

Pathway2Targets: an open-source pathway-based approach to repurpose therapeutic drugs and prioritize human targets

Mauri Dobbs Spendlove¹, Trenton M. Gibson¹, Shaney McCain¹, Benjamin C. Stone¹, Tristan Gill² and Brett E. Pickett¹

¹ Microbiology and Molecular Biology, Brigham Young University, Provo, UT, United States of America

² Carlsbad, California, United States

ABSTRACT

Background: Recent efforts to repurpose existing drugs to different indications have been accompanied by a number of computational methods, which incorporate protein-protein interaction networks and signaling pathways, to aid with prioritizing existing targets and/or drugs. However, many of these existing methods are focused on integrating additional data that are only available for a small subset of diseases or conditions.

Methods: We have designed and implemented a new R-based open-source target prioritization and repurposing method that integrates both canonical intracellular signaling information from five public pathway databases and target information from public sources including [OpenTargets.org](#). The Pathway2Targets algorithm takes a list of significant pathways as input, then retrieves and integrates public data for all targets within those pathways for a given condition. It also incorporates a weighting scheme that is customizable by the user to support a variety of use cases including target prioritization, drug repurposing, and identifying novel targets that are biologically relevant for a different indication.

Results: As a proof of concept, we applied this algorithm to a public colorectal cancer RNA-sequencing dataset with 144 case and control samples. Our analysis identified 430 targets and ~700 unique drugs based on differential gene expression and signaling pathway enrichment. We found that our highest-ranked predicted targets were significantly enriched in targets with FDA-approved therapeutics for colorectal cancer (p -value < 0.025) that included EGFR, VEGFA, and PTGS2. Interestingly, there was no statistically significant enrichment of targets for other cancers in this same list suggesting high specificity of the results. We also adjusted the weighting scheme to prioritize more novel targets for CRC. This second analysis revealed epidermal growth factor receptor (EGFR), phosphoinositide-3-kinase (PI3K), and two mitogen-activated protein kinases (MAPK14 and MAPK3). These observations suggest that our open-source method with a customizable weighting scheme can accurately prioritize targets that are specific and relevant to the disease or condition of interest, as well as targets that are at earlier stages of development. We anticipate that this method will complement other approaches to repurpose drugs for a variety

Submitted 28 March 2023
Accepted 22 August 2023
Published 29 September 2023

Corresponding author

Brett E. Pickett,
brett_pickett@byu.edu

Academic editor
Fares Ali

Additional Information and
Declarations can be found on
page 15

DOI [10.7717/peerj.16088](https://doi.org/10.7717/peerj.16088)

© Copyright
2023 Dobbs Spendlove et al.
Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

of indications, which can contribute to the improvement of the quality of life and overall health of such patients.

Subjects Bioinformatics, Computational Biology, Molecular Biology, Clinical Trials, Data Science

Keywords Drug repurposing, Drug targets, Target prioritization, Bioinformatics, Colorectal cancer, Target, Pathways, Prediction

INTRODUCTION

Substantial effort and resources have been devoted to identifying therapeutic treatments for many human diseases and conditions. The maladies could be caused by autoimmunity, uncontrolled cell growth, genetics, infection, and other acute or chronic ailments. Since moving a candidate treatment through the process of approval by the US Food and Drug Administration (FDA) is risky ([Zhong et al., 2018](#)), often taking many years, and requiring a substantial financial investment; researchers have expanded their development efforts to drug repurposing ([Hernandez et al., 2017](#); [Parvathaneni et al., 2019](#)). Traditional methods of drug discovery have involved using low- or high-throughput screens to identify inhibitors or activators of a given target ([Thakur et al., 2021](#); [Olsen, 2019](#); [Glanz et al., 2020](#)). Hits that are identified in these screens are generally optimized prior to subsequent testing in cell culture, animal models, and clinical trials ([Thakur et al., 2021](#)). Alternatively, using a pathway-based approach to drug discovery involves performing experiments to better understand the underlying mechanism(s) of a given condition, and to identify relevant targets ([Deng et al., 2020](#); [Wang et al., 2020b](#); [Damale et al., 2020](#); [Chatterjee et al., 2022](#); [Liu et al., 2021](#); [Ren et al., 2010](#)). Past studies have shown that incorporating a signaling pathway approach can successfully identify proteins that can be targeted with therapeutics having sufficient efficacy and safety to warrant approval by regulators ([Ding et al., 2020](#); [Khojasteh Poor et al., 2021](#); [Choi et al., 2020](#); [Proctor, Thompson & O'Bryant, 2014](#)).

Drug repurposing is the process of getting regulatory approval for applying an existing therapeutic to a separate disease or condition (*i.e.*, indication) ([Ding et al., 2020](#); [Zali et al., 2019](#); [Harb, Lin & Hao, 2019](#)). The benefits of this approach include a potentially shorter time to approval since the therapeutic has already been deemed as “safe” by government regulatory agencies. Early repurposing efforts were focused on identifying symptom similarities or using known side-effects from patients with other conditions to treat a separate condition ([Kingsmore, Grammer & Lipsky, 2020](#); [Ballard et al., 2020](#)). Subsequent advances in understanding intracellular signaling mechanisms enabled a transition to more complex analyses that identify a candidate therapeutic for repurposing, and to develop novel therapeutics towards known targets ([Schein, 2020](#)). This is evidenced by the wide variety of drug and target discovery tools that have already been reported ([Paananen & Fortino, 2020](#); [Seleno & Emili, 2008](#); [Huang et al., 2020](#)). The majority of these modern methods take advantage of protein-protein interaction networks ([Ma et al., 2019](#); [Ozdemir et al., 2019](#); [Cheng et al., 2019](#)), gene sets ([Masoudi-Sobhanzadeh et al., 2020](#); [Tanoli, Vähä-Koskela & Aittokallio, 2021](#)), and/or signaling pathways ([Jain et al., 2021](#)) for

repurposing-based drug- and target prioritization efforts. Some methods combine one or more of these methods with artificial intelligence to further improve the pace of drug discovery ([Tanoli, Vähä-Koskela & Aittokallio, 2021](#); [Anderson et al., 2020](#); [Paul et al., 2021](#); [Gupta et al., 2021](#)).

Even with such recent advances, many prioritization algorithms rely on public or proprietary protein network data ([Emig et al., 2013](#); [Huang et al., 2014](#); [Louhimo et al., 2016](#); [Li & Lu, 2013](#); [Isik et al., 2015](#); [Carrella et al., 2014](#); [Setoain et al., 2015](#); [Duan et al., 2016](#); [Barrio-Hernandez et al., 2023](#); [Fang et al., 2019](#); [Lee et al., 2011](#); [Greene et al., 2015](#); [Huang et al., 2018](#)), with some algorithms focusing on a particular set of diseases or conditions ([Crowther et al., 2010](#); [Xu, Kong & Hu, 2021](#); [Dezső & Ceccarelli, 2020](#); [Fiscon et al., 2021](#); [Chen & Xu, 2016](#); [Regan-Fendt et al., 2020](#)). Drug repurposing and target prioritization algorithms generally apply a consistent set of parameters, which are often specific to a given indication. Such specialization makes it difficult to effectively and adequately support the efforts of researchers working in other disease areas ([Sharma et al., 2021](#); [Begley et al., 2021](#)).

Given the specialization that is prevalent among many repurposing tools, the aim of the current study was to incorporate a novel, flexible, and customizable open-source target prioritization method into the Pathway2Targets algorithm, which would increase the number of supported use cases. This updated algorithm retrieves additional target information, clinical trial data, automatically fetches the reactome pathway diagrams for the signaling pathways with the highest number of targets, and accepts reactome pathway enrichments generated by the enrichr algorithm ([Xie et al., 2021a](#)). This additional data and prioritization method are used by the updated algorithm to generate ranked lists of targets and therapeutics that can be applicable to multiple use cases ([Scott, Jensen & Pickett, 2021](#); [Gray et al., 2022](#); [Moreno et al., 2022](#); [Rapier-Sharman, Clancy & Pickett, 2022](#)). The entities in these lists can then be evaluated as candidates for condition-specific repurposing efforts based solely on the unique signaling pathway “profile” for the disease/condition of interest.

METHOD

GEO query for transcriptomics data

The Gene Expression Omnibus database was queried for a well-controlled bulk transcriptomics human colorectal cancer dataset with a sufficiently high number of samples ([GSE156451](#)) to enable confident downstream repurposing analysis ([Sayers et al., 2021](#)). The paired-end fastq files for this publicly available study were then downloaded from the Sequence Read Archive (SRA), a database within the National Center for Biotechnology Information ([Sayers et al., 2021](#)). This study consisted of 144 samples, with 72 from tumors in patients with colorectal cancer (CRC) and the other 72 from native human tissue ([Li et al., 2021](#)).

Transcriptomic preprocessing and analysis

The 144 public colorectal cancer transcriptomics samples were preprocessed using the Automated Reproducible MOdular Workflow for Preprocessing and Differential Analysis

of RNA-seq Data (ARMOR) software ([Orjuela et al., 2019](#)). Briefly, this open-source Snakemake-based workflow was used to perform read trimming on the RNA-sequencing fastq files with TrimGalore ([Köster & Rahmann, 2018](#); https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), determine quality control metrics with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/), map and quantify reads to the human GRCh38 transcriptome with Salmon ([Patro et al., 2017](#)), and calculate differential gene expression with edgeR ([Robinson, McCarthy & Smyth, 2010](#)) by comparing the CRC samples (case) to the native tissue samples (control). The Ensembl Gene IDs that were generated by the edgeR algorithm were converted to Entrez Gene IDs using a R-based application programming interface (API) to the BiomaRt database prior to pathway analysis ([Kasprzyk, 2011](#)). Similarly, the enrichr pathway enrichment software only required a gene symbol, \log_2 fold-change values, and FDR-adjusted p -values as input from the DEG list. The statistically significant differentially expressed genes (DEGs; FDR-corrected p -value < 0.05) were then subjected to signaling pathway analysis using the Signal Pathway Impact Analysis (SPIA) algorithm with 3,000 bootstrap replicates to generate a null distribution for each of over 2,000 public signaling pathways ([Tarca et al., 2009](#)), as reported previously ([Scott, Jensen & Pickett, 2021](#); [Gray et al., 2022](#); [Moreno et al., 2022](#); [Rapier-Sharman, Clancy & Pickett, 2022](#); [Ferrarini et al., 2021](#); [Scott et al., 2022](#); [Gifford & Pickett, 2022](#)). The lists of pathways were derived from publicly available versions of KEGG ([Aoki-Kinoshita & Kanehisa, 2007](#)), Reactome ([Jassal et al., 2020](#)), Pathway Interaction Database ([Schaefer et al., 2009](#)), BioCarta, and Panther ([Mi et al., 2017](#)).

