Integration of single-cell and bulk transcriptome analyses unravels a macrophage-based gene signature for prognostication and treatment in triple-negative breast cancer

Yuan Huang (✉ huangyuan@zjcc.org.cn)
Zhejiang Cancer Hospital

Yuan Yu (✉ yuyuan20191219@163.com)
Zhejiang Cancer Hospital

Huangping Zhang (✉ zhanghuanpingnbu@163.com)
Zhejiang Cancer Hospital

Ziwen Zhang (✉ zhangziwen88@126.com)
Zhejiang Cancer Hospital

Yabing Zheng (✉ zhengyb@zjcc.org.cn)
Zhejiang Cancer Hospital

Xiaojia Wang (✉ wxiaojia0803@163.com)
Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences

Xiaowei Wang (✉ wangxiaowei1983@126.com)
Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University

Research Article

Keywords:

DOI: https://doi.org/

Additional Declarations: No competing interests reported.

Integration of single-cell and bulk transcriptome analyses unravels a macrophage-based gene signature for prognostication and treatment in triple-negative breast cancer

Authors:

Yuan Huang[1,2,3], Yuan Yu[1,3], Huangping Zhang[1,3], Ziwen Zhang[1,3], Yabing Zheng[1,3,*], Xiaojia Wang[1,3,*], Xiaowei Wang[4,*]

1. Department of Breast Medical Oncology, Zhejiang Cancer Hospital, Hangzhou 310022, Zhejiang, China.

2. Key Laboratory of Head & Neck Cancer Translational Research of Zhejiang Province, Zhejiang Cancer Hospital, Hangzhou 310022, Zhejiang, China.

3. Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou 310018, Zhejiang, China.

4. Department of Colorectal Surgery, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou 310020, Zhejiang, China.

*** Corresponding author:**

Xiaowei Wang

E-mail: wangxiaowei1983@126.com

Xiaojia Wang

E-mail: wxiaojia0803@163.com

Yabing Zheng

E-mail: zhengyb@zjcc.org.cn

Abstract:

Objective: As a dominant component within the tumor microenvironment, macrophages exert an essential role in almost all aspects of triple-negative breast cancer (TNBC). This work explored macrophage-associated signature genes for prognostication and treatment for TNBC.

Methods: Single-cell (GSE180286) and bulk transcriptome profiles (TCGA-TNBC, GSE96058 and GSE45255) were analyzed by executing multiple computational approaches. The expression of signature genes was verified in breast cancer cells (MDA-MB-231, and MCF-7) and mammary epithelial cells (MCF10A) through RT-qPCR and western blot. After RNA interference or overexpression, wound scratch assay was implemented.

Results: A single-cell map of the microenvironment of non-TNBC and TNBC was depicted. Both at the single-cell and bulk levels, macrophages exhibited the higher abundance in TNBC versus non-TNBC. A macrophage-based gene signature was built, containing CTSD, CTSL, ELK4, HSPA8, and XRCC4. High risk score was predictive of worse prognostic outcomes. Based upon external

validation, the signature could reliably predict patient prognosis, notably one-year survival. High-risk patients were more responsive to immunotherapy. The aberrant expression of CTSD, CTSL, ELK4, HSPA8, and XRCC4 was proven in breast cancer cells and mammary epithelial cells. Knockdown of XRCC4 attenuated migrative abilities of MDA-MB-231, MCF-7, and MCF10A cells, with opposite findings for overexpressed CTSD, CTSL, and HSPA8.

Conclusions: Altogether, a novel macrophage-based gene signature was proposed for estimating survival outcomes and treatment response in TNBC. The aberrant expression of the signature genes contributed to tumor aggressiveness. Our findings exert a positive impact on future clinical research involving macrophages in TNBC.

Key words: triple-negative breast cancer; macrophages; signature; prognosis; immunotherapy; migration

Background:

Triple-negative breast cancer (TNBC) remains the most lethal form of breast cancer, occupying 15%~20% of all breast malignancies [1]. TNBC cells are extremely aggressive and lack of hormones and growth factor receptors [2]. Because of the absent or low expression of estrogen receptor, human epidermal growth factor

receptor 2, and progesterone receptor, TNBC presents the resistance to hormones and endocrine treatment [3]. In comparison to other breast cancer forms, TNBC remains the most challenging owing to more extensive heterogeneity, higher risk of distant metastases and recurrence, and the insufficiency of validated treatment targets [4]. At current, chemotherapy is utilized as the major method against TNBC [5]. Following the development of immunotherapy in solid tumors and verification of the immunogenicity in TNBC, immunotherapy has attracted increasing attention [6-8]. For improving the survival outcomes of TNBC populations, predictive biomarkers for immunotherapy are challenging [9]. Moreover, due to poor therapy response, novel therapy targets and predictive biomarkers for prognosis are urgently needed for TNBC.

The accumulation of myeloid cells, especially macrophages, are major components within the tumor microenvironment of TNBC [10]. The regulatory mechanisms by which macrophages impact nearly all aspects of TNBC have gained the widespread exploration. For instance, HLF modulates ferroptosis, progression and chemoresistance of TNBC via activation of cancer cell-macrophage communication [11]. Chemotherapy in combination with macrophage suppression induce the abundance of T cells and B cells

and durable regression in TNBC [12]. OTUD5-induced deubiquitination of YAP in macrophages contributes to M2 phenotype and favors TNBC development [13]. So far, limited macrophage-associated prognostic models have been conducted for TNBC patients, none of which have been applied in clinical practice [14-16]. In this work, through integration of single-cell with bulk transcriptome data, a novel macrophage-relevant gene signature was established for estimating prognostic outcomes and therapy response in TNBC. Additionally, in vitro experiments were utilized to prove the aberrant expression of the signature genes as well as their impact on tumor aggressiveness.

