Cellular heterogeneity map of diverse immune and stromal phenotypes within breast tumor microenvironment

Yuan Li [Corresp., 1], **Zuhua Chen** [2], **Long Wu** [1], **Junjie Ye** [1], **Weiping Tao** [1]

[1] Department of Oncology, Renmin Hospital of Wuhan University, Wuhan, China

[2] Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Corresponding Author: Yuan Li
Email address: liyuanpumc@gmail.com

Background: Cellular heterogeneity with tumor microenvironment plays critical roles in tumorigenesis and development. Global view of tumor-infiltrating immune and stromal cells at high resolution is greatly needed in breast tumor.

Methods: portrayed the cellular heterogeneity map of a total of 64 cell types in 1092 breast tumor and adjacent normal tissues using xCell, which digitally dissect tissue cellular heterogeneity based on gene expression.

Results: Higher proportions of immune cells were enriched in tumor compared with normal tissues. Immune inhibitory receptors (PD1, CTLA4, LAG3 and TIM3) were co-expressed on certain subtypes of T cells in breast tumor, especially that PD1 and CTLA4 both positively correlated with CD8+ Tcm and CD8+ T cells. Furthermore, CD4+ Tem, CD8+ Tcm, CD8+ T-cells, CD8+ naive T-cells and B cells were favorable whereas CD4+ naive T-cells were adverse prognostic factors for breast cancer patients. Moreover, TDRD6 and TTK were promising targets for tumor vaccines which might activate T cells and B cells. Meanwhile, Endothelial cells and fibroblasts were significantly lower in tumor tissues. Astrocytes and mesangial cells were negatively correlated with T stage. Mesangial cells and keratinocytes were favorable prognostic factors, whereas myocytes were adverse prognostic factors. A prognosis model with high performance was built for breast cancer patients. Specifically, great cellular heterogeneity was uncovered among different subtypes of breast cancer by Her2, ER, and PR status. Tri-negative patients had the highest while luminal type patients had the smallest fraction of immune cells. Different cell might have different or even opponent role in the prognosis of breast cancer patients.

Conclusions: Our analysis uncovered a unique cellular heterogeneity map of diverse immune and stromal phenotypes within breast tumor microenvironment and novel potential therapeutic targets and prediction biomarkers with prognostic utility.

Cellular heterogeneity map of diverse immune and stromal phenotypes within breast tumor microenvironment

Yuan Li[1*], Zuhua Chen[2], Long Wu[1], Junjie Ye[1], Weiping Tao[1*]

1 Department of Oncology, Renmin Hospital of Wuhan University, Wuhan, China

2 Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei Province, China.

Correspondence to: Yuan Li and Weiping Tao, Department of Oncology, Renmin Hospital of Wuhan University, No.99, Zhangzhidong Road, Wuhan, 430060, China.

Tel: +86-27-88041911-82291.

Email: liyuanpumc@gmail.com; taowpwp@sina.com

Abstract

Background: Cellular heterogeneity with tumor microenvironment plays critical roles in tumorigenesis and development. Global view of tumor-infiltrating immune and stromal cells at high resolution is greatly needed in breast tumor.

Methods: portrayed the cellular heterogeneity map of a total of 64 cell types in 1092 breast tumor and adjacent normal tissues using xCell, which digitally dissect tissue cellular heterogeneity based on gene expression.

Results: Higher proportions of immune cells were enriched in tumor compared with normal tissues. Immune inhibitory receptors (PD1, CTLA4, LAG3 and TIM3) were co-expressed on certain subtypes of T cells in breast tumor, especially that PD1 and CTLA4 both positively correlated with CD8+ Tcm and CD8+ T cells. Furthermore, CD4+ Tem, CD8+ Tcm, CD8+ T-cells, CD8+ naive T-cells and B cells were favorable whereas CD4+ naive T-cells were adverse prognostic factors for breast cancer patients. Moreover, TDRD6 and TTK were promising targets for tumor vaccines which might activate T cells and B cells. Meanwhile, Endothelial cells and fibroblasts were significantly lower in tumor tissues. Astrocytes and mesangial cells were negatively correlated with T stage. Mesangial cells and keratinocytes were favorable prognostic factors, whereas myocytes were adverse prognostic factors. A prognosis model with high performance was built for breast cancer patients. Specifically, great cellular heterogeneity was uncovered among different subtypes of breast cancer by Her2, ER, and PR status. Tri-negative patients had the highest while luminal type patients had the smallest fraction of immune cells. Different

cell might have different or even opponent role in the prognosis of breast cancer patients.

Conclusions: Our analysis uncovered a unique cellular heterogeneity map of diverse immune and stromal phenotypes within breast tumor microenvironment and novel potential therapeutic targets and prediction biomarkers with prognostic utility.

Key words: Breast cancer; Immune; Stromal; Cellular heterogeneity

Introduction

Breast cancer is one of the predominant type of tumors in women and despite great progress has been achieved in the early diagnosis and treatment recently, drug resistance and distal metastasis remain major causes of mortality (Cassetta & Pollard 2017). Tumors are complex evolving environments, composed of not only malignant cells but also immune and stromal infiltrates. Growing evidence suggests that tumor microenvironment plays a fundamental role in the initiation to malignancy as well as resistance to therapy (Noy & Pollard 2014). Tumor-infiltrating cells can demonstrate either tumor suppressive or tumor-promoting effects, depending on the cancer type. For instance, regulatory T cells (Tregs) and tumor associated macrophages (TAMs) were reported to be associated with pro-tumor functions (De Palma & Lewis 2013; Nishikawa & Sakaguchi 2014; Noy & Pollard 2014), whereas CD8+ T cells were reported to be with improved clinical outcomes and response to immunotherapy (Tumeh et al. 2014). Basic research in cancer immunology paved the way for the development and approval of checkpoint blockers. These drugs, which augment T cell activity by blocking cytotoxic lymphocyte antigen-4 (CTLA4), programmed cell death protein 1 (PD1), or PD1 ligand (PDL1), show remarkable clinical effects.

Understanding the cellular heterogeneity within the tumor microenvironment may discover predictive biomarkers and improve existing treatments or develop novel therapeutic strategies. Traditional approaches for dissecting the cellular heterogeneity, including flow cytometry and immunohistochemistry, were extremely difficult to apply to

65 solid tumors with limited throughput (Gentles et al. 2015). Advances in bioinformatics
66 provide novel methods to dissect cellular heterogeneity based on gene expression
67 profiles (Abbas et al. 2009; Newman et al. 2015; Rooney et al. 2015; Shen-Orr & Gaujoux
68 2013), one representative method is CIBERSORT, which could provide an estimation of
69 the abundances of 22 immune cell types (Newman et al. 2015). However, rare subsets
70 of immune cells and stromal cells are ignored by CIBERSORT, which are now recognized
71 to be important both in promoting and inhibiting of tumor growth, invasion, and metastasis
72 (Galon et al. 2006; Hanahan & Coussens 2012). Recently, xCell has been reported to
73 provide portrays of 64 cell types, including immune cells and stroma cells, within tissues
74 (Aran et al. 2017a; Aran et al. 2017b). Here, we applied this novel approach to digitally
75 portray the cellular heterogeneity map within breast tumor microenvironment, aiming to
76 provide novel insights into potential interactions and uncover predictive biomarkers or
77 therapeutic targets.

78 **Methods and materials**

79 **Data curation**

80 RNA-seq data and matching clinical parameters of a total of 1,092 patients with
81 breast cancer were downloaded from The Cancer Genome Atlas (TCGA) data portal
82 (https://portal.gdc.cancer.gov/). 276 known Cancer/Testis (CT) genes was downloaded
83 from CTDatabase (http://www.cta.lncc.br/).

84 **Bioinformatics analysis**

85 xCell (http://xcell.ucsf.edu/) is a gene-signature based method for cell type

enrichment of up to 64 cell types, including immune and stroma cells at high resolution. Deconvolution of cellular heterogeneity within breast tumor microenvironment from RNA sequencing data was performed using xCell R package (Aran et al. 2017b) (Beta version) from GitHub in R (version 3.3.1). Tests for differences and correlations were performed accordingly.

Survival analysis

In searching for survival associated genes, univariate and multivariate COX regression were then applied. To compare the ability of the prognostic predictors, survivalROC package (Version 1.0.3) in R, which allows for time dependent ROC curve estimation with censored data (Heagerty et al. 2000), was used to generate the area under the curve (AUC) of the receiver-operator characteristic (ROC) curve for each parameters.

Statistical analysis

Differentially enriched cell types between different groups were compared with Student's t-test (two groups) or One-way ANOVA analysis (Three groups). Correlation analysis were performed by Spearman method. The survival curves were compared using Kaplan-Meier method and log-rank test. All tests were two sided, and a P value of less than 0.05 was considered as statistical significance unless stated otherwise. Data were analyzed using R (version 3.4.4).

Results

64 cell types were characterized in 1092 breast cancer patients

A total of 1092 breast patients were enrolled in this study, among which 112 paired adjacent normal tissues were available. We portrayed the cellular heterogeneity in 1092 breast tumor tissues and 112 normal tissues using xCell method. Generally, the 64 cell types were divided into four groups, including 34 immune cells, 13 stromal cells, 9 stem cells and 8 other cells (**Table 1**). Over half of the 64 cell types were immune cells, providing a full view of innate and adaptive immune status with detailed cell subtypes such as CD4+ naive cells, CD4+ T-cells, CD4+ Tcm and CD4+ Tem. Meanwhile, important stromal cells, such as fibroblast, osteoblast and pericyte were also included.

Breast tumor tissues had higher fractions of immunes cells than normal tissues.

