San Antonio Breast Cancer Symposium – December 9-12, 2025 #PD2-04

The ROYAL MARSDEN NHS Foundation Trust

PD2-04 Tumor Heterogeneity in the transATAC Breast Cancer Spatial Single Cell Atlas

ICR The Institute of Cancer Research

E. F. Schuster[1,2], I. T. Kleijn[1,2], K. Dunne[1,2], R. Sardinha[1,2], E. O'Neill[1,2], G. Latifi[3], C. Kennedy-Dietrich[3], N. B. Jamieson[3], J. Cuzick[4], A. R. Brentnall[4], M. Dowsett[1], S. R. Johnston[1,5], N. C. Turner[1,2,5]

[1]Division of Breast Cancer Research, The Institute of Cancer Research, [2]Ralph Lauren Breast Cancer Research Centre, Royal Marsden & ICR, [3]School of Cancer Sciences, University of Glasgow,
[4]Wolfson Institute of Population Health, Queen Mary University of London, [5]Breast Unit, The Royal Marsden NHS Foundation Trust.

BACKGROUND

Primary breast cancer (BC) is highly heterogeneous and this drives divergent clinical outcomes. Many molecular features have significant prognostic value in hormone receptor positive (HR+) BC, and several commercial multigene prognostic assays currently guide chemotherapy decision making.

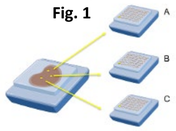
The translational study from the monotherapy arms of the Arimidex, Tamoxifen, Alone or in Combination trial (transATAC), has been pivotal for the validation[1] and regulatory approval of multiple prognostic assays. The TransATAC cohort consists of 1,231 early-stage breast cancers from postmenopausal women with a median follow up of 120 months[1].

We assessed whether single cell spatial transcriptomic analysis of a tissue micro-assay (TMA) core, sampling of ~1000 cells per tumor, is sufficient to recapitulate the major features of the tumor.

METHODS

Samples: TMA blocks of 0.6mm diameter cores were generated from tumors with 3 spatially distant cores per tumor (Fig. 1). CosMx spatial transcriptomic data (6108 genes) was available for 464 cores with 424 from 340 patients and 19 cores from 7 control samples that passed QC based on cell number (>200 cells), average transcript count (200 per cell) and FOV technical effects/biases. Cells were filtered by area (25-400μm^2), transcript count (≧100), and density (≧1.0 transcripts/μm^2), leaving 420,343 cells for analysis. Matching cores from 84 tumor and control samples were available to test intratumor heterogeneity (Fig. 1).

Fig. 1



Data integration and annotation: Data integration across six slides was performed using Harmony[2] on 50 principal components. Leiden clustering was performed with Seurat[3]. Major tumor microenvironment (TME) cell types were manually identified from individual Leiden clusters based on marker gene expression and all cells in the remaining clusters were annotated as epithelial/cancer cells.

Pseudobulk analysis: For all samples that passed QC, pseudobulk expression for whole core, epithelial/cancer (>9 cells) and TME (>9 cells) was calculated as the total count for each transcript across sample/cell-type, normalized to counts per million (CPM), transformed to log2(CPM+.1) and batch-corrected using ComBat.

Molecular data from surgical excision: OncotypeDX gene expression and recurrence scores (n=263) from ER+HER2- surgery excisions were available[4]. Estimates of OncotypeDX recurrence scores were based on median centering pseudobulk log2 cpm values to RT-PCR values from matched surgical excisions.

RESULTS

TME cell types are shared across many patients, but tumors tend to be patient specific

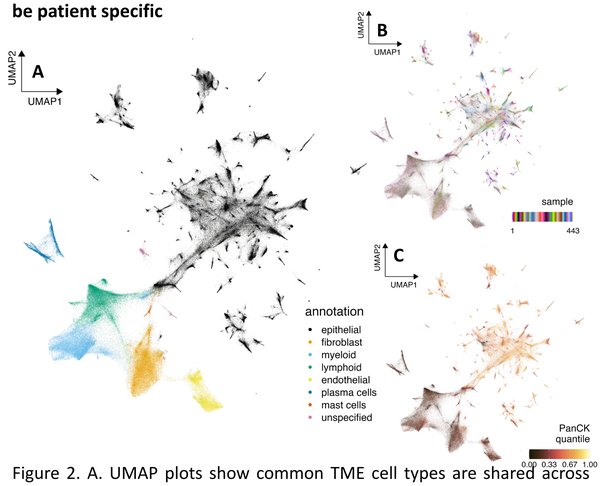


Figure 2. A. UMAP plots show common TME cell types are shared across many different patients. B. Many cancer cell clusters are patient specific. C. high pan-cytokeratin (PanCK) expression in epithelial/cancer clusters.

TME expression more variable than cancer cell expression in matching cores taken from the same tumor

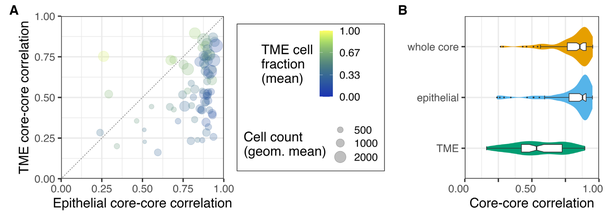


Figure 3. A. Pearson correlations of matching cores (n=84) for pseudobulk epithelial/cancer cells (x-axis) and TME specific gene expression (y-axis). B. Boxplot of Pearson correlations of matching cores for pseudobulk whole cores, epithelial/cancer cells and TME cells.

