

Appendix files

A single cell RNA atlas of human breast spanning normal, preneoplastic and tumorigenic states

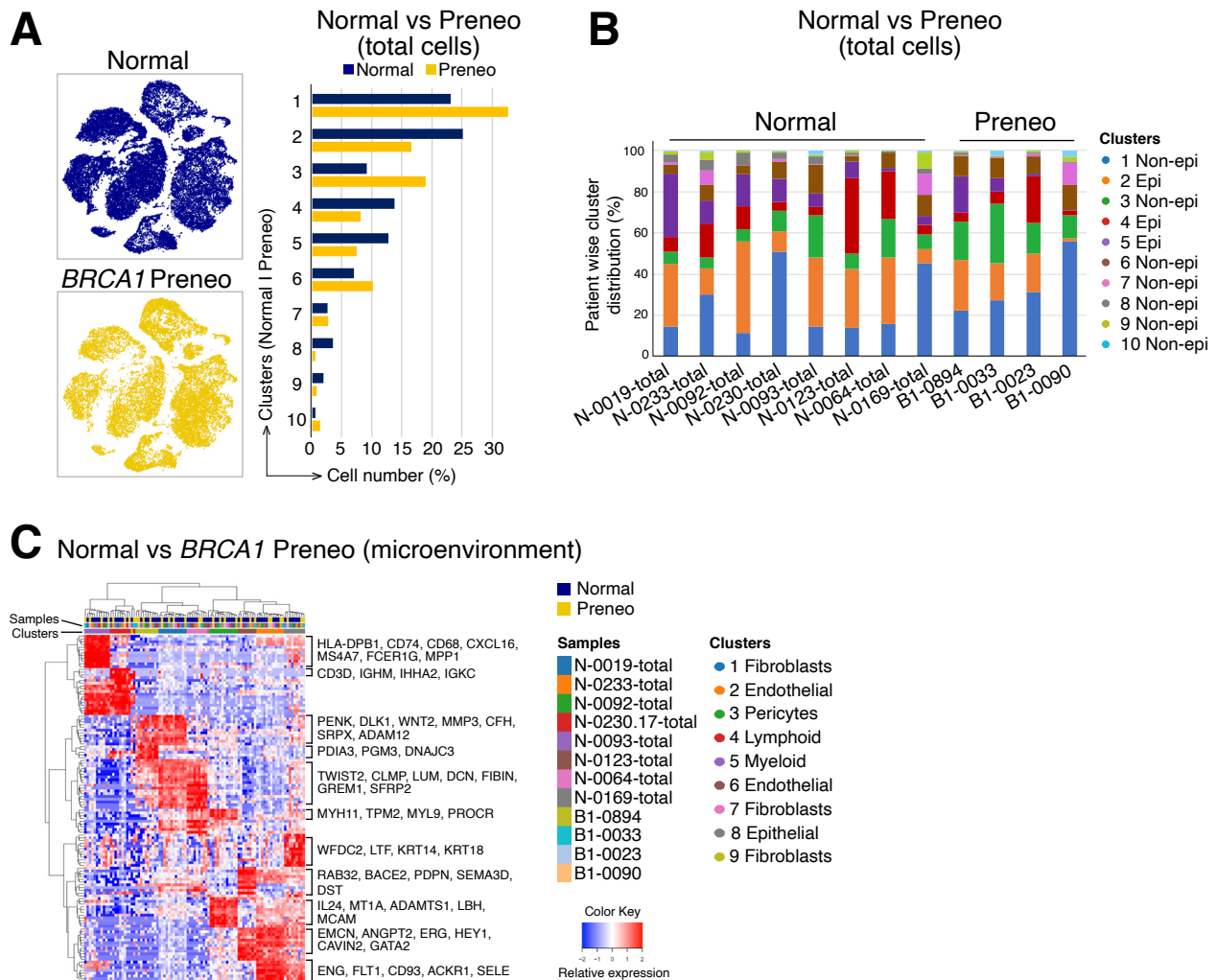
Bhupinder Pal, Yunshun Chen, François Vaillant, Bianca D. Capaldo, Rachel Joyce, Xiaoyu Song, Vanessa Bryant, Jocelyn S. Penington, Leon Di Stefano, Nina Tubau Ribera, Stephen Wilcox, G. Bruce Mann, kConFab, Anthony T. Papenfuss, Geoffrey J. Lindeman, Gordon K. Smyth and Jane E. Visvader.

