RESEARCH





Majid Mokhtari^{1*}, Samane Khoshbakht^{1,2}, Mohammad Esmaeil Akbari³ and Sayyed Sajjad Moravveji¹

Abstract

Background In recent years, drug screening has been one of the most significant challenges in the field of personalized medicine, particularly in cancer treatment. However, several new platforms have been introduced to address this issue, providing reliable solutions for personalized drug validation and safety testing. In this study, we developed a personalized drug combination protocol as the primary input to such platforms.

Methods To achieve this, we utilized data from whole-genome expression profiles of 6173 breast cancer patients, 312 healthy individuals, and 691 drugs. Our approach involved developing an individual pattern of perturbed gene expression (IPPGE) for each patient, which was used as the basis for drug selection. An algorithm was designed to extract personalized drug combinations by comparing the IPPGE and drug signatures. Additionally, we employed the concept of drug repurposing, searching for new benefits of existing drugs that may regulate the desired genes.

Results Our study revealed that drug combinations obtained from both specialized and non-specialized cancer medicines were more effective than those extracted from only specialized medicines. Furthermore, we observed that the individual pattern of perturbed gene expression (IPPGE) was unique to each patient, akin to a fingerprint.

Conclusions The personalized drug combination protocol developed in this study offers a methodological interface between drug repurposing and combination drug therapy in cancer treatment. This protocol enables personalized drug combinations to be extracted from hundreds of drugs and thousands of drug combinations, potentially offering more effective treatment options for cancer patients.

Keywords Drug repurposing, Gene expression, Breast cancer, Microarray, Oncology

*Correspondence:

Majid Mokhtari

majid.mokhtari@ut.ac.ir

¹ Department of Bioinformatics, Kish International Campus, University of Tehran, Kish Island, Iran

² Duke Molecular Physiology Institute, Duke University School

of Medicine-Cardiology, Durham, NC 27701, USA

³ Cancer Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Introduction

Precision medicine and personalized medicine are emerging concepts in the field of healthcare that are revolutionizing the way diseases are diagnosed and treated [1]. These approaches can stratify patients and treat them according to their molecular characteristics [2]. Although precision medicine and personalized medicine are two different concepts, they complement each other in their application. Precision medicine is used to develop the accuracy of diagnoses and to refine patient classification



© The Author(s) 2023, corrected publication 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data. [3]. It focuses on detecting the actual state of disease using laboratory tests to better understand pathological mechanisms. On the other hand, personalized medicine is a therapeutic approach that prepares or optimizes a specific therapy for an individual [4–6]. It uses computational frameworks such as drug repurposing, systems biology, and pharmacogenetics to discover appropriate individual treatments [7, 8]. The merging of these two approaches can lead to the era of personalized precision medicine [9, 10].

One of the key elements in precision and personalized medicine is understanding the unique gene expression patterns involved in diseases [11]. These patterns can be detected when studying patient populations, and they can be used to identify the most appropriate medicines that can regulate the genes involved in perturbed expression patterns in disease [12–14]. In particular, drug repurposing studies have shown great promise in accelerating the adoption of existing approved medicines that are in clinical use. This is because these medicines have already passed the rigorous safety tests required by regulatory agencies like the Food and Drug Administration [15–18].

Moreover, studies have shown that combination drug therapy can be more efficient in reducing drug resistance due to the control of parallel biological pathways. However, the heterogeneity of cancer leads to different responses to similar treatments. Therefore, it is essential to identify the gene expression pattern of each patient to determine the most effective combination drug therapy [19, 20].

To this end, we developed a bioinformatics protocol for personalized precision medicine by integrating the concepts of drug repurposing and drug combination. The protocol elucidates one or more drug combinations that are precisely adapted to individual gene expression profiles. This bioinformatics protocol helps the physician identify the most effective drug combination by precisely monitoring the patient's gene expression status.

In this study, we aimed to apply our bioinformatics protocol to breast cancer patients. Firstly, we identified the differentially expressed genes (DEGs) in breast cancer patients compared to healthy controls. Secondly, we calculated a healthy gene expression interval for each DEG by assessing gene expression in healthy individuals. Thirdly, we compared a patient's gene expression profile with the gene health intervals (normal gene expression levels) to create an individual pattern of perturbed gene expression (IPPGE). Fourthly, we identified one or more drug combinations for the patient by simultaneously analyzing the IPPGE and the drug signatures through the use of a network-based algorithm. Finally, we reconstructed a directed differential network (DDN) using biological Page 2 of 15

pathway data to predict the effect on gene expression of the identified drug combinations (Fig. 1).

In summary, the aim of this study was to demonstrate the feasibility and effectiveness of our bioinformatics protocol for personalized precision medicine. We believe that this protocol can be applied to other diseases and can contribute to the development of personalized precision medicine in the near future.

Methods

The BMC3PM protocol was developed to extract a personalized combination of drugs that could effectively treat a patient's disease. BMC3PM is the integration of previously established methods along with the construction of new bioinformatics tools (Fig. 1).

Data sets and Metadata information

The study analyzed gene expression data from 6173 untreated breast cancer patients and 312 healthy controls. The raw expression files (CEL files) for breast cancer patients and healthy controls were obtained using the GEOquery package in R, and all datasets were from the Affymetrix platform [21]. Supplementary Table S1 presents the sample size and platform information for each dataset. The study also used 691 drug perturbation gene expression profiles, which were downloaded from the Comparative Molecular Profiles of Approved Drugs (CMAP) databases and the National Center for Biotechnology Information (NCBI) databases (Supplementary Table S2).

Pre-processing and normalization

Frozen robust multiarray analysis (fRMA) was used for background correction and normalization of the datasets. Unlike RMA normalization, fRMA is more suitable for clinical settings and can be applied to individual or small batch normalization [22]. The dataset was adjusted for downstream analyses by removing batch effects and unwanted variation among the GSEs using a combat function [23].