Target data acquisition and integration

The only input file required for the Pathway2Targets software was the tabular output file containing the significant signaling pathways (Bonferroni-corrected p -value < 0.05) generated by SPIA, although an output file from the enrichr algorithm would have also been compatible ([Xie et al., 2021a](#)). The Pathway2Targets software then programmatically retrieved the gene products that were members of each significant pathway from the five pathway databases mentioned previously, and then obtained the UniProt protein identifiers for each Ensembl ID using the BiomaRt API ([Scott, Jensen & Pickett, 2021](#); [UniProt Consortium, 2019](#)). A GraphQL query was automatically generated and submitted through the Open Targets Platform API to access the relevant drug and target information for each of the UniProt protein identifiers in each pathway ([Ochoa et al., 2021](#)). These additional data for each target included the number of associated diseases, tractability, subcellular location, safety, number of unique drugs, number of signaling pathways, number of FDA-approved therapeutics, number of therapeutics in phase-three clinical trials, number of therapeutics in phase-two clinical trials, number of therapeutics in phase-one clinical trials, and number of therapeutics in phase-four clinical trials. This information was then automatically integrated with the significant results from the signaling pathway enrichment analysis described above in a single table to facilitate downstream target scoring and prioritization.

Table 1 Attributes and weighting values used to prioritize targets for repurposing (default) and for early/novel target prediction.

Target attribute	Default weighted score	Weighted score for novel targets
Number targets in pathways	1	0.5
Tractability	1	0.5
Number approved drugs	1	0.5
Number safety liabilities	-2	-2
Number unique drugs	1	0.5
Number associated diseases	1	0.01
Phase 1 drug	0.5	10
Phase 2 drug	1	4
Phase 3 drug	1.5	0.5
Phase 4 drug	2	0.01

Target weighting factors

A logical and customizable weighting scheme was constructed in the algorithm that would compile and analyze the data for all existing therapeutics for each pathway member to facilitate target prioritization (Table 1). The default weights for each target attribute were specifically chosen in a way that would prioritize targets present in multiple pathways, a high number of associated disease, and a higher number of therapeutics further along in clinical trials. However, these default weighted values could also be easily adjusted to customize the output based on individual prioritization preferences or desired outcome. As a proof of concept for adjusting the default weighted values, these values were adjusted to facilitate discovery of novel and/or early-stage targets (Table 1). A table of prioritized targets as well as a separate table of prioritized therapeutics, which uses the same target prioritization weighting scheme was then generated as output from the Pathway2Targets software. An option to automatically download the graphical representations for the most common Reactome pathways, *via* the Reactome API, was also provided (Jassal *et al.*, 2020).

RESULTS

Original software implementation

Our original implementation of the Pathway2Targets algorithm, implemented in R, required only a single input file containing the output from a SPIA-based signaling pathway enrichment analysis. This original version of the Pathway2Targets algorithm then iterated over each significant signaling pathway name (p -value < 0.05), together with the associated source databases, and retrieved all gene products that are members of each pathway. The diverse public identifiers used by each pathway database were then automatically converted and submitted to the Open Targets database to retrieve a smaller subset of the available target information. The data for ~10 fields associated with each target were then programmatically retrieved and stored for processing. These fields included target name, target symbol, target Ensembl identifier, therapeutic name, therapeutic type, modulation of the target, total tractability, maximum clinical phase, and

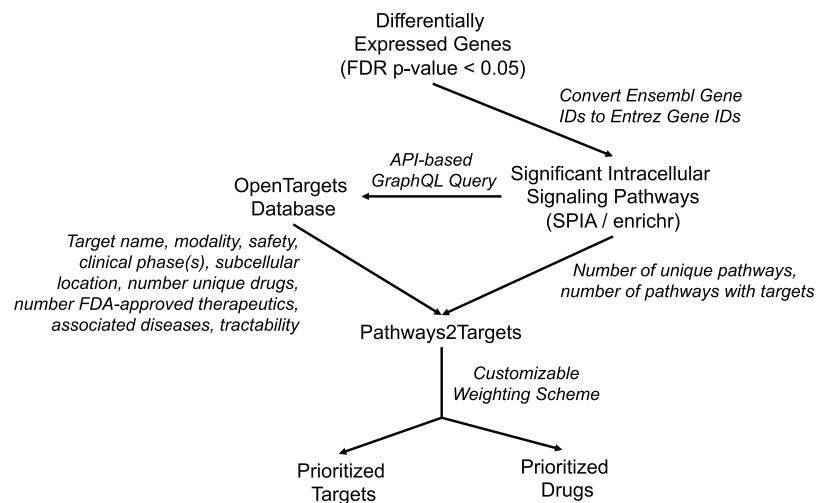


Figure 1 Diagram representing Pathway2Targets workflow, various data sources, and outputs. Upstream processing steps include differential gene expression and signaling pathway enrichment analysis. The algorithm uses significant pathways and target data from the open targets platform to perform drug- and target prioritization.

[Full-size](#) DOI: 10.7717/peerj.16088/fig-1

approval status. The single output file combined these fields with the relevant pathway information for each target (Table S1). Unfortunately, we found that the file generated by the original software required substantial downstream manual review and analysis to facilitate biologically relevant interpretation. Importantly, while the earlier version of the software did identify relevant targets from the significant signaling pathways, it did not collect sufficient data from the Open Targets database to enable the prioritization of targets directly from the output.

Novel software enhancements

In order to address the observed inefficiencies in the original algorithm and to facilitate improved target prioritization, we made substantial enhancements to the original Pathway2Targets algorithm. The updated software still only requires a tabular input file containing the significant signaling pathways. However, the updated version now retrieves data for ~20 additional fields from the Open Targets Platform, calculates additional metrics, and incorporates a customizable prioritization weighting scheme (Fig. 1). Specifically, we updated the software to retrieve additional Open Targets data including tractability for each therapeutic modality, subcellular location, number of unique drugs, number of signaling pathways, number of therapeutics in phase-three clinical trials, number of therapeutics in phase-two clinical trials, number of therapeutics in phase-one clinical trials, and number of therapeutics in phase-four clinical trials. We also updated the software to calculate the number of targets that are present in each pathway.

These enhancements to Pathway2Targets enabled us to incorporate a set of logical, flexible, and customizable weighting factors that directly support target prioritization efforts. To do so, we assigned customizable weights to each of 10 target attributes that are commonly used by researchers to rank/prioritize pathway-based targets (Table 1). We also implemented new functionality to automatically retrieve pathway diagrams from

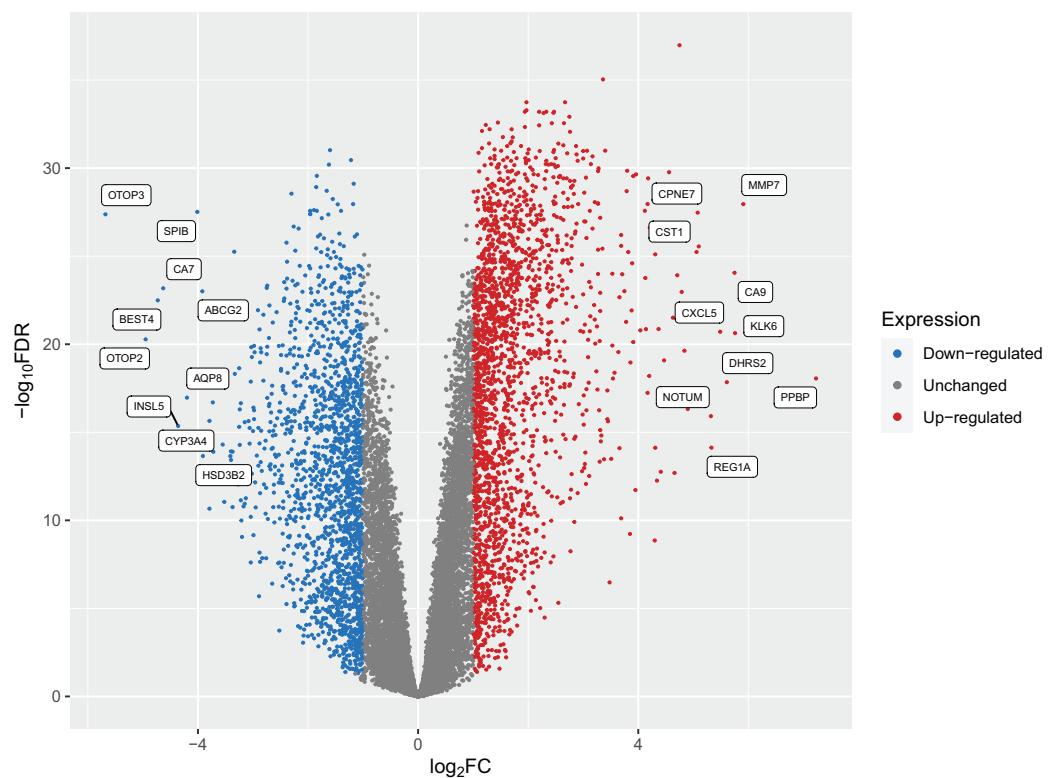


Figure 2 A volcano plot of the differentially expressed genes (DEGs) calculated from tumors from patients with colorectal cancers compared to native patient tissue. Dots represent individual genes and colors indicate down-regulation, up-regulation, or expression that did not surpass the threshold of $|\log_2 \text{fold-change}| < 1$ (blue, red, and gray respectively). [Full-size](#) DOI: [10.7717/peerj.16088/fig-2](https://doi.org/10.7717/peerj.16088/fig-2)

Reactome and to support target prioritization efforts by using the enrichr algorithm on the Reactome database.

Application to a colorectal cancer use case: gene expression & pathways

After implementing these enhancements into Pathway2Targets, we next decided to validate the biological relevance of the results from this algorithm. To do so, we tested it on the significantly enriched signaling pathways from the public RNA-sequencing dataset in colorectal cancer. To perform this analysis, we downloaded and preprocessed the fastq files for this dataset using the ARMOR automated workflow that performed quality control, trimmed and mapped reads, and calculated 12,159 differentially expressed genes (FDR-adjusted p -value < 0.05) (Fig. 2; Table S2). Examples of the most significant DEGs included KRT80, AJUBA, TRIP13, PAICS, and RNASEH2A with FDR-adjusted p -values ranging from 1×10^{-37} to 1×10^{-34} . Significant DEGs that were up-regulated include MMP7, PPBP, and CXCL5; while a subset of down-regulated DEGs included OTOP3, OTOP2, SPIB, and INSL5 (Table 2; Table S3).

To identify significant signaling pathways in this public colorectal cancer dataset we then applied the existing signaling pathway impact analysis (SPIA) algorithm to the list of

Table 2 The top 20 differentially expressed genes, ranked in descending order by the absolute value of the \log_2 fold-change values.

