Pathway Interactions Based on Drug-Induced Datasets

Shinuk Kim

Department of Civil Engineering, Sangmyung University, Cheonan, Republic of Korea.

ABSTRACT: In this study, we identified enrichment pathway connections from MCF7 breast cancer epithelial cells that were treated with 87 drugs. We extracted drug-treated samples, where the sample size was greater than or equal to 5. The drugs included 17-allylamino-geldanamycin, LY294002, trichostatin A, valproic acid, sirolimus, and wortmannin, which had sample sizes of 11, 8, 7, 7, 7, and 5, respectively. We found meaningful pathways using gene set enrichment analysis and identified intradrug and interdrug pathway interactions, which implied the influence of drug combination. Among the top 20 enrichment pathways that were wortmannin induced, there were a total of 37 intradrug pathway interactions via common genes. Thirty-seven pathway interactions were induced by valproic acid, 11 induced by trichostatin A, 20 induced by LY294002, and 59 induced by sirolimus, all via common genes. The number of interdrug-induced pathway interactions ranged from one pair of pathways to 23. The pair of ERBB_SIGNALING and INSULIN_SIGNALING pathways showed the highest score from a pair of 2 individual drugs. The highest number of pathway interactions was observed between the drugs 17-allylamino-geldanamycin and LY294002.

KEYWORDS: drug-induced pathway, cancer, gene set analysis, interdrug and intradrug pathway interactions

Introduction

Even after an enormous financial and time-consuming investment of typically 10-15 years, the approval rate of new drugs remained stagnant at 2.01%, on average, from 2000 to 2008.[1,2] As a result, drug repositioning using existing drugs for nontarget diseases has become an alternative methodology for deriving new drugs to market.[3–6] One method for this is a computational approach based on chemical-protein interactions, chemical-disease interactions, and omic interactions.[3,7–10] A major advantage of computational drug repositioning is reducing the development risk as it relates to finances and time for bringing drugs to market.

Lamp[11,12] introduced the concept of connectivity maps, which refers to the functional connections among diseases, genetic abnormalities, and drug efficacy. Connectivity maps are obtained from information about small molecules, chemicals, and drugs. Classical drug studies have focused on one gene, one drug, and one disease target. However, these single-target studies overlook the possibility of unintended beneficial effects and synergy effects of multitarget interacting drugs. Regarding this, Cheng et al[13] proposed network-based drug repurposing using drug-induced samples and protein-protein interactions. Yoo et al[14] introduced the drug signatures database (DSigDB) of drug and chemical compounds for drug repurposing. Takayuki[9] reported the drug repositioning for dengue fever using a multiomic analysis including genomics, proteomics, and interactomics.

Pathways are functional gene sets that work together to control cell processes and inform molecular interactions and/or reactions provided by the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway,[15] BioCarta,[16] and pathway ontology.[17] Abnormal activation of genes in the pathway may lead to cancer or other diseases. Therefore, finding enrichment pathways or gene sets from drug-induced datasets can help to discover a promising target candidate for treatment of these diseases. In addition, uncovering a pathway interaction can interpolate the insight of the drug efficacy systemically.

In this study, we are interested in the drug molecules associated with gene sets referred to in a pathway instead of the individual gene, noting interpathway and intrapathway interactions using drug-induced datasets. An intrapathway interaction refers to interactions from enrichment pathways derived from a single-drug analysis. An interpathway interaction means one pathway was discovered from one drug analysis and another pathway was discovered from an analysis based on a different drug.