Deregulated genes identification

After data pre-processing, differentially expressed genes (DEGs) were identified using the Limma package in R [24]. Deregulated genes were divided into two distinct groups: upregulated genes (URGs) and downregulated genes (DRGs) based on the LogFC. The LogFC of URGs was defined as greater than 0.45, and the LogFC of DRGs was defined as less than -0.45. The health intervals for each gene in the URG and DRG groups were calculated. The health interval refers to the range of gene expression in which members of the disease group are rarely seen within it [25] (Fig. 2 A and B).



Fig. 1 BMC3PM workflow. Bioinformatics Multidrug Combination Protocol for Personalized Precision Medicine (BMC3PM) is a set of various methods used in a bioinformatics protocol to extract personalized drug combinations



Fig. 2 Data processing and analysis concepts. A The concept of the health interval for the upregulated genes. B The concept of the health interval for the downregulated genes. C The concept of the differential network is illustrated in cancer computational biology

Upregulated Gene Health Interval (URGHI)

The quartiles of gene expression for the patient and control groups were calculated separately (without outliers) to identify the upregulated gene health intervals. Two main conditions were used to assess the quartiles of gene expression for the patient and control groups. Under the first condition, some genes' quartiles did not show overlap between the patient and control groups. Therefore, the range of gene expression between the minimum and maximum in the control group was selected as the URGHI. Under the second condition, some genes' quartiles showed overlap between the patient and control groups. Therefore, the range of gene expression between the minimum and third quartile in the control group was selected as the URGHI (Fig. 2A).

Downregulated Gene Health Interval (DRGHI)

Two observed conditions were used to determine the downregulated gene health interval. Similar to the first condition for URGHI, the quartiles of some genes did not show overlap between the patient and control groups. Therefore, the range of gene expression between the minimum and maximum in the control group was selected as the DRGHI. In the second condition, the quartiles of some genes showed overlap between the patient and control groups. Therefore, the range of gene expression between the first quartile and maximum in the control group was selected as the DRGHI (Fig. 2B).

Individual Pattern of Perturbed Gene Expression (IPPGE)

Each patient's disease-leading genes (individual deregulated genes) were identified using the health intervals calculated in the previous step and the individual's gene expression profile. Accordingly, each gene that was not within the health interval was determined as an individual dysregulation gene for a given individual. The simultaneous expression of these individual deregulated genes in each patient forms that patient's IPPGE.

Primary Health Matrix (PHM)

The primary health matrix (PHM) is a binary matrix that serves as a tool to identify deregulated genes and assess the effectiveness of drugs in restoring their function to a healthy state. The PHM was created through the synchronization of the Institute of Public Health Post-Graduate Program (IPPGE) and Comparative Molecular Profiles of Approved Drugs (CMAP) databases [26] with the health interval. The PHM has the CMAP drugs as columns and individual deregulated genes as rows. A PHM value of 1 signifies that the gene falls within the healthy interval post-administration of the drug.

Drug Combination Algorithm (DC algorithm)

The Drug Combination Algorithm (DC algorithm) is an algorithm that uses the PHM to determine the most appropriate drug combination for an individual patient. The algorithm has been implemented in several steps (Supplementary Table S3). Firstly, the drug that moved the most genes into their respective health intervals was selected. If several drugs equally moved the most genes into health intervals, the drug with the fewest gene expression interactions among drugs (GEIADs) was selected. GEIADs mean that a chosen drug along with previously selected drugs simultaneously affects a gene. Although the lack of GEIADs is inevitable, the algorithm can detect the fewest number of GEIADs. To identify GEIADs, we reconstructed a drug-drug network in which the drugs were defined as the nodes and GEI-ADs as edges. The network is detailed in Supplementary Table S3. If several drugs had the least GEIADs in equal quantity, we introduced separate drug combinations in parallel. Finally, the algorithm was completed when most genes were within health intervals with the least identified GEIADs. Pseudocodes for the DC algorithm can be found in Supplementary Table S3.

Directed network reconstruction

To have a reference signaling pathway, all KGML Homo sapiens files were downloaded from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and merged to create a directed network using GraphPad and the Igraph package in R [27–29]. KEGG node identifiers were annotated as gene symbols.

Gene coexpression network reconstruction

The gene coexpression networks were individually reconstructed for patients and healthy controls using the Spearman correlation. To create a differential signaling network, the directed differential coexpression network for two conditions, disease vs healthy, was detected and mapped to the reference signaling network (Fig. 2C).

Identification of downstream targets

To identify the downstream targets of drug combinations extracted by the BMC3PM algorithm, the gene targets of the drug combinations for different patients were mapped to the differential signaling network [30]. This approach allowed us to identify the genes and pathways affected by the drug combinations and further understand the mechanism of action of these drugs.

Visualization

All network visualizations were implemented in Cytoscape software [31], which allows for the interactive exploration and analysis of complex networks. The visualizations of the directed network and the gene coexpression network provided a comprehensive view of the molecular pathways involved in disease and the effects of drugs on these pathways.

Results

Individual Pattern of Perturbed Gene Expression (IPPGE)

A total of 225 genes analysis were used to construct the individual pattern of perturbed gene expression (Fig. 3A). The gene clustering and functional enrichment results were illustrated in Fig. 3B and C. A gene health interval was determined for each DEG based on comparing the gene expression between the healthy and the cancer groups. The results showed that each DEG had a unique health interval. A number of pivotal cancer-related genes, including HER2, CCND1, PIK3CA, and CDKN2A, were pinpointed within the top 500 genes in both the differential signaling network (Fig. 3D) and the analysis of differentially expressed genes. Remarkably, there were 225 genes that appeared in both analyses. As a result, certain well-known genes were left out of the study because they were not found in both analyses. The importance of genes identified as biomarkers in biological databases and previous studies is highlighted in Supplementary Table S4. Within the patient population, it was observed that the mean gene expression levels of all DEGs fell outside the health interval, as depicted in Fig. 4A. Conversely, in the healthy population, the mean gene expression of all DEGs resided within the health interval, as illustrated in



Fig. 3 Data processing and analysis. **A** The DEGs detected from the merged data are illustrated; the total 225 genes consisted of 130 downregulated and 95 upregulated genes. These genes were used to construct the IPPGEs. **B** The heatmap shows the DEG clustering in the patient and healthy samples. The Manhattan method was used for row and column clustering. **C** The significant pathways are illustrated. The numbers on the bars indicated the minus logarithm of q-value. **D** The induced signaling pathway which was extracted from the KEGG database using the DEGs



Fig. 4 Expression pattern. A The expression pattern of DEGs in the population. B The expression pattern of DEGs in individual patients

Fig. 4B. In Fig. 4, each cell symbolizes one of 225 DEGs in the pattern (Fig. 3A).

