scientific reports

Check for updates

OPEN

Exploring potential therapeutic combinations for castration-sensitive prostate cancer using supercomputers: a proof of concept study

Draško Tomić[2]✉, Jure Murgić[1], Ana Fröbe[1], Karolj Skala[2], Antonela Vrljičak[1], Branka Medved Rogina[2], Branimir Kolarek[2] & Viktor Bojović[2]

To address the challenge of finding new combination therapies against castration-sensitive prostate cancer, we introduce Vini, a computational tool that predicts the efficacy of drug combinations at the intracellular level by integrating data from the KEGG, DrugBank, Pubchem, Protein Data Bank, Uniprot, NCI-60 and COSMIC databases. Vini is a computational tool that predicts the efficacy of drugs and their combinations at the intracellular level. It addresses the problem comprehensively by considering all known target genes, proteins and small molecules and their mutual interactions involved in the onset and development of cancer. The results obtained point to new, previously unexplored combination therapies that could theoretically be promising candidates for the treatment of castration-sensitive prostate cancer and could prevent the inevitable progression of the cancer to the incurable castration-resistant stage. Furthermore, after analyzing the obtained triple combinations of drugs and their targets, the most common targets became clear: ALK, BCL-2, mTOR, DNA and androgen axis. These results may help to define future therapies against castration-sensitive prostate cancer. The use of the Vini computer model to explore therapeutic combinations represents an innovative approach in the search for effective treatments for castration-sensitive prostate cancer, which, if clinically validated, could potentially lead to new breakthrough therapies.

Prostate cancer is the most commonly diagnosed cancer and the second leading cause of cancer death in men in the United States. Worldwide, it is the second most common cancer and the fifth leading cause of cancer death in men[1]. Recently, the US Food and Drug Administration approved enzalutamide, a non-steroidal androgen receptor (AR) inhibitor, for the treatment of metastatic and non-metastatic castration-resistant prostate cancer and castration-sensitive metastatic prostate cancer at high risk of metastasis. The European Medicines Agency (EMA) recently also approved enzalutamide in combination with the PARP inhibitor talazoparib for patients with metastatic castration-resistant prostate cancer. Castration-resistant prostate cancer remains a clinical entity with high unmet medical need, and new therapies are needed that could overcome castration resistance. Combinatorial therapies may be more successful in both castration-sensitive and castration-resistant prostate cancer, as they induce profound androgen suppression and eradicate castration-sensitive tumor clones more thoroughly[2]. However, not only AR-targeted therapy can provide an optimal combination. There are also a number of clinical trials in which combination therapy with new innovative drug classes, including immunotherapy, play a key role and are in preparation or already in phase 1. Some examples are the immunocytokine M9241 in combination with docetaxel[3]; nivolumab in combination with BMS-986253, a fully human monoclonal antibody that inhibits interleukin-8[4]; neoantigen DNA vaccine in combination with nivolumab/ipilimumab and PROSTVAC, a vaccine designed to enable the immune system to recognize and attack prostate cancer cells[5]; cabozantinib and abiraterone with checkpoint inhibitor immunotherapy for metastatic castration-sensitive prostate cancer[6]. Recent data indicate that an increasing number of patients are being diagnosed with advanced stage prostate cancer. Prostate specific antigen (PSA) is a commonly used test for the detection of prostate cancer. Attention has focused primarily on the use of PSA in screening asymptomatic patients, but the diagnostic accuracy of PSA for

[1]Department of Oncology and Nuclear Medicine, Sisters of Charity Hospital, 10000 Zagreb, Croatia. [2]Centre for Informatics and Computing, Rudjer Boskovic Institute, 10000 Zagreb, Croatia. ✉email: drasko@irb.hr

prostate cancer in patients with symptoms is compromised by its low specificity, as many other benign conditions can cause PSA elevation[7]. The Gleason scoring system[8] is the most commonly used system for grading prostate cancer. To assess the degree of tumor differentiation, the most important prognostic factor in all cancers, the pathologist looks at how the cancer cells are arranged in the prostate and assigns a score on a scale of 3 to 5 at two different locations. Cancer cells that look similar to healthy cells are given a low score.

Despite the high success rate of the above-mentioned diagnostic procedures for prostate cancer, inaccuracies are possible, so that effective combination therapies are of great importance for a successful treatment outcome. The number of possible combinations of existing drugs is already large for combination therapies with 2 or 3 drugs, and clinical trials are expensive, time-consuming and often end in failure. Therefore, the theoretical model could help to create potentially more effective therapeutic combinations. It is therefore important to identify candidates for successful combination therapies already in the preclinical phase. In addition to the experience and knowledge of clinical and experimental oncologists, computer tools can also be of great help here. Computational modeling is proving to be a powerful tool for mastering the complexity of cancer biology and drug-drug interactions. There are numerous computer models that help in performing virtual drug screening[9], drug discovery[10] and drug repurposing[11]. Most of them are based on a single target and aim to find drug candidates that inhibit one or perhaps two targets. However, switching from a single-target to a multi-target approach can provide better and more accurate results and contribute to the development of new therapies[12]. While there are numerous computational methods for predicting drug combinations, such as those discussed in[13], there are not as many tools based specifically on multi-drug and metabolic pathway approaches, and the Vini in silico model for cancer is one of them. It has already proven that it can accurately calculate whether or not a particular drug is effective against a particular type of cancer. In a comprehensive study investigating how accurately Vini can predict whether a particular drug is effective or not, the model achieved 79.3% agreement with results from in vivo studies for 16 cancer types and 100% agreement with experimental data for the prostate cancer cell lines DU-145 and PC3[14]. In another study, the predictive performance of Vini was evaluated in assessing the efficacy of two-drug combinations against pancreatic ductal adenocarcinoma (PDAC), small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). The results showed good agreement between the predicted and in vivo studies, highlighting the model's ability to identify effective multidrug cancer therapies[15]. Vini has also demonstrated its versatility beyond the calculation of cancer therapies by being used in the early stages of the pandemic outbreak in the search for effective therapies against COVID-19[16]. This underlines its potential as a versatile tool for multidisciplinary drug discovery and optimization of therapies in different disease contexts.

Results

Testing the correlation between the calculated and experimental data

As a first step, we tested the accuracy of Vini in predicting the efficacy of 132 FDA-approved cancer drugs in kidney, prostate, NSCLC, colorectal, acute myeloid leukemia (AML), melanoma, breast, ovarian and brain tumors (glioma). Vini used GI50 values from the NCI-60 database as reference values. These are experimental values that indicate the concentration that causes 50% inhibition of cell division. Vini then calculated the efficacy of only those anticancer drugs for which GI50 data were available and the Pearson correlation between them and the calculated values. For this reason, the number of drugs tested varies by cancer cell line. Figure 1 shows the cancer cell lines, the type of cancer they represent, and the correlations between the calculated and GI50 values.

