Synthetic dosage lethality in the human metabolic network is highly predictive of tumor growth and cancer patient survival

Wout Megchelenbrink[a,b,c,1], Rotem Katzir[d,1], Xiaowen Lu[b], Eytan Ruppin[d,e,2,3], and Richard A. Notebaart[f,2,3]

[a]Institute for Computing and Information Science, Radboud University, Nijmegen, 6525 EC Nijmegen, The Netherlands; [b]Centre for Molecular and Biomolecular Informatics, Radboud University Medical Center, 6525 GA, Nijmegen, The Netherlands; [c]Centre for Systems Biology and Bioenergetics, Radboud University Medical Center, 6525 GA, Nijmegen, The Netherlands; [d]Center for Bioinformatics and Computational Biology, Department of Computer Science, University of Maryland, College Park, MD 20742; [e]School of Computer Science and School of Medicine, Tel-Aviv University, Tel Aviv, Israel, 69978; and [f]Department of Internal Medicine, Radboud University Medical Center, 6525 GA, Nijmegen, The Netherlands

Edited by Stanley Fields, Howard Hughes Medical Institute, University of Washington, Seattle, WA, and approved August 3, 2015 (received for review May 1, 2015)

Synthetic dosage lethality (SDL) denotes a genetic interaction between two genes whereby the underexpression of gene A combined with the overexpression of gene B is lethal. SDLs offer a promising way to kill cancer cells by inhibiting the activity of SDL partners of activated oncogenes in tumors, which are often difficult to target directly. As experimental genome-wide SDL screens are still scarce, here we introduce a network-level computational modeling framework that quantitatively predicts human SDLs in metabolism. For each enzyme pair (A, B) we systematically knock out the flux through A combined with a stepwise flux increase through B and search for pairs that reduce cellular growth more than when either enzyme is perturbed individually. The predictive signal of the emerging network of 12,000 SDLs is demonstrated in five different ways. (*i*) It can be successfully used to predict gene essentiality in shRNA cancer cell line screens. Moving to clinical tumors, we show that (*ii*) SDLs are significantly underrepresented in tumors. Furthermore, breast cancer tumors with SDLs active (*iii*) have smaller sizes and (*iv*) result in increased patient survival, indicating that activation of SDLs increases cancer vulnerability. Finally, (*v*) patient survival improves when multiple SDLs are present, pointing to a cumulative effect. This study lays the basis for quantitative identification of cancer SDLs in a model-based mechanistic manner. The approach presented can be used to identify SDLs in species and cell types in which "omics" data necessary for data-driven identification are missing.

systems biology | synthetic dosage lethality | human metabolism | cancer | genetic interactions

Synthetic lethality (SL) occurs when the combined loss of two nonessential genes renders a lethal phenotype (1). SLs have been studied by using experimental (2, 3) and computational approaches (4–6) to address various questions of cell function and evolution. The potential of SLs for cancer therapy has been recognized and accelerated the development of many SL screens (7–11). (See refs. 12–14 for reviews of SLs applied in the context of cancer research.)

Less studied are the so-called synthetic dosage lethality (SDL) interactions. An SDL is a genetic interaction between two genes whereby the underexpression of gene A ($A^{\downarrow}$) together with the overexpression of gene B ($B^{\uparrow}$) is lethal (15). The observation that an interaction with an overexpressed gene can be lethal makes it particularly interesting for targeting cancer cells with (over) expressed oncogenes. This is because many oncogenes that drive tumor growth are essential to cell function and thus difficult to target directly. Targeting the oncogenes' SDL partner, which is a nonessential gene in normal cells, may nevertheless kill cancer cells. That SDLs can have important implications for cancer research, for instance to aid in the design of new therapies, has also been recognized (12, 16–18). Moreover, it has been shown that the overexpression of specific genes can be detrimental to cancer cell growth (19). Recently, a data-mining approach was used that identifies SLs and SDLs by analyzing large volumes of cancer genomic data (20). Here we aim to complement data-driven computational efforts with a biological network model approach to identify SDLs. This has recently become feasible in the realm of metabolism, with the advent of genome-scale metabolic modeling. We introduce a method that uses a constraint-based genome-scale model of metabolism (GSMM) (21–25) to predict metabolic SDLs. GSMMs have successfully resolved a wide range of research questions in model organisms (23, 26–30) and have been the basis for many computational studies of cancer (7, 8, 31–34). Furthermore, they have contributed to a systematic understanding of the underlying mechanisms leading to lethality and SL (3–7). A major advantage of a model-based approach is that it can provide insights into the underlying network mechanisms causing SDLs. Furthermore, the modeling approach presented is general and can be used to identify SDLs in species and cell types in which "omics" data are missing.

We introduce a computational approach for identifying dosage lethality effects (IDLE) in metabolism. IDLE predicts enzymatic SDLs from a GSMM with application to cancer. For each enzyme pair (A, B) in the human GSMM, we systematically

Significance

Synthetic dosage lethality (SDL) denotes a genetic interaction whereby an underexpression of gene A combined with an overexpression of gene B kills the cell. Although many overexpressed oncogenes driving tumor growth are difficult to target directly, targeting their SDL partners may kill cancer cells. We present what is, to our knowledge, the first network-level modeling approach that is able to predict metabolic SDLs. As expected, we find that the predicted SDLs are less frequently active in tumors to avoid lethality. Cancer tumors with more and stronger SDLs have smaller tumor size and lead to increased patient survival. Beyond facilitating the development of novel anticancer therapies, model-based identification of metabolic SDLs can be used to model pathogenic bacteria and provide leads to new antibiotic targets.