Personalized studies require a unique patient profile; for this purpose, four patients in two subtypes were studied, luminal A (ER+/PR+/HER2-) and triple-negative (ER-/PR-/HER2-). The basis for constructing an IPPGE for each patient was to compare the patient's gene expression with each DEG's health interval. Four patients were chosen at random in each subtype further to illustrate the patterns' results in this paper. The patients' gene expression patterns showed that while the first and second patients were the same in terms of subtype (ER+/PR+/ HER2+), they differed in terms of the observed IPPGE. In patient 1, 181 genes were observed to be outside of the health interval, while in patient 2, 154 genes were outside of the health interval. The third and fourth patients also had the same subtype (ER-/PR-/HER2-). In patient 3, 165 genes were outside of the health, while 169 genes were outside of the health interval in patient 4. Notably, sometimes, two patients observed an equal number of genes, but various genes were identified. Therefore, not only a significant difference was observed between the numbers of genes outside of the health interval among patients, but there was also a difference in the type of genes observed. The results indicated that patients could have very different IPPGEs despite having the same cancer subtype. Most genes in each patient were outside of the health interval, but the patients had some genes within the health interval (Fig. 4B). This study examined 6173 patients and 312 healthy people to determine their IPPGE. The results indicated that each individual had a particular IPPGE, like a fingerprint, and no two IPPGEs were identical. For some samples, the expression of 31 genes could not be assessed due to differences in the

number of genes on the different platforms used in the laboratories.

Drug combinations

The IPPGEs of the abovementioned four patients and connectivity map (CMAP) data were used as input to our designed protocol to extract personalized drug combinations. The CMAP database reports the effect of drugs on gene expression. Any drug or drug combination that could affect several IPPGE genes to bring their expression back into the health interval was extracted, forming a personalized drug combination. Given the possibility of obtaining better comparisons between breast cancer-specific drugs and between nonspecific drugs, drug combinations were extracted from these two major drug groups. The first group consisted of drugs approved for breast cancer, called the Alpha group. The second group included the Alpha group drugs and all other FDAapproved drugs and was called the Beta group. Overall, 27 combinations of drugs were extracted across the two groups for the four patients. The protocol extracted 3, 5, 3, and 2 personalized drug combinations from the Alpha group for the first, second, third, and fourth patients, respectively. From the Beta group, 3, 5, 3, and 3 personalized drug combinations were extracted for the first through fourth patients, respectively (Table 1). The third drug combination from the Alpha group for the first patient consisted of four drugs, but all of the other combinations included five drugs.

Remarkably, each patient had a different number of genes affected by one drug. For example, gene data computations made by the algorithm indicate that in patient 1, doxorubicin affected 26 genes, while in patient 2, it affected 18 genes. In patient 2, the eight genes that doxorubicin might affect were already in the healthy interval before taking the drug. As a result, the use of this drug in patient 2 would not be efficient because it would not affect those eight genes. The difference in the effect of a drug on two separate patients was due to differences in each patient's IPPGE (Table 1).

Table 1 Extracted personalized drug combinations. Only specialized cancer drugs were included in the Alpha group and, thus, in all drug combinations extracted from this group, while the Beta group contained both specialized and nonspecialized drugs. Nonspecialized drugs are marked with †. The numbers within the Bracket display the number of perturbed genes for each patient that were influenced by the drug