Materials and Methods

Materials

We downloaded a total of 22 268 probes, 182 drug-induced samples, and 36 control samples from the MCF7 breast cancer cell line in March 2018 from the Gene Expression Omnibus website (updated 2017). Gene expression data were generated using the Affymetrix HT human genome U133A array (GPL3921) platform by Lamb.[11] From among the 182 drug-induced samples containing 87 unique drugs, we extracted the following 6 drugs where the sample size was greater than or equal to 5: 11 samples of 17-allylamino-geldanamycin (17-AAG; concentration: 1 μM); 8 samples of LY294002 (10 and 0.1 μM); 7 samples each of sirolimus (0.1 μM), trichostatin A (0.1 and 1 μM), and valproic acid (10 mM, 1 mM, 500 μM, 2 mM, and 50 μM); and 5 samples of wortmannin (0.01 and
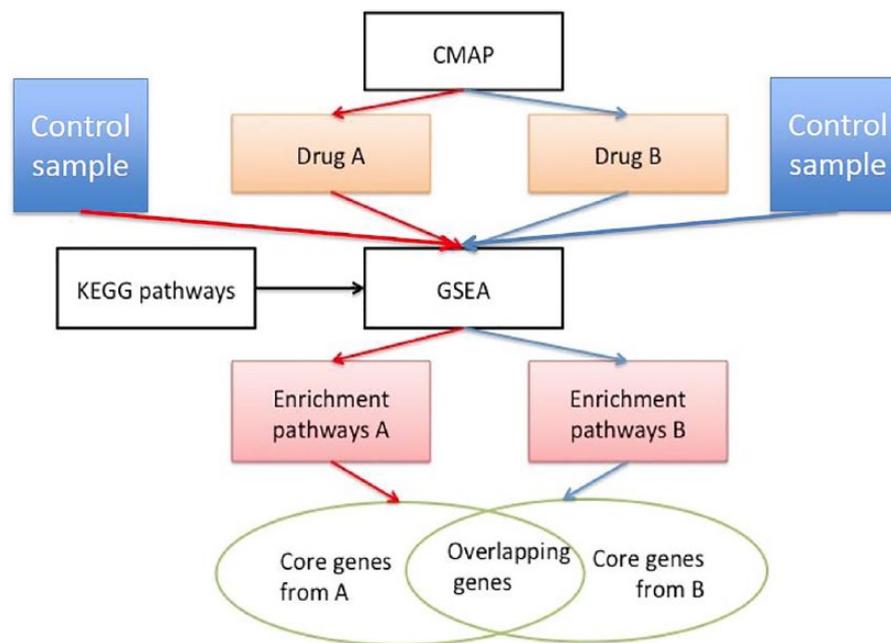
Figure 1. The overall flowchart for the method of drug-induced pathway interactions. First, different kinds of drug-induced samples are extracted and compared with the control samples via GSEA. Enrichment pathways from each drug are selected. Red arrows represent drug A performance, while blue arrows represent drug B performance. GSEA indicates gene set enrichment analysis.

1 μM). The sample sizes of the remainder of the drugs were less than 5. Although the drugs were administered at different concentrations, we ignored the induced concentration. Here, we briefly describe 6 of the study drugs. Sirolimus, also known as rapamycin, blocks the transcriptional activation of cytokines, thereby inhibiting cytokine production. Wortmannin is a steroid metabolite and a nonspecific, covalent inhibitor of PI3Ks. LY294002 is an inhibitor of phosphatidylinositol 3-kinase. Valproic acid is a fatty acid with anticonvulsant properties used in the treatment of epilepsy, with unknown mechanisms of its therapeutic path.[18] 17-AAG, also known as tanespimycin, was developed for treating cancer and solid tumors.

We also downloaded gene set enrichment analysis (GSEA) tools and molecular signatures database v.6 (MSigDB) from http://software.broadinstitute.org/gsea.

Methods

To obtain abnormal gene sets by comparing individual drug-induced samples to 36 control samples, we performed GSEA.[19] GSEA generates enrichment pathways ordered by false discovery ratio or normalized enrichment score, and leading-edge gene sets refer to core genes obtained by using the highest enrichment score for each pathway calculated by the Kolmogorov-Smirnov-like (KS) test, as in Subramanian et al.[19,20]

Here, we extracted the top 20 enrichment pathways from GSEA, ordered by normalized enrichment score, and extracted the leading-edge genes (or core genes) of each enrichment pathway. After we identified the enrichment pathways with core genes, we determined the strength pathway interaction using core genes that referred to the drug combination score

(DCSC) as a summation of the absolute values of the signal-to-noise of common genes. Computational drug repositioning approaches are usually based on similarities and shared properties.[2] Using abnormally expressed common gene sets is one of the valuable methods for presenting similarities:

$$DCSC = |s_A| + |s_B|, S_A$$
$$= \frac{M_c + M_n}{N_c - N_n} \cdot S_B = \frac{M_c + M_n}{N_c - N_n},$$

where s_A and s_B are the signal-to-noise of core genes from pathways identified from drug A and drug B, respectively. The signal-to-noise is calculated by dividing the average signal by the standard deviation of the noise. $M_c, M_n, N_c,$ and N_n represent the average of cancer datasets, average of normal datasets, total standard deviation of cancer, and standard deviation of normal, respectively.

