Harnessing synthetic lethality to predict the response to cancer treatment

Joo Sang Lee1,2, Avinash Das1, Livnat Jerby-Arnon3, Rand Arafeh4, Noam Auslander1,2, Matthew Davidson5, Lynn McGarry5, Daniel James5, Arnaud Amzallag6,7,8, Seung Gu Park1, Kuoyuan Cheng1,2, Welles Robinson1,2, Dikla Atias9, Chani Stossel9, Ella Buzhor9, Gidi Stein10, Joshua J. Waterfall11, Paul S. Meltzer11, Talia Golan9,10, Sridhar Hannenhalli1, Eyal Gottlieb5, Cyril H. Benes6,7, Yardena Samuels4, Emma Shanks5 & Eytan Ruppin1,2,3,10

While synthetic lethality (SL) holds promise in developing effective cancer therapies, SL candidates found via experimental screens often have limited translational value. Here we present a data-driven approach, ISLE (identification of clinically relevant synthetic lethality), that mines TCGA cohort to identify the most likely clinically relevant SL interactions (cSLi) from a given candidate set of lab-screened SLi. We first validate ISLE via a benchmark of large-scale drug response screens and by predicting drug efficacy in mouse xenograft models. We then experimentally test a select set of predicted cSLi via new screening experiments, validating their predicted context-specific sensitivity in hypoxic vs normoxic conditions and demonstrating cSLi’s utility in predicting synergistic drug combinations. We show that cSLi can successfully predict patients’ drug treatment response and provide patient stratification signatures. ISLE thus complements existing actionable mutation-based methods for precision cancer therapy, offering an opportunity to expand its scope to the whole genome.

1 Center for Bioinformatics and Computational Biology, University of Maryland Institute of Advanced Computer Science (UMIACS) & Department of Computer Science, University of Maryland, College Park, MD 20742, USA. 2 Cancer Data Science Lab, National Cancer Institute, National Institute of Health, Bethesda, MD 20892, USA. 3 The Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv 6997801, Israel. 4 Department of Molecular Cell Biology, Weizmann Institute, Rehovot 7610001, Israel. 5 Cancer Research UK, Beatson Institute, Switchback Road, Glasgow G61 1BD Scotland, UK. 6 Massachusetts General Hospital Center for Cancer Research, Charlestown, MA 02129, USA. 7 Harvard Medical School, Boston, MA 02114, USA. 8 PatientsLikeMe, 160 Second Street, Cambridge, MA 02142, USA. 9 Division of Oncology, Sheba Medical Center Tel-Hashomer, Ramat-Gan 5262100, Israel. 10 Genetics Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. These authors contributed equally: Joo Sang Lee, Avinash Das. Correspondence and requests for materials should be addressed to E.R. (email: eytan.ruppin@nih.gov)
The success of precision oncology depends on its ability to translate accumulating genomic data into actionable treatment options tailored for individual patients. This requires identifying a genomic signature from patient tumor samples, then matching it with the most effective therapeutic options. With the ever-increasing volume of genomic data, the bottleneck now lies on how to extract patient-specific vulnerabilities from the data and connect them to patient prognosis and drug response. One promising way to tackle this challenge is based on the concept of synthetic lethality (SL). SL describes the relationship between two genes whereby an individual inactivation of either gene results in a viable phenotype, while their combined inactivation is lethal. SL has long been considered a foundation for the development of selective anticancer therapies, which aim to inhibit the SL partner of a gene that is inactivated de novo in the cancer cells. Beyond guiding the development of novel selective cancer therapies, it has been noted that the network of SL interactions can provide a bird’s eye view on the genomic state of a given tumor that can be leveraged to identify tumor-specific vulnerabilities and develop effective synergistic drug combination therapies in a precision-based manner.

Given the importance of SL, considerable work has been devoted to identifying such interactions in cancer—both experimentally and computationally. Experimentally, extensive efforts have been made to tease out the wiring of genetic interactions in cancer cells based on single cell lines or large-scale knockout screens. Computationally, various machine learning methodologies have been applied to predict genetic interactions in different species and cancer (by utilizing yeast SL), utilizing metabolic modeling, evolutionary characteristics, transcriptomic profiles, and more recently cancer patient data. Nevertheless, so far the utility of SL in the clinic has been primarily limited to SLs in DNA damage pathways, and as we show further below, many of the SLs identified in current screens manifest a poor predictive signal in actual patients’ data.

Here we present a statistical approach for identifying clinically relevant SL interaction (SLI) in a genome-scale manner, termed identification of clinically relevant synthetic lethality (ISLE). ISLE takes lab-screened SL interactions as inputs and analyzes tumor molecular profiles, patient clinical data, and gene phylogeny relations to identify SLs that are predictive of patients’ drug response. The ISLE-identified SL interactions are shown to predict drug response to a wide variety of drugs both in vitro and in vivo, providing a basis for rational design of synergistic drug combinations. We first benchmark ISLE with large-scale in vitro drug response screens and in mouse xenograft models. We then experimentally test and validate predictions involving context-specific gene essentiality and drug efficacy, and synergistic drug combinations in patient-derived cell lines. We finally show that cSLs, which are inferred from mining untreated patients’ data, can successfully predict treatment outcomes in cancer patients without any need for training on specific patient cohorts of drug response data. Taken together, these results offer a novel approach for precision-based cancer therapy from the patients’ tumor data.

**Results**

**Identifying clinically relevant SL interactions via ISLE.** ISLE takes the initial pool of lab-identified candidate SL pairs as input and identifies among those the subset that is more likely to be clinically relevant, that is, supported by tumor data (see below). The initial pool is determined either by direct isogenic (or double-knockout) cell line screens (Initial Set I, Supplementary Data 1) or guilt-by-association using large-scale single gene knock-out experiments (Initial Set II), creating a pool of total 16 million candidate SL pairs (see Methods, Supplementary Note 1). The two initial SL input sets show significant overlap, confirming findings of Wang et al. (hypergeometric *P* < 1.4E-28) (see Supplementary Data 2). To identify putative cSLs from the initial pool, ISLE analyzes molecular, and survival data of patient tumor samples from The Cancer Genome Atlas (TCGA) and evaluates the extent to which clinical data support in vitro screens. Conceptually, ISLE selects the clinically relevant SL pairs that satisfy all of the three conditions outlined below. From a computational standpoint, however, it applies them in a specific sequential manner that minimizes the computational cost of their identification (Fig. 1a, see Methods for full details):

- First, ISLE mines gene expression and SCNA data of the input patient tumor samples to identify under-represented candidate gene pairs, whose co-inactivation is significantly less frequent than expected by their individual inactivation frequencies, testifying that their co-inactivation is under negative selection (using hypergeometric test; see Methods, Supplementary Note 1). Second, ISLE selects SL pairs where the tumor samples with a given pair in a co-inactive state exhibit better patient’s survival than the samples where it is not, testifying that this SL pair is likely to reduce tumor fitness when co-inactive. ISLE uses a stratified Cox proportional hazard model to establish this association while controlling for confounding factors including cancer type, genomic instability and patients’ gender, age, and ethnicity (see Methods).

