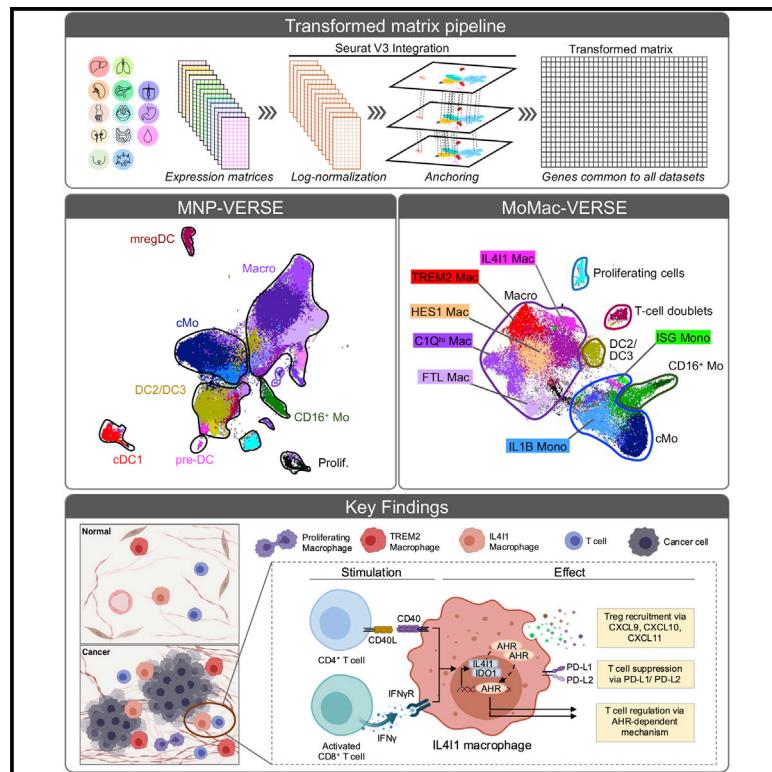


# Immunity

## Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease

### Graphical abstract



### Highlights

- Cross-tissue integration of scRNA from monocytes and macrophages in health and disease
- Conserved gene signatures of mononuclear phagocyte populations in human tissues
- IL4I1<sup>+</sup>PD-L1<sup>+</sup>IDO1<sup>+</sup> and TREM2<sup>+</sup> TAM subsets accumulate in human tumors
- IL4I1<sup>+</sup>PD-L1<sup>+</sup>IDO1<sup>+</sup> TAM in the tumor periphery exhibit immunosuppressive characteristics

### Authors

Kevin Mulder, Amit Ashok Patel,  
Wan Ting Kong, ..., Camille Blériot,  
Charles-Antoine Dutertre,  
Florent Ginhoux

### Correspondence

charles-antoine.dutertre@inserm.fr (C.-A.D.),  
florent\_ginhoux@immunol.a-star.edu.sg (F.G.)

### In brief

Mulder et al. integrate 178,651 human mononuclear phagocytes (MNPs) from 13 tissues across 41 datasets to generate a MNP single-cell RNA compendium (MNP-VERSE) that enables the definition of conserved gene signatures of MNP populations. This integrated approach provides a robust, online-available platform (<https://gustaveroussy.github.io/FG-Lab/>) for uniform annotation and dissection of specific macrophage functions in healthy and pathological states.



## Resource

# Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease

Kevin Mulder,<sup>1,2,3,23</sup> Amit Ashok Patel,<sup>1,2,23</sup> Wan Ting Kong,<sup>3,23</sup> Cécile Piot,<sup>3,24</sup> Evelyn Halitzki,<sup>3,4</sup> Garrett Dunsmore,<sup>1,2</sup> Shabnam Khalilnezhad,<sup>3</sup> Sergio Erdal Irac,<sup>3</sup> Agathe Dubuisson,<sup>1,2</sup> Marion Chevrier,<sup>3</sup> Xiao Meng Zhang,<sup>3</sup> John Kit Chung Tam,<sup>5</sup> Tony Kiat Hon Lim,<sup>6</sup> Regina Men Men Wong,<sup>7</sup> Rhea Pai,<sup>7</sup> Ahmed Ibrahim Samir Khalil,<sup>8</sup> Pierce Kah Hoe Chow,<sup>9</sup> Suny Z. Wu,<sup>10,11</sup> Ghamdan Al-Eryani,<sup>10,11</sup> Daniel Roden,<sup>10,11</sup>

(Author list continued on next page)

<sup>1</sup>Gustave Roussy Cancer Campus, Villejuif, France

<sup>2</sup>Institut National de la Santé Et de la Recherche Médicale (INSERM) U1015, Equipe Labellisée—Ligue Nationale contre le Cancer, Villejuif, France

<sup>3</sup>Singapore Immunology Network (SIgN), A\*STAR, 8A Biomedical Grove, Immunos Building, Singapore 138648, Singapore

<sup>4</sup>Life and Medical Sciences Institute, University of Bonn, Bonn, Germany

<sup>5</sup>Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, 1E Kent Ridge Road, Singapore 119228, Singapore

<sup>6</sup>Department of Pathology, Singapore General Hospital, 20 College Road, Singapore 169856, Singapore

<sup>7</sup>Genome Institute of Singapore, A\*STAR, 60 Biopolis Street, Genome, #02-01, Singapore 138672, Singapore

<sup>8</sup>Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

<sup>9</sup>Division of Surgical Oncology, National Cancer Centre, Singapore 169610, Singapore

<sup>10</sup>The Kinghorn Cancer Centre and Cancer Research Division, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia

<sup>11</sup>St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Sydney, NSW 2052, Australia

<sup>12</sup>Department of Reproductive Medicine, KK Women's and Children's Hospital, Singapore 229899, Singapore

<sup>13</sup>Academic Clinical Program in Obstetrics and Gynaecology, Duke-NUS Medical School, Singapore 169857, Singapore

<sup>14</sup>Experimental Fetal Medicine Group, Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, 1E Kent Ridge Road, Singapore 119228, Singapore

<sup>15</sup>Program in Emerging Infectious Disease, Duke-NUS Medical School, 8 College Road, Singapore 169857, Singapore

<sup>16</sup>Translational Immunology Institute, SingHealth/Duke-NUS Academic Medical Centre, the Academia, 20 College Road, Discovery Tower Level 8, Singapore 169856, Singapore

(Affiliations continued on next page)

## SUMMARY

Mononuclear phagocytes (MNPs) encompass dendritic cells, monocytes, and macrophages (MoMac), which exhibit antimicrobial, homeostatic, and immunoregulatory functions. We integrated 178,651 MNPs from 13 tissues across 41 datasets to generate a MNP single-cell RNA compendium (MNP-VERSE), a publicly available tool to map MNPs and define conserved gene signatures of MNP populations. Next, we generated a MoMac-focused compendium that revealed an array of specialized cell subsets widely distributed across multiple tissues. Specific pathological forms were expanded in cancer and inflammation. All neoplastic tissues contained conserved tumor-associated macrophage populations. In particular, we focused on IL4I1<sup>+</sup> CD274(PD-L1)<sup>+</sup>IDO1<sup>+</sup> macrophages, which accumulated in the tumor periphery in a T cell-dependent manner via interferon- $\gamma$  (IFN- $\gamma$ ) and CD40/CD40L-induced maturation from IFN-primed monocytes. IL4I1\_Macs exhibited immunosuppressive characteristics through tryptophan degradation and promoted the entry of regulatory T cell into tumors. This integrated analysis provides a robust online-available platform for uniform annotation and dissection of specific macrophage functions in healthy and pathological states.

## INTRODUCTION

The mononuclear phagocyte system (MPS) was introduced almost 50 years ago and proposed a new classification of macrophages (Macs), monocytes (Mo), and their precursor cells (van Furth et al., 1972). Following this, the identification of dendritic

cells (DCs) (Steinman and Cohn, 1973), led to a further addition to the MPS. Dichotomies have thus emerged with classical versus non-classical Mo, plasmacytoid versus classical DC (cDC), and pro-inflammatory M1 versus anti-inflammatory M2 Macs (Geissmann et al., 2003; Merad et al., 2013; Mills et al., 2000; Passlick et al., 1989; Ziegler-Heitbrock et al., 2010).



Alexander Swarbrick,<sup>10,11</sup> Jerry Kok Yen Chan,<sup>12,13,14,15</sup> Salvatore Albani,<sup>16</sup> Lisa Derosa,<sup>1,17</sup> Laurence Zitvogel,<sup>1,2,17,18,19</sup> Ankur Sharma,<sup>7,20,21</sup> Jinmiao Chen,<sup>3</sup> Aymeric Silvin,<sup>1,2</sup> Antonio Bertolletti,<sup>15</sup> Camille Blériot,<sup>1,2</sup> Charles-Antoine Dutertre,<sup>1,2,3,16,23,\*</sup> and Florent Ginhoux<sup>3,16,22,23,25,\*</sup>

<sup>17</sup>Cancer Medicine Department, Gustave Roussy, Villejuif, France

<sup>18</sup>Université Paris-Saclay, Ile-de-France, France

<sup>19</sup>Center of Clinical Investigations in Biotherapies of Cancer (BIOTHERIS), 1428 Villejuif, France

<sup>20</sup>Harry Perkins Institute of Medical Research, QEII Medical Centre and Centre for Medical Research, the University of Western Australia, PO Box 7214, 6 Verdun Street, Nedlands, Perth, WA 6009, Australia

<sup>21</sup>School of Pharmacy and Biomedical Sciences, Curtin Health Innovation Institute, Curtin University, Perth, WA 6102, Australia

<sup>22</sup>Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

<sup>23</sup>These authors contributed equally

<sup>24</sup>Present address: Immunobiology Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

<sup>25</sup>Lead contact

\*Correspondence: charles-antoine.dutertre@inserm.fr (C.-A.D.), florent\_ginhoux@immunol.a-star.edu.sg (F.G.)

<https://doi.org/10.1016/j.immuni.2021.07.007>

However, the application of high dimensional approaches such as mass cytometry and single-cell RNA sequencing (scRNA-seq) has revealed finer levels of heterogeneity, leading to the identification of subpopulations in both healthy and diseased states and across various tissues (Passlick et al., 1989; See et al., 2017; Villani et al., 2017; Ziegler-Heitbrock et al., 2010).

In particular, scRNA-seq has revolutionised our understanding of immune cell heterogeneity by providing in-depth snapshots of individual cell activity (Giladi and Amit, 2018; Stubbington et al., 2017). These technologies have been exploited to generate accessible atlases of the mouse (The Tabula Muris Consortium, 2018; Van Hove et al., 2019) and human (Regev et al., 2017; Rosenblatt-Rosen et al., 2017; Svensson et al., 2020) immune cells, in addition to organ-specific databases (Peters et al., 2020). Nevertheless, comparison between these different studies and extraction of granular detail is difficult owing to the distinct sequencing protocols and analytical pipelines used. A unified analytical approach has the potential to reveal synergy between published datasets and to avoid duplication of effort and irregularities of nomenclature.

This need for transcriptomic data unification is particularly clear for studies of the MNP, especially in disease settings. MNPs are critical for immune defense and homeostasis: patients who are deficient in these cell populations are highly susceptible to recurrent infections (Bigley et al., 2019); while in cancer, tumor-associated macrophages (TAMs) are implicated in promoting tumor progression, angiogenesis, and metastasis (Pollard, 2004), and DC are thought to prime antitumor T cells (Wculek et al., 2020). Multiple studies examining the same or different tissues have defined similar MNP subpopulations, but designated various names for them as seen in the past for the case of DC (Guilliams et al., 2014). To achieve clarity and to extract fundamental commonalities of potential therapeutic importance, there is a need to unify the multiple identities assigned to MNPs across healthy and pathological tissues.

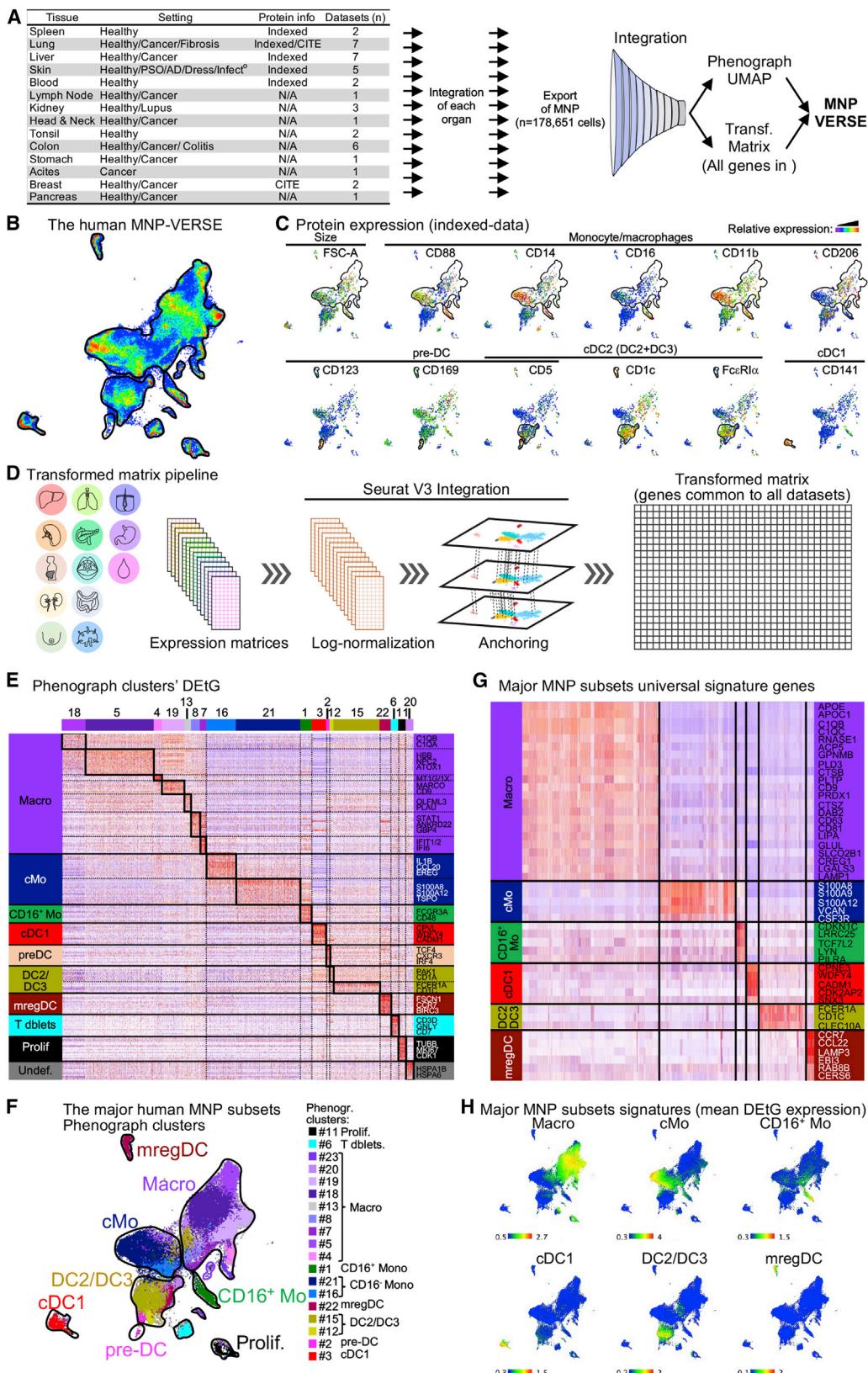
Here, we generated an integrated meta-analysis of multiple transcriptomic studies on MNP to elucidate system-wide characteristics of the MPS in health and disease. We integrated scRNA-seq datasets from 41 studies to build a comprehensive view of the human MNP in health and disease states (the MNP-VERSE). All major previously defined MNP subsets were delineated, including classical and CD16<sup>+</sup> Mo, conventional dendritic cells (cDC) subsets (cDC1, cDC2

[DC2+DC3]), as well as their progenitors (pre-DC) and terminally differentiated mature DCs enriched in immunoregulatory molecules (mregDC) (Dutertre et al., 2019; Maier et al., 2009; See et al., 2017; Ziegler-Heitbrock et al., 2010). We next focused on monocyte and Macs and generated a “MoMac-VERSE.” Using the Azimuth algorithm (Hao et al., 2021), we demonstrated that these VERSEs can be used for *de novo* mapping of datasets that were not initially integrated, a process that can provide a robust annotation tool of MNP subsets for any other human scRNA-seq dataset. The MoMac-VERSE revealed subsets of TAM that were enriched in all of the cancers studied. In particular, a population of IL4I1<sup>+</sup> IDO1<sup>+</sup> Macs appeared to be programmed through interactions with CD40L<sup>+</sup> CD4<sup>+</sup> T cells and activated interferon- $\gamma$ <sup>+</sup> (IFN- $\gamma$ <sup>+</sup>) CD8<sup>+</sup> T cells. In addition, IL-4I1<sup>+</sup> Macs could contribute to tryptophan degradation through the IL-4I1-induced activation of the aryl hydrocarbon receptor (AHR), leading to an accumulation of regulatory T (Treg) cells, thereby establishing an immunosuppressive environment in tumors. This work provides a resource to explore MNPs and more specifically Mo and Macs across human healthy and diseased tissues and can be explored through an online platform (<https://gustaveroussy.github.io/FG-Lab/>).

