

Integrating Multi-Datasets Analysis with Molecular Docking To Explore Potential Therapeutic Compound and Its Mechanism and Target Sites in Gastric Carcinoma

Shuheng Bai

Xi'an Jiaotong University Medical College First Affiliated Hospital

Ling Chen

Xi'an Jiaotong University Medical College First Affiliated Hospital

Aimin Jiang

Xi'an Jiaotong University Medical College First Affiliated Hospital

Yanli Yan

Xi'an Jiaotong University Medical College First Affiliated Hospital

Rong Li

Xi'an Jiaotong University Medical College First Affiliated Hospital

Haojing Kang

Xi'an Jiaotong University Medical College First Affiliated Hospital

Zhaode Feng

Xi'an Jiaotong University Medical College First Affiliated Hospital

Guangzu Li

Xi'an Jiaotong University Medical College First Affiliated Hospital

Wen Ma

Xi'an Jiaotong University School of Medicine

Jiangzhou Zhang

Xi'an Jiaotong University School of Medicine

Xiaozi Zhang

Xi'an Jiaotong University Medical College First Affiliated Hospital

Juan Ren (✉ renjuan88@mail.xjtu.edu.cn)

Xi'an Jiaotong University Medical College First Affiliated Hospital <https://orcid.org/0000-0002-8360-8286>

Primary research

Keywords: Gastric carcinoma, Bazedoxifene, UHRF1, GSEA, molecular docking

Posted Date: June 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-608336/v1>

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

Abstract

Background

Gastric carcinoma still threatens public health with high morbidity and mortality, especially in Eastern Asia. Target therapy possesses a vast value and makes significant progress in cancer treatment, and it is essential to explore more potential target drugs in gastric carcinoma.

Methods

Gene expression data was obtained from TCGA (The Cancer Genome Atlas) and GEO (Gene Expression Omnibus) datasets and the interaction networks of chemical-genes was obtained from CTD (Comparative Toxicogenomics Database). Then, GSEA (Gene Set Enrichment Analysis) was utilized to identify the significant compounds, which related to gene expression in gastric carcinoma. Moreover, PPI (Protein-Protein Interaction) network, screening of hub-genes, KEGG (Kyoto Encyclopedia of Genes and Genomes), and GO (Gene Ontology) analysis were employed to explore potential mechanisms and potential targets of these compounds. Finally, molecular docking was used to analyze the stable binding between the compounds and target proteins.

Results

Bazedoxifene was selected by integrating datasets and bioinformatics analysis as a potential therapeutic compound. UHRF1, MCM10, HELLS, and DTL with high expression in cancer tissues may be new targets in gastric carcinoma. Furthermore, Bazedoxifene may target UHRF1 and the pathway of ubiquitin-like protein transferase in gastric carcinoma and has potential therapeutic value.

Conclusion

It is inferred that Bazedoxifene may target on UHRF1 in gastric carcinoma and has the therapeutic potential.

Introduction

Gastric carcinoma is the fifth common cancer worldwide, with one million new diagnosis cases annually [1]. It is the third most common cause of cancer death, with 784000 deaths globally in 2018, which may be caused by its frequently advanced stage at the first diagnosis [2]. So, gastric carcinoma seriously threatens the public health.

Traditional treatment which includes surgery, chemotherapy, and radiotherapy have reached bottleneck period because of their great limitations of gastric carcinoma. Because of atypical symptoms of gastric carcinoma in its early stage, many patients have already been in the middle and late stage at first diagnosis and even with metastasis, thus missing the best opportunity for surgical treatment. Chemotherapy is the main treatment measure for patients with advanced stage gastric carcinoma. Although chemotherapy can prolong some patients' overall survival, it has adverse reactions such as nausea, vomiting, decreased white blood cells and platelet counts [3]. The curative ratio is still unsatisfactory and curative effect of chemotherapy is very limited. In recent years, along with the mechanism of occurrence and development of gastric carcinoma has been gradually elucidated, molecular targeted therapy, including targets on HER2 and VEGF, has gradually made great progress [4, 5]. However, the number of patients who meet the clinical indications of targeted drugs is limited, and patients are prone to develop the drug resistance [6]. Therefore, finding new targets for gastric carcinoma treatment and developing related therapeutic drugs is still an urgent mission.

Comparative Toxicogenomics Database (CTD) is a dataset which including datas of chemical-gene/protein interactions, chemical disease, and gene-disease relationships. Many researchers have utilized it to explore the effect of chemicals-gene expression, based on GSEA method [7, 8]. AutoDock is a suite of automated docking tools to predict how small molecules bind to a known 3D structure of receptor. It has been widely used for drug prediction and verification [9, 10]. This study aims to explore the potential drugs and related targets in gastric carcinoma based on CTD, TCGA, and GEO dataset and some bioinformatics methods, including GSEA and Autodock.

Methods

Gene expression data obtained from TCGA and GEO

The Cancer Genome Atlas (TCGA) is a public database supported by the National Cancer Institute (NCI) and National Human Genome Research Institute (NHGRI) of USA. It is a landmark cancer genomics program, molecularly characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types [11]. In this work, normalized gene expression data of stomach adenocarcinoma (TCGA-STAD-FPKM) and paired normal tissue by utilizing GDC DATA PORTAL (<https://portal.gdc.cancer.gov/>), official website of TCGA, were obtained.

NCBI-GEO (<https://www.ncbi.nlm.nih.gov/gds>) is also an international public repository, including next-generation sequencing and microarray/gene profiles. This study obtained 3 mRNA microarray expression profiles of stomach tumor tissues and paired normal tissues. Limma package (version 4.0) was utilized to normalize these expression data, which is an essential R/Bioconductor software to process next-generation sequencing or microarray/gene profiles [12].

Chemical-genes interaction network obtained

Comparative Toxicogenomics Database (CTD, <http://ctdbase.org/about/>) is a robust, publicly available database that aims to fully understand how environmental exposures affect human health. It provides manually curated information about chemical–gene/protein interactions, chemical–disease and gene–disease relationships [13, 14]. Data of chemical-gene interaction and summarized chemical-gene interaction network were downloaded using the DPLYR package (version 0.7.8) of R software.

Identification of chemical compounds related to gastric carcinoma by GSEA

Gene Set Enrichment Analysis (GSEA) is a computational method that determines whether a priori defined set of genes shows statistically significant, concordant differences between two biological datasets [15]. And some studies had utilized it to explore the relationship between defined datasets (such as KEGG dataset, immune infiltration cells dataset, and chemical-related genes dataset) and own dataset of gene expression [14, 16]. In this study, GSEA was conducted to explore the relationships between chemicals and dataset of each gene expression. Moreover, $p < 0.05$ was considered statistically significant for the results of chemicals related gene. And, venn diagram was used to summarize all chemicals, that obtained from GSEA.

KEGG and GO analysis for potential functions and pathways of the compound

The compound was obtained, and paired the genes related to chemical were also selected. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) are datasets that aims to establish links between a collective set of genes in the genome and some functions of the cell and the organism [17]. CusterProfiler package (version 3.16) was used to analyze and visualize functional profiles (GO and KEGG) of selected chemical-related genes. $p < 0.05$ was defined as statically significant.

Establishment of PPI network and identification of hub-genes

The Search Tool for the Retrieval of Interacting Genes database (STRING, version 11.0) provides credible information in interactions between proteins and supplies detailed annotation [18]. A PPI network of compound-related genes was constructed by using STRING. Cytoscape (version 3.7) is a general-purpose, open-source software which is capable of massively integrate data of molecular interaction network. Molecular Complex Detection (MCODE) is an app in Cytoscape, which aims to find densely connected regions of a large molecular interaction network [19]. The most important modules consisting of hub-genes and several relatively important modules in the PPI networks were identified by MCODE with selected criteria as following: degree cut-off=2, node score cut-off=0.2, Max depth=100 and k-score=2.

The expression difference of hub-genes between tumor and normal tissue

The Gene Expression Profiling Interactive Analysis (GEPIA) is a web server to analyze the RNA sequencing expression data of 9,736 tumors and 8,587 normal tissues from the TCGA and the GTEx projects standard processing pipeline [20]. It provides vital interactive and customizable functions, including differential expression analysis, profiling plotting, correlation analysis, patient survival analysis [21]. In this study, it was utilized to explore the difference of hub-genes expression between gastric carcinoma tissues and normal tissues. The result was statistically significant when conformed $\text{Log}_2\text{FC} > 2$ and $p < 0.05$.

