

Analysis of the Prognosis Related Genes in HER2+ Breast Cancer Based on Weighted Gene Co-expression Network Analysis

Yujie WENG

Inner Mongolia Medical University <https://orcid.org/0000-0002-5426-8147>

RONG JIA

Inner Mongolia Medical University

ZHONGXIAN LI

Inner Mongolia Medical University

WEI LIANG

Inner Mongolia Medical University

YUCHENG JI

Inner Mongolia Medical University

YING LIANG

Inner Mongolia Medical University

PENGFEI NING (✉ ningpengfei@immu.edu.cn)

Inner Mongolia Medical University

Research

Keywords: Bioinformatics, WGCNA, HER2+ breast cancer, COX regression models

Posted Date: July 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-704894/v1>

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

[Read Full License](#)

Abstract

Background: Breast cancer is one of the malignant tumors that threaten women's health, with HER2+ breast cancer being more aggressive. In this study, bioinformatics methods were used to find potential key genes in HER2 + for diagnosis and treatment.

Methods: Datasets of HER2+ breast cancer and normal tissue samples retrieved from TCGA databases were subjected to DEGs analysis using R software. Then WGCNA is constructed for DEGs. The key gene co-expression modules were then subjected to GO and KEGG pathway enrichment analyses, as well as construction of PPI networks using the STRING database for identifying key genes. Finally, key genes were further validated by survival analysis, protein expression, and COX regression models.

Results: We identified 2063 DEGs and 4 gene co-expression modules. Functional enrichment analysis showed that these key co-expression modules were mainly associated with extracellular matrix organization, extracellular matrix structural constituent and neuroactive ligand–receptor interaction. PPI network visualization identified 100 key genes, 3 of which were not present in the other subtypes of breast cancer. UTS2 DRD4 and GLP1R are key genes specific to the HER2+ subtype. Survival analysis showed that UTS2 are prognosis-related key genes in HER2+ breast cancer. Finally, UTS2 combined with clinical data to construct Cox regression model.

Conclusions: Combined with the two screening methods, 3 key genes closely related to HER2 + breast cancer were identified. UTS2 is a new potential key gene and may become a new therapeutic target for HER2 + breast cancer.

Background

Breast cancer is one of the important diseases leading to female death, which seriously threatens women's health. With the progress of medical treatment, there are many methods in the treatment of breast cancer, such as drug therapy, hormone therapy and so on. Among them, molecular targeted therapy greatly improves the therapeutic effect of breast cancer, so it is of great significance to find important molecular markers[1–2]. According to the differences of estrogen receptor (ER), human peripheral growth factor receptor 2 (HER2), and progesterone receptor (PR), breast cancer can be divided into four subtypes: triple-negative (TN), luminal A, luminal B, and HER2-positive[3].

HER2 + breast cancer accounts for 15% – 20% of breast cancer, which is caused by erbB2 / neu oncogene amplification or overexpression of HER2 transmembrane receptor protein [4–6]. HER2 is a member of transmembrane receptor tyrosine kinase family (EGFR / HER1, HER2, HER3, HER4), and they are involved in cell proliferation, motility, anti apoptosis, invasiveness and angiogenesis [7]. HER2 + breast cancer has higher grade, more aggressive phenotype and worse prognosis[8]. Therefore, from the perspective of bioinformatics, this study aims to find the key genes related to its development and prognosis, and provide new ideas for its treatment.

DNA microarray and sequencing have become essential and efficient methods in cancer research following the development of high-throughput sequencing (HTS) technology in recent years[9–10]. Weighted Gene Co-Expression Network Analysis (WGCNA) can represent microarray data more completely by considering the relationship between transcripts measured, which can be evaluated by pairwise correlation between gene expression profiles. WGCNA starts from the level of thousands of genes to identify the gene modules of clinical interest. Finally, it uses the connectivity and gene importance to identify the key genes in the disease pathway for further verification.

In this study, we analyzed DNA microarray data retrieved from The Cancer Genome Atlas (TCGA) databases using bioinformatics methods to provide a theoretical basis for exploring potential targets for the treatment and prognostic assessment of HER2 + breast cancer. It is hoped that this study will provide assistance in the treatment of HER2 + breast cancer.

Methods

Datasets

The TCGA database (<https://portal.gdc.cancer.gov/>) includes clinical sample information and sequencing data [11]. The transcriptome data on TCGA was sequenced using whole transcriptome sequencing and we used RNAseq measured expression mRNA. We downloaded 134 breast cancer samples using R's TCGAblinks toolkit, including 47 HER2 + breast cancer samples and 87 normal tissue samples [12].

Identification of DEGs

The gene expression data were divided into HER2 + breast cancer and normal groups, and differential analysis of the data was performed using the limma package in R language [13]. The final differentially expressed genes were filtered using $|\logFC| > 2.0$ and $P < 0.01$.

Co-expression Module Identification

The WGCNA analysis method can search for synergistically expressed gene modules and explore the association between gene networks and phenotypes. In this study, we analyzed the differentially expressed genes by WGCNA [14–17]. The WGCNA package in R was used to analyze the differential genes obtained in the previous step and to normalize the data. Clinical information was used for the type of information of the data, including normal and cancer groups. The soft threshold $\beta = 24$ is obtained by picksoftthreshold function to construct co expression matrix.

Functional Enrichment Analysis

In order to further explore the genes in the co-expression network obtained in the previous step, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis using the clusterProfiler package in R. The GO analysis separated the function of the genes into

three components: biological process (BP), cellular component (CC), and molecular function (MF) [18–19]. $P < 0.05$ was used as a filter for both GO and KEGG.

PPI networks and Hub Genes

We used the STRING(<https://string-db.org>) online tool to construct PPI networks of key co-expression modules[20–21]. The PPI networks were then imported into Cytoscape software for visualization and further analysis and used the MCC algorithm of the CytoHubba plugin to search for 100 key genes[22–23]. In order to differentiate HER2 + breast cancer from other breast cancer subtypes, we performed the same differential analysis of mRNA data for the other three breast cancer subtypes (basal, lumA and lumB), which were also downloaded from the TCGA database, using the same methodology using the limma package in R. The 100 key genes obtained by MCC were compared with the differential genes of the other three subtypes, which yielded key genes that were present only in the HER2 + subtype, but not in the other subtypes, as key genes that distinguish the HER2 + subtype from the other subtypes[24].

Survival analysis of key genes and Validation of Protein Expressions

The online tool PROGgeneV2 (<http://genomics.jefferson.edu>) was used to view the relationship between UTS2, DRD4 and overall survival of HER2 + patients. Divide the data into high and low expression groups based on the median gene expression value and look at the relationship between the high and low expression groups of these two genes and survival time, respectively[25–26]. In addition, the Human Protein Atlas (HPA, <https://www.proteinatlas.org/>) database was used to view the expression of key gene in tissues, cells, and tumors[27–28].

Cox Proportional Hazards Regression Model

The Cox proportional hazards regression model uses survival time as a strain, which allows simultaneous analysis of the effect of multiple factors on survival. In this study, we used R language survival package combined with clinical data including age and tumor grade as well as the selected key gene to analyze the effect of these factors on the 2-year and 4-year survival of HER2 + breast cancer patients, which can further validate the role of key gene in HER2 + breast cancer patients.

Results

Identification of DEGs

After making differential genes for 134 sets of samples, 2063 differential genes were obtained according to the screening conditions $|\log_{2}FC| > 2.0$ and $P < 0.01$, of which 897 genes were up-regulated and 1166 genes were down-regulated. The differential genes were displayed in a volcano diagram as in Fig. 1.