- Third, ISLE selects SL pairs composed of genes having high phylogenetic similarity, motivated by the observation that functionally interacting genes tend to co-evolve. ISLE analyzed SCNA, gene expression, mutation and survival data of 8749 clinical samples across 28 cancer types, to identify a pan-cancer clinical-SL-network (Fig. 1a). The initial input set consisted of 16,375,526 candidate SLs, out of which 9.1% passed the first step, 0.4% of the initial pool pass the second step, and 0.1% of the initial pool pass the third step and constitute the final cSL network. The resulting cSL network is comprised of 8511 genes and 21,534 interactions (Supplementary Fig. 1, Supplementary Data 3, all networks are available online via an interactive interface, see Supplementary Note 2, the core SL network is presented in Fig. 1b, where a stricter FDR < 0.1 is used to obtain a smaller network for visualization purposes). The cSL genes are highly enriched in cell proliferation, migration, apoptosis, and signal transduction (FDR-corrected hypergeometric *P* < 1E-8, Supplementary Data 4). The identified cSL interactions are enriched with physical protein interactions (hypergeometric *P* < 6.6E-3), consistent with previous observations in the literature.

- The final cSL pairs successfully predict the survival response in an independent cohort of breast cancer patients, as expected (Supplementary Note 1).

Throughout the manuscript, we term the candidate SL pairs that withstand the three ISLE steps as clinically relevant SL (cSL), while those that are filtered out are termed non-clinically relevant SL (ncSL). Surprisingly, the ISLE analysis shows that only a small fraction of the gold standard initial pool (Initial Set I) is cSL: (1) only 12.5% of the pairs identified in the isogenic or double-knockdown screens show evidence of negative selection in tumor samples (i.e., their co-downregulation is significantly less frequent than expected by chance, Step 1). An even smaller selection of 0.2% of the original candidate pairs is associated with improved patient survival when co-downregulated (as would be expected
from cSLs, Step II). Importantly, to estimate the clinical relevance of these interactions, we used mutation data for isogenic screens; and for double knockdown screens, we used copy number and transcriptomics data (see Supplementary Note 1 for details). Furthermore, in addition to pan-cancer tests, we also performed the clinical estimations outlined above in the specific cancer types of relevance, aiming to uncover the potential clinical relevance of the SLs screened in as comprehensive manner as possible (see Methods, Supplementary Note 1 for details). (2) A similar trend emerges when evaluating the clinical relevance of the SLi inferred via analyzing single knockdown screens (Initial Set II), with only 2.7% of the pairs from initial pools showing evidence of negative selection (Step I) and 0.4% of pairs showing better prognosis (Step II), respectively (see Supplementary Data 5). Beyond these tests, our analysis shows that ncSL pairs are not predictive of drug response in cancer patients, and their predictive signal of in vitro drug response is markedly lesser than that of cSLs (see below).

Validating ISLE-identified SL interactions. We performed three validation steps to evaluate the capability of ISLE to identify clinically relevant SLs. First, to validate individual ISLE-inferred SL interactions, we analyzed an in vitro large-scale drug response screen\(^{58}\) covering 24 drugs in 500 cancer cell lines. We tested whether a drug is more effective in cell lines where the ISLE-inferred cSL partner(s) of its targets are downregulated (according to the cell lines gene expression), thus confirming the existence of pertaining SLi (Methods). The cSLi inferred by ISLE are markedly more predictive (Area under the curve (AUC) = 0.75, Fig. 2a) than the ncSL pairs (AUCs of 0.47, and precision-recall analysis in Supplementary Fig. 2). ISLE’s performance level is also markedly superior to that of DAISY (AUCs of 0.37), and to that obtained by randomly shuffled networks (empirical P < 1E-3).

Second, we tested whether ISLE-identified SLi predict single and combination drug response in patient-derived mouse xenograft (PDX) models. We analyzed 375 samples of mouse models of 15 cancer types, which were treated with 36 single drugs\(^{39}\). As the first step in this analysis, ISLE was applied to identify the relevant drug-cSL interactions; that is, to identify the SL partners of each of the given drug targets (Methods; this results in an extended set of interactions compared to those identified in the overall cancer cSL network as the space of hypotheses that needs to be corrected for is now obviously much smaller). Second, to make the drug response predictions for each drug, we define its cSL-score in each sample, by denoting the number of its predicted drug-cSL partners that are downregulated in that sample according to its gene expression data divided by the number of drug targets (Eq. (4) in Methods). As expected, responders show significantly higher cSL-scores than non-responders for five out of seven drugs that have sufficient tumor response data available (using Wilcoxon rank sum test; Fig. 2b, Methods). Furthermore, the cSL network successfully predicts the progression-free survival in mouse models (defined as the time for the tumor size to grow to twice the size of the baseline, see Methods, Supplementary Note 1) (overall log rank P < 2.90E-6, ΔAUC = 0.15), which remains significant after controlling for cancer types (Cox hazard ratio = 1.28 (P < 3.82E-3)). Neither DAISY nor ncSL partners are predictive of any of the drugs considered, and ISLE’s performance is significantly superior to that of randomly assigned SL partners (empirical P < 1E-3).

Third, to benchmark the cSL-based drug response predictions, we analyzed the DREAM7 challenge data\(^{60}\). We focused on 15 of the 28 drugs that have specific targets and applied ISLE to infer their cSL partners (see Methods, Supplementary Note 1). Drug efficacy in a cell line is predicted via its cSL-score inferred from the cell-line’s transcriptomic profiles (as described in the previous paragraph). The resulting ISLE-based prediction is comparable to the top 5 supervised algorithms (trained in this specific data) reported in DREAM\(^7\)\(^{60}\) (Fig. 2c (left columns)). To predict
double-drug response, we hypothesized that a pair of drugs will be synergistic if there exists a strongly predicted SL between their gene targets (Methods). We used the ISLE inferred SL strength between the drug targets (cSL-pair-score) as the predictor of synergism, focusing on 12 inhibitor compounds (66 combinations) excluding two non-specific drugs from the data. ISLE prediction accuracy ranks higher than the best five supervised algorithms reported in the DREAM761 (Fig. 2c (right columns)). Importantly, ISLE does not use any drug response data from the training set cell lines or post-treatment transcriptomic profiles that were used to train all the other competing approaches in the original challenge. This makes ISLE less prone to the risk of overfitting and broadly applicable for drug response prediction without the need of further training. The performance of randomly selected, DAISY-identified and most importantly, ncSL partners is markedly inferior to that of ISLE (Supplementary Fig. 3).

