nature immunology

Identification of cDC1- and cDC2-committed DC progenitors reveals early lineage priming at the common DC progenitor stage in the bone marrow

Andreas Schlitzer[1], V Sivakamasundari[2,6], Jinmiao Chen[1,6], Hermi Rizal Bin Sumatoh[1], Jaring Schreuder[3], Josephine Lum[1], Benoit Malleret[1,4], Sanqian Zhang[1], Anis Larbi[1], Francesca Zolezzi[1], Laurent Renia[1], Michael Poidinger[1], Shalin Naik[3,5], Evan W Newell[1], Paul Robson[2,6] & Florent Ginhoux[1,6]

Mouse conventional dendritic cells (cDCs) can be classified into two functionally distinct lineages: the CD8α$^+$ (CD103$^+$) cDC1 lineage, and the CD11b$^+$ cDC2 lineage. cDCs arise from a cascade of bone marrow (BM) DC-committed progenitor cells that include the common DC progenitors (CDPs) and pre-DCs, which exit the BM and seed peripheral tissues before differentiating locally into mature cDCs. Where and when commitment to the cDC1 or cDC2 lineage occurs remains poorly understood. Here we found that transcriptional signatures of the cDC1 and cDC2 lineages became evident at the single-cell level from the CDP stage. We also identified Siglec-H and Ly6C as lineage markers that distinguished pre-DC subpopulations committed to the cDC1 lineage (Siglec-H$^-$Ly6C$^-$ pre-DCs) or cDC2 lineage (Siglec-H$^-$Ly6C$^+$ pre-DCs). Our results indicate that commitment to the cDC1 or cDC2 lineage occurs in the BM and not in the periphery.

Dendritic cells (DCs) are antigen-sensing and antigen-presenting cells that are essential for effective immunity and tolerance. Many developmentally and functionally defined subsets of DC have been identified[1,2] with important and unique roles in the initiation of immune responses during infection with pathogens or auto-immunity, as well as during vaccination and cancer therapy.

DCs found in steady-state secondary lymphoid tissues, such as the spleen and lymph nodes, are called 'lymphoid-resident DCs', while non–lymphoid-tissue-resident DCs migrating to the lymph nodes are called 'migratory DCs'. Lymphoid-resident DC subsets in the spleen that express either the coreceptor CD8α or the coreceptor CD4 are defined as conventional DCs (cDCs), to distinguish them from plasmacytoid DCs (pDCs)[2], characterized by their unique ability to produce massive amounts of type I interferons in response to viral stimulation. Migratory DC populations can be defined by mutually exclusive surface expression of the integrins CD103 (α_E) and CD11b (α_M), although DC populations in the intestinal lamina propria that express both of these markers have been reported[1,2]. Genetic and functional studies have revealed that CD8α$^+$ and CD103$^+$ DCs represent a distinct cDC lineage (cDC1) functionally specialized in antigen cross-presentation, polarization into the T_H1 subset of helper T cells and secretion of interferon-λ in response to stimulation via Toll-like receptor 3, which emphasizes their crucial role in acting against intracellular pathogens. Furthermore, development of the cDC1 subset is dependent on the transcription factors IRF8, Id2 and Batf3. The CD4$^+$ and CD11b$^+$ DC subsets represent a separate cDC lineage (cDC2) specialized in the presentation of antigen to CD4$^+$ T cells and with the unique ability to favor polarization toward T_H2 or T_H17 responses, which emphasizes their importance during immune responses to extracellular pathogens[1,3].

The development of DC subsets, its transcriptional requirements and where and when transcriptional priming of the functional specialization of DC subsets occurs are not fully understood. DC development depends on the growth factor Flt3L ('fms-like tyrosine kinase 3 ligand') and its receptor Flt3 (CD135)[4], expressed on a continuum of Flt3$^+$ DC precursors that gradually differentiate in the BM. The macrophage DC progenitor (MDP)[5], which give rise to both monocytes and DCs, is thought to differentiate into the common DC progenitors (CDP), which has lost monocyte- and/or macrophage-differentiation potential and gives rise only to cells belonging to the DC lineage, including the cDC1 and cDC2 lineages and pDCs[6–8]. CDPs differentiate into pre-DCs, which give rise mainly to the cDC1 and cDC2 lineages after they migrate out of the BM into the blood and seed lymphoid and non-lymphoid tissues such as the spleen, lymph nodes and lungs[9,10].

Published studies have indicated that the pre-DC population might be heterogeneous[6]. *In vitro* clonal analysis of CDP potential has suggested that lineage priming toward the cDC1 and cDC2 subsets starts to emerge at this stage[6], through the use of differential expression of heat-stable antigen (CD24) to identify cDC1- and cDC2-biased

[1]Singapore Immunology Network, Agency for Science, Technology and Research, Singapore. [2]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore. [3]Molecular Medicine Division, Walter and Eliza Hall Institute of Medical Research, Parkville, Australia. [4]Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore, National University Health System, Singapore. [5]Department of Medical Biology, University of Melbourne, Parkville, Australia. [6]These authors contributed equally to this work. Correspondence should be addressed to F.G. (florent_ginhoux@immunol.a-star.edu.sg).

Received 4 February; accepted 18 May; published online 8 June 2015; doi:10.1038/ni.3200

pre-DCs in the spleen[6]. However, the extent to which the priming of cDC subsets is defined in a cell-intrinsic way in the BM precursors or is conferred in the tissue microenvironment following exit from the BM has remained unknown. Heterogeneity of progenitor subsets has been described for CDPs, among which two subsets can be identified on the basis of differential expression of the receptor CSF-1R for the cytokine CSF-1 ('colony-stimulating factor 1')[6,7,11]. CSF-1R$^+$ CDPs give rise mainly to the cDC1 and cDC2 subsets, while CSF-1R$^-$ CDPs give rise mainly to pDCs[11]. In addition, a fraction of BM pre-DCs express the marker Siglec-H together with the cDC-lineage transcription factor ZBTB46, and this expression correlates with the potential to give rise to all cDC subsets but not pDCs[12]. Together these observations suggest that the heterogeneity of DC progenitors may be underappreciated and such heterogeneity might reflect intrinsic priming toward the generation of specific DC subsets early during differentiation, in BM progenitors and before exposure to peripheral tissues.

To define the extent and nature of cellular heterogeneity of BM DC precursors, we performed single-cell mRNA sequencing of Lin$^-$CD11c$^-$MHCII$^-$CD135$^+$CSF-1R$^+$CD117^{hi} MDPs, Lin$^-$CD11c$^-$MHCII$^-$CD135$^+$CSF-1R$^+$CD117^{int} CDPs and Lin$^-$CD11c$^+$MHCII$^-$CD135$^+$CD172α$^-$ pre-DCs. Such technology enables global and unbiased identification of the gene-expression signature of individual cells and has the potential to reveal the transcriptomic basis for the diversification of DC subsets within DC precursor populations. This approach, combined with traditional techniques, including bulk transcriptomics and flow cytometry, revealed that the transcriptional priming of either the cDC1 lineage or the cDC2 lineage was already imprinted at the CDP stage, which established that commitment to these lineages was defined in the BM and not in peripheral tissues. Furthermore, this analysis allowed us to identify previously unknown transcriptional programs that control the transition among MDPs, CDPs and pre-DCs and to discover a previously unrecognized link between proliferation and differentiation during the differentiation of MDPs to CDPs and CDPs to the pre-DC stage, which sheds light on the role of cell-cycle control and specific transcription factors during the development of DC lineages.

RESULTS

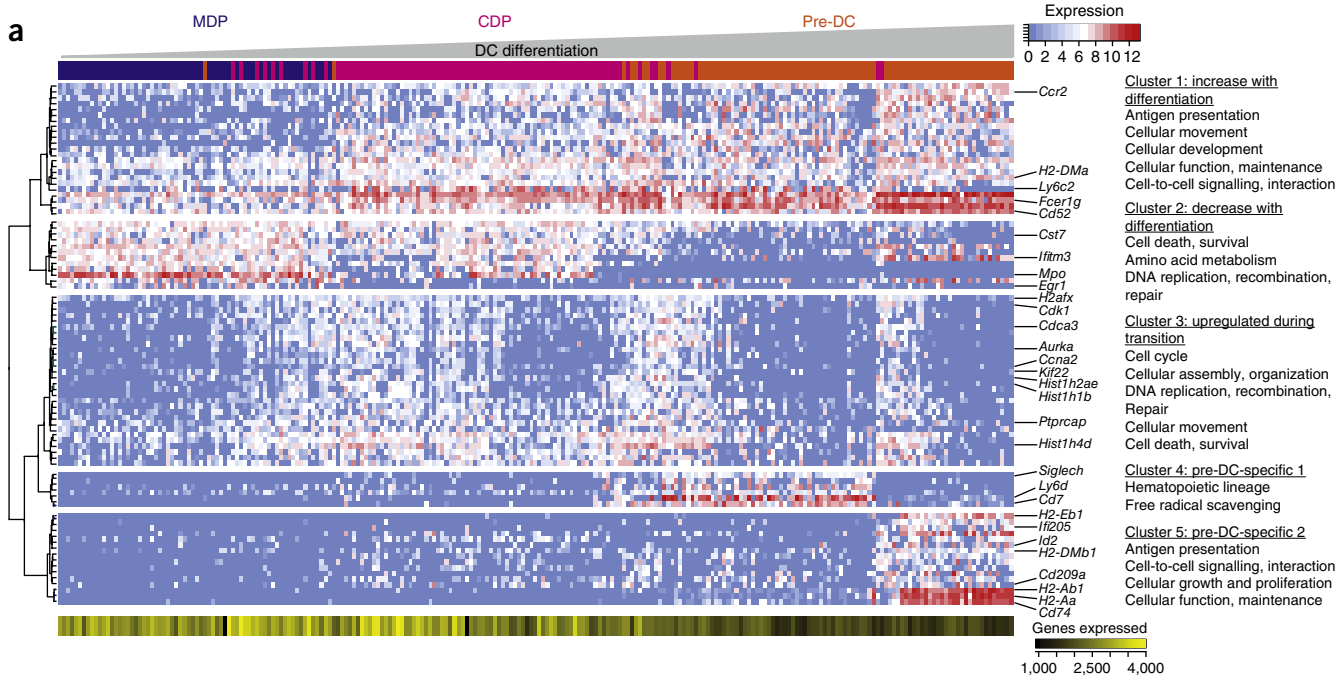
The transcriptional continuum of DC development

We isolated mouse BM DC progenitors, including 59 Lin$^-$CD11c$^-$MHCII$^-$CD135$^+$CSF-1R$^+$CD117^{hi} MDPs, 96 Lin$^-$CD11c$^-$MHCII$^-$CD135$^+$CSF-1R$^+$CD117^{int} CDPs and 96 Lin$^-$CD11c$^+$MHCII$^-$CD135$^+$CD172α$^-$ pre-DCs, by flow cytometry (workflow and quality control, **Supplementary Fig. 1**) and assessed the single-cell transcriptome by microfluidic single-cell mRNA sequencing (quantification of expressed genes detected in individual cells, **Supplementary Table 1**; sequencing-depth comparison, **Supplementary Fig. 2a**,**b**). Because CSF-1R$^+$ CDPs have been described as 'preferentially' developing into cDCs[6,7,11], we focused our analysis on this population. Subsequently, we used analysis of variance (ANOVA) and hierarchical clustering to identify differences among single MDPs, CDPs and pre-DCs in their transcriptomes to identify molecules involved in the differentiation of MDPs into CDPs and of CDPs into pre-DCs (hierarchical clustering based on all expressed genes, **Supplementary Fig. 2c**). Unsupervised single-cell hierarchical clustering of single MDP, CDP and pre-DC transcriptomes based on genes selected by ANOVA showed that 27% of single MDPs clustered with CDPs and that 28% of CDPs clustered with pre-DCs (**Supplementary Fig. 3a**). Such significant overlap in gene-expression profiles revealed transcriptional similarity between MDPs that clustered near CDPs and CDPs that clustered near pre-DCs. This indicated that although progenitor populations retained expression of surface markers at the protein level associated with the respective specific progenitor stages, individual cells within the MDP, CDP or pre-DC pool had already shifted transcriptionally toward the next step in differentiation. These observations highlighted the heterogeneity of each progenitor populations defined on the basis of the expression of a limited panel of surface markers and supported the concept that a developmental continuum applies for DC differentiation.

