



OPEN

Exploring gene knockout strategies to identify potential drug targets using genome-scale metabolic models

Abhijit Paul¹, Rajat Anand¹, Sonali Porey Karmakar¹, Surender Rawat¹, Nandadulal Bairagi² & Samrat Chatterjee¹✉

Research on new cancer drugs is performed either through gene knockout studies or phenotypic screening of drugs in cancer cell-lines. Both of these approaches are costly and time-consuming. Computational framework, e.g., genome-scale metabolic models (GSMMs), could be a good alternative to find potential drug targets. The present study aims to investigate the applicability of gene knockout strategies to be used as the finding of drug targets using GSMMs. We performed single-gene knockout studies on existing GSMMs of the NCI-60 cell-lines obtained from 9 tissue types. The metabolic genes responsible for the growth of cancerous cells were identified and then ranked based on their cellular growth reduction. The possible growth reduction mechanisms, which matches with the gene knockout results, were described. Gene ranking was used to identify potential drug targets, which reduce the growth rate of cancer cells but not of the normal cells. The gene ranking results were also compared with existing shRNA screening data. The rank-correlation results for most of the cell-lines were not satisfactory for a single-gene knockout, but it played a significant role in deciding the activity of drug against cell proliferation, whereas multiple gene knockout analysis gave better correlation results. We validated our theoretical results experimentally and showed that the drugs mitotane and myxothiazol can inhibit the growth of at least four cell-lines of NCI-60 database.

Chronic progressive diseases such as cancer do not derive from abnormal functioning of a single gene or a single pathway but reflect the disorder of complex intracellular and intercellular networks that link tissues and organs¹. Various tissues and organs like breast, central nervous system (CNS), colon, lung, ovary, prostate, renal are affected by cancer. Millions of people are suffering from cancer and the number is increasing²⁻⁴. Available drugs that are used for treating cancer, unfortunately, have many side effects⁵. Therefore, there is an increasing demand for new therapies with better therapeutic windows, implying that the drug will target a particular cell type (such as tumour cells) with no or minimum negative effects on healthy cells. The search for such suitable therapeutic windows is an important and challenging problem in the case of cancer. Another problem with the existing cancer drugs is that a particular drug shows different responses when applied to different individuals. This is because the effects of a drug on a patient depend not only on the interaction with its targets but also on the activities of many other enzymes which form a complex network of metabolic reactions in which the products of a reaction become the substrates of other reactions⁶. This leads to the emerging field of personalized medicine and personalized drug-choice⁷. Thus there is a need for developing new anti-cancer drugs taking care of the above problems and demands for an in-depth mechanistic understanding of cancer^{8,9}.

Gene knockout study in cancer cell-lines is used to see the effect of an existing cancer drug or to develop new cancer drug¹⁰⁻¹⁴. Another approach is the phenotypic screening of drugs in cancer cell-lines to find its effect on the cell growth^{10,15-18}. Both of these processes are very costly and time consuming¹⁹⁻²¹. Therefore, a computational method, like metabolic networks, could be a good alternative to find drugs having better selective ability in killing cancerous cells²²⁻²⁴. One tool which is particularly suitable to deal with problems like personalized medicine and finding a larger therapeutic window is genome-scale metabolic models (GSMMs)⁶. GSMM is the reconstructions of metabolic networks that connect gene-protein-metabolic reactions^{25,26}. Different reactions

¹Complex Analysis Group, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, 3rd milestone, Faridabad-Gurgaon Expressway, Faridabad 121001, India. ²Centre for Mathematical Biology and Ecology, Department of Mathematics, Jadavpur University, Kolkata 700032, India. ✉email: samrat.chatterjee@thsti.res.in

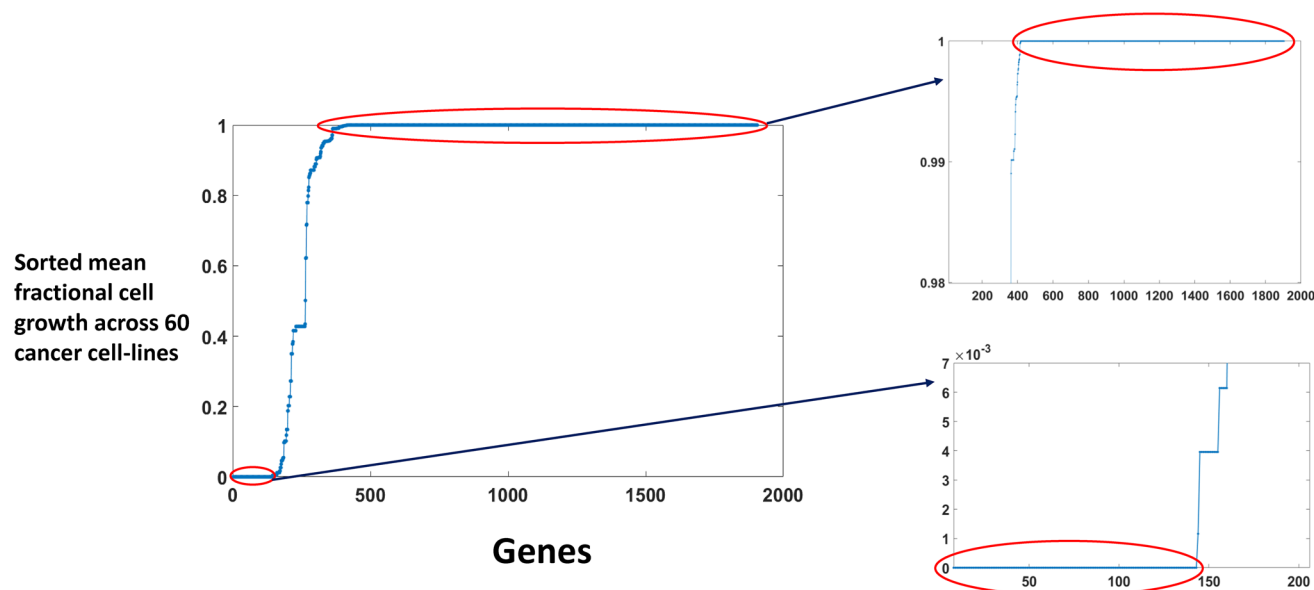


Figure 1. Gene knockout simulation result. Mean value of the fractional cell growth (FCG) across 60 cancer metabolic models for each individual of 1905 metabolic genes. The right lower panel represents the 143 genes which give very low growth rate after knocking out across all 60 cancer models (mean value $< 10^{-6}$ & s.d. $< 1.2874 \times 10^{-6}$) and the upper panel represents the 1488 genes which show no change in the growth rate after knocking out across all 60 cancer models (mean value > 0.99995 & s.d. $< 3.2838 \times 10^{-6}$).

containing stoichiometric coefficients are used to build a stoichiometric matrix to obtain quantitative information on how different metabolites are linked to each reaction in the network. Through flux balance analysis (FBA) on these GSMMs, one can evaluate the metabolic capabilities of the cell, e.g., the capability of synthesizing biomass building blocks. GSMMs have successfully been used in cancer drug development^{27–31}. For example, Folger et al.²⁷ made a generic metabolic model of cancer and predicted some growth-supporting genes for cancer, which were validated with experimental shRNA data. GSMMs were also used to predict the putative effects of drugs of DrugBank database⁶, but the results were not compared with the experimental data.

The present study is an attempt to explore gene knockout strategies that apply GSMMs in finding possible targets and mechanisms related to cancer disease. For the analysis, we considered the GSMMs of the NCI-60 cell-lines built by Yizhak et al.²⁹ using Personalized Reconstruction of Metabolic models (PRIME) approach. They established their cell-specific model based on molecular and phenotypic data. It is shown that the models can find drug targets that inhibit the proliferation of specific cell-lines. Their models can also infer on the prognosis of breast and lung cancer. In this study, we applied their model for a more comprehensive study on the single and multiple gene knockout effects on the growth rate of the cancer cell-line and compared the results with the online experimental database. We also attempted to capture the underlying mechanisms associated with the observed growth reduction rate due to gene knockout. We further analysed our top-ranked genes to get potential targets which were then validated experimentally. It is also observed that multiple knockout tests give a better correlation with experimental observation than single-gene knockout results.

Results

Single gene knockout ranking based on their influence on cancer cell proliferation using the genome-scale metabolic models.

