Robust temporal map of human in vitro myelopoiesis using single-cell genomics

Clara Alsinet1,2✉, Maria Nascimento Primo 1,2, Valentina Lorenzi 1, Erica Bello 1,2, Iva Kelava1,
Carla P. Jones1, Roser Vilarrasa-Blasi1, Carmen Sancho-Serra1, Andrew J. Knights1, Jong-Eun Park3,
Beata S. Wyspianska2,4, Gosia Trynka 1,2, David F. Tough4, Andrew Bassett1,2, Daniel J. Gaffney1,2✉,
Damiana Alvarez-Errico5,5✉ & Roser Vento-Tormo1,3,4,5✉

Myeloid cells are central to homeostasis and immunity. Characterising in vitro myelopoiesis protocols is imperative for their use in research, immunotherapies, and understanding human myelopoiesis. Here, we generate a >470K cells molecular map of human induced pluripotent stem cells (iPSC) differentiation into macrophages. Integration with in vivo single-cell atlases shows in vitro differentiation recapitulates features of yolk sac hematopoiesis, before definitive hematopoietic stem cells (HSC) emerge. The diversity of myeloid cells generated, including mast cells and monocytes, suggests that HSC-independent hematopoiesis can produce multiple myeloid lineages. We uncover poorly described myeloid progenitors and conservation between in vivo and in vitro regulatory programs. Additionally, we develop a protocol to produce iPSC-derived dendritic cells (DC) resembling cDC2. Using CRISPR/Cas9 knock-outs, we validate the effects of key transcription factors in macrophage and DC ontogeny. This roadmap of myeloid differentiation is an important resource for investigating human fetal hematopoiesis and new therapeutic opportunities.
Macrophages perform a variety of functions, from tissue homoeostasis to immune surveillance and from the response to infection to the resolution of inflammation. They originate during both development and adulthood and acquire tissue-specific functions. Despite the commonalities within mammals, there are important differences between humans and rodent models. Establishing and characterising the current human in vitro models is essential to fully exploit their research and therapeutic potential.

During development, myeloid cells originate from at least two waves of progenitors: a first wave involving myeloid-biased progenitors from the yolk sac (yolk-sac myeloid progenitors, YSMP) and a second wave through definitive hematopoietic stem cells (HSC). YSMP are thought to appear during the first 2 weeks of development in humans and are responsible for producing primordial blood. HSC are not generated until 3–4 post-conceptional weeks (PCW) in the gonad-aorta-mesonephros. HSC and myeloid progenitors derived from YSMP colonise the liver, making this foetal organ the main site of hematopoiesis until mid-pregnancy. Later, HSC are restricted to the bone marrow, the only hematopoietic site during adulthood. In mice, YSMPs generate a wide range of myeloid cells, including monocytes and neutrophils, and are thought to be the main myeloid precursors during development. In humans, we have limited knowledge about the progression and regulatory mechanisms defining YSMP.

In vitro models of macrophage differentiation hold promise to not only answer these biological questions but also to become therapeutic tools, particularly immunotherapies, and for high-throughput screening, including drug testing. Macrophages derived from human induced pluripotent stem cells (iPSC) show tissue-resident phenotypes and are an attractive alternative to adult monocyte-derived macrophage cultures. A current protocol, developed by van Wilgenburg et al., is a straightforward, feeder-free process consisting of 3 steps using constant concentrations of 1 to 3 cytokines. It provides long-term, scalable production of macrophage precursors without fluorescence-activated cell sorting (FACS) since the cells of interest continuously expand and detach from the culture. Despite this being an established in vitro macrophage model, the exact intermediate populations produced are unclear. This restricts its applications and limits our true understanding of the cells obtained. Thus, a thorough analysis of the cell identities and dynamics emerging during in vitro differentiation is imperative to establish their likely in vivo counterparts and fully exploit this technology.

Here, we profiled the single-cell transcriptome and open chromatin of >400k and >70k cells, respectively, during iPSC–myeloid differentiation with the van Wilgenburg protocol. We uncover a wide range of cell states, their ontology and underlying transcription factor (TF) networks, that accurately map foetal myelopoiesis in the YS. We demonstrate the versatility of the current in vitro protocol by modifying the media used and using CRISPR/Cas9 edited iPSCs. Altogether, we demonstrate that macrophage differentiation from iPSC is a robust system to study the early stages of human myelopoiesis, and that macrophages obtained are able to acquire definitive tissue-resident identities. Our data sets can be visualised and downloaded from www.HiPImmuneAtlas.org.

**Results**

**iPSC-derived cells have human yolk sac myelopoiesis features.** We profiled the full differentiation of iPSC into macrophages from 6 individuals using single-cell RNA and ATAC sequencing (scRNAseq and scATACseq) (Fig. 1a and Supplementary Data 1). The differentiation protocol consists of 3 steps: (i) spin-embryoid body (EB) formation from day 1 to day 4, (ii) EB myeloid differentiation from day 5 onwards (the latest sample used in this study is from day 31), and (iii) macrophage differentiation using non-adherent cells from EB myeloid differentiation phase and lasting for 7 days, in this study we used cells from day 31 (day 31 to day 31 + 7). To characterise the robustness of our results, we generated two independent scRNAseq data sets. The first data set (referred to as Discovery data set) included scRNAseq data from 3 donors at 20 timepoints (Fig. 1a, b). The second data set (hereafter, Validation data set) included scRNAseq and scATACseq data from 6 donors at 7 and 6 time points, respectively (Fig. 1a, b). The three donors from the Discovery data set were also used in the Validation data set, thus generating biological replicates.

After quality control, the Discovery data set contained a total of 135,000 cells (Fig. 1c and Supplementary Fig. 1). To annotate the cell types in an unbiased manner, we built logistic regression (LR) models trained on publicly available single-cell transcriptomics data sets and projected the data into our in vitro data set (Fig. 1b). We used multiple human developmental data sets to train our models, including a full gastrula and the main foetal hematopoietic organs: yolk sac, liver and thymus. Cell type labels were assigned based on the mean LR prediction probability of each cell cluster (Fig. 1c, Supplementary Fig. 2 and Supplementary Data 3). Marker gene expression analysis further supported the LR cell-type annotation (Fig. 1d).

Cell type label transfer from the Discovery data set into the Validation and scATACseq data sets confirmed the presence of the main cell populations in all 3 data sets (Fig. 1e, f, Supplementary Fig. 1 and Supplementary Data 4).

The majority of cells at the initial EB formation stage (day 1–4) matched gastrulation cell populations (Fig. 1c and Supplementary Fig. 3). We found primitive streak-like cells, emergent and advanced mesoderm, and the initial appearance of hemogenic endothelium. Despite using cytokines that induce hematopoietic mesoderm (Fig. 1a), we also observed populations related to other germ layers (i.e. neural crest and endoderm, Fig. 1c).

