DOI:10.1158/1078-0432.CCR-12-1856

Clinical Cancer Research

Predicting Drug Targets and Biomarkers of Cancer via Genome-Scale Metabolic Modeling

Livnat Jerby and Eytan Ruppin

Clin Cancer Res 2012;18:5572-5584. Published online October 14, 2012.

DOI:10.1158/1078-0432.CCR-12-1856

Predicting Drug Targets and Biomarkers of Cancer via Genome-Scale Metabolic Modeling

Livnat Jerby[1] and Eytan Ruppin[1,2]

Abstract

The metabolism of cancer cells is reprogrammed in various ways to support their growth and survival. Studying these phenomena to develop noninvasive diagnostic tools and selective treatments is a promising avenue. Metabolic modeling has recently emerged as a new way to study human metabolism in a systematic, genome-scale manner by using pertinent high-throughput omics data. This method has been shown in various studies to provide fairly accurate estimates of the metabolic phenotype and its modifications following genetic and environmental perturbations. Here, we provide an overview of genome-scale metabolic modeling and its current use to model human metabolism in health and disease. We then describe the initial steps made using it to study cancer metabolism and how it may be harnessed to enhance ongoing experimental efforts to identify drug targets and biomarkers for cancer in a rationale-based manner. *Clin Cancer Res; 18(20); 5572–84. ©2012 AACR.*

Introduction

Aberrant metabolism is one of the main driving forces in the initiation and development of cancer (1, 2). During carcinogenesis, selective pressures lead to diverse metabolic alterations, imposed by multiple molecular mechanisms (3–5). These metabolic adaptations enable the cancer cells not only to proliferate and cope with high energetic demands but also to avoid apoptosis, evade the immune system (6), and control the rate of mutagenesis (3, 7). Several metabolic abnormalities are quite general and have been observed in many cancer types. Notable among these is the preference to metabolize glucose by aerobic glycolysis (8, 9). This phenomenon, termed the Warburg effect, is accompanied by lactate production and increased glucose uptake. As proliferation requires a constant supply of macromolecular precursors that are generated in the tricarboxylic acid (TCA) cycle, cancer cells often use glutamine to replenish the cycle (anapleurosis). Glutamine has also been shown to support lipid synthesis in cancer cells through reductive carboxylation by the reverse activity of the TCA reaction isocitrate dehydrogenase (IDH; refs. 10, 11). Nonetheless, cancer metabolism is heterogeneous and reprogrammed in various ways. Mutations in several TCA metabolic enzymes promote specific types of cancer: Loss-of function mutation in fumarate hydratase (FH) causes leiomyoma, leiomyosarcoma, or renal cell carcinoma, whereas such mutations in succinate dehydrogenase lead to the development of paraganglioma or pheochromocytoma (12); on the other hand, gain-of-function mutations in IDH promote glioblastoma and acute myeloid leukemia (13, 14). It is yet to be elucidated why certain metabolic mutations lead to one type of cancer and not another and how metabolism promotes cancer through its interactions with other cellular processes. However, these findings imply that characterizing the unique metabolic dependencies of different cancer cells can potentially pave the way toward the development of selective treatments and diagnostic tools (15, 16).

The rapid technologic advancements in obtaining high-throughput omics data, combined with the development of the metabolic modeling methodology, has recently enhanced our ability to study metabolism on a genome-wide scale. *In silico* metabolic modeling has been shown to provide an appropriate platform to address various research questions related to metabolism and predict an array of cellular metabolic phenotypes (17–25). Here, we describe how using and developing this paradigm to study cancer metabolism can elucidate the metabolic alterations that accompany cancer progression and aid in the identification of drug targets and metabolic biomarkers. Importantly, drugs that target metabolic enzymes are especially promising because metabolism is evolutionarily more conserved than other biologic processes that have been targeted in cancer, such as signaling (26). Therefore, cancer cells are less prone to evolve resistance to these drugs by developing alternative pathways.

Genome-Scale Metabolic Modeling

In silico models of metabolism are based upon a representation of metabolism as a network. Mathematical

Authors' Affiliations: [1]The Blavatnik School of Computer Science, and [2]The Sackler School of Medicine, Tel Aviv University, Israel

Corresponding Authors: Eytan Ruppin, The Blavatnik School of Computer Sciences & School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel. Phone: 97-23-6406528; Fax: 97-23-6409357; E-mail: ruppin@post.tau.ac.il; and Livnat Jerby. Phone: 97-23-6405378; E-mail: livnatje@post.tau.ac.il

doi: 10.1158/1078-0432.CCR-12-1856

DOI:10.1158/1078-0432.CCR-12-1856

modeling of cellular metabolism has been traditionally conducted via kinetic modeling techniques, operating based on a set of differential equations that describe the changes in metabolite concentrations over time (27). These models provide an informative dynamic description of metabolism. However, their scope is still limited to small-scale systems, as they require detailed information on kinetic constants and on enzyme and metabolite levels (28, 29). An alternative computational approach that has emerged in recent years, termed constraint-based modeling (CBM), bypasses this hurdle as it does not depend on detailed kinetic information. Instead, it accounts for a set of constraints that govern cellular metabolism: (i) the mass–balance constraints that maintain a constant concentration of inner-cellular metabolites; (ii) thermodynamic constraints that dictate reaction directionality; and (iii) enzyme capacity constraints that bind the maximal flux rate of the metabolic reactions. CBM can hence be applied to analyze genome-scale metabolic models (GSMM), which consist of a collection of metabolic reactions, including their stoichiometry, and an accompanying genes to proteins to reactions (gene–protein–reaction, or GPR) mapping (Fig. 1). The GPR mapping associates between metabolic reactions and the genes that encode their catalyzing enzymes. There are several types of GPR associations. For example, if a reaction is catalyzed by a protein complex, then its activity depends on the expression of all the genes that encode this complex. Conversely, a gene can encode a promiscuous enzyme that catalyzes different reactions; in this case, the expression of the gene will affect more than one reaction. There are also isozymes, which are different enzymes that catalyze the same reaction. GPR associations enable the mapping of transcriptomics or proteomics to the level of reactions. We refer to the latter as reaction expression, which reflects for each reaction, the expression of its enzymes or enzyme-encoding genes, if it is inferred from proteomics or transcriptomics, respectively.

Model reconstruction is often based on various data types, such as gene content and expression, protein abundance, metabolomics, and fluxomics (i.e., flux rate measurements; see text under "GSMM of Cancer and Drug Target Identification" heading). Each reconstruction provides a complementary source of evidence that can be prioritized according to its accuracy and proximity to the metabolic phenotype. The quality of the GSMM depends on the data that have been used to construct it, the methodology by which it has been constructed, and the level of manual curation it has been subject to. It can be examined by its ability to capture known metabolic functionalities and recapitulate experimental results. Following its validation, a GSMM can be used to explore the metabolic state under different conditions via CBM methods. To do so, additional optimization criteria, referred to as the objective functions, may be used to determine the pertinent metabolic phenotype more accurately. A frequently used objective function when simulating proliferating cells is the maximization of biomass production (a close proxy of cellular growth or metabolic yield), as done in flux balance analysis (FBA), in which only metabolic states with maximal biomass production are considered (30). Another approach is to identify metabolic states that maximize the fit to experimental data. Measurements that are closer to the metabolic phenotype, such as fluxomics and metabolomics, are preferable. However, the former are rather scarce, small-scale, and are taken mostly from cell lines. The latter require inferring the effects metabolite concentrations have on enzyme activity by incorporating the measurements in kinetic rate equations or by accounting for thermodynamic principles (31). Transcriptomics and proteomics, which are becoming increasingly more accurate and accessible, can also provide important insights into the regulation of metabolic flux. Assuming that there is some correlation between mRNA or protein abundance and flux rates, one can constrain the model to account for these dependencies, for example, by mapping the data to the level of reactions (based on the GPR associations) and constraining as many of the lowly expressed reactions to be inactive and vice versa (ref. 32; Fig. 1).