Co-expression Module Identification

The co-expression network was constructed for the resulting 2063 differential genes, yielding a total of four modules, excluding the gray module. The different colors represent different modules, as shown in

Fig. 2A. Then a correlation heat map of the module and sample traits was drawn, as shown in Fig. 2B, with rows representing modules and columns representing traits, and traits including both HER2 + and normal. According to the correlation heat map, it can be seen that the module with the highest correlation with HER2 + breast cancer is the turquoise module with a correlation of 0.94. The turquoise module is a key co-expression module with a gene count of 760.

Functional Enrichment Analysis

The 760 genes in the key co-expression module obtained in the previous step were subjected to GO and KEGG pathway analysis in order to further understand the functions of these genes, as shown in Fig. 3. In the GO analysis, the BP of the key co-expression modules were focused on extracellular matrix organization, extracellular structure organization and regulation of ion transmembrane transport, CC mainly include collagen-containing extracellular matrix, transporter complex, apical part of cell and DNA packaging complex, and MF is mainly related to extracellular matrix structural constituent, glycosaminoglycan binding, and receptor ligand activity. KEGG pathway analysis of key co-expression modules revealed that these key genes primarily function in pathways including neuroactive ligand – receptor interaction, cytokine – cytokine, receptor interaction, tyrosine metabolism and phenylalanine metabolism.

PPI networks and Hub Genes

A PPI network was constructed for 760 key genes of key co-expression modules and 100 key genes were found using the MCC algorithm of the CytoHubba plugin (Fig. 4). Differential gene analysis was also used to obtain differential genes for the other three subtypes of breast cancer, with the same screening method $|\log_{2}FC| > 2.0$ and $P < 0.01$. The differential gene results were as follows, with basal 2093, lumA 1365, and lumB 2084. The 100 key genes obtained by the MCC algorithm were compared two-by-two with the differential genes of the other three subtypes to identify genes present only in the HER2 subtype and not in the other subtypes, DRD4 UTS2 and GLP1R, respectively, as shown in Fig. 5. DRD4 and UTS2 were up-regulated and GLP1R was down-regulated. To verify the accuracy of the key genes, we plotted their expression values in box plots across the four subtypes, as shown in Fig. 6, where UTS2 is expressed higher in the HER2 + subtype than in the other subtypes. DRD4 was most highly expressed in the lumA subtype, but its $\log_{2}FC$ value was less than 2, which did not meet the screening criteria for differential genes, as shown in Table 1. The $\log_{2}FC$ value for DRD4 in the HER2 subtype was 2.11, which exceeded the screening criteria by 2. GLP1R is less expressed in the HER2 subtype than in the other subtypes. It is therefore concluded that DRD4 UTS2 and GLP1R were key genes that distinguish the HER2 + subtype from other subtypes.

Table 1
results of differential expression of DRD4 UTS2 and GLP1R

genesymbol	subtype	logFC	t	P.Value	adj.P.Val
DRD4	lumA	1.9667676	13.53841	2.60E-34	2.68E-33
DRD4	lumB	1.4850377	8.286951	1.53E-14	5.42E-14
DRD4	basal	1.7474764	9.56844	7.61E-18	3.49E-17
DRD4	HER2+	2.1192462	9.945784	9.10E-18	5.89E-17
UTS2	lumA	1.0425104	6.531077	2.11E-10	5.49E-10
UTS2	lumB	1.4346405	7.702268	5.68E-13	1.82E-12
UTS2	basal	1.7886484	7.733012	6.77E-13	2.21E-12
UTS2	HER2+	2.2715777	8.116018	2.84E-13	1.22E-12
GLP1R	lumA	-1.56017	-7.32302	1.48E-12	4.40E-12
GLP1R	lumB	-1.80133	-7.09725	2.06E-11	5.89E-11
GLP1R	basal	-0.7034848	-2.191190	0.029700	0.04073
GLP1R	HER2+	-2.09328	-8.33133	8.62E-14	3.89E-13

Survival analysis of key genes and Validation of Protein Expressions

The PROGgeneV2 database was used to view the effect of UTS2 DRD4 and GLP1R expression on the prognosis of HER2 + patients. The data used for the survival analysis was GSE6130 from the GEO database, and the data was divided into high and low expression groups according to median gene expression to see its impact on survival in HER2 + positive patients. The results showed that patients with low expression of DRD4 and GLP1R had a lower survival time than those with high expression, while UTS2 had a lower survival rate with high expression, as shown in Fig. 7. In addition, in the survival analysis of UTS2, the p-value was 0.0025, which was statistically significant. From the above analysis, it is evident that high expression of UTS2 is significantly associated with low survival in HER2 + patients. To further validate UTS2, the HPA database was used to view its expression in normal breast tissues as well as breast cancer tissues, and the results are shown in Fig. 8. The results showed that the protein expression level of UTS2 was significantly higher in breast cancer than in normal tissues, further validating the accuracy of the results.

Cox Proportional Hazards Regression Model

In the Cox regression model, clinical data were downloaded from the TCGA database for HER2 + patients, where age was the actual age of the patient and tumor grade was represented by 1–4 for T1-T4 grade, respectively. In the expression of UTS2, 0 represented low expression and 1 represented high expression.

The Cox regression model was constructed using the data as above and combined with the patient's survival time, and the model results were displayed as a nomogram, as shown in Fig. 9A. By adding up the points corresponding to each index, which corresponds to the total points, we can get the survival rate of 2 years and 4 years for different patients. It can be seen that higher expression of UTS2 corresponds to higher points, which corresponds to lower survival rate. In addition, age and tumor grade are also positively correlated with lower survival rates. To assess the accuracy of the nomogram, we plotted the calibration curves, which were in good agreement for both predictions, as shown in Fig. 9B,C.

Discussion

In this study, microarray datasets of HER2 + breast cancer retrieved from TCGA databases were subjected to differential gene expression analysis by bioinformatics methods. The gene co-expression networks were constructed from the DEGs of TCGA with the WGCNA package. GO and KEGG pathway enrichment analyses demonstrated that these significantly DEGs are closely related to extracellular matrix organization, collagen-containing extracellular matrix, extracellular matrix structural constituent and neuroactive ligand – receptor interaction, etc. The PPI network identified 100 key genes. Compared with the difference genes of lumA, lumB and basal subtypes, 3 key genes, UTS2, DRD4 and GLP1R, were found only in HER2 subtypes. The survival analysis using the PROGgeneV2 online tool showed that the results of UTS2 showed statistical significance ($P < 0.05$) and the protein levels of the UTS2 gene was significantly higher in tumor tissues compared with normal tissues based on the HPA database. COX regression analysis also showed that high expression of UTS2 was negatively associated with 2-year and 4-year survival in HER2 + patients. UTS2 is a potential prognosis-related key genes in HER2 + breast cancer.

DRD4 is a G protein coupled receptor coding gene, which encodes the D4 subtype of the dopamine receptor and the receptors is activated by the neurotransmitter dopamine. Studies have shown that dopamine receptor is related to the death and migration of tumor cells, and can regulate various death modes of tumor cells[29–30]. DRD4 was highly expressed in breast cancer, neuroblastoma and small cell lung cancer, and its expression level was related to the survival of patients[31]. We found that DRD4 was highly expressed in all four subtypes of breast cancer, the most obvious of which was lumA subtype, but its logFC value did not exceed the threshold. The logFC value of HER2 + subtype was more than 2, which met the threshold requirements. Survival analysis showed that high expression of DRD4 did not show lower survival rate in HER2 + breast cancer, so we did not further analyze it.

