



Construction of Bone Metastasis-Specific Regulation Network Based on Prognostic Stemness-Related Signatures in Breast Invasive Carcinoma

OPEN ACCESS

Edited by:

Dianwen Song,
Shanghai First People's Hospital,
China

Reviewed by:

Zhang Hui,
Fujian Medical University, China
Peng-Chan Lin,
National Cheng Kung University,
Taiwan
Haoran Lv,
Second Affiliated Hospital of
Guangzhou Medical University, China

*Correspondence:

Zongqiang Huang
gzhuangzq@163.com
Bin Wang
wangbin12048@163.com
Penghui Yan
877492185@qq.com
Daoke Yang
15903650068@163.com

[†]These authors have contributed
equally to this work and share
first authorship

Specialty section:

This article was submitted to
Molecular and
Cellular Oncology,
a section of the journal
Frontiers in Oncology

Received: 02 October 2020

Accepted: 26 November 2020

Published: 27 January 2021

Citation:

Huang R, Li Z, Zhang J, Zeng Z,
Zhang J, Li M, Wang S, Xian S, Xue Y,
Chen X, Li J, Cheng W, Wang B,
Yan P, Yang D and Huang Z
(2021) Construction of Bone
Metastasis-Specific Regulation
Network Based on Prognostic
Stemness-Related Signatures in
Breast Invasive Carcinoma.
Front. Oncol. 10:613333.
doi: 10.3389/fonc.2020.613333

Runzhi Huang^{1,2,3†}, Zhenyu Li^{3†}, Jiayao Zhang^{4†}, Zhiwei Zeng¹, Jiaqi Zhang¹,
Mingxiao Li¹, Siqao Wang³, Shuyuan Xian³, Yuna Xue¹, Xi Chen¹, Jie Li¹,
Wenjun Cheng¹, Bin Wang^{5*}, Penghui Yan^{1*}, Daoke Yang^{6*} and Zongqiang Huang^{1*}

¹ Department of Orthopedics, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, ² Division of Spine, Department of Orthopedics, Tongji Hospital Affiliated to Tongji University School of Medicine, Shanghai, China, ³ Tongji University School of Medicine, Tongji University, Shanghai, China, ⁴ Tongji University School of Mathematical Sciences, Tongji University, Shanghai, China, ⁵ Department of General Surgery, Changzheng Hospital, Second Military Medical University, Shanghai, China, ⁶ Department of Radiotherapy, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Background: Bone is the most common metastatic site of Breast invasive carcinoma (BRCA). In this study, the bone metastasis-specific regulation network of BRCA was constructed based on prognostic stemness-related signatures (PSRSs), their upstream transcription factors (TFs) and downstream pathways.

Methods: Clinical information and RNA-seq data of 1,080 primary BRCA samples (1,048 samples without bone metastasis and 32 samples with bone metastasis) were downloaded from The Cancer Genome Atlas (TCGA). The edgeR method was performed to identify differential expressed genes (DEGs). Next, mRNA stemness index (mRNAsi) was calculated by one-class logistic regression (OCLR). To analyze DEGs by classification, similar genes were integrated into the same module by weighted gene co-expression network analysis (WGCNA). Then, univariate and multivariate Cox proportional hazard regression were applied to find the PSRSs. Furthermore, PSRSs, 318 TFs obtained from Cistrome database and 50 hallmark pathways quantified by GSEA were integrated into co-expression analysis. Significant co-expression patterns were used to construct the bone metastasis-specific regulation network. Finally, spatial single-cell RNA-seq and chromatin immunoprecipitation sequence (ChIP-seq) data and multi-omics databases were applied to validate the key scientific hypothesis in the regulation network. Additionally, Connectivity Map (CMap) was utilized to select the potential inhibitors of bone metastasis-specific regulation network in BRCA.

Results: Based on edgeR and WGCNA method, 43 PSRSs were identified. In the bone metastasis-specific regulation network, MAF positively regulated CD248 ($R = 0.435$, $P < 0.001$), and hallmark apical junction was the potential pathway of CD248 ($R = 0.353$, $P < 0.001$). This regulatory pattern was supported by spatial single-cell RNA sequence, ChIP-

seq data and multi-omics online databases. Additionally, alexidine was identified as the possible inhibitor for bone metastasis of BRCA by CMap analysis.

Conclusion: PSRSs played important roles in bone metastasis of BRCA, and the prognostic model based on PSRSs showed good performance. Especially, we proposed that CD248 was the most significant PSRS, which was positively regulated by MAF, influenced bone metastasis *via* apical junction pathway. And this axis might be inhibited by alexidine, which providing a potential treatment strategy for bone metastasis of BRCA.

Keywords: breast invasive carcinoma, bone metastasis, apical junction, MAF, CD248, mRNA stemness index (mRNAsi), weighted gene co-expression network analysis (WGCNA), spatial transcriptome

INTRODUCTION

Breast Invasive Carcinoma (BRCA) was the most common tumor in female, which originated from ducts and acinar epithelium at all levels of the breast, and most patients suffered from malignant epithelial tumor. And the BRCA can be classified into several types according to the state of progesterone receptor (PR), estrogen receptor (ER), and ERBB2 receptor (HER2) in histological stratification, which was applied in clinical practice (1). Estimated by American Cancer Society (ACS), there were 279,100 new cases, and 42,690 new death BRCA patients in 2020 (2). Besides, the five-year survival rate for BRCA in stage I, II, III, and IV were 98, 92, 75, and 27% according to the statistic from 2009 to 2015, respectively (3). Although the five-year survival rate of primary BRCA was high, the five-year survival rate of bone metastasis was only 20%, and patients were trapped in a vicious cycle between osteolytic degeneration and proliferation of cancer cells (4). Besides, the osteolytic lesions like pain in bone, fractures, spinal compression and hypercalcemia led to poor survival quality and death (4).

Machine learning based on high-throughput data played an important role in prognosis of cancer, and new characters defined by algorithm like mRNA stemness index (mRNAsi) and stemness-related gene provided a new way for analysis (5). Recently, analysis on triple-negative breast cancer were launched to explore the key gene related to the stemness and find the target for further therapy (6). Therefore, molecules participated in the bone metastasis required to be explored, and corresponded

biomarkers and possible mechanism needed to be drug for clinical strategy and therapy.

In this study, data of RNA-seq in BRCA were identified by edgeR, and differential expressed genes (DEGs) were performed by machine-learning algorithm to define the mRNAsi. Then, the profiling of DEGs were integrated into weighted gene co-expression network analysis (WGCNA) to classify similar genes into multiply modules and outline the phenotypic characteristics of modules. Besides, the mRNAsi and Hallmark gene sets were qualified to annotate modules. Then, the key module and genes most associated with mRNAsi were selected. Univariate and multivariate Cox regression analysis were applied to access the prognostic value of genes. What is more, based on the Pearson analysis for TF, genes, and Hallmark gene sets, a bone metastasis-specific network was constructed. The scientific hypothesis was determined by the correlation coefficient. Moreover, the CMap analysis was applied to find potential inhibitors for signal axis. Finally, spatial single-cell RNA sequence and chromatin immunoprecipitation sequence (ChIP-seq) data and multi-omics online databases were applied to validate the key scientific hypothesis in the regulation network. The bone metastasis-specific regulation network and inhibitors provided potential treatment strategy for bone metastasis of BRCA.

METHODS

Data Acquisition

Based on the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>), RNA-seq data of 1,080 primary BRCA samples were downloaded, including 1,048 samples without bone metastasis and 32 samples with bone metastasis. And the bone metastasis was diagnosed by imaging examinations like CT or PET-CT. Besides, demographics like age and gender, tumor information like TNM stage and grade, and follow-up data of all patients were also obtained from the TCGA database. Besides, samples without follow-up information were excluded.

Differentially Expressed Genes Identification

The edgeR package was used to screen the RNA-seq data to define DEGs between primary BRCA patients with and without

Abbreviations: BRCA, Breast invasive carcinoma; ACS, American Cancer Society; TF, Transcription Factor; TCGA, The Cancer Genome Atlas; mRNAsi, mRNA stemness index; PSRS, prognostic stemness-related signature; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene-Set Enrichment Analysis; GSVA, Gene Set Variation Analysis; CCL, Cancer Cell Line Encyclopedia; GEPIA, Gene Expression Profiling Interactive Analysis; GTEx, Genotype-Tissue Expression; ROC, The Receiver Operator Characteristic; AUC, Area Under the Curve; FDR, False Discovery Rate; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEG, Differential Expressed Gene; FC, Fold change; mRNAsi, mRNA Expression-Based Stemness Index; GS, gene significance; MS, module significance; ME, module eigengenes; MM, module membership; PCA, principal component analysis; PC, principal component; t-SNE, t-distributed Stochastic Neighbor Embedding; UMAP, Uniform Manifold Approximation and Projection; HE, Haematoxylin and Eosin; MMP9, matrix metalloproteinase 9; EMT, epithelial-mesenchymal transition.

bone metastasis, and the criteria must fit following two points at the same time: the absolute value of log Fold Change (log FC) must more than 1, and the False Discovery Rate (FDR) must less than 0.05. Then, Gene Oncology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were utilized to annotate DEGs.

mRNAsi

The mRNAsi was calculated by the one-class logistic regression (OCLR) machine-learning algorithm (7) based on the RNA-seq data of 1,080 primary BRCA samples.

