Figure S1

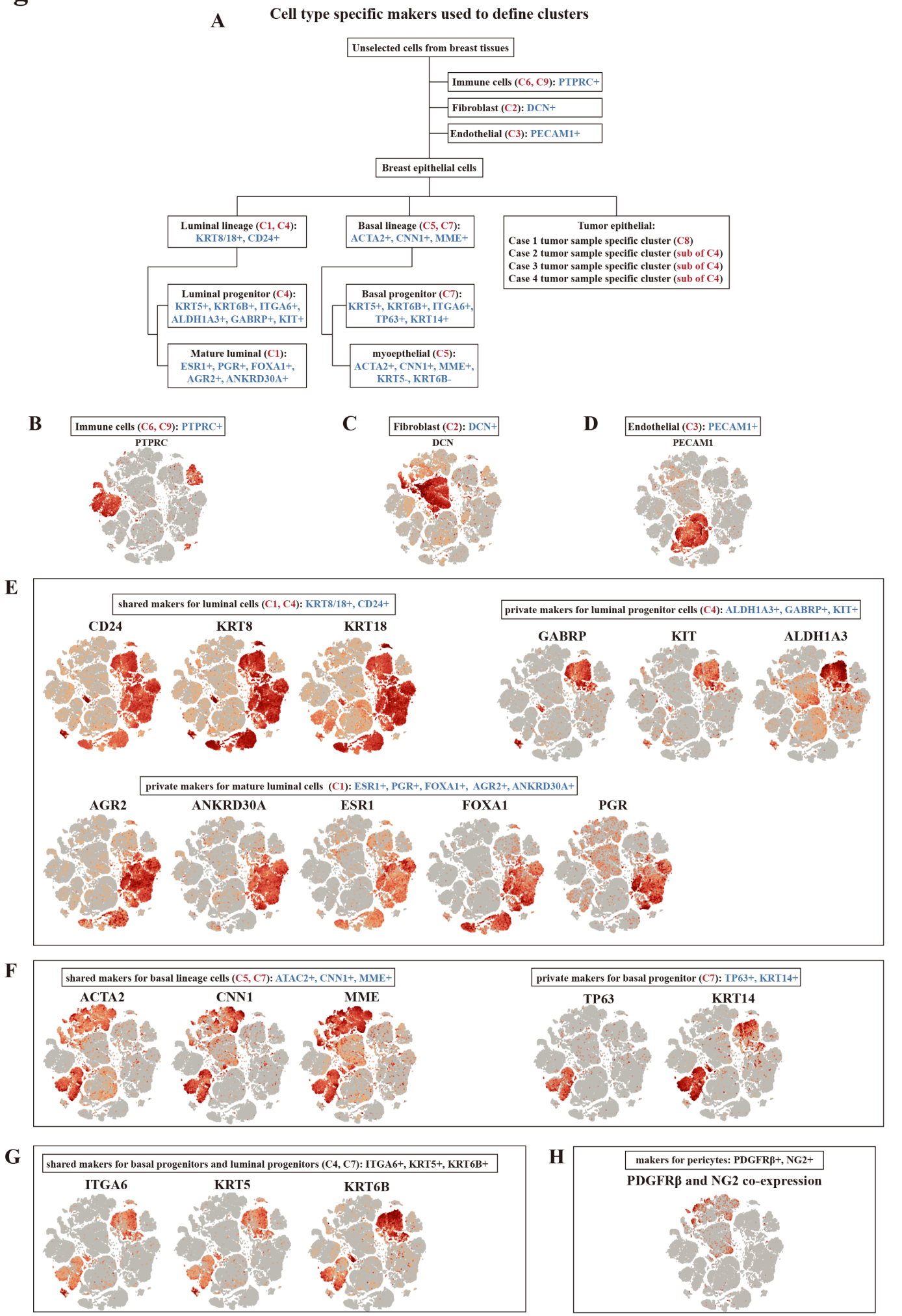

Figure S1. Known cell type specific markers used to define clusters

(A) Flow chart for cell cluster definition. (B-H) t-SNE plots showing the expression of known cell type specific markers used to define clusters.