Vini predicts the relative efficacy of a drug by calculating the SLEM (second largest eigenvalue in magnitude) value of a graph whose nodes represent the free binding energies between a KEGG receptor and a particular drug and the edges represent the interactions between the receptors. The rationale for using SLEM as a value that describes how quickly a particular system transitions from one state to another can be found in[17], where a lower SLEM value indicates a faster transition from one state to another and vice versa. This implies that a higher SLEM value of a graph representing a certain cancer system under the influence of a drug indicates a slower progression of the cancer and better efficacy of the drug and vice versa. The negative PCC (Pearson's Correlation Coefficient) between SLEM and GI50 values is due to the fact that the SLEM values are positive, while the logarithms of the GI50 values used by Vini have a negative sign. One of the main reasons for the moderate PCC values is the limited accuracy of the Rosetta and Autodock Vina tools used by Vini to calculate the free binding energies. The use of MD (Molecular Dynamics) simulation tools such as NAMD and Amber could significantly increase the accuracy of the results, but this also means a significant increase in the computational resources required to run the simulations. To overcome this challenge, we plan to implement AI (artificial intelligence) in Vina in the future. One AI model will aim to increase the accuracy of the calculated efficacy of drugs and their combinations and will be trained on data from existing clinical trials. Another AI model will be developed with the aim of increasing the accuracy of the calculated free binding energies between mAb drugs and proteins. It will be trained using existing data on experimentally measured free binding energies that can be found in external databases, for example the PDBbind database.

Analysis of the effectiveness of triple combinations of drugs on hormone-sensitive prostate *cancer*

The combinations of three drugs were selected in several steps. The main objective was to identify the most effective combinations while avoiding possible adverse drug interactions. The input list of 70 small molecule anticancer drugs was compiled based on their known activity against prostate cancer and based on input from cancer physicians. It was then entered into the Vini model. Calculating the efficacy of all triple combinations would consume too many computing resources and exceed the capacity of our customized facilities. As a first step, Vini therefore calculated the efficacy of all individual drugs on the list. The 22 drugs with the highest efficacy were selected for further analysis, and then the efficacy of their combinations was calculated. Finally, Vini ranked

Cancer cell line	Cancer type – KEGG pathway	Pearson correlation coefficient	Number of drugs validated	T-Statistics	p-value	Adjusted p-value
A498	renal – hsa05211	-0,585630079	56	5,406576335	0,000001372	0,000012808
ACHN	renal – hsa05211	-0,411194780	58	3,435439329	0,001099423	0,001231354
CAKI-1	renal – hsa05211	-0,432449642	59	3,68400037	0,000500432	0,000636914
SN12C	renal – hsa05211	-0,391557384	59	3,268594545	0,001805351	0,001872216
PC-3	prostate – hsa05215	-0,604154739	57	5,724001778	0,000000405	0,000009062
DU-145	prostate – hsa05215	-0,539154145	57	3,498542828	0,000914741	0,001067198
HOP-62	NSCLC - hsa05223	-0,417441436	58	3,498542828	0,000905440	0,001067198
HOP-92	NSCLC - hsa05223	-0,446885390	58	3,804394302	0,000344363	0,000507483
NCI-H23	NSCLC - hsa05223	-0,405304651	59	3,405454487	0,001193964	0,001285808
EKVX	NSCLC - hsa05223	-0,545674513	58	4,959125431	0,000006513	0,000026053
HCC-2998	colon – hsa05210	-0,452858255	58	3,868251809	0,000279982	0,000435528
HCT-116	colon – hsa05210	-0,519125744	59	4,665368437	0,000018146	0,000050809
KM12	colon – hsa05210	-0,496525386	59	4,393763033	0,000047038	0,000109756
HT-29	colon – hsa05210	-0,466679348	59	4,053057818	0,000149978	0,000262462
K-562	AML – hsa05221	-0,511551955	58	4,534015683	0,000029510	0,000075116
LOXIMVI	melanoma – hsa05218	-0,525569761	59	4,745196034	0,000013657	0,000042487
SK-MEL-2	melanoma – hsa05218	-0,534829828	59	4,861895067	0,000008985	0,000031446
SK-MEL-5	melanoma – hsa05218	-0,434192548	59	3,702288934	0,000472070	0,000629426
M14	melanoma – hsa05218	-0,442489692	58	3,757805908	0,000400060	0,000560084
MCF-7	breast – hsa05224	-0,496082228	57	4,313536995	0,000064533	0,000138993
T47D	breast – hsa05224	-0,564210277	54	5,021712126	0,000005935	0,000026053
MDA-MB-468	breast – hsa05224	-0,406086787	41	2,845399364	0,006893634	0,006893634
OVCAR-5	ovarian – hsa05200	-0,558253567	59	5,168350067	0,000002943	0,000020598
OVCAR-8	ovarian – hsa05200	-0,486119766	59	4,272788776	0,000071339	0,000142678
SF-268	glioma – hsa05214	-0,471567599	59	4,107571534	0,000124939	0,000233220
SF-295	glioma – hsa05214	-0,543848024	59	4,977903891	0,000005904	0,000026053
SF-539	glioma – hsa05214	-0,455596947	59	3,931207493	0,000224656	0,000370022
SNB75	glioma – hsa05214	-0,587416052	59	5,575326025	0,000000647	0,000009062

-0.3 -0.4 -0.5 -0.6

Figure 1. Analysis of the correlation between SLEM values and GI50 values in NCI-60 cancer cell lines. The table contains the names of the cancer cell line names, the cancer types, the Pearson correlation coefficient (PCC) values, the T-statistics, the corresponding p-values, and the adjusted p-values. We applied the false discovery rate (FDR) method and used the Benjamini–Hochberg algorithm to calculate the adjusted p-values. The number of drugs analyzed varies by cell line, as for some of them no GI50 value is available in the NCI-60 database. The results show a range of PCC values from −0.391557384 to −0.604154379. All p-values are highly significant ($p < 0.05$), indicating that the observed inverse correlations are statistically significant. Furthermore, all adjusted p-values are greater than the p-values, except in the case of the cancer cell line MDA-MB-468, which is equal to the p-value. This robust evidence supports the reliability of SLEM values as meaningful indicators of drug efficacy in our virtual drug screening process. Further validation will be performed with the ALMANAC and DrugComb databases to test the accuracy of Vini in predicting the efficacy of double and triple combinations of cancer drugs.

the combinations according to their efficacy: those with the highest negative SLEM value at the top of the list and those with the lowest at the bottom. The 100 triple combinations with the highest efficacy and no adverse effects due to drug-drug interactions were then selected using Medscape's drug-drug interaction software[18]. To obtain an even higher level of safety, Drugs.com's drug interaction software[19] was applied to these 100 combinations, leaving 18 combinations. These combinations, together with their calculated SLEM values and the targets on which they act, are shown in Fig. 2.

Analysis of the targets that occur most frequently in the triple combinations of drugs obtained

The drugs involved in these combinations, the number of their occurrences in the combinations and the targets on which they act are listed in Fig. 3.