Weighted Gene Co-Expression Network Analysis

With the aim to holistically analyze DEGs, modules were classified by WGCNA (8) R package (<http://www.r-project.org/>), every single module gathered highly similar DEGs. Then based on RNA-seq profiling, the Pearson correlation analysis was utilized to construct the gene co-expression network. Additionally, the power function was applied to build a weighted adjacency matrix:

$$a_{ij} = |s_{ij}|^{\beta}$$

s_{ij} represented the Pearson correlation between gene i and j , a_{ij} represented the weighted network adjacency between gene i and j . And β equaled to 4 was the soft-threshold parameter set by pickSoft Threshold from WGCNA R package. What's more, the application of soft-threshold parameter of weighted network make it possible to show the continuous variety of co-expression information in [0,1], and it might promote the idea of scale-free co-expression network come true. In addition, the correlation coefficients were utilized to construct hierarchical clustering, and a topological overlap method was performed to classify DEGs with the similar expression patterns into same module. Besides, the capacity of module must more than 20 genes, and when a certain module with less than 20 genes, similar modules will be merged.

Based on the H Collection of Molecular Signatures Database (MSigDB) v7.0 (<https://www.gsea-msigdb.org/gsea/msigdb/genesets.jsp?collection=H>) (9), 50 Hallmark gene sets were qualified by the computational approach named Gene Set Variation Analysis (GSVA). And these Hallmark gene sets were related to biological process and states. Then, with the aim to annotate the specific phenotypic traits for module, 50 Hallmark gene sets and mRNAsi were defined as phenotypes for a co-analysis with modules. Besides, gene significance (GS) coefficients and p values were illustrated the correlation between DEGs and phenotypes. Similarly, module significance (MS) coefficients and p values showed the correlation between modules and phenotypes, and the MS was calculated from the average absolute GS for all genes in every single module. Moreover, the first principle component of module genes showed the gene expression level in the certain module, and module eigengenes (MEs) represented the first principle component. And module membership (MM) represented the correlation between gene to MEs. Next, mRNAsi was the key phenotype to choose the stemness-related module, and the

largest MS with p value less than 0.05 were applied to determine the key module for next exploration. What's more, the Hallmark gene set significantly related to the key module and p value less than 0.05 was selected as the key Hallmark gene set, and the key Hallmark gene set was defined as the key pathway for stemness-related signatures (SRSs). Besides, SRSs with MM >0.300 and GS >0.300 were obtained from the key module for further analysis.

Multivariate Prognosis Model Construction

The univariate and multivariate Cox proportional hazard regression were applied to find prognostic stemness-related signatures (PSRSs). SRSs with p value <0.001 in univariate Cox proportional hazard regression were defined as PSRSs. Then, the LASSO regression analysis was applied to avoid the over-fitness. Then, residual PSRSs were integrated into the multivariate Cox regression model, and risk score for each BRCA patient was calculated by the formula:

$$Risk\ score_m = \beta_1 \times gene_1 + \beta_2 \times gene_2 + \beta_3 \times gene_3 \dots \dots + \beta_n \times gene_n$$

In the formula, "m" represented the number of each patient, "n" represented prognostic PSRS, and "β" represented the coefficient of each prognostic PSRS. Then, patients with BRCA were divided into low- and high-risk groups according to risk score. And the efficiency of risk score of the model was detected by Kaplan-Meier survival analysis. What's more, and accuracy was detected by the Receiver Operating Characteristic Curve (ROC) curve and C-index. Finally, the demographic information, TNM stage and risk score were applied for correction, and univariate and multivariate Cox regression were performed to validate the independent prognostic value.

Potential Signal Axis Identification

Based on the Cistrome database (<http://cistrome.org/>) (10), the list of 318 Translate factors (TFs) were downloaded. And edgeR was utilized to find differential expressed TFs. Then co-expression analysis for TFs and PSRSs by Pearson correlation analysis, and the significant paired TF-PSRSs were selected.

Aim to identify the significantly co-expressed pathway, the absolute quantification of 50 Hallmark gene sets between non-metastasis and metastasis patients was screened by GSVA. And to explore the up- and down-regulated pathways between non-metastasis and metastasis patients, Gene Set Enrichment Analysis (GSEA) was conducted based on the 50 gene sets of Hallmark (11). In addition, the intersection of GSVA, GSEA, and module phenotypic traits was defined as the key pathway. Then, the Pearson correlation analysis was utilized to analysis the interaction between Hallmark gene sets and PSRSs.

Eventually, the network based on TFs, PSRSs and Hallmark gene sets was constructed. And String database (12) was applied to plot a protein-protein interaction (PPI) network. Besides, the criteria for TF-PSRS paired was the absolute value of the correlation coefficient more than 0.400 and p value less than 0.05, for PSRS-Hallmark gene set was the absolute value of correlation coefficient more than 0.300 and p value less than 0.05.

Connectivity Map Analysis

To expand the application of potential signal axis, inhibitors of the signal pathway were selected by the Connectivity Map (build 02) (CMap) (<https://portals.broadinstitute.org/cmap/>) (13). Then, inhibitors for BCRA were identified. Besides, the information like chemical structural formula and biologic function of inhibitor compounds were available from the mechanism of actions (MoA) (<http://clue.io/>) (14). Ultimately, the key inhibitor was found according to the TF, PSRS, and Hallmark gene set.

Assay for Targeting Accessible-Chromatin With High-Throughput Sequencing-Seq and Chromatin Immunoprecipitation Sequence Validation

Assay for Targeting Accessible-Chromatin with high-throughput sequencing (ATAC-seq) and ChIP-seq data were used to validate the regulation mechanism of the network. ATAC-seq and ChIP-seq data were obtained from TCGA GDC (<https://gdc.cancer.gov/about-data/publications/ATACseq-AWG>) and Cistrome database (<http://cistrome.org/>), respectively (10, 15). WashU Epigenome Browser and Gviz package were used to visualize the binding peaks (16, 17).

Spatial Transcriptome Validation

The regulatory relationship between TF and PSRS, PSRS and Hallmark gene set required to be validated the direct mechanism in the molecular experiment. When linked with single-cell RNA sequence data, spatial transcriptome data can validate the cell subtype localization of the key genes (https://support.10xgenomics.com/spatial-gene-expression/datasets/1.1.0/V1_Breast_Cancer_Block_A_Section_1). For quality control, only fit the following standards at the same time can be selected: genes must express in more than 3 single cells, gene counts more than 1, cell transcripts range from 1,500 to 100,000.

To integrate data analysis, the Seurat method was performed (18). Then, the “sctransform” algorithm was utilized for normalization. In addition, to identify variable genes and spatial-specific genes, “vst” and “markvariogram” method were utilized, respectively. Next, the principal component analysis (PCA) was performed based on variable genes (18). In addition, the jackstraw analysis was utilized to select the principal components (PCs), and *p* value must less than 0.05. Then, the further t-distributed Stochastic Neighbor Embedding (t-SNE), Haematoxylin and Eosin (HE) and UMAP (Uniform Manifold Approximation and Projection) were applied to identify the cell sub-cluster based on the PCs (resolution = 0.50) (19). And in sub-cluster, DEGs were filtered when the absolute value of log FC less than 0.5 and FDR more than 0.05. Moreover, the location and expression of DEGs were demonstrated in feature plots and violin plots, respectively. Besides, every cluster was annotated by scMatch (20), singleR (21), and CellMarker (22) databases. Aim to annotate single cells, 50 hallmark gene sets were performed to absolutely quantify the signaling pathway activity in each single cell.

Multidimension Validation

To decrease the inherent defects of analysis in silicon, multiply online databases were utilized to validate the scientific hypothesis

in several aspects. And top five genes in the key pathway selected by GeneCard (<https://www.genecards.org/>) were also validated with TFs and PSRSs.

Gene Expression Profiling Interactive Analysis (GEPIA) (23), Oncomine (24), PROGeneV2 (25), UALCAN (26), Linkedomics (27), SurvExpress (28), cBioportal (29), Genotype-Tissue Expression (GTEx) (30) and UCSC xena (31) validated in gene level based, Cancer Cell Line Encyclopedia (CCLE) (32) validated in cancer cell line level, The human protein atlas (33) validated in tissue level in BRCA patient. Finally, String database (12) was utilized to construct the Protein-Protein Interaction network.

Statistics Analysis

The R software (www.r-project.org; version 3.6.1; Institute for Statistics and Mathematics, Vienna, Austria) was applied in all statistics analysis in our study, and two-sided *p* value <0.05 was determined as statistically significant.

