

Comparative analysis of metabolic network of pathogens

Kumar Gaurav, Yasha Hasija (✉)

Department of Biotechnology, Delhi Technological University, Delhi – 110042, India

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2017

BACKGROUND: Metabolic networks are complex and system of highly connected chemical reactions and hence it needs a system level computational approach to identify the genotype- phenotype relationship. The study of essential genes and reactions and synthetic lethality of genes and reactions plays a crucial role in explaining functional links between genes and gene function predictions.

METHODS: Flux balance analysis (FBA) has been developed as a powerful method for the *in silico* analyses of metabolic networks. In this study, we present the comparative analysis of the genomic scale metabolic networks of the four microorganisms i.e. *Salmonella typhimurium*, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, and *Helicobacter pylori*. The fluxes of all reaction were obtained and the growth rate of the organism was calculated by setting the biomass reaction as the objective function.

RESULTS & CONCLUSIONS: The average lethality fraction of all the four organisms studied ranged from 0.2 to 0.6. It was also observed that there are very few metabolites which are highly connected. Those metabolites that are highly connected are supposed to be the ‘global players’ similar to the hub protein in the protein – protein interaction network.

Keywords essential genes, synthetic lethal genes, metabolite connectivity, robustness analysis

Introduction

As a complex interactive nonlinear system, much attention has been gained by the study related to the metabolic networks in recent years. Efforts have also been made for the analysis of the structure of whole metabolic networks (Feist and Palsson, 2008). The use of flux balance analysis in order to predict the flux distribution in an entire metabolic network under certain physiological conditions is to analyze and compare the metabolic networks related to different organisms. There is already the availability of genome scale reconstructed models of the different organisms like *Helicobacter pylori*, *Mycobacterium tuberculosis* etc. These metabolic networks are obtained from different metabolic models databases (Hamilton and Reed, 2012). Flux Balance Analysis (FBA) is a widely used constraint based approach for studying biochemical metabolic networks, mainly the genomic scale metabolic network that have been reconstructed in the past decade. It uses the concept of linear programming. Linear programming obtains the maximum potential of the objective function that we are looking at, and

therefore, when using flux balance analysis, a single solution is obtained for the optimization problem (Raman and Chandra, 2009). FBA calculates the flow of metabolites taking place in the metabolic network, hence making it possible to predict the growth speed of an organism or the rate of production of a biochemically important metabolite (Edwards and Palsson, 2002). Essential genes comprise of genes whose specific deletion is lethal under a specific environmental condition. Identification of essential genes in the metabolic network of microorganisms helps to identify potential drug targets and in understanding of minimal requirements for a synthetic cell. However, experimental assessment of essentiality of the coding genes of metabolic network is resource intensive and not viable for all bacterial organisms, in particular if they are infective (Plaimas et al., 2010; Chowdhury et al., 2015). A constraint-based analysis of reconstructed metabolic networks has proved to be quite useful in various applications such as metabolic engineering prediction of outcomes of gene deletions, drug-target identification and in the elucidation of cellular regulatory networks (Raman et al., 2005). By analogy, Synthetic Lethals (SLs) refer to the pair of non-essential genes whose simultaneous deletion is proved to be lethal. Synthetic gene lethality can arise for different types of reasons. For example, two protein products can be interchanged with respect to an essential function that act in similar pathway or function in

Received August 30, 2016; accepted December 20, 2016

Correspondence: Yasha Hasija

E-mail: yashahasija@gmail.com

two distinct pathways with redundant or complementary essential functions. The study of synthetic lethality plays an important role in explaining functional links between genes and the prediction of the gene functions (Suthers et al., 2009; Chowdhury et al., 2015; Raman et al., 2015). Occurrence of synthetic lethality can be seen between genes and small molecules, and it can be used to explain the mechanism of action of drugs.

The true potential of synthetic lethality has been widely studied in yeast (Nijman, 2011). Biological systems are prone to mutation, to the variation in environmental conditions and to random variations in the abundance of component molecules of the biological systems. Many physiological and developmental systems are robust to such disturbances and hence immune to any such fluctuations (Masel and Siegal, 2010). Biological systems that have been experimentally proved of being robust to the significant changes in their environmental conditions need mathematical models for robustness analysis of the metabolic network. These mathematical models should themselves be robust. The necessary condition for the model robustness is that the dynamics of the model should be insensitive to small variations in the parameters of the model. The model dynamics may be very much sensitive to simultaneous parameter variations (Kim et al., 2006). Phenotype phase plane analysis can study the optimal utilization of the metabolic network as a function of the constraints. At present metabolic flux maps are typically calculated for the single growth condition, hence gives a limited view of the metabolic phenotype – genotype relation. But phenotypic phase plane analysis maps all the growth conditions characterized by two environmental variables presented in the single plane (Edwards et al., 2001). Fast-SL is an algorithm which computes combinations of reactions, which when deleted, leads to the elimination of the growth of the organism due to the modification in the metabolic pathway. It achieves this by a combination of limiting the search space and exhaustively iterating through the remaining combinations (Pratapa et al., 2015). The functional *in silico* model of *Mycobacterium tuberculosis*, iNJ6611, contains 661 genes, 1025 reactions and 826 metabolites. Genome scale models can be used as hypothesis generating tools and also for analysis and discovery which will expectantly support the rational drug development process (Jamshidi and Palsson, 2007). The functional *in silico* model of bacterium *Staphylococcus aureus*, iSB619, consists of 619 genes by which 640 metabolic reactions are catalyzed. The reaction list is the most complete till date for this pathogen (Becker and Palsson, 2005; Heinemann et al., 2005). The reconstructed network, iIT341 represents a detailed review of the current literature about *Helicobacter pylori* as it integrates biochemical and genomic data in a comprehensive framework. In total, it contains 341 metabolic genes, 476 intracellular reactions, 78 exchange reactions, and 485 metabolites (Schilling et al., 2002; Palsson et al., 2005). *Salmonella enterica* subspecies I serovar Typhimurium is a human pathogen which causes

different diseases and its increasing antibiotic resistance poses many public health problems. Different diseases caused by *Salmonella* are mainly Typhoid fever, Paratyphoid fever and Food poisoning. The functional *in silico* model of *Salmonella typhimurium*, STM_v1_0, consists of 1271 genes, 1802 metabolites and 2545 reactions (McClelland et al., 2001; Thiele et al., 2011). Methods of Constraint Based Reconstruction and Analysis (COBRA) have been successfully applied in the field of microbial metabolic engineering and are being extended to model transcriptional and signaling networks and in the field of public health. The COBRA approach focuses on using physicochemical and biological constraints to analyze the set of feasible phenotypic states of a reconstructed biological network under a given condition (Schellenberger et al., 2011). SBMLToolbox, a toolbox that facilitates importing and exporting models presented in the Systems Biology Markup Language (SBML) format in the MATLAB environment. It also provides functionality that helps an experienced user of either SBML or MATLAB to combine the computing power of MATLAB with the exchangeability and portability of an SBML model. SBMLToolbox supports all levels and versions of SBML (Keating et al., 2006). In this work, we describe the similarities and dissimilarities of metabolic networks of *Salmonella typhimurium* (STM_v1_0), *Helicobacter pylori* (iIT341), *Mycobacterium tuberculosis* (iNJ661) and *Staphylococcus aureus* (iSB619) by pairwise comparison with the help of flux balance analysis and gene deletion method.

Materials and methods

The comparative study of metabolic network for finding the similarity and dissimilarities is conducted for the following four pathogenic bacteria.

- 1) *Mycobacterium tuberculosis* (iNJ661),
- 2) *Staphylococcus aureus* (iSB619),
- 3) *Salmonella typhimurium* (STM_v1_0),
- 4) *Helicobacter pylori* (iIT341).

