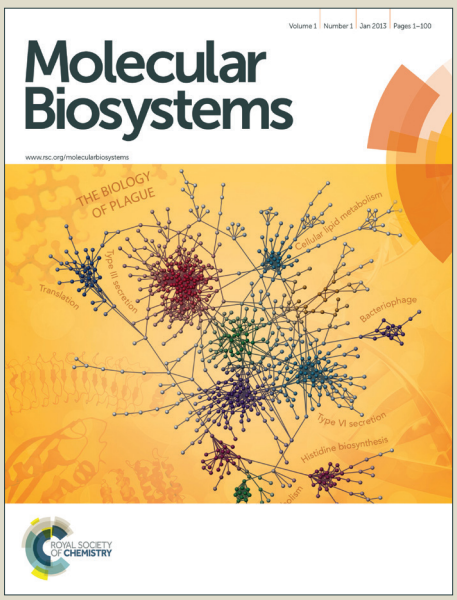
# Molecular BioSystems

Accepted Manuscript

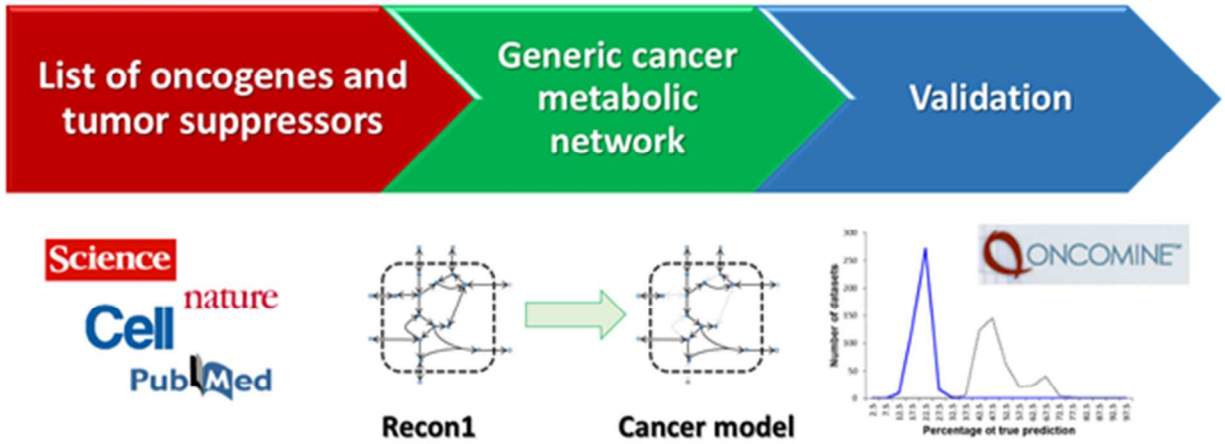




ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Graphical abstract:




Novelty of work:

We introduce a generic constraint-based model of cancer metabolism, which is able to successfully predict the metabolic phenotypes of cancer cells.

Molecular BioSystems

RSCPublishing

ARTICLE

Biomass production rate

# Reconstruction of a generic metabolic network model of cancer cells

Mahdieh Hadi[1], Sayed-Amir Marashi[1,2]

## Abstract

A promising strategy for finding new cancer drugs is to use metabolic network models to investigate the essential reactions or genes in cancer cells. In this study, we present a generic constraint-based model of cancer metabolism, which is able to successfully predict the metabolic phenotypes of cancer cells. This model is reconstructed by collecting the available data on tumor suppressor genes. Notably, we show that the activation of oncogene related reactions can be explained by the inactivation of tumor suppressor genes. We show that in a simulated growth medium similar to the body fluids, our model outperforms the previously proposed model of cancer metabolism in predicting expressed genes.



[1]Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran
[2]Corresponding author. E-mail address: Marashi@ut.ac.ir (S.-A. Marashi)

## 1. Introduction

Because of the increased incidence of cancer, it is becoming increasingly important to find ways to prevent and cure this disease. A promising strategy for finding cancer drug targets is to use metabolic networks to investigate the essential reactions or genes in cancer cells and to find cancer drugs with minimal side effects[1-4].

The Metabolism of cancer cells differs from that of healthy cells. For example, cancerous cells require more oxygen, resulting in decreased oxygen levels in their microenvironments. Consequently, angiogenesis occurs to supply the required oxygen and nutrients[5].

According to Warburg effect, while healthy cells supply 90% of their energy requirements by oxidative phosphorylation, cancer cells do so for only 50%. The remaining 50% of energy is obtained by anaerobic glycolysis. Reduction in the oxygen uptake rate and malfunctions in the citric acid cycle forces the cancer cells to get their energy by glycolysis[6]. However, even if the required oxygen is provided by angiogenesis, cancer cells still obtain their energy by glycolysis because of the problems in the citric acid cycle. In contrast, healthy cells avoid the use of glycolysis to generate energy, as this pathway produces large amounts of lactic acid, which is lethal for healthy cells. Cancer cells, on the other hand, pump out lactic acid to their microenvironment which causes damage to healthy cells[7]. In general, glycolysis is preferred when rapid cell proliferation is needed. Therefore, cancerous cells use glycolysis, to maintain the high level of glycolytic intermediates which are necessary for proliferating cells[8].

A number of metabolic differences distinguish normal and cancerous cells. Unlike normal cells, cancerous cells effectively uptake glutamine and convert it to glutamate, which is eventually used for lipid synthesis and amino acid production[9]. The activity of lactate dehydrogenase (LDH) and phosphoglycerate dehydrogenase (PHGDH) is increased in cancer cells[9, 10]. LDH converts pyruvate to lactate. The increased activity of LDH in cancer cells is required for the increased glycolytic rates and glucose uptake, as well as cancer cell growth[9]. Biosynthesis of serine and glycine from 3-phosphoglycerate is initiated by PHGDH. This pathway is increased in cancer cells and PHGDH is over-expressed in some tumors[10, 11]. Other differences between healthy and cancer cells include the increased lipid metabolism in cancer cells, which has an important role in growth and malignancy of cancer cells; and increased activity of Acyl-CoA synthetase, associated with the increased activity of this pathway[12].

There are further reports on the increased activity of other metabolic pathways in cancer cells compared to healthy cells. Some of these pathways include fatty acid metabolism [13], nucleotide metabolism[14], inositol phosphate metabolism[15-17], NAD metabolism[18], glycine, serine, and threonine metabolism[19], cholesterol metabolism[20], glycerophospholipid metabolism[21], arginine and proline metabolism[22] and eicosanoid metabolism [23]. Furthermore,

cancer cells, in general, have higher nutrient uptake rates in comparison to healthy cells[24].

