

Drug Repositioning Study in Search of Potent Inhibitors Blocking Multiple Pathways for the Treatment of Breast Cancer

Sushil Kumar Kashaw (✉ sushilkashaw@gmail.com)

Dr Hari Singh Gour University <https://orcid.org/0000-0003-3969-3388>

Shivangi Agarwal

Dr Hari Singh Gour University

Anshuman Dixit

Institute of life Sciences

Research article

Keywords: Differential gene expression, Biological pathway, Repurposing of drugs, breast cancer

Posted Date: July 16th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-42884/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Breast cancer is most malignant cancer in women worldwide. The efforts are going on for improvement in patient's survival and treatment strategies. The goal of the present work is to find out connections among drugs-genes-breast cancer and to re-purpose approved drugs for the treatment of breast cancer. In this context, the gene expression data of breast cancer samples was analyzed to identify the upregulated/downregulated genes in different clinical stages of breast cancer. A large number of genes were found to be upregulated/downregulated in different clinical stages of breast cancer. Some of the genes were found to be stage specific and some were common to all stages. The biological pathways were studied in early and late stages of breast cancer, which indicated that pathways as Methylglyoxal Degradation I, Catecholamine biosynthesis and Serotonin/Melatonin biosynthesis were enriched in early stage and pathways as Matrix metalloproteases, Airway pathology in COPD, Glycogen degradation II and Glycogen degradation III were enriched in late stages of breast cancer. The drug repurposing analysis revealed that the different classes of drugs as enzyme inhibitors, CNS agents, Glucocorticoids, insulin sensitizers, tubulin inhibitors, adhesion inhibitors showed strong connections with MCF7 cell line. Further, it was found that the drugs as Nortriptyline, AZD-6482, Acitretin and GW-507 were found to target multiple genes in interleukin pathway, enzymatic pathway and GPCR signaling pathway and drugs as Caffeine, Canertinib and Triciribine was found to target multiple genes which were dysregulated and involved in interleukin pathway, enzymatic pathway, GPCR signaling pathway and metabolic pathways.

1. Introduction

Breast cancer is the most frequently diagnosed cancer in women and is the leading cause of death worldwide¹. As per ACS statistics 2017, approximately 2 lakhs 50 thousand new cases of breast cancer were detected. It is one of the prevalent cancers in women and its heterogeneity is one of the major challenges in the treatment of breast cancer²⁻⁴. So, the need of the today is to choose the crucial and effective therapy for breast cancer treatment. The ultimate goal of the cancer research is to connect the disease to the genes/protein underlying it and finally the chemical modulator/drug which can be used to treat them. Connectivity Map⁵ is an online tool developed by Golub *et al.*, that provide generic solution to the problem by establishing the disease-gene-drug connections and accelerate the pace of drug development. The genomic signatures can be used to identify the drugs with common action and to discover mechanism of actions of unannotated small molecules and recognize the potential drugs. CMAP initially includes only 164 drugs tested on few cell lines⁵. CLUE (CMAP linked user environment) is the updated version with 1000-fold increase of CMAP pilot data⁶.

Drug repurposing is a method to identify new therapeutic uses of already existing/approved drugs. It significantly reduces the escalating cost and length of time needed for bringing new drugs to the market. Drug repurposing is also known as drug re-positioning and is the application of treating pathological conditions with the known drugs which is other than their use in original conditions. It is an alternative method to conventional drug discovery pipeline which takes 10–15 years for the discovering de novo drug compounds. Nowadays in United States, around 30% of newly marketed drugs coming through drug repositioning method. Despite recent advances in various technologies, the number of newly approved drugs does not increase. Only 5–10% of the drugs entering the Phase I clinical trials comes to market and rest fails due to their significant toxicity or suboptimal pharmacological response. Moreover, it is not possible to develop de novo drug compounds for more than 8000 orphan diseases with the huge research and development costs and time. In 2004, Ashburn and Thorn introduced an approach known as drug repurposing in which new targets and indications of existing drugs are explored and analysed using bioinformatics facilities^{7,8}. The strategy for drug repositioning involves the computational method, algorithms and tools which have been extensively used and called as computational drug repurposing. Various in-silico methods and techniques help in identifying the relationship among drug, targets and diseases. The polypharmacological compounds acting on multiple targets can be screened through high-throughput screening and can be used in treatment of multiple diseases. The data mining tools and bioinformatics analysis can be used to derive some meaningful results and conclusions required for repurposing of drugs. There are various drugs which have been repurposed for example Nelfinavir for cancer, Tamoxifen for bipolar disorder, Gleevec for rheumatoid arthritis, Pentylentetrazol for Down Syndrome, Astemizole for malaria, Lipitor for Alzheimer's and Metformin for cancer⁹. In CLUE, the authors have tested various molecular perturbagens against different cell lines belonging to breast cancer, prostate cancer, leukemia, melanoma etc. and scored them in terms of negative to positive connections. The query in CLUE involves the set of differentially expressed genes belonging to a particular pathological condition of interest. CLUE then connects those lists of genes to the drugs and gives a connectivity score ranging from +100 to -100 in positive (positive connectivity) or negative (negative connectivity). All instances were then ranked based on their connectivity scores, those at top are the positive connections and those at bottom are the negative connections.