To identify those cells with a gene-expression profile intermediate between that of two contiguous progenitor stages, we developed the NBOR algorithm ('neighborhood-based ordering of single cells'). This algorithm can objectively determine the position of a given cell in a developmental continuum and calculate the similarity of each single cell's gene-expression profile to a defined gene set of a particular cell population (landmark) and then order each cell according to such similarity score into a spatial continuum around those landmarks (**Supplementary Fig. 3b**). We used NBOR to generate an unsupervised visualization of the single-cell mRNA profiles of all MDPs, CDPs and pre-DCs sequenced (**Supplementary Fig. 3a**,**b**) into a linear developmental order from the most undifferentiated (MDP) stage to most differentiated (pre-DC) stage (**Fig. 1a**). This analysis enabled visualization of gene expression in the sorted DC progenitors and identified five distinct clusters of genes whose expression was regulated differentially during DC maturation (full list of expressed genes, **Supplementary Table 2**; pathway-enrichment analysis with IPA software (Ingenuity Pathway Analysis) and the KEGG database (Kyoto Encyclopedia of Genes and Genomes), **Supplementary Table 3**). Two distinct gene clusters (1 and 2) marked the maturation of precursors from MDP to pre-DC. Cluster 1 comprised genes whose expression was increased during maturation from MDP to pre-DC precursors, including those encoding products involved in pathways of cellular development, cellular movement and antigen presentation (**Fig. 1a**,**b**, **Supplementary Fig. 3a** and **Supplementary Tables 2** and **3**). In contrast, the genes in cluster 2 underwent decreased expression during maturation and included genes encoding products linked to cell death and survival, DNA recombination and amino acid metabolism (**Fig. 1a**,**b**, **Supplementary Fig. 3a** and **Supplementary Tables 2** and **3**). Cluster 3 included genes whose expression was upregulated during periods of transition from one progenitor stage to the next and included several genes encoding products involved in cell cycle (**Fig. 1a**,**b**, **Supplementary Fig. 3a** and **Supplementary Tables 2** and **3**), which showed that cells proliferated during differentiation into more restricted progenitors. Cluster 4 included genes whose expression was increased upon transition from CDP to pre-DC and consisted of genes encoding products involved in free-radical scavenging, and genes such as *Ly6d* and *Siglech* (**Fig. 1a**, **Supplementary Fig. 3a** and **Supplementary Tables 2** and **3**). Genes in this cluster were predicted by IPA to be regulated by the transcription factor NF-κB, which mediates multiple aspects of cell survival and maturation[13]. The genes in cluster 5 had high expression in more-mature pre-DCs; among these were genes encoding products involved in the major histocompatibility complex class II (MHCII) presentation pathway, and these genes were predicted by the pathway-enrichment analysis described above (IPA and KEGG) to be regulated by the transcriptional coactivator CIITA, which controls this presentation pathway[14] (**Fig. 1a**,**b**, **Supplementary Fig. 3a** and **Supplementary Tables 2** and **3**). Together these data identified the transcriptional pathways that regulated the differentiation of DC progenitors at the single-cell level.

Considerable heterogeneity in the DC precursor compartment

To identify putative DC subset–primed progenitor cells within the pool of individual DC precursors, we compared the transcriptomic signatures of each MDP, CDP and pre-DC with signatures specific for total DCs, macrophages and the cDC1 and cDC2 subsets obtained from the Immunological Genome Project[15,16] or from our own transcriptomic database, by connectivity map (CMap) analysis, which is an extension of the gene-set–enrichment analysis algorithm[17] (analysis workflow, **Supplementary Fig. 4**; gene signatures, **Supplementary Tables 4–7**). The CMap analysis generates scores (as scaled dimensionless quantities) that indicate the degree of 'closeness' of one cell subset to a defined signature gene set. This analysis revealed that transcripts characteristic of the cDC lineage were already present in the transcriptomes of some single MDPs, CDPs or pre-DCs,



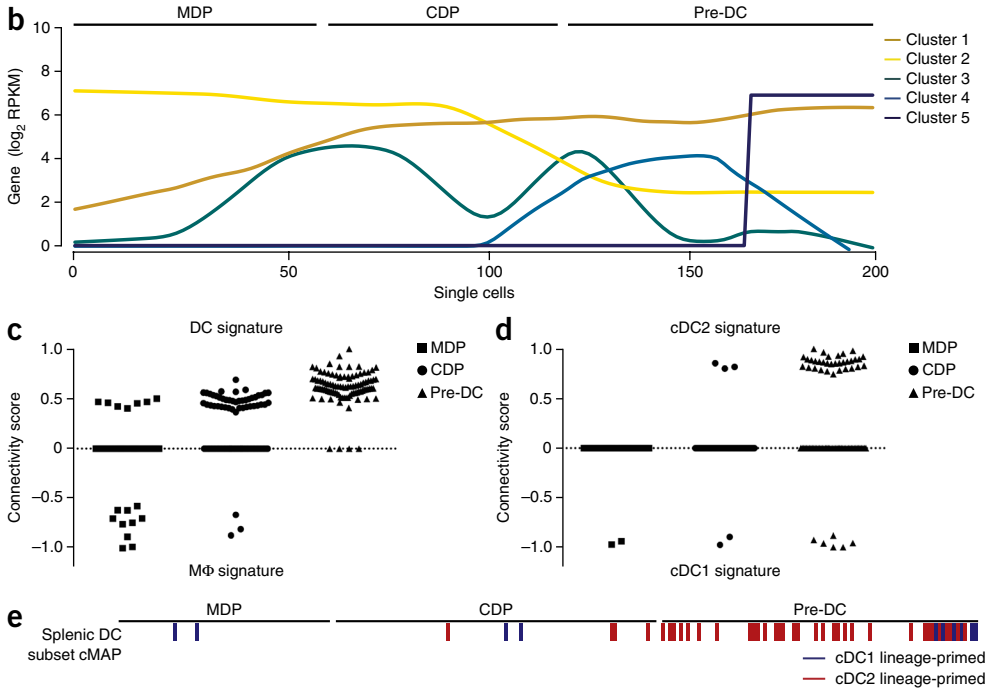


Figure 1 The transcriptional landscape of DC development. (**a**) Expression of the 87 genes (rows) with the greatest difference in expression in each single MDP, CDP and pre-DC (columns) relative to that in other MDPs, CDPs and pre-DCs, defined by data-analysis software (ANOVA followed by multiple-test correction by the Benjamini-Hochberg method), followed by analysis across each single MDP, CDP and pre-DC with the NBOR algorithm, to identify five clusters of genes encoding products in various pathways (far right). Left margin, hierarchical clustering; bottom, quantification of genes expressed in each cell (key, bottom right; RPKM (reads per kilobase of exon model per million mapped reads) of >1 for each); right margin, select gene symbols. (**b**) Expression of genes in clusters 1–5 (key; defined at right in **a**) across all single MDPs, CDPs and pre-DCs. (**c**,**d**) CMap analysis of single MDPs, CDPs and pre-DCs showing their enrichment for DC or macrophage (MΦ) gene sets (**c**) and splenic cDC1 or cDC2 gene sets (**d**) (gene sets, **Supplementary Tables 4–7**): a positive or negative score indicates connectivity to the corresponding gene set; a score of '0' (dotted line) indicates no commitment. Each symbol represents an individual cell (median values); thickening along horizontal axis indicates clustering of data. (**e**) Alignment of each single MDP, CDP and pre-DC found to show enrichment (as in **c**,**d**) for the splenic cDC1 gene set or cDC2 gene set (key), to the NBOR-generated DC-development continuum; each small vertical line in the plot indicates a single cell; CMap permutation *P* value, <0.05). Data are from one experiment with 1,000 permutations (**c**,**d**).

as judged by their closeness to a general DC lineage–specific signature. Such cDC 'imprinting' began at the MDP stage and gradually increased towards the pre-DC stage, at which point 95.8% of all pre-DCs sequenced showed closeness to the cDC signature, by CMap analysis (**Fig. 1c**). As expected, a decreasing frequency of DC progenitors along the DC-differentiation pathway (MDPs, 16.9%; CDPs, 3.1%; pre-DCs, 0%) aligned with the macrophage signature, in the CMap analysis (**Fig. 1c**), since CDPs and pre-DCs do not exhibit the potential to develop into macrophages[18].

While analyzing the closeness to cDC1- and cDC2-specific gene sets by CMap in single transcriptomes of MDPs, CDPs and pre-DCs, we found that the cDC2-lineage gene-set signature ($CD4^+$ DC and $CD11b^+$ DC subsets) was first detectable in individual CDPs (3.1% of all CDPs) and increased in the pre-DC cell pool (34% of all pre-DCs), whereas a transcriptomic signature closer to the cDC1-lineage gene set, in particular $CD8\alpha^+$ DC specific gene sets, was reliably detectable only at the pre-DC stage (6.25% of all pre-DCs) (**Fig. 1d**). We then identified MDPs, CDPs and pre-DCs whose transcriptomes showed closeness to either cDC1-specific gene sets or cDC2-specific gene sets, as revealed by CMap, in the developmental continuum generated by NBOR analysis (**Fig. 1a**). This analysis revealed a progressive clustering of cDC1 lineage– or cDC2 lineage–primed cells mostly during the pre-DC stage of maturation (**Fig. 1e**). These results indicated that individual cells within the DC precursor