## RESULTS

### The human MNP-VERSE defines conserved gene signatures of major MNP subsets

To establish an atlas of human MNP across healthy and pathologic tissues, we selected 41 scRNA-seq datasets from studies that included MNP from healthy and pathological tissues (Figure 1A; Table S1). Datasets were initially integrated tissue by tissue using the Seurat V3 pipeline (Stuart et al., 2019) to generate subatlases in which MNPs were identified by the expression of canonical markers, including S100A8 and S100A9 for Mo; C1QA, C1QB and CD68 for Macs; CADM1, CLEC9A, and XCR1 for classical DC1 (cDC1); FCER1A, CD1C, and CD1E for cDC2; and the recently described mature DCs enriched in immunoregulatory molecules (mregDC) genes (Maier et al., 2020) (Figures 1A and S1A). Of note, we excluded plasmacytoid DC (pDC) from the MNP-VERSE as we and others have recently demonstrated the lymphoid origin of these cells (Dress et al., 2019; Rodrigues et al., 2018).



(legend on next page)

A total of 178,651 MNPs were extracted from 13 tissues and integrated into a common universe—the MNP-VERSE (Figure 1B)—that was mapped in a uniform manifold approximation and projection (UMAP) space (Becht et al., 2019) (Figure S1B). Importantly, we included in-house-indexed-SMARTseq2 scRNA-seq data of 1,830 cells from 5 different tissues (spleen, lung, liver, skin, and tonsil), which allowed us to broadly identify the major MNP populations based on surface protein expression, namely CD88<sup>+</sup>CD16<sup>+/−</sup>CD14<sup>+</sup>CD11b<sup>+</sup>CD206<sup>+/−</sup> Mo/Mac, CD123<sup>+</sup>CD5<sup>+</sup>CD169<sup>+</sup> pre-DC, CD141<sup>+</sup> classical cDC1, and CD1c<sup>+</sup>cDC2, encompassing both cDC2 and DC3 as recently reported (Bourdely et al., 2020; Dutertre et al., 2019) (Figures 1C and S1C; Table S2). The populations were validated by analyzing independently our in-house-indexed-SMARTseq2 data that included 5 different tissues (tonsil, spleen, blood, lung, and liver). We identified the major DC and monocyte/macrophage populations based on differentially expressed genes (DEGs) and indexed-data protein expression (Figures S2A and S2B) and back-mapped these populations onto the MNP-VERSE. The annotation aligned with the MNP-VERSE annotations (Figure S2C), validating the approach of integrating SMARTseq2 datasets with 10X and other droplet-based scRNA-seq datasets.

To analyze the gene expression of the 41 datasets, we generated a transformed matrix (transf.matrix) that included all of the cells in the MNP-VERSE (Figure 1D). Of the 41 datasets, 6 datasets included relatively fewer genes and were consequently excluded from the transf.matrix (Figure S1F). To confirm the annotation of the major MNP populations at the gene expression level, we used the Phenograph algorithm (Levine et al., 2015) and calculated the differentially expressed “transformed” genes (DEtGs) for all Phenograph clusters using the transf.matrix (Figures 1E and 1F; Table S3). This analysis allowed us to identify 6 major MNP subsets defined by uniquely expressed DEtGs (UETGs): cDC1, cDC2 (cDC2 and DC3), mregDC, classical Mo (cMo), non-classical/intermediate Mo (CD16<sup>+</sup> Mo), and Macs, confirming the annotation obtained with our indexed data (Figures 1G, 1H, and S1D; Table S4). In addition, a strong correlation between transf.matrix and original counts was observed in selected genes, validating the integrated approach (Figure S1E). Thus, these conserved signatures allowed us to precisely define all of the major MNP subsets from scRNA-seq data across human tissues. These compiled data define the “MNP-VERSE” and provide a platform to assemble findings from the existing literature. In line with other reports, the MNP-VERSE showed that cDC1 and mregDC were separated from the main body of the MNP and formed discrete populations, while Mo and cDC2 were clustering more closely, as were the Mo and Macs. These results support and validate our pipeline, allowing further downstream analysis of Mo and Macs across healthy and diseased tissues.

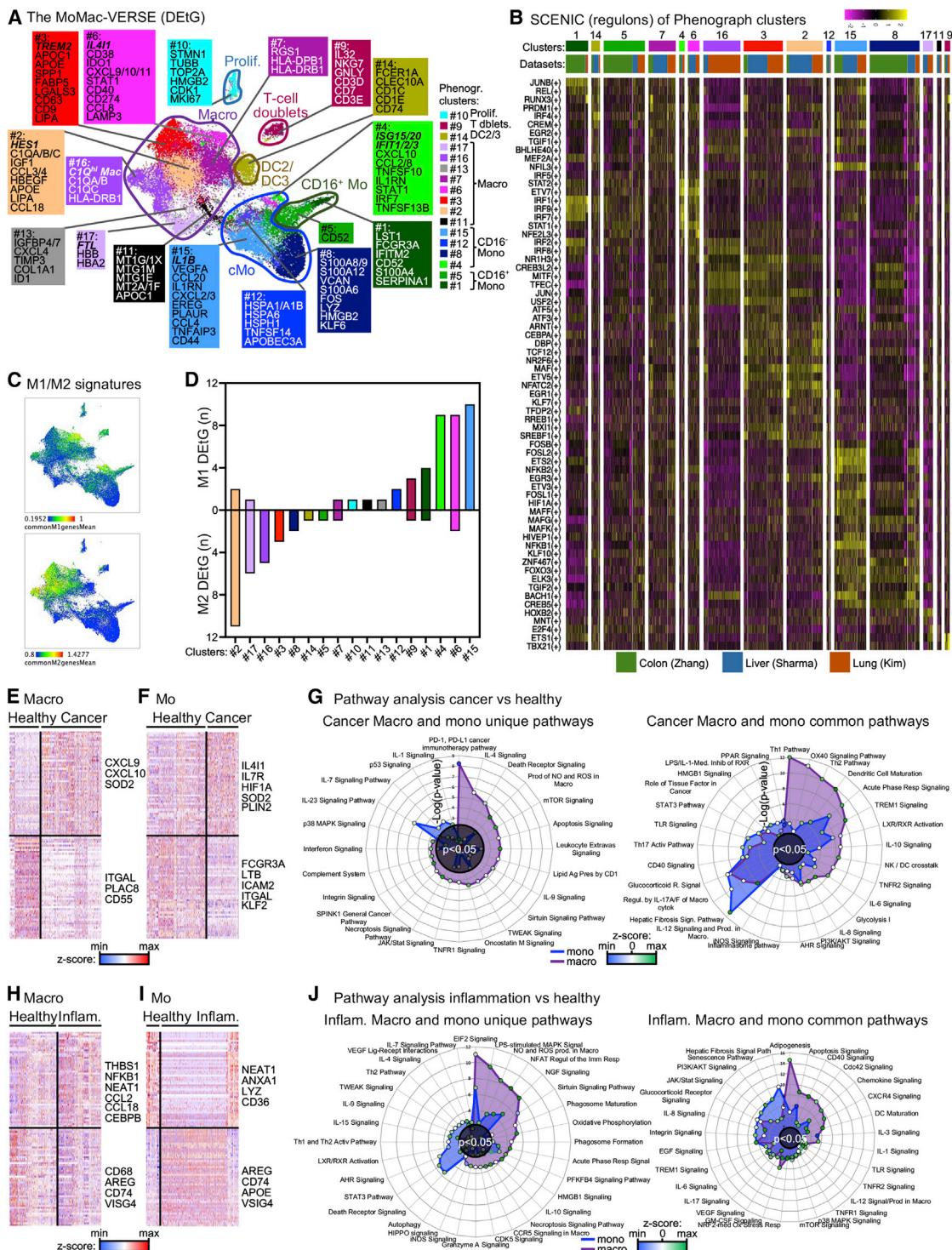
### The human MoMac-VERSE establishes conserved monocyte and macrophage states and identifies patterns of tissue-specific imprinting in health and disease

Having established a broad overview of the integrated MNP data, we extracted Macs and Mo (from healthy and pathologic tissues) and re-integrated them to establish the MoMac-VERSE (Figure 2A). As before, we mapped cells into a UMAP space and calculated DEtGs between the different Phenograph clusters using the transf.matrix (Figures 2A and S3A; Table S3). We identified 2 distinct cell populations comprising Mo (CD16<sup>+</sup> Mo, clusters #1 and #5, and CD16<sup>−</sup> Mo, ISG15<sup>+</sup>ISG20<sup>+</sup>IFIT1<sup>+</sup>IFIT2<sup>+</sup>IFIT3<sup>+</sup> “ISG” #4, #8, #12, and “IL1B” #15) and Macs (“HES1” #2, “TREM2” #3, “IL4I1” #6, #7, #13, #16, and “FTL” [ferritin light chain] #17), respectively. Mac #16 strongly expressed C1QA/B/C and major histocompatibility complex class II (MHC class II) transcripts and comprised lung alveolar Macs (Alv. Mac; Figures S3B and S3C). Interestingly, Mo #4 (ISG\_Mo) and IL4I1\_Mac (#6) shared a similar gene expression signature (Figure S3A; Table S3), suggestive of a close relationship that we explore later in the article. Of note, DEtGs associated with TREM2\_Mac (#3) shared similarities with signatures recently described in murine TREM2 Macs (Figure S3D) (Katzenelenbach et al., 2020; Zhou et al., 2020), suggesting a potential for the MoMac-VERSE to generate cross-species comparisons. We were also able to identify the presence of non-MoMac cells: cluster #14 corresponded to contaminating DC2 and cluster #9 represented MNP/T cell doublets. A third outlying population (#10) was identified as proliferating cells. As above for the MNP-VERSE, we confirmed the relationship between transf.matrix of the MoMac-VERSE and the transcripts per million (TPM) of the original datasets. Six datasets (Kim et al., 2020; Lee et al., 2020; Sharma et al., 2020; Smillie et al., 2019; Zhang et al., 2020; Zheng et al., 2017) were chosen, which adequately represented the UMAP space of the MoMac-VERSE. DEGs were selected across phenograph clusters from the 6 datasets using their original TPM. The DEGs identified in the 6 original datasets closely resembled to the DEtGs found between phenograph cluster within the transf.matrix (Figure S3E; Table S3).

Having generated a unified overview of Mo and Macs across tissues at the gene expression level, we next used the combined dataset to establish common and specific features of key subpopulations at the gene regulatory network level of the transcription factors. Using single-cell regulatory network inference and clustering (SCENIC) analysis (Aibar et al., 2017), we identified on 3 datasets the differentially expressed regulons (DERs; sets of transcription factors and genes predicted to be regulated by them) common to cells from the colon (Zhang et al., 2020), liver

**Figure 1. The MNP-VERSE reveals signatures of major MNP subsets across human tissues**

- (A) Summary of the data that were integrated and scheme of the meta-analysis. An integration was first performed at the organ level and MNPs were extracted.
- (B) Organ-restricted MNPs were then integrated, generating the human MNP-VERSE.
- (C) Meaning plots of protein expression data from in-house-indexed-SMARTseq2 data overlayed onto the MNP-VERSE.
- (D) Pipeline of the transf.matrix generation.
- (E) Heatmap of DEtGs ( $\log_{2}FC > 0.25$ ) in Phenograph clusters. Selected genes are highlighted.
- (F) Visualization of Phenograph clusters on the MNP-VERSE UMAP.
- (G and H) Heatmap showing relative expression levels of MNP signature genes across human tissues (G) and visualization of (H) their mean expression overlayed onto the MNP-VERSE UMAP space.



(Sharma et al., 2020), and lung (Kim et al., 2020) (Figure 2B; Table S5). As observed in the DEtG analysis (Figures 2A and S3A; Table S3), ISG\_Mo (#4) and IL4I1\_Mac (#6) shared a similar DER profile, while TREM2\_Mac (#3) shared DER with IL4I1\_Mac (#6) and HES1\_Mac (#2). The regulon analysis could also be used to refine subpopulation functions, as exemplified by the nuclear factor κB subunit 1 (NFKB1) and NFKB2 DER of IL1B\_Mo #15, which confirms their classification as inflammatory Mo.

To address the commonly used M1/M2 classification of Macs, we defined the number of M1- or M2-associated genes (Martinez et al., 2006) that were common to the DEtGs of each MoMac-VERSE cluster (Figures 2C, 2D, and S3F). ISG\_Mo (#4), IL4I1\_Mac (#6), and inflam\_Mono (#15) were enriched in M1 genes, while the other macrophage subsets such as HES1\_Mac (#2), Cluster\_#17, C1Qhi\_mac (#16), and TREM2\_Mac (#3) expressed mostly M2 genes. Of note, although the IL-4I1\_Mac (#6) exhibited a strong M1 program, these cells did not appear to express genes such as IL12B (Figure S3G), which has been described as a prototypic M1 cytokine of *in vitro* monocyte-derived Macs.

We next evaluated the reprogramming in pathologic settings by analyzing matched cells from individuals who are healthy and have cancer and various inflammatory diseases (Figure S3H). Globally, we found that the proportion of monocyte-related cells increased in cancer and inflammation as compared to healthy tissues (Figure S3I). We generated DEtGs comparing total Macs and total Mo obtained from matched healthy tissues and 6 different cancers (lung, colon, liver, breast, stomach, and pancreas) (Figures 2E, 2F, and S3J; Table S3). Pathway analysis revealed that Macs in tumors specifically expressed genes involved in lipid metabolism and inflammation pathways, while Mo expressed genes involved in pathways triggered by their stimulation by inflammatory cytokines (Figure 2G; Table S6). Mo and Macs commonly expressed genes involved in shared pathways related to their maturation and their interaction with T cells.

We then conducted a similar analysis comparing inflamed tissues from colitis and lupus nephritis patients with matched healthy tissues (Figures 2H and 2I). Inflammatory disease Mo expressed genes involved in pathways such as T helper 1 (Th1) and Th2 activation and their stimulation by cytokines, while macrophage-specific pathways were related to oxidative phosphorylation, proliferation, and cellular stress (Figure 2J). During inflammation, Mo and Macs also shared expression pathways involved in interleukin-6 (IL-6) and tumor necrosis factor (TNF)-mediated inflammation.

#### MoMac-VERSE as a resource to study MNP across tissues and specific pathologies

The Azimuth algorithm (Hao et al., 2021) was validated by a pre-integrated “query” dataset (Sharma et al., 2020) that was mapped onto the MoMac-VERSE (Figure 3A), which demonstrated

a high correlation of the initially integrated MoMac-VERSE and the *de novo* mapped UMAP coordinates (Figure 3B) and percentage Phenograph clusters of Sharma et al. (2020) on the MoMac-VERSE (Figure 3C).

To further validate and extend the use of the MoMac-VERSE as a resource, we used Azimuth to map 3 new “query” datasets (arthritic diseases synovial tissues, coronavirus disease 2019 [COVID-19] blood, and bronchoalveolar lavage [BAL]) on the MoMac-VERSE (Kuo et al., 2019; Liao et al., 2020; Silvin et al., 2020). This approach allowed unsupervised recapitulation of the major findings of these 3 studies in terms of macrophage heterogeneity by simply mapping their data onto the MoMac-VERSE. Of note, Mo and Macs were initially identified in these datasets using our defined signatures (Figure 1G). This is necessary as cells that are neither Macs nor Mo will be forced to be mapped onto the MoMac-VERSE (data not shown) and can consequently be misinterpreted. Heparin binding EGF-like growth factor-positive (HBEGF<sup>+</sup>) (Kuo et al., 2019, cluster 1), MER proto-oncogene, tyrosine kinase-positive (MERTK<sup>+</sup>) (Kuo et al., 2019, cluster 2), and IFN/signal transducer and activator of transcription (STAT) (Kuo et al., 2019, cluster 4) populations identified in the study of Kuo et al. (2019) in osteoarthritic and rheumatoid arthritis patients corresponded to HES1\_Mac (#2), TREM2\_Mac (#3), and IL4I1\_Mac (#6) of the MoMac-VERSE, respectively (Figures 3D–3F). In addition, we also identified cells that mapped to the FTL\_Mac (#17) population, which were not pre-defined in Kuo et al. (2019), thus highlighting how further heterogeneity can be uncovered with the MoMac-VERSE.