Fourth, we analyzed large-scale compound synergy screens from a more recent AstraZeneca-Sanger Drug Combination Prediction DREAM Challenge 201562, which covers a set of pairwise combinations of 69 drugs (169 drug combinations and 535 candidate gene pairs total) in 85 cancer cell lines. We predict synergistic drug combinations based on the working hypothesis that if there is a predicted cSL between the gene targets of each of the two drugs, they are likely to be synergistic as described above. We labeled a drug pair as overall synergistic if it experimentally exhibits synergism across many cell lines (see Methods). The simple, unsupervised cSL-based predictor described above provides a fairly accurate prediction (AUC = 0.79, Fig. 2d, Methods), significantly higher than the random (empirical P < 1E-3), DAISY-SL, or ncSL networks (Fig. 2d, Supplementary Fig. 4). We further tested whether the ISLE-identified cSL predict combinations’ drug response in PDX models, which includes 375 samples of mouse models of 15 cancer types, which were treated with 26 double-drug combinations39. Our unsupervised cSL-based predictor classifies drug pairs as synergistic vs. non-synergistic (AUC = 0.78, Fig. 2d, Supplementary Fig. 5).

Experimentally testing cSL via phenotypic screens. We performed three layers of experiments to test the performance of
ISLE in predicting gene essentiality and drug response for single agents and drug combinations.

First, we used ISLE-identified SL interactions to predict gene context-specific essentiality\(^6\). To this end, we conducted a large-scale shRNA knockdown (KD) screening in two different environmental conditions (21% \(O_2\), “normoxia” and 0.1% \(O_2\), “hypoxia”) in a liv7k oral cancer cell line (Supplementary Data 6). As described in the previous section, we computed the cSL-score of each gene targeted, by identifying how many of its SL partners are downregulated in the condition studied (Methods). The computed scores are predictive of the individual KD effects on cellular growth (FDR-corrected \(t\)-test \(P < 0.2\) between every bin, Fig. 3a), and of the context-specific differential KD effects in the two conditions (predicted correctly for 71% of the genes analyzed, Fig. 3b). Both context-generic and specific predictions were inferior when using either randomly selected, DAISY-identified, or ncSL partners (Supplementary Note 1).

Second, we examined the cSL network’s ability to predict in vitro drug response. To this end, we performed two sets of experiments. (1) We treated liv7k oral cancer cell line studied above with 463 drugs (including non-cancer drugs) under hypoxia and normoxia conditions (Methods). We labeled the drugs within the top 33% observed growth inhibition in the experiments as effective and computed the drug-cSL-scores of their targets in the two specified conditions as before, but divided by the number of targets per each drug (Supplementary Data 7, Methods). Drugs with high cSL-scores show stronger drug response in each condition separately (AUCs of 0.75, AUPRC = 0.71, empirical \(P < 1E-3\)), and randomly permuted SL networks, DAISY, and ncSL network failed to obtain a predictive signal (Methods, Supplementary Fig. 6A). Analogous to the gene knockout experiment, the cSL-score successfully predicted in which condition (hypoxia vs normoxia) the drug would be more effective for 95% of the drugs tested (Spearman \(R = 0.36\) (\(P < 0.01\), Supplementary Note 1). (2) We additionally conducted a large-scale drug response screen covering 142 small molecule inhibitors in seven breast cancer cell lines of different subtypes (Supplementary Data 8). cSL-scores inferred from the transcriptomic profiles of these cell lines successfully predicted drug response (AUC = 0.73, AUPRC = 0.48, Supplementary Fig. 6B),

**Fig. 3** New experiments to test ISLE-based predictions on growth inhibition and drug combinations. **a, b** cSL-based prediction of growth inhibition in a knockdown screen in oral cancer. a Growth rate prediction: The number of downregulated cSL partners of a gene (X-axis; cSL-score) is associated with the percentage of growth inhibition observed after its knockdown (Y-axis; quantified as percent-growth inhibition compared to control, mean and s.e.m). Each bin of cSL-score shows significant differences (FDR-corrected \(t\)-test \(P < 0.2\); see Supplementary Note 1). b cSL-based context-specific prediction of growth inhibition in hypoxic vs normoxic conditions. The fold change of cSL-score (X-axis in logscale) in normoxia vs hypoxia shows a positive correlation with the corresponding growth inhibition fold change (Y-axis in logscale), correctly predicting the differentially observed growth inhibition in more than 71% of the 38 cases observed experimentally (marked as red dots in the 1st and 3rd quadrants). **c, d** The figures depict the representative dose response curves (see Supplementary Fig. 7B, C for other cell lines) of the predicted synergistic drug combinations of **c** ABT263 (BCL2L1 inhibitor) and the OSI906 (IGFR inhibitor) and **d** GDC0941 (PIK3CA inhibitor) and the MK2206 (AKTI1 inhibitor). The percentage of cell line survival (Y-axis) was measured at varying doses of OSI906 (respectively MK2206), with and without ABT263 (respectively GDC0941) treatment at SuM (X-axis). The dashed lines denote the percentage of cell line survival at varying levels of OSI906 (MK2206) without ABT263 (GDC0941) treatments, and the solid lines denote the percentage of cell line survival at varying levels of OSI906 (MK2206) in the presence of 5 \(\mu\)M of ABT263 (GDC0941). The combined drug treatments are significantly more effective than the single treatments based on the analysis of variance (\(P < 3.17E-11\) (c) and \(P < 2.71E-8\) (d)). The Fa–CI curve for all drug combinations are in Supplementary Fig. 7D, and the full experimental measurements are presented in Supplementary Data 11, 12.
Table 1 SL partners and corresponding drug combinations tested in patient-derived melanoma cell lines

| SL interaction          | Drug combinations | Cell line 1 | Cell line 2 | Cell line 3 | Cell line 4 |
|-------------------------|-------------------|-------------|-------------|-------------|-------------|
| BCL2L1-IGF1R            | ABT263-Os1906     | 0.47        | 0.54        | 0.68        | 1.12        |
| PIK3CA-AKT1             | GDC0941-MK2206    | 0.75        | 0.23        | 0.77        | 0.59        |

The table shows the top two predicted cSL interactions tested in melanoma, the corresponding drugs tested and the experiments’ outcome as evaluated by the combination index (CI) at 50% of cells are affected (CI< 0.7: strong synergism (blue), CI< 1: moderate synergism (green), and CI> 1: no synergism (white)).

while randomly selected partners are not predictive. DAISY-SL or ncSL partners are also not predictive for this task (AUC = 0.50 and 0.51, respectively in Supplementary Fig. 6).