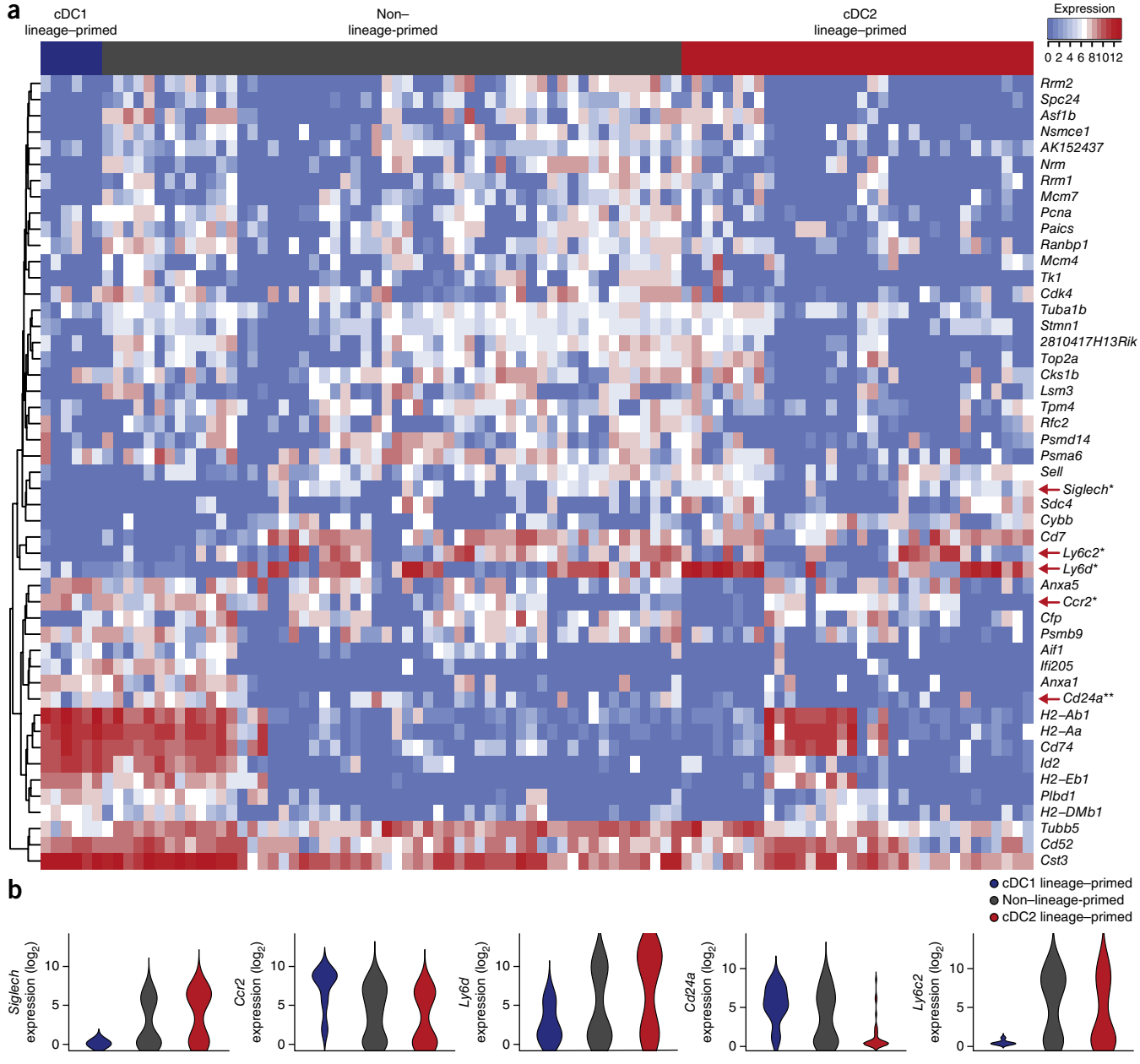


Figure 2 Heterogeneity in the pre-DC compartment is detectable at the mRNA level. (**a**) Expression of ANOVA-selected genes in single pre-DCs found to show enrichment (as in **Fig. 1c**,**d**) for the splenic cDC1 or cDC2 gene set or neither (color bars at top). Left margin, hierarchical clustering; red arrows (right margin), genes used as markers in subsequent analyses. $*P < 0.05$ and $**P < 0.01$ (ANOVA followed by multiple-test correction by the Benjamini-Hochberg method). (**b**) Expression of cell-surface protein–encoding genes selected by ANOVA in **a**, in cDC1 lineage–primed cells, cDC2 lineage–primed cells or non–lineage-primed cells (key), presented as 'violin' plots (height, gene expression; width, abundance of cells expressing the gene).

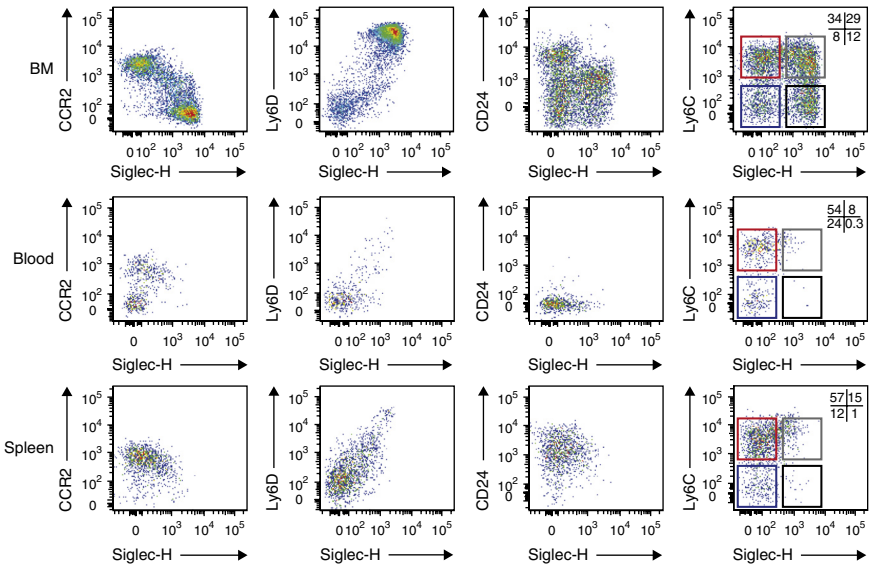


Figure 3 BM, blood and splenic pre-DCs have heterogeneous surface protein profiles. Expression of the surface proteins CCR2, Ly6D, CD24, Ly6C and Siglec-H (encoded by genes identified in **Fig. 2**) on pre-DC populations isolated from the BM, blood and spleen, analyzed by flow cytometry. Numbers in top right corners (corresponding to outlined areas in far right plots) indicate percent Ly6C$^+$Siglec-H$^-$ pre-DCs (red), Ly6C$^+$Siglec-H$^+$ pre-DCs (gray), Ly6C$^-$Siglec-H$^+$ pre-DCs (black) or Ly6C$^-$Siglec-H$^-$ pre-DCs (blue). Data are representative of three independent experiments with two replicates per condition in each.

population exhibited transcriptional priming toward the cDC1 and cDC2 lineages at the CDP and early pre-DC stage.

We then sought to determine whether those cDC1- and cDC2-primed pre-DCs could be identified on the basis of detectable differences in the expression of membrane proteins encoded by genes that were differentially expressed in cDC1- or cDC2-primed pre-DCs or non–lineage-primed pre-DCs. We began by identifying genes with significantly different ($P < 0.05$) expression in one pre-DC population relative to their expression in another pre-DC population that were identified by CMap as being transcriptionally primed to differentiate into the cDC1 or cDC2 subset or were not found to be primed to either of those subsets (**Fig. 2a** and **Supplementary Table 8**; analysis workflow, **Supplementary Fig. 4**). We found that the mRNA of several genes encoding surface markers, such as *Siglech*, *Ly6d*, *Cd24a*, *Ccr2* and *Ly6c2*, were expressed differentially by cells identified by CMap to be primed to either the cDC1 lineage or the cDC2 lineage or to be not primed at all (**Fig. 2a** and **Supplementary Table 8**; analysis workflow, **Supplementary Fig. 4**). In particular, we found high expression of *Siglech* and *Ly6d* in pre-DCs identified by CMap analysis as being not primed for either DC lineage or primed for the cDC2 lineage (**Fig. 2**), in agreement with published reports that Siglec-H marks a subset of pre-DCs that gives rise to both the cDC1 lineage and the cDC2 lineage[12,19] and that depletion of Siglec-H$^+$ cells in a Siglec-H–DTR mouse model results in loss of both the cDC1 lineage and the cDC2 lineage[12]. In addition, *Ccr2* was expressed by all cDC1-primed pre-DCs and showed bimodal expression in cDC2-primed and non–subset-primed pre-DCs identified by CMap (**Fig. 2**). All cDC1-primed pre-DCs expressed *Cd24a*, while the cDC2-primed pre-DCs had no *Cd24a* expression, and non–lineage-primed pre-DCs showed heterogeneous *Cd24a* expression (**Fig. 2**). This was in agreement with data suggesting that CD24 expression on splenic pre-DCs discriminates pre-DCs with cDC1 potential from those with cDC2 potential[6]. Finally, *Ly6c2* had no expression in cDC1-primed pre-DCs, and we detected bimodal *Ly6c2* expression in cDC2-primed pre-DCs and non–lineage-primed pre-DCs (**Fig. 2**). These results suggested that *Siglech* represents a marker for non–subset-primed pre-DCs, *Cd24a* and *Ccr2* may represent markers for commitment to the cDC1 lineage, and *Ly6c2* might be a marker for cDC2 subsets.

Heterogeneous expression of Siglec-H and Ly6C by pre-DCs

Next we used flow cytometry to examine expression of the surface markers CCR2, Ly6D, CD24, Ly6C and Siglec-H on the surface of Lin$^-$CD11c$^+$MHCII$^-$CD135$^+$CD172α$^-$ pre-DCs from BM, blood and spleen. We used Siglec-H as a marker for non–subset-primed pre-DCs, as suggested by CMap analysis and as described before[12]. Each marker combination resolved pre-DCs into distinct subpopulations and showed that the heterogeneity of BM pre-DCs was evident at both the mRNA level and the protein level. CCR2 was expressed exclusively on the Siglec-H$^-$ fraction of the BM pre-DCs and splenic pre-DCs, whereas ~30% of blood pre-DCs expressed CCR2 (**Fig. 3**). Conversely, Ly6D was expressed only on the Siglec-H$^+$ fraction of BM and splenic pre-DCs, whereas we detected no Ly6D expression on blood pre-DCs (**Fig. 3**). We detected intermediate expression of CD24 on the Siglec-H$^+$ fraction of BM pre-DCs (**Fig. 3**). The Siglec-H$^-$ fraction of BM and splenic pre-DCs showed bimodal distribution of CD24 expression, and we detected no CD24 expression on blood pre-DCs (**Fig. 3**). The expression of Siglec-H and Ly6C on pre-DCs from all tested tissues segregated these cells into four distinct subpopulations (Siglec-H$^+$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$, Siglec-H$^-$Ly6C$^+$ and Siglec-H$^-$Ly6C$^-$), with resolution superior to that of any other marker combinations (**Fig. 3** and **Supplementary Fig. 5a**). This observation suggested the existence of intermediate-phenotype populations within the BM Siglec-H$^+$ pre-DC fraction. All four BM pre-DC subsets defined by expression of Siglec-H and Ly6C (as described above) had similar morphology (**Supplementary Fig. 5b**). However, as assessed through the use of mice with transgenic expression of green fluorescent protein to report the expression of the transcription factors noted below, the four pre-DC subsets described above had very different expression of the DC-associated transcription factors ZBTB46 and Id2 and the chemokine receptor CX3CR1. Expression of ZBTB46, which marks DC precursor cells with definite commitment to give rise to the cDC1 or cDC2 subset but not pDCs, was highest in Siglec-H$^-$Ly6C$^+$ and Siglec-H$^-$Ly6C$^-$ pre-DCs (**Supplementary Fig. 5c**). In accordance with the published observation that ZBTB46$^+$ pre-DCs give rise only to the cDC1 and cDC2 subsets[20], not pDCs, we classified progenitor cells expressing ZBTB46 as potentially primed to develop into either the cDC1 lineage or the cDC2 lineage but not the pDC lineage. Expression of Id2, which has been shown to be a major developmental regulator of the cDC1 lineage[10], was restricted largely to Siglec-H$^-$Ly6C$^-$ pre-DCs (**Supplementary Fig. 5c**), which suggested that these cells might be dedicated precursors of the cDC1 subset, as suggested by the CMap analysis (**Fig. 1d**). These observations supported the notion that expression of Siglec-H and Ly6C distinguished pre-DC subpopulations committed to the cDC1 lineage (Siglec-H$^-$Ly6C$^-$ pre-DCs) or the cDC2 lineage (Siglec-H$^-$Ly6C$^+$ pre-DCs).