Gene symbol	Gene description	Ensembl gene ID	Entrez gene ID	\log_2 fold-change	FDR p-value
PPBP	Pro-platelet basic protein	ENSG00000163736	5473	7.23	8.75E-19
MMP7	Matrix metallopeptidase 7	ENSG00000137673	4316	5.91	1.11E-28
KLK6	Kallikrein related peptidase 6	ENSG00000167755	5653	5.76	2.35E-21
CA9	Carbonic anhydrase 9	ENSG00000107159	768	5.75	8.76E-25
OTOP3	Otopetrin 3	ENSG00000182938	347741	-5.68	4.20E-28
DHRS2	Dehydrogenase/reductase 2	ENSG00000100867	10202	5.61	1.43E-18
CXCL5	C-X-C motif chemokine ligand 5	ENSG00000163735	6374	5.49	1.95E-21
REG1A	Regenerating family member 1 alpha	ENSG00000115386	5967	5.33	7.47E-15
NOTUM	Notum, palmitoleoyl-protein carboxylesterase	ENSG00000185269	147111	5.32	1.21E-16
CST1	Cystatin SN	ENSG00000170373	1469	5.1	2.80E-26
CPNE7	Copine 7	ENSG00000178773	27132	5.08	3.38E-28
SLC35D3	Solute carrier family 35 member D3	ENSG00000182747	340146	5.06	5.77E-26
OTOP2	Otopetrin 2	ENSG00000183034	92736	-4.95	5.27E-21
KRT6B	Keratin 6B	ENSG00000185479	3854	4.9	4.76E-17
CLDN2	Claudin 2	ENSG00000165376	9075	4.84	2.34E-20
COL11A1	Collagen type XI alpha 1 chain	ENSG00000060718	1301	4.79	1.07E-23
KRT80	Keratin 80	ENSG00000167767	144501	4.75	1.05E-37
BEST4	Bestrophin 4	ENSG00000142959	266675	-4.73	3.22E-23
FOXQ1	Forkhead box Q1	ENSG00000164379	94234	4.71	1.21E-24
TACSTD2	Tumor associated calcium signal transducer 2	ENSG00000184292	4070	4.7	4.96E-27

DEGs in CRC. The SPIA algorithm uses permutation and bootstrapping to generate a null distribution for each pathway, and then applies a Bonferroni p -value correction to reduce the number of false-positive results. This pathway enrichment analysis identified 63 statistically significant (Bonferroni-adjusted p -value < 0.05) and biologically relevant intracellular signaling pathways including the inhibition of HIF-1 (Bonferroni p -value 0.0017), activation of cell cycle (Bonferroni p -value 0.003), activation of DNA replication (Bonferroni p -value 0.0008), activation of PLK1 (Bonferroni p -value 0.000089), and inhibition of T-cell receptor signaling pathways (Bonferroni p -value 0.001) (Table 3).

Application to a colorectal cancer use case: late-stage target prioritization from signaling pathways

After identifying significantly affected signaling pathways, we then wanted to determine whether our updated Pathway2Targets software could be used for two purposes: (1) to predict existing late-stage therapeutics from the enriched signaling pathway data, as well as those that could potentially be repurposed for treating colorectal cancer and (2) to generate a ranked list of prioritized targets that were specific to colorectal cancer. To evaluate these outcomes, we used the new Pathway2Targets software to analyze our significant pathway results from the CRC dataset.

Table 3 Top 10 most significant intracellular signaling pathways that are predicted to be affected in colorectal cancer.

Name	# Proteins in pathway	# DEGs in pathway	Bonferroni-adjusted p-value	Pathway status	Source database
Major pathway of rRNA processing in the nucleolus and cytosol	169	162	7.19E-13	Activated	Reactome
rRNA processing in the nucleus and cytosol	178	169	3.55E-12	Activated	Reactome
rRNA processing	183	173	6.45E-12	Activated	Reactome
Gene expression	1,598	1,307	3.45E-08	Activated	Reactome
Eukaryotic translation elongation	90	86	7.11E-08	Activated	Reactome
Peptide chain elongation	86	82	1.64E-07	Activated	Reactome
Processing of capped intron-containing pre-mRNA	228	202	2.61E-07	Activated	Reactome
Translation	152	137	1.01E-06	Activated	Reactome
Non-coding RNA metabolism	49	48	1.86E-06	Activated	Reactome
snRNP assembly	49	48	1.86E-06	Activated	Reactome

Our Pathway2Targets analysis generated a list of 430 targets (Table S4) and approximately 700 unique drugs (Table S5). To determine which signaling pathways were enriched in the list of targets, we ran an enrichment in Reactome and visualized the result for the “Oxygen-dependent proline hydroxylation of Hypoxia-inducible Factor Alpha” Reactome pathway (Fig. 3).

Application to a colorectal cancer use case: significant enrichment of targets for indication

We next wanted to quantify how many of the predicted targets were approved for either colorectal cancer and/or any cancer. To do so, we manually reviewed the top 50 targets, together with the attributes and default weighted scores. This analysis showed that the top 50 prioritized targets, using the default weighting parameters, included three approved CRC targets and 11 targets for other cancers. Given that many labs prefer to validate a much smaller number of results in the wet-lab, we reduced the top 50 to a list of 15 potential candidates for drug repurposing. After applying this filtering process, we observed three CRC targets (VEGFA, EGFR, and PTGS2) and five targets that had undergone one or more phase IV trials in any indication (TP53, VEGFA, EGFR, ESR1, and PTGS2) were all found in the top 15 results from this list (Table 4).

To determine whether our selected parameters were able to accurately predict therapeutic candidates that would be specific for colorectal cancer or all cancers, we performed a hypergeometric statistical analysis of the top 15 targets. This analysis showed a significant enrichment of CRC targets in the top 15 predicted results (p -value < 0.025), while there was no significant enrichment for all cancers (p -value = 0.13). This result suggests that our approach is capable of predicting biologically relevant targets based on the specific signaling pathway profile of the disease/system being evaluated.

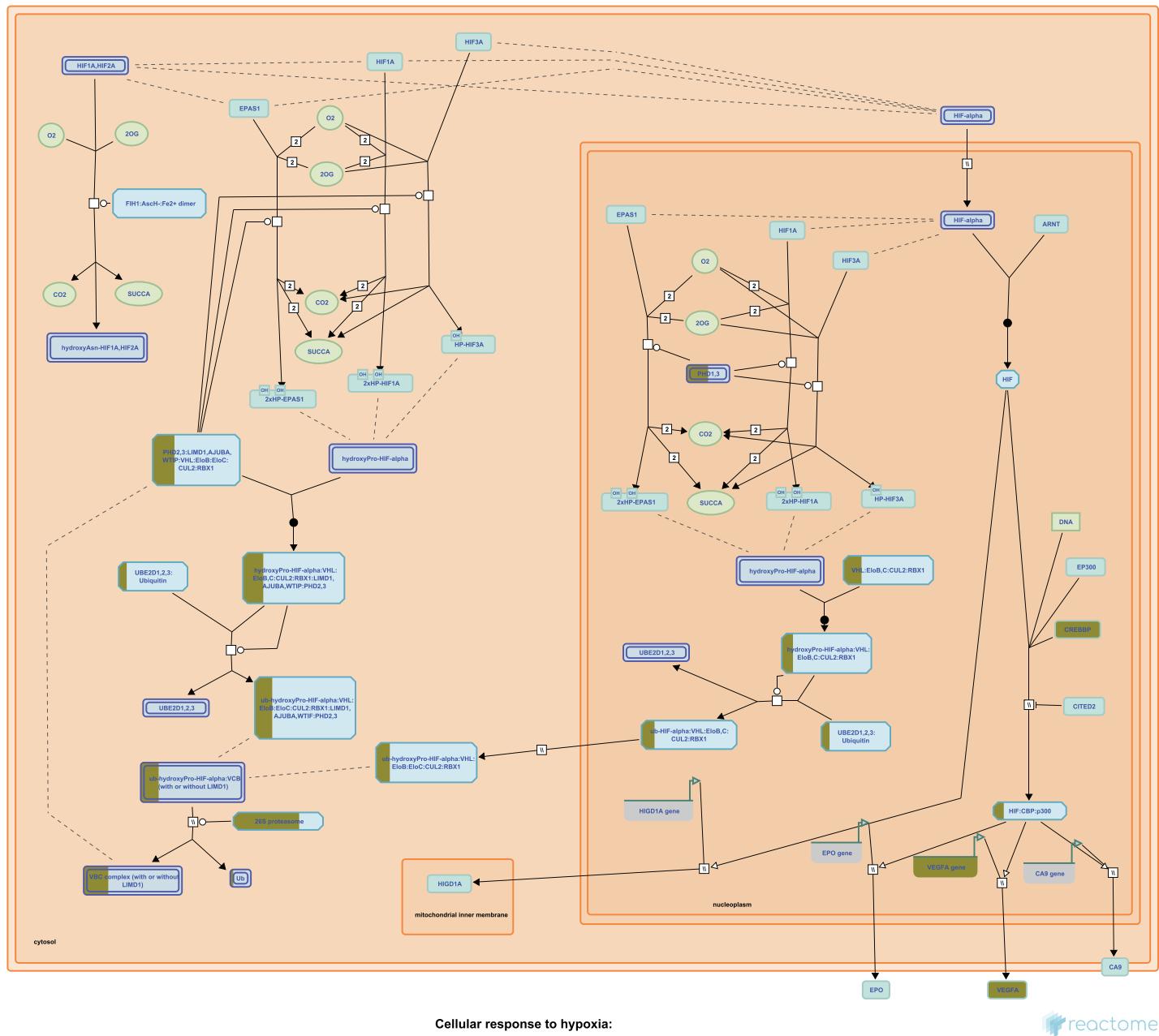


Figure 3 The “Oxygen-dependent proline hydroxylation of Hypoxia-inducible Factor Alpha” signaling pathway from Reactome. Portions of rectangular nodes are shaded yellow to represent the fraction of components within each node that were included identified as targets in our analysis.

Full-size DOI: 10.7717/peerj.16088/fig-3

Application to a colorectal cancer use case: early-stage target prioritization from signaling pathways

Since the flexible and customizable design of our updated software is capable of supporting multiple use cases, we next adjusted the weighting scheme to mimic a scenario where the default weights would be adjusted to enable the prediction of early-stage or novel targets. To achieve such a scenario, we decreased the weight for targets in pathways, tractability

Table 4 Ranked list of attributes for three approved colorectal cancer drug targets, predicted using our weighted factors.

Target symbol	VEGFA	EGFR	PTGS2
Target ID	ENSG00000112715	ENSG00000146648	ENSG00000073756
Target name	Vascular endothelial growth factor A	Epidermal growth factor receptor	Prostaglandin-endoperoxide synthase 2
Associated disease count	2,188	1,839	1,585
Tractability count	5	7	4
sm (is approved)	FALSE	TRUE	TRUE
sm (is in advanced trial)	FALSE	FALSE	FALSE
sm (is in phase 1)	FALSE	FALSE	FALSE
ab (is approved)	FALSE	TRUE	TRUE
ab (is in advanced trial)	TRUE	TRUE	FALSE
ab (is in phase 1)	FALSE	FALSE	FALSE
pr (is approved)	FALSE	FALSE	FALSE
pr (is in advanced trial)	FALSE	FALSE	FALSE
pr (is in phase 1)	FALSE	FALSE	FALSE
oc (is approved)	TRUE	FALSE	TRUE
oc (is in advanced trial)	FALSE	TRUE	FALSE
oc (is in phase 1)	FALSE	FALSE	FALSE
Subcellular location	Secreted	Cell membrane	Microsome membrane
Safety liabilities	0	2	25
Number unique drugs	11	72	75
# Pathways with target	2	1	1
# Approved therapeutics	4	14	12
# Therapeutics in phase 3	1	2	0
# Therapeutics in phase 2	0	0	0
# Therapeutics in phase 1	0	0	0
# Therapeutics in phase 4	4	14	12
Weighted score	2,219.5	1,960	1,651
Prioritized rank	3	6	12

Note:

sm, small molecule; ab, antibody; pr, proteolysis targeting chimeras; oc, other clinical modalities.

count, number of approved drugs, number of unique drugs, and phase 3 from the default value of 1 to a new value of 0.5. We also reduced the weight for both associated diseases and phase 4 from the default of 1 to 0.01, increased phase 2 from the default of 1 to 4, and increased phase 1 from the default of 0.5 to 10. The results of this analysis showed that EGFR was the only CRC-approved target that remained in the top-10 ([Table 5](#)), while VEGFA and PTGS2 decreased in rank (to 41 and 425 respectively) ([Tables S6, S7](#)). Interestingly, this modified weighting scheme resulted in MTOR, MAPK14, and TP53 being present in the top 10 results. This observation suggests that our customizable weighting approach could be capable of predicting relevant potential drug repurposing candidates even when minimal approved targets are known.