The genome scale metabolic networks of the four pathogenic organisms were collected from different sources like BiGG database, BioModel database and TheModelSeed database. The MATLAB compliant files of the four given organisms were downloaded from the BiGG Database.

Flux balance analysis

The constraints in FBA have the form $\alpha \leq v_i \leq \beta$ where vector v_i represents all the fluxes in the given metabolic network, α and β are the lower and upper limits. Thermodynamic constraints according to the reversibility and irreversibility of a reaction can be applied by setting the lower limit (α) for the equivalent flux to 0 if the reaction is irreversible and -1000 if the reaction is reversible. The upper limit is set to 1000 for reversible and irreversible reactions. The COBRA toolbox was used to run the FBA. For constraint based modeling,

COBRA toolbox is used which has a collection of MATLAB scripts that are run from within the MATLAB environment. The MATLAB compliant metabolic network files of each of the four organisms were loaded in the MATLAB. A specific metabolic flux distribution within the feasible set was found by using linear programming (LP). The Gurobi solver package was used to perform linear programming in COBRA Toolbox. The fluxes of all reaction were obtained and the growth rate of the organism was calculated by setting the biomass reaction as the objective function.

Simulating gene knockouts

Gene-protein-reaction associations are embodied in *rxnGeneMat* matrix in the MATLAB files, which is a matrix with rows equal to number of reactions in the model and columns equal to number of genes in the model. The function – *singleGeneDeletion* and *doubleGeneDeletion* performs the single and double gene deletion respectively.

Performing synthetic lethality analysis

It was performed using the Fast – SL algorithm that is implemented using the MATLAB.

Code to perform the synthetic lethal reaction analysis-fastSL (*model, cutoff, order, eliList, atpm*)

Code to perform the synthetic lethal gene analysis-fastSLgenes (*model, cutoff, order, flag*)

where *model* is metabolic network file to be analyzed, *cutoff* is cutoff percentage value for lethality and have default value of 0.01, *order* is order of synthetic lethals and have default value of 2 and maximum value 3, *eliList* is the list of the reactions to be excluded from the analysis, *flag* is set to 1 for rigorous search and have default value 0, *atpm* is ATP Maintenance Reaction Id in model.rxns if other than ATP Maintenance Reaction.

Robustness analysis

Robustness analysis uses the FBA to analyze the metabolic network properties. In this step, the flux was varied through one reaction and the optimal objective value was calculated as a function of this value. This tells about the sensitivity of the objective value to a particular reaction. We have checked the effect of varying glucose and oxygen uptake on the growth rate. COBRAToolbox has the function *robustnessAnalysis* to perform this method.

Phenotypic phase plane analysis

While performing the Robustness analysis, we varied one parameter and made another parameter to be constant. However, in phenotypic phase plane analysis, we varied two parameters glucose and oxygen uptake rate simultaneously and the results were plotted as phenotypic phase plane. It revealed the interaction between two reactions. COBRA Toolbox has the built in function *phenotypePhasePlane* to perform this step.

Metabolite connectivity

Metabolite connectivity is the number of reactions in which the metabolites occur in. In this step the Stoichiometric matrix *S* of the metabolic model was converted into the binary matrix *S_{bin}*. It replaces all nonzero elements in the matrix *S* with ‘1’ in the binary matrix. Then all ‘1s’ in each row of the *S_{bin}* was added up to determine metabolite connectivity.

Reaction essentiality and metabolite connectivity

Correlation between essentiality of reaction and metabolite connectivity can be estimated in the metabolic network. In this step, all the reactions which were associated with the metabolites were deleted and checked whether it could still produce the biomass or not. Reactions associated with the metabolites could be found out by scanning through the corresponding rows in the Stoichiometric *S* matrix. After then all the associated genes were knocked out and FBA was used to predict the possibility of the growth. Lethality fraction was calculated based on the above result. The semi log graph was plotted with this lethality fraction vs. metabolite connectivity.

Results

Calculation of growth rate using FBA

The growth rates of the four organisms were calculated using flux balance analysis. The maximization of biomass was set as the objective function to calculate the growth rate. The growth rate is calculated in the unit mmol/(gDW·h) (millimole/(gramDryWeight·hour)). The growth rates of four organisms are shown in Table 1.

The table showing growth rates, are the growth rates calculated for the wild type strain of each of the model organism. The objective function for each of the organism was

Table 1 Growth rates of all the four organisms using FBA

Organism	Number of reactions	Number of metabolites	Number of genes	Growth rate (mmol/(gDW·h))
<i>Mycobacterium tuberculosis</i>	1025	826	661	0.0522
<i>Staphylococcus aureus</i>	743	655	619	0.0687
<i>Salmonella typhimurium</i>	1076	982	801	0.4779
<i>Helicobacter pylori</i>	554	485	339	0.6928

different. Apart from this, growth rates under the alternative substrates like glucose, succinate, pyruvate etc. were also calculated. The calculation of growth rates was done under both aerobic and anaerobic conditions. The growth rates calculated under the alternative substrates were different compared to the wild type ones. For most of the substrate, growth rate under the anaerobic condition was found to be 0 for each of the organism studied. Alternative substrates under which growth rates were calculated are the important substrates that play an important role in the major metabolic pathways like glycolysis, TCA cycle, oxidative phosphorylation, amino acid metabolism etc. The growth rates calculated under alternative substrates are shown in the given below tables for each of the organism separately. The maximum substrate uptake rate was set to $-20 \text{ mmol}/(\text{gDW} \cdot \text{h})$ for every substrate. The calculated growth rates under the alternative substrates in the metabolic networks of each of the four different organisms have been shown in the tables from Tables 2 to 5.

Retrieving essential reactions and essential genes from gene knockouts

If a single gene is associated with multiple reactions, the deletion of that gene results in the removal of all reactions

Table 2 Growth rates of *Mycobacterium tuberculosis* under the alternative substrates

Substrate	Aerobic (mmol/(gDW·h))	Anaerobic (mmol/(gDW·h))
Glucose	0.3060	0
Succinate	0.1393	-2.0512e-26
Aspartate	0.0680	3.9109e-20
Phosphate	0.1440	-1.7094e-27
Glycerol	0.3060	6.9372e-26
Pyruvate	0.3060	-9.6864e-27

Table 3 Growth rates of *Staphylococcus aureus* under the alternative substrates

Substrate	Aerobic (mmol/(gDW·h))	Anaerobic (mmol/(gDW·h))
Glucose	0.0141	2.2559e-26
Succinate	Absent	Absent
Aspartate	0.0123	0
Phosphate	0.0103	0
Glycerol	-9.6865e-18	-4.1655e-18
Pyruvate	Absent	Absent

Table 4 Growth rates of *Helicobacter pylori* under the alternative substrates

Substrate	Aerobic(mmol/(gDW·h))	Anaerobic(mmol/(gDW·h))
Glucose	4.1510	0
Succinate	5.3325	0
Aspartate	5.2489	0
Phosphate	4.1510	0
Glycerol	Absent	Absent
Pyruvate	4.9698	1.5179e-29

Table 5 Growth rates of *Salmonella typhimurium* under the alternative substrates