In the present work, we have used bioinformatics approach to relate the FDA approved drugs to the genes underlying breast cancer. In context of this, we have initially analysed the gene expression data of breast cancer and adjacent normal samples and evaluated the genes and biological pathways dysregulated in different clinical stages of breast cancer. Further, the genes were related to molecular perturbagen/drug with the help of connectivity map and biological pathways of target genes were studied. The drugs were screened by mapping to biological pathways and finally selected using Venn analysis.

2. Material And Methods

The flow of the present employed computational approach is given in Fig. 1.

2.1 Data collection and classification

The expression data of transcripts was downloaded from Cancer RNA-seq Nexus¹⁰. The data contained transcript per million values (TPM) values of different coding transcripts in breast invasive carcinoma tissues and adjacent normal tissues. It also had information about clinical stages of breast cancer which enabled us to classify the data as per the clinical stages as I, Ia, Ib, II, IIa, IIb, IIIa, IIIb, IIIc and IV. The mean of expression values of transcripts belonging to a particular gene was computed.

2.2 Differential gene expression analysis

The differential gene expression was computed in different stages as stage I, Ia, Ib, II, IIa, IIb, IIIa, IIIb, IIIc and IV of breast invasive carcinoma. The fold change was calculated using transcript expression values in carcinoma and adjacent normal samples and fold change values were converted to logarithm scale. The genes with Log2 fold change (FC) greater than 1 and adjusted p-value < 0.05 were considered to be significant. The genes with Log2 FC > 1 and adjusted p-value < 0.05 were considered to be significantly upregulated and genes with Log2 FC < -1 and adjusted p-value < 0.05 were considered to be significantly downregulated.

2.3 Biological pathway analysis

The biological pathways were analyzed using Ingenuity Pathway analysis (IPA) software (<https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/>). The top 200 upregulated and downregulated genes were considered for pathway analysis from all the stages. The pathways were analysed in early and late stage of breast invasive carcinoma. The early stage included stages I, Ia, Ib, II, IIa, IIb and late stage included stages as IIIa, IIIb, IIIc and IV. The top pathways based on p-value were considered to be significant pathways.

2.4 Re-purposing of approved drugs

CLUE is an online webtool connecting different cell lines such as MCF7, PC3, VCAP, A375, A549, HA1E, HCC515, HT29, HEPG2 to the FDA approved drugs ⁶. The query in CLUE involves the set of differentially expressed genes belonging to a particular pathological condition of interest. Clue connects those lists of genes to the drugs and gives a connectivity score ranging from +100 to -100 in terms of positive or negative connectivity. All instances were then ranked based on their connectivity scores. In the present study, we have mapped specific breast cancer genes to approved drugs with the help of CLUE. The drug targets were then clubbed, and the enrichment of biological pathways were studied using IPA software. The drugs were mapped against specific gene targets in pathway maps. The drugs affecting multiple genes in pathways were further studied.

3. Results And Discussion

3.1 Differential expression analysis

The differential expression study was done in different stages of breast invasive carcinoma as Stage I, Ia, Ib, II, IIa, IIb, IIIa, IIIb, IIIc and IV. A huge number of genes were found to be dysregulated in different stages (supplementary table S1a to S1j). Some of the genes were found to be stage specific and some were common to all stages as there were about 1984 genes which were specific to early stage and around 1753 genes that were specific to late stages of breast cancer. (Fig. 2a and 2b, supplementary table S2). The number of upregulated and downregulated genes in each stage are shown in Table 1. Additionally, we found some genes that were specific to breast cancer cells and were not expressed in adjacent normal cells (Table 2). The hierarchical clustering showed that genes of stage I and Ia were clustered together based on differential expression which further clustered with genes belonging to IIIb, IIIc and IV and genes of stage IIb and IIIa were clustered together (Fig. 3a). Additionally, the clustering of genes common to all stages showed clustering of stage I & II and III & IV (Fig. 3b).