The following articles emphasize the importance of ALK[20–22], mTOR[23–25], DNA[26] and GnRH[27–29] in the development and progression of prostate cancer. However, the results of the studies on the efficacy of the mTOR inhibitor temsirolimus are partly contradictory. While the above studies confirmed the importance of mTOR

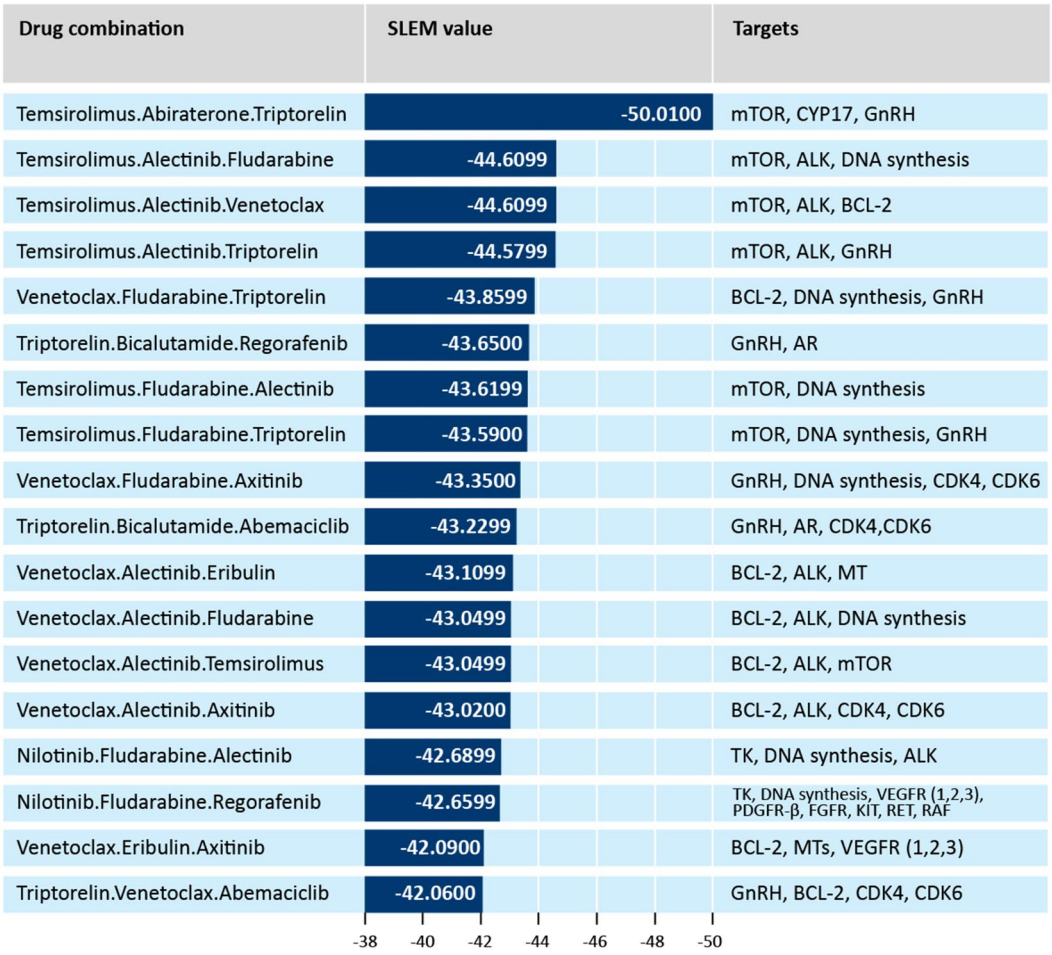

Figure 2. The final list of 3-drug combinations showing the highest efficacy in the DU-145 cell line with no drug-drug interactions. The abbreviations of the targets: mTOR—mammalian target of rapamycin; CYP1a7—17 α-hydroxylase/C17,20-lyase, GnRH- gonadotropin-releasing hormone; ALK—anaplastic lymphoma kinase, BCL-2—Proteins in the B-cell CLL/lymphoma 2 family; AR—androgen receptor; CDK4—cyclin-dependent kinases 4; CDK6—cyclin-dependent kinases 6; MT – microtubules, VEGFR-1—vascular endothelial growth factor receptor-1; VEGFR-2—vascular endothelial growth factor receptor-2; VEGFR-3—vascular endothelial growth factor receptor-3; PDGFR-β—growth factor receptor beta; FGFR—fibroblast growth factor receptors; KIT—KIT gene, receptor tyrosine kinase; RET—rearranged during transfection receptor tyrosine kinase; RAF—rapidly accelerated fibrosarcoma gene.

inhibition in slowing the progression of prostate cancer[30], found insufficient clinical activity of temsirolimus in men with mCRPC despite transient CTC improvements in some men. However, it was noted that future studies should focus on combination approaches or novel PI3K pathway inhibitors. In fact, Vini ranked temsirolimus 12th among the 70 drugs in terms of efficacy, while it is very dominant in triple combinations. This fact could further confirm the value of the multi-drug and multi-target approach used in this study.

Discussion

The fundamental contribution of this research is the identification of key targets involved in the onset and development of castration-sensitive prostate cancer and its possible transition to a castration-resistant form. These targets were identified by applying the Vini in silico cancer model to a list of existing anti-cancer drugs that may positively affect the treatment of this form of cancer. Among the most promising targets are the ALK pathway, DNA synthesis, BCL-2, mTOR and the androgen axis.

An additional contribution of the research is the identification of triple combinations of drugs that most effectively inhibit one or more of these targets without causing potential adverse effects due to drug-drug interactions. This research was conducted on a set of 70 small molecule cancer drugs using Vini on the HPC Vega supercomputer. The list of cancer drugs was compiled based on the expert knowledge of clinical oncologists and their opinions on the potential impact of these drugs. Exploring a larger number of drugs would potentially yield more relevant targets, requiring additional, more powerful computing resources, and the application for this is in progress.

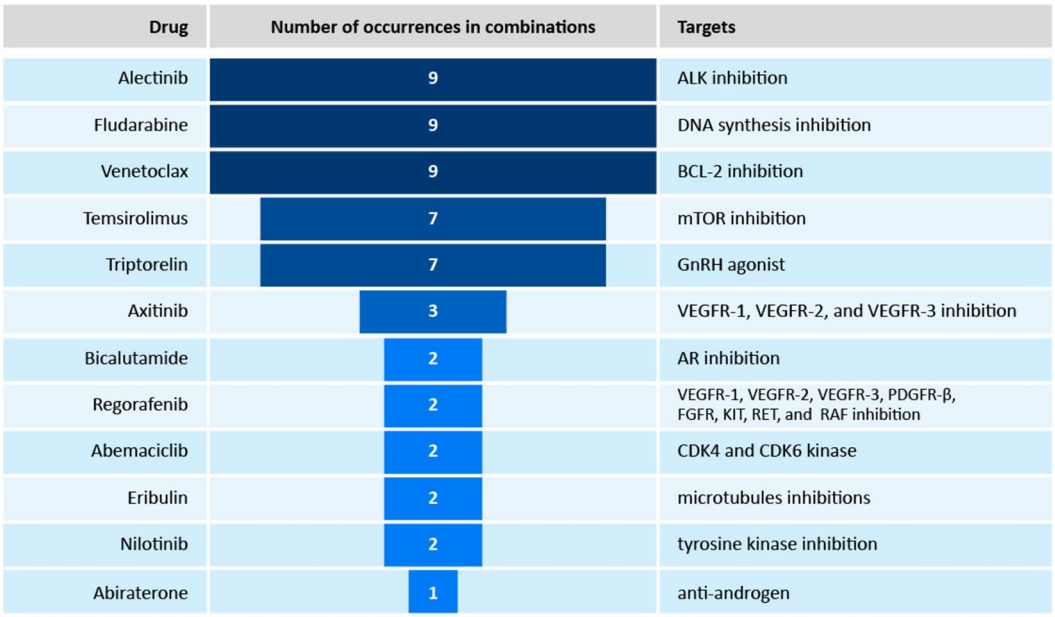

Figure 3. Individual drugs occurring in combinations, the total number of their occurrences and the targets on which they act are shown. The most common targets of the triple drug combinations obtained are ALK, DNA, BCL-2, mTOR and GnRH.

