frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 December 2022
DOI 10.3389/fgene.2022.1015531

OPEN ACCESS

EDITED BY
Rongling Wu,
The Pennsylvania State University (PSU), United States

REVIEWED BY
Shyam Nyati,
Henry Ford Health System, United States
Carlo Cesar Cortes,
National Cancer Institute, Mexico

*CORRESPONDENCE
Suzan Arslanturk,
suzan.arslanturk@wayne.edu

SPECIALTY SECTION
This article was submitted to Statistical Genetics and Methodology, a section of the journal Frontiers in Genetics

RECEIVED 09 August 2022
ACCEPTED 15 November 2022
PUBLISHED 13 December 2022

CITATION
Munj SA, Taz TA, Arslanturk S and Heath EI (2022), Biomarker-driven drug repurposing on biologically similar cancers with DNA-repair deficiencies. *Front. Genet.* 13:1015531. doi: 10.3389/fgene.2022.1015531

Biomarker-driven drug repurposing on biologically similar cancers with DNA-repair deficiencies

Seeya Awadhut Munj[1], Tasnimul Alam Taz[1], Suzan Arslanturk[1]* and Elisabeth I. Heath[2,3]

[1]Department of Computer Science, Wayne State University, Detroit, MI, United States, [2]Department of Oncology, Wayne State University, Detroit, MI, United States, [3]Molecular Therapeutics Program, Barbara Ann Karmanos Cancer Institute, Detroit, MI, United States

Similar molecular and genetic aberrations among diseases can lead to the discovery of jointly important treatment options across biologically similar diseases. Oncologists closely looked at several hormone-dependent cancers and identified remarkable pathological and molecular similarities in their DNA repair pathway abnormalities. Although deficiencies in Homologous Recombination (HR) pathway plays a significant role towards cancer progression, there could be other DNA-repair pathway deficiencies that requires careful investigation. In this paper, through a biomarker-driven drug repurposing model, we identified several potential drug candidates for breast and prostate cancer patients with DNA-repair deficiencies based on common specific biomarkers and irrespective of the organ the tumors originated from. Normalized discounted cumulative gain (NDCG) and sensitivity analysis were used to assess the performance of the drug repurposing model. Our results showed that Mitoxantrone and Genistein were among drugs with high therapeutic effects that significantly reverted the gene expression changes caused by the disease (FDR adjusted p-values for prostate cancer =1.225e-4 and 8.195e-8, respectively) for patients with deficiencies in their homologous recombination (HR) pathways. The proposed multi-cancer treatment framework, suitable for patients whose cancers had common specific biomarkers, has the potential to identify promising drug candidates by enriching the study population through the integration of multiple cancers and targeting patients who respond poorly to organ-specific treatments.

KEYWORDS

drug repurposing, DNA repair, personalized medicine, multi cancer treatment, mitoxantrone, homologous recombination

1 Introduction

Developing a new drug for a condition can take around 10–13 years and close to 2.8 billion dollars (DiMasi et al., 2016). Despite this, 90% of the drug candidates entering clinical trials fail (Sun et al., 2022). Human body is a complex system, with myriad interactions taking place simultaneously, interdependent on each other. The same pathway or mechanism involving certain genes, may be responsible for different diseases. A drug developed for a particular condition, therefore, could be a potential candidate for another condition. Drug repurposing can drastically reduce the time and cost of developing new drugs by searching for FDA-approved drugs, drugs under trial, or other chemicals that have a therapeutic effect on conditions outside the scope of the original medical indication (Pushpakom et al., 2019). Drug repurposing minimizes the chances of failure in clinical trials and reduces time for approval.

Similar molecular and genetic aberrations among diseases can lead to the discovery of jointly important treatment options across biologically similar diseases. Oncologists have closely looked at prostate, ovarian and breast cancers and identified that the tumors arising from these cancers are typically hormone-dependent and have remarkable underlying pathological and molecular similarities in their DNA repair pathway abnormalities (Risbridger et al., 2010). Analyzing patient data from biologically similar cancers together provides insights into their similarities as well as knowledge about individual cancers, which may not have been possible by analyzing individual cancer data separately. Zhou et al. (2021) identified jointly important biomarkers across breast, prostate and ovarian cancers by utilizing patient data from the three cancers using a cross-cancer learning approach. This reiterates that the same pathway or a gene is responsible for multiple diseases. These biological similarities have led to remarkably similar treatment options. For instance, combining the androgen deprivation therapy (ADT) with PARP inhibitors (i.e. drugs already used in breast cancer treatment) showed to be an effective approach in reducing the progression and recurrence of prostate cancer. Several single agent activity PARP inhibitors (PARPi) were recently approved for treating certain ovarian and breast cancers (Asim et al., 2017). The US Food and Drug Administration (FDA) approved the first multi-cancer treatment (Keytruda®), for patients whose cancers had a common specific biomarker. FDA, for the first time, approved a drug based on a common biomarker, instead of the organ the tumor had originated. Despite this, majority of studies still consider each cancer disease in isolation from the rest and identify the treatment options that are cancer-type specific. Hence, the critical need is to discover multi-cancer treatment options through the exploitation of cancers with similar molecular and genetic aberrations.

Mutations in several genes within the homologous recombination (HR) pathway occur in around 20%–25% of advanced prostate cancers (Marshall et al., 2019). There is accumulating evidence that depicts a considerable proportion of individuals with metastatic breast cancer are HR deficient with mutations in *BRCA1/BRCA2* genes (den Brok et al., 2017). Base excision repair (BER) pathway genes limit the ability of DNA repair in prostate cancer (PCa, henceforth) patients, which leads to an increased risk of PCa. (Mittal et al., 2012). Further, *APEX1*, which is a BER gene, has shown a compelling effect indicating an increased risk of breast cancer through a gene-gene interactivity analysis (Kim et al., 2013). In an effort to understand the effect of mismatch repair (MMR) genes in the progression of PCa, gene expression-based analysis were conducted within the cancer cell lines and in tumor specimens, which indicated a loss of *MSH2* and *MLH1* genes in different cell lines (Chen et al., 2001). The deficiency of MMR genes was observed across most of the subtypes of breast cancers with high-grade tumor-infiltrating lymphocyte counts (Cheng et al., 2020). All these findings confirmed that there were significant commonalities across breast and prostate cancers in their DNA repair pathway abnormalities that could lead to common and jointly important treatment options.

Drug repurposing strategies can be classified into drug-based and disease-based, depending on the substantial availability of data and the intent of the research (Jarada et al., 2020) (Dudley et al., 2011). Several computational approaches proposed in recent years have used both disease and drug data (Peyvandipour et al., 2018) (Sirota et al., 2011) (Chiang and Butte, 2009) (Gottlieb et al., 2011). In a systems biology approach proposed by Peyvandipour et al. (2018) a drug-disease network (DDN) was constructed by considering drug targets, disease-related genes and all signalling pathways that were then integrated with disease gene expression signatures and drug-exposure gene expression signatures to discover novel therapeutic roles for established drugs. Nafiseh et al. used a machine learning approach to find anti-similarities between drugs and disease (Saberian et al., 2019). In their approach, they used drug exposure gene expression data, disease gene expression data and the associations between FDA-approved drugs and diseases. They used a distance metric learning (DML) algorithm where disease and the associated FDA-approved drugs had smaller distances compared to drugs not associated with disease. Luo et al. (2016) proposed a novel approach that computed the similarity between drugs and diseases. In particular, they constructed a heterogeneous network consisting of drug and disease similarity networks and drug–disease interactions and then used a Bi-Random walk (BiRW) algorithm to rank the drugs (Xie et al., 2012). Hu and Agarwal. (2009) generated a disease-drug network based on extensive drug and disease gene expression profiles which was used for identifying new indications for drugs and side effects of drugs.

In this paper, we used several state-of-the-art drug repurposing approaches to determine potential drug

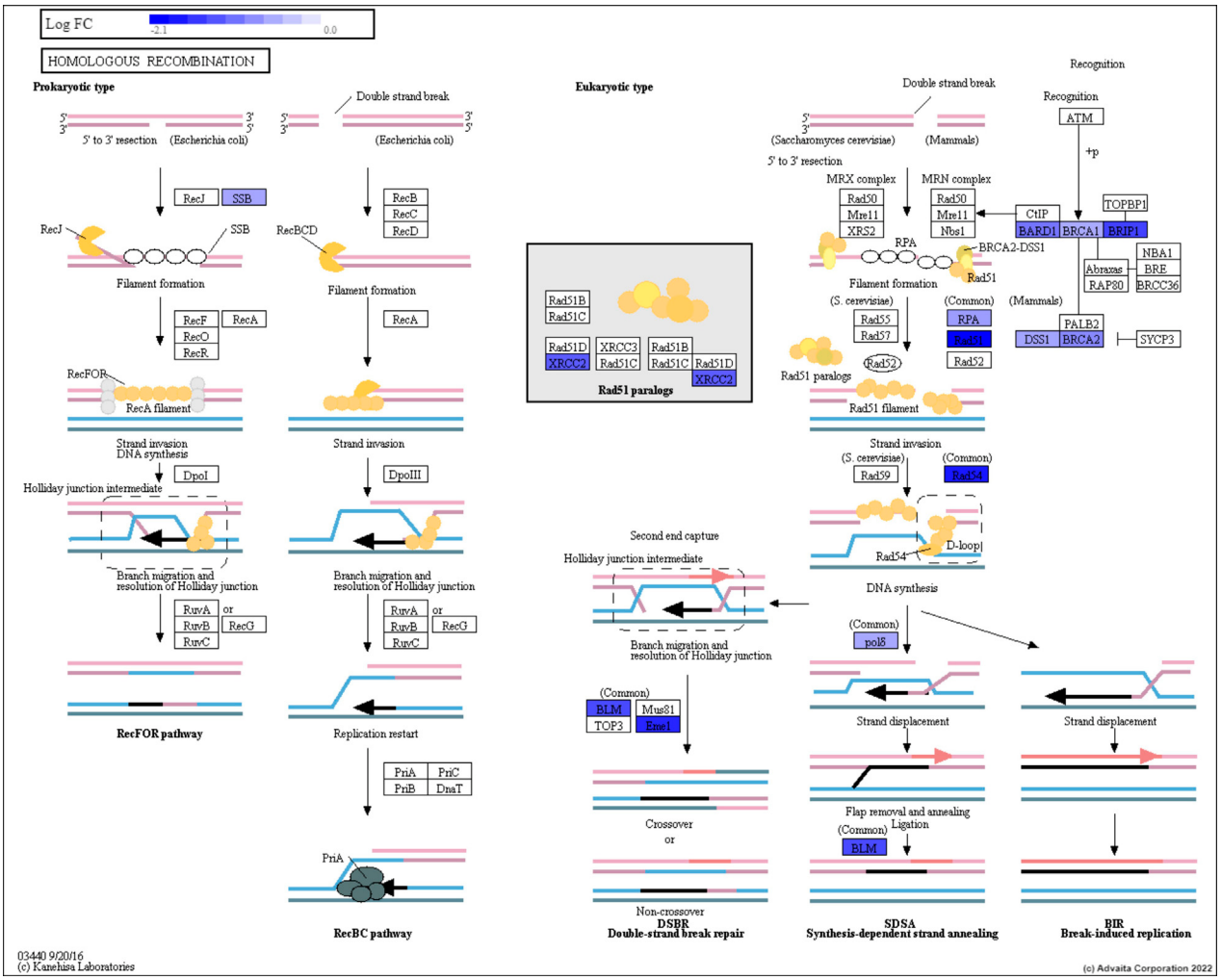


FIGURE 1
The Homologous Recombination pathway. The genes are represented in the rectangular boxes, with the shades of blue representing down-regulated genes for prostate cancer patients.