GLP1R (Glucagon Like Peptide 1 Receptor) is a Protein Coding gene which is a member of G protein coupled receptor family. GLP-1R is an important target when treating of type 2 diabetes and obesity[32–34]. Recent studies have shown that GLP-1R activation has a cytoreductive effect on human pancreatic cancer cells[35]. We found that GLP-1R was low expressed in HER2 +, lumA and lumB subtypes, and it was most obvious in HER2 + breast cancer. Survival analysis showed that the low expression of GLP-1R showed a lower survival rate, indicating that GLP-1R is an important gene related to the survival of HER2 + patients.

UTS2 (urotensin 2) has 3 transcripts (splice variants), 251 orthologues and is a member of 1 Ensembl protein family. It is a Protein Coding gene. Diseases associated with UTS2 include Portal Hypertension and Congestive Heart Failure. Among its related pathways are RET signaling and Signaling by GPCR[36]. Gene Ontology (GO) annotations related to this gene include signaling receptor binding and hormone activity[37]. Studies have found that urotensin-II could contribute to breast carcinogenesis and Thr21Met polymorphism can be an important risk factor in developing breast tumors[38]. UTS2 might be used as potential targets to improve diagnosis and immunotherapy biomarkers for renal cell carcinoma and colon cancer [39–40].

Conclusions

Our study found that UTS2 was highly expressed in HER2 + breast cancer, but not in lumA, lumB and basal. Additionally, the high expression of UTS2 is also significantly associated with the low survival rate of patients with HER2 + breast cancer. Nomogram proves it too. Therefore, UTS2 may be a key target for the diagnosis and treatment of HER2 + breast cancer.

Even though our results require experimental validation, the identified genes are potential targets for the treatment and prognostic assessment of HER2 + breast cancer.

List Of Abbreviations

HER2, Human epidermal growth factor receptor 2;

KEGG, Kyoto Encyclopedia of Genes and Genomes;

OS, Overall survival;

TCGA, The Cancer Genome Atlas;

WGCNA, Weighted correlation network analysis;

GO, Gene Ontology;

BP, Biological process;

CC, Cellular component;

MF, Molecular function.

PPI, Protein-protein interaction

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

TCGA database (<https://portal.gdc.cancer.gov/>)

Competing interests

Not applicable

Funding

Not applicable

Authors' contributions

Pengfei Ning, Zhongxian Li, Wei Liang designed the study;

Yujie WENG, Yucheng JI researched literature;

Rong Jia , Ying Liang wrote the main manuscript text;

All authors reviewed the manuscript.

Acknowledgements

Not applicable

References

1. Malvika Solanki, Daniel Visscher. Pathology of breast cancer in the last half century. *Hum Pathol.* 2020 Jan;95:137-148.
2. Anbok Lee, Byung In Moon, Tae Hyun Kim. BRCA1/ BRCA2 Pathogenic Variant Breast Cancer: Treatment and Prevention Strategies. *Ann Lab Med.* 2020 Mar;40(2):114-121.
3. Deepti Tomar, Amit S Yadav, Dhiraj Kumar, Garima Bhadauriya, Gopal C Kundu. Non-coding RNAs as potential therapeutic targets in breast cancer. *Biochim Biophys Acta Gene Regul Mech.* 2020 Apr;1863(4):194378.
4. Rebecca L Siegel, Kimberly D Miller, Ahmedin Jemal. Cancer statistics, 2016. *CA Cancer J Clin.* Jan-Feb 2016;66(1):7-30

5. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, Deming SL, Geradts J, Cheang MC, Nielsen TO, Moorman PG, Earp HS and Millikan RC. Race, breast cancer subtypes, and survival in the carolina breast cancer study. *JAMA* 2006; 295: 2492.
6. Cronin KA, Harlan LC, Dodd KW, Abrams JS and Ballard-Barbash R. Population-based estimate of the prevalence of HER-2 positive breast cancer tumors for early stage patients in the US. *Cancer Invest* 2010; 28: 963-968.
7. Giuliano AE and Hurvitz SA. Carcinoma of the female breast. In: Papadakis MA, McPhee SJ, Rabow MW, editors. *Current medical diagnosis and treatment 2020*. New York, NY: McGraw-Hill Education; 2020.
8. Ross JS, Slodkowska EA, Symmans WF, Pusztai L, Ravdin PM and Hortobagyi GN. The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine. *Oncologist* 2009; 14: 320-368.
9. Xiang Jin, Lijun Xu, Yinghui Guan, Zhen Zhang, Hongyu Li. Bioinformatics Analysis of Microarray Datasets to Identify Prognostic Factors in Lung Adenocarcinoma. *DNA Cell Biol.* 2020 Jun;39(6):965-974.
10. Mahsa Gharibi Marzancola, Abootaleb Sedighi, Paul C H Li. DNA Microarray-Based Diagnostics. *Methods Mol Biol.* 2016;1368:161-78.
11. Katarzyna Tomczak, Patrycja Czerwińska, Maciej Wiznerowicz. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn).* 2015;19(1A):A68-77.
12. Antonio Colaprico, Tiago C Silva, Catharina Olsen, et. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016 May 5;44(8):e71.
13. Ligang Bao, Ting Guo, Ji Wang, Kai Zhang, Maode Bao. Prognostic genes of triple-negative breast cancer identified by weighted gene co-expression network analysis. *Oncol Lett.* 2020 Jan;19(1):127-138.
14. Zhen-Hu Ren, Gao-Pan Shang, Kun Wu, Chuan-Yu Hu, Tong Ji. WGCNA Co-Expression Network Analysis Reveals ILF3-AS1 Functions as a CeRNA to Regulate PTBP1 Expression by Sponging miR-29a in Gastric Cancer. *Front Genet.* 2020 Feb 14;11:39.
15. Kun-Hao Bai, Si-Yuan He, Ling-Ling Shu, et. Identification of cancer stem cell characteristics in liver hepatocellular carcinoma by WGCNA analysis of transcriptome stemness index. *Cancer Med.* 2020 Jun;9(12):4290-4298.
16. Min Chen, Jianying Yan, Qing Han, Jinying Luo, Qinjian Zhang. Identification of hub-methylated differentially expressed genes in patients with gestational diabetes mellitus by multi-omic WGCNA basing epigenome-wide and transcriptome-wide profiling. *J Cell Biochem.* 2020 Jun;121(5-6):3173-3184.
17. Haijun Wang, Jia Liu, Jinsong Li, et. Identification of gene modules and hub genes in colon adenocarcinoma associated with pathological stage based on WGCNA analysis. *Cancer Genet.* 2020 Apr;242:1-7.

