Briefings in Bioinformatics, 2025, **26(3)**, bbaf255
https://doi.org/10.1093/bib/bbaf255
Problem Solving Protocol

TARGET-SL: precision essential gene prediction using driver prioritisation and synthetic lethality

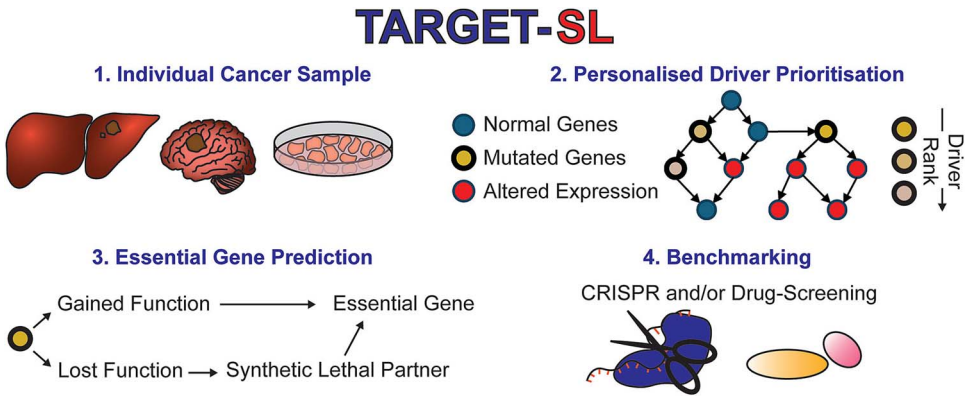
Rhys Gillman[1,2], Matt A. Field[1,2,3,4], Ulf Schmitz[1,2,5], Lionel Hebbard[1,2,6,7,‡,*]

[1]Department of Biomedical Sciences and Molecular and Cell Biology, College of Medicine and Dentistry, College of Science and Engineering, James Cook University, 1 James Cook Drive, Townsville, Queensland, Australia
[2]Centre for Tropical Bioinformatics and Molecular Biology, James Cook University, 14-88 McGregor Road, Smithfield, QLD 4878, Australia
[3]Immunogenomics Lab, Garvan Institute of Medical Research, 384 Victoria St, Darlinghurst, NSW 2010, Australia
[4]Menzies School of Health Research, Charles Darwin University, Red 9, Casuarina campus, Univ Drive North, Casuarina, NT 0811, Australia
[5]Faculty of Medicine & Health, The University of Sydney, Camperdown, Camperdown Campus, Parramatta Road, Sydney, NSW 2006, Australia
[6]Storr Liver Centre, Westmead Institute for Medical Research, Westmead Hospital and University of Sydney, 176 Hawkesbury Rd, Westmead, NSW 2145, Australia
[7]Australian Institute for Tropical Health and Medicine, 1 James Cook Drive, Townsville, QLD 4811, Australia
*Corresponding author. Department of Biomedical Sciences and Molecular and Cell Biology, College of Medicine and Dentistry, College of Science and Engineering, James Cook University, 1 James Cook Drive, Townsville, QLD 4811, Australia. E-mail: lionel.hebbard@jcu.edu.au
‡Present Address: Discipline of Molecular and Cell Biology Townsville, Queensland 4811, Australia

Abstract

The ability to identify patient-specific vulnerabilities to guide cancer treatments is a vital area of research. However, predictive bioinformatics tools are difficult to translate into clinical applications due to a lack of *in vitro* and *in vivo* validation. While the increasing number of personalised driver prioritisation algorithms (PDPAs) report powerful patient-specific information, the results do not easily translate into treatment strategies. Critical in addressing this gap is the ability to meaningfully benchmark and validate PDPA predictions. To address this, we developed Tumour-specific Algorithm for Ranking GEnetic Targets via Synthetic Lethality (TARGET-SL), which utilises PDPA predictions to produce a ranked list of predicted essential genes that can be validated *in vitro* and *in vivo*. This framework employs a novel strategy to benchmark PDPAs, by comparing predictions with ground truth gene essentiality data from large-scale CRISPR-knockout and drug sensitivity screens. Importantly TARGET-SL identifies vulnerabilities that are more exclusive to individual tumours than predictions based on canonical driver genes. We further find that TARGET-SL is better at identifying sample-specific vulnerabilities than other similar tools.

Graphical Abstract



Keywords: driver gene prediction; drug prediction; bioinformatics; personalised medicine; cancer

Received: January 29, 2025. **Revised:** April 17, 2025. **Accepted:** May 12, 2025

Introduction

Cancer drug development is increasingly focusing on personalised treatments targeting patient-specific cancer vulnerabilities that arise from genetic alterations. This requires clinicians to match patients to therapies based on curated lists of actionable mutations and their corresponding treatment. However, these lists are often short and based on frequently recurrent mutations in each cancer type, considered canonical 'driver genes'. Resultingly, if a patient's mutation is not on the list, they will unlikely benefit from personalised treatment. Furthermore, this approach assumes that mutations in canonical driver genes must be driver mutations, though this is not necessarily the case [1].

Thus, there is a clear need for algorithmic approaches to prioritise patient-specific target genes based on individual tumour molecular profiles. Personalised driver prioritisation algorithms (PDPAs) attempt to achieve this by combining genomic and transcriptomic data to link genetic alterations with the cancer phenotype, prioritising impactful driver mutations [2]. These driver mutations result in increased activity through gain-of-function (GOF) mutations in oncogenes, or decreased activity via loss-of-function (LOF) mutations in tumour suppressor genes (TSGs). Despite this critical difference, PDPAs do not distinguish between GOF and LOF mutations.

Evaluating driver gene predictions in a personalised context requires prohibitively extensive experimental validation. This has led to difficulties in performing accurate benchmarking, with most approaches suffering key limitations [3]. Firstly, PDPAs are typically evaluated based on their ability to identify canonical driver genes, such as the Cancer Gene Census (CGC) [4] genes, in a publicly accessible cancer cohort. This strategy contradicts the personalised nature of these algorithms, involving no patient-specific ground truth. Secondly, a rank-aggregate-evaluate strategy is often utilised, where the predicted drivers from a cohort are aggregated into a single consensus list before comparison with the reference set, biassing the results by increasing the likelihood of finding commonly occurring driver genes. To address this, some authors have suggested a rank-evaluate-aggregate (REA) strategy [5], where precision is calculated before averaging the results. Finally, the varying input requirements of various PDPAs makes cross-algorithm comparisons challenging. This includes weighted or unweighted, and directed or undirected gene interaction networks (GINs), and the need for both tumour and normal data. These limitations have resulted in poor utilisation and development of these tools.

The Cancer Cell Line Encyclopedia (CCLE) [6, 7] offers a potential solution to this problem. CCLE includes over a thousand cell lines with genomic and transcriptomic sequencing, as well as drug-sensitivity screening and genome-wide CRISPR knock-out (CRISPR-KO) gene essentiality screening, which identify genes critical to cell growth. No previous PDPA evaluation strategy has utilised gene essentiality screening. This is because not all driver genes are essential. Indeed, while GOF driver mutations promote cell growth, usually making them essential genes, LOF mutations occur in tumour-suppressor genes, which normally suppress cell growth. This means that LOF drivers themselves are not essential genes. However, given that synthetic lethality (SL) describes a lethal relationship upon loss of two partner genes, we hypothesise that the SL partners of LOF drivers are essential genes. Moreover, extensive research has resulted in a wealth of quality databases of both predicted and experimentally validated SL pairs [8–15], many of which have been previously reviewed [16].

To test our hypothesis, we designed the Tumour-specific Algorithm for Ranking GEnetic Targets via Synthetic Lethality (TARGET-SL), which uses PDPA rankings to predict essential genes and drug sensitivity based on known or predicted SL databases and variant effect databases. Using the TARGET-SL benchmarking mode, predictions are compared to ground truth information for CCLE essentiality and drug sensitivity screens, which are evaluated using a REA approach. Here, we present TARGET-SL as a novel tool for gene essentiality and drug sensitivity prediction, which is highly specific to individual target tumours or cell lines, while also being an evaluation framework for PDPAs.

