**CDD**press

Check for updates

# A single-cell transcriptomic landscape of innate and adaptive intratumoral immunity in triple negative breast cancer during chemo- and immunotherapies

Laura Carpen[1,2], Paolo Falvo[1], Stefania Orecchioni[1], Giulia Mitola[1], Roman Hillje[3], Saveria Mazzara[4], Patrizia Mancuso[1], Stefano Pileri[1], Alessandro Raveane [1,5,6 ✉] and Francesco Bertolini [1,6 ✉]

Breast cancer (BC) constitutes a major health problem worldwide, making it the most common malignancy in women. Current treatment options for BC depend primarily on histological type, molecular markers, clinical aggressiveness and stage of disease. Immunotherapy, such as αPD-1, have shown combinatorial clinical activity with chemotherapy in triple negative breast cancer (TNBC) delineating some therapeutic combinations as more effective than others. However, a clear overview of the main immune cell populations involved in these treatments has never been provided.

Here, an assessment of the immune landscape in the tumor microenvironment (TME) of two TNBC mouse models has been performed using single-cell RNA sequencing technology. Specifically, immune cells were evaluated in untreated conditions and after treatments with chemotherapy or immunotherapy used as single agents or in combination. A decrease of Treg was found in treatments with in vivo efficacy as well as γδ T cells, which have a pro-tumoral activity in mice. Focusing on Cd8 T cells, across all the conditions, a general increase of exhausted-like Cd8 T cells was confirmed in pre-clinical treatments with low efficacy and an opposite trend was found for the proliferative Cd8 T cells. Regarding macrophages, M2-like cells were enriched in treatments with low efficacy while M1-like macrophages followed an opposite trend. For both models, similar proportions of B cells were detected with an increase of proliferative B cells in treatments involving cisplatin in combination with αPD-1. The fine-scale characterization of the immune TME in this work can lead to new insights on the diagnosis and treatment of TNBC.

## INTRODUCTION

Breast cancer (BC) is the most common malignancy in women, and the fifth leading cancer death worldwide [1]. The triple negative breast cancer (TNBC) subtype accounts for 15-20% of BC cases and has the worst prognosis [2].

Until recently, the backbone of therapy against TNBC has been chemotherapy, including alkylating agents such as cyclophosphamide [3] and cisplatin [4], anti-microtubules such as taxanes [5] and antineoplastic agents such as doxorubicin [6]. The immune checkpoint inhibitors (ICIs) anti-PD-1 and anti-PD-L1 have been recently approved for TNBC therapy in combination with chemotherapy, but they are so far clinically active only in a minority of patients and for a limited timeframe [7–10].

PD-1 predominantly regulates effector T cell activity within tissue and tumors by binding the programmed cell death ligand 1 (PD-L1). In turn, this binding inhibits kinases involved in T cell activation [11]. In physiological conditions, the interaction of PD-1 with its ligands has been shown to play an important role in the maintenance of the balance between autoimmunity and peripheral tolerance [12, 13]. In the tumor microenvironment (TME), PD-1 and its ligand PD-L1 perform a vital role in progression and survival of cancer cells; the overexpression of PD-L1 by tumor cells is used as self-defense by the tumor against the cytotoxic T cells which contribute to cell killing [14]. PD-L1 expression on many tumors is a component of a suppressive microenvironment that leads to T cell dysfunction and exhaustion [15]. This state of exhaustion is characterized by the progressive loss of proinflammatory cytokines production, the loss of the cytotoxic activity, the decrease in the proliferative potential and an increase in apoptosis [16]. As a consequence, blocking the PD-1/PD-L1 inhibitory pathway can re-activate T cells in the TME with the release of inflammatory cytokines and cytotoxic granules to eliminate tumor cells. PD-1 is also highly expressed on regulatory T cells (Tregs), where it may enhance their proliferation in the presence of PD-L1 [17]. As Tregs suppress effector immune responses, blockade of the PD-1 pathway may enhance antitumor immune responses.

TNBC is believed to be an immunogenic BC subtype, but it is currently unclear which cell populations are involved in the immune response within the TME during specific conditions as well as their proportion in specific treatments [18].

Single-cell RNA sequencing (scRNA-seq) gives the possibility to differentiate among cell populations that are not distinguishable by

[1]Laboratory of Hematology-Oncology, European Institute of Oncology IRCCS, Milan, Italy. [2]Human Technopole, Milan, Italy. [3]The Hyve, Utrecht, Netherlands. [4]Hematopathology Unit, European Institute of Oncology IRCCS, Milan, Italy. [5]Present address: Human Technopole, Milan, Italy. [6]These authors contributed equally: Alessandro Raveane, Francesco Bertolini. ✉email: alessandro.raveane@fht.org; francesco.bertolini@ieo.it

SPRINGER NATURE

CDDpress

**Table 1.** Condition overview.

| Label | Type of treatment | Chemistry 10X for 4T1 | Chemistry 10X for EMT6 |
|---|---|---|---|
| ctr | control | v2 | v2 |
| αPD-1 | Immunocheckpoint Inhibitor (ICI) | v2 | v2 |
| C140 (Cyclophosphamide) | Alkylating agents | v3 | v3 |
| C140 + V (Vinorelbine) | Alkylating agents + anti-microtubule agents | v3 | v3 |
| C140 + αPD-1 | Alkylating agents + ICI | v2 | v2 |
| C140 + V + αPD-1 | Alkylating agents + anti-microtubule agents + ICI | v3 | v2 |
| D (Doxorubicin) + αPD-1 | Antineoplastic agents + ICI | v2 | v3 |
| P (Cisplatin) + αPD-1 | Antineoplastic agents + ICI | v2 | v2 |
| T (Paclitaxel) + αPD-1 | Anti-microtubule agents + ICI | v3 | v2 |

Schematic view of conditions (samples) with the corresponding treatment for the 4T1 and EMT6 cell line TMEs with the 10 X Genomics chemistry version used for each condition.

cell surface markers and morphology alone, opening the possibility of identifying previously uncharacterized cellular populations, phenotypes and transitional states. This approach has revolutionized our ability to study the immune system and allow us to break through the bottleneck of immunology studies [19, 20].

In this work, we have investigated at the single cell level, report and discuss in detail the transcriptome of innate and adaptive intratumoral immune cells in two syngeneic, immune competent, orthotopic murine models of local and metastatic TNBC. Mice were treated with ICIs and several different types of chemotherapeutics, alone or in combination. From previous reports, capecitabine (alone or with ICIs) was the less effective drug [21]. While platinum, doxorubicin and taxanes showed synergy with ICIs and had superimposable activity, intermittent, medium dosage cyclophosphamide (C140) plus vinorelbine and ICIs was the most active combinatorial therapy. Vinorelbine activated antigen presenting cells and C140 generated new T cell clones including stem cell-like TCF1 + CD8 + T cells [22]. The fine characterization of almost 50 000 immune cells extracted from the TME of these two mouse models helped in creating a catalogue of the immune response to several drugs and aimed to investigate specific cellular subtypes useful for future therapeutic approaches.

## MATERIALS AND METHOD
### Cell lines and treatments
As this work focuses on the computational analysis of scRNA-seq data previously published, in vivo and in vitro experiments were performed as mentioned in [22]. Briefly the laboratory procedures included the injection of two TNBC cell lines (4T1 and EMT6) in the mammary fat pad as in [21]. Tumor-bearing mice were treated with either vehicle or with different drugs used as single agents or in combination as described in [22] for a total of eight treatments and one untreated control for each cell line (Table 1). In almost all treatments for each cell line, checkpoint inhibitors alone or in combination with chemotherapeutic drug led to tumor shrinking. In fact, anti-PD-1 treatment reduced on average tumor mass of 68% and 75% in 4T1 and EMT6 model, respectively [22]. Drug usages were based on literature data associated with no or acceptable toxicity [22, 23].