Molecular docking to explore potential binding proteins of compound

ZINC (<http://zinc.docking.org/>) is a free database of commercially available compounds for virtual screening, which contains over 230 million purchasable compounds in their 3D formats [22]. The 3D structures of chemical compounds were downloaded from ZINC. UniProt is a web dataset which can provide a comprehensive, high-quality and freely accessible protein sequence and functional information [23]. The Protein Data Bank (PDB, <https://www.rcsb.org/>) was established as the first open-access digital data resource of biology and medicine, which including vast 3D structure data for large biological molecules (proteins, DNA, and RNA) [24]. So, the 3D structure data of potential binding proteins were obtained from UniPort or PDB dataset.

AutoDock is a suite of automated docking tools designed to predict how small molecules bind to a known 3D structure receptor. AutoDock Vina is a new generation of docking software. It achieves significant improvements in the average accuracy of the binding mode predictions while also being up to two faster orders than AutoDock 4 [25]. So, AutoDock (version 4.2) was used for the preparation process, in which water molecules were removed, gasteiger charges were computed, polar hydrogens were added, and non-polar hydrogens were merged. Then it was used to construct a centered grid to ensure coverage of the binding site. AutoDock Vina was utilized to perform docking simulations, generate conformations of a ligand in a complex with the receptor, ranking based on binding energy. Besides, the most conformations (based on binding energy) were visualized in the PyMol (version 2.4) software. And the conformation, which conformed to that docking affinity $< -7.5\text{Kcal/mol}$, was recognized as correct prediction [26, 27].

Statistical analysis and visualization

All packages, mentioned above were carried out based on R software (version 3.61). $P < 0.05$ was defined as statically significant in this study. Venn diagram was constructed by FunRich software (version 3.13).

Results

The overview of Gastric carcinoma gene profile datasets

The expression data of gastric carcinoma from TCGA contains 343 tumor cases and 30 normal cases. GSE26899 dataset is an array expression data that contains 96 tumor cases and 12 normal cases and recruited by Baylor college of medicine. GSE54129 dataset is also an array profile dataset, including 111 tumor samples and 21 normal samples, recruited by Ruijin hospital. Stefan S. Nicolau Institute of Virology recruited GSE103236 dataset, which contains 10 tumor cases and 9 normal cases.

The chemical compound associated with gastric carcinoma

By using GSEA analysis for TCGA dataset, 149 chemical compounds related to gastric carcinoma were identified ($p < 0.05$). The top 10 chemical compounds were shown in table 1. 50 compounds, 160 compounds and 76 compounds were obtained respectively from GSE26899, GSE54129 and GSE103236 datasets ($p < 0.05$), and the top 10 chemical compounds were shown in table 1, respectively. Then, the results of all these datasets were summarized and an intersection was taken. Only one gastric carcinoma-related chemical compound, Bazedoxifene, was selected by Venn diagram (Fig 1). And the details of GSEA results about Bazedoxifene was can be seen in table 2. Taken together, these results suggest that Bazedoxifene is a chemical compound associated with gastric carcinoma. And what signal pathways it can affect in gastric cancer will be explored in the next step.

The Go and KEGG analysis

Bazedoxifene was selected as a gene related compound of gastric carcinoma. And the enrichment genes, which were related to Bazedoxifene in gastric carcinoma, were also selected by GSEA, as shown in table 2. So, the union of these Bazedoxifene-related genes were screened, and it were conducted that GO and KEGG analysis to explore the potential biological signaling pathway and functions of Bazedoxifene in gastric carcinoma. The GO result showed that Bazedoxifene was mostly linked with ubiquitin like protein transferase activity, ubiquitin-protein transferase activity, and ubiquitin-like protein ligase activity. The KEGG result showed that it was related to MAPK pathways, Human T-cell leukemia virus infection and Viral carcinogenesis. The detailed result was shown in Figure 2. Overall, these results indicate that Bazedoxifen act roles in many fronts in gastric carcinoma, and screened the main target site and function is essential.

Construction of PPI network and selection of hub-genes

Function analysis of Bazedoxifene-related genes were conducted above. A PPI network, including 13 nodes and 18 edges, was constructed by String, as shown in Figure 3-A. Then, two sub-networks were screened from it through applying Cytoscape MCODE algorithm. And total of eight hub genes (EGR1, EGR2, FOS, DUSP1, UHRF1, MCM10, HELLS, and DTL) were obtained, as shown in Figure 3-B and 3-C.

Meanwhile, GEPIA website was utilized to explore the difference of these hub-genes expression level between gastric carcinoma groups and control groups. As can be seen from Figure 4, there were significant differences between gastric carcinoma tissue and normal tissue in the expression levels of UHRF1, MCM10, HELLS, and DTL. The expression levels of all these genes were significantly higher in tumor tissues than in normal tissues. Up to now, potential target genes of Bazedoxifene were identified by a comprehensive analysis of gene profile. Whether Bazedoxifene can bind to proteins of these genes-transcription stably will be tested based on 3D structures of proteins.

Molecular docking

UHRF1, MCM10, HELLS, and DTL, which had distinct expression differences between tumor and normal tissues, were selected from hub-genes to explore whether they could stably bound with Bazedoxifene or not. Firstly, the 3D structures of proteins, which translated from these selected genes, were obtained from Uniport and PDB datasets. The structures of UHRF1 and DTL were coded as 4GY5, 6QCO in PDB dataset, while, MCM10 and HELLS were named as Q7L590, Q9NRZ9 in Uniport dataset. Structure of Bazedoxifene was got from ZINC dataset. The detailed structures were shown in Figure 5.

Autodock Vina was utilized to analyze the binding site and energy for the mentioned proteins and Bazedoxifene. PyMol visualized each protein's best model (based on binding energy), as Figure 6 showed. It was shown that only UHRF1 (affinity: -7.8 kcal/mol) had stable binding sites (affinity < -7.5 kcal/mol). The best model of MCM10, HELLS, and DTL were regarded as unstable, because their affinity were -6.2 kcal/mol, -7.2 kcal/mol, -6.6 kcal/mol respectively, which were all lower than 7.5 kcal/mol.

Discussion

Gastric carcinoma still threatens public health with higher morbidity and mortality, especially in Eastern Asia, Eastern Europe, and South Africa [2, 28]. Recently target therapy represents a vast value and made great progress in cancer treatment, such as in lung adenocarcinoma and colorectal adenocarcinoma [29, 30]. However, up to now, excepting the most major target-HER2 and VEGF/VEGFR, other target options, including EGFR and HGF/MET, have emerged but not been proved effective in prolonging the survival time. The homodimers or heterodimers formed in HER family pairs can activate MAPK, PI3K and other signaling pathways, thereby regulating cell proliferation, adhesion, migration, and differentiation [31, 32]. Trastuzumab is a type of anti-HER-2 drug that binds to the extracellular domain IV of HER2 and inhibits homodimers' formation. It has shown beneficial in survival, combined with cisplatin and a fluoropyrimidine, for the treatment of gastric carcinoma patients with HER2 positive expression [33]. Remizumab is a type of monoclonal antibody and can bind to extracellular domain of VEGFR-2 to suppress angiogenesis. Clinical trial demonstrated that it effectively prolonged survival time when combined with chemotherapy for patients with advanced gastric carcinoma [34]. Cetuximab and Panitumumab are also anti-EGFR antibodies. Clinical trials stated that there had none improvement in survival rates for advanced gastric carcinoma when combined with Cetuximab, and Panitumumab not only failed in improving the survival time, but also increased toxicities [35–37]. Meanwhile, Rilotumumab, which can anti-HGF, could improve the prognosis for gastric cancer patients with higher expression of MET, but was proved a worse safety [38, 39]. Depending on this circumstance, a string of targets and drugs are needed in gastric carcinoma.