Material and methods

Overview

TARGET-SL is a gene-essentiality and drug-sensitivity prediction tool (Fig. 1a), and an evaluation framework for PDPAs (Fig. 1b), taking ranked, sample-specific driver gene predictions, and converting them into essential gene and drug-sensitivity predictions. In benchmark mode, it uses gene essentiality and drug screening data as ground-truth for evaluation.

Benchmark data

Input data

CCLE data was sourced from DepMap (www.depmap.org, release 22Q4) [7], including expression data [transcripts-per-million (TPM) and RSEM expected counts], somatic mutations, and gene-level copy number. X-chromosome copy numbers were corrected for sex (Supplementary Methods and Supplementary Fig. S1). STRING database (v11) [17] was used as a GIN for all algorithms. Data were filtered (Supplementary Methods and Supplementary Fig. S2) such that all PDPAs had access to identical data, amounting to 1290 cell lines across 22 cell types with 3773 genes. The final GIN contained 30 319 unique edges and 3773 nodes.

Gene Effect

Project Achilles genome-wide CRISPR-KO gene effect data were retrieved from DepMap (22Q4), whereby greater negative magnitude means more restricted cell growth. This was filtered to genes with mean TPM > 5. Gene effect was further converted into a gene uniqueness index (UI_G) (Supplementary Methods), whereby a more negative value indicates less cell growth relative to other cell lines, particularly those of the same cell type.

Drug sensitivity

Drug sensitivity data came from PRISM log-fold change (LFC) viability (DepMap 23Q2), and GDSC $lnIC_{50}$ (www.cancerrxgene.org, Release 1 and 2) [18]. These were filtered to drugs with known genetic targets and inhibitory action in the Drug-Gene Interaction Database (DGIdb) [19] by filtering for the following drug-types: antibody, antisense oligonucleotide, blocker, cleavage, inhibitor, inverse agonist, and negative modulator. The final list of drugs included in our analysis and their targets are listed in Supplementary Table S1. GDSC $lnIC_{50}$ was further converted into a drug uniqueness index (UI_D) (Supplementary Methods), whereby a more negative value indicates greater toxicity relative to other cell lines, particularly those of the same cell type.

Ground truth essential gene sets

Pickles (v3) [20] was used to create ground truth essential gene sets for CCLE lines. We took a consensus of essential genes from the Avana [21] and Score [22] essentiality screens from the three

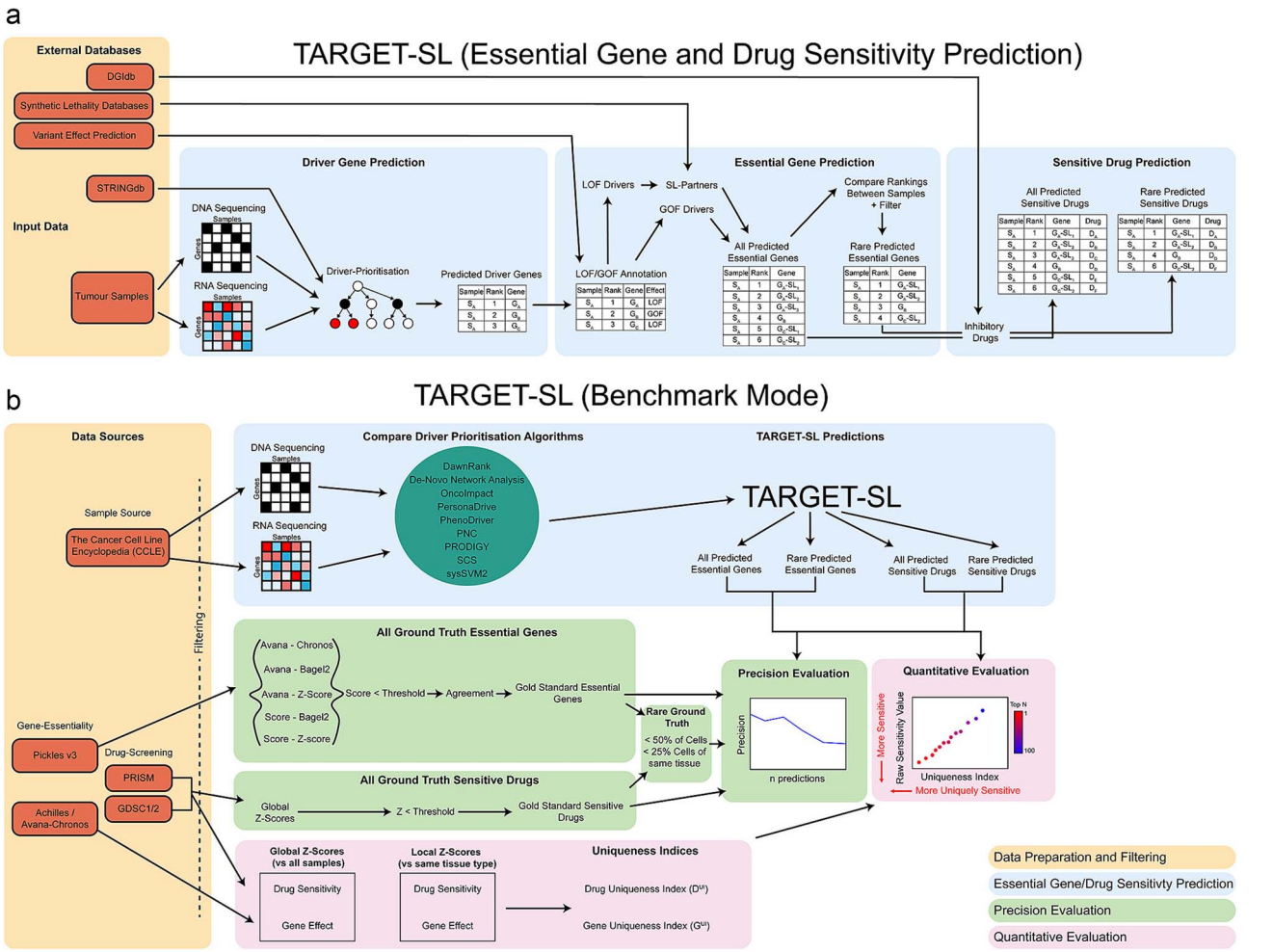


Figure 1. TARGET-SL overview. (a) PDPAs predict driver genes from DNA and RNA sequencing data. Based on variant-effect annotation and SL partner identification using external databases, essential genes are predicted, and further filtered to predict rare essential genes. Finally, known small molecule inhibitors of predicted essential genes are predicted as sensitive drugs. (b) In benchmark mode, TARGET-SL uses PDPAs to predict essential genes and drug sensitivity for CCLE cell lines, with identical input data. Predictions are evaluated with ground-truth data from essentiality screens (PicklesV3 and Achilles/Avana-Chronos) and drug sensitivity screens (PRISM and GDSC1/2).

available methods using the following thresholds: CHRONOS [23] score < -0.5; BAGEL2 [24] score > 8; and z [20] score < -3. This gene set for each cell line was called 'All Ground Truth Essential Genes'. This set was then filtered to only include genes that are essential in <50% of cells globally, and <25% of cells of the same tissue type. We call this set 'Rare Ground Truth Essential Genes'.

Ground truth sensitive drug sets

$\ln IC_{50}$ values (GDSC1, GDSC2) or LFC viabilities (PRISM) were converted into global z-scores based on the mean and standard deviation for all cell lines in the CCLE. Negative values reflected a more detrimental effect to cell growth. Drugs with a z-score < -1 were considered as 'All Ground Truth Sensitive Drugs'. This list was filtered to retain drugs that were sensitive in <50% of cells globally and <25% of cells of the same tissue type, called the 'Rare Ground Truth Sensitive Drugs'.