RESULTS

Differentially Expressed Genes Identification

The expression profiling of 1,048 primary BRCA samples without bone metastasis and 32 primary BRCA samples with bone metastasis were obtained from TCGA database, and all patients' demographics information was summarized in **Table 1**. And all analysis processes were illustrated in **Figure 1**.

TABLE 1 | Baseline information of 813 patients diagnosed with breast invasive carcinoma.

Variables	Total Patients (N = 813)
Age, years	
Mean ± SD	57.54 ± 12.66
Median (Range)	58 (26–90)
T stage	
T0	0
T1	222 (27.31%)
T2	477 (58.67%)
T3	88 (10.82%)
T4	26 (3.20%)
N stage	
N0	670 (82.41%)
N1	0
N2	92 (11.32%)
N3	51 (6.27%)
M stage	
M0	798 (98.15%)
M1	15 (1.85%)
Stage	
Stage I	152 (18.70%)
Stage II	468 (57.56%)
Stage III	178 (21.89%)
Stage IV	15 (1.85%)
Metastasis	
Bone metastasis	5 (0.62%)
Other metastasis	10 (1.23%)
No metastasis	798 (98.15%)

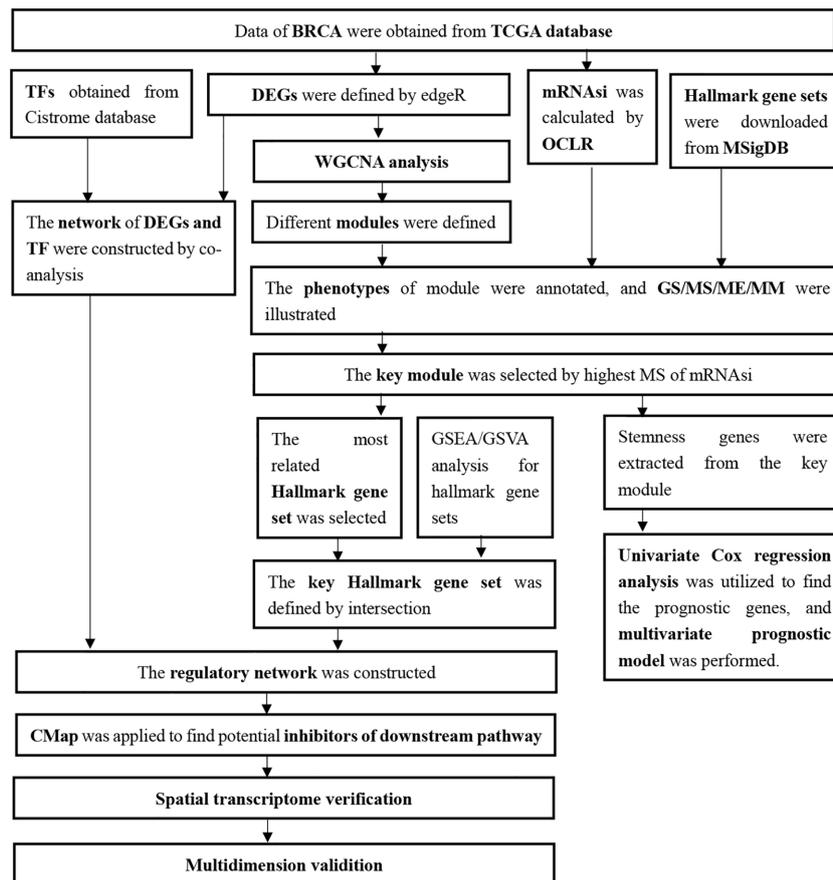


FIGURE 1 | Flow chart of analysis.

RNA-seq data from TCGA database was screened by edgeR to filter DEGs, and the information of BRCA (**Figure 2A**), the heatmap of RNA expression level in BRCA samples (**Figure 2B**) and volcano plots of DEGs and non-DEGs (**Figure 2C**) were launched. And 31 DEGs were founded. Besides, GO (**Figure 2D**) enrichment analysis for DEGs was performed, and cell-cell adhesion *via* plasma-membrane adhesion molecules (BP, GeneRatio = 0.048, $p = 0.001$, count = 21), contractile fiber (CC, GeneRatio = 0.045, $p < 0.001$, count = 20), receptor ligand activity (MF, GeneRatio = 0.066, $p = 0.007$, count = 27) were the most significant GO items. And KEGG (**Figure 2E**) enrichment analysis for DEGs showed Neuroactive ligand-receptor interaction (GeneRatio = 0.125, $p = 0.001$, count = 23) was the most significant KEGG item.

WGCNA

Based on WGCNA package, seven modules were defined (**Figures 3A, B**). Aim to annotate the phenotype of modules, 50 Hallmark gene sets and mRNAsi were co-analysis with

modules in the heatmap plot (**Figure 3C**). Module turquoise (MS = 0.550; $p < 0.001$) was the module most relevant to mRNAsi (MS = 0.670; $p < 0.001$) and 125 SRSs in turquoise were integrated into the further analysis. In addition, three Hallmark gene sets were highly correlated with module turquoise: hallmark apical junction (MS = 0.670; $p < 0.001$), hallmark myogenesis (MS = 0.66; $p < 0.001$), and hallmark IL6-JAK-STAT3 signaling (MS = 0.660; $p < 0.001$) (**Figure 3D**).

Multivariate Prognostic Model Construction

The heatmap (**Figure 4A**) and volcano plot (**Figure 4B**) demonstrated the results of DEG analysis of 125 SRSs. And univariate Cox regression analysis was utilized to find the prognostic genes, and 43 PSRSs were identified. In addition, forest plot (**Figure 4C**) showed CD248 (HR = 1.00004, 95%CI (1.00001-1.00006), $p = 0.007$) was significantly associated with prognosis of BRCA. Then, the PSRSs were integrated into

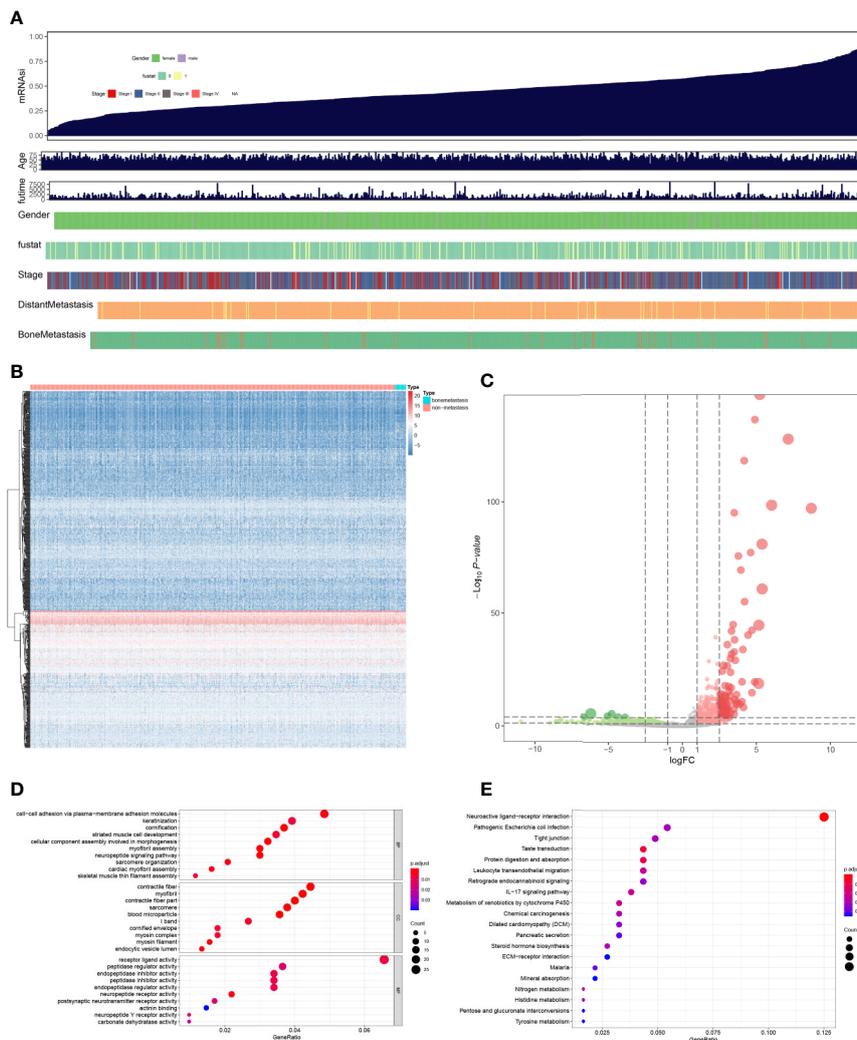


FIGURE 2 | The summary of mRNA-seq (A), heatmap plot (B), volcano plot (C), GO (D), and KEGG (E) enrichment analysis of differential expressed genes. Cell-cell adhesion via plasma-membrane adhesion molecules, contractile fiber, receptor ligand activity were the most significant GO items, and Neuroactive ligand-receptor interaction was the most significant KEGG item.

multivariate Cox regression analysis, and a prognostic model was constructed. The scatter plot (Figure 5A) and risk line plot (Figure 5B) were drawn to illustrate the risk distribution of each patient. In addition, in terms of model diagnosis, area under curve (AUC) of ROC curve was 0.711 (Figure 5C). Based on the median risk score calculated by multivariate model, low- and high-risk groups were accessed by Kaplan-Meier survival analysis (Figure 5D), and the result displayed a significant difference ($p < 0.001$). Finally, the risk score was co-analyzed with age, T stage, N stage, M stage, stage, the univariate (Figure 5E) ($HR = 167.019$, 95%CI (39.677–703.066), $p < 0.001$) and multivariate (Figure 5F) ($HR = 1.050$, 95%CI (1.033–1.067), $p < 0.001$) Cox regression, suggesting that the risk score was an independent prognostic factor.