Based on this GSEA analysis, which utilized multiple transcriptome data sets and CTD dataset, Bazedoxifene was found closely related with some specific genes which highly expressed in gastric carcinoma, and it may be a new drug for patients with gastric carcinoma. Bazedoxifene is the third-generation selective estrogen receptor modulator (SERM), which received FDA approval to prevent postmenopausal osteoporosis and moderate to severe vasomotor symptoms associated with menopause [40]. It mainly targeted on estrogen receptor alpha (ESR1) and estrogen receptor alpha (ESR2), which are all nuclear hormone receptor and can activate the expression of reporter genes, containing estrogen response elements (ERE) [41].

Furthermore, some researches demonstrated that Bazedoxifene can be a novel and useful drug in cancer treatment. Jia et al found that Bazedoxifene significantly inhibit colorectal cancer cell proliferation and migration by blocking IL-11/GP130 signaling pathway in vivo and in vitro [42]. SJW et al also demonstrated that Bazedoxifene is a novel inhibitor of GP130 signaling pathway, and it can generate synergism when combined with conventional chemotherapy in human pancreatic cancer cells [43]. The study discovered that Bazedoxifene combined with paclitaxel exhibited more potent inhibition of cell viability, colony formation, and cell migration, induced more apoptosis in vitro, and generated stronger inhibition of tumor growth of breast cancer in vivo than either drug alone [44]. Furthermore, a study also implied that it suppressed the development of hepatocellular carcinoma [45].

Interestingly, IL-11/GP130 and ERE related signaling pathway were related with the progression of gastric cancer. Some researches had identified that deregulation of gp130-dependent STAT1/STAT3 signal pathway, leading by the crucial cytokine IL-11, involved in the human *Helicobacter pylori* infection and pathogenesis of gastric cancer [46–48]. It has been identified that decreasing of estrogen receptor binding affinity to the estrogen response element could decrease the expression of the TFF1 gene, which may be involved in development of gastric cancer [49].

In this study, it was shown that UHRF1 was a potential target site for Bazedoxifene in gastric carcinoma. As this study showed, it was screened as a hub-genes from the PPI network and identified as an apparent overexpressed gene and demonstrated a stable binding site for Bazedoxifene in protein level. UHRF1 is a nuclear protein gene involved in tumorigenesis and development. It encodes a member of a subfamily of RING-finger type E3 ubiquitin ligases. UHRF1 can participate in DNA methylation modification and histone post-translational modification through specific domains, regulate gene expression, and participate in malignant tumors' occurrence and development [50]. That was consistent with our pathway analysis, which indicated that Bazedoxifene mainly targets ubiquitin-like protein transferase activity, ubiquitin-protein transferase activity, viral carcinogenesis MAPK pathway in gastric carcinoma. Meanwhile, UHRF1 is up-regulated in various cancers, including lung cancer, cervical cancer, pancreatic cancer, and so on. Researches reported that knocking down UHRF1 reduced tumor cell proliferation and promote apoptosis, so it is therefore considered to be a therapeutic target [50–53]. In gastric carcinoma, The UHRF1 expression in serum and tissue were significantly higher than health control and may represent a novel biomarker for diagnosis and prognosis [54]. Meanwhile, Zhou L et al also reported that UHRF1 was overexpressed, and it was an independent and significant predictor. Moreover, downregulation of UHRF1 suppressed cell proliferation in vitro and in vivo, and UHRF1 upregulation showed opposite effects. Also, UHRF1 overexpression promoted gastric carcinoma proliferation and carcinogenesis by inhibiting apoptosis and increasing the G1/S transition and inducing the hypermethylation of 7 tumor-suppress-genes (CDX2, CDKN2A, RUNX3, FOXO4, PPARG, BRCA1, and PML) [55]. Other research also indicated that UHRF1 could promote the growth, migration, and invasion of gastric carcinoma cells and inhibited apoptosis via a ROS-associated pathway [56]. So, Bazedoxifene may act at UHRF1 to play a role in treating patients with gastric carcinoma.

Some other genes may also be targets for gastric carcinoma, including MCM10, HELLS, and DTL, as this study shown. MCM10 is a subtype of Minichromosome Maintenance proteins (MCM) family, and is essential for replication origin firing [57]. MCM10 can combine with other proliferative markers, initiate DNA replication of cell cycle, thereby mediating cell proliferation, and recruit DNA polymerase- α and a catalytic subunit, preventing the damage of DNA [58]. Moreover, MCM10 was overexpressed in a series of cancers, including pancreatic cancer, cervical cancer, and esophageal cancers. It promoted tumor progression and maybe a prognostic biomarker and potential therapeutic target [59]. HELLS is the primary member of the SNF2 chromatin remodeling enzyme family, and is related to the processes of DNA replication, repair, recombination, and transcription, and involved in typically interacts with DNA methyltransferases [60]. Besides that, HELLS also involved in cancer development, including hepatocellular carcinoma, head and neck cancer, lung cancer, and breast cancer [61, 62]. Then, DTL is identified as an adapter of E3 ubiquitin-protein ligase complex required for cell cycle control, DNA damage response, and translesion DNA synthesis [63]. It can enhance cancer cells' motility and proliferation through PDCD4 ubiquitin-dependent degradation to promote the development of cancers [64].

Conclusion

This study conducted an integrated analysis for gene expression data, and chemical -gene interaction networks. GSEA was utilized to identify Bazedoxifene is a stomach-cancer-related compound. This study also underwent a functional analysis and constructed a PPI network to implied the potential function of Bazedoxifene in gastric carcinoma. Moreover, some hub-genes were selected and verify whether they can bind with Bazedoxifene stably. Then, it was found that Bazedoxifene may play a role in the anti-tumor effect by targeting UHRF1 in gastric carcinoma. Furthermore, this present method can be used to analyze other chemicals and malignant disease, which helps assess the potential drugs and target sites.

Abbreviations

TCGA
The Cancer Genome Atlas
GEO
Gene Expression Omnibus
NHGRI
National Human Genome Research Institute
CTD
Comparative Toxicogenomics Database
GSEA
Gene Set Enrichment Analysis
KEGG
Kyoto Encyclopedia of Genes and Genomes
STRING

The Search Tool for the Retrieval of Interacting Genes database
MCODE
Molecular Complex Detection
GEPIA
The Gene Expression Profiling Interactive Analysis
PDB
The Protein Data Bank
OS
Overall survival
PPI
Protein-Protein Interaction
ESR
estrogen receptor

Declarations

Funding: This manuscript is supported by the National Natural Science Foundations of China (Juan Ren, 81772793/H1621, Juan Ren, 31201060/C0709; Juan Ren, 30973175/H1621; and Juan Ren, 81772793/H1621); Supported by Scientific and Technological Research Foundation of Shaan'xi Province (Juan Ren, 2020JM-368); Supported by Program for New Century Excellent Talents in University (Juan Ren, NCET-12-0440);

Competing Interests: The authors have declared that no competing interest exists.

Availability of data and materials: The datasets used and analyzed during the current study are available from the TCGA and GEO database.

Authors' contributions: JR conceived and supervised the study; JR, SHB, LC, AMJ, YYL, HJK, ZDF, GZL, WM, RL and ZJJ analyzed data; JR and SHB wrote the manuscript; JR, SHB and XZZ made manuscript revisions. All authors have read and approved the final version of this submission.

Ethics approval and consent to participate: This study was approved by the Ethics Committees of Xi'an Jiaotong University.

Acknowledgements: Not applicable.