Table 1
The number of upregulated and downregulated genes in different clinical stages of breast cancer.

Stage	Upregulated (>2 fold change)	Downregulated (>2 fold change)	Expressed only in adjacent normal	Expressed only in cancer cells
I	2718	2632	03 (AX747287, OR6C2 & OR6C70)	04 (ELSPBP1, IFNL2, KRTAP5-6, SPANXN3)
Ia	2805	2816	04 (AX747287, KRTAP7-1, LINC00895 & OR6T1)	03 (IFNL2, KRTAP5-6, SPANXN3)
Ib	697	3419	298	-
II	188	1023	918	-
IIa	3666	2581	-	44
IIb	3415	2565	-	24
IIIa	3066	2846	02 (OR6C70, TECTB)	09 (AK056267, IFNL2, KRTAP5-6, LOC146481, OR8D2, PDCL2, PRAMEF1, SLC22A24)
IIIb	1943	4014	65	-

Table 2
The list of gene expressed only in breast cancer cells in different clinical stages

IIIc	2716	2996	03(BEND2, KRT25 & STATH)	02 (KRTAP5-6, SPANXN3)
IV	1684	3774	69	-

Gene symbol	Expression in Clinical stages	Gene symbol	Expression in Clinical stages
IFNL2	I,IA,IIA,IIB,IIIA	KRTAP10-3	IIA
KRTAP5-6	I,IA,IIA,IIB,IIIA,IIIC	LOC146481	IIA, IIB, IIIA
SPANXN3	I,IA,IIA,IIB,IIIA,IIIC	MBD3L1	IIA, IIB
ELSPBP1	I, IIB	MBD3L5	IIA, IIB
ACTL9	IIA	OFCC1	IIA, IIB
AK056267	IIA, IIIA	OR2G6	IIA
AK093214	IIA	OR2W5	IIA
AK131021	IIA	OR5T2	IIA, IIB
APCS	IIA	OR6M1	IIA
ATOH1	IIA	PDCL2	IIA, IIB, IIIA
AX747578	IIA	PRODH2	IIA
CGB2	IIA	RNASE9	IIA
CHAT	IIA	SLC22A24	IIA, IIB, IIIA
CT45A3	IIA	SNAR-B2	IIA, IIB
ELP4	IIA	SSX7	IIA
FAM25C	IIA	TBC1D3P2	IIA
FTHL17	IIA	TCP10	IIA
GAGE1	IIA, IIB	TINAG	IIA
GAGE12D	IIA	CELA2A	IIB
GAGE12J	IIA	CGB1	IIA
GAGE2A	IIA	LGALS14	IIB
GAGE2D	IIA, IIB	PLGLA	IIB
GAGE4	Ila, IIB	PRAMEF1	IIB, IIIA
IFNA10	IIA	SI	IIB
IL1F10	IIA	SNORA51	IIB
KIF2B	IIA	TRIM53AP (TRIM53)	IIB
KRMP1	IIA, IIB	OR8D2	IIB

To compare the genomic variations among the clinical stages of breast cancer, we have plotted the differentially expressed genes as per their fold change values in a circos plot. The outermost track (track 0) is circular ideogram representing chromosomal number. Extending inside, is the heatmap plots representing different stages of breast cancer from stage I to IV (Fig. 4). There are ten circular tracks corresponding to stage I, IA, IB, II, IIA, IIB, IIIA, IIIB, IIIC and IV. From the plot, it is inferred that in early stages (stage I to II), chromosomes 1, 3, 6, 8, 10, 16, 17, 19 and 22 were affected and in late stages (stage III to IV), most of the chromosomes were affected as chromosome 1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 18 and 19.