18. Kévin Rue-Albrecht, Paul A McGettigan, Belinda Hernández, et al. GOexpress: an R/Bioconductor package for the identification and visualisation of robust gene ontology signatures through supervised learning of gene expression data. *BMC Bioinformatics*. 2016 Mar 11;17:126.
19. Bernhard Mlecnik, Jérôme Galon, Gabriela Bindea. Automated exploration of gene ontology term and pathway networks with ClueGO-REST. *Bioinformatics*. 2019 Oct 1;35(19):3864-3866.
20. Helen Victoria Cook, Nadezhda Tsankova Doncheva, Damian Szklarczyk, Christian von Mering, Lars Juhl Jensen. STRING: A Virus-Host Protein-Protein Interaction Database. *Viruses*. 2018 Sep 23;10(10):519.
21. Damian Szklarczyk, John H Morris, Helen Cook, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017 Jan 4;45(D1):D362-D368.
22. Paul Shannon, Andrew Markiel, Owen Ozier, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003 Nov;13(11):2498-504.
23. Nadezhda T Doncheva, John H Morris, Jan Gorodkin, Lars J Jensen. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res*. 2019 Feb 1;18(2):623-632.
24. Maud S Jurgens, Johannes W G Jacobs, Maarten Boers, et al. Alternative Ways to Quantify Sustained Remission: Applying the Continuity Rewarded Score and Patient Vector Graph. *Arthritis Care Res (Hoboken)*. 2015 Oct;67(10):1471-4.
25. Chirayu Pankaj Goswami, Harikrishna Nakshatri. PROGgeneV2: enhancements on the existing database. *BMC Cancer*. 2014 Dec 17;14:970.
26. De-Zhong Wang, Guan-Yang Chen, Yi-Feng Li, Neng-Wei Zhang. Comprehensive analysis of long non-coding RNA and mRNA expression profile in rectal cancer. *Chin Med J (Engl)*. 2020 Jun 5;133(11):1312-1321.
27. Peter J Thul, Cecilia Lindskog. The human protein atlas: A spatial map of the human proteome. *Protein Sci*. 2018 Jan;27(1):233-244.
28. Zhou Zhou, Yian Cheng, Yinan Jiang, et al. Ten hub genes associated with progression and prognosis of pancreatic carcinoma identified by co-expression analysis. *Int J Biol Sci*. 2018 Jan 12;14(2):124-136.
29. Ya-Ting Wen, Alexander Th Wu, Oluwaseun Adebayo Bamodu, Li Wei, Chien-Min Lin, Yun Yen et al. A Novel Multi-Target Small Molecule, LCC-09, Inhibits Stemness and Therapy-Resistant Phenotypes of Glioblastoma Cells by Increasing miR-34a and Deregulating the DRD4/Akt/mTOR Signaling Axis. *Cancers (Basel)*. 2019 Sep 26;11(10):1442.
30. Sonam Dolma, Hayden J Selvadurai, Xiaoyang Lan, Lilian Lee, Michelle Kushida, Veronique Voisin, et al. Inhibition of Dopamine Receptor D4 Impedes Autophagic Flux, Proliferation, and Survival of Glioblastoma Stem Cells. *Cancer Cell*. 2016 Jun 13;29(6):859-873.
31. Xu Wang, Zhi-Bin Wang, Chao Luo, Xiao-Yuan Mao, Xi Li, Ji-Ye Yin, et al. The Prospective Value of Dopamine Receptors on Bio-Behavior of Tumor. *J Cancer*. 2019 Mar 3;10(7):1622-1632.

32. He-Jun Zhao, Xia Jiang, Li-Juan Hu, Lei Yang, Lian-Dong Deng, Ya-Ping Wang, et al. Activation of GLP-1 receptor enhances the chemosensitivity of pancreatic cancer cells. *J Mol Endocrinol*. 2020 Feb;64(2):103-113.
33. Yuanyuan Yue, Yuping Xu, Lirong Huang, Donghui Pan, Zhicheng Bai, Lizhen Wang, et al. Evaluation of A Novel GLP-1R Ligand for PET Imaging of Prostate Cancer. *Anticancer Agents Med Chem*. 2019;19(4):509-514.
34. Hejun Zhao, Rui Wei, Liang Wang, Qing Tian, Ming Tao, Jing Ke, et al. Activation of glucagon-like peptide-1 receptor inhibits growth and promotes apoptosis of human pancreatic cancer cells in a cAMP-dependent manner. *Am J Physiol Endocrinol Metab*. 2014 Jun 15;306(12):E1431-41.
35. Bendong Chen, Wenyan Zhou, Wenchao Zhao, Peng Yuan, Chaofeng Tang, Genwang Wang, et al. Oxaliplatin reverses the GLP-1R-mediated promotion of intrahepatic cholangiocarcinoma by altering FoxO1 signaling. *Oncol Lett*. 2019 Aug;18(2):1989-1998.
36. Betül Ozan, Seniz Demiryürek, Muhammad Safdar, Yusuf Inanc, Abdullah Tuncay Demiryürek. Lack of association between urotensin-II (UTS2) gene polymorphisms (Thr21Met and Ser89Asn) and migraine. *Bosn J Basic Med Sci*. 2017 Aug 20;17(3):268-273.
37. Robyn G Langham, Darren J Kelly. Urotensin II and the kidney. *Curr Opin Nephrol Hypertens*. 2013 Jan;22(1):107-12.
38. Onder Yumrutas, Serdar Oztuzcu, Hakan Büyükhatipoglu et al. The role of the UTS2 gene polymorphisms and plasma Urotensin-II levels in breast cancer. *Tumour Biol*. 2015 Jun;36(6):4427-32.
39. Feng Li, Peiyuan Guo, Keqin Dong, Peng Guo, Haoyuan Wang, Xianqiang Lv. Identification of Key Biomarkers and Potential Molecular Mechanisms in Renal Cell Carcinoma by Bioinformatics Analysis. *J Comput Biol*. 2019 Nov;26(11):1278-1295.
40. Haitao Chen, Jun Luo, Jianchun Guo. Development and validation of a five-immune gene prognostic risk model in colon cancer. *BMC Cancer*. 2020 May 6;20(1):395.