In this study, we identified 2 types of interactions: intrapathway and interpathway interactions using drug-induced datasets. An intrapathway interaction refers to interactions from enrichment pathways derived from a single-drug analysis. We first extracted enrichment pathways and then created pairs of the enrichment pathways based on common genes. Finally, we calculated DCSC by summation of the signal-to-noise of the common genes of both pathways. An interpathway interaction refers to an interaction between one pathway discovered from one drug analysis and another pathway being discovered from analysis of a second drug.

A schematic diagram of the interpathway interaction method is presented in Figure 1. Red arrows represent

generation of enrichment pathway A based on drug A, and blue arrows represent generation of enrichment pathway B based on drug B. Red and blue lines represent intrapathway interactions.

Here, we focused on different pathway interactions from different drug-induced samples. Although not all genes were commonly identified as core genes from the same pathways, and these pathways were generated from different drugs, the role of the drug in the pathway is functionally the same. Therefore, we ignored the efficacy of 2 different drugs generating the same pathway.

Results

Intrapathway interactions

To identify the pathway interactions, we selected 20 enrichment pathways that were generated by comparing each drug-induced sample to the control samples. Among the obtained pathway interactions using common genes, we presented one pair of pathways discovered by the most common genes. The result of 17-AAG compared with the control datasets derived 20 enrichment pathways. Using those 20 pathways, we obtained 21 pathway interactions via common genes. One of the 21 pairs of pathways is GLYCOSAMINOGLYCAN_DEGRADATION and GLYCAN_STRUCTURES_DEGRADATION, via 5 common genes (*ARSB, HPSE2, HYAL1, IDUA*, and *SPAM1*). For LY294002, we have a total of 20 pathway interactions, including the pair, APOPTOSIS and INSULIN SIGNALING PATHWAY, which had the highest number of common genes, 12: *AKT* (2, 3), *PIK3* (CA, CB, CG, R3, R5), and *PRKA* (CA, CB, R1B, R2B, R3B). The PIK3 family (as PI3K protein) is known to be inhibited by LY294002 drugs[21] (from Selleckchem). Functional treatment of LY294002 through PIK3 achieves balance of apoptosis, which plays an important role in cancer.[22,23] According to STITCH 5,[7] the false discovery ratios (FDRs) of the insulin signaling pathway and apoptosis discovered by LY294002 are 1.11e-14 and 1.7e-9, respectively, with 5 common genes: *AKT1* and *PIK3* (CA, CB, CD, CG).

For trichostatin A, a total of 11 pathway interactions were obtained. Among them, the pair INSULIN SIGNALING PATHWAY and ERBB SIGNALING PATHWAY shared 11 common genes of *AKT3, CBL, CBLB, KRAS, MAPK10, MAPK8, PIK* (3CD, R2, R5), and *SHC* (2, 3). There were 37 pathway interactions obtained for valproic acid, one of them being ARACHIDONIC ACID METABOLISM and LINOLEIC ACID METABOLISM with 12 common genes: *ALOX15 CYP2C* (18, 19, 8, 9) and *PLA2G* (2A, 2D, 2F, 3, 4A, 5, 6). The pathway interactions of sirolimus totaled 59, one of them being the ERBB SIGNALING PATHWAY and GLIOMA via 24 common genes including *AKT3*. Those 2 pathways also appeared in STITCH 5[7] with *MTOR* and *AKT1* common genes. For wortmannin, a total of 37 pathway interactions were obtained, one being HISTIDINE METABOLISM

and TYROSINE METABOLISM via 5 common genes (*AOC2, AOC3, DDC, MAOA*, and *PRMT8*). For detailed information, refer to the supplements.

Interpathway interaction

Among 15 drug pairs obtained from the 6 extracted drugs, in Table 1, we presented a pair of drugs and a pair of pathways with the largest DCSC pathway interaction. Here, we describe the top 3 DCSC pathway interactions in more detail.

With 17-AAG and LY294002, 20 pairs of pathway interactions were derived, 8 of which interacted along the same functional pathway, while the remaining 12 pairs interacted along 2 different pathways from each drug. One pair of these 12 pairs of pathways is ERBB SIGNALING PATHWAY and INSULIN SIGNALING PATHWAY via 13 common genes with DCSC 8.19.