Third, we conducted new drug combination experiments in patient-derived melanoma cell lines. Focusing on key melanoma drivers63, we selected the top target pairs among the strongest ISLE-predicted cSLi between them—AKT1 and PIK3CA, and BCL2L1 and IGF1R (Table 1, Supplementary Data 9). We performed the experiments in cell lines where each gene of a selected pair was highly expressed (Supplementary Data 10), successfully validating seven out of the eight predicted cases (Table 1, see examples in Fig. 3c, d; see Supplementary Note 1 for details, Supplementary Fig. 7D for full Fa–CI curves and Supplementary Data 11, 12 for measurement data).

cSL-based prediction of drug response in patients. We next inferred the drug-cSL-network of clinically approved cancer drugs (Methods, Supplementary Fig. 8) and tested its capacity to predict drug response in three different patients’ datasets of different tumor types64,65,67. Following the procedure described before regarding in vitro drug prediction, for each drug we predict its response in each sample by computing its cSL-score; the cSL-score of a drug denotes the number of its target genes’ SL partners that are downregulated in the specific sample divided by the number of the drug’s target genes (this number is determined from the sample’s transcriptomics data; Eq. (4) in Methods).

The first dataset is composed of 508 tumor biopsies of HER2-negative invasive breast cancer patients before treatment with taxane-anthracycline chemotherapy66. Patients with high cSL-scores have significantly longer distant relapse-free survival (DRFS) rates compared to those predicted as non-responders, as expected (cox hazard ratio 1.40 (P < 6.0E-3), log rank P of 9.2E-4, Fig. 4a). Indeed, patients annotated as responders show significantly higher cSL-scores than non-responders (Wilcoxon rank sum P < 0.028 with breast cancer subtypes controlled). The cSL partners of other drugs with the same degree distribution had no predictive signal, and so do ncSLs and DAISY based predictions (log rank P > 0.37).

Second, we applied the network to predict the response of 25 patients with recurrent or metastatic non-small cell lung cancer (NSCLC) to the EGFR-inhibitor erlotinib65,66. All patients had EGFR wild-type tumors. The number of downregulated EGFR-cSL partners is a marker of better prognosis (Fig. 4b) and shows significant association with patient survival (Cox hazard ratio = 2.15 (P < 6.5E-3), Supplementary Fig. 9 A, B), which is significantly better than the predictive performance of the randomly permuted networks (empirical P < 0.05). The cEGFR-cSL partners predict specific response to erlotinib, as opposed to merely predicting patient survival: It failed to predict patients’ response in an independent arm of the same trial in which 37 NSCLC patients were treated with sorafenib, a VEGFR inhibitor (log rank P > 0.95, Supplementary Fig. 9C, D). The performance of the ISLE is superior in predicting disease progression after eight weeks of treatment (t-test $P =$ 1.5E-3) compared to the original EMT signature (t-test $P =$ 0.052)67, and on par with other supervised clinical drug response predictors68 (see Supplementary Note 1). Finally, DAISY or ncSL networks failed to predict the clinical response to erlotinib (log rank $P > 0.56$).

Third, we analyzed an International Cancer Genome Consortium (ICGC) cohort of 80 ovarian cancer patients treated with taxane-cisplatin chemotherapy67. The corresponding cSL partners are underexpressed in responders compared to non-responders, as expected (Wilcoxon rank sum $P < 9.1E-3$, Fig. 4c). Reassuringly, the ISLE-based prediction is significantly better than random or shuffled SL partners (empirical $P < 0.04$ and 0.03, respectively); DAISY-SL or ncSL partners were not predictive either (Wilcoxon rank sum $P > 0.17$).

Finally, we predicted drug response in the TCGA compendium64. We considered six drugs that have response evaluation criteria in solid tumors (RECIST) information following treatment for at least 12 patients (for available cancer types). Four out of these six drugs show significantly higher cSL-scores in responders’ tumors (blue bars) than in those of non-responders (red bars) (Fig. 4d, empirical $P < 0.05$). (Notably, the cSL network was inferred only from the samples that have no drug treatment record (N = 6268) so that the testing is performed strictly on unseen samples, Methods.) The predictive signal is completely lost when using random, shuffled, DAISY or ncSL partners (both experimentally identified and inferred).

Discussion

Many SL lab screens have led to the discovery of numerous important leads for further in vivo follow-up. However, our analysis revealed that only a small fraction of the SLs candidates emerging from such in vitro screens analysis show an evidence of negative selection or effect on patients’ prognosis in the TCGA cohort (Supplementary Note 1). This realization has led us to develop ISLE, a method for identifying the subset of these in vitro SL pairs that are more likely to be clinically relevant. As we have shown throughout the paper, by and large cSL pairs enable successful prediction of response to a wide panel of targeted therapies in cancer patients, which is markedly superior to the predictive power of ncSL pairs.

The clinical significance estimation of the initial pool of candidate SLs should be distinguished between two different cases: (1) First, in the case of the specific candidate SLs that arise from isogenic (or double-knockout) cell lines screens, we specifically performed their clinical significance estimation by considering
the corresponding mutation (or copy number and expression) data in the tumors and also by scanning the specifically relevant cancer types in which they have been inferred, so we believe that their clinical significance estimations are quite robust. (2) Second, the clinical significance estimations of the candidate SLs emerging from the single-knockdown screens are however probably less tight, as they are both tested and inferred via a pan-cancer analysis (Methods), while some of these candidate SLs may still be of clinical value in specific cancer types and contexts (see Supplementary Note 1 for details).

The cSLs inferred are identified by analyzing a large cohort of untreated cancer samples, and then utilized to predict the response to cancer drugs without any further training, showing strong predictive power in a wide range of different data types (Supplementary Data 14). The absence of specific training on post-treatment data reduces the risk of over-fitting and poor generalization and enables the prediction of response to new, untested drug candidates. ISLE’s performance level is on par with numerous supervised prediction approaches employed for drug response prediction in the DREAM challenge (Fig. 2c). Importantly, it provides straightforward predictive cSL signatures for each drug, which can be used to guide the selection of cell lines and models for further experimental screens, and for future patient stratification.