Drug Group	Patient Label	Combination Number	First Drug	Second Drug	Third Drug	Fourth Drug	Fifth Drug	Number of Genes Affected by the Combination	Number of Up/Down Genes	Percentage of Combination Effect	Drug Interactions
Alpha	Patient 1	1	Doxorubicin [26]	Lapatinib [16]	Tamoxifen [12]	Vinblastine [7]	Methotrexate [1]	62	181	34.3	Monitor Closely
		2	Doxorubicin [26]	Tamoxifen [12]	Cyclophosphamide[9]	Epirubicin [8]	Paclitaxel [6]	61	181	33.7	Monitor Closely
		3	Doxorubicin [26]	Megestrol [17]	Tamoxifen [12]	Raloxifene [4]		59	181	32.6	NO
	Patient 2	4	Doxorubicin [18]	Megestrol [10]	Raloxifene [10]	Tamoxifen [7]	Fulvestrant [2]	47	154	30.5	NO
		5	Doxorubicin [18]	Megestrol [10]	Tamoxifen [7]	Methotrexate [2]	Fulvestrant [2]	39	154	25.3	NO
		6	Doxorubicin [18]	Paclitaxel [9]	Epirubicin [4]	Cyclophosphamide [3]	Letrozole [3]	37	154	24.0	Monitor Closely
		7	Doxorubicin [18]	Tamoxifen [7]	Lapatinib [4]	Epirubicin [4]	Cyclophosphamide [3]	36	154	23.4	Monitor Closely
		8	Doxorubicin [18]	Epirubicin [4]	Cyclophosphamide [3]	Fulvestrant [3]	Vinblastine [2]	30	154	19.5	Monitor Closely
	Patient 3	9	Tamoxifen [24]	Exemestane [11]	Megestrol [6]	Fulvestrant [3]	Vinblastine [2]	46	165	27.9	NO
		10	Tamoxifen [24]	Exemestane [11]	Megestrol [6]	Methotrexate [3]	Foretinib [1]	45	165	27.3	NO
		11	Tamoxifen [24]	Exemestane [11]	Epirubicin [3]	Doxorubicin [2]	Cyclophosphamide [1]	40	165	24.2	Monitor Closely
	Patient 4	12	Tamoxifen [24]	Exemestane [13]	Fulvestrant [3]	Vinblastine [3]	Methotrexate [1]	44	169	26.0	NO
		13	Tamoxifen [24]	Exemestane [13]	Epirubicin [3]	Doxorubicin [2]	Cyclophosphamide [1]	43	169	25.4	Monitor Closely
Beta	Patient 1	14	Ciclopirox [28] †	Tacrolimus [14] †	Mesalazine [8] †	Benperidol [5] †	Vorinostat [4] †	60	181	33.1	NO
		15	Ciclopirox [28] †	Tacrolimus [14] †	Atovaquone [9] †	Irinotecan [3]	Albendazole [2] †	56	181	30.9	Monitor Closely
		16	Ciclopirox [28] †	Tamoxifen [14]	Doxorubicin [3]	Epirubicin [2]	Cyclophosphamide [2]	51	181	28.2	Monitor Closely
	Patient 2	17	Doxorubicin [18]	Mafenide [7] †	Tamoxifen [6]	Irinotecan [3]	Ciclopirox [3] †	37	154	24.0	NO
		18	Doxorubicin [18]	Mafenide [7] †	Celecoxib [6] †	Ciclopirox [4] +	verteporfin [2] †	37	154	24.0	NO
		19	Doxorubicin [18]	Mafenide [7] †	Celecoxib [6] †	Ciclopirox [4] †	cyclopenthiazide [2]	37	154	24.0	Monitor Closely
		20	Doxorubicin [18]	Mafenide [7] †	Celecoxib [6] †	Cyclophosphamide [3]	Epirubicin [2]	36	154	23.4	Monitor Closely
		21	Doxorubicin [18]	Mafenide [7] †	Tacrolimus [5] †	metformin [3] †	tretinoin [2] †	35	154	22.7	Monitor Closely
	Patient 3	22	Tamoxifen [24]	Mesalazine [14] †	Vorinostat [8] †	pantothenic acid [5] †	tretinoin [2] †	53	165	32.1	NO
		23	Tamoxifen [24]	Mesalazine [14] †	Vorinostat [8] †	Atovaquone [3] †	tretinoin [2] †	51	165	30.9	NO
		24	Tacrolimus [15]	Mesalazine [16] †	Vorinostat [9] †	pantothenic acid [6] †	Ciclopirox [3] †	49	165	29.7	NO
	Patient 4	25	Mesalazine [26] †	Tacrolimus [14] †	Vorinostat [8] †	Benperidol [3] †	Ciclopirox [3] †	52	169	30.8	NO
		26	Mesalazine [26] †	Tacrolimus [14] †	Rifabutin [8] †	Digoxin [3] †	Deferoxamine [2] †	51	169	30.2	Serious
		27	Mesalazine [26] †	Tamoxifen [14]	Doxorubicin [3]	Epirubicin [2]	Cyclophosphamide [2]	45	169	26.6	Monitor Closely

Another intriguing observation was that after the first drug was selected, the second drug's effect under the influence of the first drug was altered. For example, tamoxifen was selected as the first drug in patient 4 in drug combination No. 13; it affected 24 genes and brought the patient's gene expression into the health interval. In comparison, in combination No. 27, tamoxifen was selected as the second drug in the same patient and was effective for 14 genes. Mesalazine, which affected 26 genes, was selected as the first drug in combination No. 27; however, tamoxifen could affect 10 of 26 genes. Consequently, genes that are jointly affected by drugs are calculated only in the drug selected in the initial step. In this way, the number of genes in a drug combination affected by each drug does not include duplicate genes (Table 1).

IPPGEs to assess drug combinations

The IPPGE of patients was analyzed computationally to understand the effect of the extracted drug combinations. The results showed that personalized drug combinations extracted from the Beta group were more effective than combinations extracted from the Alpha group in the second, third, and fourth patients even though all Alpha group drugs were approved breast cancer medications. Additionally, in patient 1, the most effective drug combination was from the Alpha group; however, the results showed that the combinations extracted from the Beta group had a similar effect as those from the Alpha group. The Drug Combinations Profile illustrates the effects of drug combinations on patient gene expression (Fig. 5A, B, C, D, E, F, G, and H). Moreover, Basing our analysis on DrugBank information, we classified drug interactions' side effects into three categories: serious, closely monitored, and no side effects. The outcomes demonstrated that the Beta group exhibited considerably fewer instances of side effects and drug interactions when compared to the Alpha group.

Biological pathways to assess drug combinations

To investigate the effect of the extracted drug combinations on biological pathways, a differential network was reconstructed from the coexpression network, which included DEGs from both healthy controls and patients (Fig. 2C). The differential network was reconstructed from an absolute weight value greater than 6.5 with 269 genes as the network node and 422 edges as the gene relationship. Using KEGG pathway data, 166 nodes and 317 edges of the differential network were able to reconstruct the directed network (Fig. 3D). For example, directional network studies for drug combinations for the four patients showed that the Beta group drug combinations were more effective than the Alpha group combinations. More efficient combinations were extracted from the Beta group, as it contains approved breast cancer-specific drugs and nonspecialized drugs. Drug combination No. 22 extracted from the Beta group for patient number 3 included tamoxifen, which is used in hormone therapy for breast cancer, and it is also included in the Alpha group. Combination No. 22 affected 44.57% of the breast cancer directional network. In contrast, drug combination No. 9 extracted from the Alpha group for the same patient could affect 37.34% of the network (Fig. 6A and B). We observed that different drugs from the same class of drugs can affect different genes in the same pathway.

The results showed that the unique IPPGEs of the patients caused different effects of each drug; thus, it was necessary to use drug combinations that were extracted based on the patient's IPPGE to achieve the best drug effect. Although some of the drug combinations contained a joint drug, these drugs had different effects when used in different combinations. For example, ciclopirox was included in drug combinations extracted from the Beta group for patients 1, 2, and 4. Additionally, in the drug combinations extracted from the Alpha group, tamoxifen was included for all four patients. In the Beta group combinations, tamoxifen was included for patients 2 and 3 (Figs. 5, 6, and Table 1).