The good accuracy of the Vini model in determining drug efficacy has been confirmed by previous studies and further validated in this work by a correlation between the calculated values and the experimental GI50 values listed in the NCI-60 database. The accuracy of the Vini model is currently limited by the Autodock Vina and Rosetta computational chemistry tools used for the binding energy calculations. The future goal is to further increase the accuracy of the Vini model by incorporating AI (Artificial Intelligence) and molecular dynamics simulation tools such as NAMD[31] and Amber[32].

This study was performed using the castration-sensitive prostate cancer cell line DU-145, the KEGG hsa05215 cancer pathway and COSMIC gene expression and mutation data for this cell line. The mutations and gene expression in prostate cancer patients are partially different from those of the DU-145 line. Therefore, our future research will focus on the analysis of the genomic imprint of each individual patient and consequently on a personalized approach to prostate cancer therapy.

The role of the Vini cancer model, as with many other computer models of complex diseases, is to assist experimental oncologists and clinicians in selecting the best possible candidates for cancer therapy. Even with triplets of 70 drugs, the number of possible combinations is high, totalling 54,720 according to the formula for combinations:

$$\mathrm{nCr} = \mathrm{n!}/(\mathrm{r!} * (\mathrm{n} - \mathrm{r})!) = 70!/(3! * (70 - 3)!) = 54,720 \quad (1)$$

Therefore, effective and accurate screening is crucial when deciding which drugs to use in in vitro and in vivo research, potentially increasing the percentage of positive results from clinical trials.

Monoclonal antibody therapy, either in combination with small molecule drugs[33] or as antibody–drug conjugates[34], is a promising approach that could further increase the efficacy of prostate cancer therapies. Accordingly, our future research will focus on combined therapies against castration-sensitive prostate cancer that include mAb drugs and conjugates.

Notes for clinical consideration

Targeting castration-sensitive prostate cancer, a disease of significant clinical importance, represents a potential improvement in the design of drug trials.

Clinical applicability may be useful to broaden the discussion of how proposed therapeutic combinations could be translated into clinical practice, including potential challenges and considerations for implementation.

Limitations

Based on our previous experience with the model and the clinical research approach, we recognize the potential limitations:

a. *Dependence on high performance computing resources*: the study makes extensive use of the HPC Vega supercomputer, suggesting that the analysis of drug combinations is highly dependent on high computing power. This dependence may limit the reproducibility of the study in environments where such resources are not readily available.

b. *Use of a specific cell line*: The study focuses on the castration-sensitive prostate cancer cell line DU-145. The genetic and expression profiles of prostate cancer may differ significantly from patient to patient. Therefore, the results of the study may not be directly applicable to all prostate cancer cases, emphasizing the need for further validation in different genetic backgrounds, cell lines and preferably in clinical trials.
c. *Limitations of the prediction model*: Although the accuracy of the Vini model is good, it is still limited by the computer programs used for the binding energy calculations, such as Autodock Vina and Rosetta. While there are plans to improve the accuracy of Vini using AI, quantum algorithms, and molecular dynamics simulation tools, the current accuracy could influence the study results. However, Vini's independence from extensive training datasets is advantageous in scenarios where data is sparse or unavailable, allowing for quicker implementation in emerging drug discovery projects.
d. *Scope of drug combinations*: The study is limited to a set of 70 small molecule anticancer drugs, suggesting that the exploration of potential therapeutic combinations is not exhaustive. While the study acknowledges that a larger number of drugs need to be explored, the initial selection may exclude viable therapeutic options that were not considered as part of the study.
e. *Generalizability of the results*: The results of the study are based on silico models and require clinical validation. The efficacy of the identified drug combinations in the clinical setting remains uncertain as long as it has not been validated by clinical trials. This gap between computational predictions and clinical applicability is a critical limitation that needs to be addressed to ensure the practical relevance of the research. Despite the varying Pearson correlations reported in different ML methods, the consistent performance of Vini across 28 cancer cell lines highlights its utility in providing mechanistic insights and prioritizing candidates for further study. Our research was mainly based on the data from COSMIC and NCI-60 databases. In future studies, we will include data from The Cancer Cell Line Encyclopedia. This will allow us to cross-check our results and possibly increase the robustness and generalizability of the Vini model to a larger number of cancer cell lines.
f. We recognize the importance of detailed dosage and toxicity studies, and our future research will include comprehensive experimental validation of the proposed combinations. This will involve testing the efficacy and safety of these combinations in relevant preclinical models to confirm their synergistic or additive effects and to optimize their dosages for maximal therapeutic benefit.

These inferred limitations highlight the challenges of relying on computational models for drug discovery and the importance of validating findings through clinical trials and broader drug selection.

Materials and methods

Vini collects the necessary data from external databases: structures of small molecule drugs and FASTA sequences of monoclonal antibodies from DrugBank[35], structures of compounds from PubChem[36], gene expressions and mutations from COSMIC[37], KEGG protein structures from Protein Data Bank[38], their FASTA sequences from UniProt[39], GI50 data from NCI-60[40] and disease pathways from KEGG[41]. A high-level representation of how the Vini model works can be found in Fig. 4.

For data processing, Vini uses various computational chemistry tools, including Autodock Vina[42], Rosetta[43], AlphaFold[44], Open Babel[45] and Reduce[46]. A more detailed scheme of the Vini model and how it works can be found in Fig. 5.