Figures

Volcano map

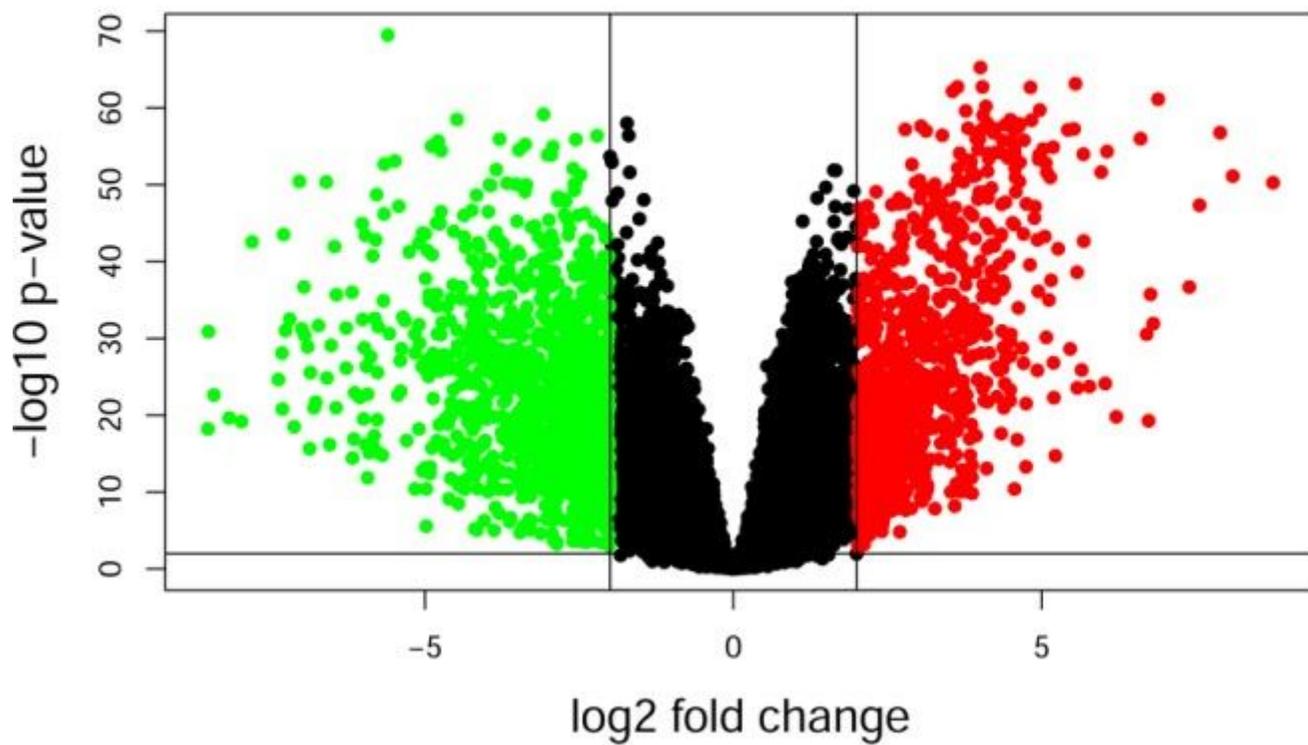


Figure 1

Volcano plot of DEGs in the TCGA dataset.

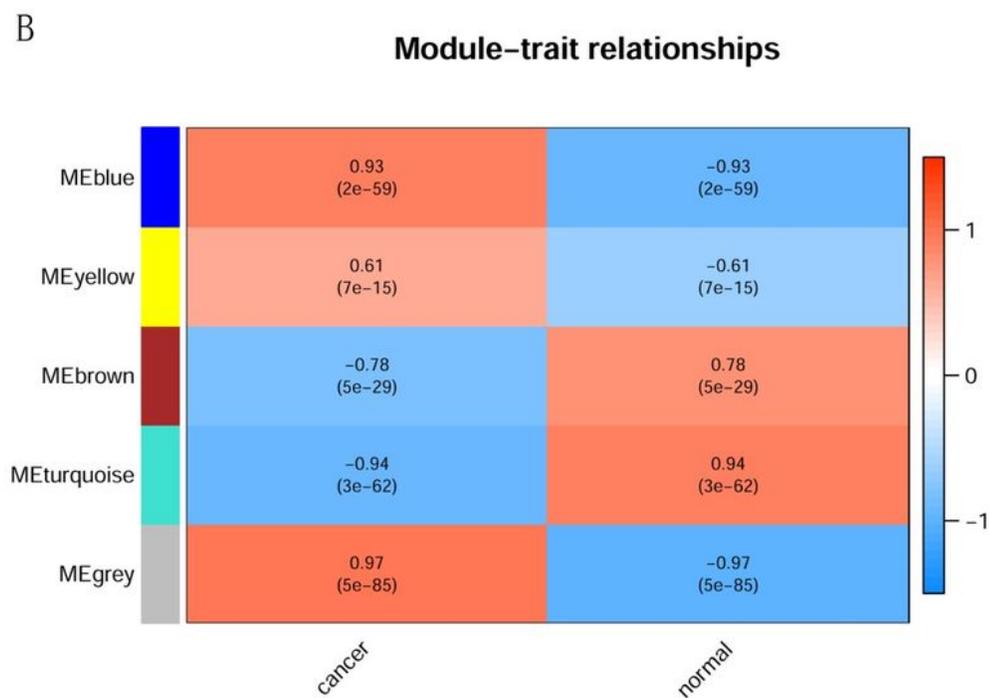
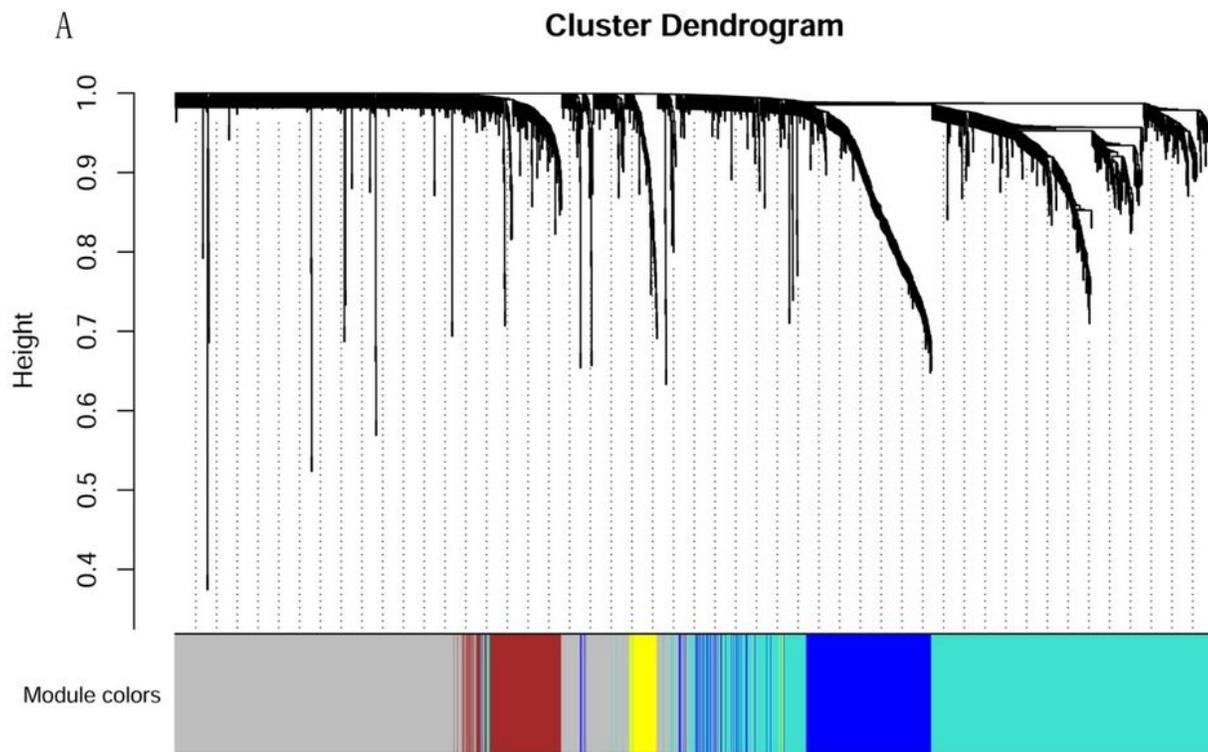


Figure 2

WGCNA of DEGs. (A) The Cluster dendrogram of co-expression network modules. (B) Module-trait relationships.