We used the study of Yizhak et al.²⁹ to obtain the GSMMs of cancer from different tissue types. The networks were made using the molecular (gene expression) and phenotypic data (proliferation rate) of cancer cell-lines by applying PRIME method. These data were used to constrain the bounds on the flux values of the corresponding reactions in metabolic networks. The data for NCI-60 collection were taken from the study by Lee et al.³², in which RPMI-1640 was used to grow the cell-lines experimentally. We used 60 cancer metabolic networks across 9 tissue types of NCI-60 panel to find cancer drug targets which inhibit cell growth across all cell-lines. We aimed to rank metabolic genes according to their growth inhibitory effect in cancer cell-lines.

Cancer cell-lines metabolic models used in our work contains 1905 genes. We simulated the models via MOMA^{33,34} and predicted the growth rate of cancer cell-lines after knocking out each gene one by one. Taking the average of fractional cell growth (FCG) of 60 cell-lines, we obtained the mean FCG of each 1905 genes (Fig. 1). Using a cut off value of 10^{-6} (lower circled portion in Fig. 1) on the sorted mean FCG, we obtained 143 genes that are responsible for very low growth rate in our knockout cancer models (see Supplementary File S1). On the other hand, we got 1488 genes, using a cut off value of 0.99995 (the upper circled portion in the Fig. 1), which show a negligible effect on the growth rate (see Supplementary File S1). Interestingly, looking at the reactions corresponding to the 143 genes, we found that all of them, except one gene, is associated with the coupled reactions (Supplementary Fig. S1).

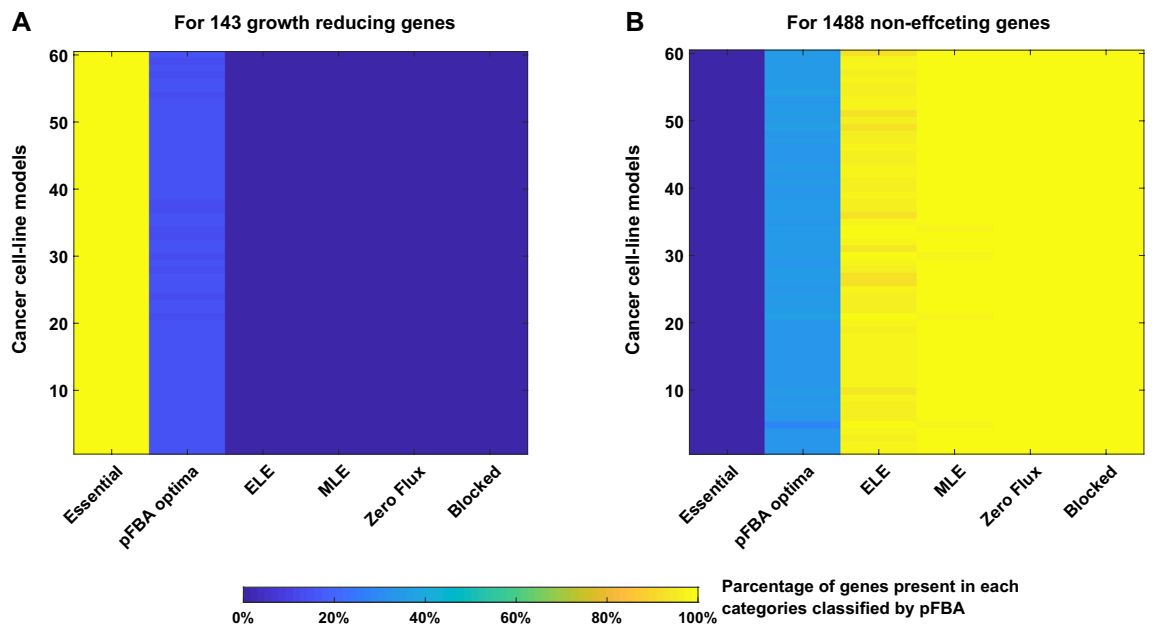


Figure 2. Coverage of pFBA classes by the 143 growth reducing and 1488 non-effecting genes set. For each of the 60 models, pFBA was applied to classify the genes into six categories and then the percentage of involvement for each class into the (A) 143 growth reducing genes set and (B) 1488 non-effecting genes set was calculated.

Mechanistic insight into the genes giving a low growth rate after knockout. We looked for the underlying mechanisms associated with the observed growth reduction rate due to gene knockout. We applied parsimonious enzyme usage FBA (pFBA)³⁵ using COBRA Toolbox³⁶. pFBA classifies each gene into six categories depending on the optimal growth solutions: essential genes, pFBA optima, enzymatically less efficient (ELE), metabolically less efficient (MLE), zero flux genes and blocked genes. There are 71 essential genes, 470–530 pFBA optima, 230–280 ELE, 545–577 MLE, 82 zero flux genes and 427 blocked genes across 60 cancer cell-line models. We looked for the classification of the 143 growth reducing genes and found that these genes contain all the 71 essential genes and the rests are pFBA optima (Fig. 2A). On the other hand, all zero flux and blocked genes, with almost all ELE and MLE genes belong to the 1488 non-effecting genes set (Fig. 2B). Though there are some genes from the 1488 set, which are present in pFBA optima class, but no essential genes are there in the 1488 gene set.

The production fluxes of the metabolites involved in the biomass reaction is changed due to gene knockout. The biomass reaction in the GSMMs uses 43 metabolites as substrate, termed as biomass metabolites (see Supplementary Table S1). We measured the fold changes in the production flux of these biomass metabolites under individual gene knockout condition with respect to wild-type condition. The biomass metabolites whose production flux are reduced by more than twofolds are shown in Fig. 3A. The upper panel of this figure shows the number of biomass metabolites associated with 143 genes responsible for growth reduction, and the lower panel shows the number of biomass metabolites associated with 1488 genes, which do not affect the growth rate. One can observe that the number of biomass metabolites associated with the 143 genes is much higher than that of 1488 genes. To confirm the association of 143 genes with the biomass reaction, we introduced a biomass reduction score (BRS) for each gene (see “Method” for details). A gene with higher BRS has more knockout effect on the biomass reaction. It was observed that BRS of 143 genes are much higher than 1488 genes (Fig. 3B), confirming that they are more effective in reducing the flux of biomass reaction.

Finally, we looked for the biomass metabolites that are associated specifically with the 143 growth reducing genes. We observed that the production flux of 37 biomass metabolites are reduced by knocking out different genes from 143 genes set (Fig. 4A), while different genes from 1488 genes set reduced production flux of 27 biomass metabolites (Fig. 4B). Calculating the set difference, we obtained 16 biomass metabolites that are specifically associated with different genes from the 143 gene set. It is observed that 12 out of 16 biomass metabolites showed association with most of the 143 genes (see Fig. 4C and Supplementary File S1). Next, we looked for these 16 metabolites whether they are flux coupled or not. If they are flux coupled, then one can expect that the corresponding genes become essential for the production of both for the metabolites. We observed that only 4 metabolites out of 16 are flux coupled (Fig. 4D). L-Aspartate (asp-L[c]) is produced from L-Glutamate (glu-L[c]) by the enzyme kinetic reactions Aspartate Transaminase (ASPTA) but there is another transport reaction L-aspartate transport via Na, H symport and K antiport (ASPT6) in which influx of cytosolic L-Aspartate happens from the extracellular space. For the three cell-lines (SNB-75, HOP-9 and SK-OV-3), flux values of another transport reaction aspartate-glutamate mitochondrial shuttle (ASPLUM) was observed, which transfer L-Aspartate from mitochondria to cytosol, and as a consequence, cytosolic L-Glutamate enters into the mitochondria. There is another reaction sterol O-acyltransferase (SOAT11) which uses cholesterol (chsterol[c]) to produce cholesterol ester (xolest_hs[c]).