3.2 Biological Pathway analysis

The top up and downregulated genes in each stage were further studied for biological pathways. In early stages, 127 biological pathways were found and in late stage 69 pathways were found (supplementary table S3a, S3b). Figure 5a and 5b shows the bar plots of biological pathways with ratio and $-\log(p\text{-value})$. In early stage of breast invasive carcinoma, pathways as Methylglyoxal Degradation I pathway, Catecholamine biosynthesis pathway and Serotonin and Melatonin biosynthesis pathway were found and in late stage, pathways as Inhibition of Matrix metalloproteases, Airway pathology in COPD, Glycogen degradation II and Glycogen degradation III were found (supplementary figure S1a to S1f). The literature reports suggest that these pathways have been involved in development or progression of breast cancer by direct or indirect means. The knockdown of glyoxalase I in MCF-7 significantly reduced tumor-associated properties such as migration and proliferation. Study supports that Glyoxalase inhibitors may be used as anti-cancer drugs.¹¹⁻¹⁴ The chronic stress can enhance breast cancer disease progression by upregulating catecholamines level and signaling of β -adrenergic receptors¹⁵. Serotonin regulates

homeostasis in breast, also blocks ER-alpha^{16,17}. It was found that the increased level of 5-HT signaling favor malignant progression of human breast cancer cells¹⁶.

3.3 Repurposing of drugs

The genes which were expressed only in breast cancer cells were considered for drug re-purposing. The inhibitors of these genes were mapped using CLUE. Around 177 drugs (inhibitors) were mapped against those genes with connectivity score cutoff – 95.0 (supplementary table S4). The inhibitors belonged to different classes as enzyme inhibitors, CNS agents, antioxidant, glucocorticoid receptor agonist, hormonal agents, diuretic, sigma receptor agonist, DNA alkylating agents, mucolytic agents, angiogenesis inhibitors, reuptake inhibitors, platelet aggregation inhibitor, retinoid receptor agonist etc. (Fig. 6a). These inhibitors were found to target various other genes which were enlisted (supplementary table S5). The biological pathways of the all these genes were studied using IPA software. A total of 421 pathways were found (supplementary table S6), out of which 22 pathways were significantly enriched (p-value < 0.05) (Fig. 6b and Table 3), such as fatty acid biosynthesis pathway, melatonin degradation pathway, tryptophan degradation pathway, IL-22 signalling, Bupropion degradation, role of JAK family kinase in IL-6 type cytokine signaling, PXR/RXR activation pathway, EGF signaling, IL-15 signaling, cAMP mediated signaling, G-protein coupled receptor signaling etc. The drugs were mapped against specific genes in pathway maps (supplementary figures S2a to S2q) and their frequency was calculated (Fig. 6c) using pathway matrix table (supplementary table S7). Further, 22 pathways were categorized into four broader categories as interleukin pathway, enzymatic pathway, GPCR signaling pathway and metabolic pathways. A venn analysis indicated that the drugs as Nortriptyline, AZD-6482, Acitretin and GW-507 were found in three categories (interleukin pathway, enzymatic pathway and GPCR signaling pathway) and drugs as Caffeine, Canertinib and Triciribine were found in all four categories (interleukin pathway, enzymatic pathway, GPCR signaling and metabolic pathway) (Fig. 7a). The score of Canertinib, Nortriptyline, Triciribine, Caffeine, Acitretin, AZD-6482, GW-5074 against MCF7 cell lines was found to be -99.5, -98.98, -98.83, -98.68, -96.3, -90.5 and – 98.02 respectively. Also, They have been scored against different other cell lines as PC3, VCAP, A375, A549, HA1E, HCC15, HT29, HEPG2 (Fig. 7b, supplementary table S8).

Additionally, the genes involved in breast cancer were related to strong knockdown-overexpression (KD/OE) gene pairs using CLUE. Around 26 strong KD/OE pairs were found which indicates how the knockdown or overexpression of a gene is related to breast cancer. For example, the overexpression of query genes (input gene to CLUE) is positively connected to knockdown of ANKZF1 and negatively connected to overexpression of ANKZF1 (Fig. 7c, supplementary table S9).