Table of Contents

Appendix Figure S1. Single cell profiling of normal versus preneoplastic breast tissue from BRCA1 mutation carriers.

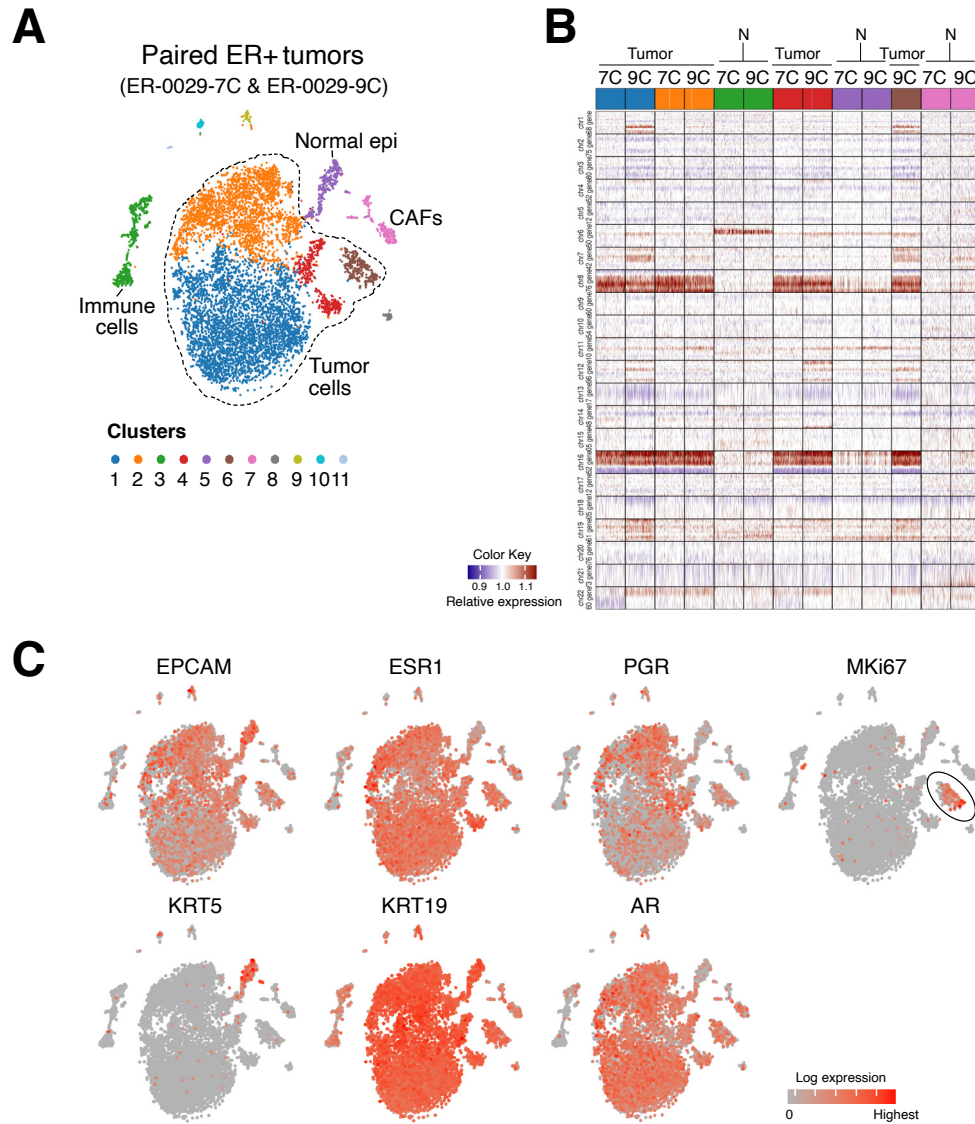
Appendix Figure S2. Tumor heterogeneity in a patient harboring two ER+ tumors.

Appendix Figure S3. Distinct tumor microenvironments define breast cancer subtypes.



