Dendritic-cell diversity in equine blood revealed by single-cell transcriptomics

Marius Botos[1,2]*, Ambre Baillou[1,2]*, Simone Oberhaensli [3†], Iva Cvitas [4††], Sigridur Jonsdottir[4†††], Anja Ziegler[4 ††††], Francisco Brito[1,2], Artur Summerfield[1,2], Eliane Marti [4], Stephanie C. Talker [1,2]

[1]*Institute of Virology and Immunology, Bern, Switzerland*
[2]*Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland*
[3]*Interfaculty Bioinformatics Unit and Swiss Institute of Bioinformatics, University of Bern, Bern, Switzerland*
[4]*Department of Clinical Research and Veterinary Public Health, Vetsuisse Faculty, University of Bern, Bern, Switzerland*

*These two authors contributed equally to this work

[†] *present affiliation: Next Generation Sequencing Platform and Institute for Fish and Wildlife Health, Vetsuisse Faculty, University of Bern, Bern, Switzerland*
[††] *present affiliation: ITPA, Vetsuisse, University of Bern*
[†††] *present affiliation: Institute for Experimental Pathology, Biomedical Center, University of Iceland, Keldur, Reykjavik, Iceland*
[††††] *present affiliation: Ennetseeklinik für Kleintiere, Hünenberg, Switzerland*

Corresponding author:
Stephanie Talker
Länggassstrasse 122, 3012 Bern, Switzerland
Tel: +41 31 684 2483
stephanie.talker@unibe.ch

ABSTRACT

Unbiased classification of equine dendritic cells (DC) is necessary to address various research questions such as the role of DC subsets in immune-mediated diseases of horses. We applied single-cell RNA sequencing (scRNA-seq) on DC enriched from blood of two horses, based on expression of CD172a, Flt3, CADM1 and CD14. All main DC subsets were detected based on key gene expression, including cDC type 1 (cDC1; *XCR1*), cDC2 (*FCER1A*, *CD1E*) and plasmacytoid DC (pDC; *TCF4*). In addition, we detected a small cluster of CD34-expressing DC progenitors. Our data confirms the previously reported phenotype of equine pDC (MHC-II^{low}CADM1^{int}CD172a^{int}), cDC1 (MHC-II^{high}CADM1^{high}CD172a$^{low-int}$) and cDC2 (MHC-II^{high}CADM1^{int}CD172a^{high}), while also highlighting considerable CD14 expression for cDC2. Among Flt3^{+} cells clustering with cDC2, we identified a cluster resembling monocytes and showing a highly pro-inflammatory signature, likely representing DC type 3 (DC3). Notably, one cDC2-associated cluster had a mixed pDC/cDC2 signature (*TCF4*, *SPIB*, *FCER1A*), indicating the presence of transitional DC (tDC), a new DC subset initially described in human and mouse, and more recently in pig. To assess cross-species conservation of DC subsets, we compared equine and porcine DC scRNA-seq datasets using SATURN, a deep learning method that combines gene expression with added biological knowledge encoded in protein language models. This enabled mapping of the most similar DC subsets between horse and pig, confirming the conservation of key transcriptomic features and supporting the identification of equine tDC. Our atlas of equine blood DC is a valuable resource for comparative analyses, and it forms the foundation for elucidating the role of DC subsets in immunological diseases such as type I hypersensitivity in horses.

Key words: equine blood dendritic cells, single-cell RNA-seq, tDC, DC3, SATURN

INTRODUCTION

Dendritic cells (DC), despite their low frequency in tissues and blood, are central players in the immune system, initiating and fine-tuning adaptive immune responses (1)Traditionally, DC have been classified as conventional DC (cDC) and plasmacytoid DC (pDC). Conventional DC are grouped into two main subtypes: cDC1, specialized in cross-presenting antigens through MHC class I molecules, enabling them to prime CD8 T cells, and cDC2, which appear specialized in promoting Th2 and Th17 responses (2). In contrast to cDC, pDC are described to be at least partly of lymphoid origin (3,4). While the antigen presenting capabilities of pDC are being debated, they are known for their potent antiviral effect by rapidly producing type I interferon (5), also in horse (6). Flow cytometry (FCM) is widely employed for the characterization of DC subsets across

