Research Article

Analysis of the mechanism underlying the effects of cyclophosphamide against triple-negative breast cancer by an integrative bioinformatics approach

Maojing Wang1, Wenwen Li2, Ding Li3 and Zhiwu Han4*

1Department of Cardiology, The Affiliated Hospital of Qingdao University, Qingdao, China
2Department of Hematology, Qingdao women and children’s Hospital Qingdao, China
3Department of TCM, The Affiliated Hospital of Qingdao University, Qingdao, China
4Department of Pharmacy, The Affiliated Hospital of Qingdao University, Qingdao, China

Abstract

Patients with Triple-Negative Breast Cancer (TNBC), a highly heterogeneous and invasive subtype of breast cancer, do not benefit from hormonal therapy or trastuzumab; therefore, chemotherapy is considered the only option. We explored the effect of the chemotherapeutic drug cyclophosphamide (Cytoxan; Cy) on TNBC by an integrated bioinformatics approach.

Methods: Pharm Mapper, Gene Cards, and Swiss Target Prediction were used to identify potential targets of Cy. Differentially expressed genes (DEGs) in TNBC were screened out from four GEO datasets. Common genes were further evaluated by a protein-protein interaction network analysis, core gene identification, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, docking assays, and survival and immune cell infiltration analyses.

Results: We collected information on 1638 drug targets and 494 DEGs in TNBC, including 267 up-regulated and 227 down-regulated genes. In total, 68 overlapping genes were identified as common targets. Ten core genes were identified in network analysis; GO and KEGG analyses revealed enrichment for DNA damage and many signaling pathways. Four core gene targets were verified by molecular docking. Kaplan–Meier analysis revealed that two core genes were significantly related to an adverse overall survival; furthermore, immune infiltration analysis suggested that Cy affects the microenvironment.

Conclusions: Our integrative bioinformatics approach revealed that the anti-TNBC effect of Cy was mediated by DNA damage-related genes and many pathways. These findings provide a basis for further functional studies aimed at improving outcomes in TNBC.

Introduction

Triple-Negative Breast Cancer (TNBC) is a breast cancer subtype characterized by no or low expression of Estrogen Receptor (ER), Progesterone Receptor (PR), and Human Epidermal Growth Factor Receptor 2 (HER2) [1]. Compared with other breast cancer subtypes, the clinical–pathological features of TNBC usually include larger tumors, a worse histologic grade, and higher proliferation index [2]. With high heterogeneity and invasiveness, a study of National Cancer Institute Surveillance, Epidemiology and End Results (SEER) from 2010 to 2012 revealed that TNBC has a worse overall survival and cause-specific survival than non–TNBC [3]. The 5-year overall survival rate of TNBC is <78.5%, and there is a
high risk of distant recurrence in the first 3 to 5 years following diagnosis [3,4].

Although the mechanisms underlying TNBC are not fully understood, approximately 10% of patients have inactivating mutations in BRCA1 or BRCA2 [5]. Transcriptomic profiling has shown that multiple genomic transformations activate the PI3K pathway, including the activation of neoplastic genes and inactivation of tumor suppressor genes, thus altering the PI3K signaling pathway [6]. PI3K changes at the genetic level result in aberrations in the AKT/mTOR pathway [7]. The Hedgehog pathway, which is a highly conserved and critical signaling pathway, plays an important role in regulating embryonic mammary gland induction and the ductal architecture and differentiation in lactation [7]. Increasing evidence suggests that upregulation of the Hh ligand sonic hedgehog is correlated with inferior overall survival in TNBC [8]. Epidermal Growth Factor Receptor (EGFR) is a growth indicator in tumors; in TNBC, the upregulation of EGFR is associated with tumor progression [9], and is always coupled with high Vascular Endothelial Growth Factor (VEGF) expression, affecting the RAS–MAPK signaling pathway [10]. Many factors involved in the proliferation of TNBC have been reported, including AMP-Activated Protein Kinase (AMPK) [11]. Mouse Double Minute 2 Homolog (MDM2) [12] and Metadherin (MTDH) [13].