When someone starts the Vini model for the first time, the specific KEGG pathway and the NCI-60 line must be specified. In addition, a list of ligands, i.e. drugs and/or chemical compounds (ligands) to be analyzed,

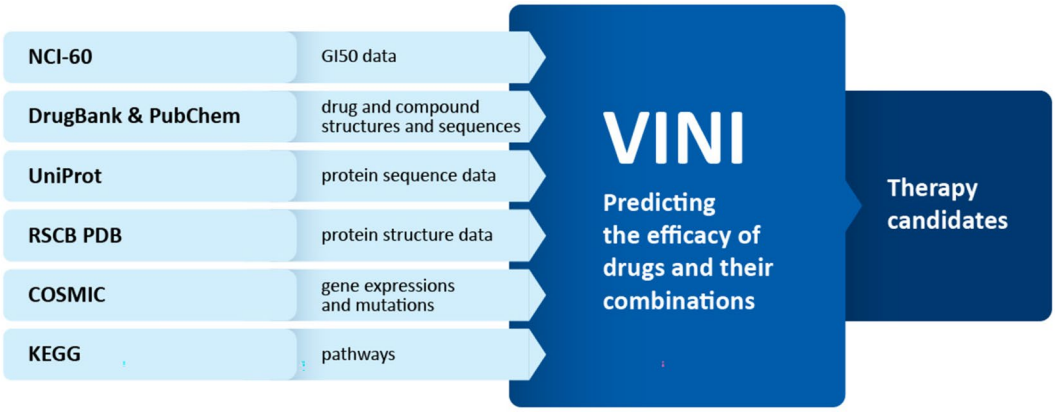

Figure 4. The overall schematic of the Vini in silico model for cancer. Vini collects the data from the external databases: KEGG (in this case the hsa05215 prostate cancer pathway), small molecule drug structures and monoclonal antibody (mAb) FASTA sequences from DrugBank, KEGG protein structures and their FASTA sequences from the RCSB PDB and UniProt databases, and gene expressions and mutations from the COSMIC database. It then calculates the efficacy of the drugs and their combinations. Finally, it sorts the drug list according to their calculated SLEM values.

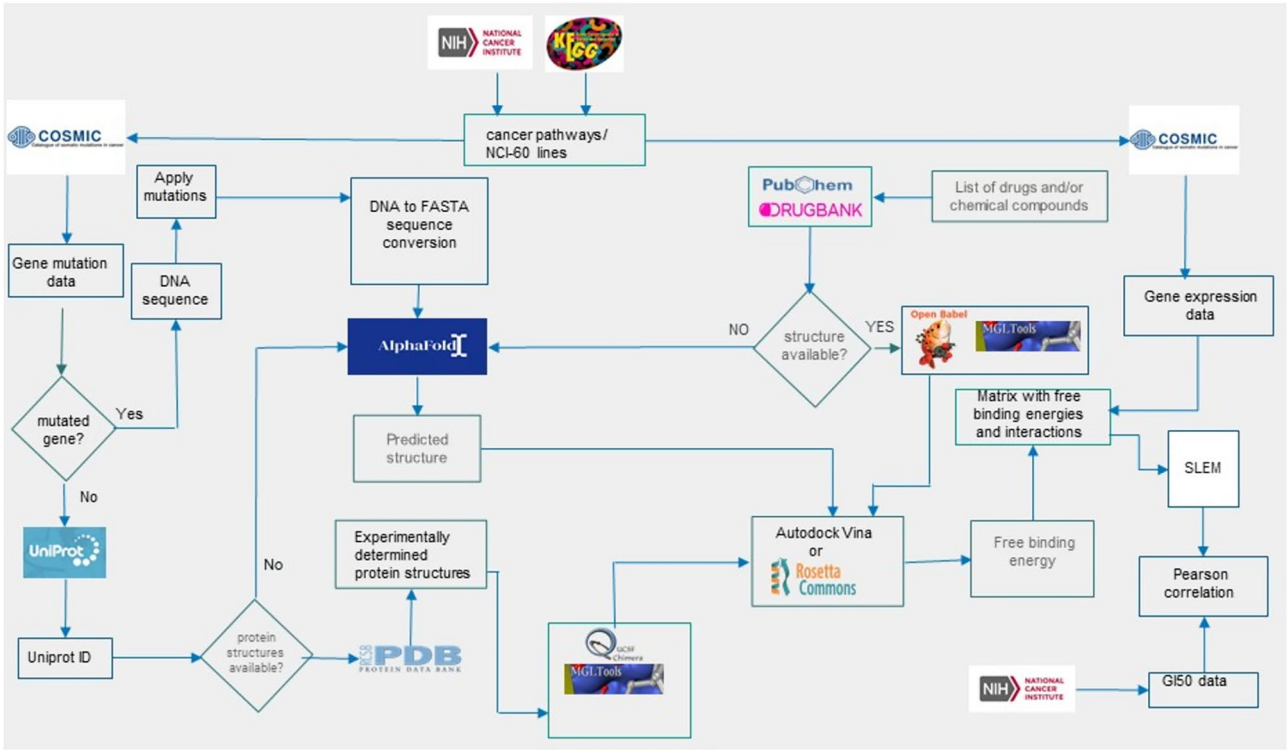

Figure 5. The detailed schematic of how Vini collects and processes data from external databases and creates SLEM. The input parameters are the specific KEGG cancer pathway, the list of cancer drugs and/or compounds to be analysed and the name of the NCI-60 cell line to be used in the simulation. 3D structures of small molecule drugs and FASTA sequences of mAb drugs are collected from DrugBank, while the structures of chemical compounds come from PubChem. Based on the name of the NCI-60 cancer cell line, Vini retrieves gene expression and mutation data from COSMIC. If the gene is not mutated, Vini retrieves the Uniprot ID of the transcribed protein and its 3D structure from the Protein Data Bank and prepares it for further simulation with UCSF Chimera. If no structure is available, it is predicted with AlphaFold. In the case of a mutated gene, its DNA sequence is pulled from COSMIC, mutations are applied to it, the mutated DNA sequence is converted to a FASTA sequence and the structure is predicted with AlphaFold. Docking is then performed between the receptor and the ligand (with Autodock Vina if the ligand is a small molecule or with Rosetta if it is a mAb) and the free binding energy is calculated. The final result is normalized with the value of the gene expression factor from COSMIC. This process is repeated for all receptors specified in the KEGG pathway, a matrix with the free binding energies and interactions is created and SLEM is calculated. This process is repeated for all drugs and compounds on the input list. If the efficacy of combined therapies is calculated, the procedure is repeated for all combinations. Finally, the output list of SLEM values is sorted in descending order of calculated efficacy, from highest to lowest.

must be specified. The next step is to select whether Vini should analyze only their efficacy or also the efficacy of their double or triple combinations. Based on the specified KEGG pathway, Vini creates a list of the receptors found in the KEGG pathway and a list of the interactions between them. The next step is to prepare the ligands. If a particular ligand is a drug, Vini checks whether it is a mAb drug or a small molecule. In the first case, Vini retrieves its FASTA sequence from DrugBank and predicts the 3D structure with AlphaFold. Otherwise, it checks whether it is a drug or a chemical compound. If it is a drug, it retrieves the SDF (Structure Data Format) file from the DrugBank and converts it to PDB (Protein Data Bank) format using Open Babel. If it is a chemical compound, Vini contacts PubChem and asks for the SDF file. If the SDF file is available, it is converted to PDB format using the Open Babel program. Otherwise, the SMILES (Simplified molecular-input line-entry system) file is retrieved and the option is offered to convert it to PDB format using the online SMILES Translator and Structure File Generator[47]. The final step in the preparation of a small molecule drug or chemical compound is the conversion of the PDB file into the PDBQT format that Autodock Vina works with, and for this purpose Vini uses the MGLTools program[48]. This process is repeated for each ligand in the input list. Vini also performs the processing of KEGG receptors in several steps. If a gene specified in the KEGG pathway for a particular cell line is mutated in COSMIC, Vini retrieves its DNA sequence from COSMIC, translates it into a FASTA sequence using the translation tool from the Expasy portal[49], and generates its 3D structure using AlphaFold. Otherwise, it contacts the Protein Data Bank and searches for the structure either without ligands or with at most one ligand. If such a structure is available, it uses UCSF Chimera[50] to clean it from ligands and water molecules and add missing residues from the Dunbrack library[51]. If the structure is not found, Vini creates the PDB structure from the FASTA sequence in Uniprot and this process is repeated for all receptors annotated in the KEGG pathway.