species (7). However, studies in veterinary species, in contrast to human and mouse, face several limitations such as limited availability of species-reactive and fluorochrome-coupled monoclonal antibodies (8). Furthermore, capturing the full spectrum of DC diversity and plasticity by a set of pre-defined markers in FCM is not feasible and prone to bias. Delineating cDC2 subsets, monocyte-like DC and monocytes is a challenging task due to overlapping phenotypes and transcriptomes, especially when studying inflammatory processes (9). Single-cell RNA-sequencing (scRNA-seq) has enabled the identification of novel DC subsets, including DC type 3 (DC3) (10) and transitional DC (tDC) in human (11), mouse (12,13), and pig (14). Across species investigated so far, DC3 have been described to share transcriptomic signatures with both cDC2 and monocytes (12,15,16), reflecting their origin from monocyte-dendritic-cell progenitors (10,12,17). In contrast, tDC were shown to originate from pro-pDC (13) and to differentiate into cDC2-like cells, with improved ability to prime T cells and to produce pro-inflammatory IL-1β (13). In horse, the recent scRNA-seq study of PBMC by Patel *et al.* confirmed the identification of cDC1, cDC2, and pDC in blood, but did not extend investigations on DC heterogeneity, likely due to the small number of DC analyzed (18).

Here, we describe previously unreported heterogeneity of DC in equine blood. By enriching DC prior to scRNA-seq, we were able to detect four DC subsets (cDC1, cDC2, pDC, tDC) alongside putative DC3 and DC progenitors. Having recently described the same DC subsets in blood of pig (14), we applied the deep-learning tool SATURN (Species Alignment Through Unification of Rna and proteiNs) (19) to integrate the DC-enriched scRNA-seq datasets of pig and horse. Among available tools for (cross-species) integration (20–22), SATURN stands out by adding additional information encoded in protein-language models, and, most noteworthy, circumvents the need for overlapping gene annotations in different species. Using this tool, all DC subsets aligned across horse and pig, further supporting the identification of equine tDC. Apart from the description of novel DC subsets, our dataset is a valuable resource for advancing research in equine and comparative immunology.

MATERIALS AND METHODS

Isolation of equine PBMC

Blood (100 mL) of three horses (**Table 1**) was collected into vacutainer collection tubes (Vacuette®; Greiner, St.Gallen, Switzerland) by puncturing the jugular vein, with sodium heparin as anti-coagulant. Peripheral blood mononuclear cells (PBMC) were isolated by density gradient centrifugation over Biocoll Separating Solution (density= 1.077 kg/l; Bioswisstec AG,

95 Schaffhausen, Switzerland). Erythrocytes were sedimented for 30 min at room temperature (RT)
96 before the upper leukocyte-rich plasma was collected and overlayed on Biocoll, and then
97 centrifuged at 1,800 x g for 20 min (RT). The interphase layer with PBMC was transferred to a new
98 50 mL tube and washed twice with PBS, with first centrifugation for 10 min at 290 x g and second
99 centrifugation for 10 min at 200 x g (RT). Finally, PBMC were resuspended in PBS before cell
100 sorting. The study was approved by the Animal Experimental Committee of the Canton of Berne,
101 Switzerland (No. BE 2/17).
102
103 Table 1. Horses used for 10X single-cell RNA-sequencing.

	Breed	**Age (years)**	**Sex**	**Cells after filtering scRNA-seq data**
Horse 1	Freiberger	10	male	2621
Horse 2	Freiberger	12	female	4374
Horse 3	Icelandic	8	female	238 (excluded from analysis)