Potential Signal Axis Identification

The expression levels of Hallmark gene sets were shown in heatmap plot (Figure 6A), and differential expressed gene sets showed in volcano plots (Figure 6B). A total of 47 significant expressed Hallmark gene sets were identified by GSVA (Figure 6C), and five up-regulated and 15 down-regulated Hallmark gene sets were identified by GSEA (Figure 6D). What's more, Hallmark gene sets were co-analyzed with PSRSs by the Pearson correlation analysis.

Based on data of 318 TFs from the Cistrome database, a series of analysis on expression launched, and heatmap plot (Figure 7A) and volcano plot (Figure 7B) were applied. Moreover, 96 TFs were significantly differential expressed. Then, the Pearson correlation analysis was utilized to find the relation between TFs and PSRSs.

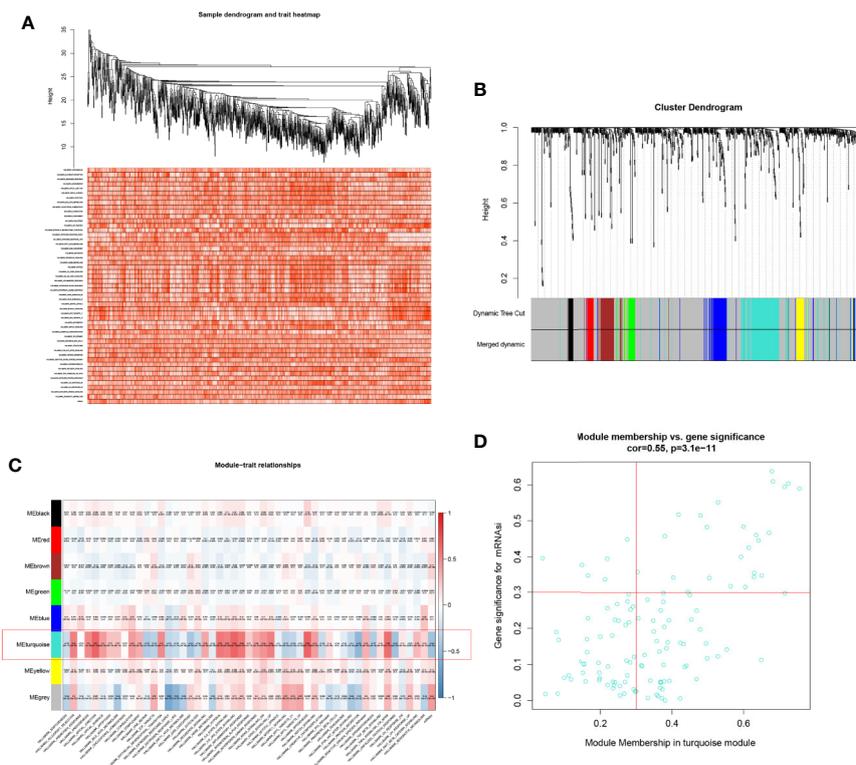


FIGURE 3 | Heatmap of sample (A), Cluster dendrogram of WGCNA (B), co-expression heatmap of modules and phenotypes (C), and the correlation between the mRNAs and the module (D). Module turquoise and hallmark apical junction was the key module and phenotype.

In addition, the intersection between co-analysis Hallmark gene sets in GSEA and significant Hallmark gene sets in GSEA was illustrated in Venn plots (Figure 7C), 22 Hallmark gene sets were found.

Next, the network (Figure 7D) of TFs, PSRSs and Hallmark gene sets were constructed based on the coefficient correlation of the Pearson correlation analysis. Therefore, key TF-PSRS paired was MAF-CD248 (Correlation coefficient = 0.435, $p < 0.001$, positive), and PSRS-Hallmark gene set paired was CD248-apical junction (Correlation coefficient = 0.353, $p < 0.001$) (Figure 7E).

In sum, the scientific hypothesis was defined: MAF positively regulated CD248, promoting apical junction pathway in BRCA, which might play a role in bone metastasis.

Cmap Analysis

To find the latent inhibitor of the bone metastasis-specific regulation network and proposed signal axis, the CMap analysis was utilized, and alexidine (enrichment = 0.6393, $p = 0.046$), clomipramine (enrichment = 0.654, $p = 0.034$), trifluoperazine (enrichment = 0.434, $p = 0.003$), thioridazine (enrichment = 0.337, $p = 0.016$) and valinomycin (enrichment = -0.639, $p = 0.041$) were significant compounds in BRCA

(Figure 8A). Based on the clue database, the detail information of trifluoperazine (Figure 8B), clomipramine (Figure 8C) and thioridazine (Figure 8D) were found, and trifluoperazine was most related to metastasis BRCA according to literature review results.

Spatial Transcriptome and Chromatin Immunoprecipitation Sequence Validation

With the aim to further explore the location of key genes in subtype cell clusters, the profiling of scRNA-seq and spatial transcriptome were co-analyzed. Fourteen clusters were identified in UMAP and t-SNE, and pare-carcinoma, invasive ductal carcinoma and intraductal carcinoma *in situ* were illustrated in HE-stained section (Figure 9A). For validation, the feature and spatial feature plots of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1 demonstrated to show the location in BRCA samples, and they were highly-expressed in the invasive ductal carcinoma tissue (cluster 2, 4, 7, 9, 12) (Figures 9B, C). Besides, a series of analysis based on cell cycle indicated that genes in cluster 4 and 7 of invasive ductal carcinoma and 10 of intraductal carcinoma highly related to phase G2M and S (Figure 9D). In addition, tumor inhibited pathways like apoptosis, p53 pathway and TNF α signaling *via*