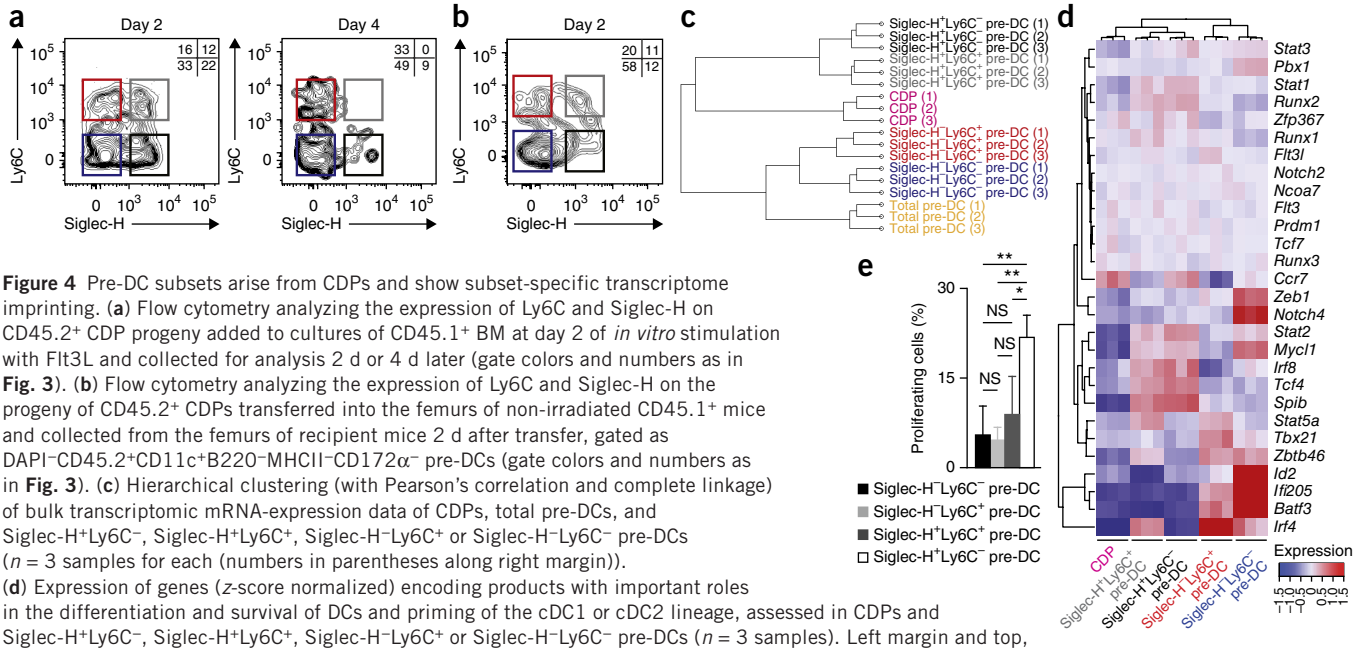


Figure 4 Pre-DC subsets arise from CDPs and show subset-specific transcriptome imprinting. (**a**) Flow cytometry analyzing the expression of Ly6C and Siglec-H on CD45.2$^+$ CDP progeny added to cultures of CD45.1$^+$ BM at day 2 of *in vitro* stimulation with Flt3L and collected for analysis 2 d or 4 d later (gate colors and numbers as in **Fig. 3**). (**b**) Flow cytometry analyzing the expression of Ly6C and Siglec-H on the progeny of CD45.2$^+$ CDPs transferred into the femurs of non-irradiated CD45.1$^+$ mice and collected from the femurs of recipient mice 2 d after transfer, gated as DAPI$^-$CD45.2$^+$CD11c$^+$B220$^-$MHCII$^-$CD172α$^-$ pre-DCs (gate colors and numbers as in **Fig. 3**). (**c**) Hierarchical clustering (with Pearson's correlation and complete linkage) of bulk transcriptomic mRNA-expression data of CDPs, total pre-DCs, and Siglec-H$^+$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$, Siglec-H$^-$Ly6C$^+$ or Siglec-H$^-$Ly6C$^-$ pre-DCs ($n = 3$ samples for each (numbers in parentheses along right margin)). (**d**) Expression of genes (*z*-score normalized) encoding products with important roles in the differentiation and survival of DCs and priming of the cDC1 or cDC2 lineage, assessed in CDPs and Siglec-H$^+$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$, Siglec-H$^-$Ly6C$^+$ or Siglec-H$^-$Ly6C$^-$ pre-DCs ($n = 3$ samples). Left margin and top, hierarchical clustering. (**e**) Frequency of proliferating cells among BM pre-DC subsets (key), as determined by expression of Azami Green in Fucci mice *in vivo*. NS, not significant; $^*P < 0.05$ and $^{**}P < 0.01$ (unpaired, two-tailed *t*-test). Data are representative of two independent experiments with one replicate (**a**) or two replicates (**b**) per condition in each, three experiments (**c**,**d**) or three independent experiments with seven mice (**e**; mean and s.e.m.).

Siglec-H$^+$ pre-DCs differentiate into Siglec-H$^-$ pre-DCs

Flow cytometry analyzing the expression of Siglec-H and Ly6C on pre-DCs from the BM, blood and spleen[18] showed that the majority of blood or spleen pre-DCs were either Siglec-H$^-$Ly6C$^+$ or Siglec-H$^-$Ly6C$^-$ (**Fig. 3** and **Supplementary Fig. 5d**), which suggested that Siglec-H$^+$ pre-DCs differentiated into Siglec-H$^-$ pre-DCs before exiting the BM. We next assessed the relationships among all four pre-DC subsets in an *in vitro* DC-differentiation protocol. We sorted each sorted pre-DC subset from the BM of CD45.2$^+$ mice and differentiated these individually *in vitro* by culture for 3 d together with Flt3L-stimulated CD45.1$^+$ BM cells. Siglec-H$^+$Ly6C$^-$ and Siglec-H$^+$Ly6C$^+$ pre-DCs gave rise to all pre-DC subsets and gave rise predominantly to Siglec-H$^-$Ly6C$^-$ pre-DCs (24%) and Siglec-H$^-$Ly6C$^+$ pre-DCs (27%), respectively (**Supplementary Fig. 5e**). Notably, Siglec-H$^-$Ly6C$^-$ and Siglec-H$^-$Ly6C$^+$ pre-DCs did not give rise to substantial populations of Siglec-H$^+$Ly6C$^-$ and Siglec-H$^+$Ly6C$^+$ pre-DC subsets (**Supplementary Fig. 5e**). Cumulative analysis suggested that the number of Siglec-H$^+$ pre-DCs generated by Siglec-H$^-$ pre-DCs was lower than the number of Siglec-H$^-$ pre-DCs generated by Siglec-H$^+$ pre-DCs (**Supplementary Fig. 5e**). These results indicated that Siglec-H$^+$Ly6C$^-$ and Siglec-H$^+$Ly6C$^+$ BM pre-DCs developed into Siglec-H$^-$Ly6C$^-$ and Siglec-H$^-$Ly6C$^+$ pre-DCs, which suggested their position developmentally upstream of Siglec-H$^-$ pre-DCs.

Subset-specific transcriptome imprinting of pre-DCs

To further understand the developmental continuity of the pre-DC subsets defined by expression of Siglec-H and Ly6C, we used a DC progenitor co-culture system (similar to that used above) to assess the kinetics of their development from upstream progenitors (CDPs). We cultured Flt3L-stimulated CD45.1$^+$ BM-derived DCs with sorted CD45.2$^+$Lin$^-$CD11c$^-$MHCII$^-$CD135$^+$CSF-1R$^+$CD117^{int} CDPs and analyzed the Siglec-H and Ly6C phenotype of their progeny by flow cytometry 2 d and 4 d later. The CD45.2$^+$CD11c$^+$B220$^-$MHCII$^-$ fraction of cultures at day 2 contained all four pre-DC populations, but by day 4, approximately 90% of CDP progeny were Siglec-H$^-$, with Siglec-H$^-$Ly6C$^-$ pre-DCs and Siglec-H$^-$Ly6C$^+$ pre-DCs representing 49% and 33% of the pre-DCs, respectively (**Fig. 4a**). These results suggested that Siglec-H$^-$Ly6C$^-$ pre-DCs and Siglec-H$^-$Ly6C$^+$ pre-DCs represented the most differentiated subsets downstream of the Siglec-H$^+$Ly6C$^-$ and Siglec-H$^+$Ly6C$^+$ pre-DC subsets within the pre-DC population.

We also sought to determine whether CDPs could generate all those pre-DC subsets *in vivo*. We isolated Lin$^-$CD11c$^-$MHCII$^-$CD135$^+$CSF-1R$^+$CD117^{int} CDPs from CD45.2$^+$ donor mice and injected the cells intra-femorally into unirradiated CD45.1$^+$ host mice, then analyzed their progeny among BM CD45.2$^+$CD11c$^+$B220$^-$MHCII$^-$ cells at 2 d after transfer. We detected Siglec-H$^+$Ly6C$^-$ (12%), Siglec-H$^+$Ly6C$^+$ (11%), Siglec-H$^-$Ly6C$^-$ (58%) and Siglec-H$^-$Ly6C$^+$ (20%) pre-DC subsets in the BM of the recipient mice 2 d after transfer (**Fig. 4b**), which suggested that CDPs were able to give rise to all four pre-DC subsets.

To gain further insight into the developmental relationships among these various DC progenitors in the BM, we performed gene-expression analysis of total CDPs, total pre-DCs and the pre-DC subsets defined by expression of Siglec-H and Ly6C. Hierarchical clustering of these subpopulations revealed that CDPs clustered together with Siglec-H$^+$Ly6C$^-$ pre-DCs and Siglec-H$^+$Ly6C$^+$ pre-DCs (**Fig. 4c**), which supported the hypothesis that these subsets were less mature than Siglec-H$^-$Ly6C$^-$ pre-DCs and Siglec-H$^-$Ly6C$^+$ pre-DCs. Conversely, Siglec-H$^-$Ly6C$^+$ pre-DCs and Siglec-H$^-$Ly6C$^-$ pre-DCs clustered more closely with total pre-DCs than with CDPs (**Fig. 4c**), which indicated that these subsets were more advanced in their maturation. We next investigated the expression of specific genes commonly associated with general DC development and DC subset identity. Siglec-H$^-$Ly6C$^+$ and Siglec-H$^-$Ly6C$^-$ pre-DCs had high expression of genes associated with a mature DC phenotype, such as *Zbtb46* and *Tbx21* (**Fig. 4d**). Genes encoding products known to have a role in cDC development, such as *Flt3*, *Stat2* and *Stat3* (ref. 15), had similar expression across all four pre-DC subsets (**Fig. 4d**).

