

Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific

Paula Michea^{1,2,11}, Floriane Noël^{1,2,3,11}, Eve Zakine^{1,2}, Urszula Czerwinska^{1,2,4,5,6}, Philémon Sirven^{1,2}, Omar Abouzid^{1,2}, Christel Goudot^{1,2}, Alix Scholer-Dahirel^{1,2}, Anne Vincent-Salomon^{1,2,7,8}, Fabien Reyal^{9,10}, Sebastian Amigorena^{1,2}, Maude Guillot-Delost^{1,2}, Elodie Segura^{1,2} and Vassili Soumelis^{1,2*}

The functions and transcriptional profiles of dendritic cells (DCs) result from the interplay between ontogeny and tissue imprinting. How tumors shape human DCs is unknown. Here we used RNA-based next-generation sequencing to systematically analyze the transcriptomes of plasmacytoid pre-DCs (pDCs), cell populations enriched for type 1 conventional DCs (cDC1s), type 2 conventional DCs (cDC2s), CD14⁺ DCs and monocytes-macrophages from human primary luminal breast cancer (LBC) and triple-negative breast cancer (TNBC). By comparing tumor tissue with non-invaded tissue from the same patient, we found that 85% of the genes upregulated in DCs in LBC were specific to each DC subset. However, all DC subsets in TNBC commonly showed enrichment for the interferon pathway, but those in LBC did not. Finally, we defined transcriptional signatures specific for tumor DC subsets with a prognostic effect on their respective breast-cancer subtype. We conclude that the adjustment of DCs to the tumor microenvironment is subset specific and can be used to predict disease outcome. Our work also provides a resource for the identification of potential targets and biomarkers that might improve antitumor therapies.

Dendritic cells (DCs) are antigen-presenting cells (APCs) specialized in triggering adaptive immune responses through the activation of T cells¹. The various subsets of DCs have been defined on the basis of their ontogeny, phenotype and anatomical location^{2,3}. Advances in high-throughput technologies have improved the classification of DCs by identifying novel subset-specific markers and molecular signatures⁴. Studies of mice and human suggest that at steady state, ontogeny is a predominant factor in defining DC subset identity^{5–8}. For example, studies of plasmacytoid pre-DCs (pDCs)⁹ and the cDC1 and cDC2 subsets of conventional DCs (CD141⁺ DCs and CD1c⁺ DCs, respectively) from human blood and tonsils have revealed that pDCs cluster first by ontogeny independently of their tissue of origin¹⁰. In contrast, cDC1s and cDC2s are more sensitive to tissue localization, as tonsil cDC1s cluster closer to tonsil cDC2s than to blood cDC1s¹⁰. Tissue imprinting also influences DC function. Gut DCs induce the homing of T cells back to the gut through a mechanism dependent on retinoic acid, the chemokine receptor CCR9 and the integrin $\alpha_4\beta_7$, but spleen DCs do not¹¹. This suggests complex interplay between ontogeny and tissue imprinting, with the relative contribution of each remaining a matter of debate.

During inflammation, complex signals must be integrated by various DC subsets, which can change their function and molecular features^{12–17}. The diversity of DC subsets itself is also modified by inflammation through the appearance of monocyte-derived inflammatory DCs, which are absent in homeostatic conditions¹⁸. In humans, inflammatory DCs have been characterized in psoriatic skin^{19,20}, ascites fluid of ovarian cancer and synovial fluid of rheumatoid arthritis²¹. DCs infiltrate most cancer types. They serve

a protective role in anti-tumor immunity through the expression of co-stimulatory molecules and inflammatory cytokines and by inducing the activation of T cells^{22,23}. Conversely, DCs also promote immunosuppression by secreting anti-inflammatory cytokines^{24–27} or by expressing negative immunological checkpoint molecules, which inhibit T cell activation and are now being targeted by promising anti-tumor therapies^{28,29}. The plasticity of DCs in various tumor microenvironments (i.e., tissue imprinting), as well as specialized ontogeny-driven DC functions, might contribute to such molecular and functional heterogeneity.

In this study, we performed a systematic comparative transcriptomics study of DC subsets in human primary breast cancer and uninvolved tissue juxtaposed to the tumor, from the same patient. We found that the transcriptional reprogramming of tumor-infiltrating DCs was DC subset specific, suggestive of complex interplay between ontogeny and tissue imprinting in conditioning DC diversity in the tumor microenvironment. Our results also provide high-quality large-scale datasets of primary tumor-infiltrating DCs that constitute a valuable resource for the biomedical community.

Results

Phenotypically distinct APCs infiltrate human breast cancer. DCs that had infiltrated breast cancer tissues were identified by multicolor flow cytometry on the basis of published studies of human DC subsets²⁰. Because this was the first in-depth characterization of DC subsets in human breast cancer, to our knowledge, we performed preliminary analyses to validate our strategy. After standard gating to eliminate debris, doublets and dead cells, we selected CD45⁺ cells to efficiently exclude CD45⁻ cells, which were mainly

¹Institut Curie, PSL Research University, Paris, France. ²INSERM, UMR 932, F-75005 Paris, France. ³Université Paris Sud, Université Paris-Saclay, Orsay, France. ⁴INSERM, U900, F-75005 Paris, France. ⁵Mines ParisTech, Paris, France. ⁶Centre de Recherches Interdisciplinaires, Paris Descartes, Paris, France. ⁷Institut Curie, Service de Pathologie, Paris, France. ⁸INSERM, U934, F-75005 Paris, France. ⁹INSERM, UMR 932, Département de Recherche Translationnelle, Residual Tumor & Response to Treatment Laboratory (RT2Lab), Paris, France. ¹⁰Institut Curie, Département de Chirurgie, Paris, France.

¹¹These authors contributed equally: Paula Michea, Floriane Noël. *e-mail: vassili.soumelis@curie.fr

Table 1 | Variables that influence the disease-free survival of patients with breast cancer

	LBC		TNBC	
	HR	P value	HR	P value
pDC				
NPI > 5.4	1	-	-	-
NPI ≤ 5.4	0.31	$7.5 \times 10^{-13}^*$	-	-
High signature ratio	1	-	-	-
Low signature ratio	1.37	0.0072*	-	-
cDC2				
NPI > 5.4	1	-	-	-
NPI ≤ 5.4	0.3	$1.8 \times 10^{-13}^*$	-	-
High signature ratio	1	-	-	-
Low signature ratio	1.27	0.041*	-	-
cDC1e				
NPI > 5.4	1	-	1	-
NPI ≤ 5.4	0.29	$7.6 \times 10^{-14}^*$	0.27	$1.1 \times 10^{-9}^*$
High signature ratio	1	-	1	-
Low signature ratio	1.39	0.0041*	1.76	0.0058*
MonoMac				
NPI > 5.4	1	-	1	-
NPI ≤ 5.4	0.31	$5.9 \times 10^{-13}^*$	0.28	$3.9 \times 10^{-9}^*$
High signature ratio	1	-	1	-
Low signature ratio	0.77	0.025*	0.67	0.049*

Multivariate Cox regression analysis of predictors of disease-free survival that influence the disease-free survival of patients with LBC or TNBC (top), showing the Nottingham prognostic index (NPI) and 'subset-specific signature z-score' for each cell type (left column), with the hazard ratio (HR) and P value for each (*P < 0.05, Cox model likelihood test).

tumor cells and fibroblasts (Supplementary Fig. 1a). We used a panel of lineage markers (Lin) to exclude CD3⁺ T cells, CD19⁺ B cells and CD56⁺ cells (Supplementary Fig. 1a). We analyzed expression of the co-receptor CD14 independently of the lineage channel to efficiently identify CD14⁺ DCs, which have been reported in patients with cancer^{20,21,30–32}. Among Lin⁻ cells, we next gated on CD11c⁺HLA-DR^{hi} cells to exclude CD11c⁺HLA-DR^{neg-lo} myeloid-derived suppressor cells³³. HLA-DR⁺CD123⁺ pDCs were identified in the CD11c⁻ gate (Supplementary Fig. 1a).

In the Lin⁻CD45⁺ gate, we identified four distinct CD11c⁺ cell populations defined by their expression of the antigen-presenting molecule CD1c and CD14 (Fig. 1a). On the basis of published standardized nomenclature for blood DC subsets³⁴, CD1c⁺CD14⁻ cells matched the definition of cDC2s, the CD1c⁻CD14⁻ cell population included cDC1s, and CD1c⁻CD14⁺ cells were monocytes-macrophages (called 'MonoMacs' here) (Fig. 1a). We also identified a CD1c⁺CD14⁺ cell population that co-expressed markers of monocytes and macrophages, such as CD14, CD64 and CD163, and cDC2 markers, such as CD1c, CD206 and FcεRI (Fig. 1b and Supplementary Fig. 1b). Because these CD1c⁺CD14⁺ cells were phenotypically distinct from MonoMacs, and because they had not been systematically distinguished in published studies³⁴, we call them 'CD14⁺ DCs' here. CD56⁺CD14⁺ cells were reported to be interferon-producing killer DCs in the context of cancer³⁵.

and were subsequently shown to correspond to activated natural killer cells³⁶. A similar CD56⁺CD14⁺ phenotype has been described for fraction of blood monocytes from healthy donors³⁷. We detected CD56⁺CD14⁺ cells in breast-cancer samples (18% of live CD45⁺CD3⁻CD19⁻ cells) (Supplementary Fig. 1c). Because of their controversial nature, we excluded them through the use of antibody to the adhesion molecule CD56 in our 'lineage cocktail' (Supplementary Fig. 1c).

The C-type lectin-like receptor Clec9A could not be used to identify cDC1s, as it was degraded during enzymatic digestion of the tissue (Supplementary Fig. 1d). Thrombomodulin (CD141 or BDCA3) was expressed promiscuously by all DCs, including pDCs and MonoMacs (Fig. 1b). However, CD141^{hi} cells were found only in the CD1c⁻CD14⁻ population (Fig. 1b); hence, the CD1c⁻CD14⁻ population showed considerable enrichment for cDC1s. Because CD141^{hi} cells were too few in number (<100 cells per sample) and rare (5–50% of CD141^{hi} cells among CD1c⁻CD14⁻ cells in only half of the patients) to allow further separation into subsets, we designated the CD1c⁻CD14⁻ cell subset 'cDC1-enriched' (cDC1e) cells and used it for further molecular characterization. MonoMacs, CD11c⁺HLA-DR^{neg-lo} cells, CD14⁺ DCs, cDC2s and cDC1e cells did not express the natural killer cell receptor CD16 (FcγRIII) (Fig. 1b and data not shown). MonoMacs, CD14⁺ DCs and cDC2s had high expression of CD32B, which has been reported on a non-inflammatory subset of cDC2s in the blood³⁸, but cDC1e cells did not. The receptor tyrosine kinase AXL, which is expressed by precursors of blood DCs and by cDC2s, was expressed mainly by cDC2s, CD14⁺ DCs and MonoMacs in breast tumors (Fig. 1b). This indicated a clear discrepancy between blood and breast tissue in terms of DC markers.

To investigate the morphology of tumor APCs, we sorted them and analyzed their cytological features. pDCs had a typical plasma-cytoid morphology⁹, while cDC2s, cDC1e cells and CD14⁺ DCs had a dendritic morphology with high ratio of nucleus to cytoplasm and a less-basophilic cytoplasm than that of pDCs (Fig. 1c). MonoMacs had an acidophilic cytoplasm with abundant vacuoles (Fig. 1c), as is commonly observed in this population.

We quantified the distinct APC subsets across 22 luminal breast cancer (LBC) samples. MonoMacs were the most abundant cells (6.1% (median value) of CD45⁺ cells), followed by CD14⁺ DCs, and pDCs (0.5% and 0.3%, respectively, of CD45⁺ cells). cDC1e cells and cDC2s were the least abundant APCs (0.2% of CD45⁺ cells) (Fig. 1d). This phenotypic analysis identified and quantified five APC populations that infiltrated human breast cancer: MonoMacs, cDC2s, CD14⁺ DCs, pDCs and cDC1e cells.

