Constrain-based analysis of gene deletion on the metabolic flux redistribution of *Saccharomyces Cerevisiae*

Zi-Xiang Xu & Xiao Sun

State Key Laboratory of Bioelectronics, Southeast University, Nanjing 210096, China. Correspondence should be addressed to Z. X. Xu.(xzx21c@163.com, xzxnpu@sina.com).

ABSTRACT

Based on the gene-protein-reaction (GPR) model of *S. cerevisiae*_iND750 and the method of constraint-based analysis, we first calculated the metabolic flux distribution of *S. cerevisiae*_iND750. Then we calculated the deletion impact of 438 calculable genes, one by one, on the metabolic flux redistribution of *S. cerevisiae*_iND750. Next we analyzed the correlation between *v* (describing deletion impact of one gene) and *d* (connection degree of one gene) and the correlation between *v* and V_{gene} (flux sum controlled by one gene), and found that both of them were not of linear relation. Furthermore, we sought out 38 important genes that most greatly affected the metabolic flux distribution, and determined their functional subsystems. We also found that many of these key genes were related to many but not several subsystems. Because the *in silico* model of *S. cerevisiae*_iND750 has been tested by many experiments, thus is credible, we can conclude that the result we obtained has biological significance.

Keywords: Metabonomics, Metabolic engineering, Metabolic networks, Gene deletion, Genome-scale simulation, Flux balance analysis, Gene-protein- reaction (GPR) model, Constraint-based analysis

1. INTRODUCTION

Since various 'omics' datasets are becoming available, biology has transited from a data-poor to a data-rich environment. This has underscored the need for systems analysis in biology and systems biology has become a rapidly growing field as well [1].

A change in mathematical modeling philosophy, *i.e.*, building and validating *in silico* models is also necessitated. Modern biological models need to meet new sets of criteria: organism-specific, data-driven, easily scalable, and so on. Many modeling approaches, such as kinetic,

stochastic and cybernetic approaches, are currently being used to model cellular processes. Owing to the computational complexity and the large number of parameters needed, it is currently difficult to use these methods to model genome-scale networks. To date, genome-scale models of metabolism have only been analyzed with the constraint-based modeling philosophy [2, 3]. Genome-scale network models of diverse cellular processes such as signal transduction, transcriptional regulation and metabolism have been generated. Gene-protein-reaction (GPR) associated models can keep track of associations between genes, proteins, and reactions [4], and there have been several genome-scale GPR models, such as *E. col i* [4,5] , *S. aureus* [6], *H. pylori* [7], *M. barkeri*[8], *S. cerevisiae* [9] and *B. subtilis* [10]. A reconstruction is herein defined as the list of biochemical reactions occurring in a particular cellular system and the associations between these reactions and relevant proteins, transcripts and genes [2]. A reconstruction can include the assumptions necessary for computational simulation, such as maximum reaction rates and nutrient uptake rates.

Computer simulations of complex biological systems began essentially as soon as the computational capability became available. As reconstructed networks have been made publicly available, researchers around the world have undertaken new computational studies using these networks. Many researches apply a core set of basic *in silico* methods and often also describe novel methods to investigate different models. An extensive set of methods for analyzing these genome-scale models have been developed and have been applied to study a growing number of biological problems. As we have mentioned above, genome-scale models of metabolism have only been analyzed with the constraint-based philosophy [2,3].

The *in silico* models can be applied to generate novel, testable and often quantitative predictions of cellular behavior. The impact of a gene deletion experiment on cellular behavior can be simulated in a manner similar to linear optimization of growth. The results can be used to guide the design of informative confirmation experiments and will be helpful to metabolic engineering. Some gene deletion studies on genome-scale *in silico* organism models have been made [4-10], most of them are from the standpoints of distinguishing lethal and

non-lethal genes or growth rate [4-10, 15-19]. The impact of gene deletion on flux redistribution, the characters and functions of key genes are still lacked research.

In this paper, there are four parts. In part two, as a base for later research, we firstly calculate flux distribution of *S. cerevisiae*_iND750. Then we will calculate the deletion impact of 438 calculable genes, one by one, on themetabolic flux redistribution of *S. cerevisiae_* iND750. Next we will analyze the correlation between *v* (describing deletion impact of one gene) and *d* (connection degree of one gene) and the correlation between *v* and v_{gene} (flux sum controlled by one gene). Furthermore, we will seek out those important genes that most greatly affected the metabolic flux distribution, determine their functional subsystems. Because the *in silico* model of *S. cerevisiae*_iND750 has been tested by many experiments [5] , we can conclude that the result we got has biological significances; In part three, we introduce the GPR model, some properties of the *in silico* model of *S. cerevisiae*_iND750 (SBML format) and the method of constraint-based analysis; Part four is a simple discussion.