candidates for patients with breast or prostate cancers with common specific biomarkers. More specifically, we identified drugs with potential therapeutic effects on patients with DNA repair deficiencies.

Our contribution in this study is three-fold: 1) We initially developed a data-driven approach able to enrich the study population by integrating data from biologically similar cancers and using patient subpopulations with different types of DNA repair deficiencies which will enable personalized treatment strategies. We then used an existing approach referred to as drug-disease similarity to come up with novel treatments on the integrated data by identifying drugs that may have a therapeutic effect on patients irrespective of their cancer type. 2) We revisited our previously published deep cross cancer learning approach to identify jointly important biomarkers among breast, prostate and ovarian cancers. These biomarkers were used to identify common treatment options among those cancers through network interactions-based drug repositioning. 3) We presented the associations between the proposed drug target genes and biological functions (e.g., cell cycle) and investigated the drug target genes within the HR pathway and their interactions with the proposed drugs.

2 Materials and methods

2.1 Data preparation

The variant data and the disease gene expression data for breast and prostate cancers were obtained from The Cancer

TABLE 1 The number of breast and prostate cancer patients with deficiencies in their DNA repair pathways. Note that, different types of DNA-repair deficiencies has formed several subpopulations, that were analyzed separately.

DNA repair pathway	Number of patients	
	Breast cancer	Prostate cancer
Homologous Recombination (HR)	36	14
Base Excision Repair (BER)	23	7
Mismatch Repair (MMR)	73	31
Nucleotide Excision Repair (NER)	55	23
Non-Homologous End Joining (NHEJ)	23	6
Total	210	147

Genome Atlas (TCGA). The number of samples for breast and prostate tumors were 1,091 and 495, respectively with 120 and 53 samples with adjacent normal tissues. All expression datasets were log2 transformed. We obtained the signalling pathways from Kyoto Encyclopedia of Genes Genomics (KEGG) (Kanehisa et al., 2016). The signalling pathways are represented in the form of a directed graph, where each node represents the genes (or proteins) and the associations including activation, inhibition, etc. between the genes were represented by the edges. The large scale drug-exposure gene expression data were obtained from the Connectivity Map and the Library of Integrated Network-Based Cellular Signatures (LINCS) (Subramanian et al., 2017).

We initially identified all genes within each DNA repair pathway separately using the KEGG database. The DNA repair pathways used were: homologous recombination (HR), base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER) and non-homologous end joining pathway (NHEJ). As an example, the set of genes (or proteins) that exist within the HR pathway can be seen in Figure 1. Using the variant data collected from TCGA, a subset of breast and prostate cancer patients with mutations in any of their DNA repair genes were identified and grouped according to their type of DNA repair deficiency. This resulted in multiple cohorts of homogeneous subpopulations with common biomarkers. Table 1 shows the distribution of the breast and prostate cancer patients within each cohort. Note that, the same patient may fall into multiple cohorts.

Next, we identified the differentially expressed genes (DEGs) through a moderated *t*-test by comparing the tumor samples with their adjacent normal tissues on each cohort separately. The resulting *p*-values were FDR adjusted to correct for multiple comparisons. Including ovarian cancer samples would have been optimal as ovarian cancer is known to also have biological similarities with breast and prostate cancers. However, due to not having access to TCGA ovarian cancer gene expression data of *adjacent normal tissue*, we were unable to run the differential expression analysis on ovarian cancer samples in this study. An alternative approach we considered was to run experiments on ovarian cancer data collected from different data sources, however this requires extensive preprocessing due to different representation, distribution, scale, and density of data.

Our previously published deep cross cancer learning approach discussed in Section 3.3 identified jointly important biomarkers among breast, prostate, and ovarian cancers (Zhou et al., 2021). We were then able to identify drug candidates common among the three cancers using the proposed biomarkers. As this was a multi-label classification based neural network, we were able to conduct the analysis without the presence of ovarian normal tissue.

The methodology used for data preparation described above has been shown in Figures 2A,B. Prediction of drugs using drug-disease similarity and validation shown in Figure 2C has been described in subsequent sections.

2.2 The prediction of drugs using drug-disease similarities

Sirota et al. (2011) proposed a systematic computational drug repurposing approach to predict novel therapeutic indications by understanding drug and disease relationships. The association between every pairing of drug and disease is represented by a similarity score ranging from +1 to −1, with +1 indicating perfect correlation and −1 indicating an opposite effect. The largest negative score representing a reverse set of changes with exposure to a drug, indicates that the drug may have a therapeutic effect on the disease.

Here, we used the preprocessed expression data as discussed in Section 2.1 for breast and prostate cancer and the drug expression signatures from CMap to calculate the similarity scores. We only considered those drugs with FDR-adjusted *p*-values less than 0.05. This shortened list was then arranged in the ascending order based on the enrichment scores. The largest negative score implied the best drug candidates with highest therapeutic effects.

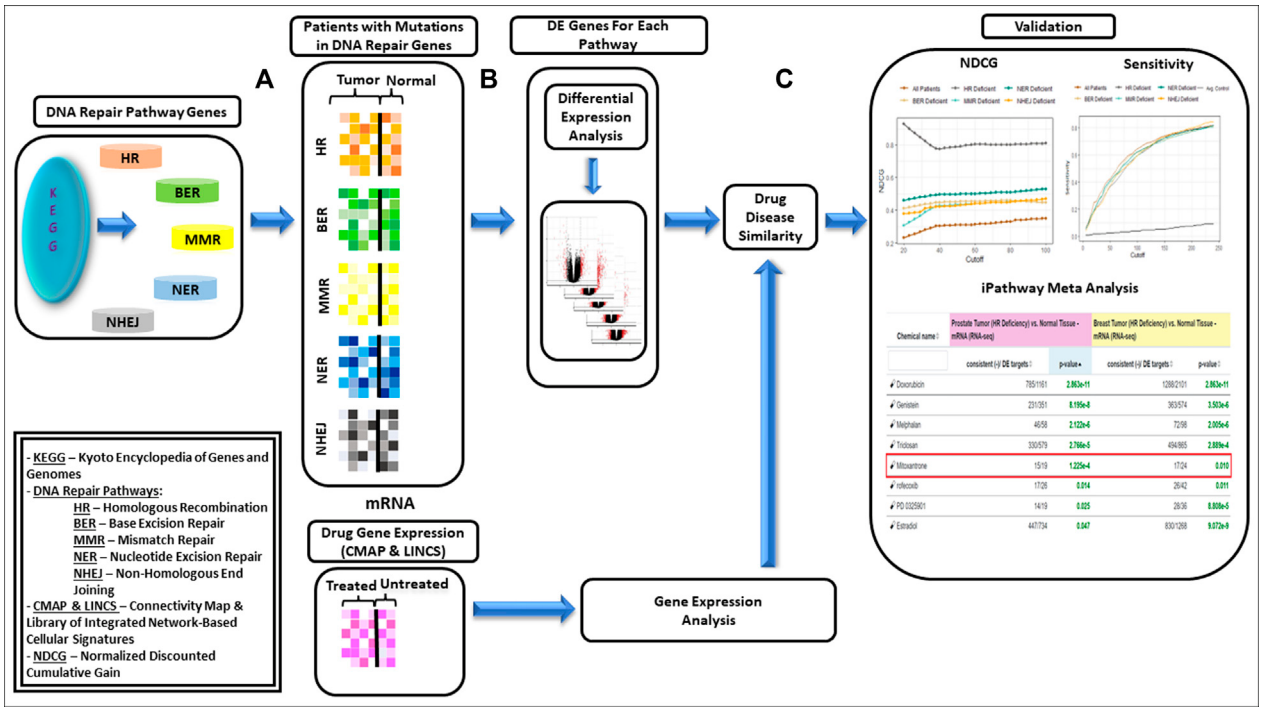


FIGURE 2
Framework proposed for data-driven drug repurposing for biologically similar cancers—**(A)**: Genes within each of the DNA repair pathways, i.e., HR (Homologous Recombination), BER (Base Excision Repair), MMR (Mismatch Repair, NER(Nucleotide Excision Repair) and NHEJ (Non-Homologous End Joining) were identified using KEGG database. Subset of breast and prostate cancer patients with mutations in DNA repair genes were identified and grouped based on DNA repair deficiency. **(B)**: Differentially expressed genes (DEGs) were identified on each cohort separately. **(C)**: Drugs for each cohort were identified using Drug-Disease Similarity. Framework was validated using NDCG (Normalized Discounted Cumulative Gain) and sensitivity scores; and network interaction analysis was used for validating the utility of the drugs.