Enriched Canonical Pathways	-log(p-value)	Ratio	Number of genes from dataset	Total number of genes in canonical pathways	Biomarkers
G-Protein Coupled Receptor Signaling	92.100	0.338	93.000	275.000	AKT, ERK1/2(2), IKB-NFkB, NF-kB(3), PI3K, RAS, STAT3(2)
cAMP-mediated signaling	70.000	0.324	73.000	225.000	ERK1/2, ERK1/2,Stat3(5), stat3 dimer(2)
Nitric Oxide Signaling in the Cardiovascular System	29.000	0.296	32.000	108.000	AKT(3), AKT-CAM-HSP-90-Nos-3, ERK1/2,PI3K,VEGF, VEGFR
tRNA Splicing	25.000	0.500	21.000	42.000	NOT FOUND
EGF Signaling	22.500	0.338	23.000	68.000	AKT,EGFR(6),ERK1/2(2),Mtor,p70 S6K, PI3K, STAT3(7), Stat(2),Stat3dimer(2)
IL-15 Signaling	21.200	0.303	23.000	76.000	AKT,ERK1/2(3),ERK1/2,IL-24,IL-6,NF-Kb(2),P38mapk,p38mapk,PI3K,Ras, Stat3, TNF-alpha
PXR/RXR Activation	20.400	0.333	21.000	63.000	AKT(2), FOXO2, IGBP1, IL-6(2),NF-Kbp65, TNF-alpha
Serotonin Receptor Signaling	16.800	0.381	16.000	42.000	Serotonin
The Visual Cycle	16.200	0.632	12.000	19.000	NOT FOUND
Inhibition of Angiogenesis by TSP1	14.100	0.406	13.000	32.000	AKT, Caspase3, KDR, P38mapk, TGF-B1, TSP1(3),VEGFA
Retinoate Biosynthesis I	12.800	0.387	12.000	31.000	Retinal dehydrogenase
IL-17A Signaling in Gastric Cells	10.900	0.400	10.000	25.000	CXCL10,EGFR,ERK1/2(2),NF-Kb1-RelA,p38MAPK(2),RELA,TNF-alpha(3)
IL-22 Signaling	9.580	0.375	9.000	24.000	AKT,ERK1/2,P38mapk,stat1/3/5(6),STAT1/3/5
Bupropion Degradation	9.580	0.375	9.000	24.000	Unspecified monooxygenase
Acetone Degradation I (to Methylglyoxal)	8.730	0.310	9.000	29.000	Unspecified monooxygenase
Role of JAK family kinases in IL-6-type Cytokine Signaling	7.940	0.320	8.000	25.000	ERK1/2,IL-6,P38MAPK,STAT3(6),STAT1/3(6),STAT1/3/5(6),STAT1/3dimer(2),STAT1/3/5DIMER(2),STAT
DNA damage-induced 14-3-3f Signaling	7.500	0.368	7.000	19.000	AKT,BRCA1
Tryptophan Degradation X (Mammalian, via Tryptamine)	5.890	0.300	6.000	20.000	Aldehyde dehydrogenase (NAD)
Melatonin Degradation II	2.690	0.500	2.000	4.000	NOT FOUND
Melatonin Degradation III	1.730	1.000	1.000	1.000	NOT FOUND
Palmitate Biosynthesis I (Animals)	1.430	0.500	1.000	2.000	3-hydroxyoctanoyl-[acyl-carrier-protein] dehydratase(5),3-hydroxypalmitoyl-acyl-carrier-protein dehydratase, 3-oxoacyl-acyl-carrier-protein reductase(7), 3-oxoacyl-acyl-carrier-protein synthase, crotonyl-acyl-carrier-protein hydratase, enoyl acyl carrier protein reductase (NADPH2,A-specific, (7),oleoyl-acyl-carrier-protein hydrolase

Enriched Canonical Pathways	-log(p-value)	Ratio	Number of genes from dataset	Total number of genes in canonical pathways	Biomarkers
Fatty Acid Biosynthesis Initiation II	1.430	0.500	1.000	2.000	3-oxoacyl-carrier-protein synthase, acyl-carrier-protein S-acetyltransferase

4. Conclusion

The differential gene expression analysis identified genes that were upregulated and genes that were downregulated in clinical stages of breast cancer from stage I to stage IV. The biological pathway analysis revealed the significant pathways in early (stage I and II) and late stages (stage III and IV) of breast cancer. The connections among breast cancer genes and approved drugs (for other pharmacological uses) was studied. The different categories of drugs as enzyme inhibitors, CNS agents, antioxidants, DNA & cell wall synthesis inhibitors showed good connectivity scores against breast cancer. The drug targets showed enrichment of fatty acid biosynthesis, melatonin degradation, tryptophan degradation, role of JAK family kinases, IL-17A & IL-22 signaling, nitric oxide signaling, serotonin biosynthesis pathways. The frequency of drugs against different pathways was calculated and it was found that the drugs as Nortriptyline, AZD-6482, Acitretin and GW-507 were found to target multiple genes in interleukin pathway, enzymatic pathway & GPCR signaling and drugs as Caffeine, Canertinib and Triciribine were found to target multiple genes in interleukin pathway, enzymatic pathway, GPCR signaling and metabolic pathways. These drugs showed strong connections to other cancers as well such as prostate cancer, lung cancer, liver cancer, colon cancer, kidney cancer, melanomas. Our approach seeks to maximize the use of datasets, tools and techniques to understand role of approved drugs in breast cancer. We hope that a single or combination of few of these drugs will contribute to drug design and discovery against breast cancer.