GSMMs are hence a platform to integrate and bridge between different data sources, based on the well-established biochemical knowledge and principles they store. This platform makes it possible to infer the production, secretion, and uptake rates of different metabolites; to determine which reactions are active or inactive; assess reaction rates; and to determine gene and enzyme essentiality for proliferation or survival. By incorporating gene expression data, GSMMs can be used to identify reactions that have been subject to posttranscriptional regulation and specify whether their rate has been posttranscriptionally increased or decreased (32, 33). As further elaborated in the following sections, when experimental data are collected from 2 types of cells, GSMMs can be used to identify knockouts (KO) that will be lethal only to one of the cells or KOs that will transform the metabolism of one of the cells to be as akin as possible to that of the other, as done via metabolic transformation analysis (MTA; Table 1). Overall, there are by now more than a hundred different algorithmic approaches to build and analyze GSMMs (34), which have been applied to study the metabolism of hundreds of species. Several reviews describe the GSMM approach and its numerous applications in more length (34–37).

Metabolic Modeling of Human Metabolism

Initially, GSMMs have been extensively used to study the metabolism of bacteria, successfully addressing both basic scientific questions and applied research goals (17–19, 21, 46, 47). Eukaryotic and human modeling studies are now advancing at an accelerating pace (Table 2). Earlier network-level computational studies of human metabolism have focused on characterizing distinct human metabolic pathways and organelles (48–50). In 2007, 2 generic genome-scale human metabolic models were constructed, based on an extensive evaluation of genomic and bibliomic data: Recon1 (23) and the Edinburg Human Metabolic

DOI:10.1158/1078-0432.CCR-12-1856

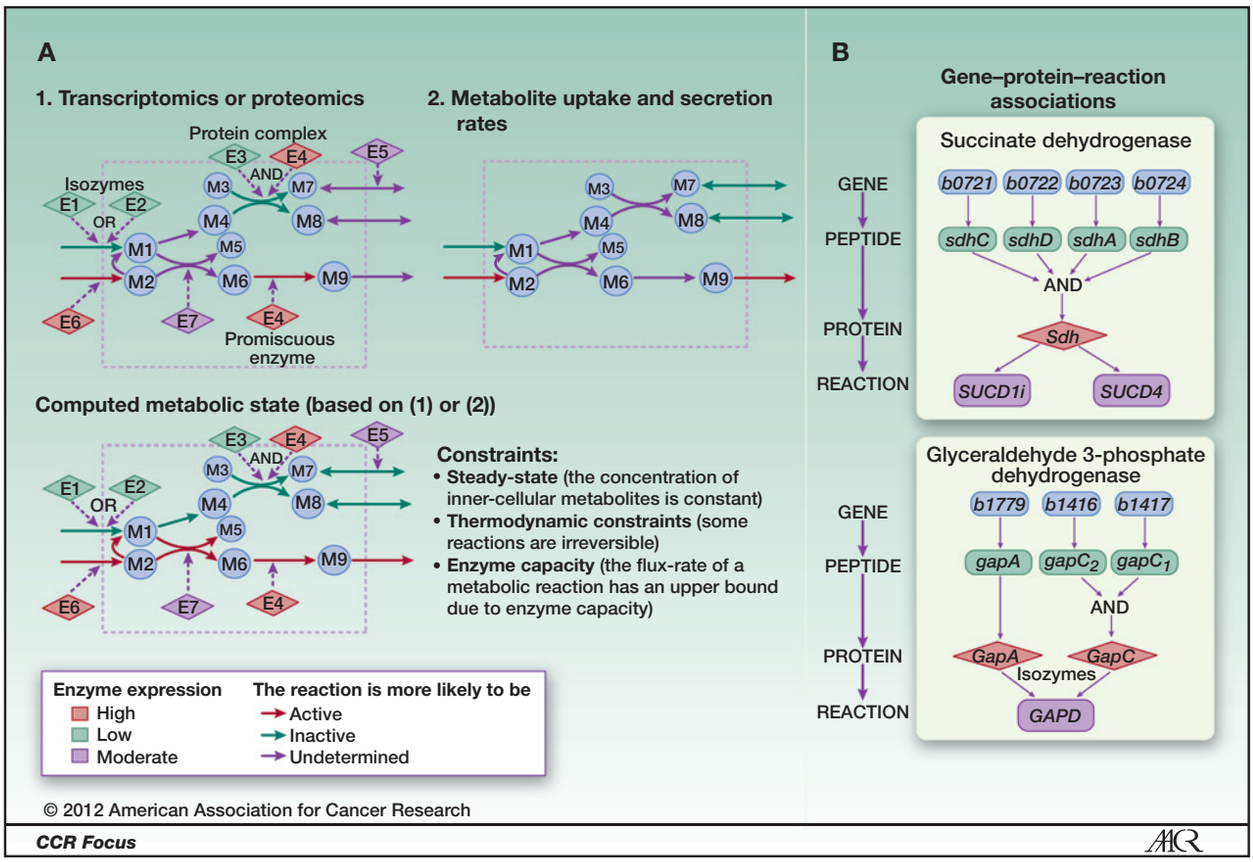

Figure 1. GSMM as a platform to integrate experimental data, with biochemical knowledge, and thermodynamic principles. A, schematic illustration of a metabolic network, using a toy example. Circular nodes represent metabolites, whereas diamond nodes represent enzymes. For enzymes, purple, red, and green represent moderate, significantly high, and significantly low expression of the enzyme, or enzyme-encoding genes, respectively. Solid edges represent metabolic reactions, and broken edges associate enzymes with the reactions they catalyze. To incorporate transcriptomics or proteomics data in the model, GPR associations are used to infer the expression state of the metabolic reactions. The latter, unlike direct metabolic flux measurements, only indicate the more likely activity state of the reactions, as the enzyme levels affect the metabolic flux. By accounting for additional constraints that govern cellular metabolism, the most probable metabolic state is deduced, estimating the flux rate and activity state of the metabolic reactions (reactions that are predicted as active are colored red, whereas those that are predicted to be inactive are colored green). Additional data types as flux rate measurements are also applicable to adjust the metabolic model. The measurements are mostly obtained for exchange reactions. These are reactions that transport metabolites in or out of the cell. B, an accompanying GPR mapping is included in the model, enabling one to simulate perturbations on both the genes and reactions levels, as desired, and map proteomics and transcriptomics to the reaction level. Examples of different types of GPR associations are shown, where the top level is the gene locus, the second level is the translated peptide, the third level is the functional protein, and the bottom level is the reaction.

Network (EHMN; ref. 51). These GSMMs consist of the biochemical reactions that are known to take place in different tissues and cell types in the human body. Recently, the human metabolic reaction (HMR) database has been published (38), containing elements of previously published generic genome-scale human metabolic models (23, 51, 52) and of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (53). The potential clinical use of Recon1 has been shown in numerous studies (refs. 23, 24, 54; Table 2). By incorporating high-throughput data, the generic GSMMs have been tailored, automatically and manually, to model different cells and tissues, including the liver (25, 55), kidney (56), brain (57), and the alveolar macrophage (58). Recently, a multi-tissue modeling approach was developed to simulate the metabolic interdependencies between the adipocytes, hepatocytes, and myocytes (59). To account for intertissue metabolic interactions, it is necessary to incorporate the different tissue-specific GSMMs into a unifying, multi-tissue one. However, this undertaking is still a serious open challenge (see Future Directions).

