

# Personalized Multi-Drug Combinations in Breast Cancer Treatment Using Individual Pattern of Perturbed Gene Expression and Drug Signatures

**Majid Mokhtari**

University of Tehran

**Samane Khoshbakht**

University of Tehran

**Mohammad Esmail Akbari**

Shahid Beheshti university of Medical Sciences

**Sayyed Sajjad Moraveji**

University of Tehran

**Ali Masoudi-Nejad** (✉ [amasoudin@ut.ac.ir](mailto:amasoudin@ut.ac.ir))

University of Tehran

---

## Short Report

**Keywords:** Systems Biology, Differential Gene Expression Analysis

**Posted Date:** August 26th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-812373/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The greatest challenge in recent years in cancer treatment has been drug screening. Few platforms presented reliable solutions for personalized drug validation and safety testing. Here, we constructed a personalized drug combination protocol as the primary input to such platforms. We used public data from whole-genome expression profiles of 6173 breast cancer patients, 312 healthy individuals, and 691 drugs. We developed an individual pattern of perturbed gene expression (IPPGE) for each patient. A protocol was designed to extract personalized drug combinations by comparing the IPPGE and drug signatures. We tried to use the concept of drug repurposing, which searches for the new benefits of the existing medicines that may perturb the desired genes. The potential for treatment effectiveness was more significant for drug combinations extracted from specialized and nonspecialized cancer medicines than specialized medicines. Thus, effective treatments can be provided through the approach of drug repurposing and combination drug therapy.

## Implications

The protocol allows personalized drug combinations to be extracted from hundreds of drugs and thousands of drug combinations. This can be used as a methodological interface between drug repurposing and combination drug therapy in cancer treatment.

## Introduction

The gene expression involved in diseases typically shows a unique pattern, which can be detected when studying patient populations<sup>17</sup>. Drug repurposing studies aim to identify the most appropriate medicines that can regulate the genes involved in perturbed expression patterns in disease. Moreover, studies have shown that combination drug therapy can be more efficient in reducing drug resistance due to the control of parallel biological pathways<sup>18</sup>. Heterogeneity of cancer leads to different responses to similar treatments<sup>6</sup>. Therefore, it is essential to identify the gene expression pattern of each patient to determine the most effective combination drug therapy<sup>10</sup>.

We developed a protocol for 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 protocol helps the physician prescribe the most effective drug combination by precisely monitoring the patient's gene expression status.

The following were carried out in this study: (i) Breast cancer oncogenes were identified through differentially expressed gene analysis by comparing patients and healthy controls. (ii) A healthy gene expression interval was calculated for each oncogene (i.e., the oncogene health interval) by assessing gene expression in healthy individuals. (iii) A patient's gene expression profile was compared with the oncogene health intervals to create a individual pattern of perturbed gene expression (IPPGE). (iv) One or more drug combinations were identified for the patient by simultaneously analyzing the IPPGE and the

drug signatures through the use of a network-based algorithm. (v) A directed differential network (DDN) was reconstructed using biological pathway data to predict the effect on gene expression of the drug combinations (**Visual Overview**).

## Materials And Methods

### Data sets and Metadata Information

The raw expression files (CEL files) of breast cancer patients were downloaded using the GEOquery package in R <sup>4</sup>. All the datasets were from the Affymetrix platform. **Supplementary Table S1** shows the sample size and platform information of each dataset. We analyzed 6173 untreated breast cancer patients and 312 healthy controls. A total of 691 gene expression profiles of drugs were used and downloaded from the CMAP and the national center for biotechnology information (NCBI) databases (**Supplementary Table S2**). Full details of all bioinformatics analyses for pre-processing, normalization, differentially expressed genes (DEGs), upregulated oncogenes (UROs), downregulated oncogenes (DROs), and oncogene health interval detection have been described in **Supplementary Materials and Methods**.

### Oncogene Detection

To identify oncogenes, samples from 6173 patients and 312 healthy individuals pre-processed, followed by the analysis of differentially expressed genes (DEGs). In total, 95 genes with a LogFC greater than 0.45 were identified as upregulated oncogenes (UROs), and 130 genes with a LogFC less than - 0.45 were identified as downregulated oncogenes (DROs). A total of 225 oncogenes from the DEG analysis were used to construct the individual pattern of perturbed gene expression (Fig. 1A). The oncogene clustering and functional enrichment results were illustrated in Fig. 1B and 1C; please see **Supplementary Materials and Methods** for full details.