Like any other genome-wide computational prediction method, ISLE has several limitations that should be acknowledged: (1) first, the tumor data are noisy, both on the molecular and the survival side. To achieve a strong and robust predictive signal, we infer a common pan-cancer network by combining data from different tumor types, while obviously there is

Fig. 4 Drug-cSL-network predicts treatment outcomes in cancer patients. a The KM plot of predicted responders (blue) vs non-responders (red) to taxane-anthracycline chemotherapy. We divided the patients into responders vs non-responders based on the median value of their cSL-scores. b The gene expression of the cSL partners in patients treated with erlotinib, ordered according to their months-to-progression (on-top). As predicted, patients with many downregulated cSL partners progressed slower. c The responders (blue) to taxane-cisplatin therapy for ovarian cancer show significantly higher ISLE cSL-scores than the non-responders (red) (Wilcoxon rank sum $P < 9.1E-3$). The X-axis shows the different groups of partners studied, and Y-axis provides their cSL-scores in responders vs non-responders. d TCGA patients with a large number of downregulated drug-cSL partners in their tumors show better response based on RECIST criteria. X-axis shows six drugs that have considerable (>12 samples) drug response information in TCGA, and Y-axis represents the cSL-score (divided by total number of SL partners to guide visualization, mean and s.e.m) of their drug targets, where cancer types are controlled for (* marks the four drugs that are significantly predicted after multiple hypothesis corrections (FDR-corrected Wilcoxon rank sum $P < 0.2$; and epirubicin FDR < 0.23), drugs are listed in order of significance). Blue (red) bars denote the cSL-scores of the responders (non-responders), and the numbers marked in blue (red) below the figure indicate the number of responders (non-responders) for each drug. All the analyses were performed using the drug-cSL-network (presented in Supplementary Fig. 8) based on the drug-target mapping available listed in Supplementary Data 13.
variability in SLi at different cancer types. (2) Second, the current version of ISLE considers multitude types of evidence to increase the signal and controls for many potential confounders. However, additional ones could be further considered in the future when more data accumulates. (3) Third, by requiring that cSLi pass both molecular and survival filters we may miss true cSLs. Nevertheless, we prefer to take this more conservative approach that minimizes the number of cSLs that turn out not to be clinically relevant even if it incurs a higher rate of missed true interactions. Indeed, the results show that cSLi are markedly more predictive than non-cSLi across the board.

Drugs targeting such genes that are high degree hubs in the cSL network may kill different sub-clones harboring different inactivated cSL partners and may reduce the likelihood of emerging resistance in heterogeneous tumors. With the accelerated accumulation of new patients' data, ISLE may be further improved by analyzing new types of omics data (e.g., more extensive sequence and epigenetic data) and lead to the generation of cancer type/context-specific SL networks. Taken together, ISLE complements existing mutation-based targeted approaches by extending their scope to the whole genome.

**Methods**

**Gene activity and FDR threshold settings.** Throughout the paper, we define the following unless specified otherwise: (A) A gene to be inactive in a sample if its gene expression (or SCNA) is below 1/3-quantile across samples in each cancer type. This is to account for the distinct basal expression levels of the genes in different cancer types (see Supplementary Note 1). We used both transcriptomic and SCNA profiles for identifying SLi using ISLE, but used only transcriptomic profiles for drug response prediction throughout the paper. (B) A candidate SL pair to be significant when FDR-corrected p-value < 0.05 in identifying the cSL network. In the three steps of ISLE, the significant pairs were selected and passed to the next step. In all multiple hypothesis correction, we consistently used a uniform FDR threshold of 0.2.

ISLE. The first step incorporates the union of candidate SL pairs derived from two separate procedures. Initial Set I: the experimentally identified SL were collected from 17 screens, each performed in a single cell line25–27. The screens spanned altogether 154,707 potential pairwise interactions tested in eight different cancer types, among which 6033 pairs compose the positive set (about 4%, see Supplementary Note 1). We excluded from the sets those interactions that focus on activating driver mutation or copy number amplification, such as KRAS14,17, PTTG127, and MYC28, resulting in 4252 unique candidate SL pairs. Initial Set II: the SLi were inferred from five large-scale shRNA/siRNA single gene knockout screens29–31, spanning a total of 9,253,974 measurements in 315 cancer cell lines from 19 different tissue of origin, through ‘guilt-by-association’30,32. By definition, it is expected that gene A will be essential only when its SL partner gene B is inactive in a given cancer cell line. Using a set of input genome-wide shRNA/siRNA screens in a reference collection of cell lines, ISLE examines all current candidate SL pairs to identify those pairs that show conditional essentiality. Gene A is defined as conditionally essential with gene B if its essentiality is significantly higher in the samples where gene B is underexpressed (defined above) using Wilcoxon rank sum test. SCNA based conditional essentiality is determined analogously. The candidate SL pairs that show significance either by mRNA or SCNA data selected in the set (see Supplementary Note 1). The union of Initial Set I and II has resulted in a total of 163,773,526 candidate SL pairs.

First, we identify under-represented SL pairs by analyzing tumor molecular data. The step searches for gene pairs whose co-inactivation are under negative selection. We mine gene expression and SCNA data of input tumor samples to identify gene pairs A and B whose co-inactivation (defined above) is significantly less frequent than expected (identified via a hypergeometric test). Formally, if N is the total number of tumor samples, nA (nB) is the number of samples with gene A (gene B) underactive, respectively, and nAB is the number of samples with co-inactivation of A and B, the significance of depletion was determined by hypergeometric(p עובד, nA, N, nAB) (see Supplementary Note 1). The depletion of the activity state using SCNA is inferred analogously.

Second, we identify survival-informative pairs by analyzing patient survival data. The step selects a gene pair A and B as SL if tumor samples with co-inactive A-B exhibit significantly better patient's survival than tumor samples without the co-inactivation. Specifically, ISLE uses the following stratified Cox proportional hazard model to check this association, while controlling for various confounding factors including the effect of respective genes, cancer types, genomic instability30, sex, age, and race (shown here for expression analysis and a similar model is used to analyze SCNA data):

\[ h(t, \text{patient}) \sim h_0(t) \exp(\beta g(A) + \beta g(B) + \beta_{\text{age}} + \beta_{\text{GII}}). \]

where g is an indicator variable over all possible combinations of patients’ stratifications based on cancer type, race, and sex, \( h(t) \) is the hazard function (defined as the risk of death per patients at each event), and \( h_0(t) \) is the baseline-hazard function at time t of the g-stratification. The model contains four covariates: (i) \( G(A, B) \): indicator variable if the SL is functionally active in the patient’s tumor; (ii) g (A) and (iii) g(B): gene expression of A and B, (iv) age: age of the patient, (v) GII: genomic instability index. GII measures the relative amplification or deletion of geno in SL pairs based on the SCNA. Given b, j the absolute of log ratio of SCNA of gene i in a sample relative to normal control, GII of the sample is given as in Bilal et al.40.

The \( \beta \)s are the regression coefficient parameters of the covariates, which quantify the effect of covariates on the survival. All covariates are normalized to 0 (1). The \( \beta \)s are determined by standard likelihood maximization of the model50 using the R-package “Survival”. The significance of \( \beta \), which is a measure of the interaction term is determined by comparison of the likelihood of the model with the null model without the interaction indicator \( G(A, B) \) followed by a likelihood ratio test70, i.e.,

\[ h_{\text{null}}(t, \text{patient}) \sim h_0(t) \exp(\beta g(A) + \beta g(B) + \beta_{\text{age}} + \beta_{\text{GII}}). \]

The p-value obtained by the likelihood ratio test is corrected for multiple hypotheses. Pairs exhibiting the significant survival improvement either in gene expression or SCNA are passed on to the next screen.