In an effort to evaluate the results obtained through the drug-disease similarity model, we performed sensitivity-based validation only (SV) and calculated the normalized discounted cumulative gain (NDCG). The best strategy for analytic validation of drug repurposing is through sensitivity based validation techniques. Sensitivity and specificity based validation, although ideal, is not practical to assess the model performance due to the lack of access to true negatives (TNs) as discussed by Adam et al. (Brown and Patel, 2018). The discounted cumulative gain was constructed under the assumption that top rank drugs were more relevant and more likely to be of interest (Schuler et al., 2022). The NDGC score was calculated as follows:

$$DCG = \sum_{i=1}^{p} \frac{2^{\mathrm{rel}_i - 1}}{\log_2 (i+1)} \quad (1)$$

$$IDCG = \sum_{i=1}^{|REL_p|} \frac{2^{\mathrm{rel}_i - 1}}{\log_2 (i+1)} \quad (2)$$

$$NDGC = DCG/IDCG \quad (3)$$

where i is the rank of the drug of interest, up to rank p, and rel_i denotes the relevance of the drug to the indication, 0 indicating non-relevance and 1 indicating relevance, REL_p is the list of associated drugs in the set up to a cutoff position of p, and $|REL_p|$ is the cardinality of the list.

2.3 The validation of proposed drugs using network interactions

Here, we used a drug repurposing analysis module to identify FDA-approved drugs that could be used to revert a given pattern of gene expression changes caused by a disease. The prediction of upstream Chemicals, Drugs, Toxicants (CDTs) is based on two types of information: 1) the enrichment of differentially expressed genes from the experiment and 2) a network of interactions from the Advaita Knowledge Base (AKB v2006). The network is a directed graph in which the source node represents either a chemical substance or compound, a drug, or a toxicant. The edges represent known effects that these CDTs have on various genes. A signed edge in this graph consists of a source CDT, a target gene, and a sign to indicate the type of effect: activation (+) or inhibition (−). To generate the network, the analysis selects only those edges observed in the literature with at least a medium confidence. The analysis considers two

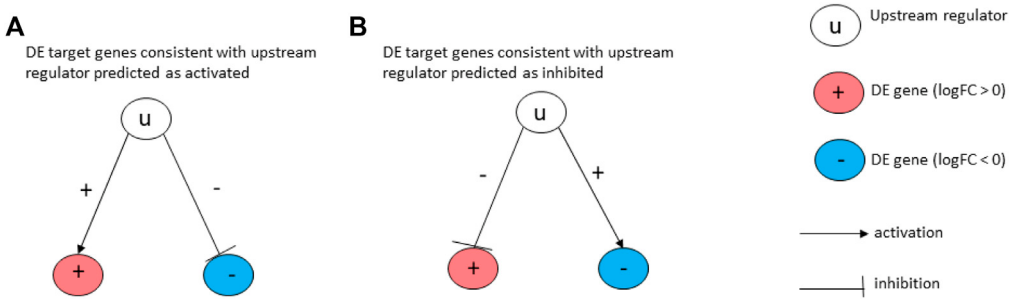


FIGURE 3
Target genes consistent with the hypothesis considered: In **(A)**, the signs of the DE genes shown in red (+) and blue (–) match the signs of their respective incoming edges, suggesting that the upstream regulator u is activated. In **(B)**, the signs of the DE genes shown in red (+) and blue (–) are opposite to the signs of their edges, suggesting that the upstream regulator u is inhibited.

hypotheses: HA: The upstream regulator is activated in the condition studied. HI: The upstream regulator is inhibited in the condition studied. The set of genes from National Center for Biotechnology Information (NCBI) Gene database is divided into many subsets by the analysis based on the measurements from the experiment and the definitions shown in Figure 3. The (+) sign in the figure indicates up-regulated genes while (−) sign indicates down-regulated genes. If a gene has at least one incoming edge, then it is considered as a target gene in the network. The gene g is consistent with hypothesis HA if there is an incoming edge e and if sign(g) = sign(e). This implies that when upstream regulator is activated, the signal is an activation and gene is up-regulated or signal is an inhibition, and the gene is down-regulated. (see Figure 3A). The gene g is consistent with hypothesis HI if there is an incoming edge e and if sign(g) does not match sign(e). This implies that when upstream regulator is inhibited the signal is inhibition and gene is up-regulated or signal is activation and gene is down-regulated. (see Figure 3B).

Herein, we focused on drugs that could reverse the changes induced by the disease. For this purpose, we hypothesized that the disease is considered as a state in which the changes are associated with the absence of a drug. Given the interactions between a specific drug A and its downstream DE genes, the Z-score was computed as follows:

$$z(A) = \frac{\sum_{e,g} w(g).s(e).s(g)}{\sqrt{\sum |w(g)|^2}} \quad (4)$$

where $s(e)$ represents the type of the edge (−1 for inhibition and +1 for activation), $s(g)$ is the sign of expression change of the gene (−1 for down-regulated and +1 for up-regulated), and $w(g)$ the confidence score of the edge g. The Z-score p-value for each drug was then calculated by mapping the z-score on a p-value using the normal distribution. (Draghici et al., 2020).

Note that, the drugs identified through drug-disase similarities as discussed in Section 2.2, though powerful, do not consider the network of interactions between drugs and their associated downstream genes. On the other hand, the network interactions as discussed in this section may still not be able to detect all significant drugs as only direct interactions between drug and disease is considered, rather than investigating indirect interactions due to co-expressions of genes. Hence in order to identify drugs with high therapeutic effects, we relied on the intersecting drugs among multiple approaches.

3 Results

3.1 Drug-disease similarity results

The results obtained through the drug-disease similarity analysis are shown in Table 2. Initially, all breast and prostate cancer patients were included in the analysis which resulted in a list of drugs presented in the first column of the table (see column: All Patients). In essence, a good repurposing approach on a truly homogeneous data should place the already FDA-approved drugs (i.e., the gold standard) at the very top of the list for that particular disease. Note that, since we focussed on multiple biologically similar diseases, we expected to see drugs approved for either or both of the conditions at the very top of the list.

Results showed that six investigational drugs (two of which are under investigation for breast and prostate cancers, and four of which are under investigation for breast cancer only) and no FDA-approved drugs appeared within the top 10 ranked drugs. Cancer being a heterogeneous disease with large genetic diversity even between tumors of the same cancer types, it is common for the patients to have significant differences between their molecular profiles (Arslanturk et al., 2020). Our results clearly showed that the data needed to be further refined to identify more homogeneous subpopulations for more optimal and targeted treatment decisions. Hence, as the next step, we

TABLE 2 The list of top ranked drugs identified through the drug-disease score analysis for subsets of patients with different types of DNA repair deficiencies. The cells highlighted in green, grey, blue and pink are the FDA-approved drugs, investigational drugs for breast and prostate cancers, investigational drugs for prostate cancer, and investigational drugs for breast cancer, respectively along with their respective similarity scores that was calculated. Results demonstrated that although there are certain drugs that are common across subpopulations, the top ranked drugs differed between different DNA-repair pathways. Hence, the identification of biomarkers associated with a specific subpopulation can change the course of treatment and enable personalized treatment strategies among individuals.

		Patients with DNA Repair Deficiencies			
All Patients	Similarity Score	Homologous Recombination	Similarity Score	Base Excision Repair	Similarity Score
GSM1738425_decitabine	-0.2777	GSM1742795_palbociclib	-0.3093	GSM1743216_canertinib	-0.4304
GSM1740982_pyrazolanthrone	-0.2743	GSM1739132_tranylcypromine	-0.2848	GSM1746670_foretinib	-0.3975
GSM1744649_linifanib	-0.2669	GSM1744653_linifanib	-0.2804	GSM1747119_mitoxantrone	-0.3804
GSM1746995_radicicol	-0.2638	GSM1744013_selumetinib	-0.2778	GSM1743879_trametinib	-0.3800
GSM1737390_motesanib	-0.2570	GSM1743929_dasatinib	-0.2725	GSM1742795_palbociclib	-0.3796
GSM1746047_foretinib	-0.2547	GSM1741703_radicicol	-0.2689	GSM1743011_imatinib	-0.3754
GSM1742374_sorafenib	-0.2415	GSM1743780_alvocidib	-0.2651	GSM1739677_dabrafenib	-0.3718
GSM1744868_roscovitine	-0.2409	GSM1739016_mocetinostat	-0.2598	GSM1747001_geldanamycin	-0.3714
GSM1744016_selumetinib	-0.2328	GSM1746633_dovitinib	-0.2547	GSM1746921_radicicol	-0.3657
GSM1739132_tranylcypromine	-0.2289	GSM1744512_saracatinib	-0.2544	GSM1741739_sirolimus	-0.3564
GSM1737918_rocilinostat	-0.2283	GSM1737624_entinostat	-0.2518	GSM1744905_erlotinib	-0.3521
GSM1747012_mitoxantrone	-0.2283	GSM1740698_dabrafenib	-0.2461	GSM1744041_gefitinib	-0.3436
GSM1743003_gefitinib	-0.2278	GSM1738635_azacitidine	-0.2411	GSM1743959_linifanib	-0.3429
GSM1745427_pelitinib	-0.2277	GSM1740982_pyrazolanthrone	-0.2384	GSM1743114_fostamatinib	-0.3404
GSM1739411_rucaparib	-0.2242	GSM1742407_fostamatinib	-0.2381	GSM1737853_iniparib	-0.3393
GSM1738010_pracinostat	-0.2227	GSM1746979_geldanamycin	-0.2356	GSM1746633_dovitinib	-0.3357
GSM1743929_dasatinib	-0.2223	GSM1739042_pracinostat	-0.2356	GSM1742708_alvocidib	-0.3336
GSM1742795_palbociclib	-0.2181	GSM1737914_rocilinostat	-0.2316	GSM1742848_afatinib	-0.3325
GSM1744029_lapatinib	-0.2167	GSM1741596_mitoxantrone	-0.2283	GSM1738293_azacitidine	-0.3307
GSM1742406_fostamatinib	-0.2150	GSM1740862_serdemetan	-0.2249	GSM1741787_vorinostat	-0.3296