GSMM of Cancer and Drug Target Identification

The first steps in applying the CBM methodology to study cancer metabolism have naturally been made by developing small-scale metabolic models of central metabolism in cancer. Such a model has been shown to accurately predict growth rates measured in HeLa cells (61); it was then used to identify reactions with a strong influence on cancer cell growth. In consistency with the literature, it identified a set of pivotal drug targets,

DOI:10.1158/1078-0432.CCR-12-1856

Table 1. Constraint-based modeling methods for constructing GSMMs, integrating high-throughput data and identifying potential drug targets and biomarkers

Method	Data	Output		References
Model construction				
MBA or INIT	Different types of omics data	A context-specific GSMM		(25, 38)
GIMME	Gene or protein expression and objective function(s)			(39)
PRIME	Gene expression and phenotypic data (e.g., growth rates) of different samples with subtle transcriptomics differences	A context-specific GSMM for each sample		Yizhak et al. (unpublished data)
Inferring context-dependent metabolic states				
iMAT	Gene or protein expression	A GSMM with context-specific constraints	The number of highly (lowly) expressed reactions that are active (inactive) is maximized	(32)
E-flux	Gene expression		The flux magnitude of every reaction is adjusted on the basis of its expression level	(40)
QP	Fluxomics		The flux magnitude of every reaction is set on the basis of its measured flux rate	
MADE	Gene or protein expression from a series of time points along a metabolic transformation	The metabolic state at each time point is inferred, such that the differences between the metabolic states along the transformation are as similar as possible to those observed in the data		(41)
Drug target identification				
FBA	Cancer GSMM, healthy tissue GSMM, objective function	Drug targets whose KO is selectively lethal		(30)
ROOM or MOMA	Cancer GSMM, healthy tissue GSMM	Candidate drug targets; the metabolic state inferred following the drug target KO is as close as possible to the state without the KO		(42, 43)
MTA	Gene expression of a source and a target state	Candidate drug targets whose KO can transform the source to the target		Yitzhak et al. (unpublished data)
Biomarker identification				
Flux variability analysis	Context-specific GSMM	Biomarker predictions based on changes in the upper and lower bounds of the metabolic reactions' flux rates		(44)
Sampling	Context-specific GSMM	Biomarker predictions based on changes in the distribution of the metabolic reactions' flux rates		(45)
MPA	Gene or protein expression of patients from 2 clinical groups	The metabolic phenotype of each sample is explored to predict potential biomarkers		(33)

Abbreviations: FVA, flux variability analysis; GIMME, gene inactivity moderated by metabolism and expression; iMAT, integrative metabolic analysis tool; INIT, integrative network inference for tissues; MADE, metabolic adjustment by differential expression; MBA, model building algorithm; MOMA, minimization of metabolic adjustment; MPA, metabolic phenotypic analysis; PRIME, personalized reconstructIon of metabolic models; QP, quadratic programming; ROOM, regulatory on/off minimization.

DOI:10.1158/1078-0432.CCR-12-1856

Table 2. Human GSMM studies

Tissue/cell type	Year	Construction method	Application	References
Generic				
Recon1	2007	Extensive evaluation of genomic and bibliomic data	Capture metabolic alterations in diabetes; identify reactions related to hemolytic anemia and potential drug targets for treating hypercholesterolemia	(23, 54)
EHMN	2007	Extensive evaluation of genomic and bibliomic data	Study disease–related genes	(51)
HMR	2012	Integration of previous generic reconstructions and KEGG with context-specific omics data	Reconstruction of cell-type and cancer-type–specific GSMMs	(38)
Tissue-specific				
Red blood cell	2002	Manual curation	Study metabolic loads in red blood cell metabolism	(48)
Mitochondria	2004	Manual curation	Study mitochondrial metabolism	(49)
Fibroblasts	2007	Manual curation	Study Leigh syndrome	
Hepatocytes (HepatoNet)	2010	Manual curation	Investigate hepatic enzyme deficiencies	(55)
Hepatocytes	2010	MBA	Predict flux and biomarkers of hepatic metabolic disorders	(25)
Kidney	2010	GIMME	Predict causal drug off-targets that impact renal function	(56)
Brain (3 neuron types and astrocytes)	2010	Manual curation	Study the role of cholinergic neurotransmission in Alzheimer disease	(57)
Alveolar macrophage	2010	Automated (GIMME, iMAT) and manual	Study host–pathogen interactions with *Mycobacterium tuberculosis*	(58)
Multitissue (hepatocytes, myocytes, and adipocytes)	2011	Automated (GIMME, iMAT) and manual	Study metabolic alterations in diabetes	(59)
Cancer				
Generic cancer	2011	MBA	Identify drug targets for cancer based on synthetic lethal interactions	(22)
FH-deficient renal cancer	2011	MBA	Identify drug targets based on synthetic lethal interaction with FH1	(60)
69 cell types 16 cancer types	2012	INIT	Identify key metabolic features of cancer cells	(38)

Abbreviations: GIMME, gene inactivity moderated by metabolism and expression; HMR, human metabolic reaction; iMAT, integrative metabolic analysis tool; INIT, integrative network inference for tissues; MBA, model building algorithm.

including lactate dehydrogenase and pyruvate dehydrogenase, consistent with the literature (61). In accordance with the Warburg effect, the analysis showed that at a fixed glucose uptake rate, a decrease in the rate of pyruvate dehydrogenase actually increased biomass production capacity. Several other CBM studies have been dedicated to elucidation of the presumably counterintuitive Warburg effect (62, 63). These studies maintained that although aerobic glycolysis is less efficient than mitochondrial respiration in terms of ATP yield per glucose uptake, it is more efficient in terms of the required solvent capacity. Hence, overall, the shift to anaerobic metabolism results in an increased biomass production and proliferation rate. This observation has been shown both by using a small-scale model of ATP production (62) and by using the generic GSMM Recon1 (63), accounting for