Materials and methods:

Single-cell and bulk transcriptome data acquisition

Through the Gene Expression Omnibus (GEO), raw single-cell RNA sequencing (scRNA-seq) data from four primary TNBC specimens were acquired from the GSE180286 dataset [17]. Three TNBC cohorts with bulk transcriptome profiling and clinical features were obtained from The Cancer Genome Atlas (TCGA) (n=115), GSE96058 (n=3409) [18] and GSE45255 (n=95) [19].

Quality control and preprocessing of scRNA-seq

By executing DropletUtils toolkit [20], empty droplets were

distinguished from all cells, followed by removal. Based upon Scater toolkit [21], cells with proportion of mitochondrial genes >10% and proportion of ribosome genes <10% were further removed. The filtered scRNA-seq data were normalized with Seurat toolkit [22].

Principal component analysis (PCA), cell clustering and annotation

The top 2000 highly variable genes were screened utilizing Seurat toolkit, and expression profiling was linearly scaled, followed by PCA. Next, principal components (PCs) with large standard deviation were chosen, which were subsequently used for cell clustering. Uniform manifold approximation and projection (UMAP) was then implemented [23]. Marker genes in each cell cluster were determined based upon the criteria of average log2 fold change (FC) ≥0.1, cell population expression ratio ≤0.25, and adjusted p≤0.05.

Immune infiltration analysis

CIBERSORTX [24] was adopted make labels based upon the identification results of single cells as reference expression matrix. In accordance with bulk expression matrix, the proportion of identified cells in each sample was computed.

Cell-cell communication

Through executing cellchat package [25], cell-cell interactions were evaluated by use of ligand-receptor pairs. The cell-cell

communication networks were visualized through Cytoscape software [26].

Functional enrichment analysis

Gene Ontology (GO) or Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were executed by use of clusterProfiler package [27]. KEGG pathways were visualized via pathview web [28]. Gene set enrichment analysis (GSEA) [29] was adopted for determining gene sets with significant difference between groups.

Differential expression analysis

SCENIC computational approach [30] was utilized for guiding the identification of transcription factors. Differentially expressed transcription factors were screened between TNBC and non-TNBC macrophages via limma method under the criteria of $p \leq 0.05$ and $|t| \geq 2$ [31]. In addition, differentially expressed genes (DEGs) between the groups were selected under the threshold of $|\log2FC| \geq 0.585$ and $q \leq 0.05$.

Least absolute shrinkage and selection operator (LASSO) analysis

Differentially expressed transcription factors and DEGs related to TNBC macrophages were selected for univariate cox regression analysis via survival package. Genes with $p \leq 0.05$ were included for

LASSO. TCGA TNBC samples were randomized into training or test cohort. By executing glmnet [32], signature genes were selected. The risk score was computed based upon regression coefficients combined with expression of the signature genes. Low- or high-risk patients were defined under the median risk score. The LASSO model was externally verified in the GSE96058 and GSE45255 cohorts.

Genetic mutation evaluation

Somatic mutation data and cancer-testis antigen (CTA) number information of TNBC samples were acquired from the TCGA dataset. Through implementing maftools package, somatic mutation was evaluated and visualized [33].

Treatment response estimation

T-cell inflamed score [34], TIDE score [35] as well as expression of immune checkpoint molecules [34] were separately computed for reflecting the response to immunotherapy. Based upon the GDSC2 database, IC50 value of drugs was estimated to infer drug response by use of oncoPredict package [36].

Cell culture

Normal human mammary epithelial cells (MCF10A) as well as human breast cancer cells (MDA-MB-231, and MCF-7) from the Cell Bank of Type Culture Collection of the Chinese Academy of Sciences

(China) were cultivated in Dulbecco's Modified Eagle Medium (Gibco, USA) with supplementation of 10% fetal bovine serum (Gibco) as well as 1% penicillin–streptomycin in a 5% CO_2 atmosphere at 37 °C.

Real-time quantitative PCR (RT-qPCR)

Total RNA isolation was conducted by use of RNAiso Plus reagent (Takara, China), with complementary DNA synthesis via HiScript III RT SuperMix reagent (Vazyme, China). The primers used were as follows: CTSL, 5'-CTTTTGCCTGGGAATTGCCTC-3' (forward primer), 5'-CATCGCCTTCCACTTGGTC-3' (reverse primer); CTSD, 5'-TGCTCAAGAACTACATGGACGC-3' (forward primer), 5'-CGAAGACGACTGTGAAGCACT-3' (reverse primer); ELK4, 5'-TGGACCTCTAATGATGGGCAG-3' (forward primer), 5'-AGGCTTGTTCTTGCGAATCCC -3' (reverse primer); XRCC4, 5'-ATGTTGGTGAACTGAGAAAAGCA-3' (forward primer), 5'-GCAATGGTGTCCAAGCAATAAC-3' (reverse primer); HSPA8, 5'-ACCTACTCTTGTGTGGGTGTT-3' (forward primer), 5'-GACATAGCTTGGAGTGGTTCG-3' (reverse primer); GAPDH, 5'-ACAACTTTGGTATCGTGGAAGG-3' (forward primer), 5'-GCCATCACGCCACAGTTTC-3' (reverse primer). RT-qPCR was conducted via ChamQ Universal SYBR qPCR Master Mix (Vazyme). The relative mRNA level was computed with $2^{-\Delta\Delta Ct}$.

Western blot

All protein extraction was executed through RIPA buffer (Cell Signaling Technology, USA), followed by protein quantification via BCA reagent (Cell Signaling Technology). Protein was separated through SDS-PAGE, with subsequent transference onto PVDF membrane (Millipore, Germany). Following blockade in 5% BCA (Yeasen, China) and incubation with specific antibody against CTSL (1/2000; ab200738), CTSD (1/2000; ab75852), XRCC4 (1/1000; ab213729), HSPA8 (1/500; ab51052) or GAPDH (1/2500; ab9485). The bands were developed via enhanced chemiluminescence detection kit (Yeasen).