During EB myeloid differentiation (day 5–31), the myeloid and stromal cell compartments emerged (Fig. 1c and Supplementary Fig. 3). The myeloid populations included a wide range of cell types, such as erythrocytes, megakaryocytes, mast cells, neutrophil myeloid progenitors (NMP), monocyte DC precursors (MDP), monocytes and macrophages. Being so unexpected, we validated the presence of the small population of mast cells by FACS (Supplementary Fig. 4). Most in vivo counterparts for these cell types were found in the foetal liver and thymus. However, we did not find any cluster in the Discovery data set that corresponded to the in vivo HSC found in the human developing liver (Supplementary Fig. 2B). Instead, there was a distinct cluster of myeloid progenitors (MP) expressing CD34, low levels of SPINK2 and PTPRC, but not HOXA genes, which are required to generate definitive HSC, suggesting in vitro myelopoiesis recapitulates YS differentiation. In vivo YSMP and macrophages LR models captured more than one cell type within the in vitro data set (Fig. 1g, h). To explore this further, we performed the opposite exercise: we trained models on our in vitro cell types and projected them onto the in vivo YS data set. As expected a subset of cells within the published YSMP cluster showed a high prediction probability with the in vitro NMP-trained model and a subset of in vivo macrophages were captured by the MDP-trained model (Supplementary Fig. 5B). Using the LR results, we annotated the NMP and MDP cell types within the embryonic YS data set (Supplementary Fig. 5B).
To quantitatively characterise the macrophage foetal-like profile observed, we projected adult and foetal macrophages data from the human decidual–placental interface into our data set. This unique tissue setting includes both adult/maternal monocyte-derived macrophages and foetal/placental YS-derived macrophages (Hofbauer cells), thus avoiding the technical confounders of adult+foetal data set integration. The Hofbauer cells LR model had a higher mean prediction probability for iPSC-derived macrophages than any of the adult macrophage subtypes identified in the placenta (Supplementary Fig. 5D, E). As an exception, day31 +1 macrophages presented a higher score with adult macrophages LR models (Supplementary Fig. 5D, E).
This was likely because day31 + 1 macrophages showed an activated state, upregulating inflammatory cytokines such as CXCL8, CCL7 or IL1B, (Supplementary Fig. SF) also found on monocyte-derived macrophages in the decidua (adult/maternal tissue). Next, we used a hepatocarcinoma cell data set27 and projected several tumour-associated macrophages (TAMs) LR models on our iPSC-derived data set (Supplementary Fig. 5G, H). We found the foetal-like FOLR2+ TAMs model showed a higher prediction score among end-stage macrophages (day31 + 7) while the SPP1+ TAMs LR model markedly captured macrophages on the activated state (day31 + 1, Supplementary Fig. 5H). Overall, this indicated that macrophages produced in the iPSC protocol have a strong foetal phenotype, and this could be relevant for their application as in vitro TAM models.

**Trajectory analysis and underlying regulatory programmes.**

Myelopoiesis is shaped by transcriptional programmes, including TFs, epigenetic regulators and post-transcriptional mechanisms28. We set out to reconstruct the main developmental pathways underlying in vitro myelopoiesis and the transcriptional networks mediating them. The high-throughput single-cell approach used, the high density of time points collected and trajectory analysis using scVelo29, allowed us to reconstruct the differentiation pathways giving rise to the wide range of cell types observed (Fig. 2a and Supplementary Fig. 6A). In parallel, we compared the underlying regulatory programmes mediating such transitions in vivo and in vitro. To this end, we measured TF activities by looking at the expression of consensus TF targets30 (Fig. 2b).

In step 1 of in vitro differentiation, iPSC differentiated into the primitive streak, which subsequently gave rise to either endoderm or emergent and advanced mesoderm (Fig. 2c). The split of primitive streak cells observed (Fig. 2c) suggested there is heterogeneity within this cell type. Later, advanced mesoderm differentiated into hemogenic endothelium, which was the precursor of myeloid cells (Fig. 2c). Primitive streak and mesoderm were transient populations that disappeared by day16, while endoderm and hemogenic endothelium were detectable at least until day31, the latest time point analysed (Fig. 2a and Supplementary Fig. 6A). The transition from mesoderm to hemogenic endothelium has also been reported during the gastrulation period31. TF activity analysis showed high conservation of the TF modules in these earlier stages of development (Fig. 2d and Supplementary Data 5). Pluripotency TFs activity (POUSF1, NANOG and SOX2) decreased upon differentiation into mesoderm and endothem. As expected, in both in vivo and in vitro settings, mesoderm activated SMAD4, HOXA9 or SRF while endoderm activated FOXA2 and HNF4A. Later, hemogenic endothelium activated TFs relevant for hematopoiesis including RUNX1, SPI1, RBPJ, MEF2A and MEF2C. GATA1 was also activated in this transition but it showed the highest activation levels in erythrocytes (Fig. 2d).

In vitro, myelopoiesis started very early in the EB formation stage, but the wide range of myeloid cell types appeared almost simultaneously starting at day14, and they all endured until day31 (Fig. 2a and Supplementary Fig. 6A). The first MP appeared at day9–11 followed by erythrocytes on day11–14, and full myelopoiesis was achieved on day16–18. In addition to myeloid cells, advanced mesoderm also differentiated into an intermediate stage of early fibroblasts (day7), giving rise to fibroblasts by day9. Trajectory analysis on the sample at day21 reconstructed all myelopoiesis differentiation steps. Hemogenic endothelial cells, derived from the mesoderm, differentiated into MP, which gave rise to both megakaryocytes and NMP (Fig. 2e and Supplementary Fig. 6A). NMP gave rise to MDP, which differentiated into either monocytes or macrophages (Fig. 2e). The MDP to macrophage trajectory was consistently observed from day21 to day31 (Supplementary Fig. 6B). The differentiation pathway of macrophages through MP and bypassing the monocytes is consistent with the first waves of myelopoiesis emerging in the YS3.

Throughout all stages of myelopoiesis, we found high similarity between the regulatory programmes activated in vivo (YS and foetal liver) and in vitro (Fig. 2f, g and Supplementary Data 5). The transition from hemogenic endothelium to MP was characterised by the activation of multiple TFs including RUNX1, SPI1 and GATA1 (Fig. 2f, g). The MP to NMP transition showed further activation of SPI1 and CEBPA. On the contrary, a large number of endothelial TFs, such as SOX2 and the ETV family, pluripotency factors, such as POUSF1 and NANOG, or lymphoid lineage-promoting factors, such as MEF2C, were inactivated32. MEF2C is a TF characteristic of definitive HSC that drives lymphoid fate choice32. The low MEF2C activity in in vitro MPs, further supported the HSC-independent features of this system (Fig. 2f).

iPSC-derived differentiation towards monocytes and macrophages also shared transcriptional programmes with its in vivo YS and foetal liver counterparts. Among the few TFs specifically activated in MDP from NMP was RFX5, which regulates MHC-II transcription and is responsible for a rare hereditary immunodeficiency33. We also observed activation of TFs controlling inflammatory programmes in monocytes and macrophages, such as JUN, RELA and NFKB1. In macrophages, we found key similarities related to their myeloid identity such as MAF34 and CEBPB35 and potentially underlying tissue-specific programmes, such as the alveolar macrophage programme characterised by PPARG36,37 (Fig. 2f, g).
Finally, we evaluated the chromatin accessibility dynamics of iPSC–macrophage differentiation using scATACseq. The number of accessible cell peaks decreased alongside the trajectories identified, suggesting a more restrictive chromatin landscape as cells differentiate. As an exception, hemogenic endothelium had a higher median of accessible peaks per cell (n = 11,074) than its mesoderm progenitors ('Emergent Mesoderm' n = 9759, 'Advanced Mesoderm' n = 8158). These differences were not associated with the number of expressed genes ('Hemogenic endothelium' n = 2514, 'Emergent Mesoderm' n = 3463, 'Advanced Mesoderm' n = 2725). Another exception was the macrophages after the macrophage differentiation phase ('Macrophages_Day31plus7'), which had more accessibility.
peaks than the macrophages from the EB myeloid differentiation phase (‘Macrophages’ $n = 6142$ vs ‘Macrophages_Day31plus7’ $n = 10,801$) despite also having a similar number of expressed genes (‘Macrophages’ $n = 2312$ vs ‘Macrophages_Day31plus7’ $n = 2219$). Finally, NMPs had a very low number of accessible peaks ($n = 4071$), in line with the low number of genes expressed in these cells ($n = 1782$) (Fig. 2h and Supplementary Fig. 6C).