Figure S2

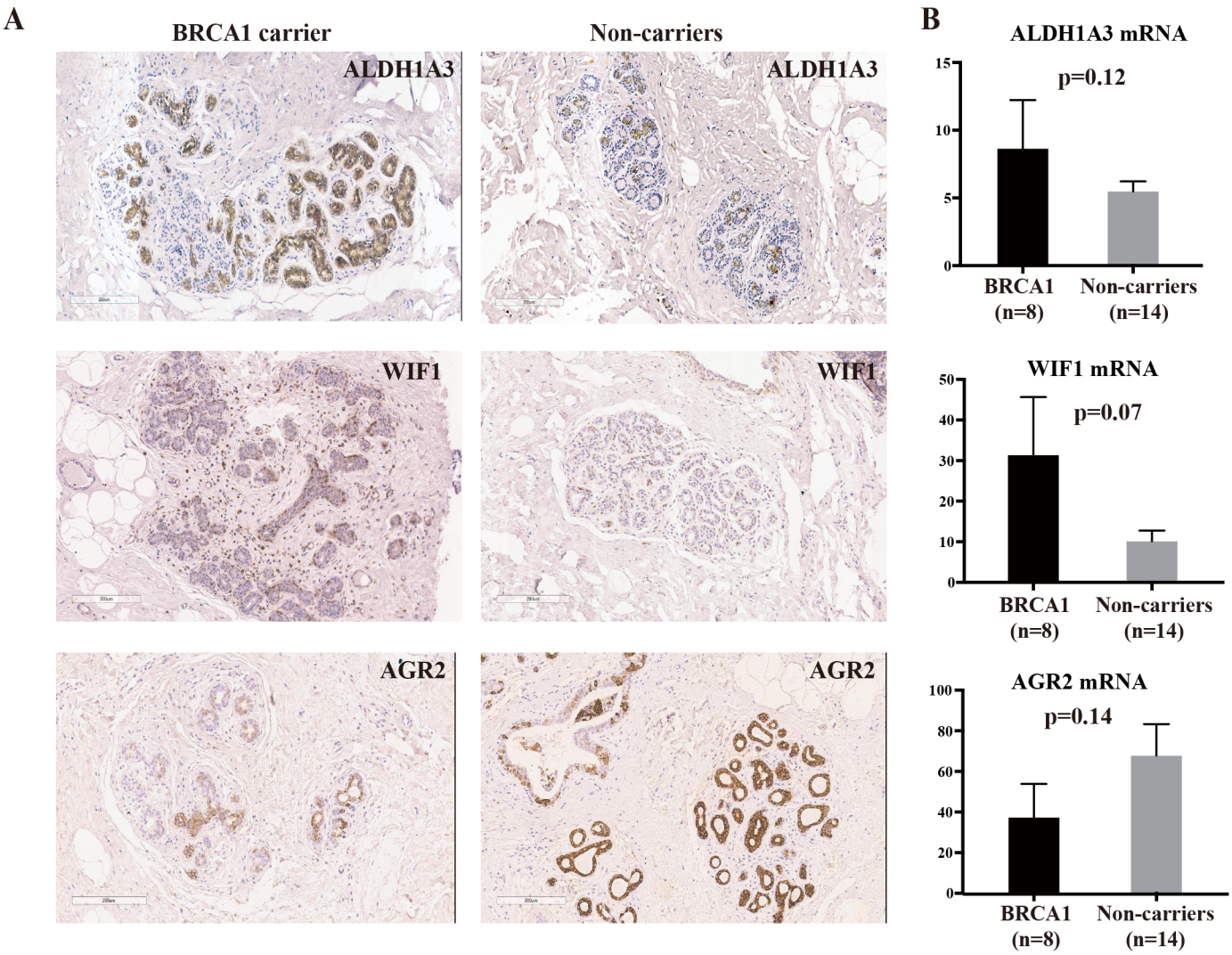

Figure S2. The aberrant proportions of epithelial cells in normal breast tissues from BRCA1 mutation carriers validated by immunohistochemistry staining and bulk RNA-seq data.
(A) Immunohistochemistry staining for ALDH1A3 (a marker of luminal progenitor cells), WIF1 (a marker of basal progenitor cells), and AGR2 (a marker of mature luminal cells) in normal breast tissue from 7 BRCA1 mutation carriers (4 scRNA-seq cases and 3 additional cases) and 10 matched non-carriers. (B) Comparison of the expression of ALDH1A3, AGR2 and WIF1 in normal breast tissues from 8 additional BRCA1 mutation carriers and 14 matched non-carriers based on bulk RNA-seq data. Data are showed as the mean ± SEM and analyzed for statistical significance using t-test or Mann-Whitney test, where appropriate. P-values less than 0.05 were considered to be statistically significant.

