H2S	Hydrogen sulfide
H2SO3	Sulfite
H3MCoA	(S)-3-Hydroxy-3-methylglutaryl-CoA
H3MCoAm	(S)-3-Hydroxy-3-methylglutaryl-CoAM
HACNIm	But-1-ene-1,2,4-tricarboxylateM
HACCA	(3S)-3-Hydroxyacyl-CoA
HAN	3-Hydroxyanthranilate
HBA	4-Hydroxy-benzyl alcohol
HCIT	2-Hydroxybutane-1,2,4-tricarboxylateM
HCITm	Homocysteine
HCYS	H ₊ EXT
HEXT	L-Histidyl-tRNA(His)
HHtRNA	(S)-3-Hydroxyisobutyrate
HB	(S)-3-Hydroxyisobutyryl-CoA
HIBCOA	HomoisocitrateM
HICITm	L-Histidine
HIS	L-Histidinol
HISOL	L-Histidinol phosphate
HISOLP	3-Hydroxykynurenine
HKYN	

5

10

15

20

25

30

35

40

45

50

55

Abbreviation	Metabolite
Hm	H+M
HMB	Hydroxymethylbilane
HOMOGEN	Homogenisate
HPRO	trans-4-Hydroxy-L-proline
HSER	L-Homoserine
HtRNA	tRNA(His)
HYXAN	Hypoxanthine
IAC	Indole-3-acetate
IAD	Indole-3-acetamide
IBCOA	2-Methylpropanoyl-CoA
ICIT	Isocitrate
ICITm	IsocitrateM
IDP	IDP
IDPm	IDPM
IGP	Indoleglycerol phosphate
IGST	4,4-Dimethylcholesta-8,14,24-trienol
IMZYMST	Intermediate_Methylzymosterol_II
IZYMST	Intermediate_Zymosterol_II
ILE	L-isoleucine
ILEm	L-isoleucineM
IMACP	3-(imidazol-4-yl)-2-octopropyl phosphate
IMP	IMP
IMZYMST	Intermediate_Methylzymosterol_I
INAC	Indoleacetate
INS	Inosine
IPC	Inositol phosphorylceramide
IPPMAL	2-Isopropylmalate
IPPMALm	2-IsopropylmalateM
IPPP	Isopentenyl diphosphate
ISUCC	α -inosuccinate
ITCCOAm	Itaconyl-CoAM
ITCm	ItaconateM
ITP	ITP

(continued)

5

10

15

20

25

30

35

40

50

55

(continued)

Abbreviation	Metabolite
ITPM	ITPM
IVCOA	3-Methylbutanoyl-CoA
IZYMST	Intermediate_Zymosterol_I
K	Potassium
KYN	L-Kynurenine
LAC	(R)-L-lactate
LACALm	(S)-LactaldehydeM
LACm	(R)-LactateM
LCCA	a Long-chain carboxylic acid
LEU	L-Leucine
LEUm	L-LeucineM
LGT	(R)-S-Lactoylglutathione
LGTrm	(R)-S-LactoylglutathioneM
LIPIV	2,3,2',3'-tetraakis(3-hydroxytetradecanoyl)-D-glucosaminyl-1,6-beta-D-glucosamine 1,4'-bisphosphate
LIPOn	LipoamideM
LIPX	Lipid X
LLACm	(S)-LactateM
LLCT	L-Cystathione
LLtRNA	L-lysyl-tRNA(Lys)
LLtRNAm	L-lysyl-tRNA(Lys)M
LNST	Lanosterol
LtRNA	tRNA(Lys)
LtRNAm	tRNA(Lys)M
LYS	L-Lysine
LYSm	L-LysineM
MAACOA	α-Methylacetoacetyl-CoA
MACAC	4-Maleylacetoacetate
MACOA	2-Methylprop-2-enoyl-CoA
MAL	Malate
MALACP	Malonyl-[acyl-carrier protein]
MALACPm	Malonyl-[acyl-carrier protein]M
MALCOA	Malonyl-CoA

5

10

15

20

25

30

35

40

45

50

55

(continued)

Abbreviation	Metabolite
MALm	MalateM
MALT	Malonate
MAL Tm	MalonateM
MAN	alpha-D-Mannose
MAN1P	alpha-D-Mannose 1-phosphate
MAN2PD	beta-D-Mannosyldiacetylglucosyldiphosphodolichol
MAN6P	D-Mannose 6-phosphate
MANNAN	Mannan
MBCOA	Methylbutyryl-CoA
MCCOA	2-Methylbut-2-enoyl-CoA
MRCOA	2-Methylbut-2-enoyl-CoA
MDAP	Meso-diaminopimelate
MELI	Melibiose
MELT	Melibioitol
MET	L-Methionine
METH	Methanethiol
METHF	5,10-Methenyltetrahydrofolate
METHFm	5,10-MethenyltetrahydrofolateM
METTHF	5,10-Methylenetetrahydrofolate
METTHFm	5,10-MethylenetetrahydrofolateM
MGCOA	3-Methylglutaconyl-CoA
MHIS	N(pai)-Methyl-L-histidine
MHVCOA	a-Methyl-β-hydroxyvaleryl-CoA
MI	myo-Inositol
MI1P	1L-myoinositol 1-phosphate
MIP2C	Inositol-mannose-P-inositol-P-ceramide
MIPC	Mannose-inositol-P-ceramide
MK	Menaquinone
MLT	Maltose
MMCOA	Methylmalonyl-CoA
MMET	S-Methylmethionine
MMS	(S)-Methylmalonate semialdehyde
MNT	D-Mannitol
MNT6P	D-Mannitol 1-phosphate
MTHF	5-Methyltetrahydrofolate

5

10

15

20

25

(continued)

Abbreviation	Metabolite
MTHFm	5-MethyltetrahydrofolateM
MTHGXL	Methylglyoxal
MTHN	Methane
MTHPTGLU	MethanemM
MTHNm	5-Methyltetrahydropteroyltri-L-glutamate
MTRNAm	L-Methionyl-tRNAM
MVL	(R)-Mevalonate
MVLm	(R)-MevalonateM
MYO _l	myo-Inositol
MZYMST	4-Methylzymosterol
N4HBZ	3-Nonaprenyl-4-hydroxybenzoate
NA	Sodium
NAAD	Deamino-NAD ⁺
NAADM	Deamino-NAD+M
NAC	Nicotinate
NACm	NicotinatemM
NAD	NAD+
NADH	NADH
NADHm	NADHm
NADm	NAD+M
NADP	NADP+
NADPH	NADPH
NADPHm	NADPHM
NADPm	NADP+M
NAG	N-Acetylglucosamine
NAGA1P	N-Acetyl-D-glucosamine 1-phosphate
NAGA6P	N-Acetyl-D-glucosamine 6-phosphate
NAGLUm	N-Acetyl-L-glutamateM
NAGLUUpm	N-Acetyl-L-glutamate 5-phosphateM
NAGLUUsm	N-Acetyl-L-glutamate 5-semialdehydeM
NAM	Nicotinamide
NAMm	NicotinamideM
NAMN	Nicotinate D-ribonucleotide

5

10

15

20

25

		(continued)
NAMNm	Nicotinate D-ribonucleotideM	
NAORNm	N2-Acetyl-L-ornithineM	
NH3		
NH3m		
NH4	NH4+	
NPP	all-trans-Nonaprenyl diphosphate	
NPPm	all-trans-Nonaprenyl diphosphateM	
NPRAN	N-(5-Phospho-D-ribofuranosyl)anthranilate	
O2	Oxygen	
O2m	OxygenM	
OA	Oxaloacetate	
OACOA	3-Oxoacyl-CoA	
OAHSER	O-Acetyl-L-homoserine	
OAm	OxaloacetateM	
OBUT	2-Oxobutanoate	
OBUTm	2-OxobutanoateM	
OPF	Oxidized flavoprotein	
OGT	Oxidized glutathione	
OHB	2-Oxo-3-hydroxy-4-phosphobutyrate	
OHm	HO-M	
OICAP	3-Carboxy-4-methyl-2-oxopentanoate	
OICAPm	3-Carboxy-4-methyl-2-oxopentanoateM	
OIVAL	(R)-2-Oxovalerate	
OIVALm	(R)-2-OxovalerateM	
OMP	Orotidine 5'-phosphate	
OMVAL	3-Methyl-2-oxobutanoate	
OMVALm	3-Methyl-2-oxobutanoateM	
OPEP	Oligopeptide	
ORN	L-Ornithine	
ORNm	L-OrnithinemM	
OROA	Orotate	
OSLHSER	O-Succinyl-L-homoserine	
OSUC	Oxalosuccinate	

5

10

15

20

25

30

35

40

50

55

(continued)

Abbreviation	Metabolite
OSUCm	OxalosuccinateM
OTHIO	Oxidized thioredoxin
OTHIOm	Oxidized thioredoxinM
OXA	Oxaloglutarate
OXAm	OxaloglutarateM
P5C	(S)-1-Pyrroline-5-carboxylate
P5Cm	(S)-1-Pyrroline-5-carboxylateM
P5P	Pyridoxine phosphate
PA	Phosphatidate
PABA	4-Aminobenzoate
PAC	Phenylacetic acid
PAD	2-Phenylacetamide
PALCOA	Palmitoyl-CoA
PAm	PhosphatidateM
PANT	(R)-Pantoate
PANTm	(R)-PantoateM
PAP	Adenosine 3',5'-bisphosphate
PAPS	3'-Phosphoadenylylsulfate
PBG	Porphobilinogen
PC	Phosphatidylcholine
PC2	Sirohydrochlorin
PCHO	Choline phosphate
PDLA	Pyridoxamine
PDLA5P	Pyridoxamine phosphate
PDME	Phosphatidyl-N-dimethyl-ethanolamine
PE	Phosphatidylethanolamine
PEm	PhosphatidylethanolamineM
PEP	Phosphoenolpyruvate
PEPD	Peptide
PEPm	PhosphoenolpyruvateM
PEPT	Peptide
PETHM	Ethanolamine phosphate
PGm	PhosphatidylglycerolM
PGPm	PhosphatidylglycerophosphateM

5

10

15

20

25

30

35

40

50

55

(continued)

Abbreviation	Metabolite
PHC	L-1-Pyrroline-3-hydroxy-5-carboxylate
PHE	L-Phenylalanine
PHEN	Prephenate
PHP	3-Phosphoopyruvate
PHPYR	Phenylpyruvate
PHSER	O-Phospho-L-homoserine
PHSP	Phytosphingosine 1-phosphate
PHT	O-Phospho-4-hydroxy-L-threonine
PI	Orthophosphate
Plm	OrthophosphateM
PIME	Pimelic Acid
PINS	1-Phosphatidyl-D-myo-inositol
PINS4P	1-Phosphatidyl-1D-myo-inositol 4-phosphate
PINSP	1-Phosphatidyl-1D-myo-inositol 3-phosphate
PL	Pyridoxal
PL5P	Pyridoxal phosphate
PMME	Phosphatidyl-N-methyl ethanolamine
PMVL	(R)-5-Phosphomevalonate
PNTO	(R)-Pantothenate
PPHG	Protoporphyrinogen IX
PPHGM	Protoporphyrinogen IX M
PPI	Pyrophosphate
PPIm	PyrophosphateM
PPIXn	ProtoporphyrinM
PPMAL	2-Isopropylmaleate
PPMVl	(R)-5-Diphosphomevalonate
PRAM	5-Phosphoribosylamine
PRBAMP	N1-(5-Phospho-D-ribosyl)-AMP
PRBATP	N1-(5-Phospho-D-ribosyl)-ATP
PRFICA	1-(5'-Phosphoribosyl)-5-formamido-4-imidazolecarboxamide
PRFP	5-(5-Phospho-D-ribosylaminoformimino)-1-(5'-phosphoribosyl)-4-imidazolecarboxamide
PRLP	L-Proline
PRO	

5

10

15

20

25

30

35

(continued)

Abbreviation	Metabolite
PROm	L-ProlineM
PROPCOA	Propanoyl-CoA
PRPP	5-Phospho-alpha-D-ribose 1-diphosphate
PRPPm	5-Phospho-alpha-D-ribose 1-diphosphateM
PS	Phosphatidylserine
PSm	PhosphatidylserineM
PSPH	Phytosphingosine
PTHm	HemeM
PTRC	Putrescine
PTRSC	Putrescine
PUR15P	Pseudouridine 5'-phosphate
PYR	Pyruvate
PYRDX	Pyridoxine
PYRm	PyruvateM
Q	Ubiquinone-9
QA	Pyridine-2,3-dicarboxylate
QAm	Pyridine-2,3-dicarboxylateM
QH2	Ubiquinol
QH2m	UbiquinolM
Qm	Ubiquinone-9M
R1P	D-Ribose 1-phosphate
R5P	D-Ribose 5-phosphate
RADP	4-(1-D-Ribitylamino)-5-amino-2,8-dihydroxypyrimidine
RAF	Raffinose
RFP	Reduced flavoprotein
RGT	Glutathione
RGTrm	GlutathioneM
RIB	D-Ribose
RIBFLAVm	RiboflavinM
RIBOFLAV	Riboflavin
RIPm	alpha-D-Ribose 1-phosphateM
RL5P	D-Ribulose 5-phosphate
RMN	D-Rhamnose

5

10

15

20

25

30

35

40

45

50

55

(continued)

Abbreviation	Metabolite
RTHIO	Reduced thioredoxin
RTHIOM	Reduced thioredoxinM
S	Sulfur
S17P	Sedoheptulose 1,7-bisphosphate
S23E	(S)-2,3-Epoxysqualene
S7P	Sedoheptulose 7-phosphate
SACP	N6-(L-1,3-Dicarboxypropyl)-L-lysine
SAH	S-Adenosyl-L-homocysteine
SAHm	S-Adenosyl-L-homocysteineM
SAICAR	1-(5'-Phosphoribosyl)-5-amino-4-(N-succinocarboxamide)-imidazole
SAM	S-Adenosyl-L-methionine
SAMm	S-Adenosyl-L-methionineM
SAMOB	S-Adenosyl-4-methylthio-2-oxobutanoate
SAPm	S-AminomethyldihydrolipoylproteinM
SER	L-Serine
SERm	L-SerineM
SLF	Sulfate
SLFm	SulfateM
SME	Shikimate
SMESP	Shikimate 3-phosphate
SOR	Sorbose
SOR1P	Sorbose 1-phosphate
SOT	D-Sorbitol
SPH	Sphinganine
SPMD	Spermidine
SPRM	Spermine
SPRMD	Spermidine
SQL	Squalene
SUC	Sucrose
SUCC	Succinate
SUCCm	SuccinateM
SUCCOAm	Succinyl-CoAM

5

10

15

20

25

30

35

(continued)

Abbreviation	Metabolite
SUCCSAL	Succinate semialdehyde
T3P1	D-Glyceraldehyde 3-phosphate
T3P2	Glycerone phosphate
T3P2m	Glycerone phosphateM
TAG16P	D-Tagatose 1,6-bisphosphate
TAG6P	D-Tagatose 6-phosphate
TAGLY	Triacylglycerol
TCOA	Tetradecanoyl-CoA
TGLP	N-Tetradecanoylglycylpeptide
THF	Tetrahydrofolate
THFG	Tetrahydrofolyl-[Glu](n)
THFm	TetrahydrofolateM
THIAMIN	Thiamin
THMP	Thiamin monophosphate
THPTGLU	Tetrahydropteroyltri-L-glutamate
THR	L-Threonine
THRm	L-ThreonineM
THY	Thymine
THZ	5-(2-Hydroxyethyl)-4-methylthiazole
THZP	4-Methyl-5-(2-phenoxyethyl)-thiazole
TPI	D-myo-inositol 1,4,5-trisphosphate
TPP	Thiamin diphosphate
TPPP	Thiamin triphosphate
TRE	alpha,alpha'-Trehalose
TRE6P	alpha,alpha'-Trehalose 6-phosphate
TRNA	tRNA
TRNAG	tRNA(Glu)
TRNAGm	tRNA(Glu)M
TRNAm	tRNAM
TRP	L-Tryptophan
TRPm	L-TryptophanM
TRPTRNAm	L-Tryptophanyl-tRNA(Trp)M
TYR	L-Tyrosine
UDP	
UDPG	UDPGlucose

55

5

10

15

20

25

Abbreviation	Metabolite
UDPG23A	UDP-2,3-bis(3-hydroxytetradecanoyl)glucosamine
UDPG2A	UDP-3-O-(3-hydroxytetradecanoyl)-D-glucosamine
UDPG2AA	UDP-3-O-(3-hydroxytetradecanoyl)-N-acetylglucosamine
UDPGAL	UDP-D-galactose
UDPNAG	UDP-N-acetyl-D-galactosamine
UDPP	Undecaprenyl diphosphate
UGC	(-)-Ureidoglycolate
UMP	Uridylate
UPRG	Uroporphyrinogen III
URA	Uracil
UREA	Urea
UREAC	Urea-1-carboxylate
URI	Uridine
UTP	UTP
VAL	L-Valine
X5P	D-Xylose-5-phosphate
XAN	Xanthine
XMP	Xanthosine 5'-phosphate
XTSNE	Xanthosine
XTSN	Xanthosine
XUL	D-Xyulose
XYL	D-Xylose
ZYMST	Zymosterol

[0056] Depending upon the particular environmental conditions being tested and the desired activity, a reaction network data structure can contain smaller numbers of reactions such as at least 200, 150, 100 or 50 reactions. A reaction network data structure having relatively few reactions can provide the advantage of reducing computation time and resources required to perform a simulation. When desired, a reaction network data structure having a particular subset of reactions can be made or used in which reactions that are not relevant to the particular simulation are omitted. Alternatively, larger numbers of reactions can be included in order to increase the accuracy or molecular detail of the methods of the invention or to suit a particular application. Thus, a reaction network data structure can contain at least 5 300, 350, 400, 450, 500, 550, 600 or more reactions up to the number of reactions that occur in or by *S. cerevisiae* or that are desired to simulate the activity of the full set of reactions occurring in *S. cerevisiae*. A reaction network data structure that is substantially complete with respect to the metabolic reactions of *S. cerevisiae* provides the advantage 10 of being relevant to a wide range of conditions to be simulated, whereas those with smaller numbers of metabolic reactions are limited to a particular subset of conditions to be simulated.

[0057] A *S. cerevisiae* reaction network data structure can include one or more reactions that occur in or by *S. cerevisiae* and that do not occur, either naturally or following manipulation, in or by another prokaryotic organism, such as *Escherichia coli*, *Haemophilus influenzae*, *Bacillus subtilis*, *Helicobacter pylori* or in or by another eukaryotic organism, such as 15 *Homo sapiens*. Examples of reactions that are unique to *S. cerevisiae* compared at least to *Escherichia coli*, *Haemophilus influenzae*, and *Helicobacter pylori* include those identified in Table 4. It is understood that a *S. cerevisiae* reaction network data structure can also include one or more reactions that occur in another organism. Addition of such heterologous reactions to a reaction network data structure of the invention can be used in methods to predict the consequences 20 of heterologous gene transfer in *S. cerevisiae*, for example, when designing or engineering man-made cells or strains.

25

30

35

40

45

50

55

Table 4. Reactions specific to *S. cerevisiae* metabolic network

5	glk1_3, hxr1_1, hxr2_1, hxr1_4, hxr2_4, pfk1_3, idh1, idp1_1, idp1_2, idp2_1, idp3_1, idp2_2, idp3_2, 1sc1R, pyc1, pyc2, cyb2, dld1, ncp1, cytr_, cyto, atpl, pma1, pma2, pmp1, pmp2, cox1, rbl2, ach1_1, ach1_2, sfa1_1R, unkrx11R, pdc1, pdc5, pdc6, lys20, adh1R, adh3R, adh2R, adh4R, adh5R, sfa1_2R, psa1, pfk26, pfk27, fbp26, gal7R, mel1_2, mel1_3, mel1_4R, mel1_5R, mel1_6R, mell_7R, fsp2b, sor1, gsy1, gsy2, fks1, fks3, gsc2, tps1, tps3, ts1, tps2, ath1, nth1, nth2, fdh1, tfo1a, tfo1b, dur1R, dur2, nit2, cyr1, guk1_3R, ade2R, pde1, pde2_1, pde2_2, pde2_3, pde2_4, pde2_5, apa2, apa1_1, apa1_3, apa1_2R, ura2_1, ura4R, ura1_1R, ura10R, ura5R, ura3, npkR, furl, fcy1, tdk1, tdk2, urk1_1, urk1_2, urk1_3, deoa1R, deoa2R, cdd1_1, cdd1_2, cdc8R, dut1, cdc21, cmka2R, dcd1R, ura7_2, ura8_2, deglR, pus1R, pus2R, pus4R, ura1_2R, ara1_1, ara1_2, gna1R, pcm1aR, qri1R, chs1, chs2, chs3, put2_1, put2, glt1, gdh2, cat2, yat1, mht1, sam4, ecm40_2, cpa2, ura2_2, arg3, spe3, spe4, amd, amid2_1, atrna, msr1, rnas, ded81, hom6_1, cys4, gly1, agtR, gcv2R, sah1, met6, cys3, met17_1, metl7hR, dph5, met3, met14, met17_2, met17_3, lys21, lys20a, lys3R, lys4R, lys12R, lys12bR, amitR, lys2_1, lys2_2, lys9R, lys1aR, krs1, msk1, pro2_1, gps1R, gps2R, pro3_3, pro3_4, pro3_1, pro3_5, dal1R, dal2R, dal3R, his4_3, hts1, hmt1, tyr1, cta1, ct1, ald6, ald4_2, ald5_1, tdo2, kfor_, kynu_1, kmo, kynu_2, bnal, aaaa, aaab, aaac, tyrdega, tyrdegB, tyrdegC, trydegD, ms wl, amd2_2, amd2_3, spra, sprb, sprc, sprd, spre, dys1, leu4, leu1_2R, pclig, xapa1R, xapa2R, xapa3R, ynk1_6R, ynk1_9R, udpR, pyrh1R, pyrh2R, cmpg, usha1, usha2, usha5, usha6, usha11, gpx1R, gpx2R, hyr1R, ecm38, nit2_1, nit2_2, nmt1, nat1, nat2, bg12, exg1, exg2, spr1, thi80_1, thi80_2, unkrxn8, pho11, fmn1_1, fmn1_2, pdx3_2R, pdx3_3R, pdx3_4R, pdx3_1, pdx3_5, bio1, fol1_4, ftfa, ftfb, fol3R, met7R, rma1R, met12, met13, mis1_2, ade3_2, mtdl, fmt1, Typell_1, Typell_2, Typell_4, Typell_3, Typell_6, Typell_5, Typell_9, Typell_8, Typell_7, c100sn, c180sy, c182sy, faa1R, faa2R, faa3R, faa4R, fox2bR, pot1_1, erg10_1R, erg10_2R, Gat1_2, Gat2_2, ADHAPR, AGAT, slc1, Gat1_1, Gat2_1, cho1aR, cho1bR, cho2, op13_1, op13_2, cki1, pct1, cpt1, eki1, ect1, ept1R, inol, impal, pis1, tor1, tor2, vps34, pik1, sst4, fab1, mss4, plc1, pgs1R, crd1, dpp1, lpp1, hmgsR, hmg1R, hmg2R, erg12_1, erg12_2, erg12_3, erg12_4, erg8, mvd1, erg9, erg1, erg7, unkrxn3, unkrxn4, cdiso4, erg11_1, erg24, erg25_1, erg26_1, erg11_2, erg25_2, erg26_2, erg11_3, erg6, erg2, erg3, erg5, erg4, lcb1, lcb2, tsc10, sur2, csyna, csynb, scs7, aurl, csg2, sur1, ipt1, lcb4_1, lcb5_1, lcb4_2, lcb5_2, lcb3, ysr3, dpl1, sec59, dpml, pmt1, pmt2, pmt3, pmt4, pmt5, pmt6, kre2, ktr1, ktr2, ktr3, ktr4, ktr6, yurl, hor2, rhr2, cda1, cda2, daga, dak1, dak2, gpd1, nadg1R, nadg2R, npt1, nadi, mnadphs, mnadg1R, mnadg2R, mnpt1, mnadi, hem1, bet2, coq1, coq2, cox10, ram1, rer2, srt1, mo2R, mco2R, methR, mmthnR, mn3R, mthfR, mmthfR, mserR, mglyR, mcbhR, moicapR, mproR, mcnpR, macR, macar_, mcar_, maclacR, mactcR, moivalR, momvalR, mpmalRR, msif, mthrR, maka, aac1, aac3, pet9, mir1aR, mir1dR, dic1_2R, dic1_1R, dic1_3, mmrlR, moabR, ctp1_1R, ctp1_2R, ctp1_3R, pyrcaR, mlacR, gcaR, gcb, ort1R, crc1, gut2, gpd2, mt3p, mgl3p, mfad, mriboR, mdtbR, mmcoaR, mmvIR, mpaR, mpntR, madR, mprppR, mdhfR, mqaR, moppR, msamR, msahR, sfc1, odc1R, odc2R, hxt1_2, hxt10_2, hxt11_2, hxt13_2, hxt15_2, hxt16_2, hxt17_2, hxt2_2, hxt3_2, hxt4_2, hxt5_2, hxt6_2, hxt7_2, hxt8_5, hxt9_2, sucup, akmupR, sorupR, arbup1R, gltupb, gal2_3, hxt1_1, hxt10_1, hxt11_1, hxt13_1, hxt15_1, hxt16_1, hxt17_1, hxt2_1, hxt3_1, hxt4, hxt4_1, hxt5_1, hxt6_1, hxt7_1, hxt8_4, hxt9_1, stl1_1, gaupR, mmp1, mltup, mntup, nagup, rmnup, ribup,
10	
15	
20	
25	
30	
35	
40	
45	