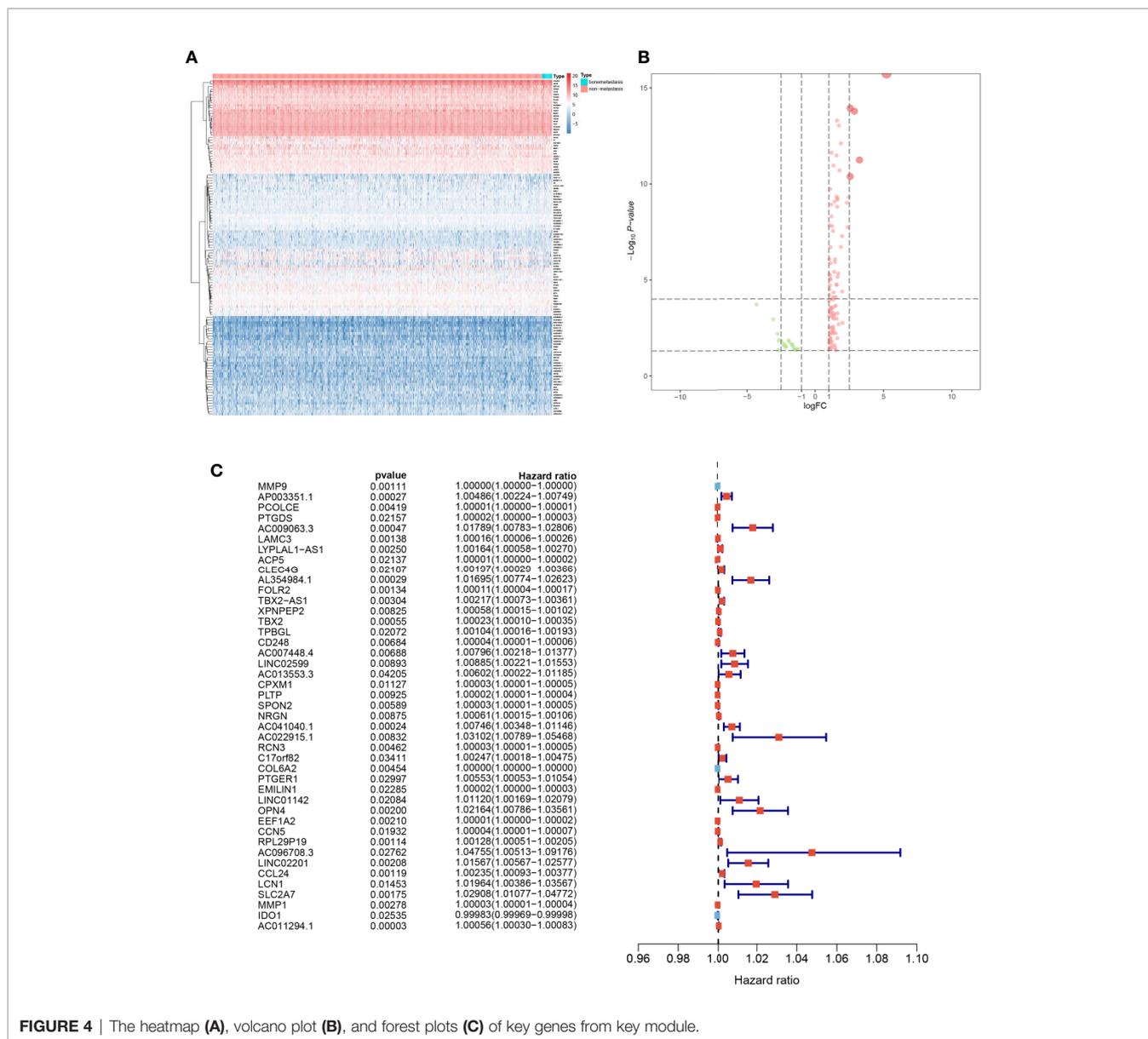


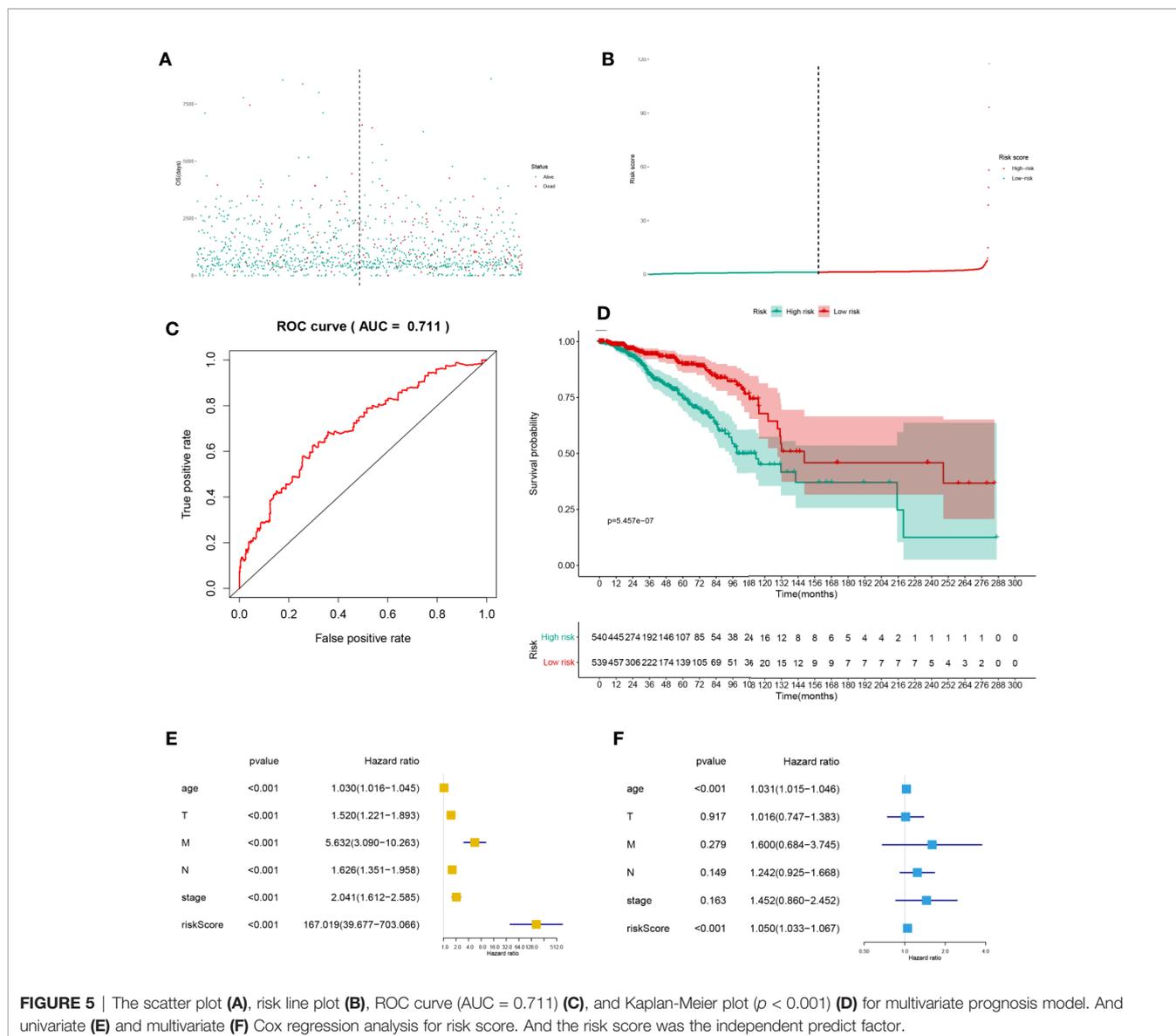
FIGURE 4 | The heatmap (A), volcano plot (B), and forest plots (C) of key genes from key module.

NF- κ B were down-regulated while tumor genesis related pathways G2M checkpoint and E2F targets were up-regulated (Figure 9E). In ChIP-seq analysis, binding peaks were illustrated in CD248 sequence (Figure 10). Moreover, genes in our hypothesis were validated in ATAC-seq (Figure 11), and they were all regulated in BRCA samples. Therefore, patterns of direct transcriptional regulatory between TF-PSRS interaction pairs were identified.

Multidimensional Validation

The correlation of key genes in signal axis based on cBioportal database was summarized in Table S1. And top five genes in apical junction were GJA1, LAMA3, TJP1, LAMC2, and

COL17A1. Several databases were applied to validate the expression level (Table S2) and prognosis value (Table S3) of key genes in hypothesis signal axis. Besides, details were demonstrated in Figure S1-12. MAF showed down-regulated in BRCA, and CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1 showed up-regulated in primary BRCA. Besides, MAF was highly-expressed in the metastasis sample. What's more, MAF (Figure S3C, $p = 0.013$), CD248 (Figure S3B, $p = 0.036$), GJA1 (Figure S3A, $p = 0.005$; Figure S6B, $p = 0.003$), LAMA3 (Figure S3A, $p = 0.008$; Figure S6B, $p = 0.006$), TJP1 (Figure S6B, $p = 0.019$) and LAMC2 (Figure S3B, $p = 0.018$) showed significantly related metastasis; MAF (Figure S3C, $p = 0.002$; S6C, $p = 0.048$), CD248 (Figure S3D, $p = 0.021$; S6C, $p = 0.015$), GJA1 (Figure S3E, $p = 0.001$; S6A, $p < 0.001$), LAMA3