Essential gene and drug sensitivity prediction

Driver prioritisation

A total of nine PDPAs were utilised in this study (Supplementary Methods and Supplementary Fig. S3). PDPAs were run using default or recommended parameters except for the changes outlined in the Supplementary Methods. A 'PDPA consensus' of DawnRank, OncoImpact, PersonaDrive, and sysSVM2 was created using a modified Borda Count, as well as randomised controls for comparison (randomDriver and randomDrug) (Supplementary Methods).

Tumour-specific algorithm for ranking genetic targets via synthetic lethality

Ranked driver genes were annotated as LOF or GOF drivers using several existing databases of variant effect annotation [4, 25–30] combined to a consensus confidence score (Supplementary Methods and Tables S2 and S3). LOF drivers were replaced with their SL partners (Supplementary Table S4), which were again combined from several existing databases (Supplementary Methods) [8–15] in order of decreasing confidence score, with duplicated essential genes predictions removed. This resulted in a rank-ordered list for each algorithm called 'All Predicted Essential Genes'. These were further filtered to genes with significantly higher ranking in a cell line compared with other cell lines of the same tissue type using a Mann–Whitney U test (P value <.05), creating a list called 'Rare Predicted Essential Genes'. Drug sensitivity predictions were made from these two lists based on drug-gene pairs in our 'ground truth

drug sets' prioritising drugs with the fewest known gene targets. This resulted in the sets 'All Predicted Sensitive Drugs' and 'Rare Predicted Sensitive Drugs', respectively.

Results

Driver prioritisation algorithms produce highly variable predictions

Nine PDPAs were included in our comparison: CSN_NCUA [31], DawnRank [32], OncoImpact [33], PersonaDrive [5], PhenoDriverR [34], PNC [35], PRODIGY [36], SCS [37], and sysSVM2 [38]. These algorithms were used to predict driver genes in 1290 CCLE cell lines using identical input data with STRING (v11) [17] as the underlying GIN.

We assessed the similarity between PDPAs by considering the exclusivity and intersection of their top 1 and top 10 predictions, respectively (Fig. 2a and b). For individual algorithms we considered the proportion of their predictions that were not shared by any other PDPA (exclusive predictions). For the top prediction, the most unique algorithms were PhenoDriver, PRODIGY, and sysSVM2, with 77%, 71%, and 71% of their predictions being entirely exclusive, respectively, while exclusive drivers amounted to just 15% and 12% for PersonaDrive and DawnRank (Fig. 2a). Expectedly, these values were much lower when considering the top 10 predictions, though still at 45% and 37% for SysSVM2 and PRODIGY, respectively (Fig. 2b). For similarity between multiple PDPAs, we considered the proportion of their combined total predictions that were shared (inclusive intersection). The top prediction was shared 64% of the time between PersonaDrive and DawnRank, both of which also had high similarity with OncoImpact (Fig. 2a). This overlap remained high at 33% for the top 10 predictions, matched by a 33% overlap between PNC and CSN_NCUA (Fig. 2b). In summary, the predictions from PDPAs using the same input data were highly variable, with the highest similarity between PersonaDrive and DawnRank, as well as PNC and CSN_NCUA. sysSVM2, PRODIGY, and PhenoDriver gave the most unique predictions.

Additionally, PDPAs varied significantly in the number of reported drivers per sample, evident in their varying set sizes (Fig. 2a and b). For example, PhenoDriver failed to report any driver for 47% of samples, while all the other algorithms generally reported at least one. SCS, OncoImpact, and PRODIGY reported 4.3, 5.4, and 11.5 drivers per sample on average, while the remaining PDPAs reported an average of 25.8 to 102 drivers per sample.

Previously, PDPAs have been benchmarked against canonical driver lists. Thus, we performed a similar analysis using identical input for each PDPA. Predictions were compared against a list of 370 canonical driver genes from the CGC. The top 10 predicted driver genes were assessed incrementally by precision (Fig. 2a) following the approach of previous studies [5, 32, 33, 35–37]. Cell lines with fewer than 10 mutated genes in the CGC list were removed.

PersonaDrive, OncoImpact, and DawnRank had the highest mean precision for their top predictions, which decreased for all PDPAs over increasing numbers of predictions. Given the ground-truth set makes up a large proportion (~10%) of the considered gene list, we suspected that precision may be over-inflated. To address this, we compared predictions with a completely random selection of driver genes (randomDriver) repeated 10 times for every cell line ($n = 12\,900$) with the results averaged. We found that randomDriver predicted CGC driver genes with ~15% precision, consistently lower than all PDPAs.

Given the low agreement between algorithms, we evaluated the performance of a PDPA consensus. Specifically, we evaluated the combination of DawnRank, PersonaDrive, OncoImpact, and SysSVM2. While the PDPA consensus yielded only a slight improvement to PersonaDrive alone (Fig. 2a), the difference was significant after 10 predictions, and all PDPAs performed significantly better than randomDriver (Supplementary Fig. S4).

The algorithms differed significantly in terms of runtime and memory utilisation (Supplementary Fig. S5). DawnRank had the shortest runtime completing analysis of the largest cell line groups in under 2 min, and small sample sizes taking <20 s (Supplementary Fig. S5B). PRODIGY, SCS, and OncoImpact tended to have the longest runtime, taking 15 min to 2 h depending on sample size. OncoImpact had very low memory requirements (maximum 100 MiB), due to its extensive generation of temporary files rather than storing data in memory. The other PDPAs varied from 100 MiB to 10 GiB depending on cohort size (Supplementary Fig. S5A).

TARGET-SL identifies tumour-specific gene targets

TARGET-SL predicts essential genes based on the functional impact of driver gene mutations, by identifying SL-partners of LOF drivers. In benchmarking mode, TARGET-SL allows evaluation of PDPAs by comparing predicted essential genes with drug sensitivity and gene-essentiality screens.

To test our approach, we evaluated whether the PDPA step of the TARGET-SL pipeline offered improvement over simply focusing on canonical CGC drivers. We used TARGET-SL to make gene essentiality and drug sensitivity predictions for each cell line in the CCLE using the top driver prediction from our PDPA consensus. Separately, we used TARGET-SL to make the same predictions but using mutated tier 1 CGC driver genes in each cell line. Finally, as a comparison, we randomly sampled background genes and drugs for each cell line that weren't predicted to be essential or sensitive by either approach. Gene effect and drug sensitivity ($\ln IC_{50}$) was compared between these three groups. In all cases, values with greater negative magnitude indicate higher essentiality or sensitivity. Using the TARGET-SL pipeline, both CGC tier 1 drivers and the PDPA consensus predicted genes and drugs with greater gene effect (Fig. 3a) or sensitivity (Fig. 3b) than the background, respectively, with the CGC tier 1 drivers performing slightly better. Thus, TARGET-SL can convert driver gene-predictions into essential gene predictions, though driver prioritisation did not outperform canonical driver genes in this sense.

However, we hypothesised that PDPAs should lead to predictions that were more unique to individual cell lines. To test this, we converted the gene effect and drug-sensitivity data into uniqueness indices for genes (UI_G) and drugs (UI_D), respectively, which can be interpreted as z-scores. These measures indicate how much more essential or sensitive a prediction was compared to other cell lines, particularly those of the same tissue type, indicating a targeted effect. Based on these metrics, CGC Tier 1-based predictions had very little uniqueness, while the consensus driver prioritisation approach produced significantly more unique predictions (Fig. 3c and d).