Table 5 The top 10 early/novel targets in colorectal cancer, based on the customized weights defined in **Table 1**.

Target symbol	Target name	Subcellular location	Weighted score
EGFR	Epidermal growth factor receptor	Cell membrane	61.43
MAPK14	Mitogen-activated protein kinase 14	Cytoplasm	60.12
RPS6KB1	Ribosomal protein S6 kinase B1	Synapse	59.89
MTOR	Mechanistic target of rapamycin kinase	Endoplasmic reticulum membrane	54.94
PIK3R1	Phosphoinositide-3-kinase regulatory subunit 1	Cytosol	54.53
TP53	Tumor protein p53	Cytoplasm	52.83
AURKB	Aurora kinase B	Nucleus	50.9
PIK3R2	Phosphoinositide-3-kinase regulatory subunit 2	No data	49.29
CD40	CD40 molecule	Cell membrane	48.63
MAPK3	Mitogen-activated protein kinase 3	Cytoplasm	48.6

DISCUSSION

This study reports the capabilities of the updated Pathway2Targets algorithm to predict, prioritize, and/or repurpose targets for a particular disease or condition by combining pathway information with additional public data and a weighting scheme in a flexible and customizable way. This approach improves on our prior work to incorporate the significant intracellular signaling pathways that best represent the underlying transcriptomic differential expression data into the target prediction and prioritization ([Scott, Jensen & Pickett, 2021](#); [Gray et al., 2022](#); [Moreno et al., 2022](#); [Rapier-Sharman, Clancy & Pickett, 2022](#); [Ferrarini et al., 2021](#); [Scott et al., 2022](#); [Gifford & Pickett, 2022](#)).

Our target prediction and prioritization approach relies on the underlying significant mechanistic pathways, based on large amounts of -omics data. This design makes our software disease-agnostic, and capable of prioritizing targets across a wide range of diseases based solely on the signaling pathway profile. In brief, our algorithm accurately predicted three approved targets with high specificity for CRC, using only signaling pathway results, relevant target information, and the default weighting scheme. It was also able to prioritize relatively early-stage and novel targets for CRC by adjusting the weight parameters. These findings suggest that our pathway-based repurposing approach could be relevant for other oncological, rare disease, or other indications. Such repurposing could potentially predict therapeutics that could still work for a given indication but that may be more economical—thereby reducing the cost of treatment in developing regions of the world.

An extensive array of drug- and target prioritization approaches have been developed and reported previously. In particular, some algorithms specifically focus on performing these tasks for particular diseases or conditions ([Di et al., 2019](#); [Dwane et al., 2021](#); [Yang et al., 2021b](#); [Behan et al., 2019](#); [Mejía-Pedroza, Espinal-Enríquez & Hernández-Lemus, 2018](#); [Urán Landaburu et al., 2020](#); [Tsuiji et al., 2021](#)). Many other implemented methods are informed by protein-protein interaction networks and/or pathway information for either target prioritization or drug repurposing ([Ma et al., 2019](#); [Isik et al., 2015](#); [Fiscon & Paci, 2021](#); [Aguirre-Plans et al., 2019](#); [Napolitano et al., 2018](#); [Wang et al., 2019a](#); [Malas](#)

et al., 2019), but are not designed to perform both tasks simultaneously as ours does. A subset of these algorithms even incorporate machine learning to improve their results (*Tsuji et al., 2021; Malas et al., 2019*). Our approach differs from these prior tools by only requiring a list of enriched signaling pathways from either SPIA or enrichr as input, implementing a flexible weighting scheme for target prioritization, and should therefore be more broadly applicable to diverse indications.

As a proof-of-concept, our secondary analysis of publicly available colorectal cancer data unsurprisingly identified thousands of DEGs. While the signaling pathways that were significantly enriched by these DEGs were used as input to the Pathway2Targets algorithm, we believe it is important to validate the findings that we observed upstream of the signaling pathway data. Some of the most highly up-regulated gene products in our analysis included MMP7, PPBP, and CXCL5. Prior work by others has identified the MMP7 (*Gao et al., 2019; Huang et al., 2021; Vočka et al., 2019*), PPBP (*Chen et al., 2022; Kothalawala & Győrffy, 2023; Feng et al., 2022*), and CXCL5 gene products to be extremely useful in the mechanisms and diagnosis of colorectal cancer (*Chen et al., 2019; Zhang et al., 2021; Novillo et al., 2020; Zhang et al., 2020*). We also identified multiple down-regulated genes in colorectal cancer that are supported by recent studies on the gene products of the OTOP2 (*Qu et al., 2019; Yang & Sakharkar, 2022*), OTOP3 (*Yang & Sakharkar, 2022*), SPIB (*Zhao et al., 2021; Liu et al., 2019*), and INSL5 genes (*Yang et al., 2021c; Sun et al., 2019*). Similarly, a subset of the signaling pathways that we observed have previously been shown to be relevant for this indication including “HIF-1 signaling pathway” (*Lamberti et al., 2019; Seo et al., 2021*) and “PLK1 signaling events” (*Yu et al., 2021; Xie et al., 2021b*).

We predicted the targets from our Pathway2Targets algorithm using only the significant signaling pathways as input, with all other required data retrieved programmatically. Although our initial target prioritization analysis identified hundreds of targets, many researchers tend to focus only on the highest-ranking results for follow-up validation work. Our observation that the top 15 highest-scoring targets were significantly enriched in approved drugs (EGFR, VEGFA, and PTGS2) demonstrates that our flexible weighting scheme can accurately focus subsequent wet-lab validation on “hits” that are more likely to be effective, as has been done previously (*Gray et al., 2022; Moreno et al., 2022; Rapier-Sharman, Clancy & Pickett, 2022; Zeng et al., 2020; Madhukar et al., 2019*). It has not escaped our notice that many of the highest scoring prioritized targets predicted by our algorithm are also well-established biomarkers of cancer including p53, IL-6, VEGFA, EGFR, NF-KB, ESR1, ERBB2, and others (*Penthaloudakis et al., 2019; Wang et al., 2020a; Liebl & Hofmann, 2021; Vainer, Dehlendorff & Johansen, 2018; Wang et al., 2019b; Li et al., 2020; Lupo et al., 2020; Ye et al., 2020; Topi et al., 2022; Bitar et al., 2021*).

Our second target prioritization analysis simulated a scenario that involved prioritizing targets when no/minimal approved drugs exist. Our customizable weighting scheme made this possible by quickly changing the weight parameters in the software. Interestingly, the results from this analysis identified many targets that are relevant to oncology-related indications including MTOR (*Hillmann & Fabbro, 2019; du Rusquec et al., 2020; Lengyel et al., 2020*), MAPK14 (*Fang & Richardson, 2005; Chen et al., 2013; Wang et al., 2022*), and TP53 (*Jovanović et al., 2018; Ciccarese et al., 2017; Yang et al., 2021a*). Each of these targets

have multiple therapeutics at early stages in the pipeline, and consequently rose to the top of the list. This suggests that our approach can accurately identify significant signaling pathways that represent a profile that is both unique to a given disease/condition as well as overlapping with relevant diseases/conditions.

The novelty of our Pathway2Targets target prediction algorithm consists of multiple factors such as open-source code, a single input file of significant pathways, a flexible weighting scheme, incorporation of signaling pathways, output of prioritized targets and therapeutics for repurposing open combines significant signaling pathway enrichment results with publicly available data and our custom weighting scheme can be integrated with safety and efficacy data to further complement ongoing prioritization and repurposing efforts. While the default weighting scheme appears to be capable of identifying relevant targets for CRC in our first analysis, the default parameters may not be universally applicable to all possible use cases. We envision that modifying the weighting scheme should adequately support a wider set of use cases.

We believe that these predictions of approved and novel targets for CRC show promise and that the software warrants additional investigation in other indications. Additional work will be required to determine the extent to which this software can be applied to other -omics technologies such as whole genome sequencing, chromatin immunoprecipitation sequencing (ChIP-seq), shotgun proteomics, *etc*. It is imperative that the user provides high-quality data with minimal numbers of confounding variables, such as timepoints, as input. Although our predictions for CRC are enriched in approved targets, it is important to note that any target predictions made by this software should be confirmed in well-controlled registered clinical trials prior to use in the clinic. Users of the software should also keep in mind that the algorithm only identifies existing therapeutics that are present in the Open Targets platform. As such, it is unrealistic to expect the software to predict drugs that have not been publicly disclosed or that it performs docking predictions to identify novel molecules that bind to novel targets. Similarly, those applying the results to early-stage research should account for toxicity or off-target effects that are associated with a given therapeutic. Even with these caveats, we believe our algorithm can be applied across a broad variety of human diseases and conditions. To do so, users would only need to generate a list of DEGs from their disease of interest, run a pathway enrichment analysis using either SPIA or enrichr, and then use the pathway results as input to the Pathway2Targets software. Future collaborative work will be required to determine how Pathway2Targets can be integrated with precision medicine, artificial intelligence, or other approaches.

We expect that researchers will be interested in applying this open-source target prioritization algorithm to predict the most relevant targets for a disease space regardless of competition in the field. However, some users may need to modify the default weights for one or more parameters in order to yield improved results for any given use case. Although additional future analyses are needed to confirm that this approach remains relevant in other human diseases and conditions, we expect that our flexible and customizable weighting approach will facilitate such efforts from gene expression and

pathway data. In conclusion, we expect that this improved algorithm will facilitate future predictions of therapeutic targets that could be repurposed for other indications.

Operating system

This software is platform-independent and has been successfully tested on 64-bit RedHat Linux and on Mac OS 12.0 and 13.0.

Programming language

This software is written in R (version 3.6.1 or later).

ACKNOWLEDGEMENTS

We gratefully acknowledge the Office of Research Computing at Brigham Young University for providing the computational infrastructure, training, and technical support needed to successfully complete this work.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received no funding for this work.

Competing Interests

Brett E. Pickett is an Academic Editor for PeerJ.

Author Contributions

- Mauri Dobbs Spendlove performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Trenton M. Gibson performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Shaney McCain performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Benjamin C. Stone performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Tristan Gill conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Brett E. Pickett conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

This software is available at GitHub and Zenodo:

- <https://github.com/bpickett/Pathway2Targets>.
- bpickett. (2023). bpickett/Pathway2Targets: v2.26 (v2.26). Zenodo. <https://doi.org/10.5281/zenodo.7693815>.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.16088#supplemental-information>.