Substrate	Aerobic (mmol/(gDW·h))	Anaerobic (mmol/(gDW·h))
Glucose	1.9317	0.4500
Succinate	1.5507	0.0525
Aspartate	1.4971	0.2959
Phosphate	0.4779	0.0525
Glycerol	1.1227e-12	-9.3396e-13
Pyruvate	1.1584e-12	-4.6855e-13

associated to the genes. On the other hand, a reaction that can be catalyzed by multiple gene products which do not interact with one other will not be removed in a single gene deletion. The essential reactions were extracted by reaction knockout method. The essential genes were extracted by gene knockout method. By the aid of the COBRA toolbox, we can retrieve the essential reactions in the metabolic network. After gene knockout, a mutant strain of the organism is generated and it also provides the grRatio (gene-reaction ratio) of the mutant strain and wild type strain of the organisms. The grRatio is the ratio between the grRatio of wild type strain and the grRatio of the mutated strain. Those genes and reactions are the essential genes and essential reactions of the organisms, if the grRatio is less than 0.05. Key metabolites affected by gene deletion were 3PG(3-phosphoglycerate), AcCoA (acetyl CoA), ADP, AKG(α -ketoglutaric acid), ATP, CoA(coenzyme-A), E4P(erythrose-4-phosphate), F6P(fructose-6-phosphate), G3P(glycerol-3-phosphate), G6P(glucose-6-phosphate), Gln-L(L-glutamine), Glu-L(L-glutamic acid), H₂O, H(hydrogen), NAD, NADH, NADP, NADPH, OAA (oxaloacetic acid), PEP(phosphoenol pyruvate), Pi(inorganic phosphate), pyr(pyruvate), R5P(ribose-5-phosphate). Number of essential reactions and essential genes present in metabolic network of each organism has been shown in Table 6. The essential reactions thus obtained for each of the organism can be compared pairwise with essential reactions of each of the organism. The pairwise comparison was performed with the help of Compare plugin in the Notepad++. Common essential reactions thus obtained by the pairwise comparison are shown in the below given Table 7.

Table 6 Essential genes and essential reactions present in metabolic networks of given organisms.

Organism	<i>S. aureus</i>	<i>M. tuberculosis</i>	<i>S. typhimurium</i>	<i>H. pylori</i>
Number of essential reactions	244	314	345	256
Number of essential genes	168	188	201	186

In the recent years, genome-scale metabolic networks have been reconstructed for many organisms (Thiele et al., 2005, 2011). These networks have been studied using tools such as Flux Balance Analysis (FBA) (Kauffman et al., 2003), for the

Table 7 Common essential reactions compared among four different organisms.

Pairwise combination of organisms	<i>M. tuberculosis</i>	<i>M. tuberculosis</i>	<i>M. tuberculosis</i>	<i>S. aureus</i>	<i>S. aureus</i>	<i>S. typhimurim</i>
	<i>S. typhimurim</i>	<i>S. aureus</i>	<i>H. pylori</i>	<i>S. typhimurim</i>	<i>H. pylori</i>	<i>H. pylori</i>
No. of common essential reactions	48	47	96	60	93	58

identification of targets for metabolic engineering (Alper et al., 2005) and to understand the robustness of organisms through systematic experimental evaluation of gene knock-outs (Deutscher et al., 2006). Another important aspect of analyzing these reconstructed metabolic networks is the identification of combinations of genes, which when simultaneously deleted, abolish growth *in silico* (Deutscher et al., 2006; Suthers et al., 2009).

Essential genes and essential reactions obtained by the gene knockout have been listed in the Table 8 and Table 9 respectively. Since this is a quantitative study, the list contains only selected genes which were found to have scientific evidence in the previous studies. They have been analyzed in the past for prediction of novel genetic interactions and analyzing the extent of robustness of biological networks (Raghunathan et al., 2009).

Table 8 List of essential genes

Organism	<i>Mycobacterium tuberculosis</i>	<i>Staphylococcus aureus</i>	<i>Helicobacter pylori</i>	<i>Salmonella typhimurim</i>
Essential genes	<i>fadD26</i>	<i>murA</i>	<i>acpS</i>	<i>rpiA</i>
	<i>ppsA</i>	<i>murB</i>	<i>aroQ</i>	<i>yedJ</i>
	<i>ppsE</i>	<i>fabA</i>	<i>dcd</i>	<i>talA</i>
	<i>drrA</i>	<i>fabB</i>	<i>eda</i>	<i>talB</i>
	<i>pks1</i>	<i>fabG</i>	<i>edd</i>	<i>talc</i>
	<i>pks10</i>	<i>fabD</i>	<i>fbp</i>	<i>sera</i>
	<i>pks7</i>	<i>fabH</i>	<i>glcD</i>	<i>glyA</i>
	<i>mmpL1</i>	<i>fabZ</i>	<i>gpt</i>	<i>purN</i>
	<i>mmpL2</i>	<i>accA</i>	<i>mgo</i>	<i>purT</i>
	<i>mmpL3</i>	<i>accD</i>	<i>murB</i>	<i>tdk</i>
	<i>mmpL12</i>	<i>hisA</i>	<i>murF</i>	<i>dut</i>
	<i>mmpL4</i>	<i>hisB</i>	<i>oorA</i>	<i>epimerase</i>
	<i>mmpL11</i>	<i>hisC</i>	<i>oorB</i>	<i>argB</i>
	<i>mmpL6</i>	<i>hisI</i>	<i>oorC</i>	<i>rpe</i>
	<i>fadD26</i>	<i>hisG</i>	<i>oorD</i>	<i>argD</i>
	<i>drrA</i>	<i>hisD</i>	<i>porA</i>	<i>aroE</i>
	<i>drrC</i>	<i>aroA</i>	<i>porB</i>	<i>prsA</i>
	<i>papA5</i>	<i>aroB</i>	<i>porC</i>	<i>cytosine deaminase</i>
	<i>mas</i>	<i>aroC</i>	<i>porD</i>	<i>speB</i>
	<i>fadD28</i>	<i>aroE</i>	<i>pyre</i>	<i>speC</i>
	<i>mmpL7</i>	<i>aroQ</i>	<i>speE</i>	<i>speF</i>
	<i>mmpL5</i>	<i>trpA</i>	<i>tal</i>	<i>tktA</i>
	<i>mmpL8</i>	<i>trpB</i>	<i>trxA</i>	<i>tktB</i>
	<i>mmpL9</i>	<i>trpF</i>	<i>trxB</i>	<i>nadA</i>
	<i>mmpL10</i>	<i>aroK</i>		<i>nadB</i>
		<i>aroL</i>		<i>thrC</i>
		<i>aroP</i>		<i>gsk</i>
		<i>aroH</i>		<i>kbl</i>
		<i>trpD</i>		
		<i>pheC</i>		
		<i>cysG</i>		
		<i>hemC</i>		
		<i>hemB</i>		
	<i>hemD</i>			
	<i>metK</i>			

Gene knockout performed on each of the organism can be visualized using the MATLAB. Double gene deletion was performed for all the genes present in each of the organism. The lethality of double gene deletion can be visualized where the scale of lethality varies from 0 to 1. Scale of 1 refers to highly lethal deletion for the particular two genes and scale of 0 refers the lethal effect to be normal.

Synthetic lethality analysis

It was performed using the algorithm Fast-SL algorithm. It is MATLAB implemented algorithm using COBRAToolbox. The synthetic gene lethality and the synthetic reaction lethality were performed separately. For synthetic gene lethality, both single and double gene lethality was performed. Similarly for synthetic reaction lethality, both single and double reaction lethality was performed. In single gene lethality there was deletion of single gene and similarly in double gene lethality, pair of genes was deleted. Similar procedure was followed for reaction lethality also. The result for gene lethality has been shown in the Table 10. The result for Reaction Lethality has been shown in the Table 11. The table shows the number of single lethal reactions and double lethal reactions. Single lethal reactions were identified by performing exhaustive single reaction deletions. Double lethal reactions were identified by performing pairwise deletions.