### ScRNA-seq library preparation and sequencing
At 28 or 70 days, depending on the efficacy of the treatment and on the cell line (details in [22]), tumor resection was performed as in [23].

Tumors of three mice were dissociated and pooled together to generate the single cell suspension. We took care to pool tumors of similar size and with no sign of necrosis and ulcers [22]. The cell suspensions were then prepared for cell sorting with FACS Fusion sorter (BD bioscience). Cd45+ DAPI- (alive immune cells) were

sorted and purity evaluated; as previously showed [22], post sorting purity was assessed to be higher than 90%. The percentages of Cd45+ DAPI⁻ was of 28.3 ± 4.3% for 4T1 and 9.3 ± 3.3 % for EMT6 tumor models, respectively. At least 5,000 cells per condition underwent scRNA-seq library preparation following the 10X Genomic protocol and using two different chemistries (v2 and v3) (Table 1). The sequencing was performed with NovaSeq™ 6000 Illumina® sequencer at a sequencing depth of 50 000 read pairs/cell.

### Alignment and quality control
FASTQ files were converted to digital gene-cell count matrices using a Singularity- dependent Snakemake pipeline [24] employing the Cell Ranger v4 software. As reference, the *Mus musculus* reference genome mm10 (GENCODE vM23/Ensembl98) was used. Market Exchange Format (MEX) for sparse matrices generated from the pipeline were loaded, merged, processed and analyzed using *Seurat* package v4 [25].

To have a comparable number of cells for each experiment, we excluded from the analyses all the conditions with a number of cells lower than 500. Then, in order to include only cells that are of high quality, we exclude cells with 500 or less transcript, 50,000 or more transcripts, having fewer than 250 expressed genes, a complexity score (log10 genes per UMI) lower than 0.80 and more than 15% mitochondrial transcripts as in [22]. At the gene-level, all genes expressed in less than five cells were filtered out.

To detect doublets, the *scDblFinder()* function of *scDblFinder* package [26] was used. This package works only with SingleCellExperiment (SCE) objects [27], therefore, a conversion from *Seurat* object was performed. Finally, doublets were excluded using the *subset()* function of the *Seurat* package [25].

During the quality control (QC), one condition (C140 + V + αPD-1), out of the nine reported here for the 4T1 cell line, was removed from further analyses because its number of cells did not exceed the minimum filtering threshold (500 cells) used by us to consider a condition suitable to explore the whole immune populations. For the EMT6, all conditions passed this quality control filter. The number of transcripts and number of genes were evaluated for the remaining conditions. The majority of the cells had more than 1 000 UMI indicating high quality cells (Fig. S1). Furthermore, in the 4T1, the C140, C140 + V and T + αPD-1 treatments had the highest number of mitochondrial transcripts; while for the EMT6 it was C140 + V, C140 + V + αPD-1 and P + αPD-1. This might be related to differences in toxicity of these specific agents in different microenvironments [28–30] due to the fact that high percentage of mtDNA transcript is usually associated with apoptotic, stressed and low-quality cells.

**Proliferation status, normalization and batch effect removing**
Cell cycle score variation was evaluated with a Principal Component Analysis (PCA) on normalized and scaled data using 2000 genes on *Seurat* v4 package after assigning a cell cycle score to each cell with the *CellCycleScoring()* function and the S and G2M specific gene reference downloaded from the *Ensemldb* R package [31]. No large differences were observed among cell cycle phases between the two cell line TMEs, therefore, we did not regress out the cell cycle variation in the following normalization step.

A batch effect is an unwanted source of variation resulting in different cells having specific profiles, not because of their biological features but because of technical differences. Our data presented a strong batch effect due to the two different types of chemistry used during the single cell library preparations (Fig. S2A). Recently, *Seurat* introduced the scRNA-seq integration workflow, a set of methods to match shared cell populations across different batches [25]. These methods identify cross-batch pairs of cells that are in a matched biological state ('anchors'). In detail, we applied the integration workflow that included splitting of the raw transcript count matrix by chemistry, normalization using the *SCTransform()* function, selection of the most variable features (genes) using *SelectIntegrationFeature()* function, preparation of the *Seurat* object for the interrogations with *PrepSCTInte-gration()* function, canonical correlation analysis (CCA) with *FindIntegrationAnchors()* function and final integration across conditions with *IntegrateData()*. These steps corrected the unwanted source of variations (Fig. S2B) and were applied before each clustering analysis.

**Dimensionality reduction, visualization and clustering**
The Uniform Manifold Approximation and Projection (UMAP) method for visualization was employed on the first 30 principal components using the *RunUMAP()* function of the *Seurat* package. In order to identify known (previously identified cell populations) or uncharacterized cell types, the *Seurat* v4 graph-based clustering approach, which exploits a K-nearest neighbour (KNN) graph, was applied. We determined the k-nearest neighbor graph with *FindNeighbors()* function and then performed the clustering with the *FindClusters()* function from resolution 0.4 to 1 in steps of 0.2. The resolution 0.8 and 0.6 were evaluated as best for the 4T1 and EMT6 TMEs respectively, according to the number of cell populations (clusters) that were possible to detect. The choice of the best granularity parameters was evaluated by visual inspection and with the aid of the *Clustree* package [32].

**Cell-type annotation**
The *SingleR* package [33] in combination with the *ImmGen()* reference transcriptome dataset [34], containing 253 fine labels generated from 830 microarray samples of sorted cell populations, was used for automatic cell type assignment. We inspected the confidence of the predicted labels using the delta values: the difference between the score for the assigned label and the median across all labels for each cell. Using the *PruneScores()* function, we marked potentially poor-quality or ambiguous assignments based on the delta value. Moreover, we uniformed the label name of the *ImmGen()* dataset according to the wanted level of resolution by using the cell ontology label present in the *Celldex* package [33]. For example, two of the several Cd4 T cell labels were T.CD4.24H (CL:0000624) and T.CD4.CTR (CL:0000624), therefore, we searched for the cell ontology label in Ontology Lookup Service (OLS) repository (https://www.ebi.ac.uk/ols/ontologies/cl) and established "T cells Cd4" as a common label. We verified the assignment using two procedures: (i) exploring the expression of known cell gene markers; ii) evaluating the top differential expressed genes (DEG) between cell clusters on *PanglaoDB* [35].

Differentially expressed genes were retrieved using the *FindAllMarkers()* function in the *Seurat* package with a MAST test [36]. Only genes expressed on 25% of cells and with a log fold-change higher than 1.5 were considered. For this analysis, the normalized data (not integrated) was used as suggested by the *Seurat* developers (https://github.com/satijalab/seurat/issues/2014#issuecomment-629358390).

Beside using *Seurat*, DEG analysis was also performed using the *SingleCellExperiment* R package (SCE) [27] that, differently from *Seurat*, allows a block on batch. This block, necessary in our dataset, would reveal biologically relevant genes to be preserved within the batch (Tables S1, 4T1; S2, EMT6).

A label name, mirroring the cell composition, was assigned to each set of cells under the same group (cluster). If multiple cell populations were present in a cluster the first name refers to the most abundant type of cell.

**Sub-clustering**
Cells in the macro-clusters of interest (Cd4 T cells, Cd8 T cells and NK cells, Macrophages and B cells) were extracted according to their label name using the *subset()* function in *Seurat* package. The filtered transcript counts were re-normalized as before using the integration workflow or the classic normalization depending on the purpose of the analysis.

PCA and UMAP methods were applied as before, with the only difference being that we evaluated the best number of PCs to use for the clustering workflow with the *maxLikGlobalDimEst()* function of the *intrinsicDimension* package [37] as used in [38].