A generic metabolic network of cancer cells has been previously reconstructed [4] to investigate the differences between healthy cells and cancer cells. This network will be referred to as "S2011" in this text. The reconstruction was based on the inclusion of reactions which are known to be active in the metabolism of cancer cells in RPMI-1640 medium. Then, a minimum number of reactions were added to the model so that the biomass components could be produced.

A number of issues may limit the usefulness of the above-mentioned strategy for model reconstruction. These issues arise due to the differences between the metabolisms of cancer cells in the synthetic medium vs. biological growth conditions. For example, we mentioned that cancer cells uptake large amounts of glutamine, while RPMI-1640 medium provides only limited amounts of glutamine[9]. Another example is the inconsistency between biological and predicted LDH and PHGDH activity levels. The S2011 model predicted much less LDH and PGDH activities compared to the healthy cell model, which is inconsistent with experimental data.

In the present study, we use a different approach to reconstruct the metabolic network of cancer cells. First, we collect a list of oncogenes and tumor suppressor genes. Next, we reconstruct the generic metabolic model of cancer by mapping tumor suppressor gene data on a human metabolic network model, Recon1[4, 25]. We show that our model, which will be called "B2014" throughout the text, better predicts the experimentally observed properties of cancer cells compared to S2011.

## 2. Experimental

### 2.1. Modeling cell metabolism and growth

Constraint-based modeling of metabolic networks [26] has been previously used as the modeling framework to analyze metabolic fluxes in healthy and cancer cells [2-4, 25, 27]. We used COBRA Toolbox v2.0 [28] for constraint-based modeling of metabolic networks, with GLPK as the linear programming solver (http://www.gnu.org/software/glpk). Briefly, metabolic fluxes are assumed to be under certain constraints. A stoichiometric matrix **S** is defined such that element $S_{ij}$ represents the stoichiometric coefficient of metabolite $i$ in reaction $j$. A biomass-producing reaction is often added to the network to model the biomass production rate. In flux balance analysis (FBA) [26], it is assumed that the biomass-production rate ($v_{\text{biomass}}$) is maximized, subject to the stoichiometric constraint ($\mathbf{S}\cdot\mathbf{v}=\mathbf{0}$) and capacity constraints (with the general form of $a_i \leq v_i \leq b_i$ for all reactions $i$). The stoichiometric constraint forces the system to stay in steady-state conditions, i.e., net production or consumption of metabolites is not allowed. Capacity constraints are related to maximum and minimum theoretical values known for reactions. For example, irreversible reactions always have non-negative fluxes ($0 \leq v_i$), the uptake flux of metabolites which are absent in the growth medium is zero ($v_i=0$), and the oxygen uptake rate in aerobic conditions is between zero and a maximum possible value, say $d$, which is controlled by the diffusion rate ($0 \leq v_{\text{oxygen}} \leq d$).

### 2.2. Using body fluid as the growth medium of S2011

The S2011 model [4] had been reconstructed assuming that the cells are grown in RPMI-1640 medium. Instead, we used the growth medium by which Recon1 has been evaluated (representing body fluid[25]). In this medium, input and output reactions allow unlimited production or consumption of metabolites. This is in accordance to cancer cell biology, as they induce angiogenesis to get as much nutrients as needed. The constituents of this medium are represented in supplementary file S1.

The growth medium for S2011 was changed so that both models are tested in the same medium. This allowed us to directly compare the correctness of model predictions in S2011 and B2014.

### 2.3. Gene deletion and reaction deletion studies

In addition to changing the biological or growth medium conditions of S2011, we did gene deletion and reaction deletion analysis to see if changing the growth medium affected the essentiality of any of the reactions or genes in the model.

To model the deletion of the reaction $r$ (or deletion of the gene responsible for reaction $r$), we added another constraint, $v_r = 0$, to the previous set of constraints. Then, the maximum biomass-production rate was recomputed. In our study, modeling gene deletion and reaction deletion were performed by "deleteModelGenes" and "removeRxns" functions in the COBRA Toolbox[28].

### 2.4. Comparing the behavior of S2011 in RPMI-1640 medium and body fluid

We performed single gene deletion and single reaction deletion analysis on the S2011 model in RPMI-1640 medium and body fluid growth conditions. Additionally, we investigated the effects of 811 drugs[4, 29] on the growth phenotypes predicted by the model in the two growth media. The complete list of drugs and their target genes are given in Supplementary file S2. A biomass reaction with a previously defined composition[4, 27] was added to the model. Although it is not easy to determine the objective of human cells, biomass production can be used as a proxy cellular objective, reflecting the cell requirement to replenish metabolites continuously[30]. Therefore, biomass production rate (i.e., growth rate) was considered as the objective function in FBA during the simulations.

We used a special scoring scheme for quantifying the effect of perturbation (i.e., gene deletion, reaction deletion or imposing drug inhibition) on S2011 model:

- Cytostatic score in cancer model (CS): ratio of growth rate after perturbation to the growth rate before perturbation.
- Side effect score in Recon1 (SE): ratio of growth rate after perturbation to the growth rate before perturbation.

This scoring scheme is shown in Table 1. We used these scores to computationally study the behavior of the S2011 model when grown in RPMI-1640 medium compared to the body fluid.

**2.5. Reconstruction of B2014, a generic metabolic network for cancer cells**

In the first step of the reconstruction of the cancer metabolic model, a comprehensive list of tumor suppressor genes and oncogenes were collected from four high-throughput datasets [31-35] and a number of relevant publications. The complete list of tumor suppressor genes and oncogenes are presented in Supplementary File S3 (81 tumor suppressor genes and 141 oncogenes). We compared the genes in our oncogene and tumor-suppressor set to the generic model of human metabolism, Recon1[25] , to identify oncogene-related and tumor suppressor-related reactions. Reactions which are related to tumor suppressor genes (TS-reactions) and reactions which are related to oncogenes (O-reactions) were found. Supplementary File S4 includes the complete list of TS-reactions and O-reactions (288 O-reactions, 46 TS-reactions). Next, we used F2C2 software [36] to determine the blocked reactions of Recon1. The set of blocked reactions were omitted from the list of TS-reactions and O-reactions. Finally, to obtain the final cancer cell metabolic network we used Recon1 and the complete list of TS-reactions were deactivated in Recon1. In Supplementary file S5, the B2014 model is presented in SBML format. This model includes 2766 metabolites and 3788 reactions, which represent 1905 genes. Following the reconstruction of the cancer cell metabolic network, we investigated whether model predictions of S2011 and B2014 models were consistent with biological observations.