[0058] A reaction network data structure or index of reactions used in the data structure such as that available in a metabolic reaction database, as described above, can be annotated to include information about a particular reaction. A reaction can be annotated to indicate, for example, assignment of the reaction to a protein, macromolecule or enzyme that performs the reaction, assignment of a gene(s) that codes for the protein, macromolecule or enzyme, the Enzyme Commission (EC) number of the particular metabolic reaction or Gene Ontology (GO) number of the particular metabolic reaction, a subset of reactions to which the reaction belongs, citations to references from which information was obtained, or a level of confidence with which a reaction is believed to occur in *S. cerevisiae*. A computer readable medium or media of the invention can include a gene database containing annotated reactions. Such information can be obtained during the course of building a metabolic reaction database or model of the invention as described below.

[0059] As used herein, the term "gene database" is intended to mean a computer readable medium or media that contains at least one reaction that is annotated to assign a reaction to one or more macromolecules that perform the reaction or to assign one or more nucleic acid that encodes the one or more macromolecules that perform the reaction. A gene database can contain a plurality of reactions some or all of which are annotated. An annotation can include, for

example, a name for a macromolecule; assignment of a function to a macromolecule; assignment of an organism that contains the macromolecule or produces the macromolecule; assignment of a subcellular location for the macromolecule; assignment of conditions under which a macromolecule is being expressed or being degraded; an amino acid or nucleotide sequence for the macromolecule; or any other annotation found for a macromolecule in a genome database such as those that can be found in *Saccharomyces* Genome Database maintained by Stanford University, or Comprehensive Yeast Genome Database maintained by MIPS.

[0060] A gene database of the invention can include a substantially complete collection of genes and/or open reading frames in *S. cerevisiae* or a substantially complete collection of the macromolecules encoded by the *S. cerevisiae* genome. Alternatively, a gene database can include a portion of genes or open reading frames in *S. cerevisiae* or a portion of the macromolecules encoded by the *S. cerevisiae* genome. The portion can be at least 10%, 15%, 20%, 25%, 50%, 75%, 90% or 95% of the genes or open reading frames encoded by the *S. cerevisiae* genome, or the macromolecules encoded therein. A gene database can also include macromolecules encoded by at least a portion of the nucleotide sequence for the *S. cerevisiae* genome such as at least 10%, 15%, 20%, 25%, 50%, 75%, 90% or 95% of the *S. cerevisiae* genome. Accordingly, a computer readable medium or media of the invention can include at least one reaction for each macromolecule encoded by a portion of the *S. cerevisiae* genome.

[0061] An *in silico* *S. cerevisiae* model according to the invention can be built by an iterative process which includes gathering information regarding particular reactions to be added to a model, representing the reactions in a reaction network data structure, and performing preliminary simulations wherein a set of constraints is placed on the reaction network and the output evaluated to identify errors in the network. Errors in the network such as gaps that lead to non-natural accumulation or consumption of a particular metabolite can be identified as described below and simulations repeated until a desired performance of the model is attained. An exemplary method for iterative model construction is provided in Example I.

[0062] Thus, the invention provides a method for making a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions in a computer readable medium or media. The method includes the steps of: (a) identifying a plurality of *S. cerevisiae* reactions and a plurality of *S. cerevisiae* reactants that are substrates and products of the *S. cerevisiae* reactions; (b) relating the plurality of *S. cerevisiae* reactants to the plurality of *S. cerevisiae* reactions in a data structure, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product; (c) making a constraint set for the plurality of *S. cerevisiae* reactions; (d) providing an objective function; (e) determining at least one flux distribution that minimizes or maximizes the objective function when the constraint set is applied to the data structure, and (f) if *at least* one flux distribution is not predictive of *S. cerevisiae* physiology, then adding a reaction to or deleting a reaction from the data structure and repeating step (e), if *at least* one flux distribution is predictive of *S. cerevisiae* physiology, then storing the data structure in a computer readable medium or media.

[0063] Information to be included in a data structure of the invention can be gathered from a variety of sources including, for example, the scientific literature or an annotated genome sequence of *S. cerevisiae* such as the Genbank, a site maintained by the NCBI (ncbi.nlm.gov), the CYGD database, a site maintained by MIPS, or the SGD database, a site maintained by the School of Medicine at Stanford University, etc.

[0064] In the course of developing an *in silico* model of *S. cerevisiae* metabolism, the types of data that can be considered include, for example, biochemical information which is information related to the experimental characterization of a chemical reaction, often directly indicating a protein(s) associated with a reaction and the stoichiometry of the reaction or indirectly demonstrating the existence of a reaction occurring within a cellular extract; genetic information which is information related to the experimental identification and genetic characterization of a gene(s) shown to code for a particular protein(s) implicated in carrying out a biochemical event; genomic information which is information related to the identification of an open reading frame and functional assignment, through computational sequence analysis, that is then linked to a protein performing a biochemical event; physiological information which is information related to overall cellular physiology, fitness characteristics, substrate utilization, and phenotyping results, which provide evidence of the assimilation or dissimilation of a compound used to infer the presence of specific biochemical event (in particular translocations); and modeling information which is information generated through the course of simulating activity of *S. cerevisiae* using methods such as those described herein which lead to predictions regarding the status of a reaction such as whether or not the reaction is required to fulfill certain demands placed on a metabolic network.

[0065] The majority of the reactions occurring in *S. cerevisiae* reaction networks are catalyzed by enzymes/proteins, which are created through the transcription and translation of the genes found on the chromosome(s) in the cell. The remaining reactions occur through non-enzymatic processes. Furthermore, a reaction network data structure can contain reactions that add or delete steps to or from a particular reaction pathway. For example, reactions can be added to optimize or improve performance of a *S. cerevisiae* model in view of empirically observed activity. Alternatively, reactions can be deleted to remove intermediate steps in a pathway when the intermediate steps are not necessary to model flux through the pathway. For example, if a pathway contains 3 nonbranched steps, the reactions can be combined or added together to give a net reaction, thereby reducing memory required to store the reaction network data structure and the

computational resources required for manipulation of the data structure. An example of a combined reaction is that for fatty acid degradation shown in Table 2, which combines the reactions for acyl-CoA oxidase, hydratase-dehydrogenase-epimerase, and acetyl-CoA C-acyltransferase of beta-oxidation of fatty acids.

[0066] The reactions that occur due to the activity of gene-encoded enzymes can be obtained from a genome database that lists genes or open reading frames identified from genome sequencing and subsequent genome annotation. Genome annotation consists of the locations of open reading frames and assignment of function from homology to other known genes or empirically determined activity. Such a genome database can be acquired through public or private databases containing annotated *S. cerevisiae* nucleic acid or protein sequences. If desired, a model developer can perform a network reconstruction and establish the model content associations between the genes, proteins, and reactions as described, for example, in Covert et al. Trends in Biochemical Sciences 26:179-186 (2001) and Palsson, WO 00/46405.

[0067] As reactions are added to a reaction network data structure or metabolic reaction database, those having known or putative associations to the proteins/enzymes which enable/catalyze the reaction and the associated genes that code for these proteins can be identified by annotation. Accordingly, the appropriate associations for some or all of the reactions to their related proteins or genes or both can be assigned. These associations can be used to capture the non-linear relationship between the genes and proteins as well as between proteins and reactions. In some cases, one gene codes for one protein which then performs one reaction. However, often there are multiple genes which are required to create an active enzyme complex and often there are multiple reactions that can be carried out by one protein or multiple proteins that can carry out the same reaction. These associations capture the logic (i.e. AND or OR relationships) within the associations. Annotating a metabolic reaction database with these associations can allow the methods to be used to determine the effects of adding or eliminating a particular reaction not only at the reaction level, but at the genetic or protein level in the context of running a simulation or predicting *S. cerevisiae* activity.

[0068] A reaction network data structure of the invention can be used to determine the activity of one or more reactions in a plurality of *S. cerevisiae* reactions independent of any knowledge or annotation of the identity of the protein that performs the reaction or the gene encoding the protein. A model that is annotated with gene or protein identities can include reactions for which a protein or encoding gene is not assigned. While a large portion of the reactions in a cellular metabolic network are associated with genes in the organism's genome, there are also a substantial number of reactions included in a model for which there are no known genetic associations. Such reactions can be added to a reaction database based upon other information that is not necessarily related to genetics such as biochemical or cell based measurements or theoretical considerations based on observed biochemical or cellular activity. For example, there are many reactions that are not protein-enabled reactions. Furthermore, the occurrence of a particular reaction in a cell for which no associated proteins or genetics have been currently identified can be indicated during the course of model building by the iterative model building methods of the invention.

[0069] The reactions in a reaction network data structure or reaction database can be assigned to subsystems by annotation, if desired. The reactions can be subdivided according to biological criteria, such as according to traditionally identified metabolic pathways (glycolysis, amino acid metabolism and the like) or according to mathematical or computational criteria that facilitate manipulation of a model that incorporates or manipulates the reactions. Methods and criteria for subdividing a reaction database are described in further detail in Schilling et al., J. Theor. Biol. 203:249-283 (2000). The use of subsystems can be advantageous for a number of analysis methods, such as extreme pathway analysis, and can make the management of model content easier. Although assigning reactions to subsystems can be achieved without affecting the use of the entire model for simulation, assigning reactions to subsystems can allow a user to search for reactions in a particular subsystem, which may be useful in performing various types of analyses. Therefore, a reaction network data structure can include any number of desired subsystems including, for example, 2 or more subsystems, 5 or more subsystems, 10 or more subsystems, 25 or more subsystems or 50 or more subsystems.

[0070] The reactions in a reaction network data structure or metabolic reaction database can be annotated with a value indicating the confidence with which the reaction is believed to occur in *S. cerevisiae*. The level of confidence can be, for example, a function of the amount and form of supporting data that is available. This data can come in various forms including published literature, documented experimental results, or results of computational analyses. Furthermore, the data can provide direct or indirect evidence for the existence of a chemical reaction in a cell based on genetic, biochemical, and/or physiological data.

[0071] The invention further provides a computer readable medium, containing (a) a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product, and (b) a constraint set for the plurality of *S. cerevisiae* reactions.

[0072] Constraints can be placed on the value of any of the fluxes in the metabolic network using a constraint set. These constraints can be representative of a minimum or maximum allowable flux through a given reaction, possibly resulting from a limited amount of an enzyme present. Additionally, the constraints can determine the direction or reversibility of any of the reactions or transport fluxes in the reaction network data structure. Based on the *in vivo* environment where *S. cerevisiae* lives the metabolic resources available to the cell for biosynthesis of essential molecules for can

be determined. Allowing the corresponding transport fluxes to be active provides the *in silico S. cerevisiae* with inputs and outputs for substrates and by-products produced by the metabolic network.

[0073] Returning to the hypothetical reaction network shown in Figure 1, constraints can be placed on each reaction in the exemplary format, shown in Figure 3, as follows. The constraints are provided in a format that can be used to constrain the reactions of the stoichiometric matrix shown in Figure 2. The format for the constraints used for a matrix or in linear programming can be conveniently represented as a linear inequality such as

$$\beta_j \leq v_j \leq \alpha_j : j = 1 \dots n \quad (\text{Eq. 1})$$

where v_j is the metabolic flux vector, β_j is the minimum flux value and α_j is the maximum flux value. Thus, α_j can take on a finite value representing a maximum allowable flux through a given reaction or β_j can take on a finite value representing minimum allowable flux through a given reaction. Additionally, if one chooses to leave certain reversible reactions or transport fluxes to operate in a forward and reverse manner the flux may remain unconstrained by setting β_j to negative infinity and α_j to positive infinity as shown for reaction R₂ in Figure 3. If reactions proceed only in the forward reaction β_j is set to zero while α_j is set to positive infinity as shown for reactions R₁, R₃, R₄, R₅, and R₆ in Figure 3. As an example, to simulate the event of a genetic deletion or non-expression of a particular protein, the flux through all of the corresponding metabolic reactions related to the gene or protein in question are reduced to zero by setting α_j and β_j to be zero. Furthermore, if one wishes to simulate the absence of a particular growth substrate, one can simply constrain the corresponding transport fluxes that allow the metabolite to enter the cell to be zero by setting α_j and β_j to be zero. On the other hand if a substrate is only allowed to enter or exit the cell via transport mechanisms, the corresponding fluxes can be properly constrained to reflect this scenario.

[0074] The *in silico S. cerevisiae* model and methods described herein can be implemented on any conventional host computer system, such as those based on Intel.RTM. microprocessors and running Microsoft Windows operating systems. Other systems, such as those using the UNIX or LINUX operating system and based on IBM.RTM., DEC.RTM. or Motorola.RTM. microprocessors are also contemplated. The systems and methods described herein can also be implemented to run on client-server systems and wide-area networks, such as the Internet.

[0075] Software to implement a method or model of the invention can be written in any well-known computer language, such as Java, C, C++, Visual Basic, FORTRAN or COBOL and compiled using any well-known compatible compiler. The software of the invention normally runs from instructions stored in a memory on a host computer system. A memory or computer readable medium can be a hard disk, floppy disc, compact disc, magneto-optical disc, Random Access Memory, Read Only Memory or Flash Memory. The memory or computer readable medium used in the invention can be contained within a single computer or distributed in a network. A network can be any of a number of conventional network systems known in the art such as a local area network (LAN) or a wide area network (WAN). Client-server environments, database servers and networks that can be used in the invention are well known in the art. For example, the database server can run on an operating system such as UNIX, running a relational database management system, a World Wide Web application and a World Wide Web server. Other types of memories and computer readable media are also contemplated to function within the scope of the invention.

[0076] A database or data structure of the invention can be represented in a markup language format including, for example, Standard Generalized Markup Language (SGML), Hypertext markup language (HTML) or Extensible Markup language (XML). Markup languages can be used to tag the information stored in a database or data structure of the invention, thereby providing convenient annotation and transfer of data between databases and data structures. In particular, an XML format can be useful for structuring the data representation of reactions, reactants and their annotations; for exchanging database contents, for example, over a network or internet; for updating individual elements using the document object model; or for providing differential access to multiple users for different information content of a data base or data structure of the invention. XML programming methods and editors for writing XML code are known in the art as described, for example, in Ray, Learning XML O'Reilly and Associates, Sebastopol, CA (2001).

[0077] A set of constraints can be applied to a reaction network data structure to simulate the flux of mass through the reaction network under a particular set of environmental conditions specified by a constraints set. Because the time constants characterizing metabolic transients and/or metabolic reactions are typically very rapid, on the order of milliseconds to seconds, compared to the time constants of cell growth on the order of hours to days, the transient mass balances can be simplified to only consider the steady state behavior. Referring now to an example where the reaction network data structure is a stoichiometric matrix, the steady state mass balances can be applied using the following system of linear equations

$$S \bullet v = 0 \quad (\text{Eq. 2})$$

where S is the stoichiometric matrix as defined above and v is the flux vector. This equation defines the mass, energy, and redox potential constraints placed on the metabolic network as a result of stoichiometry. Together Equations 1 and 2 representing the reaction constraints and mass balances, respectively, effectively define the capabilities and constraints of the metabolic genotype and the organism's metabolic potential. All vectors, v , that satisfy Equation 2 are said to occur in the mathematical nullspace of S . Thus, the null space defines steady-state metabolic flux distributions that do not violate the mass, energy, or redox balance constraints. Typically, the number of fluxes is greater than the number of mass balance constraints, thus a plurality of flux distributions satisfy the mass balance constraints and occupy the null space. The null space, which defines the feasible set of metabolic flux distributions, is further reduced in size by applying the reaction constraints set forth in Equation 1 leading to a defined solution space. A point in this space represents a flux distribution and hence a metabolic phenotype for the network. An optimal solution within the set of all solutions can be determined using mathematical optimization methods when provided with a stated objective and a constraint set. The calculation of any solution constitutes a simulation of the model.

[0078] Objectives for activity of *S. cerevisiae* can be chosen to explore the improved use of the metabolic network within a given reaction network data structure. These objectives can be design objectives for a strain, exploitation of the metabolic capabilities of a genotype, or physiologically meaningful objective functions, such as maximum cellular growth. Growth can be defined in terms of biosynthetic requirements based on literature values of biomass composition or experimentally determined values such as those obtained as described above. Thus, biomass generation can be defined as an exchange reaction that removes intermediate metabolites in the appropriate ratios and represented as an objective function. In addition to draining intermediate metabolites this reaction flux can be formed to utilize energy molecules such as ATP, NADH and NADPH so as to incorporate any growth dependent maintenance requirement that must be met. This new reaction flux then becomes another constraint/balance equation that the system must satisfy as the objective function. Using the stoichiometric matrix of Figure 2 as an example, adding such a constraint is analogous to adding the additional column V_{growth} to the stoichiometric matrix to represent fluxes to describe the production demands placed on the metabolic system. Setting this new flux as the objective function and asking the system to maximize the value of this flux for a given set of constraints on all the other fluxes is then a method to simulate the growth of the organism.

[0079] Continuing with the example of the stoichiometric matrix applying a constraint set to a reaction network data structure can be illustrated as follows. The solution to equation 2 can be formulated as an optimization problem, in which the flux distribution that minimizes a particular objective is found. Mathematically, this optimization problem can be stated as:

30

$$\text{Minimize } Z \quad (\text{Eq. 3})$$

35

$$\text{where } (\text{Eq. 4})$$

$$z = \sum c_i \cdot v_i$$

40 where Z is the objective which is represented as a linear combination of metabolic fluxes v_i using the weights c_i in this linear combination. The optimization problem can also be stated as the equivalent maximization problem; i.e. by changing the sign on Z . Any commands for solving the optimization problem can be used including, for example, linear programming commands.

[0080] A computer system of the invention can further include a user interface capable of receiving a representation of one or more reactions. A user interface of the invention can also be capable of sending at least one command for modifying the data structure, the constraint set or the commands for applying the constraint set to the data representation, or a combination thereof. The interface can be a graphic user interface having graphical means for making selections such as menus or dialog boxes. The interface can be arranged with layered screens accessible by making selections from a main screen. The user interface can provide access to other databases useful in the invention such as a metabolic reaction database or links to other databases having information relevant to the reactions or reactants in the reaction network data structure or to *S. cerevisiae* physiology. Also, the user interface can display a graphical representation of a reaction network or the results of a simulation using a model of the invention.

[0081] Once an initial reaction network data structure and set of constraints has been created, this model can be tested by preliminary simulation. During preliminary simulation, gaps in the network or "dead-ends" in which a metabolite can be produced but not consumed or where a metabolite can be consumed but not produced can be identified. Based on the results of preliminary simulations areas of the metabolic reconstruction that require an additional reaction can be identified. The determination of these gaps can be readily calculated through appropriate queries of the reaction network

data structure and need not require the use of simulation strategies, however, simulation would be an alternative approach to locating such gaps.

[0082] In the preliminary simulation testing and model content refinement stage the existing model is subjected to a series of functional tests to determine if it can perform basic requirements such as the ability to produce the required biomass constituents and generate predictions concerning the basic physiological characteristics of the particular organism strain being modeled. The more preliminary testing that is conducted the higher the quality of the model that will be generated. Typically the majority of the simulations used in this stage of development will be single optimizations. A single optimization can be used to calculate a single flux distribution demonstrating how metabolic resources are routed determined from the solution to one optimization problem. An optimization problem can be solved using linear programming as demonstrated in the Examples below. The result can be viewed as a display of a flux distribution on a reaction map. Temporary reactions can be added to the network to determine if they should be included into the model based on modeling/simulation requirements.

[0083] Once a model of the invention is sufficiently complete with respect to the content of the reaction network data structure according to the criteria set forth above, the model can be used to simulate activity of one or more reactions in a reaction network. The results of a simulation can be displayed in a variety of formats including, for example, a table, graph, reaction network, flux distribution map or a phenotypic phase plane graph.

[0084] Thus, the invention provides a method for predicting a *S. cerevisiae* physiological function. The method includes the steps of (a) providing a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product; (b) providing a constraint set for the plurality of *S. cerevisiae* reactions; (c) providing an objective function, and (d) determining at least one flux distribution that minimizes or maximizes the objective function when the constraint set is applied to the data structure, thereby predicting a *S. cerevisiae* physiological function.