Transfection

For RNA interference, the transfection of small interfering RNAs (siRNAs) against XRCC4 (si-XRCC4) and negative control siRNA (si-NC) (Invitrogen, USA) into cells was carried out by use of Lipofectamine 2000 transfection reagent (Invitrogen). For gene overexpression, the CTSL, CTSD or HSPA8 overexpressing plasmids named OE-CTSL, OE-CTSD or OE-HSPA8 were transfected into cells.

Wound scratch assay

Cells were planted onto a 6-well plate and grown until confluent, and the monolayer cells were scraped in a straight line utilizing a 10 μL pipette tip. Next, the plate was washed by PBS for removing detached cells. Photographs were acquired at 0, or 24 h after

scratching under an Olympus IX71 optical microscope.

Statistical analysis

All the analyses were executed by use of R software (version 4.0.3) or GraphPad Prism (version 9.0.1). Difference between two groups was assessed with Student's t test or one-way analysis of variance. Correlation analyses were conducted via Pearson test or Spearman test. Survival curves of overall survival (OS) or disease-free survival (DFS) were visualized via Kaplan–Meier approach, with log-rank test. Receiver operator characteristic curves (ROCs) were plotted via pROC package. $P \leq 0.05$ was considered statistically significant.

Results:

Single-cell and bulk transcriptome analyses unravel cellular heterogeneity in TNBC

This study reconstructed a single-cell landscape of TNBC based upon scRNA-seq data from four primary TNBC specimens. Firstly, single cells with empty droplets or low quality were of removal, with 2599 / 3267 cells in GSM5457199 sample, 3872 / 4161 cells in GSM5457205 sample, 3755 / 4064 cells in GSM5457208 sample, 6233 / 7521 cells in GSM5457211 sample being retained in our analysis (**Supplementary figure 1A-L**). Next, the retained scRNA-

seq data were scaled based upon PCA, with PC=9 (**Supplementary figure 2A-D**). By use of UMAP approach, the selected single cells were clustered into 14 clusters, with remarkable cellular heterogeneity between TNBC and non-TNBC (**Supplementary figure 3A-C**). Also, marker genes in each cell cluster were determined (**Supplementary figure 3D, E**). In combination with the known marker genes of cell types, nine cell populations were classified, composed of B cells (n=698), dendritic cells (n=387), endothelial cells (n=495), epithelial cells (n=9984), fibroblasts (n=350), macrophages (n=732), monocytes (n=32), plasmablasts (n=1412), and T cells (n=2369) (**Figure 1A**). The marker genes were specifically expressed in the corresponding cell populations: MS4A1 for B cells, CD1C, and FCER1A for dendritic cells, PECAM1, VWF, CDH5, SELE, and CD34 for endothelial cells, EPCAM, CDH1, and KRT18 for epithelial cells, COL1A1 and PDGFRB for fibroblasts, APOE, CD68, MRC1, MSR1, and CXCL2 for macrophages, FCN1, LILRA5, and S100A8 for monocytes, JCHAIN for plasmablasts, and CD3D, CD3E, CD3G, and CD2 for T cells (**Figure 1B**). These cell populations were notably different between TNBC and non-TNBC, with the higher cell ratio of B cells, dendritic cells, fibroblasts, macrophages, plasmablasts, and T cells, and the lower cell ratio of endothelial cells and epithelial cells in TNBC versus non-TNBC

(**Figure 1C**). We also determined novel marker genes of each cell population (**Figure 1D**). We also gathered bulk transcriptome profiling of TNBC specimens from the TCGA dataset. By executing CIBERSORTx, the reference matrix of cell markers was established based upon scRNA-seq results, and the relative cell ratios of cell populations were estimated in bulk tissues (**Figure 1E**). Consistent with scRNA-seq findings, the higher cell ratio of macrophages was found in bulk TNBC than non-TNBC tissues (**Figure 1F, G**). Thus, macrophages were active in the TNBC microenvironment.

Cell-cell interactions in the TNBC and non-TNBC microenvironment

Next, we evaluated cell-cell interactions based upon ligand-receptor pairs in non-TNBC and TNBC, respectively. In comparison to non-TNBC, more active cell-cell interactions were investigated in TNBC, especially macrophages with other cell populations (**Figure 2A, B**).

Transcriptional activity of macrophages in TNBC

Totally, 104 transcription factors presented differential expression in TNBC macrophages in comparison to non-TNBC macrophages (**Figure 2C; Supplementary table 1**), which might transcriptionally regulate macrophage activity in TNBC. In addition, 110 DEGs were determined between TNBC and non-TNBC

macrophages (**Figure 2D, E**), of which 63 were up-regulated (**Supplementary table 2**) and 47 were down-regulated (**Supplementary table 3**) in TNBC macrophages versus non-TNBC macrophages. The DEGs were notably in relation to signaling transduction, cell surface receptor signaling pathway, response to cytokine, cellular response to cytokine stimulus, etc. (**Figure 2F**). In addition, immune-related pathways were notably enriched by the DEGs, especially antigen processing and presentation (**Figure 2G, H**). Altogether, the functional enrichment analysis unraveled the significance of the DEGs in modulating transcriptional activity of macrophages in the microenvironmental niche.

Construction of a novel macrophage-relevant prognostic signature for TNBC

The differentially expressed transcription factors and DEGs related to TNBC macrophages were included for univariate-cox regression analysis. Consequently, eight genes exhibited significant associations with TNBC prognosis ($p \leq 0.05$), composed of C12orf60, CTSD, CTSL, ELK4, FCGR2A, FOLR2, HSPA8, and XRCC4. The genes were utilized for construction of a LASSO model. We randomized TCGA TNBC samples into training or test cohort. In the training cohort, LASSO analysis was executed for selecting signature genes with regression coefficient $\neq 0$. Under the lambda

minimum = 0.0267. As a result, five signature genes were eventually determined, comprising CTSD, CTSL, ELK4, HSPA8, and XRCC4 (**Figure 3A, B**). The macrophage-based prognostic signature was constructed based upon the formula: risk score = 0.859575100676907 * CTSD expression + 0.0210700891980921 * CTSL expression + (-0.644138418956012) * ELK4 expression + 0.307340530719732 * HSPA8 expression + 1.31660312733179 * XRCC4 expression (**Figure 3C, D**). Patients with risk score > median risk score were defined as high risk, while those with risk score ≤ median risk score were defined as low risk (**Figure 3E**). High-risk group owned more dead or recurred/progressed status versus low-risk group (**Figure 3F, G**). Survival analysis demonstrated the significantly shorter OS time for high-risk patients in the training cohort (**Figure 3H**). Such survival difference was proven in the test and entire cohorts (**Figure 3I, J**). In addition, ROCs were plotted for appraising the predictive efficacy of the model. In the training (**Figure 3K**), test (**Figure 3L**), and entire (**Figure 3M**) cohorts, the model was excellently predictive of patients' one-year survival (AUC>0.9).