**2.6. Investigating the accuracy of predictions by flux variability analysis**

Flux variability analysis (FVA) [37] is a method to find those reactions which have variable fluxes when growth rate has its maximum value. Before FVA, we should first find the optimal growth rate, say $v^{max}$, by FBA. Then, for FVA of a reaction $r$, we maximize and minimize the flux through this reaction, $v_r$, subject to stoichiometric and capacity constraints, and the additional constraint $v_{biomass} = v^{max}$.

The purpose of comparing the FVA results for the three models (Recon1, S2011 and B2014) is to investigate which O-reactions show increased activity in cancer cells compared to healthy cells and how many of the TS-reactions show decreased activity in cancer cells compared to healthy cells. Based on FVA, decreased activities are categorized either as "high confidence decrease" or "decrease", while increased activities are categorized either as "high confidence increase" or "increase" (Figure 1).

### 2.7. Uniform random sampling to compare the accuracy of predictions

We performed uniform random sampling analysis of the flux space for Recon1, S2011 and B2014 when biomass production rate is fixed to the maximum value (computed by FBA). The flux sampling results for O-reactions in cancer models were then compared with the fluxes of the same reactions in Recon1. One-sided Mann-Withney test was used to evaluate the statistical significance of the differences between fluxes.

### 2.8. Evaluation of the model with gene expression data

We used the uniform random sampling results of the three metabolic models to identify reactions which show increased activities in cancer cells compared to normal cells. The results are compared with the gene expression data of 428 cancer tissues reported in Oncomine database[38] (available from: www.oncomine.org). The fold change score is the $\log_2$ transformation of gene expression in cancerous tissues over gene expression in healthy tissues. Briefly, fold change scores above 0.5 were deemed as significantly increased expression, between -0.5 to 0.5 as constant expression, and below -0.5 as significantly decreased expression.

## 3. Results

### 3.1. Different behaviors of S2011 in RPMI-1640 medium and body fluid:

We first analyzed the predictions of the S2011 model in two growth media, namely, simulated body fluid and RPMI-1640. Then, we studied the effects of gene deletion, reaction deletion, and imposing drug inhibitions on the predictions of the model.

Figure 2A presents the CS scores for gene deletion analysis of genes with SE scores 1 or 2. Figure 2B and 2C show these scores for reaction deletion analysis and drug inhibition analysis, respectively. We should explain that the SE score which is presented in table1, shows how much the gene deletion, reaction deletion, or drug inhibition can affect the growth rate of healthy cells. Obviously, this effect is not desirable and we prefer those genes, reactions or drugs which have little side effect on healthy cells. Such

SE scores are either in categories 1 or 2. On the other hand, the CS score shows how much gene deletion, reaction deletion, or drug inhibition can affect the growth of cancerous cells (B2014, S2011). Clearly, reactions, genes and drugs with higher CS values are preferred, which means categories 3 or 4, so in Figure 2, CS values (and not SE values) are shown, but for those drugs, genes and reactions with SE score = 1 or 2. From these figures, one can observe that the frequency of genes, reactions or drugs in each category strongly depends on the applied medium. Therefore, predicting the phenotypes of S2011 in the simulated body fluid could potentially give rise to different results compared to the phenotypes of the simulations in RPMI-1640 medium. On the other hand, it was observed that some genes which are deleted from S2011 (Figure 2A) have a cytostatic score of 1 in the simulated body fluid, but when the model is analyzed in RPMI-1640 these genes have a cytostatic score of 2 or higher. Comparable trends were observed in the case of reaction deletion and drug inhibition analysis. In general, one may conclude by network perturbation that the cytostatic score is generally greater when simulations are performed in nutrition-limited medium (RPMI-1640) compared to the rich medium (body fluid). Presumably, this observation is due to the fact that some uptake reactions are inactive in RPMI-1640. Consequently, small perturbations in cell metabolism may unrealistically stop growth in the cell model, as the alternative reactions/pathways, which can normally compensate for the perturbations, are inactive in the model. For this reason, we suggest that it is critical to simulate cancer metabolism in body fluid.

### 3.2. FVA of O-reactions

Figure 3 shows that the number of O-reactions which show elevated flux values in B2014 is considerably higher than this number in S2011. This is a notable result, because B2014 was constructed by simple inactivation of TS-reactions in Recon1. This means that the procedure of inactivating TS-reactions (without imposing further constraints on O-reactions) can result in a redistribution of metabolic fluxes such that O-reactions carry higher flux values. This observation suggests that the activation of oncogene-related reactions may be explained by inactivation of tumor suppressor genes.

Figure 3 shows that about 65.6% of O-reactions in S2011 can carry no flux. In contrast, this value drops to only 12.5% in B2014. On the other hand, more than 52% of O-reactions show increased activity in B2014 compared to Recon1, while this value is only about 10% in S2011. Moreover, we observe decreased activity (i.e., false prediction) for about 3.8% of reactions in S2011, while this value drops to 1.7% in B2014. Altogether, in comparison with S2011, we show that B2014 better predicts the activation of O-reactions. Supplementary file S6 presents the list of O-reactions in addition to the results of FVA of these reactions in Recon1, S2011 and B2014. From these results, one can observe that those reactions and enzymes that are mentioned in the Introduction of the present article (e.g., lactate dehydrogenase, phosphoglycerate dehydrogenase and acyl-CoA synthetase) show higher activities in B2014 in comparison to S2011. Supplementary file S6 also lists the subsystems or pathways in which these

reactions are categorized. These pathways are known to be more active in cancer cells, including include fatty acid metabolism [13], nucleotide metabolism[14], inositol phosphate metabolism[15-17], NAD metabolism[18], glycine, serine, and threonine metabolism[19], cholesterol metabolism[20], glycerophospholipid metabolism[21], arginine and proline metabolism[22] and eicosanoid metabolism [23] and transporter reactions.

### 3.3. FVA of TS-reactions

In Figure 4, the behaviors of TS-reactions in S2011 and B2014 are compared. Decreased activity of TS-reactions is expected in cancerous models in relation to Recon1. Since TS- reactions are deactivated in B2014 by design, these reactions may only belong to two categories: either blocked in both healthy and cancer cell models, or showing high-confidence decrease due to the inactivation in the cancer model, but not the healthy model. On the other hand, in S2011 about 9% of TS-reactions show increased activity (i.e., false prediction), while about 17.5% are unchanged or uncategorized. In conclusion, B2014 outperforms S2011 in predicting both activated and inactivated reactions in cancerous cells.