104 **Phenotyping of equine PBMC and DC enrichment by FACS**

105 Freshly isolated equine PBMC were stained in 96-well U-bottom plates with 10 × 10^6 cells per
106 sample (phenotyping) or in 5 mL FACS tubes (sorting experiments) with 80 x 10^6 cells per tube
107 (sorting experiments). Antibodies against CD172a, CADM1, CD14, and CD163 (phenotyping only)
108 were used, alongside recombinant Flt3L. Bovine recombinant Flt3L-His, cross-reactive with
109 equine Flt3, was produced in house as previously described (6,23). All incubations were
110 performed for 20 min at 4°C, followed by washing with Cell Wash (BD Biosciences) and
111 centrifugation at 500 x g for 5 min. Prior to the staining, all samples were incubated with equine
112 IgG (Biorad; PEP001) to block Fc receptors. For phenotyping, cells were first incubated with anti-
113 CD163 (LND68A; mIgG1) and anti-CADM1 (3E1; IgY), followed by incubation with matching
114 secondary antibodies anti-mIgG1-PerCPeF710 and anti-IgY-biot. In the next step, ChromPure
115 mouse-IgG (Jackson ImmunoResearch) was added to block remaining binding sites of the anti-
116 mIgG1 secondary antibody, alongside Streptavidin-BV421 (BD Horizon, 563259) and Flt3L-His
117 (Miltenyi Biotec GmbH). In the last step, ZenonAF488-labeled anti-CD14 (105; mIgG1) was added
118 alongside ZenonAF647-labeled anti-CD172a (DH59B; mIgG1), anti-His-PE (mIgG1; Miltenyi
119 Biotec GmbH), and LIVE/DEAD™ Fixable Near-IR stain (Thermo Fisher Scientific). Zenon labeling
120 kits were derived from Thermo Fisher Scientific. For enrichment of DC by FACS, cells were first
121 incubated with anti-CD172a (DH59B; mIgG1) and with anti-CADM1 (3E1; IgY), followed by
122 incubation with corresponding secondary antibodies anti-IgG1-A647 (Invitrogen, A-21240) and
123 anti-IgY-biotin (MILAN ANALYTICA, 103-065-155). Thereafter, cells were incubated with

124 ChromPure mouse-IgG (Jackson ImmunoResearch), Streptavidin-BV421 (BD Horizon, 563259)
125 and Flt3L-His (Miltenyi Biotec GmbH), and in a final incubation step with ZenonAF488-labeled
126 anti-CD14 (105; mIgG1), anti-His-PE (mIgG1; Miltenyi Biotec GmbH), and LIVE/DEAD™ Fixable
127 Near-IR stain (Thermo Fisher Scientific). Zenon labeling kits were derived from Thermo Fisher
128 Scientific. Samples were acquired with a FACSCanto II flow cytometer (BD Biosciences)
129 equipped with three lasers (405, 488, and 633 nm). Compensation was calculated by FACSDiva
130 software using single-stained samples. Data were analyzed using FlowJo software (v10.10, BD
131 Biosciences). Sorting (FACS) was performed on a MoFlo ASTRIOS EQ equipped with five lasers
132 (355, 405, 488, 561, and 640 nm). A minimum of 30'000 cells was sorted. Cell counting and
133 viability assessments were conducted on a TC20 automated cell counter (Biorad, 1450102) using
134 Trypan Blue dye 0.40 % (Biorad, 1450021).

135 **Single-cell RNA sequencing**

136 Library preparation was done at the Next Generation Sequencing (NGS) Platform at the University
137 of Bern. The three libraries were stored at minus 80°C, before being pooled and sequenced in one
138 run. GEM generation & barcoding, reverse transcription, cDNA amplification and 3' gene
139 expression library generation steps were all performed according to the Chromium Single Cell 3'
140 Reagent Kits v3 User Guide (10x Genomics CG000183 Rev B) with all stipulated 10x Genomics
141 reagents. Specifically, cell suspensions of between 700-1100 cells / µL were prepared for the
142 GEM generation and barcoding step with a targeted cell recovery of 5'000 cells. After the GEM-
143 reverse transcription incubation, a clean-up step was performed and then 12 cycles of cDNA
144 amplification. The resulting cDNA was assessed for quantity and quality using a Thermo Fisher
145 Scientific Qubit 2.0 fluorometer with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific,
146 Q32854) and an Advanced Analytical Fragment Analyzer System using a Fragment Analyzer NGS
147 Fragment Kit (Agilent, DNF-473), respectively. Thereafter, 3' gene expression libraries were
148 constructed using a sample index PCR step of 11 cycles. At the end of the protocol, an additional
149 0.7x bead-based clean-up of the libraries was done. The produced cDNA libraries were tested for
150 quantity and quality using fluorometry and capillary electrophoresis as defined above. The cDNA
151 libraries were pooled and sequenced with a loading concentration of 300 pM, paired end and
152 single indexed, on an illumina NovaSeq 6000 sequencer with a shared NovaSeq 6000 S2 Reagent
153 Kit (100 cycles; illumina 20012862). The read set-up was as follows: read 1: 28 cycles, i7 index: 8
154 cycles, i5: 0 cycles and read 2: 91 cycles. The quality of the sequencing runs was assessed using
155 illumina Sequencing Analysis Viewer (illumina version 2.4.7) and all base call files were
156 demultiplexed and converted into FASTQ files using illumina bcl2fastq2 conversion software

157 version 2.19.0. An average of 543’231’771 reads/library were obtained. All steps were performed
158 at the Next Generation Sequencing Platform, University of Bern.