Gene essentiality benchmarking of driver prioritisation algorithms

PDPAs were compared in two ways, using a traditional precision-based approach and a novel quantitative benchmark. Firstly, predictions were compared against a ground-truth set of essential genes for each cell line from Pickles (All Ground Truth Essential

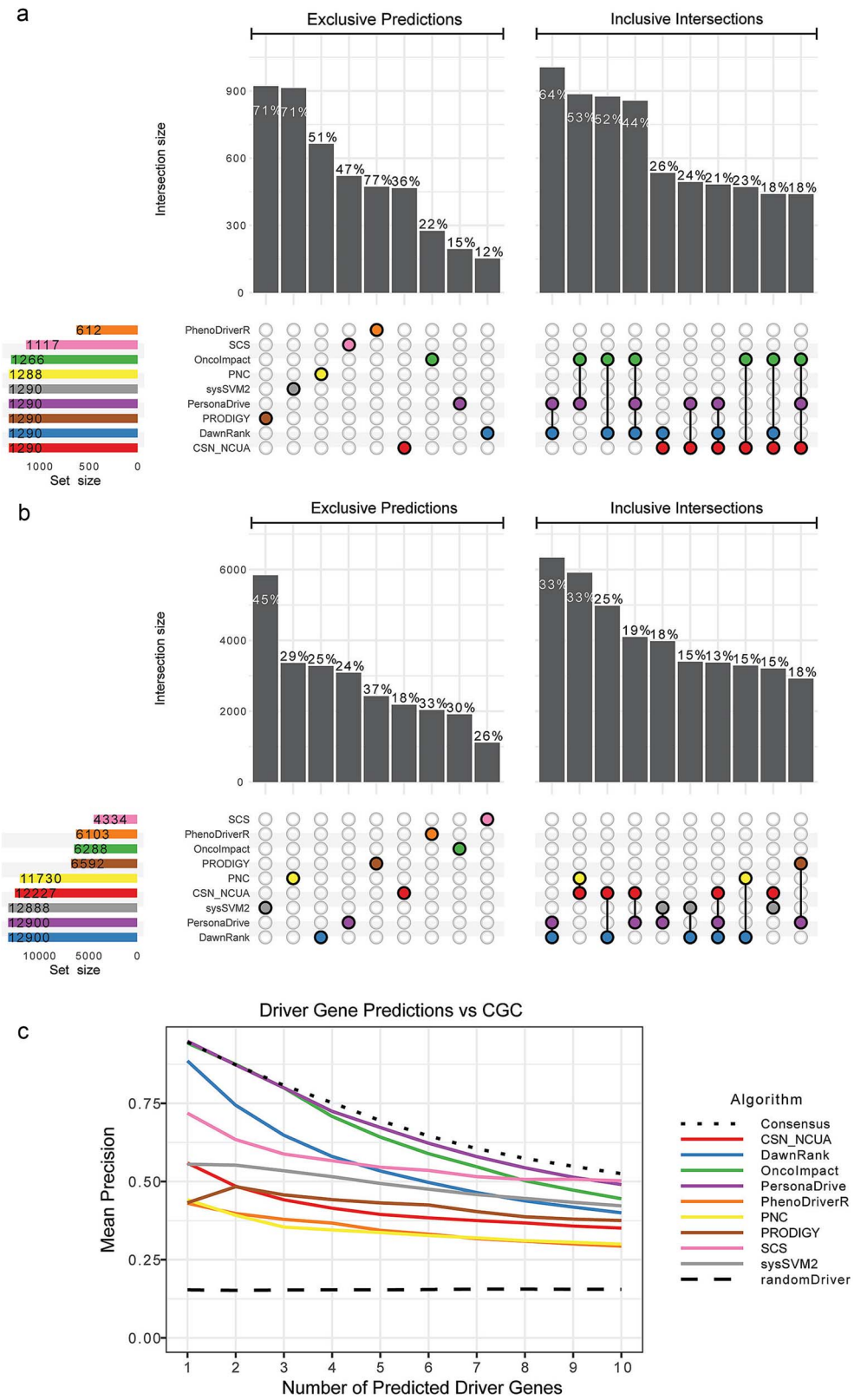


Figure 2. Similarity and traditional benchmark of results. PDPAs were used to predict driver genes in cancer cell lines. (a) Similarity of top driver prediction in each cell line for each algorithm, showing exclusive predictions and inclusive intersections of two or more PDPAs. (b) the same is shown for the top 10 predictions. (c) Mean precision of PDPA top 10 driver predictions using the CGC as ground truth.

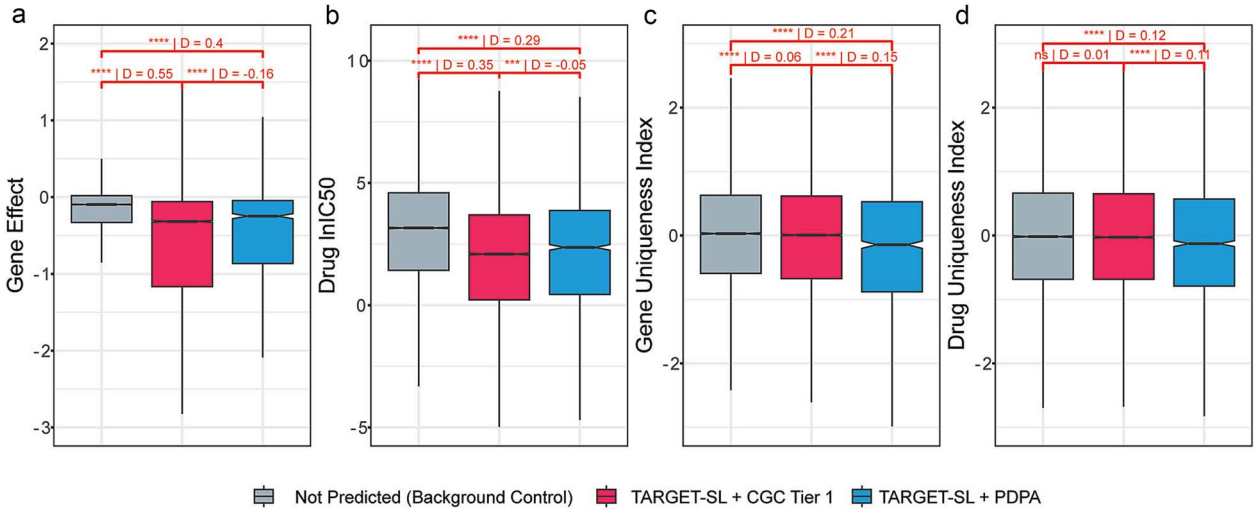


Figure 3. TARGET-SL proof of concept. The TARGET-SL framework was used to predict essential genes (a and c) and sensitive drugs (b and d) using either mutated canonical driver genes from the CGC (Tier 1 only) or the consensus top driver gene prediction for each cell line. As a comparison, cell-gene pairs that were not mutated nor predicted to be essential were collected. Gene effect (a), drug $lnIC_{50}$ (b), gene uniqueness (c) and drug uniqueness (d) indices were compared. Asterisks indicate significance for a Wilcox Rank Sum test (Benjamini–Hochberg corrected), and *D* values indicate Cohen's D effect size. Notches in boxplots indicate ~95% confidence intervals around the median.

Genes). To focus on 'rare predictions', we also compared against a subset of rarer sample-specific essential genes, which we term 'Rare Ground Truth Essential Genes' (see methods for details). Predictions of randomDriver were also processed with TARGET-SL to predict essential genes, making it a control comparison for the PDPA step.

As expected, precision was higher for PDPAs when considering the entire ground truth set (Fig. 4a) as opposed to the smaller rare ground truth set (Fig. 4b). However, in both cases, PDPAs usually outperformed randomDriver. When detecting all essential genes, for the topmost prediction, PDPAs were closely clustered with sysSVM2 and PNC performing slightly better than other algorithms. However, when considering greater numbers of predictions, OncoImpact performed better than other PDPAs, reaching over 50% precision. For predicting rarer essential genes, OncoImpact, PersonaDrive, and sysSVM2 had the best performance, though in all cases their precision decreased rapidly over increasing predictions. The PDPA consensus rarely outperformed any individual algorithm but had consistent performance across increasing numbers of predictions. Statistical comparisons of the top 10 predictions showed that all algorithms performed significantly better than randomDriver for predicting all ground truth genes (Supplementary Fig. S6A). Only the PDPA consensus, CSN_NUCA, PNC, and sysSVM2 were significantly higher for the rare ground truth predictions (Supplementary Fig. S6B).