Author contributions: W.M., E.R., and R.A.N. designed research; W.M. and R.K. performed research; W.M., R.K., and X.L. analyzed data; R.K. and X.L. contributed to method development; and W.M., E.R., and R.A.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]W.M. and R.K. contributed equally to this work.

[2]E.R. and R.A.N. contributed equally to this work.

[3]To whom correspondence may be addressed. Email: ruppin@post.tau.ac.il or richard.notebaart@radboudumc.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508573112/-/DCSupplemental.

knock out the enzyme flux through A combined with a stepwise flux increase through enzyme B and quantify the level of growth reduction. Pairs in which the growth is significantly more reduced than when either enzyme is perturbed individually are ranked as SDLs ($A^{\downarrow}$, $B^{\uparrow}$) with a corresponding value of "strength."

We demonstrate the predictive power of our approach in five different ways: First, by analyzing genome-wide experimental shRNA screens, we show that $A^{\downarrow}$ in predicted SDLs ($A^{\downarrow}$, $B^{\uparrow}$) is indeed more likely essential in an overexpressed enzyme $B^{\uparrow}$ background than when B is not overexpressed. When A is underexpressed and B is overexpressed in a predicted SDL in a given tumor sample, we denote that SDL as "active," that is, bearing potential functional effects on the tumor growth and the patient's survival. Second, we show that SDLs are less frequently active across patients with cancer compared with randomly selected enzyme pairs, indicating that tumor cells select against the presence of SDLs to avoid cell death. Third, we illustrate that tumor size in patients with breast cancer (BC) with one or more active SDLs is significantly smaller than in patients expressing randomly selected enzyme pairs. Fourth, we show that the predicted impeding effect of active SDLs on tumor growth correlates with a significantly longer patient survival time. These results become even more pronounced when one includes only highly ranked active SDLs (that show a stronger $A^{\downarrow}$, $B^{\uparrow}$ pattern at the transcriptional level), illustrating that our method successfully identifies the clinical impact of SDLs. Finally, we report that observed effects become stronger when more active SDLs are present in a given tumor, pointing to the cumulative effect of active SDLs in clinical tumors.

Results

Overview of IDLE Algorithm. The IDLE method (Fig. 1 and *SI Appendix*, *1*) computes the effect on cell growth when an enzyme B increases its activity (referred to here as the reference GSMM) compared with its activity in a KO GSMM in which, additionally, enzyme A is knocked out. The objective of IDLE is to find enzyme pairs (A, B) in which this differential growth effect is marked, searching over the space of all possible pairs. For a given pair (A, B), we define a reference WT GSMM and compute the maximum growth (biomass, μ_{max}) with flux balance analysis (35). Similarly, μ_{max} is computed for the KO GSMM, whereby reaction A is knocked out. In both models, the maximum flux through B is computed without any constraint on μ (i.e., lower bound is zero; Fig. 1 *A* and *B* show the reference and KO GSMM, respectively). Now, the lower bound of the biomass reaction is increased stepwise (by using $n = 10$ steps) toward μ_{max} in both the reference (Fig. 1*C*) and KO (Fig. 1*D*) model. For each increase, the maximal allowable flux through reaction B is computed again. The increasing growth pressure may affect the allowable flux through reaction B and, if so, it must decrease. The basic idea behind IDLE is that this argument is reversible: if the growth requirement constrains the maximum allowable flux through B, then a further flux increase through B must decrease growth. This effect is quantified and expressed as a vector (Fig. 1*E*). The angle θ between the reference and KO vectors measures the difference between the effects on cellular growth of overexpressing enzyme B in the WT (A, $B^{\uparrow}$) and after KO of enzyme A ($A^{\downarrow}$, $B^{\uparrow}$). If growth reduction is stronger in the KO situation ($A^{\downarrow}$, $B^{\uparrow}$), then we define θ positive and the enzymes (A, B) form an SDL. SDLs with the largest angle are predicted to have the maximum effect and are termed "high-impact" SDLs. We can therefore rank-order SDLs based on the computed angle θ (Fig. 1*F*).

The Metabolic SDL Network. Our method discovered 12,447 SDL interactions (*SI Appendix*, *2*, and Dataset S1). Reassuringly, the ranked list of SDLs significantly matches the top-ranked metabolic SDLs identified by the data mining synthetic lethality identification pipeline (DAISY), an approach for data-driven inference of genetic interactions in cancer that is based on the discovery of underrepresented gene pairs in cancer genomic data (20) (Wilcoxon rank-sum $P < 0.0038$).

SDLs are asymmetrical by definition, i.e., $A^{\downarrow}$, $B^{\uparrow}$ denotes a different interaction than $A^{\uparrow}$, $B^{\downarrow}$, and each may have a very different magnitude; in the first interaction, enzyme A is the KO partner, whereas, in the second interaction, it is the overexpressed partner. Surprisingly, we discovered that six enzymes are major "master" hubs, being the KO partners of many other overactivated $B^{\uparrow}$ in the SDL network (*SI Appendix*, Fig. S2). These major hubs (TPI, ENO, PGM, PYK, PGK, and GAPD) all reside in the glycolysis pathway. Interestingly, when examining the hub partners, we observed that the $B^{\uparrow}$ partners are the same for ∼80% of the SDLs. The metabolic pathways that are enriched for these overexpressed partners are shown in *SI Appendix*, Table S3.