Declarations

Conflict of interest

No potential conflict of interest was reported by the author(s).

Acknowledgement

SA wishes to thanks to Department of Science and Technology-Innovation in Scientific Pursuit for Inspired Research (DST-INSPIRE), New Delhi for providing fellowship.

Author contribution

The study was devised by SA, AD and SKK. The manuscript was written by SA and reviewed by AD and SKK.

References

1. Waks AG, Winer EP. JAMA. 2019;321:288–300.
2. Tong CWS, Wu M, Cho WCS, To KKW. Front Oncol. 2018;8:227.
3. Calaf GM, Zepeda AB, Castillo RL, Figueroa CA, Arias C, Figueroa E, Farias JG. Int J Oncol. 2015;47:437–45.
4. Szostakowska M, Trebinska-Stryjewska A, Grzybowska EA, Fabisiwicz A. Breast Cancer Res Treat. 2019;173:489–97.
5. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. Science. 2006;313:1929–35.
6. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu X, Zhao WN, Read-Button W, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, D. E. Root, B. Wong and T. R. Golub. Cell. 2017;171:1437–52 e1417.
7. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, Doig A, Guilleams T, Latimer J, McNamee C, Norris A, Sanseau P, Cavalla D, Pirmohamed M. Nat Rev Drug Discov. 2019;18:41–58.

8. Xue H, Li J, Xie H, Wang Y. *Int J Biol Sci.* 2018;14:1232–44.
9. Park K. *Translational and Clinical Pharmacology*, 2019, 27, 59–63.
10. Li JR, Sun CH, Li W, Chao RF, Huang CC, Zhou XJ, Liu CC. *Nucleic Acids Res.* 2016;44:D944–51.
11. Hutschenreuther A, Bigl M, Hemdan NY, Debebe T, Gaunitz F, Birkenmeier G. *Int J Mol Sci*, 2016, 17.
12. Guo Y, Zhang Y, Yang X, Lu P, Yan X, Xiao F, Zhou H, Wen C, Shi M, Lu J, Meng QH. *Cancer Biol Ther.* 2016;17:169–80.
13. Fonseca-Sanchez MA, Rodriguez Cuevas S, Mendoza-Hernandez G, Bautista-Pina V, Arechaga Ocampo E, Hidalgo Miranda A, Quintanar Jurado V, Marchat LA, Alvarez-Sanchez E, Perez Plasencia C, Lopez-Camarillo C. *Int J Oncol.* 2012;41:670–80.
14. Rulli A, Carli L, Romani R, Baroni T, Giovannini E, Rosi G, Talesa V. *Breast Cancer Res Treat.* 2001;66:67–72.
15. Lamkin DM, Sung HY, Yang GS, David JM, Ma JC, Cole SW, Sloan EK. *Psychoneuroendocrinology.* 2016;51:262–70.
16. Pai VP, Marshall AM, Hernandez LL, Buckley AR, Horseman ND. *Breast Cancer Res.* 2009;11:R81.
17. Kopparapu PK, Tinzl M, Anagnostaki L, Persson JL, Dizeyi N. *Anticancer Res.* 2013;33:363–70.

Figures

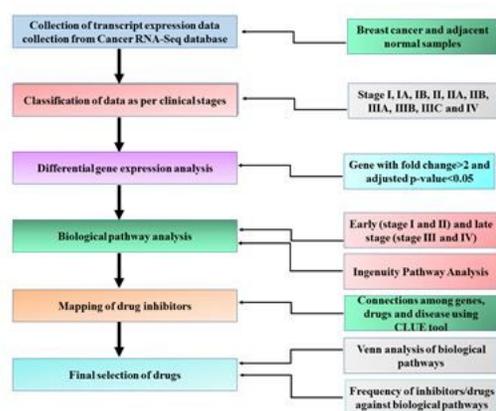


Figure 1

The flow of present employed computational approach

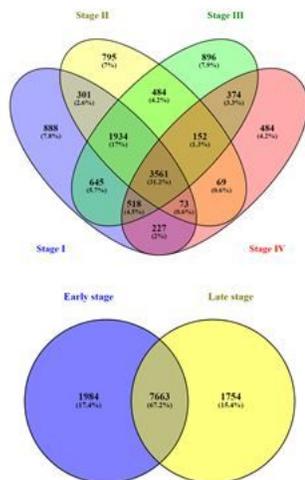


Figure 2

Venn diagram showing common and specific genes in stages I, II, III and IV a) and early and late stages (b) of breast cancer.