Discussion

Before a new drug treatment can be employed in clinical settings, it undergoes rigorous testing and evaluation multiple times to secure approval from health organizations. However, when it comes to personalized medicine tailored for a single patient, its impact is profound for that individual alone. Consequently, the standard approval processes, which rely on tests conducted on different organisms or other individuals, are not suitable for therapies designed exclusively for one person. A recent study by Pauli and colleagues introduced a platform to address this challenge. For drug treatment validation and safety testing, this platform first applies patient-derived tumour organoid (PDTO) cultures followed by patientderived xenograft (PDX) models [32-35]. However, analyzing the hundreds of drugs and thousands of possible drug combinations using this platform remains a considerable and complex challenge [36]. Our new bioinformatics protocol, called BMC3PM, enabled us to extract personalized drug options that could potentially be introduced as the primary input for personalized drug treatment validation platforms.

Designing IPPGE with a personalized medicine approach

In the present study, we created an IPPGE for each patient by comparing the gene expression interval of healthy individuals and breast cancer patients. The



Combinations of specialized and non-specialized breast cancer drugs (Beta Drug Group): B, D, F, H

Fig. 5 The effect of drug combinations on IPPGE. The effects of drug combinations on the gene expression of patients are illustrated. A, C, E, and G indicate the effect of specialized breast cancer-related drugs reported in Table 1 as the Alpha group. The effect of drug combinations extracted from the Beta group of drugs is shown in B, D, F, and H. The Beta group includes specialized and nonspecialized cancer drugs



Fig. 6 The effect of drug combinations on the directed differential coexpression network. This figure indicates the effect of the drug combinations on the signaling network. A Illustration of the effects of drug combinations from the Alpha group, as reported in Table 1. The Alpha group contains specialized breast cancer-related drugs. B The effects of the Beta group drug combinations, as reported in Table 1. Specialized and nonspecialized cancer drugs are included in the Beta group

resulting IPPGE was unique to each patient, like a fingerprint (Fig. 4A and B). A previous study reported variations in gene expression patterns in different cancer phenotypes and also in patients [37–39]. Importantly, although patients have many common perturbed genes, gene expression compared to the health interval plays a pivotal role in drug selection [17]. Thus, the drugs and their gene targets in this study were selected in such a way as to target the most perturbed genes and not to have a useless effect on genes that were within the health interval. However, the personalized drug combinations were observed to have drugs in common among the patients to control common perturbed genes (Table 1). Therefore, each patient's unique IPPGE appeared to be effective in extracting accurate personalized drug combinations.

Personalized medicine approach to drug combinations

Cancers are complex diseases regulated by the interaction of multiple signaling pathways through crosstalk. A single drug is thought to be capable of targeting only one signaling pathway for a disease; however, an alternative signaling pathway can be activated to maintain tumor development. A combination of drugs has been recommended to prevent drug resistance and to make the treatment more effective [40-43]. The current study also found that each medication affected a certain number of a patient's DEG genes; the algorithm extracted complementary drugs that affected the largest number of DEGgenes. We found that drug combinations derived from patient IPPGEs had stronger treatment potential due to their more targeted effects on DEG genes. In addition to the drug combination extracted for each patient, the effect of each drug alone was also recorded (Table 1). The use of drug combinations with the personalized medicine approach can lead to the identification of drug combinations that have the potential to produce a more significant effect in the patient.

Personalized medicine approach to drug repurposing

Several drug repurposing studies have reported significant anticancer efficacy for nonspecialized drugs. One of the first drug repurposing studies showed that the antiulcer drug cimetidine to be a therapeutic candidate for the treatment of adenocarcinoma of the lung [44–46]. Subsequent studies have found that combination drug therapy increases the success of drug repurposing [15, 25, 47, 48]. One goal of the current study was to create a personalized medicine approach for drug repurposing. We found that drug combinations extracted from the Beta group had an equal or greater potential than those from the Alpha group for patient treatment (Figs. 5, 6, and Table 1). This finding has two distinct interpretations. First, nonspecialized drugs may be used as adjunctive treatments in addition to specialized medications. Tamoxifen, which is used for hormone therapy in breast cancer, was the principal drug in combination No. 17. Mesalazine, vorinostat, pantothenic acid, and tretinoin were also included in this combination, and these drugs had the highest therapeutic potential as adjuvant therapy to tamoxifen for patient 3. Although these four drugs are not cancer-specific, previous reports have confirmed their anticancer effects [49–52]. Drug combinations 18 to 24 were similar to this type. These findings indicate that the therapeutic effect of the extracted personalized drug combinations can contain specialized cancer drugs in combination with nonspecialized drugs.

Second, according to the observations, some nonspecialized cancer drugs were identified as the potential main treatment. For example, ciclopirox was the first drug extracted in combinations 14, 15, and 16 for the first patient. However, this drug is a synthetic antifungal, and it is necessary to explain why none of the specific cancer drugs were found in combination No. 14. By contrast, combinations 15 and 16 contained tamoxifen and irinotecan, which are routinely used in cancer treatment [13, 53, 54]. Mesalazine was the first drug for the fourth patient in combinations 25 to 27; this drug is used to treat inflammatory bowel disease. Several other studies have reported the anticancer effects of ciclopirox and mesalazine [49, 55, 56]. Therefore, the main drug that is selected in personalized medicine may not always be a specific cancer drug-it can be extremely different based on the patient's IPPGE. The present study shows that the IPPGE can be highly unique to each patient, and the IPPGE can play an important role in extracting a personalized combination of drugs. Notably, personalized drugs with high therapeutic potential for a particular patient can include nonspecialized drugs.