**Modifications in the last stage produce diverse macrophages.** Non-adherent cells at day31 were collected and plated in a fresh medium with cytokines for 7 days as step 3 of differentiation (Fig. 3a). We analysed the changes over time (time points: day31, day31 + 1, day31 + 4, day31 + 7, day31 + 7) compared to the effect of multiple cytokines (cytokines: M-CSF, GM-CSF, GM-CSF + IL-34) and tested the effect of using foetal bovine serum (FBS) vs. defined medium (media: RPMI + FBS, StemPro34). We annotated the cell types using LR from the foetal liver data set22 and they clustered together across experiments (Fig. 3a, b). Next, we performed a neighbourhood expression analysis38 and observed that time points and media composition induced the highest diversity within macrophages (homogeneity ROGUE scores: time points 0.53, media composition 0.45 and cytokines 0.7, Fig. 3d–f and Supplementary Fig. 7A).

For the time point experiment, we analysed samples at day31, day31 + 1, day31 + 4 and day31 + 7 using M-CSF standard stimulation (Fig. 3a, b). The non-adherent cells collected at day31 from the EB myeloid differentiation phase were already mostly macrophages, alongside the main myeloid populations and a small subset of fibroblasts (Fig. 3c). By day31 + 7 there was an enrichment in macrophages, representing a total of 94.1% of cells (Fig. 3c), which was consistent with the CD14 + /CD64 + cells analysed by FACS (Supplementary Fig. 7B). We observed broad transcriptional differences between day31 and day31 + 1 macrophages, as well as between day31 + 1 and day31 + 4/7, while macrophages from day31 + 4 and day31 + 7 overlapped (Fig. 3b).

We also analysed the changes in TF activity over time30,39 (Supplementary Data 6). From day31 to day31 + 1, the increased activity of JUN, FOS and NFKB1 confirmed the transient but robust immune activation identified earlier (Supplementary Fig. 5F). This transcriptional activation decreased during day31 + 4 and day31 + 7 and some TF activities returned to basal levels (e.g. STAT1, NFKB1 and RELA, Fig. 3d). Despite being globally similar transcriptomically, we observed few but relevant differences between day31 and day31 + 7, including increased MITF and decreased SRF activities, which regulate phagocytosis40,41, in addition to decreased SREBF1 and SREBF2 activity, involved in lipid metabolism and macrophage polarisation42,43.

To assess whether the transient immune activation affected chromatin structure, we analysed the scATACseq macrophages data. We found that 113 TF motifs were significantly enriched on day31 + 7 ATAC peaks, compared to day31, and we obtained TF transcriptional activity scores for 55 of these (Fig. 3e and Supplementary Data 6). Importantly, the top 11 enriched motifs (enrichment score $>5$) at day31 + 7 corresponded to TFs that had been transcriptionally activated at day31 + 1 but were no longer activated at day31 + 7 (Fig. 3e). Indeed, the global TF motif enrichment profile at day31 + 7 correlated with TF activities at day31 + 1 (Pearson correlation: $r = 0.45$, $p < 0.0003$) but not at day31 + 7 (Pearson correlation: $r = 0.11$, $p = 0.38$) (Fig. 3e). In short, this analysis suggested that activation on day31 + 1 was transient transcriptionally but seemed to induce chromatin changes still present on day31 + 7. Therefore, earlier macrophages may represent a more naive cell state, amenable to further reprogramming in response to polarisation cues, which could have distinct applications than later macrophages.

Activated macrophages can be classified as M1 or M2 depending on whether they kickstart inflammation or resolve it, and the cytokines M-CSF and GM-CSF have been generally used to induce these phenotypes, respectively44. We found that specific TFs, such as MAF, ERG and LYL1, had reduced activity scores in GM-CSF vs. M-CSF macrophages (Fig. 3f and Supplementary Data 6). In contrast, RFX5 showed increased activation in GM-CSF macrophages33. Despite GM-CSF being largely known to promote PPARG activation36, this was not present in GM-CSF + IL-34 culture conditions (Fig. 3f). Of note, IL-34 is essential for the development of microglia from embryonic myeloid precursors45, and GM-CSF + IL-34 induces the microglial phenotype on monocytes in vitro46. Both the knockdown and pharmacological antagonism of PPARG promotes lipopolysaccharide (LPS)-stimulated transition from the M1 to the M2 phenotype in primary microglia, with the concomitant upregulation of markers such as CD206, TGFb and IL-447.

Furthermore, we stimulated the 4 macrophage subsets with LPS for 6 h to study their TLR4 stimulation response (Fig. 4a). These included the macrophages obtained at the EB myeloid differentiation phase, as well as those obtained after the 7-day differentiation with distinct cytokines (M-CSF, GM-CSF and GM-CSF + IL-34). As a readout, we evaluated their single-cell transcriptomics profile and cytokine protein levels in the supernatant. Both analyses showed a clear distinction between control and LPS-stimulated conditions (Fig. 4b). Known LPS-induced genes (e.g. IL6 or TNF) were overexpressed in all macrophage populations (Fig. 4c and Supplementary Data 7). Similarly, TNFa and IL-6 protein levels were also upregulated (Fig. 4d). Regarding LPS specific effects, macrophages from the EB myeloid differentiation phase (equivalent to day31)
overexpressed metallothionein and matrix metalloproteinase genes after LPS stimulation (e.g. MT1E and MMP10), along with increased IL-10 protein levels (Fig. 4c, d). Macrophages after 7 days in GM-CSF showed transcriptomic upregulation of CCL1 and CXCL9 cytokines with LPS (Fig. 4c). Macrophages after 7 days in GM-CSF + IL-34 upregulated the inflammesome complex member NLRP3, which is involved in the cleavage and activation of IL-1B. Accordingly, IL1B was upregulated both at the RNA and protein levels in this macrophage subset upon LPS stimulation (Fig. 4c, d).

Finally, we tested whether using a defined medium (StemPro34 serum-free media, SP-SFM) at this step would prevent the
transient macrophage activation. TF activity analysis showed SP-SFM induced similar activation signals (e.g. JUN, FOS and NFKB1). However, most TF activities were only partially recovered by day+7, including PPARG activity which did not significantly decrease from day+1 level. One additional difference was the maintained activity of the SREBF1 and SREBF2 TFs in SP-SFM, which link lipid metabolism to macrophages' inflammatory response (Supplementary Fig. 7A and Supplementary Data 6). These findings showed how media affects macrophage metabolism and function, which can expand the use of this model42,43.

**GM-CSF + FLT3L iPSC differentiation produces dendritic cells.** Conventional dendritic cells (cDC) present antigens to T cells and act as messengers between innate and adaptive immunity48. Protocols to induce DC differentiation in vitro are based on supplementing factors, including GM-CSF and FLT3L, that act cooperatively on cell precursors to drive cDC generation49,50. To produce iPSC-derived DCs we used GM-CSF + FLT3L in the EB myeloid differentiation phase (instead of M-CSF + IL3) and GM-CSF + IL4 in the last phase of differentiation on non-adherent cells from day31 (instead of M-CSF) (Fig. 5a). We then performed single-cell transcriptomic analysis and annotated cells using LR classifiers that were trained on gastrulation23, YS5 as well as foetal liver and thymus21,22 data sets (Fig. 5a, Supplementary Fig. 8 and Supplementary Data 8).