		Patients with DNA Repair Deficiencies			
Mismatch Repair	Similarity Score	Nucleotide Excision Repair	Similarity Score	Non Homologous End Joining	Similarity Score
GSM1744013_selumetinib	-0.2501	GSM1745351_dovitinib	-0.2526	GSM1741596_mitoxantrone	-0.1966
GSM1740698_dabrafenib	-0.2481	GSM1741655_radicicol	-0.2504	GSM1742801_palbociclib	-0.1908
GSM1741655_radicicol	-0.2408	GSM1742795_palbociclib	-0.2369	GSM1741703_radicicol	-0.1874
GSM1741596_mitoxantrone	-0.2241	GSM1741669_mitoxantrone	-0.2310	GSM1745758_roscovitine	-0.1718
GSM1738074_resveratrol	-0.2167	GSM1743222_erlotinib	-0.2198	GSM1747003_geldanamycin	-0.1675
GSM1747114_geldanamycin	-0.2148	GSM1743717_gefitinib	-0.2159	GSM1740858_serdemetan	-0.1571
GSM1743478_saracatinib	-0.2138	GSM1738639_azacitidine	-0.2146	GSM1744653_linifanib	-0.1533
GSM1737697_rucaparib	-0.2135	GSM1742421_saracatinib	-0.2145	GSM1738293_azacitidine	-0.1485
GSM1746224_trametinib	-0.2122	GSM1740517_veliparib	-0.2145	GSM1745640_enzastaurin	-0.1336
GSM1738291_azacitidine	-0.2112	GSM1745251_brivanib	-0.2144	nabumetone_5428	-0.1252
GSM1743162_palbociclib	-0.2092	GSM1744305_linifanib	-0.2079	GSM1743082_sorafenib	-0.1251
GSM1744099_alvocidib	-0.2065	GSM1743781_alvocidib	-0.2065	GSM1745910_neratinib	-0.1222
GSM1745118_roscovitine	-0.2039	GSM1743465_fostamatinib	-0.2054	GSM1742648_gefitinib	-0.1198
GSM1738554_olaparib	-0.2020	GSM1743691_selumetinib	-0.2052	GSM1746569_crizotinib	-0.1160
GSM1742182_neratinib	-0.1962	GSM1745328_enzastaurin	-0.2032	GSM1744103_alvocidib	-0.1153
GSM1745351_dovitinib	-0.1959	GSM1745118_roscovitine	-0.2025	GSM1740644_ponatinib	-0.1142
GSM1745226_ruxolitinib	-0.1958	GSM1742182_neratinib	-0.2020	GSM1746600_vemurafenib	-0.1121
GSM1742374_sorafenib	-0.1955	GSM1742860_canertinib	-0.1983	GSM1745215_nilotinib	-0.1107
GSM1740333_dasatinib	-0.1922	GSM1744937_trametinib	-0.1970	GSM1738748_belinostat	-0.1106
GSM1739131_tranylcypromine	-0.1898	GSM1742988_lapatinib	-0.1939	GSM1746833_brivanib	-0.1102

investigated potential treatment options based on common biomarkers, specifically for patients with aberrations in genes within different DNA repair mechanisms. Results showed Palbociclib, an endocrine-based chemotherapeutic agent approved for treating HER2-negative and HR-positive advanced or metastatic breast cancers (McCain, 2015) (Walker et al., 2016) (Beaver et al., 2015), appeared at the top of the list for patients with HR-deficiencies. Results further suggested that tranylcypromine, a monoamine oxidase inhibitor, mainly approved for the treatment of major depressive episodes without melancholia (Ricken et al., 2017), showed promise as a multi-cancer treatment, specifically for

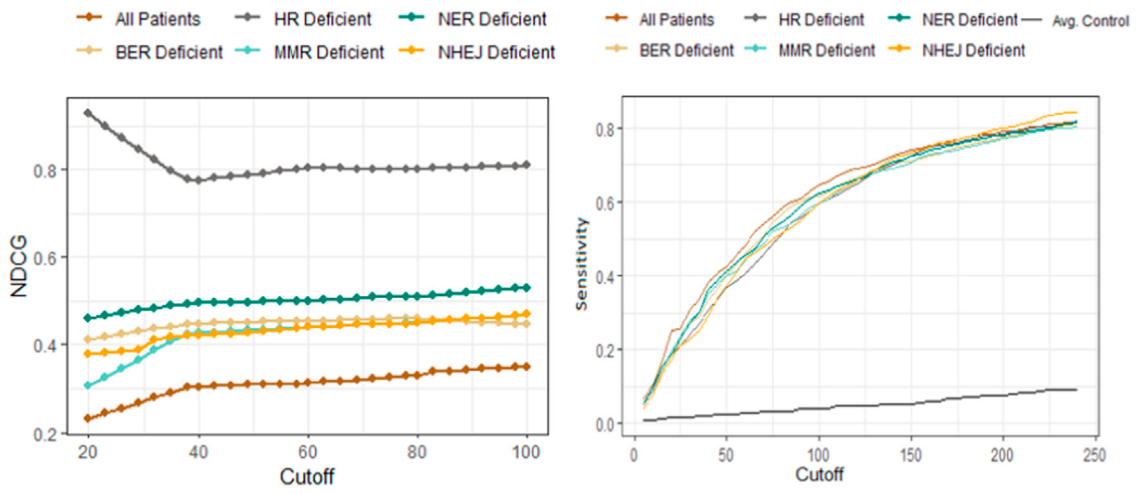


FIGURE 4
Performance comparison of the drug-disease similarity model on DNA-repair deficient patient subpopulations using NDCG (left) and sensitivity analysis (right). The NDCG/sensitivity values (vertical axes) of all drug–indication associations using different DNA repair deficient subpopulations are shown according to different cutoff values (horizontal axis). The NDCG results clearly demonstrate that the HR-deficient subpopulations result in drugs that are clinically more relevant with more FDA-approved/investigational drugs compared with other DNA-repair pathway deficiencies. The plot has further shown that identifying homogeneous subpopulations through common biomarkers result in better performances when compared to all patients combined. The sensitivity values demonstrate that the list of breast/prostate cancer drugs retrieved for all cutoff levels are clinically relevant and indicates an overall better performance relative to random controls (shown as the black curve).

breast and prostate cancers. The top ranked drugs further consisted of several chemotherapy drugs including linifanib, selumetinib and dasatinib. The top ranked drugs for all other DNA repair deficient patients are listed in Table 2. A detailed description of all the top ranked drugs for each pathway along with their clinical relevance is reported in the discussion section of the paper.

The sensitivity and NDCG scores of the proposed drugs are shown in Figure 4. The sensitivity values of all drug-disease associations for different subsets of patients based on their types of DNA repair deficiencies were compared with several random control runs. The sensitivity values were reported for different rank/cutoff levels. The SV results as shown in Figure 4B demonstrates that the list of drugs retrieved for all cutoff levels for breast and prostate cancer patients were clinically relevant and indicated an overall better performance relative to random controls. The NDCG scores as shown in Figure 4A show that the identification of homogeneous sub populations with common biomarkers resulted in drugs that were clinically more relevant with more FDA-approved/investigational drugs appearing at the very top of the list when compared with all patients combined. Results further showed that drugs proposed for patients with aberrations in their HR pathway outperformed all other pathways. This is mainly due to hormone driven cancers' significant molecular similarities within HR pathways (Toh and Ngeow, 2021) (Watkins et al., 2014). Less is known about the similarities between those cancers in other DNA-repair pathways.

3.2 Drugs proposed through network interactions

The drugs proposed through network interactions using iPathwayGuide (Advaita) are listed in Table 3. Note that, this table includes only the drugs that have a significant therapeutic effect ($p < 0.05$) on both breast and prostate cancers. The number of DE genes that would be reverted by each drug is listed. For instance, the 15/19 notation next to mitoxantrone demonstrates that there were 19 downstream genes that mitoxantrone is interacting with that were DE for prostate cancer (vs. adjacent normal tissue), 15 of which were consistent with our hypothesis as described in Section 2.3.

The SV and NDCG are metrics used to evaluate the drug repurposing models' ability to identify clinically relevant treatment options. In order to validate the utility of the drugs proposed, we investigated the mechanisms through which the drugs act on genes measured to be DE for the disease studied. Figure 5 generated using network interactions shows the mechanisms of mitoxantrone on the DE genes for prostate cancer. Mitoxantrone was able to activate the down-regulated genes and inhibit the up-regulated genes 15 out of 19 times ($p < 1.225e-4$) as described in Section 2.3 In an effort to confirm the changes in the downstream genes, we have reported the fold-changes of those genes using cell lines treated with Mitoxantrone as shown in Figure 5B. The upregulated genes are highlighted in red, and the downregulated genes are highlighted in blue.

TABLE 3 The top eight drugs proposed for repurposing using the network interactions approach. The table shows the *p*-values (sorted based on the prostate tumor vs. adjacent normal tissue experiment), as well as the number of DE genes that would be reverted by each drug (i.e., the number of genes consistent with the hypothesis) for patients with HR–deficiencies. Doxorubicin slows or stops the growth of cancer cells, and is used to treat certain neoplastic conditions such as acute lymphoblastic leukemia, soft tissue and bone sarcomas, breast carcinoma and ovarian carcinoma. Genistein is currently under clinical trials for the treatment of prostate cancer. Melphalan and Estradiol are also among drugs used to treat certain cancers. Mitoxantrone is highlighted as a promising drug candidate as it appears to be a top drug using both network interactions and drug-disease similarity scores.