Clustering was performed as above and the best resolution was evaluated with the *Clustree* package [32]. Also depending on the results and the expression of known gene markers the granularity was chosen accordingly (0.2 and 0.3 for Cd4 and Cd8 respectively in 4T1 cell line, 0.1 and 0.2 for B in the 4T1 and EMT6 cell line respectively, 0.4 for macrophages in EMT6 cell line). The sub-clusters cell assignment was performed only with manual curation by choosing a known set of genes from relevant studies that focus on the same cell populations [39–42] in similar mouse models and evaluating their expression in the sub-clusters.

**Trajectory analysis**
Dynamic changes in gene expression were evaluated by performing a trajectory analysis using the *Slingshot* package [43]. To give a finer definition of cell states and unknown cell populations the trajectory analyses were performed only on the cluster subsets.

The *Slingshot()* function was used on the *Seurat* object converted into SCE dataset [27], then the embedding of trajectory in new space was performed with the *embedCurves()* function and finally the *slingCurves()* assessed each curve in each sub-clustering.

## RESULTS
**Total immune cellular landscape in the tumor microenvironment of two TNBC mouse cell lines**
After the QC, the resulting total number of cells and genes for the 4T1 were 22,403 and 18,124 respectively, while for the EMT6, 26,245 cells and 18,637 genes were obtained (Tables S3 and S4). The treatments having the highest number of cells after the QC corresponded to T + αPD-1, C140 and C140 + αPD-1 for the 4T1; while the C140 + V, P + αPD-1 and D + αPD-1 were the treatments with the highest values in EMT6 cell line. Therefore, the analyses focused on a total of 48,648 immune cells and 17 conditions in the TMEs of two TNBC mouse cell lines.

The best granularity resolution (see Materials and method) in the two TMEs identified a total of 20 and 22 groups of cells sharing similar gene expression, for the 4T1 and EMT6 respectively (Figs. 1 and S3A, B, for details see Materials and method and Supplementary Materials). Subsequently, for both the cell lines, *Cd3e*, *Cd4*, *Cd8b1* genes were manually evaluated for T cells, *Ncr1* for NK cells, *Cd19* for B cells, *Csf3r* for neutrophils, *Adgre1* and *Cd68* for macrophages and *Basp1* for DCs (Fig. S3C, D). The known gene marker expression in each cluster was in accordance with the
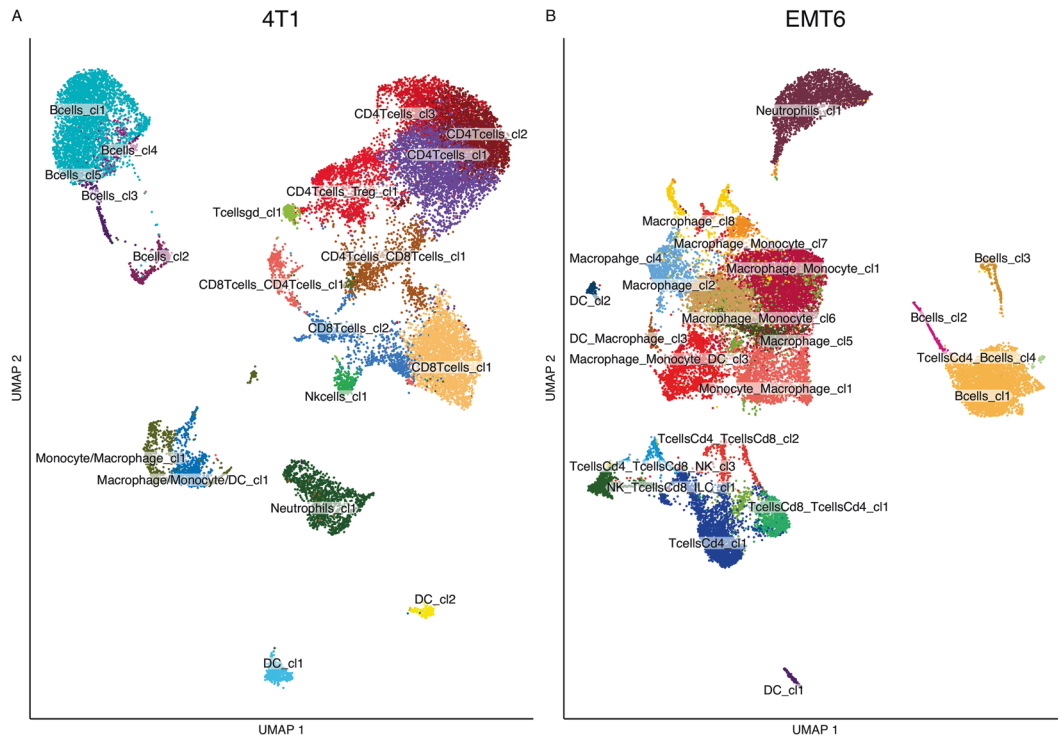
**Fig. 1 Immune cellular landscape in the two mouse TNBC TMEs. A** UMAP of 22,403 immune cells in 4T1 TME grouped in 20 clusters. Each dot refers to a cell. The colors refer to the label assigned to the clusters. **B** UMAP of 26,245 immune cells in EMT6 TME grouped in 22 clusters. Each dot refers to a cell. The colors refer to the label assigned to the clusters.

highest frequencies of cell populations automatically assigned to that cluster. An additional confirmation was obtained by analyzing DEG for each cluster compared to all the others (Tables S1 and S2). These genes were further evaluated on PangloaDB and once again, the results confirmed the automatic assignment.

Final label assignment resulted in the 20 4T1 clusters being classified in 5 B, 5 Cd4 T, 3 Cd8 T, 1 NK, 1 T γδ, 2 DC, 2 macrophage and 1 neutrophil cell clusters (Fig. 1A). On the other hand, the EMT6 cell line presented 22 clusters labelled as 8 macrophage, 1 monocyte, 1 neutrophil, 3 B, 3 DC, 4 Cd4 T, 1 Cd8 T and 1 NK cell clusters (Fig. 1B).

At first glance, a strong difference in the immune cell population composition was found among the two tumoral cell line TMEs. Specifically, the 4T1 TME recorded a prevalence of cells belonging to the lymphoid lineage: 15 clusters contained T, NK and B cells while only 5 were named as macrophages, DCs and neutrophils. On the other hand, in the EMT6 TMEs, most of the cells fall into myeloid clusters while only few cells were assigned to lymphoid lineage. Interestingly, among the lymphoid lineage the number of cells belonging to B cell and neutrophil clusters were comparable between the two tumor types.

Since a fine-scale characterization of the immune landscape wanted to be reached, a filtering and a new sub-clustering for the major immune components of the lymphoid and myeloid lineages, where an informative number of cells could be retrieved, was performed. Due to the differences found in cell population composition, here will be reported the results of Cd4, regulatory, γδ, Cd8 T and NK cells sub-clustering for the 4T1 cell line TME; while of macrophages for the EMT6. Moreover, a comparison of the equally represented B cell clusters between the two types of tumor is also presented and discussed.

## Cd4 T cell-like sub-clustering in 4T1 TMEs reveals pro-tumoral activity of mouse specific T cell population

Cells belonging to clusters that mainly contained Cd4, regulatory and γδ T cells, based on the label assignment, were retrieved and re-clustered (see Materials and method). The procedure resulted in 7 clusters (Fig. 2A) which can be presented as follows: two progenitor-like Cd4 T cell clusters (0 and 2), characterized by the high expression of *Sell*, *Ccr7*, *Lef1* and *Tcf7* genes (Fig. 2B); a Cd8-like cluster (cluster 3) showing high expression of *Cd8a* gene, these cells are likely a subset of cells deriving from a cluster presenting a mixed cell composition and retrieved because included Cd4 cell population; a Treg-like cell cluster (cluster 1) that presented a high level of *Foxp3* and *Ikzf2* genes; a γδ T cell cluster (cluster 4) that expressed the *Trdc* gene; and finally, two exhausted-like Cd4 T cell clusters (5 and 6) that had a high expression of *Nr4a1* and *Tox* gene markers. Among them, cluster 5 exhibited a more active profile due the high expression of *Cd40lg* gene but at the same time a higher expression of *Il7r* gene that encodes for a receptor whose ligand was related to tumor progression in γδ T cells [44].