GM-CSF + FLT3L + IL4 stimulation produced DCs with a transcriptomic profile similar to cDC2 and a small DC population resembling CCR7 + DCs. These populations were exclusive to this protocol and were not found in the Discovery data set (M-CSF + IL-3 followed by M-CSF, Fig. 1c). All other cell types in the myeloid and stromal cell compartments, and their temporal dynamics (Supplementary Fig. 9A), remained the same as the macrophage protocol (Fig. 5b and 1c). Marker gene expression analysis supported the cell type annotations, and cDC2 expressed bona fide DC markers (e.g., HLA-DR, CD1C, CLEC10A)48 (Fig. 5c). The identity of cDC2 was further explored using LR models built on adult tonsil cDC2s and two subsets of monocye-derived DCs (moDCs) from adult ascites51. To match the cell types present in the adult data set we created a mixed manifold including the samples from day31 to day31 + 7 from both macrophages and DC protocols. The tonsil cDC2s LR model showed the highest prediction probability on the cells identified as cDC2s on the mixed model (tonsil cDC2 model AUC 0.837, Supplementary Fig. 9B, C), thus supporting our annotation. One of the moDCs subsets from ascites, which showed blood cDC2 features, overlapped with our in vitro cDC2 cells with a lower AUC (blood cDC2-moDC AUC 0.756 Supplementary Fig. 9B, C). This result demonstrated our in vitro cells have cDC2 features, although we cannot fully discard a moDC identity. The small subpopulation annotated as CCR7 + DCs showed a strong and specific prediction using a foetal thymus21 CCR7 DC LR model (Supplementary Fig. 8) and the expected marker expression (e.g., CCR7, LAMP3 and lack of CD14; Fig. 5b, c and Supplementary Fig. 9D). This result showed the ability of iPSC-derived DCs to adopt known activated DC profiles described in vivo52. Finally, cells obtained with the DC protocol showed dendrite-like structures in contrast to iPSC-derived macrophages (Supplementary Fig. 9E).

As in the presence of M-CSF + IL-3 (macrophage protocol) (Fig. 2f, g), myeloid cells appearing in the presence of GM-CSF + FLT3L (DC protocol) activated similar TFs when compared to their YS and foetal liver in vivo counterparts (Fig. 5d and Supplementary Data 9). Moreover, in vitro cDC2s activated PU.1 (SPI1 gene)53 and KLF454 TF networks relevant for in vivo cDC2 identity (Fig. 5d). We also observed increased RFX5 activity, which regulates MHC II gene expression55. A recent study postulated a role for CEBPB in the control of DC maturation and later stages of DC commitment56. Our results showed reduced CEBPB activity in cDC2 cells compared to monocytes (in vivo and in vitro, Fig. 5d), which indicates that the in vitro phenotype shares features with a functionality mature DC subset characterised by upregulation of costimulatory and MHC class II molecules.

We analysed step 3 of the DC protocol in detail, from the non-adherent cells produced during EB myeloid differentiation until the end of DC differentiation phase. A mean of 47.5% (standard deviation = 3.67) of the cells produced in the last three time points were cDC2 cells (Fig. 5f). Thus, this differentiation was less efficient than the macrophage protocol, where macrophages represented 94.1% of cells by day31 + 7 (Fig. 3c). In contrast to what was observed with macrophages, the proportion of cDC2 remained stable (Figs. 3c and 5f). Regarding the TF activity analysis, the in vitro activation of cDC2 and macrophages induced shared regulatory programmes, including activation of NFKB1, RELA or JUN on day+1 of this phase (Figs. 3c and 5e and Supplementary Data 9). Interestingly, the particular profile of TF networks induced in cDC2 on that day (including JUN, REL, SPI1 and HIF1A) did not clearly return to basal levels by day31 + 7 (Fig. 5f and Supplementary Data 9). This was in contrast to what was observed in the macrophage differentiation phase (Fig. 5d and Supplementary Data 6).

To confirm the DC phenotype of the iPSC-derived cells in our new protocol, we analysed canonical DC markers by FACS and functionally interrogated the cells using an antigen processing (DQ-OVA) and T cell activation assays. We observed positive cells for DC markers (CD1C, CD209, CD11c, HLA-DR, CD86) and low levels of the macrophage marker CD14, thus validating the DC phenotype46 (Fig. 5g and Supplementary Fig. 9F). Functionally, DQ-OVA antigen processing is higher in adult moDCs at earlier time points (15 to 30 min), whereas cDC processing capacity peaks at 60 min57. Accordingly, iPSC-derived cDC2 showed DQ-OVA
**Fig. 4 LPS stimulation of distinct subtypes of macrophages.**

a  Schematic illustration of the in vitro differentiation protocol and cell-type annotation analysis, steps before the macrophage differentiation phase are hidden. Additional experimental conditions (LPS stimulation at distinct time points) are highlighted in red. UMAP projections of scRNAseq analysis of LPS stimulated samples and matched controls for 4 populations of macrophages. (top) Dot plot of genes overexpressed >3 log fold-change in either of the 4 LPS stimulated samples vs their control, (bottom) Dot plot of genes significantly overexpressed in 1 of the 4 populations analysed. Cytokines protein expression levels in supernatants after LPS stimulation (n = 2) and controls (n = 1) for the 4 experimental conditions analysed. Media was collected from samples processed for scRNAseq (shown in b, c). Source data are provided in the Source data file.
processing at longer time points only (45 and 60 min) (Fig. 5h). While this time distinction is not commonly used to distinguish moDCs from cDC2s, it shows that the antigen processing behaviour of our iPSC-derived DCs resembles that of cDC2s. Finally, the cells produced at the end of the DC and macrophage differentiation were co-cultured with human CD4+ T cells purified from peripheral blood mononuclear cells. iPSC-derived DCs showed a stronger ability to induce T cell proliferation (Fig. 5i). These results demonstrated the DCs generated in our newly developed protocol were transcriptionally and functionally similar to in vivo cDC2, and could be used as a model to study DC function in vitro.
GWAS immune phenotype genes shape iPSC-derived myeloid cells. Dysfunctional myeloid differentiation and signalling can lead to immune-related disorders. To dissect potential myeloid contributions involved in these pathologies, we selected four genes linked to immune-related GWAS hits (ICAM1, LSP1, PRKCB and ZEB2) based on the existing literature. The ICAM1 and LSP1 loci contain SNPs linked to autoimmune inflammatory diseases by GWAS (https://www.ebi.ac.uk/gwas/home) and interact with each other\(^{19}\). PRKCB is a protein kinase associated with inflammatory diseases and blood cell counts in GWAS and is involved in myeloid DC differentiation\(^{60}\). Finally, ZEB2 is found in GWAS for blood phenotypes, and regulates hematopoiesis in mice\(^{61}\) and DC cell fate decisions\(^{62}\). To study their involvement in myeloid differentiation and identity, we generated knock-out (KO) iPSC lines using CRISPR/Cas9 in one of our donor cell lines (kolf_2) (Supplementary Fig. 10).

KO iPSC lines were differentiated into macrophages and DC alongside wild-type (WT) isogenic lines. Single-cell transcriptomic analyses were performed at day0 (IPS stage) and day31 (EB myeloid differentiation phase) (Fig. 6a, b). Though we observed all cell populations in all conditions (Fig. 6c, d), some of the KOs affected the cell type proportions (Supplementary Fig. 11A). As expected, knocking out ZEB2 reduced myeloid cells to 5.5% versus 65.6% in WT lines in the macrophage protocol (12-fold decrease), and 2.1% versus 20.4% in WT lines in the DC protocol (10-fold decrease) (Supplementary Data 10). The overall reduction in cell numbers in the ZEB2 KO vs the other lines suggested that ZEB2 absence either blocked myeloid differentiation or induced apoptosis as described for myeloid leukaemic cells\(^{63}\). On the contrary, PRKCB KO increased the proportion of myeloid cells in the DC differentiation protocol (20.4% in WT versus 78% in PRKCB KO, fourfold increase) (Supplementary Data 10). The other KOs did not markedly influence myeloid cell proportions.