Tumor-infiltrating DCs show enrichment for human blood DC signatures. Because the number of APCs obtained from primary breast cancer samples after sorting was very low (range, 2–12,000 cells), we adapted a protocol aimed at obtaining robust transcriptomes from rare cell populations by RNA-based next-generation sequencing (RNA-seq) (Supplementary Fig. 1e). We analyzed only those cell populations with more than 100 events. We generated RNA-seq profiles for pDCs, cDC2s, cDC1e cells, CD14⁺ DCs and MonoMacs from 13 patients with LBC (Supplementary Table 1), with 44 transcriptomes passing all quality-control criteria (Supplementary Table 2). On average, 60.5% of reads were mapped to the reference transcriptome across all samples. After filtering and normalizing the raw RNA-seq data, we obtained an average of 14,417 expressed genes.

To verify the identity of each of the subsets at the RNA level relative to the flow-cytometry analysis, we checked the expression of genes encoding various subset-specific and shared DC markers (Fig. 1e). As expected, pDCs had high expression of *IL3RA*, *CLEC4C* and *TLR9*; cDC2s had high expression of *CD1A*, *CD1B* and *FCER1A*; *CLEC9A*, *XCR1* and *BATF3* (all markers of cDC1s)

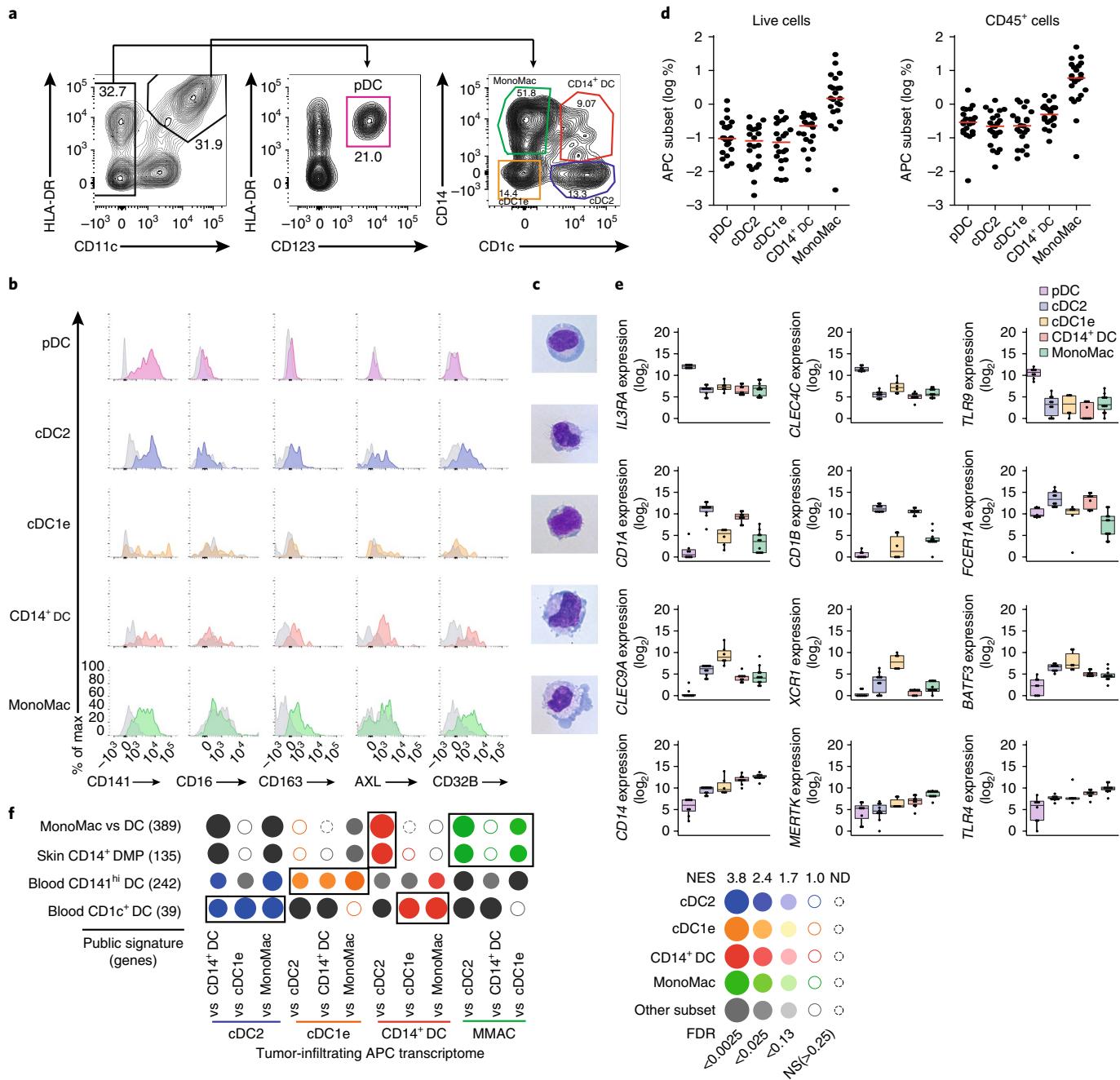


Fig. 1 | Phenotypic and molecular characterization of innate APCs that infiltrate breast cancer tissue. **a**, Flow cytometry showing the gating strategy used to distinguish DC subsets from MonoMacs in breast-cancer samples. Numbers in or adjacent to outlined areas indicate percent cells in each (subset designations included in plots here). **b**, Expression of CD141, CD16, CD163, AXL and CD32B by various APC subsets (left margin) in breast-cancer samples ($n=3$ donors, with similar results), presented as mean fluorescent intensity. **c**, Microscopy of Giemsa-stained cytopsin preparations, showing the morphology of APCs sorted by flow cytometry from tumors ($n=3$ donors, with similar results). Original magnification, $\times 100$. **d**, Frequency of APC subsets (horizontal axis) among total live cells (left) or $CD45^+$ cells (right), assessed by flow cytometry. Each symbol represents an individual donor ($n=22$); small red horizontal lines indicate the median. **e**, Expression of genes encoding DC-selective markers, by APCs (key) isolated from tumors, presented as read counts + 1 (\log_2 values). Each symbol represents an individual sample ($n=8$ (pDC), $n=10$ (cDC2), $n=6$ (cDC1e), $n=9$ (CD14⁺ DC) and $n=11$ (MonoMac)); middle line indicates the median, box limits indicate the first and third quartiles, and 'whiskers' indicate 'extreme' data points no more than 1.5 \times the length of the box beyond the box limit. **f**, Enrichment for various APC public signatures⁵⁹ (left margin) in pairwise comparisons (below plot) of the transcriptomes of APCs (key color) isolated from tumors (n values as in **e**), plotted with the BubbleMap module of BubbleGUM software; color intensity and symbol size indicate the normalized enrichment score (NES) and FDR, respectively (key at right); outlined areas indicate the expected signature-enrichment analysis. DMP, dermal mononuclear phagocyte.

were 'preferentially' expressed by cDC1e cells; MonoMacs had high expression of *CD14*, *MERTK* and *TLR4*; and CD14⁺ DCs shared the expression of *FCER1A* and *CD14* with cDC2s and MonoMacs,

respectively (Fig. 1e). Gene set-enrichment analyses using public datasets indicated that breast-cancer cDC2s had the highest normalized enrichment score (NES) with the blood cDC2 signature;

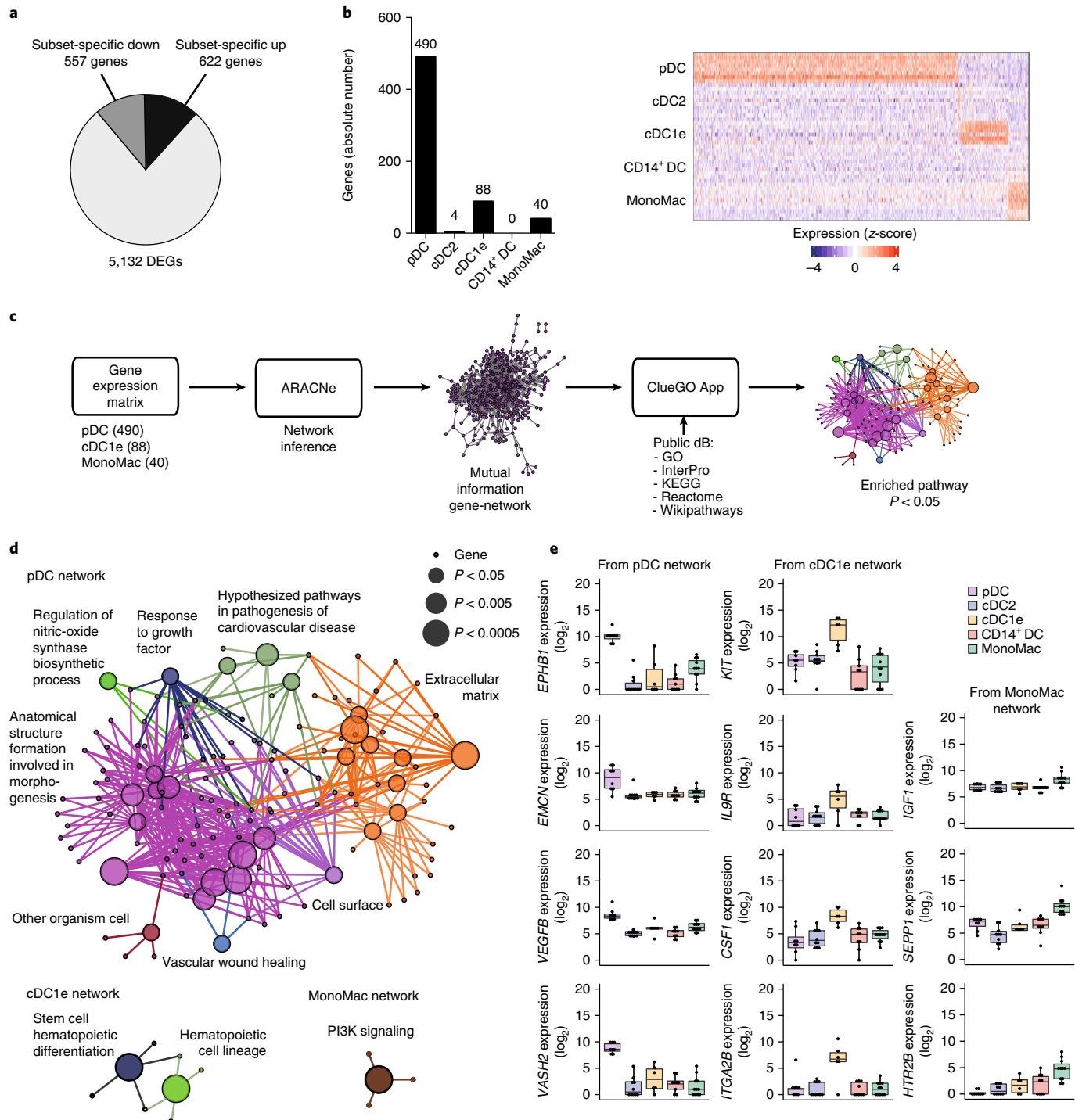


Fig. 2 | Subset-specific signatures that define tumor APCs. **a**, Proportion of genes upregulated (up) or downregulated (down) among DEGs in APC subsets isolated from tumors. $P < 0.05$ (one-way analysis of variance (ANOVA) and Tukey's post-hoc correction). **b**, Quantification of genes upregulated in one subset (horizontal axis) relative to their expression in all other subsets (left), and expression (z-score; key) of each gene (one per column) in the APC subsets (right margin; one sample per row) (right). Numbers above bars (left) indicate specific values. **c**, Bioinformatics pipeline used for functional inference (in **d**) from subset-specific gene signatures; numbers below plot at far left indicate the number of genes in each group. ARACNe, algorithm for the reconstruction of accurate cellular networks; ClueGO App, plug-in for the Cytoscape software platform for visualizing complex networks; dB, database (GO, gene ontology knowledgebase; InterPro, database of protein families, domains and functional site; KEGG, Kyoto encyclopedia of genes and genomes; Reactome, pathway database; Wikipathways, database of biological pathways). **d**, Functional network inference showing the biological pathways (along perimeter) most significantly overrepresented (FDR < 0.05) in the gene signatures of pDCs, cDC1e cells and MonoMacs (n values as in **e**) from tumors: color indicates pathway; node size indicates P value; smallest symbols indicate the pathway-associated genes (key). PI3K, phosphatidylinositol-3-OH kinase. **e**, Expression of genes in the pathways most significantly enriched in **d**, for pDCs ($n=8$ samples), cDC1e cells ($n=6$ samples) and MonoMacs ($n=11$ samples) (above plots), presented as in Fig. 1e.

breast-cancer cDC1e cells had the highest NES with the blood cDC1 (CD141^{hi}) signature; and breast-cancer MonoMacs had the highest NES with the CD14⁺ dermal mononuclear phagocyte and MonoMac signatures, compared with all the other gene signatures. Finally, breast-cancer CD14⁺ DCs shared the highest NES with the blood cDC and MonoMac skin CD14⁺ dermal mononuclear phagocyte signatures (Fig. 1f). Hence, robust transcriptional profiles confirmed the identity of the main DC subsets and MonoMacs that infiltrated breast cancer.