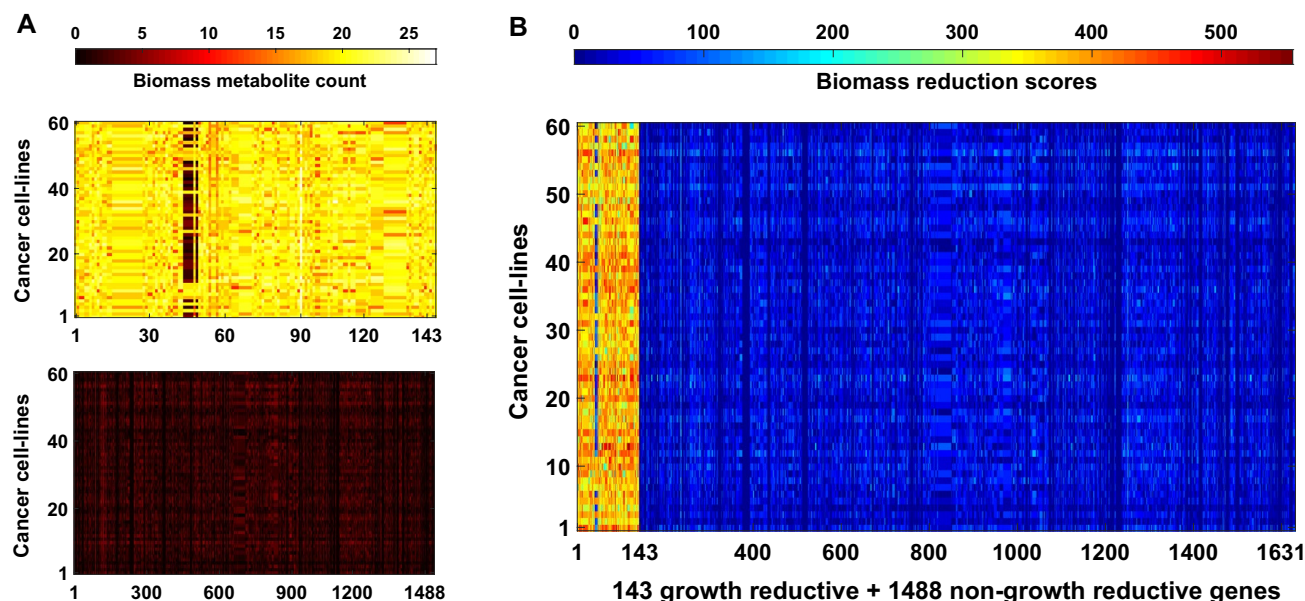


Figure 3. Effect of metabolic genes on biomass function. (A) The upper and lower panels indicate the number of biomass metabolites whose production flux at least twofolds decreased by each of the 143 growth reducing genes and 1488 non-growth reducing genes, respectively. For the details of these genes, see Supplementary File S1. (B) Biomass reduction scores (BRS) of genes following synergic effect. BRS is high for first 143 growth reducing genes and it is low for 1488 non-growth reducing genes.

Finding potential cancer drug targets from the top-ranked genes. Our top-ranked genes can be a potential drug target if their knockout does not significantly affect the growth of non-cancerous cells. So, we need to see the effect of our growth reducing genes on normal cell model. The 60 cell-line panel covers nine different tissues. So, we considered nine models²⁶ built on different cell-type from these nine tissues representing their normal condition. The cell-specific models²⁶ were built from a global reconstruction (Recon 2), which contains 7440 reactions and 2194 transcripts, using protein expression data from the Human Protein Atlas³⁷. These models consist of 2426 ± 467 reactions (\pm s.d.) and 1262 ± 204 transcripts. They applied a published algorithm^{38,39} to predict the flux activity states of the genes by applying an optimization method. The method maximizes the consistency between gene expression and the corresponding enzyme activity. A comparison of the result of normal cells with that of the cancer cell-lines will help us to find targets that can reduce the growth of cancer cells but has minimal effect on the normal cells. Our gene knockout simulation result gave us 143 genes whose knockout can reduce the growth rate in cancer cell-line metabolic models. Some of these genes have multiple isoforms. After removing those isoforms, the gene list reduced to 121 unique genes and all of them could be potential drug targets. However, to be a potential drug target, a gene should show a minimal activity in normal cells. So, we looked for the activity of these 121 genes on normal cell models and found from the literature²⁶ that only 13 genes (Table 1) out of 121 are inactive across all the 9 normal cell models. Thus we have 13 genes that reduce the growth in 60 cancer cell-lines models but have no effect on the flux state of the normal cell models. Therefore, these 13 genes can be considered as potential drug targets against cancer for these 9 tissues with minimal side effects. Interestingly, each of these 13 genes has very high BRS (> 353) (see Table 1). Most of the genes (UQCR11, CYC1, UQCRQ, UQCR10, MT-CYB, UQCRB, UQCRC1, UQCRC2, UQCRFS1 and UQCRH) given in Table 1 have same BRS because they are associated with the same reactions “Ubiquinol-6 Cytochrome C Reductase, Complex III” with AND combinations. This reaction is catalysed by cytochrome bc1 complex (EC 1.10.2.2), which is the third complex in the electron transport chain in mitochondria and the subunit proteins are encoded by these ten genes. It plays a critical role in ATP generation process by catalysing electron transfer from ubiquinol to cytochrome c, coupled to proton transport from the matrix space to the intermembrane space of mitochondria^{40,41}.

Experimental validation of the identified potential drug targets. To validate our simulation results for identifying potential drug targets, we searched for the inhibitors of these 13 genes and used two commercially available inhibitors, i.e., mitotane (SOAT1 inhibitor)⁴² and myxothiazol (CYTB inhibitor)⁴³ for in-vitro studies. Mitotane is reported to show anticancer activity in some cell-lines such as NCI-H295, HeLa, HepG2, IMR-32 and HEK293⁴². So we considered four cell-lines (HCT116, K562, HL60 and A549) from NCI-60 cell-line panel, which are different from the cell-lines reported in Ref.⁴². The effect of these two inhibitors on the cell viability was studied by adding the inhibitor at different concentrations and measuring the growth rate of cell-lines. Fold change in the growth rate (i.e., cell viability) was calculated for each cell-line at different drug concentrations for both the drugs and was plotted in Fig. 5. An EC₅₀ value (concentration of drug at 50% fold change cell viability) was calculated for each cell-line from the resultant curve. The EC₅₀ value of mitotane was 37.83 μ M (for HCT116), 60.85 μ M (for K562), 38.51 μ M (for HL60) and 57.99 μ M (for A549) and for myxothiazol, it was