Associations of the macrophage-based prognostic model with more advanced status

Further analysis was conducted for evaluation of the correlations

between the macrophage-based prognostic model and TNBC clinicopathological traits. TCGA patients with more advanced T, N, M stages, and pathological stages presented the notably higher risk score (**Figure 4A-D**). In addition, the risk score displayed a negative association with tumor purity (**Figure 4E**). Both in the GSE96058 and GSE45255 cohorts, patients with more advanced histological grades exhibited the prominently higher risk score (**Figure 4F, G**). Altogether, the macrophage-based prognostic model was in relation to more advanced status of TNBC patients.

External verification of the macrophage-based prognostic model

The GSE96058 and GSE45255 cohorts were adopted for independently proving the efficacy of the macrophage-relevant signature in patient survival. In the GSE96058 cohort, it was proven that high-risk cases owned worse OS outcomes versus low-risk cases (**Figure 4H**). Also, the model enabled to accurately predict one-year survival (**Figure 4I**). In the GSE45255 cohort, shorter DFS time was investigated in high-risk cases, with the reliable efficacy in DFS prediction (**Figure 4J, K**). These findings proved the generalizability of the macrophage-based prognostic model.

Heterogeneous somatic mutations between low- and high-risk TNBC patients

The study quantified TMB score across TNBC samples, with the median TMB of 1.34/MB (**Figure 5A**). Overall, low-risk samples occurred relatively higher TMB score versus high-risk samples (**Figure 5B**). TP53 had the highest mutated frequency both in low- and high-risk samples (**Figure 5C**). OBSCN, UTP20, and KMT2D had significantly higher mutated frequency in high- versus low-risk cases, while FLG occurred significantly lower mutated frequency in low- versus high-risk cases (**Figure 5D, E**). In addition, more frequent co-occurred mutations were investigated in high- than low-risk samples (**Figure 5F, G**).

Molecular mechanisms underlying the macrophage-based prognostic model

For biological processes, mitotic spindle assembly checkpoint, and negative regulation of mitotic metaphase anaphase transition and mitotic sister chromatid separation were notably enriched in high-risk samples, while positive regulation of antigen receptor-mediated signaling pathway, and antigen processing and presentation of endogenous antigen, and positive T cell selection were prominently enriched in low-risk samples (**Figure 6A**). For cellular components, preribosome, large subunit precursor, and preribosome were notably enriched in high-risk group, with the significant enrichment of immunoglobulin complex and T cell

receptor complex in low-risk group (**Figure 6B**). For molecular functions, 3’-5’ DNA helicase activity, single-stranded DNA helicase activity, and WNT-activated receptor activity were enriched in high-risk cases, with the prominent enrichment of peptide antigen binding, C-C chemokine binding, and immunoglobulin receptor binding in low-risk cases (**Figure 6C**). Moreover, for KEGG pathways, DNA replication, RNA polymerase, and mannose type O-glycan biosynthesis presented the notable enrichment in high-risk group, with the notable enrichment of type I diabetes mellitus, allograft rejection, and graft-versus-host disease in low-risk group (**Figure 6D**).

High-risk patients with higher response to immunotherapy

High-risk group exhibited the higher T-cell inflamed score and lower TIDE score in comparison to low-risk group (**Figure 7A, B**). In addition, the higher ratios of responders to immunotherapy were investigated in high- versus low-risk group (**Figure 7C**). In **Figure 7D**, CTA number was remarkably higher in low- than high-risk group. Most immune checkpoints comprising CD80, CD86, IDO1, LAG3, LAIR1, PDCD1, HAVCR2, and LGALS3 presented the notably higher expression in high- versus low-risk group (**Figure 7E**). Above findings proved that high-risk patients might respond to immunotherapy.

Heterogeneous sensitivity to drugs between low- and high-risk TNBC patients

Low-risk samples presented the notably lower IC50 of BI-2536, and UMI-77, indicative of higher sensitivity to the drugs (**Figure 7F**). Meanwhile, high-risk individuals presented stronger sensitivity to 5-Fluorouracil, AZD2014, GSK2606414, AZD1332, Temozolomide, Oxaliplatin, Epirubicin, Taselisib, Entospletinib, Camptothecin, Dabrafenib, BMS-536924, PF-4708671, Topotecan, AZD8055, Rapamycin, Mitoxantrone, Dactolisib, Vinorelbine, Pictilisib, Foretinib, Ribociclib, GNE-317, Dactinomycin, Buparlisib, Teniposide, Irinotecan, JAK1_8709, Alisertib, AZD5363, AZ960, Palbociclib, LGK974, Gemcitabine, VX-11e, Uprosertib, KU-55933, Trametinib, SCH772984, JAK_8517, and Gallibiscoquinazole.