Some myeloid cell types showed transcriptomic differences between KO lines. LSP1 and ICAM1 KOs monocytes from the DC protocol had a transcriptomic profile with intermediate monocytes features. They were characterised by the upregulation of the HLA genes and downregulation of the S100 gene family\(^{64}\) (Fig. 6e). Interestingly, KO monocytes from the macrophage protocol did not show this profile (Supplementary Fig. 11B). An increased population of intermediate monocytes has been described in many autoimmune diseases such as active Crohn’s disease\(^{65}\) and rheumatoid arthritis\(^{66}\). Monocytes from the PRKCB KO cell line after the DC differentiation protocol had a myeloid-derived suppressor profile, including low expression of HLA-DR and CD74 with high levels of CCL2 and MMP9\(^{77-78}\) (Fig. 6f). This was consistent with the observation that myeloid-derived suppressor cells have decreased levels of PRKCB, which dampens DC differentiation and function in vivo\(^ {60}\).

Macrophages generated from iPSCs deficient in PRKCB, LSP1 or ICAM1 exhibited a mixed anti-inflammatory and anti-fibrotic phenotype. KO macrophages upregulated the suppressors of the NFκB-dependent inflammatory pathway KLF3\(^ {71}\) and ATF3\(^ {72}\) (Fig. 6g). They also decreased the expression of M2 profibrotic phenotype genes (e.g. FN1, GRN and SPP1, Supplementary Data 11), as well as decreased activity of M2-promoting transcription factors (e.g. MAF\(^ {34}\) and PPAR\(^ {23}\), Fig. 6g). We observed a connection between PRKCB, LSP1 and ICAM1, as ICAM1 was downregulated in PRKCB and LSP1 KOs (Supplementary Data 11). This suggested that the 3 genes are part of a regulatory network that fine-tunes macrophage phenotype and represses tissue healing both by promoting REL/p65-mediated inflammation and controlling the expression of profibrotic M2 genes. Notably, these KO lines generated a macrophage population with regenerative potential due to the concomitant suppression of the tissue remodelling programmes (Fig. 6g) and the NFκB pathway. Our data illustrate how such broad defects in macrophage polarisation could impair the proper resolution of inflammation, leading to in vivo autoimmune inflammatory disorders\(^ {74}\).

Altogether, we found deletion of genes associated with autoimmunity mediated early hematopoiesis (e.g. ZEB2) and influenced the inflammatory potential (e.g. ICAM1, LSP1 and PRKCB) of iPSC-derived myeloid cells, coinciding with the expected phenotype. Therefore, iPSC-differentiation protocols can be powerful tools to functionally interrogate GWAS hits and other genes of interest in vitro.

**Discussion**

The full characterisation and assessment of the robustness, accuracy and efficiency of in vitro protocols are essential to utilising them as models for disease as well as leveraging them in the search for novel therapeutic targets. Myeloid cells are central to immunity and participate in major inflammatory and autoimmune disorders. Experimental protocols to generate macrophages that are easy to replicate and amenable to scaling up are paramount to studying human macrophage ontogeny, genetics and function in health and disease. Here, we profiled more than 470k single cells across a commonly used, straightforward differentiation process from human iPSC to terminally differentiated macrophages. We reconstructed, using cell trajectories, the in vitro sequence of events, which parallel foetal hematopoiesis prior to the establishment of HSC. Moreover, we show this protocol is a valuable resource in studying different macrophage cell states, multiple myeloid...
populations, including erythrocytes, megakaryocytes and mast cells that spontaneously arise, as well as DC, induced by adjusting the media composition. Finally, we used this model to interrogate the functional effect of genes associated with inflammatory and autoimmune disorders and interpreted the results in relation to their in vivo counterparts.

To quantitatively assess the accuracy of our in vitro models, we used machine learning tools. We built LR models trained on scRNAseq data from developmental atlases mapping the formation of the immune system and projected them onto the in vitro data sets. The computational framework we have established could be adapted to annotate multiple iPSC differentiation protocols.
Following this strategy, we found that the initial phases of iPSC–macrophage differentiation faithfully recapitulate YS foetal hematopoiesis and generate foetal-like FOLR2+ macrophages. The lack of an HSC cluster in our data, the activation of master regulator RUNX1 in the endothelial-to-hematopoietic transition, and the lack of expression of HOXA genes in the myeloid progenitors all suggest the iPSC-derived macrophages protocol closely recapitulate YS differentiation prior to the establishment of definitive hematopoiesis. Thus, we propose this model as a unique system for interrogating the early stages of hematopoietic differentiation in humans, which are largely unexplored.

Our study shows that iPSC–macrophage differentiation generates a wide range of myeloid cells and presents a detailed list of TFs that mediate the generation of distinct myeloid progenitors in vitro. We also observed that EB stimulation with GM-CSF and FLT3L produces cell populations resembling cDC2. However, these in vitro DCs may also be related to monocytes, due to the presence of monocytes in the culture and the resemblance with an in vivo moDC subset, which shows features of blood cDC2. Additionally, a small subpopulation of DCs in the culture recapitulates the CCR7+ in vivo activation state2, highlighting the functional potential of iPSC-derived DCs. Nevertheless, the protocol for DC generation had a lower efficiency than the macrophages protocol, which is consistent with the known bias of YSMP towards macrophages.

Interestingly, we found erythrocytes in our data following a decline in RUNX1 and SPI1 TF activity and a rise in GATA1 expression of genes associated to myeloid-derived suppressor cells in the monocytes produced by each KO and the wild type in the iPSC-to-DC protocol. The average expression of genes associated to myeloid-derived suppressor cells in the monocytes produced by each KO and the wild type in the iPSC-to-DC protocol.

Methods
Reagent/material details can be found in the Supplementary Information.

Human induced pluripotent stem cell lines. All iPSC lines used in the study were generated by the HIPSICI project. Details on their generation are available at http://www.hipsci.org. Briefly, we used kolf_2, yemz_1 and vass_1 in the Discovery and DC data sets, and we added ceik_1, esb_1 and wegi_1 for the Validation data sets. All cells in the KO data set are derived from kolf_2 as a parental line, which was subcloned prior to editing (kolf_2_C1). All HIPSICI samples were collected from consenting research volunteers recruited from the NIHR Cambridge BioResource (http://www.cambridgebioresource.org.uk), initially under existing ethics rules for iPSC derivation (Regional Ethics Committee (REC) reference 09/H0304/77, v.2, 4
In vitro differentiation to macrophages and dendritic cells. We used an adapted version of the method to generate human iPSC lines were cultured in E8 (StemCell Technologies) on vitronectin-coated plates (Life Technologies). For the embryoid body (EB) formation, step 1, a single-cell suspension of hiPSC was plated in 100 µl of EB medium – E8 + SCF (20 ng/ml, PeproTech) + VEGF (50 ng/ml, PeproTech) + BMP-4 (50 ng/ml, PeproTech) + ROCK inhibitor (10 µM, Sigma) – at a density of 10,000 cells per well in round bottomed 96-well plates (Corning). After 2 days, we changed half the media (50 µl) and replaced it with fresh EB media. At day 4, EB myeloid differentiation started, step 2, when EBs were plated in gelatin-coated 6-well plates (Sigma Aldrich) at a density of 8 × 10^5 cells per well of a 6-well plate. For the media experiment (Fig 3b and Supplementary Fig. 7A), fully differentiated EBs were plated at 6 × 10^5 cells per well of a 6-well plate. For the media experiment, the EB medium was changed every 4 to 5 days (days 1, 3, 5, non-adherent cells were collected by centrifugation from the culture media and 1 × 10^6 cells were cultured in 10 cm tissue culture plates for 7 days in macrophage differentiation medium – RPMI (ThermoFisher) + 10% heat-inactivated FBS + M-CSF (100 ng/ml, ThermoFisher).