REFERENCES

- Aguirre-Plans J, Piñero J, Sanz F, Furlong LI, Fernandez-Fuentes N, Oliva B, Guney E. 2019. GUILDFy v2.0: a tool to identify molecular networks underlying human diseases, their comorbidities and their druggable targets. *Journal of Molecular Biology* 431(13):2477–2484 DOI 10.1016/j.jmb.2019.02.027.
- Anderson E, Havener TM, Zorn KM, Foil DH, Lane TR, Capuzzi SJ, Morris D, Hickey AJ, Drewry DH, Ekins S. 2020. Synergistic drug combinations and machine learning for drug repurposing in chordoma. *Scientific Reports* 10(1):12982 DOI 10.1038/s41598-020-70026-w.
- Aoki-Kinoshita KF, Kanehisa M. 2007. Gene annotation and pathway mapping in KEGG. *Methods in Molecular Biology* 396:71–91 DOI 10.1007/978-1-59745-515-2.
- Ballard C, Aarsland D, Cummings J, O'Brien J, Mills R, Molinuevo JL, Fladby T, Williams G, Doherty P, Corbett A, Sultana J. 2020. Drug repositioning and repurposing for Alzheimer disease. *Nature Reviews Neurology* 16(12):661–673 DOI 10.1038/s41582-020-0397-4.
- Barrio-Hernandez I, Schwartzenbuber J, Shrivastava A, Del-Toro N, Gonzalez A, Zhang Q, Mountjoy E, Suveges D, Ochoa D, Ghousaini M, Bradley G, Hermjakob H, Orchard S, Dunham I, Anderson CA, Porras P, Beltrao P. 2023. Network expansion of genetic associations defines a pleiotropy map of human cell biology. *Nature Genetics* 55:389–398 DOI 10.1038/s41588-023-01327-9.
- Begley CG, Ashton M, Baell J, Bettess M, Brown MP, Carter B, Charman WN, Davis C, Fisher S, Frazer I, Gautam A, Jennings MP, Kearney P, Keeffe E, Kelly D, Lopez AF, McGuckin M, Parker MW, Rayner C, Roberts B, Rush JS, Sullivan M. 2021. Drug repurposing: misconceptions, challenges, and opportunities for academic researchers. *Science Translational Medicine* 13(612):eabd5524 DOI 10.1126/scitranslmed.abd5524.
- Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, Santos R, Rao Y, Sassi F, Pinnelli M, Ansari R, Harper S, Jackson DA, McRae R, Pooley R, Wilkinson P, van der Meer D, Dow D, Buser-Doeppner C, Bertotti A, Trusolino L, Stronach EA, Saez-Rodriguez J, Yusa K, Garnett MJ. 2019. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* 568(7753):511–516 DOI 10.1038/s41586-019-1103-9.
- Bitar L, Zouein J, Haddad FG, Eid R, Kourie HR. 2021. HER2 in metastatic colorectal cancer: a new to target to remember. *Biomarkers in Medicine* 15(2):133–136 DOI 10.2217/bmm-2020-0491.
- Carrella D, Napolitano F, Rispoli R, Miglietta M, Carissimo A, Cutillo L, Sirci F, Gregoretti F, Di Bernardo D. 2014. Mantra 2.0: an online collaborative resource for drug mode of action and repurposing by network analysis. *Bioinformatics* 30(12):1787–1788 DOI 10.1093/bioinformatics/btu058.
- Chatterjee A, Paul S, Bisht B, Bhattacharya S, Sivasubramaniam S, Paul MK. 2022. Advances in targeting the WNT/β-catenin signaling pathway in cancer. *Drug Discovery Today* 27(1):82–101 DOI 10.1016/j.drudis.2021.07.007.
- Chen YJ, Cheng YJ, Hung AC, Wu YC, Hou MF, Tyan YC, Yuan SSF. 2013. The synthetic flavonoid WYC02-9 inhibits cervical cancer cell migration/invasion and angiogenesis via MAPK14 signaling. *Gynecologic Oncology* 131(3):734–743 DOI 10.1016/j.ygyno.2013.10.012.

- Chen Y, Xu R.** 2016. Drug repurposing for glioblastoma based on molecular subtypes. *Journal of Biomedical Informatics* **64**(12):131–138 DOI [10.1016/j.jbi.2016.09.019](https://doi.org/10.1016/j.jbi.2016.09.019).
- Chen C, Xu ZQ, Zong YP, Ou BC, Shen XH, Feng H, Zheng MH, Zhao JK, Lu AG.** 2019. CXCL5 induces tumor angiogenesis via enhancing the expression of FOXD1 mediated by the AKT/NF- κ B pathway in colorectal cancer. *Cell Death & Disease* **10**(3):178 DOI [10.1038/s41419-019-1431-6](https://doi.org/10.1038/s41419-019-1431-6).
- Chen D, Ye Z, Lew Z, Luo S, Yu Z, Lin Y.** 2022. Expression of NMU, PPBP and GNG4 in colon cancer and their influences on prognosis. *Translational Cancer Research* **11**(10):3572–3583 DOI [10.21037/tcr-22-1377](https://doi.org/10.21037/tcr-22-1377).
- Cheng F, Lu W, Liu C, Fang J, Hou Y, Handy DE, Wang R, Zhao Y, Yang Y, Huang J, Hill DE, Vidal M, Eng C, Loscalzo J.** 2019. A genome-wide positioning systems network algorithm for in silico drug repurposing. *Nature Communications* **10**(1):3476 DOI [10.1038/s41467-019-10744-6](https://doi.org/10.1038/s41467-019-10744-6).
- Choi HS, Kim SL, Kim JH, Lee DS.** 2020. The FDA-approved anti-asthma medicine ciclesonide inhibits lung cancer stem cells through hedgehog signaling-mediated SOX2 regulation. *International Journal of Molecular Sciences* **21**(3):1014 DOI [10.3390/ijms21031014](https://doi.org/10.3390/ijms21031014).
- Ciccarese C, Massari F, Blanca A, Tortora G, Montironi R, Cheng L, Scarpelli M, Raspollini MR, Vau N, Fonseca J, Lopez-Beltran A.** 2017. Tp53 and its potential therapeutic role as a target in bladder cancer. *Expert Opinion on Therapeutic Targets* **21**(4):401–414 DOI [10.1080/14728222.2017.1297798](https://doi.org/10.1080/14728222.2017.1297798).
- Crowther GJ, Shanmugam D, Carmona SJ, Doyle MA, Hertz-Fowler C, Berriman M, Nwaka S, Ralph SA, Roos DS, Van Voorhis WC, Agüero F.** 2010. Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. *PLOS Neglected Tropical Diseases* **4**(8):e804 DOI [10.1371/journal.pntd.0000804](https://doi.org/10.1371/journal.pntd.0000804).
- Damale MG, Pathan SK, Shinde DB, Patil RH, Arote RB, Sangshetti JN.** 2020. Insights of tankyrases: a novel target for drug discovery. *European Journal of Medicinal Chemistry* **207**:112712 DOI [10.1016/j.ejmech.2020.112712](https://doi.org/10.1016/j.ejmech.2020.112712).
- Deng L, Meng T, Chen L, Wei W, Wang P.** 2020. The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduction and Targeted Therapy* **5**(1):11 DOI [10.1038/s41392-020-0107-0](https://doi.org/10.1038/s41392-020-0107-0).
- Dezső Z, Ceccarelli M.** 2020. Machine learning prediction of oncology drug targets based on protein and network properties. *BMC Bioinformatics* **21**(1):104 DOI [10.1186/s12859-020-3442-9](https://doi.org/10.1186/s12859-020-3442-9).
- Di J, Zheng B, Kong Q, Jiang Y, Liu S, Yang Y, Han X, Sheng Y, Zhang Y, Cheng L, Han J.** 2019. Prioritization of candidate cancer drugs based on a drug functional similarity network constructed by integrating pathway activities and drug activities. *Molecular Oncology* **13**(10):2259–2277 DOI [10.1002/1878-0261.12564](https://doi.org/10.1002/1878-0261.12564).
- Ding C, Song Z, Shen A, Chen T, Zhang A.** 2020. Small molecules targeting the innate immune cGAS-STING-TBK1 signaling pathway. *Acta Pharmaceutica Sinica B* **10**(12):2272–2298 DOI [10.1016/j.apsb.2020.03.001](https://doi.org/10.1016/j.apsb.2020.03.001).
- du Rusquec P, Blonz C, Frenel JS, Campone M.** 2020. Targeting the PI3K/Akt/mTOR pathway in estrogen-receptor positive HER2 negative advanced breast cancer. *Therapeutic Advances in Medical Oncology* **12**:1758835920940939 DOI [10.1177/1758835920940939](https://doi.org/10.1177/1758835920940939).
- Duan Q, Reid SP, Clark NR, Wang Z, Fernandez NF, Rouillard AD, Readhead B, Tritsch SR, Hodos R, Hafner M, Niepel M, Sorger PK, Dudley JT, Bavari S, Panchal RG, Ma'ayan A.** 2016. L1000CDS: LINCS L1000 characteristic direction signatures search engine. *NPJ Systems Biology and Applications* **2**:16015 DOI [10.1038/npjbsa.2016.15](https://doi.org/10.1038/npjbsa.2016.15).

- Dwane L, Behan FM, Gonçalves E, Lightfoot H, Yang W, van der Meer D, Shepherd R, Pignatelli M, Iorio F, Garnett MJ. 2021. Project score database: a resource for investigating cancer cell dependencies and prioritizing therapeutic targets. *Nucleic Acids Research* 49(D1):D1365–D1372 DOI 10.1093/nar/gkaa882.
- Emig D, Ivliev A, Pustovalova O, Lancashire L, Bureeva S, Nikolsky Y, Bessarabova M. 2013. Drug target prediction and repositioning using an integrated network-based approach. *PLOS ONE* 8(4):e60618 DOI 10.1371/journal.pone.0060618.
- Fang JY, Richardson BC. 2005. The MAPK signalling pathways and colorectal cancer. *The Lancet Oncology* 6(5):322–327 DOI 10.1016/S1470-2045(05)70168-6.
- Fang H, ULTRA-DD Consortium, De Wolf H, Knezevic B, Burnham KL, Osgood J, Sanniti A, Lledó Lara A, Kasela S, De Cesco S, Wegner JK, Handunnetthi L, McCann FE, Chen L, Sekine T, Brennan PE, Marsden BD, Damerell D, O'Callaghan CA, Bountra C, Bowness P, Sundström Y, Milani L, Berg L, Göhlmann HW, Peeters PJ, Fairfax BP, Sundström M, Knight JC. 2019. A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nature Genetics* 51(7):1082–1091 DOI 10.1038/s41588-019-0456-1.
- Feng W, Zhang Y, Liu W, Wang X, Lei T, Yuan Y, Chen Z, Song W. 2022. A prognostic model using immune-related genes for colorectal cancer. *Frontiers in Cell and Developmental Biology* 10:813043 DOI 10.3389/fcell.2022.813043.
- Ferrarini MG, Lal A, Rebollo R, Gruber AJ, Guerracino A, Gonzalez IM, Floyd T, de Oliveira DS, Shanklin J, Beausoleil E, Pusa T, Pickett BE, Aguiar-Pulido V. 2021. Genome-wide bioinformatic analyses predict key host and viral factors in SARS-CoV-2 pathogenesis. *Communications Biology* 4(1):590 DOI 10.1038/s42003-021-02095-0.
- Fiscon G, Conte F, Farina L, Paci P. 2021. SAVeRUNNER: a network-based algorithm for drug repurposing and its application to COVID-19. *PLOS Computational Biology* 17(2):e1008686 DOI 10.1371/journal.pcbi.1008686.
- Fiscon G, Paci P. 2021. SAVeRUNNER: an R-based tool for drug repurposing. *BMC Bioinformatics* 22(1):150 DOI 10.1186/s12859-021-04076-w.
- Gao Y, Nan X, Shi X, Mu X, Liu B, Zhu H, Yao B, Liu X, Yang T, Hu Y, Liu S. 2019. SREBP1 promotes the invasion of colorectal cancer accompanied upregulation of MMP7 expression and NF-κB pathway activation. *BMC Cancer* 19(1):685 DOI 10.1186/s12885-019-5904-x.
- Gifford KTL, Pickett BE. 2022. Comparative meta-analysis of host transcriptional response during *Streptococcus pneumoniae* carriage or infection. *Microbial Pathogenesis* 173:105816 DOI 10.1016/j.micpath.2022.105816.
- Glanz A, Chawla K, Fabry S, Subramanian G, Garcia J, Jay B, Ciricillo J, Chakravarti R, Taylor RT, Chattopadhyay S. 2020. High throughput screening of FDA-approved drug library reveals the compounds that promote IRF3-mediated pro-apoptotic pathway inhibit virus replication. *Viruses* 12(4):442 DOI 10.3390/v12040442.
- Gray M, Guerrero-Arguero I, Solis-Leal A, Robison RA, Berges BK, Pickett BE. 2022. Chikungunya virus time course infection of human macrophages reveals intracellular signaling pathways relevant to repurposed therapeutics. *PeerJ* 10(7):e13090 DOI 10.7717/peerj.13090.
- Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, Zhang R, Hartmann BM, Zaslavsky E, Sealfon SC, Chasman DI, FitzGerald GA, Dolinski K, Grosser T, Troyanskaya OG. 2015. Understanding multicellular function and disease with human tissue-specific networks. *Nature Genetics* 47(6):569–576 DOI 10.1038/ng.3259.
- Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. 2021. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular Diversity* 25(3):1315–1360 DOI 10.1007/s11030-021-10217-3.