Previous work by Silvin et al. (2020) has shown that monocyte populations in the blood of severe COVID-19 patients had decreased human leukocyte antigen-DR beta 1 (HLA-DRB1) gene expression, and increased NFKB inhibitor alpha (NFKBIA) and C5AR1 gene expression. When these data were projected onto the MoMac-VERSE, the majority of cells fell within the CD16<sup>+</sup>\_Mono (#1, #5) and CD16<sup>-</sup>\_Mono (#8, #15) clusters of the MoMac-VERSE, with expression of HLA-DRB1, NFKBIA, and C5AR1 following the observations of Silvin et al. (2020) (Figures 3G–3I). We were also able to identify Macs from the BAL of COVID-19 patients in the MoMac-VERSE by mapping them with Azimuth. Cells from the COVID-19 patients fell within the clusters of the TREM2\_Mac (#3) ISG\_Mono (#4) and IL4I1\_Mac (#6). In mild COVID-19 samples, mapped macrophages fell within the TREM2\_Mac (#3) cluster. In contrast, the BAL Macs from severe COVID-19 patients mapped mostly to IL4I1\_Mac (#6), which were in association with stronger C-X-C motif chemokine ligand 10 (CXCL10) gene expression and IFN-stimulated gene (ISG) expression. These results are in agreement with the previous observations (Liao et al., 2020; Zhang et al., 2021). The link between COVID-19 BAL ISG\_Mo (#4) and IL4I1\_Mac (#6) further strengthens the connection between these 2 cell subsets (Figures 3J–3M). The results of mapping of COVID-19 patient BAL and blood with Azimuth reveals the capacity of the

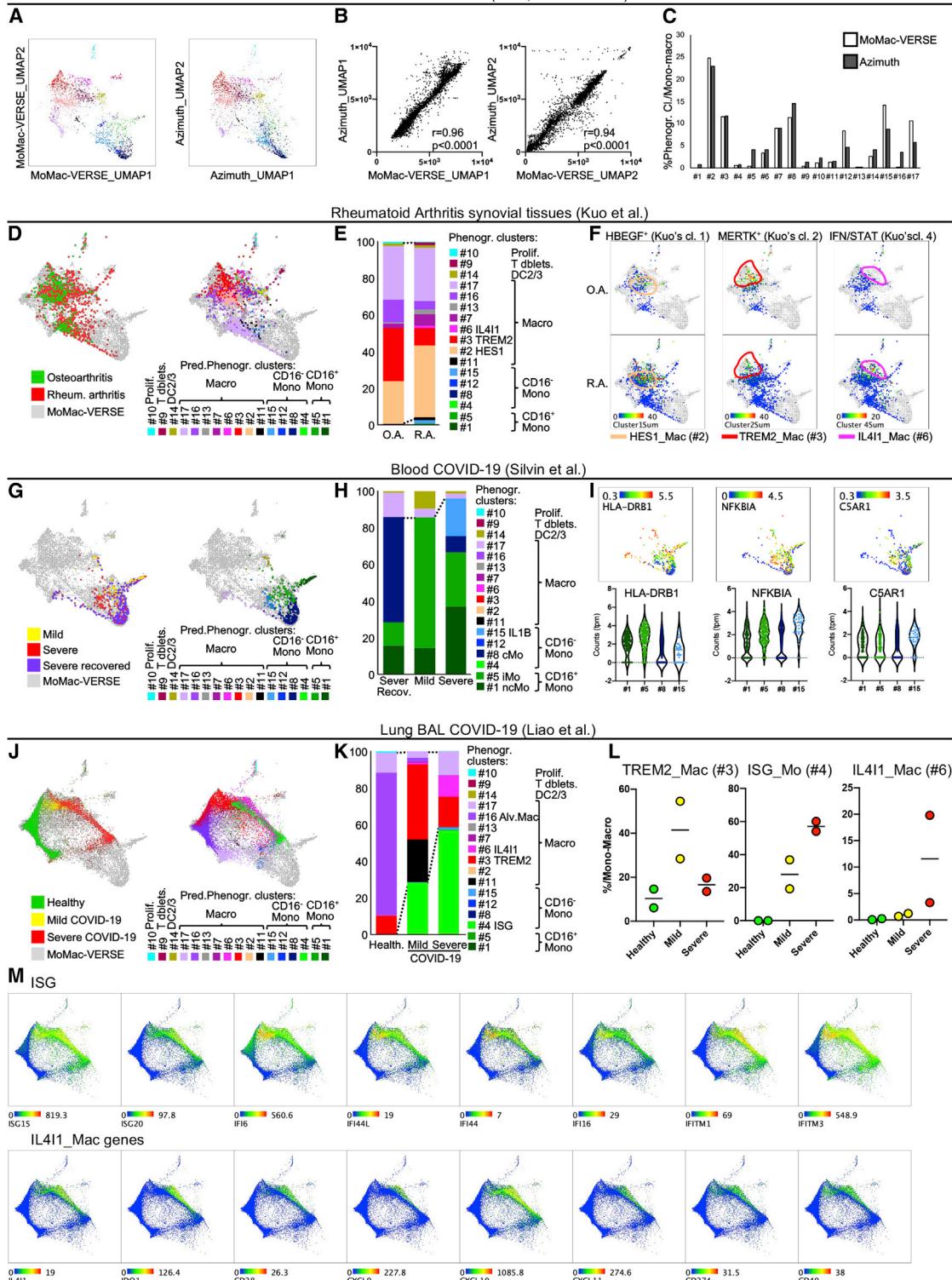
(E and F) Heatmap showing relative expression levels of DEtGs in the whole (E) macrophage or (F) monocyte populations comparing matched cancer and healthy tissues.

(G) Radar plot showing pathways obtained from DEtGs in (E) and (F).

(H and I) Heatmap showing relative expression levels of DEtGs in the whole (H) macrophage or (I) monocyte populations comparing matched inflamed and healthy tissues.

(J) Radar plot showing pathways obtained from DEtGs in (H) and (I).

Validation of Azimuth (Liver; Sharma et al.)



**Figure 3. De novo mapping of rheumatoid arthritis synovial tissue and of COVID-19 lung macrophages on the MoMac-VERSE**

(A) Validation of the Azimuth algorithm. Projection of (left panel) initially integrated and of (right panel) *de novo* “mapped query cells” (liver, Sharma et al. dataset) onto the MoMac-VERSE using (B).

(B) Correlation of the initially integrated MoMac-VERSE and the *de novo* mapped UMAP coordinates. The Pearson correlation coefficient ( $r$ ) was calculated alongside the  $p$  value.

(C) Percentage of Phenograph clusters of liver (Sharma) on the MoMac-VERSE.

(legend continued on next page)

MoMac-VERSE to recapitulate findings with increased specificity to well-conserved and well-identified macrophage/monocyte populations. Note also that blood Mo mapped only to the MoMac-VERSE Mo, while BAL Macs mapped to Macs. These results show that the MoMac-VERSE is capable of both recapitulating findings and adding increased depth to observations about cell phenotypes.

### The MoMac-VERSE identifies disease-specific monocyte and macrophage states

We next looked at cellular reprogramming events in individual pathologies (Figures 4 and S4A–S4H). Liver cirrhosis was associated with an accumulation of CD16<sup>+</sup> Mo (#1) when compared to healthy liver (Figures 4A and 4B). When deconvoluting data based on patients, including only healthy liver (excluding “normal adjacent” samples from cancer patients and including only patients for which at least 100 cells were sequenced), cells from cluster #1 also trended to increase. In the kidneys of patients with lupus nephritis, cells with a monocyte-derived DC genotype (#7) and MNP/T cell doublets (#9) were increased (Figures 4C and 4D), while classical S100A8/A9/A12<sup>hi</sup> Mo accumulated in the inflamed colons of patients presenting with colitis (Figures 4E and 4F), this latter observation also being confirmed at the patient level (Figure 4F).

We also looked into individual cancers and observed an accumulation of HES1\_Mac (#2), IL4I1\_Mac (#6), and proliferating\_Mac (#10) both in the tumor and metastatic lymph nodes (LNs) of lung cancer patients, while TREM2\_Mac (#3) also accumulated but only in the tumors (Figures 4G–4J and S4I–S4J). Colon and liver tumors also contained increased proportions of TREM2\_Mac (#3), IL4I1\_Mac (#6), and proliferating\_Mac (#10). Liver tumors exhibited increased HES1\_Mac (#2) and C1Q<sup>hi</sup>\_Mac (#16), as recently described (Sharma et al., 2020).

In summary, TREM2\_Mac (#3), IL4I1\_Mac (#6), and proliferating\_Mac (#10) were accumulated in all of the tumors from the 6 cancer types included in this study (Figures 4K, 4O, S4K, and S4L). Among these cancer types, liver tumors were unique in their accumulation of monocyte-like C1Q<sup>hi</sup>\_Mac (#16) and were the only tumors in which the inflammatory IL1B\_Mo #15 were not accumulating (Figures 4L and 4M). In certain datasets in which patient information was available, patients for whom >35 cells were sequenced in both the normal adjacent and the tumor were selected for the analysis. At the patient level, TREM2\_Mac (#3), IL4I1\_Mac (#6), and proliferating\_Mac (#10) were also significantly enriched in the tumor (Figures 4N, S4M, and S4N). These results show that the MoMac-VERSE provides a platform to identify unique and conserved cell types and states across tissues and pathologies.

### Identification of long-term resident and monocyte-derived-like macrophages across human tissues

A recent breakthrough in the field of monocyte and macrophage biology was the clarification of tissue-resident macrophage ontogeny (Blériot et al., 2020). Based on murine studies, it has been shown that in most healthy tissues, the majority of tissue-resident Macs do not derive from adult circulating Mo but directly from embryonic precursors seeded in tissues early during development and constitute a pool of “long-term resident” Macs (Ginhoux et al., 2010; Mass et al., 2016), while monocyte-derived macrophages accumulate notably during inflammation or carcinogenesis (Franklin et al., 2014; Zhu et al., 2017).

We next asked whether the MoMac-VERSE could be used to indicate the likely “long-term resident” versus the recently differentiated adult monocytic source of distinct Mac populations from various tissues, based on similarities in the gene expression profile. We overlayed cells expressing the mean conserved CD16<sup>-</sup> monocyte signature derived from Figure 1G onto the MoMac-VERSE UMAP space (Figures 5A and S5A). This revealed that monocyte-derived DC-like macrophages (#7), TREM2\_Mac (#3), and IL4I1\_Mac (#6) expressed monocyte signature genes, whereas HES1\_Mac (#2) did not. We recently described that most murine tissues were populated by 2 phenotypically distinct populations of interstitial Macs that can be discriminated based on differential LYVE1 expression (Chakarov et al., 2019). In mice, both these populations were initially derived from embryonic precursors and then replaced by Mo at a tissue-specific pace. Here, we also observed that differential LYVE1 expression separated the total human macrophage population, but interestingly, we noticed that LYVE1<sup>+</sup> cells strongly overlap with fetal liver Macs included in the MoMac-VERSE (Figure 5A). This suggested either their common embryonic origin or that a full reprogramming of Mo had occurred during the establishment of long-term tissue residency. In addition, these LYVE1<sup>+</sup> Macs fell mostly within the HES1\_Mac (#2) and the FTL gene expressing Macs (FTL\_Mac), which is in line with our recent description of the fetal-like reprogramming of HES1<sup>+</sup>FOLR2<sup>+</sup> Macs in liver tumors (Sharma et al., 2020) and the iron metabolism-related program detected in human embryonic Macs (Bian et al., 2020). Of note, similar to what has been reported in mouse studies (Franklin et al., 2014; Zhu et al., 2017), we observed that Macs accumulating in cancer tissues had a more prominent monocyte signature (Figure 5B). Notably, apart from liver tumors, where the “long-term resident”-like HES1\_Mac (#2) accumulated, all of the macrophage populations that were increased in cancer corresponded to the putative monocyte-derived Macs characterized in Figure 4, highlighting the cancer type-specific imprinting on certain MNP populations with the recruitment of monocyte-derived TAM.

(D–M) Mapping of (D–F) rheumatoid arthritis (Kuo et al., 2019) and of the blood of COVID-19 patients (Silvin et al., 2020) (G–I) and (J–M) BAL (Liao et al., 2020) new datasets onto the MoMac-VERSE.

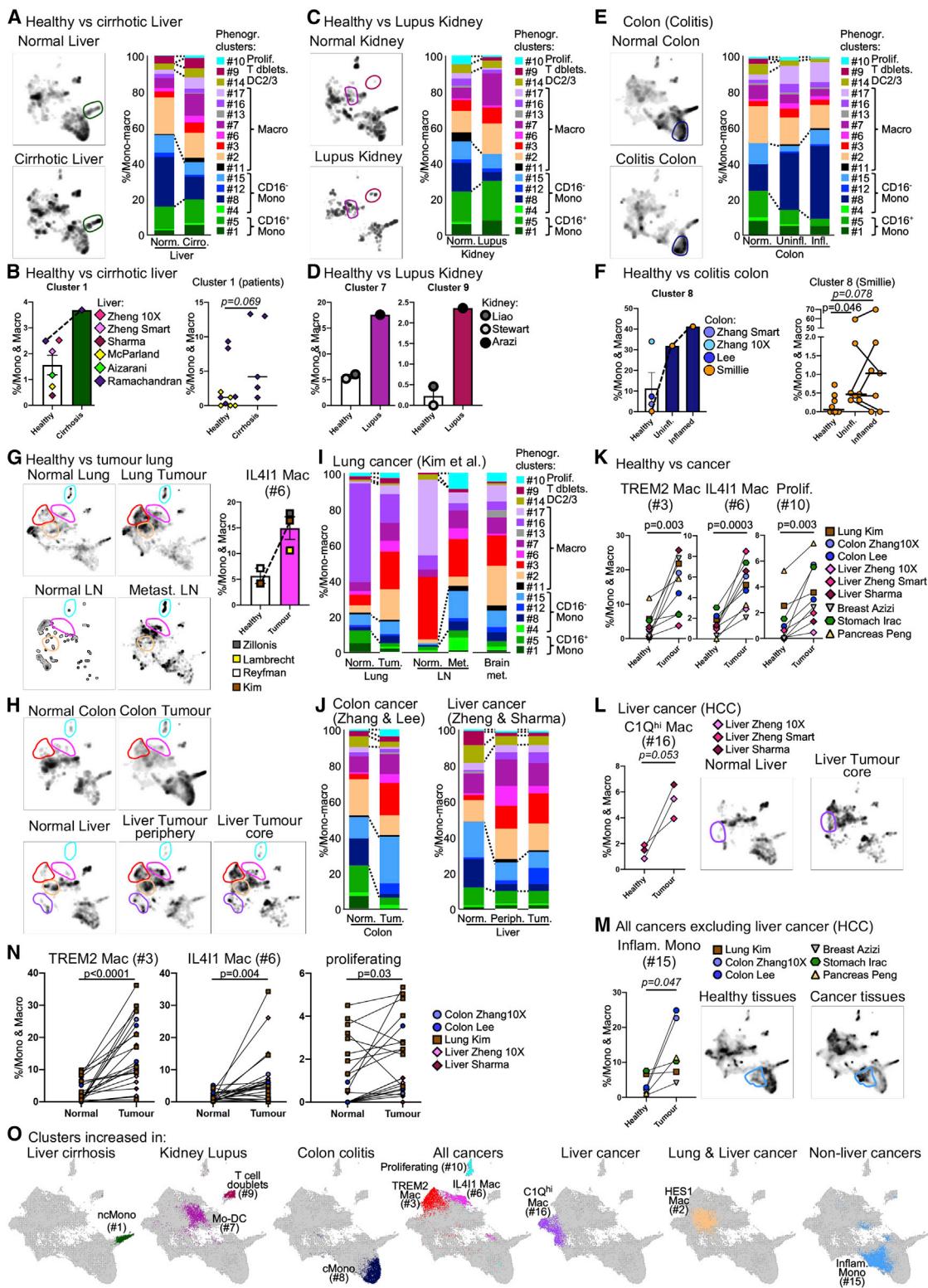
(D, E, G, H, J, and K) Mapping of cells from different patient groups (D, G, and J) and (E, H, K) their corresponding frequencies within the MoMac-VERSE Phenograph clusters.

(F) Mapping of HBEGF<sup>+</sup>, MERTK<sup>+</sup>, and IFN/STAT clusters identified in the study of Kuo et al. (2019) in osteoarthritis and rheumatoid arthritis patients.

(I) Mapping of HLA-DRB1<sup>hi</sup> cells (increased in mild COVID-19 patients) and of NFKBIA<sup>hi</sup> and C5AR1<sup>hi</sup> (increased in severe COVID-19 patients) identified in the study of Silvin et al. (2020).