[0085] As used herein, the term "physiological function," when used in reference to *S. cerevisiae*, is intended to mean an activity of a *S. cerevisiae* cell as a whole. An activity included in the term can be the magnitude or rate of a change from an initial state of a *S. cerevisiae* cell to a final state of the *S. cerevisiae* cell. An activity can be measured qualitatively or quantitatively. An activity included in the term can be, for example, growth, energy production, redox equivalent production, biomass production, development, or consumption of carbon, nitrogen, sulfur, phosphate, hydrogen or oxygen. An activity can also be an output of a particular reaction that is determined or predicted in the context of substantially all of the reactions that affect the particular reaction in a *S. cerevisiae* cell or substantially all of the reactions that occur in a *S. cerevisiae* cell. Examples of a particular reaction included in the term are production of biomass precursors, production of a protein, production of an amino acid, production of a purine, production of a pyrimidine, production of a lipid, production of a fatty acid, production of a cofactor, or transport of a metabolite. A physiological function can include an emergent property which emerges from the whole but not from the sum of parts where the parts are observed in isolation (see for example, Palsson Nat. Biotech 18:1147-1150 (2000)).

[0086] A physiological function of *S. cerevisiae* reactions can be determined using phase plane analysis of flux distributions. Phase planes are representations of the feasible set which can be presented in two or three dimensions. As an example, two parameters that describe the growth conditions such as substrate and oxygen uptake rates can be defined as two axes of a two-dimensional space. The optimal flux distribution can be calculated from a reaction network data structure and a set of constraints as set forth above for all points in this plane by repeatedly solving the linear programming problem while adjusting the exchange fluxes defining the two-dimensional space. A finite number of qualitatively different metabolic pathway utilization patterns can be identified in such a plane, and lines can be drawn to demarcate these regions. The demarcations defining the regions can be determined using shadow prices of linear optimization as described, for example in Chvatal, Linear Programming New York, W.H. Freeman and Co. (1983). The regions are referred to as regions of constant shadow price structure. The shadow prices define the intrinsic value of each reactant toward the objective function as a number that is either negative, zero, or positive and are graphed according to the uptake rates represented by the x and y axes. When the shadow prices become zero as the value of the uptake rates are changed there is a qualitative shift in the optimal reaction network.

[0087] One demarcation line in the phenotype phase plane is defined as the line of optimality (LO). This line represents the optimal relation between respective metabolic fluxes. The LO can be identified by varying the x-axis flux and calculating the optimal y-axis flux with the objective function defined as the growth flux. From the phenotype phase plane analysis the conditions under which a desired activity is optimal can be determined. The maximal uptake rates lead to the definition of a finite area of the plot that is the predicted outcome of a reaction network within the environmental conditions represented by the constraint set. Similar analyses can be performed in multiple dimensions where each dimension on the plot corresponds to a different uptake rate. These and other methods for using phase plane analysis, such as those described in Edwards et al., Biotech Bioeng. 77:27-36(2002), can be used to analyze the results of a simulation using an *in silico* *S. cerevisiae* model of the invention.

[0088] A physiological function of *S. cerevisiae* can also be determined using a reaction map to display a flux distribution.

A reaction map of *S. cerevisiae* can be used to view reaction networks at a variety of levels. In the case of a cellular metabolic reaction network a reaction map can contain the entire reaction complement representing a global perspective. Alternatively, a reaction map can focus on a particular region of metabolism such as a region corresponding to a reaction subsystem described above or even on an individual pathway or reaction. An example of a reaction map showing a subset of reactions in a reaction network of *S. cerevisiae* is shown in Figure 4.

[0089] The invention also provides an apparatus that produces a representation of a *S. cerevisiae* physiological function, wherein the representation is produced by a process including the steps of: (a) providing a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product; (b) providing a constraint set for the plurality of *S. cerevisiae* reactions; (c) providing an objective function; (d) determining at least one flux distribution that minimizes or maximizes the objective function when the constraint set is applied to the data structure, thereby predicting a *S. cerevisiae* physiological function, and (e) producing a representation of the activity of the one or more *S. cerevisiae* reactions.

[0090] The methods of the invention can be used to determine the activity of a plurality of *S. cerevisiae* reactions including, for example, biosynthesis of an amino acid, degradation of an amino acid, biosynthesis of a purine, biosynthesis of a pyrimidine, biosynthesis of a lipid, metabolism of a fatty acid, biosynthesis of a cofactor, transport of a metabolite and metabolism of an alternative carbon source. In addition, the methods can be used to determine the activity of one or more of the reactions described above or listed in Table 2.

[0091] The methods of the invention can be used to determine a phenotype of a *S. cerevisiae* mutant. The activity of one or more *S. cerevisiae* reactions can be determined using the methods described above, wherein the reaction network data structure lacks one or more gene-associated reactions that occur in *S. cerevisiae*. Alternatively, the methods can be used to determine the activity of one or more *S. cerevisiae* reactions when a reaction that does not naturally occur in *S. cerevisiae* is added to the reaction network data structure. Deletion of a gene can also be represented in a model of the invention by constraining the flux through the reaction to zero, thereby allowing the reaction to remain within the data structure. Thus, simulations can be made to predict the effects of adding or removing genes to or from *S. cerevisiae*. The methods can be particularly useful for determining the effects of adding or deleting a gene that encodes for a gene product that performs a reaction in a peripheral metabolic pathway.

[0092] A drug target or target for any other agent that affects *S. cerevisiae* function can be predicted using the methods of the invention. Such predictions can be made by removing a reaction to simulate total inhibition or prevention by a drug or agent. Alternatively, partial inhibition or reduction in the activity a particular reaction can be predicted by performing the methods with altered constraints. For example, reduced activity can be introduced into a model of the invention by altering the α_j or β_j values for the metabolic flux vector of a target reaction to reflect a finite maximum or minimum flux value corresponding to the level of inhibition. Similarly, the effects of activating a reaction, by initiating or increasing the activity of the reaction, can be predicted by performing the methods with a reaction network data structure lacking a particular reaction or by altering the α_j or β_j values for the metabolic flux vector of a target reaction to reflect a maximum or minimum flux value corresponding to the level of activation. The methods can be particularly useful for identifying a target in a peripheral metabolic pathway.

[0093] Once a reaction has been identified for which activation or inhibition produces a desired effect on *S. cerevisiae* function, an enzyme or macromolecule that performs the reaction in *S. cerevisiae* or a gene that expresses the enzyme or macromolecule can be identified as a target for a drug or other agent. A candidate compound for a target identified by the methods of the invention can be isolated or synthesized using known methods. Such methods for isolating or synthesizing compounds can include, for example, rational design based on known properties of the target (see, for example, DeCamp et al., Protein Engineering Principles and Practice, Ed. Cleland and Craik, Wiley-Liss, New York, pp. 467-506 (1996)), screening the target against combinatorial libraries of compounds (see for example, Houghten et al., Nature, 354, 84-86 (1991); Dooley et al., Science, 266, 2019-2022 (1994), which describe an iterative approach, or R. Houghten et al. PCT/US91/08694 and U.S. Patent 5,556,762 which describe a positional-scanning approach), or a combination of both to obtain focused libraries. Those skilled in the art will know or will be able to routinely determine assay conditions to be used in a screen based on properties of the target or activity assays known in the art.

[0094] A candidate drug or agent, whether identified by the methods described above or by other methods known in the art, can be validated using an *in silico* *S. cerevisiae* model or method of the invention. The effect of a candidate drug or agent on *S. cerevisiae* physiological function can be predicted based on the activity for a target in the presence of the candidate drug or agent measured *in vitro* or *in vivo*. This activity can be represented in an *in silico* *S. cerevisiae* model by adding a reaction to the model, removing a reaction from the model or adjusting a constraint for a reaction in the model to reflect the measured effect of the candidate drug or agent on the activity of the reaction. By running a simulation under these conditions the holistic effect of the candidate drug or agent on *S. cerevisiae* physiological function can be predicted.

[0095] The methods of the invention can be used to determine the effects of one or more environmental components

or conditions on an activity of *S. cerevisiae*. As set forth above, an exchange reaction can be added to a reaction network data structure corresponding to uptake of an environmental component, release of a component to the environment, or other environmental demand. The effect of the environmental component or condition can be further investigated by running simulations with adjusted α_j or β_j values for the metabolic flux vector of the exchange reaction target reaction to reflect a finite maximum or minimum flux value corresponding to the effect of the environmental component or condition. The environmental component can be, for example an alternative carbon source or a metabolite that when added to the environment of *S. cerevisiae* can be taken up and metabolized. The environmental component can also be a combination of components present for example in a minimal medium composition. Thus, the methods can be used to determine an optimal or minimal medium composition that is capable of supporting a particular activity of *S. cerevisiae*.

[0096] The invention further provides a method for determining a set of environmental components to achieve a desired activity for *S. cerevisiae*. The method includes the steps of (a) providing a data structure relating a plurality of *S. cerevisiae* reactants to a plurality of *S. cerevisiae* reactions, wherein each of the *S. cerevisiae* reactions includes a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating the substrate and the product; (b) providing a constraint set for the plurality of *S. cerevisiae* reactions; (c) applying the constraint set to the data representation, thereby determining the activity of one or more *S. cerevisiae* reactions (d) determining the activity of one or more *S. cerevisiae* reactions according to steps (a) through (c), wherein the constraint set includes an upper or lower bound on the amount of an environmental component and (e) repeating steps (a) through (c) with a changed constraint set, wherein the activity determined in step (e) is improved compared to the activity determined in step (d).

[0097] The following examples are intended to illustrate but not limit the present invention.

EXAMPLE I

Reconstruction of the metabolic network of *S. cerevisiae*

[0098] This example shows how the metabolic network of *S. cerevisiae* can be reconstructed.

[0099] The reconstruction process was based on a comprehensive search of the current knowledge of metabolism in *S. cerevisiae* as shown in Figure 5. A reaction database was built using the available genomic and metabolic information on the presence, reversibility, localization and cofactor requirements of all known reactions. Furthermore, information on non-growth-dependent and growth-dependent ATP requirements and on the biomass composition was used.

[0100] For this purpose different online reaction databases, recent publications and review papers (Table 5 and 9), and established biochemistry textbooks (Zubay, Biochemistry Wm.C. Brown Publishers, Dubuque, IA (1998); Stryer, Biochemistry W.H. Freeman, New York, NY (1988)) were consulted. Information on housekeeping genes of *S. cerevisiae* and their functions were taken from three main yeast on-line resources:

- The MIPS Comprehensive Yeast Genome Database (CYGD) (Mewes et al., Nucleic Acids Research 30(1): 31-34 (2002));
- The *Saccharomyces* Genome Database (SGD) (Cherry et al., Nucleic Acids Research 26(1): 73-9 (1998));
- The Yeast Proteome Database (YPD) (Costanzo et al., Nucleic Acids Research 29(1): 75-9 (2001)).

[0101] The following metabolic maps and protein databases (available online) were investigated:

- Kyoto Encyclopedia of Genes and Genomes database (KEGG) (Kanehisa et al., Nucleic Acids Research 28(1): 27-30 (2000));
- The Biochemical Pathways database of the Expert Protein Analysis System database (ExPASy) (Appel et al., Trends Biochem Sci. 19(6): 258-260 (1994));
- ERGO from Integrated Genomics (www.integratedgenomics.com)
- SWISS-PROT Protein Sequence database (Bairoch et al., Nucleic Acids Research 28(1): 45-48 (2000)).

[0102] Table 5 lists additional key references that were consulted for the reconstruction of the metabolic network of *S. cerevisiae*.

Table 5

Amino acid biosynthesis

Strathern et al., The Molecular biology of the yeast *Saccharomyces* : metabolism and gene expression Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1982))

Lipid synthesis

- Daum et al., Yeast 14(16): 1471-510 (1998);
 Dickinson et al., The metabolism and molecular physiology of *Saccharomyces cerevisiae* Taylor & Francis, London; Philadelphia (1999);
 Dickson et al., Methods Enzymol. 311:3-9 (2000);
 Dickson, Annu Rev Biochem 67: 27-48 (1998);
 Parks, CRC Crit Rev Microbiol 6(4): 301-41 (1978))

Nucleotide Metabolism

Strathern et al., supra (1982))

Oxidative phosphorylation and electron transport

- (Verduyn et al., Antonie Van Leeuwenhoek 59(1): 49-63 (1991);
 Overkamp et al., J. of Bacteriol 182(10): 2823-2830 (2000))

Primary Metabolism

Zimmerman et al., Yeast sugar metabolism : biochemistry, genetics, biotechnology, and applications Technomic Pub., Lancaster, PA (1997);

Dickinson et al., supra (1999);

Strathern et al., supra (1982))

Transport across the cytoplasmic membrane

Paulsen et al., FEBS Lett 430(1-2): 116-125 (1998);

Wieczorka et al., FEBS Lett 464(3): 123-128 (1999);

Regenberg et al., Curr Genet 36(6): 317-328 (1999);

Andre, Yeast 11(16): 1575-1611 (1995))

Transport across the mitochondrial membrane

Palmieri et al., J Bioenerg Biomembr 32(1): 67:77 (2000);

Palmieri et al., Biochim Biophys Acta 1459(2-3): 363-369 (2000);

Palmieri et al., J Biol Chem 274(32):22184-22190 (1999);

Palmieri et al., FEBS Lett 417(1): 114-118 (1997);

Paulsen et al., supra (1998);

Pallotta et al., FEBS Lett 428(3): 245-249 (1998);

Tzagologg et al. Mitochondria Plenum Press, New York (1982); Andre Yeast 11(16): 1575-611 (1995))

[0103] All reactions are localized into the two main compartments, cytosol and mitochondria, as most of the common metabolic reactions in *S. cerevisiae* take place in these compartments. Optionally, one or more additional compartments can be considered. Reactions located *in vivo* in other compartments or reactions for which no information was available regarding localization were assumed to be cytosol. All corresponding metabolites were assigned appropriate localization and a link between cytosol and mitochondria was established through either known transport and shuttle systems or through inferred reactions to meet metabolic demands.

[0104] After the initial assembly of all the metabolic reactions the list was manually examined for resolution of detailed biochemical issues. A large number of reactions involve cofactors utilization, and for many of these reactions the cofactor requirements have not yet been completely elucidated. For example, it is not clear whether certain reactions use only NADH or only NADPH as a cofactor or can use both cofactors, whereas other reactions are known to use both cofactors.

For example, a mitochondrial aldehyde dehydrogenase encoded by *ALD4* can use both NADH and NADPH as a cofactor (Remize et al. Appl Environ Microbiol 66(8): 3151-3159 (2000)). In such cases, two reactions are included in the reconstructed metabolic network.

[0105] Further considerations were taken into account to preserve the unique features of *S. cerevisiae* metabolism.

S. cerevisiae lacks a gene that encodes the enzyme transhydrogenase. Insertion of a corresponding gene from *Azotobacter vinelandii* in *S. cerevisiae* has a major impact on its phenotypic behavior, especially under anaerobic conditions (Niessen et al. Yeast 18(1): 19-32 (2001)). As a result, reactions that create a net transhydrogenic effect in the model were either constrained to zero or forced to become irreversible. For instance, the flux carried by NADH dependent glutamate dehydrogenase (Gdh2p) was constrained to zero to avoid the appearance of a net transhydrogenase activity through coupling with the NADPH dependent glutamate dehydrogenases (Gdh1p and Gdh3p).

[0106] Once a first generation model is prepared, microbial behavior can be modeled for a specific scenario, such as anaerobic or aerobic growth in continuous cultivation using glucose as a sole carbon source. Modeling results can then be compared to experimental results. If modeling and experimental results are in agreement, the model can be considered as correct, and it is used for further modeling and predicting *S. cerevisiae* behavior. If the modeling and experimental results are not in agreement, the model has to be evaluated and the reconstruction process refined to determine missing or incorrect reactions, until modeling and experimental results are in agreement. This iterative process is shown in Figure 5 and exemplified below.

EXAMPLE II

Calculation of the P/O ratio

[0107] This example shows how the genome-scale reconstructed metabolic model of *S. cerevisiae* was used to calculate the P/O ratio, which measures the efficiency of aerobic respiration. The P/O ratio is the number of ATP molecules produced per pair of electrons donated to the electron transport system (ETS).

[0108] Linear optimization was applied, and the *in silico* P/O ratio was calculated by first determining the maximum number of ATP molecules produced per molecule of glucose through the electron transport system (ETS), and then interpolating the *in silico* P/O ratio using the theoretical relation (i.e. in *S. cerevisiae* for the P/O ratio of 1.5, 18 ATP molecules are produced).

[0109] Experimental studies of isolated mitochondria have shown that *S. cerevisiae* lacks site I proton translocation (Verduyn et al., Antonie Van Leeuwenhoek 59(1): 49-63 (1991)). Consequently, estimation of the maximum theoretical or "mechanistic" yield of the ETS alone gives a P/O ratio of 1.5 for oxidation of NADH in *S. cerevisiae* grown on glucose (Verduyn et al., *supra* (1991)). However, based on experimental measurements, it has been determined that the net *in vivo* P/O ratio is approximately 0.95 (Verduyn et al., *supra* (1991)). This difference is generally thought to be due to the use of the mitochondrial transmembrane proton gradient needed to drive metabolite exchange, such as the proton-coupled translocation of pyruvate, across the inner mitochondrial membrane. Although simple diffusion of protons (or proton leakage) would be surprising given the low solubility of protons in the lipid bilayer, proton leakage is considered to contribute to the lowered P/O ratio due to the relatively high electrochemical gradient across the inner mitochondrial membrane (Westerhoff and van Dam, Thermodynamics and control of biological free-energy transduction Elsevier, New York, NY (1987)).

[0110] Using the reconstructed network, the P/O ratio was calculated to be 1.04 for oxidation of NADH for growth on glucose by first using the model to determine the maximum number of ATP molecules produced per molecule of glucose through the electron transport system (ETS) (YATP,max=12.5 ATP molecules/glucose molecule via ETS *in silico*). The *in silico* P/O ratio was then interpolated using the theoretical relation (i.e. 18 ATP molecules per glucose molecule are produced theoretically when the P/O ratio is 1.5). The calculated P/O ratio was found to be close to the experimentally determined value of 0.95. Proton leakage, however, was not included in the model, which suggests that the major reason for the lowered P/O ratio is the use of the proton gradient for solute transport across the inner mitochondrial membrane. This result illustrates the importance of including the complete metabolic network in the analysis, as the use of the proton gradient for solute transport across the mitochondrial membrane contributes significantly to the operational P/O ratio.

EXAMPLE III

Phenotypic phase plane analysis

[0111] This example shows how the *S. cerevisiae* metabolic model can be used to calculate the range of characteristic phenotypes that the organism can display as a function of variations in the activity of multiple reactions.

[0112] For this analysis, O₂ and glucose uptake rates were defined as the two axes of the two-dimensional space. The optimal flux distribution was calculated using linear programming (LP) for all points in this plane by repeatedly solving

the LP problem while adjusting the exchange fluxes defining the two-dimensional space. A finite number of quantitatively different metabolic pathway utilization patterns were identified in the plane, and lines were drawn to demarcate these regions. One demarcation line in the phenotypic phase plane (PhPP) was defined as the line of optimality (LO), and represents the optimal relation between the respective metabolic fluxes. The LO was identified by varying the x-axis (glucose uptake rate) and calculating the optimal y-axis (O_2 uptake rate), with the objective function defined as the growth flux. Further details regarding Phase-Plane Analysis are provided in Edwards et al., Biotechnol. Bioeng. 77:27-36 (2002) and Edwards et al., Nature Biotech. 19:125-130 (2001)).

[0113] As illustrated in Figure 6, the *S. cerevisiae* PhPP contains 8 distinct metabolic phenotypes. Each region (P1-P8) exhibits unique metabolic pathway utilization that can be summarized as follows:

[0114] The left-most region is the so-called "infeasible" steady state region in the PhPP, due to stoichiometric limitations.

[0115] From left to right:

[0116] P1: Growth is completely aerobic. Sufficient oxygen is available to complete the oxidative metabolism of glucose to support growth requirements. This zone represents a futile cycle. Only CO_2 is formed as a metabolic by-product. The growth rate is less than the optimal growth rate in region P2. The P1 upper limit represents the locus of points for which the carbon is completely oxidized to eliminate the excess electron acceptor, and thus no biomass can be generated.

[0117] P2: Oxygen is slightly limited, and all biosynthetic cofactor requirements cannot be optimally satisfied by oxidative metabolism. Acetate is formed as a metabolic by-product enabling additional high-energy phosphate bonds via substrate level phosphorylation. With the increase of O_2 supply, acetate formation eventually decreases to zero.

[0118] P3: Acetate is increased and pyruvate is decreased with increase in oxygen uptake rate.

[0119] P4: Pyruvate starts to increase and acetate is decreased with increase in oxygen uptake rate. Ethanol production eventually decreases to zero.

[0120] P5: The fluxes towards acetate formation are increasing and ethanol production is decreasing.

[0121] P6: When the oxygen supply increases, acetate formation increases and ethanol production decreases with the carbon directed toward the production of acetate. Besides succinate production, malate may also be produced as metabolic by-product.

[0122] P7: The oxygen supply is extremely low, ethanol production is high and succinate production is decreased. Acetate is produced at a relatively low level.

[0123] P8: This region is along the Y-axis and the oxygen supply is zero. This region represents completely anaerobic fermentation. Ethanol and glycerol are secreted as a metabolic by-product. The role of NADH-consuming glycerol formation is to maintain the cytosol redox balance under anaerobic conditions (Van Dijken and Scheffers Yeast 2(2): 123-7 (1986)).

[0124] **Line of Optimality:** Metabolically, the line of optimality (LO) represents the optimal utilization of the metabolic pathways without limitations on the availability of the substrates. On an oxygen/glucose phenotypic phase plane diagram, LO represents the optimal aerobic glucose-limited growth of *S. cerevisiae* metabolic network to produce biomass from unlimited oxygen supply for the complete oxidation of the substrates in the cultivation processes. The line of optimality therefore represents a completely respiratory metabolism, with no fermentation by-product secretion and the futile cycle fluxes equals zero.

[0125] Thus, this example demonstrates that Phase Plane Analysis can be used to determine the optimal fermentation pattern for *S. cerevisiae*, and to determine the types of organic byproducts that are accumulated under different oxygenation conditions and glucose uptake rates.

EXAMPLE IV

Calculation of line of optimality and respiratory quotient

[0126] This example shows how the *S. cerevisiae* metabolic model can be used to calculate the oxygen uptake rate (OUR), the carbon dioxide evolution rate (CER) and the respiration quotient (RQ), which is the ratio of CER over OUR.

[0127] The oxygen uptake rate (OUR) and the carbon dioxide evolution rate (CER) are direct indicators of the yeast metabolic activity during the fermentation processes. RQ is a key metabolic parameter that is independent of cell number.