We calculated the median fractions for each cell type in adjacent normal and breast tumor tissues, respectively. As results, the relative proportions of the 64 cells were quite different between breast tumor and normal tissues (**Figure 1A**), breast tumor tissue had higher fraction of immune cells with red to light blue labels whereas normal tissue had larger proportions of stem and stromal cells with blue to red labels (**Figure 1A**). Unsupervised cluster analysis revealed that breast tumor tissues and adjacent normal tissues were generally clustered into different groups, meanwhile, immune cells were also largely clustered into several subgroups (**Figure 1B**), indicating that the cellular heterogeneity in tumor vs. normal tissues was much greater than that in individual sample. Accordingly, dimensionality reduction and visualization by t-Distributed Stochastic Neighbor Embedding (t-SNE) also suggested clear gap between tumor and

adjacent normal tissues (**Figure 1C**). Next, we compared the fractions of each cell type between breast tumor and normal tissues. As shown in **Figure 1D-J**, a number of cell types showed dramatic changes between the tumor and normal tissues. Generally, greater diversity was seen in immune cells compared with stem or stromal cells. For innate immune cells, neutrophils were significantly higher in normal tissue whereas eosinophils were higher in tumor tissues (**Figure 1D**). Meanwhile, difference of DC cells between normal and tumor tissues was not significant, but iDC was significantly lower whereas pDC and aDC was significantly higher in tumor tissues (**Figure 1E**). Same phenomenon was also seen in macrophages, which was macrophage M1 was higher while macrophage M2 was lower in tumor tissues (**Figure 1F**). For adaptive immune cells, CD4+ Tcm was significantly lower in tumor tissues, CD4+ Tem, CD8+ naïve T cells and CD8+ Tcm were significantly higher in tumor tissues (**Figure 1G**). Moreover, plasma cells, pro B cells, Tgd, Th1, Th2 cells and Tregs were also significantly higher in tumor tissues (**Figure 1H and I**). For stromal cells, representative cells such as endothelial cells and fibroblasts were significantly lower in tumor tissues (**Figure 1J**).

Inhibitory receptors were co-expressed on certain subtypes of T cells

Inhibitory receptors expressed on T cells, including PD1, CTLA4, LAG3 and TIM3, often lead to T-cell exhaustion, providing tumors with potential to escape immune control (Huang et al. 2017; Nirschl & Drake 2013). Blockage of CTLA4 or PD1 with specific antibodies has shown significant promise in overcoming immune suppression and

mediate tumor regression (Brahmer et al. 2012; Callahan et al. 2010).

We next investigated the potential correlations among these inhibitory receptors and CD4+/CD8+ T cells. In **Figure 2A and B**, heatmaps suggested that expression patterns of these inhibitory receptors were correlated with certain subsets of T cells with commonalities and differences between tumor (**Figure 2A**) and normal tissues (**Figure 2B**). Correlation analyses also demonstrated that in tumor tissues, CD8+ T-cells, CD8+ Tcm, CD8+ naive T-cells, CD4+ memory T cells and CD4+ naïve T cells were all positively correlated with expressions of these inhibitory receptors, especially with PD1 and CTLA4 expression ($P<0.05$, **Figure 2C and E**), whereas CD8+ Tem, CD4+ Tcm and CD4+ T-cells were not strongly correlated with them (**Figure 2C**). However, in normal tissues, only a few T cells were significantly correlated with these inhibitory receptors, especially that TIM3 expression was negatively correlated with CD4+ Tcm (**Figure 2D and G**). More importantly, significant correlation among the expression of inhibitory receptors was also observed (**Figure 2F and H**).

Cancer/testis genes TDRD6 and TTK were promising targets for breast cancer

Cancer/Testis (CT) genes are a cluster of tumor-associated proteins that are normally expressed in normal germ cells and diverse types of cancers, but not in usual somatic cells (Scanlan et al. 2002). Due to this limited expression process, the CT genes are considered striking points for cancer biomarkers and immunotherapy.

We next focus on how antitumor immunity responses to antigens generated by

cancer/Testis (CT) genes. First, we examined 276 known CT genes, which was downloaded from CTDatabase, for association with immune components. We listed all the significant correlations between immune cells and CT genes with cutoff of P value less than 0.001 in **Figure 3A**. The majority of the adaptive immune cells were significantly correlated with CT genes. We noticed that T cells such as CD8+ T cells, and aDC, which belonged to adaptive and innate immune responses, were positively correlated with most of the CT genes (**Figure 3A-C**). Moreover, two CT genes, TDRD6 and TTK, were positively correlated with a number of immune cells, especially the CD4+/CD8+ T cells (**Figure 3D and E**), implying strong host immune reactions to these two cancer antigens.

Cellular heterogeneity was correlated with clinic-pathology of breast cancer

Cellular heterogeneity is an important part of tumor microenvironment which plays critical roles in initiation and development of tumor. We asked whether there were certain cell types were significantly correlated with clinical parameters, including age, sex, T stage, N stage, M stage and TNM stage. As shown in **Figure 4A**, generally, a number of cell types were significantly correlated with clinical parameters, especially T stage and M stage. For T stage, astrocytes, mesangial cells, and mast cells were negatively correlated with T stage, whereas plasma cells were positively correlated with T stage (**Figure 4B**). For M stage, CD4+ Tcm, CD4+ Tem, mv endothelial cells, NKT, and MSC were all significantly higher in patients with distal metastasis (**Figure 4C**). For N stage, one representative cell type was CLP, patients with lymphnode metastasis had significantly

191 higher CLP (**Figure 4D**). For TNM stage, Th1 cells and MSC were both positively
192 correlated with TNM stage (**Figure 4E**). Interestingly, we noticed that there were 12 male
193 breast cancer patients, who tended to have higher proportion of CLP and NKT, compared
194 with female breast cancer patients (**Figure 4F**).
195
196 **Prognostic model was built with survival associated cell types**
197 Emerging evidence suggests that the amount of tumor infiltrating lymphocytes (TILs)
198 of primary tumors consistently predicts favorable outcomes in a number of tumor types,
199 including breast cancer. Therefore, survival analyses were performed aiming to find
200 survival associated cell types within tumor microenvironment, which were shown in
201 **Figure 5A**. Generally, immune cells were more intensively involved in patients overall
202 survival, especially the CD4+ and CD8+ T cells (**Figure 5A**). Moreover, most T cells,
203 including CD8+ T cells, CD8+ Tcm, CD8+ naïve T cells and CD4 Tem, were favorable
204 prognostic factors, except that high CD4+ naïve T cells were associated with worse
205 overall survival (**Figure 5A-M**). Furthermore, NKT, class switched memory B cells, NK
206 cells, cDC and pDC were also significantly associated with overall survival (**Figure 5A-**
207 **M**). More importantly, a number of stromal cells were also prognostic factors. For
208 example, mesangial cells and keratinocytes were favorable prognostic factors, whereas
209 myocytes were adverse prognostic factors (**Figure 5A-M**). Multivariate COX regression
210 revealed that CD8+ T cells, keratinocytes, NKT and Class switched memory B cells were
211 independent prognostic factors. We then built a prognosis predictor with all these

212 significant prognosis factors, and the prediction model performed well in distinguishing
213 good or poor survival of breast cancer patients with the highest AUC of ROC of 0.708
214 compared to all the factors separately (**Figure 5N and O**).
215
216 **Subtypes of breast cancer had diverse phenotypes of cellular heterogeneity**
217 Emerging evidence suggests that the breast cancer transcriptome has a wide range
218 of intratumoral heterogeneity, as well as genomic heterogeneity simply based on ER, PR
219 and Her2, which is shaped by the tumor cells and immune cells in the surrounding
220 microenvironment (Chung et al. 2017). We further explored the cellular heterogeneity
221 among different subtypes of breast cancer by Her2, ER, and PR status. According the
222 clinicopathological parameters provided by TCGA, 1,092 breast cancer patients were
223 classified into five groups, which including 30 Her2+_HR- patients, 59 Her2+_HR+
224 patients, 426 Luminal type (Her2-_HR+) patients, 97 triple negative (Tri-negative)
225 patients, and 480 unknown patients. As shown in **Figure 6**, the relative proportion of
226 different cells varied a lot among these five subtypes. Tri-negative patients had the
227 highest fraction of immune cells while luminal type patients had the smallest fraction of
228 immune cells, especially the CD4+ and CD8+ T cells (**Figure 6A and Supplementary**
229 **Figure S1**). Cluster analysis based on ImmuneScore and StromalScore and by heatmap
230 showed that there might not be certain driven cell types to distinguish these five subtypes
231 (**Figure 6B-C**). Furthermore, t-SNE cluster analysis highlighted that there was huge
232 difference of tumor-infiltrating cells among these five subtypes (**Figure 6D**). Specifically,

B cells, T cells, macrophages, Th cells and stromal cells including keratinocytes were significantly differentially enriched in these subtypes (**Figure 6E-F and Supplementary Figure S2**). For example, Tri-negative breast cancer tissues had highest fractions of plasma cells, pro B cells, macrophages M1, Th1 and Th2 cells whereas lowest fraction of macrophages M2 cells (**Figure 6E**). Meanwhile, keratinocytes, sebocytes, pericytes were high whereas MSC cells were low in Tri-negative breast cancer (**Figure 6F**). Lastly, survival analyses uncovered rich and interesting findings among these five subtypes (**Figure 7**). Each subtype of breast cancer had its unique pattern of survival associated tumor-infiltrating cells and different type of cell might have different or even opponent roles in the prognosis of breast cancer patients. Notably, keratinocytes and neurons were favorable and adverse prognosis factors in luminal type patients, respectively (**Figure 7A**). However, keratinocytes and neurons predicted worse and better overall survival in Tri-negative patients, respectively (**Figure 7D**). Taken together, these diversity of cellular heterogeneity among different subtypes of breast cancer suggested that the tumor-infiltrating cells within tumor microenvironment had essential roles in shaping the intratumor heterogeneity of breast cancer.