DOI:10.1158/1078-0432.CCR-12-1856

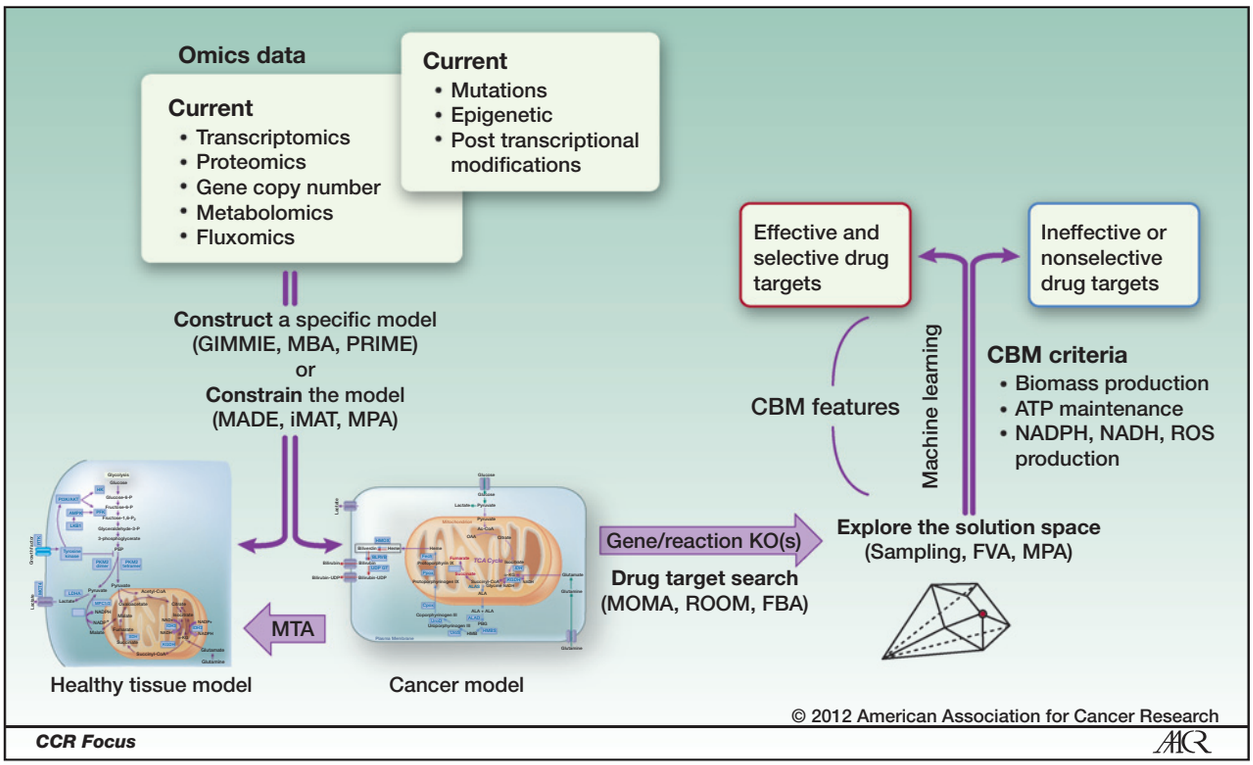

Figure 2. A pipeline for predicting drug targets via metabolic modeling. First, GSMMs that describe the healthy and cancerous tissue are obtained. Some tissue-specific models have been previously published and validated (Table 2) and could be used to describe the healthy tissue. Drug targets are then simulated by the inhibition of the pertinent reactions in the healthy and cancer models, and their functional effect on the cells' metabolic capabilities is computed. The latter can be used directly, or as features for machine-learning algorithms, to infer the effectiveness and selectivity of the drug target. FVA, flux variability analysis; GIMME, Gene inactivity moderated by metabolism and expression; iMAT, integrative metabolic analysis tool; MADE, metabolic adjustment by differential expression; MBA, model building algorithm; MOMA, minimization of metabolic adjustment; MPA, metabolic phenotypic analysis; PRIME, personalized reconstruction of metabolic models; QP, quadratic programming; ROOM, regulatory on/off minimization; ROS, reactive oxygen species.

stoichiometric and enzyme solvent capacity considerations. The latter study also captured a 3-phase metabolic behavior that has been observed experimentally during oncogenic progression and the high glutamine uptake of cancer cells.

To move toward a genome-scale investigation of cancer metabolism, the initial, yet crucial, step is to obtain a GSMM that depicts the metabolism of the tumor (Fig. 2). One approach to tackle this challenge is to apply model construction methods and use cancer-specific omics data to build a cancer GSMM that is then directly amenable to further intervention simulations (Table 1). Alternatively, a generic GSMM such as Recon1 can be adjusted to capture the metabolism of the tumor by requiring an optimal or suboptimal fit to pertinent experimental omics data (see methods for inferring context-dependent metabolic states in Table 1). The first approach yields a cancer model *per se* (a fixed subset of the human reactions that are active in cancer), whereas the second approach retains the global scope of the human generic model but specifies a set of metabolic states that best fit the cancer data. The second approach may be preferable as the construction or constraining of the model is based on data obtained under certain conditions, whereas the model is often used to estimate the metabolic state under different conditions, following some perturbation for example. To capture the cascade of changes that are triggered by perturbations, some limited deviation from the initial, unperturbed state is enabled. As the second approach still accounts for the initial state and the environmental conditions that nonetheless limit the set of possible metabolic modifications, it could potentially obtain more realistic estimators of new metabolic states.

Both descriptions can be used to identify potential drug target enzymes by simulating the effects of their inhibition (36, 64). The simulation is conducted by restricting the flux through the reactions that are catalyzed by the drug target and exploring the implications *in silico* by using CBM methods (30, 42–45). Obviously, candidate drug targets are those whose inhibition disrupts the viability of the cancer model, having as small an effect as possible on the viability and functionality of healthy tissue models. Viability is often estimated by the capacity to activate a set of essential metabolic functions (such as ATP and NADPH production). Drug target selectivity has also been pursued by targeting synthetic lethal genes (pairs of genes whose

DOI:10.1158/1078-0432.CCR-12-1856

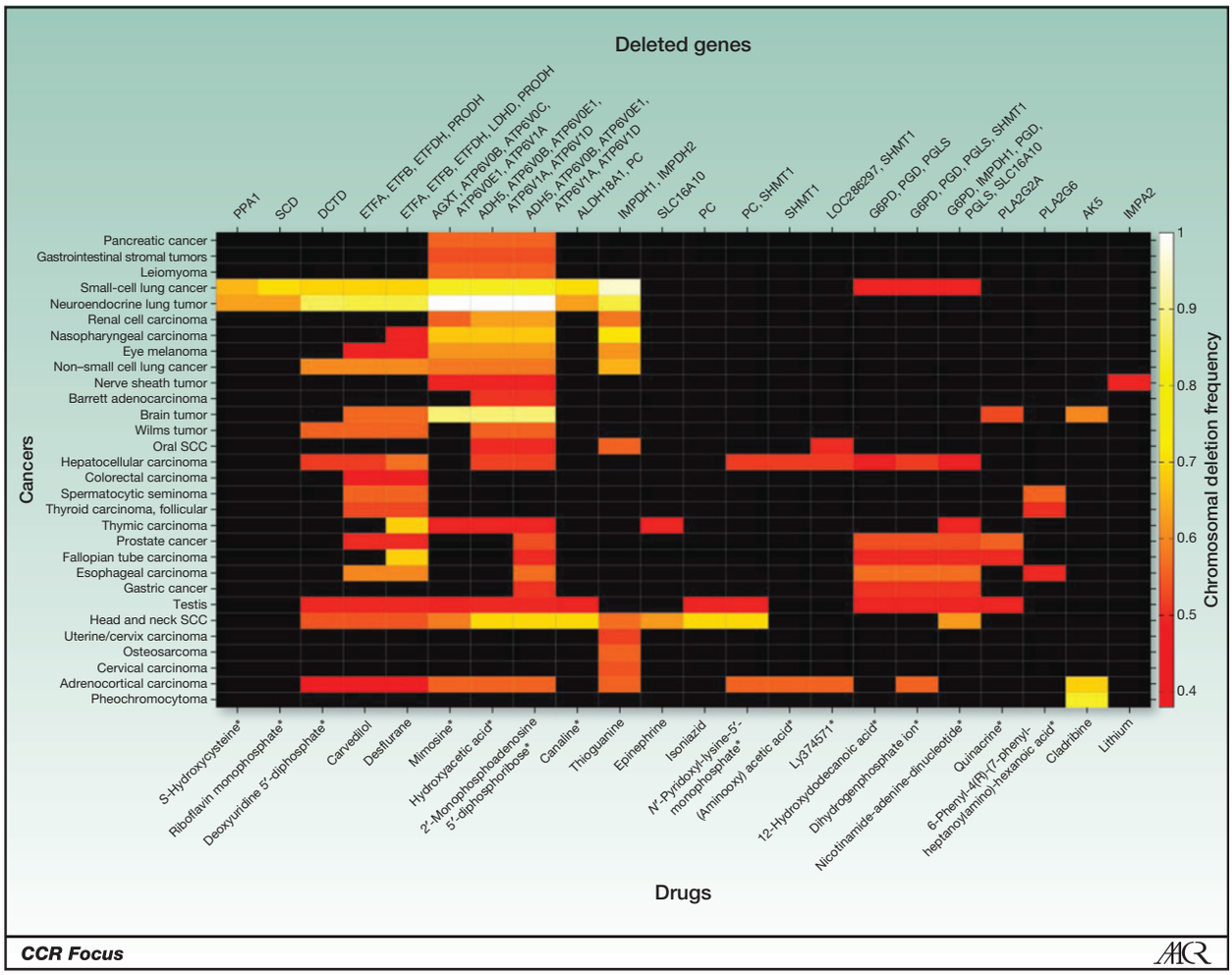

Figure 3. Potential drug targets identified in a previous GSMM study. Drugs predicted to target specific cancer types based on chromosomal loss of synthetic lethal participant genes. Cancer types that show a high frequency (in yellow and white) of chromosomal deletions of specific genes are susceptible to drugs inhibiting the genes' synthetic lethal complements. Experimental drugs are followed by an asterisk. SCC, squamous cell carcinoma. Adapted from Folger et al. (22).

combined but not individual KO is lethal; refs. 65–67). Because genetic and epigenetic mutations often silence the expression of specific genes exclusively in the cancer cells, targeting the remaining synthetic lethal pair gene of the inactivated gene(s) may selectively kill the cancer cells while sparing the healthy tissue, where the drug target gene has not lost its backup (60).