Experimental verification of the expression of the signature genes

We validated the expression of the signature genes in mammary epithelial cells (MCF10A) and two breast cancer cells (MDA-MB-231, and MCF-7). RT-qPCR results showed that CTSL expression was significantly lower in MCF-7 not MDA-MB-231 cells versus MCF10A cells (**Figure 8A**). No significant difference in CTSD expression was detected between MCF10A cells and MDA-MB-231/MCF-7 cells (**Figure 8B**). In comparison to MCF10A cells, lower ELK4

expression was found in MDA-MB-231, and MCF-7 cells (**Figure 8C**). XRCC4 expression was remarkably higher in MDA-MB-231 not MCF-7 cells than MCF10A cells (**Figure 8D**). In addition, HSPA8 was found to be remarkably lower in MCF-7 not MDA-MB-231 cells versus MCF10A cells (**Figure 8E**). Western blot results demonstrated that CTSL, and CTSD expression was notably down-regulated and XRCC4 expression was notably up-regulated both in MDA-MB-231, and MCF-7 cells versus MCF10A cells (**Figure 8F**). HSPA8 expression was modestly down-regulated in MDA-MB-231/MCF-7 cells versus MCF10A cells. However, ELK4 expression was not detected in the three cell lines.

Knockdown of XRCC4 attenuates migration of MCF10A, MDA-MB-231, and MCF-7 cells

For observing the role of XRCC4 in tumor cell migration, MCF10A, MDA-MB-231, and MCF-7 cells were transfected with specific siRNAs of XRCC4. Consequently, XRCC4 expression was significantly lowered in MCF10A, MDA-MB-231, and MCF-7 cells, notably si-XRCC4#1 and si-XRCC4#2 (**Figure 9A-C**). Wound scratch assay results showed that MCF10A, MDA-MB-231, and MCF-7 cells with si-XRCC4#1 and si-XRCC4#2 had the significantly lower migration rate in comparison to those with si-NC transfection (**Figure 9D-I**). Therefore, silencing XRCC4 attenuated migrative

capacities of MCF10A, MDA-MB-231, and MCF-7 cells.

Overexpressed CTSL, CTSD, and HSPA8 mitigate migration of MCF10A, MDA-MB-231, and MCF-7 cells

CTSL, CTSD, and HSPA8 were significantly overexpressed by the corresponding gene overexpression plasmids in MCF10A, MDA-MB-231, and MCF-7 cells (**Figure 10A-C**). As illustrated in wound scratch assay results, overexpressed CTSL, CTSD, and HSPA8 notably attenuated the migrative rate of MCF10A, MDA-MB-231, and MCF-7 cells (**Figure 10D-I**). Thus, overexpressed CTSL, CTSD, and HSPA8 mitigated migrative capacities of MCF10A, MDA-MB-231, and MCF-7 cells.

Discussion:

TNBC is an aggressive subtype characterized by widespread intratumoral heterogeneity [37]. Recent technological development allows for increasingly reliable and integrative single-cell analysis of the tumor microenvironment at the transcriptional level, which facilitates the observation of cell populations and cell-cell crosstalk [38]. In this work, we utilized scRNA-seq in combination with bulk transcriptome data for systematically analyzing the cell components within the tumor microenvironment of TNBC. Tumorigenesis is governed both by genetically altered tumor cells and non-malignant

host cells within the tumor microenvironment, extensively impacting tumor progression, metastases, and therapeutic outcomes [39]. We reported that the tumor microenvironment of TNBC was composed of B cells, dendritic cells, endothelial cells, epithelial cells, fibroblasts, macrophages, monocytes, plasmablasts, and T cells. Most of them (B cells, dendritic cells, fibroblasts, macrophages, plasmablasts, and T cells) presented the higher abundance in TNBC versus non-TNBC. In combination with bulk transcriptome analysis, macrophages were significantly abundant in the TNBC microenvironment. In addition, our work provided two maps of cellular interactions within the microenvironment of non-TNBC and TNBC. There was more active cell-cell crosstalk in the TNBC microenvironment, especially macrophages with other cell populations, unravelling the crucial implication of macrophages in TNBC.

Inter-patient and intra-tumor heterogeneity complicate the identification of predictive biomarkers and efficient treatment for TNBC [40]. This work proposed a novel macrophage-based prognostic model composed of CTSD, CTSL, ELK4, HSPA8, and XRCC4. High risk score was predictive of poor prognostic outcomes. Based upon external verification, the model enabled to reliably predict TNBC patient prognosis, especially one-year survival.

Somatic mutations were extensively heterogeneous between low- and high-risk TNBC patients. Additionally, high-risk patients presented higher responses to immunotherapy in accordance with higher T-cell inflamed score, lower TIDE score, up-regulated immune checkpoint molecules (CD80, CD86, IDO1, LAG3, LAIR1, PDCD1, HAVCR2, and LGALS3). It was also investigated the heterogeneous sensitivity to drugs between low- and high-risk patients. Thus, the macrophage-relevant gene signature displayed the potential in estimating prognostic outcomes and therapy response in TNBC.

High XRCC4 expression correlates to undesirable progression-free survival following radiotherapy for TNBC patients. In vitro, silencing XRCC4 sensitizes TNBC cells to ionizing radiation [41]. Herein, XRCC4 was proven to be up-regulated in TNBC cells versus healthy cells. Knockdown of XRCC4 contributed to the impaired migrative capacities of MCF10A, MDA-MB-231, and MCF-7 cells. Nuclear CTSL has been determined as a positive biomarker of TNBC [42]. BRCA1 deficiency activates CTSL-induced degradation of 53BP1 in TNBC cells. Nuclear levels of CTSL, vitamin D receptor, and 53BP1 act as a triple biomarker signature for stratifying patients with BRCA1 mutation tumors and TNBC, possessing prediction significance for drug response. High CTSD expression is linked with

poor recurrence-free survival of TNBC patients, and extracellular CTSD can be measured within the tumor microenvironment, but not in normal breast stroma [43]. A 9-kDa matricellular SPARC fragment released by CTSD displays pro-tumor activity in the TNBC microenvironment [44]. Immunotherapy with CTSD-targeting antibody has been determined as a promising therapeutic regimen against TNBC. HSPA8 has been determined to associate with TNBC survival [45]. CTSL, CTSD, and HSPA8 were down-regulated in breast cancer cells, and their overexpression attenuated migrative abilities of MCF10A, MDA-MB-231, and MCF-7 cells. Altogether, the aberrantly expressed signature genes (XRCC4, CTSL, CTSD, and HSPA8) were in relation to tumor aggressiveness.