Discussion

Distinct tumor-infiltrating cell types were observed within tumor microenvironment, and the abundance and activation status of those cell types draw great attention to researchers and are being explored by novel bioinformatic techniques. Tumor-infiltrating

cells are now recognized to play important roles in the regulation of tumor proliferation, metastasis and invasion (Galon et al. 2006; Hanahan & Coussens 2012). With rapid accumulation of high-throughput data and evolution of bioinformatics algorithms, it is now possible to digitally dissect interactions tumors cells and tumor-infiltrating cells, including immune cells and stromal cells (Aran et al. 2017b; Hackl et al. 2016). Utilizing this high-throughput approach could provide novel insights into complexity of tumor microenvironment and innovations to breast cancer treatment and prognosis.

In this study, we portrayed the landscape of cellular heterogeneity within breast tumor and normal tissues by digital deconvolution using xCell approach. A total of 64 cell types were characterized with more than 30 immune cell types at high resolution, which were also the most studied set of cell types especially the tumor-infiltrating lymphocytes (TILs). Huge differences between breast tumor tissues and adjacent normal tissues were discovered with polarized enrichment of certain cell types. We focused on immune cell types, especially the CD4+/CD8+ T cells. Our results demonstrated that expression of inhibitory receptors (including PD1, CTLA4, LAG3 and TIM3) were positively correlated and were correlated with certain types of T cells, especially with CD8+ Tcm and CD8+ T cells, especially in tumor tissues. Furthermore, CD4+ Tem, CD8+ Tcm, CD8+ T-cells, CD8+ naive T-cells and B cells were associated with better prognosis whereas CD4+ naive T-cells were adverse prognostic factor for breast cancer patients. Moreover, active immune responses to tumor antigens were widely generated by innate and adaptive immune cells, including T cells, B cells and DC, while TDRD6 and TTK were promising

targets for cancer vaccines which could activated a number of immune cells, especially T cells and B cells. Meanwhile, stromal cells were also widely involved in the development of breast cancer. Endothelial cells and fibroblasts were significantly lower in tumor tissues. Astrocytes and mesangial cells were negatively correlated with T stage. Mesangial cells and keratinocytes were favorable prognostic factors, whereas myocytes were adverse prognostic factors. Taken together, we built a prognosis predictor with survival associated cell types and the prediction model performed well in distinguishing good or poor overall survival of breast cancer patients. Last but not least, cellular heterogeneity was also profiled in different subtypes of breast cancer based on Her2, ER and PR status. Five subtypes of breast cancer demonstrated diver phenotypes and different cell might had different or even opponent roles in each subtype of breast cancer.

Immunotherapies, especially the immune checkpoint blockers as well as therapeutic vaccines and engineered T cells, are being intensively investigated nowadays (Schumacher & Schreiber 2015), one important issue is how tumor cells interact with immune cells. The investigation of tumor-immune cell interaction poses considerable challenges, since the development of cancer and the immune surveillance by innate and adaptive immune cells with plasticity ad memory are both evolving ecosystems. The complex interplay between solid tumors and host immunity has been widely studied but remains incompletely understood. In multiple tumor types, tumor infiltrating lymphocytes (TILs) have been associated with clinical outcomes (Anagnostou & Brahmer 2015; Schoenfeld 2015). For example, CD8+ TILs have been shown to be favorable prognostic

296 in melanoma, colorectal, ovarian, and non-small cell lung cancer. In selected tumors, it
297 has been demonstrated that these CD8+ TILs are able to specially kill tumor cells (Yee
298 et al. 2002). When comes to immunity in breast cancer which remains largely unstudied,
299 only a few preliminary evaluations about prognosis value of CD4+/CD8+ T lymphocytes
300 has been reported. Presence of TILs has been shown to be potentially predictive and
301 prognostic in specific breast cancer subtypes. Specially in patients with human epidermal
302 growth factor receptor 2 positive and triple-negative breast cancer. Large adjuvant studies
303 have shown that higher levels of TILs in primary biopsies are associated with improved
304 overall survival and fewer recurrences, regardless of therapy (Adams et al. 2014; Dieci et
305 al. 2015; Loi et al. 2013).

306 In our study, we provided detailed information about immune cells in breast cancer
307 with numerous novel findings. Firstly, inhibitory receptors were expressed on certain
308 types of T cells, on which CD8+ T cells and CD8+ Tcm were preferred; secondly, Co-
309 expression of PD1, CTLA4, LAG3 and TIM3 were more commonly observed in tumor
310 tissues compared with normal tissues, which might be an explanation of limited effects of
311 single immune checkpoint inhibitor and the basis of combinatorial strategies. Ongoing
312 clinical trials of Simultaneous inhibition of PD1 and CTLA4 (Wolchok et al. 2013) or TIM3
313 (Fourcade et al. 2010) in advanced melanoma patients have shown enhanced efficacy.
314 Lastly, not all T cells were protective factors, CD8+ naive T cells rather than CD4+ naive
315 T cells were favorable prognostic factor for breast cancer patients overall survival. Taken
316 together, these results suggested that upregulated co-expression of multiple immune

317 inhibitory receptors might contribute to immune suppression and more attention should
318 be paid to subtypes of T cells when using immune checkpoint blockers, since immune
319 cells were highly conditional and might play different or even opposing roles in responses
320 to tumor cells.

321 Growing evidence suggests that not only immune cells, but also tumor cell-extrinsic
322 factors, including fibroblasts, endothelial cells, adipocytes with in tumor
323 microenvironment, have important roles in inhibiting apoptosis, enabling immune evasion,
324 and promoting proliferation, angiogenesis, invasion and metastasis (Whiteside 2008). In
325 our analysis, endothelial cells were significantly higher in adjacent normal tissues (Figure
326 1J), moreover, breast cancer patients with metastasis had higher fraction of mv
327 endothelial cells (Figure 4C) and high level of mv endothelial cells were significantly
328 associated with worse overall survival (Figure 5A). Recent study showed that endothelial
329 cells could promote triple-negative breast cancer cell metastasis via PAI-1 and CCL5
330 signaling (Zhang et al. 2018), and presence of endothelial cells significantly enhanced the
331 angiogenic activity of breast cancer cells (Buchanan et al. 2012). These results suggest
332 that our analysis was largely reliable, and in-depth study of the clinical relevance of these
333 cell types might provide novel insights into the initiation and progression of breast cancer.

334 This was a descriptive analysis of potential roles different tumor-infiltrating cells and
335 the major limitation was that in-depth analysis and experimental validations were greatly
336 needed in the future since a lot of 64 types of cells were profiled in this study.

337 In summary, our study, for the first time, revealed the landscape of cellular

heterogeneity at high resolution and provided novel insights into cell interactions within tumor microenvironment in breast cancer. Development of future therapeutic and predictive strategies should focus on subtypes of immune cells and stromal cells.

Acknowledgements

Not applicable.

Conflicts of interest

The authors declared that there is no conflict of interest.

Funding

This work was supported by National Natural Science Foundation of China under Grant 81902369.

Reference:

Abbas AR, Wolslegel K, Seshasayee D, Modrusan Z, and Clark HF. 2009. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS One* 4:e6098. 10.1371/journal.pone.0006098

Adams S, Gray RJ, Demaria S, Goldstein L, Perez EA, Shulman LN, Martino S, Wang M, Jones VE, Saphner TJ, Wolff AC, Wood WC, Davidson NE, Sledge GW, Sparano JA, and Badve SS. 2014. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. *J Clin Oncol* 32:2959-2966. 10.1200/JCO.2013.55.0491

Anagnostou VK, and Brahmer JR. 2015. Cancer immunotherapy: a future paradigm shift in the treatment of non-small cell lung cancer. *Clin Cancer Res* 21:976-984. 10.1158/1078-0432.CCR-14-1187

Aran D, Hu Z, and Butte AJ. 2017a. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* 18:220. 10.1186/s13059-017-1349-1

Aran D, Hu Z, and Butte AJ. 2017b. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *bioRxiv*. 10.1101/114165

Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, Drake CG, Camacho LH, Kauh J, Odunsi K, Pitot HC, Hamid O, Bhatia S, Martins R, Eaton K, Chen S, Salay TM, Alaparthy S, Grosso JF, Korman AJ, Parker SM, Agrawal S, Goldberg SM, Pardoll DM, Gupta A, and Wigginton JM. 2012. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 366:2455-2465. 10.1056/NEJMoa1200694

Buchanan CF, Szot CS, Wilson TD, Akman S, Metheny-Barlow LJ, Robertson JL, Freeman JW, and Rylander MN. 2012. Cross-talk between endothelial and breast cancer cells regulates reciprocal expression of angiogenic factors in vitro. *J Cell Biochem* 113:1142-1151. 10.1002/jcb.23447

Callahan MK, Wolchok JD, and Allison JP. 2010. Anti-CTLA-4 antibody therapy: immune monitoring during clinical development of a novel immunotherapy. *Semin Oncol* 37:473-484. 10.1053/j.seminoncol.2010.09.001

Cassetta L, and Pollard JW. 2017. Repolarizing macrophages improves breast cancer therapy. *Cell Res* 27:963-964. 10.1038/cr.2017.63

Chung W, Eum HH, Lee HO, Lee KM, Lee HB, Kim KT, Ryu HS, Kim S, Lee JE, Park YH, Kan Z, Han W, and Park WY. 2017. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* 8:15081. 10.1038/ncomms15081

De Palma M, and Lewis CE. 2013. Macrophage regulation of tumor responses to anticancer therapies. *Cancer Cell* 23:277-286. 10.1016/j.ccr.2013.02.013

Dieci MV, Mathieu MC, Guarneri V, Conte P, Delaloge S, Andre F, and Goubar A. 2015. Prognostic and predictive value of tumor-infiltrating lymphocytes in two phase III randomized adjuvant breast cancer trials. *Ann Oncol* 26:1698-1704. 10.1093/annonc/mdv239

Fourcade J, Sun Z, Benallaoua M, Guillaume P, Luescher IF, Sander C, Kirkwood JM, Kuchroo V, and Zarour HM. 2010. Upregulation of Tim-3 and PD-1 expression is associated with tumor antigen-specific CD8+ T cell dysfunction in melanoma patients. *J Exp Med* 207:2175-2186. 10.1084/jem.20100637

Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, Tosolini M, Camus M, Berger A, Wind P, Zinzindohoue F, Bruneval P, Cugnenc PH, Trajanoski Z, Fridman WH, and Pages F. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 313:1960-1964. 10.1126/science.1129139

Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, Diehn M, West RB, Plevritis SK, and Alizadeh AA. 2015. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 21:938-945. 10.1038/nm.3909

Hackl H, Charoentong P, Finotello F, and Trajanoski Z. 2016. Computational genomics tools for dissecting tumour-immune cell interactions. *Nat Rev Genet* 17:441-458. 10.1038/nrg.2016.67

Hanahan D, and Coussens LM. 2012. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21:309-322. 10.1016/j.ccr.2012.02.022

Heagerty PJ, Lumley T, and Pepe MS. 2000. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 56:337-344.