To better understand the putative mechanisms underlying the workings of these SDLs, we conducted a further model-based analysis. First, we charted SL interactions of the six master hubs, i.e., searched for genetic interacting pairs involving these six hub reactions in which the growth reduction after their combined KO is larger compared with that observed after the single KOs. We were surprised to see that, although these SDL hub reactions are highly sensitive to a synthetic dosage load (each being essential for ∼500 overexpressed partners), they have only very few SL partners (a list of these reactions and their pathways is shown in *SI Appendix*, Table S4).

Examining the SDL partners of the six central glycolytic hubs, we find that they are quite distributed across the metabolic network in 10 different pathways that are significantly enriched with the SDL partners (*SI Appendix*, Table S5). When further investigating these SDLs, we discovered that glycogen production is decreased by (on average) 60% when such SDLs are active compared with the WT

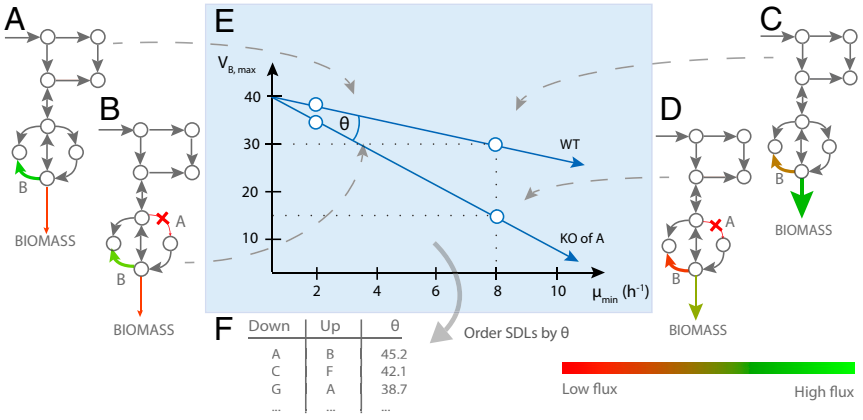

Fig. 1. Conceptual overview of the IDLE method. (*A*) The maximum flux through enzyme B is computed when there is no biomass pressure (i.e., lower bound flux is zero). (*B*) This process is repeated for the KO model. (*C*) The biomass pressure is increased in stepwise fashion and the maximum flux through enzyme B is computed at each step. (*D*) This is repeated for the KO model. (*E*) The maximum relative flux of B ($V_{B,max}$) is plotted at each biomass step (μ_{max}) and the angle θ between the reference and KO vector is computed. (*F*) SDL pairs are ranked based on their growth impact, quantified by their angle θ (*SI Appendix*, Fig. S1).

and KO conditions. Interestingly, it has recently been shown that glycogen metabolism and its initial accumulation is a key pathway induced by hypoxia, and its activity is necessary for optimal glucose utilization in tumors (36).

SDL Is Predictive of in Vitro shRNA Essentiality Screens. We expect that a knockdown of enzyme A ($A^{\downarrow}$) will be lethal in a $B^{\uparrow}$ background in the case of an SDL ($A^{\downarrow}$, $B^{\uparrow}$). To study this, we exploited gene essentiality at a genome-wide scale in cancer cell lines by using experimental shRNA screens (37) and matched it with gene-expression profiles (38). In a typical shRNA screen in a given cell line, each gene is individually knocked down by targeting its mRNA (inhibiting and degrading it) by specific shRNAs that bind to it. Then, the effect of each individual gene knockdown on cell growth is measured, from which scores are calculated that indicate gene essentiality [a $P = 0.05$ cutoff was used to consider a gene essential (37)].

For each cancer cell line, we divided SDLs into two groups: group 1 consists of SDLs in which at least one of the B enzymes that form an SDL with enzyme A is overexpressed ($B^{\uparrow}$) and group 2 consists of SDLs in which none of the B enzymes are overexpressed (*Materials and Methods* provides the definition of overexpression and *SI Appendix*, *3*, provides mapping genes to reactions). Then, the number of essential and nonessential A enzymes observed experimentally in the shRNA screen was compared between group 1 and group 2 in each cell line (one-tailed Fisher exact test). Using a $P = 0.05$ cutoff, we counted the number of cell lines in which enzymes A from group 1 are more frequently essential compared with these enzymes in group 2. This procedure was also repeated 5,000 times for a set of random enzyme pairs of equal size. As expected, the number of cell lines in which essentiality of A in a $B^{\uparrow}$ background is enriched (group 1) is significantly higher for SDL than for random pairs (empirical $P = 0.002$).

Cancer Cells Select Against SDL. Cancer cells are expected to select against the negative effect that SDLs have on (tumor) growth. Thus, when the enzyme pair (A, B) is an SDL, underexpression of enzyme A and overexpression of enzyme B should occur less frequently than for random enzyme pairs. We analyzed a gene expression dataset of 7,362 patients from the TCGA cohort (39) and determined for each gene whether it is underexpressed ($^{\downarrow}$), overexpressed ($^{\uparrow}$), or unchanged compared with expression levels in normal tissue samples (40) (*SI Appendix*, *4*). We then computed for all SDLs the number of patients, F_{SDL}, with an active SDL ($A^{\downarrow}$, $B^{\uparrow}$) relative to those patients having only enzyme A underexpressed ($A^{\downarrow}$, B) or having only enzyme B overexpressed (A, $B^{\uparrow}$; *SI Appendix*, *4*). This was repeated for 5,000 randomly constructed enzyme pair sets of equal size (F_{RANDOM}). As expected, F_{SDL} is significantly smaller than F_{RANDOM}, illustrating that an underexpression of A combined with an overexpression of B when A and B have an SDL relation occurs significantly less frequently than when the enzyme pair have no SDL relation (Fig. 2). In fact, when the angle θ increases, the fraction of patients that have an active SDL approaches zero, testifying to the strong negative selection exerted on such SDLs.