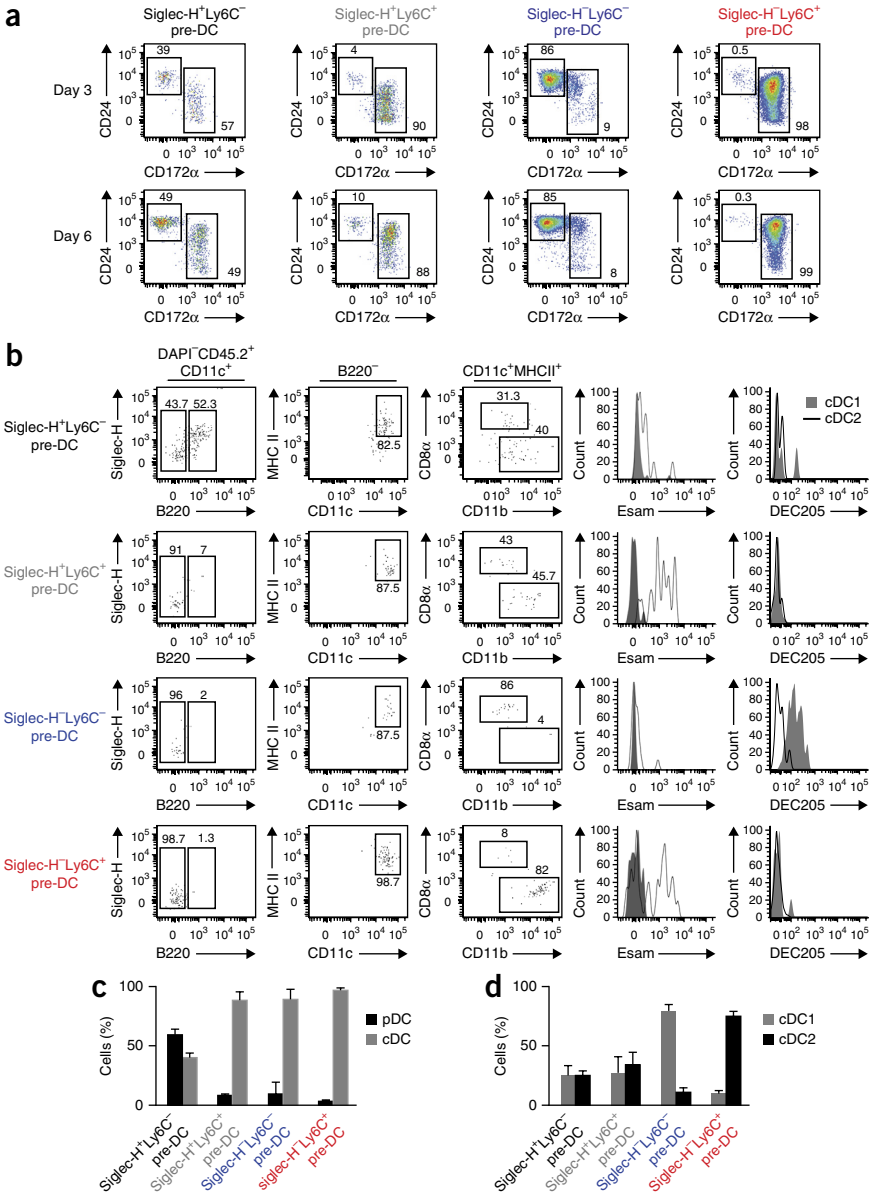


Figure 5 Identification of DC subset–specific progenitors. (**a**) Flow cytometry analyzing the expression of CD24 and CD172α on the progeny of CD45.2+ pre-DC subsets (defined by Siglec-H and Ly6C (above plots) as in **Fig. 4**) added to cultures of CD45.1+ BM at day 2 of *in vitro* stimulation with Flt3L, followed by incubation for 3 d or 6 d; gated on DAPI− CD45.2+CD11c+MHCII+B220− DCs. Numbers adjacent to outlined areas indicate percent CD24+CD172α− cells (top left) or CD24hi–loCD172α+ cells (bottom right). (**b**) Flow cytometry of donor pre-DCs (defined by Siglec-H and Ly6C (left margin) as in **Fig. 4**) obtained from the spleen of non-irradiated CD45.1+ host mice 3 d after transfer (into the femurs of the recipient mice) of pre-DC subsets isolated from the BM of CD45.2+ donor mice (gating above plots). Numbers adjacent to outlined areas (left half) indicate percent Siglec-Hhi–loB220− cells (left) or Siglec-Hhi–loB220+ cells (right) (far left column), MHCII+CD11c+ cells (middle column), or CD8α+CD11b− cells (top left) or CD8α−CD11b+ cells (bottom right) (right column). Right, expression of Esam (left) and DEC205 (right) on cDC1 and cDC2 cells (key). (**c**) Frequency of donor-derived CD11c+B220+Siglec-H+ pDCs and CD11c+B220−Siglec-H−MHCII+ cDCs in the spleen of host mice as in **b** at 3 d after transfer, presented as CD11c+MHCII+ cells among donor-derived CD45.2+ cells. (**d**) Contribution of CD45.2+ pre-DCs (defined by Siglec-H and Ly6C (horizontal axis) as in **Fig. 4**) in the spleen of host mice (as in **c**) to splenic cDC1 or cDC2 cells in the CD11c+MHCII+Siglec-H−B220− population also in the spleen. Data are representative of three independent experiments with two replicates per condition in each (**a**), four independent experiments with one replicate per condition in each (**b**) or two independent experiments with three replicates per condition in each (**c**,**d**; mean and s.e.m.).

Irf4, which encodes a major regulator of the cDC2 lineage[21,22], had high expression in Siglec-H−Ly6C+ pre-DCs, whereas cDC1 lineage–specific transcripts, including *Id2*, *Batf3*, *Notch4* and *Ifi205*, had their most abundant expression in Siglec-H−Ly6C− pre-DCs[10,15] (**Fig. 4d**). Both Siglec-H+Ly6C− pre-DCs and Siglec-H+Ly6C+ pre-DCs had much higher expression of the pDC-associated transcripts *Tcf4* and *Spib*[23] than that of Siglec-H−Ly6C+ pre-DCs or Siglec-H−Ly6C− pre-DCs (**Fig. 4d**), which indicated that these pre-DCs might have retained some pDC potential.

Because DC progenitors have been shown to decrease their proliferation as they become more differentiated toward the pre-DC stage[18], we assessed the proliferation of Siglec-H+Ly6C−, Siglec-H+Ly6C+, Siglec-H−Ly6C− and Siglec-H−Ly6C+ pre-DCs *in vivo* in the BM, blood and spleen of Fucci reporter mice ('fluorescent ubiquitination–based cell-cycle indicator'), in which the green-emitting fluorescent protein Azami Green labels the protein geminin, whose expression is associated with cells in the S, G2 and M phases of the cell cycle[24]. Siglec-H+Ly6C− pre-DCs and Siglec-H+Ly6C+ pre-DCs proliferated considerably more than Siglec-H−Ly6C+ pre-DCs or Siglec-H−Ly6C− pre-DCs did (**Fig. 4e** and **Supplementary Fig. 6a**), which indicated that Siglec-H+Ly6C− pre-DCs and Siglec-H+Ly6C+ pre-DCs were less differentiated than Siglec-H−Ly6C+ pre-DCs or Siglec-H−Ly6C− pre-DCs were. Together these data indicated that Siglec-H+Ly6C+ pre-DCs and Siglec-H+Ly6C− pre-DCs were non–subset-primed cells, Siglec-H+Ly6C− pre-DCs were probably cDC1 subset–primed cells and Siglec-H−Ly6C+ pre-DCs were primed toward the cDC2 lineage.

Identification of dedicated DC subset–specific progenitors

To assess the developmental potential of Siglec-H+Ly6C−, Siglec-H+Ly6C+, Siglec-H−Ly6C− and Siglec-H−Ly6C+ pre-DCs individually, we purified those subsets from CD45.2+ BM by flow cytometry and added them into individual Flt3L-stimulated CD45.1+ BM cultures, then analyzed their progeny 3 d and 6 d later. Because CD8α and CD11b are not reliably expressed by *in vitro* BM-derived DC cultures[5], we identified cDC1 cells here as CD24+CD172α− and identified cDC2 cells as CD24CD172α+ by gating on CD45.2+CD11c+MHCII+B220− cells in the total DC population. We found that Siglec-H+Ly6C+ pre-DCs and Siglec-H+Ly6C− pre-DCs gave rise efficiently to cDC1

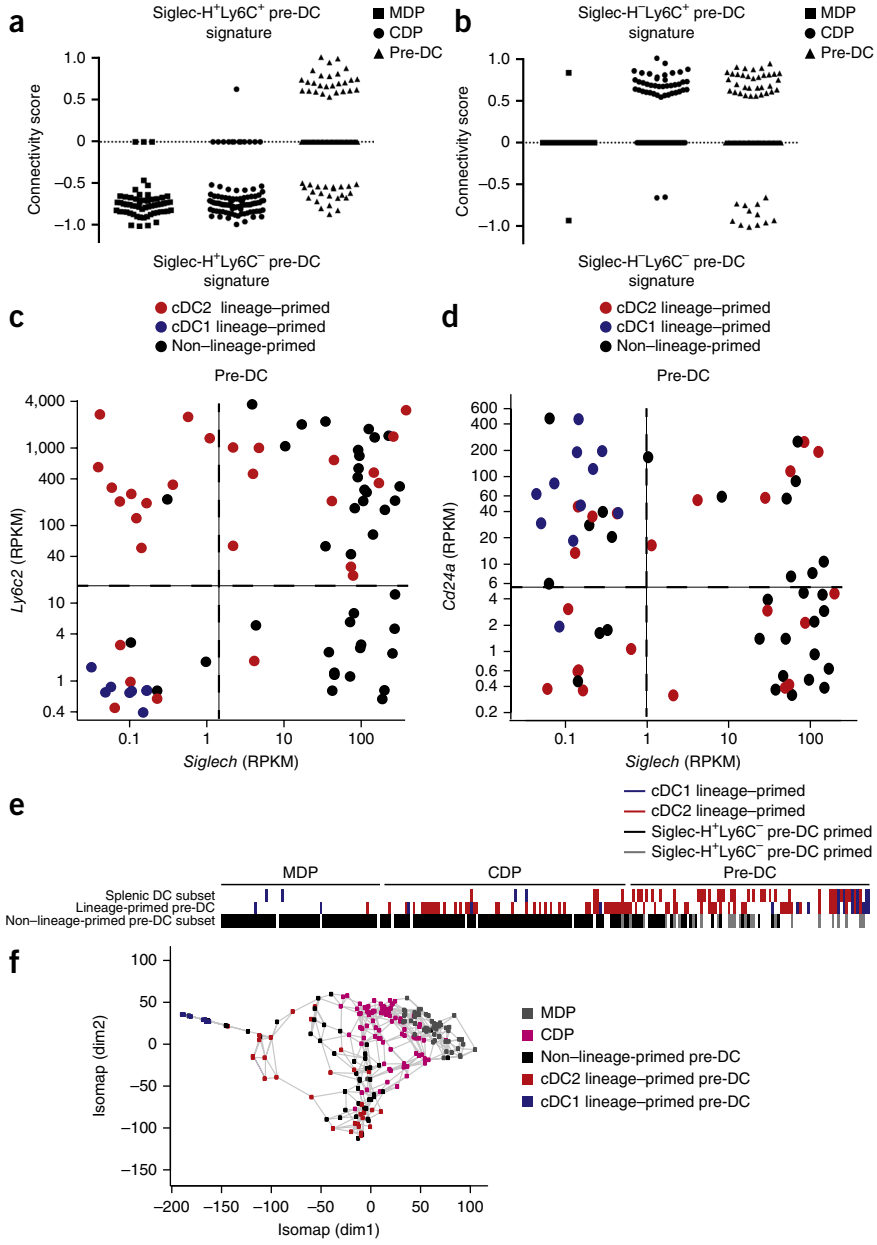


Figure 6 Analysis of single MDPs, CDPs and pre-DCs facilitates visualization of dedicated cDC1 lineage– and cDC2 lineage–primed pre-DCs. (**a**,**b**) CMap analysis of single DC precursors showing their enrichment for the gene signatures of Siglec-H$^+$Ly6C$^+$ or Siglec-H$^+$Ly6C$^-$ pre-DC populations (**a**) or Siglec-H$^-$Ly6C$^+$ or Siglec-H$^-$Ly6C$^-$ pre-DC populations (**b**) (presented as in **Fig. 1c**,**d**). (**c**,**d**) Expression (as RPKM) of *Siglech* and *Ly6c2* (**c**) or *Siglech* and *Cd24a* (**d**) in single pre-DC transcriptomes, overlaid with enrichment for the cDC1- or cDC2-specific gene set identified in **b**, or no enrichment in either (key); distribution of CMap-identified cDC1 lineage– or cDC2 lineage–primed or non–lineage-primed single pre-DCs versus a null distribution assessed by Ψ2 test. (**e**) Alignment of cDC1 lineage– or cDC2 lineage–primed pre-DCs identified in **b** to the continuum of single MDPs, CDPs and pre-DCs (presented as in **Fig. 1e**). (**f**) Isomap visualization of the developmental continuum of DC progenitors: MDPs, CDPs and pre-DC populations identified in **b** (key); dim1 and dim2 (along axes), dimensions 1 and 2. Data are from one experiment with 1,000 permutations.

and cDC2 cells on days 3 and 6 of culture (**Fig. 5a**). Siglec-H$^-$Ly6C$^-$ pre-DCs gave rise predominantly to cDC1 cells, whereas Siglec-H$^-$Ly6C$^+$ pre-DCs gave rise 'preferentially' to cDC2 cells (**Fig. 5a**). We also sought to determine whether any of the four pre-DC subsets was able to generate pDCs *in vitro*. Siglec-H$^+$Ly6C$^+$ pre-DCs and Siglec-H$^+$Ly6C$^-$ pre-DCs produced CD45.2$^+$CD11c$^+$B220$^+$ pDC progeny by days 3 and 6 of culture (**Supplementary Fig. 6b**). In contrast, by day 3, Siglec-H$^-$Ly6C$^-$ pre-DCs gave rise to few pDCs, and although the proportion of pDCs in the culture had increased by day 6, it remained considerably lower than that in Siglec-H$^+$Ly6C$^-$ or Siglec-H$^+$Ly6C$^+$ pre-DC cultures (**Supplementary Fig. 6b**). We did not detect pDCs in Siglec-H$^-$Ly6C$^+$ pre-DC cultures on either day 3 or day 6 (**Supplementary Fig. 6b**).