Tumor-infiltrating DC harbor subset-specific signatures. We performed analysis comparing gene expression in pDCs, cDC2s, cDC1e cells, CD14⁺ DCs and MonoMacs and identified 5,132 genes that were expressed differentially in at least one subset relative to their expression in all other APCs ('differentially expressed genes' (DEGs)) (Fig. 2a). We then applied a post-hoc test to extract the genes upregulated in each type of APC relative to their expression in all other subsets, which we defined as its subset-specific signature. From a total of 662 subset-specific genes, 490 corresponded to pDCs, 88 corresponded to cDC1e cells, 40 corresponded to MonoMacs and 4 corresponded to cDC2s (Fig. 2b). We found no genes specifically upregulated in CD14⁺ DCs (Fig. 2b).

Among the ten DEGs with the highest significance, genes encoding the oncoprotein TCL1A and the anti-apoptotic molecule ZFAT were found in the pDC signature; genes encoding the glutamate receptor GRIP and the cytokines CCL22 and IL-29 (*IFNL1*) were found in the cDC2 signature; genes encoding the plasma-membrane receptors IL-33R (ST2; *IL1RL1*) and XCR1 were found in the cDC1e cell signature; and genes encoding the fatty acid-biosynthesis enzymes ASAHI and ME1 were found in the MonoMac signature (Supplementary Table 3).

We then identified functions linked to each subset-specific signature (Fig. 2c,d). From a total of 29 pathways (false-discovery rate (FDR), <0.05) found in the pDC gene network, the most significantly enriched pathway was 'anatomical structure involved in morphogenesis' (FDR = 2.7×10^{-07}), and this included *EPHB1*, *VEGFB* and *VASH2* (Fig. 2e,f). The cDC1e cell gene network showed enrichment for two pathways, both linked to hematopoiesis, and this included *KIT*, *IL9R*, *CSF1* (which encodes the cytokine M-CSF) and *ITGA2B* (Fig. 2e,f). The MonoMac gene signature showed enrichment for only the pathway 'PI3K signaling', which included *IGF1*, *SEPP1* and *HT2RB* (Fig. 2e,f). Thus, subset-specific genes were identified for LBC-infiltrating pDCs, cDC2s, cDC1e cells and MonoMac. Notably, none of those subsets showed differential enrichment for any pathway directly linked to immunological function.

DC plasticity in the tumor microenvironment is subset specific. To determine how tumor-infiltrating APCs adapt to their microenvironment, we analyzed non-malignant tissue juxtaposed to tumor tissue ('juxta-tumoral' tissue) from eight donors (pairing tumor tissue with juxta-tumoral tissue from the same patient). The pDC, cDC2, cDC1e, CD14⁺ DC and MonoMac populations that we described in the tumors were also identified in the juxta-tumoral tissue, but with a lower frequency among the CD45⁺ cells than in the tumor, a result that was significant for pDCs ($P=0.078$) and cDC1e cells ($P=0.039$ (likelihood ratio test); Fig. 3a and Supplementary Fig. 2a). We generated transcriptional profiles for each APC subset in the juxta-tumoral tissue using the RNA-seq workflow used for the tumor DC subsets; the transcriptomes were generated in parallel, were run in the same batch as their tumoral counterpart and were matched for each donor (Supplementary Fig. 2b). We compared the transcriptome of each APC subset in the tumor with that of the juxta-tumoral sample (Supplementary Fig. 2b). We identified 607 DEGs for pDCs, 348 DEGs for CD14⁺ DCs, 236 DEGs for MonoMacs, 45 DEGs for cDC1e cells and 22 DEGs for cDC2s, which resulted in a total of 1,258 DEGs (FDR < 0.05 , and

a change in expression of over onefold (log₂ values)) that were used for further analysis (Fig. 3b). DEGs from all DC subsets had higher expression in the tumor than in the juxta-tumoral tissue (Fig. 3b). We identified seven genes with the highest significance (FDR = 1.72×10^{-17} to 4.1×10^{-10}) among DEGs of tumor CD14⁺ DCs and juxta-tumoral CD14⁺ DCs compared with DEGs from other APC subsets; these included genes encoding the secretoglobulins TFF1 and TFF3, which have a function in mucosal healing. Conversely to DCs, DEGs from MonoMacs were upregulated mostly in juxta-tumoral samples (195 DEGs) rather than tumor samples (41 DEG). Among the genes most significantly upregulated in juxta-tumoral MonoMacs was the gene encoding the scavenger receptor ligand CD163L, which is associated with M2 polarization (Fig. 3b).

Among the five transcripts whose expression was the most increased in tumor APCs relative to their expression in juxta-tumoral APCs ('top five'), we detected transcript encoding the negative regulator CD5 in pDCs (Fig. 3c) and transcripts encoding the secretoglobulins SCGB2A2 and SCGB1D2 in cDC2s. SCGB2A2 was also among the top five DEGs of CD14⁺ DCs and pDCs in the tumor-versus-juxta-tumoral comparison (Fig. 3c and Supplementary Fig. 2b). The gene encoding TACI (*TNFRSF13B*), a member of the cytokine TNF receptor superfamily, was among the top five DEGs upregulated in tumor cDC1e cells relative to their expression in juxta-tumoral cDC1e cells, whereas the gene encoding the chemokine CCL7 was substantially upregulated in tumor MonoMacs relative to its expression in juxta-tumoral MonoMacs (Fig. 3b). The gene encoding AGR2 (a protein disulfide isomerase needed for mucin folding) was among the genes with the most significant upregulation in tumor cDC2s, CD14⁺ DCs and MonoMacs relative to their expression in the juxta-tumoral counterparts of those cells (Fig. 3b).

We next analyzed whether the genes expressed differentially by tumor APCs relative to their expression in juxta-tumoral APCs were shared across subsets. Strikingly, most of the genes were expressed differentially exclusively in one subset (1,074 genes) or two subsets (184 genes) (Fig. 3d). Only 21 DEGs were shared with two other subsets and none were shared with three or four other subsets (Fig. 3d,e). This indicated that the tumor-induced transcriptional reprogramming of APCs was subset specific.

The differential expression of SCGB2A2, a gene previously associated with mammary epithelial tumor cells^{39,40}, raised questions about its tumor specificity versus its immune-cell specificity⁴¹. We excluded the possibility of contamination by tumor-cell mRNA by our stringent gating strategy (Fig. 1a and Supplementary Fig. 1) and by the observation that epithelium-specific mRNA, such as *EPCAM*, was not detected among the DEGs in tumor pDCs (Supplementary Fig. 2b). Given that SCGB2A2 was detected in a transcriptome analysis of blood pDCs from healthy donors⁴², these observations suggested that pDCs might express SCGB2A2 mRNA endogenously at steady state and in inflammatory conditions. These observations indicated that DCs adapted to the tumor microenvironment in a subset-specific manner.

Immunological pathways are absent from APC 'tumor-emerging genes'. For each APC separately, we analyzed the functions linked to molecules encoded by 'tumor-emerging genes' (DEGs upregulated in tumor cells relative to their expression in juxta-tumoral cells), meaning those functional pathways for which in the tumor APC showed enrichment, relative to their presence in the corresponding juxta-tumoral APC. Pathway-enrichment analysis identified the pathways 'actomyosin structure organization' and 'proteinaceous extracellular matrix' in pDCs; 'receptor protein tyrosine kinase signaling' in CD14⁺ DCs; and 'kinetochore' in MonoMacs (Fig. 4a). The major molecules driving the enriched pathways included the growth factor CTGF in pDCs, AGR2 in CD14⁺ DCs, and the mitotic checkpoint BUB1 in MonoMac (Fig. 4b). Because

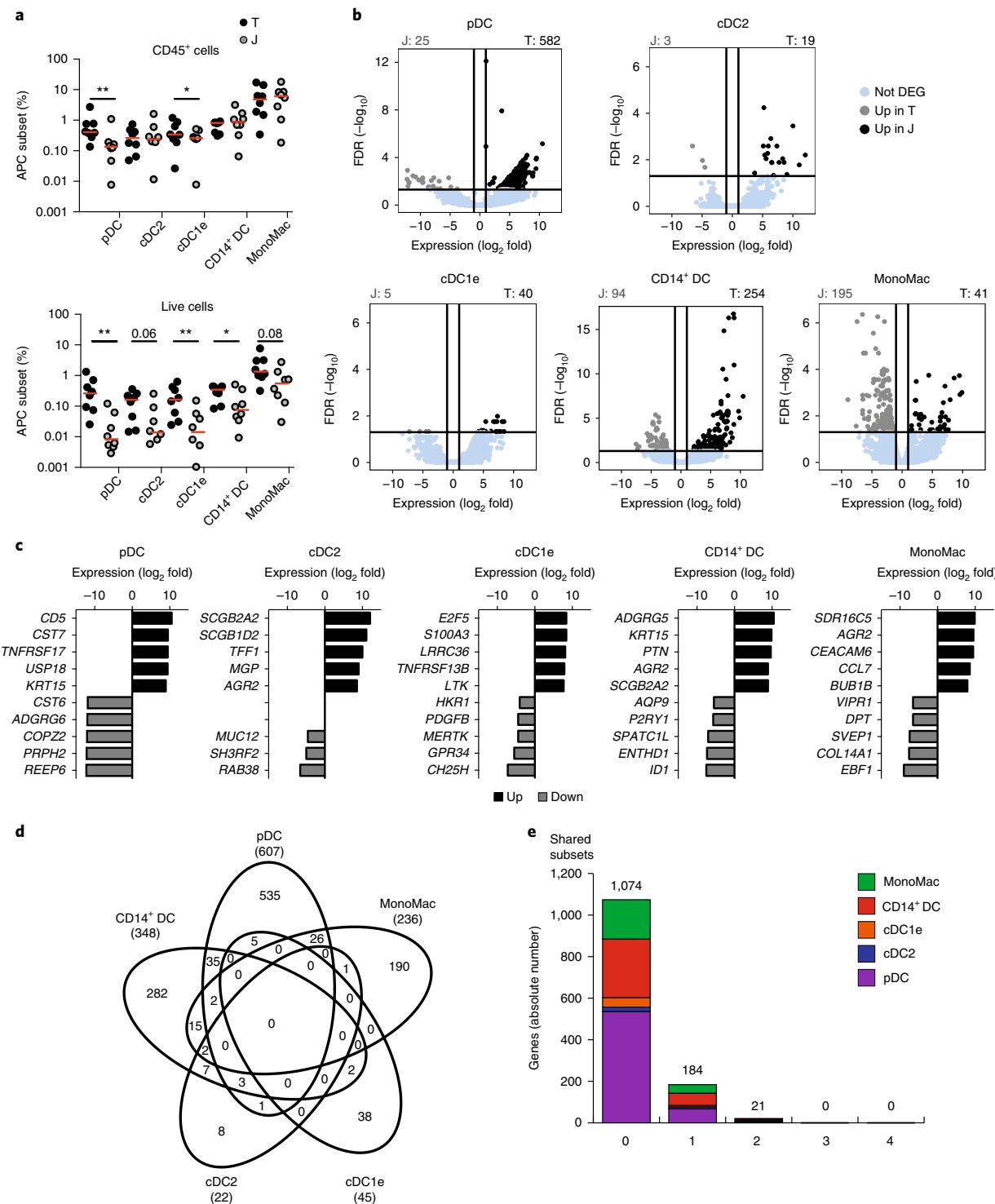


Fig. 3 | 'Tumor-emergent genes' from innate APC are subset specific. **a**, Frequency of APC subsets (horizontal axis) among CD45⁺ cells (top) or total live cells (bottom) in tumor samples (T) and juxta-tumoral samples (J) (key) from patients with LBC, assessed by flow cytometry. Each symbol represents an individual donor ($n=8$; samples paired by donor); small red horizontal lines indicate the median. $*P < 0.05$ and $**P < 0.01$ (two-tailed Wilcoxon-test).