- Harb J, Lin PJ, Hao J.** 2019. Recent development of Wnt signaling pathway inhibitors for cancer therapeutics. *Current Oncology Reports* 21(2):12 DOI 10.1007/s11912-019-0763-9.
- Hernandez JJ, Pryszlak M, Smith L, Yanchus C, Kurji N, Shahani VM, Molinski SV.** 2017. Giving drugs a second chance: overcoming regulatory and financial hurdles in repurposing approved drugs as cancer therapeutics. *Frontiers in Oncology* 7:273 DOI 10.3389/fonc.2017.00273.
- Hillmann P, Fabbro D.** 2019. PI3K/mTOR pathway inhibition: opportunities in oncology and rare genetic diseases. *International Journal of Molecular Sciences* 20(22):5792 DOI 10.3390/ijms20225792.
- Huang JK, Carlin DE, Yu MK, Zhang W, Kreisberg JF, Tamayo P, Ideker T.** 2018. Systematic evaluation of molecular networks for discovery of disease genes. *Cell Systems* 6(4):484–495.e5 DOI 10.1016/j.cels.2018.03.001.
- Huang A, Garraway LA, Ashworth A, Weber B.** 2020. Synthetic lethality as an engine for cancer drug target discovery. *Nature Reviews Drug Discovery* 19(1):23–38 DOI 10.1038/s41573-019-0046-z.
- Huang X, Lan Y, Li E, Li J, Deng Q, Deng X.** 2021. Diagnostic values of MMP-7, MMP-9, MMP-11, TIMP-1, TIMP-2, CEA, and CA19-9 in patients with colorectal cancer. *Journal of International Medical Research* 49(5):3000605211012570 DOI 10.1177/0300605211012570.
- Huang L, Li F, Sheng J, Xia X, Ma J, Zhan M, Wong STC.** 2014. DrugComboRanker: drug combination discovery based on target network analysis. *Bioinformatics* 30(12):i228–i236 DOI 10.1093/bioinformatics/btu278.
- Isik Z, Baldow C, Cannistraci CV, Schroeder M.** 2015. Drug target prioritization by perturbed gene expression and network information. *Scientific Reports* 5(1):17417 DOI 10.1038/srep17417.
- Jain AS, Prasad A, Pradeep S, Dharmashekar C, Achar RR, Ekaterina S, Victor S, Amachawadi RG, Prasad SK, Pruthvish R, Syed A, Shivamallu C, Kollur SP.** 2021. Everything old is new again: drug repurposing approach for non-small cell lung cancer targeting MAPK signaling pathway. *Frontiers in Oncology* 11:741326 DOI 10.3389/fonc.2021.741326.
- Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, Loney F, May B, Milacic M, Rothfels K, Sevilla C, Shamovsky V, Shorser S, Varusai T, Weiser J, Wu G, Stein L, Hermjakob H, D'Eustachio P.** 2020. The reactome pathway knowledgebase. *Nucleic Acids Research* 48(D1):D498–D503 DOI 10.1093/nar/gkz1031.
- Jovanović KK, Escure G, Demonchy J, Willaume A, Van de Wyngaert Z, Farhat M, Chauvet P, Facon T, Quesnel B, Manier S.** 2018. Dereulation and targeting of TP53 pathway in multiple myeloma. *Frontiers in Oncology* 8:665 DOI 10.3389/fonc.2018.00665.
- Kasprzyk A.** 2011. BioMart: driving a paradigm change in biological data management. *Database* 2011:bar049 DOI 10.1093/database/bar049.
- Khojasteh Poor F, Keivan M, Ramazii M, Ghaedrahmati F, Anbiyaee A, Panahandeh S, Khoshnam SE, Farzaneh M.** 2021. Mini review: the FDA-approved prescription drugs that target the MAPK signaling pathway in women with breast cancer. *Breast Disease* 40(2):51–62 DOI 10.3233/BD-201063.
- Kingsmore KM, Grammer AC, Lipsky PE.** 2020. Drug repurposing to improve treatment of rheumatic autoimmune inflammatory diseases. *Nature Reviews Rheumatology* 16(1):32–52 DOI 10.1038/s41584-019-0337-0.
- Köster J, Rahmann S.** 2018. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* 34(20):3600 DOI 10.1093/bioinformatics/bty350.

- Kothalawala WJ, Győrffy B.** 2023. Transcriptomic and cellular content analysis of colorectal cancer by combining multiple independent cohorts. *Clinical and Translational Gastroenterology* 14(2):e00517 DOI [10.14309/ctg.0000000000000517](https://doi.org/10.14309/ctg.0000000000000517).
- Lamberti MJ, Rettel M, Krijgsveld J, Rivarola VA, Rumie Vittar NB.** 2019. Secretome profiling of heterotypic spheroids suggests a role of fibroblasts in HIF-1 pathway modulation and colorectal cancer photodynamic resistance. *Cellular Oncology* 42(2):173–196 DOI [10.1007/s13402-018-00418-8](https://doi.org/10.1007/s13402-018-00418-8).
- Lee I, Blom UM, Wang PI, Shim JE, Marcotte EM.** 2011. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Research* 21(7):1109–1121 DOI [10.1101/gr.118992.110](https://doi.org/10.1101/gr.118992.110).
- Lengyel CG, Altuna SC, Habeeb BS, Trapani D, Khan SZ.** 2020. The potential of PI3K/AKT/mTOR signaling as a druggable target for endometrial and ovarian carcinomas. *Current Drug Targets* 21(10):946–961 DOI [10.2174/138945012066191120123612](https://doi.org/10.2174/138945012066191120123612).
- Li QL, Lin X, Yu YL, Chen L, Hu QX, Chen M, Cao N, Zhao C, Wang CY, Huang CW, Li LY, Ye M, Wu M.** 2021. Genome-wide profiling in colorectal cancer identifies PHF19 and TBC1D16 as oncogenic super enhancers. *Nature Communications* 12(1):6407 DOI [10.1038/s41467-021-26600-5](https://doi.org/10.1038/s41467-021-26600-5).
- Li J, Lu Z.** 2013. Pathway-based drug repositioning using causal inference. *BMC Bioinformatics* 14(Suppl 16):S3 DOI [10.1186/1471-2105-14-S16-S3](https://doi.org/10.1186/1471-2105-14-S16-S3).
- Li QH, Wang YZ, Tu J, Liu CW, Yuan YJ, Lin R, He WL, Cai SR, He YL, Ye JN.** 2020. Anti-EGFR therapy in metastatic colorectal cancer: mechanisms and potential regimens of drug resistance. *Gastroenterology Report* 8(3):179–191 DOI [10.1093/gastro/goa026](https://doi.org/10.1093/gastro/goa026).
- Liebl MC, Hofmann TG.** 2021. The role of p53 signaling in colorectal cancer. *Cancers* 13(9):2125 DOI [10.3390/cancers13092125](https://doi.org/10.3390/cancers13092125).
- Liu Q, Deng J, Wei X, Yuan W, Ma J.** 2019. Integrated analysis of competing endogenous RNA networks revealing five prognostic biomarkers associated with colorectal cancer. *Journal of Cellular Biochemistry* 120(7):11256–11264 DOI [10.1002/jcb.28403](https://doi.org/10.1002/jcb.28403).
- Liu Z, Wang P, Wold EA, Song Q, Zhao C, Wang C, Zhou J.** 2021. Small-molecule inhibitors targeting the canonical WNT signaling pathway for the treatment of cancer. *Journal of Medicinal Chemistry* 64(8):4257–4288 DOI [10.1021/acs.jmedchem.0c01799](https://doi.org/10.1021/acs.jmedchem.0c01799).
- Louhimo R, Laakso M, Belitskin D, Klefström J, Lehtonen R, Hautaniemi S.** 2016. Data integration to prioritize drugs using genomics and curated data. *BioData Mining* 9(1):21 DOI [10.1186/s13040-016-0097-1](https://doi.org/10.1186/s13040-016-0097-1).
- Lupo B, Sassi F, Pinnelli M, Galimi F, Zanella ER, Vurchio V, Migliardi G, Gagliardi PA, Puliafito A, Manganaro D, Luraghi P, Kragh M, Pedersen MW, Horak ID, Boccaccio C, Medico E, Primo L, Nichol D, Spiteri I, Heide T, Vatsiou A, Graham TA, Élez E, Argiles G, Nuciforo P, Sottoriva A, Dienstmann R, Pasini D, Grassi E, Isella C, Bertotti A, Trusolino L.** 2020. Colorectal cancer residual disease at maximal response to EGFR blockade displays a druggable Paneth cell-like phenotype. *Science Translational Medicine* 12(555):aax8313 DOI [10.1126/scitranslmed.aax8313](https://doi.org/10.1126/scitranslmed.aax8313).
- Ma J, Wang J, Ghoraie LS, Men X, Haibe-Kains B, Dai P.** 2019. A comparative study of cluster detection algorithms in protein-protein interaction for drug target discovery and drug repurposing. *Frontiers in Pharmacology* 10:109 DOI [10.3389/fphar.2019.00109](https://doi.org/10.3389/fphar.2019.00109).
- Madhukar NS, Khade PK, Huang L, Gayvert K, Galletti G, Stogniew M, Allen JE, Giannakakou P, Elemento O.** 2019. A Bayesian machine learning approach for drug target identification using diverse data types. *Nature Communications* 10(1):5221 DOI [10.1038/s41467-019-12928-6](https://doi.org/10.1038/s41467-019-12928-6).