Huang RY, Francois A, McGray AR, Miliotto A, and Odunsi K. 2017. Compensatory upregulation of PD-1, LAG-3, and CTLA-4 limits the efficacy of single-agent checkpoint blockade in metastatic ovarian cancer.

Oncoimmunology 6:e1249561. 10.1080/2162402X.2016.1249561

Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, Rouas G, Francis P, Crown JP, Hitre E, de Azambuja E, Quinaux E, Di Leo A, Michiels S, Piccart MJ, and Sotiriou C. 2013. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J Clin Oncol* 31:860-867. 10.1200/JCO.2011.41.0902

Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, and Alizadeh AA. 2015. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 12:453-457. 10.1038/nmeth.3337

Nirschl CJ, and Drake CG. 2013. Molecular pathways: coexpression of immune checkpoint molecules: signaling pathways and implications for cancer immunotherapy. *Clin Cancer Res* 19:4917-4924. 10.1158/1078-0432.CCR-12-1972

Nishikawa H, and Sakaguchi S. 2014. Regulatory T cells in cancer immunotherapy. *Curr Opin Immunol* 27:1-7. 10.1016/j.coi.2013.12.005

Noy R, and Pollard JW. 2014. Tumor-associated macrophages: from mechanisms to therapy. *Immunity* 41:49-61. 10.1016/j.immuni.2014.06.010

Rooney MS, Shukla SA, Wu CJ, Getz G, and Hacohen N. 2015. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160:48-61. 10.1016/j.cell.2014.12.033

Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, and Chen YT. 2002. Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. *Immunol Rev* 188:22-32.

Schoenfeld JD. 2015. Immunity in head and neck cancer. *Cancer Immunol Res* 3:12-17. 10.1158/2326-6066.CIR-14-0205

Schumacher TN, and Schreiber RD. 2015. Neoantigens in cancer immunotherapy. *Science* 348:69-74. 10.1126/science.aaa4971

Shen-Orr SS, and Gaujoux R. 2013. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* 25:571-578. 10.1016/j.coi.2013.09.015

Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, Chmielowski B, Spasic M, Henry G, Ciobanu V, West AN, Carmona M, Kivork C, Seja E, Cherry G, Gutierrez AJ, Grogan TR, Mateus C, Tomasic G, Glaspy JA, Emerson RO, Robins H, Pierce RH, Elashoff DA, Robert C, and Ribas A. 2014. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515:568-571. 10.1038/nature13954

Whiteside TL. 2008. The tumor microenvironment and its role in promoting tumor growth. *Oncogene* 27:5904-5912. 10.1038/onc.2008.271

Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, Segal NH, Ariyan CE, Gordon RA, Reed K, Burke MM, Caldwell A, Kronenberg SA, Agunwamba BU, Zhang X, Lowy I, Inzunza HD, Feely W, Horak CE, Hong Q, Korman AJ, Wigginton JM, Gupta A, and Sznol M. 2013. Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med* 369:122-133. 10.1056/NEJMoa1302369

Yee C, Thompson JA, Byrd D, Riddell SR, Roche P, Celis E, and Greenberg PD. 2002. Adoptive T cell therapy using antigen-specific CD8+ T cell clones for the treatment of patients with metastatic melanoma: in vivo persistence, migration, and antitumor effect of transferred T cells. *Proc Natl Acad Sci U S A* 99:16168-16173. 10.1073/pnas.242600099

Zhang W, Xu J, Fang H, Tang L, Chen W, Sun Q, Zhang Q, Yang F, Sun Z, Cao L, Wang Y, and Guan X. 2018. Endothelial cells promote triple-negative breast cancer cell metastasis via PAI-1 and CCL5 signaling. *FASEB*

445 *J* 32:276-288. 10.1096/fj.201700237RR

446

447

448

449 **Table 1. Descriptions of 64 cell types in this study.**

Full name	Abbreviations	Group
Activated dendritic cells	aDC	Immune
B-cells	/	Immune
Basophils	/	Immune
CD4+ naive T-cells	/	Immune
CD4+ T-cells	/	Immune
Central memory CD4+ T Cell	CD4+ Tcm	Immune
Effector memory CD4+ T cell	CD4+ Tem	Immune
CD8+ naive T-cells	/	Immune
CD8+ T-cells	/	Immune
Central memory CD8+ T Cell	CD8+ Tcm	Immune
Effector memory CD8+ T cell	CD8+ Tem	Immune
Conventional dendritic cells	cDC	Immune
Class-switched memory B-cells	/	Immune
Dendritic cells	DC	Immune
Eosinophils	/	Immune
Immature dendritic cells	iDC	Immune
Macrophages	/	Immune
Inflammatory (M1) macrophages	Macrophages M1	Immune
Reparative (M2) macrophages	Macrophages M2	Immune
Mast cells	/	Immune
Memory B-cells	/	Immune
Monocytes	/	Immune
naive B-cells	/	Immune
Neutrophils	/	Immune
Nature killer cells	NK cells	Immune
Natural killer T cells	NKT	Immune
Plasmacytoid dendritic cells	pDC	Immune
Plasma cells	/	Immune
pro B-cells	/	Immune
Gamma delta T cells	Tgd cells	Immune
Regulatory T cells	Tregs	Immune
Type 1 T helper (Th1) cells	Th1 cells	Immune
Type 2 T helper (Th2) cells	Th2 cells	Immune
CD4+ memory T-cells	/	Immune
Astrocytes	/	Others
Epithelial cells	/	Others
Hepatocytes	/	Others
Keratinocytes	/	Others

Full name	Abbreviations	Group
Melanocytes	/	Others
Mesangial cells	/	Others
Neurons	/	Others
Sebocytes	/	Others
Common lymphoid progenitor	CLP	Stem
Common myeloid progenitor	CMP	Stem
Granulocyte-macrophage progenitor	GMP	Stem
Hematopoietic stem cells	HSC	Stem
Megakaryocytes	/	Stem
Multipotent progenitors	MPP	Stem
Erythrocytes	/	Stem
Megakaryocyte-erythroid progenitor	MEP	Stem
Platelets	/	Stem
Adipocytes	/	Stromal
Chondrocytes	/	Stromal
Endothelial cells	/	Stromal
Fibroblasts	/	Stromal
Lymphatic endothelial cells	ly Endothelial cells	Stromal
Mesenchymal stem cells	MSC	Stromal
Microvascular endothelial cells	mv Endothelial cells	Stromal
Myocytes	/	Stromal
Osteoblast	/	Stromal
Pericytes	/	Stromal
Preadipocytes	/	Stromal
Skeletal muscle	/	Stromal
Smooth muscle	/	Stromal

Figure legends:

Figure 1. Differences of cellular heterogeneity between breast tumor tissue and normal tissues. A, Median fraction of 64 cell types in breast tumor and normal tissues. 64 cell types were grouped into four groups: Immune, stem, stromal and other cells. B, Heatmap of fractions of 64 cell types in 1,092 breast tumor tissues and 112 adjacent normal tissues. C, Dimensionality reduction and visualization by t-Distributed Stochastic

457 Neighbor Embedding (t-SNE) clustering. D to H, Dot plots of fractions of certain cell types
458 in breast tumor and normal tissues. Lines between dots indicated paired tissues from the
459 same breast cancer patient. *, P<0.05. **, P<0.01. ***, P<0.001.
460
461 **Figure 2. Expression patterns of inhibitory receptors on CD4+/CD8+ T cells.** A and
462 B, Heatmaps of expression of inhibitory receptors, including PD1, CTLA4, LAG3 and
463 TIM3, and fractions of CD4+/CD8+ T cells in tumor tissues (A) and normal tissues (B).
464 Data were transformed by rank. C and D, Clustered correlation matrixes among inhibitory
465 receptors and CD4+/CD8+ T cells in tumor tissues (C) and normal tissues (D). E, Dot plot
466 of correlations between PD1 expression and fractions of CD8+ Tcm in tumor tissues. F,
467 Dot plot of correlations between PD1 expression and CTLA4 expression in tumor tissues.
468 G, Dot plot of correlations between TIM3 expression and fractions of CD4+ Tcm in normal
469 tissues. H. Dot plot of correlations between PD1 expression and LAG3 in normal tissues.
470
471 **Figure 3. Correlations between cancer/testis genes and immune cells.** A, Significant
472 correlations between cancer/testis (CT) genes and immune cells. Scaled color dots
473 represented significant correlations between CT genes and immune cells (P<0.001) and
474 red dots represented positive correlations while blue dots represent negative correlations.
475 B and C, CD8+ naïve T-cells and aDC were positively correlated with most of the CT
476 genes. D and E, TDRD6 and TTK were positively correlated with a number of immune
477 cells.

478

479 **Figure 4. Involvement of cellular heterogeneity in clinic-pathology of ESCC.** A, A
480 number of cell types were significantly correlated with clinical parameters. B to F,
481 Examples of significant correlations between different cell types and clinical parameters.
482 *, P<0.05. **, P<0.01. ***, P<0.001.