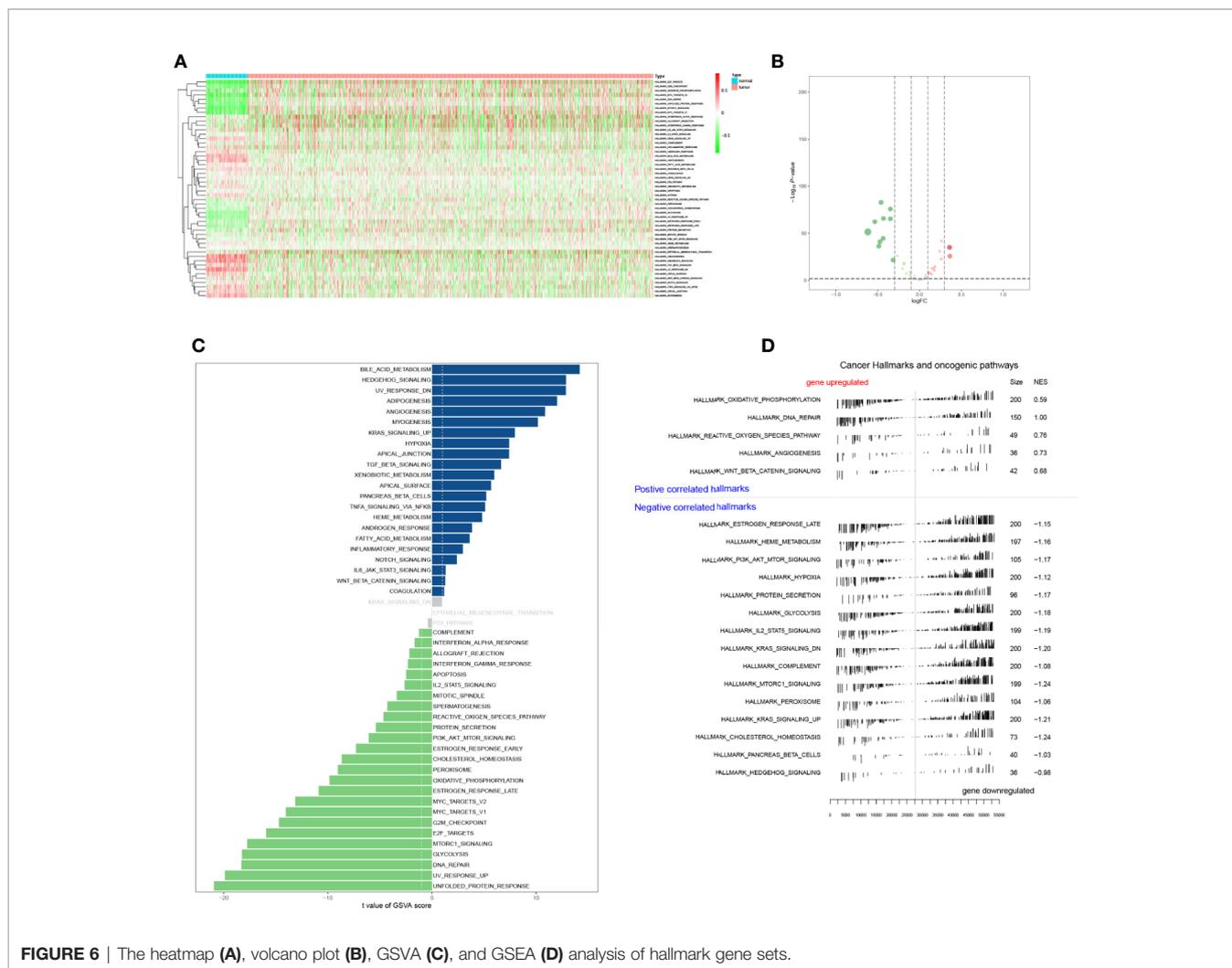
(S6A, $p = 0.047$), TJP1 (Figure S3D, $p = 0.031$; Figure S6A, $p = 0.044$; Figure S7A, $p = 0.027$), LAMC2 (Figure S5F, $p = 0.038$) and COL17A1 (Figure S1R, $p = 0.014$; Figure S3D, $p = 0.007$; Figure S5G, $p = 0.001$; Figure S6A, $p = 0.048$) showed significantly related prognosis; LAMC2 (Figure S7C, $p = 0.047$) and COL17A1 (Figure S7C, $p = 0.001$) were significantly related to progression free.

DISCUSSION

BRCA was a common tumor in female, and patients with bone metastasis suffered from the pain and the risk of fracture and

even death (4). What’s worse, tumor genesis and osteolytic damage could be mutually reinforcing (4). As consequence, the mechanism of bone metastasis must be expounded for early diagnosis and precise therapy.

In the recent study, a total of 813 primary BRCA samples were analyzed. Based on WGCNA method and univariate Cox regression analysis, several modules were annotated by mRNAsi and Hallmark gene sets to find the key module and correspond PSRSs. And the multivariate Cox model was constructed. In addition, multivariate Cox model and risk score were accessed by ROC curve and Kaplan-Meier survival analysis. And the risk score was an independent predict factor. Then, a metastasis-specific regulation network was constructed by the Pearson analysis, and MAF, CD248 and apical junction were significant. Moreover, the regulatory pattern was supported by



spatial single-cell RNA sequence and ChIP-seq data and multi-omics online databases. Based on CMap analysis, trifluoperazine was identified as the possible inhibitor for bone metastasis of BRCA.

MAF was MAF basic leucine zipper transcription factor, which belonged to AP-1 super family, and it regulated the terminal differentiation (34). Besides, MAF was crucial in promoting osteoblast differentiation of bone marrow stromal cell (35). What' more, MAF directly regulated osteoblast-specific promoter Bglap1, and co-regulated the osteoblast differentiation with RUNX2 (35). In addition, Milica Pavlovic et al. found that high-expressed MAF was significant related to bone metastasis instead of visceral metastasis based on Genomic copy number distortion analysis and immunohistochemistry, and MAF was also correlated to overall survival (36). Moreover, MAF played a role in promoting bone metastasis rather than cancer cell proliferation *in vivo* (36), which was consistent with our scientific hypothesis. Additionally, MAF also potentially controlled biological processes like migration, adhesion, and osteoclast differentiation in bone metastasis (36).

CD248 highly-expressed in tumor tissue, especially in BRCA, and it was related to the prognosis of patients (37). And the transcription product of CD248 was endosialin, which expressed on the cell surface of fibroblasts and pericytes in tumor instead of tumor endothelium (38). Further, Carmen Viski et al. found CD248 played a pivotal role by promoting the step of infiltration from primary to circulatory system *via* pericytes in metastasis, which affected on tumor microenvironment (39). Besides, CD248 enhanced the adhesion to the extracellular matrix, and activated the matrix metalloproteinase 9 (MMP9) in tumor metastasis (40). Moreover, CD248 expressed in osteoblast instead of osteoclast, and it has negative effects on osteoblast maturation and ossification (41).

Although none of the study reported the correlation of MAF and CD248, we proposed that MAF positively regulated the transcription of CD248: promoted the function of the transcript in cell adhesion, invasion and migration, and regulated osteoblast function (36, 40, 41).

Apical junction was a structure of apical domain of epithelial cells, and it linked adjacent epithelial cells by tight and adherent

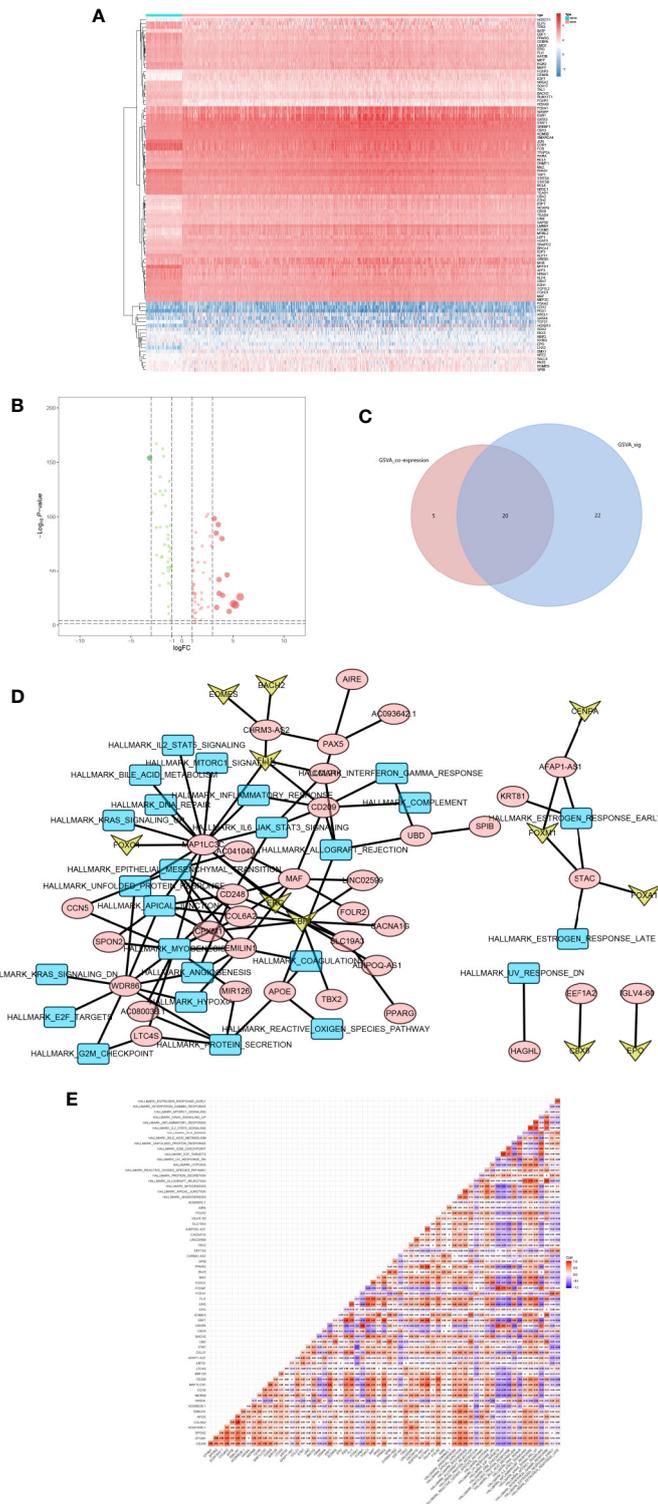


FIGURE 7 | The heatmap (A) and volcano plot (B) of TFs. And the venn plot of hallmark gene sets (C). The network plot of TFs, DEGs and hallmark gene sets (D). And the co-analysis result for TFs, DEGs and hallmark gene sets (E).