Figure S3

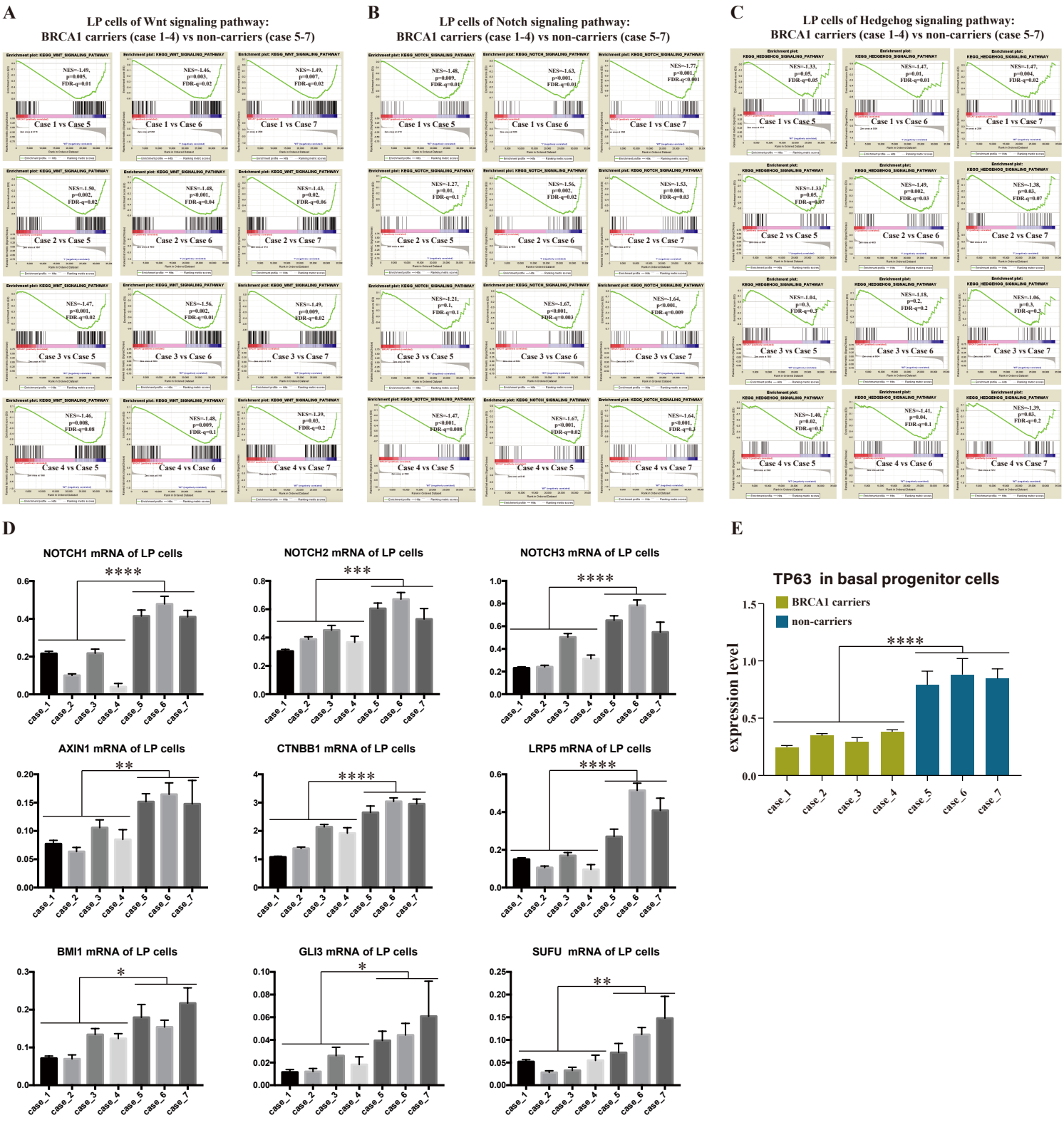

Figure S3. Impaired differentiation processes of luminal and basal/myoepithelial lineages in normal breast tissue from BRCA1 mutation carriers.

(A-C) GSEA plots showing downregulated breast epithelial differentiation pathways in luminal progenitor cells from BRCA1 mutation carriers compared with those from each non-carrier, including the wnt, notch and hedgehog signaling pathways. LP, luminal progenitor. (D) Comparison of the mRNA levels of key genes of these pathways (Notch1, Notch2 and Notch3 for Notch signaling pathway; AXIN1, CTNBB1 and LRP5 for Wnt signaling pathway; BMI1, GLI3 and SUFU for Hedgehog signaling pathway) in luminal progenitor cells from BRCA1 mutation carriers (cases 1-4, 5035 cells) with those from non-carriers (cases 5-7, 775 cells). (E) Comparison of TP63 expression (a key gene for basal/myoepithelial lineage differentiation) between basal progenitor cells from BRCA1 mutation carriers (cases 1-4, 4485 cells) and those from non-carrier (cases 5-7, 353 cells). Data represent the mean ± SEM. P-values less than 1e-5 were considered to be statistically significant due to a large number of cells (Mann-Whitney test). *P < 1e-5, **P <1e-10, ***P < 1e-15 and ****P < 1e-20