### 3.4. Increase and decrease of fluxes based on uniform random sampling

We used uniform random sampling of the flux space for analyzing the fluxes of the O- and TS-reactions in healthy cell model, as well as S2011 and B2014 models. The results of this analysis are summarized in Figure 5. Similar to FVA results, one can observe that the predictions of B2014 are better than S2011. About 56.6% of O-reactions in Figure 5A show increased activity in B2014, while this value is only about 13.9% in S2011. On the other hand, about 36.5 % of O-reactions in S2011 show decreased activity, which represents false predictions. This value drops to 25.35% in B2014. In Figure 5B, it can be seen that about 9% of TS-reactions in S2011 incorrectly show increased activity. In contrast, as a result of the reconstruction approach used in B2014, we see that no false increased activity can occur in B2014. The O-reactions which show higher activities in B2014 in comparison to S2011 are listed in Supplementary file S7.The subsystems and pathways in which these reactions are active are listed in supplementary file S7. These subsystems and pathways are in accordance with the above-mentioned pathways. The inositol phosphate metabolism[15-17] is one of the pathways showing increased flux in cancer cells compared to healthy cells. Therefore, the reactions of this pathway are expected to be more active in the cancer model compared to Recon1. Some of the reactions of this pathway that are O-reactions are highlighted in the Figure 6. These reactions have no flux in S2011 but they are active in B2014, which is in agreement with the biological data.

### 3.5. Comparing the flux sampling results with gene expression data

To evaluate the usefulness of our model in predicting gene expression data in cancer cells, we compared our flux sampling results with gene expression data from Oncomine database[38]. Figure 7 shows the frequency of true predictions in each model based on the gene expression

dataset. The frequency of true predictions of B2014 is significantly greater than S2011 ($p$-value= $1.8\times10^{-141}$). We also compared the predictions of cancer models (S2011 and B2014) independently with different cancer gene expression data (Supplementary file S8). The results clearly show that compared to S2011, predictions of B2014 are much more consistent with gene expression of different cancer tissues ($p$-value< 0.05 in all cases). While the true predictions of B2014 are generally highly significant (especially in case of breast, colon, brain and lung cancers), in a few cancer subtypes such as adrenal and uterus cancers the significance level is not as high. We hope that further improvement of the model improves the prediction of these cancer subtypes.

## 4. Discussion

In this study, we present a generic metabolic network model of cancer cells. Based on our findings, we suggested that simulated body fluid should be used instead of RPMI-1640 for metabolic modeling to obtain biomedically relevant results. We propose that focusing on the core metabolism of cancer cells may result in overlooking alternative reactions that can potentially be activated. Therefore, it is not surprising to see that S2011, when grown in RPMI-1640, leads to potential drug candidates that target the core metabolism[4]. This may explain the in vivo failure of some drugs, which can successfully inhibit cancer cells in vitro.

On the other hand, in this paper, we present a new strategy to reconstruct the generic metabolic networks of cancer. The resulting cancer model, B2014, is essentially a subnetwork of the healthy cell model, Recon1. In this network, reactions which are related to tumor suppressor genes are deactivated. Strikingly, we observed that simple inactivation of these genes led to the increase of flux values through the oncogene-related reactions. This observation may shed light on a new mechanism of gene regulation in cancer.

By using flux variability analysis and uniform random sampling of the flux space, we compared the reaction fluxes in cancer models vs. healthy model. We showed that our cancer model (B2014) significantly outperforms the previous cancer model (S2011) in predicting activity of reactions and gene expression data. In addition to gene expression profile, most of the pathways which should be more active in cancer cells in comparison to healthy cells (including include fatty acid metabolism [13], nucleotide metabolism[14], inositol phosphate metabolism[15-17], NAD metabolism[18], glycine, serine, and threonine metabolism[19], cholesterol metabolism[20], glycerophospholipid metabolism[21], arginine and proline metabolism[22] and eicosanoid metabolism [23]) are far better predicted by B2014 in comparison to S2011. This improvement is presumably related, at least partly, to the high-quality and complete generic model of human cells, Recon1, which was used as the starting point for our model reconstruction. In conclusion, we believe that our model has the potential to be used for suggesting novel cancer drugs.

## Acknowledgements

We would like to acknowledge the financial support of University of Tehran for this research under grant number

28791/1/2. We would also like to thank R. S. Mirhassani (University of Tehran) and R. Behrouzi (Harvard Medical School) for editing the English text of this manuscript.

† Electronic supplementary information (ESI) available:

Supplementary file S1: The maximum allowable uptake flux of the components which are used in the model as "body fluid" medium. The first column shows the metabolites name and the second one is Maximum Uptake which its unit is mmol / gDW / hour.

Supplementary file S2: The complete list of drugs and their target genes which are used in the present study to investigate the effect of changing the medium composition on the drugs effect. The first column includes the drug IDs which are extracted from DrugBank database, the second column is the generic name of these drugs, the third column represents the Entrez IDs of metabolic targets of the drugs, and the fourth column shows whether the drug is an anticancer drug or not.

Supplementary file S3: The complete list of tumor suppressor genes and oncogenes are presented in two sheets. First sheet includes tumor suppressor genes, and the second sheet includes oncogenes. The first column of each sheet shows the gene IDs in the model, while the second column represents the Entrez IDs of the genes.

Supplementary file S4: The complete list of TS-reactions which have been deactivated to reconstruct B2014 and O-reactions which have been used to compare the percent of correct predictions between S2011 and B2014. The first column specifies if the reaction is O-reaction or TS-reaction, the second column shows the reaction numbers in the models, the third column shows the reaction names, and the fourth column represents the references in which these reactions are introduced.

Supplementary file S5: This file is the cancer metabolic network model (B2014). The metabolic network is in the standard SBML format.

Supplementary file S6: The list of O-reactions in B2014 which shows true predictions based on FVA results. The first column shows the reaction numbers in the models, the second column shows the reaction IDs, the third column represents the complete reaction names, the fourth column is the subsystem or pathway to which each reaction belongs to. The fifth and the sixth columns show the minimunm and maximum fluxes of each of these reactions in B2014 obtained by flux variability analysis (FVA). Similarly, the seventh and the eighth columns show the minimunm and maximum fluxes of S2011, while the ninth and the tenth columns show the minimunm and maximum fluxes of Recon1.

Supplementary file S7: The list of O-reactions in B2014 which shows true prediction based on uniform random sampling results.

Supplementary file S8: Accuracy of the predictions of S2011 and B2014 based on cancer gene expression. Each graph shows, for a certain type of cancer, the percentage of true predictions of S2011 and B2014. The p-values represent the significance level of the comparison of the two distributions using one-sided Mann-Whitney U test.

## References:

1. O. Resendis-Antonio, C. Gonzalez-Torres, G. Jaime-Munoz, C. E. Hernandez-Patino and C. F. Salgado-Munoz, *Seminars in cancer biology*, 2014.
2. A. Mardinoglu, F. Gatto and J. Nielsen, *Biotechnology journal*, 2013, **8**, 985-996.
3. N. E. Lewis and A. M. Abdel-Haleem, *Frontiers in physiology*, 2013, **4**, 237.