[0128] As illustrated in Figure 7, if the *S. cerevisiae* is grown along the line of optimality, LO, its growth is at optimal aerobic rate with all the carbon sources being directed to biomass formation and there are no metabolic by-products secreted except CO_2 . The calculated RQ along the LO is a constant value of 1.06; the RQ in P1 region is less than 1.06; and the RQ in the remaining regions in the yeast PhPP are greater than 1.06. The RQ has been used to determined the cell growth and metabolism and to control the glucose feeding for optimal biomass production for decades (Zeng et al. Biotechnol. Bioeng. 44:1107-1114 (1994)). Empirically, several researchers have proposed the values of 1.0 (Zigova, J Biotechnol 80: 55-62 (2000). Journal of Biotechnology), 1.04 (Wang et al., Biotechnol & Bioeng 19:69-86 (1977)) and 1.1 (Wang et al., Biotechnol. & Bioeng. 21:975-995 (1979)) as optimal RQ which should be maintained in fed-batch or continuous production of yeast's biomass so that the highest yeast biomass could be obtained (Dantigny et al., Appl.

Microbiol. Biotechnol. 36:352-357 (1991)). The constant RQ along the line of optimality for yeast growth by the metabolic model is thus consistent with the empirical formulation of the RQ through on-line measurements from the fermentation industry.

5 EXAMPLE V

Computer simulations

[0128] This example shows computer simulations for the change of metabolic phenotypes described by the yeast PhPP.

[0129] A piece-wise linearly increasing function was used with the oxygen supply rates varying from completely anaerobic to fully aerobic conditions (with increasing oxygen uptake rate from 0 to 20 mmol per g cell-hour). A glucose uptake rate of 5 mmol of glucose per g (dry weight)-hour was arbitrarily chosen for these computations. As shown in Figure 8A, the biomass yield of the *in silico* *S. cerevisiae* strain was shown to increase from P8 to P2, and become optimal on the LO. The yield then started to slowly decline in P1 (futile cycle region). At the same time, the RQ value declines in relation to the increase of oxygen consumption rate, reaching a value of 1.06 on the LO1 and then further declining to become less than 1.

[0130] Figure 8B shows the secretion rates of metabolic by-products; ethanol, succinate, pyruvate and acetate with the change of oxygen uptake rate from 0 to 20 mmol of oxygen per g (dry weight)-h. Each one of these by-products is secreted in a fundamentally different way in each region. As oxygen increases from 0 in P7, glycerol production (data not shown in this figure) decreases and ethanol production increases. Acetate and succinate are also secreted.

EXAMPLE VI

Modeling of phenotypic behavior in chemostat cultures

[0131] This example shows how the *S. cerevisiae* metabolic model can be used to predict optimal flux distributions that would optimize fermentation performance, such as specific product yield or productivity. In particular, this example shows how flux based analysis can be used to determine conditions that would minimize the glucose uptake rate of *S. cerevisiae* grown on glucose in a continuous culture under anaerobic and under aerobic conditions.

[0132] In a continuous culture, growth rate is equivalent to the dilution rate and is kept at a constant value. Calculations of the continuous culture of *S. cerevisiae* were performed by fixing the *in silico* growth rate to the experimentally determined dilution rate, and minimizing the glucose uptake rate. This formulation is equivalent to maximizing biomass production given a fixed glucose uptake value and was employed to simulate a continuous culture growth condition. Furthermore, a non growth dependent ATP maintenance of 1 mmol/gDW, a systemic P/O ratio of 1.5 (Verduyn et al. Antonie Van Leeuwenhoek 59(1): 49-63 (1991)), a polymerization cost of 23.92 mmol ATP/gDW, and a growth dependent ATP maintenance of 35.36 mmol ATP/gDW, which is simulated for a biomass yield of 0.51 gDW/h, are assumed. The sum of the latter two terms is included into the biomass equation of the genome-scale metabolic model.

[0133] Optimal growth properties of *S. cerevisiae* were calculated under anaerobic glucose-limited continuous culture at dilution rates varying between 0.1 and 0.4 h⁻¹. The computed by-product secretion rates were then compared to the experimental data (Nissen et al. Microbiology 143(1): 203-18 (1997)). The calculated uptake rates of glucose and the production of ethanol, glycerol, succinate, and biomass are in good agreement with the independently obtained experimental data (Figure 9). The relatively low observed acetate and pyruvate secretion rates were not predicted by the *in silico* model since the release of these metabolites does not improve the optimal solution of the network.

[0134] It is possible to constrain the *in silico* model further to secrete both, pyruvate and acetate at the experimental level and recompute an optimal solution under these additional constraints. This calculation resulted in values that are closer to the measured glucose uptake rates (Figure 9A). This procedure is an example of an iterative data-driven constraint-based modeling approach, where the successive incorporation of experimental data is used to improve the *in silico* model. Besides the ability to describe the overall growth yield, the model allows further insight into how the metabolism operates. From further analysis of the metabolic fluxes at anaerobic growth conditions the flux through the glucose-6-phosphate dehydrogenase was found to be 5.32% of the glucose uptake rate at dilution rate of 0.1 h⁻¹, which is consistent with experimentally determined value (6.34%) for this flux when cells are operating with fermentative metabolism (Nissen et al., Microbiology 143(1): 203-218 (1997)).

[0135] Optimal growth properties of *S. cerevisiae* were also calculated under aerobic glucose-limited continuous culture in which the Crabtree effect plays an important role. The molecular mechanisms underlying the Crabtree effect in *S. cerevisiae* are not known. The regulatory features of the Crabtree effect (van Dijken et al. Antonie Van Leeuwenhoek 63(3-4):343-52 (1993)) can, however, be included in the *in silico* model as an experimentally determined growth rate-dependent maximum oxygen uptake rate (Overkamp et al. J. of Bacteriol 182(10): 2823-30 (2000))). With this additional constraint and by formulating growth in a chemostat as described above, the *in silico* model makes quantitative predictions

about the respiratory quotient, glucose uptake, ethanol, CO₂, and glycerol secretion rates under aerobic glucose-limited continuous condition (Fig. 10).

EXAMPLE VII

5

Analysis of deletion of genes involved in central metabolism in *S. cerevisiae*

[0136] This example shows how the *S. cerevisiae* metabolic model can be used to determine the effect of deletions of individual reactions in the network.

10 [0137] Gene deletions were performed *in silico* by constraining the flux(es) corresponding to a specific gene to zero. The impact of single gene deletions on growth was analysed by simulating growth on a synthetic complete medium containing glucose, amino acids, as well as purines and pyrimidines.

15 [0138] *In silico* results were compared to experimental results as supplied by the *Saccharomyces* Genome Database (SGD) (Cherry et al., Nucleic Acids Research 26(1):73-79 (1998)) and by the Comprehensive Yeast Genome Database (Mewes et al., Nucleic Acids Research 30(1):31-34 (2002)). In 85.6% of all considered cases (499 out of 583 cases), the *in silico* prediction was in qualitative agreement with experimental results. An evaluation of these results can be found in Example VIII. For central metabolism, growth was predicted under various experimental conditions and 81.5% (93 out of 114 cases) of the *in silico* predictions were in agreement with *in vivo* phenotypes.

20 [0139] Table 6 shows the impact of gene deletions on growth in *S. cerevisiae*. Growth on different media was considered, including defined complete medium with glucose as the carbon source, and minimal medium with glucose, ethanol or acetate as the carbon source. The complete reference citations for Table 6 can be found in Table 9.

[0140] Thus, this example demonstrates that the *in silico* model can be used to uncover essential genes to augment or circumvent traditional genetic studies.

25

Table 6

Defined Medium	Complete	Minimal	Minimal	Minimal	References:
Carbon Source	Glucose	Glucose	Acetate	Ethanol	(Minimal media)
Gene	<i>in silico</i> / <i>in vivo</i>				
<i>ACO1</i>	+/+	-/-			(Gangloff et al., 1990)
<i>CDC19#</i>	+/-	+/-			(Boles et al., 1998)
<i>CIT1</i>	+/+	+/+			(Kim et al., 1986)
<i>CIT2</i>	+/+	+/+			(Kim et al., 1986)
<i>CIT3</i>	+/+				
<i>DAL7</i>	+/+	+/+	+/+	+/+	(Hartig et al., 1992)
<i>ENO1</i>	+/+				
<i>ENO2\$\$</i>	+/-	+/-			
<i>FBA1*</i>	+/-	+/-			
<i>FBP1</i>	+/+	+/+		-/-	(Sedivy and Fraenkel, 1985; Gancedo and Delgado, 1984)
<i>FUM1</i>	+/+				
<i>GLK1</i>	+/+				
<i>GND1##</i>	+/-	+/-			
<i>GND2</i>	+/+				
<i>GPM1 </i>	+/-	+/-			
<i>GPM2</i>	+/+				
<i>GPM3</i>	+/+				
<i>HXK1</i>	+/+				
<i>HXK2</i>	+/+				
<i>ICL1</i>	+/+	+/+			(Smith et al., 1996)
<i>IDH1</i>	+/+	+/+			(Cupp and McAlister-Henn, 1992)
<i>IDH2</i>	+/+	+/+			(Cupp and McAlister-Henn, 1992)
<i>IDP1</i>	+/+	+/+			(Loftus et al., 1994)
<i>IDP2</i>	+/+	+/+			(Loftus et al., 1994)

(continued)

Defined Medium	Complete Glucose	Minimal Glucose	Minimal Acetate	Minimal Ethanol	
Carbon Source					
5 <i>IDP3</i>	+/-				
<i>KGD1</i>	+/-	+/-			(Repetto and Tzagoloff, 1991)
<i>KGD2</i>	+/-	+/-			(Repetto and Tzagoloff, 1991)
<i>LPD1</i>	+/-				
10 <i>LSC1</i>	+/-		+/-	+/-	(Przybyla-Zawislak et al., 1998)
<i>LSC2</i>	+/-		+/-	+/-	(Przybyla-Zawislak et al., 1998)
<i>MAE1</i>	+/-	+/-		+/-	(Boles et al., 1998)
<i>MDH1</i>	+/-	+/-	+/-		(McAlister-Henn and Thompson, 1987)
<i>MDH2</i>	+/-		+/-	+/-	(McAlister-Henn and Thompson, 1987)
15 <i>MDH3</i>	+/-				
<i>MLS1</i>	+/-	+/-	+/-	+/-	(Hartig et al., 1992)
<i>OSM1</i>	+/-				
<i>PCK1</i>	+/-				
20 <i>PDC1</i>	+/-	+/-			(Flikweert et al., 1996)
<i>PDC5</i>	+/-	+/-			(Flikweert et al., 1996)
<i>PDC6</i>	+/-	+/-			(Flikweert et al., 1996)
<i>PFK1</i>	+/-	+/-			(Clifton and Fraenkel, 1982)
<i>PFK2</i>	+/-	+/-			(Clifton and Fraenkel, 1982)
25 <i>PGI1* &</i>	+/-	+/-			(Clifton et al., 1978)
<i>PGK1*</i>	+/-	+/-			
<i>PGM1</i>	+/-	+/-			(Boles et al., 1994)
<i>PGM2</i>	+/-	+/-			(Boles et al., 1994)
30 <i>PYC1</i>	+/-	+/-	+/-	+/-	(Wills and Melham, 1985)
<i>PYC2</i>	+/-				
<i>PYK2</i>	+/-	+/-		+/-	(Boles et al., 1998; McAlister-Henn and Thompson, 1987)
<i>RKI1</i>	-/-				
35 <i>RPE1</i>	+/-				
<i>SOL1</i>	+/-				
<i>SOL2</i>	+/-				
<i>SOL3</i>	+/-				
<i>SOL4</i>	+/-				
40 <i>TAL1</i>	+/-	+/-			(Schaaff-Gerstenschläger and Zimmermann, 1993)
<i>TDH1</i>	+/-				
<i>TDH2</i>	+/-				
<i>TDH3</i>	+/-				
45 <i>TKL1</i>	+/-	+/-			(Schaaff-Gerstenschläger and Zimmermann, 1993)
<i>TKL2</i>	+/-				
<i>TPI1*,\\$</i>	+/-				

50

55

(continued)

Defined Medium	Complete	Minimal	Minimal	Minimal	
Carbon Source	Glucose	Glucose	Acetate	Ethanol	
5 ZWF1	+/-	+/-	(Schaaff Gerstenschläger and Zimmermann, 1993)		

+/- Growth/no growth

The isoenzyme Pyk2p is glucose repressed, and cannot sustain growth on glucose.

* Model predicts single deletion mutant to be (highly) growth retarded.

\$ Growth of single deletion mutant is inhibited by glucose.

& Different hypotheses exist for why Pgi1p deficient mutants do not grow on glucose, e.g. the pentose phosphate pathway in *S. cerevisiae* is insufficient to support growth and cannot supply the EMP pathway with sufficient amounts of fructose-6-phosphate and glyceraldehydes-3-phosphate (Boles, 1997).|| The isoenzymes Gpm2p and Gpm3p cannot sustain growth on glucose. They only show residual *in vivo* activity when they are expressed from a foreign promoter (Heinisch et al., 1998).

Gnd1p accounts for 80% of the enzyme activity. A mutant deleted in GND1 accumulates gluconate-6-phosphate, which is toxic to the cell (Schaaff-Gerstenschläger and Miosga, 1997).

\$\$ ENO1 plays central role in gluconeogenesis whereas ENO2 is used in glycolysis (Müller and Entian, 1997).

EXAMPLE VIII**Large-scale gene deletion analysis in *S. cerevisiae***

[0141] A large-scale *in silico* evaluation of gene deletions in *S. cerevisiae* was conducted using the genome-scale metabolic model. The effect of 599 single gene deletions on cell viability was simulated *in silico* and compared to published experimental results. In 526 cases (87.8%), the *in silico* results were in agreement with experimental observations when growth on synthetic complete medium was simulated. Viable phenotypes were predicted in 89.4% (496 out of 555) and lethal phenotypes are correctly predicted in 68.2% (30 out of 44) of the cases considered.

[0142] The failure modes were analyzed on a case-by-case basis for four possible inadequacies of the *in silico* model: 1) incomplete media composition; 2) substitutable biomass components; 3) incomplete biochemical information; and 4) missing regulation. This analysis eliminated a number of false predictions and suggested a number of experimentally testable hypotheses. The genome-scale *in silico* model of *S. cerevisiae* can thus be used to systematically reconcile existing data and fill in knowledge gaps about the organism.

[0143] Growth on complete medium was simulated under aerobic condition. Since the composition of a complete medium is usually not known in detail, a synthetic complete medium containing glucose, twenty amino acids (alanine, arginine, asparagine, aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophane, tyrosine, valine) and purines (adenine and guanine) as well as pyrimidines (cytosine and thymine) was defined for modeling purposes. Furthermore, ammonia, phosphate, and sulphate were supplied. The *in silico* results were initially compared to experimental data from a competitive growth assay (Winzeler et al., Science 285:901-906 (1999)) and to available data from the MIPS and SGD databases (Mewes et al., Nucleic Acids Research 30(1):31-34 (2002); Cherry et al., Nucleic Acids Research 26(1):73-79 (1998)). Gene deletions were simulated by constraining the flux through the corresponding reactions to zero and optimizing for growth as previously described (Edwards and Palsson, Proceedings of the National Academy of Sciences 97(10):5528-5533 (2000)). For this analysis, a viable phenotype was defined as a strain that is able to meet all the defined biomass requirements and thus grow. Single gene deletion mutants that have a reduced growth rate compared to the wild type simulation are referred to as growth retarded mutants.

[0144] The analysis of experimental data was approached in three steps:

■ The initial simulation using the synthetic medium described above, referred to as simulation 1.

■ False predictions of simulation 1 were subsequently examined to determine if the failure was due to incomplete information in the *in silico* model, such as missing reactions, the reversibility of reactions, regulatory events, and missing substrates in the synthetic complete medium. In simulation 2, any such additional information was introduced into the *in silico* model and growth was re-simulated for gene deletion mutants whose *in silico* phenotype was not in agreement with its *in vivo* phenotype.

■ A third simulation was carried out, in which dead end pathways (i.e. pathways leading to intracellular metabolites that were not further connected into the overall network), were excluded from the analysis (simulation 3).

[0145] The effect of single gene deletions on the viability of *S. cerevisiae* was investigated for each of the 599 single

gene deletion mutants. The *in silico* results were categorized into four groups:

1. True negatives (correctly predicted lethal phenotype);
2. False negatives (wrongly predicted lethal phenotype);
3. True positives (correctly predicted viable phenotypes);
4. False positives (wrongly predicted viable phenotypes).

[0146] In simulation 1, 509 out of 599 (85%) simulated phenotypes were in agreement with experimental data. The number of growth retarding genes in simulation 1 was counted to be 19, a surprisingly low number. Only one deletion, the deletion of *TPI1*, had a severe impact on the growth rate. Experimentally, a deletion in *TPI1* is lethal (Ciriacy and Breitenbach, J Bacteriol 139(1):152-60 (1979)). *In silico*, a *tpi1* mutant could only sustain a specific growth rate of as low as 17% of the wild type. All other growth retarding deletions sustained approximately 99% of wild type growth, with the exception of a deletion of the mitochondrial ATPase that resulted in a specific growth rate of approximately 90% of wild type.

[0147] Predictions of simulation 1 were evaluated in a detailed manner on a case-by-case basis to determine whether the false predictions could be explained by:

1. Medium composition used for the simulation;
2. The biomass composition used in the simulation;
3. Incomplete biochemical information; and
4. Effects of gene regulation.

[0148] Analysis of the false predictions from simulation 1 based on these possible failure modes resulted in model modifications that led to 526 out of 599 correctly predicted phenotypes (87.8%), i.e. simulation 2.

[0149] Simulation 3 uncovered some 220 reactions in the reconstructed network that are involved in dead end pathways. Removing these reactions and their corresponding genes from the genome-scale metabolic flux balance model, simulation 3 resulted in 473 out of 530 (89.6%) correctly predicted phenotypes of which 91.4% are true positive and 69.8% are true negative predictions.

[0150] Table 7 provides a summary of the large-scale evaluation of the effect of *in silico* single gene deletions in *S. cerevisiae* on viability.

Table 7

Simulation	1	2	Genes involved in dead end pathways	3
35	Number of deletion	599	599	530
	Predicted Total	509	526	475
40	True positive	481	496	445
	True negative	28	30	30
	False positive	63	59	42
	False negative	27	14	13
45	Overall Prediction	85.0%	87.8%	89.6%
	Positive Prediction	88.4%	89.4%	91.4%
	Negative Prediction	50.9%	68.2%	69.8%

[0151] A comprehensive list of all the genes used in the *in silico* deletion studies and results of the analysis are provided in Table 8. Table 8 is organized according to the categories true negative, false negative, true positive and false positive predictions. Genes highlighted in grey boxes, such as *INO1*, corresponded initially to false predictions (simulation 1); however, evaluation of the false prediction and simulation 2 identified these cases as true predictions. ORFs or genes that are in an open box, such as *TRR2* were excluded in simulation 3, as the corresponding reactions catalysed steps in dead end pathways.

Table 8**False Positive**

5 ACS2 [AUR1] BET2 CDC19 CDC21 CDC8 CYR1 [DED81] DFR1 [DIM1] DUT1 [DYS1] ENO2 ERG10 ERG13 FADI
 FMN1 FOL1 FOL2 FOL3 GFA1 GPM1 [HEM1] [HEM12] [HEM13] [HEM15] [HEM2] [HEM3] [HEM4] HIP1 [HTS1] ILV3
 ILV5 [KRS1] LCB1 LCB2 MSS4 NAT2 NCP1 [NMT1] PCMI PET9 [PGS1] [PIK1] PMA1 PRO3 QNS1 QRI1 RER2 RIB5
 SEC59 STT4 THI80 TOR2 TPI TSC10 UGP1 URA6 YDR341C YGL245W

False Negative

10 ADE3 ADK1 CHO1 CHO2 DPP1 ERG3 ERG4 ERG5 ERG6 INMI MET6 OPI3 PPT2 YNK1

True Negative

15 ACC1 [ADE13] CDS1 DPM1 ERG1 ERG7 ERG8 ERG9 ERG11 ERG12 ERG20 ERG25 ERG26 ERG27 FBA1 [GLN1]
 GUK1 IDII IPP1 MVD1 PG11 PGK1 PIS1 PMI40 PSA1 RK11 SAH1 SEC53 TRR1 YDR531W

True Positive

20 AAC1 AAC3 AAH1 AAT1 AAT2 ABZ1 ACO1 ACS1 ADE1 ADE12 ADE16 ADE17 ADE2 ADE4 ADE5 ADE6 ADE7
 ADE8 ADH1 ADH2 ADH3 ADH4 ADH5 ADK2 AGP1 [AGP2] AGP3 ALD2 ALD3 ALD4 ALD5 ALD6 ALP1 ASP1
 ATH1 ATP1 BAP2 BAP3 BAT1 BAT2 BGL2 [BIO2] [BIO3] [BIO4] [BIO5] [BNA1] CAN1 CAR1 CAR2 CAT2 CDA1 CDA2
 CDD1 [CEM1] CHA1 CHS1 CHS2 CHS3 CIT1 CIT2 CIT3 CK11 COQ1 COQ2 [COQ3] [COQ5] [COQ6] COXI COX10
 CPA2 [CPT1] CRC1 [CRD1] CSG2 CTA1 CTP1 CTT1 [CYB2] CYS3 CYS4 DAK1 DAK2 DAL1 DAL2 DAL3 DAL4
 25 DAL5 DAL7 DCD1 DEG1 DIC1 DIP5 DLD1 [DPH5] DPL1 DURI DUR3 ECM17 [ECM31] ECM40 ECT1 [EK11]

EN01 [EPT1] ERG2 ERG24 ERR1 ERR2 EXG1 EXG2 FAA1 FAA2 FAA3 FAA4 FAB1 [FAS1] FBP1 FBP26 FCY1
 FCY2 FKS1 FKS3 FLX1 [FMT1] FOX2 FRDS FU11 FUM1 FUN63 FURI FUR4 GAD1 GAL1 GAL10 GAL2 GAL7
 30 GAP1 GCV1 GCV2 GDH1 GDH2 GDH3 GLC3 GLK1 [GLO1] [GLO2] [GLO4] GLRI GLT1 GLY1 GNA1 GND1 GND2
 GNP1 GPD1 GPD2 GPH1 GPM2 GPM3 GPX1 GPX2 GSC2 GSH1 GSH2 GSY1 GSY2 GUA1 GUT1 GUT2 [HEM14]
 HIS1 HIS2 HIS3 HIS4 HIS5 HIS6 HIS7 HMG1 HMG2 [HMT1] HM1 HOM2 HOM3 HOM6 HOR2 HPT1 HXK1
 HXK2 HXT1 HXT10 HXT11 HXT13 HXT14 HXT15 HXT16 HXT17 HXT2 HXT3 HXT4 HXT5 HXT6 HXT7 HXT8
 35 HXT9 HYR1 ICL1 ICL2 IDH1 IDP1 IDP2 IDP3 ILV1 ILV2 [INO1] [PT1] ITR1 ITR2 JEN1 KGD1 KRE2 KTR1 KTR2
 KTR3 KTR4 KTR6 LCB3 LCB4 LCB5 [LEU1] LEU2 LEU4 [LPD1] [LPP1] LSC1 LSC2 LYP1 LYS1 LYS12 LYS2 LYS20
 LYS21 LYS4 LYS9 MAE1 MAK3 MAL12 MAL31 MAL32 MDH1 MDH2 MDH3 MEL1 MEP1 MEP2 MEP3 [MET1]
 MET10 MET12 MET13 MET14 MET16 MET17 MET2 MET22 MET3 MET7 MHT1 MIR1 MJS1 MLS1 [MMP1] [MSE1]
 40 MSK1 MSR1 [MSW1] MTD1 MUP1 MUP3 NAT1 NDH1 NDH2 ND11 [NHA1] [NIT2] NPT1 NTA1 NTH1 NTH2 OAC1
 ODC1 ODC2 ORT1 OSM1 PAD1 PCK1 [PCT1] PDA1 PDC1 PDC5 PDC6 PDE1 PDE2 PDX3 PFK1 PFK2 PFK26
 PFK27 PGM1 PGM2 PHA2 PHO8 PHO11 [PHO84] [PLC1] PMA2 PMP1 PMP2 PMT1 PMT2 PMT3 PMT4 PMT5
 PMT6 [PNC1] PNP1 POS5 [POT1] PPA2 PRM4 PRM5 PRM6 PRO1 PRO2 PRS1 PRS2 PRS3 PRS4 PRS5 [PSD1]
 PSD2 PTR2 PUR5 PUS1 PUS2 PUS4 PUT1 PUT2 PUT4 PYC1 PYC2 PYK2 QPT1 RAM1 RBK1 RHR2 RIB1 RIB4
 RIB7 RMA1 RNR1 RNR3 RPE1 SAM1 SAM2 SAM3 SAM4 [SCS7] SDH3 SER1 SER2 SER3 SER33 SFA1 SFC1 SHM1
 45 SHM2 SLC1 SOL1 SOL2 SOL3 SOL4 [SOR1] SPE1 SPE2 SPE3 SPE4 SPR1 SRT1 STL1 SUC2 SUL1 SUL2 SUR1
 SUR2 TAL1 TAT1 TAT2 TDH1 TDH2 TDH3 THI20 THI21 THI22 THI17 THI7 THM2 THM3 THR1 THR4 TKL1 TKL2
 TOR1 TPS1 [TPS2] TPS3 [TRK1] TRP1 TRP2 TRP3 TRP4 TRP5 [TRR2] TSL1 TYR1 UGA1 UGA4 [URA1] [URA2] [URA3]
 URA4 URA5 URA7 URA8 URA10 URH1 URK1 UTR1 VAP1 VPS34 XPT1 YAT1 YSR3 YUR1 ZWF1 [YBL098W]
 50 YBR006W YBR284W YDL100C YDR111C YEL041W YER053C YFL030W YFR055W YGR012W YGR043C YGR125W
 YGR287C [YIL145C] YIL167W YJL070C YJL200C YJL216C YJL218W YJR078W [YLR089C] YLR231C YLR328W
 YML082W YMR293C

[0152] The following text describes the analysis of the initially false predictions of simulation 1 that were performed, leading to simulation 2 results.