The next step is to dock a single ligand to all receptors specified in the KEGG pathway and calculate the free binding energies. A matrix is then formed with the elements on the main diagonal representing these energies and the elements outside the main diagonal representing the interaction between the receptors. Finally, the SLEM value of this matrix is calculated. The procedure is repeated for all ligands from the input list and for all combinations (double or triple) if they were specified for the analysis. The GI50 values of the drug combinations are not reported by the NCI and therefore Vini cannot determine the synergistic or additive effect of the calculated combinations. Various other computational methods, for example, based on the similarity comparison between the queried and known combinations[52–54], physicochemical characteristics of drugs and network features[55,56], gene expression data[57] and networks[58–60], have been proposed. However, most of them, except those in[61] which is the combination of transcriptomic and network approach applied to the specific type of cancer under analysis, were built based on data from multiple diseases without considering context specificity. Vini also calculates the efficacy of individual drugs and their combinations using a combined method based on the KEGG network, which describes the metabolic process of a given disease and defines the genes that participate in its onset and development, as well as the expression of these genes, i.e. their transcription and translation into proteins. The validity of such an approach in calculating the efficacy of combination therapies was confirmed when Vini predicted that the most effective triple combinations of therapies against castration-resistant prostate cancer are those that most frequently act on ALK, DNA, BCL-2, mTOR and the androgen axis. This study made extensive use of databases such as KEGG, DrugBank, Pubchem, Uniprot, NCI-60 and COSMIC, which indicates a thorough approach to data collection and utilization and enhances the credibility of the results. The study was conducted with a set of 70 small molecule anticancer drugs on the HPC Vega supercomputer at IZUM in Maribor, Slovenia. All results of this study were obtained using the in silico cancer model Vini, Medscape and Drugs.com drug interaction screening software. Vini is an open-source application and is available on Github[62], while the Medscape and Drugs.com interaction screening programs are also freely available and can be used. The Vini model requires a Linux operating system and a SLURM job scheduler installed. The efficacy results of triple-drug combinations were determined using the KEGG hsa05215 prostate cancer pathway and COSMIC data on gene expression and mutations in the DU-145 cell line. 28 cancer cell lines and 9 KEGG cancer pathways were used to calculate the Pearson correlation. Figure 6 illustrates how the 18 best triple combinations of drugs that have no unfavorable side effects due to drug interactions were identified.

Data use and access: We confirm that the molecular structures, genomic and transcriptomic data and GI50 data used in this study were obtained from publicly available sources, in particular the NCI (National Cancer Institute), KEGG, RCSB, Uniprot and COSMIc. We adhere to the conditions set by the data providers to ensure responsible and lawful use.

No experiments on living organisms: We assure that no experiments on living organisms were conducted as part of this research. The study uses only existing data sets and does not involve direct manipulation or experimentation on biological samples.

No use of patient data: We confirm that no patient data was used in this study. The research focuses exclusively on publicly available GI50 data for NCI-60 cancer cell lines and does not include access to personal or identifiable patient data.

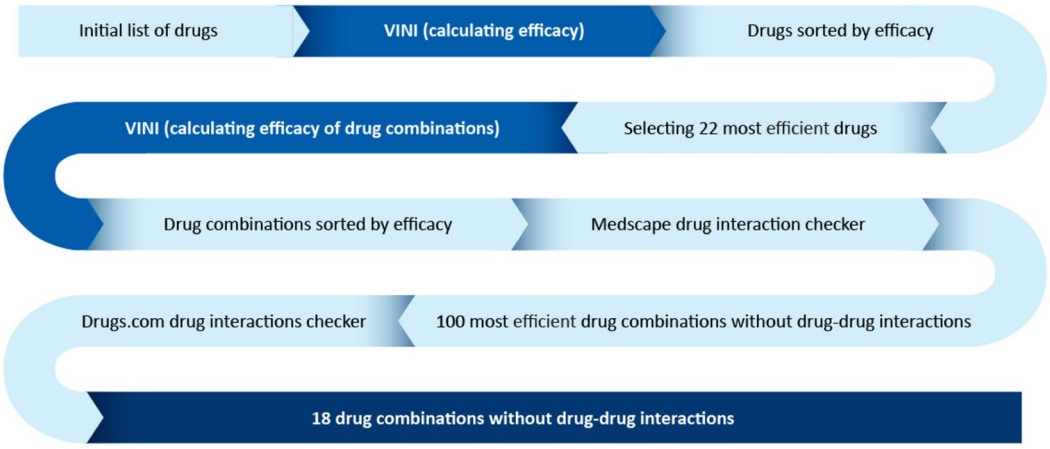

Figure 6. The Vini in silico model for cancer and the Drug Interaction Checker software from Medscape and Drugs.com were used to obtain the combinations with the highest efficacy without the risk of unwanted interactions between drugs within the combinations. Drugs.com's drug interaction checker was found to be more restrictive than Medscape's. Of the 100 combinations that Medscape declared interaction-free, only 18 were classified as interaction-free by Drugs.com. The remaining drug combinations were found to have either major and minor interactions, replication (two or more drugs acting on the same target, increasing the possibility of side effects), or a combination of the above.

Conclusion

The article describes a study in which the Vini computer model was used to identify effective drug combinations for the treatment of castration-sensitive prostate cancer. By analyzing data from various databases, the study finds promising triple-drug combinations targeting ALK, BCL-2, mTOR, DNA and the androgen axis to prevent cancer progression to a castration-resistant form. Research supported by high-performance computing demonstrates the potential of computer models in predicting effective cancer therapies and suggests a future focus on personalized treatments and exploring new drug combinations. Ideally, these findings should be validated on different cell lines and eventually in clinical trials.

While acknowledging varying Pearson correlations reported for different ML methods, the study emphasizes Vini's unique strengths in providing detailed mechanistic insights into drug-target interactions and prioritizing candidates for further investigation. Vini's independence from extensive training datasets enhances its versatility and speed of implementation, particularly in drug discovery projects where new targets and compounds frequently emerge. Our ongoing integration of ML and quantum computing algorithms aims to further improve Vini's predictive performance and extend its applicability across diverse cancer types and patient populations.

Data availability

The list of drugs used in this study, the calculated effectiveness of single, double and triple combination therapies in the DU-145 line and the values of the Pearson correlation coefficients calculated for various NCI-60 lines can be found at the link https://data.fulir.irb.hr/islandora/object/irb%3A375.