Third, we identify phylogenetically linked SL pairs by analyzing phylogenetic profile of SL pairs based on to be conserved across different species.52–55. Based on this notion, we further filter and select SL pairs composed of genes having high phylogenetic similarity. This is done by comparing the phylogenetic profile of the two genes A and B in a candidate SL pair across 86 species (adopting the method of Tabach et al.40,56). We then calculate the phylogenetic similarity between A and B using a single-novice matrix from NanoMint57, which measures evolutionary distance while taking into account their phylogeny. To determine the threshold used for this step, we used the median phylogenetic similarity of the pairs in Initial Set I as it optimally separates the positive and negative sets among these pairs (see Supplementary Note 1).

ISLE-based drug response prediction. For drug response and essentiality predictions, we applied ISLE to a limited set of candidate SL pairs between the inactivated genes (either by drug or gene knockout) and all protein-coding genes, performing multiple hypothesis correction at each drug or gene level with a single uniform FDR threshold of 0.2 throughout the paper. The drugs were mapped to their targets primarily based on DrugBank58, and we referred other sources such as CCLE59, GDSNC60 and the literature with exception to target genes whose mechanism of action is explicitly denoted as an agonist in DrugBank (Supplementary Data 13). We excluded those drugs whose target information is not concrete from such databases and the literature. As an example, Supplementary Fig. 8 depicts the drug-cSL-network that was used for predicting drug response in patients, that covers 232 SLi involving 14 drug targets (blue), their corresponding 16 drugs (green), and their 207 SL partners (red).

Throughout the manuscript, we made predictions for drug response to (1) single agents and (2) drug combinations using ISLE. We focused on inhibitory compounds with specific targets because only targeted inhibitors are relevant to SL. For single agents, we use the following procedure, which we term Procedure (1). We hypothesize that a drug will be more effective in cell lines/samples that have more underactive cSL partners of its targets. To this end, we (i) identified the SL partners of the genes targeted by each of the drugs (using ISLE pipeline as described above), (ii) defined cSL-score for individual samples/cell lines (defined below) and (iii) used cSL-score to predict the drug response and evaluated the prediction accuracy by comparing it with experimentally/clinically measured drug response data.

\[ \text{cSL-score} = \sum_{i=1}^{N} \left( \frac{\text{mRNA}_{i}}{p_{i}} \right) / N_{\text{T}}, \]

where cSL-score, denotes the cSL-score of the jth sample, mRNA, is the pre-treatment gene expression of jth SL partners in the jth sample; \( p_{i} \) is the 1/3-quantile threshold of the cancer type to which the jth sample belongs, \( f(t) \) is an hazard model to check this association, while controlling for various confounding factors including the effect of respective genes, cancer types, genomic instability30, sex, age, and race (shown here for expression analysis and a similar model is used to analyze SCNA data):
For drug combinations, we use the following procedure, which we term Procedure (2). ISLE SL-network was applied to predict synergistic drug combinations based on the working hypothesis that a synergy between two compounds arises from an underlying SLI between the gene targets of each of the two drugs. Accordingly, (i) we applied ISLE to all possible pairs between the drug-target genes, (ii) predicted the compound synergy using the best cSL-pair-score (defined below) between their target genes, and (iii) evaluated the prediction accuracy versus experimentally measured synergy.

cSL-pair-score = \( r_{\text{mean}} + r_{12} + r_{13} + r_{23} \)

(5)

where cSL-pair-score is a qualitative measure combining the significance levels at the initial candidate SL pool and the subsequent three statistical tests, \( r_{\text{mean}} \) denotes the significance based on the experimental in vitro mixtures, \( r_{12}, r_{13}, \) and \( r_{23} \) denote the rank-normalized values (between 0 and 1, with 1 representing a pair with the highest significance and 0 with the lowest) of the statistical significance levels across all genes, respectively. We tested our hypothesis using the Wilcoxon rank sum test with FDR < 0.2 and fold change >0.35

Validating ISLE-based drug combination screens. We followed Procedure 2 (see ‘ISLE-based drug response prediction’ in Methods above) and evaluated the prediction accuracy based on ROC analysis and precision-recall statistics. We applied this approach to two recent DREAM challenges57,58, which contains drug combinations that involve 535 target gene pairs in 85 cancer cell lines (AstraZeneca-Sanger Drug Combination Prediction DREAM challenge 2015, different from the DREAM7 challenge). In this dataset, the single and double-drug response were measured across multiple dose conditions, and a synergism was defined by the difference between the actual combination effect and the lowest response from the respective single agents. This was implemented based on Loewe model, summarized into a single synergy score57,58,59 (see the Data Description of the AstraZeneca-Sanger Drug Combination Prediction DREAM challenge60). We labeled a drug pair to be synergistic based on its maximal synergy score.

For drug combination, we use the same approach to a collection of in vivo mouse xenograft models that cover 26 drug combinations in 375 samples59. We define a drug pair is synergistic if it shows synergism in more than 2/3-quantile of the samples it was tested. In each sample, a drug pair is synergistic, if the double treatment effect is greater than the summation of the single treatments. Specifically, we used the BestAverageResponse (BestAvgResponse), which measures the averaged value of the maximal tumor size reduction over multiple time points (see Methods in Gao et al59) to quantify the effect of the double and single treatment.

Testing cSL network with functional and drug screens. We use liv7k cell line, an aggressive (T3N2b), HPV-negative cell line derived from a primary tumor of the tongue, in a patient who received no chemo/radiotherapy before tumor excision. Generously gifted by Professor Richard Shaw (Head and Neck Cancer Consultant, Liverpool, UK). Cells are maintained in keratinocyte SFM (Life Technologies), supplemented with 2 mM L-glutamine, 0.2 ng/ml epidermal growth factor (EGF) and 25 µg/ml bovine pituitary extract (BPE). The cell line was tested for mycoplasma contamination and authenticated by exome sequencing and CNV.

For drug repurposing screen, we used a panel of drug candidates from FDA-approved and clinically tested agents (NIH Clinical Collection (Evotec, San Francisco, CA), LOPAC Pfizer (Sigma Aldrich), and the FDA-Approved Drug Library (Selleck Chemicals)) were tested at a single concentration of 10 µM dimethyl sulfoxide (DMSO). We focused on inhibitor compounds that have single target genes (\( N = 237 \)). Compounds were first diluted 1:50 in serum-free media, and then further diluted 1:20 into 96-well plates containing 5000 cells in 190 µl of keratinocyte SFM as described. Vehicle control (0.1% DMSO) was used as a negative control, and 1 µM staurosporine (Sigma Aldrich) was used as a positive cell lethality control in all plates. Each drug was tested for 72h under 21 and 1% O2. Drugs were fixed using 4% paraformaldehyde and stained with lactate dehydrogenase (Sigma Aldrich). We assess viability by quantifying the number of stained nuclei using the Operetta High-Content Imaging System (Perkin Elmer).