### Individual Pattern of Perturbed Gene Expression (IPPGE)

Each patient's disease-leading genes (individual oncogenes) were identified using the health intervals calculated in the previous step and the individual's gene expression profile (**Supplementary Materials and Methods**). The health interval refers to the range of gene expression in which members of the disease group are rarely seen within it (Fig. 1D). Accordingly, each gene that was not within the health interval was determined as an individual oncogene for a given individual. The simultaneous expression of these individual oncogenes in each patient forms that patient's IPPGE.

### Drug Combination Algorithm (DC algorithm)

The primary health matrix (PHM) was created through synchrony between the IPPGE and CMAP drugs with the health interval <sup>9</sup>. The PHM is a binary matrix with the CMAP drugs as columns and individual oncogenes as rows. A value of 1 indicates that the oncogene lies within the health interval after the patient receives a drug. The DC algorithm uses the PHM to determine the most appropriate drug

combination for an individual patient. We implemented the DC algorithm in several steps. Pseudocodes for the DC algorithm are in **Supplementary Table S3**.

## Directed Network Reconstruction

All KGML *Homo sapiens* files were downloaded from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and merged to create a directed network as a reference signaling pathway. The gene coexpression networks were individually reconstructed for patients and healthy controls using the Spearman correlation. To create a differential signaling network, the differential network for two conditions, disease vs. healthy, was detected and mapped to the reference network (Fig. 1E and 1F). Finally, to identify the downstream targets of drug combinations extracted by our protocol, the gene targets of the drug combinations for different patients were mapped to the differential signaling network<sup>19</sup>. All the network visualizations were implemented in Cytoscape software<sup>15</sup>.

## Results

### Individual Pattern of Perturbed Gene Expression (IPPGE)

An oncogene health interval was determined for each oncogene based on a comparison of the gene expression between the healthy group and cancer group. The results showed that each oncogene had a unique health interval. In the patient population, all oncogenes were found to have a mean gene expression level outside of the health interval. By contrast, in the healthy population, the mean gene expression of all oncogenes was within the health interval (Fig. 1G). In Fig. 1G, each cell symbolizes one of 225 oncogenes in the healthy population (each oncogene was previously identified in Fig. 1A). 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 oncogene's health interval. 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 oncogenes were observed to be outside of the health interval, while in patient 2, 154 oncogenes were outside of the health interval. The third and fourth patients also had the same subtype (ER-/PR-/HER2-). In patient 3, 165 oncogenes were outside of the health, while 169 oncogenes were outside of the health interval in patient 4. Notably, sometimes an equal number of oncogenes were observed in two patients, but various oncogenes were identified. Therefore, not only a significant difference was observed between the numbers of oncogenes outside of the health interval among patients, but there was also a difference in the type of oncogenes observed. The results indicated that patients could have very different IPPGEs despite having the same cancer subtype. Most oncogenes in each patient were outside of the health interval, but the patients had some oncogenes within the health interval (Fig. 2A and 2B). 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 oncogenes 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 oncogenes 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 FDA-approved 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 patient, 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.

## IPPGEs to Assess Drug Combinations

The effect of the extracted drug combinations on the IPPGE of patients was examined. 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 (Fig. 2C, 2D, 2E, 2F, 2G, 2H, 2I and 2J). Moreover, side effects and drug interactions in the Beta group were observed less often than in 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 (shown in Fig. 1A). 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. 1F). 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 (**Supplementary Fig. S1A and 3B**).

## Discussion

Every new drug treatment must be tested several times to obtain approval from health organizations before it can be used in the clinic. However, if a personalized treatment designed for a single patient, it will have a significant effect only for that individual. Therefore, the routine approval processes for therapies based on treating other organisms and individuals will not provide the appropriate treatment response. 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 tumor organoid (PDTO) cultures followed by patient-derived xenograft (PDX) models<sup>3,5,13,14</sup>. However, analyzing the hundreds of drugs and thousands of possible drug combinations using this platform remains a considerable and complex challenge<sup>20</sup>. Our protocol, enabled us to extract personalized drug options that could potentially be introduced as the primary input for that of Pauli's platform.

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 resulting IPPGE was unique to each patient, like a fingerprint (Fig. 2A and 2B). A previous study reported variations in gene expression patterns in different cancer phenotypes and also in the patients<sup>11</sup>. Importantly, although patients have many common perturbed genes, the gene expression compared to the health interval plays a pivotal role in drug selection. 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.