The first step in cancer genome-scale metabolic modeling was to develop a generic GSMM of cancer, aiming to capture the metabolic characteristics that are shared by different types of cancer (22). This conceptually parallels the first step done in human metabolic modeling, where generic models representing the collection of all human metabolic reactions have been constructed first. The model has been shown to correctly identify gene essentiality across an array of cancer cell lines and was then used to predict selective synthetic lethal gene pairs. The synthetic lethal predictions have been validated using drug efficacy and gene expression measurements across the NCI-60 cancer cell line collection. The synthetic lethal pairs were mapped to drug targets of approved drugs (not necessarily anticancer drugs) with known metabolic targets, and gene loss events that occur frequently in specific cancers involving these predicted pairs were identified. This combined analysis hence provided a set of cancer-specific selective drug target candidates (Fig. 3).

A type-specific model may generate more accurate drug target predictions for a specific type of cancer. Accordingly, a specific metabolic model of hereditary leiomyomatosis and renal cell cancer (HLRCC) has been constructed (60). HLRCC is caused by a germline mutation in the gene encoding FH, followed by a somatic mutation in its second allele. Through analysis of the specific metabolic model of the FH-deficient cells, the survival mechanism that enables the cells to operate the mitochondrial electron transport chain despite the mutation was unraveled. According to the computational predictions, the FH-deficient cells produce NADH, the driving force of the electron transport chain, by activating a linear metabolic pathway beginning with

DOI:10.1158/1078-0432.CCR-12-1856

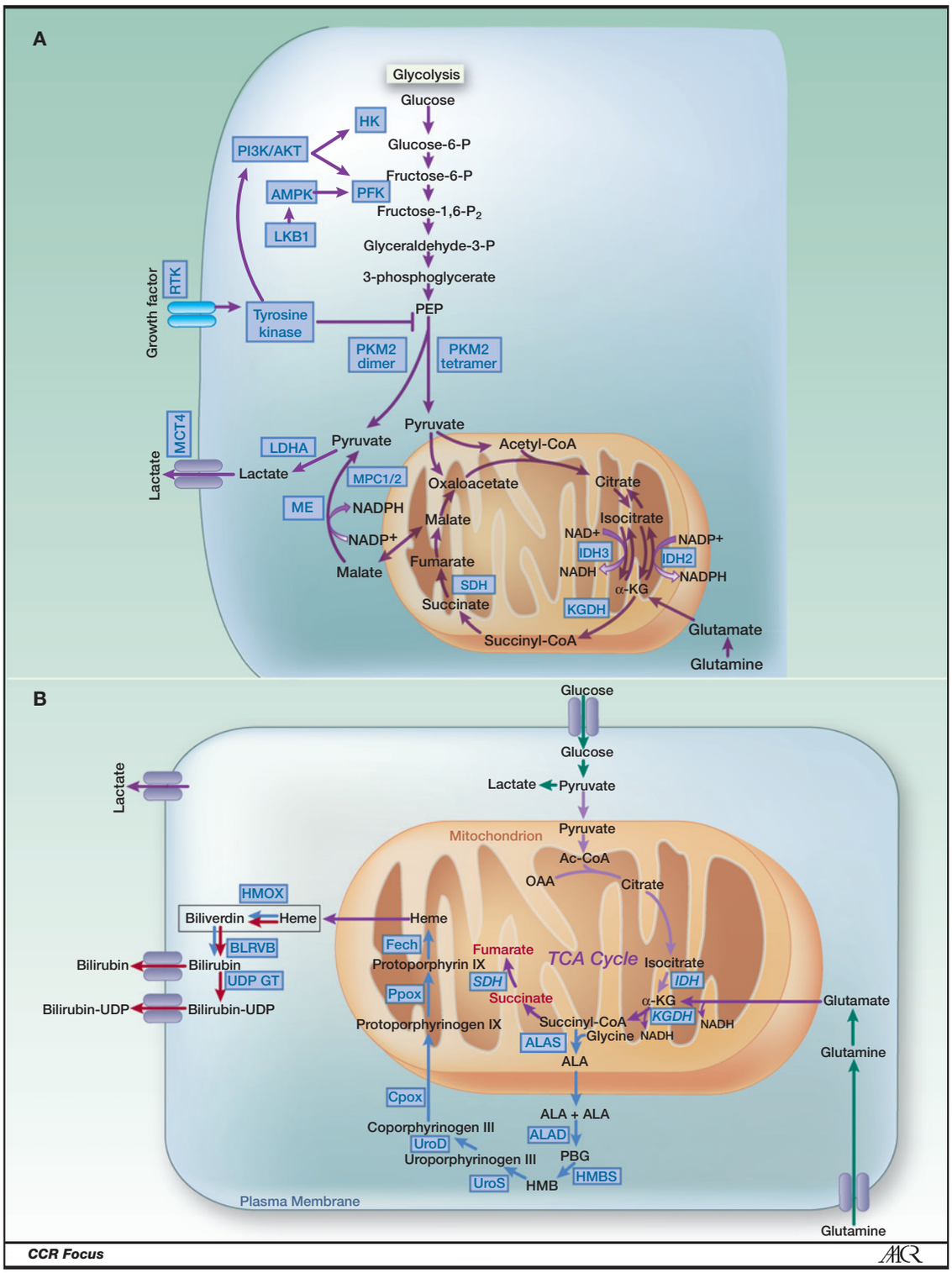

Figure 4. Drug targets identified in a GSMM study of renal cancer. Schematic representation of central metabolism of (A) healthy cells, as opposed to (B) the metabolic flux rearrangement observed in FH-deficient cells, based on model-driven predictions and experimental validations (59). Blue arrows indicate FH synthetic lethal metabolic reactions predicted by the metabolic model; red arrows indicate genes and reactions found to be upregulated in FH-deficient cells. The scheme also shows the truncation of the TCA cycle observed in FH-deficient cells. Fumarate and succinate are significantly accumulated (in red). The flux through the first part of the TCA cycle is reduced in FH-deficient cells due to decreased pyruvate entry and absence of recycling of metabolites through the TCA cycle. Glutamine uptake and glycolytic production of lactate (in green) are induced in FH-deficient cells. Figure 4B is adapted from Frezza et al. (60).

DOI:10.1158/1078-0432.CCR-12-1856

glutamine uptake and ending with bilirubin excretion (Fig. 4). This pathway, through the biosynthesis and degradation of heme, permits FH-deficient cells a partial mitochondrial NADH production, as it prevents the lethal accumulation of TCA cycle metabolites. In agreement, according to the model, numerous synthetic lethal pairs of FH are located along the heme biosynthesis pathway (Fig. 4). These synthetic lethal predictions have been confirmed experimentally *in vitro*, showing that targeting a key enzyme on this pathway (HMOX) renders only the FH-deficient cells nonviable, selectively sparing wild-type cells. This provides a new potential target for treating patients with HLRCC with a drug that is potentially selective and has minimal side effects on healthy renal cells.

Recently, an array of 69 normal and 16 cancer cell-type GSMMs has been automatically generated via the integrative network inference for tissues (iNIT) algorithm (38). The models have been constructed based on the HMR database, according to cell-type–specific protein abundances data obtained from the Human Proteome Atlas. Several metabolites, along with their associated reactions, have been found to appear significantly more often in the cancer models than in the healthy models. Among them are polyamines, isoprenoid, prostaglandins, and leukotrienes. These metabolites are tightly linked to oxidative stress, prenylation of oncogenes, and inflammation, respectively (68–70). Indeed attempts have been made to treat cancer by reducing the level of these metabolites in the tumor through inhibition of their production or uptake (71–73).