- Malas TB, Vlietstra WJ, Kudrin R, Starikov S, Charrout M, Roos M, Peters DJM, Kors JA, Vos R, 't Hoen PAC, van Mulligen EM, Hettne KM.** 2019. Drug prioritization using the semantic properties of a knowledge graph. *Scientific Reports* **9**(1):6281 DOI [10.1038/s41598-019-42806-6](https://doi.org/10.1038/s41598-019-42806-6).
- Masoudi-Sobhanzadeh Y, Omidi Y, Amanlou M, Masoudi-Nejad A.** 2020. Drug databases and their contributions to drug repurposing. *Genomics* **112**(2):1087–1095 DOI [10.1016/j.ygeno.2019.06.021](https://doi.org/10.1016/j.ygeno.2019.06.021).
- Mejía-Pedroza RA, Espinal-Enríquez J, Hernández-Lemus E.** 2018. Pathway-based drug repositioning for breast cancer molecular subtypes. *Frontiers in Pharmacology* **9**:905 DOI [10.3389/fphar.2018.00905](https://doi.org/10.3389/fphar.2018.00905).
- Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD.** 2017. PANTHER version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research* **45**(D1):D183–D189 DOI [10.1093/nar/gkw1138](https://doi.org/10.1093/nar/gkw1138).
- Moreno C, Bybee E, Tellez Freitas CM, Pickett BE, Weber KS.** 2022. Meta-analysis of two human RNA-seq datasets to determine periodontitis diagnostic biomarkers and drug target candidates. *International Journal of Molecular Sciences* **23**(10):5580 DOI [10.3390/ijms23105580](https://doi.org/10.3390/ijms23105580).
- Napolitano F, Carrella D, Mandriani B, Pisonero-Vaquero S, Sirci F, Medina DL, Brunetti-Pierri N, di Bernardo D.** 2018. gene2drug: a computational tool for pathway-based rational drug repositioning. *Bioinformatics* **34**(9):1498–1505 DOI [10.1093/bioinformatics/btx800](https://doi.org/10.1093/bioinformatics/btx800).
- Novillo A, Gaibar M, Romero-Lorca A, Gilsanz MF, Beltrán L, Galán M, Antón B, Malón D, Moreno A, Fernández-Santander A.** 2020. Efficacy of bevacizumab-containing chemotherapy in metastatic colorectal cancer and expression: six case reports. *World Journal of Gastroenterology* **26**(16):1979–1986 DOI [10.3748/wjg.v26.i16.1979](https://doi.org/10.3748/wjg.v26.i16.1979).
- Ochoa D, Hercules A, Carmona M, Suveges D, Gonzalez-Uriarte A, Malangone C, Miranda A, Fumis L, Carvalho-Silva D, Spitzer M, Baker J, Ferrer J, Raies A, Razuvayevskaya O, Faulconbridge A, Petsalaki E, Mutowo P, Machlitt-Northen S, Peat G, McAuley E, Ong CK, Mountjoy E, Ghousaini M, Pierleoni A, Papa E, Pignatelli M, Koscielny G, Karim M, Schwartzentruber J, Hulcoop DG, Dunham I, McDonagh EM.** 2021. Open Targets Platform: supporting systematic drug-target identification and prioritisation. *Nucleic Acids Research* **49**(D1):D1302–D1310 DOI [10.1093/nar/gkaa1027](https://doi.org/10.1093/nar/gkaa1027).
- Olgen S.** 2019. A prospective overview of drug repurposing in drug discovery and development. *Current Medicinal Chemistry* **26**(28):5338–5339 DOI [10.2174/092986732628191025094454](https://doi.org/10.2174/092986732628191025094454).
- Orjuela S, Huang R, Hembach KM, Robinson MD, Soneson C.** 2019. ARMOR: an automated reproducible modular workflow for preprocessing and differential analysis of RNA-seq data. *G3 Genes|Genomes|Genetics* **9**(7):2089–2096 DOI [10.1534/g3.119.400185](https://doi.org/10.1534/g3.119.400185).
- Ozdemir ES, Halakou F, Nussinov R, Gursoy A, Keskin O.** 2019. Methods for discovering and targeting druggable protein-protein interfaces and their application to repurposing. *Methods in Molecular Biology* **1903**:1–21 DOI [10.1007/978-1-4939-8955-3](https://doi.org/10.1007/978-1-4939-8955-3).
- Paananen J, Fortino V.** 2020. An omics perspective on drug target discovery platforms. *Briefings in Bioinformatics* **21**(6):1937–1953 DOI [10.1093/bib/bbz122](https://doi.org/10.1093/bib/bbz122).
- Parvathaneni V, Kulkarni NS, Muth A, Gupta V.** 2019. Drug repurposing: a promising tool to accelerate the drug discovery process. *Drug Discovery Today* **24**(10):2076–2085 DOI [10.1016/j.drudis.2019.06.014](https://doi.org/10.1016/j.drudis.2019.06.014).
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C.** 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14**(4):417–419 DOI [10.1038/nmeth.4197](https://doi.org/10.1038/nmeth.4197).

- Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. 2021.** Artificial intelligence in drug discovery and development. *Drug Discovery Today* **26**(1):80–93
[DOI 10.1016/j.drudis.2020.10.010](https://doi.org/10.1016/j.drudis.2020.10.010).
- Pentheroudakis G, Mavroelidis L, Papadopoulou K, Kolioi GA, Bamia C, Chatzopoulos K, Samantas E, Mauri D, Efstratiou I, Pectasides D, Makatsoris T, Bafaloukos D, Papakostas P, Papatsibas G, Bombolaki I, Chrisafi S, Kourea HP, Petraki K, Kafiri G, Fountzilas G, Kotoula V. 2019.** Angiogenic and antiangiogenic VEGFA splice variants in colorectal cancer: prospective retrospective cohort study in patients treated with irinotecan-based chemotherapy and bevacizumab. *Clinical Colorectal Cancer* **18**(4):e370–e384 [DOI 10.1016/j.clcc.2019.07.007](https://doi.org/10.1016/j.clcc.2019.07.007).
- Proctor AE, Thompson LA, O'Bryant CL. 2014.** Vismodegib: an inhibitor of the Hedgehog signaling pathway in the treatment of basal cell carcinoma. *Annals of Pharmacotherapy* **48**(1):99–106 [DOI 10.1177/1060028013506696](https://doi.org/10.1177/1060028013506696).
- Qu H, Su Y, Yu L, Zhao H, Xin C. 2019.** Wild-type p53 regulates OTOP2 transcription through DNA loop alteration of the promoter in colorectal cancer. *FEBS Open Bio* **9**(1):26–34
[DOI 10.1002/2211-5463.12554](https://doi.org/10.1002/2211-5463.12554).
- Rapier-Sharman N, Clancy J, Pickett BE. 2022.** Joint secondary transcriptomic analysis of non-Hodgkin's B-cell lymphomas predicts reliance on pathways associated with the extracellular matrix and robust diagnostic biomarkers. *Journal of Bioinformatics and Systems Biology* **5**(4):119–135.
- Regan-Fendt K, Li D, Reyes R, Yu L, Wani NA, Hu P, Jacob ST, Ghoshal K, Payne PRO, Motiwala T. 2020.** Transcriptomics-based drug repurposing approach identifies novel drugs against sorafenib-resistant hepatocellular carcinoma. *Cancers* **12**(10):2730
[DOI 10.3390/cancers12102730](https://doi.org/10.3390/cancers12102730).
- Ren X, Duan L, He Q, Zhang Z, Zhou Y, Wu D, Pan J, Pei D, Ding K. 2010.** Identification of niclosamide as a new small-molecule inhibitor of the STAT3 signaling pathway. *ACS Medicinal Chemistry Letters* **1**(9):454–459 [DOI 10.1021/ml100146z](https://doi.org/10.1021/ml100146z).
- Robinson MD, McCarthy DJ, Smyth GK. 2010.** edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1):139–140
[DOI 10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616).
- Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W, Marchler-Bauer A, Landrum M, Lathrop S, Lu Z, Madden TL, O'Leary N, Phan L, Rangwala SH, Schneider VA, Skripchenko Y, Wang J, Ye J, Trawick BW, Pruitt KD, Sherry ST. 2021.** Database resources of the national center for biotechnology information. *Nucleic Acids Research* **49**(D1):D10–D17 [DOI 10.1093/nar/gkaa892](https://doi.org/10.1093/nar/gkaa892).
- Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, Buetow KH. 2009.** PID: the pathway interaction database. *Nucleic Acids Research* **37**(Database issue):D674–D679
[DOI 10.1093/nar/gkn653](https://doi.org/10.1093/nar/gkn653).
- Schein CH. 2020.** Repurposing approved drugs on the pathway to novel therapies. *Medicinal Research Reviews* **40**(2):586–605 [DOI 10.1002/med.21627](https://doi.org/10.1002/med.21627).
- Scott TM, Jensen S, Pickett BE. 2021.** A signaling pathway-driven bioinformatics pipeline for predicting therapeutics against emerging infectious diseases. *F1000Research* **10**:330
[DOI 10.12688/f1000research.52412.2](https://doi.org/10.12688/f1000research.52412.2).
- Scott TM, Solis-Leal A, Lopez JB, Robison RA, Berges BK, Pickett BE. 2022.** Comparison of intracellular transcriptional response of NHBE cells to infection with SARS-CoV-2 Washington and New York strains. *Frontiers in Cellular and Infection Microbiology* **12**:1009328
[DOI 10.3389/fcimb.2022.1009328](https://doi.org/10.3389/fcimb.2022.1009328).