Robustness analysis

Robustness analysis was performed to study the effect of the glucose and oxygen uptake on the growth rates. This analysis was performed for metabolic network of each of the organism. The sensitivity of the growth of the metabolic network was checked against the varying glucose and oxygen uptake rate. To determine the effect of varying glucose uptake on growth, oxygen uptake rate was fixed at 17 mmol/(gDW · h) which is a realistic uptake of oxygen. To determine the effect of varying oxygen uptake on growth, glucose uptake rate was fixed at 10 mmol/(gDW · h) which is a realistic uptake of glucose. Results for the robustness analysis of the metabolic networks of each of the four organisms have been shown in the figures from Fig. 1 to Fig.8. Figures 1 and 2 represent the Robustness study of metabolic network of *Mycobacterium tuberculosis*. In Fig. 1, with oxygen uptake rate kept constant, growth rate remained 0/h until a glucose uptake of around 2.5 mmol/(gDW · h). Growth reached to the maximum level of 0.3/h at the glucose uptake of around 5 mmol/(gDW · h). After the maximum growth rate of 0.3/h growth rate became constant.

Table 9 List of essential reactions

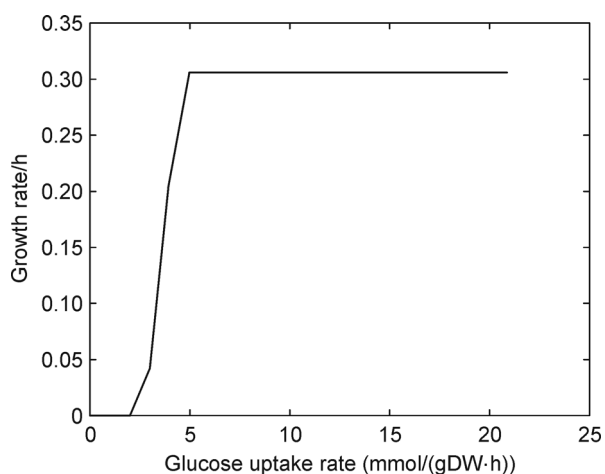
Organism	<i>Mycobacterium tuberculosis</i>	<i>Staphylococcus aureus</i>	<i>Helicobacter pylori</i>	<i>Salmonella typhimurium</i>
Essential reactions	'AACPS10'	'ACCOAC'	A5PISO'	'3HAD100'
	'AACPS11'	'ACGK'	'ACCOACr'	'3HAD121'
	'AACPS3'	'ACHBS'	'ADCL'	'3HAD141'
	'ACACT1r'	'ACLS'	'ADK1'	'3HAD161'
	'ACCC'	'ACOTA'	'ADMDCr'	'3HAD181'
	'ACCOACr'	'ADCL'	'ADSL1r'	'3HAD40'
	'ACChex'	'ADCS'	'ADSL2r'	'3HAD60'
	'ACGAMT'	'ADK1'	'ADSS'	'3HAD80'
	'ACGK'	'ADKd'	'AGMHE'	'3OAR100'
	'ACHBS'	'ADMDCr'	'AGMT'	'3OAR121'
	'ACLS'	'ADPT'	'AHCYSNS'	'3OAR140'
	'ACOTA'	'ADSK'	'AICART'	'3OAR141'
	'ADSK'	'ADSL1'	'AIRC2r'	'3OAR161'
	'ADSL1r'	'ADSL2r'	'AIRC3'	'3OAR181'
	'ADSL2r'	'ADSS'	'ALAALAr'	'3OAR40'
	'ADSS'	'AGPR'	'AMAOTr'	'3OAR60'
	'AFE'	'AICART'	'AMMQT6'	'3OAR80'
	'AFTA'	'AIRC2'	'ANPRT'	'3OAS100'
	'AGPAT160'	'AIRC3'	'AOXSr'	'3OAS121'
	'AGPAT160190'	'AKGDa'	'ARGDC'	'3OAS140'
	'AGPAT190'	'AKGDb'	'ARGN'	'3OAS141'
	'AGPR'	'ALAR'	'ARGt2r'	'3OAS161'
	'AHC'	'ALATA_D'	'ASAD'	'3OAS181'
	'ALAALAr'	'ANPRT'	'ASNS2'	'3OAS60'
	'ALAR'	'ANS'	'ASP1DC'	'3OAS80'
	'ANPRT'	'APRAUR'	'ASPCT'	'A5PISO'
	'ANS'	'ASAD'	'ASPK'	'ACCOAC'
	'ARABF'	'ASNS1'	'ASPO2'	'ACGK'
	'ARABI'	'ASPKi'	'ASPTA'	'ACGS'
	'ARACHTA'	'ASPTA'	'BC10_new'	'ACHBS'
	'ASAD'	'ATPPRT'	'BTS2'	'ACLS'
	'ASNS1'	'BPNT'	'BiomassHP_published'	'ACODA'
	'ASPCT'	'CHORM'	'C140SN'	'ACONTa'
	'ASPK'	'CHORS'	'C160SN'	'ACONTb'
	'ASPTA'	'CLPNS_SA'	'C180SN'	'ACOTA'
	'ATPPRT'	'CO2t'	'C181SN'	'ADCL'
	'BDH'	'CPPPGO'	'C190cSN'	'ADCS'
	'BPNT'	'CSNt2'	'CBPS'	'ADMDC'
	'CBPS'	'CYSS'	'CDPMEK'	'ADSK'
	'CDPMEK'	'CYSTL'	'CHCOAL'	'ADSL1r'
	'CDPPT160190'	'CYTBD'	'CHORM'	'ADSL2r'
	'CHORM'	'CYTD'	'CHORS'	'ADSS'
	'CHORS'	'CYTK1'	'CLPNS_HP'	'AGMHE'
	'CHRPL'	'DAPDC'	'CTPS1'	'AGPAT140'
	'CLPNS160190'	'DAPE'	'CYOO_HP'	'AGPAT160'
	'CMCBTFL'	'DASYN_SA'	'CYSS'	'AGPAT161'
	'CYSS'	'DB4PS'	'CYTK1'	'AGPAT180'
	'CYSTL'	'DDPA'	'DAPDC'	'AGPAT181'
	'DAPDC'	'DGK1'	'DAPE'	'AGPR'
	'DAPE'	'DHAD1'	'DASYN_HP'	'AHCYSNS'
	'DASYN160'	'DHAD2'	'DBTSr'	'AICART'
	'DASYN160190'	'DHDPRy'	'DDPA'	'AIRC2'
	'DASYN190190'	'DHDPS'	'DHDPRy'	'AIRC3'
	'DDPA'	'DHFR'		'ALAALAr'

Table 10 Synthetic gene lethality.

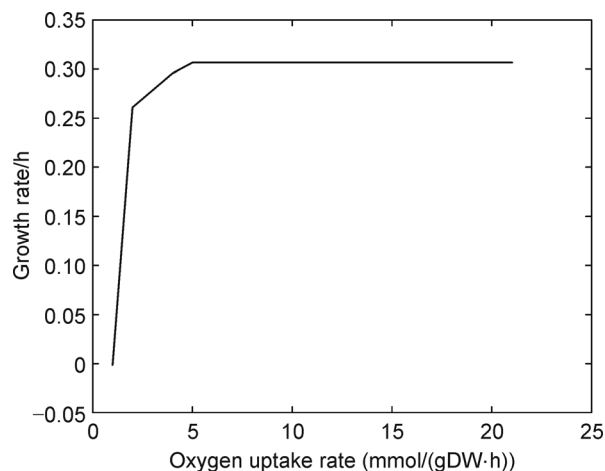
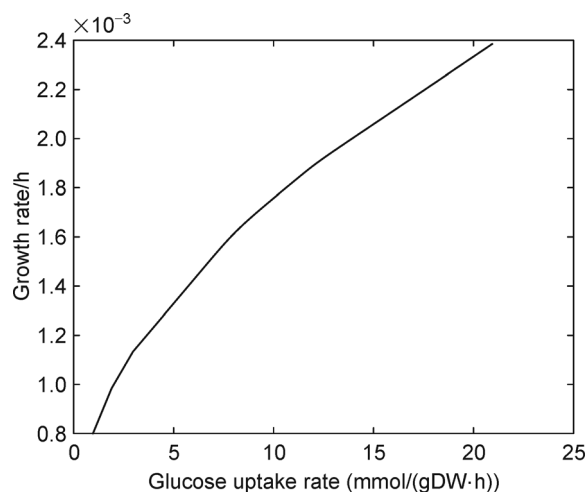
Organisms	Number of genes for single lethality	Number of genes for double lethality
<i>Mycobacterium tuberculosis</i>	188	188
<i>Staphylococcus aureus</i>	168	168
<i>Salmonella typhimurium</i>	202	204
<i>Helicobacter pylori</i>	168	168

Table 11 Synthetic reaction lethality

Organism	Number of reactions for single lethality	Number of reactions for double lethality
<i>Mycobacterium tuberculosis</i>	314	314
<i>Staphylococcus aureus</i>	244	244
<i>Salmonella typhimurium</i>	346	346
<i>Helicobacter pylori</i>	135	244

**Figure 1** Robustness analysis of metabolic network of *Mycobacterium tuberculosis* with oxygen uptake rate fixed at 17 mmol/(gDW·h).