Combining additional computational methods with GSSM has been shown to further improve drug target predictions in a synergistic manner. It has been shown that applying structural bioinformatics methods to infer drug off-target enzymes with GSMM can aid in identifying drug side effects (56). This combined approach was applied to study the side effect of the drug torcetrapib in the context of renal function. Torcetrapib was developed to treat cardiovascular diseases and was withdrawn from phase III clinical trials due to its observed side effect of fatal hypertension. A metabolic kidney model was generated in which torcetrapib treatment was simulated by accounting for both its main target and off-targets. The latter were predicted to bind the drug based on the structure of their ligand-binding sites. Based on this simulation, causal drug off-targets were predicted, capturing the observed implications of the drug in patients with renal disorders.

Machine learning approaches can also be used to integrate CBM-based and other important characteristics of metabolic enzymes to determine their potential as drug targets. This approach has been shown by predicting new targets for approved anticancer drugs based on their enzyme structure and their cell line–specific flux state across the NCI-60 cell lines (74). First, a drug reaction network was constructed, providing a global view of drug reaction and drug pathway interactions. Then, 2 metrics of similarities between reactions were developed and used: structural similarity based on the structure of the enzymes that catalyze the reactions, and a functional similarity, computed according to the flux state of the reactions in each of the NCI-60 cell lines (the latter was predicted via a GSMM, given the cell lines' gene expression). Integrating these 2 similarity metrics to predict drug targets for approved cancer drugs yielded fairly accurate prediction performance (with an area under the curve of 0.92) and novel predictions. The same approach can be used to predict the anticancer effect of other approved drugs (not necessarily anticancer ones) based on the similarity of their targets to the targets of anticancer drugs.

Identification of Cancer Biomarkers via Metabolic Modeling

The aberrant metabolism of tumors enables their diagnosis by detecting increased glucose uptake via F-deoxyglucose positron emission tomography (PET). However, the differential uptake of other metabolites such as ^{11}C-choline, ^{11}C-acetate, ^{11}C-methionine, and ^{18}F-labeled amino acid analogues, was shown in some human cancers, testifying to the heterogeneity of cancer metabolism. A pending challenge in cancer diagnosis is the identification of metabolic biomarkers in the biofluids, forming noninvasive, cost-effective means for early diagnosis and monitoring treatment efficiency (75, 76).

The first GSMM method for predicting biomarkers was applied to predict biomarkers for inborn errors of metabolism (IEM), showing a fairly accurate level of prediction (24). However, its applicability is limited to the realm of IEMs, where the loss of functionality of specific metabolic genes can be simulated via *in silico* KOs. Identifying biomarkers for diseases such as cancer, where the metabolic rerouting results from more elaborate genetic and epigenetic alterations, is more complex. As described above, the modeling of cancer metabolism is based on integrating pertinent high-throughput data within the model. These methods can be used to infer cancer biomarkers by incorporating gene expression data of clinical samples in the model and inferring the exchange rates of the different metabolites for each individual sample (Fig. 5). Metabolites that significantly differentiated between 2 clinical groups of interest are then marked as candidate biomarkers. We have recently applied this approach by utilizing a new method. The method, "Metabolic Phenotypic Analysis" (MPA), gauges the adaptive potential of cells to produce metabolites of essence in a given context (33). It was first validated by predicting amino acid biomarkers for breast cancer and confirming them based on measured plasma-free amino acid profiles of breast cancer patients and control subjects. It was then used to predict novel biomarkers for metastatic breast cancer, highlighting the potential role of choline-containing metabolites. Indeed, choline is a known potential PET marker for imaging breast cancer (77).

Future Directions

Despite the encouraging achievements of GSMMs, the approach has its caveats and limitations. First, the

DOI:10.1158/1078-0432.CCR-12-1856

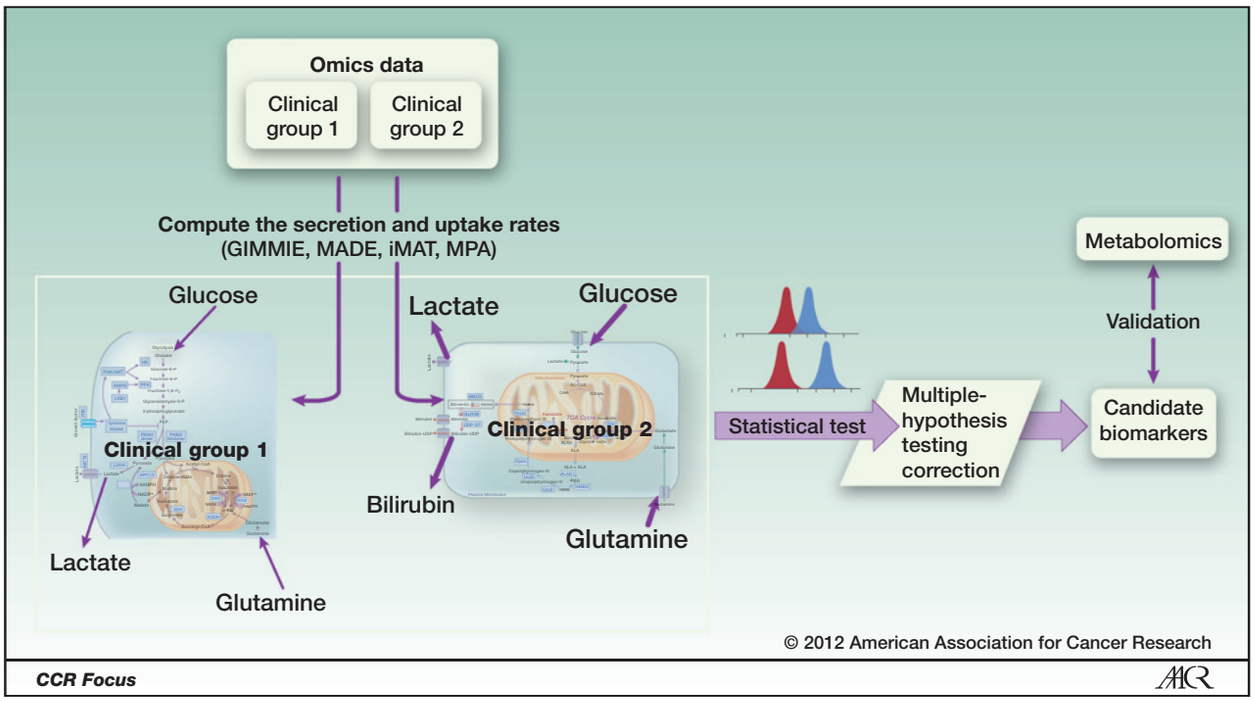

Figure 5. Identifying biomarkers via GSMM. Given high-throughput data of 2 clinical groups of interest (e.g., cancerous vs. healthy tissue or cancer tissues of different grading or staging), the rate of metabolite uptake and secretion can be computed for each individual sample. Metabolites whose computational net uptake or secretion rate significantly differentiates between the groups can then be selected as potential biomarkers, ranked by their predictive power. These predictions can be further tested and filtered on the basis of relevant metabolomics. Gene Inactivity Moderated by Metabolism and Expression (GIMME), integrative Metabolic Analysis Tool (iMAT), Metabolic Adjustment by Differential Expression (MADE).

curation and testing of metabolic models is far more complex when it comes to multicellular organisms. Unlike bacteria, in which *in silico* simulations can be directly compared with genome-scale experiments, the ability to test tissue-specific models is more qualitative and requires *in vivo* experimental systems. This limitation is somewhat alleviated when studying cancer metabolism, as the cancer cells can be grown *in vitro*, and genome-scale experiments, as those measuring drug efficacy and gene essentiality across numerous cell lines and conditions are available to calibrate the model. Second, current GSMMs describe metabolism as operating independently of other cellular systems. Constructing unified models that account for the interactions of metabolism with other cellular processes such as transcriptional regulation and signaling remains a cardinal, nontrivial challenge. This challenge has been addressed by integrating GSMMs with regulatory and signaling networks in microorganisms (46, 78, 79). However, to date, these methods have not been applied for human metabolism, mainly due to the lack of sufficient biologic data. Third, the incorporation of omics data in GSMMs is pivotal. However, it is often done by estimating the most probable connection between gene, protein, and flux rate. Various CBM methods differ in the type of gene-to-protein-to-flux rate connections they assume. Integrative experimental measurements of these hierarchical regulatory levels in unison under various conditions are required to rigorously substantiate our understanding and ability to deduce metabolic flux from gene or protein expression. Although these types of studies have been done in *Escherichia coli* (80), they have not yet been conducted in human cells.