Cyclophosphamide (Cytoxan; Cy) has been a widely used chemotherapeutic agent for immunotherapy in various cancer types since it was first synthesized. It is usually combined with platinum-based drugs [14], taxane-based drugs [15] and anthracycline-based drugs [16] as Neoadjuvant Chemotherapy (NACT) in TNBC. It has been reported to function by three distinct antitumor mechanisms, namely, the inhibition of nuclear DNA replication, prevention of DNA synthesis and RNA transcription from the affected DNA, and mispairing of nucleotides, leading to mutations.

Cy often leads to chemo-toxicity; however, in some cases, it is the only available therapeutic choice to reduce and prevent tumor relapse. To reveal the mechanisms by which Cy exerts anti-TNBC effects, we used a network pharmacology approach to study its targets and related pathways in the treatment of TNBC, with validation by survival analyses. The workflow is shown in Figure 1.

Material and methods

Cy target information

Cy target information was extracted from Pharm Mapper (http://lilab-ecust.cn/pharmmapper/), GeneCards (https://www.genecards.org/), and Swiss Target Prediction (http://www.swisstargetprediction.ch/).

TNBC target information from GEO

Four gene expression datasets were downloaded from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov). All datasets were separated into two groups (TNBC and non-TNBC) based on the original data, and the basal–like subtype was classified as TNBC [17]. Table 1 provides detailed information for the four GEO datasets.

Identification of differentially expressed genes

Using the R computing environment (R×64 3.5.3, https://cran.r-project.org/), a Robust Rank Aggregation (RRA) analysis was performed to evaluate the four datasets [18,19].

The limma package in R http://www.bioconductor.org/packages/release/bioc/html/limma.html was used to identify differentially expressed genes (DEGs) between the TNBC and non–TNBC groups. Hub genes with |Log2FC| > 1 and adjusted-P < 0.05 were regarded as DEGs.

Overlap between TNBC-related DEGs from GEO and Cy targets

The DEGs in TNBC and potential targets of Cy were uploaded to E Venn (http://www.ehbio.com/test/venn/#/) to obtain 68 overlapping genes, which are visually displayed in a Venn diagram in Figure 2. All 68 overlapping genes were considered common targets and included in further analyses.

Constructions of PPI networks

Protein–protein interaction (PPI) networks of common targets were constructed using STRING (https://string-db.org/), which includes protein interaction information covering a large number of taxa; Homo sapiens was selected as the study object. Then, the STRING network results were exported to Cytoscape (version 3.8.2, https://cytoscape.org/), The MCODE package was used to visualize and identify core genes in PPI networks based on the following criteria: MCODE score >5 and number of nodes >20.
GO and KEGG enrichment analyses

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of core genes were performed using "ClusterProfiler" (http://bioconductor.org/pages/release/bioc/html/clusterProfiler.html) in R.

Molecular docking

Molecular docking was used to explore the binding interactions between drugs and small molecule compounds, such as proteins. The structure of Cy was transformed into a three-dimensional structure using Chem3D Ultra 14.0 (CambridgeSoft Corporation, Cambridge, MA, USA). The MM2 method was used for energy minimization, and results were saved in mol format and exported as a pdbqt ligand parameter file using Autodocktools-1.5.6.

The protein structures UBE2C (PDB ID: 1i7k), CHEK1 (PDB ID: 2e9u), MKI67 (PDB ID: 1r21), TYMS (PDB ID: 6qy1), CCNA2 (PDB ID: 1oi9), BIRC5 (PDB ID: 3ued), TTK (PDB ID: 3gfw), CENPF (PDB ID: 6m4i), PLK (PDB ID: 2ogq), EZH2 (PDB ID: 4mi5) of core genes were evaluated using MGLTools v1.5.6, including hydrogenation, circuit calculations, and merging nonpolar hydrogens. All data were exported as pdb files and saved in pdbqt format.