Figure S4

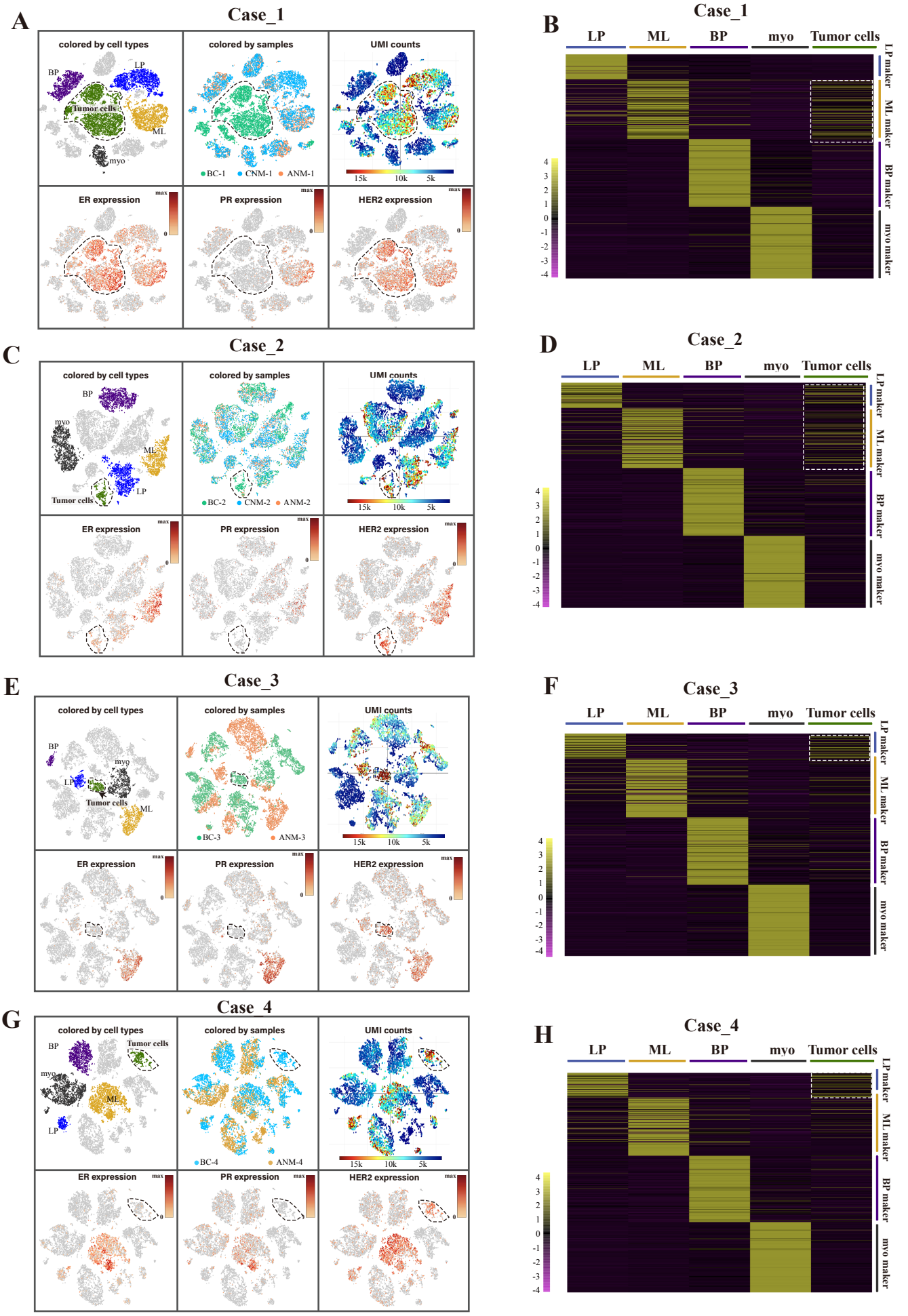

Figure S4. Identification and characterization of the tumor cells in BRCA1 mutation carriers.
(A, C, E, G) Identification of the tumor cells in each BRCA1 mutation carrier (cases 1-4). t-SNE plots of high-quality cells from all samples of each BRCA1 mutation carrier colored by cell type, sample, unique molecular identifiers (UMI) count and ER/PR/HER2 expression level. (B, D, F, H) Comparison of the expression patterns of tumor cells and distinct types of epithelial cells in normal breast tissue from each BRCA1 mutation carrier (case 1-4).