483

484 **Figure 5. Survival associated tumor-infiltrating cells in breast cancer.** A, Forrest plot
485 of hazard ratios of survival associated cell types. B to M, Kaplan-Meier curves of survival
486 associated cell types. Red lines indicated high fraction while blue lines indicated low
487 fraction of each cell types, respectively. N, Kaplan-Meier curves of predictor built with
488 significant prognostic factors (P<0.01). O, ROC curves of prognostic predictors.

489

490 **Figure 6. Differences of cellular heterogeneity among different subtypes of breast**
491 **cancer.** A, Median fraction of 64 cell types in five subtypes of breast tumor. 1,092 breast
492 cancer patients were classified into five groups, which including 30 Her2+_HR- patients,
493 59 Her2+_HR+ patients, 426 Luminal type (Her2-_HR+) patients, 97 triple negative (Tri-
494 negative) patients, and 480 unknown patients. B, Cluster analysis by ImmuneScore and
495 StromalScore, which were calculated by summing up the fractions of immune and stromal
496 cells, respectively. C, Dimensionality reduction and visualization by t-Distributed
497 Stochastic Neighbor Embedding (t-SNE) clustering. D, Heatmap of fractions of 64 cell
498 types in five subtypes of breast cancer. E and F, Box plots with dots of fractions of certain

499 cell types in five subtypes of breast cancer. *, P<0.05. **, P<0.01. ***, P<0.001, ****,
500 P<0.0001.

501

502 **Figure 7. Survival associated tumor-infiltrating cells in five subtypes of breast**
503 **cancer.** A to E, Forrest plots of hazard ratios of survival associated cell types in five
504 subtypes of breast cancer.

505

506 **Supplementary Figure S1. Median fractions of 64 types of cells in five subtypes of**
507 **breast cancer.**

508

509 **Supplementary Figure S2. Diverse differences of 64 types of cells among the five**
510 **subtypes of breast cancer.**

511

512

513

Figure 1

Differences of cellular heterogeneity between breast tumor tissue and normal tissues.

A, Median fraction of 64 cell types in breast tumor and normal tissues. 64 cell types were grouped into four groups: Immune, stem, stromal and other cells. B, Heatmap of fractions of 64 cell types in 1,092 breast tumor tissues and 112 adjacent normal tissues. C, Dimensionality reduction and visualization by t-Distributed Stochastic Neighbor Embedding (t-SNE) clustering. D to H, Dot plots of fractions of certain cell types in breast tumor and normal tissues. Lines between dots indicated paired tissues from the same breast cancer patient. *, $P<0.05$. **, $P<0.01$. ***, $P<0.001$.

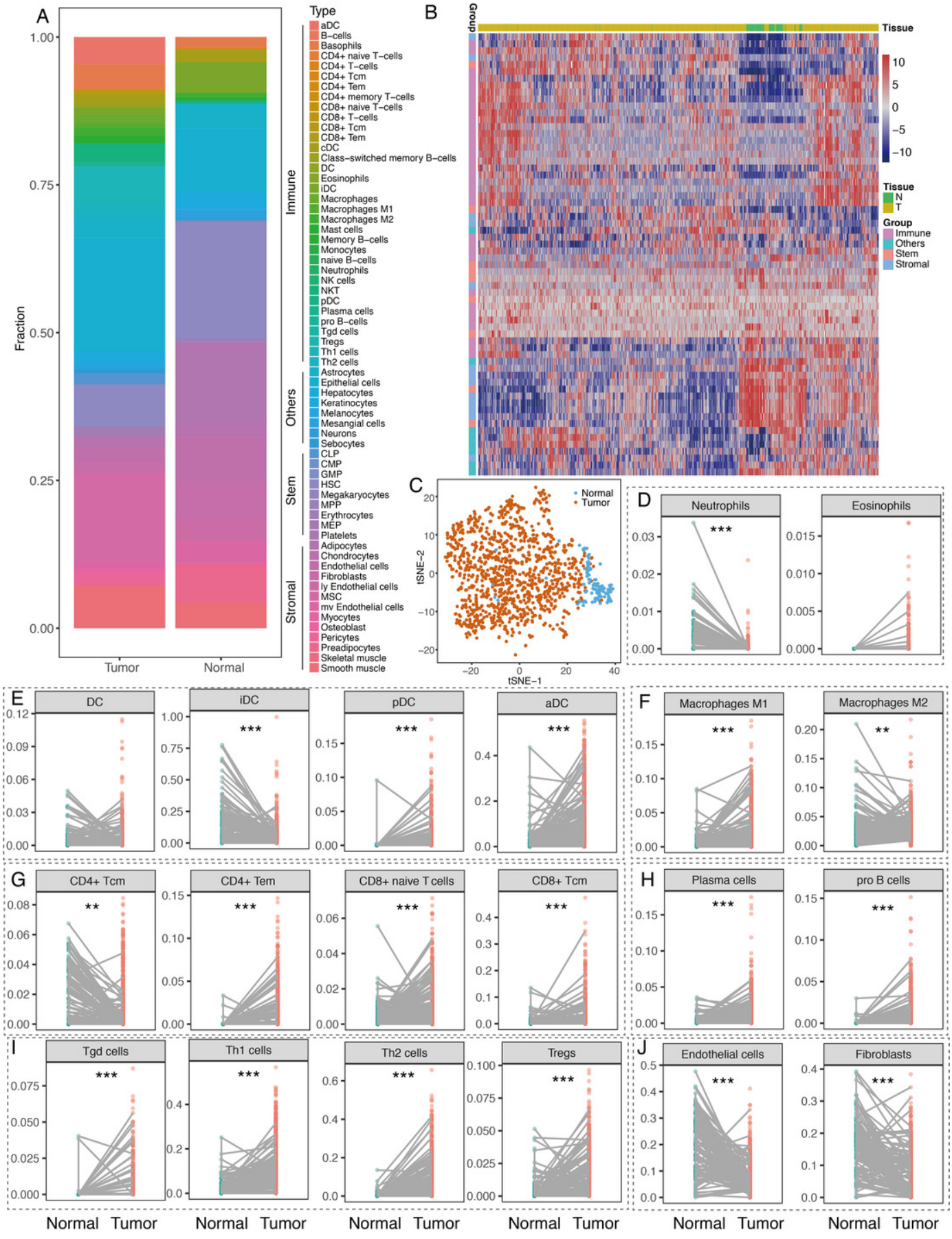
A
Fraction
1.00
0.75
0.50
0.25
0.00
Tumor
Normal
Type
aDC
B−cells
Basophils
CD4+ naive T−cells
CD4+ T−cells
CD4+ Tcm
CD4+ Tem
CD4+ memory T−cells
CD8+ naive T−cells
CD8+ T−cells
CD8+ Tcm
CD8+ Tem
cDC
Class−switched memory B−cells
DC
Eosinophils
iDC
Macrophages
Macrophages M1
Macrophages M2
Mast cells
Memory B−cells
Monocytes
naive B−cells
Neutrophils
NK cells
NKT
pDC
Plasma cells
pro B−cells
Tgd cells
Tregs
Th1 cells
Th2 cells
Astrocytes
Epithelial cells
Hepatocytes
Keratinocytes
Melanocytes
Mesangial cells
Neurons
Sebocytes
CLP
CMP
GMP
HSC
Megakaryocytes
MPP
Erythrocytes
MEP
Platelets
Adipocytes
Chondrocytes
Endothelial cells
Fibroblasts
ly Endothelial cells
MSC
mv Endothelial cells
Myocytes
Osteoblast
Pericytes
Preadipocytes
Skeletal muscle
Smooth muscle
Immune
Others
Stem
Stromal
B
Group
Tissue
10
5
0
−5
−10
Tissue
N
T
Group
Immune
Others
Stem
Stromal
C
Normal
Tumor
tSNE−2
tSNE−1
D
Neutrophils
Eosinophils
E
DC
iDC
pDC
aDC
F
Macrophages M1
Macrophages M2
G
CD4+ Tcm
CD4+ Tem
CD8+ naive T cells
CD8+ Tcm
H
Plasma cells
pro B cells
I
Tgd cells
Th1 cells
Th2 cells
Tregs
J
Endothelial cells
Fibroblasts
Normal
Tumor

Figure 2

Expression patterns of inhibitory receptors on CD4+/CD8+ T cells.

A and B, Heatmaps of expression of inhibitory receptors, including PD1, CTLA4, LAG3 and TIM3, and fractions of CD4+/CD8+ T cells in tumor tissues (A) and normal tissues (B). Data were transformed by rank. C and D, Clustered correlation matrixes among inhibitory receptors and CD4+/CD8+ T cells in tumor tissues (C) and normal tissues (D). E, Dot plot of correlations between PD1 expression and fractions of CD8+ Tcm in tumor tissues. F, Dot plot of correlations between PD1 expression and CTLA4 expression in tumor tissues. G, Dot plot of correlations between TIM3 expression and fractions of CD4+ Tcm in normal tissues. H. Dot plot of correlations between PD1 expression and LAG3 in normal tissues.