Autodock Vina 1.1.2 was used for molecular docking; the conformation with the highest affinity was chosen as the final docking conformation, and results were visualized using Maestro 11.9.

Associations between core genes and survival and immune infiltration Kaplan–Meier plotter (http://kmplot.com/analysis/) was used to determine the influence of core genes on prognosis [20]. The data sources included GEO, EGA, and TCGA. The primary purpose of the tool is meta-analysis–based discovery and validation of survival biomarkers.

The Tumor Immune Estimation Resource database (TIMER, https://cistrome.shinyapps.io/timer/) applies a deconvolution method to infer the abundance of tumor-infiltrating immune cells from gene expression profiles. We used TIMER to analyze the characteristics of ten core genes that were closely related to the anti-TNBC effects of Cy. Correlations between expression levels of ten core genes and the abundance of immune cells were evaluated by Spearman’s correlation coefficients, setting the cancer type to BRCA (Breast Invasive Carcinoma)-Basal. The infiltration level for each Somatic Copy Number Alteration (SCNA) category was compared with the normal level by a two-sided Wilcoxon rank–sum test.

Results

Collection of Cy-related drug target information

We obtained 305 drug targets from Pharm Mapper, 1235 drug targets from Gene Cards, and 98 drug targets from Swiss Target Prediction.

Identification of DEGs in TNBC based on GEO datasets

The RRA method was used for background correction and normalization. According to the criterion of 1.3 items, 494 DEGs were screened out, including 267 up-regulated and 227 down-regulated genes in TNBC. Figure 3 shows the top 20 up-regulated and down-regulated genes in the four datasets from GEO.

Construction of PPI networks

We constructed a PPI network with 68 nodes (based on the overlap between drug targets and DEGs) and 311 edges using STRING (Figure 4). The average node degree was 9.15, the average local clustering coefficient was 0.539, and the PPI enrichment p-value was <1.0e-16. Ten core genes were screened out using Cytoscape, including UBE2C, CHEK1, MKI67, TYMS, CCNA2, BIRC5, TTK, CENPF, PLK1, and EZH2 (Figure 5).

Functional enrichment analyses of candidate genes

We evaluated all 68 targets of Cy with GO and KEGG pathway enrichment analyses and had the following key findings (Figure 6):

1. The anti–TNBC effect of Cy involved the following biological processes: neutrophil chemotaxis, regulation of mitotic nuclear division, neutrophil migration, granulocyte chemotaxis, regulation of nuclear division, negative regulation of cell cycle phase transition, regulation of chromosome organization, mitotic nuclear division, response to ionizing radiation, and granulocyte migration.

2. The anti–TNBC effect of Cy involved the following cellular components: pronucleus, condensed chromosome, chromosomal region, immunological
Figure 3: The hot map of top 20 up-regulated and down-regulated DEGs between TNBC and non-TNBC groups. The red indicates DEGs up expression, and the green indicates DEGs down expression, the number exhibited in the figure indicates the logarithmic fold change of genes in each dataset.

Figure 4: Constructed the PPI network though SRTING. The nodes represent common gene and the lines represent the connect of each targets.

synapse, condensed chromosome outer kinetochore, kinetochore, basolateral plasma membrane, anaphase-promoting complex, an intrinsic component of the external side of the plasma membrane, and the external side of the plasma membrane.

3. The anti-TNBC effect of Cy involved the following molecular functions: CXCR chemokine receptor binding, chemokine activity, chemokine receptor binding, microtubule-binding, cytokine receptor binding, cytokine receptor binding, signaling receptor activator activity, G protein-coupled receptor binding, growth factor receptor binding, cytokine activity.

4. KEGG enrichment analysis revealed four signaling pathways related to the anti-TNBC effects of Cy, including the p53 signaling pathway, PI3K-Akt signaling pathway, MAPK signaling pathway, and chemokine signaling pathway.