2. RESULTS

2.1. Metabolic flux distribution and of *S. cerevisiae*_iND750

As a base for the later compared research, we here calculated the flux distribution of *S. cerevisiae*_iND750 [9]. What we use is *Sc*_iND750_GlcMM.xml, the SBML file that is presented with the reconstruction of *S. cerevisiae*_iND750 [11]. The computational method we used is flux balance analysis (FBA) [11], one of the most fundamental genome-scale phenotypic calculations, which can simulate cellular growth. FBA is based on linear optimization of an objective function, which typically is bio-mass formation. Given an uptake rate for key nutrients and the biomass composition of the cell (usually in mmol component gDW^{-1} and defined in the biomass objective function), the maximum possible growth rate of the cells can be predicted *in silico*. We use the COBRA toolbox [11] to carry out this computation of FBA. The flux distribution of *S. cerevisiae*_iND750 is illustrated in **Figure 1**.
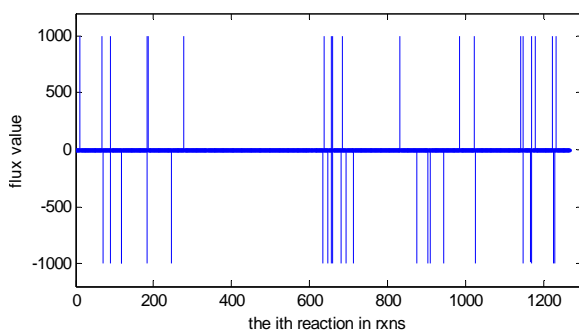

Figure 1. Flux distribution of *S. cerevisiae* _iND750. *X*-axis indicating every reaction in **rxns** (the order is as the same as in **rxns**, total 1266) and *y*-axis indicating the value of its corresponding flux (unit is mmol gDW^{-1}h^{-1}). **Rxns** is the reaction set in the model.

2.2 Impact of gene deletion on the metabolic flux redistribution and key genes

2.2.1 Impact of gene deletion on the metabolic flux redistribution

There are 750 genes in the model of *S. cerevisiae*_iND750, but we can not calculate the impact of every gene deletion. If a single gene is associated with multiple reactions, the deletion of that gene will result in the removal of all associated reactions. On the other hand, a reaction that can be catalyzed by multiple non-interacting gene products will not be removed in a single gene deletion. Among 750 genes of *S. cerevisiae_* iND750, there are 438 genes which have no "OR" relationship with other genes in every reaction of *S. cerevisiae*_iND750, and by the aid of the COBRA toolbox [11], we can calculate the impact of their deletion. We define the impact of one gene deletion on the whole metabolic flux redistribution as *v*

$$v = \sum_{i}^{R} (v'_i - v_i)^2 \qquad (1)$$

Where v_i and v'_i are respectively the flux value of *i*-th reaction of the model of *S. cerevisiae*_iND750 before and after a single gene deleting, and *R* is the whole reaction set.

Figure 2 shows the deletion impact of these 438 genes. From the figure, we can't see what are important genes which greatly affect, if deleted, the metabolic redistribution of *S. cerevisiae_* iND750. This remains as a problem, and we will settle it in section 4). In the following, we will analyze the relationship between the impact of every gene deletion (*v*) and the connection degree of every gene (*d*) and the connection degree of every gene (v_{gene}).