Received: 10 March 2024; Accepted: 9 August 2024
Published online: 13 August 2024

References

1. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin.* **72**(1), 7–33. https://doi.org/10.3322/caac.21708 (2022) (**Epub 2022 Jan 12**).
2. Posdzich, P. *et al.* Metastatic prostate cancer-a review of current treatment options and promising new approaches. *Cancers (Basel).* **15**(2), 461. https://doi.org/10.3390/cancers15020461 (2023).
3. Atiq, M. O. *et al.* Combining IL-12 immunocytokine (M9241) with docetaxel in metastatic prostate cancer: A phase I study. *J. Clin. Oncol.* **40**(16), 55. https://doi.org/10.1200/JCO.2022.40.16_suppl.e17033
4. Ignacio, B. M. *et al.* Phase 1b/2 study of nivolumab in combination with an anti–IL-8 monoclonal antibody, BMS-986253, in a biomarker-enriched population of patients with advanced cancer. *J. Clin. Oncol.* **36**(15_suppl). https://doi.org/10.1200/JCO.2018.36.15_suppl.TPS3109
5. Koral, S. *et al.* A pilot trial of neoantigen DNA vaccine in combination with nivolumab/ipilimumab and prostvac in metastatic hormone-sensitive prostate cancer (mHSPC). *J. Clin. Oncol.* **39**(6_suppl). https://doi.org/10.1200/JCO.2021.39.6_suppl.TPS192
6. Melissa, A. R. *et al.* A phase 1b clinical trial of cabozantinib (CABO) and abiraterone (ABI) with checkpoint inhibitor immunotherapy (CPI) in metastatic hormone-sensitive prostate cancer (mHSPC) (CABIOS Trial). *J. Clin. Oncol.* **40**(6_suppl). https://doi.org/10.1200/JCO.2022.40.6_suppl.TPS214
7. Merriel, S. W. D. *et al.* Systematic review and meta-analysis of the diagnostic accuracy of prostate-specific antigen (PSA) for the detection of prostate cancer in symptomatic patients. *BMC Med.* **20**(1), 54. https://doi.org/10.1186/s12916-021-02230-y (2022).
8. Chen, N. & Zhou, Q. The evolving Gleason grading system. *Chin J Cancer Res.* **28**(1), 58–64. https://doi.org/10.3978/j.issn.1000-9604.2016.02.04 (2016).
9. Lavecchia, A. & Di Giovanni, C. Virtual screening strategies in drug discovery: A critical review. *Curr Med Chem.* **20**(23), 2839–2860. https://doi.org/10.2174/09298673113209990001 (2013).
10. Sliwoski, G., Kothiwale, S., Meiler, J. & Lowe, E. W. Jr. Computational methods in drug discovery. *Pharmacol Rev.* **66**(1), 334–395. https://doi.org/10.1124/pr.112.007336 (2013).
11. Park, K. A review of computational drug repurposing. *Transl Clin Pharmacol.* **27**(2), 59–63. https://doi.org/10.12793/tcp.2019.27.2.59 (2019).
12. Medina-Franco, J. L., Giulianotti, M. A., Welmaker, G. S. & Houghten, R. A. Shifting from the single to the multitarget paradigm in drug discovery. *Drug Discov Today.* **18**(9–10), 495–501. https://doi.org/10.1016/j.drudis.2013.01.008 (2013May) (**Epub 2013 Jan 20**).
13. Kong, W. *et al.* Systematic review of computational methods for drug combination prediction. *Comput Struct Biotechnol J.* **1**(20), 2807–2814. https://doi.org/10.1016/j.csbj.2022.05.055 (2022).
14. Tomic, D. *et al.* Evaluation of the efficacy of cancer drugs by using the second largest eigenvalue of metabolic cancer pathways. *J Comput Sci Syst Biol* **11**, 4. https://doi.org/10.4172/jcsb.1000280 (2018).
15. Tomic, D., Pirkic, B., Skala, K., Kranjcevic, L. Predicting the effectiveness of multi-drug cancer therapies. In *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, 2019, pp. 375–380, https://doi.org/10.23919/MIPRO.2019.8757131.
16. Tomic, D. *et al.* The screening and evaluation of potential clinically significant HIV drug combinations against the SARS-CoV-2 virus. *Inform Med Unlocked.* **23**, 100529. https://doi.org/10.1137/S0036144503423264 (2021).
17. Boyd, S., Diaconis, P., Xiao, L. Fastest mixing markov chain on a graph. *SIAM Rev.* **46**(4) (2004).
18. https://reference.medscape.com/drug-interactionchecker
19. https://www.drugs.com/drug_interactions.html
20. Patel, R. A. *et al.* Comprehensive assessment of anaplastic lymphoma kinase in localized and metastatic prostate cancer reveals targetable alterations. *Cancer Res Commun.* **2**(5), 277–285. https://doi.org/10.1158/2767-9764.crc-21-0156 (2022).
21. Unno, K. *et al.* Activated ALK cooperates with N-myc via Wnt/β-catenin signaling to induce neuroendocrine prostate cancer. *Cancer Res.* **81**(8), 2157–2170. https://doi.org/10.1158/0008-5472.CAN-20-3351 (2021).
22. Carneiro, B. A. *et al.* Anaplastic lymphoma kinase mutation (ALK F1174C) in small cell carcinoma of the prostate and molecular response to alectinib. *Clin Cancer Res.* **24**(12), 2732–2739. https://doi.org/10.1158/1078-0432.CCR-18-0332 (2018).
23. Morgan, T. M., Koreckij, T. D. & Corey, E. Targeted therapy for advanced prostate cancer: inhibition of the PI3K/Akt/mTOR pathway. *Curr Cancer Drug Targets.* **9**(2), 237–249. https://doi.org/10.2174/156800909787580999 (2009).
24. Roudsari, N. M. *et al.* Inhibitors of the PI3K/Akt/mTOR pathway in prostate cancer chemoprevention and intervention. *Pharmaceutics.* **13**(8), 1195. https://doi.org/10.3390/pharmaceutics13081195 (2021).
25. Statz, C. M., Patterson, S. E. & Mockus, S. M. mTOR inhibitors in castration-resistant prostate cancer: A systematic review. *Target Oncol.* **12**(1), 47–59. https://doi.org/10.1007/s11523-016-0453-6 (2017).