## Potential of drug combinations derived from patient IPPGEs

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<sup>6,7,2</sup>. The current study also found that each medication affected a certain number of a patient's oncogenes; the protocol extracted complementary drugs that affected the largest number of oncogenes. We found that drug combinations derived from patient IPPGEs had stronger treatment potential due to their more targeted effects on oncogenes. 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.

## Drug Repurposing

Several drug repurposing studies have reported significant anticancer efficacy for nonspecialized drugs. One of the first drug repurposing studies showed that the anti-ulcer drug cimetidine to be a therapeutic candidate for the treatment of adenocarcinoma of the lung <sup>1,16</sup>. Subsequent studies have found that combination drug therapy increases the success of drug repurposing <sup>8,12,18</sup>. 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 (**Table 1** and **Supplementary Fig. S1**). This finding has two distinct interpretations; First, nonspecialized drugs may be used as adjunctive treatments in addition to specialized medications. Second, according to the observations, some nonspecialized cancer drugs were identified as the potential main treatment. 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.

## Conclusion

The present study extracts personalized drug combinations from hundreds of drugs and thousands of drug combinations. 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. <sup>13</sup>. The protocol can be used as a methodological interface between drug repurposing activities and combination drug therapy.

## Declarations

Conflict of Interest Disclosure Statement: The authors declare no potential conflicts of interest.

## References

1. Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov.* 2004;3(8):673–683. doi:10.1038/nrd1468
2. Aust S, Schwameis R, Gagic T, et al. Precision Medicine Tumor Boards: Clinical Applicability of Personalized Treatment Concepts in Ovarian Cancer. *Cancers.* 2020;12(3):548. doi:10.3390/cancers12030548

3. Bleijs M, van de Wetering M, Clevers H, Drost J. Xenograft and organoid model systems in cancer research. *EMBO J*. 2019;38(15):e101654. doi:10.15252/embj.2019101654
4. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*. 2007;23(14):1846–1847. doi:10.1093/bioinformatics/btm254
5. Dugger SA, Platt A, Goldstein DB. Drug development in the era of precision medicine. *Nat Rev Drug Discov*. 2018;17(3):183–196. doi:10.1038/nrd.2017.226
6. Housman G, Byler S, Heerboth S, et al. Drug Resistance in Cancer: An Overview. *Cancers*. 2014;6(3):1769–1792. doi:10.3390/cancers6031769
7. Huang L, Li F, Sheng J, et al. DrugComboRanker: drug combination discovery based on target network analysis. *Bioinformatics*. 2014;30(12):i228–i236. doi:10.1093/bioinformatics/btu278
8. 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. doi:10.1142/s0219720010004732
9. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *science*. 2006;313(5795):1929–1935. doi:10.1126/science.1132939
10. Li YY, Jones SJ. Drug repositioning for personalized medicine. *Genome Med*. 2012;4(3):27. doi:10.1186/gm326
11. Nevins JR, Potti A. Mining gene expression profiles: expression signatures as cancer phenotypes. *Nat Rev Genet*. 2007;8(8):601–609. doi:10.1038/nrg2137
12. Padhy BM, Gupta YK. Drug repositioning: Re-investigating existing drugs for new therapeutic indications. *J Postgrad Med*. 2011;57(2):153. doi:10.4103/0022-3859.81870
13. Pauli C, Hopkins BD, Prandi D, et al. Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. *Cancer Discov*. 2017;7(5):462–477. doi:10.1158/2159-8290.CD-16-1154
14. Picco G, Garnett MJ. A Road Map for Precision Cancer Medicine Using Personalized Models. *Cancer Discov*. 2017;7(5):456–458. doi:10.1158/2159-8290.CD-17-0268
15. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504. doi:10.1101/gr.1239303
16. Sirota M, Dudley JT, Kim J, et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci Transl Med*. 2011;3(96):96ra77–96ra77. doi:10.1126/scitranslmed.3001318
17. del Sol A, Balling R, Hood L, Galas D. Diseases as network perturbations. *Curr Opin Biotechnol*. 2010;21(4):566–571. doi:10.1016/j.copbio.2010.07.010
18. Sun W, Sanderson PE, Zheng W. Drug combination therapy increases successful drug repositioning. *Drug Discov Today*. 2016;21(7):1189–1195. doi:10.1016/j.drudis.2016.05.015
19. Tesson BM, Breitling R, Jansen RC. DiffCoEx: a simple and sensitive method to find differentially coexpressed gene modules. *BMC Bioinformatics*. 2010;11(1):497. doi:10.1186/1471-2105-11-497

20. Zhao H, Gu S, Bao S, et al. Mechanistically derived patient-level framework for precision medicine identifies a personalized immune prognostic signature in high-grade serous ovarian cancer. *Brief Bioinform*. Published online 2020. doi:10.1093/bib/bbaa069

## Tables

Table 1 is included in the Supplementary Files section.