In this study, we used FDA-approved drugs to assess the safety and interactions of the drug combinations, but there is great therapeutic potential among other small molecules for use as treatment options to extract future personalized drug combinations if reliable mechanisms are identified to assess drug safety and identify their interactions. This research is confined to the realm of bioinformatics, employing sophisticated in-silico methodologies for the analysis and prediction of outcomes. While our study rigorously explores computational algorithms and utilizes diverse datasets to generate insights, it is imperative to acknowledge the inherent limitations. The findings presented here have not been confirmed through laboratory experiments, thus warranting cautious interpretation. In subsequent phases of this research, comprehensive laboratory validations are indispensable to affirm the reliability and applicability of our

computational predictions. These experiments are crucial steps forward, bridging the gap between theoretical analyses and practical applications, and are fundamental for the advancement of this study into tangible real-world solutions.

Conclusion

The BMC3PM protocol was constructed in this research to extract personalized drug combinations from hundreds of drugs and thousands of drug combination. The combinations extracted included both specialized and nonspecialized cancer medications. These drug combinations can be used as the primary input for personalized medicine platforms, such as that of Pauli et al. The BMC3PM protocol can be used as a methodological interface between drug repurposing activities and combination drug therapy in a personalized precision medicine approach.

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s12920-023-01745-y.

Additional file 1: Table S1. Samples characteristics.

Additional file 2: Table S2. Drugs and datasets information.

Additional file 3: Table S3. Algorithms' psedocode.

Additional file 4: Table S4. IPPGE gene characterization.

Acknowledgements

We would like to express our gratitude for the support and advice we received during this research from Dr. Ali Masudi-Nejad's laboratory, and no other substantial contributions were received from non-authors.

Authors' contributions

M.M. contributed significantly to the study's conceptualization, formal analysis, methodology, and original draft writing, and also participated in the manuscript's revision and editing. S.K. contributed to the methodology of the study, as well as the review and editing of the manuscript. M.A. provided guidance and supervision in the conceptualization of the study. S.M. contributed to the review and editing of the manuscript. All authors reviewed and approved the final manuscript.

Funding

The authors did not receive support from any organization for the submitted work.

Availability of data and materials

A majority of the data generated or analyzed during this study are included in this published article. The datasets generated and/or analyzed during this study can be obtained from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing Interests

The authors declare no competing interests.

Received: 11 May 2023 Accepted: 20 November 2023 Published online: 12 December 2023

References

- 1. Valdes R, Yin DT. Fundamentals of pharmacogenetics in personalized precision medicine. Clin Lab Med. 2016;36(3):447–59.
- Fountzilas E, Tsimberidou AM, Vo HH, Kurzrock R. Clinical trial design in the era of precision medicine. Genome Med. 2022;14(1):101. https://doi. org/10.1186/s13073-022-01102-1.
- Xu H, Jiao D, Liu A, Wu K. Tumor organoids: applications in cancer modeling and potentials in precision medicine. J Hematol Oncol J Hematol Oncol. 2022; 15 (1): 58. https://doi.org/10.1186/s13045-022-01278-4.
- Popova A A, Levkin PA. Precision medicine in oncology: in vitro Drug Sensitivity and Resistance Test (DSRT) for selection of personalized anticancer therapy. Adv Ther 2020; 3(2):1900100. https://doi.org/10.1002/adtp.20190 0100.
- Ho D, Quake SR, McCabe ERB, Chng WJ, Chow EK, Ding X, Gelb BD, Ginsburg GS, Hassenstab J, Ho C-M, Mobley WC, Nolan GP, Rosen ST, Tan P, Yen Y, Zarrinpar A. Enabling technologies for personalized and precision medicine. Trends Biotechnol. 2020;38(5):497–518. https://doi.org/10. 1016/j.tibtech.2019.12.021.
- Knowles L, Luth W, Bubela T. Paving the road to personalized medicine: recommendations on regulatory, intellectual property and reimbursement challenges. J Law Biosci. 2017;4(3):453–506. https://doi.org/10. 1093/jlb/lsx030.
- Juengst E, McGowan ML, Fishman JR, Settersten RA. From, "Personalized" to "Precision" Medicine: the ethical and social implications of rhetorical reform in genomic medicine. Hastings Cent Rep. 2016;46(5):21–33. https://doi.org/10.1002/hast.614.
- Klein ME, Parvez MM, Shin J-G. Clinical implementation of pharmacogenomics for personalized precision medicine: barriers and solutions. J Pharm Sci. 2017;106(9):2368–79. https://doi.org/10.1016/j.xphs.2017.04. 051.
- Faulkner E, Holtorf AP, Walton S, Liu C. Y, Lin H, Biltaj E, Brixner D Barr C, Oberg J, Shandhu G, Siebert U, Snyder S. R, Tiwana S, Watkins J, Uzerman MJ, Payne K. Being precise about precision medicine: what should value frameworks incorporate to address precision medicine? A report of the personalized precision medicine special interest group. Value Health. 2020; 23 (5), 529–539. https://doi.org/10.1016/j.jval.2019.11.010.
- Carretero-Puche C, García-Martín S, García-Carbonero R, Gómez-López G, Al-Shahrour F. How can bioinformatics contribute to the routine application of personalized precision medicine? Expert Rev Precis Med Drug Dev. 2020;5(3):115–7. https://doi.org/10.1080/23808993.2020.1758062.
- Freimer JW, Shaked O, Naqvi S, Sinnott-Armstrong N, Kathiria A, Garrido CM, Chen AF, Cortez JT, Greenleaf WJ, Pritchard JK, Marson A. Systematic discovery and perturbation of regulatory genes in human T cells reveals the architecture of immune networks. Nat Genet. 2022;54(8):1133–44. https://doi.org/10.1038/s41588-022-01106-y.
- del Sol A, Balling R, Hood L, Galas D. Diseases as network perturbations. Curr Opin Biotechnol. 2010;21(4):566–71. https://doi.org/10.1016/j. copbio.2010.07.010.
- Moravveji SS, Khoshbakht S, Mokhtari M, Salimi M, Lanjanian H, Nematzadeh S, Torkamanian-Afshar M, Masoudi-Nejad A. Impact of 5HydroxyMethylCytosine (5hmC) on Reverse/Direct association of cellcycle, apoptosis, and extracellular matrix pathways in gastrointestinal cancers. BMC Genomic Data. 2022;23(1):49. https://doi.org/10.1186/ s12863-022-01061-x.
- Aquilina-Reid C, Brennan S, Curry-Hyde A, Teunisse GM. The Nygc Als consortium, null; Janitz, M. Circular RNA expression and interaction patterns are perturbed in amyotrophic lateral sclerosis. Int J Mol Sci. 2022; 23(23):14665. https://doi.org/10.3390/ijms232314665.
- Sun W, Sanderson PE, Zheng W. Drug combination therapy increases successful drug repositioning. Drug Discov Today. 2016;21(7):1189–95. https://doi.org/10.1016/j.drudis.2016.05.015.