(L) Frequency of the TREM2\_Mac (#3), ISG\_Mono (#4), and IL4I1\_Mac (#6) in individual healthy, mild COVID-19, and severe COVID-19 patients included in the COVID-19 BAL analysis.

(M) Visualization of the mean expression of cells for ISGs and IL4I1\_Mac (#6) genes in the study of Liao et al. (2020) and overlaid on the MoMac-VERSE.



**Figure 4. Definition of monocyte and macrophage states triggered in inflamed and cancerous tissue**

(A–M) Density plots and quantification of the different clusters in different conditions: (A and B) cirrhosis, (C and D) lupus, (E and F) colitis, and (G–M) cancer. For each condition, the quantification of each cluster is provided across datasets and (B and F) for individual patients for whom at least 100 cells per sample were

(legend continued on next page)

We were particularly interested in the IL4I1\_Mac (TAM, #6) that accumulated in all tumor types and noted that they exhibited a marked overlap in DEtGs and DER with ISG\_Mo (#4) (Figures 2B and S3A). This degree of both DEtG and DER overlap was higher than for the other Phenograph cluster comparisons, suggesting a close relationship between the IL4I1\_Mac (#6) and ISG\_Mo (#4) and a potential differentiation path between the ISG\_Mo (#4) to the IL4I1\_Mac (#6) (Figure 5C). When we dissected the expression of genes that may regulate directly the functions of ISG\_Mo (#4) and IL4I1\_Mac (#6) DEtGs (defined in Figures 2A and S3A), we identified common key transcription factors among DERs of both cell types that could be responsible for their specific activation program (Figure 5D): STAT1, STAT2, ETS variant transcription factor 7 (ETV7), IFN regulatory factor 1 (IRF1), and IRF7 could be involved in the regulation of indoleamine 2,3-dioxygenase 1 (IDO1) and IL4I1 genes. Although IDO1 has been considered a master regulator of AHR activation, a recent study demonstrated that IL4I1 was a more potent activator of AHR and thus acts as a metabolic immune checkpoint that can promote tumor progression through the catabolism of tryptophan (Sadik et al., 2020). Of note, we noticed that IL4I1\_Mac (#6) had a higher expression of IDO1 and IL4I1 in tumors than in matched healthy adjacent tissues (Figures S5D and S5E). We also observed that many shared DEtGs between ISG\_Mo (#4) and IL4I1\_Mac (#6) or IL4I1\_Mac (#6)-specific DEtGs have also recently been described as highly expressed by mature DCs enriched in immunoregulatory molecules (mregDC), suggesting that similar events and interactions could drive both IL4I1\_Mac (#6) and mregDC programming in tumors (Maier et al., 2020) (Figure 5E).

We recently developed a new fate mapping mouse model using the expression of the *Ms4a3* gene to label monocyte-derived cells (Liu et al., 2019). To further validate the monocytic origin of the IL4I1\_Mac (#6), we integrated in-house-generated liver data from *Ms4a3cre-RosatdTomato* mice with publicly available mouse liver scRNA-seq data (Remmerie et al., 2020; Seidman et al., 2020) (Figures 5F–5H and S5F). We confirmed that mouse *Trem2<sup>+</sup>Spp1<sup>+</sup>* Macs were monocyte derived, while *Hes1<sup>+</sup>Folr2<sup>+</sup>* Macs were mostly of non-monocytic origin (Figure 5G) (Sharma et al., 2020). Strikingly, a minor subset of cells that shared genes with the human IL4I1\_Mac (#6) (see Figure 5E) could also be detected in the mouse liver (Figure 5H). Mouse *Il4i1<sup>+</sup>*-like Macs were all dTomato<sup>+</sup> and thus derived from Mo. On the whole, human IL4I1\_Mac (#6) expressed monocyte genes, shared DEtGs and DER with ISG\_Mono (#4), and shared similarities with a monocyte-derived murine orthologous population, which argues that the human IL4I1\_Mac (#6) population is of monocytic origin.

Finally, we aimed to extend the ontogenical characterization of the various macrophage populations included in the MoMac-VERSE. We evaluated the commonalities between signatures of macrophage subsets from different studies and the top 5 DEtGs of the MoMac-VERSE Phenograph clusters that were either UEtGs or DEtGs with at least a 2-fold increase in log fold change (logFC) expression as compared to DEtGs from other

clusters (Figure S5G). We grouped the MoMac-VERSE clusters into 5 main categories: Mo (both classical CD16<sup>−</sup> Mo and CD16<sup>+</sup> Mo), monocyte-derived Macs (moMac), macrophages (Mac), proliferating macrophages (Prolif), and cells with low viability (Dead). We then compared these 5 main groups with the published cluster annotations: monocyte (blue) or macrophage (green) and inflammatory (pink), anti-inflammatory (light green), or cycling (light blue) (Figure 5I). We observed that most populations reported as “inflammatory Macs” had commonalities with monocyte clusters of the MoMac-VERSE, arguing for their monocytic origin. In addition, the MoMac-VERSE wide integration allows one to clarify the annotation of MNP subpopulations. For example, while some studies have used the term “macrophages” to annotate cells belonging to the S100A8<sup>+</sup>S100A9<sup>+</sup>S100A12<sup>+</sup> classical Mo (cluster #8), others have called them Mo. However, within the MoMac-VERSE, all of these differently annotated but highly similar populations belong to a common tissue monocyte subpopulation (Figure 5I).

### Exploring the MoMac-VERSE allows in-depth characterization of IL4I1 tumor-associated macrophages

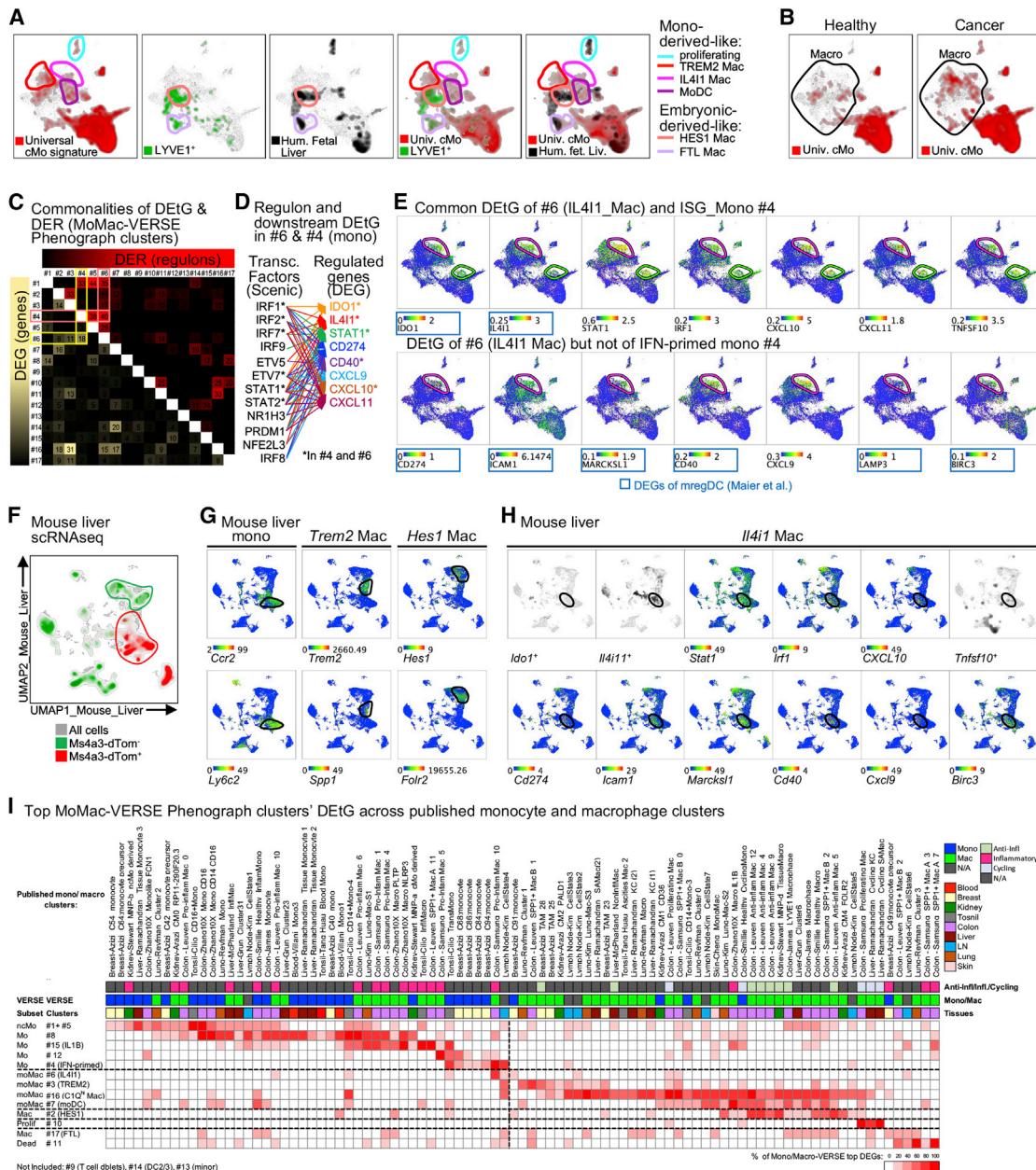
We next aimed to clarify the functions of the different TAM subpopulations. Using DEtGs (Figures 2A and S3A), we examined the pathways of the TAM populations that were increased in all included cancer studies (Figure 6A; Table S6) and the ISG\_Mo (#4), as we hypothesized that they could give rise to IL4I1\_Mac (#6). This analysis confirmed the inflammatory potential of IL1B\_Mac (#15) and suggested that TREM2\_Mac (#3) was involved in lipid metabolism (Jaitin et al., 2019), and that ISG\_Mo (#4) was IFN primed. IL4I1\_Mac (#6) was associated with antigen presentation, interaction with both Th2 and Th1 T cells, T cell exhaustion, and, importantly, tryptophan degradation, again linking this population to tumor progression through the catabolism of tryptophan (Sadik et al., 2020). IL4I1\_Mac (#6) was the only Mac that also expressed genes involved in phagosome maturation. Therefore, similar to the mregDC, with which they share a strong common specific gene expression profile (Figure 5E), the IL4I1\_Mac (#6) could have their gene expression program modulated following phagocytosis (Maier et al., 2020). We next looked at the predicted upstream regulators of these cell subsets. We confirmed that IFN-I was the top predicted upstream regulators of ISG\_Mo (#4) and also observed that IFNG was among the top predicted upstream regulators of both ISG\_Mo (#4) and IL4I1\_Mac (#6) (Figure 6B).

Next, we focused on 3 studies of the MoMac-VERSE containing cells isolated from 3 distinct anatomical sectors: the normal adjacent tissue, the tumor periphery, and the tumor core. Contrary to the other subsets, ISG\_Mo (#4) and IL4I1\_Mac (#6) were detected in greater proportions in the tumor periphery (Figure 6C). This observation supported the hypothesis of a recruitment of Mo that mature in TAM within the tissue microenvironment (TME). Given the strong IFNG imprinting in IL4I1\_Mac

sequenced. (K–M) Graphical representation of the relative proportion of clusters increased in (K) all cancer studies, (L) only in liver cancer, or (M) in all cancers, excluding liver cancer.

(N) Quantification of TREM2\_Mac (#3), IL4I1\_Mac (#6), and proliferating cells (#10) in healthy tissues and corresponding tumors in each patient of liver, lung, and colon datasets. The included patients had at least 35 cells in healthy or in tumor tissues.

(O) Summary of clusters increased in the different diseased tissues. p values were calculated and reported using a paired t test.



**Figure 5. Characterization of human long-term resident and of monocyte-derived macrophages**

(A) Overlay of the conserved cMo mean signature from Figures 1G and 1H, of lymphatic vessel endothelial hyaluronan receptor (LYVE)-expressing cells and of human fetal liver cells on the MoMac-VERSE UMAP.

(B) Overlay of the conserved cMo mean signature in healthy and cancerous tissues.

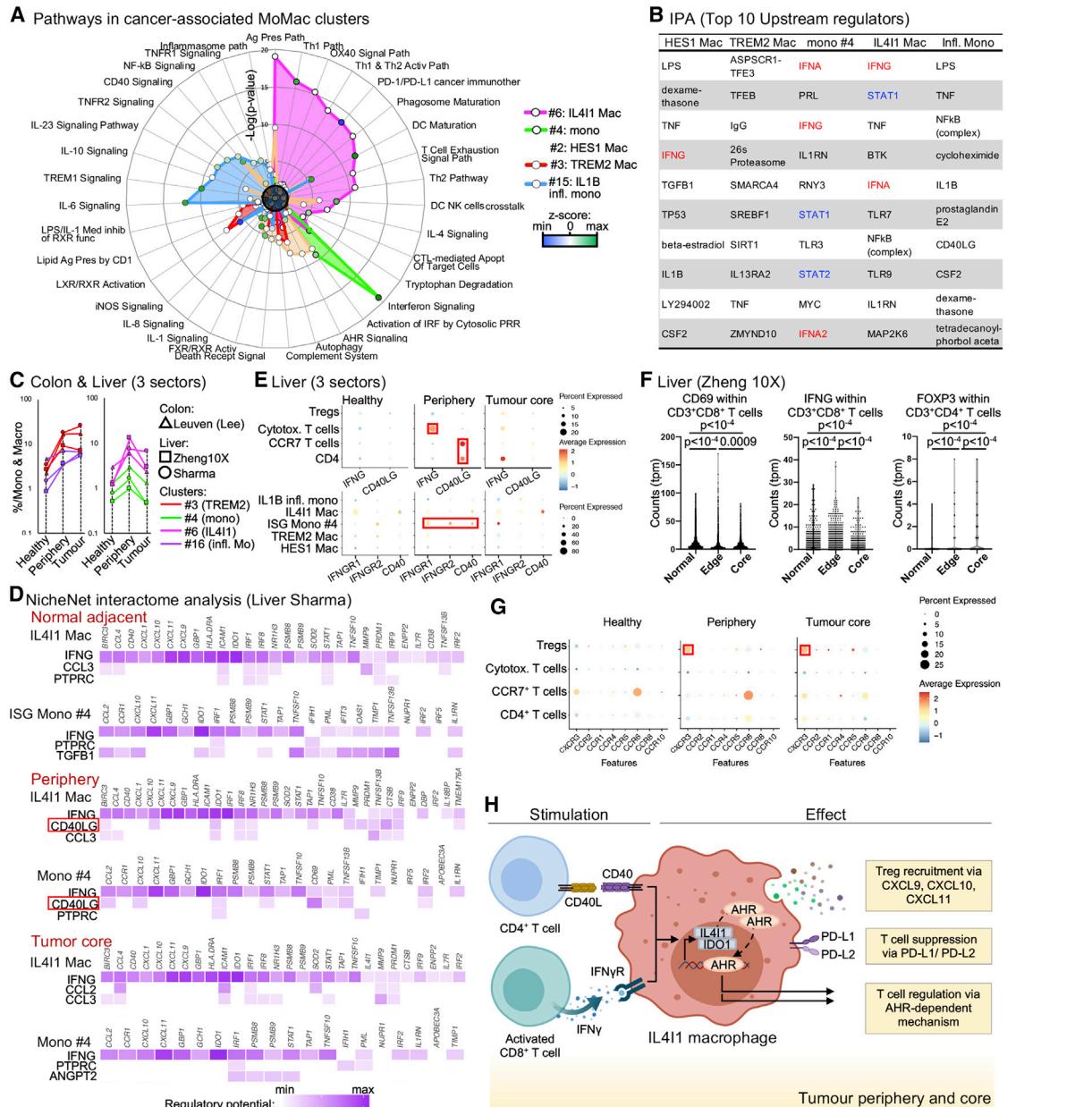
(C) Heatmap showing relative expression levels of the number of common DETGs and DER between Phenograph clusters of the MoMac-VERSE.

(D) Visualization of DEtGs predicted as regulated by transcription factors of IL4I1\_Mac (#6) and ISG\_Mo (#4) DER.

(E) Overlay of the expression on indicated genes on the MoMac-VERSE. Representative common DETGs between ISG\_Mo (#4) and IL4I1\_Mac (#6; up) and DETGs specific for IL4I1\_Mac (#6; down) are displayed. DETGs identified as specific for mregDC ([Maier et al., 2020](#)) are indicated by the blue frames.  
(F) UMAP obtained following integration of in-house-generated liver data originating from Ms4a3cre-RosatdTmato mice and publicly available data ([Remmerie et al., 2020](#); [Seidman et al., 2020](#)). Briefly, 2 libraries have been generated separately on tdTomato<sup>-</sup> and tdTomato<sup>+</sup> cells and integrated with available data coming from the literature. Ms4a3cre-Ms4a3cre<sup>+</sup> cells are highlighted in red.