## Figures

Figure 1

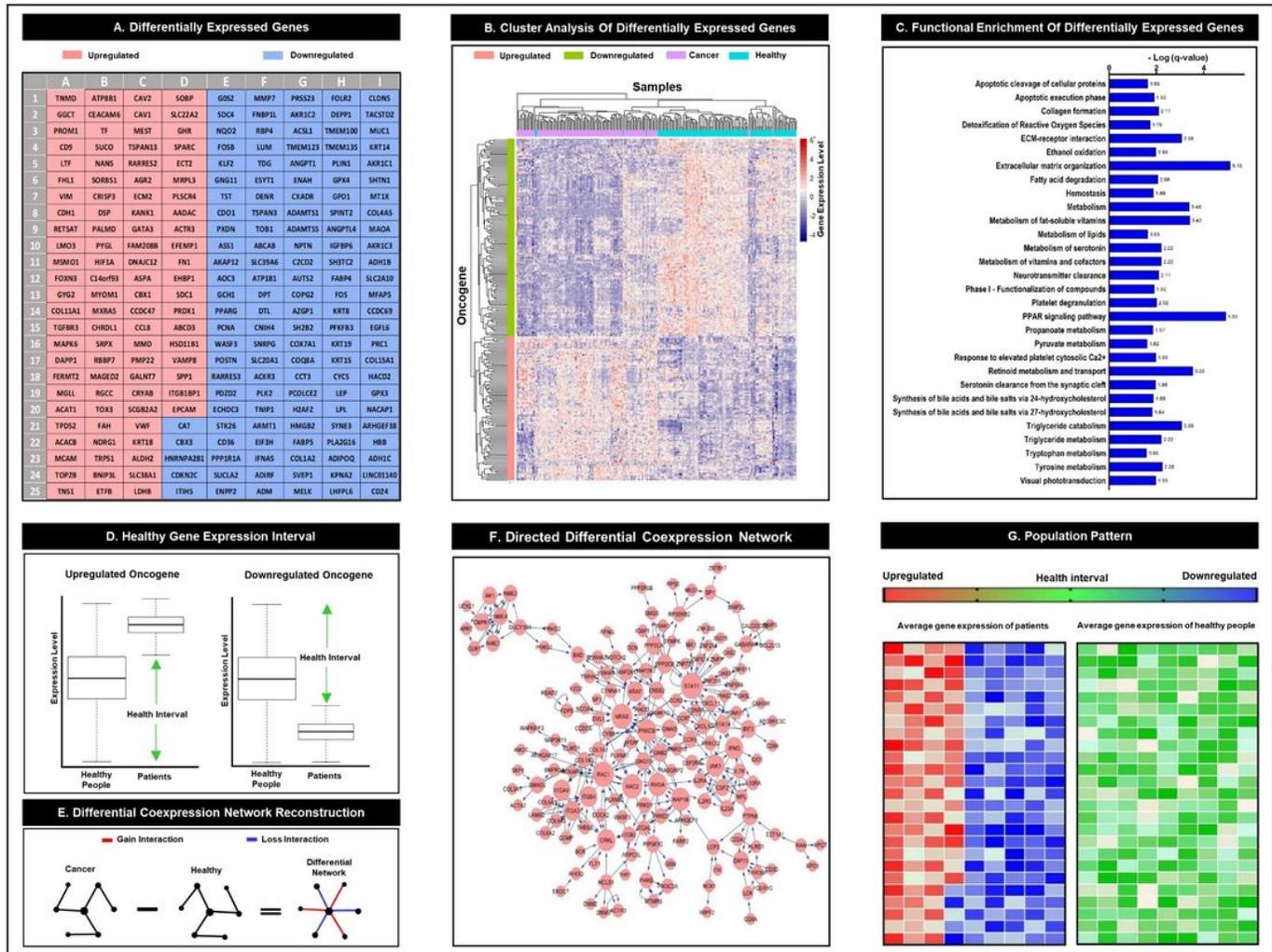


Figure 1

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 concept of the health interval

is illustrated in the upregulated and the downregulated genes. E. The concept of the differential network is illustrated in cancer computational biology. F. The induced signaling pathway which was extracted from the KEGG database using the DEGs. G. The expression pattern of DEGs in the population.