Nevertheless, more experiments are needed to verify the roles of the signature genes in TNBC aggressiveness. In addition, the efficacy of the macrophage-based gene signature in survival prediction requires to be proven in larger prospective cohorts.

Conclusion:

Collectively, based upon the integration of single-cell and bulk transcriptome analyses, this work proposed a novel macrophage-based gene signature (containing CTSD, CTSL, ELK4, HSPA8, and XRCC4) for prediction of survival outcomes and therapy response in

TNBC patients. The signature genes were proven to associate with tumor aggressiveness. These findings might positively impact on future clinical research involving macrophages in TNBC.

Abbreviations:

TNBC: triple-negative breast cancer; scRNA-seq: single-cell RNA sequencing; GEO: Gene Expression Omnibus; TCGA: The Cancer Genome Atlas; PCA: principal component analysis; PCs: principal components; UMAP: uniform manifold approximation and projection; FC: fold change; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; GSEA: gene set enrichment analysis; DEGs: differentially expressed genes; LASSO: least absolute shrinkage and selection operator; CTA: cancer-testis antigen; siRNAs: small interfering RNAs; OS: overall survival; DFS: disease-free survival; ROCs: receiver operator characteristic curves.

Acknowledgements:

Not applicable.

Funding:

This research was supported by The Natural Science Foundation of Zhejiang Province, China (Grant No. LY23H160004, LY23H160006),

The Medical and Health Research Project of Zhejiang Province (Grant No. 2022KY635, 2023RC007).

Availability of data and material:

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Authors' contributions:

Xiaowei Wang, Xiaojia Wang conceived and designed the study. Yuan Huang, Yuan Yu and Yabing Zheng conducted most of the experiments and data analysis, and wrote the manuscript. Huangping Zhang, Ziwen Zhang participated in collecting data and helped to draft the manuscript. All authors reviewed and approved the manuscript.

Ethics approval and consent to participate:

Not applicable.

Consent for publication:

Not applicable.

Conflicts of Interest:

The authors declare that they have no competing interests.

References:

1. Liao L, Zhang YL, Deng L, Chen C, Ma XY, Andriani L, et al. Protein Phosphatase 1 Subunit PPP1R14B Stabilizes STMN1 to Promote Progression and Paclitaxel Resistance in Triple-Negative Breast Cancer. Cancer Res 2023; 83:471-484.
2. Zhang TM, Liao L, Yang SY, Huang MY, Zhang YL, Deng L, et al. TOLLIP-mediated autophagic degradation pathway links the VCP-TMEM63A-DERL1 signaling axis to triple-negative breast cancer progression. Autophagy 2023; 19:805-821.
3. Yang F, Xiao Y, Ding JH, Jin X, Ma D, Li DQ, et al. Ferroptosis heterogeneity in triple-negative breast cancer reveals an innovative immunotherapy combination strategy. Cell Metab 2023; 35:84-100.e8.
4. Manoochehri M, Borhani N, Gerhäuser C, Assenov Y, Schönung M, Hielscher T, et al. DNA methylation biomarkers for noninvasive detection of triple-negative breast cancer using liquid biopsy. Int J Cancer 2023; 152:1025-1035.
5. Zhang Y, Chen H, Mo H, Hu X, Gao R, Zhao Y, et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer.

Cancer Cell 2021; 39:1578-1593.e8.

6. Schmid P, Rugo HS, Adams S, Schneeweiss A, Barrios CH, Iwata H, et al. Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2020; 21:44-59.

7. Loibl S, Untch M, Burchardi N, Huober J, Sinn BV, Blohmer JU, et al. A randomised phase II study investigating durvalumab in addition to an anthracycline taxane-based neoadjuvant therapy in early triple-negative breast cancer: clinical results and biomarker analysis of GeparNuevo study. Ann Oncol 2019; 30:1279-1288.

8. Loibl S, Schneeweiss A, Huober J, Braun M, Rey J, Blohmer JU, et al. Neoadjuvant durvalumab improves survival in early triple-negative breast cancer independent of pathological complete response. Ann Oncol 2022; 33:1149-1158.

9. Zhu Y, Zhu X, Tang C, Guan X, Zhang W. Progress and challenges of immunotherapy in triple-negative breast cancer. Biochim Biophys Acta Rev Cancer 2021; 1876:188593.

10. Hey J, Halperin C, Hartmann M, Mayer S, Schönung M, Lipka DB, et al. DNA methylation landscape of tumor-associated

macrophages reveals pathways, transcription factors and prognostic value relevant to triple-negative breast cancer patients. Int J Cancer 2023; 152:1226-1242.

11. Li H, Yang P, Wang J, Zhang J, Ma Q, Jiang Y, et al. HLF regulates ferroptosis, development and chemoresistance of triple-negative breast cancer by activating tumor cell-macrophage crosstalk. J Hematol Oncol 2022; 15:2.

12. Singh S, Lee N, Pedroza DA, Bado IL, Hamor C, Zhang L, et al. Chemotherapy Coupled to Macrophage Inhibition Induces T-cell and B-cell Infiltration and Durable Regression in Triple-Negative Breast Cancer. Cancer Res 2022; 82:2281-2297.

13. Zhang Y, Fan Y, Jing X, Zhao L, Liu T, Wang L, et al. OTUD5-mediated deubiquitination of YAP in macrophage promotes M2 phenotype polarization and favors triple-negative breast cancer progression. Cancer Lett 2021; 504:104-115.

14. Ye Y, Ma J, Zhang Q, Xiong K, Zhang Z, Chen C, et al. A CTL/M2 macrophage-related four-gene signature predicting metastasis-free survival in triple-negative breast cancer treated with adjuvant radiotherapy. Breast Cancer Res Treat 2021; 190:329-341.