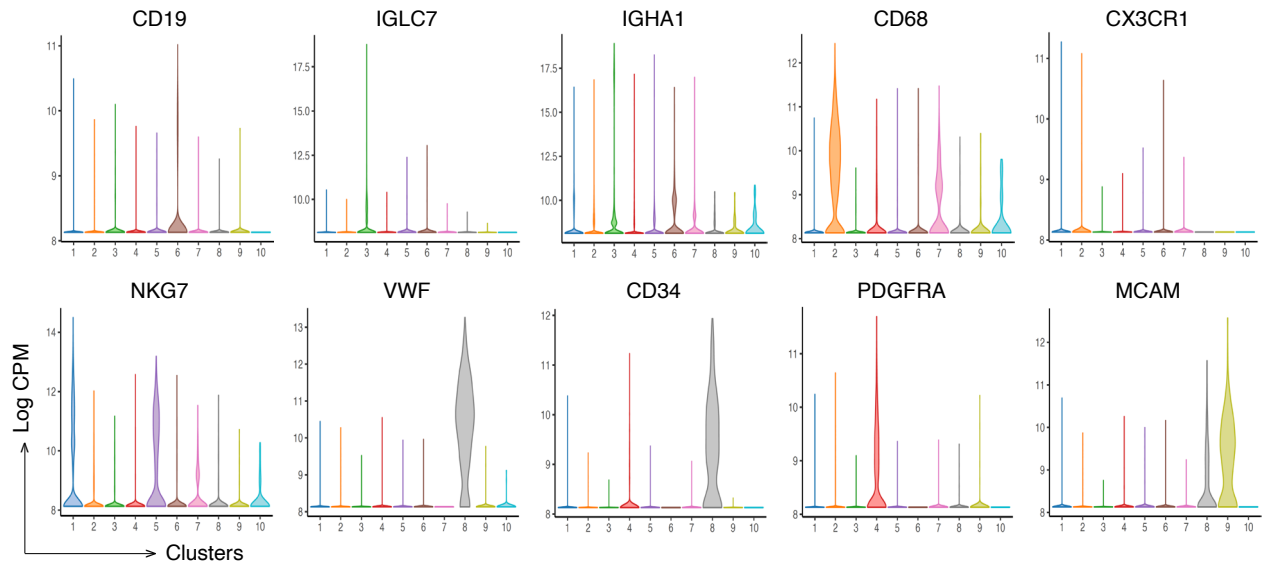
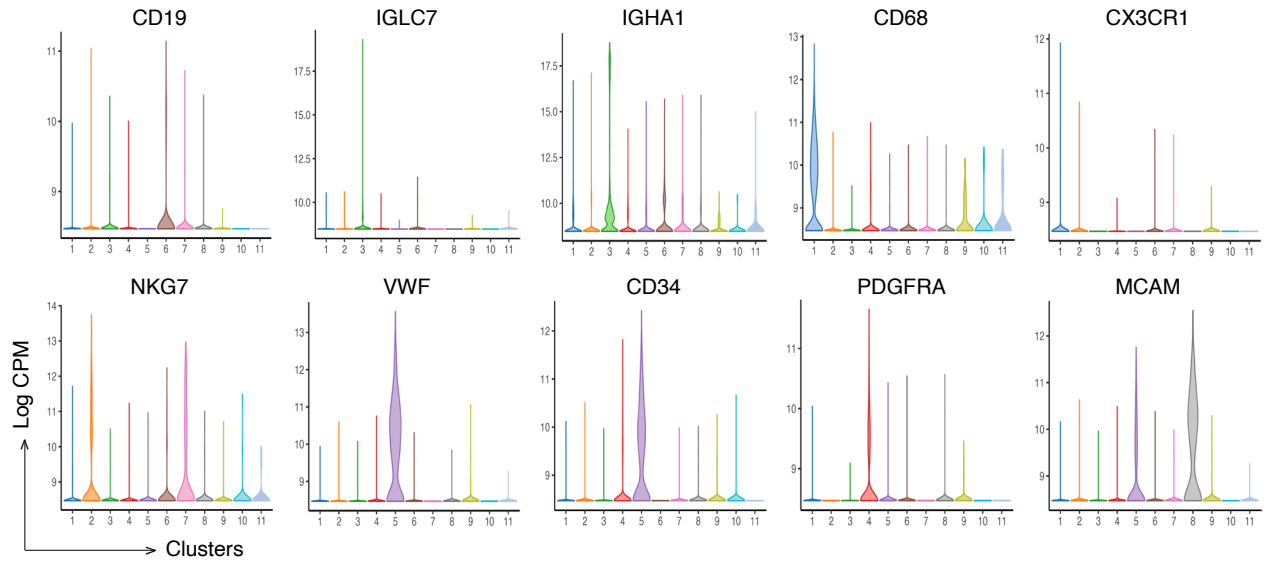
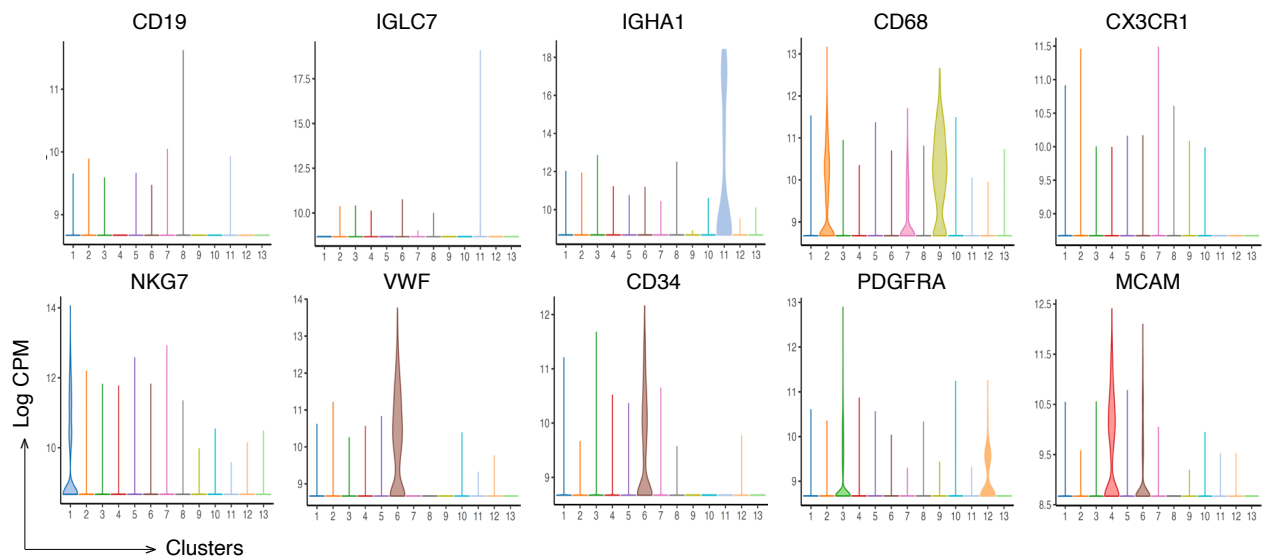
Appendix Figure S1. Single cell profiling of normal versus preneoplastic breast tissue from *BRCA1* mutation carriers.