Secondly, to retain quantitative gene effect data, we compared the gene effect scores of each ranked prediction (Fig. 4c). This was done in terms of the average gene effect itself (*y*-axis) and the average UI_G (*x*-axis), determined cumulatively with increasing numbers of predictions. Here, the best performing PDPAs produce points in the bottom left of the graph, indicating a strong and targeted effect on cell growth. As expected, randomDriver produced points in a single cluster that showed no uniqueness, however, it still produced predictions with negative gene effect as expected due to the TARGET-SL pipeline. PersonaDrive, the PDPA consensus, DawnRank, and OncoImpact predictions had the greatest magnitude UI_G scores, indicating better ability to find sample-specific essential genes. In contrast, CSN_NCUA, PNC, and sysSVM2 produced predictions with a stronger gene effect, but with weaker UI_G scores, indicating that these genes are essential in many other cell lines as well. Considering both measures, PersonaDrive, DawnRank, sysSVM2, OncoImpact, and the PDPA consensus produced the best predictions. Again, all algorithms showed improved performance over randomDriver.

Drug sensitivity benchmarking of driver prioritisation algorithms

Using the DGIdb, TARGET-SL predicts drug sensitivity by identifying known inhibitors of predicted essential genes, prioritising drugs with the fewest known target genes to maximise specificity. In benchmarking mode, precision is calculated based on ground truth information from GDSC1/2 and PRISM. We also compared drug-sensitivity predictions with their quantitative $lnIC_{50}$ and UI_D from GDSC1/2. Finally, we introduced an additional control, randomDrug, which makes random drug-predictions for each cell line.

We compared our results to PanDrugs2 [39], a tool which uses a similar approach to TARGET-SL for drug prediction and can use the same input data from the CCLE cell lines. Using PanDrugs2 we made drug predictions for each cell line, filtered these to the same drug list available for TARGET-SL (inhibitory drugs only).

Overall, precision was worse for drug predictions than essential gene predictions, as was expected given drug-gene interactions are far less specific than CRISPR-Cas9 resulting in more off-target interactions. However, all PDPAs in combination with TARGET-SL showed improved performance over randomDrug (Fig. 5a and b). sysSVM2 had the best performance across all numbers of predictions, and together with PersonaDrive were the only PDPAs to show consistently better performance than randomDriver. This suggests that TARGET-SL successfully predicts drug-sensitivity, but most PDPAs generate only a marginal improvement. PanDrugs2 predictions were consistently lower and similar to randomDrug at higher numbers of predictions.

In predicting rare ground truth sensitive drugs (Fig. 5b), PanDrugs2 initially showed high precision for its top prediction, but then rapidly decreased in precision with increased predictions. No single PDPA performed best across all numbers of predictions in this comparison, however OncoImpact, SCS, and sysSVM2

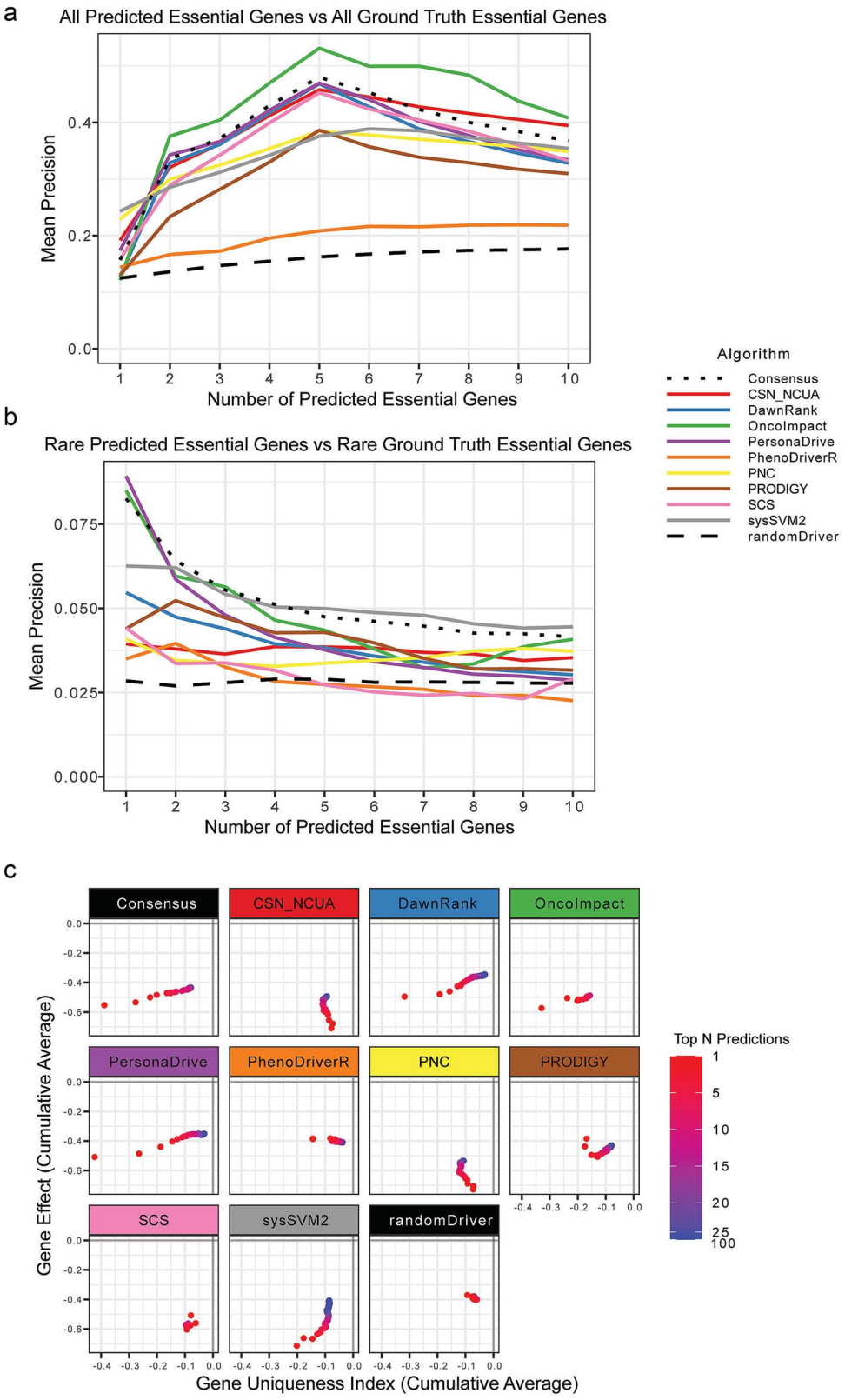


Figure 4. Gene-level comparison of driver prioritisation algorithms using TARGET-SL. (a) Essential genes were predicted for cell lines using each PDPA and compared against all ground truth essential genes. (b) Rare essential gene predictions were compared against the rare ground truth essential genes. (c) The cumulative average of the gene effect and the gene-uniqueness index (UI_G) were compared for the top 100 rare essential gene predictions.