More work needs to be done to fully exploit GSMMs to study human metabolism in general and cancer metabolism in particular. As cancer metabolism is heterogeneous, more personalized approaches are required to model it. Recently, we addressed this task in 2 ways. By applying MPA, we described the metabolic state of different patients with breast cancer, providing a system-level view of generic and subtype-specific metabolic characteristics of breast cancer (33). We used MPA to assess growth rates, lipid production capacities, posttranscriptional regulation, and metabolic biomarkers in breast cancer, obtaining highly accurate results. However, MPA and other CBM methods are inapplicable when the similarity in expression patterns between samples is high, as they define the metabolic reactions as active or inactive. To account for more subtle differences we developed personalized reconstructIon of metabolic models (PRIME; Yizhak et al., unpublished data). PRIME integrates individual gene expression and phenotypic data (e.g., growth rates) within a generic human model to generate a tailor-made model for each sample by varying the reactions' bounds, rather than excluding them from the model.

DOI:10.1158/1078-0432.CCR-12-1856

Once tissue-specific GSMMs will be sufficiently accurate and applicable, the next challenging and worthy endeavor is the development of a multitissue GSMMs. Such a model could be used to model the tumor in the context of whole-body physiology. It could improve both drug target and biomarker identification by accounting for intertissue effects and identifying biomarkers in a biofluid-specific manner. Further advancements can also be obtained by combining GSMMs with machine-learning techniques, structural biology tools, and genomic and epigenetic information. For example, cancer loss-of-function mutations in metabolic genes can be used to identify their synthetic lethal pairs as selective drug targets. An alternative approach for identifying candidate drugs in cancer is to seek drug targets whose targeting would not necessarily kill the cancer cells but would, instead, work to transform their metabolism back to a nonproliferative, noncancerous state. Such methods could be applied to reverse the Warburg effect, as has already been attempted experimentally (81, 82). Finally, all GSMM methods described here are currently restricted to identification of drug targets that are targeted by enzyme inhibition. Because many drugs act by augmenting the activity of different enzymes, developing next-generation GSMM methods for predicting the outcome of enzyme overactivity is required. In summary, given the current status of genome-scale metabolic modeling and the perspectives of upcoming developments, this approach shows promise for enhancing the identification of drug targets and biomarkers in a rationale-based manner.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: L. Jerby, E. Ruppin
Writing, review, and/or revision of the manuscript: L. Jerby, E. Ruppin
Study supervision: E. Ruppin

Grant Support

L. Jerby is funded by the Dan David Foundation and the Edmond J. Safra Bioinformatics Center. E. Ruppin's research is supported by the Israeli Centers of Research Excellence, Gene Regulation in Complex Human Disease Center (41/11), and by grants from the Israeli Science Foundation and the Israeli Cancer Research Fund.

Received June 6, 2012; revised August 10, 2012; accepted August 28, 2012; published online October 15, 2012.

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646–74.
2. Ward PS, Thompson CB. Metabolic reprogramming: a cancer hallmark even Warburg did not anticipate. Cancer Cell 2012;21:297–308.
3. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer 2011;11:85–95.
4. Tamada M, Suematsu M, Saya H. Pyruvate kinase M2: multiple faces for conferring benefits on cancer cells. Clin Cancer Res 2012;18:5554–61.
5. Miller DM, Thomas SD, Islam A, Sedoris K. c-Myc and cancer metabolism. Clin Cancer Res 2012;18:5546–53.
6. Prendergast GC. Cancer: why tumours eat tryptophan. Nature 2011;478:192–4.
7. Sotgia F, Martinez-Outschoorn U, Lisanti M. Mitochondrial oxidative stress drives tumor progression and metastasis: should we use antioxidants as a key component of cancer treatment and prevention? BMC Med 2011;9:62.
8. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 2009;324:1029–33.
9. Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. Cell 2008;134:703–7.
10. Mullen AR, Wheaton WW, Jin ES, Chen P-H, Sullivan LB, Cheng T, et al. Reductive carboxylation supports growth in tumour cells with defective mitochondria. Nature 2012;481:385–8.
11. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature 2011;481:380–4.
12. King A, Selak MA, Gottlieb E. Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. Oncogene 2006;25:4675–82.
13. Dang L, Jin S, Su SM. IDH mutations in glioma and acute myeloid leukemia. Trends Mol Med 2010;16:387–97.
14. Yang H, Ye D, Guan KL, Xiong Y. *IDH1* and *IDH2* mutations in tumorigenesis: mechanistic insights and clinical perspectives. Clin Cancer Res 2012;18:5562–71.
15. Vander Heiden MG. Targeting cancer metabolism: a therapeutic window opens. Nat Rev Drug Discov 2011;10:671–84.
16. Meijer TWH, Kaanders JHAM, Span PN, Bussink J. Targeting hypoxia, HIF-1 and tumor glucose metabolism to improve radiotherapy efficacy. Clin Cancer Res 2012;18:5585–94.
17. Papp B, Notebaart RA, Pál C. Systems-biology approaches for predicting genomic evolution. Nat Rev Genet 2011;12:591–602.
18. Wessely F, Bartl M, Guthke R, Li P, Schuster S, Kaleta C. Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. Mol Syst Biol 2011;7:515.
19. Trawick JD, Schilling CH. Use of constraint-based modeling for the prediction and validation of antimicrobial targets. Biochem Pharmacol 2006;71:1026–35.
20. Deutscher D, Meilijson I, Kupiec M, Ruppin E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. Nat Genet 2006;38:993–8.
21. Burgard AP, Pharkya P, Maranas CD. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol Bioeng 2003;84:647–57.
22. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting selective drug targets in cancer through metabolic networks. Mol Syst Biol 2011;7:501.
23. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, et al. Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc Natl Acad Sci U S A 2007;104:1777–82.
24. Shlomi T, Cabili MN, Ruppin E. Predicting metabolic biomarkers of human inborn errors of metabolism. Mol Syst Biol 2009;5:263.
25. Jerby L, Shlomi T, Ruppin E. Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. Mol Syst Biol 2010;6:401.
26. Locasale Jason W, Cantley Lewis C. Metabolic flux and the regulation of mammalian cell growth. Cell Metab 2011;14:443–51.
27. Garfinkel D, Hess B. Metabolic control mechanisms. VII.A detailed computer model of the glycolytic pathway in ascites cells. J Biol Chem 1964;239:971–83.
28. Lee I-D, Palsson BO. A comprehensive model of human erythrocyte metabolism: extensions to include pH effects. Biomed Biochim Acta 1991;49:771–89.
29. Bakker B, van Eunen K, Jeneson JA, van Riel NA, Bruggeman FJ, Teusink B. Systems biology from micro-organisms to human