SDL Correlates with Smaller BC Tumor Size. As SDL negatively affects growth in cancer cell lines, we expect that the tumor size will be smaller for patients with at least one active SDL compared with those who do not. To address this, we used a dataset in which gene expression and matched tumor size data are available for 1,587 patients with BC (41). We divided the patients in this heterogeneous dataset based on the estrogen receptor (ER) sensitivity of their tumor (key properties of the data set are provided in *SI Appendix*, *5*).

We analyzed whether the tumor size of patients with an active SDL ($A^{\downarrow}$, $B^{\uparrow}$) is significantly smaller compared with patients who

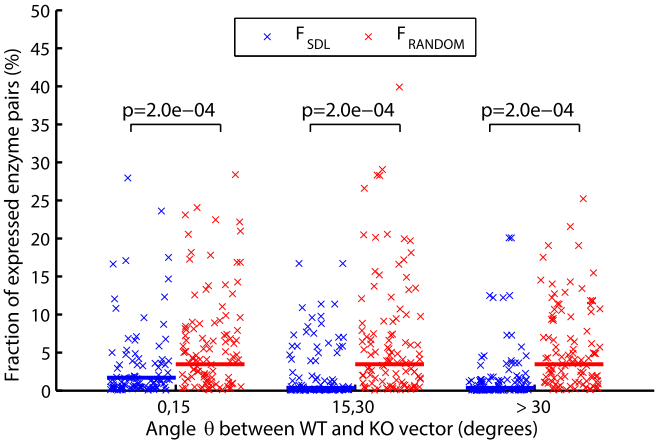

Fig. 2. Percentage of active enzyme pairs (i.e., A, B) with A underexpressed and B overexpressed. When the angle θ increases, the fraction of active SDLs approaches zero. SDLs are significantly less frequently active than randomly chosen enzyme pairs. For all cutoffs, the *P* values obtain their maximum (1+1/5,000+1) significance.

have one of the single effects, meaning only an under- ($A^{\downarrow}$, B) or overexpression (A, $B^{\uparrow}$) of enzyme A or B, respectively. To investigate $A^{\downarrow}$, $B^{\uparrow}$ in relation to $A^{\downarrow}$, B, we separated patients into two groups: patients whose tumor overexpresses enzyme B (*Materials and Methods* provides the definition of overexpression) with varying underexpression of enzyme A (σ between 0 and 3 given the underlying gene expression distribution) and patients whose tumor does not overexpress enzyme B with varying underexpression of enzyme A. When comparing $A^{\downarrow}$, $B^{\uparrow}$ with A, $B^{\uparrow}$, we also separated the patients into two groups: patients who have enzyme A underexpressed (*Materials and Methods* provides the definition of underexpression) with varying overexpression of enzyme B (σ between 0 and 3 given the underlying gene expression distribution) and patients who have enzyme A not underexpressed with varying overexpression of enzyme B. Finally, we created random enzyme pairs ($n = 5{,}000$) to serve as control for testing the specific effects of the SDLs. Statistical significance for all comparisons was computed with a signed Wilcoxon rank-sum test, analogous to the signed Kaplan–Meier test as previously defined (20) (*SI Appendix*, *6*).

As expected, we observed in ER^{+} BC that patients with (at least one active) SDL have significantly smaller tumors compared with patients with only overexpression of enzyme B ($P < 4 \times 10^{-8}$; Fig. 3). We found for ER^{-} disease that the tumor sizes in patients with SDL are also significantly smaller compared with patients with only overexpression of enzyme B ($P < 5 \times 10^{-5}$), as well compared with those with only underexpression of enzyme A ($P < 7 \times 10^{-5}$). Moreover, smaller tumors are observed for patients with ER^{-} and ER^{+} disease with active SDLs compared with patients with randomly selected enzyme pairs with the $A^{\downarrow}$, $B^{\uparrow}$ pattern active ($P < 2 \times 10^{-3}$).

SDL Correlates with Increased Cancer Survival. As SDLs decrease BC tumor size, we hypothesized that their presence also affects patient survival. For the BC data, matched survival times were available such that we could correlate them to the level of SDL activation (41). We hence performed a survival analysis analogous to the tumor size analysis described earlier. The significance of the results obtained for SDL were compared with the single effects and random pairs by a modified signed Kaplan–Meier test introduced in a previous publication (20) (*SI Appendix*, *6*).

As expected, we found that patients with ER^{+} BC with at least one active SDL have significantly better survival times compared with patients with only an underexpression of enzyme A ($P < 4 \times 10^{-3}$, Fig. 4 *A* and *B*). Patients with activated highly ranked SDLs