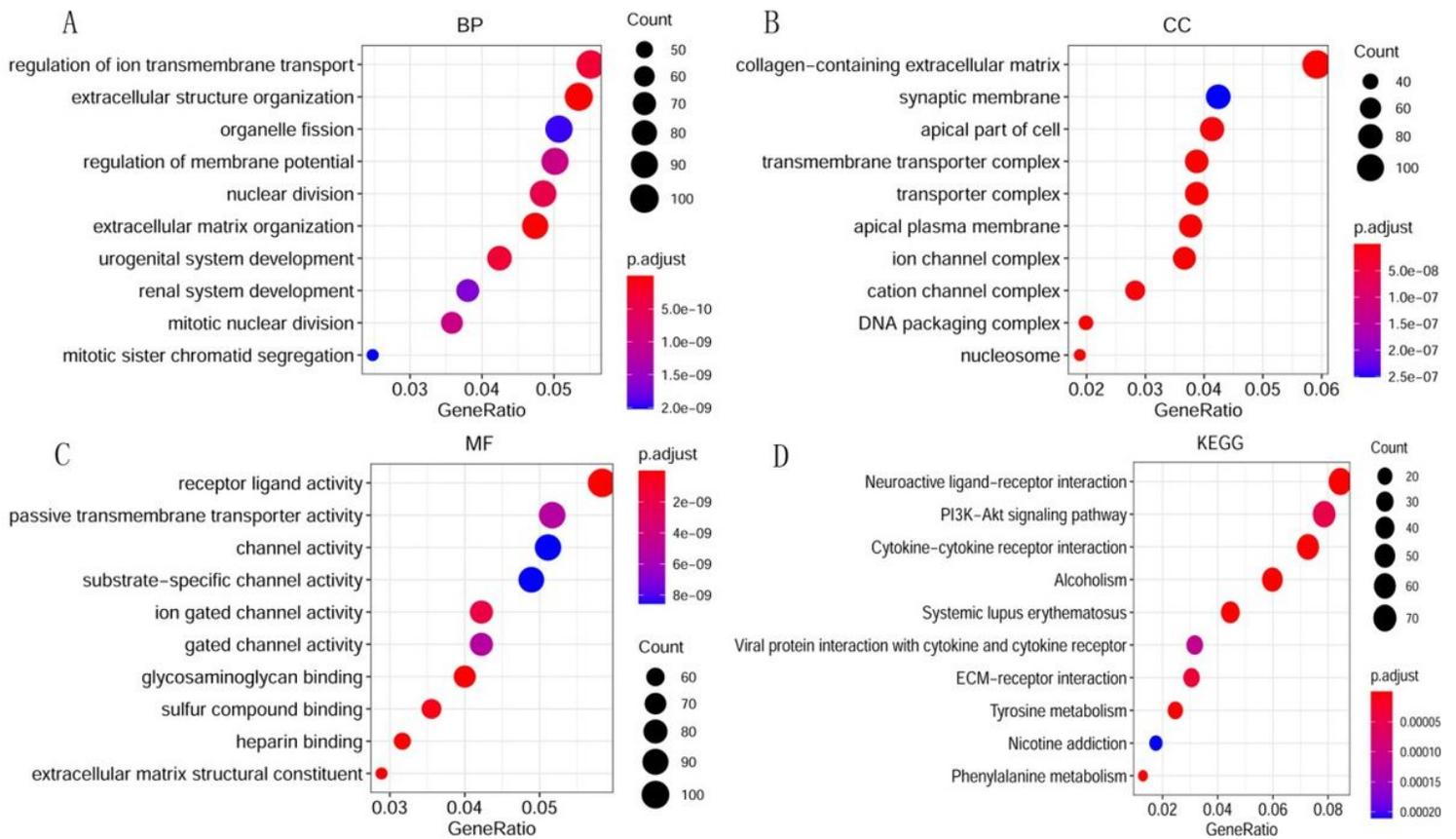


Figure 3

Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis for the genes in the turquoise module. (A) biological process of the turquoise module. (B) cellular component of the turquoise module. (C) molecular function of the turquoise module. (D) Kyoto Encyclopedia of Genes and Genomes of the turquoise module.

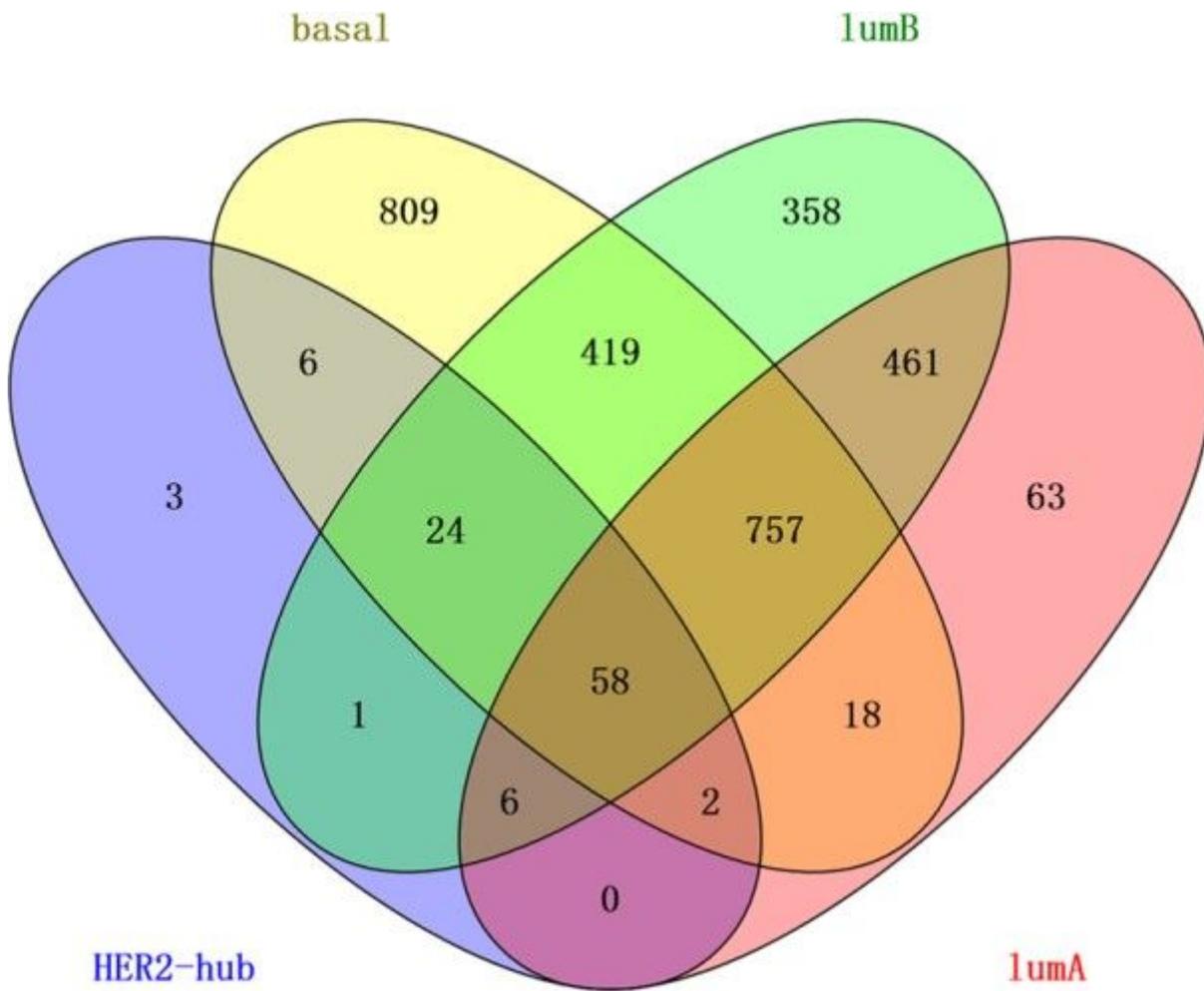


Figure 5

The Venn diagram of genes among DEGs of three subtypes—basal lumA lumB—and hub gene in HER2+. The three genes (DRD4 UTS2 and GLP1R) in the blue part are the key genes that distinguish HER2+ from other subtypes.