We individually discovered the ERBB (epidermal growth factor receptors [EGFR]) signaling pathway (with FDR 1.4e-4) from 17-AAG and the insulin signaling pathway (with 1.11e-14) from LY294002 with common gene *AKT1* sourced from STITCH 5.[7]

For 17-AAG and sirolimus, we discovered 12 pairs of pathway interactions, 9 of which followed different functional pathway interactions: one pair is ERBB_SIGNALING_PATHWAY and GLIOMA via 12 common genes with 7.53 DCSC, and another interesting pair of pathways is ERBB_SIGNALING_ PATHWAY and INSULIN_SIGNALING_PATHWAY, via 10 common genes with 6.63 DCSC. The ERBB signaling pathway from 17-AAG and glioma and the insulin signaling pathway from sirolimus were also found from STITCH 5[7] with FDRs of 1.1e-4, 1.04e-2, and 2.77e-9, respectively.

Among 6 pairs of different functional pathways out of a total of 11 pairs from 17-AAG and valproic acid, a pair is LINOLEIC ACID METABOLISM and ARACHIDONIC ACID METABOLISM, via 7 common genes with 5.88 DCSC.

Among 4 pairs of different functional pathways out of a total of 5 pairs from wortmannin and sirolimus (rapamycin), a pair of pathways is EPITHELIAL_CELL_SIGNALING_ IN_HELICOBACTER_PYLORI_INFECTION and ERBB_SIGNALING_PATHWAY, via 6 common genes with DCSC 3.35. Those pathways were not discovered in STITCH.

With LY294002 and sirolimus, we found that a pair containing an insulin signaling pathway and an ERBB signaling pathway has the greatest DCSC (9.16) of 22 pairs of pathways (with 17 common genes) and confirmed that the insulin signaling pathway (FDR 1.11e-14) from LY294002 and the ERBB signaling pathway (FDR 3.89e-6) from sirolimus were in STITCH 5 with 3 common genes, *MTOR, AKT1*, and *RPS6KB1*.

Table 1 shows that the interesting pair of insulin signaling pathway and ERBB signaling pathway was present in 5 pairs of

Table 1. Pairs of pathway interactions from 2 different drugs with the greatest DCSC score.

DRUGS	T_INT/D_INT	PATHWAY INTERACTION WITH HIGHEST DCSC	DCSC	NO. OF COMMON GENES
17-AAG_ly24	**20/12**	**ERBB SIGNALING PATHWAY/INSULIN SIGNALING PATHWAY**	8.19	13
17-AAG_sir	12/9	ERBB SIGNALING PATHWAY/GLIOMA	7.53	12
17-AAG_tri	**6/3**	**ERBB SIGNALING PATHWAY/INSULIN SIGNALING PATHWAY**	4.61	8
17-AAG_VPA	11/6	LINOLEIC ACID METABOLISM/ARACHIDONIC ACID METABOLISM	5.88	7
17-AAG_wort	3/1	ERBB SIGNALING PATHWAY/EPITHELIAL CELL SIGNALING IN HELICOBACTER PYLORI INFECTION	3.94	6
ly24_sir	**24/22**	**INSULIN SIGNALING PATHWAY/ERBB SIGNALING PATHWAY**	9.16	17
ly24_tri	**6/2**	**INSULIN SIGNALING PATHWAY/ERBB SIGNALING PATHWAY**	5.19	10
ly24_VPA	10/6	LINOLEIC ACID METABOLISM/ARACHIDONIC ACID METABOLISM	5.95	7
ly24_wort	3/1	APOPTOSIS/ SMALL CELL LUNG CANCER	2.55	6
sir_tri	**11/10**	**INSULIN SIGNALING PATHWAY/ERBB SIGNALING PATHWAY**	4.88	10
sir_VPA	15/11	ARACHIDONIC ACID METABOLISM/LINOLEIC ACID METABOLISM	5.70	8
tri_VPA	2/2	ERBB SIGNALING PATHWAY/T CELL RECEPTOR SIGNALING PATHWAY	3.53	7
wort_sir	5/4	EPITHELIAL CELL SIGNALING IN HELICOBACTER PYLORI INFECTION/ ERBB SIGNALING PATHWAY	3.35	6
wort_tri	2/0	No different pathway interaction	—	—
wort_VPA	1/0	No different pathway interaction	—	—

The first column presents the pair of drugs, the second column indicates total pairs of pathway interactions (T_int) and number of pairs of different pathway interactions (D_int) generated from each drug, the fourth column represents the score of the pathway interaction, and the fifth column represents the number of common pathways.
Abbreviations: 17-AAG, 17-allylamino-geldanamycin; DCSC, drug combination score; ly24, LY294002; sir, sirolimus; tri, trichostatin A; VPA, valproic acid; wort, wortmannin.

drugs. The result suggests that even though 5 pairs of drugs demonstrated the same functional interaction, the strength of the pair LY294002 and sirolimus is greater than the strength of the remaining pairs. That implies that LY294002 and sirolimus are better treatments for breast cancer than the remaining pairs.