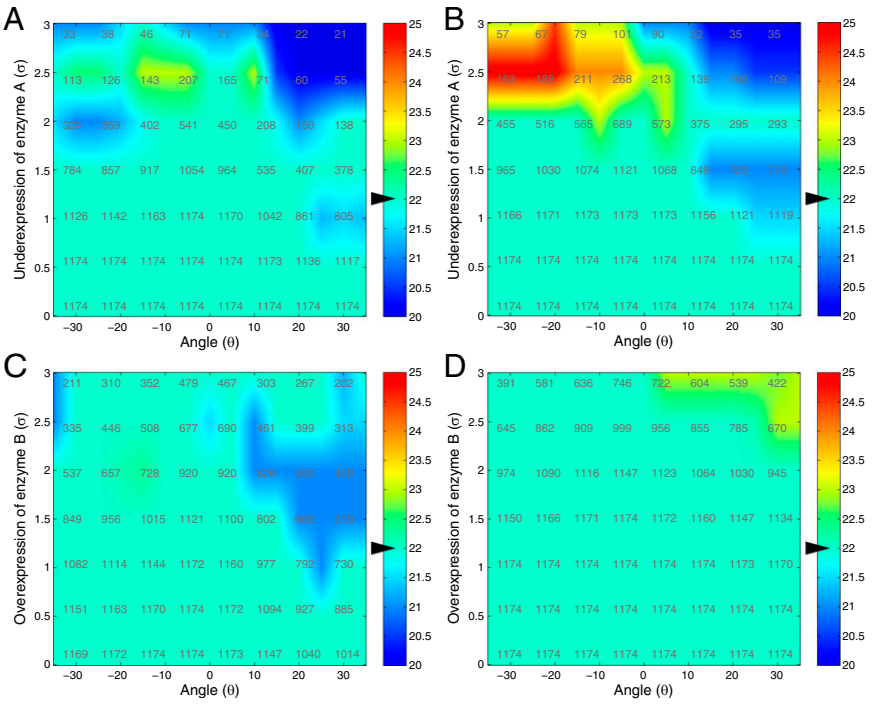

Fig. 3. Median BC tumor size (in millimeters) for patients with ER$^+$ disease. Arrowheads denote the median tumor size for all patients with ER$^+$ BC (22 mm). The number of patients that express at least one enzyme pair are denoted inside the figures. (*A*) Patients with at least one active SDL (A$^{\downarrow}$, B$^{\uparrow}$) with constant overexpression of enzyme B. (*B*) Patients whose disease only underexpresses enzyme A (A$^{\downarrow}$, B) of the SDL. (*C*) Patients with at least one active SDL (A$^{\downarrow}$, B$^{\uparrow}$) with constant underexpression of enzyme A. (*D*) Patients whose disease only overexpresses enzyme B of the SDL (A, B$^{\uparrow}$).

show the longest ER$^+$ BC survival times, as long as a median of more than 12 y (Fig. 4*A*). In line with expectation, the survival time of patients with active SDL is significantly better compared with patients with only enzyme B overexpressed ($P < 3 \times 10^{-4}$; Fig. 4 *C* and *D*). Moreover, significantly longer survival is also observed for patients with SDLs compared with those with random enzyme pairs with the A$^{\downarrow}$, B$^{\uparrow}$ pattern active ($P < 1 \times 10^{-3}$). *SI Appendix*, *7*, provides survival analysis of ER$^-$ patients. As overexpression of enzyme B is generally not beneficial when enzyme A is not underexpressed, we wondered whether underexpressing enzyme B alone would be beneficial. *SI Appendix*, Fig. S5, indicates that this is not the case. In particular, severe underexpression of enzyme B correlates with increased tumor sizes (*SI Appendix*, Fig. S4 *A* and *C*) and decreased survival times (*SI Appendix*, Fig. S4 *B* and *D*) in patients with ER$^+$ and ER$^-$ BC.

SDLs predicted by IDLE are not expected to be specific for BC. To examine their predictive power in another cancer type, we analyzed a large cancer type-specific cohort of 921 patients diagnosed with serous epithelial ovarian cancer (OC) (42) with

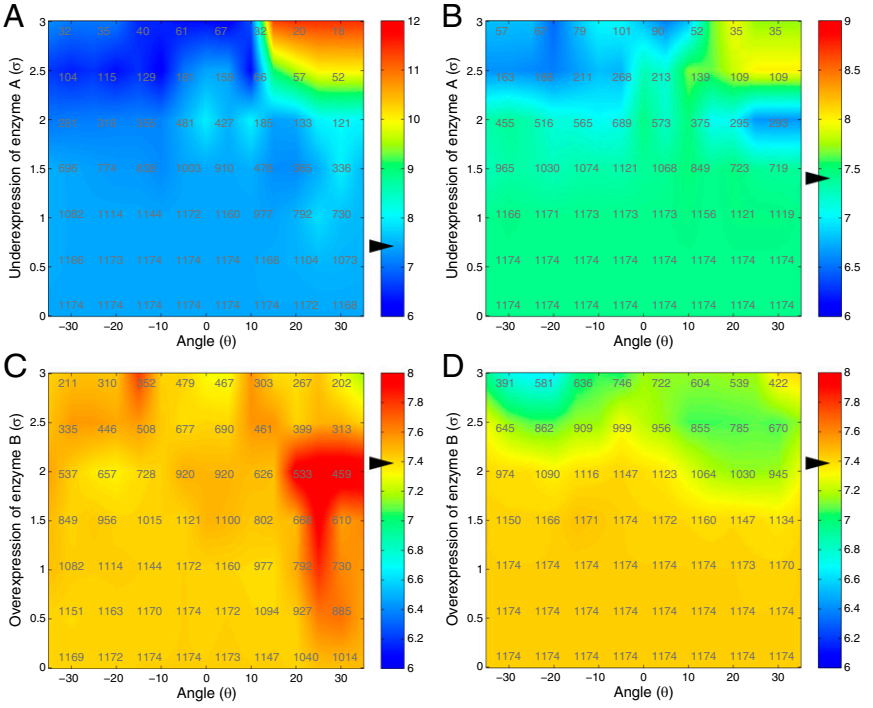

Fig. 4. Median ER$^+$ BC survival time (in years). Arrowheads denote the median survival for all patients with ER$^+$ BC (7.4 y). The numbers of patients whose disease expresses at least one enzyme pair are denoted inside the figures. Note that the axis of figure a scales differently. (*A*) Patients with at least one active SDL (A$^{\downarrow}$, B$^{\uparrow}$) with constant overexpression of enzyme B. (*B*) Patients whose disease only underexpresses enzyme A (A$^{\downarrow}$, B) of the SDL. (*C*) Patients with at least one active SDL (A$^{\downarrow}$, B$^{\uparrow}$) with constant underexpression of enzyme A. (*D*) Patients whose disease only overexpresses enzyme B of the SDL (A, B$^{\uparrow}$).