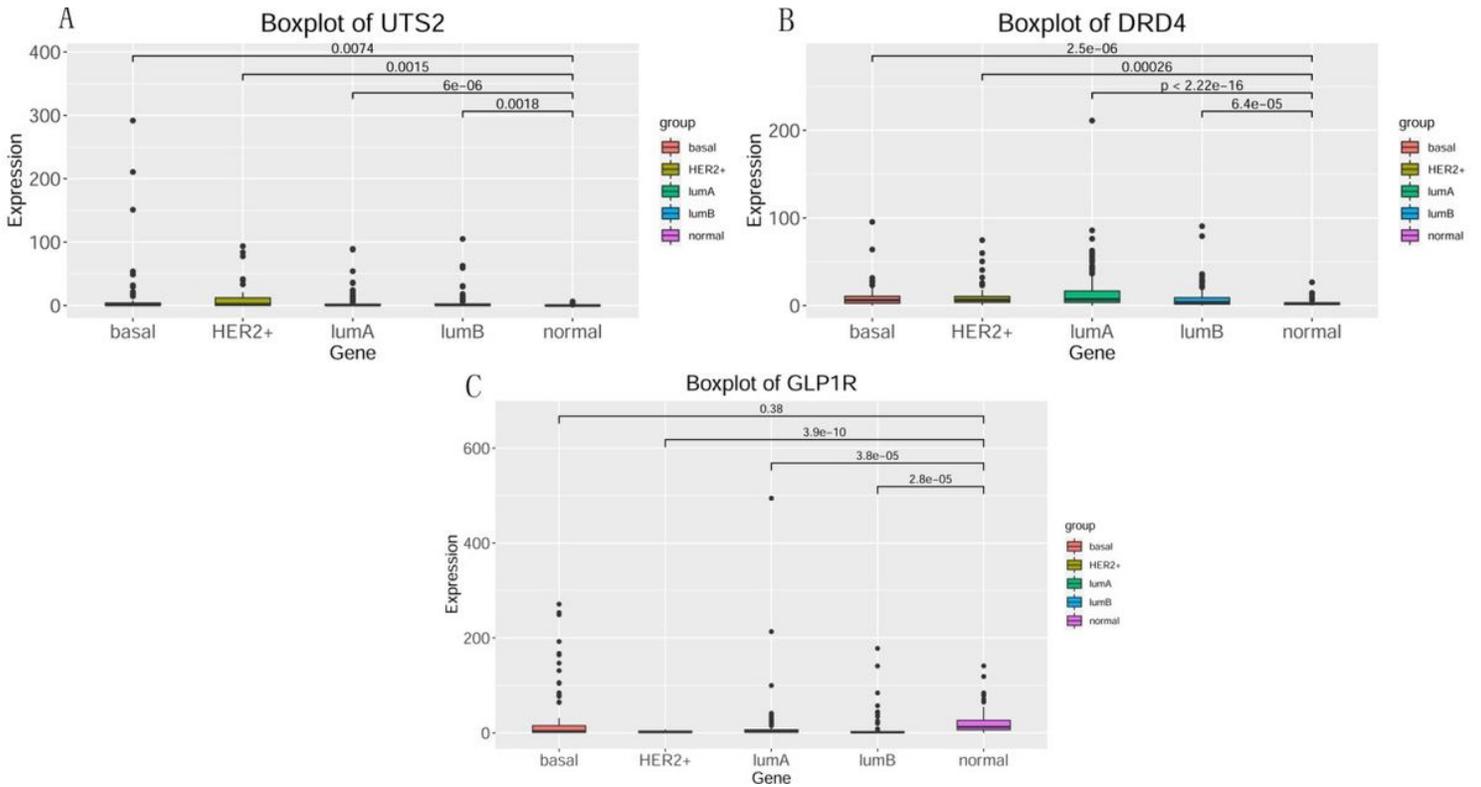


Figure 6

Validation of expression levels of the 2 hub genes among basal, HER2+, lumA, lumB and normal tissues from the TCGA database. (A) Gene expression value UTS2 among samples of TCGA. (B) Gene expression value DRD4 among samples of TCGA. (C) Gene expression value GLP1R among samples of TCGA.

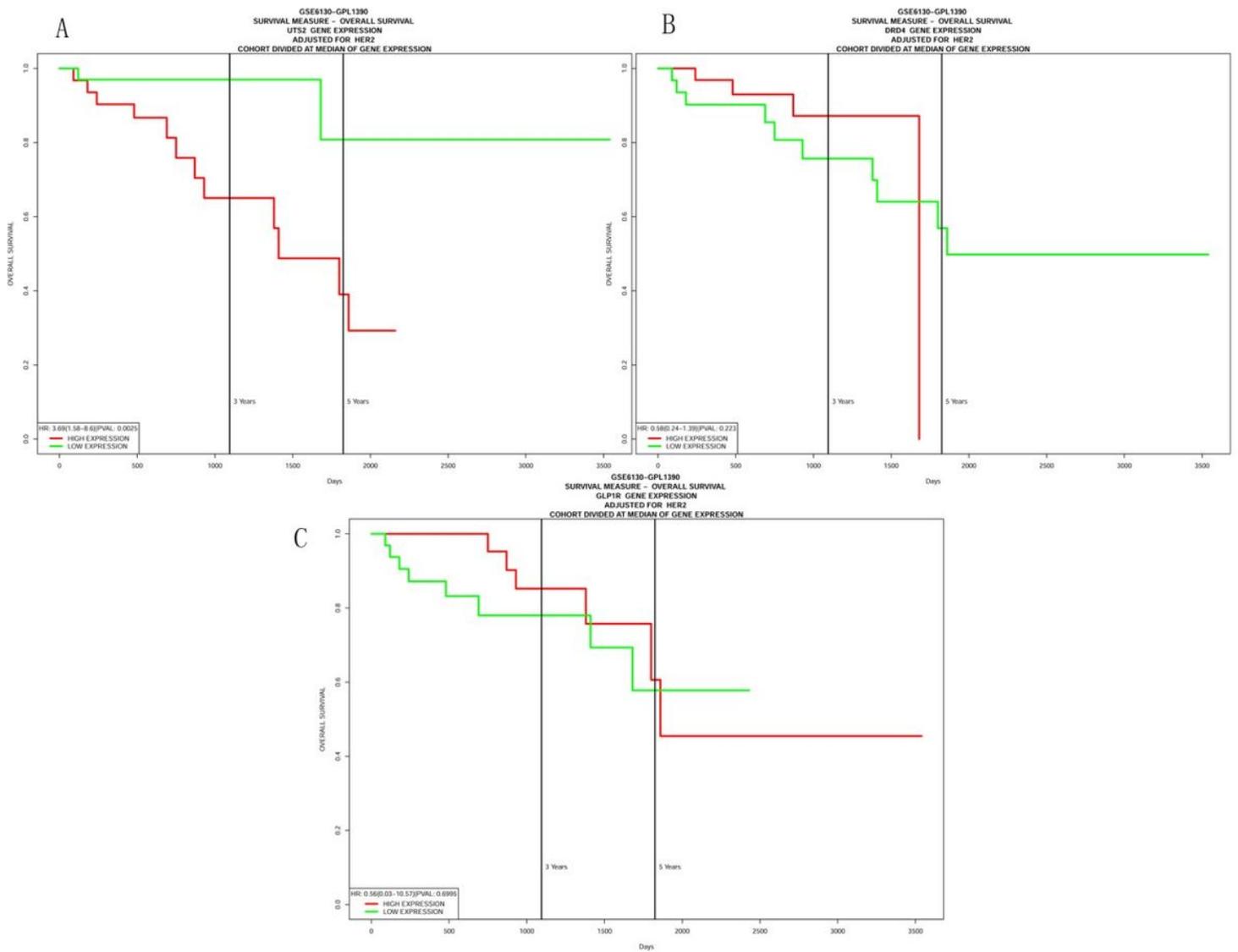


Figure 7

Overall survival (OS) analysis of 2 hub genes in HER2+ patients. (A) Survival analysis for UTS2 in HER2+. (B) Survival analysis for DRD4 in HER2+. (C) Survival analysis for GLP1R in HER2+.