2.2.2. Correlation between *v* and *d* (connection degree of every gene)

We compute out the related reaction number *d* of every gene in those 438 genes of the *S. cerevisiae*_iND750 model, as illustrated in **Figure 3**. From the figure, we can find that some but not many genes have high *d* value, but we don't know whether they affect metabolic flux distribution greatly.
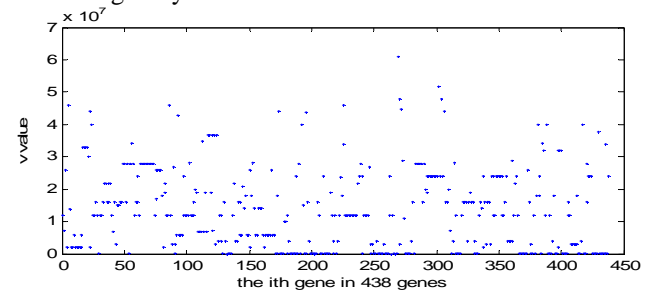

Figure 2. The deletion impact of calculable 438 genes of the *S. cerevisiae*_iND750 model. *X*-axis indicating every gene in 438 genes (the order is as the same as in **genes**, total 438) and *y*-axis indicating its impact *v*.
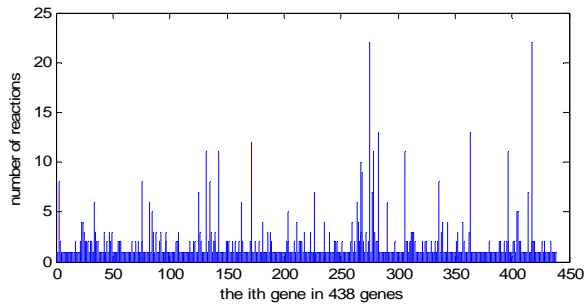

Figure 3. The related reaction number of every gene in 438 genes of the *S. cerevisiae*_iND750 model. *X*-axis indicating every gene in 438 genes (the order is as the same as in **genes**, total 438) and *y*-axis indicating the number of its related reactions.
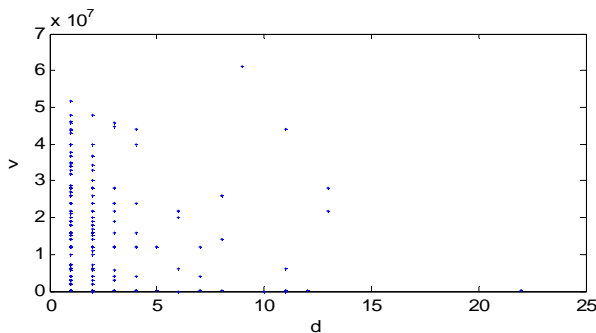

Figure 4. The scatters diagram (*d, v*). *X*-axis indicating *d* (connection degree of every gene) and *y*-axis indicating the corresponding gene impact *v*.

Figure 4 is the scatters diagram (*d, v*), total 438 data pairs. From the diagram, we can easily find that the relationship between *d* and *v* is not of linear relation. So high-*d* genes and low-*d* genes are equally important to the metabolism of *S. cerevisiae*_iND750.

2.2.3. Correlation between *v* and v_{gene} (flux sum controlled by every gene)

We define the flux sum controlled by every gene as

$$v_{gene} = \sum_{j}^{R_{gene}} |v_j| \qquad (2)$$

Where v_j is the flux value of *j*-th reaction of the model of *S. cerevisiae*_iND750 before a single gene deleting, and where R_{gene} is the reaction set controlled by the given gene. We can easily compute out the flux sum v_{gene} of every gene in those 438 genes of the *S. cerevisiae*_iND750 model, as illustrated in **Figure 5**. From the figure, we can find that some but not many genes have high v_{gene}, but will they affect metabolic flux distribution greatly.

Figure 6 is the scatters diagram (v_{gene}, *v*), total 438 data pairs. From the diagram, we can easily find that the relationship between v_{gene} and *v* is also not of linear relation. So high-v_{gene} genes and low-v_{gene} genes are equally important to the metabolism of *S. cerevisiae*_
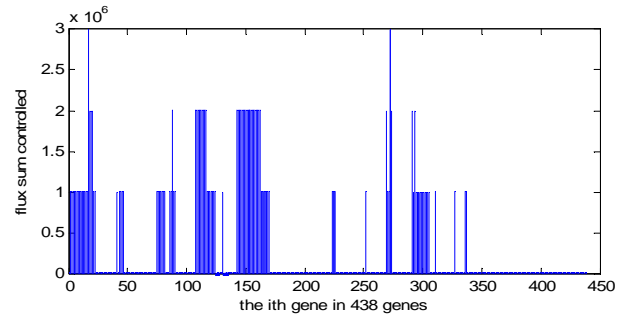