Wortmannin was suggested for treatment of Helicobacter pylori in gastric epithelial cells.[24] The ERBB signaling pathway is involved in the development of human cancer and pursued as the target of the treatment.[25] Figure 2 presents our findings with those of previous studies. Sirolimus is suggested to target the mammalian target of rapamycin via the mechanistic target of rapamycin (mTOR) signaling pathway in breast cancer,[28] presented as red arrows (Figure 2A). Chiu suggested that sirolimus is the inhibitor of both the mTOR and EGFR signaling pathways in pancreatic neuroendocrine tumors,[26] shown as blue arrows (Figure 2B). Hynes[25,27,29] reported that lapatinib is used for breast cancer treatment via the ERBB signaling pathway, shown as a green arrow in Figure 2C. Comprehensively, our findings that sirolimus plays an important role in breast cancer treatment via the ERBB signaling pathway matches the results of previous studies.

An insulin receptor has arisen as a new target for cancer therapy via PI3K.[30] A recently published study states that PI3K inhibitor LY294002 resists pancreatic cancer cells[31] through the AKT signaling pathway. In our findings, the 17 common genes between the ERBB signaling pathway and insulin signaling pathway based on sirolimus and LY294002 include *PIK3* (CD, CG, R3, R5).

Conclusions

In this study, we analyzed statistical findings of pathway interactions using drug-induced datasets. Of the 6 extracted drugs and 15 analyzed pairs of drugs, we were especially interested in the connections of different pathways from different drug pairs, which imply drug combination effects. The highest DCSC interactions were selected from 5 out of 13 pairs of drugs involving interactions of the ERBB signaling pathway and the insulin signaling pathway. Among the 5 pairs, the highest DCSC of 9.16, indicating the strongest combination efficacy of the drugs, was for sirolimus and LY294002.

Using open source STITCH5, some pathway interactions were confirmed. However, as we used disease-specific datasets, some pathways we found did not appear in the previous computational results. Our computational approach supports the exploration of pathway interactions based on drug pairs to discover the interactions of drugs, pathways, and diseases and to further identify suitable drugs for repurposing/repositioning.
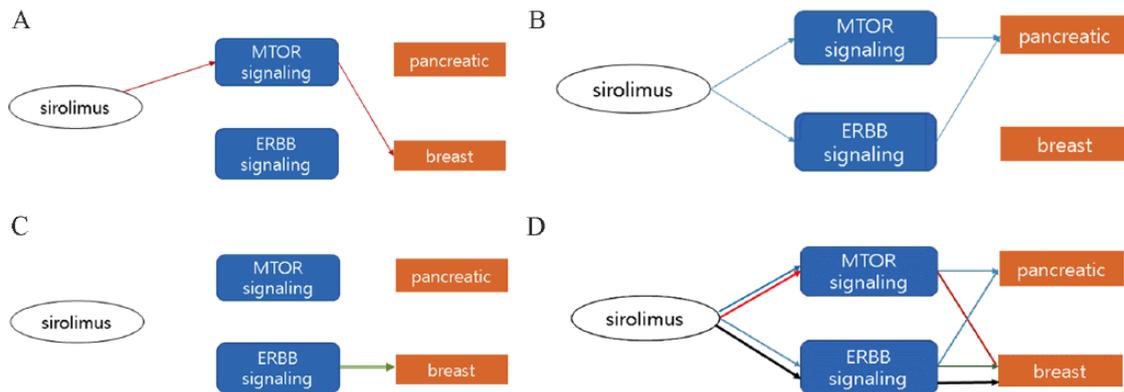
Figure 2. Interactions of drugs, pathways, and diseases in our and previous studies. (A) Red arrows present the drug to cancer pathway via MTOR as in Chiu et al.[26] (B) Blue arrows present sirolimus to pancreatic cancer pathways via MTOR and ERBB as in Hynes and MacDonald.[27] (C) The ERBB signaling pathway is the primary choice for the breast cancer treatment.[28–30] (D) Description of our findings (black arrows) compared with previous studies.