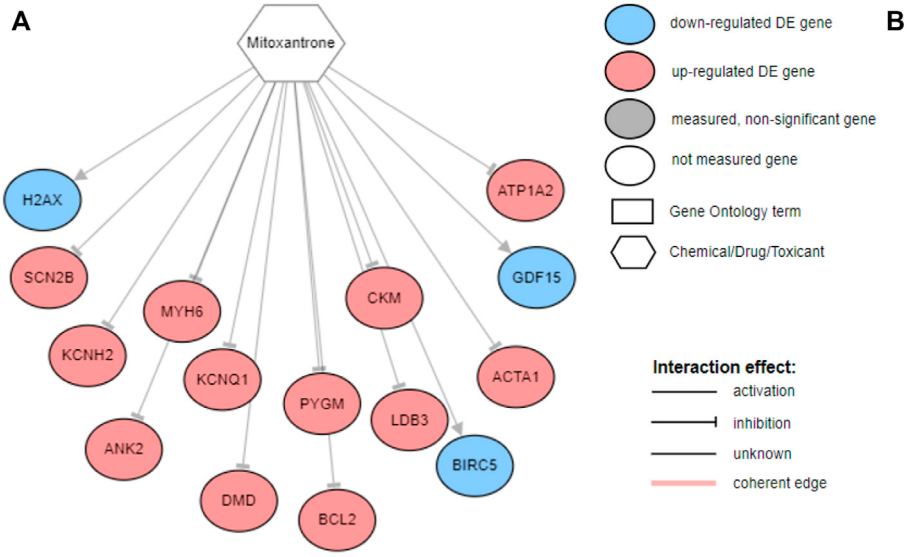
Chemical name	Prostate tumor (HR deficiency) vs. Normal tissue - mRNA (RNA-seq)		Breast tumor (HR deficiency) vs. Normal tissue - mRNA (RNA-seq)	
	Consistent (-)/DE targets	*p*-value	Consistent (-)/DE targets	*p*-value
Doxorubicin	785/1161	2.863e-11	1288/2101	2.863e-11
Genistein	231/351	8.195e-8	363/574	3.503e-6
Melphalan	46/58	2.122e-6	72/98	2.005e-6
Triclosan	330/579	2.766e-5	494/865	2.889e-4
Mitoxantrone	15/19	1.225e-4	17/24	0.01
rofecoxib	17/26	0.014	26/42	0.011
PD 0325901	14/19	0.025	28/36	8.808e-5
Estradiol	447/734	0.047	830/1268	9.072e-9

3.3 Drugs proposed using novel biomarkers discovered using cross cancer learning approach

We utilized our previously published approach that discovered jointly important novel biomarkers across breast, prostate and ovarian cancers through a data-driven, deep learning approach referred to as cross-cancer learning (Zhou et al., 2021). This approach exploited patient data from multiple cancers to discover prostate cancer biomarkers and jointly important biomarkers across breast, prostate and ovarian cancers by leveraging pathological and molecular similarities in their DNA repair pathways. Different cancers share common genomic instabilities. Exploring cancers having similarities can help discover previously unknown biomarkers and pathways. In addition, this helps in alleviating the problem of limited patient samples availability and underestimation of various genes previously not known to be involved. This cross cancer learning framework utilized a multi-label classification autoencoder (MLC-AE) that used lower dimensional latent representation of the mRNA gene expression profiles to predict the tissue type (breast, prostate, ovarian) and the disease state (solid tumor vs. adjacent normal tissue) as separate output layers. To explain and interpret the MLC-AE model, SHapley Additive exPlanations (SHAP) was used. This method uses SHAP values to extract feature importance across three cancers. SHAP method used each feature to calculate the change in performance in the presence and absence of each feature. The features whose absence lead to reduction in the performance were given the highest score. The cross cancer framework has been shown in Figure 6. Figure 7 A shows the most significant genes based on their contribution towards prediction using breast, prostate, and ovarian tissues. The biomarkers discovered using this approach were further used to find disrupted pathways using the impact analysis. The drugs identified using cross cancer genes are listed in Figure 7B and are discussed in detail in the Discussion section.

In order to validate our results further, additional experiments were conducted using the cell lines obtained from CMap. Table 4 shows the fold changes that were calculated using the cell lines treated with the drugs shown on each column. Specifically, the drugs investigated were Genistein, Mitoxantrone, Palbociclib, Tranylcypromine, Linifanib and Selumetinib. Threshold parameters used for the analysis were an absolute fold-change greater than 0.6 and false discovery rate (FDR) adjusted *p*-value less than 0.05. All genes presented in the table are differentially expressed, with the genes associated with DNA repair pathways being color-coded. Specifically, the red, yellow, green, blue, and gray colors represent significant changes in the genes associated with BER, HR, MMR, NER and a combination of multiple DNA repair pathways, respectively.

Note that there are no differentially expressed genes involved in the DNA repair process for Genistein. However, expression changes obtained from CMap includes an arbitrary selection of patients and is not filtered based on homogeneous subpopulations identified through specific DNA repair deficiencies. Instead, our proposed drug candidates have been derived by filtering a list of patients with specific types of DNA repair deficiencies, and therefore, is a preprocessed dataset with a more homogenous population than the CMap patient set.



Gene	Fold Change
H2AX	-1.014
SCN2B	1.115
KCNH2	1.137
ANK2	1.17
MYH6	1.045
KCNQ1	1.286
DMD	1.671
BCL2	1.388
PYGM	1.745
CKM	2.513
LDB3	2.262
BIRC5	-1.953
ACTA1	1.964
GDF15	-2.001
ATP1A2	2.269

FIGURE 5
(A) The mechanism through which Mitoxantrone act on the genes measured to be DE for prostate cancer. Note that, out of the 19 downstream DE genes that Mitoxantrone is interacting with, 15 were consistent with the hypothesis, i.e., the drug was able to revert the expression changes caused by disease 15 out of 19 times. All 15 genes were shown on the figure with three down-regulated genes (blue circles) being activated, and 12 up-regulated genes (red circles) being inhibited with the exposure of the drug. **(B)** Fold changes reported for cell lines treated with Mitoxantrone. The upregulated genes are highlighted in red, and the downregulated genes are highlighted in blue.

Although this could explain the lack of gene changes in DNA repair pathways when Genistein is administered, additional analyses would be required to further confirm the therapeutic effect of this drug.

In order to understand the effect of our proposed drugs on the nodes within the HR pathway, we explored the drug-gene interactions. The results are shown in Figure 8. The differentially expressed genes highlighted in this figure are based on patients with HR deficient breast cancer vs. adjacent normal tissue. This figure clearly shows that several HR genes are indeed drug targets and our proposed drugs are indeed interacting with such genes.

In summary, our results showed several promising drug candidates including Mitoxantrone, Palbociclib and Genistein for multi-cancer treatment as supported by multiple approaches. Mitoxantrone appeared to be a top drug using drug-disease similarity scores and network interactions approaches, and Genistein appeared to be a top drug using cross-cancer biomarkers and network interactions.

Experiments conducted by Tang et al. (2018) and Siddiqui et al. (2021) suggest that genistein and mitoxantrone in combination with other drugs can influence the cell cycle of the cancer cells. In order to understand the effects of these drugs on cell-cycle, we ran an experiment on iPathwayGuide, to understand the associations between the downstream genes of Genistein and Mitoxantrone and their associations with biological processes including the cell cycle. Results are presented in Figure 9.

4 Discussion

DNA damage is not uncommon and results in tens of thousands of damages everyday (Jackson and Bartek, 2009; O'Connor, 2015). This genomic instability is the key feature of carcinogenesis. DNA damage response (DDR) collectively refers to all the mechanisms that are responsible for the DNA damage repair. O'Connor. (2015) discussed targeted therapies based on DNA damage response of patients to tailor targeted therapy. They further mentioned various drugs under clinical trials for different types of cancers targeting DNA repair pathways.

Homologous Recombination is responsible for the repair of DNA double stranded breaks (DSBs) during G2/M phase (Saleh-Gohari and Helleday, 2004). Li and Heyer. (2008); Al-Mugotir et al. (2021) showed that doxorubicin, and quinacrine, along with mitoxantrone were effective in HR deficient cells by recruiting *RAD52* to repair sites of DNA damage.

Table 2 shows the drugs that were identified using drug-disease score analysis for the subset of patients who had deficiencies in their DNA repair pathways for prostate cancer and breast cancer. In the list of drugs identified for HR pathway, palbociclib came as significant. Palbociclib is approved for HER2-negative and HR-positive advanced or metastatic breast cancer. It is known that *BRCA1* and *BRCA2* mutations are involved in the HR deficiency. Hence, this could be a promising drug for the prostate cancer patients who at

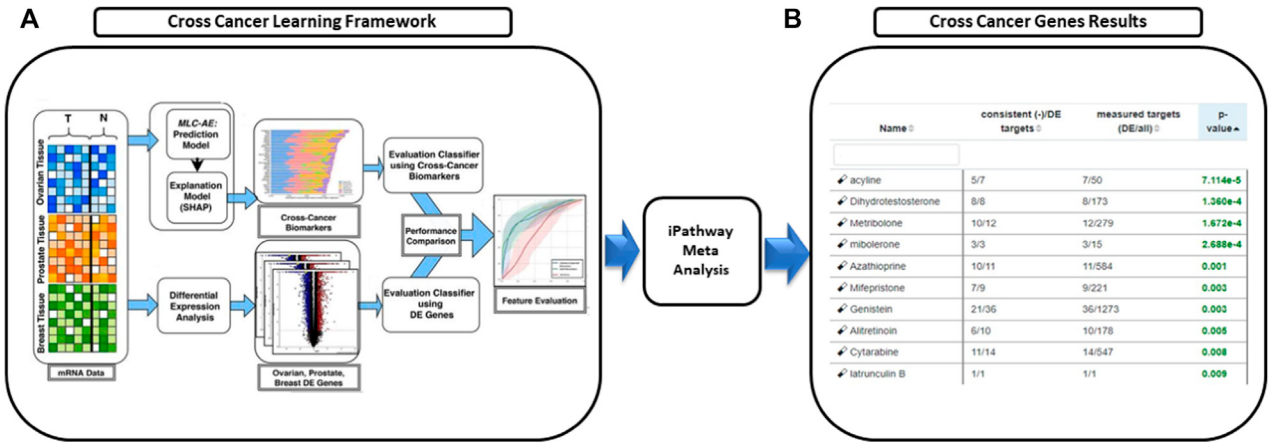
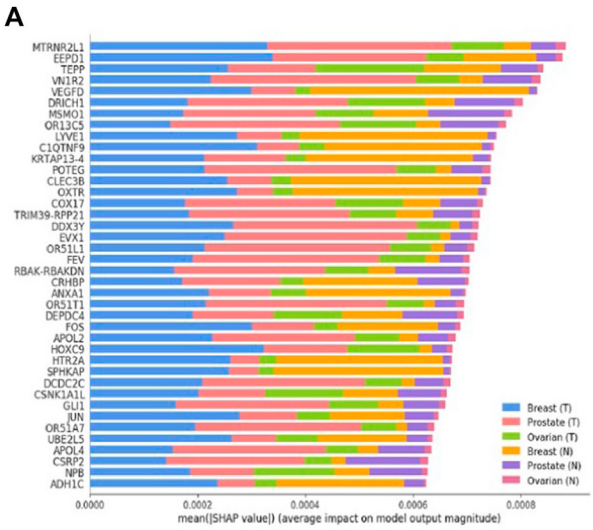


FIGURE 6
Drugs proposed using cross cancer genes. **(A)** Breast, prostate and ovarian cancer expression data was used to predict the tissue type and the disease type using multi-label classification—auto encoder (MLC-AE). SHAP Explanation model was used to identify the contribution of each gene towards the prediction using SHAP values that rank the genes. **(B)** Network interaction analysis was used to perform meta analysis and predict novel drugs.