Alternative macrophage differentiation media were used in the macrophage differentiation phase, step 3. For the cytokines experiment (Fig 3b-d), we used RPMI + 10% heat-inactivated FBS + GM-CSF (50 ng/ml, PeproTech) and RPMI + 10% heat-inactivated FBS + GM-CSF (10 ng/ml) – IL-3 (100 ng/ml, PeproTech), for the latter cells were plated at 6 × 10^5 cells per well of a 6-well plate. For the media experiment (Fig 3b and Supplementary Fig. 7A), fully differentiated medium – StemPro-34 + M-CSF (100 ng/ml) - was used.

Step 1 is shared between macrophages and DCs, while different cytokines are used in steps 2 and 3. For DC differentiation EBs at day4 are plated with EB-DC media – StemPro-34 + M-CSF (50 ng/ml) + IL-7L (100 ng/ml, PeproTech) in the same type of plates and density as the macrophage protocol. At day3, step 3, non-adherent cells were collected and plated in 10 cm tissue culture plates in DC differentiation medium – RPMI + 10% heat-inactivated FBS + GM-CSF (50 ng/ml) + IL-4 (100 ng/ml, PeproTech).

LPS stimulation of iPSC-derived cells. Four populations of cells derived from the macrophage protocol described above were stimulated with LPS (i.e. non-adherent cells from the EB myeloid differentiation phase -MacDay0-, cells at the end of the Mac differentiation phase incubated with LPS for 5 days with M-CSF -MacDay5 M-CSF-, with GM-CSF (50 ng/ml) -MacDay7 GM-CSF- and with GM-CSF (10 ng/ml) – IL-3 (100 ng/ml, MacDay7 GM-CSF, IL-3-)). For each sample, LPS (Sigma Aldrich) was added to the media in two wells for a final concentration of 2.5 µg/ml while a third well was kept as control. After 2 h, Brefeldin A (Sigma Aldrich) was added to reach 5 µg/ml to all wells and cells were incubated for an additional 4 h (total 6 h of stimulation with LPS). Media was collected for control and LPS wells for cytokine analysis using V-PLEX Proinflammatory Panel 1 Human Kit (Mesoscale) according to manufacturing instructions. In parallel, cells were collected using 10 mg/ml LPS (Sigma Aldrich) and 2 mM EDTA (ThermoFisher) solution for 5 min at 57 °C, the two wells of LPS stimulation per population were combined. Single-cell RNAseq analysis was performed aiming at 3000 cells per condition as described in ‘10X’ Genomics Chromium GEM sample preparation and sequencing’.
Marker protein and antigen processing DQ-OVA assay analysis by FACS. Macrophages and dendritic cells were detached from 10 cm plates using 10 mg/ml Lidocaine 2 mM EDTA solution for 5 min at 37 °C, collected in DPBS and spun down at 300 x g for 3 min. Samples were then fixed with BD Cytofix/Cytoperm buffers (ThermoFisher) for 20 min at room temperature and washed with DPBS + 1% FBS. Staining with fluorescent-labelled primary antibodies (see Supplementary Information for antibody details) was performed in the dark at room temperature for 30 min. After 2 washes with DPBS + 1% FBS, cells were analysed by FACS in a BD LSR Fortessa II and data was analysed using FlowJo v10. The gating strategy is provided in Supplementary Fig. 4B.

T cell activation assay by CFSE staining and FACS. Twenty thousand iPSC-derived non-adherent cells from the macrophage or DC differentiation protocol were differentiated to DC (in presence of GM-CSF and IL-4) or macrophages (in the presence of M-CSF) for 7 days. This was performed as described for the last phase of differentiation in the protocols used throughout the study but in round-bottom 96 well plates. In parallel, peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll-Paque PLUS (GE Healthcare) density gradient centrifugation. Naïve T cells were isolated from the PBMC fraction using EasySep® human naïve T cell isolation kits II (StemCell Technologies) according to the manufacturer’s instructions. After the 7 days of DC and macrophage differentiation described above, 100,000 naïve T cells, stained with CFSE (ThermoFisher), were stimulated with anti-CD3 (Tonbo Biosciences, 70-0037-U100, 100 ng/μg) and anti-CD28 (StemCell Technologies, 10971, 6μg/ml/1 millinlons). After 4 days of co-culture, T cell proliferation was measured by assessing CFSE dilution. Anti-CD3 alone was tested independently and had no effect on T cell proliferation rate. The experiment was performed in two independent experiments using two PBMC donors each time in one of the iPSC lines (kolf_2). The gating strategy is provided in Supplementary Fig. 9C.

CRISPR-Cas9 KO of human induced pluripotent stem cell lines. KO iPSC lines were generated by substituting an asymmetrical exon with a Puromycin cassette and expanding those clones with a frame-shift indel in the remaining allele. A hSpCas9 were generated by substituting an asymmetrical exon with a Puromycin cassette and CRISPR-Cas9 KO of human induced pluripotent stem cell lines using two PBMC donors each time in one of the iPSC lines (kolf_2). The gating protocol was provided in Supplementary Information for antibody details) was performed in the dark at room temperature for 30 min. After 2 washes with DPBS + 1% FBS. Staining with fluorescent-labelled primary antibodies (see Supplementary Information for antibody details) was performed in the dark at room temperature for 30 min. After 2 washes with DPBS + 1% FBS, cells were analysed by FACS in a BD LSR Fortessa II and data was analysed using FlowJo v10. The gating strategy is provided in Supplementary Fig. 4B.

KO iPSC-derived cell types analysis. Transcriptomic alterations between cell types arising in WT and KO lines were assessed using differential expression from Seurat. Genes present in 10% of the cells and with a minimal log fold change of 0.25 were selected for differential expression analysis of each cell type in each KO line vs their WT counterpart. Only genes with an FDR <0.05 were considered as significantly differentially expressed. Differential amount of cells produced was calculated using two-sided t test between ZER2 KO number of cells collected for the two clones used vs the number of cells collected for all other lines.