Figure S5

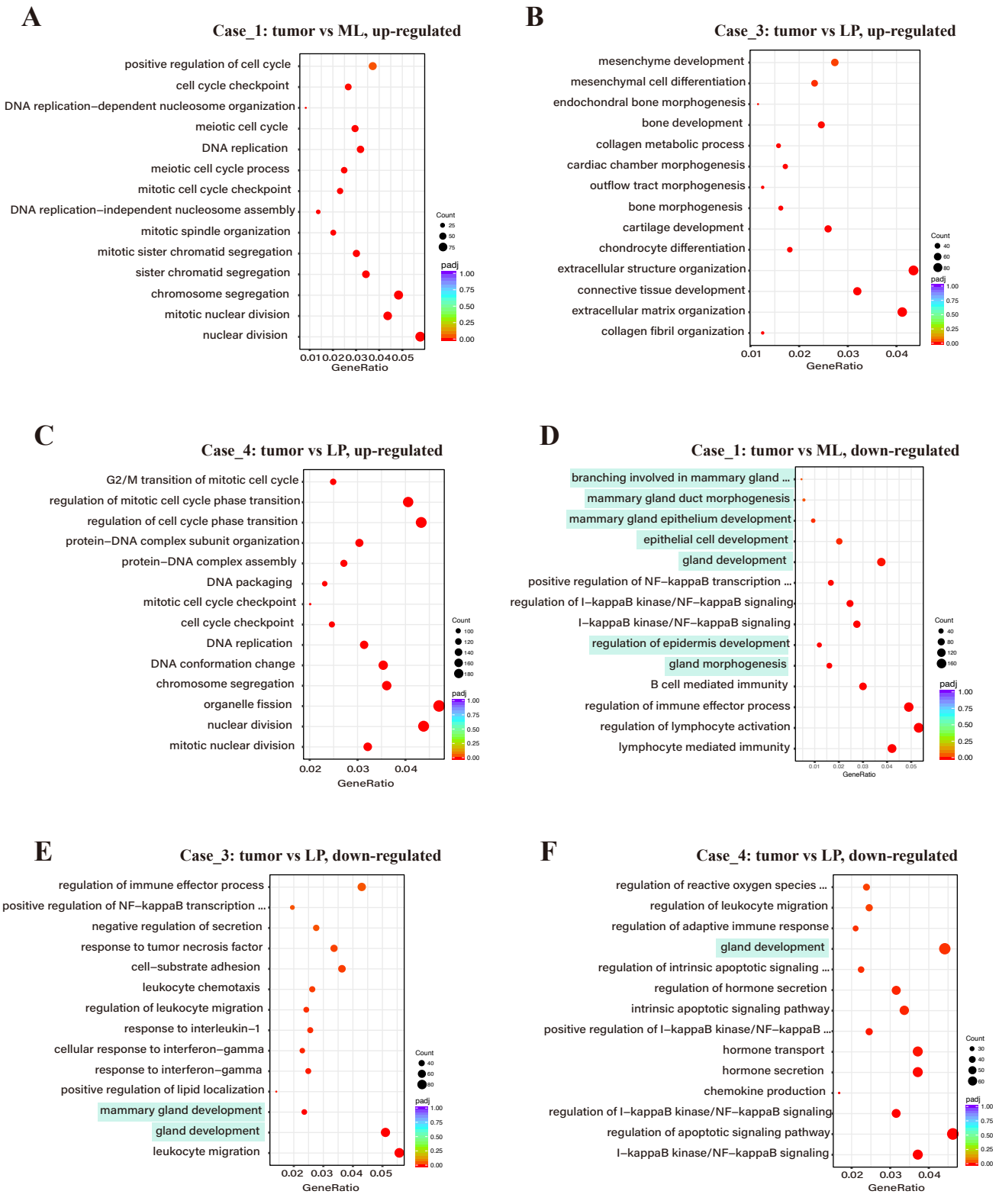

Figure S5. Upregulated and downregulated pathways in tumor cells compared with their putative cells of origin in each BRCA1 mutation carrier.

(A-C) Gene ontology analysis of upregulated pathways in tumor cells compared with their original epithelial cells in each BRCA1 mutation carrier (cases 1, 3, 4). (D-F) Gene ontology analysis of downregulated pathways in tumor cells compared with their original epithelial cells in each BRCA1 mutation carrier (cases 1, 3, 4). ML, mature luminal cells, ML, mature luminal cells.

