LETTER TO THE EDITOR

Single-cell and spatially resolved analysis uncovers cell heterogeneity of breast cancer

Si-Qing Liu1†, Zhi-Jie Gao1†, Juan Wu2†, Hong-Mei Zheng3, Bei Li2, Si Sun4*, Xiang-Yu Meng5* and Qi Wu6*

Abstract

The heterogeneity and the complex cellular architecture have a crucial effect on breast cancer progression and response to treatment. However, deciphering the neoplastic subtypes and their spatial organization is still challenging. Here, we combine single-nucleus RNA sequencing (snRNA-seq) with a microarray-based spatial transcriptomics (ST) to identify cell populations and their spatial distribution in breast cancer tissues. Malignant cells are clustered into distinct subpopulations. These cell clusters not only have diverse features, origins and functions, but also emerge to the crosstalk within subtypes. Furthermore, we find that these subclusters are mapped in distinct tissue regions, where discrepant enrichment of stromal cell types are observed. We also inferred the abundance of these tumorous subpopulations by deconvolution of large breast cancer RNA-seq cohorts, revealing differential association with patient survival and therapeutic response. Our study provides a novel insight for the cellular architecture of breast cancer and potential therapeutic strategies.

Keywords: Single-nucleus RNA sequencing, Spatial transcriptomics, Breast cancer, Tissue architecture, Heterogeneity
and MYC TARGETS), and this cluster exhibited a high metabolic state with the activation of glycolysis, fatty acid and cholesterol metabolism potentially regulated by PI3K/AKT/mTOR signaling, hypoxia and MYC (Fig. 1f, Additional file 1: Figure S1B-C). In contrast, genomic unstable luminal-A (LumA) cells only showed an activated KRAS-downregulated signaling (Fig. 1f). Next, we asked whether transcription factors (TFs) contribute to the phenotypic state of these subpopulations (Fig. 1g). First, GATA3, a key mediator during the differentiation from luminal progenitor to mature luminal cells, was active in these luminal clusters. AP-1 transcription factor complexes, including JUN, JUNB, JUND, FOS and FOSB, were enriched in ERBB4+AREG+ LumA, genomic unstable LumA, EGFR+KIT+ basal and cycling LumB clusters. Likewise, ERBB4+AREG+ LumA cluster also highly expressed SREBF1/2 which are crucial for controlling cholesterol homeostasis [6]. Pseudotime analysis revealed trajectories of luminal and basal/myoepithelial lineages (Fig. 1h). The six major cell clusters were corroborated by typical mammary lineage marker genes for basal progenitor/myoepithelium (e.g., MYLK and ACTA2), mature luminal cells (e.g., ESR1, FOXA1, PGR and AFF3) and luminal progenitor cells (e.g., EZH2) (Additional file 1: Figure S3C-D) [7]. Moreover, the myoepithelial cells stemmed from basal progenitor cells, while luminal progenitor cells gradually differentiated into mature luminal cells (Fig. 1h). Interestingly, a cluster of luminal cells expressed certain level of basal-like genes, suggesting that they may emerge as a dedifferentiated state or represent invasiveness. EGFR+TP63+ normal-like subtype was fully distributed in the left arm, while cycling LumB cluster was completely situated in the arm of the luminal progenitor (Fig. 1h). Then, we investigated the interactions between ligands and receptors across all neoplastic populations. Interestingly, ERBB4+AREG+ LumA cells secreted increasing EGF and EGF-like cytokine-like AREG, which could combine with EGF receptors such as EGFR and ERBB4 (Fig. 1i, Additional file 1: Figure S1D) (Additional file 2: Result S2).

Subsequently, we sought to annotate the intratumoral heterogeneity by integrating the snRNA-seq and ST datasets [5]. The ST dataset was classified into six main areas: the luminal region, basal region, the interfacing area between the luminal region and basal region, stroma and infiltrating lymphocyte areas based on the principal component scores across all ST spots (Fig. 2a, b, Additional file 1: S5C). We found that these luminal subtypes were principally assembled in the luminal region and the interfacing area except that the cycling LumB cells were scattered throughout the tissue. Interestingly, the ERBB4+ LumA type scarcely existed in the interfacing area (Fig. 2a, b, Additional file 1: S5B). Finally, the spatial division of the primary nontumor cells was identified in the ST dataset, showing that neutrophils were enriched in the luminal region, while B cells principally infiltrated into the basal region (Fig. 2c) (Additional file 2: Result S3).