(C and D) Expression of genes to define (C) M<sub>1</sub>, Tum2, and V<sub>1</sub> cells. Mean and (D) expression of genes to define mouse *H41*, *M1*, *Tum2*, and *V1* cells.

(G and H) Expression of genes to define (G) Mo, *Trem2*, and *Hes1* mouse Macs, and (H) expression of genes to define mouse *Il4i1* Macs.  
 (I) Re-assignment of published monocyte and macrophage clusters according to the MoMac-VERSE. The clusters as designed in the original studies are compared to the top 3 to 5 DETGs (selected either as UEtGs or DETGs with an at least 2-fold increase in logFC expression as compared to DETGs from other clusters) of the MoMac-VERSE Phenograph clusters.



**Figure 6. Potential mechanisms leading to pathogenic macrophage states in human tumors**

(A) Pathway analysis for individual TAM clusters. P-values falling within the dark circle are not significant ( $p>0.05$ ).

(B) Top 10 upstream regulators for each cluster determined by Ingenuity Pathway Analysis (IPA).

(C) Repartition of cells from the indicated clusters in cancer subcellular localizations for 1 colon cancer and the 2 liver cancer studies included.

(D) Visualization of the regulatory potential of the top 3 predicted ligands associated with DEGs and DER expression in IL411\_Mac (#6) and ISG\_Mo (#4) across liver sections in the Sharma et al. dataset as determined by NicheNet analysis.

(E) Visualization of the expression of IFNG and CD40LG by T cell subsets and of IFNγR1, IFNγR2, and CD40 by macrophage subsets in the different liver cancer sections.

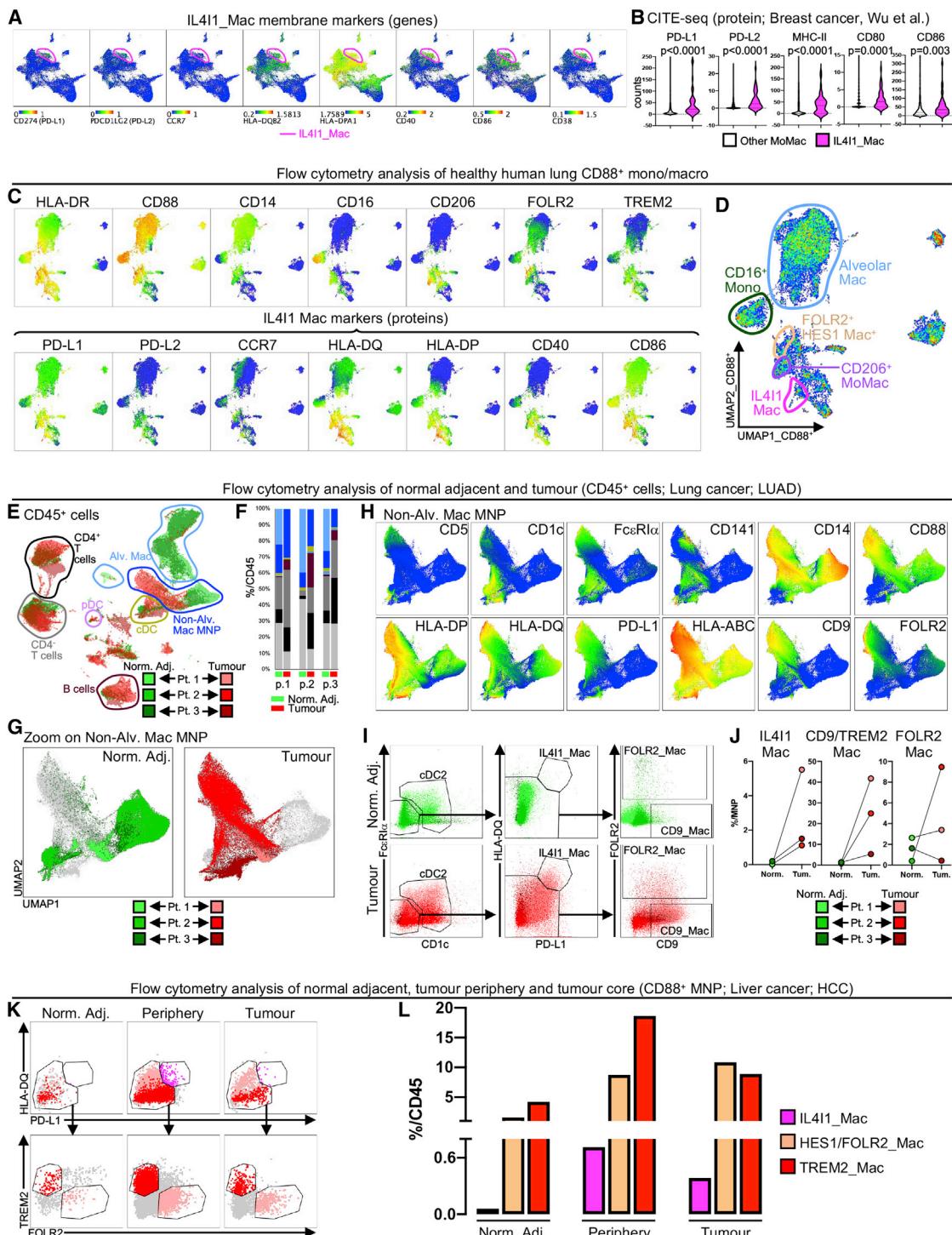
(F) Violin plots of CD69 and IFNG gene counts within CD8<sup>+</sup> T cells and of FOXP3 within CD4<sup>+</sup> T cells in real TPM of the indicated dataset.

(G) Visualization of the expression of transcripts coding for selected chemokine receptors by T cell subsets.

(H) Potential mechanism leading to IL411 macrophage programming and function. p values were calculated with the one-way ANOVA test in (F).

(#6) and that IL411\_Mac (#6) pathways revolved around antigen presentation and interaction with T cells, we hypothesized that ISG\_Mo (#4) and IL411\_Mac (#6) interact with IFN- $\gamma$ -secreting T cells present in the tumor periphery.

We carried out a NicheNet analysis to predict putative T cell interactions with monocyte/macrophage subsets in cancer (Figures 6D and S6A; Table S7). We chose the ISG\_Mo (#4) and the MoMac subsets that were increased in all of the included



**Figure 7.** Flow cytometry validation of the MoMac-VERSE macrophage subsets

(A) Expression of IL4I1 Mac (#6) marker genes overlaid onto the MoMac-VERSE.

(B) Expression of IL411 Mac (#6) marker genes revealed by CITE-seq data from S.Z.W. (unpublished data).

(C) Flow cytometry analysis of healthy human lung using antibody binding proteins of the II 411 Mac (#6) markers shown in (A)

(D) Relative expression of surface markers used to define the distinct populations

(F-I) Flow cytometry analysis of normal adjacent and tumor obtained from 3 LIAD patients

(E and F) Mapping of normal adjacent (Norm. Adj.) and tumor CD45<sup>+</sup> cells from the 3 included patients and (F) frequency among CD45<sup>+</sup> cells of the populations defined where histogram colors correspond to populations defined in (E)

(legend continued on next page)

cancer studies (#2, #3, #6, and #15) as “target” populations, and asked which T cell-related molecules and T cell subsets would be predicted to act as “stimulators.” We confirmed that IFNG was the top predicted stimulator of ISG\_Mo (#4) and IL4I1\_Mac (#6), but not of the other cell subsets. Importantly, CD40LG was also one of the top predicted stimulators of ISG\_Mo (#4) and IL4I1\_Mac (#6), but only in the tumor periphery, where these subsets were detected most abundantly (Figure 6C). Furthermore, IFNG-expressing cytotoxic CD8<sup>+</sup> T cells and CD40LG-expressing CD4<sup>+</sup> T cells were also most abundant within the tumor periphery, where higher expression of IFN $\gamma$ 1/2 and CD40 in ISG\_Mo (#4) and IL4I1\_Mac (#6) was evident (Figures 6E, 6F, S6B, and S6C). This is in agreement with a previous study demonstrating that activated CD69<sup>+</sup>CD8<sup>+</sup> T cells can induce IDO1 expression in monocyte-derived Macs in an IFN- $\gamma$ -dependent manner (Zhao et al., 2012). In addition, several studies have demonstrated that IDO1 expression can be induced *in vitro* in monocyte-derived Macs through their stimulation by CD40-ligand (CD40LG) and IFN- $\gamma$  (IFNG) (Munn et al., 1999; Zhao et al., 2012). Furthermore, CXCL9, CXCL10, and CXCL11 were among the top DETGs of IL4I1\_Mac (#6) (Figures 2A and S3A; Table S3). CXCR3, the receptor for these 3 chemokines, was strongly expressed by Tregs within the tumor periphery and core, the 2 tissue regions where Tregs were also the most abundant (Figures 6F and 6G). Therefore, our observations suggest a potential interaction of CD40L-expressing CD4<sup>+</sup> T cells and IFN- $\gamma$ -producing CD69<sup>+</sup>-activated CD8<sup>+</sup> T cells with ISG\_Mo #4, contributing to their reprogramming into immunosuppressive IDO1/IL4I1\_Mac. These Macs, through their specific chemokine production, their expression of programmed cell death ligand 1 (PD-L1) and PD-L2, and their IL4I1/AHR tryptophan degradation function may in turn suppress T cells and attract Tregs into the tumor (Figure 6H).

#### Validation of macrophage heterogeneity in human tissues

Following the establishment of the landscape of human Macs, we applied multiparametric flow cytometry to validate our findings focusing on the IL4I1\_Mac (#6). Our DETG analysis (Figures 2A and S3A; Table S3) indicated that CD274 (PD-L1), PDCD1LG2 (PD-L2), CCR7, isoforms of HLA-DP/DQ, CD40, CD86, and CD38 were highly expressed by IL4I1\_Mac (#6) (Figure 7A). The MoMac-VERSE also included a scRNA-seq/cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) breast cancer dataset (S.Z.W., unpublished data), which revealed that at the protein level, IL4I1\_Mac (#6) significantly expressed more PD-L1, PD-L2, MHC class II, CD80, and CD86 proteins as compared to the other Mo and Macs analyzed in this study, thus validating their unique genotype and associated phenotype (Figure 7B). Therefore, we combined antibodies

against the corresponding proteins encoded by these genes together with triggering receptor expressed on myeloid cells 2 (TREM2), folate receptor beta (FOLR2), CD206 (MRC1, which we recently described as a defining marker of HES1/FOLR2 Macs in liver cancer; Sharma et al., 2020), and monocyte/macrophage lineage markers, and analyzed first by flow cytometry healthy human lung a tissue in which IL4I1\_Mac (#6) could be detected even at steady state (Figures S7A and S7B; Table S2). Data were analyzed within a first UMAP space (Figures S7A and S7B), from which MNPs were extracted and re-analyzed (Figures 7C and 7D). We detected 3 major clusters of HLA-DR<sup>+</sup> MNP, a major population of Alv. Mac, a population of CD16<sup>+</sup> Mo, and a population of monocyte/macrophages. While we could not clearly identify TREM2\_Mac (#3) using an anti-TREM2 antibody in this healthy human lung sample, within monocyte/macrophages we detected FOLR2<sup>+</sup>CD206<sup>+</sup> cells corresponding to the HES1\_Mac (#2), FOLR2<sup>−</sup>CD206<sup>+</sup> Macs corresponding to Macs #16, and, notably, a minor population of PD-L1<sup>hi</sup>PL-L2<sup>hi</sup>HLA-DPhighCD40<sup>hi</sup>CD86<sup>hi</sup> cells that corresponded to the IL4I1\_Mac (#6) (Figures 7C, 7D, and S7C). We next established a classical manual gating strategy that allows the definition of these several macrophage populations in the human lung (Figures S7C–S7E).

Flow cytometry analyses were subsequently performed on matched samples obtained from the normal adjacent and tumor cells of 3 lung adenocarcinoma (LUAD) patients (Figures 7E–7J and S7F–S7I). In these samples, CD9 was used as a TREM2 surrogate marker to identify TREM2\_Mac (#3) since TREM2 could not be detected previously in healthy human lungs (Figure 7C). Data were analyzed within a first UMAP space (Figures 7E, 7F, and S7C), from which non-alveolar macrophage MNPs (Non-Alv. Mac MNP) were extracted and re-analyzed (Figures 7H–7J and S7G–S7I). While IL4I1\_Mac (#6) and TREM2\_Mac (#3) increased in tumors for all 3 patients, HES1/FOLR2\_Mac (#2) only increased in the tumors of 2 out of the 3 patients, although all 3 populations were found to be increased in the MoMac-VERSE (Figure 7J). Note also that IL4I1\_Mac had the highest expression of CD38 and HLA-DP, confirming our findings from the MoMac-VERSE at the protein level (Figure S7I).

We also performed flow cytometry analysis on cells from the normal adjacent, the tumor periphery, and the tumor core of the liver of a patient with hepatocellular carcinoma (HCC; Figures 7K, 7L, and S7J). Similar to what was observed in the MoMac-VERSE (Figure 6C), IL4I1\_Mac (#6) was retrieved mostly in the tumor periphery and was totally absent from normal adjacent liver (Figures 7K and 7L). We also confirmed our previous observation of the progressive accumulation of HES1/FOLR2\_Mac (#2) and TREM2\_Mac (#3) from normal adjacent tissue toward the liver tumor core (Figures 7K and 7L) (Sharma et al., 2020).

(G and H) For non-alveolar MNP, (G) visualization of detected in the normal adjacent, and tumor tissues of the 3 patients and (H) relative expression of surface markers defining cDC2 (CD5, CD1c, Fc $\epsilon$ R1 $\alpha$ ), cDC1 (CD141) and macrophage subsets (other markers).

(I and J) Within non-alveolar MNP, (I) gating strategy defining cDC2 and macrophage and (J) quantification of IL4I1, CD9/TREM2, and FOLR2 Macs in healthy tissues and corresponding tumors in each patient.

(K) Flow cytometry analysis of Norm. Adj. and tumor periphery and tumor core obtained from a liver cancer (HCC) patient. Cells falling within the MNP and defined as HLA-DR<sup>+</sup>CD88<sup>+</sup> Macs are displayed (see Figure S7J for the upstream gating strategy). Within Macs, PD-L1<sup>hi</sup>HLA-DQ<sup>hi</sup> IL4I1\_Mac (#6) was gated. Next, among PD-L1<sup>lo/−</sup>HLA-DQ<sup>lo/−</sup> cells, gating of TREM2<sup>+</sup>FOLR2<sup>lo/−</sup> TREM2\_Mac (#3) and of TREM2<sup>lo/−</sup>FOLR2<sup>+</sup> FOLR2\_Mac.

(L) Frequency among CD45<sup>+</sup> mononuclear cells of macrophage subsets from the analysis of (K). p values were calculated using the non-parametric Mann-Whitney in Figure 7B.

## DISCUSSION

Here, we have compiled a large number of MNP scRNA-seq datasets and built MoMac- VERSEs, which allowed us to propose a unified annotation of human Mo and Macs. Consequently, we identified immunosuppressive macrophage populations present and potentially involved in the progression of tumors. To our knowledge, this study constitutes the largest meta-analysis of human Mo and Macs, and we believe it will benefit the scientific community to better apprehend MPS diversity in normal and diseased tissues. Another study used a similar strategy to resolve lung MNP populations (Peters et al., 2020); our findings here fully corroborate this approach while extending it across additional tissues and diseases. Furthermore, the possibility offered by workflows such as Azimuth, which allow mapping of “query” datasets to our reference MoMac-VERSE, should facilitate the integration of upcoming studies by providing a unified framework (Hao et al., 2021). By projecting data from published COVID-19 datasets onto the MoMac-VERSE, we were able to recapitulate findings of previous works along with providing novel insight into the biology of macrophages and Mo. The analysis of COVID-19 Mo and Macs highlighted an increased ISG signature in Mo as observed by Silvin et al. (2020). In addition, we observed an increased abundance of IL4I1\_Mac expressing CXCL9, CXCL10, CXCL11, IDO1, and IL4I1, as found by Zhang et al. (2021), in inflammatory diseases and severe COVID-19. Considering the relationship between IFN signaling and the presence of IL4I1 Macs, the MoMac-VERSE unifies the findings of these 2 publications to reveal a more robust analysis of the biology of monocytes and macrophages in severe COVID-19.