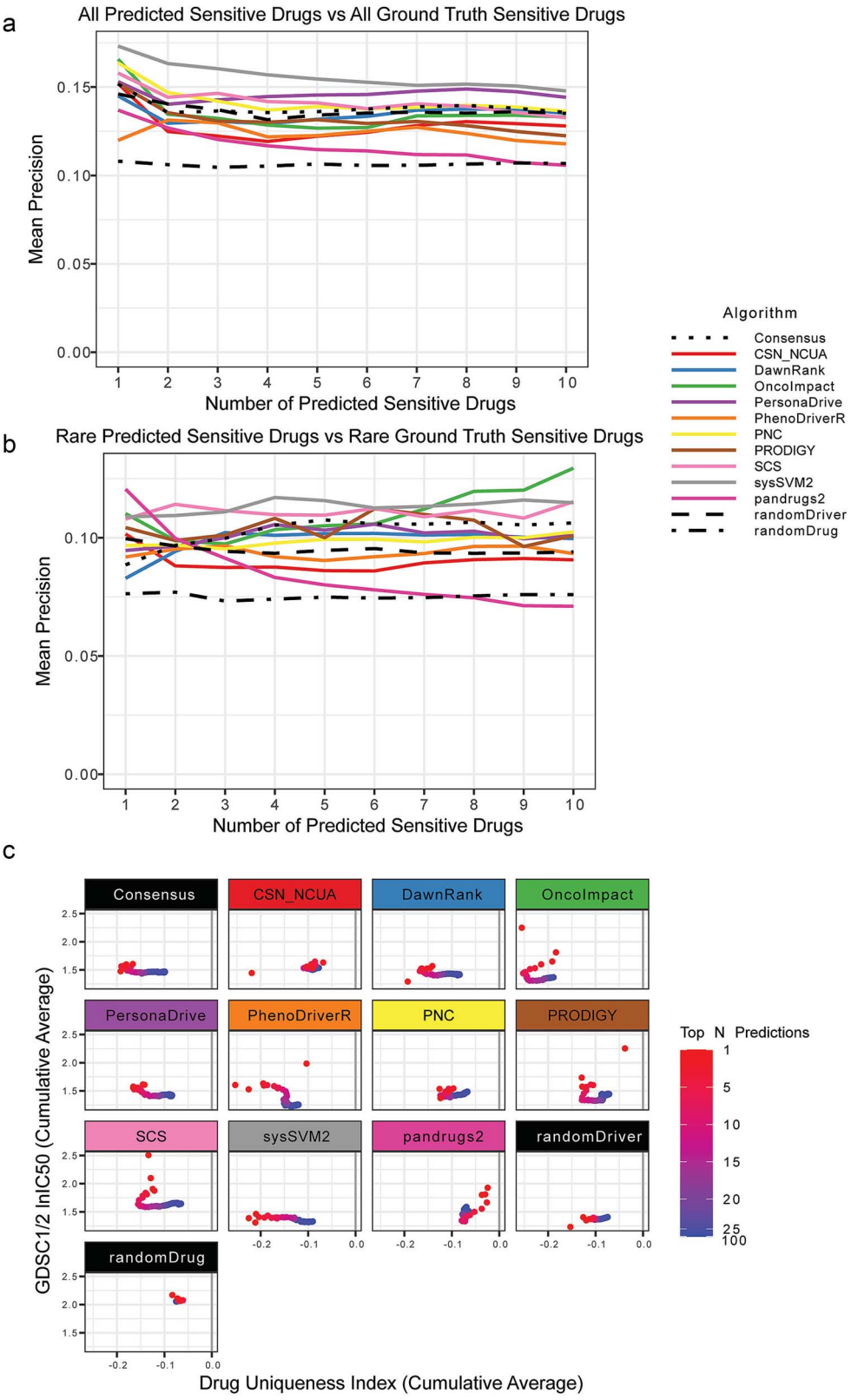


Figure 5. Drug-level comparison of driver prioritisation algorithms using TARGET-SL. (a) Drug sensitivity was predicted for cell lines using each PDPA and compared against all ground truth sensitive drugs. (b) Rare drug sensitivity predictions were compared against the rare ground truth sensitive drugs. (c) The cumulative average of the $lnIC_{50}$ and the drug uniqueness index (UI_G) were compared for the top 100 rare sensitivity predictions.

had the most improved performance over randomDriver. Overall, TARGET-SL-based predictions using PDPAs had higher precision than PanDrugs2, with PanDrugs2 failing to perform better than randomDrug after the top 7–8 predictions. Statistical comparisons for the top 10 predictions showed no significance compared with randomDrug in both precision measures (Supplementary Fig. S7).

For the quantitative UI_D analysis, compared with the gene effect-based analysis, we observed less cases of the desirable trend towards the bottom-left quadrant in these plots (Fig. 5c). However, for drug-sensitivity, the uniqueness of the effect (*x*-axis) is of greater importance than the IC_{50} of the drug, which varies for different compounds. Hence, focusing on the UI_D, OncoImpact, PhenoDriver, PersonaDrive, DawnRank, sysSVM2, and the PDPA consensus performed best in identifying targeted drug sensitivity. When comparing PDPAs and TARGET-SL with PanDrugs2, we observed that both approaches performed similar in terms of raw drug sensitivity prediction ($lnIC_{50}$), however, TARGET-SL's predictions were far more unique, with PanDrugs2 failing to outperform the randomised controls in this regard.

Top driver-essential gene pairs

We also examined frequent driver mutations associated with the highest sensitivity to TARGET-SL-predicted essential genes or drugs (Supplementary Fig. S8). Across all CCLE cells, HRAS and SIRT2 GOF mutations were most strongly associated with sensitivity to their genetic knockout, while LOF MAPK1 driver mutations incurred the strongest sensitivity to its predicted synthetic lethal partner, MTOR. Cell type-specific GOF mutations such as breast-KRAS, myeloid-BCR, biliary-BRAF, skin-PPM1D, myeloid-ABL1, skin-BIRC2, and oesophageal-ERBB2 mutations also conferred sensitivity to inhibition of their respective drivers within their corresponding cell types, while LOF SMARCA4 mutations in liver cells incurred sensitivity to SMARCA2 as a synthetic lethal partner. In most cases, strong driver-essential gene relationships did not necessarily correspond with strong sensitivity to known small-molecule inhibitors of these genes. The top cell type-specific predictions and cell type-agnostic predictions are available in Supplementary Tables S5 and S6, respectively.

Discussion

The major limitation of PDPAs is that their predictions are not readily actionable or able to be evaluated leading to their under-utilisation and underdevelopment. To address this, we developed TARGET-SL, combining data from several databases to predict essential genes and drug sensitivity based on PDPA predictions. This allows direct evaluation using LOF screening approaches like CRISPR-KO and inhibitory drug screens. This is a major advancement for the field, as TARGET-SL offers a novel method for gene-essentiality and drug-sensitivity prediction and serves as a benchmarking strategy for PDPAs.

TARGET-SL has several key innovations as a framework for PDPA benchmarking. Currently, the common practise is to compare predicted driver genes with canonical driver genes. This process misses novel drivers and ignores sample-specific information, making this a poor ground-truth for benchmarking. TARGET-SL solves this problem by utilising sample-specific ground-truth data. Additionally, by comparing algorithms with identical inputs, including reference GIN, TARGET-SL enables a fairer comparison. Additionally, we created a novel quantitative benchmark that eliminates arbitrary thresholds for true positives, offering a clearer view of the data.

To this end, we used TARGET-SL to compare nine PDPAs. Generally, all algorithms outperformed randomised drivers (randomDriver) that were subjected to the TARGET-SL pipeline. This demonstrates that driver prioritisation itself plays an important role in successfully identifying essential genes, separate from variant-effect and SL considerations. PersonaDrive, DawnRank, OncoImpact, and SysSVM2, in combination with TARGET-SL, were the most effective PDPAs, with precision dropping after the first 1 to 5 predictions. Interestingly, this list encompasses one of the most recent algorithms in the field (PersonaDrive [5]) as well as two of the oldest (DawnRank [32] and OncoImpact [33]).

We observed surprisingly little agreement between the various algorithms. The greatest similarity was between DawnRank and PersonaDrive. This was unexpected given that these two algorithms utilise very different approaches, namely different GIN types and prioritisation approaches. DawnRank only considers information from the single sample it is analysing, while PersonaDrive considers the connectedness between mutated and differentially expressed genes in other samples as well [5, 32]. On the other hand, PNC and CSN_NCUA showed expectedly high similarity given their similar approach to driver prioritisation [35].

Given the similar predictive performance of PDPAs despite their very different predictions, we used a modified Borda Count to generate a consensus driver list, in the expectation that this may yield improved results as previously achieved in similar scenarios [40]. The PDPA consensus did not drastically increase performance but provided consistent performance in all comparisons. The consensus also addresses the issue of individual PDPAs failing to provide any predictions for some samples.