References

1. Smyth EC, Nilsson M, Grabsch HI, van Grieken NCT, Lordick F. Gastric cancer. *The Lancet*. 2020;396(10251):635–48.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424.
3. Shah MA, Schwartz GK. Treatment of metastatic esophagus and gastric cancer. *Semin Oncol*. 2004;31(4):574–87.
4. Shitara K, Bang YJ, Iwasa S, Sugimoto N, Ryu MH, Sakai D, Chung HC, Kawakami H, Yabusaki H, Lee J, et al. Trastuzumab Deruxtecan in Previously Treated HER2-Positive Gastric Cancer. *N Engl J Med*. 2020;382(25):2419–30.
5. Fuchs CS, Tomasek J, Yong CJ, Dumitru F, Passalacqua R, Goswami C, Safran H, Dos Santos LV, Aprile G, Ferry DR, et al. Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): an international, randomised, multicentre, placebo-controlled, phase 3 trial. *Lancet*. 2014;383(9911):31–9.
6. Pellino A, Riello E, Nappo F, Brignola S, Murgioni S, Djaballah SA, Lonardi S, Zagonel V, Rugge M, Loupakis F, et al. Targeted therapies in metastatic gastric cancer: Current knowledge and future perspectives. *World J Gastroenterol*. 2019;25(38):5773–88.
7. Cheng S, Wen Y, Ma M, Zhang L, Liu L, Qi X, Cheng B, Liang C, Li P, Kafle OP, et al. **Identifying 5 Common Psychiatric Disorders Associated Chemicals Through Integrative Analysis of Genome-Wide Association Study and Chemical-Gene Interaction Datasets.** *Schizophr Bull* 2020.
8. Cheng S, Han B, Ding M, Wen Y, Ma M, Zhang L, Qi X, Cheng B, Li P, Kafle OP, et al. Identifying psychiatric disorder-associated gut microbiota using microbiota-related gene set enrichment analysis. *Brief Bioinform*. 2020;21(3):1016–22.
9. Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc*. 2016;11(5):905–19.
10. Umesh KD, Selvaraj C, Singh SK, Dubey VK. **Identification of new anti-nCoV drug chemical compounds from Indian spices exploiting SARS-CoV-2 main protease as target.** *J Biomol Struct Dyn* 2020:1–9.
11. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell*. 2018;173(2):400–16. e411.
12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
13. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegiers J, Wiegiers TC, Mattingly CJ. The Comparative Toxicogenomics Database: update 2019. *Nucleic Acids Res*. 2019;47(D1):D948–54.
14. Tan X, Tang H, Gong L, Xie L, Lei Y, Luo Z, He C, Ma J, Han S. Integrating Genome-Wide Association Studies and Gene Expression Profiles With Chemical-Genes Interaction Networks to Identify Chemicals Associated With Colorectal Cancer. *Front Genet*. 2020;11:385.

15. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, 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(43):15545–50.
16. Liu Z, Zhu Y, Xu L, Zhang J, Xie H, Fu H, Zhou Q, Chang Y, Dai B, Xu J. Tumor stroma-infiltrating mast cells predict prognosis and adjuvant chemotherapeutic benefits in patients with muscle invasive bladder cancer. *Oncoimmunology*. 2018;7(9):e1474317.
17. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–62.
18. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607–13.
19. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
20. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–102.
21. Bai S, Wu Y, Yan Y, Shao S, Zhang J, Liu J, Hui B, Liu R, Ma H, Zhang X, et al. Construct a circRNA/miRNA/mRNA regulatory network to explore potential pathogenesis and therapy options of clear cell renal cell carcinoma. *Sci Rep*. 2020;10(1):13659.
22. Sterling T, Irwin JJ. ZINC 15—Ligand Discovery for Everyone. *J Chem Inf Model*. 2015;55(11):2324–37.
23. UniProt C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019;47(D1):D506–15.
24. Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, Christie C, Dalenberg K, Duarte JM, Dutta S, et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res*. 2019;47(D1):D464–74.
25. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31(2):455–61.
26. Stojic N, Eric S, Kuzmanovski I. Prediction of toxicity and data exploratory analysis of estrogen-active endocrine disruptors using counter-propagation artificial neural networks. *J Mol Graph Model*. 2010;29(3):450–60.
27. Montes-Grajales D, Martinez-Romero E, Olivero-Verbel J. Phytoestrogens and mycoestrogens interacting with breast cancer proteins. *Steroids*. 2018;134:9–15.
28. Torre LA, Siegel RL, Ward EM, Jemal A. Global Cancer Incidence and Mortality Rates and Trends—An Update. *Cancer Epidemiol Biomarkers Prev*. 2016;25(1):16–27.
29. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, Paz-Ares L. Lung cancer: current therapies and new targeted treatments. *Lancet*. 2017;389(10066):299–311.
30. Piawah S, Venook AP. Targeted therapy for colorectal cancer metastases: A review of current methods of molecularly targeted therapy and the use of tumor biomarkers in the treatment of metastatic colorectal cancer. *Cancer*. 2019;125(23):4139–47.
31. Witton CJ, Reeves JR, Going JJ, Cooke TG, Bartlett JM. Expression of the HER1-4 family of receptor tyrosine kinases in breast cancer. *J Pathol*. 2003;200(3):290–7.
32. Begnami MD, Fukuda E, Fregnani JH, Nonogaki S, Montagnini AL, da Costa WL Jr, Soares FA. Prognostic implications of altered human epidermal growth factor receptors (HERs) in gastric carcinomas: HER2 and HER3 are predictors of poor outcome. *J Clin Oncol*. 2011;29(22):3030–6.
33. Wagner AD, Syn NL, Moehler M, Grothe W, Yong WP, Tai BC, Ho J, Unverzagt S. Chemotherapy for advanced gastric cancer. *Cochrane Database Syst Rev*. 2017;8:CD004064.
34. Wadhwa R, Taketa T, Sudo K, Blum-Murphy M, Ajani JA. Ramucirumab: a novel antiangiogenic agent. *Future Oncol*. 2013;9(6):789–95.
35. Lordick F, Kang YK, Chung HC, Salman P, Oh SC, Bodoky G, Kurteva G, Volovat C, Moiseyenko VM, Gorbunova V, et al. Capecitabine and cisplatin with or without cetuximab for patients with previously untreated advanced gastric cancer (EXPAND): a randomised, open-label phase 3 trial. *Lancet Oncol*. 2013;14(6):490–9.
36. Waddell T, Chau I, Cunningham D, Gonzalez D, Okines AF, Okines C, Wotherspoon A, Saffery C, Middleton G, Wadsley J, et al. Epirubicin, oxaliplatin, and capecitabine with or without panitumumab for patients with previously untreated advanced oesophagogastric cancer (REAL3): a randomised, open-label phase 3 trial. *Lancet Oncol*. 2013;14(6):481–9.
37. Tebbutt NC, Price TJ, Ferraro DA, Wong N, Veillard AS, Hall M, Sjoquist KM, Pavlakis N, Strickland A, Varma SC, et al. Panitumumab added to docetaxel, cisplatin and fluoropyrimidine in oesophagogastric cancer: ATAX3 phase II trial. *Br J Cancer*. 2016;114(5):505–9.
38. Iveson T, Donehower RC, Davidenko I, Tjulandini S, Deptala A, Harrison M, Nirni S, Lakshmaiah K, Thomas A, Jiang Y, et al. Rilotumumab in combination with epirubicin, cisplatin, and capecitabine as first-line treatment for gastric or oesophagogastric junction adenocarcinoma: an open-label, dose de-escalation phase 1b study and a double-blind, randomised phase 2 study. *Lancet Oncol*. 2014;15(9):1007–18.
39. Catenacci DVT, Tebbutt NC, Davidenko I, Murad AM, Al-Batran SE, Ilson DH, Tjulandini S, Gotovkin E, Karaszewska B, Bondarenko I, et al. Rilotumumab plus epirubicin, cisplatin, and capecitabine as first-line therapy in advanced MET-positive gastric or gastro-oesophageal junction cancer (RILOMET-1): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. 2017;18(11):1467–82.
40. Parish SJ, Gillespie JA. The evolving role of oral hormonal therapies and review of conjugated estrogens/bazedoxifene for the management of menopausal symptoms. *Postgrad Med*. 2017;129(3):340–51.