In mice, we know that Macs can be derived from adult circulating Mo or from embryonic precursors seeded in tissues during early development (Ginhoux et al., 2010; Hashimoto et al., 2013; Mass et al., 2016; Perdiguero and Geissmann, 2016; Schulz et al., 2012; Yona et al., 2013). Translation of these findings to the human is challenging, but scRNA-seq permits overcoming such challenges (Bian et al., 2020). Here, we focused our attention on 3 main subsets of Macs present across tissues, designated as TREM2, HES1, and IL4I1 Macs. The MoMac-VERSE suggests that TREM2 and IL4I1 Macs could be mostly monocyte derived, whereas HES1 Macs harbor an embryonic signature. To validate this, we used our mouse fate-mapping model to identify the equivalent murine populations and reveal their ontogeny.

These 3 macrophage populations have been separately reported where our integrated approach identified them as major conserved subsets. The M1/M2 classification initially developed for *in vitro* monocyte-derived Macs is still widely used, but as previously demonstrated, this dichotomy does not embrace macrophage diversity, especially in cancer, in which TAM differentiation from circulating Mo appears to be a distinct pathway from the assumed M2 anti-inflammatory (pro-tumoral) status (Franklin et al., 2014). For example, although IL4I1 Macs have a strong M1 program, *in vitro*-generated M1 monocyte-derived Macs may not recapitulate the primary M1-like Macs that, contrary to *in vitro* M1 Macs, do not produce IL12B, the production of which is restricted to activated cDC1 and mregDC. Therefore, the versatility of resident Macs and TAM needs to be taken into account to refine our understanding of healthy and diseased tissues. CD38+

M1-like Macs have also recently been described in the context of aging and aging-related disease (Covarrubias et al., 2020), which may correspond to the IL4I1\_Mac. Further validation is necessary to understand the role of IL4I1\_Mac in this context.

Until recently, TREM2 Macs were studied mostly in the brain due to their role in the development of neurodegenerative disorders (Krasemann et al., 2017; Parhizkar et al., 2019). However, TREM2 Macs involved in metabolic disorders and obesity have now been detected in adipose tissue (Jaitin et al., 2019) and in tumors, where they occupy a potentially immunosuppressive role (Katzenellenbogen et al., 2020; Molgora et al., 2020). In parallel, a subset of immunosuppressive IL4I1<sup>+</sup> TAM has previously been detected (Sadik et al., 2020; Zhao et al., 2012). Herein, the combination of data from multiple studies has fully validated these discoveries, where we have detected these populations in all 6 of the cancers analyzed. From a mechanistic point of view, even if it is clear that these TAMs are involved in dampening antitumoral T cell responses, the precise pathways involved are still unclear and deserve further investigation to fully decipher the T cell/TAM relationship (Gordon et al., 2017). The third population of Macs that we focused our attention on were HES1 Macs, which correspond to the FOLR2<sup>+</sup> Macs that were detected in liver cancer (Sharma et al., 2020). Interestingly, these Macs seem to reprogram toward a fetal macrophage identity during cancer development.

Within the literature, 2 major monocyte populations (CD16<sup>+</sup> and CD16<sup>-</sup>) have been recognized (Passlick et al., 1989; Ziegler-Heitbrock et al., 2010), which we also observed here. However, we observed further heterogeneity, as in the case of ISG Mo (#4) and their particular enrichment within the tumor periphery of liver and colon cancer patients. The resemblance of these cells to IL4I1\_Mac (#6) further highlighted a potential precursor-product relationship that would have been missed in separate analyses.

Finally, we validated hypotheses raised by the MoMac-VERSE using multiparametric spectral flow cytometry, quantifying protein markers validated through RNA/protein expression from the integrated CITE-seq/scRNA-seq breast cancer dataset (S.Z.W., unpublished data). Even if we fully recognize the relevance of the *in silico* approach to integrate large and widespread datasets, one could argue that such analyses require biological validation before being used in clinical settings. Data obtained from tumor biopsies here were in full accordance with the predictions from the MoMac-VERSE and thus, we believe that such results should accelerate translational research by providing markers and potential targets for new treatments.

In conclusion, we have revealed the diverse but conserved array of human monocyte and Macs populations in both health and disease. This study also provides an online tool to explore human MNPs, where this diversity should now be taken into account for further studies. The new era of single-cell transcriptomics and subsequent integration will now lead us to go beyond the establishment of atlases and descriptive catalogs of cell subsets, and will pave the way to help the scientific community better design innovative and more specific macrophage-based immunotherapy strategies.

### Limitations of the study

The transformed matrix is made in such a way that only genes common to all of the included datasets were taken forward. Consequently, some genes were lost in this process but those included

are conserved across studies and thus have a greater chance to be detected in any other study not included here. As a consequence, most genes included in MNP subsets and individual monocyte and macrophage subset gene expression signatures should also be detected in any other study. Since we only included scRNA-seq datasets that were available at the time when these VERSES were generated, the MNP- and MoMac-VERSES could continuously be improved by mapping more recent and upcoming datasets through reintegration or using Azimuth. We were also limited by the number of datasets that had separately sequenced normal adjacent tissue, periphery, and tumor core. Consequently, our findings on the accumulation of IL4I1\_Mac and their IFN-primed Mo (cluster #4) within the tumor periphery were limited to the liver and colon. Therefore, further validation is required for other tumors. Nevertheless, the interactive MNP- and MoMac-VERSES can be used to explore and align new findings to the data published here.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Human samples
- **METHOD DETAILS**
  - Flow cytometry and cell sorting
  - Algorithms for dimensionality reduction
  - MNP extraction and Seurat V3 integration
  - Generation of indexed-sorting and SMARTseq2 single-cell transcriptome data
  - Pre-processing, quality assessment and control and analysis of SMARTseq2 single-cell transcriptome data
  - Generation of transformed matrix
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Differentially expressed genes (DEGs) and “transformed” genes (DEtGs) analyses
  - Metadata analysis
  - M1 and M2 signature
  - Generation of the major MNP subset DEGs
  - Scenic gene regulatory network analyses
  - Pathway analysis
  - NicheNet analysis

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.immuni.2021.07.007>.

## ACKNOWLEDGMENTS

We thank L. Robinson of Insight Editing London for critical review and editing of the manuscript; F.G. is an EMBO YIP awardee and is supported by Singapore Immunology Network (SlgN) core funding as well as a Singapore National Research Foundation Senior Investigatorship (NRFI) NRF2016NRF-NRFI001-02. This work was supported by the ARC foundation France (ANT202001330). The bioinformatics and immunogenomics platforms are

part of the SlgN Immunomonitoring platform (supported by a BMRC IAF 311006 grant and BMRC transition funds #H16/99/b0/011).

## AUTHOR CONTRIBUTIONS

Experiments, K.M., W.T.K., A.A.P., S.K., S.E.I., and C.-A.D.; data analysis, K.M., W.T.K., C.P., A.A.P., E.H., G.D., M.C., X.M.Z., R.M.M.W., R.P., A.I.S.K., A.S., C.B., and C.-A.D.; providing human samples, J.K.C.T.; providing healthy lung samples, T.K.H.L., P.K.H.C., and A.B.; providing the liver samples, A.D.; providing the matched normal adjacent and lung tumor (LUAD) samples, L.D. and L.Z.; providing the breast cancer CITE-seq/scRNA-seq data, G.A.-E. and A.S.; writing of the manuscript, C.B., C.-A.D., and F.G.; intellectual input, A.S. and A.B.; project supervision, C.-A.D. and F.G.; study conceptualization, C.A.D. and F.G.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 6, 2020

Revised: April 5, 2021

Accepted: July 12, 2021

Published: July 30, 2021

## REFERENCES

- Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* 14, 1083–1086.
- Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kisielovas, V., Setty, M., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* 174, 1293–1308.e36.
- Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37, 38–44.
- Bian, Z., Gong, Y., Huang, T., Lee, C.Z.W., Bian, L., Bai, Z., Shi, H., Zeng, Y., Liu, C., He, J., et al. (2020). Deciphering human macrophage development at single-cell resolution. *Nature* 582, 571–576.
- Bigley, V., Cytlak, U., and Collin, M. (2019). Human dendritic cell immunodeficiencies. *Semin. Cell Dev. Biol.* 86, 50–61.
- Blériot, C., Chakarov, S., and Ginhoux, F. (2020). Determinants of Resident Tissue Macrophage Identity and Function. *Immunity* 52, 957–970.
- Bourdely, P., Anselmi, G., Vaivode, K., Ramos, R.N., Missolo-Koussou, Y., Hidalgo, S., Tosselo, J., Nuñez, N., Richer, W., Vincent-Salomon, A., et al. (2020). Transcriptional and Functional Analysis of CD1c<sup>+</sup> Human Dendritic Cells Identifies a CD163<sup>+</sup> Subset Priming CD8<sup>+</sup>CD103<sup>+</sup> T Cells. *Immunity* 53, 335–352.e8.
- Browaeys, R., Saelens, W., and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* 17, 159–162.
- Chakarov, S., Lim, H.Y., Tan, L., Lim, S.Y., See, P., Lum, J., Zhang, X.-M., Foo, S., Nakamizo, S., Duan, K., et al. (2019). Two distinct interstitial macrophage populations coexist across tissues in specific subtissular niches. *Science* 363, eaau0964.
- Cheng, J.B., Sedgewick, A.J., Finnegan, A.I., Harirchian, P., Lee, J., Kwon, S., Fassett, M.S., Golovato, J., Gray, M., Ghadially, R., et al. (2018). Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution. *Cell Rep.* 25, 871–883.
- Cillo, A.R., Kürten, C.H.L., Tabib, T., Qi, Z., Onkar, S., Wang, T., Liu, A., Duvvuri, U., Kim, S., Soose, R.J., et al. (2020). Immune Landscape of Viral- and Carcinogen-Driven Head and Neck Cancer. *Immunity* 52, 183–199.e9.
- Covarrubias, A.J., Kale, A., Perrone, R., Lopez-Dominguez, J.A., Pisco, A.O., Kasler, H.G., Schmidt, M.S., Heckenbach, I., Kwok, R., Wiley, C.D., et al. (2020). Senescent cells promote tissue NAD<sup>+</sup> decline during ageing via the activation of CD38<sup>+</sup> macrophages. *Nat. Metab.* 2, 1265–1283.

- Dress, R.J., Dutertre, C.-A., Giladi, A., Schlitzer, A., Low, I., Shadan, N.B., Tay, A., Lum, J., Kairi, M.F.B.M., Hwang, Y.Y., et al. (2019). Plasmacytoid dendritic cells develop from Ly6D<sup>+</sup> lymphoid progenitors distinct from the myeloid lineage. *Nat. Immunol.* 20, 852–864.
- Dutertre, C.-A., Becht, E., Irac, S.E., Khalilnezhad, A., Narang, V., Khalilnezhad, S., Ng, P.Y., van den Hoogen, L.L., Leong, J.Y., Lee, B., et al. (2019). Single-Cell Analysis of Human Mononuclear Phagocytes Reveals Subset-Defining Markers and Identifies Circulating Inflammatory Dendritic Cells. *Immunity* 51, 573–589.e8.
- Franklin, R.A., Liao, W., Sarkar, A., Kim, M.V., Bivona, M.R., Liu, K., Pamer, E.G., and Li, M.O. (2014). The cellular and molecular origin of tumor-associated macrophages. *Science* 344, 921–925.
- van Furth, R., Cohn, Z.A., Hirsch, J.G., Humphrey, J.H., Spector, W.G., and Langevoort, H.L. (1972). The mononuclear phagocyte system: a new classification of macrophages, monocytes, and their precursor cells. *Bull. World Health Organ.* 46, 845–852.
- Geissmann, F., Jung, S., and Littman, D.R. (2003). Blood monocytes consist of two principal subsets with distinct migratory properties. *Immunity* 19, 71–82.
- Giladi, A., and Amit, I. (2018). Single-Cell Genomics: A Stepping Stone for Future Immunology Discoveries. *Cell* 172, 14–21.
- Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R., et al. (2010). Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* 330, 841–845.
- Gordon, S.R., Maute, R.L., Dulken, B.W., Hutter, G., George, B.M., McCracken, M.N., Gupta, R., Tsai, J.M., Sinha, R., Corey, D., et al. (2017). PD-1 expression by tumour-associated macrophages inhibits phagocytosis and tumour immunity. *Nature* 545, 495–499.
- Guilliams, M., Ginhoux, F., Jakubzick, C., Naik, S.H., Onai, N., Schraml, B.U., Segura, E., Tussiwand, R., and Yona, S. (2014). Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. *Nat. Rev. Immunol.* 14, 571–578.
- Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zagar, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29.
- Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, M.B., Leboeuf, M., Becker, C.D., See, P., Price, J., Lucas, D., et al. (2013). Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity* 38, 792–804.
- Jaitin, D.A., Adlung, L., Thaiss, C.A., Weiner, A., Li, B., Descamps, H., Lundgren, P., Bleriot, C., Liu, Z., Deczkowska, A., et al. (2019). Lipid-Associated Macrophages Control Metabolic Homeostasis in a Trem2-Dependent Manner. *Cell* 178, 686–698.e14.
- James, K.R., Gomes, T., Elmentaita, R., Kumar, N., Gulliver, E.L., King, H.W., Stares, M.D., Bareham, B.R., Ferdinand, J.R., Petrova, V.N., et al. (2020). Distinct microbial and immune niches of the human colon. *Nat. Immunol.* 21, 343–353.
- Katzenelenbogen, Y., Sheban, F., Yalin, A., Yofe, I., Svetlichny, D., Jaitin, D.A., Bornstein, C., Moshe, A., Keren-Shaul, H., Cohen, M., et al. (2020). Coupled scRNA-Seq and Intracellular Protein Activity Reveal an Immunosuppressive Role of TREM2 in Cancer. *Cell* 182, 872–885.e19.
- Kim, N., Kim, H.K., Lee, K., Hong, Y., Cho, J.H., Choi, J.W., Lee, J.-I., Suh, Y.-L., Ku, B.M., Eum, H.H., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* 11, 2285.
- Krasemann, S., Madore, C., Cialic, R., Baufeld, C., Calcagno, N., El Fatimy, R., Beckers, L., O'Loughlin, E., Xu, Y., Fanek, Z., et al. (2017). The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity* 47, 566–581.e9.
- Kuo, D., Ding, J., Cohn, I.S., Zhang, F., Wei, K., Rao, D.A., Rozo, C., Sokhi, U.K., Shanaj, S., Oliver, D.J., et al. (2019). HBEGF<sup>+</sup> macrophages in rheumatoid arthritis induce fibroblast invasiveness. *Sci. Transl. Med.* 11, eaau8587.
- Lee, H.-O., Hong, Y., Etilioglu, H.E., Cho, Y.B., Pomella, V., Van den Bosch, B., Vanhecke, J., Verbandt, S., Hong, H., Min, J.-W., et al. (2020). Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat. Genet.* 52, 594–603.
- Levine, J.H., Simonds, E.F., Bendall, S.C., Davis, K.L., Amir, A.D., Tadmor, M.D., Litvin, O., Fienberg, H.G., Jager, A., Zunder, E.R., et al. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* 162, 184–197.
- Liao, M., Liu, Y., Yuan, J., Wen, Y., Xu, G., Zhao, J., Cheng, L., Li, J., Wang, X., Wang, F., et al. (2020). Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* 26, 842–844.
- Liu, Z., Gu, Y., Chakarov, S., Bleriot, C., Kwok, I., Chen, X., Shin, A., Huang, W., Dress, R.J., Dutertre, C.-A., et al. (2019). Fate Mapping via Ms4a3-Expression History Traces Monocyte-Derived Cells. *Cell* 178, 1509–1525.e19.
- MacParland, S.A., Liu, J.C., Ma, X.-Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nat. Commun.* 9, 4383.
- Maier, B., Leader, A.M., Chen, S.T., Tung, N., Chang, C., LeBerichel, J., Chudnovskiy, A., Maskey, S., Walker, L., Finnigan, J.P., et al. (2020). A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* 580, 257–262.
- Maier, T., Güell, M., and Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* 583, 3966–3973.
- Martinez, F.O., Gordon, S., Locati, M., and Mantovani, A. (2006). Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *J. Immunol.* 177, 7303–7311.
- Mass, E., Ballesteros, I., Farlik, M., Halbritter, F., Günther, P., Crozet, L., Jacome-Galarza, C.E., Händler, K., Klughammer, J., Kobayashi, Y., et al. (2016). Specification of tissue-resident macrophages during organogenesis. *Science* 353, aaf4238.
- McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv*, 1802.03426v3 <http://arxiv.org/abs/1802.03426v3>.
- Merad, M., Sathe, P., Hefti, J., Miller, J., and Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu. Rev. Immunol.* 31, 563–604.
- Mills, C.D., Kincaid, K., Alt, J.M., Heilman, M.J., and Hill, A.M. (2000). M-1/M-2 macrophages and the Th1/Th2 paradigm. *J. Immunol.* 164, 6166–6173.
- Molgora, M., Esaulova, E., Vermi, W., Hou, J., Chen, Y., Luo, J., Brioschi, S., Bugatti, M., Omodei, A.S., Ricci, B., et al. (2020). TREM2 Modulation Remodels the Tumor Myeloid Landscape Enhancing Anti-PD-1 Immunotherapy. *Cell* 182, 886–900.e17.
- Munn, D.H., Shafizadeh, E., Attwood, J.T., Bondarev, I., Pashine, A., and Mellor, A.L. (1999). Inhibition of T cell proliferation by macrophage tryptophan catabolism. *J. Exp. Med.* 189, 1363–1372.
- Parhizkar, S., Arzberger, T., Brendel, M., Kleinberger, G., Deussing, M., Focke, C., Nuscher, B., Xiong, M., Ghasemigharagoz, A., Katzmarski, N., et al. (2019). Loss of TREM2 function increases amyloid seeding but reduces plaque-associated ApoE. *Nat. Neurosci.* 22, 191–204.
- Passlick, B., Flieger, D., and Ziegler-Heitbrock, H.W. (1989). Identification and characterization of a novel monocyte subpopulation in human peripheral blood. *Blood* 74, 2527–2534.
- Perdigero, E.G., and Geissmann, F. (2016). The development and maintenance of resident macrophages. *Nat. Immunol.* 17, 2–8.
- Peters, J.M., Blainey, P., and Bryson, B.D. (2020). Consensus transcriptional states describe human mononuclear phagocyte diversity in the lung across health and disease. *bioRxiv*. <https://doi.org/10.1101/2020.08.06.240424>.
- Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9, 171–181.
- Pollard, J.W. (2004). Tumour-educated macrophages promote tumour progression and metastasis. *Nat. Rev. Cancer* 4, 71–78.

- Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al.; Human Cell Atlas Meeting Participants (2017). The Human Cell Atlas. *eLife* 6, e27041.
- Remmerie, A., Martens, L., Thoné, T., Castoldi, A., Seurinck, R., Pavie, B., Roels, J., Vanneste, B., De Prijck, S., Vanhockerhout, M., et al. (2020). Osteopontin Expression Identifies a Subset of Recruited Macrophages Distinct from Kupffer Cells in the Fatty Liver. *Immunity* 53, 641–657.e14.
- Rodrigues, P.F., Alberti-Servera, L., Eremin, A., Grajales-Reyes, G.E., Ivanek, R., and Tussiwand, R. (2018). Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells. *Nat. Immunol.* 19, 711–722.
- Rozenblatt-Rosen, O., Stubbington, M.J.T., Regev, A., and Teichmann, S.A. (2017). The Human Cell Atlas: from vision to reality. *Nature* 550, 451–453.
- Sadik, A., Somarribas Patterson, L.F., Öztürk, S., Mohapatra, S.R., Panitz, V., Secker, P.F., Pfänder, P., Loth, S., Salem, H., Prentzell, M.T., et al. (2020). IL4I1 Is a Metabolic Immune Checkpoint that Activates the AHR and Promotes Tumor Progression. *Cell* 182, 1252–1270.e34.
- Schulz, C., Gomez Perdiguero, E., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E.W., Pollard, J.W., et al. (2012). A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* 336, 86–90.
- See, P., Dutertre, C.-A., Chen, J., Günther, P., McGovern, N., Irac, S.E., Gunawan, M., Beyer, M., Händler, K., Duan, K., et al. (2017). Mapping the human DC lineage through the integration of high-dimensional techniques. *Science* 356, eaag3009.
- Seidman, J.S., Troutman, T.D., Sakai, M., Gola, A., Spann, N.J., Bennett, H., Bruni, C.M., Ouyang, Z., Li, R.Z., Sun, X., et al. (2020). Niche-Specific Reprogramming of Epigenetic Landscapes Drives Myeloid Cell Diversity in Nonalcoholic Steatohepatitis. *Immunity* 52, 1057–1074.e7.
- Sharma, A., Seow, J.J.W., Dutertre, C.-A., Pai, R., Blériot, C., Mishra, A., Wong, R.M.M., Singh, G.S.N., Sudhagar, S., Khalilnezhad, S., et al. (2020). Onco-fetal Reprogramming of Endothelial Cells Drives Immunosuppressive Macrophages in Hepatocellular Carcinoma. *Cell* 183, 377–394.e21.
- Silvin, A., Chapuis, N., Dunsmore, G., Goubet, A.-G., Dubuisson, A., Derosa, L., Almire, C., Hénon, C., Kosmider, O., Droin, N., et al. (2020). Elevated Calprotectin and Abnormal Myeloid Cell Subsets Discriminate Severe from Mild COVID-19. *Cell* 182, 1401–1418.e18.
- Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019). Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* 178, 714–730.e22.
- Steinman, R.M., and Cohn, Z.A. (1973). Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *J. Exp. Med.* 137, 1142–1162.
- Stewart, B.J., Ferdinand, J.R., Young, M.D., Mitchell, T.J., Loudon, K.W., Riding, A.M., Richoz, N., Frazer, G.L., Staniforth, J.U.L., Vieira Braga, F.A., et al. (2019). Spatiotemporal immune zonation of the human kidney. *Science* 365, 1461–1466.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888–1902.e21.
- Stubbington, M.J.T., Rozenblatt-Rosen, O., Regev, A., and Teichmann, S.A. (2017). Single-cell transcriptomics to explore the immune system in health and disease. *Science* 358, 58–63.
- Svensson, V., da Veiga Beltrame, E., and Pachter, L. (2020). A curated database reveals trends in single-cell transcriptomics. *Database* 2020, baaa073.
- Tan-Garcia, A., Wai, L.-E., Zheng, D., Ceccarello, E., Jo, J., Banu, N., Khakpoor, A., Chia, A., Tham, C.Y.L., Tan, A.T., et al. (2017). Intrahepatic CD206<sup>+</sup> macrophages contribute to inflammation in advanced viral-related liver disease. *J. Hepatol.* 67, 490–500.
- The Tabula Muris Consortium (2018). Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* 562, 367–372.
- Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* 15, 2247–2276.
- Van Hove, H., Martens, L., Scheyltjens, I., De Vlaminck, K., Pombo Antunes, A.R., De Prijck, S., Vandamme, N., De Schepper, S., Van Isterdael, G., Scott, C.L., et al. (2019). A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by ontogeny and tissue environment. *Nat. Neurosci.* 22, 1021–1035.
- Vieira Braga, F.A., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., et al. (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* 25, 1153–1163.
- Villani, A.-C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* 356, eaah4573.
- Wculek, S.K., Cueto, F.J., Mujal, A.M., Melero, I., Krummel, M.F., and Sancho, D. (2020). Dendritic cells in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* 20, 7–24.
- Xue, D., Tabib, T., Morse, C., and Lafyatis, R. (2019). Single-cell RNA sequencing reveals different subsets of macrophage and dendritic cells in human skin. *J. Immunol.* 202, 177.8.
- Yona, S., Kim, K.-W., Wolf, Y., Mildner, A., Varol, D., Breker, M., Strauss-Ayali, D., Viukov, S., Guilliams, M., Misharin, A., et al. (2013). Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* 38, 79–91.
- Zhang, L., Li, Z., Skrzypczynska, K.M., Fang, Q., Zhang, W., O'Brien, S.A., He, Y., Wang, L., Zhang, Q., Kim, A., et al. (2020). Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. *Cell* 181, 442–459.e29.
- Zhang, F., Mears, J.R., Shakib, L., Beynor, J.I., Shanaj, S., Korsunsky, I., Nathan, A., Donlin, L.T., and Raychaudhuri, S. (2021). IFN-γ and TNF-α drive a CXCL10+ CCL2+ macrophage phenotype expanded in severe COVID-19 and other diseases with tissue inflammation. *Genome Med.* 13, 64.
- Zhao, Q., Kuang, D.-M., Wu, Y., Xiao, X., Li, X.-F., Li, T.-J., and Zheng, L. (2012). Activated CD69<sup>+</sup> T cells foster immune privilege by regulating IDO expression in tumor-associated macrophages. *J. Immunol.* 188, 1117–1124.
- Zheng, C., Zheng, L., Yoo, J.-K., Guo, H., Zhang, Y., Guo, X., Kang, B., Hu, R., Huang, J.Y., Zhang, Q., et al. (2017). Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* 169, 1342–1356.e16.
- Zhou, Y., Song, W.M., Andhey, P.S., Swain, A., Levy, T., Miller, K.R., Poliani, P.L., Cominelli, M., Grover, S., Gilfillan, S., et al. (2020). Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat. Med.* 26, 131–142.
- Zhu, Y., Herndon, J.M., Sojka, D.K., Kim, K.-W., Knolhoff, B.L., Zuo, C., Cullinan, D.R., Luo, J., Bearden, A.R., Lavine, K.J., et al. (2017). Tissue-Resident Macrophages in Pancreatic Ductal Adenocarcinoma Originate from Embryonic Hematopoiesis and Promote Tumor Progression. *Immunity* 47, 597.
- Ziegler-Heitbrock, L., Ancuta, P., Crowe, S., Dalod, M., Grau, V., Hart, D.N., Leenen, P.J.M., Liu, Y.-J., MacPherson, G., Randolph, G.J., et al. (2010). Nomenclature of monocytes and dendritic cells in blood. *Blood* 116, e74–e80.
- Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerowitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* 50, 1317–1334.e10.

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Anti-goat (Polyclonal) AF488	Jackson Immuno-Research Laboratory	Cat# 705-545-147, RRID:AB_2336933
BTLA (Clone 743986) BV605	BD Biosciences	Cat# 743986, RRID:AB_2741906
CCR2 (Clone LS132.1D9) BUV563	BD Biosciences	Cat# 749076, RRID:AB_2873470
CCR7 (Clone G043H7) Biotin	BD Biosciences	Cat# 552174, RRID:AB_394352
CD11b (Clone M1/70) Biotin	BD Biosciences	Cat# 553309, RRID:AB_394773
CD11b (Clone M1/70) BV570	Biolegend	Cat# 101233, RRID:AB_10896949
CD123 (Clone 6H6) PE/Daz594	Biolegend	Cat# 306034, RRID:AB_2566450
CD123 (Clone 7G3) BUV395	BD Biosciences	Cat# 564195, RRID:AB_2714171
CD14 (Clone 63D3) SparkB550	Biolegend	Cat# 367148, RRID:AB_2832724
CD14 (Clone M5E2) BV650	BD Biosciences	Cat# 563419, RRID:AB_2744286
CD141 (Clone AD5-14H12) APC	Miltenyi	Cat# 130-113-314, RRID:AB_2733313
CD141 (Clone M81) BV421	Biolegend	Cat# 344114, RRID:AB_2562956
CD16 (Clone 3G8) APC/Cy7	Biolegend	Cat# 302017, RRID:AB_314217
CD16 (Clone 3G8) APC/Cy7	Biolegend	Cat# 302018, RRID:AB_314218
CD16 (Clone 3G8) BV650	Biolegend	Cat# 302042, RRID:AB_2563801
CD163 (Clone GHI/61) BV605	Biolegend	Cat# 333615, RRID:AB_2562712
CD163 (Clone GHI/61) BV605	Biolegend	Cat# 333616, RRID:AB_2616879
CD169 (Clone 7-239) BUV661	BD Biosciences	Cat# 750363, RRID:AB_2874538
CD169 (Clone 7-239) PE	BD Biosciences	Cat# 565248, RRID:AB_2732051
CD19 (Clone HIB19) BV650	Biolegend	Cat# 302238, RRID:AB_2562097
CD19 (Clone SJ25C1) BV650	BD Biosciences	Cat# 563226, RRID:AB_2744313
CD1a (Clone HI149) AF700	Biolegend	Cat# 300120, RRID:AB_528764
CD1c (Clone L161) BV421	Biolegend	Cat# 331525, RRID:AB_10933249
CD1c (Clone L161) BV421	Biolegend	Cat# 331526, RRID:AB_10962909
CD1c (Clone L161) SB436	Invitrogen	Cat# 62-0015-42, RRID:AB_2762426
CD20 (Clone 2H7) BV650	Biolegend	Cat# 302336, RRID:AB_2563806
CD20 (Clone 2H7) BV650	BD Biosciences	Cat# 563780, RRID:AB_2744327
CD206 (Clone 15-2) PP/Cy5.5	Biolegend	Cat# 321122, RRID:AB_10899411
CD206 (Clone 19.2) PE/CF594	BD Biosciences	Cat# 564063, RRID:AB_2732052
CD209 (Clone 9E9A8) PE/Cy7	Biolegend	Cat# 330114, RRID:AB_10719953
CD209 (eB-h209) PercP/Cy5.5	ThermoFisher Scientific	Cat# 45-2099-73, RRID:AB_1106983
CD3 (Clone SP34-2) BV650	BD Biosciences	Cat# 563916, RRID:AB_2738486
CD3 (Clone UCHT1) BUV805	BD Biosciences	Cat# 612895, RRID:AB_2870183
CD301 (Clone H037G3) PE	Biolegend	Cat# 354704, RRID:AB_11219002
CD38 (Clone HIT2) BB515	BD Biosciences	Cat# 564498, RRID:AB_2744374
CD4 (Clone SK3) CfI.YG584	Cytokine Biosciences	Cat# R7-20041, RRID:AB_2885083
CD40 (Clone 5C3) BV750	BD Biosciences	Cat# 746948, RRID:AB_2871735
CD45 (Clone 30-F11) PercP	Biolegend	Cat# 103130, RRID:AB_893339
CD45 (Clone HI30) BV510	BD Biosciences	Cat# 563204, RRID:AB_2738067
CD45 (Clone HI30) V500	BD Biosciences	Cat# 560777, RRID:AB_1937324
CD45RA (Clone 5H9) FITC	BD Biosciences	Cat# 556626, RRID:AB_396498
CD5 (Clone UCHT2) APC/R700	BD Biosciences	Cat# 565121, RRID:AB_2744433
CD5 (Clone UCHT2) BV711	BD Biosciences	Cat# 563170, RRID:AB_2738044
CD86 (Clone 2331) BV786	BD Biosciences	Cat# 740990, RRID:AB_2870657

(Continued on next page)

***Continued***

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CD86 (Clone IT2.2) BV605	Biolegend	Cat# 305429, RRID:AB_11203889
CD88 (Clone S5/1) APC/Fire750	Biolegend	Cat# 344316, RRID:AB_2750445
CD88 (Clone S5/1) PE/Cy7	Biolegend	Cat# 344307, RRID:AB_11125761
CD88 (Clone S5/1) PE/Cy7	Biolegend	Cat# 344308, RRID:AB_11126750
CD89 (Clone A59) BUV496	BD Biosciences	Cat# 750617, RRID:AB_2874749
CD9 (Clone M-L13) BV480	BD Biosciences	Cat# 746356, RRID:AB_2743675
CTLA4 (Clone 14D3) PP/eF710	Invitrogen	Cat# 46-1529-42, RRID:AB_2573718
Fc $\epsilon$ RI $\alpha$ (Clone AER-37 (CRA-1)) BV711	Biolegend	Cat# 334638, RRID:AB_2687186
Fc $\epsilon$ RI $\alpha$ (Clone AER-37 (CRA-1)) PerCP	Biolegend	Cat# 334616, RRID:AB_2168079
FOLR2 (Clone 94b/FOLR2) PE	Biolegend	Cat# 391704, RRID:AB_2721336
FOLR2 (Clone 94b/FOLR2) APC	Biolegend	Cat# 391706, RRID:AB_2721303
HLA-ABC (Clone W6/32) Pacific Blue	Biolegend	Cat# 311417, RRID:AB_493668
HLA-ABC (Clone W6/32) Spark	Biolegend	Cat# 311453, RRID:AB_2876612
HLA-DP (Clone B7/21) BUV395	BD Biosciences	Cat# 750866, RRID:AB_2874962
HLA-DP (Clone B7/21) eFluor506	BD Biosciences	Cat# 334609, RRID:AB_1227656
HLA-DQ (Clone Tu169) AF647	BD Biosciences	Cat# 564806, RRID:AB_2738963
HLA-DQ (Clone Tu169) BV510	BD Biosciences	Cat# 742609, RRID:AB_2740907
HLA-DR (Clone L243) BV785	Biolegend	Cat# 307641, RRID:AB_2561360
HLA-DR (Clone L243) BV785	Biolegend	Cat# 307642, RRID:AB_2563461
HLADR (Clone L243) APC/Fire810	Biolegend	Cat# 307673, RRID:AB_2876603
NOTCH2 (Clone MHN2-25) BUV615	BD Biosciences	Cat# 752315, RRID:AB_2875832
PD-L1 (Clone MIH1) R700	BD Biosciences	Cat# 565188, RRID:AB_2739101
PD-L1 (Clone MIH1) PE/Cy5	Invitrogen	Cat# 15-5983-42, RRID:AB_2802211
PD-L2 (Clone 24F.10Ca12) APC	Biolegend	Cat# 329608, RRID:AB_1089013
PD-L2 (Clone MIH18) BUV737	BD Biosciences	Cat# 748698, RRID:AB_2873102
PD1 (Clone EH12.1) BV750	BD Biosciences	Cat# 747446, RRID:AB_2872125
PD1 (Clone PD1.3) PE/Cy5.5	Beckman Coulter	Cat# B36123
SLAN (Clone DD-1) VioBlue	Miltenyi Biotec	Cat# 130-119-868, RRID:AB_2733608
Streptavidin (Polyclonal) BV570	Biolegend	Cat# 405227
TREM2 (Polyclonal) pure	R&D Systems	Cat# AF1828, RRID:AB_2208689
TREM2 (Clone #237920) AF647	R&D Systems	Cat# FAB17291P, RRID:AB_884528
XCR1 (Clone S15046E) FITC	Biolegend	Cat# 372612, RRID:AB_2715831
Zombie NIR Fixable Viability Kit	Biolegend	Cat# 423106
LIVE/DEAD Fixable Blue Dead Cell	Invitrogen	Cat# L23105
<b>Chemicals, peptides, and recombinant proteins</b>		
DNase I	Sigma Aldrich	Cat# 48024000
PBS	ThermoFisher	Cat# 20012-027
Collagenase IV	Sigma Aldrich	Cat# C5138-500MG
RPMI	ThermoFisher	Cat# 31870-025
Brilliant Stain Buffer	BD Biosciences	Cat# 563794
FCS	ThermoFisher	Cat# 26140079
<b>Deposited data</b>		
SMARTseq2 single cell transcriptome data of human blood and tissue MNP	This paper	<a href="#">GSE178209</a>
<b>Biological samples</b>		
Healthy Lung sample	National University Hospital (NUH), Singapore	Healthy Lung sample
Matched LUAD lung sample	Marie Lannelongue Hospital, Paris	Matched LUAD lung sample
Matched HCC liver sample	Singapore General Hospital (SGH) and National University (NUH)	Matched HCC liver sample