b, Expression of genes in tumor samples relative to their expression in juxta-tumoral samples (horizontal axis), plotted against the FDR ($FDR < 0.05$; vertical axis), for the transcriptome of each APC subset (above plot), showing genes upregulated in the tumor sample (change in expression of over 1-fold (\log_2 value)) or juxta-tumoral sample (change in expression of less than -1-fold (\log_2 value)) or unchanged (Not DEG) (key); numbers above plots indicate the number of DEGs in each group. **c**, Expression of the top five DEGs (left margin) upregulated or downregulated (key) in each APC subset in the tumor, presented relative to their expression in the juxta-tumoral sample (\log_2 value). **d**, Quantification of total DEGs in tumor samples versus juxta-tumoral sample, for each subset (in parenthesis along perimeter), and DEGs shared by various subsets (overlapping loops) or all subsets (center). **e**, Quantification of DEGs unique to a specific subset (0) or shared with one, two, three or four other subsets (horizontal axis), for each APC subset (key); numbers above bars indicate total DEGs per shared group. Number of independent matched donors with LBC (**b–e**) (likelihood ratio test from edgeR R software package): $n=3$ (pDC), $n=4$ (cDC2), $n=4$ (cDC1e), $n=3$ (CD14⁺ DCs) and $n=5$ (MonoMac).

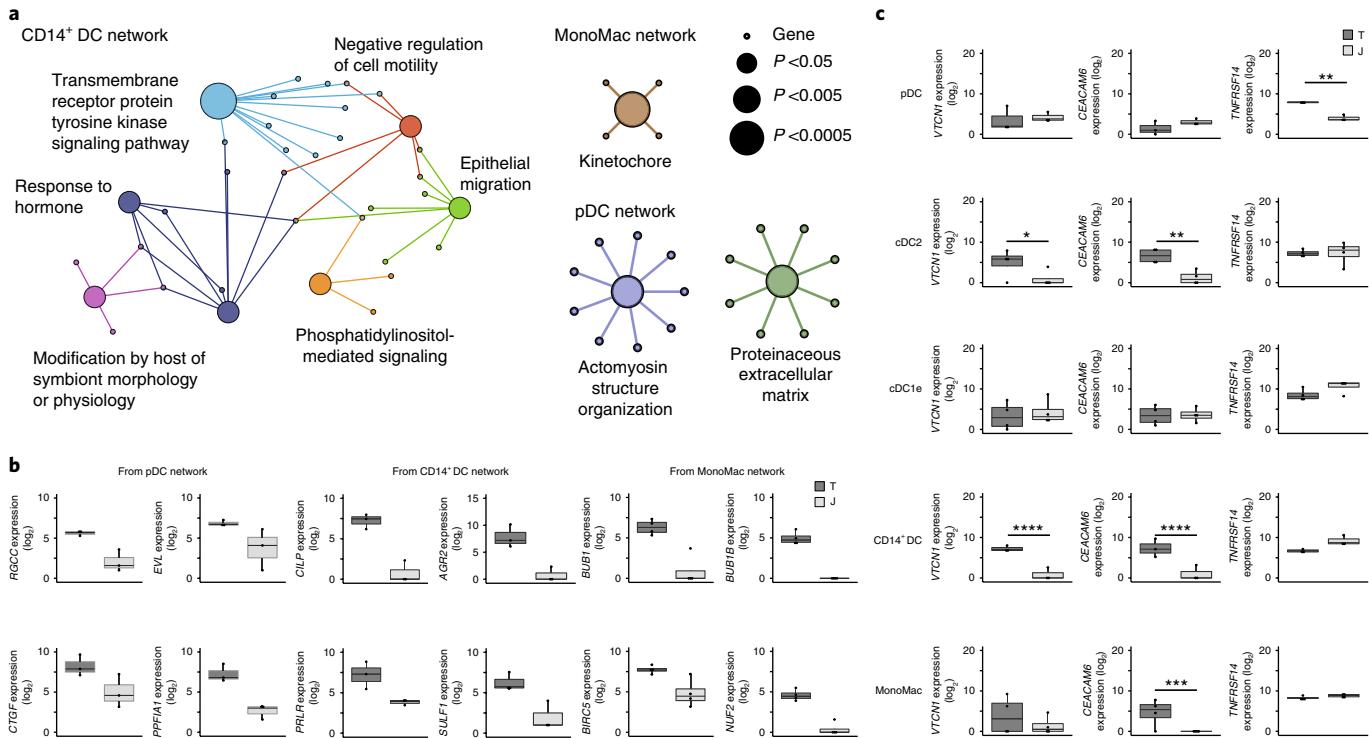


Fig. 4 | Absence of immunological function enrichment in tumor-upregulated genes. **a**, Functional network inference results for the gene signatures (FDR < 0.05) of pDCs, CD14⁺ DCs and MonoMacs (above plots) from LBC samples, presented as in Fig. 2d. **b**, Expression of genes encoding molecules in the pathways most significantly enriched (as in **a**) for pDCs, CD14⁺ DCs and MonoMacs (above plots) in tumor or juxta-tumoral samples (key) from LBC, presented as read counts + 1 (\log_2 values); box plots as in Fig. 1e. **c**, Expression of the checkpoint molecule-encoding genes VTCN1, CEACAM6 and TNFRSF14 (entire list, Methods) in tumor and juxta-tumoral samples (key) from LBC, presented as in **b**; only genes expressed differentially in at least one subset are presented. *P < 0.05, **P < 0.01, ***P < 0.005 and ****P < 0.0001 (likelihood ratio test, edgeR). Number of independent matched donors with LBC (**a–c**): n = 3 (pDC), n = 4 (cDC2), n = 4 (cDC1e), n = 3 (CD14⁺ DC) and n = 5 (MonoMac).

we did not identify any enrichment for immunological function by this unbiased approach, we specifically investigated the expression of immunological checkpoint molecules important in anti-tumor immunity^{28,29}. Out of 19 positive and 15 negative immunological checkpoint molecules, we found that genes encoding the following were expressed differentially in tumor APCs relative to their expression in juxta-tumoral APCs: HVEM (TNFRSF14) in pDCs; B7-H4 (VTCN1) and (CEACAM6) in cDC2s and CD14⁺ DCs; and CEACAM6 in MonoMacs (Fig. 4c). In conclusion, the molecules encoded by ‘tumor-emerging genes’ from LBC APCs were poorly linked to immunological functions.

The transcriptomics profile of tumor APCs depends on the breast-cancer subtype. To evaluate the effect of tumor type on the DC transcriptional profile, we generated the transcriptomes of pDCs, cDC2s and CD14⁺ DCs from four triple-negative breast cancer (TNBC) samples and of cDC1e cells and MonoMacs from four TNBC samples (Supplementary Fig. 3 and Supplementary Table 1). Principal-component analysis of tumor DC transcriptional profiles using the 500 genes with the most-variant expression indicated that DCs clustered by cancer subtype rather than by DC subset (Fig. 5a), suggestive of differential tumor imprinting on DCs. pDCs separated from the other APC subsets in both cancer types (Fig. 5a). To identify the genes upregulated in TNBC relative to their expression in LBC for each DC subset, we performed differential analysis (FDR < 0.05, and a change in expression of over onefold (\log_2 values)). MonoMacs had the greatest number of DEGs (2,930), followed by CD14⁺ DCs (2,662s) and pDC (1,434) (Fig. 5b). cDC1e cells (605 DEGs) and cDC2s (521 DEGs) were less affected by tumor

type (Fig. 5b). The majority of DEGs (65% of genes upregulated in TNBC relative to their expression in LBC) were upregulated exclusively in one DC subset (Fig. 5c). Four DEGs (IFNL1, IFNB1, ISG20 and ISG15), all associated with the interferon pathway, were upregulated in TNBC relative to their expression in LBC (Fig. 5d). These data indicated that two different types of cancer had a major effect on the transcriptomes of the infiltrating DCs and MonoMacs.

TNBC promotes a shared immune system-related signature in DCs. The pDCs had the greatest number of enriched pathways (166) relative to the number of enriched pathways in other APCs (Fig. 6a). MonoMacs, cDC2s and CD14⁺ DCs shared 49%, 36% and 29%, respectively, of their enriched pathways with at least one other subset (Fig. 6a). In contrast, cDC1e cells shared only 6% of their enriched pathways with other subsets (Fig. 6a). These results suggested that pathways for which TNBC APCs showed enrichment were mostly subset specific, indicative of functional specialization for each subset.

We then focused on the pathways for which TNBC APCs commonly showed enrichment. We identified 38 pathways, including those linked to immune system-related functions, that were shared with at least another APC subset (Fig. 6b and Supplementary Fig. 4a). In particular, ‘chemokine activity’, ‘cytokine activity’, ‘cytokine receptor binding’ and ‘IL-10 signaling’ were shared by cDC2s and CD14⁺ DCs (Supplementary Fig. 4a). All DC subsets commonly showed enrichment for type I interferon-related pathways, such as ‘IFN α / β signaling’ and ‘negative regulation of viral life cycle’ (Fig. 6b). From all type I interferon-related pathways, we selected the genes that showed significant enrichment, including IFNB1, ISG15 and