41. Komm BS, Kharode YP, Bodine PV, Harris HA, Miller CP, Lyttle CR. Bazedoxifene acetate: a selective estrogen receptor modulator with improved selectivity. *Endocrinology*. 2005;146(9):3999–4008.
42. Wei J, Ma L, Lai YH, Zhang R, Li H, Li C, Lin J. Bazedoxifene as a novel GP130 inhibitor for Colon Cancer therapy. *J Exp Clin Cancer Res*. 2019;38(1):63.
43. Wu X, Cao Y, Xiao H, Li C, Lin J. Bazedoxifene as a Novel GP130 Inhibitor for Pancreatic Cancer Therapy. *Mol Cancer Ther*. 2016;15(11):2609–19.
44. Fu S, Chen X, Lo HW, Lin J. Combined bazedoxifene and paclitaxel treatments inhibit cell viability, cell migration, colony formation, and tumor growth and induce apoptosis in breast cancer. *Cancer Lett*. 2019;448:11–9.
45. Ma H, Yan D, Wang Y, Shi W, Liu T, Zhao C, Huo S, Duan J, Tao J, Zhai M, et al. Bazedoxifene exhibits growth suppressive activity by targeting interleukin-6/glycoprotein 130/signal transducer and activator of transcription 3 signaling in hepatocellular carcinoma. *Cancer Sci*. 2019;110(3):950–61.
46. Buzzelli JN, O'Connor L, Scurr M, Chung Nien Chin S, Catubig A, Ng GZ, Oshima M, Oshima H, Giraud AS, Sutton P, et al. Overexpression of IL-11 promotes premalignant gastric epithelial hyperplasia in isolation from germline gp130-JAK-STAT driver mutations. *Am J Physiol Gastrointest Liver Physiol*. 2019;316(2):G251–62.
47. Ernst M, Najdovska M, Grail D, Lundgren-May T, Buchert M, Tye H, Matthews VB, Armes J, Bhathal PS, Hughes NR, et al. STAT3 and STAT1 mediate IL-11-dependent and inflammation-associated gastric tumorigenesis in gp130 receptor mutant mice. *J Clin Invest*. 2008;118(5):1727–38.
48. Tye H, Kennedy CL, Najdovska M, McLeod L, McCormack W, Hughes N, Dev A, Sievert W, Ooi CH, Ishikawa TO, et al. STAT3-driven upregulation of TLR2 promotes gastric tumorigenesis independent of tumor inflammation. *Cancer Cell*. 2012;22(4):466–78.
49. Moghanibashi M, Mohamadynejad P, Rasekhi M, Ghaderi A, Mohammadianpanah M. Polymorphism of estrogen response element in TFF1 gene promoter is associated with an increased susceptibility to gastric cancer. *Gene*. 2012;492(1):100–3.
50. Bronner C, Krifa M, Mousli M. Increasing role of UHRF1 in the reading and inheritance of the epigenetic code as well as in tumorigenesis. *Biochem Pharmacol*. 2013;86(12):1643–9.
51. Hu Q, Qin Y, Ji S, Xu W, Liu W, Sun Q, Zhang Z, Liu M, Ni Q, Yu X, et al. UHRF1 promotes aerobic glycolysis and proliferation via suppression of SIRT4 in pancreatic cancer. *Cancer Lett*. 2019;452:226–36.
52. Goto D, Komeda K, Uwatoko N, Nakashima M, Koike M, Kawai K, Kodama Y, Miyazawa A, Tanaka I, Hase T, et al. UHRF1, a Regulator of Methylation, as a Diagnostic and Prognostic Marker for Lung Cancer. *Cancer Invest*. 2020;38(4):240–9.
53. Chen T, Yang S, Xu J, Lu W, Xie X. Transcriptome sequencing profiles of cervical cancer tissues and SiHa cells. *Funct Integr Genomics*. 2020;20(2):211–21.
54. Ge M, Gui Z, Wang X, Yan F. Analysis of the UHRF1 expression in serum and tissue for gastric cancer detection. *Biomarkers*. 2015;20(3):183–8.
55. Zhou L, Shang Y, Jin Z, Zhang W, Lv C, Zhao X, Liu Y, Li N, Liang J. UHRF1 promotes proliferation of gastric cancer via mediating tumor suppressor gene hypermethylation. *Cancer Biol Ther*. 2015;16(8):1241–51.
56. Zhang H, Song Y, Yang C, Wu X. **UHRF1 mediates cell migration and invasion of gastric cancer.** *Biosci Rep* 2018, 38(6).
57. Yeeles JT, Deegan TD, Janska A, Early A, Diffley JF. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature*. 2015;519(7544):431–5.
58. Chattopadhyay S, Bielinsky AK. Human Mcm10 regulates the catalytic subunit of DNA polymerase-alpha and prevents DNA damage during replication. *Mol Biol Cell*. 2007;18(10):4085–95.
59. Mahadevappa R, Neves H, Yuen SM, Jameel M, Bai Y, Yuen HF, Zhang SD, Zhu Y, Lin Y, Kwok HF. **DNA Replication Licensing Protein MCM10 Promotes Tumor Progression and Is a Novel Prognostic Biomarker and Potential Therapeutic Target in Breast Cancer.** *Cancers (Basel)* 2018, 10(9).
60. Lee DW, Zhang K, Ning ZQ, Raabe EH, Tintner S, Wieland R, Wilkins BJ, Kim JM, Blough RI, Arceci RJ. Proliferation-associated SNF2-like gene (PASG): a SNF2 family member altered in leukemia. *Cancer Res*. 2000;60(13):3612–22.
61. Zhu W, Li LL, Songyang Y, Shi Z, Li D. Identification and validation of HELLS (Helicase, Lymphoid-Specific) and ICAM1 (Intercellular adhesion molecule 1) as potential diagnostic biomarkers of lung cancer. *PeerJ*. 2020;8:e8731.
62. He X, Yan B, Liu S, Jia J, Lai W, Xin X, Tang CE, Luo D, Tan T, Jiang Y, et al. Chromatin Remodeling Factor LSH Drives Cancer Progression by Suppressing the Activity of Fumarate Hydratase. *Cancer Res*. 2016;76(19):5743–55.
63. Higa LA, Banks D, Wu M, Kobayashi R, Sun H, Zhang H. L2DTL/CDT2 interacts with the CUL4/DDB1 complex and PCNA and regulates CDT1 proteolysis in response to DNA damage. *Cell Cycle*. 2006;5(15):1675–80.
64. Cui H, Wang Q, Lei Z, Feng M, Zhao Z, Wang Y, Wei G. DTL promotes cancer progression by PDCD4 ubiquitin-dependent degradation. *J Exp Clin Cancer Res*. 2019;38(1):350.

Tables

Table 1

The top 10 chemical compounds related with gene expression data of TCGA, GSE26899, GSE54129 and GSE103236 datasets respectively.

Datasets	Compounds name	Count	p-value
TCGA	DASATINIB	478	0.000
	4-BIPHENYLAMINE	70	0.000
	CPG-OLIGONUCLEOTIDE	126	0.000
	2,3,5-(TRIGLUTATHION-S-YL) HYDROQUINONE	363	0.000
	AFURESERTIB	410	0.000
	GERANIOL	351	0.000
	BETA-GLYCEROPHOSPHORIC ACID	51	0.000
	BROMOVANIN	123	0.000
	ARECOLINE	64	0.000
GSE26899 dataset	AC 93253	19	0.002
	GINGEROL	44	0.008
	AM 251	75	0.008
	ACTEOSIDE	34	0.010
	CANTHARIDIN	42	0.010
	DOXAZOSIN	22	0.011
	2,5,7,8-TETRAMETHYL-2R-(4R,8R,12-TRIMETHYLTRIDECYL) CHROMAN-6-YLOXY ACETIC ACID	22	0.011
	6-OH-BDE-47	84	0.014
	GOLD SODIUM THIOMALATE	77	0.014
GSE54129 dataset	ARSENATES	90	0.000
	ALUMINUM	52	0.000
	2-(4-AMINO-3-METHYLPHENYL)-5-FLUOROBENZOTHIAZOLE	21	0.000
	CARMUSTINE	399	0.002
	2-METHOXY-N-(3-METHYL-2-OXO-1,2,3,4-TETRAHYDROQUINAZOLIN-6-YL) BENZENESULFONAMIDE	31	0.002
	BRUCINE	15	0.002
	ELLAGIC ACID	93	0.002
	CALYCULIN A	24	0.002
	DONEPEZIL	15	0.004
BORON NITRIDE	41	0.004	
GSE103236 dataset	DIPYRIDAMOLE	38	0.000
	CPG-OLIGONUCLEOTIDE	120	0.002
	FLAVONOIDS	73	0.004
	FULVESTRANT	383	0.006
	ASCORBIC ACID	334	0.006
	ETOPOSIDE	500	0.006
	4-(2-AMINOETHYL) BENZENESULFONYLFLUORIDE	16	0.008
	ALUMINUM OXIDE	60	0.009
	2-(1H-INDAZOL-4-YL)-6-(4-METHANESULFONYLPYPERAZIN-1-YLMETHYL)-4-MORPHOLIN-4-YLTHIENO(3,2-D) PYRIMIDINE	456	0.010
	CASTICIN	101	0.010