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Software and algorithms</b>		
Anaconda 2018.12	Anaconda	<a href="https://www.anaconda.com">https://www.anaconda.com</a>
DIVA	BD Biosciences	<a href="https://www.bdbiosciences.com/en-us">https://www.bdbiosciences.com/en-us</a>
FlowJo v10.5.3	Tree Star	<a href="https://www.flowjo.com">https://www.flowjo.com</a>
GraphPad Prism 6	GraphPad	<a href="https://www.graphpad.com/scientificsoftware/prism/">https://www.graphpad.com/scientificsoftware/prism/</a>
Ingenuity Pathway Analysis v01-16	QIAGEN	<a href="https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/">https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/</a>
NicheNet	Browaeys et al., 2020	<a href="https://github.com/saeyslab/nichenet">https://github.com/saeyslab/nichenet</a>
Phenograph	Levine et al., 2015	<a href="https://github.com/JinmiaoChenLab/Rphenograph">https://github.com/JinmiaoChenLab/Rphenograph</a>
Python v3.7.1	Python Software Foundation	<a href="https://www.python.org">https://www.python.org</a>
R v4.4	The R Foundation	<a href="https://www.r-project.org">https://www.r-project.org</a>
SCENIC	Van de Sande et al., 2020	<a href="https://github.com/aertslab/pySCENIC">https://github.com/aertslab/pySCENIC</a>
SeqGeq v1.6	FlowJo	<a href="https://www.flowjo.com/solutions/seqgeq">https://www.flowjo.com/solutions/seqgeq</a>
Seurat v3	Stuart et al., 2019	<a href="https://satijalab.org/seurat/">https://satijalab.org/seurat/</a>
Seurat v4	Hao et al., 2021	<a href="https://satijalab.org/seurat/">https://satijalab.org/seurat/</a>
UMAP	McInnes et al., 2018	<a href="https://github.com/lmcinnes/umap">https://github.com/lmcinnes/umap</a>

**RESOURCE AVAILABILITY****Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Charles-Antoine Dutertre ([charles-antoine.dutertre@inserm.fr](mailto:charles-antoine.dutertre@inserm.fr)) and Florent Ginhoux ([Florent\\_Ginhoux@immunol.a-star.edu.sg](mailto:Florent_Ginhoux@immunol.a-star.edu.sg)).

**Materials availability**

Previously unpublished single-cell RNA-seq data have been deposited at Gene Expression Omnibus (GEO:[GSE178209](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178209)) and are publicly available as of the date of publication. Accession numbers are listed in the [Key resources table](#). Our analysis code has been uploaded onto github (<https://github.com/gustaveroussy/FG-Lab>).

**Data and code availability**

The MNP- and MoMac- VERSES can be explored and downloaded at <https://gustaveroussy.github.io/FG-Lab/>.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS****Human samples**

Human samples were obtained in accordance with a favorable ethical opinion from Singapore SingHealth and National Health Care Group Research Ethics Committees. Liver tissues were obtained from living donor transplantations (Asian American Liver Centre, Gleneagles Hospital, Singapore). Healthy liver perfusates were obtained from the liver grafts of donors by flushing intrahepatic veins with cold saline prior to transplantation. Explanted liver tissues from patients with advanced cirrhosis and/or HCC were obtained and perfused in a similar manner *in vitro* (using syringe and needle) and leukocytes isolated as previously described (Tan-Garcia et al., 2017). Spleen tissue was obtained from patients with tumors in the pancreas who underwent distal pancreatectomy (Singapore General Hospital, Singapore). Tonsil tissue was obtained from patients with adeno-tonsillar obstruction and who underwent adeno-tonsillectomy (KK Hospital, Singapore). Lung tissue was obtained from pancreatic cancer patients from the Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore. Tumor and healthy adjacent tissue from lung were obtained from lung adenocarcinoma patients following written informed consent (Marie Lannelongue Hospital, Paris) and ethical approval (N°ID-RCB: 2016-A00732-49). Spleen, lung, liver, and tonsil tissues were processed as previously described (Dutertre et al., 2019) for SMARTseq2. Liver samples for flow cytometry validation were processed as previously described (Sharma et al., 2020). All subjects provided IRB-approved consent.

**METHOD DETAILS****Flow cytometry and cell sorting**

Cells were thawed from liquid nitrogen and transferred into RPMI (ThermoFischer) with 20% decomplemented FCS (ThermoFischer). Samples were treated with 1mg/ml DNase I (Sigma-Aldrich) at 37 C. Cells were incubated with Live/Dead blue dye (Invitrogen) or with

Zombie NIR (Invitrogen) for 30 min at 4 °C in phosphate buffered saline (PBS) and then incubated in 5% heat-inactivated fetal calf serum (FCS) for 15 min at 4 °C (Sigma Aldrich). Cells were stained with appropriate antibodies (listed in the STAR Methods key resources table) in PBS with 2% FCS and 2mM EDTA (Sigma Aldrich) and Brilliant Stain buffer (BD) and incubated for 30 min at 4 °C, and then washed. For indexed-sorting, cells were sorted using a BD ARIALII (70 µm nozzle; BD Biosciences) and for macrophage phenotyping, cells were analyzed using a Cytek Aurora 4-laser or 5-laser spectral analyzer. Fcs files were exported and analyzed using FlowJo v10.5.3.

#### Algorithms for dimensionality reduction

For flow cytometry data, marker expression values were transformed using the auto-logicle transformation function from the flowCore R package. Uniform Manifold Approximation and Projection (UMAP) were carried out using all markers (flow cytometry) or significant PCs (based on Seurat analysis for scRNaseq data). UMAP was run using 15 nearest neighbors (*nn*), a *min\_dist* of 0.001 to 0.2 and Euclidean distance (Becht et al., 2019; McInnes et al., 2018). Phenograph clustering (Levine et al., 2015) was performed using all markers or significant PCs (based on Seurat analysis) before dimension reduction, and with the number of nearest neighbors equal to 150 and 100, for MoMac-VERSE and MNP-VERSE, respectively, 30 for flow cytometry and equal to 15 for scRNaseq analysis.

#### MNP extraction and Seurat V3 integration

The 41 datasets used (Table S1) were either at the raw count matrix or already pre-processed and at the filtered stage. We first integrated all the datasets in an organ-specific manner. Before running the datasets through the integration, we applied universal quality control to keep everything in a unified manner. Cells that expressed fewer than 500 genes and had more than 20% mitochondrial reads were filtered out. All datasets were then unified in the same expression matrix format. Integration was initiated using the Seurat V3 anchoring method (Stuart et al., 2019) and log normalized. The matrix was scaled and a Principal Component Analysis (PCA) was performed (Becht et al., 2019) from which the first 50 Principal Components (PCs) were selected for UMAP analysis. Following the identification of mononuclear phagocytes (MNPs) using canonical markers, a global integration (using 100 PCs for dimensional reduction) of monocytes and macrophages from all tissues was carried out as above.

#### Generation of indexed-sorting and SMARTseq2 single-cell transcriptome data

Cells isolated from organs were indexed-sorted using the Indexed-sorting panel (Table S2) on a BD FACS ARIALII (BD Biosciences) into 96 well plates containing 3 µL Lysis buffer (see below) using a 70 µm nozzle. Single-cell cDNA libraries were prepared using the SMARTSeq v2 protocol (Picelli et al., 2014) with the following modifications: (i) 1 mg/ml BSA Lysis buffer (Ambion Thermo Fisher Scientific, Waltham, MA, USA); and (ii) 200 pg cDNA with 1/5 reaction of Illumina Nextera XT kit (Illumina, San Diego, CA, USA). The length distribution of the cDNA libraries was monitored using a DNA High Sensitivity Reagent Kit on the Perkin Elmer Labchip (Perkin Elmer, Waltham, MA, USA). All samples were subjected to an indexed paired-end sequencing run of 2x151 cycles on an Illumina Hi-Seq 4000 system (Illumina, San Diego, CA, USA), with 300 samples/lane. Data are available through GEO (GEO: GSE178209).

#### Pre-processing, quality assessment and control and analysis of SMARTseq2 single-cell transcriptome data

Paired-end raw reads were aligned to the human reference genome (GRCh38 version 25 release; Gencode) using RSEM version 1.3.0. Transcript Per Million read (TPM) values were calculated using RSEM and used for downstream analysis. Quality control, selection of highly variable genes, PCA, and differential gene expression analysis was performed using the Seurat R package. All scRNaseq dot plots and mean plots displaying the gene expression levels or mean signature genes were generated using SeqGeq v1.6 (Flow Jo LLC).

#### Generation of transformed matrix

To allow for analyses across all datasets, a transformed matrix was generated using datasets that contain more than 10,000 common genes inclusive of FOLR2. Datasets of Cheng et al. (2018), James et al. (2020), MacParland et al. (2018), Stewart et al. (2019), Vieira Braga et al. (2019), Xue et al. (2019), and Zheng et al. (2017) were excluded as they did not meet the above criteria (Figure S1). The pipeline for the generation was adapted from Seurat V3. When loading in the selected datasets, we increased the amount of anchoring points (*anchor features*) and also the number of gene features (*nfeatures*). Both values were set to the highest number of genes (56,000) in all the datasets. The normalized matrix was extracted from the Seurat object.

## QUANTIFICATION AND STATISTICAL ANALYSIS

#### Differentially expressed genes (DEGs) and “transformed” genes (DEtGs) analyses

DEG and DEtG analyses were performed using the Seurat v3 package (Stuart et al., 2019). All DEtGs obtained from the transformed matrix were calculated on non-normalized values with a logFC threshold of 0.25 (unless specified otherwise) and DEGs obtained from real TPM/count matrixes were calculated on normalized values with a logFC threshold of 0.25. In both cases, the likelihood-ratio test for single-cell gene expression (bimodal test) was used and correction for multiple testing was carried out using the Bonferroni method. DEtGs between healthy and cancer tissues (Figures 2E and F), healthy and inflammation (Figures 2H and 2I) were calculated with a logFC threshold of 0.025. The heatmap was generated using the top 50 significant DEtGs from each cluster. All DEtGs

calculated from the transformed matrix are depicted in a red/white/blue color scheme, while DEGs calculated from original TPM are shown in a yellow/black/purple color scheme.

### Metadata analysis

Metadata analysis was performed for selected studies with paired conditions (healthy versus cancer or inflammation). The proportion of phenograph clusters were plotted for each condition as charts and density plots for the selected studies. Further analysis was performed to deconvolute at the patient level in datasets where this information was provided. We only analyzed datasets where more than 35 cells were present. Charts and density plots were made in GraphPad Prism v6 and SeqGeq v1.6, respectively. Statistical tests were performed using GraphPad Prism v6 and are specified within the figure legends.

### M1 and M2 signature

M1 and M2 signatures were derived from Table 1 published by [Martinez et al. \(2006\)](#). Those in common with DEtGs obtained across phenograph clusters were used to define M1- and M2-like signatures. The mean gene expression of M1- and M2-like gene signatures were analyzed in SeqGeq v1.6 ([Figures 2C](#) and [S3F](#)). We also graphed the number of common genes between the DEtGs from each cluster ([Figure 2D](#)).

### Generation of the major MNP subset DEGs

Six datasets ([Azizi et al., 2018](#); [Cillo et al., 2020](#); [Kim et al., 2020](#); [Smillie et al., 2019](#); [Zhang et al., 2020](#); [Zilionis et al., 2019](#)) that cover the UMAP space of the MNP-VERSE were selected for calculating DEGs across major MNP subsets using the original TPMs. Unique DEGs to each of these subsets that were common across the six datasets were identified as the universal gene signatures. Expression of these universal gene signatures on the transformed matrix (DEtGs) was plotted as a heatmap ([Figure 1G](#)). The mean expression of the universal gene signatures for each MNP subset was visualized using SeqGeq v1.6 ([Figure 1H](#)).

### Scenic gene regulatory network analyses

To infer gene regulatory networks (GRNs) from TPM-normalized expression matrices of colon ([Zhang et al., 2020](#)), liver ([Sharma et al., 2020](#)), and lung ([Kim et al., 2020](#)), a pySCENIC (single-cell regulatory network inference and clustering) v0.10.3 analysis was performed ([Van de Sande et al., 2020](#)). The analysis consisted of three main steps (GitHub/pySCENIC): generation of co-expression modules with GRNBoost2, refinement of these modules with RcisTarget and evaluation of the regulon activity with AUCCell ([Van de Sande et al., 2020](#)). Differentially expressed regulons (DER) were calculated using the Seurat pipeline with the same parameters as described above for DEG/DETG analysis (adjusted p value lower or equal to 0.05 and Log2FC cut-off of 0.25). Phenograph cluster-specific DER, as well as DER that had similar expression patterns across closely related phenograph clusters were identified and subsequently used to generate a heatmap ([Figure 2B](#)).

### Pathway analysis

Subset-specific and/or condition specific DEtGs, together with the respective fold-change and p values, were uploaded to the Ingenuity Pathway Analysis (IPA) software (QIAGEN). IPA analysis reported the p value of canonical pathways and upstream regulators. Predicted upregulated or downregulated pathways were represented by a positive or negative Z-score, respectively. Canonical pathways determined by IPA's default threshold [ $-\log_2(p \text{ value}) > 1.3$ ] were then shortlisted and radar plots were used to visualize the p values and Z-scores as previously described in [Dutertre et al. \(2019\)](#). Full lists of pathways can be found in [Table S6](#).

### NicheNet analysis

NicheNet is a computational method that uses scRNaseq data to predict interactions between cell populations ([Browaeys et al., 2020](#)). To identify potential interactions between T cells and subsets of macrophages in cancer, we used the HCC dataset from [Sharma et al. \(2020\)](#), that includes transcriptomic data for MNPs, lymphocytes and non-immune cells. The phenograph clusters that were enriched in cancer (clusters #2, #3, #4, #6 and # 15) were set as "receiver" populations, while the subsets of T cells were set as "sender" populations. Ligand-receptor pairs were included in the analysis when a minimum of 15% of the cells in the respective population expressed the protein (pct = 15 in the `get_expressed_genes` function). The genes identified by DEtGs and DER in the MoMac-VERSE were used to predict the T cell ligands activities for each phenograph cluster ("geneset" input in the `predict_ligand_activities` function). The putative T cells ligands (the output of the `predict_ligand_activities` function) were ranked according to their ability to predict the gene signatures using Pearson correlation coefficients and can be found for each MNP phenograph cluster in [Table S7](#). The top 3 ligands and their regulatory potential were represented in a heatmap following the NicheNet pipeline, using the function `get_weighted_ligand_target_links` on the top 20 predicted ligands with n = 200 (top target genes), and a cutoff of 0.33 in the function `prepare_ligand_target_visualization`.