26. Zhang, W., van Gent, D. C., Incrocci, L., van Weerden, W. M. & Nonnekens, J. Role of the DNA damage response in prostate cancer formation, progression and treatment. *Prostate Cancer Prostatic Dis.* **23**(1), 24–37. https://doi.org/10.1038/s41391-019-0153-2 (2020).
27. Cook, T. & Sheridan, W. P. Development of GnRH antagonists for prostate cancer: New approaches to treatment. *Oncologist.* **5**(2), 162–168. https://doi.org/10.1634/theoncologist.5-2-162 (2000).
28. Labrie, F. GnRH agonists and the rapidly increasing use of combined androgen blockade in prostate cancer. *Endocr Relat Cancer.* **21**(4), R301–R317. https://doi.org/10.1530/ERC-13-0165 (2014).
29. Maiti, K. *et al.* Differential effects of gonadotropin-releasing hormone (GnRH)-I and GnRH-II on prostate cancer cell signaling and death. *J Clin Endocrinol Metab.* **90**(7), 4287–4298. https://doi.org/10.1210/jc.2004-1894 (2005).
30. Armstrong, A. J. *et al.* A phase II trial of temsirolimus in men with castration-resistant metastatic prostate cancer. *Clin Genitourin Cancer.* **11**(4), 397–406. https://doi.org/10.1016/j.clgc.2013.05.007 (2013) (**Epub 2013 Jul 3**).
31. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem.* **26**(16), 1781–1802. https://doi.org/10.1002/jcc.20289 (2005).
32. Case, D. A. *et al.* AmberTools. *J. Chem. Inf. Model.* **63**(20), 6183–6191 (2023)
33. Jakobovits, A. Monoclonal antibody therapy for prostate cancer. *Handb Exp Pharmacol.* **181**, 237–256. https://doi.org/10.1007/978-3-540-73259-4_11 (2008).
34. Fu, Z., Li, S., Han, S., Shi, C. & Zhang, Y. Antibody drug conjugate: the "biological missile" for targeted cancer therapy. *Signal Transduct Target Ther.* **7**(1), 93. https://doi.org/10.1038/s41392-022-00947-7 (2022).
35. Knox, C., Wilson, M., Klinger, C. M., Franklin, M., Oler, E., Wilson, A., Pon, A., Cox, J., Chin, N. E. L., Strawbridge, S. A., Garcia-Patino, M., Kruger, R., Sivakumaran, A., Sanford, S., Doshi, R., Khetarpal, N., Fatokun, O., Doucet, D., Zubkowski, A., Rayat, D. Y., Jackson, H., Harford, K., Anjum, A., Zakir, M., Wang, F., Tian, S., Lee, B., Liigand, J., Peters, H., Wang, R. Q. R., Nguyen, T., So, D., Sharp, M., da Silva, R., Gabriel, C., Scantlebury, J., Jasinski, M., Ackerman, D., Jewison, T., Sajed, T., Gautam, V., Wishart, D. S. DrugBank 6.0: The DrugBank knowledgebase for 2024. *Nucleic Acids Res.* 2024;**52**(D1), D1265–D1275. https://doi.org/10.1093/nar/gkad976.
36. Kim, S. *et al.* PubChem 2023 update. *Nucleic Acids Res.* **51**(D1), D1373–D1380. https://doi.org/10.1093/nar/gkac956 (2023).
37. Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., & Forbes, S. A. COSMIC: The Catalogue of somatic mutations in cancer. *Nucleic Acids Res.*, **47**(D1), D941–D947.
38. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat Struct Biol.* **10**(12), 980. https://doi.org/10.1038/nsb1203-980 (2003).
39. The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**(D1), D523–D531. https://doi.org/10.1093/nar/gkac1052 (2023).
40. Shoemaker, R. H. The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer.* **6**(10), 813–823. https://doi.org/10.1038/nrc1951 (2006).
41. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30. https://doi.org/10.1093/nar/28.1.27 (2000).
42. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* **31**(2), 455–461. https://doi.org/10.1002/jcc.21334 (2010).
43. Chaudhury, S. & Gray, J. J. Conformer selection and induced fit in flexible backbone protein–protein docking using computational and NMR ensembles. *J Mol Biol.* **381**(4), 1068–1087. https://doi.org/10.1016/j.jmb.2008.05.042 (2008).
44. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature.* **596**(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2 (2021).
45. O'Boyle, N. M. *et al.* Open babel: An open chemical toolbox. *J Cheminform.* **7**(3), 33. https://doi.org/10.1186/1758-2946-3-33 (2011).
46. Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol.* **285**(4), 1735–1747. https://doi.org/10.1006/jmbi.1998.2401 (1999).
47. https://cactus.nci.nih.gov/translate/
48. Sanner, M. F. Python: A programming language for software integration and development. *J. Mol. Graphics Model.* **17**(1), 57–61 (1999).
49. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**(13), 3784–3788. https://doi.org/10.1093/nar/gkg563 (2003).
50. Pettersen, E. F. *et al.* UCSF Chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**(13), 1605–1612. https://doi.org/10.1002/jcc.20084 (2004).
51. Dunbrack, R. L. Jr. & Cohen, F. E. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci.* **6**(8), 1661–1681. https://doi.org/10.1002/pro.5560060807.PMID:9260279;PMCID:PMC2143774 (1997).
52. Li, P. *et al.* Large-scale exploration and analysis of drug combinations. *Bioinformatics (Oxford, England)* **31**(12), 2007–2016. https://doi.org/10.1093/bioinformatics/btv080 (2015).
53. Li, S. *et al.* Prediction of synergistic drug combinations for prostate cancer by transcriptomic and network characteristics. *Front. Pharmacol.* **12**, 634097. https://doi.org/10.3389/fphar.2021.634097 (2021).
54. Chen, X. *et al.* NLLSS: Predicting synergistic drug combinations based on semi-supervised learning. *PLoS Comput. Biol.* **12**(7), e1004975. https://doi.org/10.1371/journal.pcbi.1004975 (2016).
55. Zhao, X. M. *et al.* Prediction of drug combinations by integrating molecular and pharmacological data. *PLoS Comput. Biol.* **7**(12), e1002323. https://doi.org/10.1371/journal.pcbi.1002323 (2011).
56. Li, X. *et al.* Prediction of synergistic anti-cancer drug combinations based on drug target network and drug induced gene expression profiles. *Artif. Intell. Med.* **83**, 35–43. https://doi.org/10.1016/j.artmed.2017.05.008 (2017).
57. Stathias, V. *et al.* Drug and disease signature integration identifies synergistic combinations in glioblastoma. *Nature Commun.* **9**(1), 5315. https://doi.org/10.1038/s41467-018-07659-z (2018).
58. Wu, Z., Li, W., Liu, G. & Tang, Y. Network-based methods for prediction of drug-target interactions. *Front. Pharmacol.* **9**, 1134. https://doi.org/10.3389/fphar.2018.01134 (2018).
59. Cheng, F., Kovács, I. A. & Barabási, A. L. Network-based prediction of drug combinations. *Nature Commun.* **10**(1), 1197. https://doi.org/10.1038/s41467-019-09186-x (2019).
60. Zhou, Y. *et al.* Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery* **6**, 14. https://doi.org/10.1038/s41421-020-0153-3 (2020).
61. Li, S. *et al.* Prediction of synergistic drug combinations for prostate cancer by transcriptomic and network characteristics. *Front Pharmacol.* **12**(12), 634097. https://doi.org/10.3389/fphar.2021.634097 (2021).
62. https://github.com/draskot/Vini

Acknowledgements

We acknowledge EuroHPC Joint Undertaking for awarding the project ID REG-2022R03-224 for access to the HPC Vega, hosted at IZUM in Maribor, Slovenia. The Slovenian National Supercomputing Network (SLING)

and the HPC Vega staff at IZUM in Maribor, Slovenia, for their excellent support in successfully performing all necessary computations.

Author contributions

D.T. performed study design, writing, data analysis, interpretation, literature review, and computer simulations. J.M. participated in writing, data analysis, data interpretation, and literature review. K.S. participated in writing. A.F., A.V., and B.M.R. participated in the design of the study. B.K. prepared figures. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.T.

Reprints and permissions information is available at www.nature.com/reprints.

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