Author Contributions

SK designed and performed research, analyzed data, and wrote the paper.

REFERENCES

1. Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov*. 2004;3:673–683.
2. Yeu Y, Yoon Y, Park S. Protein localization vector propagation: a method for improving the accuracy of drug repositioning. *Mol Biosyst*. 2015;11:2096–2102.
3. Napolitano F, Zhao Y, Moreira VM, et al. Drug repositioning: a machine-learning approach through data integration. *J Cheminform*. 2013;5:30.
4. Walters WP, Green J, Weiss JR, Murcko MA. What do medicinal chemists actually make? A 50-year retrospective. *J Med Chem*. 2011;54:6405–6416.
5. Iorio F, Bosotti R, Scacheri E, et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci U S A*. 2010;107:14621–14626.
6. Xue H, Li J, Xie H, Wang Y. Review of drug repositioning approaches and resources. *Int J Biol Sci*. 2018;14:1232–1244.
7. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res*. 2016;44:D380–D384.
8. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P. STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Res*. 2008;36:D684–D688.
9. Amemiya T, Gromiha MM, Horimoto K, Fukui K. Drug repositioning for dengue haemorrhagic fever by integrating multiple omics analyses. *Sci Rep*. 2019;9:523.
10. Napolitano F, Carrella D, Mandriani B, et al. gene2drug: a computational tool for pathway-based rational drug repositioning. *Bioinformatics*. 2018;34:1498–1505.
11. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*. 2006;313:1929–1935.
12. Lamb J. The Connectivity Map: a new tool for biomedical research. *Nat Rev Cancer*. 2007;7:54–60.
13. Cheng F, Desai RJ, Handy DE, et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun*. 2018;9:2691.
14. Yoo M, Shin J, Kim J, et al. DSigDB: drug signatures database for gene set analysis. *Bioinformatics*. 2015;31:3069–3071.
15. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
16. Nishimura D. BioCarta. *Biotech Software Internet Rep*. 2001;2:117–120.
17. Petri V, Jayaraman P, Tutaj M, et al. The pathway ontology: updates and applications. *J Biomed Semantics*. 2014;5:7.
18. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46:D1074–D1082.
19. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545–15550.
20. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics*. 2007;23: 3251–3253.
21. Mbengue A, Bhattacharjee S, Pandharkar T, et al. A molecular mechanism of artemisinin resistance in Plasmodium falciparum malaria. *Nature*. 2015;520: 683–687.
22. Jean S, Kiger AA. Classes of phosphoinositide 3-kinases at a glance. *J Cell Sci*. 2014;127(Pt. 5):923–928.
23. Franke TF, Hornik CP, Segev L, Shostak GA, Sugimoto C. PI3K/Akt and apoptosis: size matters. *Oncogene*. 2003;22:8983–8998.
24. Kwok T, Backert S, Schwarz H, Berger J, Meyer TF. Specific entry of Helicobacter pylori into cultured gastric epithelial cells via a zipper-like mechanism. *Infect Immun*. 2002;70:2108–2120.
25. Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer*. 2005;5:341–354.
26. Chiu CW, Nozawa H, Hanahan D. Survival benefit with proapoptotic molecular and pathologic responses from dual targeting of mammalian target of rapamycin and epidermal growth factor receptor in a preclinical model of pancreatic neuroendocrine carcinogenesis. *J Clin Oncol*. 2010;28: 4425–4433.
27. Hynes NE, MacDonald G. ErbB receptors and signaling pathways in cancer. *Curr Opin Cell Biol*. 2009;21:177–184.
28. Malaguti P, Vari S, Cognetti F, Fabi A. The Mammalian target of rapamycin inhibitors in breast cancer: current evidence and future directions. *Anticancer Res*. 2013;33:21–28.
29. Wang Z. ErbB receptors and cancer. *Methods Mol Biol*. 2017;1652:3–35.
30. Malaguarnera R, Belfiore A. The insulin receptor: a new target for cancer therapy. *Front Endocrinol (Lausanne)*. 2011;2:93.
31. Wang Y, Kuramitsu Y, Baron B, et al. PI3K inhibitor LY294002, as opposed to wortmannin, enhances AKT phosphorylation in gemcitabine-resistant pancreatic cancer cells. *Int J Oncol*. 2017;50:606–612.