Received: 7 July 2021; Accepted: 6 May 2022; Published online: 24 May 2022

References
1. Okabe, Y. & Medzhitov, R. Tissue biology perspective on macrophages. Nat. Immunol. 17, 9–17 (2016).
2. Hoeksema, M. A. & Glass, C. K. Nature and nurture of tissue-specific macrophage phenotypes. Atherosclerosis 281, 159–167 (2019).
3. Lavin, Y., Mortha, A., Rahman, A. & Merad, M. Regulation of macrophage development and function in peripheral tissues. Nat. Rev. Immunol. 15, 731–744 (2015).
4. Gordon, S. & Plaßdamm, A. Tissue macrophages: heterogeneity and functions. BMC Biol. 15, 53 (2017).
5. Hoefel, G. & Ginhoux, F. Fetal monocytes and the origins of tissue-resident macrophages. Cell. Immunol. 330, 5–18 (2015).
6. Bian, Z. et al. Deciphering human macrophage development at single-cell resolution. Nature 582, 571–576 (2020).
7. Kaczanowska, S. et al. Genetically engineered myeloid cells rebalance the core immune suppression program in metastasis. Cell 184, 2033–2052.e21 (2021).
8. Frame, J. M., McGrath, K. E. & Palis, J. Erythroid-myeloid progenitors: definitive hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. Blood Cells Mol. Dis. 51, 220–225 (2013).
9. Ivanovs, A. et al. Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. J. Exp. Med. 208, 2417–2427 (2011).
10. Holt, P. G. & James, C. A. The development of the immune system during pregnancy and early life. Allergy 55, 688–697 (2000).
11. Freyer, L., Iituri, L., Biton, A. & Perdiguerio, E. O. Overlapping definitive progenitor waves divide and conquer to build a layered hematopoietic system. Preprint at bioRxiv https://doi.org/10.1101/2020.12.24.424392 (2020).
12. Lee, C. Z. W., Kozaki, T. & Ginhoux, F. Studying tissue macrophages in vitro: are iPSC-derived cells the answer? Nat. Rev. Immunol. 18, 716–725 (2018).
13. Alasoo, K. et al. Transcriptional profiling of macrophages derived from monocytes and iPSC cells identifies a conserved response to LPS and novel alternative transcription. Sci. Rep. 5, 12324 (2015).
14. Cao, X. et al. Differentiation and functional comparison of monocytes and macrophages from hiPSCs with peripheral blood derivatives. Stem Cell Rep. 12, 1282–1297 (2019).
15. van Wilgenburg, B., Browne, C., Vowles, J. & Cowley, S. A. Efficient, long term production of monocyte-derived macrophages from human pluripotent stem cells under partly-defined and fully-defined conditions. PLoS ONE 8, e71098 (2013).
16. Gubser, T. et al. Large-scale production of human iPSC-derived macrophages for drug screening. Int. J. Mol. Sci. 21, 4808 (2020).
17. Camp, J. G., Wollny, D. & Truelein, B. Single-cell genomics to guide human stem and tissue engineering. Nature Methods 15, 661–667 (2018).
18. Balan, S. et al. Large-scale human dendritic cell differentiation revealing Notch-dependent lineage bifurcation and heterogeneity. Cell Rep. 24, 1902.e6–1915.e6 (2018).
19. Garcia-Alonso, L. et al. Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. Nat. Genet. 53, 1698–1711 (2021).
20. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. Nature 563, 347–353 (2018).
21. Park, J.-E. et al. A cell atlas of human thymic development defines T cell repertoire formation. Science 367, eaay3224 (2020).
22. Popescu, D.-M. et al. Decoding human fetal liver haematopoiesis. Nature 574, 365–371 (2019).
23. Tyser, R. C. V. et al. Single-cell transcriptomic characterization of a gastrulating human embryo. Nature 600, 285–289 (2021).
24. Stuart, T. et al. Comprehensive integration of single-cell data. Cell 177, 1888.e21–1902.e21 (2019).
25. Dou, D. R. et al. Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat. Cell Biol. 18, 595–606 (2016).
26. Jung, H. S. et al. SOX17 integrates HOXA and arterial programs in hemogenic endothelium to drive definitive lympho-myloid hematopoiesis. Cell Rep. 34, 108758 (2021).
27. Sharma, A. et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. Cell 183, 1511.e21–394.e21 (2020).
28. Álvarez-Errico, D., Vento-Tormo, R., Sieweke, M. & Ballestar, E. Epigenetic control of myoid cell differentiation, identity and function. Nat. Rev. Immunol. 15, 7–17 (2015).
29. Burgin, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. 38, 1408–1414 (2020).

Data availability
The single-cell sequencing fastq files are available in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession numbers E-MTAB-11623 for single-cell RNAseq and E-MTAB-11616 for single-cell ATACseq. Processed data sets can be queried and downloaded through the web portal www.HiPImmunomeatlas.org. Single-cell sequencing data was mapped against GRCh38 http://ftp.ensembl.org/pub/release-100/. Publicly available data sets used include Gastrulation http://www.human-gastrula.net/, Foetal liver (+ kidney + skin) E-MTAB-7407, Foetal thymus E-MTAB-4851, Placenta E-MTAB-6761, HCC https://doi.org/10.17632/6wmzcskt6k.1, Yolk sac GSE133345 and adult DCs and Macs GSE115006. Source data are provided with this paper.

Code availability
All relevant codes used for data analysis are available from https://github.com/Ventolab/iPSCmyeloid.
30. Holland, C. H. et al. Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. Genome Biol. 21, 36 (2020).

31. Marcello, K. L., Goldie, L. C. & Hirsch, K. X. Regulation of endothelial cell differentiation and specification. Circ. Res. 112, 1272–1287 (2013).

32. Stehling-Sun, S., Dade, J., Nutt, S. L., DeKoter, R. P. & Camargo, F. D. Regulation of lymphoid versus myeloid fate ‘choice’ by the transcription factor Mef2c. Nat. Immunol. 10, 289–296 (2009).

33. Clausen, B. E. et al. Residual MHC class II expression on mature dendritic cells and activated B cells in REX5-deficient mice. Immunity 8, 143–155 (1998).

34. Kang, K. et al. Interferon-γ represses M2 gene expression in human macrophages by disassembling enhancers bound by the transcription factor MAF. Immunity 47, 235.e4–250.e4 (2017).

35. Heisterkamp, K. et al. The combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell. 38, 576–589 (2010).

36. Schneider, C. et al. Induction of the nuclear receptor PPAR-γ by the cytokine GM-CSF is critical for the differentiation of fetal monocytes into alveolar macrophages. Nat. Immunol. 15, 1026–1037 (2014).

37. Yu, X. et al. The cytokine TGF-β promotes the development and homeostasis of alveolar macrophages. Immunity 47, 903.e4–912.e4 (2017).

38. Liu, B. et al. An entropy-based metric for assessing the purity of single cell populations. Nat. Commun. 11, 3155 (2020).

39. Garcia-Alonso, L., Holland, C. H., Ibrahim, M. M., Turei, D. & Saez-Rodriguez, J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genes Dev. 29, 1363–1375 (2015).

40. Hume, D. A. The many alternative faces of macrophage activation. Front. Immunol. 6, 370 (2015).

41. Sullivan, A. L. et al. Serum response factor utilizes distinct promoter- and enhancer-based mechanisms to regulate cytoskeletal gene expression in human macrophages. Mol. Syst. Biol. 31, 861–875 (2017).

42. Im, S.-S. et al. Linking lipid metabolism to the innate immune response in human monocytes. Trends Immunol. 13, 540–549 (2011).

43. Guo, C. et al. Cholesterol homeostatic regulator SCAP-SREBP2 integrates NLR3 inflammasome activation and cholesterol biosynthetic signaling in macrophages. Immunity 49, 842–856 (2018).

44. Hamilton, T. A., Zhao, C., Pavicic, P. G. Jr. & Datta, S. Myeloid colony-stimulating factors as regulators of macrophage polarization. Front. Immunol. 5, 554 (2014).

45. Wang, Y. et al. IL-34 is a tissue-restricted ligand of CSFIR required for the development of Langerhans cells and microglia. Nat. Immunol. 13, 753–760 (2012).

46. Ohgihara, M. et al. Indirect induction of ramified microglia-like cells from human monocytes: dynamic microglial dysfunction in Nasu-Hakola disease. Sci. Rep. 4, 4957 (2014).

47. Ji, J. et al. Antagonizing peroxisome proliferator-activated receptor γ facilitates M2 to M1 shift of microglia by enhancing autagy via the LKB1-AMPK signaling pathway. Aging Cell 17, e12774 (2018).