We deconvoluted breast cancer profiles in the METABRIC cohort [8]. Based on the defined gene signature, all of cases were divided into three discriminated subgroups (DSGs). And DSG2, representing the HER2-over-expressing subtypes based on PAM50 was designated the genomic unstable LumA subtypes (Fig. 2d). For survival analysis, patients representing the DSG2 subtype were associated with the worst survival (Fig. 2e). Through examining public datasets of patients treated with chemotherapy[9], the proportions of the ERBB4+ LumA cells or EGFR+TP63+ normal-like cells were associated with poor response to neoadjuvant chemotherapy (NAC) while the ratio of the cycling LumB cluster significantly increased in patients who response to NAC (Fig. 2f) (Additional file 2: Result S4, Additional file 3).

In summary, we have provided a comprehensive approach to depict the heterogeneity and the architecture of breast cancer. Our study could provide novel insights...
Fig. 1 (See legend on previous page.)
Fig. 2  ST analysis of BC-A and cell type deconvolution. a Scaled deconvolution values for six epithelial subclusters overlaid onto tissue spots. b Different spatial distribution of six groups as defined in a. c Heatmap of estimated scores of immune cells in each group by MCPcounter. d Heatmap of the ssGSEA score of six gene signatures estimated in each METABRIC sample. e Kaplan–Meier survival curve for METABRIC cohort in three groups. P value was calculated with log-rank test. Log-rank p value < 0.05 was considered as statistically significant. f Box plot of the estimated proportion of six gene signatures in two NAC cohorts using CIBERSORTx. Statistical significance was determined using a two-sided t-test in a pairwise comparison of means between groups, with P values adjusted using the Benjamini–Hochberg procedure. *P < 0.05, **P < 0.005, ***P < 0.005
into the ecosystem of breast cancer and novel therapeutic strategies.

Abbreviations
snRNA-seq: Single-nucleus RNA sequencing; ST: Spatial transcriptomics; BC: Breast cancer; CNV: Copy number variation; LUMA: Luminal-A; LUMB: Luminal-B; AREG: Amphiregulin; EGF: Epidermal growth factor; TGFα: Transforming growth factor alpha; EMT: Epithelial–mesenchymal transition; PTN: Pleiotrophin; EFNAs: Ephrin; TFs: Transcription factors; LTF: Lactotransferrin; ECM: Extracellular matrix; H&E: Hematoxylin and eosin; DCIS: Ductal carcinoma in situ, IHC: Immunohistochemistry, ER: Estrogen receptor; PR: Progesterone receptor; DSG: Discriminated subgroups; NAC: Neoadjuvant chemotherapy; pCR: Pathologic complete response; RD: Residual disease.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s13045-022-01236-0.

Acknowledgements
We thank a professional language editing by American Journal Experts.

Authors’ contributions
QW contributed to conceptualization, XJM and JW were involved in methodology, SQL, ZJG and XYM wrote and prepared the original draft, SQL and ZJG took part in writing, reviewing and editing, HMZ conducted the supervision, BL participated in project administration, and SS acquired the funding. All authors read and approved the final manuscript.

Funding
This work was supported by a National Natural Science Foundation of China (NSFC) grant (Grant No: 81417181) to Prof. Sheng-Rong Sun and (Grant No: 81903166) to Dr. Si Sun. This work was also supported by a Hubei Province health and family planning scientific research project to Dr. Si Sun (Grant No: WJ2019Q044) and a project (W2019Q053) to Dr. Hong-Mei Zheng. And this work was supported by a Fundamental Research Funds for the Central Universities of China grant (Grant No: 413000480) to Dr. Qi Wu. Finally, this work was supported by the Natural Science Foundation of Hubei Province (Grant No. 2018CKXB16) to Dr. Si Sun and a project (2020CFB874) to Dr. Hong-Mei Zheng. Dr. Xiang-Yu Meng was supported by a fellowship from ITMO Cancer AVIESAN within the framework of Cancer Plan.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Department of Breast Surgery, Hubei Cancer Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, People’s Republic of China. 2 Department of Pathology, Renmin Hospital of Wuhan University, Wuhan, Hubei, People’s Republic of China.

Received: 27 December 2021 Accepted: 11 February 2022

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.