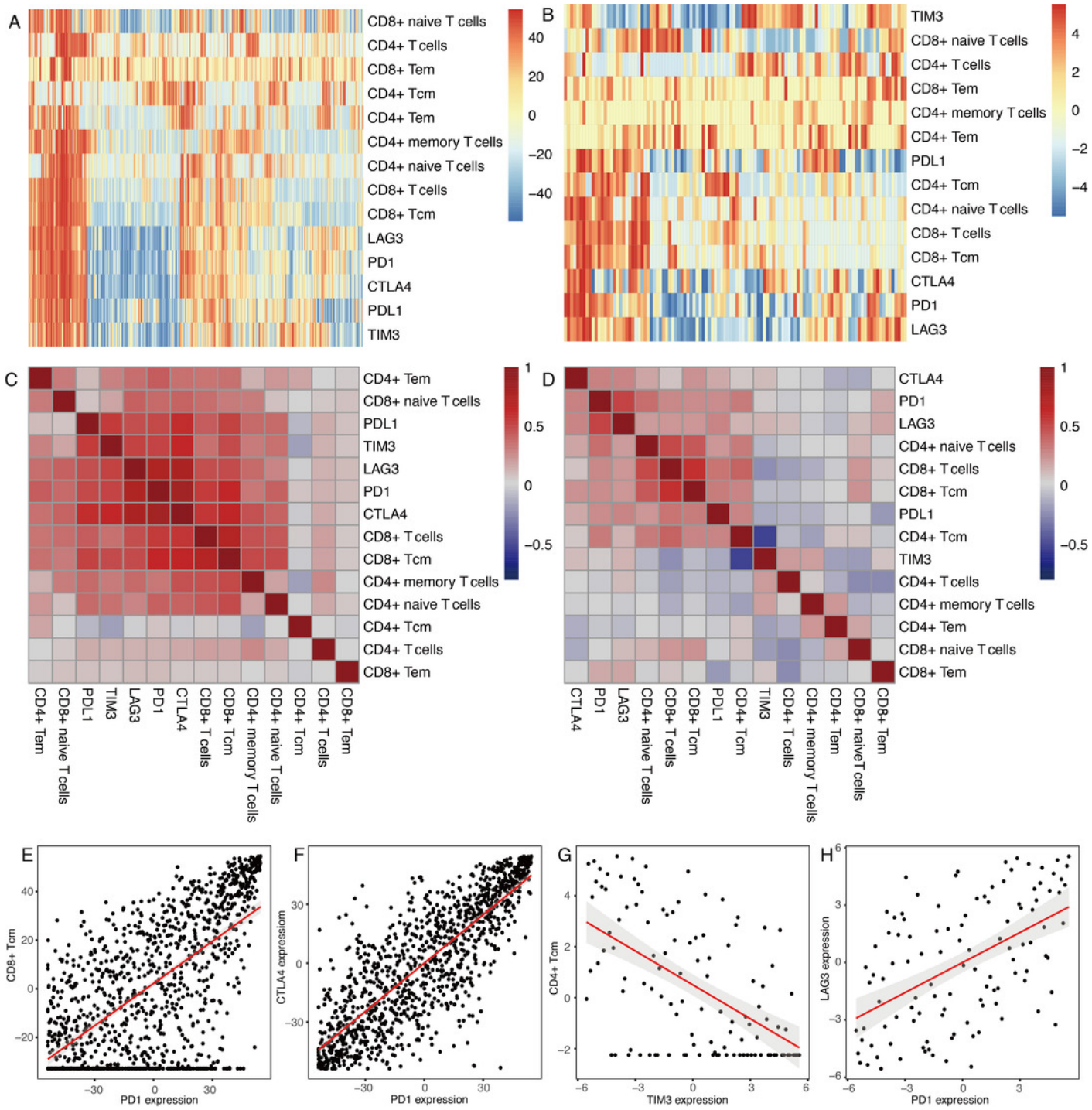
A
CD8+ naive T cells
CD4+ T cells
CD8+ Tem
CD4+ Tcm
CD4+ Tem
CD4+ memory T cells
CD4+ naive T cells
CD8+ T cells
CD8+ Tcm
LAG3
PD1
CTLA4
PDL1
TIM3
40
20
0
−20
−40
B
TIM3
CD8+ naive T cells
CD4+ T cells
CD8+ Tem
CD4+ memory T cells
CD4+ Tem
PDL1
CD4+ Tcm
CD4+ naive T cells
CD8+ T cells
CD8+ Tcm
CTLA4
PD1
LAG3
4
2
0
−2
−4
C
CD4+ Tem
CD8+ naive T cells
PDL1
TIM3
LAG3
PD1
CTLA4
CD8+ T cells
CD8+ Tcm
CD4+ memory T cells
CD4+ naive T cells
CD4+ Tcm
CD4+ T cells
CD8+ Tem
1
0.5
0
−0.5
D
CTLA4
PD1
LAG3
CD4+ naive T cells
CD8+ T cells
CD8+ Tcm
PDL1
CD4+ Tcm
TIM3
CD4+ T cells
CD4+ memory T cells
CD4+ Tem
CD8+ naive T cells
CD8+ Tem
1
0.5
0
−0.5
E
CD8+ Tcm
PD1 expression
F
CTLA4 expression
PD1 expression
G
CD4+ Tcm
TIM3 expression
H
LAG3 expression
PD1 expression

Figure 3

Correlations between cancer/testis genes and immune cells

A, Significant correlations between cancer/testis (CT) genes and immune cells. Scaled color dots represented significant correlations between CT genes and immune cells ($P<0.001$) and red dots represented positive correlations while blue dots represent negative correlations. B and C, CD8+ naïve T-cells and aDC were positively correlated with most of the CT genes. D and E, TDRD6 and TTK were positively correlated with a number of immune cells.

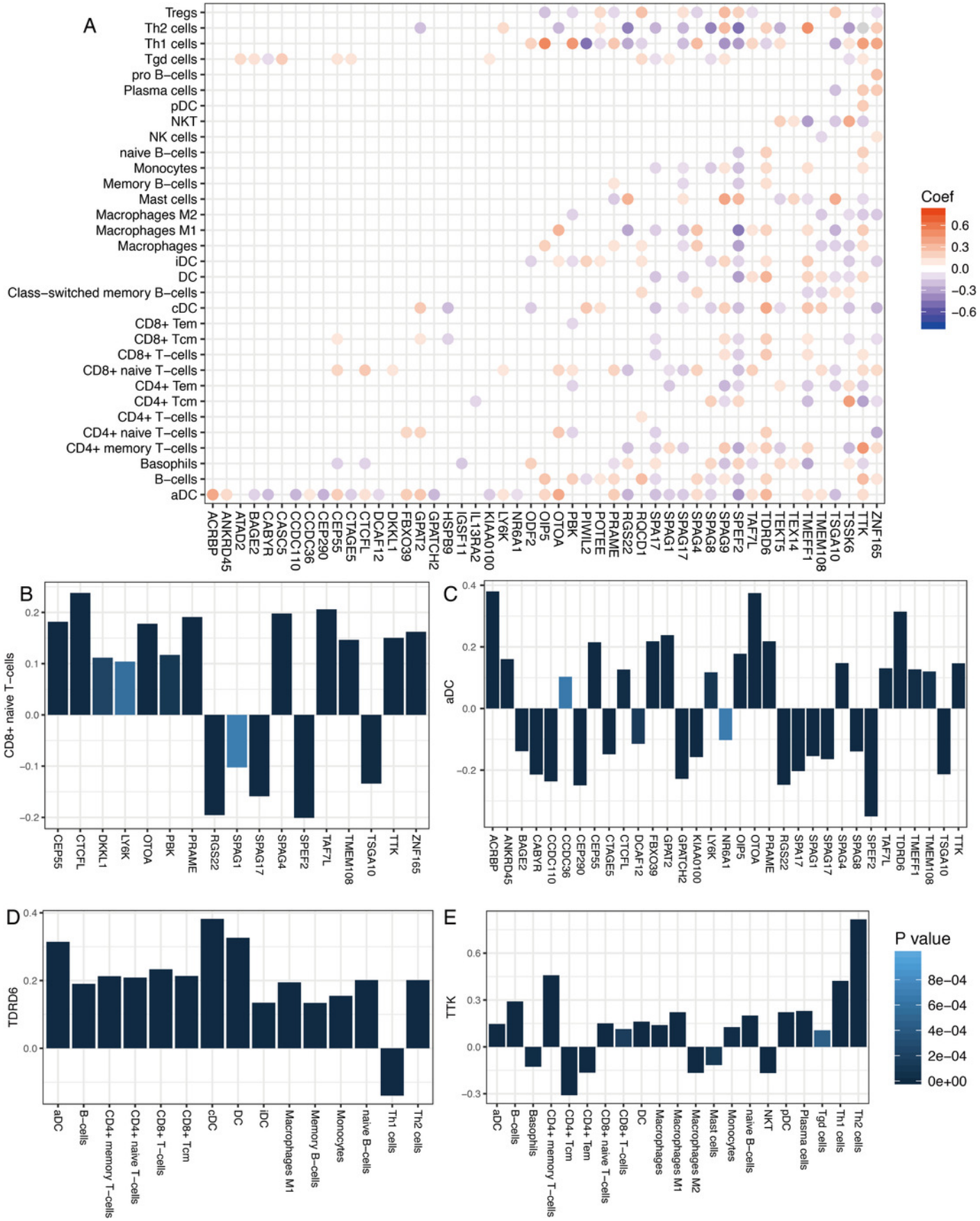
A
Tregs
Th2 cells
Th1 cells
Tgd cells
pro B−cells
Plasma cells
pDC
NKT
NK cells
naive B−cells
Monocytes
Memory B−cells
Mast cells
Macrophages M2
Macrophages M1
Macrophages
iDC
DC
Class−switched memory B−cells
cDC
CD8+ Tem
CD8+ Tcm
CD8+ T−cells
CD8+ naive T−cells
CD4+ Tem
CD4+ Tcm
CD4+ T−cells
CD4+ naive T−cells
CD4+ memory T−cells
Basophils
B−cells
aDC
ACRBP ANKRD45 ATAD2 BAGE2 CABYR CASC5 CCDC110 CCDC36 CEP290 CEP55 CTAGE5 CTCFL DCAF12 DKKL1 FBXO39 GPAT2 GPATCH2 HSPB9 IGSF11 IL13RA2 KIAA0100 LY6K NR6A1 ODF2 OIP5 OTOA PBK PIWIL2 POTEE PRAME RGS22 RQCD1 SPA17 SPAG1 SPAG17 SPAG4 SPAG8 SPAG9 SPEF2 TAF7L TDRD6 TEKT5 TEX14 TMEFF1 TMEM108 TSGA10 TSSK6 TTK ZNF165
Coef
0.6
0.3
0.0
−0.3
−0.6
B
CD8+ naive T−cells
CEP55 CTCFL DKKL1 LY6K OTOA PBK PRAME RGS22 SPAG1 SPAG17 SPAG4 SPEF2 TAF7L TMEM108 TSGA10 TTK ZNF165
C
aDC
ACRBP ANKRD45 BAGE2 CABYR CCDC110 CCDC36 CEP290 CEP55 CTAGE5 CTCFL DCAF12 FBXO39 GPAT2 GPATCH2 KIAA0100 LY6K NR6A1 OIP5 OTOA PRAME RGS22 SPA17 SPAG1 SPAG17 SPAG4 SPAG8 SPEF2 TAF7L TDRD6 TMEFF1 TMEM108 TSGA10 TTK
D
TDRD6
aDC B−cells CD4+ memory T−cells CD4+ naive T−cells CD8+ T−cells CD8+ Tcm cDC DC iDC Macrophages M1 Memory B−cells Monocytes naive B−cells Th1 cells Th2 cells
E
TTK
aDC B−cells Basophils CD4+ memory T−cells CD4+ Tcm CD4+ Tem CD8+ naive T−cells CD8+ T−cells DC Macrophages Macrophages M1 Macrophages M2 Mast cells Monocytes naive B−cells NKT pDC Plasma cells Tgd cells Th1 cells Th2 cells
P value
8e−04
6e−04
4e−04
2e−04
0e+00