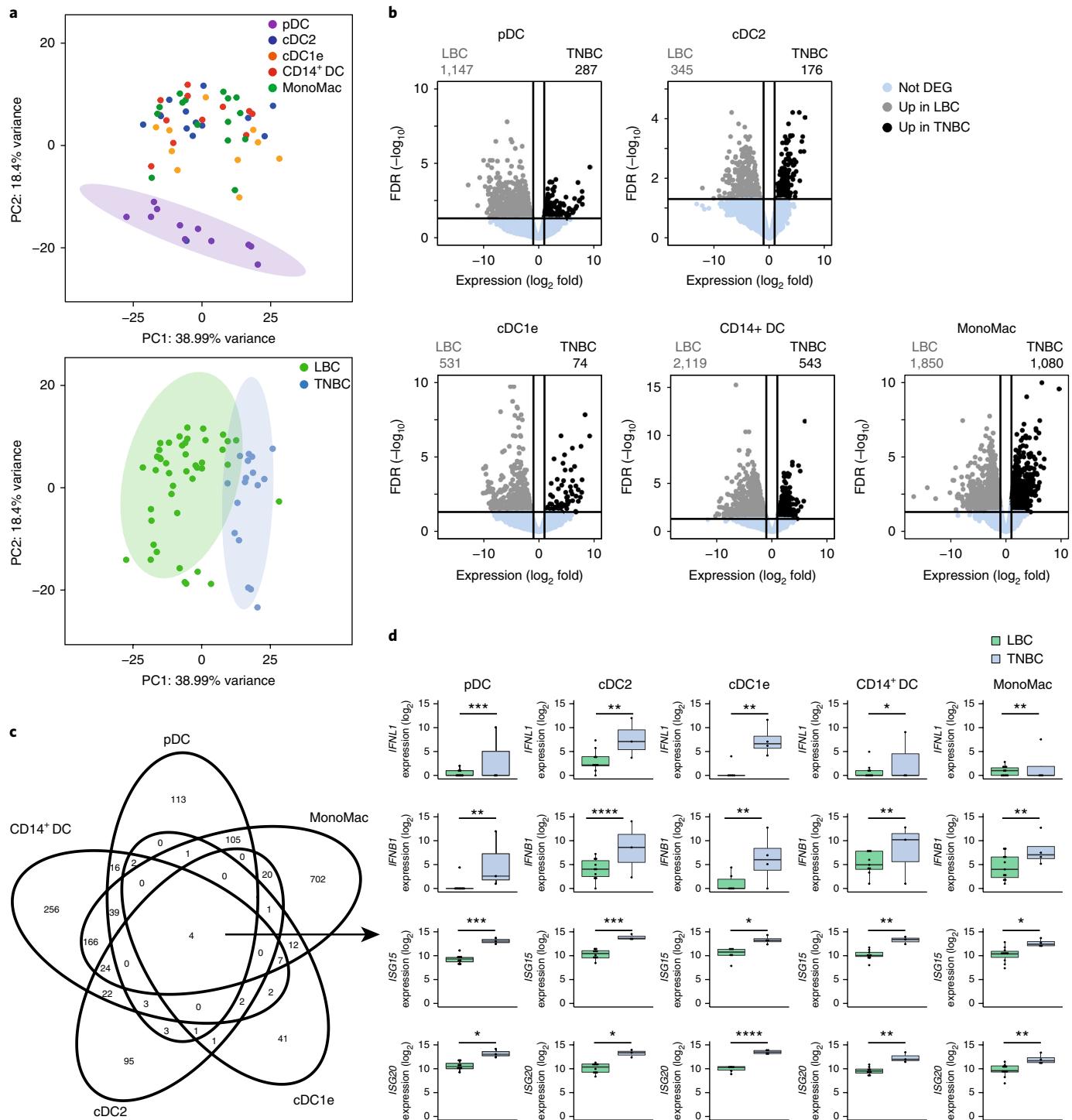


Fig. 5 | Transcriptional profile of innate APC subset is dependent on breast-cancer subtype. **a**, Principal-component analysis showing the clustering of transcriptional profiles (with the 500 most variant genes) of innate APC subsets isolated from LBC or TNBC tumors, with principal components PC1 and PC2 projected and the variance of each (along axes), presented by subset (key; top) or by breast-cancer type (key; bottom). **b**, Expression of genes in TNBC relative to their expression in LBC (horizontal axis), plotted against the FDR ($\text{FDR} < 0.05$; vertical axis), for the transcriptome of each APC subset (above plot), showing genes upregulated in TNBC (change in expression of over 1-fold (\log_2 value)) or LBC (change in expression of less than -1-fold (\log_2 value)) or unchanged (key); numbers above plots indicate the number of DEGs in each group. **c**, Quantification of DEGs shared by various subsets of APCs (overlapping loops) or all subsets (center), among DEGs upregulated in TNBC relative to their expression in LBC. **d**, Expression of the four genes upregulated in all APC subsets in TNBC (center in **c**), assessed in the five APC subsets (above plots) from LBC or TNBC (key), presented as read counts + 1 (\log_2 values); box plots as in Fig. 1e. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.005$ (likelihood ratio test, edgeR). Number of samples and donors (**a-d**): $n = 8$ (pDC), $n = 10$ (cDC2), $n = 6$ (cDC1e), $n = 9$ (CD14⁺ DC) and $n = 11$ (MonoMac), from six to ten donors with LBC; and $n = 3$ (pDC), $n = 3$ (cDC2), $n = 4$ (cDC1e), $n = 3$ (CD14⁺ DC) and $n = 4$ (MonoMac), from three to four TNBC donors.

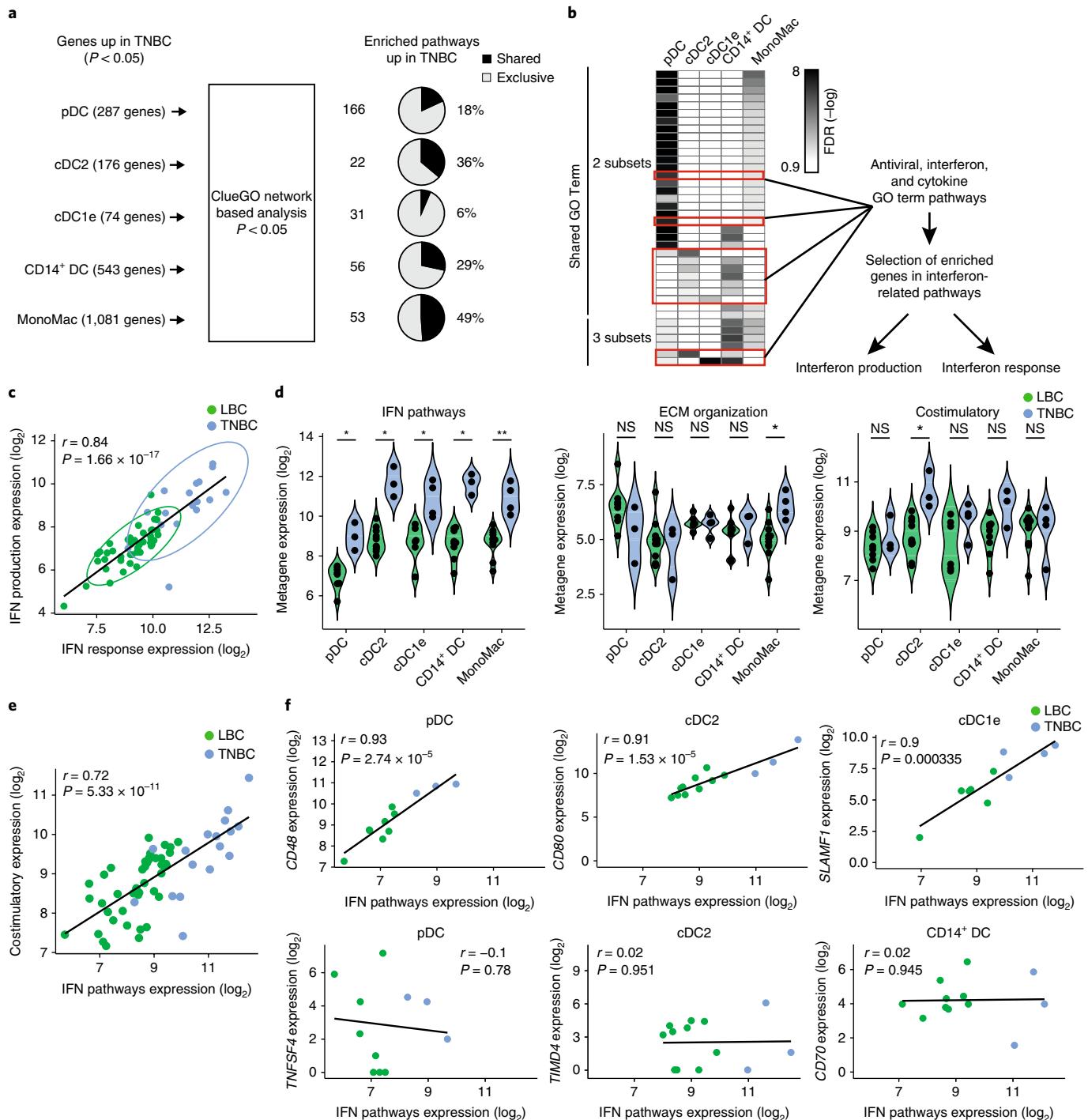


Fig. 6 | The type I interferon pathway is upregulated in all APC subsets in TNBC. **a**, Functional pathways analysis of DEGs upregulated in APCs in TNBC, showing the resultant number of genes in each (left), the number of pathways (middle) and the proportion of shared or specific pathways (key) for each APC subset (right). **b**, Significance (FDR, $-\log$ values) for enriched pathways ($FDR < 0.05$) shared by two or three (left margin) APC subsets (above plots); red outlines indicate the immunological pathways at right (interferon metagene extracted from significantly enriched pathways categorized as interferon production or interferon response). **c**, Correlation between the expression (\log_2 values) of genes encoding molecules involved in the interferon response (IFN response expression) and those involved in interferon production (IFN production expression), for all APC subsets isolated from LBC or TNBC (key). **d**, Expression of the IFN pathways metagene, the 'ECM organization' pathway metagene and the 'Costimulatory' pathway metagene (above plots) for each APC subset (horizontal axis) from LBC or TNBC (key), presented as a violin plot. NS, not significant ($P > 0.05$); * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.005$ (two-sided Wilcoxon test). **e**, Correlation between expression (\log_2 values) of the IFN pathways metagene and 'Costimulatory' pathway metagene in all APC subsets isolated from LBC or TNBC (key). **f**, Correlation between expression (\log_2 values) of the IFN pathway metagene and the costimulatory molecule-encoding genes CD48 (SLAMF2), CD80, SLAMF1, TNFSF4 (OX40L), TIMD4 and CD70 in all APC subsets (above plots) from LBC or TNBC (key). Numbers in plots (**c,d,f**) indicate the correlation coefficient (r) and P value (Pearson correlation test). Number of samples: $n = 8$ (pDC), $n = 10$ (cDC2), $n = 6$ (cDC1e), $n = 9$ (CD14⁺ DC) and $n = 11$ (MonoMAC) for LBS, and $n = 3$ (pDC), $n = 4$ (cDC1e), $n = 3$ (CD14⁺ DC) and $n = 4$ (MonoMAC) for TNBC.