Influence of media composition on simulation results:

[0153] A rather simple synthetic complete medium composition was chosen for simulation 1. The *in silico* medium contained only glucose, amino acids and nucleotides as the main components. However, complete media often used for experimental purposes, e.g. the YPD medium containing yeast extract and peptone, include many other components, which are usually unknown.

[0154] **False negative predictions:** The phenotype of the following deletion mutants: *ecm1Δ*, *yil145cΔ*, *erg2Δ*, *erg24Δ*, *fas1Δ*, *ura1Δ*, *ura2Δ*, *ura3Δ* and *ura4Δ* were falsely predicted to be lethal in simulation 1. In simulation 2, an additional supplement of specific substrate could rescue a viable phenotype *in silico* and as the supplemented substrate may be assumed to be part of a complex medium, the predictions were counted as true positive predictions in simulation 2. For example, both Ecm1 and Yil145c are involved in pantothenate synthesis. Ecm1 catalyses the formation of dehydropantoate from 2-oxovalerate, whereas Yil145c catalyses the final step in pantothenate synthesis from β-alanine and pantoate. *In vivo*, *ecm1Δ*, and *yil145cΔ* mutants require pantothenate for growth (White et al., J Biol Chem 276(14): 10794-10800 (2001)). By supplying pantothenate to the synthetic complete medium *in silico*, the model predicted a viable phenotype and the growth rate was similar to *in silico* wild type *S. cerevisiae*.

[0155] Similarly other false predictions could be traced to medium composition:

- Mutants deleted in *ERG2* or *ERG24* are auxotroph for ergosterol (Silve et al., Mol Cell Biol 16(6): 2719-2727 (1996); Bourot and Karst, Gene 165(1): 97-102 (1995)). Simulating growth on a synthetic complete medium supplemented with ergosterol allowed the model to accurately predict viable phenotypes.
- A deletion of *FAS1* (fatty acid synthase) is lethal unless appropriate amounts of fatty acids are provided, and by addition of fatty acids to the medium, a viable phenotype was predicted.
- Strains deleted in *URA1*, *URA2*, *URA3*, or *URA4* are auxotroph for uracil (Lacroute, J Bacteriol 95(3): 824-832 (1968)), and by supplying uracil in the medium the model predicted growth.

[0156] The above cases were initially false negative predictions, and simulation 2 demonstrated that these cases were predicted as true positive by adjusting the medium composition.

[0157] **False positive predictions:** Simulation 1 also contained false positive predictions, which may be considered as true negatives or as true positives. Contrary to experimental results from a competitive growth assay (Winzeler et al., Science 285: 901-906 (1999)), mutants deleted in *ADE13* are viable *in vivo* on a rich medium supplemented with low concentrations of adenine, but grow poorly (Guetsova et al., Genetics 147(2): 383-397 (1997)). Adenine was supplied in the *in silico* synthetic complete medium. By not supplying adenine, a lethal mutant was predicted. Therefore, this case was considered as a true negative prediction.

[0158] A similar case was the deletion of *GLN1*, which codes a glutamine synthase, the only pathway to produce glutamine from ammonia. Therefore, *gln1Δ* mutants are glutamine auxotroph (Mitchell, Genetics 111(2):243-58 (1985)). In a complex medium, glutamine is likely to be deaminated to glutamate, particularly during autoclaving. Complex media are therefore likely to contain only trace amounts of glutamine, and *gln1Δ* mutants are therefore not viable. However, *in silico*, glutamine was supplied in the complete synthetic medium and growth was predicted. By not supplying glutamine to the synthetic complete medium, the model predicted a lethal phenotype resulting in a true negative prediction.

[0159] *Ilv3* and *Ilv5* are both involved in branched amino acid metabolism. One may expect that a deletion of *ILV3* or *ILV5* could be rescued with the supply of the corresponding amino acids. For this, the model predicted growth. However, contradictory experimental data exists. In a competitive growth assay lethal phenotypes were reported. However, earlier experiments showed that *ilv3Δ* and *ilv5Δ* mutants could sustain growth when isoleucine and valine were supplemented to the medium, as for the complete synthetic medium. Hence, these two cases were considered to be true positive predictions.

Influence of the definition of the biomass equation

[0160] The genome-scale metabolic model contains the growth requirements in the form of biomass composition. Growth is defined as a drain of building blocks, such as amino acids, lipids, nucleotides, carbohydrates, etc., to form biomass. The number of biomass components is 44 (see Table 1). These building blocks are essential for the formation of cellular components and they have been used as a fixed requirement for growth in the *in silico* simulations. Thus, each biomass component had to be produced by the metabolic network otherwise the organism could not grow *in silico*. *In vivo*, one often finds deletion mutants that are not able to produce the original biomass precursor or building block; however, other metabolites can replace these initial precursors or building blocks. Hence, for a number of strains a wrong phenotype was predicted *in silico* for this reason.

[0161] Phosphatidylcholine is synthesized by three methylation steps from phosphatidylethanolamine (Dickinson and Schweizer, The metabolism and molecular physiology of *Saccharomyces cerevisiae* Taylor & Francis, London ; Phila-

5 delphia (1999)). The first step in the synthesis of phosphatidylcholine from phosphatidylethanolamine is catalyzed by a methyltransferase encoded by *CHO2* and the latter two steps are catalyzed by phospholipid methyltransferase encoded by *OPI3*. Strains deleted in *CHO2* or *OPI3* are viable (Summers et al., *Genetics* 120(4): 909-922 (1988); Daum et al., *Yeast* 14(16): 1471-1510 (1998)); however, either null mutant accumulates mono- and dimethylated phosphatidylethanolamine under standard conditions and display greatly reduced levels of phosphatidylcholine (Daum et al., *Yeast* 15 (7): 601-614 (1999)). Hence, phosphatidylethanolamine can replace phosphatidylcholine as a biomass component. *In silico*, phosphatidylcholine is required for the formation of biomass. One may further speculate on whether an alternative pathway for the synthesis of phosphatidylcholine is missing in the model, since Daum et al., *supra* (1999) detected small amounts of phosphatidylcholine in *cho2Δ* mutants. An alternative pathway, however, was not included in the *in silico* model.

10 15 [0162] Deletions in the ergosterol biosynthetic pathways of *ERG3*, *ERG4*, *ERG5* or *ERG6* lead *in vivo* to viable phenotypes. The former two strains accumulate ergosta-8,22,24 (28)-trien-3-beta-ol (Bard et al., *Lipids* 12(8): 645-654 (1977); Zweytick et al., *FEBS Lett* 470(1): 83-87 (2000)), whereas the latter two accumulate ergosta-5,8-dien-3beta-ol (Hata et al., *J Biochem (Tokyo)* 94(2): 501-510 (1983)), or zymosterol and smaller amounts of cholesta-5,7,24-trien-3-beta-ol and cholesta-5,7,22,24-trien-3-beta-ol (Bard et al., *supra* (1977); Parks et al., *Crit Rev Biochem Mol Biol* 34(6): 399-404 (1999)), respectively, components that were not included in the biomass equations.

20 [0163] The deletion of the following three genes led to false positive predictions: *RER2*, *SEC59* and *QIR1*. The former two are involved in glycoprotein synthesis and the latter is involved in chitin metabolism. Both chitin and glycoprotein are biomass components. However, for simplification, neither of the compounds was considered in the biomass equation. Inclusion of these compounds into the biomass equation may improve the prediction results.

Incomplete biochemical information

25 [0164] For a number of gene deletion mutants (*inm1Δ*, *met6Δ*, *ynk1Δ*, *pho84Δ*, *psd2Δ*, *tps2Δ*), simulation 1 produced false predictions that could not be explained by any of the two reasons discussed above nor by missing gene regulation (see below). Further investigation of the metabolic network including an extended investigation of biochemical data from the published literature showed that some information was missing initially in the *in silico* model or information was simply not available.

30 35 [0165] *Inm1* catalyses the ultimate step in inositol biosynthesis from inositol 1-phosphate to inositol (Murray and Greenberg, *Mol Microbiol* 36(3): 651-661 (2000)). Upon deleting *INM1*, the model predicted a lethal phenotype in contrary to the experimentally observed viable phenotype. An isoenzyme encoded by *IMP2* was initially not included in the model, which may take over the function of *INM1* and this addition would have led to a correct prediction. However, an *inm1Δ imp2Δ* *in vivo* double deletion mutant is not inositol auxotroph (Lopez et al., *Mol Microbiol* 31(4): 1255-1264 (1999)). Hence, it appears that alternative routes for the production of inositol probably exist. Due to the lack of comprehensive biochemical knowledge, effects on inositol biosynthesis and the viability of strains deleted in inositol biosynthetic genes could not be explained.

40 45 [0166] *Met6Δ* mutants are methionine auxotroph (Thomas and Surdin-Kerjan, *Microbiol Mol Biol Rev* 61(4):503-532 (1997)), and growth may be sustained by the supply of methionine or S-adenosyl-L-methionine. *In silico* growth was supported neither by the addition of methionine nor by the addition of S-adenosyl-L-methionine. Investigation of the metabolic network showed that deleting *MET6* corresponds to deleting the only possibility for using 5-methyltetrahydrofolate. Hence, the model appears to be missing certain information. A possibility may be that the carbon transfer is carried out using 5-methyltetrahydropteroyltri-L-glutamate instead of 5-methyltetrahydrofolate. A complete pathway for such a by-pass was not included in the genome-scale model.

50 55 [0167] The function of *Ynk1p* is the synthesis of nucleoside triphosphates from nucleoside diphosphates. *YNK1Δ* mutants have a 10-fold reduced *Ynk1p* activity (Fukuchi et al., *Genes* 129(1):141-146 (1993)), though this implies that there may either be an alternative route for the production of nucleoside triphosphates or a second nucleoside diphosphate kinase, even though there is no ORF in the genome with properties that indicates that there is a second nucleoside diphosphate kinase. An alternative route for the production of nucleoside triphosphate is currently unknown (Dickinson et al., *supra* (1999)), and was therefore not included in the model, hence a false negative prediction.

60 [0168] *PHO84* codes for a high affinity phosphate transporter that was the only phosphate transporter included in the model. However, at least two other phosphate transporters exist, a second high affinity and Na^+ dependent transporter *Pho89* and a low affinity transporter (Persson et al., *Biochim Biophys Acta* 1422(3): 255-72 (1999)). Due to exclusion of these transporters a lethal *pho84Δ* mutant was predicted. Including *PHO89* and a third phosphate transporter, the model predicted a viable deletion mutant.

65 70 [0169] In a null mutant of *PSD2*, phosphatidylethanolamine synthesis from phosphatidylserine is at the location of *Psd1* (Trotter et al., *J Biol Chem* 273(21): 13189-13196 (1998)), which is located in the mitochondria. It has been postulated that phosphatidylserine can be transported into the mitochondria and phosphatidylethanolamine can be transported out of the mitochondria. However, transport of phosphatidylethanolamine and phosphatidylserine over the

mitochondrial membrane was initially not included in the model. Addition of these transporters to the genome-scale flux balance model allowed *in silico* growth of a *PSD2* deleted mutant.

[0170] Strains deleted in *TPS2* have been shown to be viable when grown on glucose (Bell et al., J Biol Chem 273 (50): 33311-33319 (1998)). The reaction carried out by Tps2p was modeled as essential and as the final step in trehalose synthesis from trehalose 6-phosphate. However, the *in vivo* viable phenotype shows that other enzymes can take over the hydrolysis of trehalose 6-phosphate to trehalose from Tps2p (Bell et al., *supra* (1998)). The corresponding gene(s) are currently unknown. Inclusion of a second reaction catalyzing the final step of trehalose formation allowed for the simulation of a viable phenotype.

[0171] Strains deleted in *ADE3* (C1-tetrahydrofolate synthase) and *ADK1* (Adenylate kinase) could not be readily explained. It is possible that alternative pathways or isoenzyme-coding genes for both functions exist among the many orphan genes still present in the *S. cerevisiae*.

[0172] The reconstruction process led to some incompletely modeled parts of metabolism. Hence, a number of false positive predictions may be the result of gaps (missing reactions) within pathways or between pathways, which prevent the reactions to completely connect to the overall pathway structure of the reconstructed model. Examples include:

- Sphingolipid metabolism. It has not yet been fully elucidated and therefore was not included completely into the model nor were sphingolipids considered as building blocks in the biomass equation.
- Formation of tRNA. During the reconstruction process some genes were included responsible for the synthesis of tRNA (*DED81*, *HTS1*, *KRS1*, *YDR41C*, *YGL245W*).
- However, pathways of tRNA synthesis were not fully included.
- Heme synthesis was considered in the reconstructed model (*HEM1*, *HEM12*, *HEM13*, *HEM15*, *HEM2*, *HEM3*, *HEM4*). However no reaction was included that metabolized heme in the model.
- Hence, the incomplete structure of metabolic network may be a reason for false prediction of the phenotype of *aur1Δ*, *lcb1Δ*, *lcb2Δ*, *tsc10Δ*, *ded81Δ*, *hts1Δ*, *krs1Δ*, *ydr41cΔ*, *ygl245wΔ*, *hem1Δ*, *hem12Δ*, *hem13Δ*, *hem15Δ*, *hem2Δ*, *hem3Δ*, and *hem4Δ* deletion mutants. Reaction reversibility. The *CHO1* gene encodes a phosphatidylserine synthase, an integral membrane protein that catalyses a central step in cellular phospholipid biosynthesis. *In vivo*, a deletion in *CHO1* is viable (Winzeler et al., Science 285: 901-906 (1999)). However, mutants are auxotrophic for choline or ethanolamine on media containing glucose as the carbon source (Birner et al., Mol Biol Cell 12(4): 997-1007 (2001)).
- Nevertheless, the model did not predict growth when choline and/or ethanolamine were supplied. Further investigation of the genome-scale model showed that this might be due to defining reactions leading from phosphatidylserine to phosphatidylcholine via phosphatidylethanolamine exclusively irreversible. By allowing these reactions to be reversible, either supply of choline and ethanolamine could sustain growth *in silico*.

Gene Regulation

[0173] Whereas many false negative predictions could be explained by either simulation of growth using the incorrect *in silico* synthetic complete medium or by initially missing information in the model, many false positives may be explained by *in vivo* catabolite expression, product inhibition effects or by repressed isoenzymes, as kinetic and other regulatory constraints were not included in the genome-scale metabolic model.

[0174] A total of 17 false positive predictions could be related to regulatory events. For a deletion of *CDC19*, *ACS2* or *ENO2* one may usually expect that the corresponding isoenzymes may take over the function of the deleted genes. However, the corresponding genes, either *PYK2*, *ACS1* or *ENO1*, respectively, are subject to catabolite repression (Boles et al., J Bacteriol 179(9): 2987-2993 (1997); van den Berg and Steensma, Eur J Biochem 231(3): 704-713 (1995); Zimmerman et al., Yeast sugar metabolism : biochemistry, genetics, biotechnology, and applications Technomic Pub., Lancaster, PA (1997)). A deletion of *GPM1* should be replaced by either of the two other isoenzymes, *Gpm2* and *Gpm3*; however for the two latter corresponding gene products usually no activity is found (Heinisch et al., Yeast 14(3): 203-13 (1998)).

[0175] Falsely predicted growth phenotypes can often be explained when the corresponding deleted metabolic genes are involved in several other cell functions, such as cell cycle, cell fate, communication, cell wall integrity, etc. The following genes whose deletions yielded false positive predictions were found to have functions other than just metabolic function: *ACS2*, *BET2*, *CDC19*, *CDC8*, *CYR1*, *DIM1*, *ENO2*, *FAD1*, *GFA1*, *GPM1*, *HIP1*, *MSS4*, *PET9*, *PIK1*, *PMA1*, *STT4*, *TOR2*. Indeed, a statistical analysis of the MIPS functional catalogue (<http://mips.gsf.de/proj/yeast/>) showed that in general it was more likely to have a false prediction when the genes that had multiple functions were involved in cellular communication, cell cycling and DNA processing or control of cellular organization.

Table 9. Reference list for Table 2**[0176]**

- 5 Boles,E., Liebetrau,W., Hofmann,M. & Zimmermann,F.K. A family of hexosephosphate mutases in *Saccharomyces cerevisiae*. Eur. J. Biochem. 220, 83-96 (1994).
- 10 Boles,E. Yeast sugar metabolism. Zimmermann,F.K. & Entian,K.-D. (eds.), pp. 81-96(Technomic Publishing CO., INC., Lancaster,1997).
- 15 Boles,E., Jong-Gubbels,P. & Pronk,J.T. Identification and characterization of MAE1, the *Saccharomyces cerevisiae* structural gene encoding mitochondrial malic enzyme. J. Bacteriol. 180, 2875-2882 (1998).
- Clifton,D., Weinstock,S.B. & Fraenkel,D.G. Glycolysis mutants in *Saccharomyces cerevisiae*. Genetics 88, 1-11 (1978).
- Clifton,D. & Fraenkel,D.G. Mutant studies of yeast phosphofructokinase. Biochemistry 21, 1935-1942 (1982).
- 20 Cupp,J.R. & McAlister-Henn,L. Cloning and Characterization of the gene encoding the IDH1 subunit of NAD(+) -dependent isocitrate dehydrogenase from *Saccharomyces cerevisiae*. J. Biol. Chem. 267, 16417-16423 (1992).
- Flikweert,M.T. et al. Pyruvate decarboxylase: an indispensable enzyme for growth of *Saccharomyces cerevisiae* on glucose. Yeast 12, 247-257 (1996).
- 25 Gancedo,C. & Delgado,M.A. Isolation and characterization of a mutant from *Saccharomyces cerevisiae* lacking fructose 1,6-bisphosphatase. Eur. J. Biochem. 139,651-655 (1984).
- Gangloff,S.P., Marguet,D. & Lauquin,G.J. Molecular cloning of the yeast mitochondrial aconitase gene (ACO1) and evidence of a synergistic regulation of expression by glucose plus glutamate. Mol Cell Biol 10, 3551-3561 (1990).
- 30 Hartig,A. et al. Differentially regulated malate synthase genes participate in carbon and nitrogen metabolism of *S. cerevisiae*. Nucleic Acids Res. 20, 5677-5686 (1992).
- Heinisch,J.J., Muller,S., Schluter,E., Jacoby,J. & Rodicio,R. Investigation of two yeast genes encoding putative isoenzymes of phosphoglycerate mutase. Yeast 14, 203-213 (1998).
- 35 Kim,K.S., Rosenkrantz,M.S. & Guarente,L. *Saccharomyces cerevisiae* contains two functional citrate synthase genes. Mol. Cell Biol. 6, 1936-1942 (1986).
- Loftus,T.M., Hall,L.V., Anderson,S.L. & McAlister-Henn,L. Isolation, characterization, and disruption of the yeast gene encoding cytosolic NADP-specific isocitrate dehydrogenase. Biochemistry 33, 9661-9667 (1994).
- 40 McAlister-Henn,L. & Thompson,L.M. Isolation and expression of the gene encoding yeast mitochondrial malate dehydrogenase. J. Bacteriol. 169, 5157-5166 (1987).
- 45 Muller,S. & Entian,K.-D. Yeast sugar metabolism. Zimmermann,F.K. & Entian,K.-D. (eds.), pp. 157-170 (Technomic Publishing CO.,INC., Lancaster,1997).
- Ozcan,S., Freidel,K., Leuker,A. & Ciriacy,M. Glucose uptake and catabolite repression in dominant HTR1 mutants of *Saccharomyces cerevisiae*. J. Bacteriol. 175, 5520-5528 (1993).
- 50 Przybyla-Zawislak,B., Dennis,R.A., Zakharkin,S.O. & McCammon,M.T. Genes of succinyl-CoA ligase from *Saccharomyces cerevisiae*. Eur. J Biochem. 258, 736-743 (1998).
- 55 Repetto,B. & Tzagoloff,A. In vivo assembly of yeast mitochondrial alpha-ketoglutarate dehydrogenase complex. Mol. Cell Biol. 11, 3931-3939 (1991).
- Schaaff-Gerstenschlager,I. & Zimmermann,F.K. Pentose-phosphate pathway in *Saccharomyces cerevisiae*: anal-

- ysis of deletion mutants for transketolase, transaldolase, and glucose 6-phosphate dehydrogenase. *Curr. Genet.* 24, 373-376 (1993).
- 5 Schaaff-Gerstenschlager,I. & Miosga,T. Yeast sugar metabolism. Zimmermann,F.K. & Entian,K.-D. (eds.), pp. 271-284 (Technomic Publishing CO.,INC., Lancaster, 1997).
- Sedivy,J.M. & Fraenkel,D.G. Fructose bisphosphatase of *Saccharomyces cerevisiae*. Cloning, disruption and regulation of the FBP1 structural gene. *J. Mol. Biol.* 186, 307-319 (1985).
- 10 Smith,V., Chou,K.N., Lashkari,D., Botstein,D. & Brown,P.O. Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science* 274, 2069-2074 (1996).
- Swartz,J. A PURE approach to constructive biology. *Nat. Biotechnol.* 19, 732-733 (2001).
- 15 Wills,C. & Melham,T. Pyruvate carboxylase deficiency in yeast: a mutant affecting the interaction between the glyoxylate and Krebs cycles. *Arch. Biochem. Biophys.* 236, 782-791 (1985).