Table 2
The details of GSEA results about bazedoxifene

BAZEDOXIFENE	Count	p-value	Enrichment genes
TCGA datasets	40	0.014	HELLS,DTL,MCM10,FANCB,ZNF367,UHRF1,C1QTNF6,CDC25A,MDFI, LAMB1,SEZ6L2,RC3H2,COL12A1,PITRM1
GSE26899 dataset	35	0.024	COL12A1, EGR2,C1QTNF6,MCM10,CDC25A,UHRF1,DTL,HELLS,PITRM1,LIPC,FANCB,UBR4,KIF13A
GSE54129 dataset	39	0.011	MGP, EGR2,EGR1,COL12A1,DUSP1,FOS,ASAP1,C1QTNF6,CDC42,DTL,GREB1,SBSPON,RC3H2,RNF146,LAMB1,LIPC,CDC25A,
GSE103236 dataset	37	0.039	MDFI,COL12A1,C1QTNF6,UHRF1,MCM10,DTL,LAMB1,LIPC,HELLS, FANCB,CDH7,GPR87,

Figures

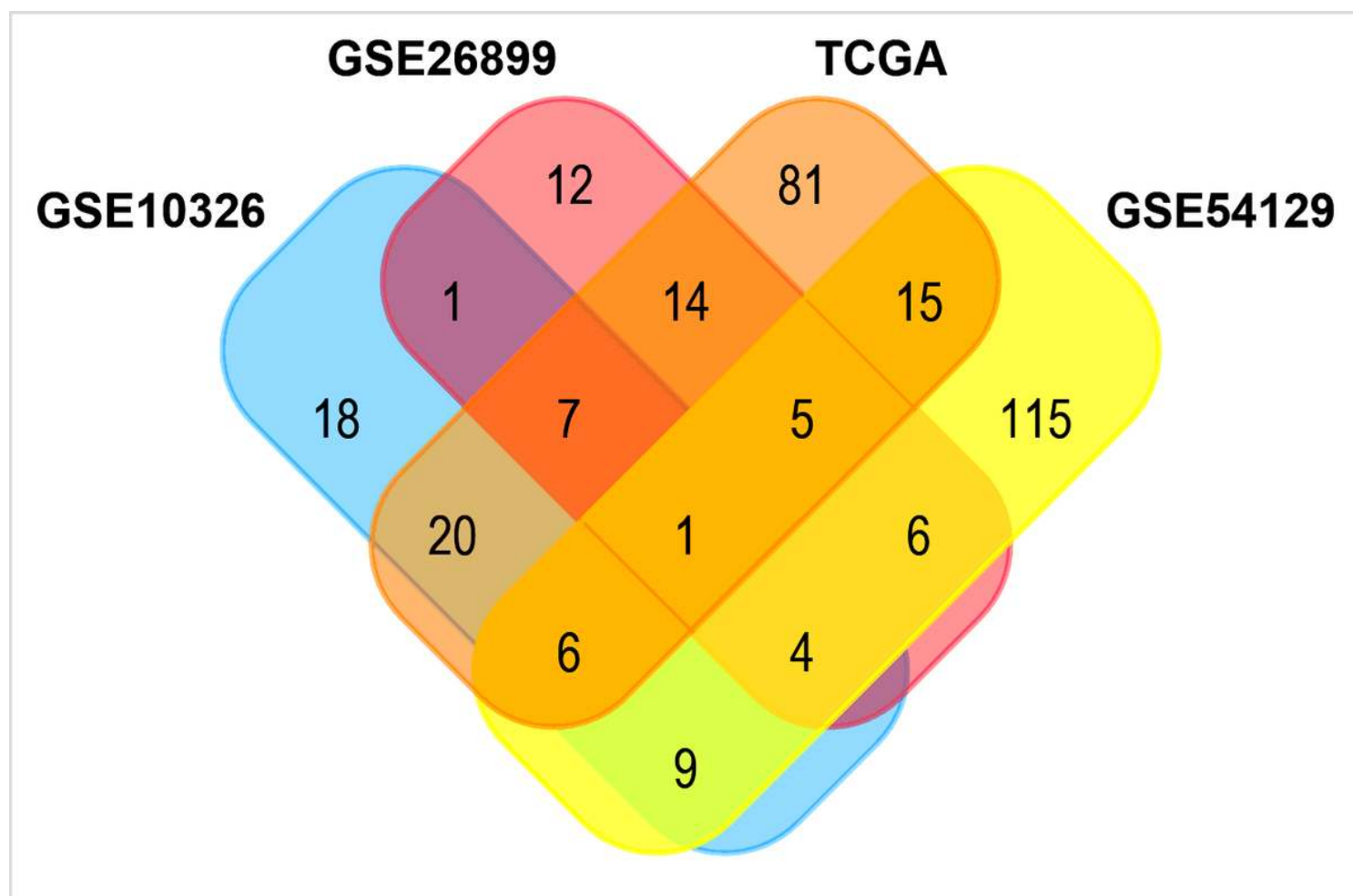


Figure 1

The Venn diagram about. Chemical compounds from TCGA, GSE26899, GSE54129 and GSE103236 datasets were taken intersection, and only one compound was screened in this diagram.