TARGET-SL also stands alone as a gene-essentiality and drug sensitivity predictor, with its key advantage being its ability to produce highly sample-specific predictions. Specifically, we demonstrated that the PDPA step of TARGET-SL, when compared with canonical tier 1 CGC drivers, significantly improved the uniqueness of the predicted gene effect (Fig. 3). TARGET-SL also offered better drug-prediction performance than PanDrugs2 [39], performing with higher precision in rare drug sensitivity prediction with most PDPAs, particularly when using sysSVM2, SCS, and OncoImpact (Fig. 5b). Specifically, TARGET-SL predictions were similar to PanDrugs2 in terms of raw $lnIC_{50}$ scores, but the TARGET-SL predictions had greater uniqueness scores. We also found that our gene-essentiality predictions outperformed that of the BROAD-DREAM [41] challenge winner (Supplementary Fig. S9). However, we excluded this from our main analysis after observing a lack of correlation between CRISPR-Cas9 essentiality and the short hairpin RNA (shRNA) essentiality screens used in the BROAD-DREAM challenge (Supplementary Fig. S10), as noted by other authors [42]. Overall, we anticipate that the uniqueness of our predictions could translate to drug predictions with minimal off-target effects, a question that requires ongoing analyses.

While the primary focus of this manuscript was the broader performance of individual PDPAs and TARGET-SL, we also examined a selection of frequent driver mutations that incurred sensitivity to their target genes, as identified by TARGET-SL. GOF drivers in well-known cancer-related genes were predominant in this analysis. However, two interesting LOF drivers show frequently showed sensitivity with their predicted SL-partners, MAPK1-MTOR and a liver-specific SMARCA4-SMARCA2 relationship. Indeed, co-targeting of MTOR and MAPK1 has previously been reported as a potential therapeutic strategy for oral cancer [43]. Additionally, the SMARCA4-SMARCA2 SL relationship has been reported frequently in other cancer types [44–46], but has not been extensively studied in a liver-specific context.

Some limitations remain, such as TARGET-SL's reliance on cell line data for benchmarking algorithms designed for tissue data and the lack of healthy controls. This could be circumvented in future implementations by including organoid datasets that can contain both tumour and normal organoids and better recapitulate *in vivo* tissue heterogeneity [1]. Additionally, TARGET-SL relies on external databases that are incomplete. The LOF/-GOF annotation stage is of particular importance as it marks the first critical decision point between two opposing treatment approaches, and most variant-effect predictors are known to perform poorly for non-LOF mutations [47]. Finally, the gene inhibition from drug-gene interactions is less specific than the CRISPR-Cas9-gene interactions, and this is compounded by a lack of small-molecule inhibitors for many gene targets. Indeed, we found almost no correlation between CRISPR gene effect scores and the $\ln IC_{50}$ scores of the drugs predicted to inhibit the same genes (Supplementary Fig. S11). However, as many personalised-medicine approaches are targeting methods to screen thousands of drugs per patient, we believe TARGET-SL offers a valuable strategy to greatly reduce the number of drugs used in *ex vivo* screening.

In summary, TARGET-SL marks an important step on the path to personalised therapy options for cancer patients. TARGET-SL predictions are highly unique to individual samples, contributed to by both PDPAs and our variant effect and SL considerations. We expect this will translate into targeted treatment approaches with fewer off-target effects and support the development of PDPAs. In future, we plan to update TARGET-SL as its dependant databases improve and to use organoid datasets as they become available.

Key Points

- Algorithms designed to predict personalised cancer driver genes lack effective means of evaluation.
- TARGET-SL extrapolates these predictions into verifiable essential gene predictions.
- Combining driver prioritisation and synthetic lethality produces highly targeted predictions.

Author contributions

Rhys Gillman (Conceptualization Formal Analysis, Investigation, Methodology, Software, Visualization, Writing—original draft), Matt Field (Supervision, Writing—review & editing, Ulf Schmitz (Supervision, Resources, Writing—review & editing), Lionel Hebbard (Supervision, Resources, Writing—review & editing)

Supplementary data

Supplementary data are available at *Briefings in Bioinformatics* online.

Conflict of interest: None declared.

Funding

This work was supported by Tour De Cure (RSP-379-FY2023 to R.G.); James Cook University Postgraduate Research Scholarship (to R.G.); Tropical Australian Academic Health Centre Limited—Research Seed Grant (SF000121 to L.H.). National Health and Medical Research Council Investigator Grants (#5121190 to M.F.; #1196405 to U.S.); and Cancer Council NSW project grant (RG20–12 to U.S.).

Data availability

The data underlying this article are available in the Zenodo Repository at https://doi.org/10.5281/zenodo.14625394

References

1. Reiter JG, Baretti M, Gerold JM. *et al.* An analysis of genetic heterogeneity in untreated cancers. *Nat Rev Cancer* 2019;**19**:639–50. https://doi.org/10.1038/s41568-019-0185-x.
2. Gillman R, Field MA, Schmitz U. *et al.* Identifying cancer driver genes in individual tumours. *Comput Struct Biotechnol J* 2023;**21**:5028–38. https://doi.org/10.1016/j.csbj.2023.10.019.
3. Tokheim CJ, Papadopoulos N, Kinzler KW. *et al.* Evaluating the evaluation of cancer driver genes. *Proc Natl Acad Sci USA* 2016;**113**:14330–5. https://doi.org/10.1073/pnas.1616440113.
4. Sondka Z, Bamford S, Cole CG. *et al.* The COSMIC cancer gene census: describing genetic dysfunction across all human cancers. *Nat Rev Cancer* 2018;**18**:696–705. https://doi.org/10.1038/s41568-018-0060-1.
5. Erten C, Houdjedj A, Kazan H. *et al.* PersonaDrive: a method for the identification and prioritization of personalized cancer drivers. *Bioinformatics* 2022;**38**:3407–14. https://doi.org/10.1093/bioinformatics/btac329.
6. Barretina J, Caponigro G, Stransky N. *et al.* The cancer cell line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;**483**:603–7. https://doi.org/10.1038/nature11003.
7. DepMap B. DepMap 22Q4 public. *figshare* 2022. https://figshare.com/articles/dataset/DepMap_22Q4_Public/21637199.
8. Barrena N, Valcarcel LV, Olaverri-Mendizabal D. *et al.* Synthetic lethality in large-scale integrated metabolic and regulatory network models of human cells. *NPJ Syst Biol Appl* 2023;**9**:32. https://doi.org/10.1038/s41540-023-00296-3.
9. Oughtred R, Rust J, Chang C. *et al.* The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci* 2021;**30**:187–200. https://doi.org/10.1002/pro.3978.
10. Han Y, Wang C, Dong Q. *et al.* Genetic interaction-based biomarkers identification for drug resistance and sensitivity in cancer cells. *Mol Ther Nucleic Acids* 2019;**17**:688–700. https://doi.org/10.1016/j.omtn.2019.07.003.
11. Yang C, Guo Y, Qian R. *et al.* Mapping the landscape of synthetic lethal interactions in liver cancer. *Theranostics* 2021;**11**:9038–53. https://doi.org/10.7150/thno.63416.
12. Guo L, Dou Y, Xia D. *et al.* SLOAD: a comprehensive database of cancer-specific synthetic lethal interactions for precision cancer therapy via multi-omics analysis. *Database (Oxford)* 2022;1–8. https://doi.org/10.1093/database/baac075.
13. Benstead-Hume G, Chen X, Hopkins SR. *et al.* Predicting synthetic lethal interactions using conserved patterns in protein interaction networks. *PLoS Comput Biol* 2019;**15**:e1006888. https://doi.org/10.1371/journal.pcbi.1006888.
14. Guo J, Liu H, Zheng J. SynLethDB: synthetic lethality database toward discovery of selective and sensitive anticancer drug targets. *Nucleic Acids Res* 2016;**44**:D1011–7. https://doi.org/10.1093/nar/gkv1108.
15. Wang J, Wu M, Huang X. *et al.* SynLethDB 2.0: a web-based knowledge graph database on synthetic lethality for novel