20 [0177] Throughout this application various publications have been referenced. The disclosures of these publications in their entireties are hereby incorporated by reference in this application in order to more fully describe the state of the art to which this invention pertains.

[0178] Although the invention has been described with reference to the examples provided above, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is only limited by the claims.

25 Aspects and features of the present disclosure are set out in the following numbered clauses which contain the subject matter of the claims of the parent application as filed.

1. A computer readable medium or media, comprising:
 - a) a data structure relating a plurality of *saccharomyces cerevisiae* reactants to a plurality of *Saccharomyces cerevisiae* reactions, wherein each of the said *Saccharomyces cerevisiae* reactions comprises a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product, wherein at least one of said *Saccharomyces cerevisiae* reactions is annotated to indicate an associated gene;
 - b) a gene database comprising information characterising said associated gene;
 - c) a constraint set for said plurality of *Saccharomyces cerevisiae* reactions, and
 - d) commands for determining at least one flux distribution that minimises or maximises an objective function wherein said constraint set is applied to said data representation, wherein said at least one flux distribution is predictive of a *Saccharomyces cerevisiae* physiological function.
2. The computer readable medium or media of clause 1, wherein at least one reactant in said plurality of *Saccharomyces cerevisiae* reactants or at least one reaction in said plurality of *Saccharomyces cerevisiae* reactions is annotated with an assignment to a subsystem or compartment.
3. The computer readable medium or media of clause 1, wherein said plurality of reactions comprises at least one reaction from a peripheral metabolic pathway.
4. The computer readable medium or media of clause 3, wherein said peripheral metabolic pathway is selected from the group consisting of an amino acid biosynthesis, amino acid degradation, purine biosynthesis, pyrimide biosynthesis, lipid biosynthesis, fatty acid metabolism, cofactor biosynthesis, cell wall metabolism and transport processes.
5. The computer readable medium or media of clause 1, wherein said *Saccharomyces cerevisiae* physiological function is selected from the group consisting of growth, energy production, redox equivalent production, biomass production, production of biomass precursors, production of protein, production of an amino acid, production of a purine, production of a pyrimidine, production of a lipid, production of a fatty acid, production of a cofactor, production of a cell wall component, transport of a metabolite, consumption of a carbon, nitrogen, sulphur, phosphate, hydrogen or oxygen.

6. The computer readable medium or media of clause 1, wherein *Saccharomyces cerevisiae* physiological function is selected from the group consisting of degradation of a protein, degradation of an amino acid, degradation of a purine, degradation of a pyrimidine, degradation of a lipid, degradation of a fatty acid, degradation of a cofactor and degradation of a cell wall component.

5

7. The computer readable or media of clause 1, wherein said data structure comprises a set of linear algebraic equations.

10

8. The computer readable or media of clause 1, wherein said data structure comprises a matrix.

15

9. The computer readable or media of clause 1, wherein said commands comprise an optimisation problem.

10. The computer readable or media of clause 1, wherein said commands comprise a linear program.

20

11. The computer readable medium or media of clause 2, wherein a first substrate or product in said plurality of *Saccharomyces cerevisiae* reactions is assigned to a first compartment and a second substrate or product in said plurality of *Saccharomyces cerevisiae* reactions is assigned to a second compartment.

25

12. The computer readable medium or media of clause 1, wherein a plurality of said *Saccharomyces cerevisiae* reactions is annotated to indicate a plurality of associated genes and wherein said gene database comprises information characterising said plurality of associated genes.

30

13. A computer readable medium or media, comprising:

35

a) a data structure relating a plurality of *Saccharomyces cerevisiae* reactants to a plurality of *Saccharomyces cerevisiae* reactions, wherein each of said *Saccharomyces cerevisiae* reactions comprises a reactant identified as a substrate of a reaction, a reactant identified as a product of a reaction and a stoichiometric coefficient relating said substrate and said product;

b) a constraint set for said plurality of *Saccharomyces cerevisiae* reactions, and

40

c) commands for determining at least one flux distribution that minimises or maximises an objective function when said constraint set is applied to said data representation, wherein said at least one flux distribution is predictive of *Saccharomyces cerevisiae* growth.

45

14. A method of predicting a *Saccharomyces cerevisiae* physiological function , comprising:

50

a) providing a data structure relating a plurality of *Saccharomyces cerevisiae* reactants to plurality of reactions, wherein each of said *Saccharomyces cerevisiae* reactions comprises a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product, wherein at least one of said *Saccharomyces cerevisiae* reactions is annotated to indicate an associated gene;

b) providing a constraint set for said plurality of *Saccharomyces cerevisiae* reactions;

c) providing an objective function, and

55

d) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said data structure, thereby predicated a *Saccharomyces cerevisiae* physiological function related to said gene.

15. The method of clause 14, wherein said plurality of *Saccharomyces cerevisiae* reactions comprises at least one reaction from a peripheral metabolic pathway.

60

16. The method of clause 14, wherein said peripheral metabolic pathway is selected from the group consisting of amino acid biosynthesis, amino acid degradation, purine biosynthesis, pyrimide biosynthesis, lipid biosynthesis, fatty acid metabolism, cofactor biosynthesis, cell wall metabolism and transport processes.

65

17. The method of clause 14, wherein said *Saccharomyces cerevisiae* physiological function is selected from the group consisting of growth, energy production, redox equivalent production, biomass production, production of biomass precursors, production of a protein, production of an amino acid, production of a purine, production of a pyrimidine, production of a lipid, production of a fatty acid, production of a cofactor, production of a cell wall component, transport of a metabolite, consumption of a carbon, nitrogen, sulphur, phosphate, hydrogen or oxygen.

18. The method of clause 14, wherein said *Saccharomyces cerevisiae* physiological function is selected from the group consisting of glycolysis, the TCA cycle, pentose phosphate pathway, respiration, biosynthesis of an amino acid, degradation of an amino acid, biosynthesis of a purine, biosynthesis of a pyrimidine, biosynthesis of a lipid, metabolism of a fatty acid, biosynthesis of a cofactor, metabolism of a cell wall component, transport of a metabolite and metabolism of a carbon source, nitrogen source, oxygen source, phosphate source, hydrogen source or sulphur source.
19. The method of clause 14, wherein said data structure comprises a set of linear algebraic equations.
20. The method of clause 14, wherein said data structure comprises a matrix.
21. The method of clause 14, wherein said flux distribution is determined by linear programming.
22. The method of clause 14, further comprising:
- 15 e) providing a modified data structure, wherein said modified data structure comprises at least one added reaction, compared to the data structure of part (a), and
 - f) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said modified data structure, thereby predicting a *Saccharomyces cerevisiae* physiological function.
23. The method of clause 22, further comprising identifying at least one participant in said at least one added reaction.
24. The method of clause 23, wherein said identifying at least one participant comprises associating a *Saccharomyces cerevisiae* protein with said at least one reaction.
25. The method of clause 24, further comprising identifying at least one gene that encodes said protein.
26. The method of clause 23, further comprising identifying at least one compound that alters the activity or amount of said at least one participant, thereby identifying a candidate drug or agent that alters a *Saccharomyces cerevisiae* physiological function.
27. The method of clause 14, further comprising:
- 30 e) providing a modified data structure, wherein said modified data structure lacks at least one reaction compounded to the data structure of part (a), and
 - f) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said modified data structure, thereby predicting a *Saccharomyces cerevisiae* physiological function.
28. The method of clause 27, further comprising identifying at least one participant in said at least one reaction.
29. The method of clause 28, wherein said identifying at least one participant comprises associating a *Saccharomyces cerevisiae* protein with said at least one reaction.
30. The method of clause 29, further comprising identifying at least one gene that encodes said protein that performs said at least one reaction.
31. The method of clause 28, further comprising identifying at least one compound that alters the activity or amount of said at least one participant, thereby identifying a candidate drug or agent that alters a *Saccharomyces cerevisiae* physiological function.
32. The method of clause 14, further comprising:
- 40 e) providing a modified constraint set, wherein said modified constraint set comprises a changed constraint for at least one reaction compared to the constraint for said at least one reaction in the data structure of part (a), and
 - f) determining at least one flux distribution that minimises or maximises said objective function when said modified constraint set is applied to said data structure, thereby predicated a *Saccharomyces cerevisiae* phys-

iological function.

33. The method of clause 32, further comprising identifying at least one participant in said at least one reaction.
- 5 34. The method of clause 33, wherein said identifying at least one participant comprises associating a *Saccharomyces cerevisiae* protein with said at least one reaction.
- 10 35. The method of clause 34, further comprising identifying at least one gene that encodes said protein.
- 15 36. The method of clause 33, further comprising identifying at least one compound that alters the activity or amount of said at least one participant, thereby identifying a candidate drug or agent that alters a *Saccharomyces cerevisiae* physiological function.
37. The method of clause 14, further comprising providing a gene database relating one or more reactions in said data structure with one or more genes or proteins in *Saccharomyces cerevisiae*.
- 20 38. A method of predicting *Saccharomyces cerevisiae* growth, comprising;
- a) providing a data structure relating a plurality of *Saccharomyces cerevisiae* reactants to plurality of *Saccharomyces cerevisiae* reactions, wherein each of said *Saccharomyces cerevisiae* reactions comprises a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product;
 - b) providing a constraint set for said plurality of *Saccharomyces cerevisiae* reactions;
 - c) providing an objective function, and
 - 25 d) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said data structure, thereby predicting *Saccharomyces cerevisiae* growth.
39. A method for making a data structure relating a plurality of *Saccharomyces cerevisiae* reactants to plurality of *Saccharomyces cerevisiae* reactions in a computer readable medium or media, comprising:
- 30 a) identifying a plurality of *Saccharomyces cerevisiae* reactions and a plurality of *Saccharomyces cerevisiae* reactants that are substrates and products of said *Saccharomyces cerevisiae* reactions;
- 35 b) relating said plurality of *Saccharomyces cerevisiae* reactants to said plurality of *Saccharomyces cerevisiae* reactions in a data structure, wherein each of the said *Saccharomyces cerevisiae* reactions comprises a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product;
- c) determining a constraint set for said plurality of *Saccharomyces cerevisiae* reactions;
- d) providing a objective function;
- e) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said data structure, and
- 40 f) if said least one flux distribution is not predictive of a *Saccharomyces cerevisiae* physiological function, then adding a reaction to or deleting a reaction from said data structure and repeating step (e), if said at least one flux distribution is predictive of a *Saccharomyces cerevisiae* physiological function, then storing said data structure in a computer readable medium or media.
- 45 40. The method of clause 39, wherein a reaction in said data structure is identified from an annotated genome.
41. The method of clause 40, further comprising storing said reaction that is identified from an annotated genome in a gene database.
- 50 42. The method of clause 39, further comprising annotation a reaction in said data structure.
43. The method of clause 42, wherein said annotation is selected from the group consisting of assignment of a gene, assignment of a protein, assignment of a subsystem, assignment of a confidence rating, reference to a genome annotation information and reference to a publication.
- 55 44. The method of clause 39, wherein step (b) further comprises identifying an unbalanced reaction in said data structure and adding a reaction to said data structure, thereby changing said unbalanced reaction to a balanced reaction.

45. The method of clause 39, wherein said adding a reaction comprises adding a reaction selected from the group consisting of an intra-system reaction, an exchange reaction, a reaction from a peripheral metabolic pathway, reaction from a central metabolic pathway, a gene associated reaction and a non-gene associated reaction.

5 46. The method of clause 45, wherein said peripheral metabolic pathway is selected from the group consisting of amino acid biosynthesis, amino acid degradation, purine biosynthesis, pyrimidine biosynthesis, lipid biosynthesis fatty acid metabolism, cofactor biosynthesis, cell wall metabolism and transport processes.

10 47. The method of clause 39, wherein said *Saccharomyces cerevisiae* physiological function is selected from the group consisting of growth, energy production, redox equivalent production, biomass production, production of biomass precursors, production of a protein, production of an amino acid, production of a purine, production of a pyrimidine, production of a lipid, production of a fatty acid, production of a cofactor, production of a cell wall component, transport of a metabolite, development, intercellular signalling, consumption of carbon, nitrogen, sulphur, phosphate, hydrogen or oxygen.

15 48. The method of clause 39, wherein said *Saccharomyces cerevisiae* physiological function is selected from the group consisting of degradation of an amino acid, degradation of a purine, degradation of a pyrimidine, degradation of a lipid, degradation of a fatty acid, degradation of a cofactor and degradation of a cell wall component.

20 49. The method of clause 39, wherein said data structure comprises a set of linear algebraic equations.

50. The method of clause 39, wherein said data structure comprises a matrix

51. The method of clause 39, wherein said flux distribution is determined by linear programming.

25 52. A data structure relating a plurality of *Saccharomyces cerevisiae* reactants, to a plurality of *Saccharomyces cerevisiae* reactions wherein said data structure is produced by a process comprising:

- 30 a) identifying a plurality of *Saccharomyces cerevisiae* reactions and a plurality of *Saccharomyces cerevisiae* reactants that are substrates and products of said *Saccharomyces cerevisiae* reactions;
- 35 b) relating said plurality of *Saccharomyces cerevisiae* reactants to said plurality of *Saccharomyces cerevisiae* reactions in a data structure, wherein each of said *Saccharomyces cerevisiae* reactions comprises a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product;
- 35 c) determining a constraint set for said plurality of *Saccharomyces cerevisiae* reactions;
- 35 d) providing an objective function;
- 35 e) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said data structure, and
- 40 f) if said at least one flux distribution is not predictive of *Saccharomyces cerevisiae* physiology, then adding a reaction to or deleting a reaction from said data structure and repeating step (e), if said at least one flux distribution is predictive of *Saccharomyces cerevisiae* physiology, then storing said data structure in a computer readable medium or media.

45 Claims

1. A method for predicting a yeast physiological function, comprising:

- 50 (a) providing a data structure relating a plurality of yeast reactants to a plurality of reactions, wherein each of said yeast reactions comprises a reactant identified as a substrate of the reaction, a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product, wherein at least one of said yeast reactions is annotated to indicate an associated gene;
- 55 (b) providing a constraint set for said plurality of yeast reactions;
- 55 (c) providing an objective function, and
- 55 (d) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said data structure, thereby predicting a yeast physiological function related to said gene.

2. The method of claim 1, wherein said plurality of yeast reactions comprises at least one reaction from a peripheral metabolic pathway, optionally wherein said peripheral metabolic pathway is selected from the group consisting of amino acid biosynthesis, amino acid degradation, purine biosynthesis, pyrimidine biosynthesis, lipid biosynthesis, fatty acid metabolism, cofactor biosynthesis, cell wall metabolism and transport processes.
- 5
3. The method of claim 2, wherein said yeast physiological function is selected from the group consisting of growth, energy production, redox equivalent production, biomass production, production of biomass precursors, production of a protein, production of an amino acid, production of a purine, production of a pyrimidine, production of a lipid, production of a fatty acid, production of a cofactor, production of a cell wall component, transport of a metabolite, consumption of a carbon, nitrogen, sulfur, phosphate, hydrogen or oxygen, glycolysis, the TCA cycle, pentose phosphate pathway, respiration, biosynthesis of an amino acid, degradation of an amino acid, biosynthesis of a purine, biosynthesis of a pyrimidine, biosynthesis of a lipid, metabolism of a fatty acid, biosynthesis of a cofactor, metabolism of a cell wall component, transport of a metabolite and metabolism of a carbon source, nitrogen source, oxygen source, phosphate source, hydrogen source or sulfur source.
- 10
4. The method of claim 1, wherein said data structure comprises a set of linear algebraic equations, or a matrix, and/or wherein said flux distribution is determined by linear programming.
- 15
5. The method of claim 1, further comprising:
- 20
- (e) providing a modified data structure, wherein said modified data structure comprises at least one added reaction, or lacks at least one reaction compared to the data structure of part (a), and
(f) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said modified data structure, thereby predicting a yeast physiological function.
- 25
6. The method of claim 5, further comprising identifying at least one participant in said at least one added reaction or at least one reaction, optionally wherein said identifying at least one participant comprises associating a yeast protein with said at least one reaction, and optionally further comprising identifying at least one gene that encodes said protein that performs said at least one reaction.
- 30
7. The method of claim 1, further comprising:
- 35
- (e) providing a modified constraint set, wherein said modified constraint set comprises a changed constraint for at least one reaction compared to the constraint for said at least one reaction in the data structure of part (a), and
(f) determining at least one flux distribution that minimises or maximises said objective function when said modified constraint set is applied to said data structure, thereby predicting a yeast physiological function.
- 40
8. The method of claim 7, further comprising identifying at least one participant in said at least one reaction.
9. The method of claim 6 or 8, wherein said identifying at least one participant comprises associating a yeast protein with said at least one reaction, and optionally further comprising identifying at least one gene that encodes said protein.
- 45
10. The method of claim 6 or 8, further comprising identifying at least one compound that alters the activity or amount of said at least one participant, thereby identifying a candidate drug or agent that alters a yeast physiological function.
11. The method of claim 1, further comprising providing a gene database relating one or more reactions in said data structure with one or more genes or proteins in yeast.
- 50
12. A method for making a data structure relating a plurality of yeast reactants to plurality of yeast reactions in a computer readable medium or media, comprising:
- 55
- (a) identifying a plurality of yeast reactions and a plurality of yeast reactants that are substrates and products of said yeast reactions;
(b) relating said plurality of yeast reactants to said plurality of yeast reactions in a data structure, wherein each of the said yeast reactions comprises a reactant identified as a product of the reaction and a stoichiometric coefficient relating said substrate and said product;
(c) determining a constraint set for said plurality of yeast reactions;
(d) providing a objective function;

(e) determining at least one flux distribution that minimises or maximises said objective function when said constraint set is applied to said data structure, and

5 (f) if said at least one flux distribution is not predictive of a yeast physiological function, then adding a reaction to or deleting a reaction from said data structure and repeating step (e), if said at least one flux distribution is predictive of a yeast physiological function, then storing said data structure in a computer readable medium or media.

13. The method of claim 12, wherein

10 i) a reaction in said data structure is identified from an annotated genome, and optionally further comprising storing said reaction that is identified from an annotated genome in a gene database,

15 ii) step (b) further comprises identifying an unbalanced reaction in said data structure and adding a reaction to said data structure, thereby changing said unbalanced reaction to a balanced reaction and/or wherein said data structure comprises a set of linear algebraic equations, or a matrix and/or wherein said flux distribution is determined by linear programming.

14. The method of claim 12, further comprising annotation of a reaction in said data structure, optionally wherein said annotation is selected from the group consisting of assignment of a gene, assignment of a protein, assignment of a subsystem, assignment of a confidence rating, reference to a genome annotation information and reference to a publication.

15. The method of claim 14, wherein said adding a reaction comprises adding a reaction selected from the group consisting of an intra-system reaction, an exchange reaction, a reaction from a peripheral metabolic pathway, reaction from a central metabolic pathway, a gene associated reaction and a non-gene associated reaction, optionally wherein said peripheral metabolic pathway is selected from the group consisting of amino acid biosynthesis, amino acid degradation, purine biosynthesis, pyrimidine biosynthesis, lipid biosynthesis fatty acid metabolism, cofactor biosynthesis, cell wall metabolism and transport processes and/or said yeast physiological function is selected from the group consisting of growth, energy production, redox equivalent production, biomass production, production of biomass precursors, production of a protein, production of an amino acid, production of a purine, production of a pyrimidine, production of a lipid, production of a fatty acid, production of a cofactor, production of a cell wall component, transport of a metabolite, development, intercellular signaling, consumption of carbon, nitrogen, sulfur, phosphate, hydrogen or oxygen, degradation of an amino acid, degradation of a purine, degradation of a pyrimidine, degradation of a lipid, degradation of a fatty acid, degradation of a cofactor and degradation of a cell wall component.