15. Su P, Peng Z, Xu B, Yang B, Jin F. Establishment and validation of an individualized macrophage-related gene signature to

predict overall survival in patients with triple negative breast cancer. PeerJ 2021; 9:e12383.

16. Luo H, Hong R, Xu Y, Zheng Q, Xia W, Lu Q, et al. Construction and validation of a macrophage polarization-related prognostic index to predict the overall survival in patients with early-stage triple-negative breast cancer. Gland Surg 2023; 12:225-242.

17. Xu K, Wang R, Xie H, Hu L, Wang C, Xu J, et al. Single-cell RNA sequencing reveals cell heterogeneity and transcriptome profile of breast cancer lymph node metastasis. Oncogenesis 2021; 10:66.

18. Brueffer C, Vallon-Christersson J, Grabau D, Ehinger A, Häkkinen J, Hegardt C, et al. Clinical Value of RNA Sequencing-Based Classifiers for Prediction of the Five Conventional Breast Cancer Biomarkers: A Report From the Population-Based Multicenter Sweden Cancerome Analysis Network-Breast Initiative. JCO Precis Oncol 2018; 2.

19. Nagalla S, Chou JW, Willingham MC, Ruiz J, Vaughn JP, Dubey P, et al. Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. Genome Biol 2013; 14:R34.

20. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, Marioni JC. EmptyDrops: distinguishing cells from empty droplets in

droplet-based single-cell RNA sequencing data. Genome Biol 2019; 20:63.

21. McCarthy DJ, Campbell KR, Lun AT, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 2017; 33:1179-1186.

22. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018; 36:411-420.

23. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol 2018.

24. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12:453-7.

25. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021; 12:1088.

26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res

2003; 13:2498-504.

27. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics 2012; 16:284-7.

28. Luo W, Pant G, Bhavnasi YK, Blanchard SG, Jr., Brouwer C. Pathview Web: user friendly pathway visualization and data integration. Nucleic Acids Res 2017; 45:W501-w508.

29. 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. Proceedings of the National Academy of Sciences 2005; 102:15545-15550.

30. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017; 14:1083-1086.

31. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43:e47.

32. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat

Softw 2010; 33:1-22.

33. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018; 28:1747-1756.

34. Hu J, Yu A, Othmane B, Qiu D, Li H, Li C, et al. Siglec15 shapes a non-inflamed tumor microenvironment and predicts the molecular subtype in bladder cancer. Theranostics 2021; 11:3089-3108.

35. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med 2018; 24:1550-1558.

36. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform 2021; 22.

37. So JY, Ohm J, Lipkowitz S, Yang L. Triple negative breast cancer (TNBC): Non-genetic tumor heterogeneity and immune microenvironment: Emerging treatment options. Pharmacol Ther 2022; 237:108253.

38. Tietscher S, Wagner J, Anzeneder T, Langwieder C, Rees M, Sobottka B, et al. A comprehensive single-cell map of T cell exhaustion-associated immune environments in human breast

cancer. Nat Commun 2023; 14:98.

39. Houthuijzen JM, de Bruijn R, van der Burg E, Drenth AP, Wientjens E, Filipovic T, et al. CD26-negative and CD26-positive tissue-resident fibroblasts contribute to functionally distinct CAF subpopulations in breast cancer. Nat Commun 2023; 14:183.

40. Henriet E, Knutsdottir H, Grasset EM, Dunworth M, Haynes M, Bader JS, et al. Triple negative breast tumors contain heterogeneous cancer cells expressing distinct KRAS-dependent collective and disseminative invasion programs. Oncogene 2023; 42:737-747.

41. Wen Y, Dai G, Wang L, Fu K, Zuo S. Silencing of XRCC4 increases radiosensitivity of triple-negative breast cancer cells. Biosci Rep 2019; 39.

42. Grotsky DA, Gonzalez-Suarez I, Novell A, Neumann MA, Yaddanapudi SC, Croke M, et al. BRCA1 loss activates cathepsin L-mediated degradation of 53BP1 in breast cancer cells. J Cell Biol 2013; 200:187-202.

43. Ashraf Y, Mansouri H, Laurent-Matha V, Alcaraz LB, Roger P, Guiu S, et al. Immunotherapy of triple-negative breast cancer with cathepsin D-targeting antibodies. J Immunother Cancer 2019; 7:29.

44. Alcaraz LB, Mallavialle A, David T, Derocq D, Delolme F, Dieryckx C, et al. A 9-kDa matricellular SPARC fragment released by cathepsin D exhibits pro-tumor activity in the triple-negative breast cancer microenvironment. Theranostics 2021; 11:6173-6192.

45. Yang Q, Sun K, Xia W, Li Y, Zhong M, Lei K. Autophagy-related prognostic signature for survival prediction of triple negative breast cancer. PeerJ 2022; 10:e12878.

Figure legends:

Figure 1. Single-cell and bulk transcriptome analyses unravel cellular heterogeneity in TNBC. (A) UMAP mapping the identified cell populations based upon scRNA-seq data. (B) The expression of the used marker genes across diverse cell populations. (C) The cell ratio of each cell population in single-cell TNBC and non-TNBC. (D) The top 10 novel marker genes of diverse cell populations. (E) The relative cell abundance of the identified cell populations in bulk TNBC and non-TNBC tissues. (F, G) Comparison of the cell ratio of each cell population between bulk TNBC and non-TNBC tissues. **p≤0.01; ***p≤0.001.

Figure 2. Cell-cell interactions and transcriptional activity of macrophages in TNBC. (A, B) Cell-cell interaction networks in (A)

non-TNBC and (B) TNBC. (C) Differentially expressed transcription factors in TNBC macrophages versus non-TNBC macrophages. (D, E) DEGs between TNBC macrophages versus non-TNBC macrophages. (F, G) The main GO and KEGG pathways enriched by the DEGs. (H) Antigen processing and presentation enriched by the DEGs.