Figure 4

Involvement of cellular heterogeneity in clinic-pathology of ESCC

A, A number of cell types were significantly correlated with clinical parameters. B to F, Examples of significant correlations between different cell types and clinical parameters. *, P<0.05. **, P<0.01. ***, P<0.001.

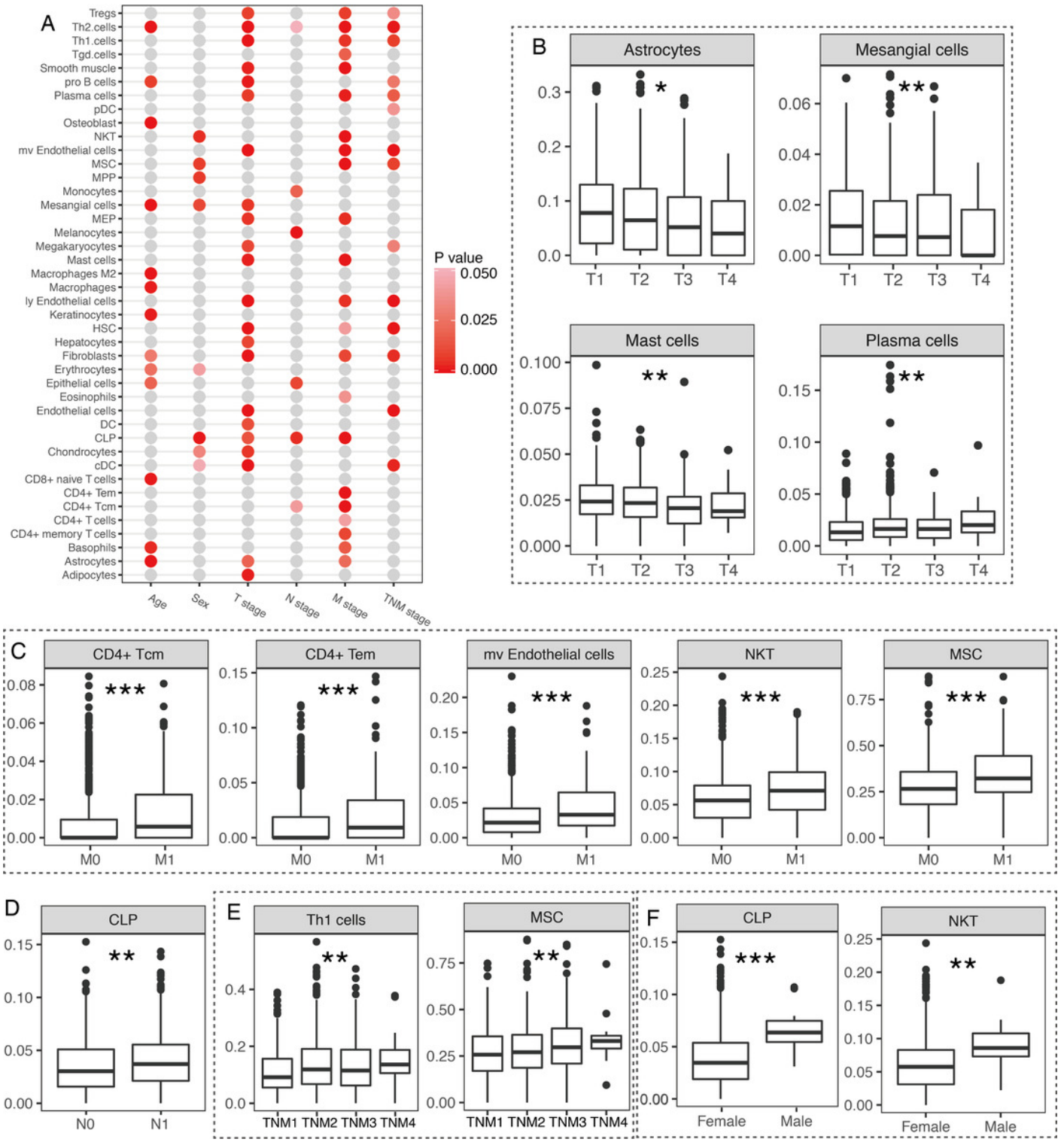
A
Tregs
Th2.cells
Th1.cells
Tgd.cells
Smooth muscle
pro B cells
Plasma cells
pDC
Osteoblast
NKT
mv Endothelial cells
MSC
MPP
Monocytes
Mesangial cells
MEP
Melanocytes
Megakaryocytes
Mast cells
Macrophages M2
Macrophages
ly Endothelial cells
Keratinocytes
HSC
Hepatocytes
Fibroblasts
Erythrocytes
Epithelial cells
Eosinophils
Endothelial cells
DC
CLP
Chondrocytes
cDC
CD8+ naive T cells
CD4+ Tem
CD4+ Tcm
CD4+ T cells
CD4+ memory T cells
Basophils
Astrocytes
Adipocytes
Age
Sex
T stage
N stage
M stage
TNM stage
P value
0.050
0.025
0.000
B
Astrocytes
Mesangial cells
Mast cells
Plasma cells
T1
T2
T3
T4
C
CD4+ Tcm
CD4+ Tem
mv Endothelial cells
NKT
MSC
M0
M1
D
CLP
N0
N1
E
Th1 cells
MSC
TNM1
TNM2
TNM3
TNM4
F
CLP
NKT
Female
Male

Figure 5

Survival associated tumor-infiltrating cells in breast cancer.

A, Forrest plot of hazard ratios of survival associated cell types. B to M, Kaplan-Meier curves of survival associated cell types. Red lines indicated high fraction while blue lines indicated low fraction of each cell types, respectively. N, Kaplan-Meier curves of predictor built with significant prognostic factors (P<0.01). O, ROC curves of prognostic predictors.