Figure 5. The controlled reaction number of every gene in 438 genes of the *S. cerevisiae*_iND750 model. *X*-axis indicating every gene in 438 genes (the order is as the same as in **genes**, total 438) and *y*-axis indicating the number of its controlled reactions.
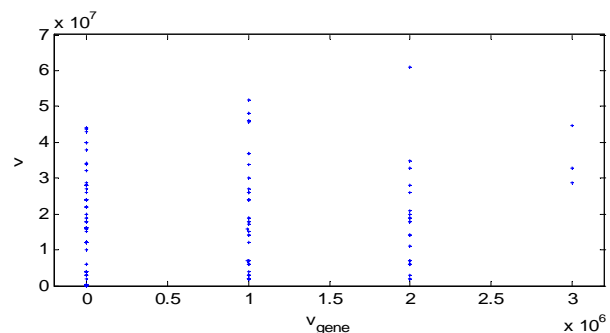

Figure 6. The scatters diagram (**v_{gene}**, *v*). *X*- axis indicating **v_{gene}** (the flux sum controlled by every gene) and *y*-axis indicating the impact *v*.

Table 1. Gene number (GN) and *v* scope

$v (\times 10^7)$	0	0-0.1	0.1-1.5	1.0-1.5	1.5-1.8
GN	4	100	80	71	46
$v (\times 10^7)$	1.8-2.0	2.0-2.5	2.5-3.0	>3.0	
GN	13	43	43	38	

iND750.

2.2.4. Key genes that affect metabolism most greatly

In this section, we seek out what are important or key genes which greatly affect the metabolic redistribution of *S. cerevisiae*_iND750, and furthermore in next section, we will give their belonged functional subsystems. **Table 1** provides the corresponding relationship between gene number (GN) and *v* scope, and as an example, GN=100 & *v*= (0-0.1)$\times 10^7$ means that there are 100 genes while the *v* scope which these genes control is (0~0.1)$\times 10^7$. We define those genes with *v*>3.0$\times 10^7$ as key genes, and there are 38 genes.

2.2.5. Functional subsystems to which these key genes belong

If a gene catalyze a reaction and while the reaction belong to a certain subsystem, we will say that the gene belong to the subsystem. Functional subsystems about important genes in the metabolic system of micro organism are seldom reported. We list the functional subsystems to which every key gene belong, and several genes

Table 2. The functional subsystems and their related genes of *S. cerevisiae*_iND750

Sub system	Purine and Pyrimidine Biosynthesis	Galactose metabolism	Transport Extracellular	Pyridoxine Metabolism	Transport Peroxisomal	Fatty Acid Bio-synthesis	Pentose Phosphate Pathway
genes	'YAR015W' 'YLR209C' 'YOR128C'	'YBR018C' 'YBR019C' 'YBR020W'	'YBR021W' 'YHL016C'	'YBR035C'	'YBR041W'	'YBR041W'	'YCR036W'
Sub system	Citric Acid Cycle	Glycolysis-Gluconeogene-sis	Tyrosine Tryptophan and Phenyla-lanine Me-tabolism	Pantothenate and CoA Bio-synthesis	Quinone Bio-synthesis	Oxidative Phos-phory-Lation	Nucleotide Salvage Pathway
genes	'YDR148C' 'YLR174W' 'YOR136W'	'YDR148C'	'YDR256C' 'YGR088W' 'YKL106W' 'YPR060C'	'YDR531W' 'YIL145C'	'YDR538W'	'YEL027W' 'YEL051W' 'YKL080W' 'YOR270C'	'YEL042W' 'YKL067W' 'YLR209C'
Sub system	Glutamate metabolism	Porphyrin and Chlorophyll Metabolism	Folate Me-tabolism	Glycerolipid Metabolism	Alanine and Aspartate Metabolism	Arginine and Proline Metabo-lism	Pyruvate Metabo-lism
genes	'YEL058W'	'YER014W' 'YOR278W'	'YGR267C'	'YIL155C'	'YKL106W'	'YKL106W' 'YOR303W'	'YLR153C'
Sub system	NAD Biosyn-thesis	Alternate Carbon Me-tabolism	Transport Mitochondrial	Thiamine Metabolism	Cysteine Me-tabolism		
genes	'YLR209C'	'YOR040W'	'YOR100C'	'YPL214C'	'YPR167C'		

appear in more than one subsystem, shown in **Table 2**. We will find that many of 38 key genes are related to 26 but not several subsystems.