- Singhal S, Maheshwari P, Krishnamurthy PT, Patil VM. Drug repurposing strategies for non-cancer to cancer therapeutics. Anticancer Agents Med Chem. 2022;22(15):2726–56. https://doi.org/10.2174/187152062266622 0317140557.
- Moravveji SS, Khoshbakht S, Mokhtari M, Salimi M, Masoudi-Nejad A. Pan-cancer analysis of biological events on cell cycle instability in gastrointestinal cancers with integrative scoring method. Genomics. 2022;114(1):253–65. https://doi.org/10.1016/j.ygeno.2021.12.005.
- Ziyaei K, Ataie Z, Mokhtari M, Adrah K, Daneshmehr MA. An insight to the therapeutic potential of algae-derived sulfated polysaccharides and polyunsaturated fatty acids: focusing on the COVID-19. Int J Biol Macromol. 2022;209:244–57. https://doi.org/10.1016/j.ijbiomac.2022.03.063.
- Li YY, Jones SJ. Drug repositioning for personalized medicine. Genome Med. 2012;4(3):27. https://doi.org/10.1186/gm326.
- Ghosh D, Sehgal K, Sodnar B, Bhosale N, Sarmah D, Datta A, Chaudhary A, Kalia K, Wang X, Bhattacharya P. Drug repurposing for stroke intervention. Drug Discov Today. 2022;27(7):1974–82. https://doi.org/10.1016/j.drudis. 2022.03.003.
- 21. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846–7. https://doi.org/10.1093/bioinformatics/btm254.
- McCall MN, Bolstad BM, Irizarry RA. Frozen Robust Multiarray Analysis (FRMA). Biostatistics. 2010;11(2):242–53. https://doi.org/10.1093/biostatist ics/kxp059.
- Walsh CJ, Hu P, Batt J, Santos CCD. Microarray meta-analysis and crossplatform normalization: integrative genomics for robust biomarker discovery. Microarrays. 2015;4(3):389–406. https://doi.org/10.3390/micro arrays4030389.
- Smyth GK. Limma: linear models for microarray data. In bioinformatics and computational biology solutions using R and Bioconductor; Gentleman, R., Carey, V. J., Huber, W., Irizarry, R. A., Dudoit, S., Eds.; Statistics for biology and health; New York : Springer; 2005. pp 397–420. https://doi. org/10.1007/0-387-29362-0_23.
- Khoshbakht S, Mokhtari M, Moravveji SS, Azimzadeh Jamalkandi S, Masoudi-Nejad A. Re-wiring and gene expression changes of AC025034.1 and ATP2B1 play complex roles in early-to-late breast cancer progression. BMC Genomic Data. 2022; 23(1):6. https://doi.org/10.1186/ s12863-021-01015-9.
- Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP. Subramanian A, Ross KN. The connectivity map: using geneexpression signatures to connect small molecules, genes, and disease. Science. 2006; 313(5795):1929–1935. https://doi.org/10.1126/science. 1132939.
- Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27.
- Zhang JD, Wiemann S. KEGGgraph: a graph approach to KEGG PATHWAY in R and Bioconductor. Bioinforma Oxf Engl. 2009;25(11):1470–1. https:// doi.org/10.1093/bioinformatics/btp167.
- 29. Csardi G, Nepusz T. The igraph software package for complex network research. InterJ Complex Syst. 2006;1695(5):1–9.
- Tesson BM, Breitling R, Jansen RC. DiffCoEx: a simple and sensitive method to find differentially Coexpressed gene modules. BMC Bioinformatics. 2010;11(1):497. https://doi.org/10.1186/1471-2105-11-497.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303.
- Pauli C, Hopkins BD, Prandi D, Shaw R, Fedrizzi T, Sboner A, Sailer V, Augello M, Puca L, Rosati R, McNary TJ, Churakova Y, Cheung C, Triscott J, Pisapia D, Rao R, Mosquera JM, Robinson B, Faltas BM, Emerling BE, Gadi VK, Bernard B, Elemento O, Beltran H, Demichelis F, Kemp CJ, Grandori C, Cantley LC, Rubin MA. Personalized in vitro and in vivo cancer models to guide precision medicine. Cancer Discov. 2017;7(5):462–77. https://doi. org/10.1158/2159-8290.CD-16-1154.
- Bleijs M, van de Wetering M, Clevers H, Drost J. Xenograft and organoid model systems in cancer research. EMBO J. 2019; 38(15): e101654. https:// doi.org/10.15252/embj.2019101654.
- Picco G, Garnett MJ. A road map for precision cancer medicine using personalized models. Cancer Discov. 2017;7(5):456–8. https://doi.org/10. 1158/2159-8290.CD-17-0268.