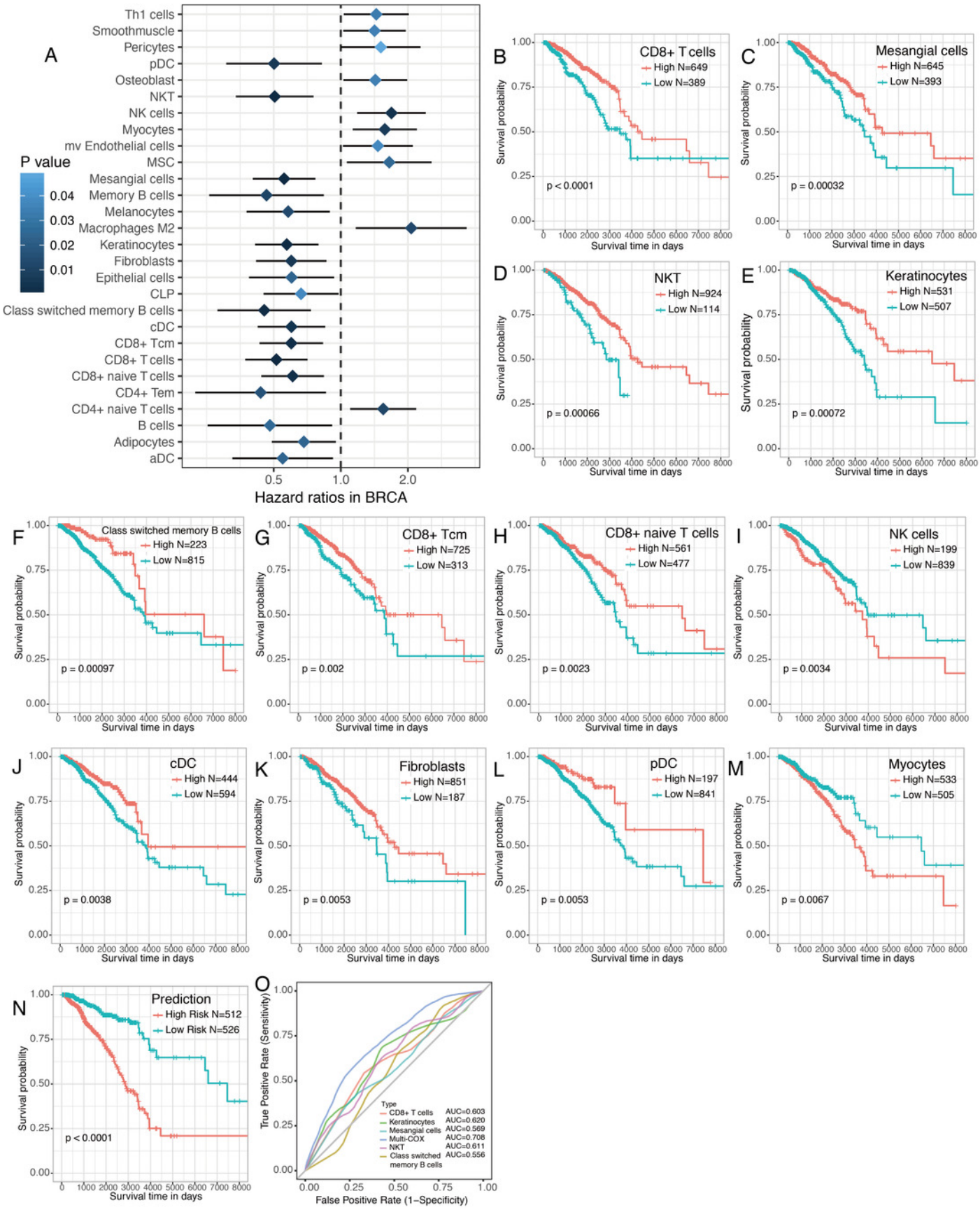
A
Th1 cells
Smoothmuscle
Pericytes
pDC
Osteoblast
NKT
NK cells
Myocytes
mv Endothelial cells
MSC
Mesangial cells
Memory B cells
Melanocytes
Macrophages M2
Keratinocytes
Fibroblasts
Epithelial cells
CLP
Class switched memory B cells
cDC
CD8+ Tcm
CD8+ T cells
CD8+ naive T cells
CD4+ Tem
CD4+ naive T cells
B cells
Adipocytes
aDC
P value
0.04
0.03
0.02
0.01
0.5
1.0
2.0
Hazard ratios in BRCA
B
CD8+ T cells
High N=649
Low N=389
p < 0.0001
C
Mesangial cells
High N=645
Low N=393
p = 0.00032
D
NKT
High N=924
Low N=114
p = 0.00066
E
Keratinocytes
High N=531
Low N=507
p = 0.00072
F
Class switched memory B cells
High N=223
Low N=815
p = 0.00097
G
CD8+ Tcm
High N=725
Low N=313
p = 0.002
H
CD8+ naive T cells
High N=561
Low N=477
p = 0.0023
I
NK cells
High N=199
Low N=839
p = 0.0034
J
cDC
High N=444
Low N=594
p = 0.0038
K
Fibroblasts
High N=851
Low N=187
p = 0.0053
L
pDC
High N=197
Low N=841
p = 0.0053
M
Myocytes
High N=533
Low N=505
p = 0.0067
N
Prediction
High Risk N=512
Low Risk N=526
p < 0.0001
O
True Positive Rate (Sensitivity)
False Positive Rate (1−Specificity)
Type
CD8+ T cells AUC=0.603
Keratinocytes AUC=0.620
Mesangial cells AUC=0.569
Multi-COX AUC=0.708
NKT AUC=0.611
Class switched memory B cells AUC=0.556
Survival probability
Survival time in days

Figure 6

Differences of cellular heterogeneity among different subtypes of breast cancer.

A, Median fraction of 64 cell types in five subtypes of breast tumor. 1,092 breast cancer patients were classified into five groups, which including 30 Her2+_HR- patients, 59 Her2+_HR+ patients, 426 Luminal type (Her2-_HR+) patients, 97 triple negative (Tri-negative) patients, and 480 unknown patients. B, Cluster analysis by ImmuneScore and StromalScore, which were calculated by summing up the fractions of immune and stromal cells, respectively. C, Dimensionality reduction and visualization by t-Distributed Stochastic Neighbor Embedding (t-SNE) clustering. D, Heatmap of fractions of 64 cell types in five subtypes of breast cancer. E and F, Box plots with dots of fractions of certain cell types in five subtypes of breast cancer. *, $P<0.05$. **, $P<0.01$. ***, $P<0.001$, ****, $P<0.0001$.

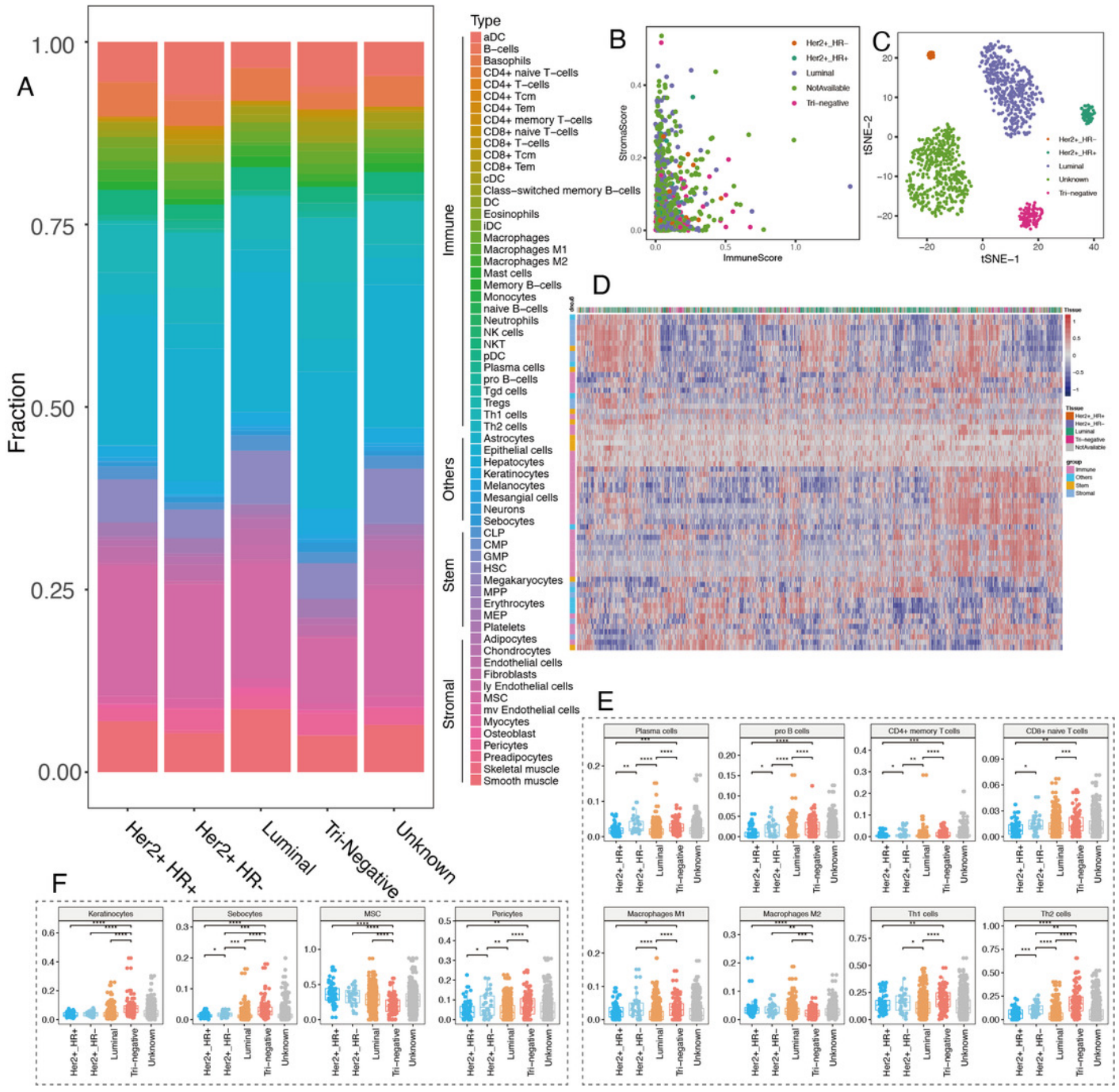
A
Fraction
1.00
0.75
0.50
0.25
0.00
Her2+ HR+
Her2+ HR-
Luminal
Tri-Negative
Unknown
Type
Immune
Others
Stem
Stromal
aDC
B−cells
Basophils
CD4+ naive T−cells
CD4+ T−cells
CD4+ Tcm
CD4+ Tem
CD4+ memory T−cells
CD8+ naive T−cells
CD8+ T−cells
CD8+ Tcm
CD8+ Tem
cDC
Class−switched memory B−cells
DC
Eosinophils
iDC
Macrophages
Macrophages M1
Macrophages M2
Mast cells
Memory B−cells
Monocytes
naive B−cells
Neutrophils
NK cells
NKT
pDC
Plasma cells
pro B−cells
Tgd cells
Tregs
Th1 cells
Th2 cells
Astrocytes
Epithelial cells
Hepatocytes
Keratinocytes
Melanocytes
Mesangial cells
Neurons
Sebocytes
CLP
CMP
GMP
HSC
Megakaryocytes
MPP
Erythrocytes
MEP
Platelets
Adipocytes
Chondrocytes
Endothelial cells
Fibroblasts
ly Endothelial cells
MSC
mv Endothelial cells
Myocytes
Osteoblast
Pericytes
Preadipocytes
Skeletal muscle
Smooth muscle
B
StromaScore
ImmuneScore
Her2+_HR−
Her2+_HR+
Luminal
NotAvailable
Tri−negative
C
tSNE−2
tSNE−1
Her2+_HR−
Her2+_HR+
Luminal
Unknown
Tri−negative
D
Tissue
group
Immune
Others
Stem
Stromal
E
Plasma cells
pro B cells
CD4+ memory T cells
CD8+ naive T cells
Macrophages M1
Macrophages M2
Th1 cells
Th2 cells
F
Keratinocytes
Sebocytes
MSC
Pericytes
Her2+_HR+
Her2+_HR−
Luminal
Tri−negative
Unknown

Figure 7

Survival associated tumor-infiltrating cells in five subtypes of breast cancer.

A to E, Forrest plots of hazard ratios of survival associated cell types in five subtypes of breast cancer.