3. MATERIALS AND METHODS

3.1. Gene-protein-reaction (GPR) associated model

The association between genes and reactions is not a one-to-one relationship. Many genes may encode suunits genes that encode so-called promiscuous enzymes that can catalyze several different reactions. So it is necessary to keep track of associations between genes, proteins, and reactions and to distinguish "&" and "OR" associations in GPR models. Examples of different types of GPR associations are illustrated in the figures of Ref. [4, 14].

3.2. GPR model structure of *S. cerevisiae*_iND750

The *in silico* model that we used is *S. cerevisiae*_iND750 [9], a metabolic reconstruction consisting of the chemical reactions that transport and interconvert metabolites within yeast. This network reconstruction was based on a previous reconstruction, termed *S. cerevisiae*_iFF708 [15]. The general features of *S. cerevisiae*_iND750 are shown in the Table 1 of Ref. [9].

A SBML format file to the model *S. cere-visiae*_iND750 can be downloaded from the supplementary information of Ref. [11]. SBML file properties are given in the supplementary of Ref. [9]. The dimensions of **rxns**, **mets**, and **genes** are respectively 1266, 1061, 750. We can use Cytoscape [12] to draw the GPR-network (**Figure 7**) of *S. cerevisiae*_ iND750, which contains 3077 nodes and 6666 edges.

The minimal media of *in silico* model is an important aspect. The computational minimal media of *S. cere-visiae*_iND750 is also included in Ref. [9]. In the method of constraint-based analysis, the biomass objective function (BOF) should be defined. The BOF was generated by defining all of the major and essential constituents that make up the cellular biomass content of *S. cerevisiae* [9]. Gene-protein-reaction associations embodied in **rxnGeneMat** matrix, which is a matrix with as many rows as there are reactions in the model and as many columns as there are genes in the model. The *i*-th row and the *j*-th column contains a one if the *j*-th gene in **genes** is associated with the *i*-th reaction in **rxns** and zero otherwise.

3.3. Methodology of constraint-based analysis

3.3.1. Constraint-based analysis

In silico modeling and simulation of genome-scale biological systems are different from that practiced in the physicochemical sciences. A network can fundamentally have many different states or many different solutions. Which states (or solutions) are picked is up to the cell and based on the selection pressure experienced, and such choices can change over time. Therefore, constraint-based approaches [2, 3] to the analysis of complex biological systems have proven to be very useful. This difference between the physicochemical sciences and the physical sciences or engineering is illustrated in Ref. [14]. All theory-based considerations (i.e., engineering and physics) lead one to attempt to seek an "exact" solution, and typically computed based on the laws of physics and chemistry. However, constraint-based considerations (as in biology) are useful. Not only can a network have many different behaviors that are picked based on the evolutionary history of the organism, but
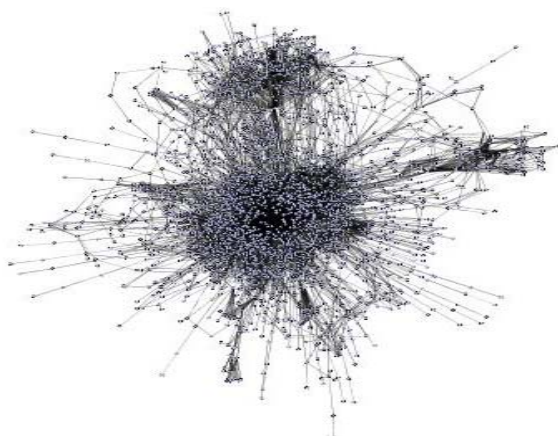
Figure 7. GPR-Network of *S. cerevisiae_* iND750, 3077 nodes and 6666 edges, Created by Cytoscape with Layout (yFiles-Organic)

also these networks can carry out the same function in many different and equivalent ways.

3.3.2. Representation of reconstructed metabolic network

Before calculation and simulation, the reconstructed metabolic network must be represented mathematically.The stoichiometric matrix, *S*, is the centerpiece of a mathematical representation of genome-scale metabolic networks. It represents each reaction as a column and each metabolite as a row, where each numerical element is the corresponding stoichiometric coefficient. A graphical form of the first few reactions of glycolysis and the corresponding stoichiometric matrix are shown in the **Figure 2** of Ref. [11].