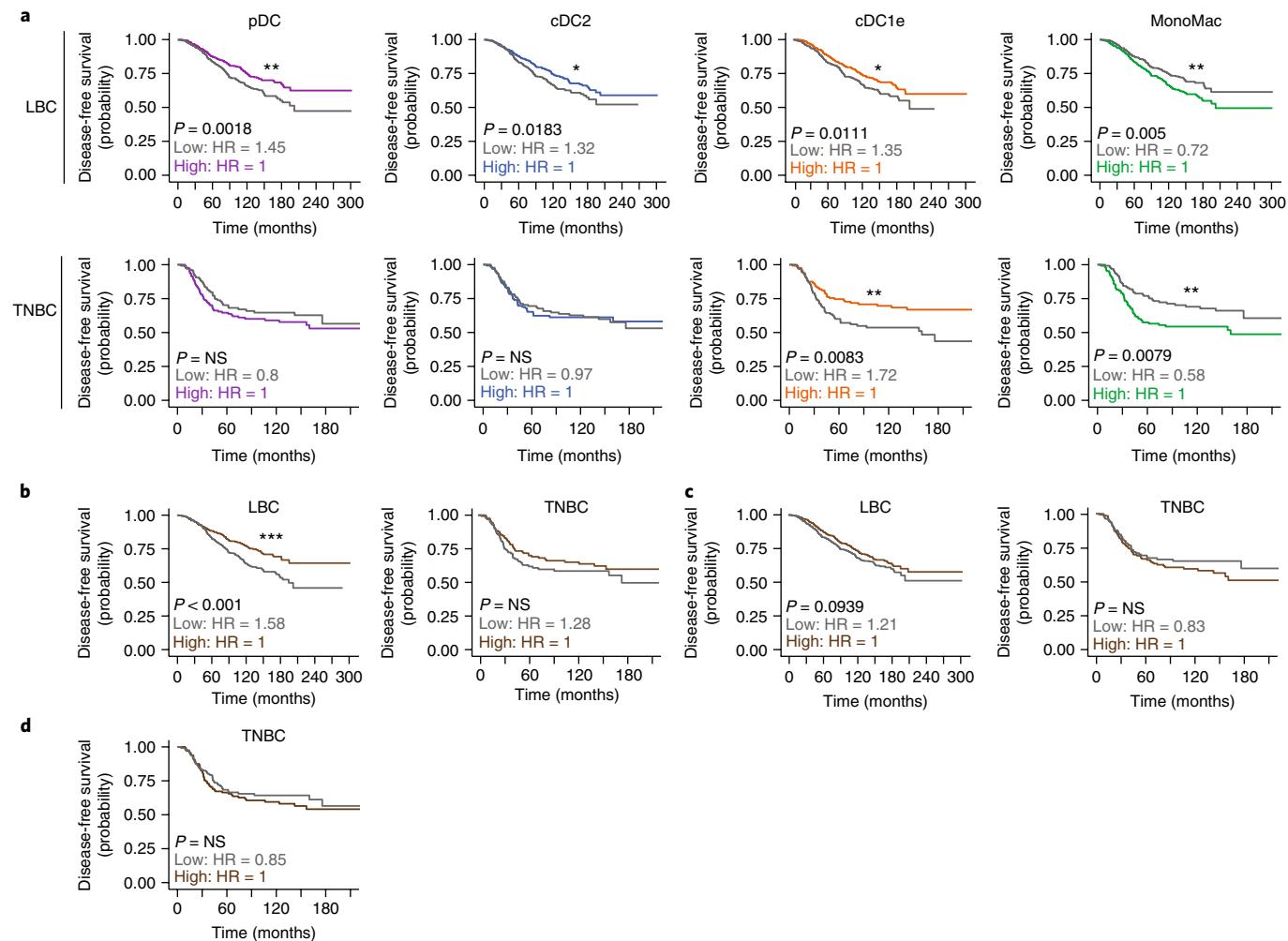


Fig. 7 | Subset-specific signatures are linked to distinct disease-free survival depending on the subset and breast cancer type. **a**, Probability of disease-free survival over time for patients with LBC (top row) or TNBC (bottom) (Kaplan–Meier plots), associated with a high or low z-score (line color) for the signature specific to each APC subset (above plots): line colors indicate the hazard ratio (bottom left in each plot) associated with a low z-score (Low: HR) or high z-score (High: HR); P values (bottom left in each plot), log-rank test. **b,c**, Probability of disease-free survival over time for patients with LBC (left) or TNBC (right), associated with the z-scores for a published CD103⁺ DC gene signature⁴⁴ (**b**) or a published pDC signature⁷ (**c**), presented as in **a**. **d**, Probability of disease-free survival over time in patients with TNBC, associated with z-scores for the APC ‘common IFN’ signature, presented as in **a**. *P < 0.05, **P < 0.005 and ***P < 0.001 (log-rank test). Data are from the METABRIC public dataset: n = 1,043 donors (LBC) and n = 259 donors (TNBC).

ISG20, and classified them as distinct ‘metagenes’ according to the contribution of the molecules encoded to either interferon production or the interferon response (Supplementary Fig. 4b). Because the amount of expression of each metagene was strongly correlated across all TNBC samples (Fig. 6c), we pooled them into a single ‘IFN pathways’ metagene, which had higher expression in all TNBC APCs than in LBC APCs (Fig. 6d). As a control, the ‘ECM organization pathway’ metagene (Supplementary Fig. 4c) had significantly higher expression only in TNBC MonoMacs (Fig. 6d). We also analyzed the expression of a ‘costimulatory’ metagene (Supplementary Fig. 4d) that had significantly higher expression in TNBC than in LBC only for cDC2s, not for other APCs (Fig. 6d) and whose expression was highly correlated with that of the ‘IFN pathways’ metagene (Fig. 6e). When analyzing the dependence of individual checkpoint molecule-encoding genes with the ‘IFN pathways’ metagene, we found that the expression of genes encoding molecules such as SLAMF2 (CD48) in pDCs, CD80 in cDC2s and SLAMF1 in cDC1e cells was highly correlated with expression of the ‘IFN pathways’ metagene (Fig. 6f). In contrast, the expression of TNFSF4 in pDCs, TIMD4 in cDC2s and of CD70 in CD14⁺ DCs

was not correlated with expression of the ‘IFN pathways’ metagene (Fig. 6f). This revealed two groups of checkpoint molecule-encoding genes that were associated differentially with the ‘IFN pathways’ metagene (Supplementary Fig. 4e). Thus, the transcriptomes of APCs in TNBC differed substantially from those of APCs in LBC, with a common ‘IFN pathways’ metagene being upregulated in all TNBC APCs, indicative of a specific contribution of TNBC to the reprogramming of APCs.

Subset-specific signatures of tumor APCs can be used to predict the survival of patients with breast cancer. To assess whether the APC subset-specific signatures might have a prognostic effect, we took advantage of the publicly available dataset from whole-breast-cancer transcriptome METABRIC, which includes patient-survival clinical annotation⁴³. Because of the differences in the APC transcriptional profiles, we investigated LBC and TNBC datasets separately. We calculated a z-score for each APC subset-specific signature⁴⁴ (Supplementary Fig. 5a). We found that high z-scores for pDCs, cDC2s and cDC1e cells were significantly predictive of disease-free survival for patients with LBC (Fig. 7a). On the contrary,

a high *z*-score for MonoMacs was linked to a bad prognosis for patients with LBC or TNBC (Fig. 7a). A high *z*-score for cDC1e cells was linked to a good prognosis for patients with TNBC, with a greater significance than that for LBC (Fig. 7a). The *z*-scores for pDCs and cDC2s had no prognostic value in TNBC (Fig. 7a). This suggested that the various signatures might have a different clinical effect according to DC subset and breast-cancer type.

A CD103⁺ DC gene signature has been reported to correlate with a good prognosis in several tumor types, including breast cancer⁴⁴. Using the METABRIC dataset, we found that the *z*-score for the CD103⁺ DC gene signature had a significant effect on the survival of patients with LBC but not on the prognosis of patients with TNBC (Fig. 7b). We then assessed the prognostic value of the blood pDC signature⁷. The score for the blood pDC signature had no significant effect on survival outcome for patients with LBC or TNBC (Fig. 7c). Hence, prognostic significance was most efficiently reached in a given tumor through the use of DC signatures generated from the same tumor type. Finally, no prognostic value associated with the signature of common DEGs associated with the interferon pathway was found for patients with TNBC (Fig. 7d), which showed that subset-specific signatures harbored more prognostic information than did a shared signature.

We then determined whether subset-specific signatures could be independently associated with survival when integrated with the Nottingham prognostic index, a reference clinical score that determines survival⁴⁵. We observed that all significant scores in univariate analysis were kept in the multivariate analysis for pDCs, cDC2s, cDC1e cells and MonoMacs in LBC and for cDC1e cells and MonoMacs in TNBC (Table 1), which indicated that subset-specific APC signatures in LBC and TNBC were independent prognostic factors associated with disease-free survival. These results demonstrated the relevance of generating subset- and breast cancer type-specific signatures in predicting clinical outcome.

Discussion

Here we used DC-specific markers to identify resident DC populations (cDC2s, cDC1s and pDCs), MonoMacs and subsets that shared many features with previously described inflammatory DCs (CD14⁺ DCs)^{2,21} to provide broad and systematic coverage of the APC subsets that have been identified in two types of breast cancer (LBC and TNBC).

Our analysis revealed that pDCs were the most distinct APC subtype, as reported before in various tissues and species^{5–7,10,46,47}. We propose that part of this pDC-specific signature is determined by ontogeny, as supported by various genes, such as *CLEC4C*, *GZMB* and *TCF4*, identified in the pDC signatures independently of tissue type^{6,7,10}. Other pDC signature genes, such as those encoding the basal membrane laminins *LAMA4*, *LAMB1* and *LAMC1*, not previously associated with a pDC-specific signature^{6,7}, might be attributed to tissue imprinting or to a combined effect of ontogeny and tissue-driven factors. In contrast to pDCs, CD14⁺ DCs and cDC2s had very close similarity with other subsets. Comparative analyses of DC subsets across multiple studies might identify conserved, ontogeny-determined signatures, rather than more plastic and environment-driven transcriptional modifications.

Among high-throughput studies of tumor-infiltrating APCs in the mouse^{48–50}, only two compared tumor tissue with non-tumor tissue^{48,50}, but such studies focused on a single APC population, such as CD11b⁺ DCs⁵¹ or macrophages^{48,50}, and did not systematically compare diverse APCs in the context of their adaptation to a tumor context. Here, by systematically comparing the transcriptome of each APC subset in tumor tissue with that in uninvolving juxta-tumoral tissue, we identified emergent features in tumor-infiltrating APCs relative to those in non-tumor-tissue APCs. This ‘imprinting’ was different for distinct APC subsets, both qualitatively and quantitatively, which indicated that in breast cancer, there is no unique

signature that can be attributed to tissue imprinting, as has been suggested for other anatomical sites^{6–8}. We propose that the effect of the tissue microenvironment on innate immune cells should be considered and interpreted in close interaction with subset-specific molecular features.

cDC1s have been proposed to be the main APC subset that drives antitumor responses in mouse tumor models in a type I interferon-dependent manner^{44,52,53}. In our study, cDC1e cells expressing genes encoding XCR1 and CLEC9A, as well as other cDC-specific markers, did not have higher expression of genes encoding molecules related to DC activation or antigen presentation than that of the other APC signatures in either LBC or TNBC. Moreover, all human APC transcriptomes from TNBC, not only cDC1e cells, showed enrichment for genes encoding molecules involved in the interferon response and interferon production; this indicated that, at least in human breast cancer, all DCs are able to upregulate an interferon signature. Further experiments are needed to determine whether cDC1s are key to antitumor immune responses in humans.

Tumors have been segregated on the basis of their infiltration by immune cells: poor versus substantial (‘cold tumors’ versus ‘hot tumors’)⁵¹. The first category is characterized by poor infiltration by T cells and an increase in angiogenic and extracellular matrix factors^{54,55}. The second category has greater infiltration by T cells and increased expression of chemokines and type I interferons^{52,54,55}. Both tumor types are associated with distinct mechanisms for escaping the immune system^{51,54,55}. The breast-cancer subtypes we investigated here, TNBC and LBC, have substantial infiltration and poor infiltration, respectively, by immune cells⁵⁶. LBC DCs, and in particular LBC pDCs, showed enrichment for the ‘vascular wound healing’ and ‘extracellular matrix’ pathways, whereas TNBC DC subsets showed enrichment for immunological signatures, including interferon pathways. Hence, our findings have identified DCs as another level for the immune system-based stratification of tumors. This could aid in studies of the differential contributions of DC subsets to the mechanisms used by different tumor types to escape the immune system.

TNBC is a rare and aggressive breast-cancer subtype⁵⁷. Clinical trials using checkpoint blockers in TNBC are ongoing, with promising results^{14,58}. Hence, there is considerable interest in precisely characterizing the immune-system compartment in such patients. Here we have provided detailed analysis of APC subsets in TNBC. In particular, signatures specific to cDC1e cells, but not those specific to pDCs or cDC2s, were predictive of survival in TNBC, in contrast to results obtained for LBC. Hence, our data can be exploited to identify TNBC-specific prognostic signatures, as well as to identify promising targets to better direct therapies targeting immunological checkpoints.

Overall, our study has provided detailed and comprehensive molecular profiling of tumor-infiltrating DC subsets and MonoMacs in human cancer, which might serve as a reference dataset to increase biological knowledge of DCs in the context of disease. Our findings have shed light on the rules that dictate DC diversity and adaptation to complex microenvironments, such as in cancer, through transcriptional reprogramming. Our data will help to delineate the individual contributions of DC subsets to anti-tumor immunity and should provide a valuable resource for the identification of potential targets and biomarkers to better direct cancer immunotherapies.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41590-018-0145-8>.