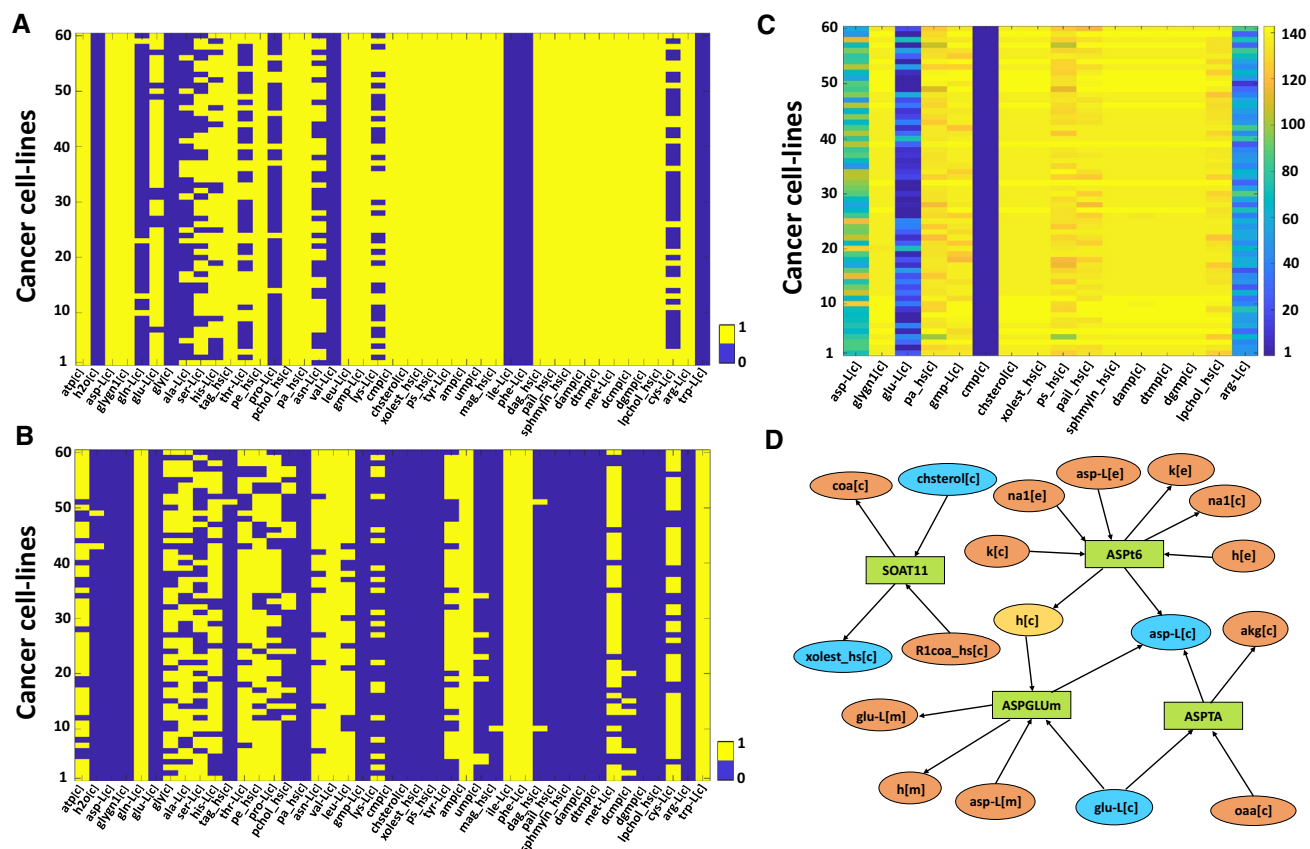


Figure 4. Biomass metabolites affected by the 143 growth reducing genes and 1488 non-affecting genes. Reduction in production fluxes for each individual biomass metabolites were captured in the gene knockout condition. (A) Represent the production flux are twofold decreased or not for any 143 genes and (B) represents the case for 1488 genes. We gave value 1 if the production flux of a metabolite is decreased in at least one gene knockout condition across each cell-lines, otherwise, the value remains zero. For 16 metabolites, all values are zero across the 60 cancer cell-lines in (B) but there are non-zero values in (A). (C) The number of genes among the set of 143 growth reducing genes whose knockout shows at least twofold reduction in the production fluxes for each of the 16 metabolites. (D) A bipartite network that represents the connection between L-Aspartate, L-Glutamate, cholesterol and cholesterol ester. Rectangular and ellipse-shaped boxes indicate the reaction and metabolite respectively and the arrow shows the flow of the flux.

No.	Gene ID	Gene symbol	Avg. of BRS
1	10975	UQCR11	379
2	1537	CYC1	379
3	27089	UQCRQ	379
4	29796	UQCR10	379
5	31	ACACA	385
6	32	ACACB	370
7	4519	MT-CYB	379
8	6646	SOAT1	383
9	7381	UQCRB	379
10	7384	UQCRC1	379
11	7385	UQCRC2	379
12	7386	UQCRES1	379
13	7388	UQCRH	379

Table 1. Gene ID, gene symbol and the corresponding average value of the biomass metabolic score (BRS) across 60 cancer cell-line models of the cancer specific drug targets.

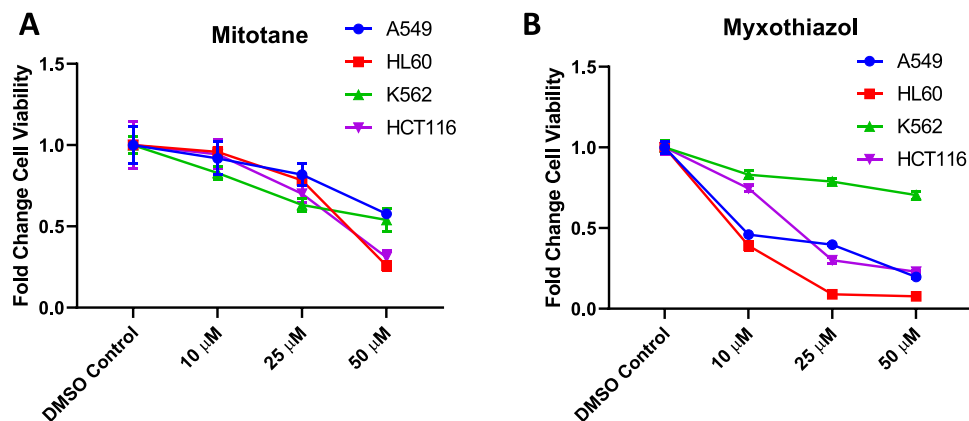


Figure 5. Effect of mitotane and myxothiazol in cell viability. Experimentally measured fold change in cell viability of four different cell-lines (A549, HL60, K562, HCT116) at different concentrations of mitotane and myxothiazol.

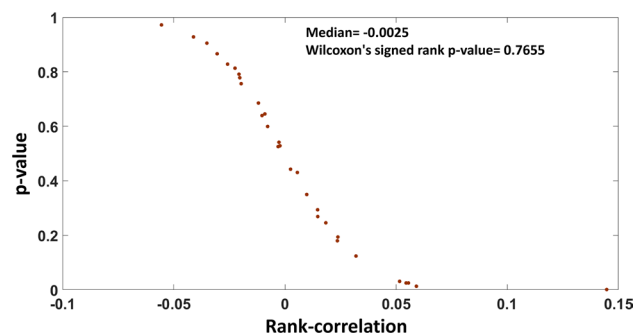


Figure 6. Comparison of gene knockout simulation results with DEMETER database. Distribution of measured Spearman rank-correlation between predicted FCG of genes using GSMM and experimental data from DEMETER for the 30 cell-lines. The p -values were measured using permutation test.

18.28 μM (for HCT116), 84.92 μM (for K562), 8.21 μM (for HL60) and 9.25 μM (for A549). Thus, mitotane and myxothiazol both are effective in inhibiting the growth of these four cell-lines.

Testing the predictive ability of GSMM for single-gene knockout. We extracted the information from DEMETER database⁴⁴ on the knockdown effect of 1444 genes in 30 cell-lines (Supplementary File S1). We calculated the Spearman rank-correlation between the predicted FCG from our GSMM and the experimental data from DEMETER database for each of the 30 cell-lines. The corresponding p -value of the rank-correlation for each cell-line was obtained by permutation test. It was observed that most of the obtained positive rank-correlation were not significant (Fig. 6). There were only 5 cell-lines, which are showing significant positive rank-correlation but their correlation value was less than 0.15. Thus, the obtained gene ranking from single-gene knockout results does not show much correlation with the experimentally observed result. So, we looked for the effect of multiple genes knockout on the growth rate.

Identification of multiple targets using DrugBank database information. We took the drugs available in the DrugBank database⁴⁵ based on the gene target information. DrugBank database contains the biochemical and pharmacological information about the drugs. We only selected 380 drugs which are inhibitory and the corresponding target genes are present in the cancer cell-lines metabolic models (see Supplementary File S1). These 380 drugs have 202 metabolic targets in the models. To observe the effect of a particular drug on the growth rate, we knocked out all the genes that were inhibited by that drug and simulated the models via MOMA. This exercise was repeated for all the 380 drugs across all the 60 cancer cell-line models. We used a cut off value of 0.5 on FCG (representing at least 50% reduction on the growth rate) to call a drug active (see Supplementary File S1) and obtained 76 drugs. 10 out of these 76 drugs have already been approved in cancer treatment and another 18 drugs are in different phases of the clinical trial.

We considered NCI-60 growth inhibition database⁴⁶ to get the GI50 values of the drugs and compared the predicted anti-proliferative activity with the experimentally measured potency of the drugs. There are 64 drugs with GI50 value in the NCI-60 growth inhibition database. Out of these 64 drugs, 23 drugs have mean log GI50

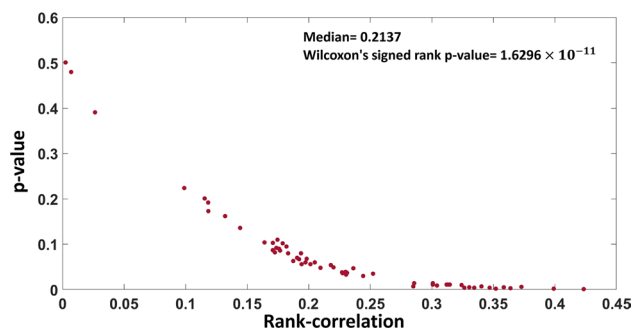


Figure 7. Validation of predicted cell-line specific responses of drugs. Distribution of measured Spearman rank-correlation between predicted FCG of drugs using GSMM and their log GI50 values for the 60 cell-lines. The p -values were measured using permutation test.

value less than -5 across 60 cell-lines and therefore considered to be active against cancer for most of the cell-lines. Comparing these 23 drugs with our list of 76 drugs, we obtained 17 drugs common in both sets.