In Fig. 2, with glucose uptake rate kept constant, growth rate steadily increased with increase in oxygen uptake. Figures 3 and 4 represent the Robustness study of metabolic network of *Staphylococcus aureus*. In Fig.3, with oxygen uptake rate kept constant, growth rate steadily increased with increase in glucose uptake rate. In Fig.4, with glucose uptake rate kept constant, growth rate increased steadily for most of the part with increase in oxygen uptake rate. At oxygen uptake of around 2 mmol/(gDW·h), growth rate remained constant at 0.6/h. Figure5 and Fig.6 represent the Robustness study of metabolic network of *Helicobacter pylori*. In Fig.5, with oxygen uptake rate kept constant, growth rate steadily increased with increase in glucose uptake rate. In Fig.6, with glucose uptake rate kept constant, growth rate increased steadily for most of the part with increase in oxygen uptake rate. Figure7 and Fig.8 represent Robustness study of metabolic network of *Salmonella typhimurium*. In Fig.7, with oxygen uptake rate kept constant, growth rate steadily

**Figure 2** Robustness analysis of metabolic network of *Mycobacterium tuberculosis* with glucose uptake rate fixed at 10 mmol/(gDW·h).**Figure 3** Robustness analysis of metabolic network of *Staphylococcus aureus* with oxygen uptake rate fixed at 17 mmol/(gDW·h).

increased with increase in glucose uptake rate up to the growth rate of around 0.62/h. Then after growth rate becomes constant. In Fig.8, with glucose uptake rate kept constant, growth rate remains 0/h for the oxygen uptake rate of around 12 mmol/(gDW·h). Then after growth rate starts increasing steadily up to the growth rate of 0.62/h and becomes constant.

Phenotypic phase plane analysis

Phenotypic phase plane analysis was performed by varying two substrates simultaneously. In this study phenotypic phase plane analysis is done by varying the glucose and oxygen uptake rate simultaneously. It assesses the effect of the simultaneous variation of the uptake rate of the glucose and oxygen on the growth rate. The effect was plotted on the 2-D graph as well as on the 3-D graph. The plot is divided into

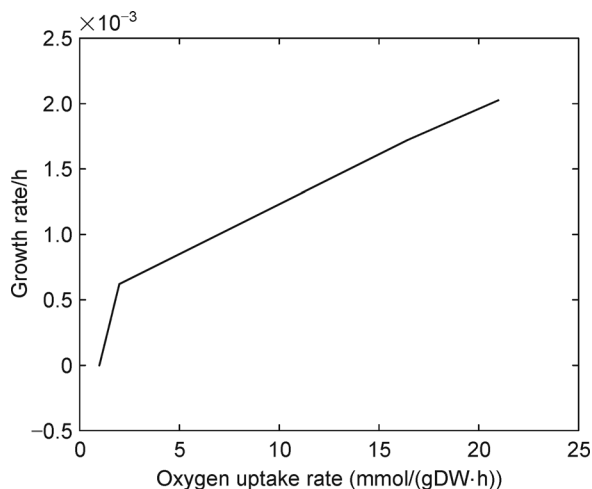


Figure 4 Robustness analysis of metabolic network of *Staphylococcus aureus* with glucose uptake rate fixed at 10 mmol/(gDW ·h).

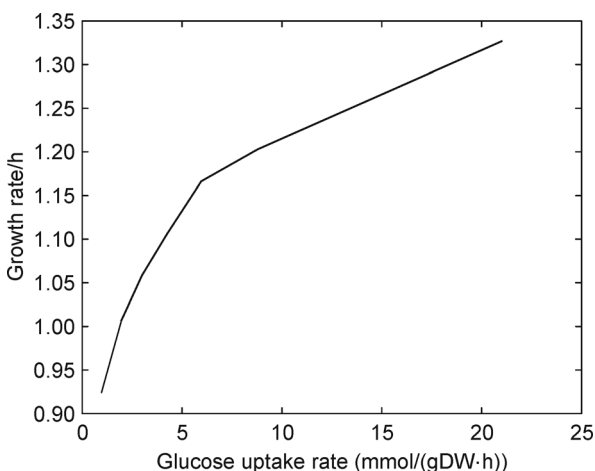


Figure 5 Robustness analysis of metabolic network of *Helicobacter pylori* with oxygen uptake rate fixed at 17 mmol/(gDW ·h).

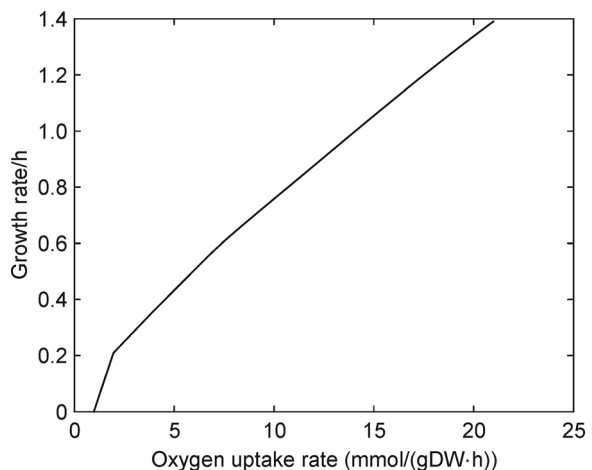


Figure 6 Robustness analysis of metabolic network of *Helicobacter pylori* with glucose uptake rate fixed at 10 mmol/(gDW ·h).

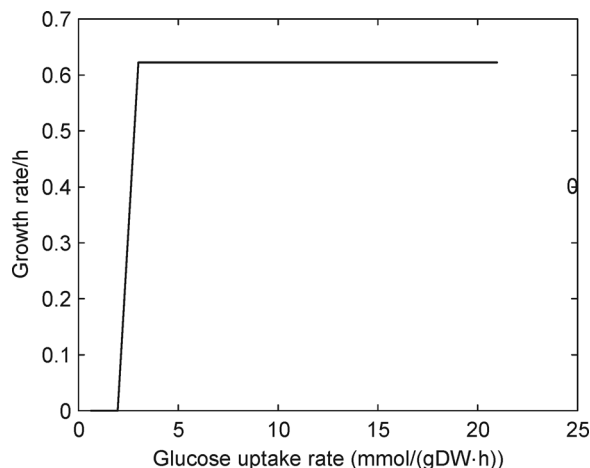


Figure 7 Robustness analysis of metabolic network of *Salmonella typhimurium* with oxygen uptake rate fixed at 17 mmol/(gDW ·h).