Received: 30 March 2017; Accepted: 11 May 2018;
Published online: 16 July 2018

References

- Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **392**, 245–252 (1998).
- Collin, M., McGovern, N. & Haniffa, M. Human dendritic cell subsets. *Immunology* **140**, 22–30 (2013).
- Mildner, A. & Jung, S. Development and function of dendritic cell subsets. *Immunity* **40**, 642–656 (2014).
- Shay, T. & Kang, J. Immunological Genome Project and systems immunology. *Trends Immunol.* **34**, 602–609 (2013).
- Miller, J. C. et al. Deciphering the transcriptional network of the dendritic cell lineage. *Nat. Immunol.* **13**, 888–899 (2012).
- Heidkamp, G. F. et al. Human lymphoid organ dendritic cell identity is predominantly dictated by ontogeny, not tissue microenvironment. *Sci. Immunol.* **1**, eaai7677 (2016).
- Haniffa, M. et al. Human tissues contain CD141hi cross-presenting dendritic cells with functional homology to mouse CD103+ nonlymphoid dendritic cells. *Immunity* **37**, 60–73 (2012).
- Watchmaker, P. B. et al. Comparative transcriptional and functional profiling defines conserved programs of intestinal DC differentiation in humans and mice. *Nat. Immunol.* **15**, 98–108 (2014).
- Liu, Y. J. IPC: professional type 1 interferon-producing cells and plasmacytoid dendritic cell precursors. *Annu. Rev. Immunol.* **23**, 275–306 (2005).
- Lindstedt, M., Lundberg, K. & Borrebaek, C. A. Gene family clustering identifies functionally associated subsets of human in vivo blood and tonsillar dendritic cells. *J. Immunol.* **175**, 4839–4846 (2005).
- Mora, J. R. et al. Selective imprinting of gut-homing T cells by Peyer's patch dendritic cells. *Nature* **424**, 88–93 (2003).
- Huang, Q. et al. The plasticity of dendritic cell responses to pathogens and their components. *Science* **294**, 870–875 (2001).
- Pulendran, B., Palucka, K. & Banchereau, J. Sensing pathogens and tuning immune responses. *Science* **293**, 253–256 (2001).
- Stagg, J. & Allard, B. Immunotherapeutic approaches in triple-negative breast cancer: latest research and clinical prospects. *Ther. Adv. Med. Oncol.* **5**, 169–181 (2013).
- Liu, Y. J. Dendritic cell subsets and lineages, and their functions in innate and adaptive immunity. *Cell* **106**, 259–262 (2001).
- Dalod, M., Chelbi, R., Malissen, B. & Lawrence, T. Dendritic cell maturation: functional specialization through signaling specificity and transcriptional programming. *EMBO J.* **33**, 1104–1116 (2014).
- Soumelis, V., Patarini, L., Michea, P. & Cappuccio, A. Systems approaches to unravel innate immune cell diversity, environmental plasticity and functional specialization. *Curr. Opin. Immunol.* **32**, 42–47 (2015).
- Segura, E. & Amigorena, S. Inflammatory dendritic cells in mice and humans. *Trends Immunol.* **34**, 440–445 (2013).
- Wollenberg, A., Haberstroh, J., Teichmann, B., Wen, S. P. & Bieber, T. Demonstration of the low-affinity IgE receptor FcεRI/CD23 in psoriatic epidermis: inflammatory dendritic epidermal cells (IDEC) but not Langerhans cells are the relevant CD1a-positive cell population. *Arch. Dermatol. Res.* **290**, 517–521 (1998).
- Zaba, L. C., Krueger, J. G. & Lowes, M. A. Resident and “inflammatory” dendritic cells in human skin. *J. Invest. Dermatol.* **129**, 302–308 (2009).
- Segura, E. et al. Human inflammatory dendritic cells induce Th17 cell differentiation. *Immunity* **38**, 336–348 (2013).
- Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat. Immunol.* **3**, 991–998 (2002).
- Bell, D. et al. In breast carcinoma tissue, immature dendritic cells reside within the tumor, whereas mature dendritic cells are located in peritumoral areas. *J. Exp. Med.* **190**, 1417–1426 (1999).
- DeNardo, D. G. & Coussens, L. M. Inflammation and breast cancer. Balancing immune response: crosstalk between adaptive and innate immune cells during breast cancer progression. *Breast Cancer Res.* **9**, 212 (2007).
- Ghiringhelli, F. et al. CD4+CD25+ regulatory T cells inhibit natural killer cell functions in a transforming growth factor-β-dependent manner. *J. Exp. Med.* **202**, 1075–1085 (2005).
- Faget, J. et al. ICOS-ligand expression on plasmacytoid dendritic cells supports breast cancer progression by promoting the accumulation of immunosuppressive CD4+ T cells. *Cancer Res.* **72**, 6130–6141 (2012).
- Ghirelli, C. et al. Breast cancer cell-derived GM-CSF licenses regulatory Th2 induction by plasmacytoid dendritic cells in aggressive disease subtypes. *Cancer Res.* **75**, 2775–2787 (2015).
- Topalian, S. L., Drake, C. G. & Pardoll, D. M. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* **27**, 450–461 (2015).
- Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1–10 (2013).
- Angel, C. E. et al. Cutting edge: CD1a+ antigen-presenting cells in human dermis respond rapidly to CCR7 ligands. *J. Immunol.* **176**, 5730–5734 (2006).
- Bakdash, G. et al. Expansion of a BDCA1+CD14+ myeloid cell population in melanoma patients may attenuate the efficacy of dendritic cell vaccines. *Cancer Res.* **76**, 4332–4346 (2016).
- McGovern, N. et al. Human dermal CD14+ cells are a transient population of monocyte-derived macrophages. *Immunity* **41**, 465–477 (2014).
- Bronte, V. et al. Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nat. Commun.* **7**, 12150 (2016).
- Guilliams, M. et al. Unsupervised high-dimensional analysis aligns dendritic cells across tissues and species. *Immunity* **45**, 669–684 (2016).
- Taieb, J. et al. A novel dendritic cell subset involved in tumor immunosurveillance. *Nat. Med.* **12**, 214–219 (2006).
- Caminschi, I. et al. Putative IKDCs are functionally and developmentally similar to natural killer cells, but not to dendritic cells. *J. Exp. Med.* **204**, 2579–2590 (2007).
- Krasselt, M., Baerwald, C., Wagner, U. & Rossol, M. CD56+ monocytes have a dysregulated cytokine response to lipopolysaccharide and accumulate in rheumatoid arthritis and immunosenescence. *Arthritis Res. Ther.* **15**, R139 (2013).
- Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
- Grünewald, K. et al. Mammaglobin gene expression: a superior marker of breast cancer cells in peripheral blood in comparison to epidermal-growth-factor receptor and cytokeratin-19. *Lab. Invest.* **80**, 1071–1077 (2000).
- Han, J. H. et al. Mammaglobin expression in lymph nodes is an important marker of metastatic breast carcinoma. *Arch. Pathol. Lab. Med.* **127**, 1330–1334 (2003).
- Kowalewska, M., Chechlińska, M., Markowicz, S., Kober, P. & Nowak, R. The relevance of RT-PCR detection of disseminated tumour cells is hampered by the expression of markers regarded as tumour-specific in activated lymphocytes. *Eur. J. Cancer* **42**, 2671–2674 (2006).
- Novershtern, N., Regev, A. & Friedman, N. Physical module networks: an integrative approach for reconstructing transcription regulation. *Bioinformatics* **27**, i177–i185 (2011).
- Curtis, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
- Broz, M. L. et al. Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity. *Cancer Cell* **26**, 638–652 (2014).
- Galea, M. H., Blamey, R. W., Elston, C. E. & Ellis, I. O. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res. Treat.* **22**, 207–219 (1992).
- Gautier, E. L. et al. Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat. Immunol.* **13**, 1118–1128 (2012).
- Robbins, S. H. et al. Novel insights into the relationships between dendritic cell subsets in human and mouse revealed by genome-wide expression profiling. *Genome Biol.* **9**, R17 (2008).
- Franklin, R. A. et al. The cellular and molecular origin of tumor-associated macrophages. *Science* **344**, 921–925 (2014).
- Ojalo, L. S., Whittaker, C. A., Condeelis, J. S. & Pollard, J. W. Gene expression analysis of macrophages that facilitate tumor invasion supports a role for Wnt-signaling in mediating their activity in primary mammary tumors. *J. Immunol.* **184**, 702–712 (2010).
- Pyfferen, L. et al. The transcriptome of lung tumor-infiltrating dendritic cells reveals a tumor-supporting phenotype and a microRNA signature with negative impact on clinical outcome. *OncolImmunology* **6**, e123565 (2016).
- Wargo, J. A., Reddy, S. M., Reuben, A. & Sharma, P. Monitoring immune responses in the tumor microenvironment. *Curr. Opin. Immunol.* **41**, 23–31 (2016).
- Fuertes, M. B. et al. Host type I IFN signals are required for antitumor CD8+ T cell responses through CD8α+ dendritic cells. *J. Exp. Med.* **208**, 2005–2016 (2011).
- Salmon, H. et al. Expansion and activation of CD103+ dendritic cell progenitors at the tumor site enhances tumor responses to therapeutic PD-L1 and BRAF inhibition. *Immunity* **44**, 924–938 (2016).
- Gajewski, T. F., Schreiber, H. & Fu, Y. X. Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* **14**, 1014–1022 (2013).
- Spranger, S. & Gajewski, T. F. Tumor-intrinsic oncogene pathways mediating immune avoidance. *OncolImmunology* **5**, e1086862 (2015).
- Stanton, S. E., Adams, S. & Disis, M. L. Variation in the incidence and magnitude of tumor-infiltrating lymphocytes in breast cancer subtypes: a systematic review. *JAMA Oncol.* **2**, 1354–1360 (2016).
- Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N. Engl. J. Med.* **363**, 1938–1948 (2010).
- Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **13**, 674–690 (2016).
- Carpentier, S. et al. Comparative genomics analysis of mononuclear phagocyte subsets confirms homology between lymphoid tissue-resident and dermal XCR1+ DCs in mouse and human and distinguishes them from Langerhans cells. *J. Immunol. Methods* **432**, 35–49 (2016).

Acknowledgements

We thank the Institut Curie Cytometry Core facility for cell sorting; INSERM U932, particularly C. Laurent and A.S. Hamy-Petit, for bioinformatics advice; and S. Alculumbe and P. Vargas for discussions. F. Noël was supported by a fellowship from the French Ministry of Research. This work was supported by funding from INSERM (BIO2012-02, BIO2014-08, HTE2016), Fondation pour la Recherche Médicale, ANR-10-IDEX-0001-02 PSL* and ANR-11-LABX-0043, European Research Council (IT-DC 281987) and CIC IGR-Curie 1428, INCA EMERG-15-ICR-1, la Ligue contre le cancer (Labellisation EL2016.LNCC/WaS). High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium), InCa from ANR ('Investissements d'Avenir' program), by the Canceropole Ile-de-France and by the SiRIC-Curie program (SiRIC Grant 'INCa-DGOS- 4654').

Author contributions

P.M. designed and performed experiments, analyzed results and wrote the manuscript; F.N. performed bioinformatics analyses and wrote the manuscript; E.Z., U.C. and C.G.

analyzed results; P.S. and O.A. performed experiments; A.S.-D. and M.G-D. contributed to project management; A.V.-S. contributed to clinical project management and pathology review and provided clinical samples; F.R. contributed to clinical project management; S.A. and E.S. provided strategic advice and revised the manuscript; and V.S. designed experiments, supervised the research and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41590-018-0145-8>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to V.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Human samples and patient characteristics. Fresh samples of tumor tissues and juxta-tumoral tissues (lacking malignant tumor cells) of untreated patients with breast cancer were obtained from Hôpital de l'Institut Curie (Paris) in accordance with Institut Curie ethical guidelines. LBC and TNBC types were included in the study according to the hormonal receptor status. Patient characteristics are summarized in Supplementary Table 2.