35

40

45

50

55

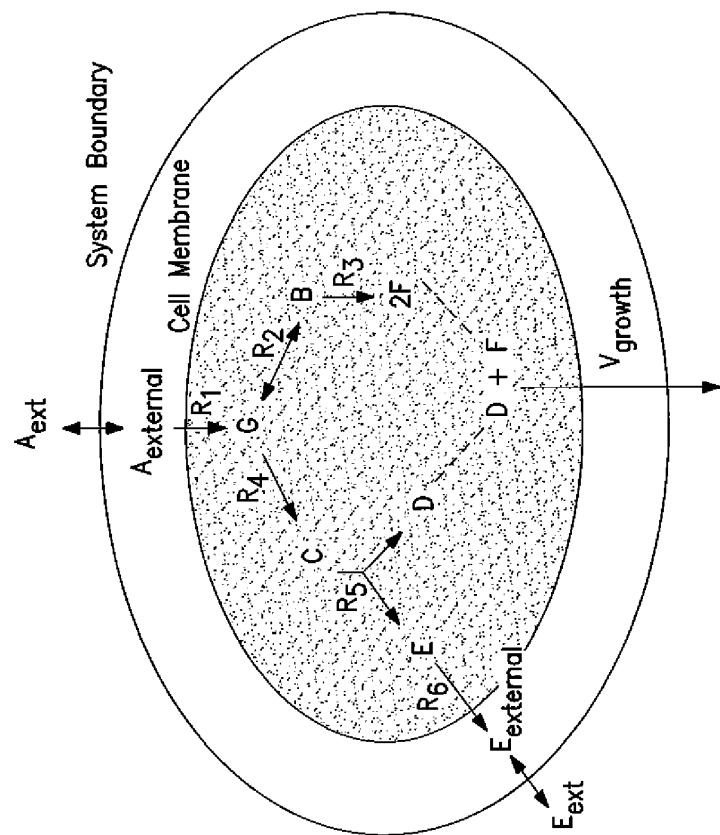


FIG. 1

$$\begin{array}{c}
 \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right] \\
 = \\
 \left[\begin{array}{c} R_1 \\ R_2 \\ R_3 \\ R_4 \\ R_5 \\ R_6 \end{array} \right] \left[\begin{array}{c} V_{\text{growth}} \\ A_{\text{xt}} \\ E_{\text{xt}} \\ 0 \\ 0 \\ 0 \end{array} \right]
 \end{array}$$

$$\begin{array}{cccccc}
 R_1 & R_2 & R_3 & R_4 & R_5 & R_6 \\
 0 & 1 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & -1 \\
 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0
 \end{array}$$

$$\begin{array}{c}
 V_{\text{growth}} \\
 A_{\text{xt}} \\
 E_{\text{xt}}
 \end{array}$$

FIG. 2

Mass Balance	Flux Constraints
B: $R_2 - R_3 = 0$	$0 \leq R_1 \leq \infty$
C: $R_4 - R_5 = 0$	$-\infty \leq R_2 \leq \infty$
D: $R_5 - V_{\text{growth}} = 0$	$0 \leq R_3 \leq \infty$
E: $R_5 - R_6 = 0$	$0 \leq R_4 \leq \infty$
F: $2 R_3 - V_{\text{growth}} = 0$	$0 \leq R_5 \leq \infty$
G: $R_1 - R_2 - R_4 = 0$	$0 \leq R_6 \leq \infty$
A external : $-A_{xt} - R_1 = 0$	$0 \leq V_{\text{growth}} \leq \infty$
E external : $R_6 - E_{xt} = 0$	$Y_1 \leq A_{xt} \leq Y_1$ $-\infty \leq E_{xt} \leq \infty$ $Y_1 = \text{const.}$
Objective Function $Z = V_{\text{growth}}$	

FIG. 3

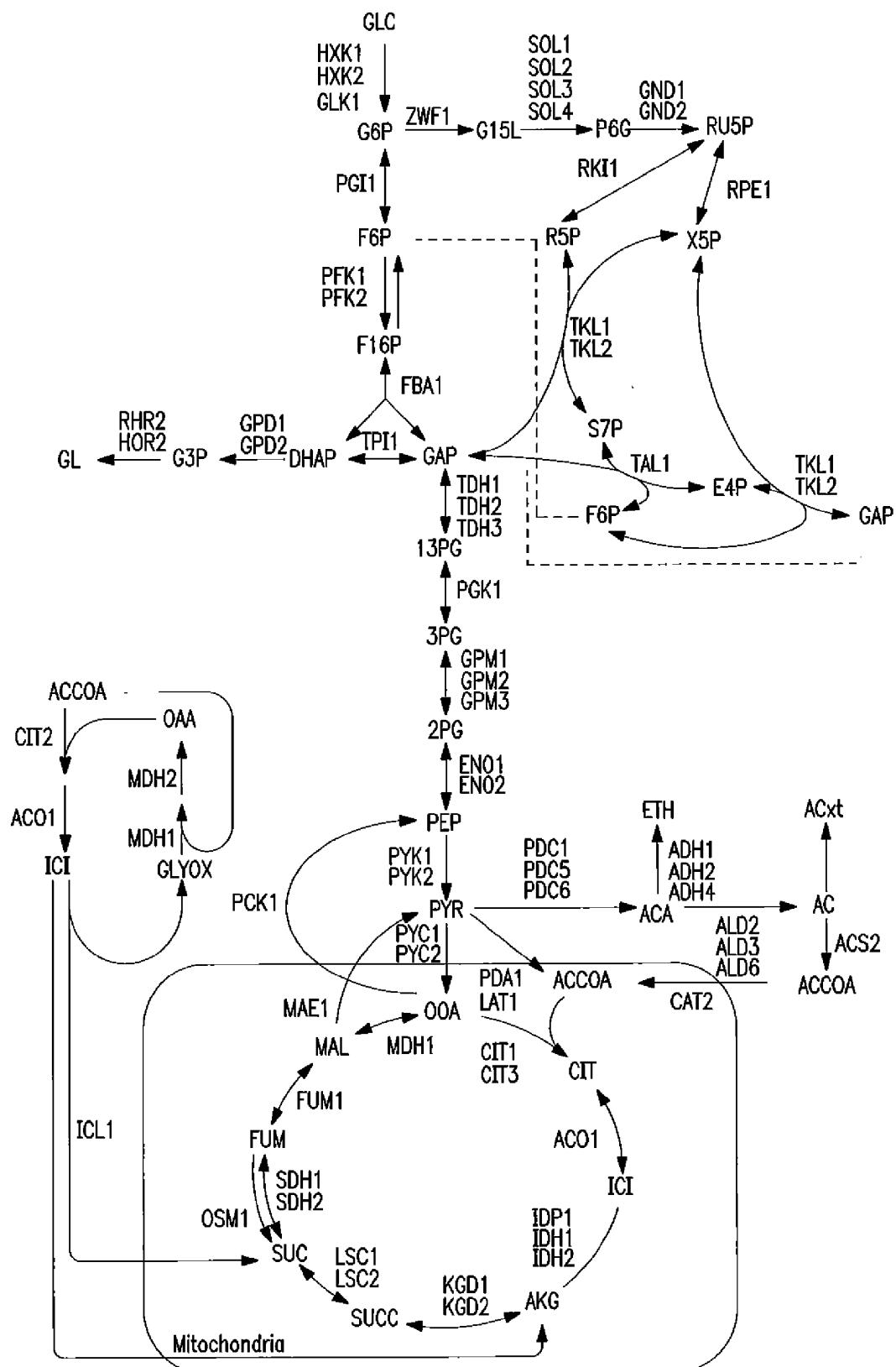
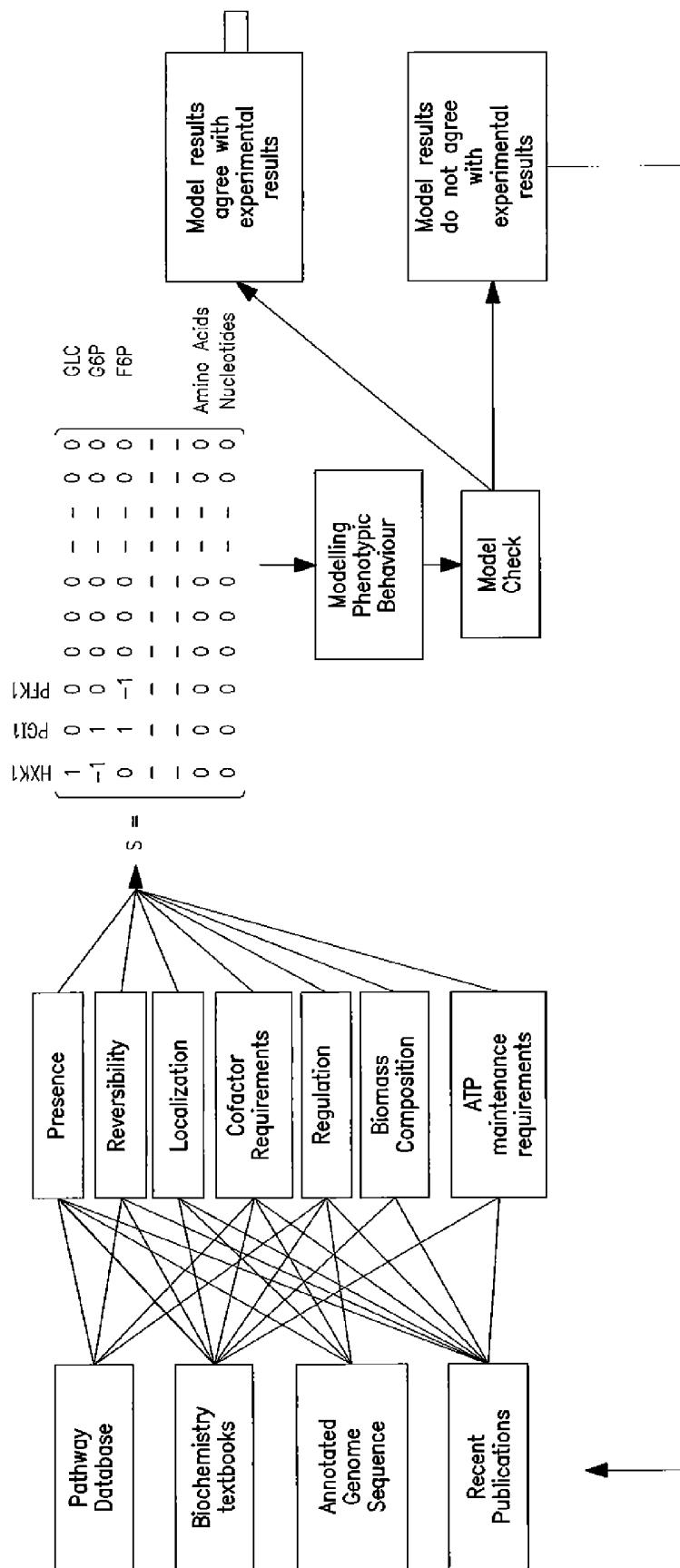
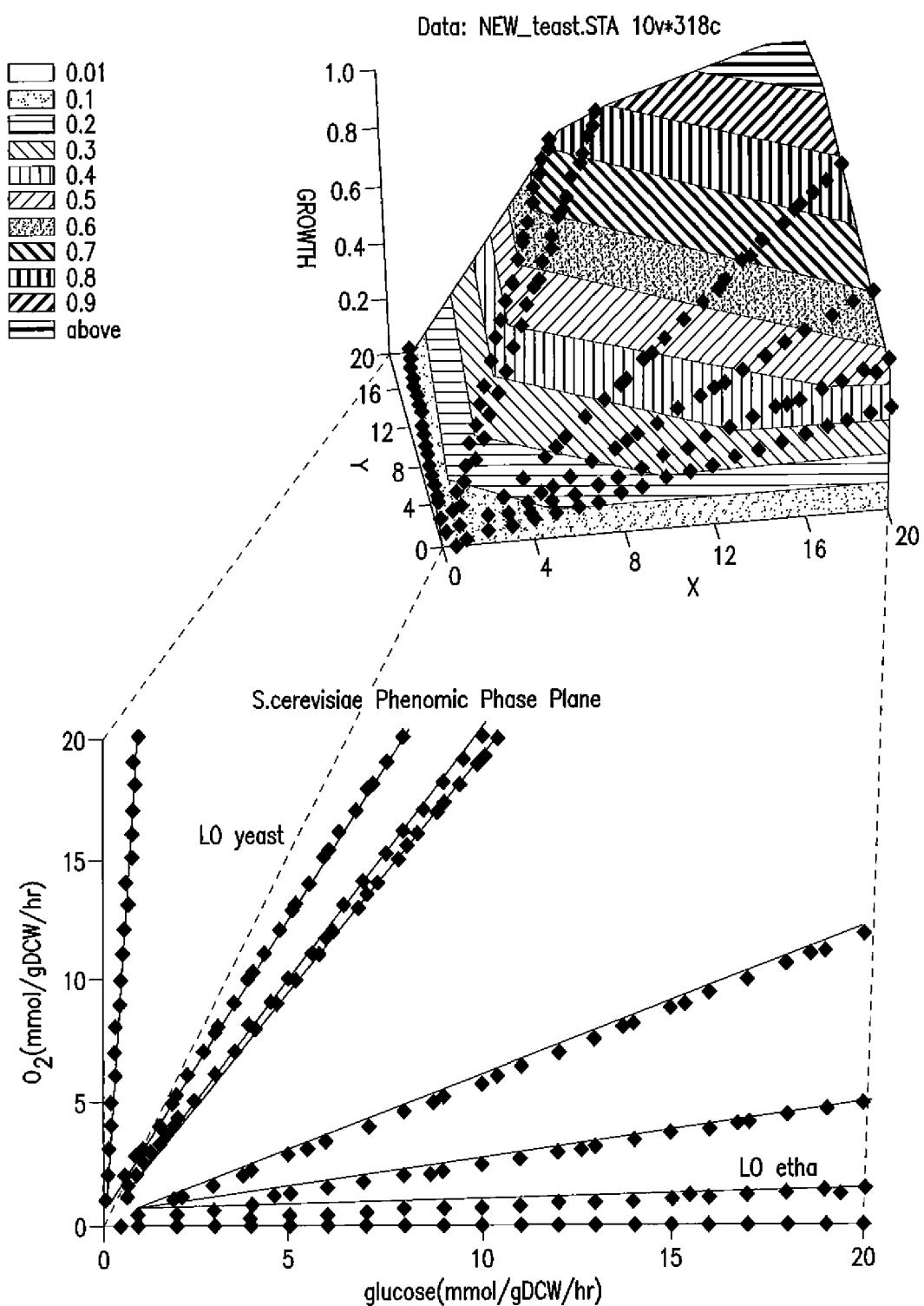
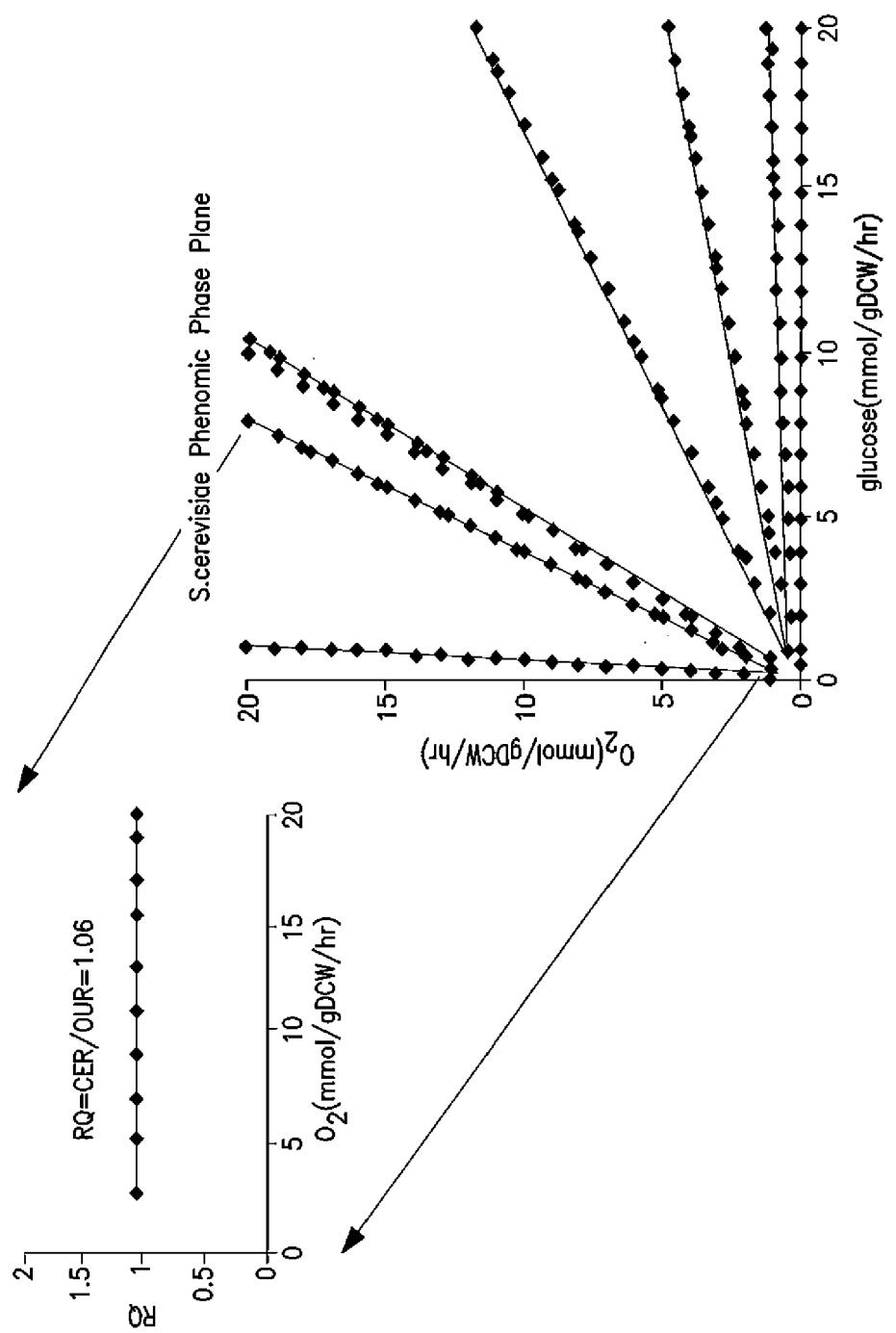
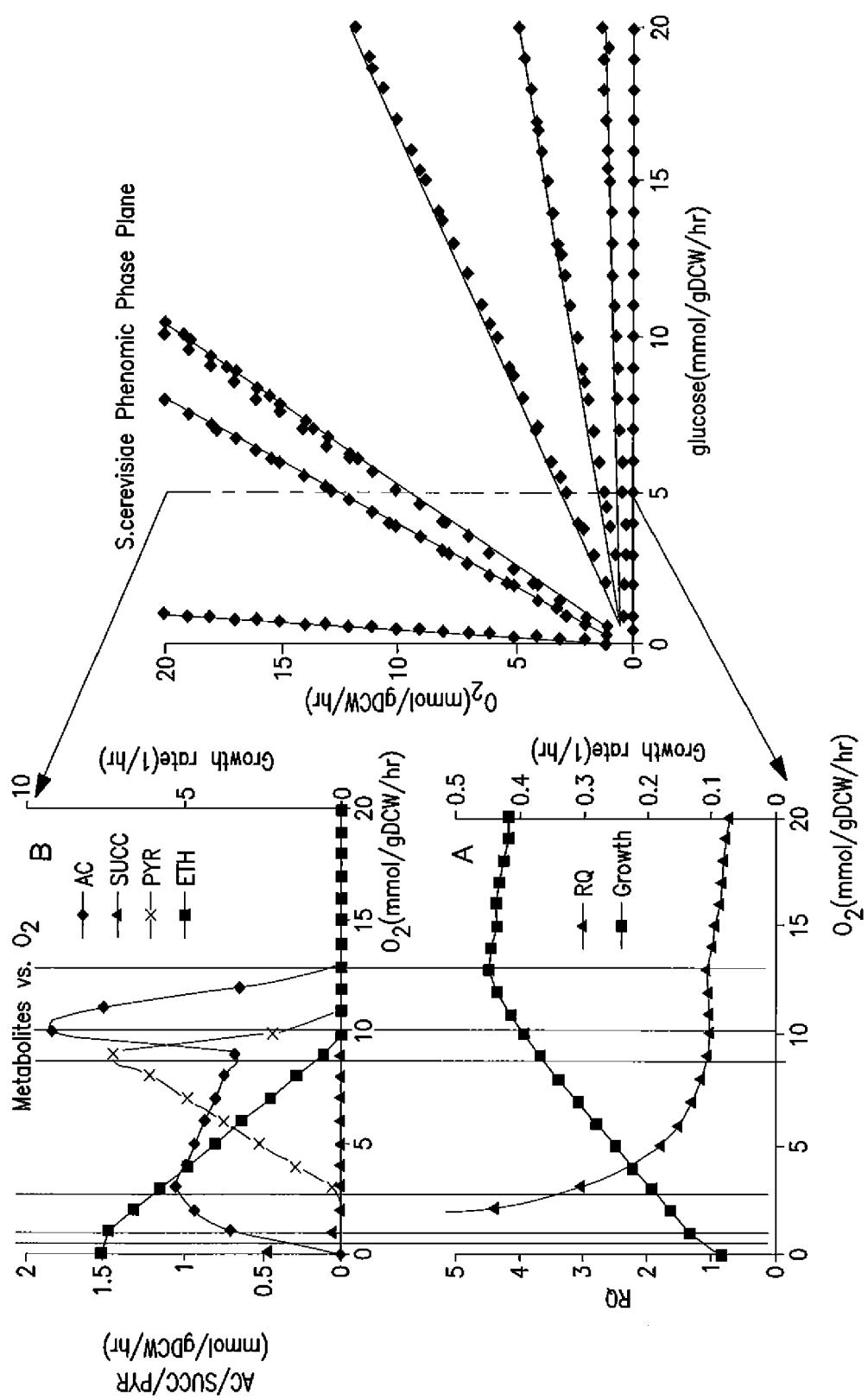