matched survival times. Indeed, the same observations were made as in the case of patients with ER^+ BC, i.e., patients with OC with at least one active SDL have significantly better survival times compared with those with the single or random effects ($P < 0.09$ vs. $A^{\downarrow}$, B and $P < 0.01$ vs. all others; *SI Appendix*, Fig. S5). These results are even more apparent in the relapse-free survival times (OC-RFS) of these patients ($P < 0.02$ vs. $A^{\downarrow}$, B and $P < 9 \times 10^{-4}$ vs. all others; *SI Appendix*, Fig. S6).

Cumulative Effect of SDLs in a Tumor Correlates to Better Survival. As SDL activity in a tumor correlates to survival prognosis, we asked if survival time would increase when patients have more SDLs active. We tested the presence of such a cumulative effect in the two largest cancer subtypes: patients with ER^+ BC ($n = 1{,}174$) and those with serous epithelial OC ($n = 921$). Patients were categorized into three groups, those having one to three, four to eight, or more than eight active SDLs in their expression profiles (over- and under-expression are defined in *Materials and Methods*).

The Kaplan–Meier survival curve in Fig. 5 shows, as expected, better survival for patients with large numbers of active SDLs compared with those with only a few active SDLs. Indeed, a log-rank test (43) revealed significantly improved survival times in both cancer types when the number of active SDLs increases ($P < 8 \times 10^{-3}$ for ER^+ BC and $P < 2 \times 10^{-3}$ for OC and OC-RFS). The largest cumulative effect in the BC survival is related to SDLs being active with enzyme A as one of the major glycolytic hubs. Interestingly, the observed cumulative effect in OC is already present for patients who have four to eight active (Fig. 5*B*). The underactivated enzymes A in these SDLs are enriched for pathways that use glutamine through glutamate metabolism, the TCA cycle, and mitochondrial transport ($P < 0.001$, hypergeometric test). It has recently been shown that severe types of OC, such as the epithelial subtype we considered, are driven by glutamine metabolism, in contrast to BC tumors that depend on an overactivity of glycolytic enzymes (44).

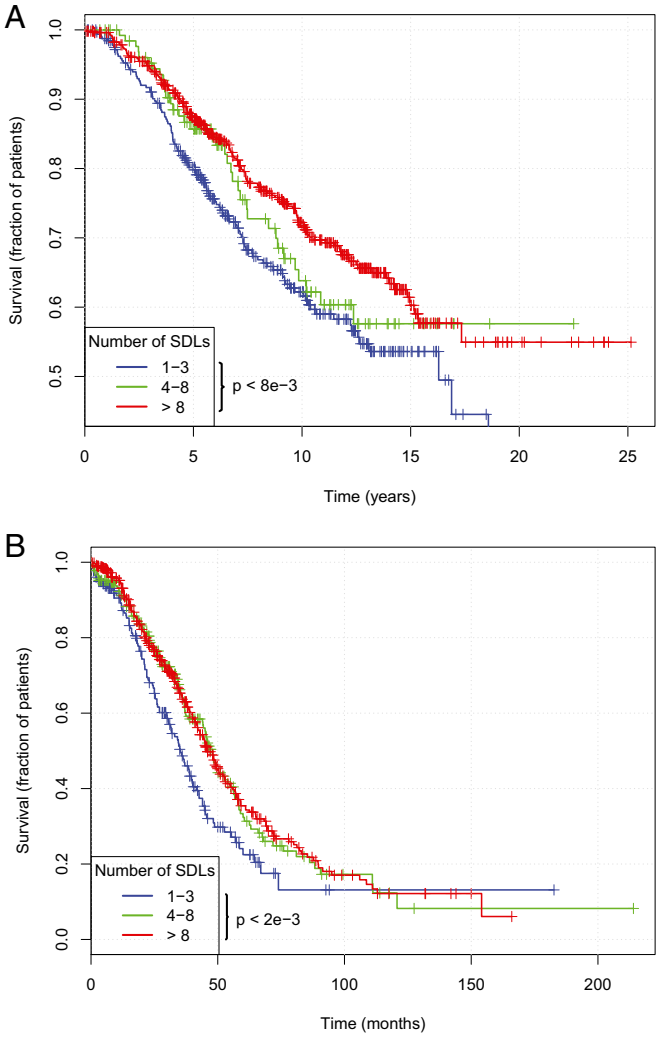

Fig. 5. Kaplan–Meier survival curves for patient groups that have one to three, four to eight, or more than eight active SDLs. (*A*) Survival times for the patients with ER^+ BC. (*B*) Survival times for patients with serous epithelial OC.

Discussion

Here we introduce what is, to our knowledge, the first computational method that captures enzymatic SDL effects in metabolic networks. Our method does not only identify SDLs that are strictly lethal to the cell, but also those that have a significant effect on tumor growth or proliferation in clinical settings (i.e., "synthetic dosage sick"). We show that our method is able to assign a measure of strength θ to each SDL, which correlates to its predictive power in an array of different tumor clinical attributes. It is therefore of interest to focus further research toward therapeutic interventions on the basis of "high-impact" pairs, which may have the largest beneficial effect on killing cancer cells.