Finally, a cell-wise comparison between predicted FCG of drugs using GSMM and their log GI50 values was performed and the Spearman rank-correlations was obtained. It is observed that around 50% cell-lines were showing significant rank-correlation and their correlations were also higher than single knockout result, see Fig. 7 (median Spearman rank-correlation = 0.2137, Wilcoxon's signed-rank p -value = 1.6296×10^{-11}).

Linking the significance of single-gene knockout ranking on the activity of drug. We used our gene knockout results of 202 target genes corresponding to those 380 drugs to test the significance of gene ranking. It was observed that 37 genes were present at the top position and 146 genes at the bottom in our gene ranking list (see Supplementary Table S2). The drugs corresponding to these 37 genes significantly reduced (mean FCG $< 10^{-6}$) the growth rate of cancer models (Fig. 8). On the contrary, those drugs whose targets belong to the set of 146 genes, placed at the bottom in our gene ranking list, showed no effect in the cancer models.

Discussion

Recently, GSMM gained a lot of attention in drug discovery. It has been used to study drugs related to cancer. Ghaffari et al.⁴⁷ explored strategy for identifying anti-growth factors for the inhibition of cell growth using GSMM on 11 cell-lines and identified potential antimetabolites that could inhibit the growth or kill any of the cell-lines. They also checked the in-silico toxicity by employing GSMMs for 83 human healthy cell-types. The same methodology had been applied by Agren et al.³⁰ in a different study to find potential drugs for hepatocellular carcinoma (HCC) by reconstructing and analysing personalized GSMMs for six HCC patients. Raškevičius et al.⁶ used GSMMs to predict the putative effects against cancer of those compounds which are structurally similar to human metabolite and also gave a concept of finding therapeutic windows through GSMMs. Turanil et al.⁴⁸ introduced a drug repositioning based method via GSMMs to predict therapeutic agents for cancer treatment. They reconstructed prostate cancer-specific GSMM by combining personalized GSMMs ($n > 450$) and proteomics data. They used drug-perturbed gene expression data of three cell-lines (PC3, HL60, and MCF7) from the ConnectivityMap2 (CMap2)⁴⁹ to reveal drug off-targets by predicting novel gene–drug interactions and evaluated in-silico cell viability. The present study aimed to develop a knockout strategies for identifying potential drug targets and the associated mechanisms using GSMM and gene expression data. This will help us to get novel drug targets as well as targets that might be used for drug repurposing. We used existing GSMMs of NCI-60 cell-lines²⁹ to predict the anti-proliferative activity of single metabolic genes as targets against cancer and ranked them accordingly. We got 143 genes whose knockout reduced the cell growth across all the metabolic models of the NCI-60 panel. We also obtained a list of 1488 genes whose knockout does not show any effect on the growth rate of any cancer model. We searched for the underlying mechanism for such reduction in the growth rate of cancer cells by 143 genes and found that the biomass reduction score (BRS) of those genes were much higher than those of 1488 genes. The synergic effect in biomass reaction is much more for 143 gene list than 1488 gene list. It is observed that there are 12 biomass metabolites which are influenced by almost all the 143 genes but not by any of the 1488 genes for all the 60 cell-lines. Glycogen is the top-ranked among those 12 biomass metabolites. It is already reported that glycogen pathway is up-regulated in various cancers^{50,51} and funneled into glycolysis to promote cell growth, invasion and metastasis⁵². Likewise, glycogen is used by cancer cells to survive under nutrient starvation condition⁵³. Considering its importance, inhibition of glycogen metabolism has become a new potential strategy for cancer treatment^{50,54–56}.

We looked for potential targets from these 143 genes which showed a significant reduction in the growth rate following knockout. To be a potential target, these genes need to show the minimal knockout effect on the normal cells. We obtained 13 such targets from 143 genes whose knockout reduced the proliferation rate of cancer cells but were inactive across all the 9 normal cell models. One of the main features of the identified targets was that they were showing their effect in all the NCI-60 cell-lines. To experimentally validate the effect of these targets on multiple cell-lines, we chose SOAT1 and CYTB and showed that its inhibition reduces the growth rate in multiple cell-lines. Inhibition of SOAT1 is known to reduce growth rate in 5 cell-lines⁴² and additionally we have

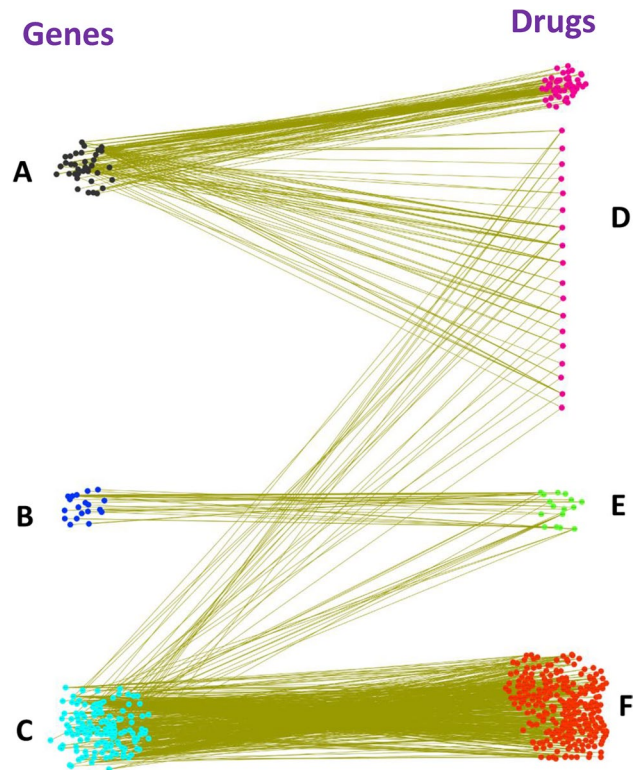


Figure 8. Robustness of gene ranking. (A–C) Gene ranking of total 202 targets of at least one of the 380 drugs (Supplementary Table S2). (A) Top 37 genes lie on the lower circled portion (mean FCG <math>< 10^{-6}</math>) in Fig. 1. (B) 19 Genes lie in between the lower and upper circled portion (mean FCG lies between

showed its effect on 4 more cell-lines. These cell-lines are taken from different tissues like blood peripheral, colon, adrenocortical gland, cervix, liver, kidney and brain. We have used mitotane to inhibit SOAT1, which is already known for the treatment of adrenocortical carcinoma and Cushing's syndrome^{57–59}. However, the inhibitor of CYTB, myxotiazol, is not known as an anticancer drug, but we have seen its growth inhibiting effect on four cancer cell-lines. This could be a potential novel repurposed drug and need further evaluation. Moreover, literature survey showed that inhibition of SOAT1 and CYTB do not have any significant influence on the growth of the normal cells^{60–62}. Literature also supports the validity of our other identified targets (given in Table 1). For example, atovaquone, a potent and selective mitochondrial inhibitor^{63,64}, has been shown to reduce proliferation in cervical cancer cell-lines⁶⁵, Du145 prostate cancer cells⁶⁶ and MCF7-derived Cancer Stem-like CSCs⁶². Another identified target is UQCRB, whose inhibitor terpestacin blocks vascular endothelial growth factor (VEGF)-induced angiogenesis in endothelial cells and is proposed to be applied as a drug for human cancer^{67,68}. Cytochrome c-1 (CYC1) is found to play an important role in breast cancer patients. Knocking down of CYC1 inhibits proliferation in human breast cancer cell-lines⁶⁹. In another study, it was observed that silencing CYC1 by shRNA transfection also inhibits proliferation in human osteosarcoma (OS) cells⁷⁰. Another identified gene target UQCRFS1 appears to be involved in the progression of gastric cancers and in the development of more aggressive phenotype of breast cancer^{71,72}. Lentivirus-mediated knockdown of UQCRC2 suppresses cell growth and colony formation in RKO and HCT116 cells, result in cell cycle arrest and induce cell apoptosis in vitro in colorectal cancer (CRC)⁷³. The knockdown of ACACA expression inhibits cell proliferation in prostate⁷⁴ and breast⁷⁵ cancer cell-lines.