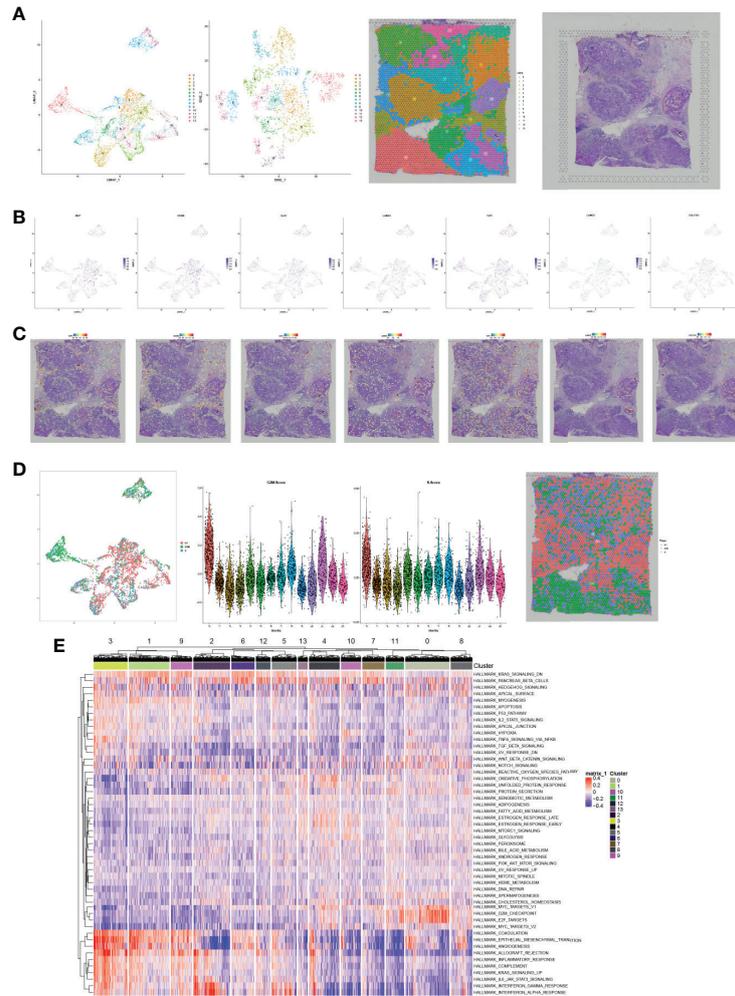


FIGURE 9 | The UMAP, t-SNE plots, and pare-carcinoma, invasive ductal carcinoma and intraductal carcinoma *in situ* were illustrated in HE-stained section **(A)**. The feature and spatial plots of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1, showing that these key genes were highly expressed in the invasive ductal carcinoma tissue **(B, C)**. UMAP plot related to cell cycle, violin plots of phase G2M and S, and phase annotated section **(D)**, and cluster 4 and 7 of invasive ductal carcinoma and 10 of intraductal carcinoma highly related to phase G2M and S. The heatmap of Hallmark gene sets in cell clusters **(E)**.

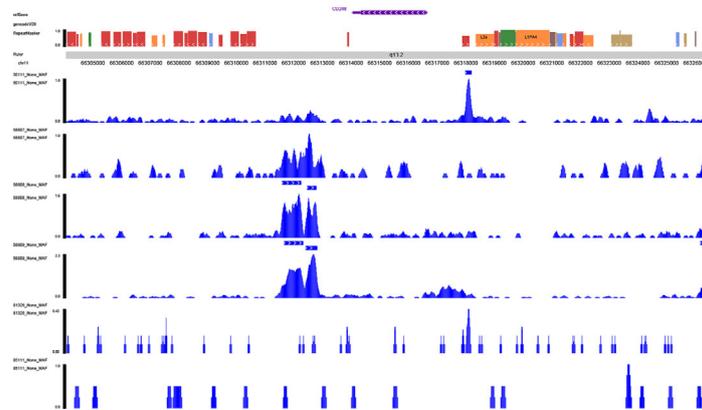


FIGURE 10 | Validation of the transcriptional regulation mechanisms of MAF-CD248 in ChIP-seq data available from Cistrome database.



FIGURE 11 | Validation of seven key genes [MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G)] in ATAC-seq data available from TCGA.

role in bone metastasis. And it could be inhibited by trifluoperazine. Besides, the MAF-CD248-apical junction signal axis was verified by spatial single-cell RNA sequence and ChIP-seq data and multi-omics online databases.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: TCGA program (<https://portal.gdc.cancer.gov>) and 10x genomics support center (<https://support.10xgenomics.com/spatial-gene-expression/datasets>).

ETHICS STATEMENT

The study was approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RH, ZL, JiayZ, ZZ, JiaqZ, ML, SW, SX, YX, XC, JL, WC, BW, PY, DY, and ZH conceived and designed the study. RH, ZL, JiayZ, ZZ, JiaqZ, ML, SW, SX, YX, XC, JL, WC, BW, PY, DY, and ZH collected and/or assembled the data. RH, ZL, JiayZ, ZZ, JiaqZ, ML, SW, SX, YX, XC, JL, WC, BW, PY, DY, and ZH analyzed and interpreted the data. RH, ZL, JiayZ, ZZ, JiaqZ, ML, SW, SX, YX, XC, JL, WC, BW, PY, DY, and ZH wrote the manuscript. RH, ZL, JiayZ, ZZ, Jiaq, ML, SW, SX, YX, XC, JL, WC, BW, PY, DY, and ZH gave the final approval of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The study was supported by the National Natural Science Foundation of China, joint fund cultivation project (Grant No. U1504822); Henan Medical Science and Technology Research Project (No. 201602031); Henan Medical Science and technology Research Plan, Joint Project of the Ministry and the Province, (No. SB201901037); Henan Provincial Department of Science and Technology, Social Development Project (No. 142102310055). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank the The Cancer Genome Atlas (TCGA) team for using their data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.613333/full#supplementary-material>

Supplementary Figure 1 | The expression level of MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G) between tumor and normal. The stage plot of MAF (H), CD248 (I), GJA1 (J), LAMA3 (K), TJP1 (L), LAMC2 (M), and COL17A1 (N) in BRCA. The Kaplan-Meier survival analysis for overall survival in MAF (O), CD248 (P), LAMA3 (Q), and COL17A1 (R), for Disease free survival in MAF (S), CD248 (T), LAMA3 (U), and COL17A1 (V).

Supplementary Figure 2 | The expression level of MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G) between tumor and normal.

Supplementary Figure 3 | The Kaplan-Meier survival analysis for MAF, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA metastasis in NK1 (A); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA metastasis in GSE2990 (B); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA metastasis in GSE11121 (C). The Kaplan-Meier survival analysis for MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA overall survival in GSE19783 (D); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA overall survival in GSE3494 (E).

Supplementary Figure 4 | The expression level of MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G) between tumor and normal.

Supplementary Figure 5 | The Kaplan-Meier survival analysis of MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G) between high- and low-expression groups. The positively (H) and negatively (I) related genes of BRCA.

Supplementary Figure 6 | The Kaplan-Meier survival analysis for MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA overall survival in Van (A); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA metastasis in Van (B); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA overall survival in TCGA (C).

Supplementary Figure 7 | The Kaplan-Meier survival analysis for MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA overall survival (A); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA disease free (B); MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, COL17A1, and integrated genes in BRCA progression free (C).

Supplementary Figure 8 | The expression level of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1 in normal people tissue.

Supplementary Figure 9 | The expression level of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1 in BRCA (A). The PCA plot of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1 (B). The heatmap of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1 (C).

Supplementary Figure 10 | The expression level of MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G) in cancer cell lines.

Supplementary Figure 11 | The expression level of MAF (A), CD248 (B), GJA1 (C), LAMA3 (D), TJP1 (E), LAMC2 (F), and COL17A1 (G) in normal and BRCA tissue.

Supplementary Figure 12 | The protein-protein interaction network of MAF, CD248, GJA1, LAMA3, TJP1, LAMC2, and COL17A1.