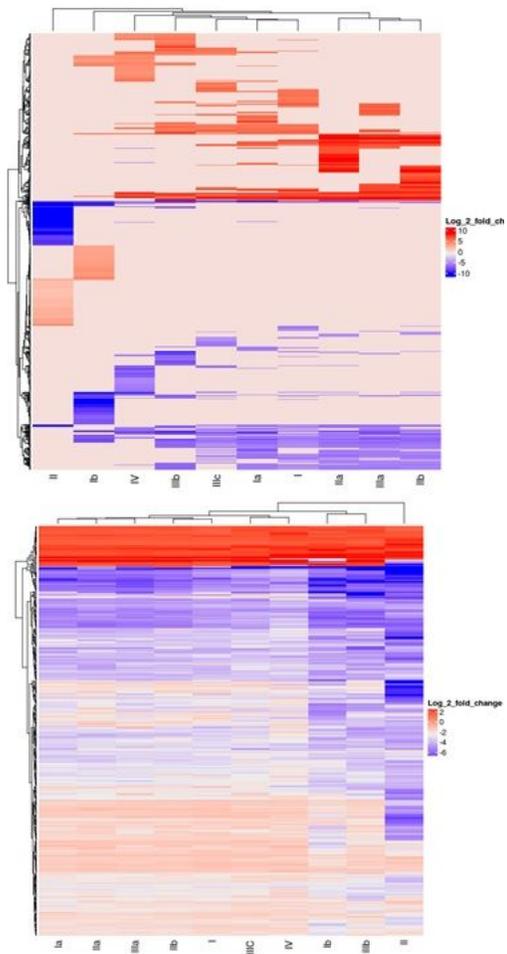


Figure 3

Heatmap showing all differentially expressed genes (a) and common genes (b) in different stages.