Figure S6

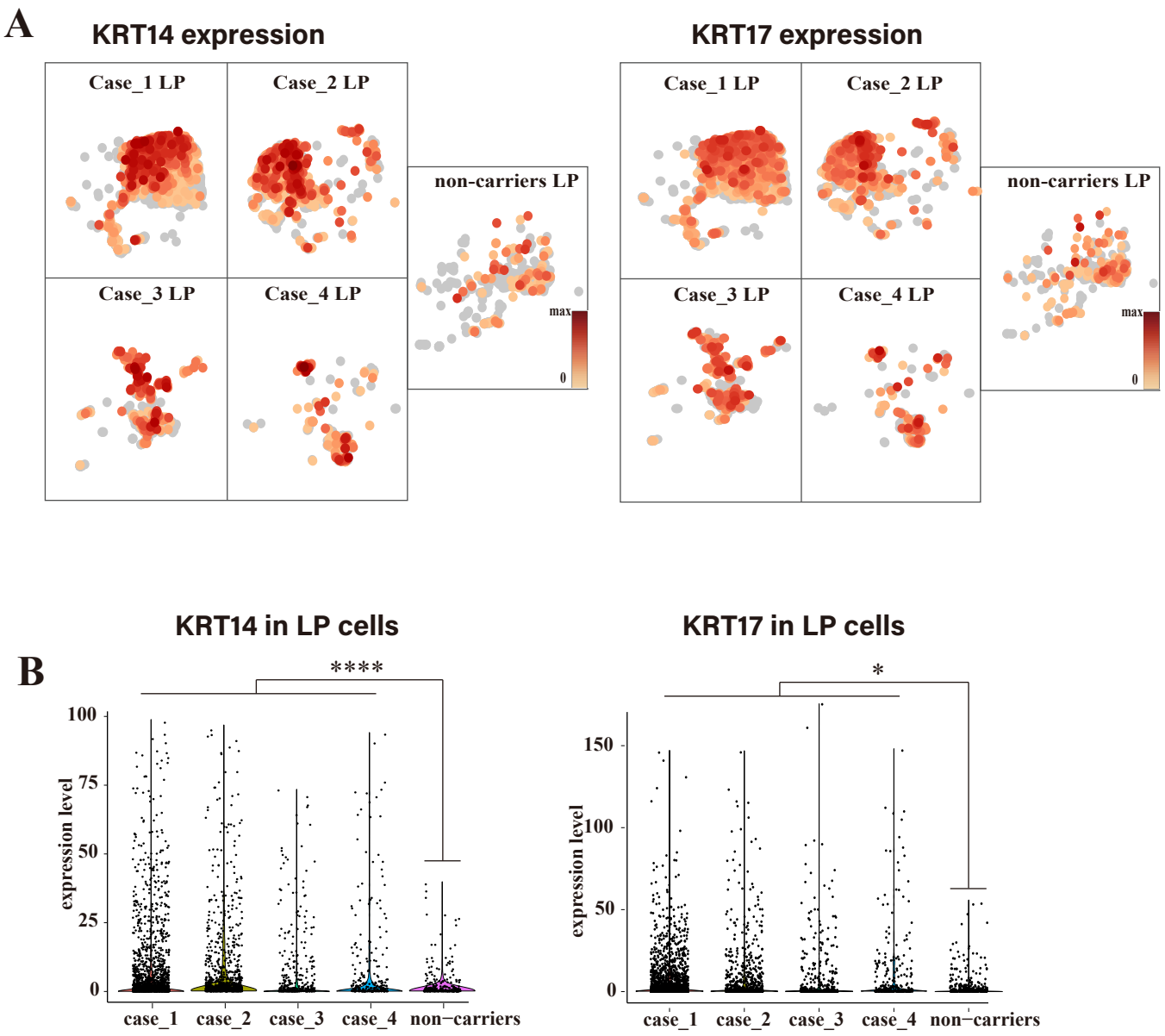

Figure S6. Comparison of the expression levels of KRT14/17 in luminal progenitor cells in normal breast tissue between BRCA1 mutation carriers and non-carriers.
(A) Feature plots showing that luminal progenitor cells in BRCA1 mutation carriers (cases 1-4) expressed higher levels of KRT14/KRT17 than luminal progenitor cells in non-carriers. (B) Violin plots showing the same information as in A. LP, luminal progenitor cells. P-values less than 1e-5 were considered to be statistically significant due to a large number of cells (Mann-Whitney test). *P < 1e-5, **P <1e-10, ***P < 1e-15 and ****P < 1e-20