B

Name	consistent (-)/DE targets	measured targets (DE/all)	p-value
acyline	5/7	7/50	7.114e-5
Dihydrotestosterone	8/8	8/173	1.360e-4
Metribolone	10/12	12/279	1.672e-4
mibolerone	3/3	3/15	2.688e-4
Azathioprine	10/11	11/584	0.001
Mifepristone	7/9	9/221	0.003
Genistein	21/36	36/1273	0.003
Alitretinoin	6/10	10/178	0.005
Cytarabine	11/14	14/547	0.008
Iatrunculin B	1/1	1/1	0.009

FIGURE 7
(A) Significant genes identified using SHAP based on contribution scores for all three tissues (breast, prostate and ovary). (T) Denotes solid tumor and (N) denotes solid normal tissue. Figure utilized from Zhou et al. (2021) **(B)** Top eight drugs proposed for repurposing using cross cancer genes.

present do not respond to the current treatment. The network interactions approach shown in Table 3 came up with interesting set of drugs. Studies have shown that Genistein affects cell cycle during G2/M phase (Zhang et al., 2013). Genistein inhibits protein-tyrosine kinase and topoisomerase-II (DNA topoisomerases, type II) and is under investigation as an anti-cancer agent. *In vivo* experiments carried out by Tang et al. (2018). have showed that Genistein when combined with AG1024 (a tyrosine kinase inhibitor) led to a decrease in tumor size in prostate cancer patients. Genistein suppressed the homologous recombination (HR) and the non-homologous end joining (NHEJ) pathways by inhibiting the expression of *Rad51* and *Ku70* (Tang et al., 2018). Genistein, an isoflavone found in soy products and an integral part of the Asian diet, was found to be effective against various cancers and responsible for lowering the prostate and breast cancer rates in

TABLE 4 This table shows the expression changes of genes when the drugs that were found to be significant in our analysis were administered. Red, yellow, green, blue, and gray colors represent significant changes in the genes associated with BER, HR, MMR, NER, and a combination of multiple DNA repair pathways, respectively.

MITOXANTRONE		GENISTEIN		PALBOCICLIB		TRANYLCYPROMINE		LINIFANIB		SELUMETINIB	
Gene	Fold Change	Gene	Fold Change	Gene	Fold Change	Gene	Fold Change	Gene	Fold Change	Gene	Fold Change
POLE2	-16.0984	TFPI2	-0.601714122	PCNA	-10	OXA1L	-10	BIRC5	-10	PARP1	-10
RFC2	-12.9617	TACC3	0.666655568	POLE2	-7.882	CCNA1	-10	CD44	-10	CD58	-10
BRCA1	-9.5644	KRT12	0.648860924	PARP1	-5.49	DNM1L	-10	CDK1	-10	IER3	-10
RFC5	-6.14226667	MRPS14	-1.007313661	DDB2	-4.78	PHKA1	-9.618	DLD	-10	DNM1L	-10
TOPBP1	-4.7233	DET1	-0.640326879	BRCA1	-4.616	TJP1	-6.588	MCM3	-10	TXNDC9	-10
PARP1	-2.8768	CRYBB1	-0.614814657	RFC2	-4.372	MSH6	-2.472	LIG1	-2.806	CCNH	0.732
MSH6	-2.0806	OSR2	-0.827627071	LIG1	-3.853	RFC5	-0.747	TOPBP1	1.414	PARP2	1.32
RAD51C	-1.7007	NCAPD3	0.632002377	PARP2	-3.061	RFC2	-0.633	PARP2	1.51	CDK7	1.33
CCNH	-1.3119	MBIP	-0.619790329	RFC5	-2.443	CDK7	0.719	RFC5	2.55	RFC5	1.506
CDK7	1.4994	ASCC2	-0.685010907	MSH6	-1.715	POLD4	0.801	DUSP22	8.1212	ERCC6	2.187
ERCC6	2.5378	DUSP1	0.618809242	TOPBP1	-1.405	RAD51C	1.05	PRR15L	8.34	CCL2	10
POLD4	4.5108	TMEM186	-0.633685938	RAD51C	-1.186	BRCA1	1.68	POLD4	8.877	SUV39H1	10
DDB2	5.9346	LIPE	-0.995489484	CCNH	-0.761	P4HA2	4.113	FBXL12	9.742	TLR4	10
ADAT1	13.0098	GABARAPL3	-0.625627357	ERCC6	0.732	CLPX	4.227	CIAO3	10	MAP7	10
DDX42	13.9626	MAP7D1	-0.715569418	CDK7	1.634	PRR15L	4.285	C2CD2	10	CLPX	10
MAPKAPK2	15.32706667	SUPT16H	-0.628213568	POLD4	8.877	NUDT9	4.683	E2F2	10	PRR7	10

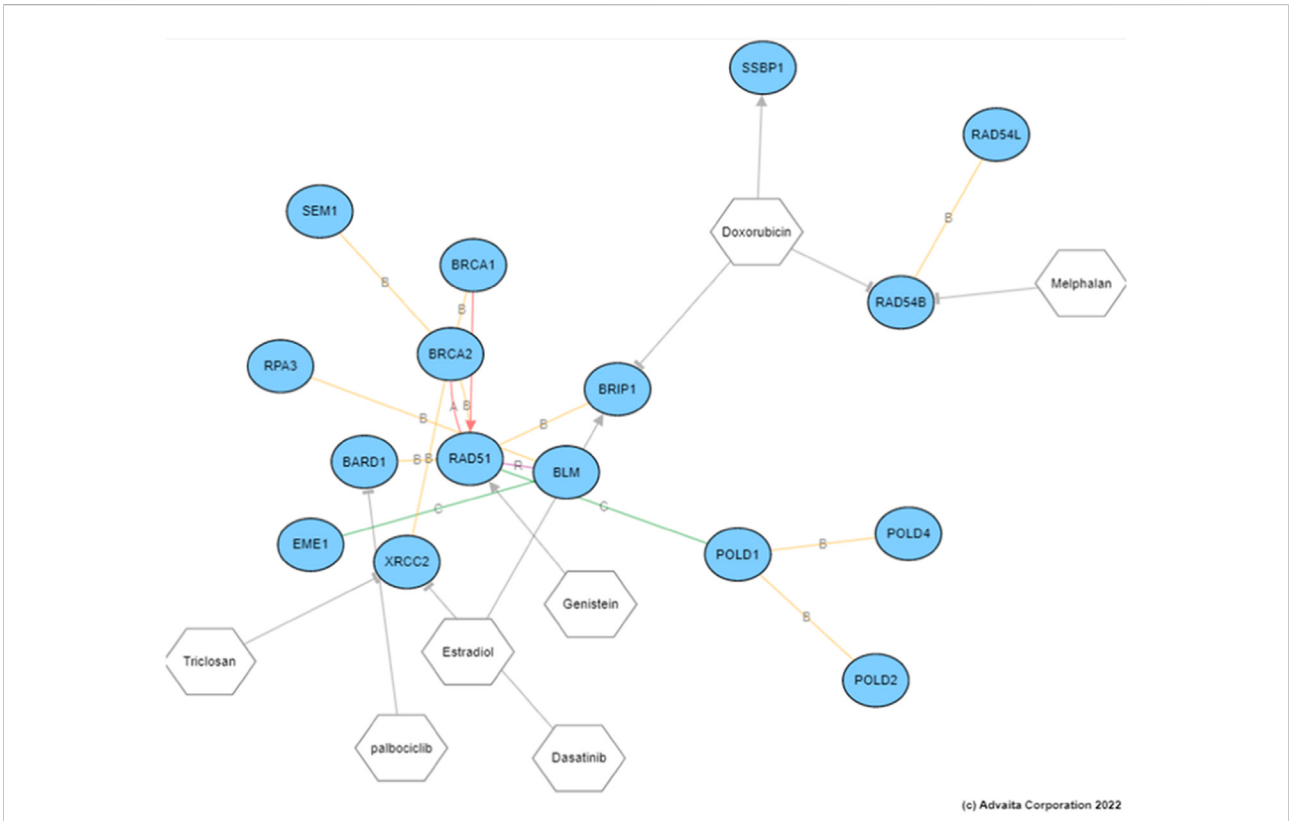


FIGURE 8
The drug target genes within the HR pathway and their interactions with the proposed drugs. The differential expression analysis here were conducted on patients with breast cancer with HR deficiencies vs. adjacent normal tissue.

Asian countries. It inhibited the cell cycle proliferation and induces apoptosis. (Banerjee et al., 2008). Khan et al. (2021) described the emerging role of natural products in cancer treatment. Among them, soy isoflavones, were reported to target *BRCA* histones for repair. Through their *in vivo* experiments, Fan et al. (2006) found that genistein along with indoole-3-carbinol targeted both *BRCA1* and *BRCA2* genes in breast and prostate cancer cells. This research is useful in