UMAP analyses can be used to project the information related to multi-branched trajectories in order to facilitate pseudo-time analysis that measures the relative progression of each of the cells along a biological process of interest without explicit time-series data [45–47].

Computationally imputed pseudotime trajectory confirmed and extended the understanding of this sub-cluster composition (Fig. 2C). As previously reported [48], three distinct trajectories or cluster differentiations always starting from the same root (cluster 2) were found. The first connects the root (cluster 2) to the γδ T cell cluster passing by the second progenitor-like cluster (cluster 0). The end of the trajectory in the second trajectory is the Cd8-like cluster (cluster 3). While in the third, cluster 2 generates exhausted-like cluster 5 passing through the other exhausted state (cluster 6) (Fig. 2C).

The percentage of cells for each cluster varied significantly among the different conditions (Figs. 2D and S4). Among the most relevant results, a decrease, compared to the control TME, in the percentage of cells belonging to Treg cluster (cluster 1) was observed in treatments with high in vivo efficacy such as
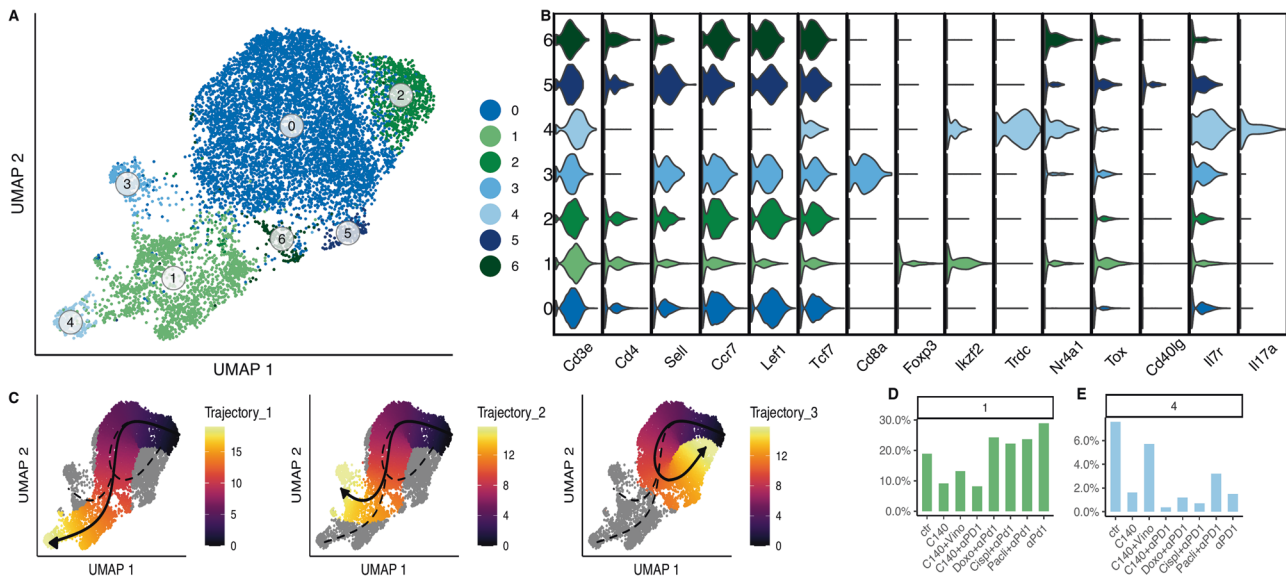
**Fig. 2  Cd4-like T cell sub-clustering analysis in 4T1. A** UMAP of Cd4-like T cell sub-clustering in 4T1 cell line TMEs. Each dot refers to a cell. The colors refer to the sub-clusters. **B** Gene signature: the violin plot shows the gene on the x-axis, while the clusters on the y-axis. The colors indicate the corresponding cluster. **C** Pseudotime trajectory: each dot represents a cell color-coded for pseudotime. **D, E** Proportions of cells in Cd4 T cell sub-cluster among different conditions. Bar graph shows on the x-axis the conditions, while the percentage of cells per cluster is plotted on the y-axis.
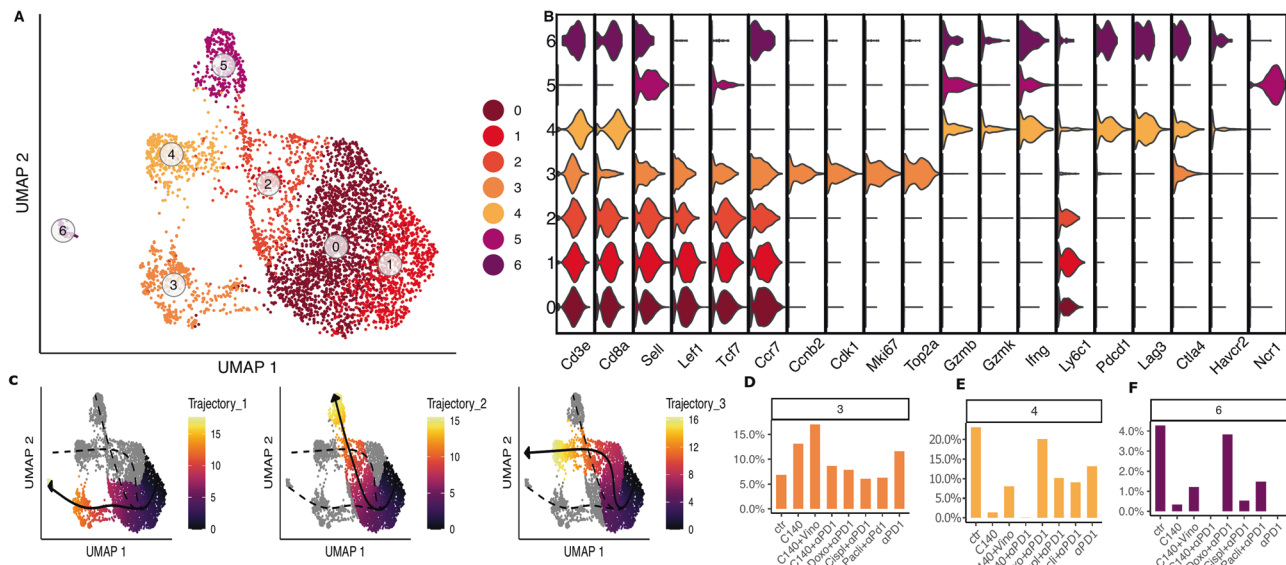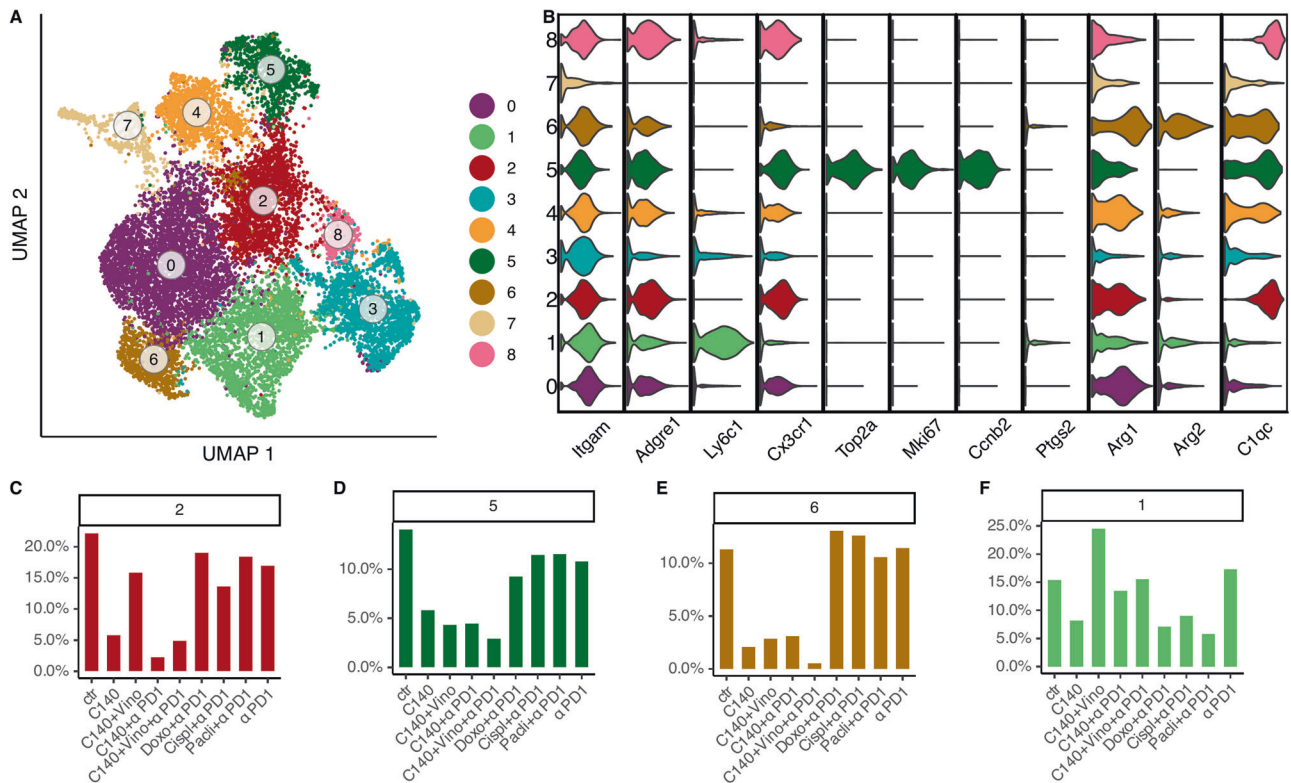