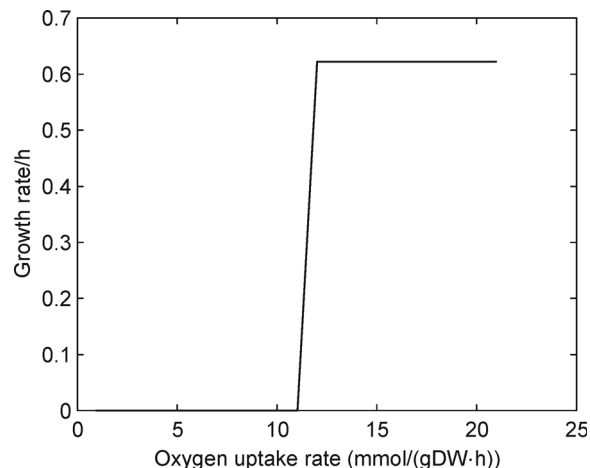


Figure 8 Robustness analysis of metabolic network of *Salmonella typhimurium* with glucose uptake rate fixed at 10 mmol/(gDW ·h).

different phases with different color coding. The 3-D graph was plotted using *surf* MATLAB function. The results for the phenotypic phase plane analysis can be visualized in the form of 3-D plot which have been shown in figures from Fig.9 to Fig.12. Figure 9 represents phenotypic phase plane study of metabolic network of *Mycobacterium tuberculosis*. Phase 1, on the far left of the plots, is characterized by 0 growth. In phase 2, growth is limited by oxygen. The line between phase 2 and phase 3 is where glucose and oxygen are perfectly balanced and growth yield is highest. In phases 3, 4, and 5, oxygen and glucose are both limiting growth. Figure 10 represents the phenotypic phase plane study of metabolic network of *Staphylococcus aureus*. Phase 1, on the far left of the plots, is characterized by negative growth. In phase 2, growth is limited by oxygen uptake rate. The line between phase 2 and phase 3 is where glucose and oxygen are perfectly balanced and growth yield is highest. In phases 3, 4, and 5, growth rate is limited by glucose uptake rate. Figure 11

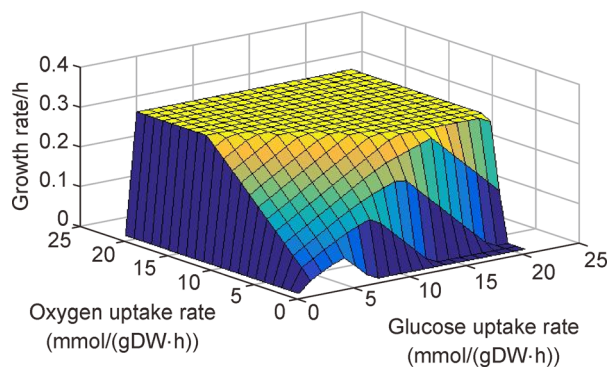


Figure 9 Phenotypic phase plane analysis of metabolic network of *Mycobacterium tuberculosis*.

represents phenotypic phase plane study of metabolic network of *Helicobacter pylori*. Phase 1, on the far left of the plots, is characterized by negative growth. In phase 2, growth is limited by oxygen uptake rate. The line between phase 2 and phase 3 is where glucose and oxygen are perfectly balanced and growth yield is highest. In phases 3, 4, and 5, growth rate is limited by glucose uptake rate. Figure 12 represents phenotypic phase plane study of metabolic network of *Salmonella typhimurium*. Phase 1, on the far left of the plots, is characterized by 0 growth. In phase 2, growth is limited by oxygen. The line between phase 2 and phase 3 is where glucose and oxygen are perfectly balanced and growth yield is highest. In phases 3, 4, and 5, oxygen and glucose are both limiting growth.

Metabolite connectivity

For each of the four networks, the number of reactions in which each possible metabolite occurs was determined. It is the measure of individual connectivity of each of the metabolite in the metabolic network. This metabolite connectivity was plotted on the log-log scale for each of the four metabolic networks. The approximate linear appearance of the curve relates to the power law distribution of the metabolic networks. The plots of four different metabolic networks show that there are very few metabolites, which are highly connected. Most of the metabolites occur only in a few reactions. The highly connected metabolites are the ‘global factors’ similar to the hub protein in the protein–protein interaction network. The least connected metabolites are the ‘local factors’ which occur in linear pathway. The power law distribution shows that the networks are scale-free. Metabolite connectivity in the metabolic network of each of the four different organisms has been shown in the figures from Fig. 13 to Fig. 16.

Reaction participation, i.e., the number of metabolites per reaction, for each of the reactions in the network was also calculated. The most common type of reaction in the metabolic network of each of the organism was the bi-linear reaction involving two substrates and two products. Average

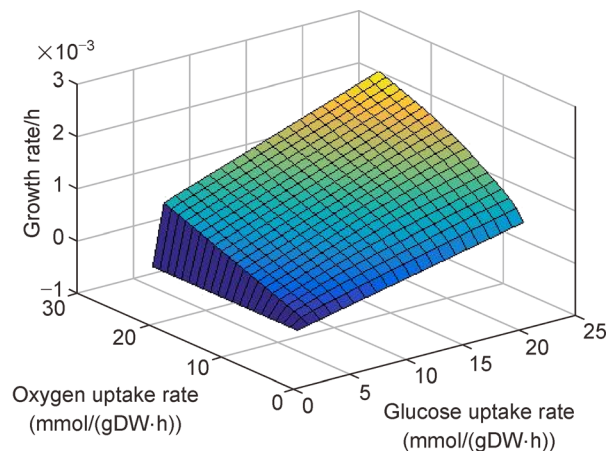


Figure 10 Phenotypic phase plane analysis of metabolic network of *Staphylococcus aureus*.

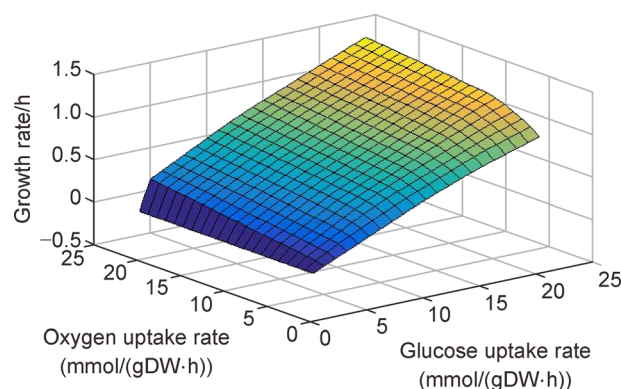


Figure 11 Phenotypic phase plane analysis of metabolic network of *Helicobacter pylori*.

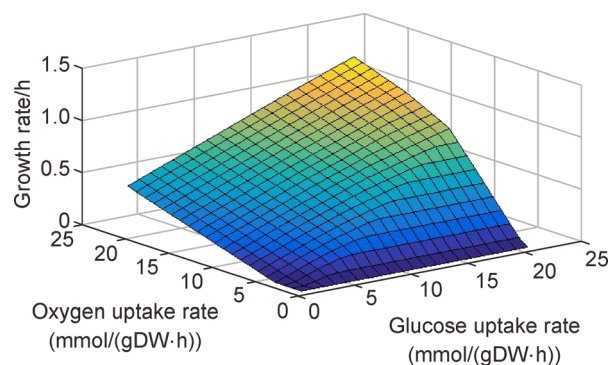


Figure 12 Phenotypic phase plane analysis of metabolic network of *Salmonella typhimurium*.

of the reaction participation in each of the network is shown in the Table 13.