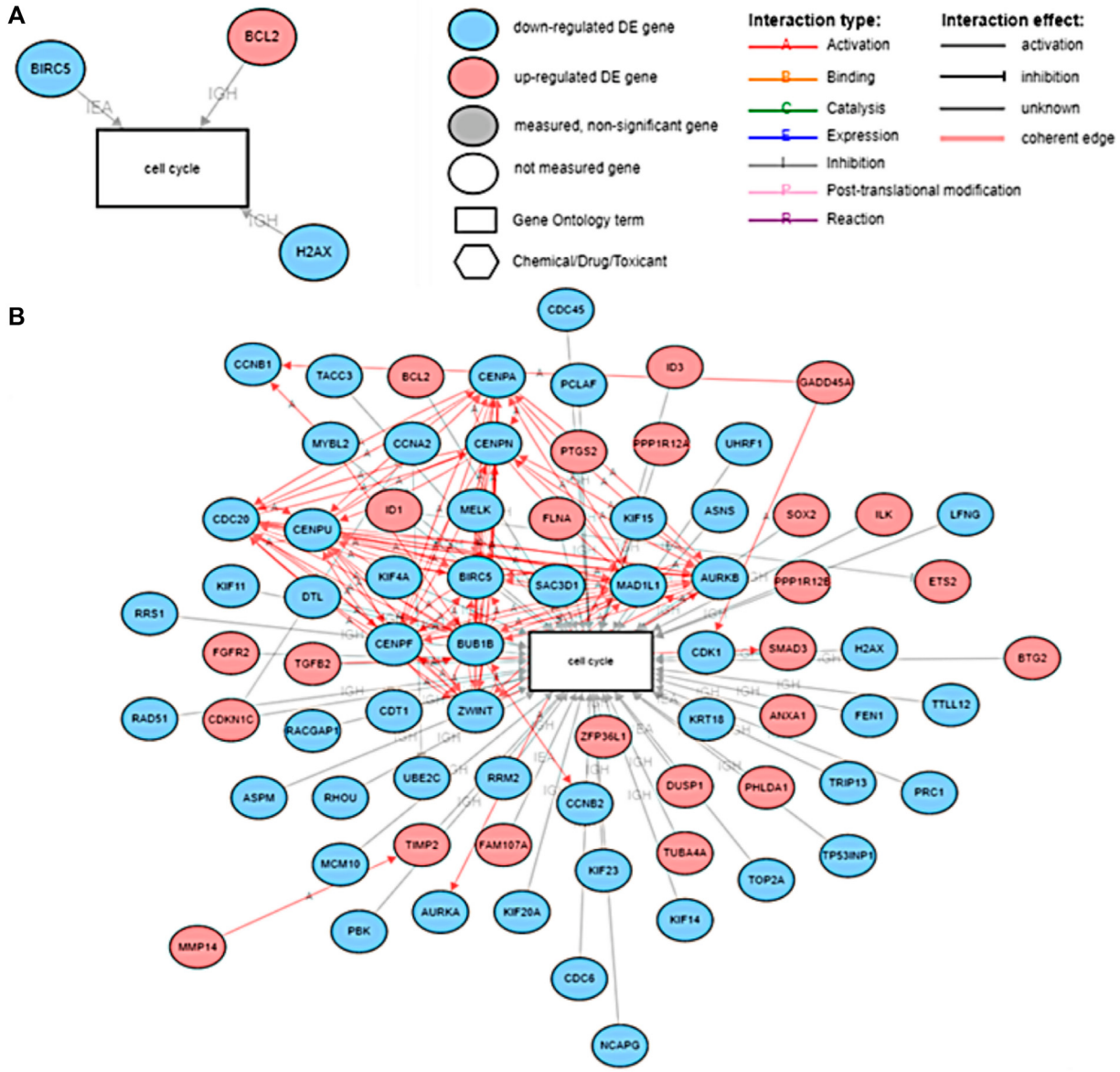


FIGURE 9
(A) The DE genes of prostate cancer associated with the cell cycle downstream of Mitoxantrone and **(B)** The DE genes of prostate cancer associated with the cell cycle downstream of Genistein.

suggesting that natural products can be potential therapeutics for cancer treatment.

Al-Mugotir et al. (2021) listed mitoxantrone as a potential drug for clinical use targeting Topoisomerase II. Siddiqui et al. (2021) showed that mitoxantrone along with imatinib could be used to suppress apoptosis. Their research specifically targeted treatment-resistant HR-proficient cancers. *RAD52*, a protein involved in the HR pathway, was found to be differentially expressed in BRCA-deficient cells. The changes in the expression of the gene *RAD52* is associated with HR activity and hence can affect the way cancer can be treated (Nogueira et al., 2019), (Lok and Powell, 2012). Al-Mugotir et al. (2021) reported that *RAD52* could be a potential target for the HR deficient cancers and further showed the effectiveness of mitoxantrone on such cancers. These findings further strengthen our proposed results of mitoxantrone as a potential candidate for patients with mutations in their HR repair pathways.

In a study conducted on COX-2 inhibitors and breast cancer patients between 1998–2004, it was shown that rofecoxib had the highest percentage (71%, $p < 0.01$) of breast cancer reduction as compared to other drugs including ibuprofen (63%) and 325 mg aspirin (49%). (Harris et al., 2014).

Estradiol is already in use for breast and prostate cancers for palliation therapy.

Figure 7B shows the drugs that were listed as siginificant for the novel biomarkers discovered using cross cancer learning approach by Zhou et al. (2021) Acyline showed as significant drug in our table. In a study conducted by Sofikerim et al. (2007) to find the hormonal predictors of the prostate cancer, follicle-stimulating hormone (FSH) was found to be significantly higher in patients with prostate cancer. Crawford et al. (2017) discussed about evidences of high levels of FSH in the advanced and metastatic prostate cancer. Christenson and Antonarakis. (2018) discussed the use of gonadotropin-releasing hormone (GnRH) agonists to inhibit FSH levels as an initial step once prostate cancer turns metastatic. In the first experiment conducted on humans, Herbst et al., (2002) found that acyline, a novel GnRH antagonist was found to suppress FSH levels. They discussed the use of acyline as a probable prostate cancer drug. O'Toole et al. (2007) discussed the potential use of acyline for breast cancer and prostate cancer. Limonta et al. (2012) discussed GnRH agonists decreasing the tumor growth and proliferation in prostate, ovarian and breast cancers. Genistein, which came up as significant for HR-deficient patients earlier, was listed as significant for cross cancer genes as well and has been discussed earlier.

Currently, there is a strong evidence that the biologically similar cancers have the same underlying genetic aberrations (Risbridger et al., 2010). Hence, providing jointly important treatments could drastically reduce the time invested in development of novel drugs as well as repurposing drugs for diseases separately. Our study exploited the prostate cancer and breast cancer patients with deficiencies in their DNA-repair pathways. There is not clear understanding of DNA repair pathways (excluding HR pathway) involved in the breast and prostate cancer, and hence may require further study. There is a strong evidence that a subset of prostate and breast cancer patients have deficiencies in their HR pathways. The drugs proposed using our approach for this pool of patients have strong evidence from literature and show strong promise.

5 Conclusion

DNA repair pathways are responsible for maintaining the genome stability by performing various mechanisms to reverse the damage caused. Failure to do so may result in various diseases, including cancer. Most malignancies arise from mutations caused by damage to the DNA that was not repaired. While some patients respond to treatments, a subset of patients do not respond to the standard treatments. This clearly concludes that there is heterogeneity within the same type of cancer that needs to be further refined.

In this paper, we identified commonalities and differences among multiple cancers by leveraging the abnormalities within the DNA repair pathways to identify potential drugs through repurposing. Often, a specific drug repurposing approach may not always provide optimal results due to its limitations. Hence, we employed multiple approaches and provided treatment options that were intersecting between the approaches.

Our multi cancer treatment model 1) integrated subsets of patients with common biomarkers in their DNA repair pathways and 2) provided promising drug candidates for patients with different DNA repair deficiencies. The results of the proposed framework can be further utilized as a personalized medicine option for patients who do not respond to regular and organ specific treatment options.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://cancergenome.nih.gov.

Author contributions

SM, EH, and SA. conceived and designed the project. SM performed the experiments. SM, TT, and SA analyzed the data and the results. SM and SA wrote the paper. All authors read and approved the final manuscript.

Funding

This research was funded by the National Science Foundation (NSF: 1948338), the Department of Defense (DoD: W81XWH-21-1-0570), and the National Institutes of Health (NIH: 2P50CA186786-06).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al-Mugotir, M., Lovelace, J. J., George, J., Bessho, M., Pal, D., Struble, L., et al. (2021). Selective killing of homologous recombination-deficient cancer cell lines by inhibitors of the rpa: Rad52 protein-protein interaction. *PloS one* 16, e0248941. doi:10.1371/journal.pone.0248941

Arslanturk, S., Draghici, S., and Nguyen, T. (2020). Integrated cancer subtyping using heterogeneous genome-scale molecular datasets. *Pac. Symp. Biocomput.* 25, 551–562.

Asim, M., Tarish, F., Zecchini, H. I., Sanjiv, K., Gelali, E., Massie, C. E., et al. (2017). Synthetic lethality between androgen receptor signalling and the parp pathway in prostate cancer. *Nat. Commun.* 8, 374. doi:10.1038/s41467-017-00393-y

Banerjee, S., Li, Y., Wang, Z., and Sarkar, F. H. (2008). Multi-targeted therapy of cancer by genistein. *Cancer Lett.* 269, 226–242. doi:10.1016/j.canlet.2008.03.052

Beaver, J. A., Amiri-Kordestani, L., Charlab, R., Chen, W., Palmby, T., Tilley, A., et al. (2015). Fda approval: Palbociclib for the treatment of postmenopausal patients with estrogen receptor–positive, her2-negative metastatic breast cancer. *Clin. Cancer Res.* 21, 4760–4766. doi:10.1158/1078-0432.CCR-15-1185

Brown, A. S., and Patel, C. J. (2018). A review of validation strategies for computational drug repositioning. *Brief. Bioinform.* 19, 174–177. doi:10.1093/bib/bbw110

Chen, Y., Wang, J., Fraig, M. M., Metcalf, J., Turner, W. R., Bissada, N. K., et al. (2001). Defects of dna mismatch repair in human prostate cancer. *Cancer Res.* 61, 4112–4121.