4. O. Folger, L. Jerby, C. Frezza, E. Gottlieb, E. Ruppin and T. Shlomi, *Molecular systems biology*, 2011, **7**, 501.
5. V. Gogvadze, S. Orrenius and B. Zhivotovsky, *Trends in cell biology*, 2008, **18**, 165-173.
6. S. M. Jeon, N. S. Chandel and N. Hay, *Nature*, 2012, **485**, 661-665.
7. R. A. Gatenby and R. J. Gillies, *Nature reviews. Cancer*, 2004, **4**, 891-899.
8. S. Y. Lunt and M. G. Vander Heiden, *Annual review of cell and developmental biology*, 2011, **27**, 441-464.
9. N. Hammoudi, K. B. Ahmed, C. Garcia-Prieto and P. Huang, *Chinese journal of cancer*, 2011, **30**, 508-525.
10. K. Hiller and C. M. Metallo, *Current opinion in biotechnology*, 2013, **24**, 60-68.
11. J. W. Locasale, A. R. Grassian, T. Melman, C. A. Lyssiotis, K. R. Mattaini, A. J. Bass, G. Heffron, C. M. Metallo, T. Muranen, H. Sharfi, A. T. Sasaki, D. Anastasiou, E. Mullarky, N. I. Vokes, M. Sasaki, R. Beroukhim, G. Stephanopoulos, A. H. Ligon, M. Meyerson, A. L. Richardson, L. Chin, G. Wagner, J. M. Asara, J. S. Brugge, L. C. Cantley and M. G. Vander Heiden, *Nature genetics*, 2011, **43**, 869-874.
12. T. Mashima, S. Sato, S. Okabe, S. Miyata, M. Matsuura, Y. Sugimoto, T. Tsuruo and H. Seimiya, *Cancer science*, 2009, **100**, 1556-1562.
13. F. P. Kuhajda, *Cancer research*, 2006, **66**, 5977-5980.
14. Y. C. Liu, F. Li, J. Handler, C. R. Huang, Y. Xiang, N. Neretti, J. M. Sedivy, K. I. Zeller and C. V. Dang, *PloS one*, 2008, **3**, e2722.
15. E. Lee and S. H. Yuspa, *Carcinogenesis*, 1991, **12**, 1651-1658.
16. T. W. Miller, B. N. Rexer, J. T. Garrett and C. L. Arteaga, *Breast cancer research : BCR*, 2011, **13**, 224.
17. G. Schramm, E. M. Surmann, S. Wiesberg, M. Oswald, G. Reinelt, R. Eils and R. Konig, *BMC medical genomics*, 2010, **3**, 39.
18. C. C. Chini, A. M. Guerrico, V. Nin, J. Camacho-Pereira, C. Escande, M. T. Barbosa and E. N. Chini, *Clinical cancer research : an official journal of the American Association for Cancer Research*, 2014, **20**, 120-130.
19. J. W. Locasale, *Nature reviews. Cancer*, 2013, **13**, 572-583.
20. C. Munoz-Pinedo, N. El Mjiyad and J. E. Ricci, *Cell death & disease*, 2012, **3**, e248.
21. V. Dolce, A. R. Cappello, R. Lappano and M. Maggiolini, *Current molecular pharmacology*, 2011, **4**, 167-175.
22. B. Perroud, J. Lee, N. Valkova, A. Dhirapong, P. Y. Lin, O. Fiehn, D. Kultz and R. H. Weiss, *Molecular cancer*, 2006, **5**, 64.
23. D. Wang and R. N. Dubois, *Nature reviews. Cancer*, 2010, **10**, 181-193.
24. R. G. Jones and C. B. Thompson, *Genes & development*, 2009, **23**, 537-548.
25. N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas and B. O. Palsson, *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**, 1777-1782.
26. J. D. Orth, I. Thiele and B. O. Palsson, *Nature biotechnology*, 2010, **28**, 245-248.
27. C. Frezza, L. Zheng, O. Folger, K. N. Rajagopalan, E. D. MacKenzie, L. Jerby, M. Micaroni, B. Chaneton, J. Adam, A. Hedley, G. Kalna, I. P. Tomlinson, P. J. Pollard, D. G. Watson, R. J. Deberardinis, T. Shlomi, E. Ruppin and E. Gottlieb, *Nature*, 2011, **477**, 225-228.
28. J. Schellenberger, R. Que, R. M. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke and B. O. Palsson, *Nature protocols*, 2011, **6**, 1290-1307.
29. D. S. Wishart, C. Knox, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B.

Gautam and M. Hassanali, *Nucleic acids research*, 2008, **36**, D901-906.

30. A. Wagner, R. Zarecki, L. Reshef, C. Gochev, R. Sorek, U. Gophna and E. Ruppin, *Proceedings of the National Academy of Sciences of the United States of America*, 2013, **110**, 19166-19171.
31. M. Zhao, J. Sun and Z. Zhao, *Nucleic acids research*, 2013, **41**, D970-976.
32. T. Santarius, J. Shipley, D. Brewer, M. R. Stratton and C. S. Cooper, *Nature reviews. Cancer*, 2010, **10**, 59-64.
33. R. Possemato, K. M. Marks, Y. D. Shaul, M. E. Pacold, D. Kim, K. Birsoy, S. Sethumadhavan, H. K. Woo, H. G. Jang, A. K. Jha, W. W. Chen, F. G. Barrett, N. Stransky, Z. Y. Tsun, G. S. Cowley, J. Barretina, N. Y. Kalaany, P. P. Hsu, K. Ottina, A. M. Chan, B. Yuan, L. A. Garraway, D. E. Root, M. Mino-Kenudson, E. F. Brachtel, E. M. Driggers and D. M. Sabatini, *Nature*, 2011, **476**, 346-350.
34. M. E. Higgins, M. Claremont, J. E. Major, C. Sander and A. E. Lash, *Nucleic acids research*, 2007, **35**, D721-726.
35. K. Xu, J. Cui, V. Olman, Q. Yang, D. Puett and Y. Xu, *PloS one*, 2010, **5**, e13696.
36. A. Larhlimi, L. David, J. Selbig and A. Bockmayr, *BMC bioinformatics*, 2012, **13**, 57.
37. J. L. Reed and B. O. Palsson, *Genome research*, 2004, **14**, 1797-1805.
38. D. R. Rhodes, J. Yu, K. Shanker, N. Deshpande, R. Varambally, D. Ghosh, T. Barrette, A. Pandey and A. M. Chinnaiyan, *Neoplasia*, 2004, **6**, 1-6.
39. M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi and M. Tanabe, *Nucleic acids research*, 2014, **42**, D199-205.
40. M. Kanehisa and S. Goto, *Nucleic acids research*, 2000, **28**, 27-30.