We next investigated the developmental potential of each of the four pre-DC subsets defined by their expression of Siglec-H and Ly6C through the use of an *in vitro* clonal development assay. We purified Siglec-H$^+$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$, Siglec-H$^-$Ly6C$^-$ and Siglec-H$^-$Ly6C$^+$ pre-DCs by flow cytometry and cultured single pre-DCs for 6 d on an OP9 stromal feeder layer in conditioned medium from BM cultures stimulated for 9 d with Flt3L. All Siglec-H$^+$Ly6C$^-$ or Siglec-H$^+$Ly6C$^+$ cells gave rise to cDCs, and 29% of Siglec-H$^+$Ly6C$^-$ pre-DC clones and 11% of Siglec-H$^+$Ly6C$^+$ pre-DC clones also generated pDCs (**Supplementary Fig. 6c**). However, none of the Siglec-H$^+$Ly6C$^-$ or Siglec-H$^+$Ly6C$^+$ single cells had exclusive pDC potential. Most Siglec-H$^+$Ly6C$^-$ pre-DC clones produced both cDC1 cells and cDC2 cells (54%), while Siglec-H$^+$Ly6C$^+$ pre-DCs clones gave rise to either cDC1 cells (68%), cDC2 cells (11%) or both subsets (20%). In line with their limited proliferative capacity, the single Siglec-H$^-$Ly6C$^-$ or Siglec-H$^-$Ly6C$^+$ pre-DCs did not produce sufficient progeny for detection within the 6-day culture period.

To confirm the observations reported above, we sorted each of those four pre-DC subsets from the BM of CD45.2$^+$ adult mice, injected each individually into the femurs of non-irradiated CD45.1$^+$ mice, and assessed the phenotypes of their progeny in the spleen of recipient mice 3 d and 5 d later. All pre-DC subsets gave rise to CD11c$^+$MHCII$^+$ cDCs, while only Siglec-H$^+$Ly6C$^-$ pre-DCs gave rise to Siglec-H$^+$B220$^+$ pDCs (**Fig. 5b**,**c** and **Supplementary Fig. 6d**). Siglec-H$^+$Ly6C$^+$ pre-DCs gave rise to both CD8α$^+$DEC205$^+$ cDC1 cells and CD11b$^+$Esam$^+$ cDC2 cells, Siglec-H$^-$Ly6C$^+$ pre-DCs gave rise only to CD11b$^+$Esam$^+$ cDC2 cells, and the Siglec-H$^-$Ly6C$^-$ pre-DC population gave rise only to CD8α$^+$DEC205$^+$ cDC1 cells (**Fig. 5b**,**d** and **Supplementary Fig. 6d**).

In addition, we did a similar experiment investigating the progeny of Siglec-H$^-$Ly6C$^+$CD24$^-$ pre-DCs and Siglec-H$^-$Ly6C$^-$CD24$^+$ pre-DCs, as we had found that *Cd24a* was expressed by all

cDC1-primed pre-DCs (**Fig. 2**), and published data suggest that CD24 expression on splenic pre-DCs can be used to discriminate pre-DCs with cDC1 potential from those with cDC2 potential[6]. We assessed the phenotypes of their progeny in the spleen of recipient mice 3 d after transfer. At this time point, Siglec-H$^-$Ly6C$^+$CD24$^-$ pre-DCs produced only cDC2 cells, whereas Siglec-H$^-$Ly6C$^-$CD24$^+$ pre-DCs produced only cDC1 cells (data not shown). Together these data identified Siglec-H$^-$Ly6C$^+$CD24$^-$ pre-DCs as cDC2 subset–specific precursors, Siglec-H$^-$Ly6C$^-$CD24$^+$ pre-DCs as cDC1 subset–specific precursors and Siglec-H$^+$Ly6C$^-$ pre-DCs as progenitors that retained pDC-differentiation potential *in vivo* (**Supplementary Fig. 6e**).

DC subset priming at the CDP stage

We next sought to determine at what stage during the differentiation of DC progenitors in the BM did the transcriptomic signatures of subset-specific pre-DCs become evident. By CMap analysis, we compared the transcriptomes of single MDPs, CDPs and pre-DCs with signature gene sets of Siglec-H$^+$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$, Siglec-H$^-$Ly6C$^+$ or Siglec-H$^-$Ly6C$^-$ pre-DCs. This analysis revealed that the majority of single MDPs (95%) and CDPs (89%) displayed a transcriptomic signature similar to that of Siglec-H$^+$Ly6C$^-$ pre-DCs, whereas pre-DCs had a transcriptomic signature shared by Siglec-H$^+$Ly6C$^-$ pre-DCs (25%) and Siglec-H$^+$Ly6C$^+$ pre-DCs (28%) (**Fig. 6a**, **Supplementary Fig. 6e** and **Supplementary Tables 9–12**), which indicated that Siglec-H$^+$Ly6C$^-$ pre-DCs were 'between' CDPs and Siglec-H$^+$Ly6C$^+$ pre-DCs. Furthermore, this analysis indicated that individual CDPs expressed a substantial number of genes specific for cDC2 lineage–primed pre-DCs, while we detected expression of genes associated with cDC1-lineage priming only in pre-DCs (**Fig. 6a**, **Supplementary Fig. 6e** and **Supplementary Tables 9–12**). Through use of the χ^2 test with *P* values representing the significance of the distribution of CMap-identified cell types versus a null distribution, we observed a significant correlation between the expression of *Ly6c2* and *Siglech* mRNA ($P = 7.07 \times 10^{-10}$) and *Cd24a* and *Siglech* mRNA ($P = 2.39 \times 10^{-3}$) in pre-DCs (**Fig. 6c**,**d**), which indicated that Ly6C, CD24 and Siglec-H could be used as markers, not only at the protein level but also at the transcriptomic level, for the identification of cDC2-primed, cDC1-primed or non-primed pre-DC subsets, respectively. Correlating results obtained by CMap analysis (**Figs. 1d** and **6b**) showed that pre-DCs with a transcriptome closer to cDC1 lineage–specific genes or cDC2 lineage–specific genes (**Fig. 1d**) overlapped cell populations enriched for Siglec-H$^-$Ly6C$^-$ pre-DCs (cDC1 lineage–primed pre-DCs) or Siglec-H$^-$Ly6C$^+$ pre-DCs (cDC2 lineage–primed pre-DCs) (**Fig. 6b**), which showed that cDC1- or cDC2-lineage priming happened in the transition phase from CDP to pre-DC (**Fig. 6e**).

Finally, we clustered gene-expression data for genes expressed by all single MDPs, CDPs and pre-DC progenitors by Isomap ('isometric feature mapping')[25], a nonlinear dimensionality reduction algorithm. Overlaying Isomap-generated clustering with the information on the priming of pre-DC to the cDC1 lineage or cDC2 lineage or to neither of these two lineages, as obtained by CMap analysis (**Fig. 1b**), allowed us to visualize the emergence of the transcriptional signature of cDC1- or cDC2-lineage priming. Tightly clustered MDPs gave rise to the CDP population, which further differentiated into pre-DCs that were not primed for the cDC1 or cDC2 lineage (**Fig. 6f**). These pre-DCs further differentiated into either cDC1 lineage–primed pre-DCs or cDC2 lineage–primed pre-DCs (**Fig. 6f**). Isomap analysis provided visualization of the complex transcriptional landscape starting from MDPs to cDC1 lineage– or cDC2 lineage–primed pre-DCs. This suggested that DC development is not a synchronized linear event but is instead a continuous emergence of dedicated precursor cells, in which commitment is acquired via several stages of maturation.

DISCUSSION

Here, single-cell mRNA sequencing enabled analysis of the developmental progression of individual cells in the BM toward the cDC1 and cDC2 subsets. By single-cell mRNA-sequencing analysis, traditional transcriptomics and flow cytometry, we identified Siglec-H$^+$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$, Siglec-H$^-$Ly6C$^-$ and Siglec-H$^-$Ly6C$^+$ pre-DC subsets within the BM pre-DC population. Siglec-H$^+$Ly6C$^-$ pre-DCs were developmentally close to late CDPs and retained the potential to differentiate into pDCs *in vivo* up until their differentiation into Siglec-H$^+$Ly6C$^+$ pre-DCs, which gave rise exclusively to cDCs. Siglec-H$^+$Ly6C$^+$ pre-DCs differentiated into either Siglec-H$^-$Ly6C$^+$ pre-DCs that were primed toward the cDC2 lineage or Siglec-H$^-$Ly6C$^-$ pre-DCs that were primed toward the cDC1 lineage. We also identified CD24 as a marker for dedicated cDC1 precursors in the BM and the spleen but not in the blood.

Published work has shown that imprinting of the DC lineage is observed at the level of early hematopoietic progenitors, such as the lymphoid-primed multipotent progenitor[26]. Our observations here allowed more detailed mapping of the developmental continuum of DCs: DC lineage–primed lymphoid-primed multipotent progenitors differentiated into MDPs, which had no monocyte or macrophage potential and in turn differentiated into CDPs and subsequently into pre-DCs. The transition from CDP to pre-DC marked the appearance of transcriptional signatures characteristic of the cDC1 and cDC2 subsets. Such 'primed' progenitors subsequently gave rise to cDC1- and cDC2-specific pre-DCs in the BM, which would then emigrate to peripheral tissues.

Published studies have shown that single CDPs cultured *in vitro* are biased to develop 'preferentially' into the cDC1 or cDC2 subset or both subsets of cDCs, which suggests that CDPs and pre-DCs might be heterogeneous[6]. We found significant heterogeneity in DC precursors *in vivo* and found that these could be defined by both a characteristic transcriptomic signature and a distinct pattern of cell-surface markers. These observations should allow better understanding of the molecular events that underlie the process of the priming of progenitors toward the cDC1 and cDC2 subsets, and revealed distinct patterns of gene expression and identified molecular switches that drive DC subset differentiation. However, it remains unclear whether the transcriptional programs that regulate priming of the cDC1 and cDC2 subsets of BM pre-DCs also operate in CDPs. In addition, the particular cell-intrinsic or BM niche–extrinsic signals that drive subset priming of DC progenitors remain to be identified. Genome-wide chromatin mapping of MDPs or CDPs, as has been done for hematopoietic progenitor cells[27] such as hematopoietic stem cells, may help elucidate how cDC1- or cDC2-subset identity is enforced in pre-DCs. Finally, our data have revealed the intimate connection between proliferation and differentiation in DC progenitors, as the induction of genes encoding cell cycle–associated products always preceded substantial changes in expression of gene clusters encoding products that drive the differentiation of DC progenitors. Such temporal coupling would suggest that the remodeling of chromatin during the cell cycle facilitates differentiation, and/or that transcriptional regulation of the cell-cycle machinery and cell differentiation are coupled[28,29]. To what extent proliferation and differentiation are interrelated in during DC differentiation remains to be investigated.

The identification of BM-resident DC subset–specific precursors should allow better assessment of the role of the microenvironment in

peripheral tissues on DC differentiation on various functional subsets. We propose that a central transcriptomic subset–specific program is imprinted in DC precursors at the CDP stage, which confers a core subset identity regardless of the final tissue destination, while peripheral tissue–dependent programming is added to that to ensure site-specific functionality and adaptation. The transcriptional regulators Id2, Batf3 and NFIL3 are examples of such central programming, as their loss affects all cDC1 subsets throughout the body[10,30,31]. How such central imprinting is regulated remains to be investigated. The maturation of hematopoietic stem cells depends on multiple layers of regulation that integrate environmental signals from the BM, the activation status of bystander cells in the niche, and the local cytokine milieu, as well as inflammatory signals received from the periphery[32–37]. Similar regulatory mechanisms may prime subset identity in pre-DCs.