DOI:10.1158/1078-0432.CCR-12-1856

metabolic diseases: the role of detailed kinetic models. Biochem Soc Trans 2010;38:1294–301.
30. Varma A, Palsson BO. Metabolic flux balancing: basic concepts, scientific and practical use. Nature Biotechnol 1994;12:994–8.
31. Yizhak K, Benyamini T, Liebermeister W, Ruppin E, Shlomi T. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. Bioinformatics 2010;26:i255–60.
32. Shlomi T, Cabili MN, Herrgard MJ, Palsson BO, Ruppin E. Network-based prediction of human tissue-specific metabolism. Nat Biotechnol 2008;26:1003–10.
33. Jerby L, Wolf L, Denkert C, Stein GY, Hilvo M, Oresic M, et al. Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. Cancer Res 2012 Sep 17. [Epub ahead of print].
34. Lewis NE, Nagarajan H, Palsson BO. Constraining the metabolic genotype–phenotype relationship using a phylogeny of *in silico* methods. Nat Rev Microbiol 2012;10:291–305.
35. Ruppin E, Papin JA, de Figueiredo LF, Schuster S. Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. Curr Opin Biotechnol 2010;21:502–10.
36. Bordbar A, Palsson BO. Using the reconstructed genome-scale human metabolic network to study physiology and pathology. J Intern Med 2012;271:131–41.
37. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nat Biotechnol 2010;28:245–8.
38. Agren R, Bordel S, Mardinoglu A, Pornputtapong N, Nookaew I, Nielsen J. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. PLoS Comput Biol 2012;8:e1002518.
39. Becker SA, Palsson BO. Context-specific metabolic networks are consistent with experiments. PLoS Comput Biol 2008;4:e1000082.
40. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, et al. Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. PLoS Comput Biol 2009;5:e1000489.
41. Jensen PA, Papin JA. Functional integration of a metabolic network model and expression data without arbitrary thresholding. Bioinformatics 2011;27:541–7.
42. Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proc Natl Acad Sci U S A 2005;102:7695–700.
43. Segrè D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. Proc Natl Acad Sci U S A 2002;99:15112–7.
44. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng 2003;5:264–76.
45. Schellenberger J, Palsson BØ. Use of randomized sampling for analysis of metabolic networks. J Biol Chem 2009;284:5457–61.
46. Chandrasekaran S, Price ND. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 2010;107:17845–50.
47. Heinemann M, Sauer U. Systems biology of microbial metabolism. Curr Opin Microbiol 2010;13:337–43.
48. Wiback SJ, Palsson BO. Extreme pathway analysis of human red blood cell metabolism. Biophys J 2002;83:808–18.
49. Vo TD, Greenberg HJ, Palsson BO. Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. J Biol Chem 2004;279:39532–40.
50. Chatziioannou A, Palaiologos G, Kolisis FN. Metabolic flux analysis as a tool for the elucidation of the metabolism of neurotransmitter glutamate. Metab Eng 2003;5:201–10.
51. Ma H, Sorokin A, Mazein A, Selkov A, Selkov E, Demin O, et al. The Edinburgh human metabolic network reconstruction and its functional analysis. Mol Syst Biol 2007;3:135.
52. Romero P, Wagg J, Green M, Kaiser D, Krummenacker M, Karp P. Computational prediction of human metabolic pathways from the complete human genome. Genome Biol 2004;6:R2.
53. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008;36 Suppl 1:D480–4.
54. Zelezniak A, Pers TH, Soares So, Patti ME, Patil KR. Metabolic network topology reveals transcriptional regulatory signatures of type 2 diabetes. PLoS Comput Biol 2010;6:e1000729.
55. Gille C, Bolling C, Hoppe A, Bulik S, Hoffmann S, Hubner K, et al. HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. Mol Syst Biol 2010;6:411.
56. Chang RL, Xie L, Xie L, Bourne PE, Palsson BØ. Drug off-target effects predicted using structural analysis in the context of a metabolic network model. PLoS Comput Biol 2010;6:e1000938.
57. Lewis NE, Schramm G, Bordbar A, Schellenberger J, Andersen MP, Cheng JK, et al. Large-scale *in silico* modeling of metabolic interactions between cell types in the human brain. Nat Biotechnol 2010;28:1279–85.
58. Bordbar A, Lewis NE, Schellenberger J, Palsson BO, Jamshidi N. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. Mol Syst Biol 2010;6:422.
59. Bordbar A, Feist A, Usaite-Black R, Woodcock J, Palsson B, Famili I. A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. BMC Syst Biol 2011;5:180.
60. Frezza C, Zheng L, Folger O, Rajagopalan KN, MacKenzie ED, Jerby L, et al. Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. Nature 2011;477:225–8.
61. Resendis-Antonio O, Checa A, Encarnación S. Modeling core metabolism in cancer cells: surveying the topology underlying the Warburg effect. PLoS One 2010;5:e12383.
62. Vazquez A, Liu J, Zhou Y, Oltvai Z. Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited. BMC Syst Biol 2010;4:58.
63. Shlomi T, Benyamini T, Gottlieb E, Sharan R, Ruppin E. Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. PLoS Comput Biol 2011;7:e1002018.
64. Kim HU, Sohn SB, Lee SY. Metabolic network modeling and simulation for drug targeting and discovery. Biotechnol J 2012;7:330–42.
65. Hartwell LH, Szankasi P, Roberts CJ, Murray AW, Friend SH. Integrating genetic approaches into the discovery of anticancer drugs. Science 1997;278:1064–8.
66. Hartman JL, Garvik B, Hartwell L. Principles for the buffering of genetic variation. Science 2001;291:1001–4.
67. Kaelin WG. The concept of synthetic lethality in the context of anticancer therapy. Nat Rev Cancer 2005;5:689–98.
68. Eisenberg T, Knauer H, Schauer A, Buttner S, Ruckenstuhl C, Carmona-Gutierrez D, et al. Induction of autophagy by spermidine promotes longevity. Nat Cell Biol 2009;11:1305–14.
69. Sebti SM, Hamilton AD. Farnesyltransferase and geranylgeranyltransferase I inhibitors and cancer therapy: lessons from mechanism and bench-to-bedside translational studies. Oncogene 2000;19:6584–93.
70. Schneider C, Pozzi A. Cyclooxygenases and lipoxygenases in cancer. Cancer Metastasis Rev 2011;30:277–94.
71. Seiler N. Thirty years of polyamine-related approaches to cancer therapy. Retrospect and prospect. Part 2. Structural analogues and derivatives. Curr Drug Targets 2003;4:565–85.
72. Dudakovic A, Tong H, Hohl R. Geranylgeranyl diphosphate depletion inhibits breast cancer cell migration. Invest New Drugs 2011;29:912–20.
73. Eruslanov E, Kaliberov S, Daurkin I, Kaliberova L, Buchsbaum D, Vieweg J, et al. Altered expression of 15-hydroxyprostaglandin dehydrogenase in tumor-infiltrated CD11b myeloid cells: a mechanism for immune evasion in cancer. J Immunol 2009;182:7548–57.
74. Li L, Zhou X, Ching W-K, Wang P. Predicting enzyme targets for cancer drugs by profiling human Metabolic reactions in NCI-60 cell lines. BMC Bioinformatics 2010;11:501.
75. Davis VW, Bathe OF, Schiller DE, Slupsky CM, Sawyer MB. Metabolomics and surgical oncology: potential role for small molecule biomarkers. J Surg Oncol 2011;103:451–9.

DOI:10.1158/1078-0432.CCR-12-1856

76. Teicher BA, Linehan WM, Helman LJ. Targeting cancer metabolism. Clin Cancer Res 2012;18:5537–45.
77. Zheng Q-H, Stone KL, Mock BH, Miller KD, Fei X, Liu X, et al. [11C]choline as a potential PET marker for imaging of breast cancer athymic mice. Nucl Med Biol 2002;29:803–7.
78. Covert MW, Xiao N, Chen TJ, Karr JR. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. Bioinformatics 2008;24:2044–50.
79. Min Lee J, Gianchandani EP, Eddy JA, Papin JA. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. PLoS Comput Biol 2008;4:e1000086.
80. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, et al. Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. Science 2007;316:593–7.
81. Sutendra G, Michelakis ED. Reversing the Warburg effect: metabolic modulation as a novel cancer therapy mitochondria and cancer. [cited 2012 Sep 26]. Available at: http://stm.sciencemag.org/content/3/94/94ra70.abstract.
82. Chan DA, Sutphin PD, Nguyen P, Turcotte S, Lai EW, Banh A, et al. Targeting GLUT1 and the Warburg effect in renal cell carcinoma by chemical synthetic lethality. Sci Transl Med 2011;3:94ra70.