**Figure legends:**

**Figure 1**. The four reaction categories which show flux decrease or increase in cancer models in comparison with Recon1. The solid arrows show the flux range in cancer models while dashed arrows show the flux range in Recon1.

**Figure 2**. (A) Gene deletion analysis of the S2011 model in two different growth media (RPMI-1640 and body fluid) when side effect score is 1 or 2, i.e., little side effect is observed in healthy cells; (B) Reaction deletion analysis of the S2011 model in the two different growth media when side effect score is 1 or 2; (C) Drugs inhibition analysis in S2011 in two different media when side effect score is 1 or 2.

**Figure 3**. Number of O-reactions which show increased/decreased fluxes in cancer models (S2011 and B2014) compared to their fluxes in the healthy cell model (Recon1) using FVA.

**Figure 4**. Number of TS-reactions which show increased/decreased fluxes in cancer models (S2011 and B2014) compared to their fluxes in the healthy cell model (Recon1) using FVA. It should be emphasized here that our reconstruction approach is based on the deactivation of TS-reactions, and therefore, it is obvious that B2014 must be fully consistent with the biological data.

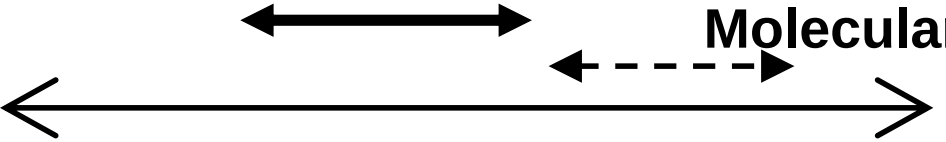
**Figure 5**. (A) Number of O-reactions which show increased/decreased fluxes in cancer models (S2011 and B2014) compared to their fluxes in the healthy cell model (Recon1) using uniform random sampling of the flux space; (B) Number of TS-reactions which show increased/decreased fluxes in cancer models (S2011 and B2014) compared to their fluxes in the healthy cell model (Recon1) using uniform random sampling of the flux space. It should be noted here that because of our reconstruction strategy (by deactivation of TS-reactions) no false prediction can occur in case of B2014.

**Figure 6.** Inositol phosphate metabolism pathway shows increased activity in cancer cells compared to healthy cells. Highlighted reactions are those reactions for which increased activity is expected, but S2011 (in contrast to B2014) fails to predict this behavior. This figure is taken from KEGG database[39, 40].

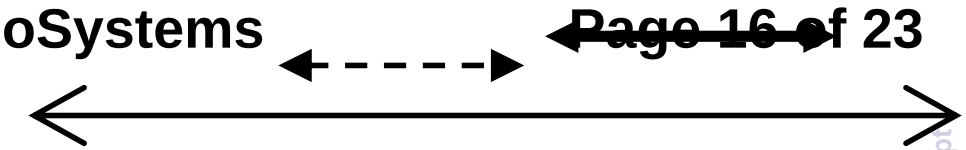
**Figure 7**. Accuracy of prediction of reactions with increased flux values in S2011 and B2014 based on different gene expression datasets.

**Table 1**. Definition of different categories. Here, x represents either cytostatic score (CS) or side effect score (SE). CS is a score that show how much a perturbation (reaction deletion, gene deletion or drug inhibition) decreases the growth rate of cancer cells. SE score means how much such a perturbation affects the growth rate of healthy cells. Therefore, SE scores in categories 1 or 2 and CS scores in categories 3 or 4 are preferred.

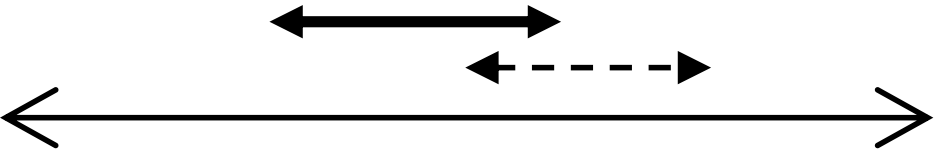
| Category number | Range of $x$ |
|---|---|
| 1 | $x \geq 0.9$ |
| 2 | $0.5 \leq x < 0.9$ |
| 3 | $0.05 \leq x < 0.5$ |
| 4 | $0 \leq x < 0.05$ |