Molecular docking results

Molecular docking results revealed the 2-dimensional and 3-dimensional docking poses of Cy with each core target. Only four core targets showed successful docking with Cy (Figure 7), indicating that Cy may act via sites in these core targets.

Validation of core genes

Based on a Kaplan–Meier analysis, 2 of 10 core genes were significantly related to an adverse overall survival in Cy–treated TNBC (Figure 8), including MKI67 (HR = 0.39 (0.19–0.18), P = 0.0081) and CENPF (HR = 0.26 (0.12–0.57), P = 0.00031).

There is documented evidence that MKI67 is a suitable genetic marker in TNBC; accordingly, we selected MKI67 as an efficient indicator of the efficacy of Cy against TNBC.

We explored the correlations between the levels of the 10 core genes and immune infiltration to evaluate the effects of Cy on the tumor microenvironment. Our results indicated that the core gene expression levels had a highly complex and variable effect on immune cell infiltration, reflecting the heterogeneity and complexity of the immune microenvironment in TNBC (Figure 9). The infiltration level for each SCNA illustrated the association with mutations in core genes in BRCA–Basal.

Figure 7: Molecular docking of Cy with core target. The left were 3-dimensional docking diagram results, the rights were 3-dimensional docking diagram results. Docking process of Cy with (A), UBE2C. (B), CCNA2. (C), TTK. (D), PLK1.
Figure 8: Validation of core genes by Kaplan–Meier analysis. Kaplan–Meier analysis of overall survival (months) of the core genes in breast cancer patients; red line indicates high expression, black line indicates low expression.

Figure 9: Correlation of core genes expression with immune infiltration. Partial correlation analysis of the expression of the 10 core genes on tumor purity and immune infiltration levels by TIMER database. (A)UBE2C, (B)CHEK1, (C)MKI67, (D)TYMS, (E)CCNA2, (F)BIRC5, (G)TTK, (H)CENPF, (I)PLK1, and (J)EZH2.

distributions of core gene subsets for each copy number status in BRCA-Basal are shown in box plots in Figure 10.

Discussion

TNBC, a highly heterogeneous and invasive subtype of breast cancer, has a poor prognosis, despite huge clinical investment. The median overall survival is less than 40 months [21]. TNBC usually occurs in young women carrying a BRAC1 mutation, especially in African-American women [22]. The aggressive character of TNBC often involves the lung and central nervous system metastasis, rather than bone metastasis. Many clinicopathologic features are closely associated with TNBC, including younger age, bigger tumor size, higher rate of node positivity, and higher parity of tumor [23]. In clinical practice, surgery, and radiotherapy are considered the first-line therapeutic strategy in both TNBC and non-TNBC. Many additional therapeutic strategies have been evaluated in TNBC, such as targeted strategies and immunotherapy; however, outcomes remain poor.

In this study, we analyzed four genome-wide datasets downloaded from GEO (based on the same platform, GPL-570) and used an integrated RRA method to reduce batch effects. We identified a substantial number of significantly upregulated or downregulated genes, some of which were known biomarkers of TNBC, such as GABRP, a potential therapeutic target [24,25]. Androgen receptor FOXC1, an important immunohistochemical marker for the classification of TNBC [26]. Homeobox transcription factor EN1, a key BRD4-S coregulator, particularly in TNBC [27]. Trefoil factor 1 (TFF1), a potential prognostic biomarker with functional significance in breast cancers [28]. And NAT1, with an important role in TNBC [29]. The functions of the other DEGs and mechanisms of action are poorly documented and should be a focus of future research.

Cy is a prodrug that is activated by biotransformation phase I enzymes. After administration, most Cy is transformed to 4-hydroxycyclophosphamide by hepatic cytochrome P450 (CYP) enzymes [30]. The major site of Cy action is the liver [31]. It may form both intra- and interstrand DNA cross-links and DNA-protein cross-links, both resulting in the inhibition of DNA replication and cell death by apoptosis [32] however, the mechanism is unclear. We utilized an integrative bioinformatics approach to explore the effect of Cy in TNBC.