**Fig. 3  Cd8-like T cell sub-clustering analysis in 4T1. A** UMAP of Cd8 T cell sub-clustering in 4T1 cell line. Each dot refers to a cell. The colors refer to the sub-clusters. **B** Gene signature: the violin plot shows the genes on the x-axis, while the clusters on the y-axis. The colors indicate the corresponding cluster. **C** Pseudotime trajectory: each dot represents a cell color-coded for pseudotime. **D–F** Proportions of cells in each 4T1 Cd8-like T cell sub-cluster among different conditions. Bar graph shows on the x-axis the conditions, while the percentage of cells per cluster is plotted on the y-axis.

C140, C140 + V, C140 + αPD-1 [22] confirming their immuno-suppressive activity. γδ T cell cluster (cluster 4) followed a similar pattern, being high in the control and lower than the control in all the treatments (Fig. 2E). This might be associated with a recently discovered mouse pro-tumoral activity of these specific cell populations [44, 49], also confirmed by the high expression of *Il17a*, a marker proved to promote the expansion of pro-tumoral γδ T cells [50]. The two exhausted-like T cell clusters, related in the pseudotime trajectory analysis, had opposite trends with the more exhausted (cluster 6) being more expanded in treatments with high efficacy than cluster 5 (Fig. S4).

**The Cd8 T cell-like composition in the TME of 4T1 TNBC**
We carried out a sub-clustering of Cd8-like T cells like as we did for the Cd4; this led to the identification of 7 clusters (Fig. 3A) that can be associated, also in this case, to different cell types on the basis of a series of gene markers (Fig. 3B). Two progenitor-like Cd8 clusters (0 and 1) were identified as confirmed by the expression of *Sell*, *Lef1*, *Tcf7* and *Ccr7* genes. Three clusters (clusters 3, 4 and 6) had terminally differentiated profiles. Specifically, cells belonging to cluster 3 showed high expression of genes associated with proliferation of phase 2 cell cycle such as *Ccnb2*, *Cdk1*, *Mki67* and *Top2a*. Cluster 4 was the most active cluster since it presented high levels of *Gzmb*, *Gzmk*, *Ifng* and *Ly6c1* genes, markers

**Fig. 4 Macrophage-like cell sub-clustering analysis in EMT6. A** UMAP of macrophages sub-clustering in EMT6 cell line. Each dot refers to a cell and the colors refer to the sub-clusters. **B** Gene signature: the violin plot shows the genes on the x-axis, while the clusters on the y-axis. The colors indicate the corresponding cluster. **C–F** Proportions of cells in each EMT6 macrophages sub-cluster among different conditions. Bar graph shows on the x-axis the conditions, while the percentage of cells per cluster is plotted on the y-axis.

characteristic of effector cells. Furthermore, cluster 6 was defined as exhausted-like Cd8 cluster since it had high levels of *Pdcd1*, *Lag3*, *Ctla4*, *Havcr2* gene expression. Finally, cluster 5 was associated with NKs as the high level of *Ncr1* gene expression suggested (Fig. 3B).

Looking at the sub-cluster cell proportions along the conditions (Fig. S5), one of the most relevant results was that the highly proliferative Cd8 T cells belonging to cluster 3 increased in TMEs treated with C140 and almost doubled their percentage, compared to the control, in C140 alone and C140 + V treated TMEs, underlining a possible anti-tumoral effect. An opposite trend was recorded for the exhausted-like Cd8 T cluster (cluster 6): their cells decreased or were not found in treatments that involved cyclophosphamide while they were present in high percentage in the untreated TME. Similarly, cluster 4 was found enriched in the untreated control while its percentage of cells dropped down in C140, C140 + V and C140 + αPD-1 treatments (Fig. 3D–F).

Additionally, phenotypic heterogeneity along the Cd8 T cell-like sub-clusters, as for that of the Cd4 cells, was visualized using trajectory analyses. As for the Cd4 sub-clustering, the trajectory of Cd8 sub-clusters (Fig. 3C) revealed three distinct lineages that share a common cluster as root. This cluster was the progenitor-like cell cluster 1. The first trajectory linked the root with the most-exhausted Cd8 cluster (cluster 6). The second related cluster 1 with the cluster referred to NKs; while the third showed a connection with the most-active Cd8 cluster (cluster 4) passing through cluster 2 as for the second trajectory. This analysis confirms both the assignment done previously and the gene expression signature for each cluster.

**M1- and M2 -like tumor associated macrophages populations in the EMT6 TMEs**
Abundance in macrophages found in EMT6 TMEs allowed a sub-clustering of these cells. This led to the identification of 9 clusters

of cells sharing similar transcriptional profiles (Fig. 4A); of which clusters 2, 5 and 6 were associated to a M2-macrophage subtype because of the expression of *Cx3cr1* gene and the negative expression of *Ly6c1* gene [51] (Fig. 4B). Specifically, cluster 6 presented moderate expression of the *Arg2* gene and less extent of *Ptgs2* gene that results in immune suppression [52]. Instead, cells belonging to cluster 5, highly expressed gene markers associated with proliferation such as *Ccnb2*, *Mki67* and *Top2a* genes defining this cluster as M2-like proliferative. While cluster 2 highly expressed *C1qa/C1qc* genes that were found upregulated in a previously reported M2-like tumor associated macrophage cluster [42].