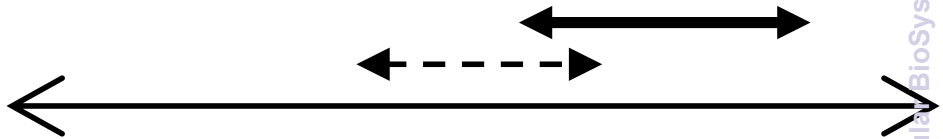
High confidence decrease



High confidence increase

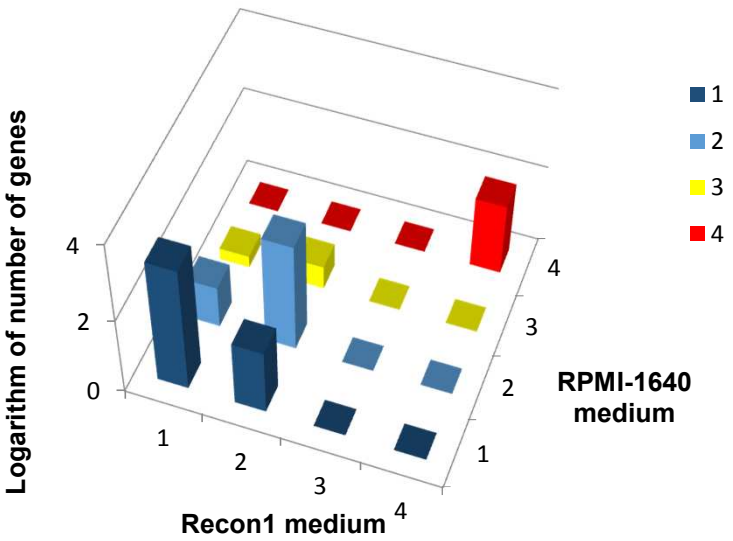


Decrease

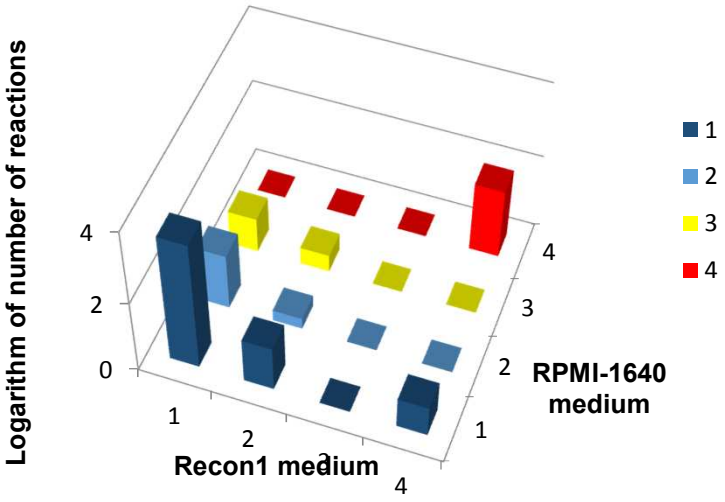


Increase

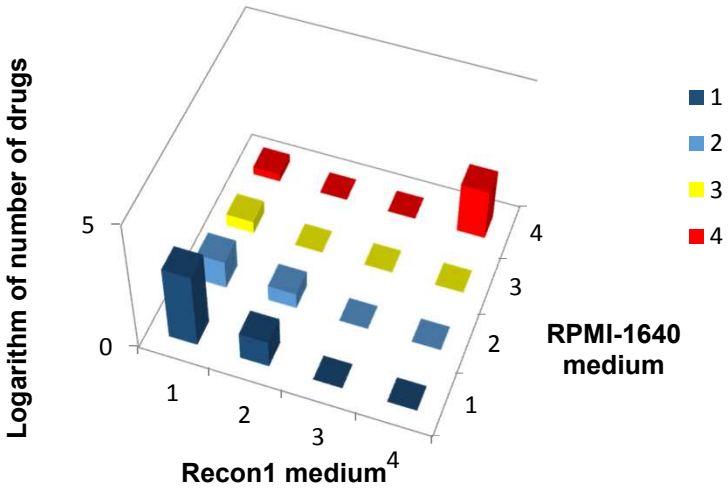
(A)

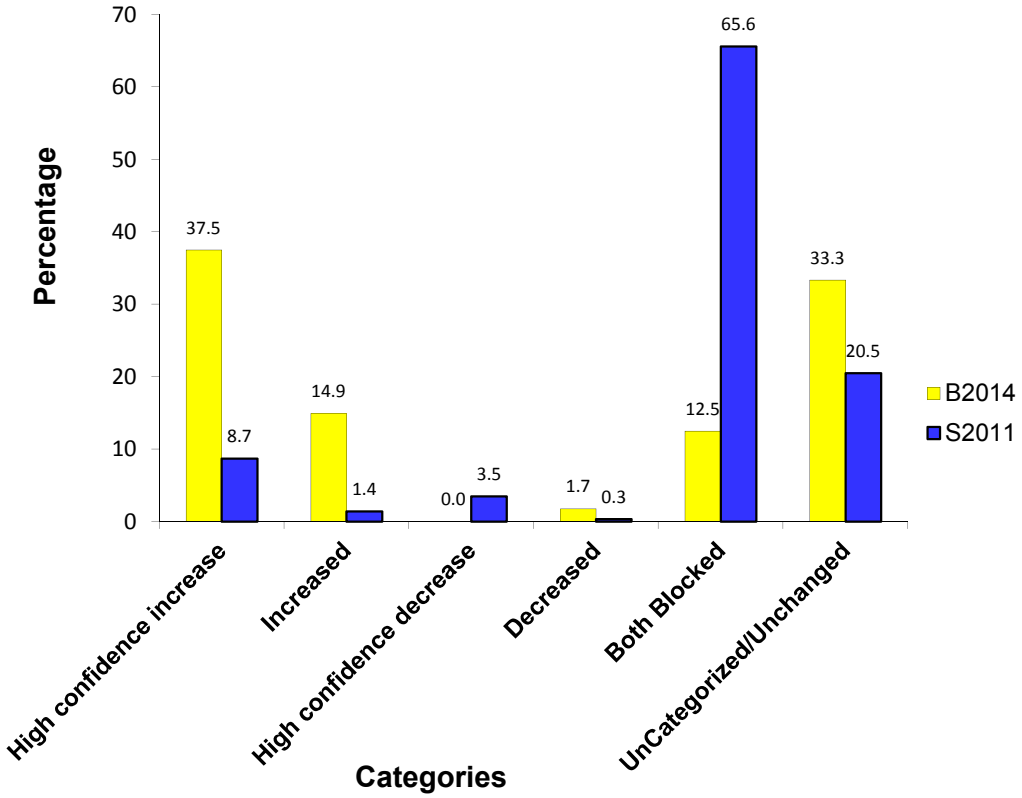


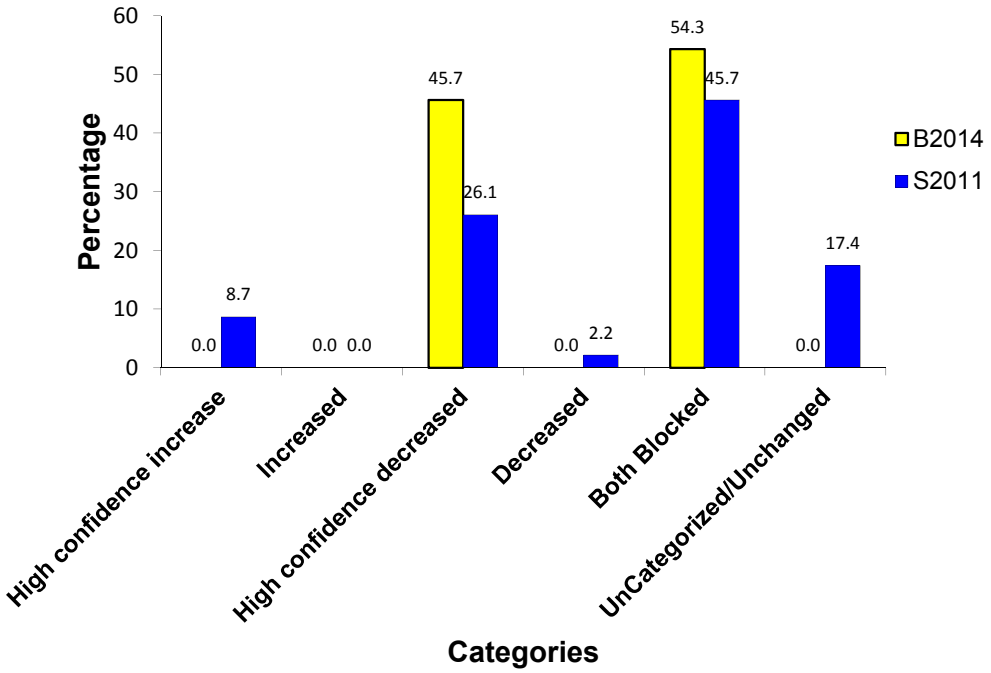

(B)