anticancer drug discovery. *Database (Oxford)* 2022;1–10. https://doi.org/10.1093/database/baac030.
16. Wang J, Zhang Q, Han J. *et al.* Computational methods, databases and tools for synthetic lethality prediction. *Brief Bioinform* 2022;**23**:1–22. https://doi.org/10.1093/bib/bbac106.
17. Szklarczyk D, Gable AL, Lyon D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;**47**:D607–13. https://doi.org/10.1093/nar/gky1131.
18. Yang W, Soares J, Greninger P. *et al.* Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 2012;**41**:D955–61. https://doi.org/10.1093/nar/gks1111.
19. Freshour SL, Kiwala S, Cotto KC. *et al.* Integration of the drug–gene interaction database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res* 2020;**49**:D1144–51. https://doi.org/10.1093/nar/gkaa1084.
20. Novak LC, Chou J, Colic M. *et al.* PICKLES v3: the updated database of pooled in vitro CRISPR knockout library essentiality screens. *Nucleic Acids Res* 2023;**51**:D1117–21. https://doi.org/10.1093/nar/gkac982.
21. Meyers RM, Bryan JG, McFarland JM. *et al.* Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* 2017;**49**:1779–84. https://doi.org/10.1038/ng.3984.
22. Behan FM, Iorio F, Picco G. *et al.* Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* 2019;**568**:511–6. https://doi.org/10.1038/s41586-019-1103-9.
23. Dempster JM, Boyle I, Vazquez F. *et al.* Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biol* 2021;**22**:343. https://doi.org/10.1186/s13059-021-02540-7.
24. Kim E, Hart T. Improved analysis of CRISPR fitness screens and reduced off-target effects with the BAGEL2 gene essentiality classifier. *Genome Med* 2021;**13**:2. https://doi.org/10.1186/s13073-020-00809-3.
25. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* 2015;**31**:761–3. https://doi.org/10.1093/bioinformatics/btu703.
26. Sevim Bayrak C, Stein D, Jain A. *et al.* Identification of discriminative gene-level and protein-level features associated with pathogenic gain-of-function and loss-of-function variants. *Am J Hum Genet* 2021;**108**:2301–18. https://doi.org/10.1016/j.ajhg.2021.10.007.
27. Stein D, Kars ME, Wu Y. *et al.* Genome-wide prediction of pathogenic gain- and loss-of-function variants from ensemble learning of a diverse feature set. *Genome Med* 2023;**15**:103. https://doi.org/10.1186/s13073-023-01261-9.
28. Ng PK, Li J, Jeong KJ. *et al.* Systematic functional annotation of somatic mutations in cancer. *Cancer Cell* 2018;**33**:450–462 e410. https://doi.org/10.1016/j.ccell.2018.01.021.
29. Repana D, Nulsen J, Dressler L. *et al.* The network of cancer genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome Biol* 2019;**20**:1. https://doi.org/10.1186/s13059-018-1612-0.
30. Lever J, Zhao EY, Grewal J. *et al.* CancerMine: a literature-mined resource for drivers, oncogenes and tumor suppressors in cancer. *Nat Methods* 2019;**16**:505–7. https://doi.org/10.1038/s41592-019-0422-y.
31. Guo WF, Yu XT, Shi QQ. *et al.* Performance assessment of sample-specific network control methods for bulk and single-cell biological data analysis. *PLoS Comput Biol* 2021;**17**:e1008962. https://doi.org/10.1371/journal.pcbi.1008962.
32. Hou JP, Ma J. DawnRank: discovering personalized driver genes in cancer. *Genome Med* 2014;**6**:56. https://doi.org/10.1186/s13073-014-0056-8.
33. Bertrand D, Chng KR, Sherbaf FG. *et al.* Patient-specific driver gene prediction and risk assessment through integrated network analysis of cancer omics profiles. *Nucleic Acids Res* 2015;**43**:1–13. https://doi.org/10.1093/nar/gku1393.
34. Li Y, Zhang SW, Xie MY. *et al.* PhenoDriver: interpretable framework for studying personalized phenotype-associated driver genes in breast cancer. *Brief Bioinform* 2023;**24**:1–13. https://doi.org/10.1093/bib/bbad291.
35. Guo WF, Zhang SW, Zeng T. *et al.* A novel network control model for identifying personalized driver genes in cancer. *PLoS Comput Biol* 2019;**15**:e1007520. https://doi.org/10.1371/journal.pcbi.1007520.
36. Dinstag G, Shamir R. PRODIGY: personalized prioritization of driver genes. *Bioinformatics* 2020;**36**:1831–9. https://doi.org/10.1093/bioinformatics/btz815.
37. Guo WF, Zhang SW, Liu LL. *et al.* Discovering personalized driver mutation profiles of single samples in cancer by network control strategy. *Bioinformatics* 2018;**34**:1893–903. https://doi.org/10.1093/bioinformatics/bty006.
38. Nulsen J, Misetic H, Yau C. *et al.* Pan-cancer detection of driver genes at the single-patient resolution. *Genome Med* 2021;**13**:12. https://doi.org/10.1186/s13073-021-00830-0.
39. Jimenez-Santos MJ, Nogueira-Rodriguez A, Pineiro-Yanez E. *et al.* PanDrugs2: prioritizing cancer therapies using integrated individual multi-omics data. *Nucleic Acids Res* 2023;**51**:W411–8. https://doi.org/10.1093/nar/gkad412.
40. Waardenberg AJ, Field MA. consensusDE: an R package for assessing consensus of multiple RNA-seq algorithms with RUV correction. *PeerJ* 2019;**7**:e8206. https://doi.org/10.7717/peerj.8206.
41. Gonen M, Weir BA, Cowley GS. *et al.* A community challenge for inferring genetic predictors of gene essentialities through analysis of a functional screen of cancer cell lines. *Cell Syst* 2017;**5**:485–497.e3. https://doi.org/10.1016/j.cels.2017.09.004.
42. Wang W, Malyutina A, Pessia A. *et al.* Combined gene essentiality scoring improves the prediction of cancer dependency maps. *EBioMedicine* 2019;**50**:67–80. https://doi.org/10.1016/j.ebiom.2019.10.051.
43. Yamaguchi K, Iglesias-Bartolome R, Wang Z. *et al.* A synthetic-lethality RNAi screen reveals an ERK-mTOR co-targeting pro-apoptotic switch in PIK3CA+ oral cancers. *Oncotarget* 2016;**7**:10696–709. https://doi.org/10.18632/oncotarget.7372.
44. Cantley J, Ye X, Rousseau E. *et al.* Selective PROTAC-mediated degradation of SMARCA2 is efficacious in SMARCA4 mutant cancers. *Nat Commun* 2022;**13**:6814. https://doi.org/10.1038/s41467-022-34562-5.
45. Lee ECY, Reichl KD, Gopalsamy A. Synthetic lethality: targeting the SMARCA2 bromodomain for degradation in SMARCA4-deficient tumors - a review of patent literature from 2019-June 2023. *Expert Opin Ther Pat* 2024;**34**:211–29. https://doi.org/10.1080/13543776.2024.2355258.
46. Ehrenhofer-Wolfer K, Puchner T, Schwarz C. *et al.* SMARCA2-deficiency confers sensitivity to targeted inhibition of SMARCA4 in esophageal squamous cell carcinoma cell lines. *Sci Rep* 2019;**9**:11661. https://doi.org/10.1038/s41598-019-48152-x.
47. Gerasimavicius L, Livesey BJ, Marsh JA. Loss-of-function, gain-of-function and dominant-negative mutations have profoundly different effects on protein structure. *Nat Commun* 2022;**13**:3895. https://doi.org/10.1038/s41467-022-31686-6.

Briefings in Bioinformatics, 2025, **26(3)**, bbaf255
https://doi.org/10.1093/bib/bbaf255
Problem Solving Protocol