Figure S7

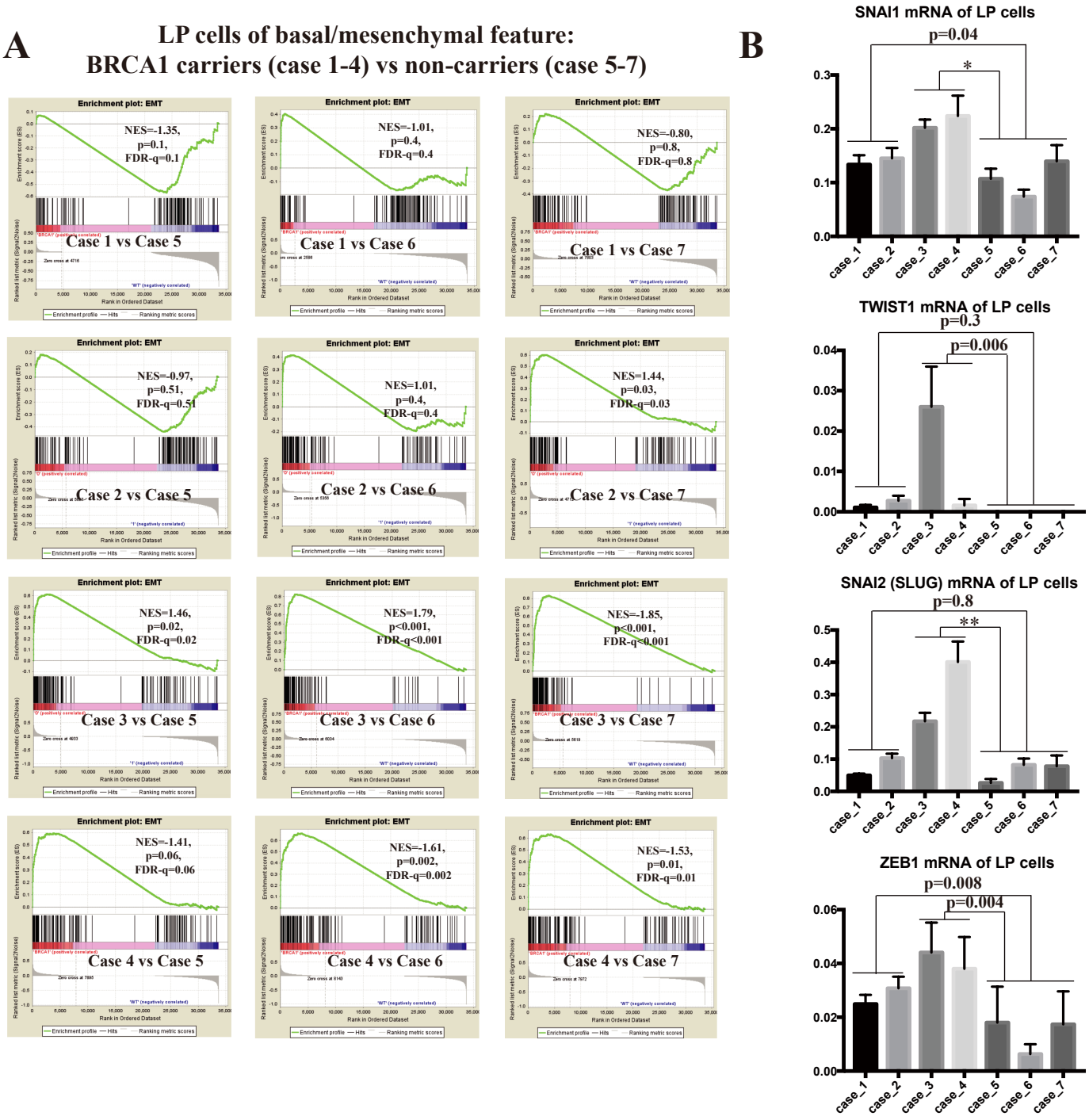

Figure S7. Comparison of the basal/mesenchymal features of luminal progenitor cells in normal breast tissue between BRCA1 mutation carriers and non-carriers.

(A) GSEA plots showing upregulated basal/mesenchymal features of luminal progenitor cells in normal breast tissue from BRCA1 mutation carriers with basal-like tumors (cases 3-4) compared with each non-carrier (cases 5-7). Luminal progenitor cells from BRCA1 mutation carriers with luminal tumors (cases 1-2) didn't show significantly upregulated basal/mesenchymal features compared with those from each non-carrier (cases 5-7). (B) Comparison of the expression of key epithelial-mesenchymal transition (EMT) transcription factors (SNAI1, SLUG, TWIST1, ZEB1) in luminal progenitor cells from normal breast tissue among BRCA1 mutation carriers with luminal tumors (cases 1-2, 4159 cells), BRCA1 mutation carriers with basal-like tumors (cases 3-4, 876 cells) and non-carriers (cases 5-7, 775 cells). LP, luminal progenitor cells. Data represent the mean ± SEM. P-values less than 1e-5 were considered to be statistically significant due to a large number of cells (Mann-Whitney test).*P < 1e-5, **P <1e-10, ***P < 1e-15 and ****P < 1e-20.