Single-cell suspensions from human samples. Tumor and juxta-tumoral tissue were cut into small pieces and digested in CO₂-independent medium (Gibco) containing 5% FBS (HyClone), 2 mg/ml collagenase I (C0130, Sigma), 2 mg/ml hyaluronidase (H3506, Sigma) and 25 µg/ml DNase (Roche) by three rounds of 15 min of incubation with agitation at 37 °C. The samples were filtered on a 40-µm cell strainer (Fischer Scientific) and were diluted in PBS 1× (Gibco) supplemented with 1% de-complemented human serum (BioWest) and EDTA 2 mM (Gibco). After centrifugation, cells were resuspended in the same medium and were counted before being assessed by flow cytometry or sorted.

Antibodies and cell sorting. For phenotypical characterization, single-cell suspensions were stained with antibodies to the following human molecules: CD3-Alexa700 (557943; clone: UCHT1), CD19-Alexa700 (557921; clone: HIB19), CD56-Alexa700 (557919; clone: B159), CD56-BUV737 (564448; clone: NCAM16.2), CD163-BV786 (741003; clone: GH1/61), CD11c-PECy5 (551077; clone:B-ly6), CD11c-PE-CF594 (562393; clone:B-ly6), CD123-BV650 (563405; clone: 7G3), HLA-DR-BUV395 (564040; clone: G46-6), and CD45 allophycocyanin-Cy7 (557833; clone: 2D1), all from BD; CD14-Qdot605 (Q10013; clone: TuK4) from Life Technologies; CD14-BV605 (301833; clone:M5E2), CD16-BV510 (302047 clone: 3G8), CD123-PE-Cy7 (306010; clone: 6h6), CD1c-PE (331506; clone: L161) and HLA-DR BV711 (307643; clone: L243), all from Biolegend; CD1c-PerCP-eFluor710 (46-0015-42; clone: L161) and FcεR1-allophycocyanin (17-5899-42; clone: AER-37), all from eBioscience; AXL-AlexaFluor488 (FAB154G; clone: 108724) and CD32B-allophycocyanin (FAB1330A; clone: 190723), both from R&D; and CD141-PE (130-098-841; clone: AD5-14H12) from Miltenyi Biotec. For DC sorting, we used the following antibodies instead of the corresponding marker: CD45-BV570 (304033; clone: HI30), from Biolegend; CD14-FITC (555527; clone: 10.1), from BD; and HLA-DR-allophycocyanin-eFluor780 (47-9956-42; clone: LN3), from eBioscience. Single-cell suspensions of tumor-digested cells were sorted in a BD FACSAria III upgrade using the purity mode, a 100-µm nozzle loop, and at low pressure (20 psi). DC subsets were sorted in Eppendorf tubes containing RPMI plus 5% FBS (HyClone) for morphological analysis. Once the morphology for each subset was confirmed, and because of the low number of tumor-infiltrating APC, we directly sorted tumor APCs in TCL buffer (Qiagen) supplemented with 1% β-mercaptoethanol (SIGMA) for RNA-seq experiments.

Morphological analysis. Sorted cells were subjected to cytocentrifuge and were stained with May-Grunwald–Giemsma stain. Images were obtained with a ProgRes SpeedXT core 5 Microscope Camera (JENOPIK) on a Leica DM 4000 B microscope.

RNA-seq. The general RNA-seq workflow is summarized in Supplementary Fig. 1. In brief, RNA from sorted cells (>100 cells) was extracted by using a Single Cell RNA Purification Kit (Norgen Bioteck), including on-column DNase digestion (Qiagen), as described by the manufacturer's protocol. RNA integrity was confirmed with an RNA 6000 Pico Kit (Agilent Technologies) in BioAnalyzer. cDNA was generated with SMARTer Ultra Low input RNA for Illumina Sequencing-HV (Clontech), following manufacturer's protocol; 14 cycles were used to amplify cDNA. The quantity of cDNA and quality of cDNA were assessed with Qubit dsDNA high sensitivity (Thermofisher) and an Agilent Bioanalyzer using nanochip (Agilent Technologies), respectively. Multiplexed pair-end libraries 50 nt in length were obtained using Nextera XT kit (Clontech). Sequencing was performed in the same batch in Illumina HiSeq 2500 using an average depth of 15 million reads; 50-nt-length reads per samples were obtained. Library, sequencing and quality control of the sequencing were performed by the NGS facility at Institut Curie.

RNA-seq data pre-processing. Reads were mapped to the human genome reference (hg19/GRCh37) using TopHat2 software version 2.0.6⁶⁰. Gene expression values were quantified as read counts using HTSeq-count⁶¹. We filtered out genes with fewer than five read counts in at least 25% of samples and normalized the raw data using RUVg method (RUVSeq R package)⁶². This method identifies technical noise based on negative control genes that should be affected by unwanted variations but not affected by biological effects of interest. We selected the 5,000 less-variant genes as negative-control genes. From the 82 samples sequenced, only two were excluded from this study, corresponding to tumor and juxta-tumoral pDC. These samples had low expression of pDC-specific markers and high expression of macrophage markers.

For exploratory analyses, we performed principal-component analysis (PCA) of the 500 most-variant genes, based on inter-quartile range method (IQR) (EMA R package)⁶³, of APC transcriptomes from LBC and TNBC tumor samples. Data were log₂-transformed, centered and scaled. PCA was performed using the FactoMineR

package. The z-score of log₂-transformed gene expression, scaled by gene, were presented in a heat-map color.

Gene set-enrichment analysis. We selected APC specific gene sets from literature⁵⁹ and performed enrichment analysis on our dataset selected LBC T samples. To do so, we used the BubbleMap module of the BubbleGUM software which perform GSEA with multiple-testing correction⁶⁴.

Statistical analysis. Significant differences in the frequency of APCs among total live cells or CD45⁺ cells were performed using ANOVA, followed by a post-hoc test. For paired samples in the tumor-versus-juxta-tumor comparison of APC, we performed a Wilcoxon test by using the GraphPad Prism 6.0.

To generate subset-specific signature of APC for each condition, we performed one-way ANOVA differential analysis test on the log₂ expression data of the five APCs. We kept only the genes differentially expressed by at least two subsets ($P < 0.05$). We then performed a Tukey post-hoc test to select genes exclusively upregulated in one subset relative to their expression in all the other subsets ($P < 0.05$). Those upregulated genes were defined as the subset-specific signature.

To identify genes whose expression varied between tumor tissues and juxta-tumoral tissues, for each APC separately, we performed pairwise comparison of gene-expression-matched samples using the generalized linear model (GLM) likelihood ratio test of EdgeR R package⁶⁵. Only DEGs with an FDR of <0.05 and a change in expression of over onefold (log₂ values) were considered 'differentially expressed'. The same analysis was applied to find genes expressed differentially TNBC samples relative to their expression in LBC samples for each subset.

Metagene expression was defined as the median expression (log₂ value) of the genes of interest in each sample. Differential expression analysis of metagenes was done using the non-paired Wilcoxon test. Correlations were assessed using the Pearson correlation test, with a threshold of $P < 0.05$.

All RNA-seq statistical analyses were performed using R software (Version 3.2.3).

Regulatory network and functional inference. We extracted the gene-expression matrix for each subset, and each comparison. The conditions were as followed: 1, one subset versus all other subsets in LBC; 2, one subset versus all other subsets in TNBC; 3, tumor tissue versus juxta-tumoral tissue for each subset separately in LBC; and 4, TNBC versus LBC, for each subset separately. We then loaded the matrix on Cytoscape software version 3.4.0. One analysis per subset was performed. Network inference was performed using ARACNE application, which is based on mutual information theory^{66,67}. The parameters used in ARACNE were Mutual Information Algorithm Type: Variable Bandwidth. We used a transcription factor (TF) list for Hub/TF Definition from the dataset Fantom⁶⁸. The mutual information threshold was 0.5. We next used the ClueGO Application⁶⁹ to determine pathway enrichment in each network. Public datasets only from 'Experimental evidence' of Gene Ontology (GO) – Biological process-GOA, - Cellular Component-GOA, - ImmuneSystemProcess-GOA, - Molecular Function-GOA, (updated date: 15.01.2017), InterPro dB: Protein Domains (updated date: 03.11.2015), Reactome (updated date: 20.01.2017), and WikiPathways (updated date: 20.01.2017) were used. The Go Term Fusion option was selected. Only pathways with a Benjamini-Hochberg (BH) adjusted P value below 0.05 were kept.

Checkpoint expression analysis. The presence of the following immunological checkpoints was analyzed among DEGs in tumor samples versus juxta-tumoral samples, for each subset. Positive checkpoint genes included those encoding CD40, CD70, CD80, CD83, TNFSF9 (4-1BB), ICOSL, SEMA4A, TIMD4, C10orf54 (VISTA or B7-H5), TNFRSF13C (BAFFR), TNFSF13 (APRIL), TNFSF13 (HVEML), CD84, CD48, TNSF4 (OX40L) and PVR (CD155). Negative checkpoint genes included those encoding CD274 (PD-L1), CD276 (B7-H3), PDCD1LG2 (PD-L2), BTLA, LGALS1, LGALS3, LGALS9, CD279 (PD1), CEACAM6 and CD209 (DC-SIGN).

Clinical outcome of subset-specific signature score in public breast cancer dataset. METABRIC is a public dataset¹³ of transcriptomics data of breast tumor samples with clinical data associated. From this dataset, we selected samples from LBC ($n = 1,043$) and TNBC ($n = 259$) according to the expression of the receptors ER, PR and HER2. To study the clinical outcome of patients, we considered those with the label 'd-d.s' and 'a' in the 'last follow up status' variable. Similar to a published report¹⁴, we calculated a z-score ratio of upregulated APC subset-specific signatures to downregulated APC subset-specific signatures that we generated from our breast cancer RNA-seq data, as follows:

$$\log_2 \text{Ratio} = \log_2 \left(\frac{\text{Signature UP}}{\text{Signature DOWN}} \right)$$

$$z\text{-score ratio} = \frac{\log_2 \text{Ratio} - \bar{\log}_2 \text{Ratio}}{\text{sd}(\log_2 \text{Ratio})}$$

To assess the predictive value of the CD103⁺ DCs reported before⁴⁴, we applied the same z-score, based on the CD103⁺ DC signature, as the 'signature UP', and the CD103⁻ DC signature as the 'signature DOWN'. The CD103⁺ DC and CD103⁻ DC signatures contained 9 genes and 16 genes, respectively⁴⁴.

To assess predictive value of the pDC signature reported before⁷, we applied the same z-score, based on pDC upregulated genes as the 'signature UP' (440 genes) and pDC downregulated genes as the 'signature DOWN' (524 genes).

To assess predictive value of the interferon signature found in TNBC APCs, we performed a z-score on the log₂ mean expression of *IFNL1*, *IFNB1*, *ISG15* and *ISG20*. We performed univariate Cox analysis to assess the link between subset-specific signatures z-score ratio expression and disease-free survival. We divided the subset-specific z-score expression in two groups: 'high' or 'low', according to the median value. Kaplan-Meier curves were generated using survminer R package. Multivariate cox analysis was performed to link subset-specific signatures and the clinical prognostic parameter, Nottingham Prognostic Index (NPI)⁴⁵, to disease-free survival.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. RNA-seq data that support the findings of this study have been deposited in the NCBI Sequence Read Archive (SRA) with the accession code PRJNA380940.

References

60. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
61. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
62. Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32**, 896–902 (2014).
63. Servant, N. et al. EMA - A R package for Easy Microarray data analysis. *BMC Res. Notes* **3**, 277 (2010).
64. Spinelli, L., Carpentier, S., Montañana Sanchis, F., Dalod, M. & Vu Manh, T. P. BubbleGUM: automatic extraction of phenotype molecular signatures and comprehensive visualization of multiple Gene Set Enrichment Analyses. *BMC Genom.* **16**, 814 (2015).
65. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
66. Margolin, A. A. et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinforma.* **7**, S7 (2006).
67. Basso, K. et al. Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **37**, 382–390 (2005).
68. Joshi, A. et al. Technical Advance: Transcription factor, promoter, and enhancer utilization in human myeloid cells. *J. Leukoc. Biol.* **97**, 985–995 (2015).
69. Bindea, G. et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).