REFERENCES

1. Yeo SK, Guan JL. Breast Cancer: Multiple Subtypes within a Tumor? *Trends Cancer* (2017) 3(11):753–60. doi: 10.1016/j.trecan.2017.09.001
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70(1):7–30. doi: 10.3322/caac.21590
3. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, et al. Breast cancer statistics, 2019. *CA Cancer J Clin* (2019) 69(6):438–51. doi: 10.3322/caac.21583
4. Shemanko CS, Cong Y, Forsyth A. What Is Breast in the Bone? *Int J Mol Sci* (2016) 17(10):1764. doi: 10.3390/ijms17101764
5. Bai KH, He SY, Shu LL, Wang WD, Lin SY, Zhang QY, et al. Identification of cancer stem cell characteristics in liver hepatocellular carcinoma by WGCNA analysis of transcriptome stemness index. *Cancer Med* (2020) 9(12):4290–8. doi: 10.1002/cam4.3047
6. Suo HD, Tao Z, Zhang L, Jin ZN, Li XY, Ma W, et al. Coexpression Network Analysis of Genes Related to the Characteristics of Tumor Stemness in Triple-Negative Breast Cancer. *BioMed Res Int* (2020) 2020:7575862. doi: 10.1155/2020/7575862
7. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* (2018) 173(2):338–54 e15. doi: 10.1016/j.cell.2018.03.034
8. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform* (2008) 9:559. doi: 10.1186/1471-2105-9-559
9. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* (2015) 1(6):417–25. doi: 10.1016/j.cels.2015.12.004
10. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res* (2019) 47(D1):D729–D35. doi: 10.1093/nar/gky1094
11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, 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. doi: 10.1073/pnas.0506580102
12. Snel B, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res* (2000) 28(18):3442–4. doi: 10.1093/nar/28.18.3442
13. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* (2017) 45(D1):D362–d8. doi: 10.1093/nar/gkw937
14. Huang RZ, Liu ZQ, Tian TL, Song DW, Yan PH, Yin HB, et al. The construction and analysis of tumor-infiltrating immune cells and ceRNA networks in metastatic adrenal cortical carcinoma. *Biosci Rep* (2020) 40:16. doi: 10.1042/bsr20200049
15. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W, et al. The chromatin accessibility landscape of primary human cancers. *Science (New York NY)* (2018) 362(6413):eaav1898. doi: 10.1126/science.aav1898
16. Li D, Hsu S, Purushotham D, Sears RL, Wang T. WashU Epigenome Browser update 2019. *Nucleic Acids Res* (2019) 47(W1):W158–w65. doi: 10.1093/nar/gkz348
17. Hahne F, Ivanek R. Visualizing Genomic Data Using Gviz and Biocductor. *Methods Mol Biol (Clifton NJ)* (2016) 1418:335–51. doi: 10.1007/978-1-4939-3578-9_16
18. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* (2018) 36(5):411–20. doi: 10.1038/nbt.4096
19. Chung NC, Storey JD. Statistical significance of variables driving systematic variation in high-dimensional data. *Bioinformatics (Oxford England)* (2015) 31(4):545–54. doi: 10.1093/bioinformatics/btu674
20. Hou R, Denisenko E, Forrest ARR. scMatch: a single-cell gene expression profile annotation tool using reference datasets. *Bioinformatics (Oxford England)* (2019) 35(22):4688–95. doi: 10.1093/bioinformatics/btz292
21. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* (2019) 20(2):163–72. doi: 10.1038/s41590-018-0276-y
22. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, et al. CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res* (2019) 47(D1):D721–d8. doi: 10.1093/nar/gky900
23. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* (2017) 14(10):979–82. doi: 10.1038/nmeth.4402
24. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* (2004) 6(1):1–6. doi: 10.1016/s1476-5586(04)80047-2
25. Goswami CP, Nakshatri H. PROGeneV2: enhancements on the existing database. *BMC Cancer* (2014) 14:970. doi: 10.1186/1471-2407-14-970
26. Chandrashekar DS, Bachel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* (2017) 19(8):649–58. doi: 10.1016/j.neo.2017.05.002
27. Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res* (2018) 46(D1):D956–D63. doi: 10.1093/nar/gkx1090
28. Aguirre-Gamboa R, Gomez-Rueda H, Martinez-Ledesma E, Martinez-Torteya A, Chacolla-Huaringa R, Rodriguez-Barrionto A, et al. SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis. *PLoS One* (2013) 8(9):e74250. doi: 10.1371/journal.pone.0074250
29. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* (2012) 2(5):401–4. doi: 10.1158/2159-8290.cd-12-0095
30. Consortium G. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (New York NY)* (2015) 348(6235):648–60. doi: 10.1126/science.1262110
31. Goldman M, Craft B, Swatloski T, Cline M, Morozova O, Diekhans M, et al. The UCSC Cancer Genomics Browser: update 2015. *Nucleic Acids Res* (2015) 43(Database issue):D812–7. doi: 10.1093/nar/gku1073
32. Ghandi M, Huang FW, Jane-Valbuena J, Kryukov GV, Lo CC, McDonald ER3rd, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* (2019) 569(7757):503–8. doi: 10.1038/s41586-019-1186-3
33. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science (New York NY)* (2015) 347(6220):1260419. doi: 10.1126/science.1260419
34. Labott AT, Lopez-Pajares V. Epidermal differentiation gene regulatory networks controlled by MAF and MAFB. *Cell Cycle* (2016) 15(11):1405–9. doi: 10.1080/15384101.2016.1172148
35. Nishikawa K, Nakashima T, Takeda S, Isogai M, Hamada M, Kimura A, et al. Maf promotes osteoblast differentiation in mice by mediating the age-related switch in mesenchymal cell differentiation. *J Clin Invest* (2010) 120(10):3455–65. doi: 10.1172/JCI42528
36. Pavlovic M, Arnal-Estape A, Rojo F, Bellmunt A, Tarragona M, Guio M, et al. Enhanced MAF Oncogene Expression and Breast Cancer Bone Metastasis. *J Natl Cancer Inst* (2015) 107(12):djv256. doi: 10.1093/jnci/djv256
37. Davies G, Cunnick GH, Mansel RE, Mason MD, Jiang WG. Levels of expression of endothelial markers specific to tumour-associated endothelial cells and their correlation with prognosis in patients with breast cancer. *Clin Exp Metastasis* (2004) 21(1):31–7. doi: 10.1023/b:clin.0000017168.83616.d0
38. MacFadyen JR, Haworth O, Roberston D, Hardie D, Webster MT, Morris HR, et al. Endosialin (TEM1, CD248) is a marker of stromal fibroblasts and is not selectively expressed on tumour endothelium. *FEBS Lett* (2005) 579(12):2569–75. doi: 10.1016/j.febslet.2005.03.071
39. Viski C, Konig C, Kijewska M, Mogler C, Isacke CM, Augustin HG. Endosialin-Expressing Pericytes Promote Metastatic Dissemination. *Cancer Res* (2016) 76(18):5313–25. doi: 10.1158/0008-5472.CAN-16-0932
40. Tomkowicz B, Rybinski K, Foley B, Ebel W, Kline B, Routhier E, et al. Interaction of endosialin/TEM1 with extracellular matrix proteins mediates cell adhesion and migration. *Proc Natl Acad Sci U S A* (2007) 104(46):17965–70. doi: 10.1073/pnas.0705647104
41. Naylor AJ, Azzam E, Smith S, Croft A, Poyser C, Duffield JS, et al. The mesenchymal stem cell marker CD248 (endosialin) is a negative regulator of bone formation in mice. *Arthritis Rheum* (2012) 64(10):3334–43. doi: 10.1002/art.34556
42. Zhu MJ, Sun X, Du M. AMPK in regulation of apical junctions and barrier function of intestinal epithelium. *Tissue Barriers* (2018) 6(2):1–13. doi: 10.1080/21688370.2018.1487249
43. Silvestrini VC, Lanfredi GP, Masson AP, Poersch A, Ferreira GA, Thome CH, et al. A proteomics outlook towards the elucidation of epithelial-mesenchymal

- transition molecular events. *Mol Omics* (2019) 15(5):316–30. doi: 10.1039/c9mo00095j
44. Hollestelle A, Peeters JK, Smid M, Timmermans M, Verhoog LC, Westenend PJ, et al. Loss of E-cadherin is not a necessity for epithelial to mesenchymal transition in human breast cancer. *Breast Cancer Res Treat* (2013) 138(1):47–57. doi: 10.1007/s10549-013-2415-3
45. Gonzalez-Mariscal L, Miranda J, Gallego-Gutierrez H, Cano-Cortina M, Amaya E. Relationship between apical junction proteins, gene expression and cancer. *Biochim Biophys Acta Biomembr* (2020) 1862(9):183278. doi: 10.1016/j.bbamem.2020.183278
46. Wang H, Yu C, Gao X, Welte T, Muscarella AM, Tian L, et al. The osteogenic niche promotes early-stage bone colonization of disseminated breast cancer cells. *Cancer Cell* (2015) 27(2):193–210. doi: 10.1016/j.ccell.2014.11.017
47. Finlayson AE, Freeman KW. A cell motility screen reveals role for MARCKS-related protein in adherens junction formation and tumorigenesis. *PLoS One* (2009) 4(11):e7833. doi: 10.1371/journal.pone.0007833
48. Liu S, Fan Y, Chen A, Jalali A, Minami K, Ogawa K, et al. Osteocyte-Driven Downregulation of Snail Restrains Effects of Drd2 Inhibitors on Mammary Tumor Cells. *Cancer Res* (2018) 78(14):3865–76. doi: 10.1158/0008-5472.CAN-18-0056

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Huang, Li, Zhang, Zeng, Zhang, Li, Wang, Xian, Xue, Chen, Li, Cheng, Wang, Yan, Yang and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.