Figure 2

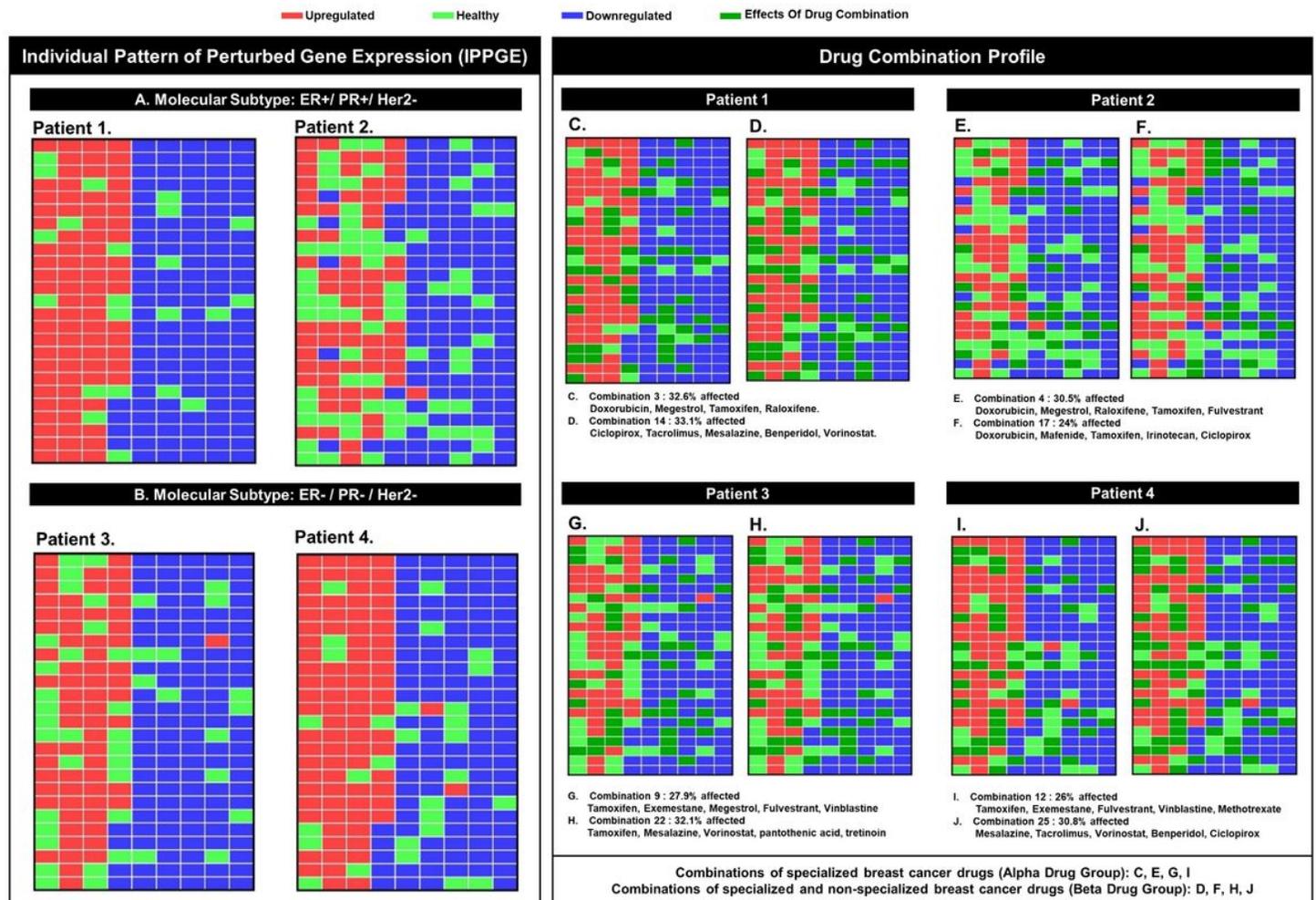


Figure 2

Individual pattern of perturbed gene expression (IPPGE). A, B. The expression pattern of DEGs in individual patients. Drug combinations profile. The effects of drug combinations on the gene expression of patients are illustrated. C, E, G, and I 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 D, F, H, and J. The Beta group includes specialized and nonspecialized cancer drugs.