48. Cabeza-Cabrerizo, M., Cardoso, A., Minutti, C. M., Pereira da Costa, M. & Faria, E. The transcription factor C/EBPα promotes conventional dendritic cell differentiation and trained innate immunity. Cell Res. 21, 127–139 (2011).

49. Li, J. et al. The EMT transcription factor Zeb2 regulates development of conventional and plasmacytoid DCs by repressing epithelial lineage genes. PLoS One 8, e58501 (2013).

50. Hossain, M., Qadir, S. M., Su, Y. & Liu, L. ICAM-1-mediated leukocyte adhesion is critical for the activation of endothelial LSP1. Am. J. Physiol. Cell Physiol. 304, C895–C904 (2013).

51. Farren, M. R. et al. Tumor-induced STAT3 signaling in myeloid cells impairs dendritic cell generation by decreasing PKCβII abundance. Sci. Signal. 7, ra16 (2014).

52. Li, J. et al. The EMT transcription factor Zeb2 controls adult murine hematopoietic differentiation by regulating cytokine signaling. Blood 129, 460–472 (2017).

53. Scott, C. L. et al. The transcription factor Zeb2 regulates development of conventional and plasmacytoid DCs by repressing Id2. Idy. J. Exp. Med. 213, 897–911 (2016).

54. Fardi, M., Mohammadi, A., Baradaran, B. & Safaei, S. ZEB2 knock-down induces apoptosis in human myeloid leukemia HL-60 cells. Carr. Gene Ther. 21, 149–159 (2021).

55. Wang, K. L. et al. Gene expression profiling reveals the defining features of the classical, intermediate, and nonclassical human monocyte subsets. Blood 118, e16–e31 (2011).

56. Griep, O., Bredberg, A., Lindgren, S. & Henriksson, G. Increased subpopulations of CD16 and CD56 blood monocytes in patients with active Crohn’s disease. Inflamm. Bowel Dis. 13, 566–572 (2007).

57. Rossol, M., Kraus, S., Pierer, M., Baerwald, C. & Wagner, U. The CD14brightCD16 monocyte subset is expanded in rheumatoid arthritis and promotes expansion of the Th17 cell population. Arthritis Rheum. 64, 861–872 (2012).

58. Menges, A. E., Gastineau, D. A. & Gustafson, M. P. The CD14+HLA-DRlow/−neg monocyte: an immune-suppressive phenotype that restrains responses to cancer immunotherapy. Front. Immunol. 10, 1147 (2019).

59. Alban, T. J. et al. Gluloblastoma myeloid-derived suppressor cells express differential macrophage migration inhibitory factor receptor profiles that can macrophages. Mol. Syst. Biol. 31, 1061 (2017).

60. Kumar, V., Patel, S., Tcygnow, E. & Gabrilovich, D. I. The nature of myeloid-derived suppressor cells in the tumor microenvironment. Trends Immunol. 37, 208–220 (2016).

61. Fleming, V. et al. Targeting myeloid-derived suppressor cells to bypass tumor-induced immunosuppression. Front. Immunol. 9, 398 (2018).

62. Knigts, A. J. et al. Kriippel-like factor 3 (KLF3) suppresses NF-κB-driven inflammation in mice. J. Biol. Chem. 295, 6080–6091 (2020).

63. Zhao, Q. et al. Activating transcription factor 3 involved in Pseudomonas aeruginosa PAO1-induced macrophage senescence. Mol. Immunol. 133, 122–127 (2021).

64. Bruewer, F. F. et al. Human yolk sac-like haemopoietesis generates RUNX1-GFI1- and/or GFI1B-dependent blood and SOX17-positive endothelium. Development 147, dev19307 (2020).

65. Willcockson, M. A. et al. Runx1促进 murine erythroid progenitor proliferation and inhibits differentiation by preventing Pu.1 downregulation. Proc. Natl. Acad. Sci. USA 116, 17841–17847 (2019).

66. Ludwig, L. S. et al. Transcriptional states and chromatin accessibility underlying human erythropoiesis. Cell Rep. 27, 3228.e7–3240.e7 (2019).

67. Ludwig, L. S. et al. Altered translation of GATA1 in Diamond-Blackfan anemia. Nat. Med. 20, 748–753 (2014).

68. Bergen, V., Solodovtov, R. A., Kharchenko, P. V. & Theis, F. J. RNA velocity—current challenges and future perspectives. Mol. Syst. Biol. 17, e10282 (2021).

69. Lavin, Y. et al. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell 159, 1312–1326 (2014).

70. Gosselin, D. et al. Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. Cell 159, 1327–1340 (2014).

71. Saaed, S. et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science 345, 1251086 (2014).

72. Gosselin, D. et al. An environment-dependent transcriptional network specifies human macrophage identity. Science 356, eaal3222 (2017).

73. Schulz, C. et al. A lineage of myeloid cells independent of Myb and hematopoietic stem cells. Science 336, 86–90 (2012).

74. Sakai, M. et al. Liver-derived signals spontaneously reprogram myeloid enhancers to initiate and maintain Kupffer cell identity. Immunity 51, 655–670.e8 (2019).

75. Ydens, E. et al. Profiling peripheral nerve macrophages reveals two macrophage subsets with distinct localization, transcriptome and response to injury. Nat. Neurosci. 23, 676–689 (2020).

76. Al-Atrash Jiee, O., Ayyangar, U., Kurbit, A. S., Aasok, D. & Baghavan, S. Unraveling the ECM-Immune Cell Crosstalk In Skin Diseases. Front. Cell Dev. Biol. 7, 68 (2019).
88. Dorrier, C. E. et al. CNS fibroblasts form a fibrotic scar in response to immune cell infiltration. Nat. Neurosci. 24, 234–244 (2021).
89. Heaton, H. et al. Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. Nat. Methods 17, 615–620 (2020).
90. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst. 8, 281.e9–291.e9 (2019).
91. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq computational analysis. We thank Jana Eliasova for her help with the figure design and Christina Usher and Aidan Maartens for their edits in the text. This work was mainly funded by the Open Targets consortium (OTAR026 and OTAR032 project) and the Wellcome Sanger core funding (WT206194) with additional support from Open Targets projects OTAR037, OTAR2065, OTAR2071. The authors are grateful to the funders for their support and additional care given to their members during the COVID-19 pandemic. D.A.-E. thanks CERCA Programme/Generalitat de Catalunya and the Josep Carreras Foundation for institutional support. This study makes use of cell lines and data generated by the HiPSci Consortium, funded by The Wellcome Trust and the MRC (Medical Research Council). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. This publication is part of the Human Cell Atlas—www.humancellatlas.org/publications.

Author contributions
C.A., D.J.G., D.A.-E. and R.V.-T. conceived and designed the experiments and analyses; C.A. and M.N.P. performed the experiments with contributions from E.B., I.K., R.V.-B., and C.S.-S.; C.P.I. performed and analysed the T cell activation experiments; A.J.K. designed and performed the ATAC single-cell sample processing; C.A. analysed the data with contributions from V.I. with the trajectories and knock-out analyses, from J.-E.P. with the thymus data set logistic regression and from B.S.W. and D.F.T. with the dendritic cells data set; C.A., R.V.-T. and D.A.-E. interpreted the data and wrote the manuscript with contributions from V.L., R.V.-B., G.T. and A.B.; D.J.G., D.A.-E. and R.V.-T. supervised the work. All authors read and approved the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-30557-4.

Correspondence and requests for materials should be addressed to Clara Alsinet, Daniel J. Gaffney, Damiana Alvarez-Errico or Roser Vento-Tormo.

Peer review information Nature Communications thanks Marc Dalod, Nico Lachmann and Manu Setty for their contribution to the peer review of this work.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022