Reaction essentiality and metabolite connectivity

To correlate the reaction essentiality and the metabolite connectivity in each of the metabolic networks, metabolite

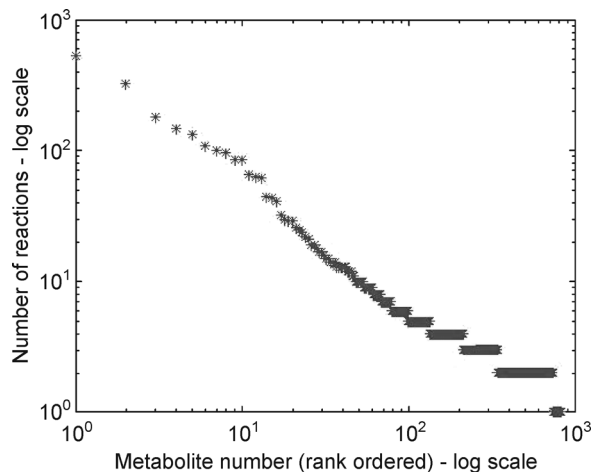


Figure 13 Metabolite connectivity in metabolic network *Mycobacterium tuberculosis*.

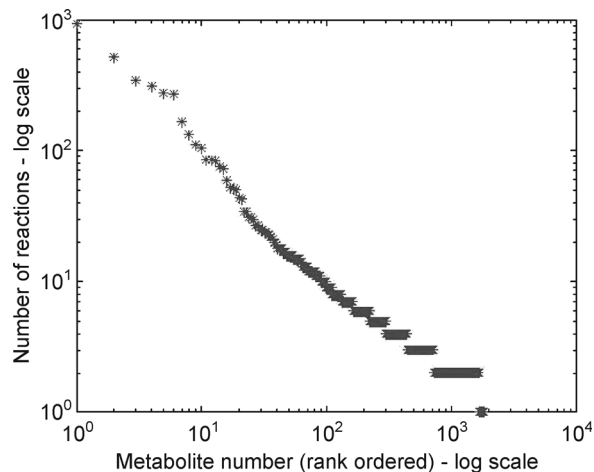


Figure 16 Metabolite connectivity in metabolic network *Salmonella typhimurium*.

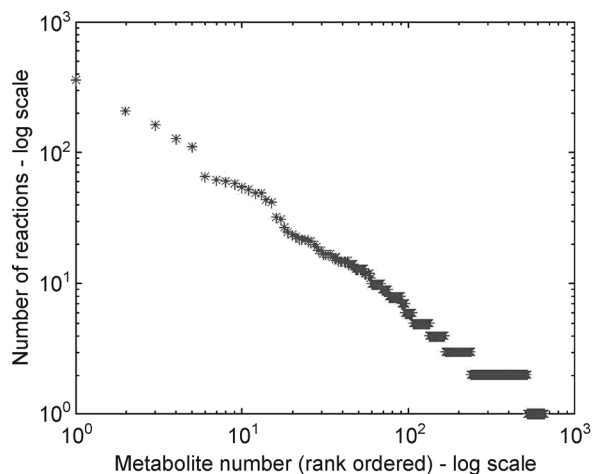


Figure 14 Metabolite connectivity in metabolic network *Staphylococcus aureus*.

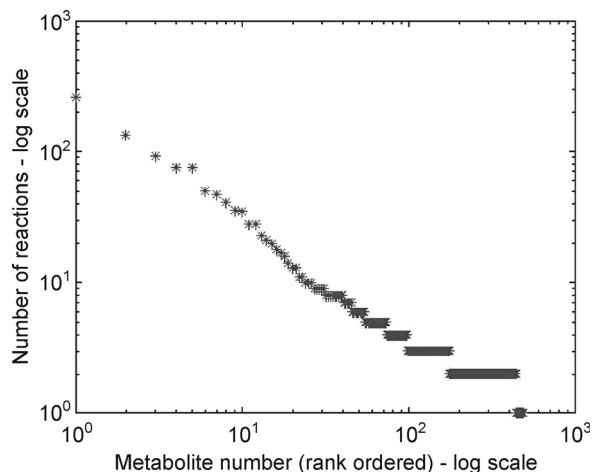


Figure 15 Metabolite connectivity in metabolic network *Helicobacter pylori*.

associated reactions were found out and finally they are deleted. After then FBA was used to predict the growth if possible. This step was performed for the metabolic networks of each of the four organisms. The results were plotted on the semi-log scale with metabolite connectivity and lethality of the network. The results for the metabolic network of four different organisms show that some less connected metabolites have a higher lethality fraction than highly connected metabolites. For the *H. pylori* core model, the average lethality fraction lies between 0.3 and 0.6 for the majority of the metabolites, regardless of their connectivity as shown in the Fig.19. For the *M. tuberculosis* core model, the average lethality fraction lies between 0.3 and 0.6 for the majority of the metabolites, regardless of their connectivity as shown in the Fig.17. For the *S. aureus* core model, the average lethality fraction lies between 0.2 and 0.6 for the majority of the metabolites, regardless of their connectivity which has been shown in the Fig.18. For the *S. typhimurium* core model, the average lethality fraction lies between 0.2 and 0.5 for the majority of the metabolites, regardless of their connectivity and is shown in the Fig.20.

Discussion

All the four metabolic network files are the reconstructed models of these organisms. Maximization of growth is very important compartmental objective to find out good FBA estimation using the kind of objective function explored. While performing the FBA analysis, growth was possible both aerobically and anaerobically for each of the four organisms. In this study, the variations in flux values were generally associated with important metabolic pathways like glycolysis, oxidative phosphorylation, and amino acid biosynthesis etc. Single gene deletion and single reaction deletion functions estimate the amount of essential genes and reactions by deleting the certain genes and reactions and

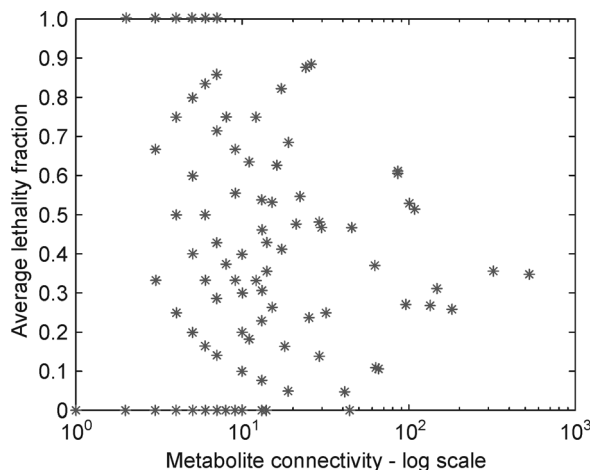


Figure 17 Correlation between metabolite connectivity and average lethality of *Mycobacterium tuberculosis*.

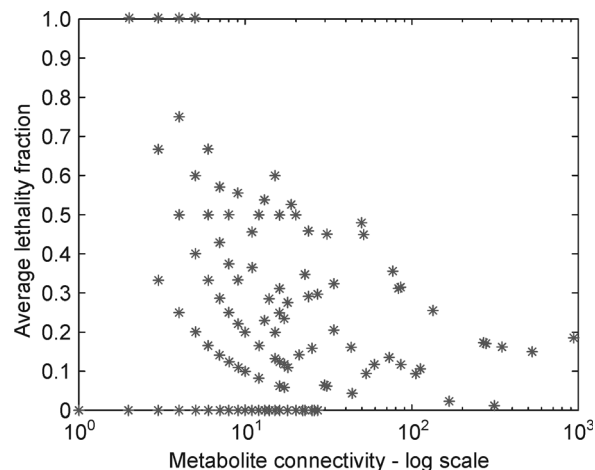


Figure 20 Correlation between metabolite connectivity and average lethality of *Salmonella typhimurium*.