- Seo J, Yun J, Fukuda J, Chun YS.** 2021. Tumor-intrinsic FABP5 is a novel driver for colon cancer cell growth via the HIF-1 signaling pathway. *Cancer Genetics* **258–259**(6):151–156 DOI [10.1016/j.cancergen.2021.11.001](https://doi.org/10.1016/j.cancergen.2021.11.001).
- Setoain J, Franch M, Martínez M, Tabas-Madrid D, Sorzano COS, Bakker A, Gonzalez-Couto E, Elvira J, Pascual-Montano A.** 2015. NFFinder: an online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. *Nucleic Acids Research* **43**(W1):W193–W199 DOI [10.1093/nar/gkv445](https://doi.org/10.1093/nar/gkv445).
- Sharma PP, Bansal M, Sethi A, Poonam, Pena L, Goel VK, Grishina M, Chaturvedi S, Kumar D, Rathi B.** 2021. Computational methods directed towards drug repurposing for COVID-19: advantages and limitations. *RSC Advances* **11**(57):36181–36198 DOI [10.1039/D1RA05320E](https://doi.org/10.1039/D1RA05320E).
- Sleno L, Emili A.** 2008. Proteomic methods for drug target discovery. *Current Opinion in Chemical Biology* **12**(1):46–54 DOI [10.1016/j.cbpa.2008.01.022](https://doi.org/10.1016/j.cbpa.2008.01.022).
- Sun G, Li Y, Peng Y, Lu D, Zhang F, Cui X, Zhang Q, Li Z.** 2019. Identification of differentially expressed genes and biological characteristics of colorectal cancer by integrated bioinformatics analysis. *Journal of Cellular Physiology* **234**(9):15215–15224 DOI [10.1002/jcp.28163](https://doi.org/10.1002/jcp.28163).
- Tanoli Z, Vähä-Koskela M, Aittokallio T.** 2021. Artificial intelligence, machine learning, and drug repurposing in cancer. *Expert Opinion on Drug Discovery* **16**(9):977–989 DOI [10.1080/17460441.2021.1883585](https://doi.org/10.1080/17460441.2021.1883585).
- Tarca AI, Draghici S, Khatri P, Hassan SS, Mittal P, Kim JS, Kim CJ, Kusanovic JP, Romero R.** 2009. A novel signaling pathway impact analysis. *Bioinformatics* **25**(1):75–82 DOI [10.1093/bioinformatics/btn577](https://doi.org/10.1093/bioinformatics/btn577).
- Thakur A, Tan Z, Kameyama T, El-Khateeb E, Nagpal S, Malone S, Jamwal R, Nwabufo CK.** 2021. Bioanalytical strategies in drug discovery and development. *Drug Metabolism Reviews* **53**(3):434–458 DOI [10.1080/03602532.2021.1959606](https://doi.org/10.1080/03602532.2021.1959606).
- Topi G, Ghatak S, Satapathy SR, Ehrnström R, Lydrup ML, Sjölander A.** 2022. Combined estrogen alpha and beta receptor expression has a prognostic significance for colorectal cancer patients. *Frontiers in Medicine* **9**:739620 DOI [10.3389/fmed.2022.739620](https://doi.org/10.3389/fmed.2022.739620).
- Tsuji S, Hase T, Yachie-Kinoshita A, Nishino T, Ghosh S, Kikuchi M, Shimokawa K, Aburatani H, Kitano H, Tanaka H.** 2021. Artificial intelligence-based computational framework for drug-target prioritization and inference of novel repositionable drugs for Alzheimer's disease. *Alzheimer's Research & Therapy* **13**(1):92 DOI [10.1186/s13195-021-00826-3](https://doi.org/10.1186/s13195-021-00826-3).
- UniProt Consortium.** 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* **47**(D1):D506–D515 DOI [10.1093/nar/gky1049](https://doi.org/10.1093/nar/gky1049).
- Urán Landaburu L, Berenstein AJ, Videla S, Maru P, Shanmugam D, Chernomorez A, Agüero F.** 2020. TDR Targets 6: driving drug discovery for human pathogens through intensive chemogenomic data integration. *Nucleic Acids Research* **48**(D1):D992–D1005 DOI [10.1093/nar/gkz999](https://doi.org/10.1093/nar/gkz999).
- Vainer N, Dehlendorff C, Johansen JS.** 2018. Systematic literature review of IL-6 as a biomarker or treatment target in patients with gastric, bile duct, pancreatic and colorectal cancer. *Oncotarget* **9**(51):29820–29841 DOI [10.18632/oncotarget.25661](https://doi.org/10.18632/oncotarget.25661).
- Vočka M, Langer D, Fryba V, Petrtyl J, Hanus T, Kalousova M, Zima T, Petruzelka L.** 2019. Serum levels of TIMP-1 and MMP-7 as potential biomarkers in patients with metastatic colorectal cancer. *The International Journal of Biological Markers* **34**(3):292–301 DOI [10.1177/1724600819866202](https://doi.org/10.1177/1724600819866202).

- Wang Z, He E, Sani K, Jagodnik KM, Silverstein MC, Ma'ayan A.** 2019a. Drug Gene Budger (DGB): an application for ranking drugs to modulate a specific gene based on transcriptomic signatures. *Bioinformatics* 35(7):1247–1248 DOI [10.1093/bioinformatics/bty763](https://doi.org/10.1093/bioinformatics/bty763).
- Wang R, Ma Y, Zhan S, Zhang G, Cao L, Zhang X, Shi T, Chen W.** 2020a. B7-H3 promotes colorectal cancer angiogenesis through activating the NF-κB pathway to induce VEGFA expression. *Cell Death & Disease* 11(1):55 DOI [10.1038/s41419-020-2252-3](https://doi.org/10.1038/s41419-020-2252-3).
- Wang D, Peng L, Hua L, Li J, Liu Y, Zhou Y.** 2022. Mapk14 is a prognostic biomarker and correlates with the clinicopathological features and immune infiltration of colorectal cancer. *Frontiers in Cell and Developmental Biology* 10:817800 DOI [10.3389/fcell.2022.817800](https://doi.org/10.3389/fcell.2022.817800).
- Wang T, Song P, Zhong T, Wang X, Xiang X, Liu Q, Chen H, Xia T, Liu H, Niu Y, Hu Y, Xu L, Shao Y, Zhu L, Qi H, Shen J, Hou T, Fodde R, Shao J.** 2019b. The inflammatory cytokine IL-6 induces FRA1 deacetylation promoting colorectal cancer stem-like properties. *Oncogene* 38(25):4932–4947 DOI [10.1038/s41388-019-0763-0](https://doi.org/10.1038/s41388-019-0763-0).
- Wang H, Xu X, Guan X, Shen S, Huang X, Kai G, Zhao S, Ruan W, Zhang L, Pang T, Mo R.** 2020b. Liposomal 9-aminoacridine for treatment of ischemic stroke: from drug discovery to drug delivery. *Nano Letters* 20(3):1542–1551 DOI [10.1021/acs.nanolett.9b04018](https://doi.org/10.1021/acs.nanolett.9b04018).
- Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, Jeon M, Ma'ayan A.** 2021a. Gene set knowledge discovery with enrichr. *Current Protocols* 1(3):e90 DOI [10.1002/cpz1.90](https://doi.org/10.1002/cpz1.90).
- Xie Y, Zhang W, Guo L, Kril LM, Begley KL, Sviripa VM, Chen X, Liu X, Lee EY, He D, Wang C, Gao T, Liu X, Evers BM, Watt DS, Liu C.** 2021b. Potent synergistic effect on C-Myc-driven colorectal cancers using a novel indole-substituted quinoline with a Plk1 inhibitor. *Molecular Cancer Therapeutics* 20(10):1893–1903 DOI [10.1158/1535-7163.MCT-20-1017](https://doi.org/10.1158/1535-7163.MCT-20-1017).
- Xu Y, Kong J, Hu P.** 2021. Computational drug repurposing for Alzheimer's disease using risk genes from GWAS and single-cell RNA sequencing studies. *Frontiers in Pharmacology* 12:617537 DOI [10.3389/fphar.2021.617537](https://doi.org/10.3389/fphar.2021.617537).
- Yang C, Huang X, Li Y, Chen J, Lv Y, Dai S.** 2021a. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: an in silico strategy towards precision oncology. *Briefings in Bioinformatics* 22(3):394 DOI [10.1093/bib/bbaa164](https://doi.org/10.1093/bib/bbaa164).
- Yang J, Li H, Wang F, Xiao F, Yan W, Hu G.** 2021b. Network-based target prioritization and drug candidate identification for multiple sclerosis: from analyzing “Omics Data” to druggability simulations. *ACS Chemical Neuroscience* 12(5):917–929 DOI [10.1021/acschemneuro.1c00011](https://doi.org/10.1021/acschemneuro.1c00011).
- Yang BY, Sakharakar MK.** 2022. Alterations in gene pair correlations as potential diagnostic markers for colon cancer. *International Journal of Molecular Sciences* 23(20):12463 DOI [10.3390/ijms232012463](https://doi.org/10.3390/ijms232012463).
- Yang X, Wei W, Tan S, Guo L, Qiao S, Yao B, Wang Z.** 2021c. Identification and verification of HCAR3 and INSL5 as new potential therapeutic targets of colorectal cancer. *World Journal of Surgical Oncology* 19(1):248 DOI [10.1186/s12957-021-02335-x](https://doi.org/10.1186/s12957-021-02335-x).
- Ye SB, Cheng YK, Deng R, Deng Y, Li P, Zhang L, Lan P.** 2020. The predictive value of estrogen receptor 1 on adjuvant chemotherapy in locally advanced colorectal cancer: a retrospective analysis with independent validation and its potential mechanism. *Frontiers in Oncology* 10:214 DOI [10.3389/fonc.2020.00214](https://doi.org/10.3389/fonc.2020.00214).
- Yu Z, Deng P, Chen Y, Liu S, Chen J, Yang Z, Chen J, Fan X, Wang P, Cai Z, Wang Y, Hu P, Lin D, Xiao R, Zou Y, Huang Y, Yu Q, Lan P, Tan J, Wu X.** 2021. Inhibition of the PLK1-coupled cell cycle machinery overcomes resistance to oxaliplatin in colorectal cancer. *Advanced Science* 8(23):e2100759 DOI [10.1002/advs.202100759](https://doi.org/10.1002/advs.202100759).

- Zali H, Golchin A, Farahani M, Yazdani M, Ranjbar MM, Dabbagh A.** 2019. FDA approved drugs repurposing of toll-like receptor4 (TLR4) candidate for neuropathy. *Iranian Journal of Pharmaceutical Research* **18**(3):1639–1647 DOI [10.22037/ijpr.2019.2394](https://doi.org/10.22037/ijpr.2019.2394).
- Zeng X, Zhu S, Lu W, Liu Z, Huang J, Zhou Y, Fang J, Huang Y, Guo H, Li L, Trapp BD, Nussinov R, Eng C, Loscalzo J, Cheng F.** 2020. Target identification among known drugs by deep learning from heterogeneous networks. *Chemical Science* **11**(7):1775–1797 DOI [10.1039/C9SC04336E](https://doi.org/10.1039/C9SC04336E).
- Zhang W, Wang H, Sun M, Deng X, Wu X, Ma Y, Li M, Shuo SM, You Q, Miao L.** 2020. CXCL5/CXCR2 axis in tumor microenvironment as potential diagnostic biomarker and therapeutic target. *Cancer Communications* **40**(2–3):69–80 DOI [10.1002/cac2.12010](https://doi.org/10.1002/cac2.12010).
- Zhang Y, Zheng S, Liao N, Huang H, Chen W, Wu Z, Wu D.** 2021. CircCTNNA1 acts as a ceRNA for miR-363-3p to facilitate the progression of colorectal cancer by promoting CXCL5 expression. *Journal of Biological Research-Thessaloniki* **28**(1):7 DOI [10.1186/s40709-021-00135-8](https://doi.org/10.1186/s40709-021-00135-8).
- Zhao X, Li L, Yuan S, Zhang Q, Jiang X, Luo T.** 2021. SPIB acts as a tumor suppressor by activating the NFkB and JNK signaling pathways through MAP4K1 in colorectal cancer cells. *Cellular Signalling* **88**(6):110148 DOI [10.1016/j.cellsig.2021.110148](https://doi.org/10.1016/j.cellsig.2021.110148).
- Zhong H, Chan G, Hu Y, Hu H, Ouyang D.** 2018. A comprehensive map of FDA-approved pharmaceutical products. *Pharmaceutics* **10**(4):263 DOI [10.3390/pharmaceutics10040263](https://doi.org/10.3390/pharmaceutics10040263).