For gene knockout screening in liv7k cell line, transfections were performed in 384-well plates in 21% O2 and 0.1% O2. A complex of siRNA (Dharmacon, GE Healthcare) and 0.125 µL RNAiMAX (Life Technologies) was added to each well and incubated for 30 min (final siRNA concentration 25 nM). 5000 cells were added per well in keratinocyte serum-free medium. The screen was conducted using siGnome RNAi pools and included All-Stars Cell Death control (Qiagen) and ON-TARGETplus Non-targeting Control Pool (Dharmacon, GE Healthcare). Cells were treated with DAPI 72h later. The number of cells per well was counted using the Operetta High-Content Imaging System (Perkin Elmer).

For drug response screen in breast cancer cell lines, a panel of 7 breast cancer cell lines was screened using the drug repurposing libraries and workflow detailed for the liv7k cell line, under normoxic conditions only. All cell lines were maintained in Dulbecco’s Modified Eagle Medium (DMEM) containing 10% FBS and 10% glutamine, with MCF10A cells receiving additional EGF (20 ng/ml), hydrocortisone (0.5 mg/ml), cholera toxin, (100 ng/ml) and insulin (10 µg/ml) supplementation. Cells were seeded at concentrations which resulted in approximately 80% well confluence after 72 h incubation (2000 > 8000).

For gene essentiality and drug response prediction, a gene was marked as underactive if its expression level is less than 1/3-quantile of its level across all CCLE cell lines (\( N > 1000 \)) (see Supplementary Note 1). The growth inhibition effect of the genes over both the conditions was measured using the Wilcoxon rank sum test with FDR < 0.2 and fold change >0.35. We thus used the cSL-pair-score of the cSL partners as the predictor for synergism. We followed Procedure 1 (see ‘ISLE-based drug response prediction’ in Methods above) using drug-cSL-network to make ISLE-based drug response predictions, and evaluated and compared the performance using weighted probabilistic concordance index (wpc-index), which is the collective measure of concordance index for each drug, taking the variance of the response to the individual drugs into account, as defined in Bansal et al36,61.
Testing cSL-based drug synergy in patient-derived cell lines. All DNA samples used in this study were derived from patient samples. Samples used for whole-exome capture were extracted from cell lines established directly from patient tumors as described previously. Briefly, a panel of pathology-confirmed metastatic melanoma tumor resections, paired with apheresis-collected peripheral blood mononuclear cells, were collected from patients enrolled in IRB-approved clinical trials at the Surgery Branch of the National Cancer Institute. Pathology-confirmed melanoma cell lines were derived from mechanically or enzymatically dispersed tumor cells, which were then cultured in RPMI 1640 + 10% FBS at 37 °C in 5% CO2 for 5–15 passages. Genomic DNA was isolated using DNeasy Blood & Tissue kit (Qiagen, Valencia, CA). For all samples, matching between germline and tumor DNA was verified by direct sequencing of 26 single nucleotide polymorphisms (SNP) at 24 loci. A subset of cell lines used in the study was derived from a panel of cell lines established directly from patient tumors as described previously. For drug response prediction, we followed Procedure 1 (see "ISLE-based drug response prediction" in Methods above). We focused on 237 and 92 single-target drugs respectively for liv7k and BC cell lines. We considered %Growth Inhibition > 33% as a positive set (responsive), and the rest as a negative set (non-responsive). We performed ROC and precision-recall analysis using the cSL-score as a prediction for drug response. We performed an analogous analysis to predict drug response in breast cancer cell lines. To predict the conditional essentiality in liv7k cell line, we compared the cSL-score of the drug targets and the drug’s effectiveness in hypoxic and normoxic conditions (see Supplementary Note 1). The prediction accuracy was quantified by simple fold change comparison between our prediction and measured drug response values associated to the feature set described above.

References

1. Hartwell, L. H., Szakasitz, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Integrating genetic and drug sensitivity data to identify drug candidates for cancer therapy. 

2. D’Amore, A., Lowe, S. W. & Helleday, T. Cancer genes and the cell cycle. 

3. Jope, R. S., Milligan, G. & Soderling, T. R. The PI3K/Akt pathway in diabetes and cancer. 

4. Awamleh, L. M., Hay, N., Thompson, J. R. & Khan, A. The cdk5 promoter is a positive regulator of cell growth and glucocorticoid receptor-regulated transcription. 

5. Ashworth, A., Lord, C. J. & Reis-Filho, J. S. Genetic interactions in cancer progression and treatment. 

6. Beijersbergen, R. L., Wessels, L. F. A. & Bernards, R. Synthetic lethality in cancer therapeutics. 

7. Senft, D., Leiserson, M. D. M., Ruppin, E. & Ronai, Z. A. Precision oncology: using machine learning and network properties. 

8. Madhukar, N. S., Elemento, O. & Pandey, G. Prediction of genetic interactions associated with the key melanoma drivers. 

9. Bommi-Reddy, A. et al. Kinase requirements in human cells: III. Altered expression of MSH2. 

10. Luo, J. et al. A genome-wide RNAi screen identifies therapeutic targets for MYC-driven cancer. 

11. Lord, C. J., McDonald, S., Swift, S., Turner, N. C. Springer & Cell 145, 30–38 (2011). 

12. Beijersbergen, R. L., Wessels, L. F. A. & Bernards, R. Synthetic lethality in cancer therapeutics. Annu Rev. Cancer Biol. 1, 141–161 (2017). 

13. Senft, D., Leiserson, M. D. M., Ruppin, E. & Ronai, Z. A. Precision oncology: using machine learning and network properties. 

14. Madhukar, N. S., Elemento, O. & Pandey, G. Prediction of genetic interactions using machine learning and network properties. 

15. Whitehurst, A. W. et al. Synthetic lethal screen identification of chemo-resistant loci in cancer cells. Nature 469, 817–819 (2007). 

16. Bommi-Reddy, A. et al. Kinase requirements in human cells: III. Altered kinase requirements in VHL−/− cancer cells detected in a pilot synthetic lethal screen. Proc Natl Acad Sci USA 105, 16484–16489 (2008). 

17. Lord, C. J., McDonald, S., Swift, S., Turner, N. C. Springer & Cell 145, 30–38 (2011). 

18. Martin, S. A. et al. Methotrexate induces oxidative DNA damage and is selectively lethal to tumour cells with defects in the DNA mismatch repair gene MSH2. 