On the other hand, cluster 1 resulted in being associated with a M1-like macrophagic population due to the expression of *Ly6c1/ Ly6c2* gene [51, 53]. The remaining clusters, due to the continuous expression of most gene markers along multiple cell populations, were of uncertain classification as also observed in [54] and [55].

Observing the proportion of cells for each cluster among the different conditions, the most relevant results were shown when observing the M2-like clusters 2, 5 and 6 (Fig. 4C–E, Fig. S6). They exhibited a similar trend; indeed, their number of cells had high frequencies in the untreated condition and decreased in all the treatments involving cyclophosphamide. On the other hand, M1-like cluster 1 exhibited a remarkable increase only with cyclophosphamide combined with vinorelbine (C140 + V) treatment (Fig. 4F).

**Common B cell sub-cluster proportions on both the cell line TMEs**
Thanks to the comparable number of cells classified as B cells in the two TNBC cell line TMEs, we explored similarities and differences of the sub-clustering of these cell populations (Fig. 5). The TMEs of the 4T1 cell line displayed five clusters, while eight were the clusters observed in the TMEs of EMT6. Strikingly, the EMT6 model had a
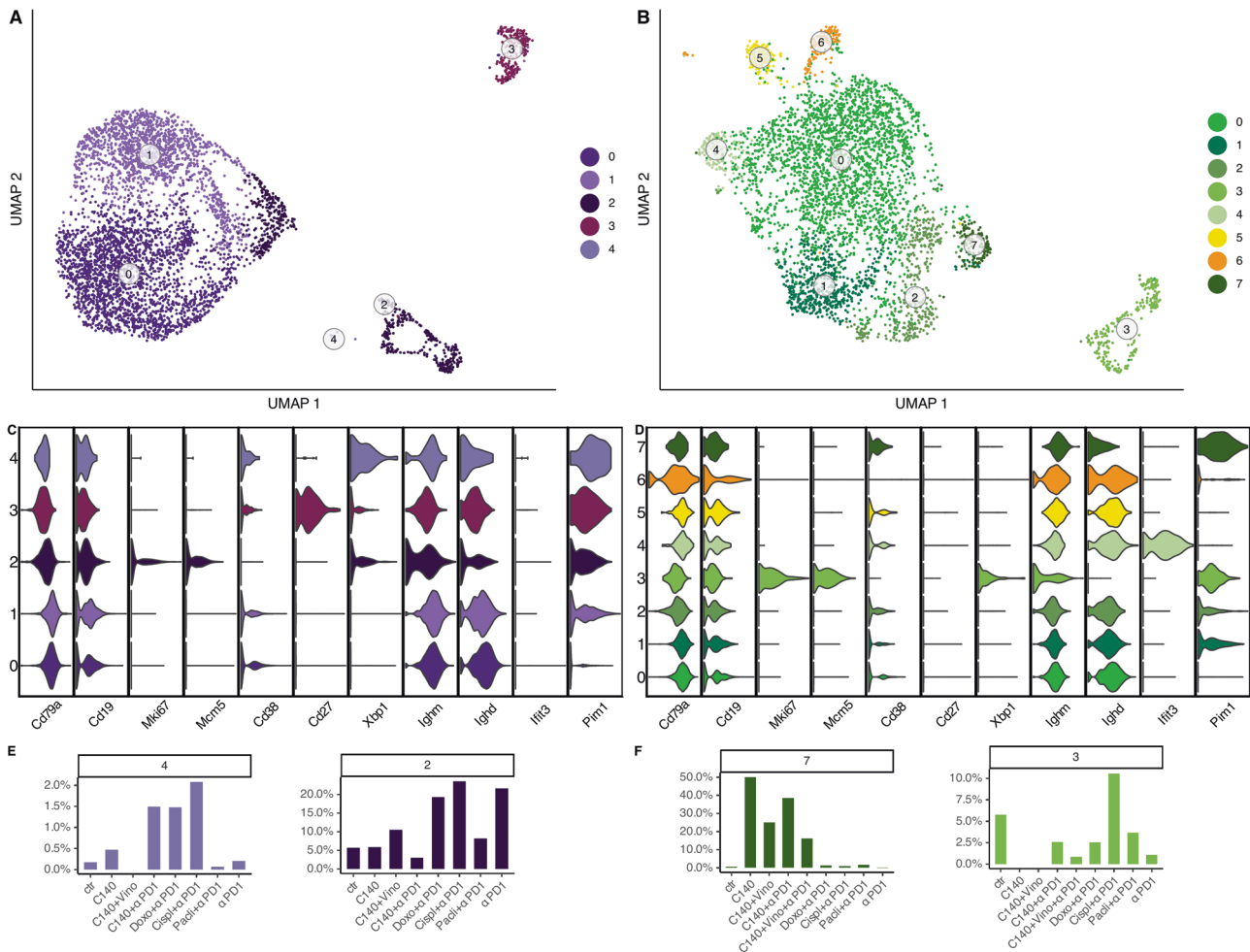
**Fig. 5   B-like cell sub-clustering analysis in 4T1 and EMT6.** UMAP of B cell sub-clustering in 4T1 **A** and EMT6 **B**. Gene signature of B cell sub-clustering in 4T1 **C** and EMT6 **D**. Proportions of cells in relevant 4T1 **E** and EMT6 **F** B cell sub-clusters among different conditions.

more structured sub-clustering revealing more sub-populations. Among these, EMT6 cluster 4 resembled a interferon-induced naive B cell (expression of *Ifit3*) and clusters 1 and 2, as part of cells belonging to cluster 1 in the 4T1, were classified as an intermediated state between proliferative and naive-like B cells as clear by the expression of *Pim1* (Fig. 5C, D) [56]. On the other hand, clusters 2 and 3, in 4T1 and EMT6 respectively, had a similar transcriptional profile attributable to a proliferative B cell population, as confirmed by the expression of *Mki67* and *Mcm5* genes [39, 57]. Similarly, cluster 4 and cluster 7, in 4T1 and EMT6 respectively, highly expressed *Cd38* gene but not *Mki67* (Fig. 5C, D). These genes have been associated with germinal B cells. Notably, one cluster on the 4T1 (cluster 3) displayed a gene signature typical of plasma B cells (high expression of *Cd27*, *Cd38*, *Xbp1*) [39, 40] that was not observed in the EMT6 counterpart despite the higher number of subpopulations retrieved. As reported from the scRNA-seq analysis of human nasopharyngeal carcinoma TMEs [25], a large number of less differentiated B cell subpopulations identified as naive-like cells were found on both the cell line TMEs. This was the case of cluster 0 for both 4T1 and EMT6, characterized by the expression of *Ighm*, *Ighd* but not *Cd27* [25, 40].

Strikingly, the proportions of cells sharing the same transcriptional profile were generally comparable between the treated and untreated TMEs of the two cell lines. Indeed, the germinal B cells (cluster 4 and 7, in 4T1 and EMT6 respectively) in the untreated sample were few while in the C140 + αPD-1 treatments increased in both the murine models (Fig. 5F–H). In addition, similarities on the proportions of the proliferative B cell clusters were found.

Specifically, cluster 2, for 4T1, decreased only in correspondence of cyclophosphamide alone or in combination with the immunotherapy compared to the control; while the percentage of cells in cluster 3 in EMT6 decreased in all the treatments but the cisplatin in combination with anti-PD-1 (Fig. 5E, G). Interestingly, Cispl + αPD-1 condition presented the highest percentage of proliferative B cells for both the clusters, suggesting a common behavior of the two tumors in response to this specific combination of treatment (Figs. 5C and S7).