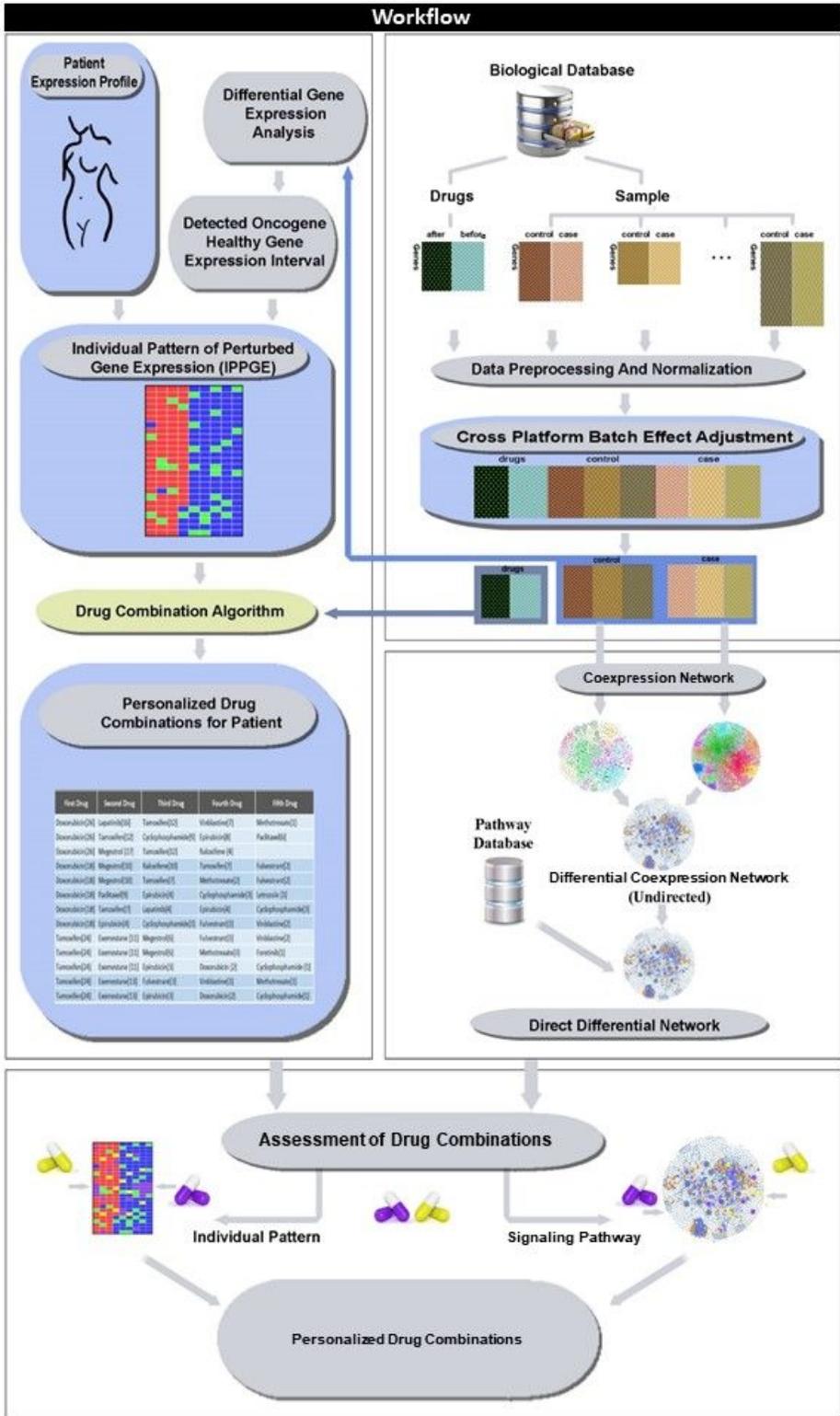


Figure 3

Visual Overview: protocol workflow. The method is a set of various methods used in a bioinformatics protocol to extract personalized drug combinations.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFiles.rar](#)
- [Table1.docx](#)