Such a multifactorial model of signal integration in DC differentiation could explain the dependence of the versatile cDC2 population on tissue-specific transcription factors such as Notch2, Relb or IRF4 (refs. 21,22,38,39), while their essential cDC2-specific functions seem to have been already programmed in the BM. The cDC2 transcriptional requirements differ according to the tissue location of DCs. Spleen Esam$^+$ cDC2 cells are dependent on Notch2, Relb and the receptor for lymphotoxin-β for their maintenance and proliferation but are not dependent on IRF4 (refs. 21,38–41), and CD103$^+$ cDC2 cells in the small intestinal lamina propria are dependent on Notch2, IRF4 and lymphotoxin-β receptor[21,22,38]; in contrast, lung cDC2 cells are independent of Notch2 (A.S. and F.G., data not shown) and Relb (data not shown) but do require IRF4 (refs. 21,22,38,39,41) for their survival. Similarly, the Epstein-Barr virus–induced G protein–coupled receptor EBI2 is crucial for the homeostasis and the correct localization of cDC2 cells in the spleen but not in lymph nodes or non-lymphoid tissues. However, EBI2 is not required for early development of DCs, as DC precursors, including MDPs, CDPs and pre-DCs, are present in normal (wild-type) numbers in the BM and spleen of mice lacking this receptor[42,43]. This further illustrates the interaction between transcriptomic imprinting of DC subset identity in the BM and functional imprinting in peripheral organs. Such data show that although all cDC2 cells originate from the same lineage of dedicated precursors and emerge from the BM with central cDC2 programming, tissue-specific factors add another layer of imprinting in the periphery[42]. This dual layer of transcriptional imprinting is biologically relevant, as it simultaneously equips cells of the cDC1 or cDC2 lineage with their essential lineage-specific functions and allows the opportunity for the tissue site and/or context to fine-tune transcriptional control of phenotype, localization, proliferation and DC function.

Collectively, we have provided evidence that a central transcriptional program for DC subset development emerged in BM DC progenitors as early as the pre-DC stage, for the cDC1 lineage, and the CDP state, for the cDC2 lineage, which established that identity of the cDC1 lineage versus that of the cDC2 lineage was imprinted before pre-DCs left the BM. Further work should investigate the effect of peripheral programming by each tissue on the final maturation, localization and functional specialization of DC subsets in the periphery. Better understanding of the origins and the mechanisms that ensure DC subset differentiation, identity and homeostasis has the potential to improve intervention and therapeutic strategies.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GEO: single-cell mRNA sequencing and microarray data, GSE60783.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank L. Robinson for critical review and editing of the manuscript; M.L. Ng, S.H. Tan, T.B. Lu, I. Low and N.B. Shadan for technical assistance; K. Murphy (University of Washington) for ZBTB46-GFP BM; and G. Belz (Walter & Eliza Hall Institute) for Id2-GFP BM. Supported by Singapore Immunology Network (F.G.), Agency for Science, Technology and Research (BMRC Young Investigator grant to A.S.), the Genomics Institute of Singapore (P.R.) and the Singapore Medical Research Council (NMRC/CBRG/0047/2013 to B.M.).

AUTHOR CONTRIBUTIONS

A.S., P.R. and F.G. conceived of the study; A.S., S.V., H.R.B.S., J.S., J.L. and B.M. performed experiments; A.S., J.C., S.Z., M.P. and F.G. analyzed data; A.L., F.Z., L.R., S.N. and E.W.N. provided reagents and intellectual guidance; and A.S., S.V., J.C. and F.G. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Schlitzer, A. & Ginhoux, F. Organization of the mouse and human DC network. *Curr. Opin. Immunol.* **26**, 90–99 (2014).
2. Merad, M., Sathe, P., Helft, J., Miller, J. & Mortha, A. 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 (2013).
3. Schlitzer, A., McGovern, N. & Ginhoux, F. Dendritic cells and monocyte-derived cells: Two complementary and integrated functional systems. *Semin. Cell Dev. Biol.* doi:10.1016/j.semcdb.2015.03.011 (2015).
4. Karsunky, H., Merad, M. & Cozzio, A. Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3$^+$ dendritic cells in vivo. *J. Exp. Med.* **198**, 305–313 (2003).
5. Fogg, D.K. *et al.* A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* **311**, 83–87 (2006).
6. Naik, S.H. *et al.* Development of plasmacytoid and conventional dendritic cell subtypes from single precursor cells derived *in vitro* and *in vivo*. *Nat. Immunol.* **8**, 1217–1226 (2007).
7. Onai, N. *et al.* Identification of clonogenic common Flt3$^+$M-CSFR$^+$ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. *Nat. Immunol.* **8**, 1207–1216 (2007).
8. Sathe, P. *et al.* Lymphoid tissue and plasmacytoid dendritic cells and macrophages do not share a common macrophage-dendritic cell-Restricted progenitor. *Immunity* **41**, 104–115 (2014).
9. Liu, K. & Nussenzweig, M.C. Development and homeostasis of dendritic cells. *Eur. J. Immunol.* **40**, 2099–2102 (2010).
10. Ginhoux, F. *et al.* The origin and development of nonlymphoid tissue CD103$^+$ DCs. *J. Exp. Med.* **206**, 3115–3130 (2009).
11. Onai, N. *et al.* A clonogenic progenitor with prominent plasmacytoid dendritic cell developmental potential. *Immunity* **38**, 943–957 (2013).
12. Swiecki, M. *et al.* Cell depletion in mice that express diphtheria toxin receptor under the control of SiglecH encompasses more than plasmacytoid dendritic cells. *J. Immunol.* **192**, 4409–4416 (2014).
13. Kucharczak, J., Simmons, M.J., Fan, Y. & Gélinas, C. To be, or not to be: NF-κB is the answer–role of Rel/NF-κB in the regulation of apoptosis. *Oncogene* **22**, 8961–8982 (2003).
14. LeibundGut-Landmann, S. *et al.* Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes. *Eur. J. Immunol.* **34**, 1513–1525 (2004).
15. Miller, J.C. *et al.* Deciphering the transcriptional network of the dendritic cell lineage. *Nat. Immunol.* **13**, 888–899 (2012).
16. Gautier, E.L. *et al.* Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat. Immunol.* **13**, 1118–1128 (2012).
17. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
18. Liu, K. *et al.* In vivo analysis of dendritic cell development and homeostasis. *Science* **324**, 392–397 (2009).
19. Schlitzer, A. *et al.* Tissue-specific differentiation of a circulating CCR9$^-$ pDC-like common dendritic cell precursor. *Blood* **119**, 6063–6071 (2012).
20. Satpathy, A.T. *et al.* Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. *J. Exp. Med.* **209**, 1135–1152 (2012).

21. Schlitzer, A. *et al.* IRF4 transcription factor-dependent CD11b$^+$ dendritic cells in human and mouse control mucosal IL-17 cytokine responses. *Immunity* **38**, 970–983 (2013).
22. Persson, E.K. *et al.* IRF4 transcription-factor-dependent CD103$^+$CD11b$^+$ dendritic cells drive mucosal T helper 17 cell differentiation. *Immunity* **38**, 958–969 (2013).
23. Cisse, B. *et al.* Transcription factor E2–2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell* **135**, 37–48 (2008).
24. Sakaue-Sawano, A. *et al.* Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487–498 (2008).
25. Tenenbaum, J.B., de Silva, V. & Langford, J.C. A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323 (2000).
26. Naik, S.H. *et al.* Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229–232 (2013).
27. Lara-Astiaso, D. *et al.* Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943–949 (2014).
28. Besson, A., Dowdy, S.F. & Roberts, J.M. CDK inhibitors: cell cycle regulators and beyond. *Dev. Cell* **14**, 159–169 (2008).
29. Meister, P., Mango, S.E. & Gasser, S.M. Locking the genome: nuclear organization and cell fate. *Curr. Opin. Genet. Dev.* **21**, 167–174 (2011).
30. Hildner, K. *et al.* Batf3 deficiency reveals a critical role for CD8α^+ dendritic cells in cytotoxic T cell immunity. *Science* **322**, 1097–1100 (2008).
31. Kashiwada, M., Pham, N.-L.L., Pewe, L.L., Harty, J.T. & Rothman, P.B. NFIL3/E4BP4 is a key transcription factor for CD8α^+ dendritic cell development. *Blood* **117**, 6193–6197 (2011).
32. Mendelson, A. & Frenette, P.S. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. *Nat. Med.* **20**, 833–846 (2014).
33. Prendergast, A.M. & Essers, M.A.G. Hematopoietic stem cells, infection, and the niche. *Ann. NY Acad. Sci.* **1310**, 51–57 (2014).
34. Essers, M.A.G. *et al.* IFNalpha activates dormant haematopoietic stem cells *in vivo*. *Nature* **458**, 904–908 (2009).
35. Baldridge, M.T., King, K.Y., Boles, N.C., Weksberg, D.C. & Goodell, M.A. Quiescent haematopoietic stem cells are activated by IFN-γ in response to chronic infection. *Nature* **465**, 793–797 (2010).
36. Challen, G.A., Boles, N.C., Chambers, S.M. & Goodell, M.A. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-β1. *Cell Stem Cell* **6**, 265–278 (2010).
37. Mercher, T. *et al.* Notch signaling specifies megakaryocyte development from hematopoietic stem cells. *Cell Stem Cell* **3**, 314–326 (2008).
38. Lewis, K.L. *et al.* Notch2 receptor signaling controls functional differentiation of dendritic cells in the spleen and intestine. *Immunity* **35**, 780–791 (2011).
39. Wu, L. *et al.* RelB is essential for the development of myeloid-related CD8α^- dendritic cells but not of lymphoid-related CD8α^+ dendritic cells. *Immunity* **9**, 839–847 (1998).
40. Kabashima, K. *et al.* Intrinsic lymphotoxin-β receptor requirement for homeostasis of lymphoid tissue dendritic cells. *Immunity* **22**, 439–450 (2005).
41. Satpathy, A.T. *et al.* Notch2-dependent classical dendritic cells orchestrate intestinal immunity to attaching-and-effacing bacterial pathogens. *Nat. Immunol.* **14**, 937–948 (2013).
42. Gatto, D. *et al.* The chemotactic receptor EBI2 regulates the homeostasis, localization and immunological function of splenic dendritic cells. *Nat. Immunol.* **14**, 446–453 (2013).
43. Yi, T. & Cyster, J.G. EBI2-mediated bridging channel positioning supports splenic dendritic cell homeostasis and particulate antigen capture. *eLife* **2**, e00757 (2013).

ONLINE METHODS

Mice. C57BL/6 (CD45.2$^+$) mice were from the Biological Resource Center, Agency for Science, Technology and Research, Singapore. Congenic C57BL/6 (CD45.1$^+$), CX3CR1-GFP and UBC-GFP mice were from the Jackson Laboratory. Fucci-492 mice[24] were from the Riken BioResource Center (Ibaraki, Japan). ZBTB46-GFP BM was provided by K. Murphy. Id2-GFP BM was provided by G. Belz. All mice were bred and maintained in the Singapore Immunology Network animal facility before use at 7–10 weeks of age. Only healthy male mice were used in this study. All experiments and procedures were approved by the Institutional Animal Care and Use Committee of the Biological Resource Center (Agency for Science, Technology and Research, Singapore) in accordance with the guidelines of the Agri-Food and Veterinary Authority and the National Advisory Committee for Laboratory Animal Research of Singapore.

Preparation of cell suspensions. Organs were digested for 30 min in Hank's balanced salt solution HBSS contaiing 10% FBS and collagenase type IV (0.2 mg/ml; working activity of 770 U/mg) (Sigma). BM was flushed from the femur and tibia of one leg and was used without any digestion. For sorting, BM cell suspensions underwent preenrichment for CD135$^+$ cells with biotin–anti-CD135 (A2F10; eBioscience) and anti-biotin microbeads (Miltenyi) and were separated on an AutoMacs (Miltenyi).