Cheng, A. S., Leung, S. C., Gao, D., Burugu, S., Anurag, M., Ellis, M. J., et al. (2020). Mismatch repair protein loss in breast cancer: Clinicopathological associations in a large British columbia cohort. *Breast Cancer Res. Treat.* 179, 3–10. doi:10.1007/s10549-019-05438-y

Chiang, A. P., and Butte, A. J. (2009). Systematic evaluation of drug–disease relationships to identify leads for novel drug uses. *Clin. Pharmacol. Ther.* 86, 507–510. doi:10.1038/clpt.2009.103

Christenson, E. S., and Antonarakis, E. S. (2018). Parp inhibitors for homologous recombination-deficient prostate cancer. *Expert Opin. Emerg. Drugs* 23, 123–133. doi:10.1080/14728214.2018.1459563

Crawford, E. D., Schally, A. V., Pinthus, J. H., Block, N. L., Rick, F. G., Garnick, M. B., et al. (2017). "The potential role of follicle-stimulating hormone in the cardiovascular, metabolic, skeletal, and cognitive effects associated with androgen deprivation therapy," in *Urologic Oncology: Seminars and original investigations* (Amsterdam, Netherlands: Elsevier), Vol. 35, 183–191.

den Brok, W. D., Schrader, K. A., Sun, S., Tinker, A. V., Zhao, E. Y., Aparicio, S., et al. (2017). Homologous recombination deficiency in breast cancer: A clinical review. *JCO Precis. Oncol.* 1, 1–13. doi:10.1200/PO.16.00031

DiMasi, J. A., Grabowski, H. G., and Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of r&d costs. *J. Health Econ.* 47, 20–33. doi:10.1016/j.jhealeco.2016.01.012

Draghici, S., Nguyen, T.-M., Sonna, L. A., Ziraldo, C., Vanciu, R., Fadel, R., et al. (2020). *Covid-19: Disease pathways and gene expression changes predict methylprednisolone can improve outcome in severe cases*. medRxiv. doi:10.1101/2020.05.06.20076687

Dudley, J. T., Deshpande, T., and Butte, A. J. (2011). Exploiting drug–disease relationships for computational drug repositioning. *Brief. Bioinform.* 12, 303–311. doi:10.1093/bib/bbr013

Fan, S., Meng, Q., Auborn, K., Carter, T., and Rosen, E. (2006). Brca1 and brca2 as molecular targets for phytochemicals indole-3-carbinol and genistein in breast and prostate cancer cells. *Br. J. Cancer* 94, 407–426. doi:10.1038/sj.bjc.6602935

Gottlieb, A., Stein, G. Y., Ruppin, E., and Sharan, R. (2011). Predict: A method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* 7, 496. doi:10.1038/msb.2011.26

Harris, R. E., Casto, B. C., and Harris, Z. M. (2014). Cyclooxygenase-2 and the inflammogenesis of breast cancer. *World J. Clin. Oncol.* 5, 677–692. doi:10.5306/wjco.v5.i4.677

Herbst, K. L., Anawalt, B. D., Amory, J. K., and Bremner, W. J. (2002). Acyline: The first study in humans of a potent, new gonadotropin-releasing hormone antagonist. *J. Clin. Endocrinol. Metab.* 87, 3215–3220. doi:10.1210/jcem.87.7.8675

Hu, G., and Agarwal, P. (2009). Human disease-drug network based on genomic expression profiles. *PloS one* 4, e6536. doi:10.1371/journal.pone.0006536

Jackson, S. P., and Bartek, J. (2009). The dna-damage response in human biology and disease. *Nature* 461, 1071–1078. doi:10.1038/nature08467

Jarada, T. N., Rokne, J. G., and Alhajj, R. (2020). A review of computational drug repositioning: Strategies, approaches, opportunities, challenges, and directions. *J. Cheminform.* 12, 46–23. doi:10.1186/s13321-020-00450-7

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). Kegg as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462. doi:10.1093/nar/gkv1070

Khan, H., Labanca, F., Ullah, H., Hussain, Y., Tzvetkov, N. T., Akkol, E. K., et al. (2021). Advances and challenges in cancer treatment and nutraceutical prevention: The possible role of dietary phenols in brca regulation. *Phytochem. Rev.* 2021, 385–400. doi:10.1007/s11101-021-09771-3

Kim, K.-Y., Han, W., Noh, D. Y., Kang, D., and Kwack, K. (2013). Impact of genetic polymorphisms in base excision repair genes on the risk of breast cancer in a Korean population. *Gene* 532, 192–196. doi:10.1016/j.gene.2013.09.069

Li, X., and Heyer, W.-D. (2008). Homologous recombination in dna repair and dna damage tolerance. *Cell. Res.* 18, 99–113. doi:10.1038/cr.2008.1

Limonta, P., Marelli, M. M., Mai, S., Motta, M., Martini, L., and Moretti, R. M. (2012). Gnrh receptors in cancer: From cell biology to novel targeted therapeutic strategies. *Endocr. Rev.* 33, 784–811. doi:10.1210/er.2012-1014

Lok, B. H., and Powell, S. N. (2012). Molecular pathways: Understanding the role of Rad52 in homologous recombination for therapeutic advancement. *Clin. Cancer Res.* 18, 6400–6406. doi:10.1158/1078-0432.CCR-11-3150

Luo, H., Wang, J., Li, M., Luo, J., Peng, X., Wu, F.-X., et al. (2016). Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm. *Bioinformatics* 32, 2664–2671. doi:10.1093/bioinformatics/btw228

Marshall, C. H., Fu, W., Wang, H., Baras, A. S., Lotan, T. L., and Antonarakis, E. S. (2019). Prevalence of dna repair gene mutations in localized prostate cancer according to clinical and pathologic features: Association of gleason score and tumor stage. *Prostate Cancer Prostatic Dis.* 22, 59–65. doi:10.1038/s41391-018-0086-1

McCain, J. (2015). First-in-class cdk4/6 inhibitor palbociclib could usher in a new wave of combination therapies for hr+, her2- breast cancer. *P T.* 40, 511–520.

Mittal, R. D., Mandal, R. K., and Gangwar, R. (2012). Base excision repair pathway genes polymorphism in prostate and bladder cancer risk in north indian population. *Mech. Ageing Dev.* 133, 127–132. doi:10.1016/j.mad.2011.10.002

Nogueira, A., Fernandes, M., Catarino, R., and Medeiros, R. (2019). Rad52 functions in homologous recombination and its importance on genomic integrity maintenance and cancer therapy. *Cancers* 11, 1622. doi:10.3390/cancers11111622

O'Connor, M. J. (2015). Targeting the dna damage response in cancer. *Mol. Cell.* 60, 547–560. doi:10.1016/j.molcel.2015.10.040

O'Toole, E., Amory, J., Bremner, W., Page, S., Adamczyk, B., Lee, A., et al. (2007). Mer-104 tablets: A dose-ranging study of an oral formulation of a gonadotropin-releasing hormone antagonist, acyline. *Mol. Cancer Ther.* 6, B83.

Peyvandipour, A., Saberian, N., Shafi, A., Donato, M., and Draghici, S. (2018). A novel computational approach for drug repurposing using systems biology. *Bioinformatics* 34, 2817–2825. doi:10.1093/bioinformatics/bty133

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., et al. (2019). Drug repurposing: Progress, challenges and recommendations. *Nat. Rev. Drug Discov.* 18, 41–58. doi:10.1038/nrd.2018.168

Ricken, R., Ulrich, S., Schlattmann, P., and Adli, M. (2017). Tranylcypromine in mind (part ii): Review of clinical pharmacology and meta-analysis of controlled studies in depression. *Eur. Neuropsychopharmacol.* 27, 714–731. doi:10.1016/j.euroneuro.2017.04.003

Risbridger, G. P., Davis, I. D., Birrell, S. N., and Tilley, W. D. (2010). Breast and prostate cancer: More similar than different. *Nat. Rev. Cancer* 10, 205–212. doi:10.1038/nrc2795

Saberian, N., Peyvandipour, A., Donato, M., Ansari, S., and Draghici, S. (2019). A new computational drug repurposing method using established disease–drug pair knowledge. *Bioinformatics* 35, 3672–3678. doi:10.1093/bioinformatics/btz156

Saleh-Gohari, N., and Helleday, T. (2004). Conservative homologous recombination preferentially repairs dna double-strand breaks in the s phase of the cell cycle in human cells. *Nucleic Acids Res.* 32, 3683–3688. doi:10.1093/nar/gkh703

Schuler, J., Falls, Z., Mangione, W., Hudson, M. L., Bruggemann, L., and Samudrala, R. (2022). Evaluating the performance of drug-repurposing technologies. *Drug Discov. Today* 27, 49–64. doi:10.1016/j.drudis.2021.08.002

Siddiqui, A., Tumiati, M., Joko, A., Sandholm, J., Roering, P., Aakko, S., et al. (2021). Targeting dna homologous repair proficiency with concomitant topoisomerase ii and c-abl inhibition. *Front. Oncol.* 11, 733700. doi:10.3389/fonc.2021.733700

Sirota, M., Dudley, J. T., Kim, J., Chiang, A. P., Morgan, A. A., Sweet-Cordero, A., et al. (2011). Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci. Transl. Med.* 3, 96ra77. doi:10.1126/scitranslmed.3001318

Sofikerim, M., Eskicorapcı, S., Oruç, Ö., and Oezen, H. (2007). Hormonal predictors of prostate cancer. *Urol. Int.* 79, 13–18. doi:10.1159/000102906

Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., et al. (2017). A next generation connectivity map: L1000 platform and the first 1, 000, 000 profiles. *Cell.* 171, 1437–1452. doi:10.1016/j.cell.2017.10.049

Sun, D., Gao, W., Hu, H., and Zhou, S. (2022). Why 90% of clinical drug development fails and how to improve it? *Acta Pharm. Sin. B* 12, 3049–3062. doi:10.1016/j.apsb.2022.02.002

Tang, Q., Ma, J., Sun, J., Yang, L., Yang, F., Zhang, W., et al. (2018). Genistein and ag1024 synergistically increase the radiosensitivity of prostate cancer cells. *Oncol. Rep.* 40, 579–588. doi:10.3892/or.2018.6468

Toh, M., and Ngeow, J. (2021). Homologous recombination deficiency: Cancer predispositions and treatment implications. *Oncologist* 26, e1526–e1537. doi:10.1002/onco.13829

Walker, A. J., Wedam, S., Amiri-Kordestani, L., Bloomquist, E., Tang, S., Sridhara, R., et al. (2016). Fda approval of palbociclib in combination with fulvestrant for the treatment of hormone receptor–positive, her2-negative metastatic breast cancer. *Clin. Cancer Res.* 22, 4968–4972. doi:10.1158/1078-0432.CCR-16-0493

Watkins, J. A., Irshad, S., Grigoriadis, A., and Tutt, A. N. (2014). Genomic scars as biomarkers of homologous recombination deficiency and drug response in breast and ovarian cancers. *Breast Cancer Res.* 16, 211–11. doi:10.1186/bcr3670

Xie, M., Hwang, T., and Kuang, R. (2012). "Prioritizing disease genes by bi-random walk," in *Pacific-asia conference on knowledge discovery and data mining* (Berlin, Germany: Springer), 292–303.

Zhang, Z., Wang, C.-Z., Du, G.-J., Qi, L.-W., Calway, T., He, T.-C., et al. (2013). Genistein induces g2/m cell cycle arrest and apoptosis via atm/p53-dependent pathway in human colon cancer cells. *Int. J. Oncol.* 43, 289–296. doi:10.3892/ijo.2013.1946

Zhou, K., Arslanturk, S., Craig, D. B., Heath, E., and Draghici, S. (2021). Discovery of primary prostate cancer biomarkers using cross cancer learning. *Sci. Rep.* 11, 10433. doi:10.1038/s41598-021-89789-x