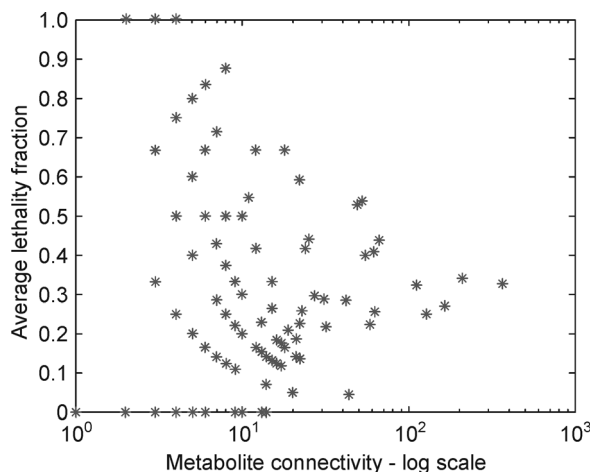


Figure 18 Correlation between metabolite connectivity and average lethality of *Staphylococcus aureus*.

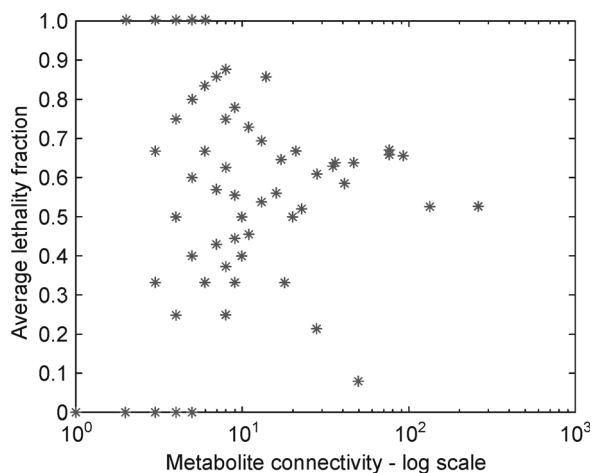


Figure 19 Correlation between metabolite connectivity and average lethality of *Helicobacter pylori*.

analyzing the impact of the deletion. These essential genes and reactions are the major players in the regulation of important pathways of the organisms. These genes and reactions are also involved in the signal transduction pathways. These genes and reactions serve as the candidate for the potential drug target against diseases caused by these organisms. The identification of synthetic lethals in organisms can be used to understand complex genetic interactions between genes and identify drug targets for combinatorial therapy. In this comparative study, it has been found that *Salmonella typhimurium*'s metabolic network contains more number of essential genes and essential reactions compared to that of other three. Similar comparison applies for the synthetic lethal genes and reactions also. Since metabolic networks of *Salmonella typhimurium* has more number of essential genes and reactions, visualization of gene deletion confirms the same. Most of the gene deletions were found to be in the lethal range having the value close to '1' on the lethality scale. Gene deletion in *Mycobacterium tuberculosis*'s and *Staphylococcus aureus*'s metabolic networks were found to be almost similar on the lethality scale. This implies that it may present more number of drug targets and hence to target the disease more effectively caused by this organism. Robustness analysis estimates up to what extent a given network is stable under the given physiological conditions. We have performed the robustness analysis by analyzing the effect of variation of the glucose and oxygen uptake rate on the growth rate of the organism. Growth rates of these two organisms were stable for most of the part of varying the glucose and oxygen uptake rate. Metabolic Metabolite connectivity enables us to understand the basic structure of these metabolic networks of four different organisms. Among the metabolic networks of all the four organisms studied, there are very few metabolites that are highly connected. Most of the metabolites are scattered. Metabolites which are highly connected serve as 'global factor' similar to the hub protein in the protein – protein interaction network. The

correlation study between metabolite connectivity and reaction essentiality gives the major insight into the lethality fraction of the metabolic networks. The lethality fraction is computed as the ratio between metabolite connectivity and reaction essentiality. The average lethality fraction of all the four organisms studied, ranged from 0.2 to 0.6.

Compliance with ethics guidelines

Kumar Gaurav and Yasha Hasija declare that they have no conflict of interest. This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Feist and Palsson (2008). The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nature Biotechnol*, 22(6): 659–667
- Alper H, Jin Y S, Moxley J F, Stephanopoulos G (2005). Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab Eng*, 7(3): 155–164
- Plaimas K, Eils R, König R. Identifying essential genes in bacterial metabolic networks with machine learning methods. *BMC systems biology*. 2010 May 3;4(1):1
- Chowdhury R, Chowdhury A, Maranas C D (2015). Using gene essentiality and synthetic lethality information to correct yeast and CHO cell genome-scale models. *Metabolites*, 5(4): 536–570
- Becker S A, Palsson B Ø (2005). Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. *BMC Microbiol*, 5(1): 8
- Schilling C H, Covert M W, Famili I, Church G M, Edwards J S, Palsson B O (2002). Genome-scale metabolic model of *Helicobacter pylori* 26695. *J Bacteriol*, 184(16): 4582–4593
- Deutscher D, Meilijson I, Kupiec M, Ruppin E (2006). Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat Genet*, 38(9): 993–998
- Edwards J S, Ramakrishna R, Palsson B O (2001). Characterizing the metabolic phenotype: a phenotype phase plane analysis. *Biotechnol Bioeng*, 77(1): 27–36
- Hamilton J J, Reed J L (2012). Identification of functional differences in metabolic networks using comparative genomics and constraint-based models. *PLoS One*, 7(4): e34670
- Heinemann M, Kummel A, Ruinatscha R, Panke S (2005). In silico genome-scale reconstruction and validation of the *Staphylococcus aureus* metabolic network. *Biotechnol Bioeng*, 92(7): 850–64
- Jamshidi N, Palsson B Ø (2007). Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. *BMC Syst Biol*, 1(1): 26
- Schellenberger J, Que R, Fleming R M, Thiele I, Orth J D, Feist A M, Zielinski D C, Bordbar A, Lewis N E, Rahmanian S, Kang J, Hyduke D R, Palsson B Ø (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc*, 6(9): 1290–1307
- Keating S M (2006). SBMLToolbox: an SBML toolbox for MATLAB users. *Bioinformatics*, 22(10): 1275–1277
- Masel J, Siegal M L (2010). Robustness: mechanisms and consequences. *Trends Genet*, 25(9): 395–403
- Raman K, Chandra N (2009). Flux balance analysis of biological systems: applications and challenges. *Brief Bioinform*, 10(4): 435–449
- Kauffman K J, Prakash P, Edwards J S (2003). Advances in flux balance analysis. *Curr Opin Biotechnol*, 14(5): 491–496
- McClelland M, Sanderson K E, Spieth J, Clifton S W, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson R K (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*, 413(6858): 852–856
- Nijman S M B (2011). Synthetic lethality: general principles, utility and detection using genetic screens in human cells. *FEBS Lett*, 585(1): 1–6
- Pratapa A, Balachandran S, Raman K (2015). Fast-SL: an efficient algorithm to identify synthetic lethal sets in metabolic networks. *Bioinformatics*, 31(20): 3299–3305
- Raghunathan A, Reed J, Shin S, Palsson B, Daefler S (2009). Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during host-pathogen interaction. *BMC Syst Biol*, 3(1): 38
- Raman K, Rajagopalan P, Chandra N (2005). Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLOS Comput Biol*, 1(5): e46
- Suthers P F, Zomorodi A, Maranas C D (2009). Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Mol Syst Biol*, 5: 301
- Thiele I, Hyduke D R, Steeb B, Fankam G, Allen D K, Bazzani S, Charusanti P, Chen F C, Fleming R M, Hsiung C A, De Keersmaecker S C, Liao Y C, Marchal K, Mo M L, Özdemir E, Raghunathan A, Reed J L, Shin S I, Sigurbjörnsdóttir S, Steinmann J, Sudarsan S, Swainston N, Thijs I M, Zengler K, Palsson B O, Adkins J N, Bumann D (2011). A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella Typhimurium* LT2. *BMC Syst Biol*, 5: 8
- Thiele I, Vo TD, Price ND, Palsson B Ø (2005). Expanded metabolic reconstruction of *Helicobacter pylori* (iIT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants. *J Bacteriol*, 187(16): 5818–5830