- A. Left panel, same t-SNE map as in Fig 4C but separated into cells from normal (blue) and preneoplastic specimens (yellow). Right panel, barplot showing relative proportions for Fig 4C clusters by condition.
- B. Cluster proportions as in (A) but by individual patient. A quasi-likelihood F-test was used to test for differences in cluster frequencies between pre- and post-menopausal samples ($P = 0.14$).
- C. Heat map of pseudo-bulk samples showing marker genes for each non-epithelial cluster. The top 15 marker genes were identified for each cluster by differential expression analysis of the pseudo-bulk read counts. Color bars at the top of the plot indicate the cluster and *BRCA1* status of each sample (blue/yellow for *WT/BRCA1*-mutant). Individual patients are also indicated.



Appendix Figure S2. Tumor heterogeneity in a patient harboring two ER+ tumors.

- Cell clusters in the combined t-SNE transcriptome profiles of the two ER+ tumors. Samples ER-0029-7C and ER-0029-9C were concurrently isolated from the same breast of patient ER-0029. Populations were identified through marker expression analyses. Dotted line represents malignant epithelial cells.
- InferCNV plots were generated from the combined transcriptomes (A) to map copy number variations (CNV) in each chromosome. N indicates normal cells.
- t-SNE plots showing expression of epithelial markers and MKI67 (marked by the circle).

A**TNBC****B****HER2+****C****ER+**

Appendix Figure S3. Distinct tumor microenvironments de ine breast cancer subtypes.

A-C. Violin plots showing expression of key immune and stromal cell markers in the non-epithelial cell clusters identified in Fig 7A-C for the different tumor subtypes. Y-axis shows log-expression (log2 counts-per-million); x-axis shows cell clusters.