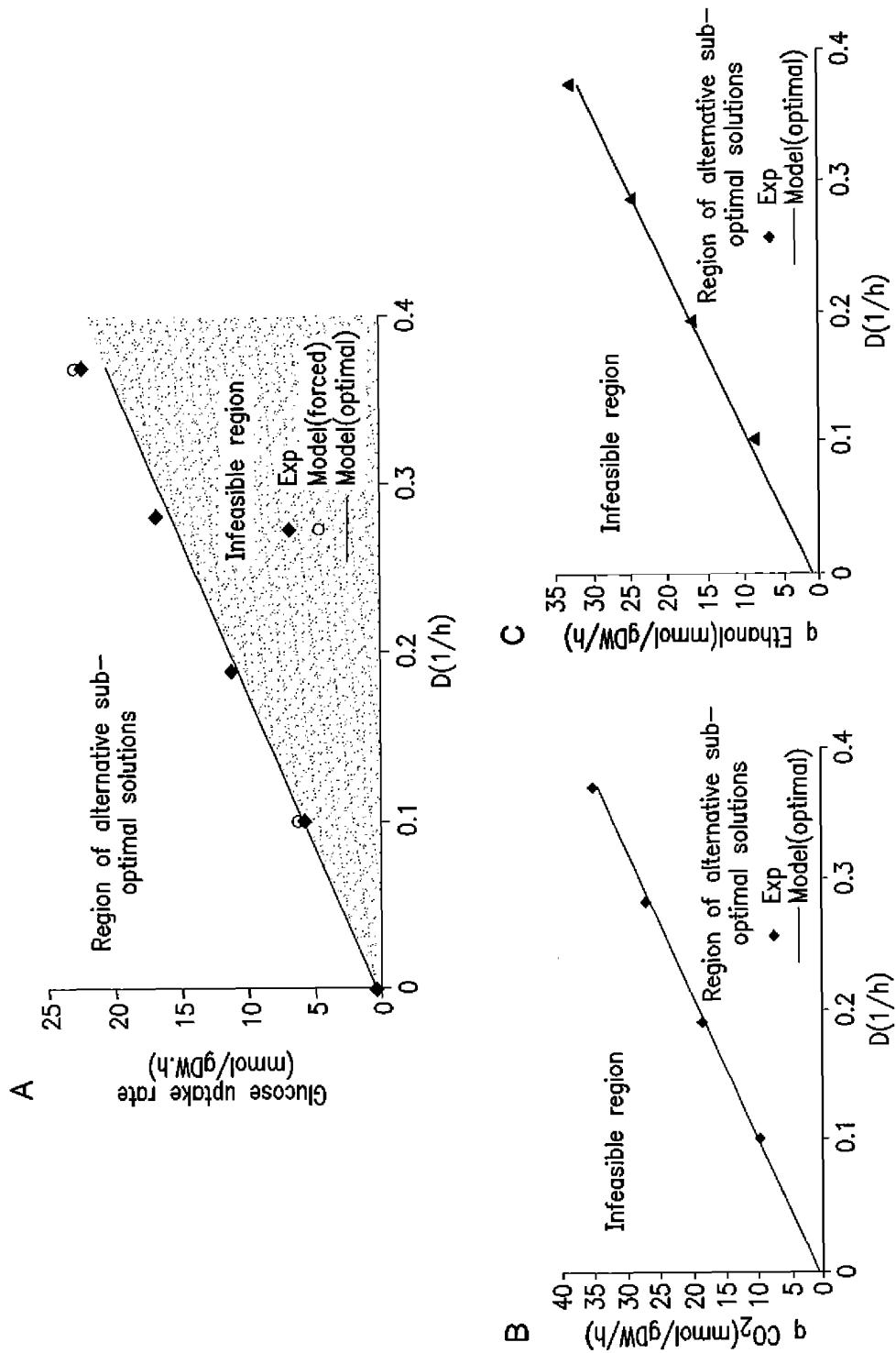
FIG. 4

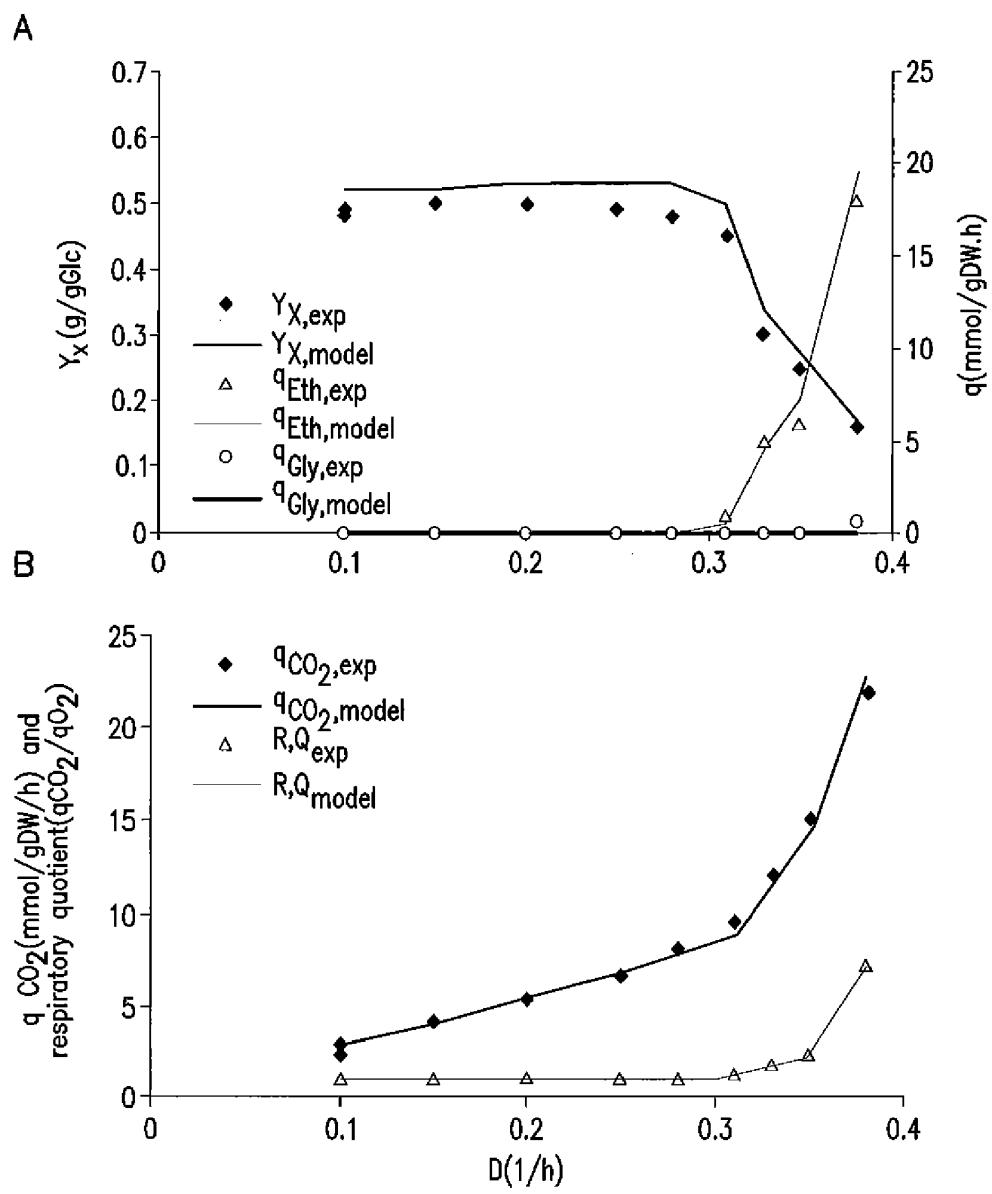
**FIG. 5**

**FIG. 6**

**FIG. 7**

**FIG. 8**

**FIG. 9**

**FIG. 10**



EUROPEAN SEARCH REPORT

Application Number

EP 11 18 2034

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (IPC)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
X	WO 00/46405 A2 (PALSSON BERNHARD [US]) 10 August 2000 (2000-08-10) * page 3, line 24 - line 26 * * page 5, line 34 - page 6, line 2 * * page 7, line 5 - line 8 * * page 7, line 19 - line 20 * * page 7, line 29 - page 8, line 5 * * page 8, line 34 - page 9, line 1 * * page 9, line 13 - line 17 * * figure 2; table 1 * -----	1-15	INV. G01N33/48
X	PRAMANIK J ET AL: "Stoichiometric model of Escherichia coli metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements", BIOTECHNOLOGY AND BIOENGINEERING, WILEY & SONS, HOBOKEN, NJ, US, vol. 56, no. 4, 1997, pages 398-421, XP002153946, ISSN: 0006-3592 * the whole document *	1-15	TECHNICAL FIELDS SEARCHED (IPC)
X	EDWARDS J S ET AL: "The escherichia coli Mg1655 in silico metabolic genotype: its definition, characteristics and capabilities", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC, US, vol. 97, no. 10, 9 May 2000 (2000-05-09), pages 5528-5533, XP002961698, ISSN: 0027-8424 * the whole document *	1-15	G06F G01N
		-/-	
The present search report has been drawn up for all claims			
1	Place of search	Date of completion of the search	Examiner
	Munich	7 May 2012	Sisk, Aisling
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			



Europäisches
Patentamt
European
Patent Office
Office européen
des brevets

EUROPEAN SEARCH REPORT

Application Number

EP 11 18 2034

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (IPC)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
X	<p>SCHILLING C H ET AL: "TOWARD METABOLIC PHENOMICS: ANALYSIS OF GENOMIC DATA USING FLUX BALANCES", BIOTECHNOLOGY PROGRESS, vol. 15, no. 5, May 1999 (1999-05), pages 288-295, XP000961124, ISSN: 8756-7938 * the whole document *</p> <p>-----</p>	1-15	
			TECHNICAL FIELDS SEARCHED (IPC)
1	The present search report has been drawn up for all claims		
	Place of search	Date of completion of the search	Examiner
	Munich	7 May 2012	Sisk, Aisling
CATEGORY OF CITED DOCUMENTS			
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			
T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 11 18 2034

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

07-05-2012

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0046405 A2 10-08-2000	AU 3482200 A EP 1147229 A2 US 2002012939 A1 US 2008270096 A1 WO 0046405 A2	25-08-2000 24-10-2001 31-01-2002 30-10-2008 10-08-2000	

EPO FORM F0159

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 0046405 A, Palsson [0066]
- US 9108694 W, R. Houghten [0093]
- US 5556762 A [0093]

Non-patent literature cited in the description

- WINZELER et al. *Science*, 1999, vol. 285, 901-906 [0003] [0143] [0157] [0172]
- COVERT et al. *Trends in Biochemical Sciences*, 2001, vol. 26, 179-186 [0066]
- SCHILLING et al. *J. Theor. Biol.*, 2000, vol. 203, 249-283 [0069]
- RAY. Learning XML. O'Reilly and Associates, 2001 [0076]
- PALSSON. *Nat. Biotech*, 2000, vol. 18, 1147-1150 [0085]
- CHVATAL. Linear Programming. W.H. Freeman and Co, 1983 [0086]
- EDWARDS et al. *Biotech Bioeng.*, 2002, vol. 77, 27-36 [0087]
- DECAMP et al. Protein Engineering Principles and Practice. Wiley-Liss, 1996, 467-506 [0093]
- HOUGHTEN et al. *Nature*, 1991, vol. 354, 84-86 [0093]
- DOOLEY et al. *Science*, 1994, vol. 266, 2019-2022 [0093]
- ZUBAY. Biochemistry. Wm.C. Brown Publishers, 1998 [0100]
- STRYER. Biochemistry. W.H. Freeman, 1988 [0100]
- MEWES et al. *Nucleic Acids Research*, 2002, vol. 30 (1), 31-34 [0100] [0138] [0143]
- CHERRY et al. *Nucleic Acids Research*, 1998, vol. 26 (1), 73-9 [0100]
- COSTANZO et al. *Nucleic Acids Research*, 2001, vol. 29 (1), 75-9 [0100]
- KANEHISA et al. *Nucleic Acids Research*, 2000, vol. 28 (1), 27-30 [0101]
- APPEL et al. *Trends Biochem Sci.*, 1994, vol. 19 (6), 258-260 [0101]
- BAIROCH et al. *Nucleic Acids Research*, 2000, vol. 28 (1), 45-48 [0101]
- STRATHERN et al. *The Molecular biology of the yeast Saccharomyces : metabolism and gene expression*, 1982 [0102]
- DAUM et al. *Yeast*, 1998, vol. 14 (16), 1471-510 [0102]
- DICKINSON et al. The metabolism and molecular physiology of *Saccharomyces cerevisiae*. Taylor & Francis, London, 1999 [0102]
- DICKSON et al. *Methods Enzymol.*, 2000, vol. 311, 3-9 [0102]
- DICKSON. *Annu Rev Biochem*, 1998, vol. 67, 27-48 [0102]
- PARKS. *CRC Crit Rev Microbiol*, 1978, vol. 6 (4), 301-41 [0102]
- VERDUYN et al. *Antonie Van Leeuwenhoek*, 1991, vol. 59 (1), 49-63 [0102] [0109] [0132]
- OVERKAMP et al. *J. of Bacteriol*, 2000, vol. 182 (10), 2823-2830 [0102]
- ZIMMERMAN et al. Yeast sugar metabolism : biochemistry, genetics, biotechnology, and applications. Technomic Pub, 1997 [0102] [0174]
- PAULSEN et al. *FEBS Lett*, 1998, vol. 430 (1-2), 116-125 [0102]
- WIECZORKE et al. *FEBS Lett*, 1999, vol. 464 (3), 123-128 [0102]
- REGENBERG et al. *Curr Genet*, 1999, vol. 36 (6), 317-328 [0102]
- ANDRE. *Yeast*, 1995, vol. 11 (16), 1575-1611 [0102]
- PALMIERI et al. *J Bioenerg Biomembr*, 2000, vol. 32 (1), 67-77 [0102]
- PALMIERI et al. *Biochim Biophys Acta*, 2000, vol. 1459 (2-3), 363-369 [0102]
- PALMIERI et al. *J Biol Chem*, 1999, vol. 274 (32), 22184-22190 [0102]
- PALMIERI et al. *FEBS Lett*, 1997, vol. 417 (1), 114-118 [0102]
- PALLOTTA et al. *FEBS Lett*, 1998, vol. 428 (3), 245-249 [0102]
- TZAGOLOGG et al. Mitochondria. Plenum Press, 1982 [0102]
- ANDRE. *Yeast*, 1995, vol. 11 (16), 1575-611 [0102]
- REMIZE et al. *Appl Environ Microbiol*, 2000, vol. 66 (8), 3151-3159 [0104]
- NIESSEN et al. *Yeast*, 2001, vol. 18 (1), 19-32 [0105]
- WESTERHOFF ; VAN DAM. Thermodynamics and control of biological free-energy transduction. Elsevier, 1987 [0109]
- EDWARDS et al. *Biotechnol. Bioeng.*, 2002, vol. 77, 27-36 [0112]
- EDWARDS et al. *Nature Biotech.*, 2001, vol. 19, 125-130 [0112]

- **VANDIJKEN ; SCHEFFERS.** *Yeast*, 1986, vol. 2 (2), 123-7 [0123]
- **ZENG et al.** *Biotechnol. Bioeng.*, 1994, vol. 44, 1107-1114 [0127]
- **ZIGOVA.** *J Biotechnol.*, 2000, vol. 80, 55-62 [0127]
- **WANG et al.** *Biotechnol & Bioeng.*, 1977, vol. 19, 69-86 [0127]
- **WANG et al.** *Biotechnol. & Bioeng.*, 1979, vol. 21, 975-995 [0127]
- **DANTIGNY et al.** *Appl. Microbiol. Biotechnol.*, 1991, vol. 36, 352-357 [0127]
- **NISSEN et al.** *Microbiology*, 1997, vol. 143 (1), 203-18 [0133]
- **NISSEN et al.** *Microbiology*, 1997, vol. 143 (1), 203-218 [0134]
- **VAN DIJKEN et al.** *Antonie Van Leeuwenhoek*, 1993, vol. 63 (3-4), 343-52 [0135]
- **OVERKAMP et al.** *J. of Bacteriol.*, 2000, vol. 182 (10), 2823-30 [0135]
- **CHERRY et al.** *Nucleic Acids Research*, 1998, vol. 26 (1), 73-79 [0138] [0143]
- **EDWARDS ; PALSSON.** *Proceedings of the National Academy of Sciences*, 2000, vol. 97 (10), 5528-5533 [0143]
- **CIRIACY ; BREITENBACH.** *J Bacteriol.*, 1979, vol. 139 (1), 152-60 [0146]
- **WHITE et al.** *J Biol Chem.*, 2001, vol. 276 (14), 10794-10800 [0154]
- **SILVE et al.** *Mol Cell Biol*, 1996, vol. 16 (6), 2719-2727 [0155]
- **BOUROT ; KARST.** *Gene*, 1995, vol. 165 (1), 97-102 [0155]
- **LACROUTE.** *J Bacteriol.*, 1968, vol. 95 (3), 824-832 [0155]
- **GUETSOVA et al.** *Genetics*, 1997, vol. 147 (2), 383-397 [0157]
- **MITCHELL.** *Genetics*, 1985, vol. 111 (2), 243-58 [0158]
- **DICKINSON ; SCHWEIZER.** The metabolism and molecular physiology of *Saccharomyces cerevisiae*. Taylor & Francis, 1999 [0161]
- **SUMMERS et al.** *Genetics*, 1988, vol. 120 (4), 909-922 [0161]
- **DAUM et al.** *Yeast*, 1998, vol. 14 (16), 1471-1510 [0161]
- **DAUM et al.** *Yeast*, 1999, vol. 15 (7), 601-614 [0161]
- **BARD et al.** *Lipids*, 1977, vol. 12 (8), 645-654 [0162]
- **ZWEYICK et al.** *FEBS Lett.*, 2000, vol. 470 (1), 83-87 [0162]
- **HATA et al.** *J Biochem*, 1983, vol. 94 (2), 501-510 [0162]
- **PARKS et al.** *Crit Rev Biochem Mol Biol*, 1999, vol. 34 (6), 399-404 [0162]
- **MURRAY ; GREENBERG.** *Mol Microbiol*, 2000, vol. 36 (3), 651-661 [0165]
- **LOPEZ et al.** *Mol Microbiol*, 1999, vol. 31 (4), 1255-1264 [0165]
- **THOMAS ; SURDIN-KERJAN.** *Microbiol Mol Biol Rev*, 1997, vol. 61 (4), 503-532 [0166]
- **FUKUCHI et al.** *Genes*, 1993, vol. 129 (1), 141-146 [0167]
- **PERSSON et al.** *Biochim Biophys Acta*, 1999, vol. 1422 (3), 255-72 [0168]
- **TROTTER et al.** *J Biol Chem*, 1998, vol. 273 (21), 13189-13196 [0169]
- **BELL et al.** *J Biol Chem*, 1998, vol. 273 (50), 33311-33319 [0170]
- **BIRNER et al.** *Mol Biol Cell*, 2001, vol. 12 (4), 997-1007 [0172]
- **BOLES et al.** *J Bacteriol*, 1997, vol. 179 (9), 2987-2993 [0174]
- **VAN DEN BERG ; STEENSMA.** *Eur J Biochem*, 1995, vol. 231 (3), 704-713 [0174]
- **HEINISCH et al.** *Yeast*, 1998, vol. 14 (3), 203-13 [0174]
- **BOLES,E. ; LIEBETRAU,W. ; HOFMANN,M. ; ZIMMERMANN,F.K.** A family of hexosephosphate mutases in *Saccharomyces cerevisiae*. *Eur. J. Biochem.*, 1994, vol. 220, 83-96 [0176]
- **BOLES,E.** Yeast sugar metabolism. Technomic Publishing CO., INC, 1997, 81-96 [0176]
- **BOLES,E. ; JONG-GUBBELS,P. ; PRONK,J.T.** Identification and characterization of MAE1, the *Saccharomyces cerevisiae* structural gene encoding mitochondrial malic enzyme. *J. Bacteriol.*, 1998, vol. 180, 2875-2882 [0176]
- **CLIFTON,D. ; WEINSTOCK,S.B. ; FRAENKEL,D.G.** Glycolysis mutants in *Saccharomyces cerevisiae*. *Genetics*, 1978, vol. 88, 1-11 [0176]
- **CLIFTON,D. ; FRAENKEL,D.G.** Mutant studies of yeast phosphofructokinase. *Biochemistry*, 1982, vol. 21, 1935-1942 [0176]
- **CUPP,J.R. ; MCALISTER-HENN,L.** Cloning and Characterization of the gene encoding the IDH1 subunit of NAD(+) -dependent isocitrate dehydrogenase from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, 1992, vol. 267, 16417-16423 [0176]
- **FLIKWEERT,M.T. et al.** Pyruvate decarboxylase: an indispensable enzyme for growth of *Saccharomyces cerevisiae* on glucose. *Yeast*, 1996, vol. 12, 247-257 [0176]
- **GANCEDO,C. ; DELGADO,M.A.** Isolation and characterization of a mutant from *Saccharomyces cerevisiae* lacking fructose 1,6-bisphosphatase. *Eur. J. Biochem.*, 1984, vol. 139, 651-655 [0176]
- **GANGLOFF,S.P. ; MARGUET,D. ; LAUQUIN,G.J.** Molecular cloning of the yeast mitochondrial aconitase gene (ACO1) and evidence of a synergistic regulation of expression by glucose plus glutamate. *Mol Cell Biol*, 1990, vol. 10, 3551-3561 [0176]
- **HARTIG,A. et al.** Differentially regulated malate synthase genes participate in carbon and nitrogen metabolism of *S. cerevisiae*. *Nucleic Acids Res.*, 1992, vol. 20, 5677-5686 [0176]

- HEINISCH,J.J. ; MULLER,S. ; SCHLUTER,E. ; JACOBY,J. ; RODICIO,R. Investigation of two yeast genes encoding putative isoenzymes of phosphoglycerate mutase. *Yeast*, 1998, vol. 14, 203-213 [0176]
- KIM,K.S. ; ROSENKRANTZ,M.S. ; GUARENTE,L. *Saccharomyces cerevisiae* contains two functional citrate synthase genes. *Mol. Cell Biol.*, 1986, vol. 6, 1936-1942 [0176]
- LOFTUS,T.M. ; HALL,L.V. ; ANDERSON,S.L. ; MCALISTER-HENN,L. Isolation, characterization, and disruption of the yeast gene encoding cytosolic NADP-specific isocitrate dehydrogenase. *Biochemistry*, 1994, vol. 33, 9661-9667 [0176]
- MCALISTER-HENN,L. ; THOMPSON,L.M. Isolation and expression of the gene encoding yeast mitochondrial malate dehydrogenase. *J. Bacteriol.*, 1987, vol. 169, 5157-5166 [0176]
- MULLER,S. ; ENTIAN,K.-D. Yeast sugar metabolism. Technomic Publishing CO.,INC, 1997, 157-170 [0176]
- OZCAN,S. ; FREIDEL,K. ; LEUKER,A. ; CIRIACY,M. Glucose uptake and catabolite repression in dominant HTR1 mutants of *Saccharomyces cerevisiae*. *J. Bacteriol.*, 1993, vol. 175, 5520-5528 [0176]
- PRZYBYLA-ZAWISLAK,B. ; DENNIS,R.A. ; ZAKHARKIN,S.O. ; MCCAMMON,M.T. Genes of succinyl-CoA ligase from *Saccharomyces cerevisiae*. *Eur. J Biochem.*, 1998, vol. 258, 736-743 [0176]
- REPETTO,B. ; TZAGOLOFF,A. In vivo assembly of yeast mitochondrial alpha-ketoglutarate dehydrogenase complex. *Mol. Cell Biol.*, 1991, vol. 11, 3931-3939 [0176]
- SCHAAFF-GERSTENSCHLAGER,I. ; ZIMMERMANN,F.K. Pentose-phosphate pathway in *Saccharomyces cerevisiae*: analysis of deletion mutants for transketolase, transaldolase, and glucose 6-phosphate dehydrogenase. *Curr. Genet.*, 1993, vol. 24, 373-376 [0176]
- SCHAAFF-GERSTENSCHLAGER,I. ; MIOSGA,T. Yeast sugar metabolism. Technomic Publishing CO.,INC, 1997, 271-284 [0176]
- SEDIVY,J.M. ; FRAENKEL,D.G. Fructose bisphosphatase of *Saccharomyces cerevisiae*. Cloning, disruption and regulation of the FBP1 structural gene. *J. Mol. Biol.*, 1985, vol. 186, 307-319 [0176]
- SMITH,V. ; CHOU,K.N. ; LASHKARI,D. ; BOTSTEIN,D. ; BROWN,P.O. Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science*, 1996, vol. 274, 2069-2074 [0176]
- SWARTZ,J. A PURE approach to constructive biology. *Nat. Biotechnol.*, 2001, vol. 19, 732-733 [0176]
- WILLS,C. ; MELHAM,T. Pyruvate carboxylase deficiency in yeast: a mutant affecting the interaction between the glyoxylate and Krebs cycles. *Arch. Biochem. Biophys.*, 1985, vol. 236, 782-791 [0176]