## DISCUSSION
Here we report a fine characterization of the immune transcriptional profiles of almost 50,000 single cells in two TNBC murine cell line TMEs. Major differences included a major macrophagic component in the EMT6 TME versus a great number of Cd4⁻ and Cd8⁻ T cells into 4T1 TME. A comparable percentage of B cells and neutrophils were observed in both models. This is in agreement with flow cytometry (FC) results [21]; however, scRNA-seq analyses include less cells versus FC; moreover, comparing populations previously characterized by FC with surface protein markers, and for which a gene expression profile is unknown, with cell populations finely characterized by gene expression profiles using scRNA-seq technique is challenging.

Sub-clustering of T cells in the TMEs of the 4T1 cell line revealed previously uncharacterized sub-populations with unique transcriptional profiles. Trajectory analyses highlighted some of these

populations as intermediate states and identified a tripartite differentiation for both Cd4- and Cd8- like T cells as also reported in human TNBC TMEs [48]. Subsets associated with regulatory and γδ T cells (expressing high level of *Il17a*) decreased in conditions with a higher pre-clinical efficacy in in vivo experiments [22]. Although the association of regulatory T cells and poor prognosis in multiple cancer types has been widely characterized, the pro-tumoral activity of γδ T cells in murine TNBC was poorly or never observed at scRNA-seq level [58], therefore, here we helped in validating their function along with the expression of *Il17a* [44, 50]. These data will help to better understand the role of this specific subset. We also confirmed the increase of exhausted-like Cd8 T cell subpopulation in pre-clinical treatments with low in vivo efficacy and in the untreated samples [22]. This is in line with tumor escape by expression of checkpoint inhibitors [15]. On the other hand, a proliferative Cd8-like T cell sub-cluster (expressing high levels of Mki67) was found to increase in correspondence of the treatments with cyclophosphamide alone or in combination with other chemotherapy/immunotherapy. This is in accordance with the association of proliferative Cd8 T cells and better outcome in cancer [59].

Macrophage-like cells on the TMEs of EMT6 TNBC revealed subtypes expressing genes related to M2-like macrophages enriched in untreated condition and in treatments with poor efficacy in in vivo experiments [22]. This confirms their known pro-tumoral activity [60]. Contrarily, clusters expressing genes related to M1-like macrophages were increased after efficient preclinical treatments.

Finally, we noted a common behavior along some conditions (C140 + αPD-1, C140 + V and Cispl+αPD-1) of two B cell sub-clusters presenting similar gene signatures in both the murine tumor cell lines. These clusters were associated with proliferative B and germinal B cells and followed opposite trends for some of the high efficacy treatments. The alkylating agent cisplatin in combination with the immunotherapy αPD-1 favored the expansion of germinal B cells. Plasma B cells, only identified in 4T1 TME, increased after most efficient treatments (Fig. S7), confirming their association with improved survival [61, 62].

Our fine-scale characterization of the immune TME could be used as a resource to novel studies with the aim to improve the choice of treatment in TNBC patients.

### Future perspectives and limitations of the study
In this context a detailed extension and validation of the analyses is required due to some limitations related to scRNA-seq. The number of cells in some sub-clusters were really low and therefore a more detailed analyses either using classic FC or a scRNA-seq only on sorted cells belonging to those specific sub-clusters is required. In addition, more replications of each condition reported in this work and the transcriptional investigation of Cd45⁻ cells populating the two TMEs with the aim to investigate also the release of specific chemokines and cytokines of the tumor cell could strengthen the results obtained here.

### DATA AVAILABILITY
Raw count matrices generated with the scRNA-seq Sankemake CellRanger v4 alignment pipeline and pre-QC *Seurat* objects are available with the accession number GSE191246 at the GEO (http://www.ncbi.nlm.nih.gov/geo/).

### CODE AVAILABILITY
The Snakemake CellRanger v4 pipeline is available at this link https://github.com/raveancic/scRNAaltas_TNBC_mm/tree/master/cl_crt_FASTQ2countmat

### REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality world-wide for 36 cancers in 185 countries. CA: A Cancer J Clin. 2021;71:209–49.
2. Kalimutho M, Parsons K, Mittal D, López JA, Srihari S, Khanna KK. Targeted therapies for triple-negative breast cancer: combating a stubborn disease. Trends Pharmacol. Sci. 2015;36:822–46.
3. Korkmaz A, Topal T, Oter S. Pathophysiological aspects of cyclophosphamide and ifosfamide induced hemorrhagic cystitis; implication of reactive oxygen and nitrogen species as well as PARP activation. Cell Biol Toxicol. 2007;23:303–12.
4. Rosenberg B, Vancamp L, Trosko JE, Mansour VH. Platinum compounds: a new class of potent antitumour agents. Nature. 1969;222:385–6.
5. Jordan MA, Wilson L. Microtubules as a target for anticancer drugs. Nat Rev Cancer. 2004;4:253–65.
6. Gewirtz D. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. Biochem. Pharmacol. 1999;57:727–41.
7. Seidel JA, Otsuka A, Kabashima K. Anti-PD-1 and anti-CTLA-4 therapies in cancer: mechanisms of action, efficacy, and limitations. Front Oncol.2018;8:86
8. Heimes A-S, Schmidt M. Atezolizumab for the treatment of triple-negative breast cancer. Expert Opin Investig Drugs. 2019;28:1–5.
9. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010;363:711–23.
10. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based che-motherapy: a single-arm, multicentre, phase 2 trial. Lancet. 2016;387:1909–20.
11. Freeman GJ, Long AJ, Iwai Y, Bourque K, Chernova T, Nishimura H, et al. Engagement of the Pd-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. J Exp Med. 2000;192:1027–34.
12. Yearley JH, Gibson C, Yu N, Moon C, Murphy E, Juco J, et al. PD-L2 expression in human tumors: relevance to anti-PD-1 therapy in cancer. Clin Cancer Res. 2017;23:3158–67.
13. Seliger B. Basis of PD1/PD-L1 therapies. J Clin Med. 2019;8:2168.
14. Han Y, Liu D, Li L. PD-1/PD-L1 pathway: current researches in cancer. Am J Cancer Res. 2020;10:727–42.
15. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012;12:252–64.
16. Yi JS, Cox MA, Zajac AJ. T-cell exhaustion: characteristics, causes and conversion. Immunology. 2010;129:474–81.
17. Francisco LM, Salinas VH, Brown KE, Vanguri VK, Freeman GJ, Kuchroo VK, et al. PD-L1 regulates the development, maintenance, and function of induced reg-ulatory T cells. J Exp Med. 2009;206:3015–29.
18. Zhu Y, Zhu X, Tang C, Guan X, Zhang W. Progress and challenges of immu-notherapy in triple-negative breast cancer. Biochim Biophys Acta (BBA) - Rev Cancer. 2021;1876:188593.
19. Chen H, Ye F, Guo G. Revolutionizing immunology with single-cell RNA sequencing. Cell Mol Immunol. 2019;16:242–9.
20. Haque A, Engel J, Teichmann SA, Lönnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. 2017;9:75.
21. Orecchioni S, Talarico G, Labanca V, Calleri A, Mancuso P, Bertolini F. Vinorelbine, cyclophosphamide and 5-FU effects on the circulating and intratumoural land-scape of immune cells improve anti-PD-L1 efficacy in preclinical models of breast cancer and lymphoma. Br J Cancer. 2018;118:1329–36.
22. Falvo P, Orecchioni S, Hillje R, Raveane A, Mancuso P, Camisaschi C, et al. Cyclophosphamide and vinorelbine activate stem-like CD8+ T cells and improve anti-PD-1 efficacy in triple-negative breast cancer. Cancer Res. 2021;81:685–97.
23. Reggiani F, Labanca V, Mancuso P, Rabascio C, Talarico G, Orecchioni S, et al. Adipose progenitor cell secretion of GM-CSF and MMP9 promotes a stromal and immunological microenvironment that supports breast cancer progression. Cancer Res. 2017;77:5169–82.
24. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, et al. Sustainable data analysis with Snakemake. F1000Res. 2021;10:33.
25. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573. e29
26. Germain P-L, Lun A, Macnair W, Robinson MD. Doublet identification in single-cell sequencing data using scDblFinder. F1000Research. 2021;10:979.
27. Amezquita RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with bioconductor. Nat Methods. 2020;17:137–45.
28. Wu J, Waxman DJ. Metronomic cyclophosphamide eradicates large implanted GL261 gliomas by activating antitumor Cd8+ T-cell responses and immune memory. OncoImmunology. 2015;4:e1005521.
29. Pfirschke C, Engblom C, Rickelt S, Cortez-Retamozo V, Garris C, Pucci F, et al. Immunogenic chemotherapy sensitizes tumors to checkpoint blockade therapy. Immunity. 2016;44:343–54.