We used Spearman rank-correlation method to compare our single-gene knockout ranking obtained from the GSMM with the experimental data. Most of the observed rank-correlations were very low and/or not significant. This was in agreement with other studies where single-gene knockout does not show the desired result^{76–78}. On the other hand, when applied multiple gene knockout strategies, we obtained a higher and significant rank-correlations with the experimental results. It leads to the conclusion that multiple genes knockout show a better result than single-gene knockout, confirming the similar observations by other studies^{77–81}. This might be because genes or proteins interact in a complex network, where alternative pathways always exist to carry the function^{77,82}. Though the single knockout results did not give the desired correlations, the gene ranking obtained using this strategy seemed to be significant. It was noticed that a drug could only be active if it has at least one target belonging to the top rank. In case, none of the targets is from the top rank then that drug is observed to be ineffective.

The developed strategy to use gene knockout in GSMMs to identify drug targets has general applicability with any novel set of gene expression data associated with a tumour. The tumour-specific GSMM can be built from the gene expression data to go for the single or multiple gene knockout strategies. We proposed to use drug information for multiple gene knockout strategies that can also serve as drug repurposing technique. This method could be used to identify novel drug targets as well as targets that might be used by existing drugs on novel kinds of tumour. It is, therefore, possible to identify genes that are more relevant to specific cancer. The developed methodology could also be used to screen for common therapy. This strategy was used in the present study where we obtained 143 genes whose knockout were showing significant growth reduction across all cell-lines (see Supplementary Fig. S2)

Finally, we want to mention that the GSMM models use the metabolic flux values of reactions for prediction. It is known that these fluxes depend on protein's post-translational state as well as corresponding metabolite levels⁸³. More detailed data of these cell-lines (proteomics and metabolomics)^{84–86} would help to construct a more accurate model of cancer cell-lines and will certainly enhance the possibility of finding new targets.

Materials and methods

Predicting cancer cell growth by gene knockout. We used minimization of metabolic adjustment (MOMA) technique³³ for observing gene knockout effect on the growth rates of 60 cancer cell-line models. To mimic the knockout condition of a gene, we changed the lower and upper bounds of flux values of the reactions associated with the gene to zero. In the case of simultaneous knockout of multiple genes, the bounds of all reactions of these participating genes were set to zero. To measure the growth rate, we used two models as input for MOMA: one representing the knockout condition and the other one representing the wild-type condition of a particular cell-line. The fractional cell growth (FCG) was then calculated by taking the ratio of growth rates in knockout condition to wild-type condition. COBRA toolbox was used from the MATLAB package³⁴ for MOMA. Though there exists several genes associated with the same reaction with non-trivial AND and OR combinations, in order to keep our analysis simple, we followed the literatures^{27,29,87,88} and assumed zero flux for the reaction associated with the knockout gene.

Gene symbol to Gene ID conversion. In GSMMs, the identity of genes is given in gene ID format and therefore the gene symbols need to be converted into gene IDs. For this purpose, we used the Uniprot database, which provides the mapping between gene IDs and gene symbols of the metabolic genes. Using this method, one can find out the metabolic genes from any given gene set having gene ID information. One can also extract genes from any given gene set, which are present in the cancer cell-line metabolic models, by taking intersection with the 1905 gene IDs present in the cancer models.

Gene knockout effect on the production flux rate of metabolite. Flux balance analysis (FBA) was applied to get the flux values of each of the metabolic reactions present in the metabolic networks under the wild-type and gene knockout condition. Then, the production flux rate of a metabolite was obtained by summing up the flux values of the reactions in which it is produced. Finally, the ratio between the production flux rates under the two different conditions was calculated.

Biomass reduction score (BRS). Total 43 metabolites are used as the substrate in the biomass reaction, and we termed them as biomass metabolites (see Supplementary Table S1). These biomass metabolites are then arranged in the decreasing order of their coefficient values in the biomass reaction and scored them accordingly. For example, the biomass metabolite having the highest coefficient was scored 43 and the biomass metabolite with the lowest coefficient was scored 1. To capture the individual gene knockout effect on the biomass reaction, we first listed out the biomass metabolites whose production flux rates are reduced by more than twofolds. We then calculated the sum of the scores of these biomass metabolites and termed them as biomass reduction score (BRS).

MTT assay. Different cancerous cell-lines were maintained at 37 °C in a humidified incubator with 5% CO₂ supply. A549 (Human lung carcinoma), HCT116 (Human colorectal carcinoma), K562 (Human chronic myelogenous leukaemia) cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Hyclone, West South Logan, USA) and HL60 (Human promyeloblast) was grown in RPMI 1640 (Gibco) supplemented with 10% fetal bovine serum (FBS). The SOAT1 inhibitor mitotane and CYTB inhibitor myxothiazol were purchased from Sigma Aldrich (St. Louis, USA). They were dissolved in dimethyl sulfoxide (DMSO) (Sigma-Aldrich, MO, United States) and stored in –20 °C as a stock concentration of 10 mM and 50 mM respectively. The effect of mitotane and myxothiazol on the proliferation and viability of all cell-lines was monitored using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. MTT powder was purchased from Merck, USA and was dissolved in PBS (5 mg/ml) prior to the assay. MTT assay is based on the principle of reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide to purple coloured formazan by the mitochondria of metabolically active cells. Cells were seeded at a density of 2×10^4 cells per well in 96 well plate for 24 h followed by addition of the drug at different concentrations. The MTT assay was performed after 48 h of incubation with the drug. Each experiment was performed three times with five replicates. Mean, standard deviation and standard error were determined using Graphpad prism.

Extraction of experimental gene knockdown information from database. We used DEMETER database⁴⁴ to obtain the experimental gene knockdown effect on cell growth. It contains a knockdown effect on

the viability of cells from 501 cancer cell-lines using the shRNA library. The data is given in the form of log fold change in cell number due to gene knockdown. A gene with a negative log fold change value implies that there is a reduction in the cell growth upon knockdown of this gene⁴⁴. The database contains the knockdown results of 17,098 genes on 501 cell-lines, out of which 30 cell-lines are present in NCI-60 cell-lines panel. In our cancer models, we got only 1444 genes from those 17,098 genes. So, we obtained the gene knockdown effect of 1444 metabolic genes on 30 cancer cell-lines.

Permutation test. Permutation test was applied to get the p -value corresponding to each Spearman rank-correlation. In the first step, Spearman rank-correlation was calculated between the predicted data and the experimental data. Next, the predicted data were randomly permuted by using the *randperm* function in MATLAB. The function *randperm* gives a randomly permuted vector of the integers from 1 to n without repeating elements. In our case, n represents the number of data points. The elements of the randomly generated vector is used as index of the original data to generate the permuted data. Spearman rank-correlations were calculated between the permuted data and the experimental data. This process was repeated for 1000 times and the p -value was obtained by using the formula $(r + 1)/1001$, where r is the number of cases for which random permuted set gave better rank-correlation than the non-permuted case.

To verify the usefulness of the permutation test we obtained the distribution of the rank-correlations for the 1000 randomly permuted set (see Supplementary Fig. S3 for three cell-lines). It was observed that they follow normal distribution.

Finding drugs with inhibitory type nature from DrugBank database. In DrugBank database⁴⁵, 6490 drugs (out of 8283) have information about its targets and the corresponding mechanism of action. These targets are given in the form of gene names. There are 1684 drugs which have gene ID information of at least one metabolic target (see method “Gene symbol to Gene ID conversion”). We then searched for the drugs that can decrease the activity of at least one metabolic target. For this purpose, we selected the following three mechanisms of actions: (i) inhibition, (ii) antagonist, (iii) inverse agonist. We found 410 drugs out of 1684 which decreased the activity of at least one metabolic gene. Out of these 410 drugs, only 380 drugs have at least one inhibitory type target on cancer cell-lines GSMMS.

Link between DrugBank database and NCI-60 growth inhibition database. DrugBank database has DrugBank ID, whereas NCI-60 growth inhibition database⁴⁶ has NSC ID. So, we first converted their IDs into a single ID for further analysis. In the DrugBank database, the conversion of drug ID to PubChem CID or CAS ID is provided and the conversion of CAS ID to NSC ID is given in the chemical data section of the NCI-60 growth inhibition database. We therefore used CAS ID information of the drugs to link these two datasets. There are 373 drugs, out of previously described 380 drugs, which have CAS ID information and out of them only 200 drugs have NSC IDs. However, all NSC IDs do not have information regarding GI50 value. So, we finally obtained 64 drugs with NSC IDs that have a GI50 score in negative log value.