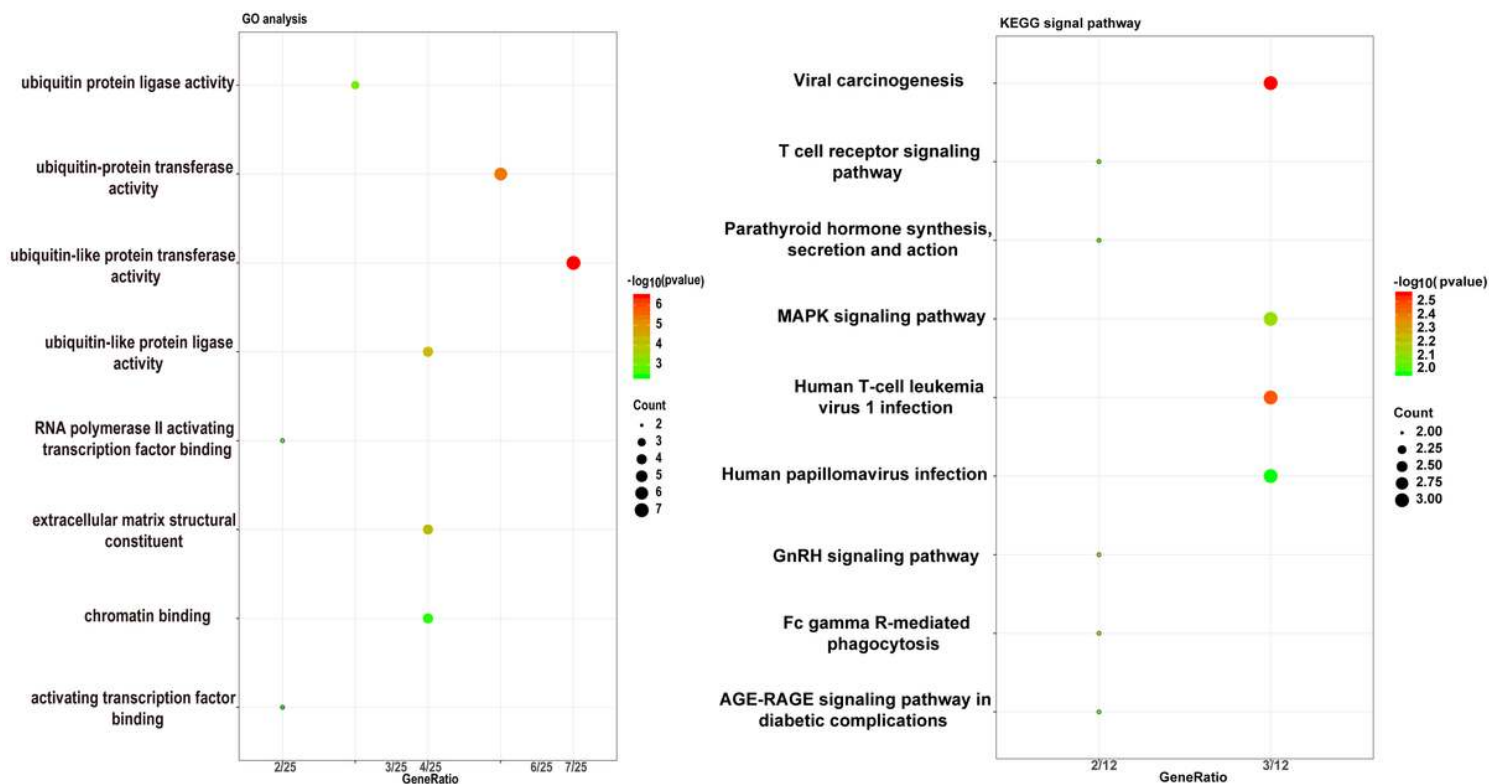


Figure 2
 GO function and KEGG pathway analyses for Bazedoxifene related genes. The color intensity of the nodes shows the confidence of this analysis. The gene-ratio is defined as the ratio of the differential genes in the entire genome. The dot size represents the count of genes in a pathway.

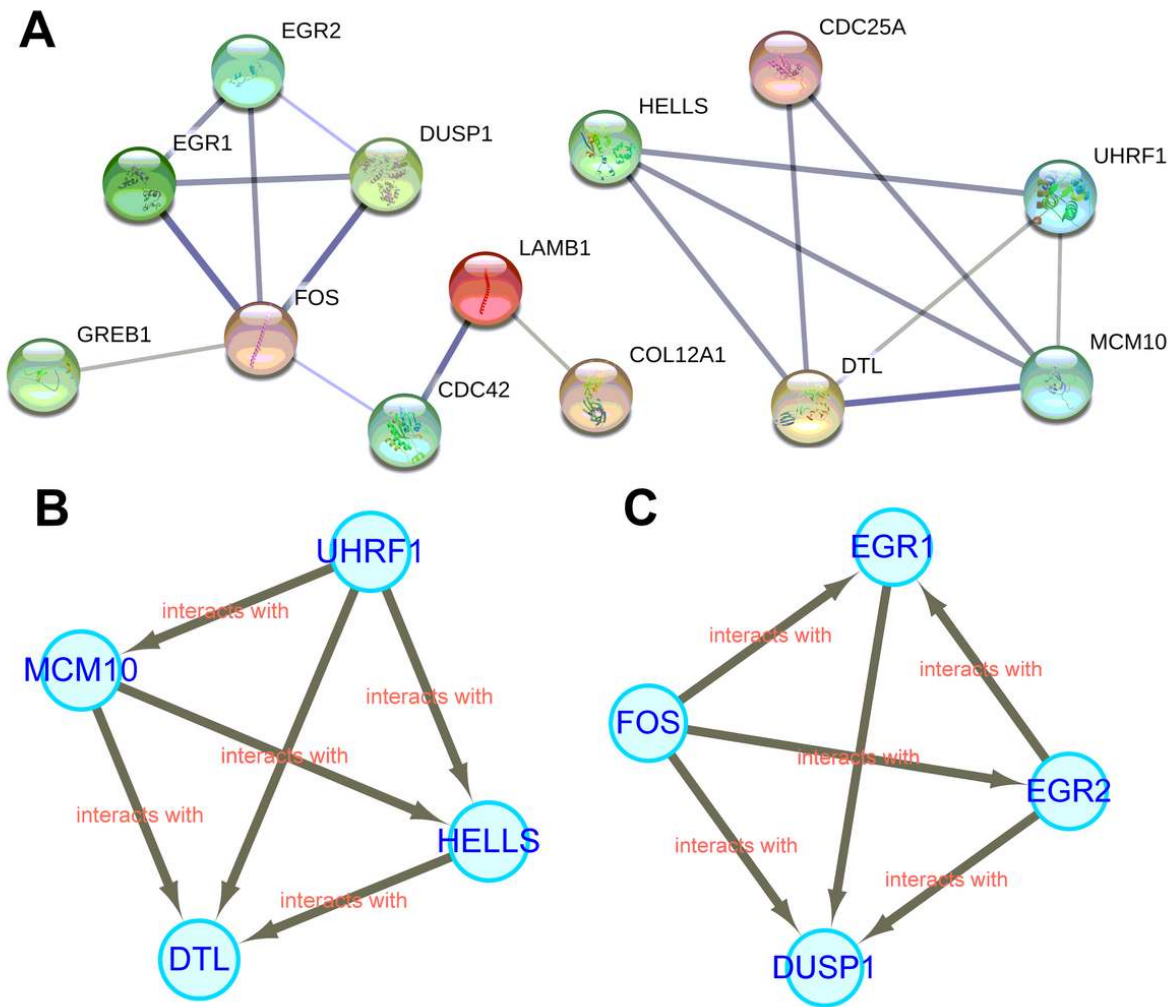


Figure 3
 PPI network and sub-network screened from it. A, a network of Bazedoxifene related genes. B and C, two sub-networks of 8 hub-genes screened from PPI network by MCODE algorithm.