Flow cytometry and sorting. Multi-parameter analyses of labeled cell suspensions were performed on an LSR II (Becton Dickinson) and data were analyzed with FlowJo software (TreeStar). A FACSAria II (Becton Dickinson) was used for flow cytometry. Fluorochrome- or biotin-conjugated monoclonal antibodies (mAbs) to the following were used: mouse IA/IE (M5/114.15.2) and CD172α (P84) (both from BD Biosciences); and CD11c (N418), CD45 (30F11), CD45.1 (A20), CD45.2 (104), CD115 (AFS98), Gr-1 (RB6-8C5), Ly6C (HK1.4), Siglec-H (440c), B220 (RA3-6B2), CD135 (A2F10), CD8α (53-6.7) and CD11b (M1/70) (all from eBioscience). The streptavidin–phycoerythrin–cyanine 7 conjugate (25-4317-82) was from eBioscience.

***In vivo* proliferation assay.** The proliferation of individual BM pre-DC subsets was investigated with the Fucci ('fluorescent ubiquitination–based cell-cycle indicator') transgenic mouse model in which the green-emitting fluorescent protein Azami Green is fused to geminin, a ubiquitination oscillator whose expression is regulated by cell-cycle-dependent proteolysis, which results in the fluorescence of cells in S, G2 and M phases and their identification by flow cytometry[24].

***In vitro* development assay for DC progenitor populations.** Flt3L-stimulated BM cultures were generated by seeding of 4.5×10^6 CD45.1$^+$ BM cells into six-well plates in RPMI medium supplemented with L-glutamine (Life Technologies), penicillin-streptomycin (Life Technologies), non-essential amino acids (Life Technologies), 10% FCS (Serana) and β-mercaptoethanol (Sigma-Aldrich). For the induction of DC development 100 ng of mouse Flt3L was added after cells were seeded. After 2 d, 5×10^3 sorted CD45.2$^+$ CDPs or pre-DCs from different subsets were added to individual wells and their CD45.2$^+$ progeny were phenotypically assessed by flow cytometry 2, 3 or 6 d after that 'spike-in'.

Single-cell clonal assay of pre-DC subsets. Irradiated OP9 stromal feeder cells (3×10^3) were cultured for 24 h in supernatants obtained from BM-derived DC cultures at day 8 of stimulation with Flt3L, followed by the addition of single green florescent protein–positive (GFP$^+$) cells (isolated from UBC-GFP mice) from different BM pre-DC subsets. Co-cultures were maintained for 6 d, after which the GFP$^+$ progeny of the single pre-DCs were analyzed for surface molecule expression patterns by flow cytometry.

Intra-femoral transfer of DC precursors. 1×10^4 to 5×10^4 CD45.2$^+$ BM DC precursors (CDPs or the four pre-DC subsets) suspended in PBS were transferred into the femurs of CD45.1$^+$ mice via am insulin syringe with a short needle (Becton Dickinson). At 2, 3 or 5 d after cell transfer, BM and spleen were collected and the resulting cell suspensions were analyzed by flow cytometry to establish the phenotype of the CD45.2$^+$ progeny of the transferred cell populations.

Scanning electron microscopy. For imaging by scanning electron microscopy, sorted cells were allowed to adhere for 15 min at room temperature to glass coverslips pretreated with poly-L-lysine (Sigma), then were fixed for 1 h at room temperature in 2.5% glutaraldehyde and 0.1 M phosphate buffer (pH 7.4) and were washed twice in PBS. After fixation for 1 h at room temperature with 1% osmium tetroxide (Ted Pella), cells were washed in deionized water and dehydrated with a graded series of ethanol immersions from 25% to 100%, and were dried to the critical point (CPD 030; Bal-Tec). The glass coverslip was then laid on adhesive film on a scanning electron microscope sample holder and was firmly touched with an adhesive sample holder. The surface on which the cells were deposited, as well as the adhesive surface, were both coated with 5 nm of gold in a high-vacuum sputtering device (SCD005 sputter coater; Bal-Tec). The coated samples were examined with a field emission scanning electron microscope (JSM-6701F; JEOL) at an acceleration voltage of 8 kV with the in-lens secondary electron detector.

Microarray. Total cellular RNA was extracted with a mirVana miRNA isolation kit (Ambion) and was prepared for microarray (Illumina Mouse WG6) according to the manufacturer's instructions (Illumina). Microarray data were processed by quantile normalization. Hierarchical clustering of samples was achieved with Pearson's correlation and the complete agglomeration method. Comparison of one subset with the rest was carried out with the Limma software package[44] to derive signature genes for each of the four pre-DC subsets. Similarly, Limma was used to select genes that were up- or downregulated in the CD4$^+$ DC subset by comparison of microarray data of thespleen CD4$^+$ DC subset with that of the spleen CD8α$^+$ DC subset (in-house microarray data).

Single-cell capture and library preparation for RNA sequencing. Cell populations isolated by flow cytomtery (MDPs, CDPs and pre-DCs) were diluted to a final concentration range of 250–400 cells per μl and were loaded onto C_1 integrated fluidic circuits IFC (5- to 10-μm chip) for cell lysis, reverse transcription with oligo (dT) primers and amplification of cDNA on a C_1 Single-cell Auto Prep System according to the mRNA-seq protocol of the manufacturer (Fluidigm). Array control RNA spikes were used (1, 4 and 7) (PN AM1781) as instructed in the mRNA-sequencing protocol of the manufacturer (Ambion). The amount of cDNA generated from single cells was quantified with a Quant-iT PicoGreen dsDNA Assay Kit (PN P11496; Life Technologies), and quality was checked with High Sensitivity DNA Reagents (PN 5067-4626) according to the manufacturer's instructions (Agilent Technologies). Only cells with high-quality cDNA were processed for subsequent library preparation. A Nextera_XT Kit (PN FC-131-1096; Illumina) with dual indices (PN FC-131-1002; Illumina) was used for the preparation of single-cell multiplexed libraries, which were sequenced as 51-bp single-end reads on the Illumina HiSeq 2000 platform. 'Reads' were mapped to the reference genome mm9 (NCBI assembly of the mouse genome).

Single-cell mRNA-sequencing data analysis. 'Reads' were mapped against the mm9 reference genome with the STAR alignment program with 'intronMotif' option, and the reference genome was generated with UCSC known gene annotation and the '–sjdbOverhang 50' option. The number of reads mapped to each UCSC known gene was quantified by htseq-count and was normalized by calculation of the RPKM value (reads per kilobase of exon model per million mapped reads)[45–52]. Genes with an RPKM value of 0 in all cells were discarded. RPKM values less than 1 were considered background, and $\log_2$ RPKM was set to 0 as a conservative background cutoff[50,51]. Positive RPKM values then underwent $\log_2$ transformation. Outlier cells were identified and visualized with the Singular Analysis Toolset with default parameters: 14 cells (1 MDP, 8 CDPs and 5 pre-DCs) of the 251 total cells (59 MDPs, 96 CDPs and 96 pre-DCs) were identified as outliers and were excluded from downstream analysis (quality control, **Supplementary Fig. 1b–g**). The entire data set was clustered with all expressed genes and the ANOVA function in the Singular Toolset was used for the identification of genes differentially expressed in MDP, CDP and pre-DC populations. 87 genes had significantly different expression in MDPs, CDPs and pre-DCs, with P values of <0.05. Hierarchical clustering of the 87 differentially expressed genes and 237 non-outlier cells was performed by the Euclidean distance and ward agglomeration method. Multiple-test correction was done by the Benjamini-Hochberg method for multiple-testing correction.

doi:10.1038/ni.3200

On the basis of the hierarchical clustering, genes were grouped into five major clusters; the median $\log_2$ RPKM values per gene cluster were calculated for each cell, and curves smoothed by the lowess ('locally weighted scatterplot smoothing') method were plotted to connect the cells in developmental order as determined by NBOR. Genes encoding products involved in biological functions, pathways and gene ontology that showed enrichment in gene clusters 1–5 were identified with IPA software and the KEGG database, and statistical robustness was tested with the standard hypergeometric test (results in **Supplementary Table 3**). Gene cluster 3, associated with cell cycle or proliferation, was excluded in nonlinear dimensionality reduction carried out via Isomap ('isometric mapping') with the Euclidean distance option[25]. Three nearest neighbors were used for approximation of the geodesic distances between cells for construction of the Isomap plot. The nearest neighbors were indicated by lines connecting cells in the Isomap plot. Expressed genes detected in each cell were quantified by counting of genes with RPKM values >1. Student's *t*-test was used for comparison of the number of genes detected in MDPs versus that detected in CDPs, and in CDPs versus pre-DCs. Unless otherwise noted, all analyses were performed with software of the R project for statistical computing, version 2.15.2 (Bioconductor).

Reordering of single cells from unsupervised clustering into the order of DC development with a neighborhood-based ordering strategy. We performed hierarchical clustering of differentially expressed genes from all single cell transcriptomes (**Supplementary Table 2**) and cells by Euclidean distance and the Ward agglomeration method[53]. Cells clustered into five main groups: four corresponding to MDPs, CDPs and two pre-DC clusters; and one comprising a mixture of MDPs, CDPs and pre-DCs. The average expression per gene for each of the first four clusters (MDP, CDP, pre-DC 1 and pre-DC 2) was calculated from all of the single cells in the respective clusters. These averages represent canonical cell states for MDPs, CDPs and pre-DCs (1 and 2), and are called 'landmarks' here (**Supplementary Fig. 3**). For each individual cell, we calculated the Euclidean distance to each landmark; the two landmarks nearest a single cell were thus called the 'nearest-neighbor pair' for that cell. To reorder single cells in the order of DC development, we placed the four landmarks MDP, CDP, pre-DC 1 and pre-DC 2 in a linear order and then allocated each single cell a place between its nearest-neighbor pair of landmarks. We decided the exact position of each single cell between the landmark pair on the basis of their Euclidean distances to the landmarks. By this reordering of single cells, we generated a heat map (**Fig. 1a**) to show the progression of gene expression along the DC development continuum from MDP to pre-DC.

CMap analysis. CMap analysis is an extension of the gene-set–enrichment analysis algorithm (provided by the Broad Institute) in which 'enrichment' of a gene set (signature genes) in another gene set can be measured. CMap scores are scaled dimensionless quantities that indicate the degree of enrichment or 'closeness' of one assessed cell subset to another[17]. cDC and macrophage signature genes were identified from the literature[15,16] and our own transcriptomic data and were used as signature genes for the respective populations for CMap analysis of each single cell[17]. The 'enrichment' of gene sets was tested with 1,000 permutations. Cells whose gene-expression profile was significantly correlated with those of signature genes were selected by a P value of <0.05 after 1,000 permutations. CMap scores were scaled to a range –1 to 1. Cells with positive CMap score are correlated with cDCs; cells with negative CMap scores are correlated with macrophages (**Fig. 1c**). Similarly, signature genes of cDC2 and cDC1 subsets were used to identify DC subset–primed cells at the MDP, CDP and pre-DC stages (**Fig. 1d**) Signatures genes of Siglec-H$^-$Ly6C$^+$, Siglec-H$^-$Ly6C$^-$, Siglec-H$^+$Ly6C$^+$ and Siglec-H$^+$Ly6C$^-$ pre-DC were derived from microarray data (**Fig. 3**) and were used for CMap analysis (**Fig. 5a**,**b**). Detailed gene lists of signatures are in **Supplementary Tables 4–7** and **9–12**.

44. Gentleman, R., Carey, V., Huber, W., Irizarry, R. & Dudoit, S. Bioinformatics and computational biology solutions using R and Bioconductor. *Brief. Bioinformatics* **8**, 136–137 (2006).
45. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
46. Adiconis, X. *et al.* Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat. Methods* **10**, 623–629 (2013).
47. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
48. Pollen, A.A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
49. Usoskin, D. *et al.* Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat. Neurosci.* **18**, 145–153 (2015).
50. Shalek, A.K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 1–22 (2014).
51. Shalek, A.K. *et al.* Single-cell transcriptomics reveals bimodality inexpression and splicing in immune cells. *Nature* **498**, 236–240 (2014).
52. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
53. Ward, J.T.J. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–244 (1963).