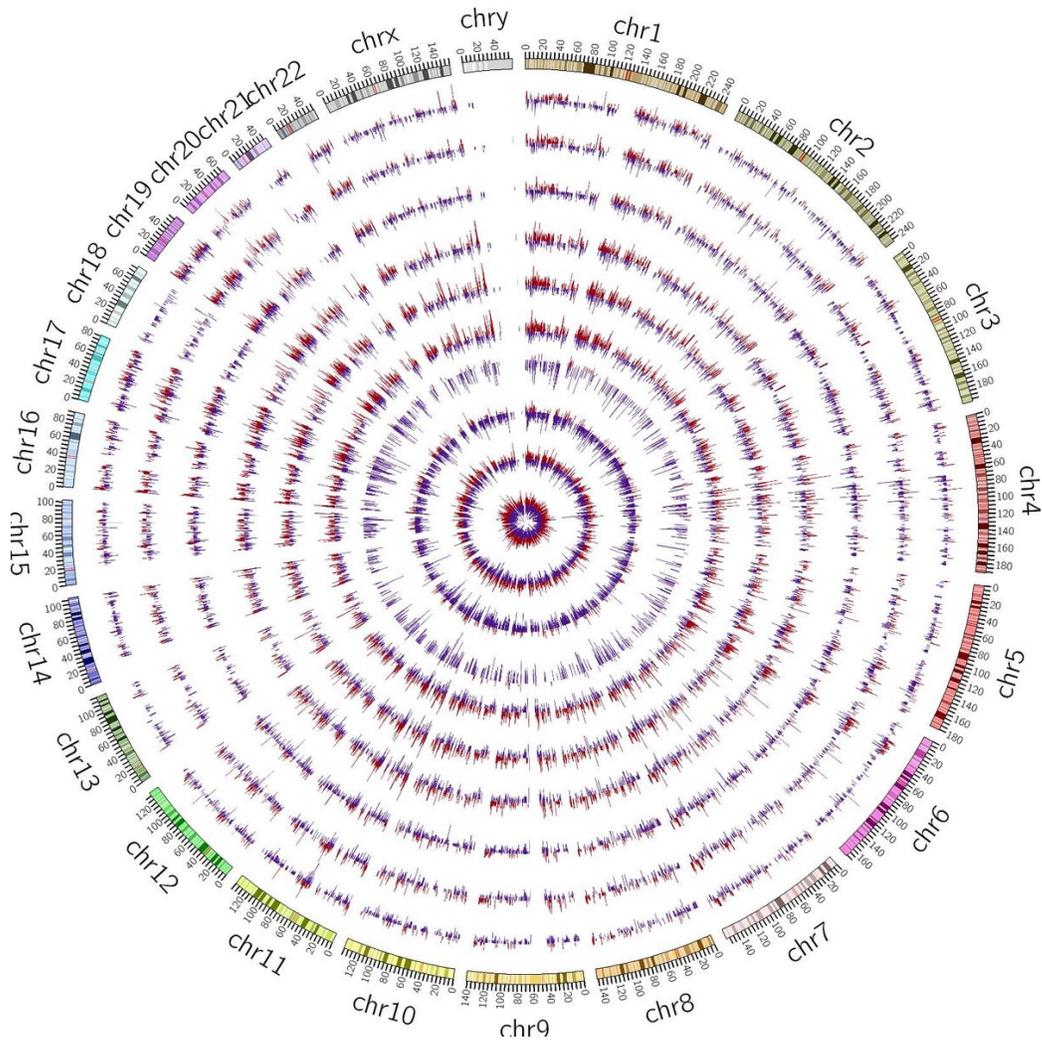


Figure 4

Heatmap showing all differentially expressed genes (a) and common genes (b) in different stages.

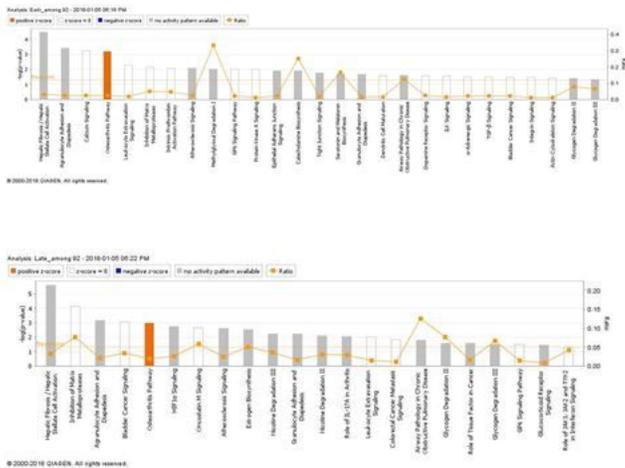


Figure 5

Bar diagrams representing biological pathways in a) early stage and b) late stage of breast cancer.

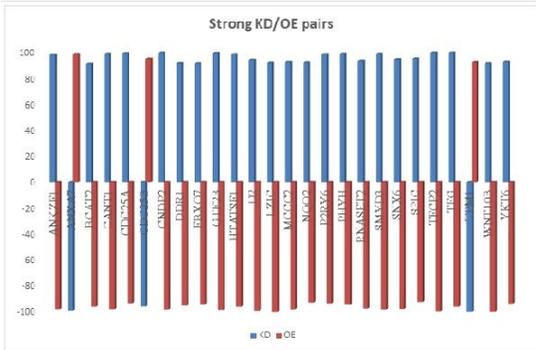
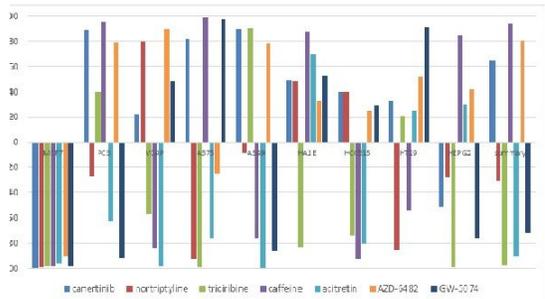
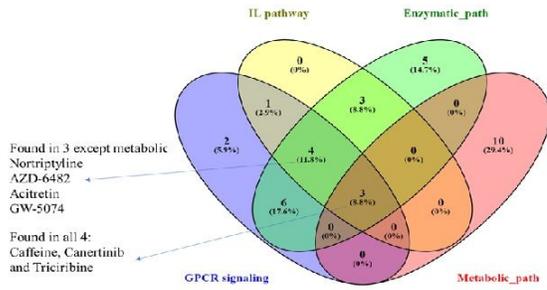


Figure 7

a) Venn analysis of drugs against different pathways. b) Connectivity scores of drugs against different cell line as MCF7, PC3, VCAP, A375, A549, HA1E, HCC515, HT29 and HEPG2. C) Strong Knockdown-overexpression (KD/OE) gene pairs.

Supplementary Files

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

- [Supplementaryfigures.docx](#)
- [Supplementarytables.xlsx](#)