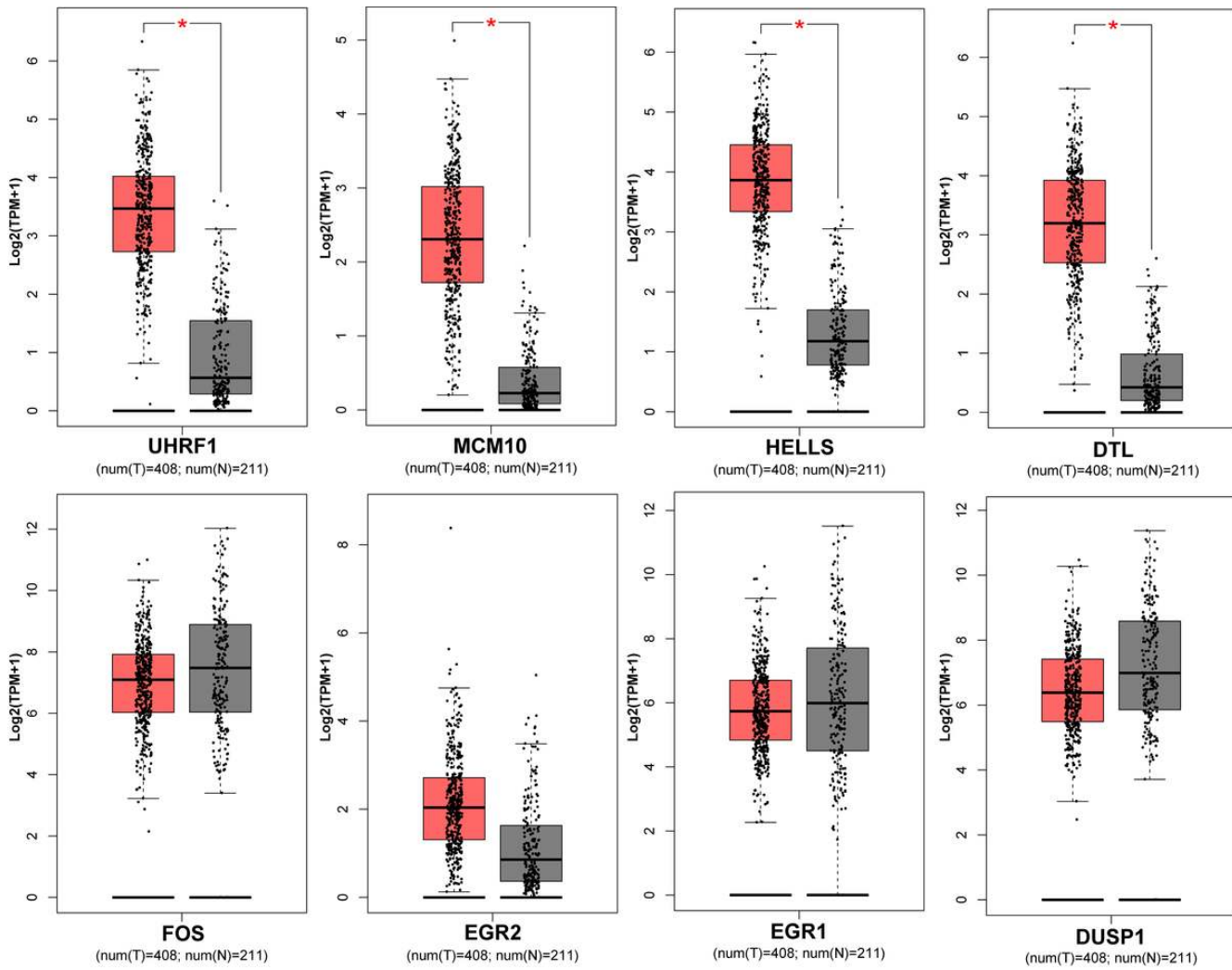


Figure 4
 Expression levels of hub-genes. The expression levels of UHRF1, MCM10, HELLS, and DTL are significant difference between cancer and normal tissue, however, there are no difference of EGR1, EGR2, FOS and DUSP1.

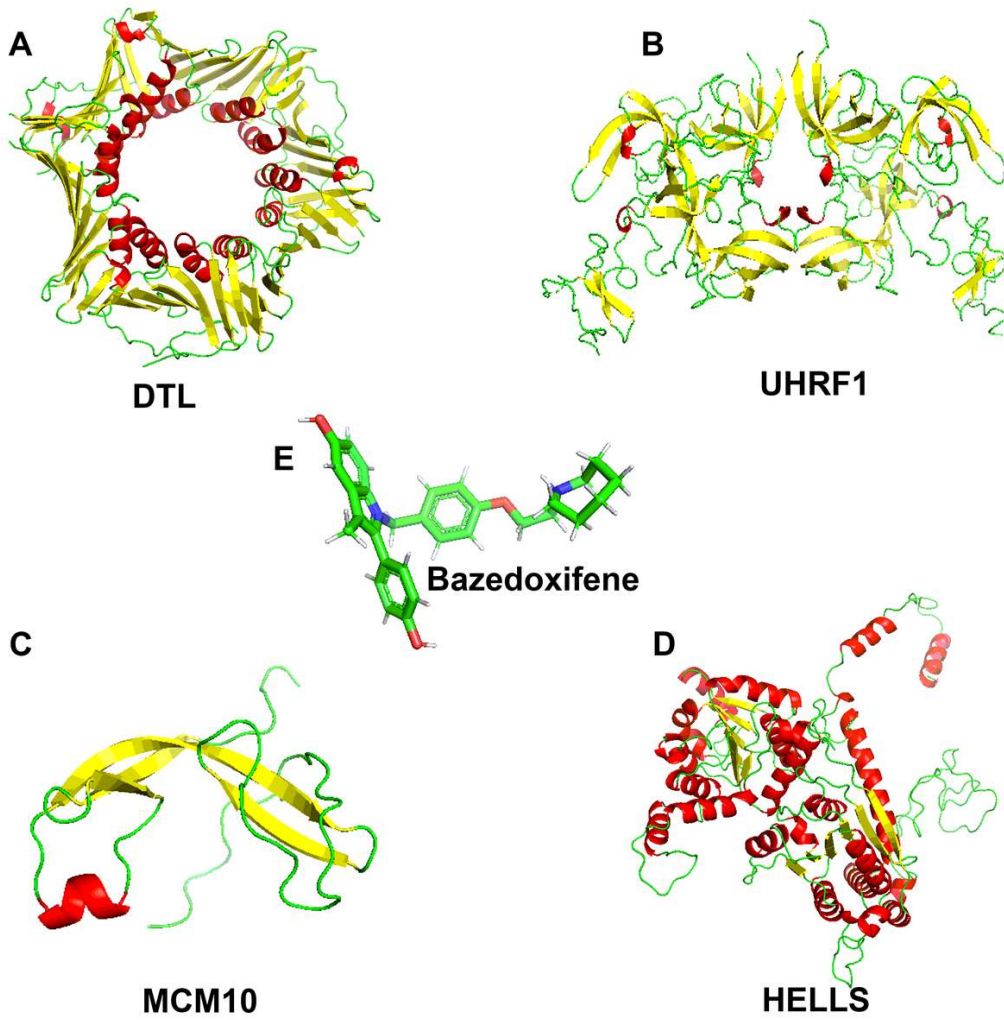


Figure 5
 The 3D-structures of Bazedoxifene, UHRF1, MCM10, HELLS, and DTL. A, structure of DTL. B, structure of UHRF1. C, structure of MCM10. D, structure of HELLS. E, structure of Bazedoxifene.

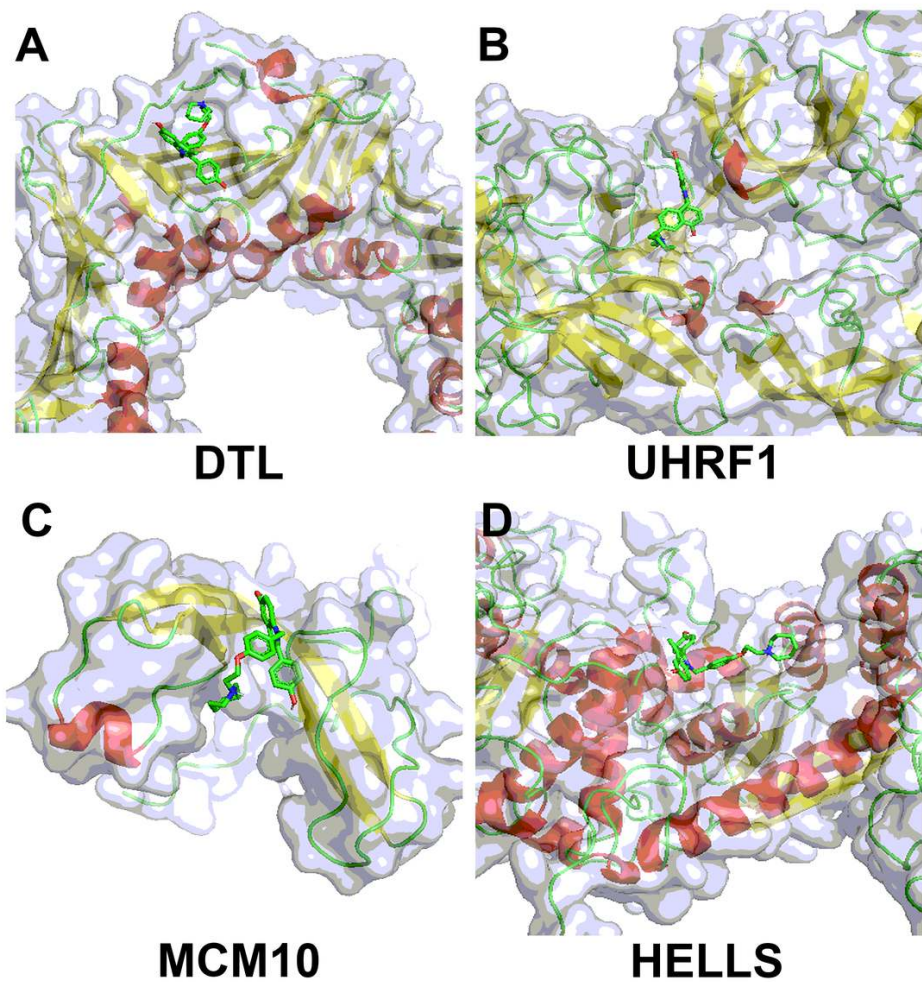


Figure 6

The results of molecular docking. A, the best model (affinity: -6.6 kcal/mol) of bazedoxifene binding to DTL. B, the best model (affinity: -7.8 kcal/mol) of bazedoxifene binding to UHRF1. C, the best model (affinity: -6.2 kcal/mol) of bazedoxifene binding to MCM10. D, the best model (affinity: -7.2 kcal/mol) of bazedoxifene binding to HELLS.