We show that SDLs are less frequently active than expected in cancer cells. This shows that rapidly expanding cancer cells select against interactions that reduce their growth rate. The activation of "high-impact" SDLs is associated with smaller tumor sizes and longer patient survival. The effect strongly depends on the extent to which SDLs are activated, but most SDLs we found do not require a complete enzyme KO to exert a functional predictive signal. Last, we demonstrated a cumulative effect of SDL presence; the more SDLs active in a tumor sample, the better this is for a patient's prognosis. This observation may shed light on targeting cancers that rely on glycolysis. Down-regulating glycolytic enzymes that are the major hubs in the SDL network is hence expected to have a large growth-inhibitory effect in tumor cells that overexpress many of the glycolytic SDL partners. As glycolysis is usually less active in normal cells, and SDL partners of glycolytic hubs are less frequently overexpressed in normal cells compared with cancer cells in the majority of tissue types (*SI Appendix*, 2.2), targeting these glycolytic SDLs may be of therapeutic interest, especially when a large number of their partners are overexpressed.

The present study, being the first of its kind of which we are aware, naturally focuses on harnessing the generic human metabolic model to identify a common core of SDLs that may be shared by many different cancer types. However, the IDLE approach is general and could be extended in the future to identify cancer type-specific SDL interactions more precisely by integrating patient- and tumor-specific omics data such as gene expression or proteomics.

The results of our metabolic network modeling do not support the hypothesis that SDLs arise as a result of draining alternative compensatory pathways that compensate for the loss of the KO enzyme. This is because we do not find that the flux in such backup reactions of the major key glycolytic enzyme hubs is reduced following the overexpression of their SDL partners. Intriguingly, we do find that disrupted glycogen metabolism is predicted to be the major mechanism by which hundreds of SDLs of key glycolytic enzymes exert their growth-inhibitory effects. Indeed, it has recently been shown that glycogen metabolism and its initial accumulation is key for optimal glucose utilization in tumors (36). Thus, SDL relations do not arise via simple proximal interactions, but are likely to be the result of complex stoichiometric network relations that withdraw flux from biomass production through activation of other pathways.

Our results testify to the potential contribution of model-based approaches to identify and uncover the mechanisms behind SDLs. Model-based SDL prediction via IDLE is widely applicable and not limited to cancer. It could be used to identify SDL networks in pathogenic bacteria or fungi, providing new antibiotic therapeutic leads. Other possible applications include metabolic engineering to increase the yield of valuable metabolic byproducts. Specifically, this may be achieved by engineering an SDL effect to inhibit the production of undesired byproducts, or inversely, neutralizing the SDL effect to force an increased flux through desired pathways. Taken together, IDLE is expected to contribute to various research fields ranging from medical sciences to biotechnology.

Materials and Methods

IDLE requires a GSMM with *m* metabolites, *n* reactions, and a well-defined cellular objective function. We used the human metabolic network (recon1) (45), supplemented with a biomass reaction to simulate growth. A rich environment was simulated by allowing a maximum metabolite uptake rate of 5.0 mmol/gram dry weight/h through all boundary reactions. The goal of IDLE is to find SDL enzyme pairs (A, B) that severely interrupt cell growth when the flux through enzyme A is decreased and the flux through B is increased (denoted as $A^{\downarrow}$, $B^{\uparrow}$). As illustrated in Fig. 1 and *SI Appendix, 1*, SDL is measured by an angle θ. Only those enzyme pairs with a significant difference between the reference and KO were analyzed, i.e., all pairs with $|\theta| \geq 2°$ were selected, resulting in a list of 12,447 putative SDLs. *SI Appendix, 1*, provides a detailed description of IDLE with an example.

In all analyses, we defined an enzyme/gene to be under- or overexpressed when its expression was below or above 0.5σ to 1.0σ from the mean in the gene expression distribution (see *Results* for references to gene expression datasets).

Detailed procedures of mapping gene expression to enzyme reaction level and calculating the fraction of SDLs in cancer cells (F_{SDL}) and descriptions of tumor size and patient survival statistics are provided in *SI Appendix, 1*.

ACKNOWLEDGMENTS. The authors thank Livnat Jerby-Arnon and Keren Yizhak for their assistance in various phases of the study. This work was supported by The Netherlands Organisation for Scientific Research Grant CSBR09/013V (to W.M. and R.A.N.), European Union FP7 large-scale integrated network Gencodys (www.gencodys.eu) Health-241995 (to X.L.), the Israeli Science Foundation (E.R.), the Israeli Cancer Research Fund (E.R.), the Israeli Center of Excellence (I-CORE) Program of the Planning and Budgeting Committee (E.R.), and Israel Science Foundation Grant No. 41/11.