- Dugger SA, Platt A, Goldstein DB. Drug development in the era of precision medicine. Nat Rev Drug Discov. 2018;17(3):183–96. https://doi.org/ 10.1038/nrd.2017.226.
- 36. Zhao H, Gu S, Bao S, Yan C, Zhang Z, Hou P, Zhou M, Sun J. Mechanistically derived patient-level framework for precision medicine identifies a personalized immune prognostic signature in high-grade serous ovarian cancer. Brief Bioinform. 2020. https://doi.org/10.1093/bib/bbaa069.
- Nevins JR, Potti A. Mining gene expression profiles: expression signatures as cancer phenotypes. Nat Rev Genet. 2007;8(8):601–9. https://doi.org/10. 1038/nrg2137.
- Ziyaei K, Abdi F, Mokhtari M, Daneshmehr MA, Ataie Z. Phycocyanin as a nature-inspired antidiabetic agent: a systematic review. Phytomedicine. 2023;119.
- Liu Y, Zhang Q, Xing B, Luo N, Gao R, Yu K, Hu X, Bu Z, Peng J, Ren X, Zhang Z. Immune phenotypic linkage between colorectal cancer and liver metastasis. Cancer Cell. 2022;40(4):424-437.e5. https://doi.org/10. 1016/j.ccell.2022.02.013.
- Housman G, Byler S, Heerboth S, Lapinska K, Longacre M, Snyder N, Sarkar S. Drug resistance in cancer: an overview. Cancers. 2014;6(3):1769–92. https://doi.org/10.3390/cancers6031769.
- Huang L, Li F, Sheng J, Xia X, Ma J, Zhan M, Wong STC. DrugCombo-Ranker: drug combination discovery based on target network analysis. Bioinformatics. 2014;30(12):i228–36. https://doi.org/10.1093/bioinforma tics/btu278.
- Sheikh Beig Goharrizi MA, Ghodsi S, Mokhtari M, Moravveji SS. Noninvasive STEMI-related biomarkers based on meta-analysis and gene prioritization. Comput Biol Med. 2023; 161:106997. https://doi.org/10. 1016/j.compbiomed.2023.106997.
- Aust S, Schwameis R, Gagic T, Müllauer L, Langthaler E, Prager G, Grech C, Reinthaller A, Krainer M, Pils D, Grimm C, Polterauer S. Precision medicine tumor boards: clinical applicability of personalized treatment concepts in ovarian cancer. Cancers. 2020;12(3):548. https://doi.org/10.3390/cance rs12030548.
- Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A, Sage J, Butte AJ. Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med. 2011; 3(96):96ra77–96ra77. https://doi.org/10.1126/scitranslmed.3001318.
- Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov. 2004;3(8):673–83. https:// doi.org/10.1038/nrd1468.
- 46. Song P, Li W, Wu X, Qian Z, Ying J, Gao S, He J. Integrated analysis of single-cell and bulk RNA-sequencing identifies a signature based on B cell marker genes to predict prognosis and immunotherapy response in lung adenocarcinoma. Cancer Immunol Immunother Cll. 2022;71(10):2341–54. https://doi.org/10.1007/s00262-022-03143-2.
- Padhy BM, Gupta YK. Drug repositioning: re-investigating existing drugs for new therapeutic indications. J Postgrad Med. 2011;57(2):153. https:// doi.org/10.4103/0022-3859.81870.
- Kotelnikova E, Yuryev A, Mazo I, Daraselia N. Computational approaches for drug repositioning and combination therapy design. J Bioinform Comput Biol. 2010;8(03):593–606. https://doi.org/10.1142/s021972001 0004732.
- Stolfi C, De Simone V, Pallone F, Monteleone G. Mechanisms of Action of Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) and mesalazine in the chemoprevention of colorectal cancer. Int J Mol Sci. 2013;14(9):17972–85. https://doi.org/10.3390/ijms140917972.
- Richon VM. Cancer biology: mechanism of antitumour action of Vorinostat (Suberoylanilide Hydroxamic Acid), a novel histone deacetylase inhibitor. Br J Cancer. 2006;95(1):S2–6. https://doi.org/10.1038/sj.bjc. 6603463.
- Hutschenreuther A, Birkenmeier G, Bigl M, Krohn K, Birkemeyer C. Glycerophosphoglycerol, Beta-Alanine, and Pantothenic acid as metabolic companions of glycolytic activity and cell migration in breast cancer cell lines. Metabolites. 2013;3(4):1084–101. https://doi.org/10.3390/metab o3041084.
- Schultze E, Buss J, Coradini K, Begnini KR, Guterres SS, Collares T, Beck RCR, Pohlmann AR, Seixas FK. Tretinoin-loaded lipid-core nanocapsules overcome the triple-negative breast cancer cell resistance to tretinoin and show synergistic effect on cytotoxicity induced by doxorubicin and 5-fluororacil. Biomed Pharmacother. 2017;96:404–9. https://doi.org/10. 1016/j.biopha.2017.10.020.

- 53. Rosen LS. Irinotecan in lymphoma, leukemia, and breast, pancreatic, ovarian, and small cell lung cancers. Oncol. 1998; 12:103–109.
- 54. Morizane C, Machida N, Honma Y, Okusaka T, Boku N, Kato K, Nomura S, Hiraoka N, Sekine S, Taniguchi H, Okano N, Yamaguchi K, Sato T, Ikeda M, Mizuno N, Ozaka M, Kataoka T, Ueno M, Kitagawa Y, Terashima M, Furuse J. Japan Clinical Oncology Group (JCOG). Effectiveness of etoposide and cisplatin vs irinotecan and cisplatin therapy for patients with advanced neuroendocrine carcinoma of the digestive system: the TOPIC-NEC phase 3 randomized clinical trial. JAMA Oncol. 2022; 8(10):1447–1455. https:// doi.org/10.1001/jamaoncol.2022.3395.
- Zhou H, Shen T, Luo Y, Liu L, Chen W, Xu B, Han X, Pang J, Rivera CA, Huang S. The antitumor activity of the fungicide ciclopirox. Int J Cancer. 2010;127(10):2467–77. https://doi.org/10.1002/ijc.25255.
- Pan C, Liu M, Li H, Wei L, Wang P, Wu K, Ji X, Gu L, Tian Y. Systematic evaluation of randomized clinical trials of Huangqin Tang in combination with Mesalazine for ulcerative colitis. Evid-Based Complement Altern Med. 2022; 2022:2175913. https://doi.org/10.1155/2022/2175913.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