Figure S8

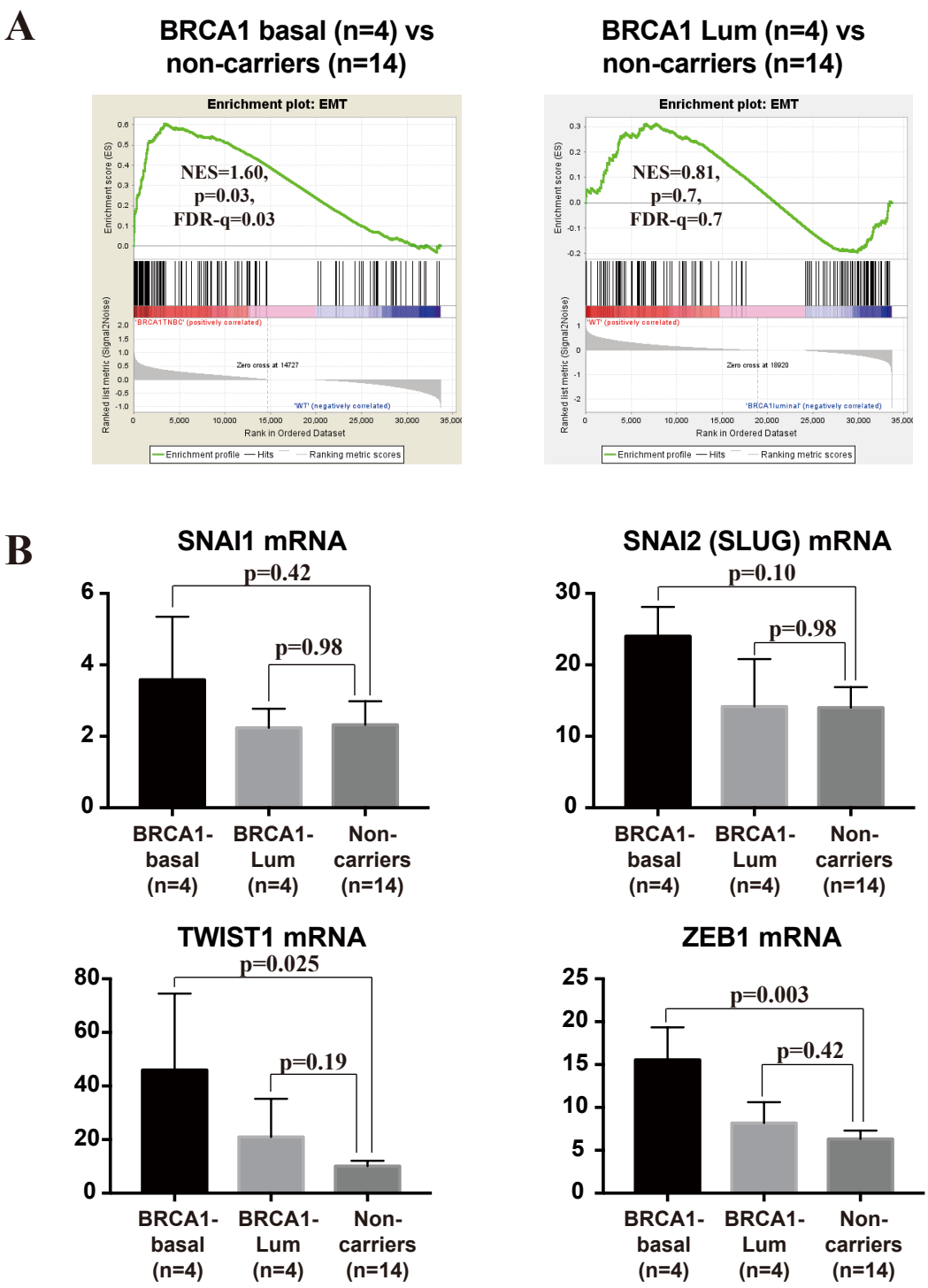

Figure S8. Validation of the heterogeneity of basal/mesenchymal features in normal mammary tissue from BRCA1 mutation carriers by bulk RNA-seq.
(A) Normal mammary tissues from 4 additional BRCA1 mutation carriers with basal-like tumors showed significantly upregulated basal/mesenchymal features compared with those from 14 matched non-carriers (left), while the normal mammary tissues from 4 additional BRCA1 mutation carriers with luminal tumors didn't show upregulated basal/mesenchymal features compared with those from 14 matched non-carriers (right). (B) Comparison of the expression of some important epithelial-mesenchymal transition (EMT) transcription factors (SNAI1, SLUG, TWIST1, ZEB1) in normal mammary tissue between BRCA1 mutation carriers with luminal tumors (n=4), BRCA1 mutation carriers with basal-like tumors (n=4) and non-carriers (n=14). Data are showed as the mean ± SEM and analyzed for statistical significance using t-test or Mann-Whitney test, where appropriate. P-values less than 0.05 were considered to be statistically significant.