An upper and lower bound for the allowable flux through each reaction also requires defining. This represents the lowest and highest reaction rate possible for each reaction. The set of upper and lower bounds is represented as two separate vectors, each containing as many components as there are columns in *S*, and in the same order. An example is shown in **Figure 2** of Ref. [11]. In many cases, reversible reactions are defined to have an arbitrary large upper bound and an arbitrarily large negative lower bound. Irreversible reactions have a lower bound that is nonnegative, usually zero.

In order to predict meaningful fluxes, setting upper and lower bounds is especially important for exchange reactions which serve to uptake compounds to the cell or secrete compounds from the cell. The lower bound of the exchange reaction column must be a finite negative number using this orientation (e.g., glucose). The upper bound of the exchange reaction column must be greater than zero. At least one of the reactions in the model must have a constrained lower/upper bound, and typically, the substrate (e.g., glucose or oxygen) uptake rates are set to experimentally measured values. The upper and lower bounds for exchange reactions are quantitative *in silico* representations of the growth media environment.

3.3.3. Biomass objective function (BOF) and minimal

media

The constraint-based approach is based on the assumption that cells strive to maximize their growth rate. This assumption which provides an acceptable starting point for many types of computations is satisfied by simulating maximal production of the molecules required to make new cells (biomass precursor molecules). In spite of their limitations, the predictive power of genome-scale models of metabolic networks has been demonstrated in diverse situations through careful experimentation [11].

The biomass objective function (the function v_{growth}, see below) is a special reaction taking as substrates of all biomass metabolites, ATP and water, and producing ADP, protons, and phosphate (as a result of the non-growth associated ATP maintenance requirement)[6].

The minimal media is determined computationally with the systematic testing of distinct inputs. Different combinations of molecules were allowed to enter the reaction network until the minimal group that allowed biomass production, or non-zero *Z* (see below), was found [6]. It is only concerned that some amount of biomass production is calculated but do not discriminate between extremely slow, inefficient growth and rapid growth.

3.3.4. Computation of phenotypic states

In genome-scale metabolic networks, the fluxes within a cell usually cannot be uniquely calculated because a range of feasible values exist when fluxes are subjected to known constraints. Flux-balance analysis (FBA) was used to find optimal growth phenotypes. Briefly, a large-scale linear programming was used to find a complete set of metabolic fluxes (*v*) that are consistent with steady-state condition (eq. 3) and reaction rate bounds (eq. 4), and at the same time maximize the biomass objective function in the defined ratio. This corresponds to the following linear programming problem [6]:

$$\text{Max } Z = v_{growth}$$

Subject to

$$S \cdot v = 0 \qquad (3)$$
$$\alpha_i < v_i < \beta_i \qquad (4)$$

where *S* is the stoichiometric matrix, and where α_i and β_i define the bounds through each reaction v_i. The flux range was set arbitrarily high for all internal reactions so that no internal reaction restricted the network, with the exception of irreversible reactions, which have a minimum flux of zero. The inputs to the system were restricted to a minimal media.

The value of *Z* computed with the above procedure can either be zero (predicting no growth) or greater than zero (corresponding to cellular growth) depending on the inputs and outputs that are allowed, according to the nutrients provided in the media.

3.3.5. Gene deletion study

The effect of a gene deletion experiment on cellular

JBiSE

growth can be simulated in a manner similar to linear optimization of growth [5, 11]. Gene–reaction associations model the logical relationship between genes and their corresponding reactions. If a single gene is associated with multiple reactions, the deletion of that gene will result in the removal of all associated reactions, i.e. to simultaneously restrict the fluxes (upper and lower flux bounds) of these reactions to zero prior to computing maximal biomass objective function. On the other hand, a reaction that can be catalyzed by multiple non-interacting gene products will not be removed in a single gene deletion. The possible results from a simulation of a single gene deletion are unchanged maximal growth (non-lethal), reduced maximal growth or no growth (lethal). Those genes were considered essential if no biomass could be produced without their usage.

4. DISCUSSION

Based on the gene-protein-reaction model of *S. cerevisiae*_iND750 and the methodology of constraint- based analysis and by the aid of the COBRA toolbox and MATLAB software, we have calculated the deletion impact of 438 calculable genes, one by one, on the metabolic flux redistribution of *S. cerevisiae*_iND750.

We found that both of the *v-d* correlation and the *v-v_{gene}* correlation were not of linear relation. Although some properties about the metabolic network of microorganisms have been reported in literatures [15-19], our research will provide further evidences to the properties about the metabolic network, because the measure we defined is different.