19. Martin, S. A. et al. Methotrexate induces oxidative DNA damage and is selectively lethal to tumour cells with defects in the DNA mismatch repair gene MSH2. EMBO Mol. Med. 1, 323–337 (2009). 

20. Luo, J. et al. Genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncoprote. Cell 137, 835–848 (2009). 

21. Steckel, M. et al. Determination of synthetic lethal interactions in KRAS oncopogene-dependent cancer cells reveals novel therapeutic targeting strategies. Cell Res. 22, 1227–1245 (2012). 

22. Toyoshima, M. et al. Functional genomics identifies therapeutic targets for MYC-driven cancer. Proc Natl Acad Sci USA 109, 9545–9550 (2012). 

23. Vitezcoumar, F. J. et al. A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. Mol. Syst. Biol. 9, 696 (2013). 

24. Wang, X., Fu, A. Q., McNerney, M. E. & White, K. P. Widespread genetic epistasis among cancer genes. Nat. Commun. 5, 4828 (2014). 

25. Bloumen, V. A. & D. R. A. Gene essentiality and synthetic lethality in haploid human cells. Science 350, 1092–1096 (2015).
20. Shen, J. P. et al. Chemogenetic profiling identifies RAD17 as synthetically lethal with checkpoint kinase inhibition. Oncotarget 6, 35755–35769 (2015).
21. Pathak, H. B. et al. A synthetic lethality screen using a focused siRNA library to identify sensitizers to dasatinib therapy for the treatment of epithelial ovarian cancer. PLoS One 10, e0144126 (2015).
22. Srivats, R. et al. A network of conserved synthetic lethal interactions for generation of new cancer therapeutics. Cell 63, 514–525 (2016).
23. Wang, T. et al. Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncopgenic Ras. Cell 168, 890–903 e15 (2017).
24. Han, K. et al. Synergistic drug combinations for cancer identiﬁed in a CRISPR screen for pairwise genetic interactions. Nat. Biotechnol. 35, 463–474 (2017).
25. Shen, J. P. et al. Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. Nat. Methods 16, 573–576 (2017).
26. Cheung, H. W. et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-speciﬁc dependencies in ovarian cancer. Proc. Natl Acad. Sci. USA 108, 12372–12377 (2011).
27. Marcotte, R. et al. Essential gene profiles in breast, pancreatic, and ovarian cancer cells. Cancer Discov. 2, 172–189 (2012).
28. Cowley, C. S. et al. Parallel genome-scale loss of function screens in 216 cancer cell lines for the identiﬁcation of context-speciﬁc dependencies. Sci. Data 1, 140035 (2014).
29. Marcotte, R. et al. Functional genomic landscape of human breast cancer drivers, vulnerabilities, and resistance. Cell 164, 293–309 (2016).
30. Auguet, A. L. et al. A genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. Cancer Discov. 6, 914–929 (2016).
31. Wong, S. et al. Combining biological networks to predict genetic interactions. Proc. Natl Acad. Sci. USA 101, 15682–15687 (2004).
32. Zhong, W. & Sternberg, P. W. Genome-wide prediction of C. elegans genetic interactions. Proc. Natl Acad. Sci. USA 109, 12132–12137 (2012).
33. Lu, X., Kensche, P. R., Huynen, M. A. & Notebaart, R. A. Genome evolution on biological networks. BMC Syst. Biol. 10, 17 (2009).
34. Boucher, B. & Jena, S. Genetic interaction networks: better understand to better predict. Front. Genet. 4, 290 (2013).
35. Lu, X., Kensche, P. R., Huynen, M. A. & Notebaart, R. A. Genome evolution predicts genetic interactions in protein complexes and reveals cancer drug targets. Nat. Commun. 4, 2124 (2013).
36. Conde-Pueyo, N., Munteanu, A., Sole, R. & Rodríguez-Caso, C. Human synthetic lethal inference as potential anti-cancer target gene detection. BMC Syst. Biol. 3, 116 (2009).
37. Folger, O. et al. Predicting selective drug targets in cancer through metabolic networks. Mol. Syst. Biol. 7, 301 (2011).
38. Frezza, C., Pollard, P. & Gottlieb, E. Inborn and acquired metabolic defects in cancer. J. Mol. Med. 89, 213–220 (2011).
39. Wang, K. et al. Genome-wide identiﬁcation of post-translational modulators of transcription factor activity in human B cells. Nat. Biotechnol. 27, 829–839 (2009).
40. Kim, J. W. et al. Characterizing genomic alterations in cancer by complementary functional associations. Nat. Biotechnol. 34, 539–546 (2016).
41. Apaolaza, I. et al. An in-silico approach to predict and exploit synthetic lethality in cancer metabolism. Nat. Commun. 8, 459 (2017).
42. Szczurek, E., Misra, N. & Vingron, M. Synthetic sickness or lethality points at candidate combination therapy targets in glioblastoma. Int. J. Cancer 133, 2123–2132 (2013).
43. Lu, X., Meghelenbrink, W., Notebaert, R. A. & Huynen, M. A. Predicting human genetic interactions from cancer genome evolution. PLoS ONE.10, e0125795 (2015).
44. Park, S. & Lehner, B. Cancer type-dependent genetic interactions between cancer driver alterations indicate plasticity of epistasis across cell types. Mol. Syst. Biol. 11, 824 (2015).
45. Jerby-Arnon, L. et al. Predicting cancer-speciﬁc vulnerability via data-driven detection of synthetic lethality. Cell 158, 1199–1209 (2014).
46. Sinha, S. et al. Systematic discovery of mutation-speciﬁc synthetic lethals by mining pan-cancer human primary tumor data. Nat. Commun. 8, 15580 (2017).
47. Lord, C. J. SpringerAmpamp; Ashworth, A. PARP inhibitors: synthetic lethality in the clinic. Science 355, 1152–1158 (2017).
48. Barbie, D. A. et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462, 108–112 (2009).
49. Cancer Genome Atlas Research, N. et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. 45, 1113–1120 (2013).
50. Bilal, E. et al. Improving breast cancer survival analysis through competition-based multidimensional modeling. PLoS Comput. Biol. 9, e1003047 (2013).
51. Ryan, C. J. et al. Hierarchical modularity and the evolution of genetic interactions across species. Mol. Cell 46, 691–704 (2012).
52. Ma, Y. & Kirschner, D. E. A comparative genomic approach for identifying synthetic lethal interactions in human cancer. Cancer Res. 73, 6128–6136 (2013).
D. E.R. wrote the paper. J.S.L. and A.D. contributed equally to each other; L.J.-A. and R. A. contributed equally to each other; and Y.S. and E.S. contributed equally to each other.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04647-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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