Finding active drugs using GI50 score. GI50 score is an important measure of drug activity. It quantifies the dosage of the drug required to inhibit the cell growth by 50%. In the NCI-60 growth inhibition database, the range of log value of GI50 score of the drugs is given between -10 to -1 . Drugs with log GI50 score less than -5 was considered to be active against cancer^{89,90}.

Data and code availability

The data and the code are provided as Supplementary File S1, Supplementary Table S1 and Supplementary File S2.

Received: 11 March 2020; Accepted: 11 December 2020

Published online: 08 January 2021

References

- Hornberg, J. J., Bruggeman, F. J., Westerhoff, H. V. & Lankelma, J. Cancer: A systems biology disease. *Biosystem* **83**, 81–90 (2006).
- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA Cancer J. Clin.* **68**, 7–30 (2018).
- Fitzmaurice, C. *et al.* Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA Oncol.* **3**, 524–548 (2017).
- Maruthappu, M. *et al.* Investments in cancer research awarded to UK institutions and the global burden of cancer 2000–2013: a systematic analysis. *BMJ Open* **7**, e013936 (2017).
- Fu, H. Y. *et al.* Protein quality control dysfunction in cardiovascular complications induced by anti-cancer drugs. *Cardiovasc. Drugs Ther.* **31**, 109–117 (2017).
- Raškevičius, V. *et al.* Genome scale metabolic models as tools for drug design and personalized medicine. *PLoS One* **13**, e0190636 (2018).
- Whirl-Carrillo, M. *et al.* Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **92**, 414–417 (2012).
- Werner, H. M., Mills, G. B. & Ram, P. T. Cancer systems biology: A peek into the future of patient care?. *Nat. Rev. Clin. Oncol.* **11**, 167–176 (2014).
- Yizhak, K., Chaneton, B., Gottlieb, E. & Ruppin, E. Modeling cancer metabolism on a genome scale. *Mol. Syst. Biol.* **11**, 817 (2015).
- Swinney, D. C. & Anthony, J. How were new medicines discovered?. *Nat. Rev. Drug Discov.* **10**, 507 (2011).
- Hopkins, A. L. Network pharmacology: The next paradigm in drug discovery. *Nat. Chem. Biol.* **4**, 682 (2008).
- Zambrowicz, B. P., Turner, C. A. & Sands, A. T. Predicting drug efficacy: Knockouts model pipeline drugs of the pharmaceutical industry. *Curr. Opin. Pharmacol.* **3**, 563–570 (2003).
- Zambrowicz, B. P. & Sands, A. T. Modeling drug action in the mouse with knockouts and RNA interference. *Drug Discov. Today TARGETS* **3**, 198–207 (2004).

14. Wahi, K. & Holst, J. Asct2: A potential cancer drug target. *Expert. Opin. Ther. Targets* **23**(7), 555–558 (2019).
15. Chen, X. *et al.* Inhibitors of plasmodium falciparum methionine aminopeptidase 1b possess antimalarial activity. *Proc. Natl. Acad. Sci.* **103**, 14548–14553 (2006).
16. Chong, C. R., Chen, X., Shi, L., Liu, J. O. & Sullivan, D. J. Jr. A clinical drug library screen identifies astemizole as an antimalarial agent. *Nat. Chem. Biol.* **2**, 415 (2006).
17. Zheng, W., Thorne, N. & McKew, J. C. Phenotypic screens as a renewed approach for drug discovery. *Drug Discov. Today* **18**, 1067–1073 (2013).
18. Parvathaneni, V., Kulkarni, N. S., Muth, A. & Gupta, V. Drug repurposing: A promising tool to accelerate the drug discovery process. *Drug Discov. Today* **24**(10), 2076–2085 (2019).
19. King, F. J. *et al.* Pathway reporter assays reveal small molecule mechanisms of action. *JALA J. Assoc. Lab. Autom.* **14**, 374–382 (2009).
20. Reaume, A. G. Drug repurposing through nonhypothesis driven phenotypic screening. *Drug Discov. Today Ther. Strateg.* **8**, 85–88 (2011).
21. Warchal, S. J., Unciti-Broceta, A. & Carragher, N. O. Next-generation phenotypic screening. *Future Med. Chem.* **8**, 1331–1347 (2016).
22. Oberhardt, M. A., Yizhak, K. & Ruppin, E. Metabolically re-modeling the drug pipeline. *Curr. Opin. Pharmacol.* **13**, 778–785 (2013).
23. Kell, D. B. Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discov. Today* **11**, 1085–1092 (2006).
24. Nielsen, J. Systems biology of metabolism: A driver for developing personalized and precision medicine. *Cell Metab.* **25**, 572–579 (2017).
25. Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci.* **104**, 1777–1782 (2007).
26. Thiele, I. *et al.* A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **31**, 419–425 (2013).
27. Folger, O. *et al.* Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* **7**, 501 (2011).
28. Agren, R. *et al.* Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using init. *PLoS Comput. Biol.* **8**, e1002518 (2012).
29. Yizhak, K. *et al.* Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer. *Elife* **3**, e03641 (2014).
30. Agren, R. *et al.* Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol. Syst. Biol.* **10**, 721 (2014).
31. Yizhak, K. *et al.* A computational study of the warburg effect identifies metabolic targets inhibiting cancer migration. *Mol. Syst. Biol.* **10**, 744 (2014).
32. Lee, J. K. *et al.* A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. *Proc. Natl. Acad. Sci.* **104**, 13086–13091 (2007).
33. Segre, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci.* **99**, 15112–15117 (2002).
34. Becker, S. A. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the cobra toolbox. *Nat. Protoc.* **2**, 727–738 (2007).
35. Lewis, N. E. *et al.* Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* **6**, 390 (2010).
36. Heirendt, L. *et al.* Creation and analysis of biochemical constraint-based models using the cobra toolbox v. 3.0. *Nat. Protoc.* **14**, 639–702 (2019).
37. Uhlen, M. *et al.* Towards a knowledge-based human protein atlas. *Nat. Biotechnol.* **28**, 1248 (2010).
38. Schellenberger, J. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: The cobra toolbox v2.0. *Nat. Protoc.* **6**, 1290 (2011).
39. Shlomi, T., Cabili, M. N., Herrgård, M. J., Palsson, B. Ø. & Ruppin, E. Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* **26**, 1003 (2008).
40. Brandt, U. & Trumppower, B. The protonmotive q cycle in mitochondria and bacteria. *Crit. Rev. Biochem. Mol. Biol.* **29**, 165–197 (1994).
41. Crofts, A. R. The cytochrome bc 1 complex: Function in the context of structure. *Annu. Rev. Physiol.* **66**, 689–733 (2004).
42. Sbiera, S. *et al.* Mitotane inhibits sterol-o-acyl transferase 1 triggering lipid-mediated endoplasmic reticulum stress and apoptosis in adrenocortical carcinoma cells. *Endocrinology* **156**, 3895–3908 (2015).
43. Von Jagow, G., Ljungdahl, P. O., Graf, P., Ohnishi, T. & Trumppower, B. An inhibitor of mitochondrial respiration which binds to cytochrome b and displaces quinone from the iron-sulfur protein of the cytochrome bc1 complex. *J. Biol. Chem.* **259**, 6318–6326 (1984).
44. Tsherniak, A. *et al.* Defining a cancer dependency map. *Cell* **170**, 564–576 (2017).
45. V, Law *et al.* Drugbank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **42**, D1091–D1097 (2013).
46. Shoemaker, R. H. The nci60 human tumour cell line anticancer drug screen. *Nat. Rev. Cancer* **6**, 813 (2006).
47. Ghaffari, P. *et al.* Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling. *Sci. Rep.* **5**, 1–10 (2015).
48. Turanli, B. *et al.* Discovery of therapeutic agents for prostate cancer using genome-scale metabolic modeling and drug repositioning. *EBioMedicine* **42**, 386–396 (2019).
49. Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
50. Zois, C. E., Favaro, E. & Harris, A. L. Glycogen metabolism in cancer. *Biochem. Pharmacol.* **92**, 3–11 (2014).
51. Rousset, M. *et al.* Growth-related enzymatic control of glycogen metabolism in cultured human tumor cells. *Cancer Res.* **44**, 154–160 (1984).
52. Curtis, M. *et al.* Fibroblasts mobilize tumor cell glycogen to promote proliferation and metastasis. *Cell Metab.* **29**, 141–155 (2019).
53. Pelletier, J. *et al.* Glycogen synthesis is induced in hypoxia by the hypoxia-inducible factor and promotes cancer cell survival. *Front. Oncol.* **2**, 18 (2012).
54. Miyashita, K. *et al.* An emerging strategy for cancer treatment targeting aberrant glycogen synthase kinase 3 β . *Anti-Cancer Agents Med. Chem. (Formerly Curr. Med. Chem. Agents)* **9**, 1114–1122 (2009).
55. Dukhande, V. V., Barot, S., Husein, S. & Palaguachi, C. Inhibition of glycogen metabolism as a potential strategy for anticancer therapy. *FASEB J.* **31**, 942–1010 (2017).
56. Barot, S., Abo-Ali, E. M., Zhou, D. L., Palaguachi, C. & Dukhande, V. V. Inhibition of glycogen catabolism induces intrinsic apoptosis and augments multikinase inhibitors in hepatocellular carcinoma cells. *Exp. Cell Res.* **381**(2), 288–300 (2019).
57. Hahner, S. & Fassnacht, M. Mitotane for adrenocortical carcinoma treatment. *Curr. Opin. Investig. Drugs (London, England: 2000)* **6**, 386–394 (2005).
58. Terzolo, M. *et al.* Adjuvant mitotane treatment for adrenocortical carcinoma. *N. Engl. J. Med.* **356**, 2372–2380 (2007).
59. Baudry, C. *et al.* Efficiency and tolerance of mitotane in Cushing's disease in 76 patients from a single center. *Eur. J. Endocrinol.* **167**, 473–481 (2012).