Furthermore, we sought out 38 important genes that most greatly affected the metabolic flux distribution, determined their functional subsystems and found that many of 38 key genes were related to many but not several subsystems. From these results, we speculate that many but not several subsystems are important subsystems in the metabolism of *S. cerevisiae* and that this may increase the robustness of the metabolic network.

As a next step, we will do similar research on other organisms and compare them with the case of *E. coli*. Although it is theoretically possible to attempt double deletion of every possible gene pair experimentally, the sheer number of possible two-gene deletions makes this virtually impossible. However, computational predicttions of double gene deletion phenotypes can be made in a matter of hours [11]. This will also become our work in the future.

ACKNOWLEDGEMENTS

REFERENCES

[1] Bork, P. (2005) Is there biological research beyond Systems Biology? A comparative analysis of terms. *Mol. Syst. Biol*, **1**, Art. No. 2005. 0012.

[2] Price ND, Papin JA, Schilling CH, Palsson BO. (2003) Genome-scale microbial *in silico* models: the constraints-based approach. *Trends Biotechnol*, **21** (4), 162-169.

[3] Price N D, Reed J L, Palsson B O. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology*, **2**(11), 886-897.

[4] J.L. Reed, T.D. Vo, C.H. Schilling, and B.O. Palsson. (2003) An expanded genome scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biology*, **4**, R54.1– R54.12.

[5] Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., Karp, P.D., Broadbelt, L.J., Hatzimanikatis, V., Palsson, B.Ø. (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*, **3**, Art. No. 121.

[6] Scott A Becker, Bernhard Ø Palsson. (2005) Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. *BMC Microbiol*, **5**, Art. No. 8.

[7] Ines Thiele, Thuy D. Vo, Nathan D. Price, and Bernhard Ø. Palsson. (2005) Expanded metabolic reconstruction of *Helicobacter pylori* (iIT341 GSM/GPR): an *in silico* genome-scale characterization of single- and double- deletion mutants. *Journal of Bacteriology*, **187** (16), 5818-5830.

[8] Adam M Feist, Johannes C M Scholten, Bernhard Ø Palsson, Fred J Brockman, Trey Ideker. (2006) Modeling methanogenesis with a genome-scale metabolic recon-struction of *Methanosarcina barkeri*. *Molecular Systems Biology*, **2**, Art. No. 2006.0004.

[9] Natalie C. Duarte, Markus J. Herrgård and Bernhard Ø. Palsson. (2004) Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome- scale metabolic model . *Genome Res*, **14** (7), 1298 -1309.

[10] Oh YK, Palsson BO, Park SM, Schilling CH, Mahadevan R. (2007) Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. *J Biol Chem*, **282** (39), 28791-28799.

[11] Becker, S.A., Feist, A.M., Mo, M. L., Hannum, G., Palsson, B.O., Herrgard, M.J. (2007) Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox. *Nat. Protocols*, **2**(3), 727-738.

[12] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel amage, Nada Amin, Benno Schwikowski and Trey Ideker. (2003) Cytoscape: a software environment for integrated models of biomo- lecular interaction networks. *Genome Res*, **13** (11), 2498-2504.

[13] E. Almaas, B. Kova´cs, T. Vicsek, Z. N. Oltvai3& A.-L. Baraba´ si. (2004) Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature*, **427** (6997), 839-843.

[14] B.O. Palsson . (2006) Systems biology properties of reconstructed networks. Cambridge university press, Cambridge, New York.

[15] Förster, J., Famili, I., Fu, P.C., Palsson, B.Ø., and Nielsen, J. (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res*, **13**(2), 244–253.

[16] R. Mahadevan and B. O. Palsson. (2005) Properties of Metabolic Networks: Structure versus Function. *Biophysical Journal*, **88**(1), L07-L09.

[17] Bala´zs Papp, Csaba Pa´l , Laurence D. Hurst. (2004) Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature*, **429**(6992), 661-664.

[18] Iman Famili, Jochen Förster, Jens Nielsen, Bernhard O. Palsson. (2003) *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(23), 13134-13139.

[19] Jochen Förster, Iman Famili, Bernhard Ø. Palsson, Jens Nielsen. (2003) Large-scale evaluation of *in silico* gene deletions in *Saccharomyces cerevisiae*. *Omics*, **7**, 193-202.