Figure 10: The correlation between copy number alterations of core genes. (A)UBE2C, (B)CHEK1, (C)MIK67, (D)TYMS, (E)CCNA2, (F)BIRC5, (G)TTK, (H)CENPF, (I)PLK1, and (J)EZH2.

UBE2c, CHEK1, MKI67, TYMS, CCNA2, BIRC5, TTK, CENPF, PLK1, and EZH2 were core genes associated with the anti-TNBC effect of Cy. UBE2c and TTK are drivers of TNBC tumorigenesis based on a study of sensitivity to paclitaxel [33]. CHEK1, a conserved serine/threonine-protein kinase (also known as CHK1), is a major factor in halting the cell cycle in response to DNA damage [34]. CHK1 inhibition by miR-320c increases oxaliplatin responsiveness in TNBC [35]. Studies on TNBC cell lines have shown that CHK1 inhibition leads to decreased cell viability via mitotic cell death and confirmed that it is a new therapeutic target [36]. MKI67 (proliferation marker protein Ki-67) is expressed in nuclei during different phases of the cell cycle, except in the Go quiescent state. A study of 2465 female patients with TNBC revealed that a Ki-67 cutoff of 30% can be used for the further classification of TNBC into two subtypes with different responses and prognoses [37]. Thymidylate Synthase (TYMS) is a rate-limiting enzyme in thymidylate biosynthesis. Song, et al. [38] demonstrated that there are significant correlations between TYMS expression and clinical parameters and survival in breast cancer, suggesting that TYMS functions in the onset and progression of breast cancer. Baculoviral inhibitor of apoptosis repeat-containing 5 (BIRC5) plays a vital role in carcinogenesis by influencing cell division and proliferation and by inhibiting apoptosis. BIRC5 expression is higher in various breast cancer subtypes than in matched controls [39]. Turne, et al. screened 1,363 drugs in patient-derived xenograft (PDX) models and found that BIRC5 is highly expressed in basal-like PDXs, cell lines, and patient tumors and is negatively associated with outcomes [40]. Polo-Like Kinase-1 (PLK1) is correlated with a poor prognosis in TNBC. PLK1 overexpression is related to Ki67 [41]. Based on the functions of the core genes, the anti-TNBC effect of Cy is related to DNA damage, involving multiple targets.

GO and KEGG pathway enrichment analyses revealed that the differentially expressed core genes are enriched for various biological processes, including the regulation of chromosome organization and negative regulation of cell cycle phase transition. They were also enriched for cellular components, including the promucleus, condensed chromosome, and chromosomal region components with the ability to promote cell cycle progression and increase invasion [42]. In the molecular function category, we detected enrichment for CXCR chemokine receptors, which regulate breast tumor metastasis and angiogenesis. The results further suggested that regulation of the cell cycle and transcriptional dysregulation are vital processes in the effect of Cy against TNBC.

Molecular docking results demonstrated the interactions between Cy and core genes. Cy functioned via different locations of different core genes, potentially conferring different pharmacological activities. However, further in vivo and in vitro analyses of the pharmacodynamic are needed to determine the utility of this finding.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References


Discover a bigger impact and Visibility of your article with Peertechz Publications

**Highlights**

- Signatory publisher of ORCID
- Signatory Publisher of DORA (San Francisco Declaration on Research Assessment)
- Articles archived in world’s renowned service providers such as Portico, CNKI, AGRIS, TDSnet, Base (Bielefeld University Library), CrossRef, Scilit, J-Gate etc.
- Journals indexed in CCIJE, SHERPA/RDMEO, Google Scholar etc.
- OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)
- Dedicated Editorial Board for every journal
- Accurate and rapid peer review process
- Increased citations of published articles through promotions
- Reduced timeline for article publication

Submit your articles and experience a new surge in publication services ([https://www.peertechzpublications.com/submission](https://www.peertechzpublications.com/submission)).

Peertechz journals wishes everlasting success in your every endeavours.