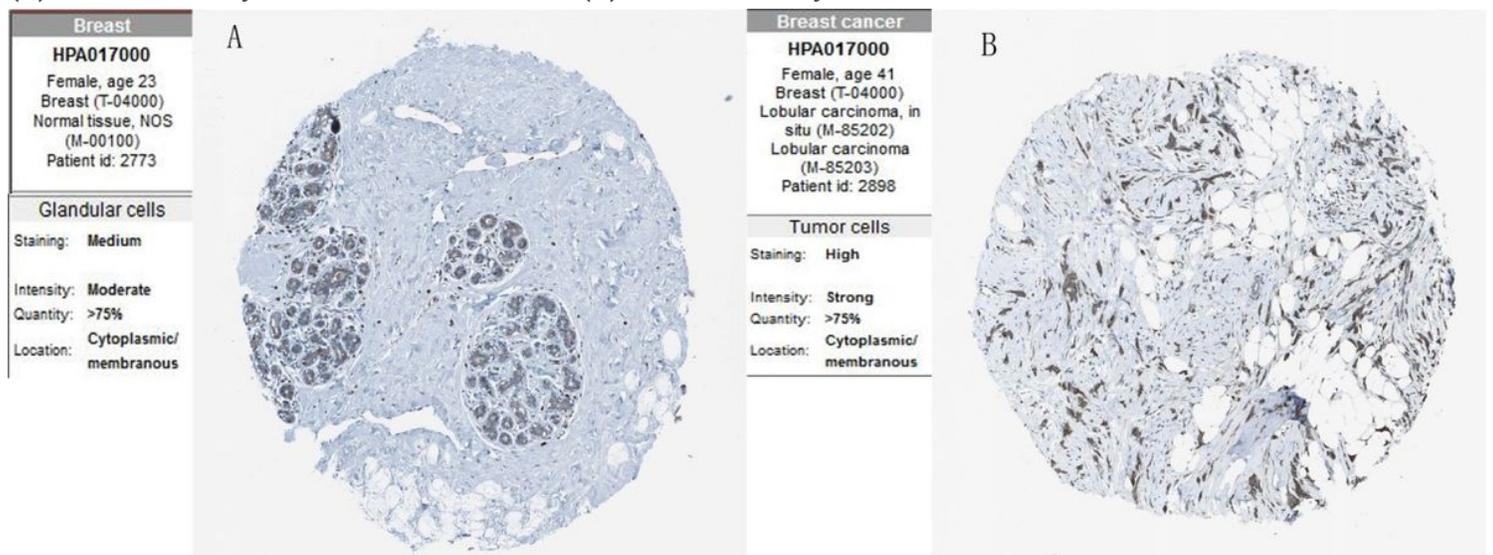


Figure 8

Immunohistochemistry of the UTS2 gene in breast cancer and normal tissues. (A) Protein levels of UTS2 in normal tissues (B) Protein levels of UTS2 in breast cancer tissues.

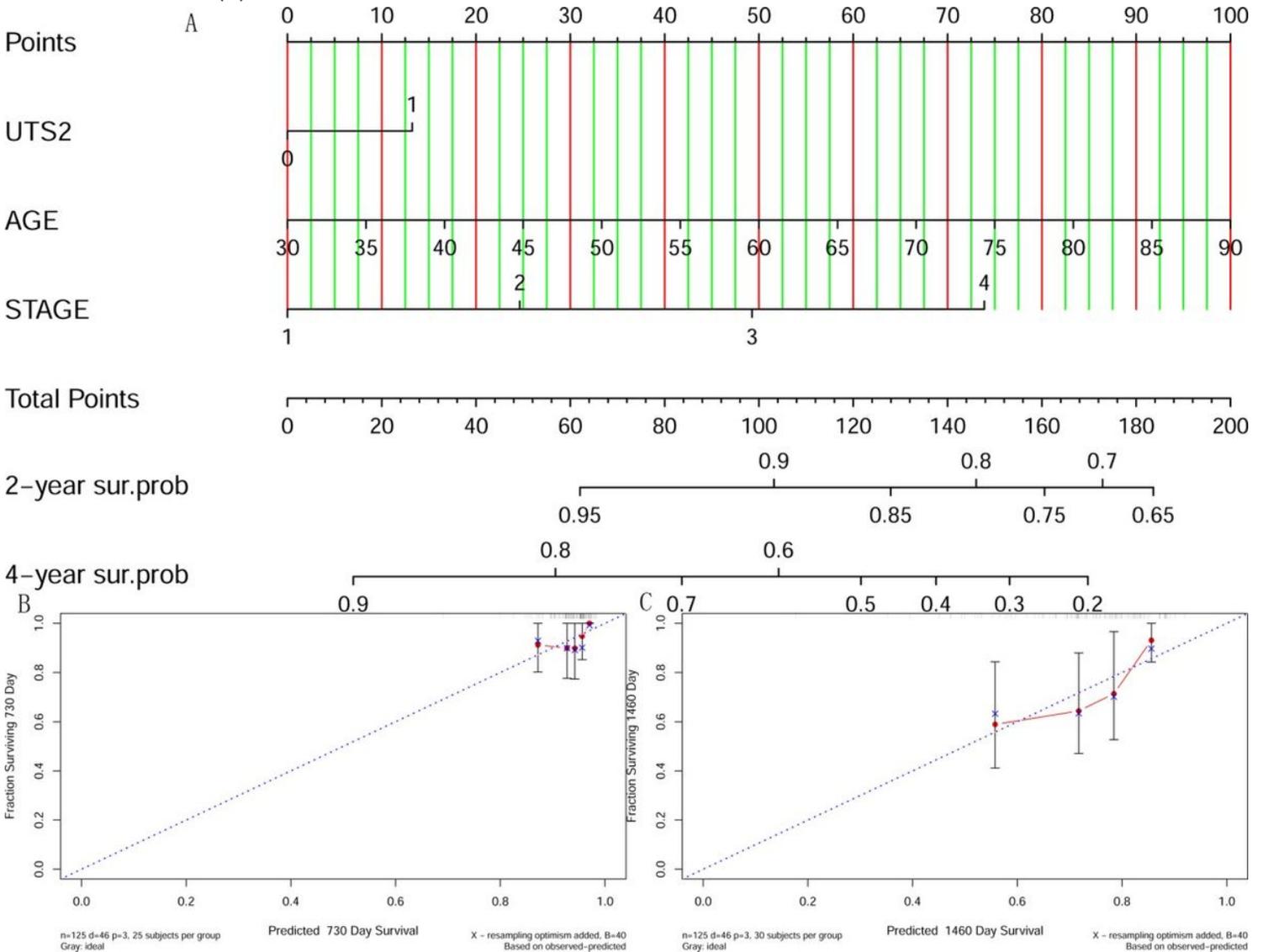


Figure 9

The nomogram can predict the prognosis probability in HER2+. (A) A nomogram used to predict 2- and 4-year survival of HER2+ patients. (B). Calibration maps used to predict the 2-year. (C) Calibration maps used to predict the 4-year survival.