60. Geng, F. *et al.* Inhibition of soat1 suppresses glioblastoma growth via blocking srebp-1-mediated lipogenesis. *Clin. Cancer Res.* **22**, 5337–5348 (2016).
61. Oni, T. E. *et al.* Soat1 promotes mevalonate pathway dependency in pancreatic cancer. *J. Exp. Medicine* **217** (2020).
62. Fiorillo, M. *et al.* Repurposing atovaquone: Targeting mitochondrial complex iii and oxphos to eradicate cancer stem cells. *Oncotarget* **7**, 34084 (2016).
63. Fry, M. & Pudney, M. Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4'-chlorophenyl) cyclohexyl]-3-hydroxy-1, 4-naphthoquinone (566c80). *Biochem. Pharmacol.* **43**, 1545–1553 (1992).
64. Srivastava, I. K., Rottenberg, H. & Vaidya, A. B. Atovaquone, a broad spectrum antiparasitic drug, collapses mitochondrial membrane potential in a malarial parasite. *J. Biol. Chem.* **272**, 3961–3966 (1997).
65. Tian, S., Chen, H. & Tan, W. Targeting mitochondrial respiration as a therapeutic strategy for cervical cancer. *Biochem. Biophys. Res. Commun.* **499**, 1019–1024 (2018).
66. Zhou, J. *et al.* Atovaquone derivatives as potent cytotoxic and apoptosis inducing agents. *Bioorg. Med. Chem. Lett.* **19**, 5091–5094 (2009).
67. Jung, H. J. *et al.* Mitochondrial uqcrb regulates vegfr2 signaling in endothelial cells. *J. Mol. Med.* **91**, 1117–1128 (2013).
68. Jung, N., Kwon, H. J. & Jung, H. J. Downregulation of mitochondrial uqcrb inhibits cancer stem cell-like properties in glioblastoma. *Int. J. Oncol.* **52**, 241–251 (2018).
69. Han, Y. *et al.* Cyc1 predicts poor prognosis in patients with breast cancer. *Dis. Markers* **2016**, (2016).
70. Li, G. *et al.* Cyc1 silencing sensitizes osteosarcoma cells to trail-induced apoptosis. *Cell. Physiol. Biochem.* **34**, 2070–2080 (2014).
71. Jun, K. H., Kim, S. Y., Yoon, J. H., Song, J. H. & Park, W. S. Amplification of the uqcrf1 gene in gastric cancers. *J. Gastr. Cancer* **12**, 73–80 (2012).
72. Ohashi, Y., Kaneko, S. J., Cupples, T. E. & Young, S. R. Ubiquinol cytochrome c reductase (uqcrf1) gene amplification in primary breast cancer core biopsy samples. *Gynecol. Oncol.* **93**, 54–58 (2004).
73. Shang, Y. *et al.* Overexpression of uqcr2 is correlated with tumor progression and poor prognosis in colorectal cancer. *Pathol. Pract.* **214**, 1613–1620 (2018).
74. Brusselmans, K., De Schrijver, E., Verhoeven, G. & Swinnen, J. V. RNA interference-mediated silencing of the acetyl-coa-carboxylase- α gene induces growth inhibition and apoptosis of prostate cancer cells. *Cancer Res.* **65**, 6719–6725 (2005).
75. Chajès, V., Cambot, M., Moreau, K., Lenoir, G. M. & Joulin, V. Acetyl-coa carboxylase α is essential to breast cancer cell survival. *Cancer Res.* **66**, 5287–5294 (2006).
76. Gu, Z. *et al.* Role of duplicate genes in genetic robustness against null mutations. *Nature* **421**, 63–66 (2003).
77. Deutscher, D., Meilijson, I., Schuster, S. & Ruppín, E. Can single knockouts accurately single out gene functions?. *BMC Syst. Biol.* **2**, 50 (2008).
78. Deutscher, D., Meilijson, I., Kupiec, M. & Ruppín, E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat. Genet.* **38**, 993–998 (2006).
79. Shlomi, T. *et al.* Systematic condition-dependent annotation of metabolic genes. *Genome Res.* **17**, 1626–1633 (2007).
80. Kuepfer, L., Sauer, U. & Blank, L. M. Metabolic functions of duplicate genes in *saccharomyces cerevisiae*. *Genome Res.* **15**, 1421–1430 (2005).
81. Behre, J., Wilhelm, T., von Kamp, A., Ruppín, E. & Schuster, S. Structural robustness of metabolic networks with respect to multiple knockouts. *J. Theor. Biol.* **252**, 433–441 (2008).
82. Yosef, N., Kaufman, A. & Ruppín, E. Inferring functional pathways from multi-perturbation data. *Bioinformatics* **22**, e539–e546 (2006).
83. Hackett, S. R. *et al.* Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* **354**, aaf2786 (2016).
84. Gholami, A. M. *et al.* Global proteome analysis of the nci-60 cell line panel. *Cell Rep.* **4**, 609–620 (2013).
85. Lima, A. R. *et al.* Discrimination between the human prostate normal and cancer cell exometabolome by GC-MS. *Sci. Rep.* **8**, 5539 (2018).
86. Jain, M. *et al.* Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **336**, 1040–1044 (2012).
87. Dai, Z. *et al.* Identification of cancer-associated metabolic vulnerabilities by modeling multi-objective optimality in metabolism. *Cell Commun. Signal.* **17**, 1–15 (2019).
88. Förster, J., Famili, L., Palsson, B. Ø. & Nielsen, J. Large-scale evaluation of in silico gene deletions in *saccharomyces cerevisiae*. *OMICS A J. Integr. Biol.* **7**, 193–202 (2003).
89. Wang, H. *et al.* Chemical data mining of the nci human tumor cell line database. *J. Chem. Inf. Model.* **47**, 2063–2076 (2007).
90. Lee, A. C., Shedden, K., Rosania, G. R. & Crippen, G. M. Data mining the nci60 to predict generalized cytotoxicity. *J. Chem. Inf. Model.* **48**, 1379–1388 (2008).

Acknowledgements

The work is supported by SERB (Govt. of India), grant no.: CRG/2019/005477. Research of AP is supported by University Grants Commission (11-04-2016-423482) Govt. of India and research of NB is supported by J.U. RUSA 2.0, Ref. No.: R-11/743/19.

Author contributions

A.P., R.A. & S.C. conceived and designed the study. A.P. performed the analysis. S.P.K. & S.R. performed the experimental validation. A.P., N.B. & S.C. interpreted the results and contribute to the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-80561-1>.

Correspondence and requests for materials should be addressed to S.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021