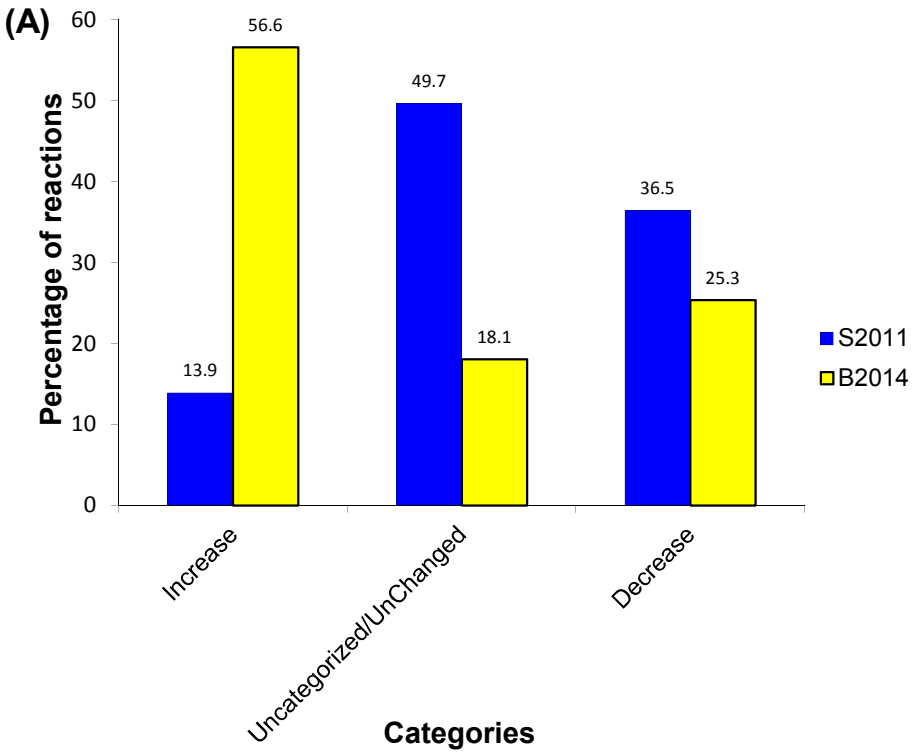


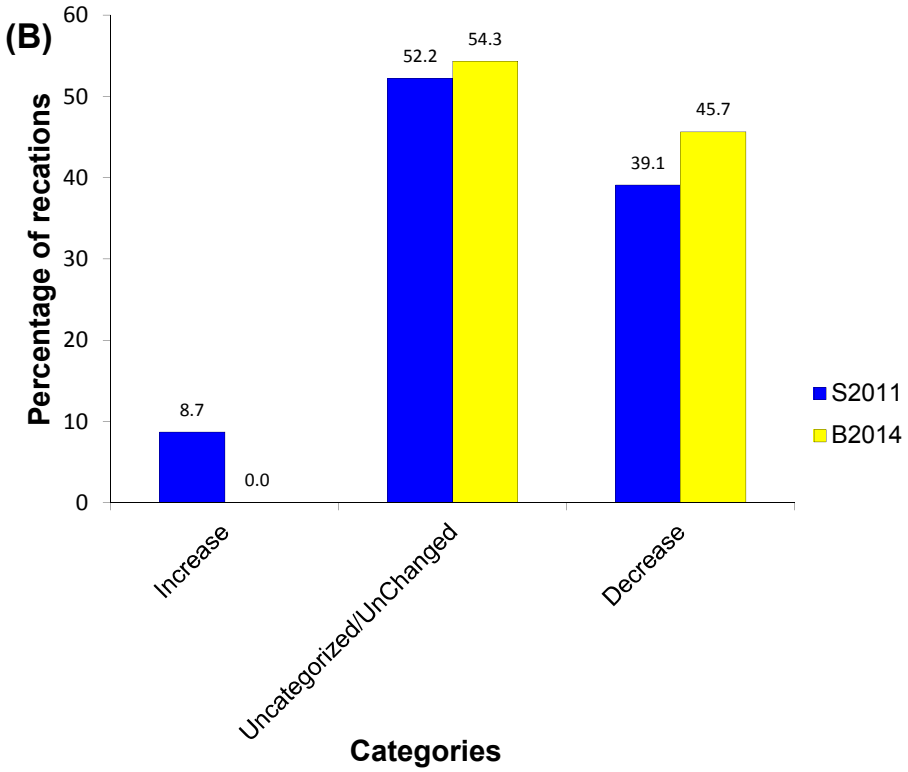

(C)

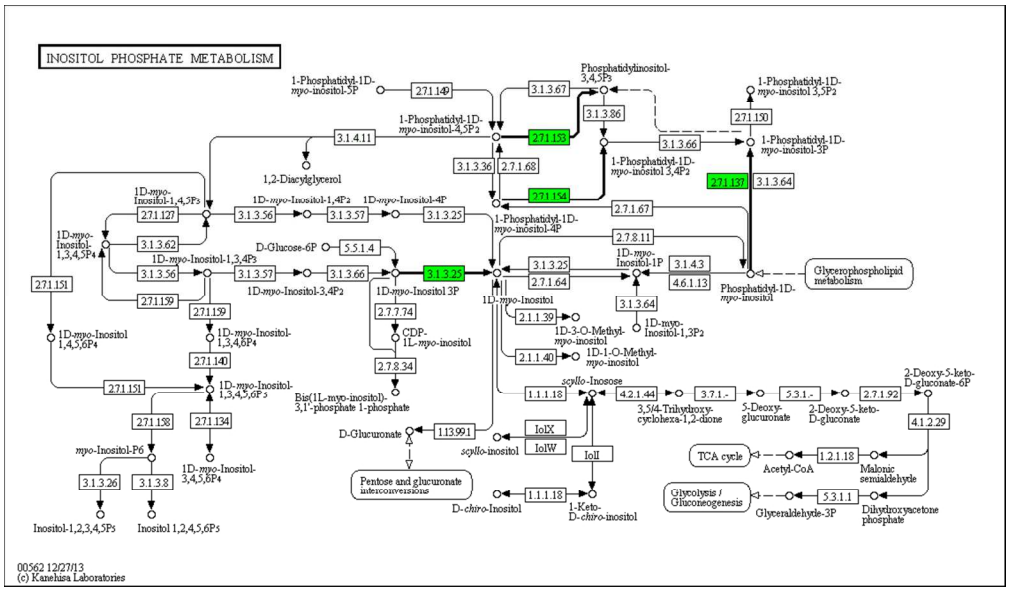











400x232mm (72 x 72 DPI)