30. Trail PA, Willner D, Bianchi AB, Henderson AJ, TrailSmith MD, Girit E, et al. Enhanced antitumor activity of paclitaxel in combination with the anticarcinoma immunoconjugate BR96-doxorubicin. Clin Cancer Res. 1999;5:3632–8.

31. Rainer J, Gatto L, Weichenberger CX. ensembldb: an R package to create and use Ensembl-based annotation resources. Bioinformatics. 2019;35:3151–3.

32. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. GigaScience. 2018;7:giy083.

33. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20:163–72.

34. Frankish A, Diekhans M, Ferreira A-M, Johnson R, Jungreis I, Loveland J, et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. 2019;47:D766–D773.

35. Franzén O, Gan L-M, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database. 2019;2019:baz046.

36. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. 2015;16:278.

37. Johnsson K Intrinsic Dimension and Cluster Analysis. *Centre for Mathematical Sciences*, Lund University, 2016. 188 p.

38. Germain P-L, Sonrel A, Robinson MD. pipeComp, a general framework for the evaluation of computational pipelines, reveals performant single cell RNA-seq preprocessing tools. Genome Biol. 2020;21:227.

39. Glaros V, Rauschmeier R, Artemov AV, Reinhardt A, Ols S, Emmanouilidi A, et al. Limited access to antigen drives generation of early B cell memory while restraining the plasmablast response. Immunity. 2021;54:2005–23. e10

40. Hu Q, Hong Y, Qi P, Lu G, Mai X, Xu S, et al. Atlas of breast cancer infiltrated B-lymphocytes revealed by paired single-cell RNA-sequencing and antigen receptor profiling. Nat Commun. 2021;12:2186.

41. Wisdom AJ, Mowery YM, Hong CS, Himes JE, Nabet BY, Qin X, et al. Single cell analysis reveals distinct immune landscapes in transplant and primary sarcomas that determine response or resistance to immunotherapy. Nat Commun. 2020;11:6410.

42. Zhang Y, Chen H, Mo H, Hu X, Gao R, Zhao Y, et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell. 2021;39:1578–93.

43. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018;19:477.

44. Rei M, Pennington DJ, Silva-Santos B. The emerging protumor role of γδ T lymphocytes: implications for cancer immunotherapy. Cancer Res. 2015;75:798–802.

45. Campbell KR, Yau C. Uncovering pseudotemporal trajectories with covariates from single cell and bulk expression data. Nat Commun. 2018;9:2442.

46. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature. 2018;560:494–8.

47. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32:381–6.

48. Bassez A, Vos H, Van Dyck L, Floris G, Arijs I, Desmedt C, et al. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. Nat Med. 2021;27:820–32.

49. Silva-Santos B, Serre K, Norell H. γδ T cells in cancer. Nat Rev Immunol. 2015;15:683–91.

50. Rei M, Gonçalves-Sousa N, Lança T, Thompson RG, Mensurado S, Balkwill FR, et al. Murine CD27(−) Vγ6(+) γδ T cells producing IL-17A promote ovarian cancer growth via mobilization of protumor small peritoneal macrophages. Proc Natl Acad Sci USA. 2014;111:E3562–E3570.

51. Shechter R, Miller O, Yovel G, Rosenzweig N, London A, Ruckh J, et al. Recruitment of beneficial M2 macrophages to injured spinal cord is orchestrated by remote brain choroid plexus. Immunity. 2013;38:555–69.

52. Grzywa TM, Sosnowska A, Matryba P, Rydzynska Z, Jasinski M, Nowis D, et al. Myeloid cell-derived arginase in cancer immune response. Front Immunol. 2020;11:938.

53. Ni G, Yang X, Li J, Wu X, Liu Y, Li H, et al. Intratumoral injection of caerin 1.1 and 1.9 peptides increases the efficacy of vaccinated TC-1 tumor-bearing mice with PD-1 blockade by modulating macrophage heterogeneity and the activation of CD8+ T cells in the tumor microenvironment. Clin Transl Immunol. 2021;10:e1335.

54. Qu Y, Wen J, Thomas G, Yang W, Prior W, He W, et al. Baseline frequency of inflammatory Cxcl9-expressing tumor-associated macrophages predicts response to avelumab treatment. Cell Rep. 2020;32:107873.

55. Samstein RM, Krishna C, Ma X, Pei X, Lee K-W, Makarov V, et al. Mutations in BRCA1 and BRCA2 differentially affect the tumor microenvironment and response to checkpoint blockade immunotherapy. Nat Cancer. 2020;1:1188–203.

56. Wu Y, Deng Y, Zhu J, Duan Y, Weng W, Wu X. Pim1 promotes cell proliferation and regulates glycolysis via interaction with MYC in ovarian cancer. Onco Targets Ther. 2018;11:6647–56.

57. Gong L, Kwong DL-W, Dai W, Wu P, Li S, Yan Q, et al. Comprehensive single-cell sequencing reveals the stromal dynamics and tumor-specific characteristics in the microenvironment of nasopharyngeal carcinoma. Nat Commun. 2021;12:1540.

58. Park JH, Lee HK. Function of γδ T cells in tumor immunology and their application to cancer therapy. Exp Mol Med. 2021;53:318–27.

59. Kamphorst AO, Pillai RN, Yang S, Nasti TH, Akondy RS, Wieland A, et al. Proliferation of PD-1+ CD8 T cells in peripheral blood after PD-1–targeted therapy in lung cancer patients. Proc Natl Acad Sci USA. 2017;114:4993–8.

60. Lin Y, Xu J, Lan H. Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. J Hematol Oncol. 2019;12:76.

61. Griss J, Bauer W, Wagner C, Simon M, Chen M, Grabmeier-Pfistershammer K, et al. B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. Nat Commun. 2019;10:4186.

62. GuhaThakurta D, Sheikh NA, Fan L-Q, Kandadi H, Meagher TC, Hall SJ, et al. Humoral immune response against nontargeted tumor antigens after treatment with Sipuleucel-T and its association with improved clinical outcome. Clin Cancer Res. 2015;21:3619–30.

## AUTHOR CONTRIBUTIONS

FB directed the study. FB and AR designed the study. LC performed the analyses under the supervision of AR and FB. SM and RH provided computational support for the analyses. PF, SO, GM, and PM provided support for the interpretation of laboratory analyses. LC, AR, and FB wrote the original draft. All the authors discussed the results and contributed to the final version of the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41420-022-00893-x.

**Correspondence** and requests for materials should be addressed to Alessandro Raveane or Francesco Bertolini.

**Reprints and permission information** is available at http://www.nature.com/reprints

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