Figure 3. Construction of a novel macrophage-based prognostic signature for TCGA TNBC. (A) Coefficient profiling in the LASSO analysis. (B) The 10-fold cross-validation results. (C) Univariate-cox regression results of the identified signature genes with TNBC survival. (D) The expression of the signature genes along the risk score. (E) Distribution of risk score among TNBC cases. (F, G) Distribution of alive/dead or disease-free/recurred/progressed status along the risk score. (H-J) OS probability of low- or high-risk patients in the training, test, and entire cohorts. (K-M) One-, three-, or five-year ROCs based upon the prognostic model.

Figure 4. Associations of the macrophage-based prognostic model with clinicopathological traits and external validation of the model. (A-C) Distribution of the risk score across diverse T, N, M stages in the TCGA TNBC cohort. (D) Distribution of the risk score across distinct pathological stages of TCGA TNBC cases. (E) Correlation between the risk score and tumor purity among TCGA TNBC cases.

(F, G) Distribution of the risk score among diverse histological grades in the GSE96058 or GSE45255 cohort. (H) OS probability of low- or high-risk patients in the GSE96058 cohort. (I) One-, three-, or five-year ROCs on the basis of the model among the GSE96058 samples. (J) DFS probability of low- or high-risk patients in the GSE45255 cohort. (K) One-, three-, or five-year ROCs based upon the model among the GSE45255 samples.

Figure 5. Heterogeneous somatic mutations between low- and high-risk TNBC patients. (A) Distribution of TMB score across TNBC samples. (B) Comparison of TMB score in high- versus low-risk groups. (C) The top ten mutated genes across low- and high-risk samples. (D, E) Comparison of the frequency of mutated genes in low- or high-risk group. (F) Co-occurred or mutually exclusive mutated genes in high-risk samples. (G) Co-occurred or mutually exclusive mutated genes in low-risk samples. NS: no significance; *$p \leq 0.05$; **$p \leq 0.01$.

Figure 6. Molecular mechanisms underlying the macrophage-based prognostic model. (A-D) GSEA showing the main biological processes, cellular components, molecular functions, and KEGG pathways that were significantly different between high- and low-risk groups.

Figure 7. High-risk patients with higher response to immunotherapy. (A, B) Comparison of (A) T-cell inflamed score and (B) TIDE score

between low- and high-risk TCGA TNBC patients. (C) Distribution of the ratio of responders or non-responders to immunotherapy in low- or high-risk group. (D) Difference in CTA number between groups. (E) Comparison of the expression of immune checkpoints between groups. (F) Difference in the sensitivity to drugs between groups. NS: no significance; *p≤0.05; **p≤0.01; ***p≤0.001.

Figure 8. Experimental verification of the expression of the signature genes. (A-E) RT-qPCR of the mRNA expression of CTSL, CTSD, ELK4, XRCC4, and HSPA8 in MCF10A, MDA-MB-231, and MCF-7 cells. (F) Western blot of the expression of CTSL, CTSD, XRCC4, and HSPA8 in MCF10A, MDA-MB-231, and MCF-7 cells. NS: no significance; **p≤0.01; ***p≤0.001; ****p≤0.0001.

Figure 9. Knockdown of XRCC4 attenuates migration of MCF10A, MDA-MB-231, and MCF-7 cells. (A-C) RT-qPCR of the mRNA expression of XRCC4 in MCF10A, MDA-MB-231, and MCF-7 cells transfected with XRCC4 siRNAs or controls. (D-I) Wound scratch assay results of (D, E) MCF10A, (F, G) MDA-MB-231, and (H, I) MCF-7 cells with XRCC4 siRNAs or controls. Bar, 20, or 50 μm. **p≤0.01; ***p≤0.001; ****p≤0.0001.

Figure 10. Overexpressed CTSL, CTSD, and HSPA8 mitigate migration of MCF10A, MDA-MB-231, and MCF-7 cells. (A-C) RT-qPCR of the mRNA expression of CTSL, CTSD, and HSPA8 in

MCF10A, MDA-MB-231, and MCF-7 cells transfected with the corresponding gene overexpression plasmids. (D-I) Wound scratch assay results of (D, E) MCF10A, (F, G) MDA-MB-231, and (H, I) MCF-7 cells transfected with empty vectors or CTSL, CTSD or HSPA8 overexpression plasmids. Bar, 20, or 50 μm. *p≤0.05; **p≤0.01; ***p≤0.001; ****p≤0.0001.

Supplementary files:

Supplementary figure 1. Quality control of scRNA-seq data. (A-L) Removal of empty droplets and single cells with proportion of mitochondrial genes >10% and proportion of ribosome genes <10% in (A-C) GSM5457199; (D-F) GSM5457205; (G-I) GSM5457208; and (J-L) GSM5457211 specimens.

Supplementary figure 2. PCA of scRNA-seq data after quality control. (A) The top 2 PCs. (B) Selection of the appropriate number of PCs. (C) PCA plots of single cells from TNBC and non-TNBC. (D) The top 20 marker genes of the top 9 PCs.

Supplementary figure 3. Single cell clustering analysis. (A) UMAP for clustering single cells into diverse clusters. (B) Distribution of TNBC and non-TNBC single cells. (C) The cell ratio of the identified clusters between TNBC and non-TNBC. (D) The top 10 marker genes identified in each cell cluster. (E) The top one marker gene of each

cell cluster.

Supplementary table 1. Differentially expressed transcription factors in TNBC macrophages versus non-TNBC macrophages.

Supplementary table 2. Up-regulated genes in TNBC macrophages versus non-TNBC macrophages.

Supplementary table 3. Down-regulated genes in TNBC macrophages versus non-TNBC macrophages.

Figures

Figure 6

Figure 7

Figure 8

Figure 9

Figure 10

Figure 11

Figure 12

Figure 13

Figure 14

Figure 15

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigure1.tif
- Supplementaryfigure2.tif
- Supplementaryfigure3.tif
- Supplementarytable1.xlsx
- Supplementarytable2.xlsx
- Supplementarytable3.xlsx