Inverse correlation between ESR1 expression and TME cell fractions of cores

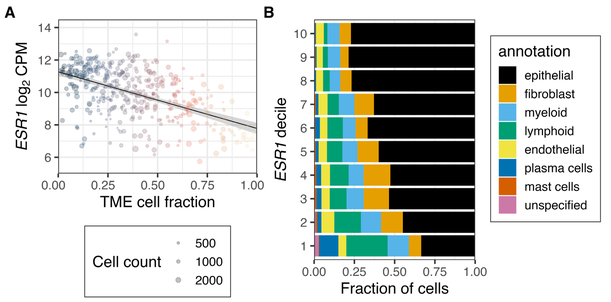


Figure 4. A. Scatter plot of pseudobulk whole core *ESR1* expression and fraction of TME cells in a core (Spearman's rho=-0.46;*p<1e-9*). The correlation is also significant when only consider ER+HER2- tumors (rho=-0.53;*p<1e-9*). B. Barplot with fraction of epithelial/cancer cells and TME cell types when divided into deciles based on ESR1 pseudobulk whole core gene expression.

Correlation of *ESR1, ERBB2 and OncotypeDX risk scores* between bulk (surgical excisions) and pseudobulk (TMA cores) gene expression

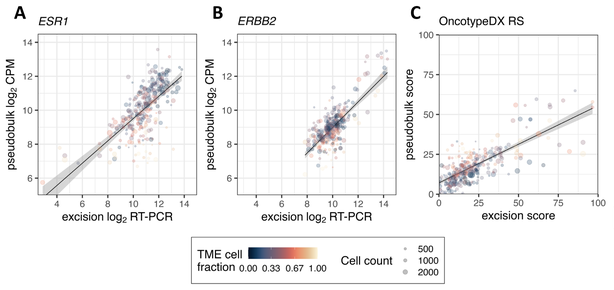


Figure 5. Scatter plots of pseudobulk whole-core *ESR1* (A) and *ERBB2* (B) gene expression and Oncotype DX (C) risk scores compared to gene expression/risk scores taken from ER+HER2+ surgical excisions (RT-PCR). Pearson correlations are significant for *ESR1* (*r*=0.65, *p<1e*-9), *ERBB2* (*r*=0.72, *p<1e*-9) and OncotypeDX (*r*=0.77, *p<1e*-9).

Correlation of *ESR1, ERBB2 and OncotypeDX risk scores for matching corestaken from the same tumor*

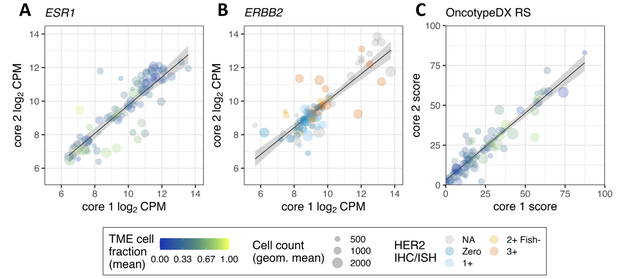


Figure 6. . Scatter plots of pseudobulk whole-core *ESR1* (*r*=0.82, *p<1e*-9) (A) and *ERBB2* (*r*=0.84, *p<1e*-9) (B) gene expression and Oncotype DX (C) risk scores (*r*=0.92, *p<1e*-9) from matched cores.

CONCLUSIONS

1. Single cell spatial analysis of transATAC TMA cores of ~1000 cells reproduced key features of tumor biology and risk as assessed in surgical excisions
2. Matching cores from the same tumor showed high correlation across multiple tumor related molecular features
3. Variation of TME expression in cores suggests multiple sampling is required to capture TME
4. Inverse association between *ESR1* expression and number of non-tumor cells in a core supports emerging evidence of immunosuppressive role of ER signaling in breast cancer[6]
5. Preliminary results suggest taking multiple small samples (~1000 cells) from many patients (goal=~1000 tumors) may be a successful strategy to understand BC biology, TME cell-types/niches and cell-cell interactions that impact prognosis.

REFERENCE

1. Cuzick, et al. Lancet Oncol. 11(12):1135-41 (2010).
2. Korsunsky, et al. Nat Methods 16(12):1289–1296 (2019).
3. Hao, et al. Nat Biotechnol 42, 293–304 (2024).
4. Sestak, et al. JAMA Oncol 4:545-553 (2018).
5. Rae, et al. JCI 15;131(24) (2021).

ACKNOWLEDGEMENTS

We thank all patients, collaborating investigators and pathologists of the Arimidex, Tamoxifen, Alone or in Combination (ATAC) trial, and Breast Cancer Now (previously Breakthrough Breast Cancer) and AstraZeneca for support and funding of the transATAC study. The transATAC Research Tissue Bank is supported by NHS funding to the Royal Marsden and Institute of Cancer Research NIHR Biomedical Research Centre. Analysis of the spatial data is supported by the Arthur Foundation and the Breast Cancer Research Foundation.

Breast Cancer Now The research & support charity

BCRF Breast Cancer Research Foundation

This presentation is the intellectual property of the author/presenter. Contact them at Gene.Schuster@icr.ac.uk for permission to reprint and/or distribute.