1. Hartwell LH, Szankasi P, Roberts CJ, Murray AW, Friend SH (1997) Integrating genetic approaches into the discovery of anticancer drugs. *Science* 278(5340):1064–1068.
2. Costanzo M, et al. (2010) The genetic landscape of a cell. *Science* 327(5964):425–431.
3. Szappanos B, et al. (2011) An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nat Genet* 43(7):656–662.
4. Suthers PF, Zomorrodi A, Maranas CD (2009) Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Mol Syst Biol* 5:301.
5. Deutscher D, Meilijson I, Kupiec M, Ruppin E (2006) Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat Genet* 38(9):993–998.
6. Segrè D, Deluna A, Church GM, Kishony R (2005) Modular epistasis in yeast metabolism. *Nat Genet* 37(1):77–83.
7. Folger O, et al. (2011) Predicting selective drug targets in cancer through metabolic networks. *Mol Syst Biol* 7:501.
8. Frezza C, et al. (2011) Haem oxygenase is synthetically lethal with the tumour suppressor fumaratehydratase. *Nature* 477(7363):225–228.
9. Laufer C, Fischer B, Billmann M, Huber W, Boutros M (2013) Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nat Methods* 10(5):427–431.
10. Luo J, et al. (2009) A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* 137(5):835–848.
11. Lu X, Kensche PR, Huynen MA, Notebaart RA (2013) Genome evolution predicts genetic interactions in protein complexes and reveals cancer drug targets. *Nat Commun* 4:2124.
12. Kaelin WG, Jr (2005) The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 5(9):689–698.
13. Rehman FL, Lord CJ, Ashworth A (2010) Synthetic lethal approaches to breast cancer therapy. *Nat Rev Clin Oncol* 7(12):718–724.
14. Chan DA, Giaccia AJ (2011) Harnessing synthetic lethal interactions in anticancer drug discovery. *Nat Rev Drug Discov* 10(5):351–364.
15. Kroll ES, Hyland KM, Hieter P, Li JJ (1996) Establishing genetic interactions by a synthetic dosage lethality phenotype. *Genetics* 143(1):95–102.
16. Bian Y, et al. (2014) Synthetic genetic array screen identifies PP2A as a therapeutic target in Mad2-overexpressing tumors. *Proc Natl Acad Sci USA* 111(4):1628–1633.
17. Sajesh BV, Guppy BJ, McManus KJ (2013) Synthetic genetic targeting of genome instability in cancer. *Cancers (Basel)* 5(3):739–761.
18. Sopko R, et al. (2006) Mapping pathways and phenotypes by systematic gene overexpression. *Mol Cell* 21(3):319–330.
19. Wagner A, et al. (2013) Computational evaluation of cellular metabolic costs successfully predicts genes whose expression is deleterious. *Proc Natl Acad Sci USA* 110(47): 19166–19171.
20. Jerby-Arnon L, et al. (2014) Predicting cancer-specific vulnerability via data-driven detection of synthetic lethality. *Cell* 158(5):1199–1209.
21. Mo ML, Jamshidi N, Palsson BØ (2007) A genome-scale, constraint-based approach to systems biology of human metabolism. *Mol Biosyst* 3(9):598–603.
22. Schellenberger J, et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0. *Nat Protoc* 6(9):1290–1307.
23. Price ND, Papin JA, Schilling CH, Palsson BO (2003) Genome-scale microbial in silico models: The constraints-based approach. *Trends Biotechnol* 21(4):162–169.
24. Thiele I, et al. (2013) A community-driven global reconstruction of human metabolism. *Nat Biotechnol* 31(5):419–425.
25. Terzer M, Maynard ND, Covert MW, Stelling J (2009) Genome-scale metabolic networks. *Wiley Interdiscip Rev Syst Biol Med* 1(3):285–297.
26. Feist AM, Palsson BØ (2008) The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol* 26(6):659–667.
27. Oberhardt MA, Palsson BØ, Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 5:320.
28. Price ND, Reed JL, Palsson BØ (2004) Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat Rev Microbiol* 2(11):886–897.
29. Nam H, et al. (2012) Network context and selection in the evolution to enzyme specificity. *Science* 337(6098):1101–1104.
30. Notebaart RA, et al. (2014) Network-level architecture and the evolutionary potential of underground metabolism. *Proc Natl Acad Sci USA* 111(32):11762–11767.
31. Jerby L, et al. (2012) Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. *Cancer Res* 72(22):5712–5720.
32. Agren R, et al. (2012) Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLOS Comput Biol* 8(5):e1002518.
33. Jerby L, Ruppin E (2012) Predicting drug targets and biomarkers of cancer via genome-scale metabolic modeling. *Clin Cancer Res* 18(20):5572–5584.
34. Gatto F, Nookaew I, Nielsen J (2014) Chromosome 3p loss of heterozygosity is associated with a unique metabolic network in clear cell renal carcinoma. *Proc Natl Acad Sci USA* 111(9):E866–E875.
35. Varma A, Palsson BO (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Appl Environ Microbiol* 60(10):3724–3731.
36. Favaro E, et al. (2012) Glucose utilization via glycogen phosphorylase sustains proliferation and prevents premature senescence in cancer cells. *Cell Metab* 16(6): 751–764.
37. Marcotte R, et al. (2012) Essential gene profiles in breast, pancreatic, and ovarian cancer cells. *Cancer Discov* 2(2):172–189.
38. Barretina J, et al. (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483(7391):603–607.
39. Chin L, Hahn WC, Getz G, Meyerson M (2011) Making sense of cancer genomic data. *Genes Dev* 25(6):534–555.
40. Lu X, Megchelenbrink W, Notebaart RA, Huynen MA (2015) Predicting human genetic interactions from cancer genome evolution. *PLoS One* 10(5):e0125795.
41. Curtis C, et al.; METABRIC Group (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486(7403):346–352, 10.1038/nature10983.
42. Gyorffy B, Lánczky A, Szállási Z (2012) Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocr Relat Cancer* 19(2):197–208.
43. Bland JM, Altman DG (2004) The logrank test. *BMJ* 328(7447):1073.
44. Yang L, et al. (2014) Metabolic shifts toward glutamine regulate tumor growth, invasion and bioenergetics in ovarian cancer. *Mol Syst Biol* 10(5):728, 10.1002/msb.20134892.
45. Duarte NC, et al. (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci USA* 104(6):1777–1782.