159 **Genome reference and alignment**

160 Raw data alignment of FASTQ files and UMI counting were performed using Cell Ranger (10X
161 Genomics, version 7.1). We used equCab3.0 (NCBI GCA_002863925.1) as the reference genome
162 and Ensembl EquCab3.0.111 as annotation. The reference obtained was used to align the FASTQ
163 files from three horse samples with the default arguments, except for "--include-introns = T". The
164 human gene names *HLA-DRA* and *HLA-DOB* found in the horse genome annotation were
165 replaced by *ELA-DRA* and *ELA-DOB* respectively, the gene names of their equine orthologs.
166 The sample from “Horse 3 “was excluded from further analysis (estimated number of cells 238).
167 CellRanger reports are given in **Supplementary Figure 1**.
168 For integration with the publicly available dataset from Patel et al. (scRNA-seq equine PBMC), our
169 dataset was re-processed using Ensembl v95 and End Sequence Analysis Toolkit (ESAT)(24), as
170 previously performed and described by Patel et al. (18).

171 **Quality control and filtering**

172 Scanpy (25) was used for the downstream analysis of the scRNA-seq data. Cells were marked as
173 outliers following single-cell best practices (26) by using lenient thresholds in order to avoid
174 excluding smaller subpopulations and filtering out cells that have median absolute deviations
175 (MAD) < 3 or > 5 for the log transformed total number of counts in a cell (sequencing depth), the
176 log transformed number of genes detected in a cell (library complexity), and the percentage of
177 total RNA counts in a cell that come from the 20 most highly expressed genes (technical artifacts
178 if too high value). Thus, cells with mitochondrial genes > 10% and a number of MAD < 5 or > 5, as
179 well as cells with hemoglobin genes > 5% and a number of MAD < 8 or > 3, were filtered out. Next,
180 each sample was individually reassessed and processed for removal of doublets using scrublet
181 (27). PCA was applied to each sample to visualize the scores and results of the filtering steps
182 before integration.

183 **Integration of equine data, clustering, and ranking of differentially expressed genes**

184 The tool scVI (Single-cell Variational Inference) (28) was used to integrate and batch-correct the
185 data based on the batch covariate (sample) with standard parameters except for the “n_latent =
186 30” and “n_layers = 2” arguments, in order to obtain a larger number of embeddings. These
187 embeddings were the input for UMAP, and clustering was performed with the Leiden algorithm at
188 different resolutions. To obtain the ranking of highly differentially expressed genes in each cluster

189 at different resolutions, Wilcoxon rank-sum was used with the scanpy default function
190 "sc.tl.rank_genes_groups", on the normalized and log transformed counts. Inclusion of the extra
191 argument "pts=True" computed the fraction of cells expressing the genes, found as column
192 names pts.1 and pts.2, reported in the differential expressed gene list.
193 For subsetting and integrating our dataset with the publicly available scRNA-seq dataset from
194 equine PBMC, we followed the code published by Patel et al. alongside their manuscript (18).
195 Briefly, we used SCT for normalization and performed integration on the 5,500 most variable
196 genes. Unsupervised graph-based clustering was performed on the top 40 principal
197 components.

198 **Gene Ontology Biological Processes & Hallmarks Over-Representation Analysis**

199 Decoupler (29) was used to find over-represented gene signatures from "gene ontology biological
200 processes (go-bp)" and "hallmarks" from the Molecular Signature Database (MSigDB) (30) with
201 the function "run_ora". Wilcoxon rank-sum test was used to select the top "go-bp" and
202 "hallmarks" gene signatures.

203 **Across-species gene signature enrichment analysis using AUC**

204 Decoupler (29) was used to calculate the enrichment score of specific gene signatures, with its
205 package function "run_aucell". The function performs area under the curve (AUC) analysis on the
206 non-transformed count values obtained from the Cellranger alignment step. Gene signature sets
207 were obtained from publicly available scRNA-seq data of DC populations from human (11),
208 mouse (31), and pig (14). We tested different parameters for the percent of top genes to be
209 selected in each single cell. Percentages were increased to 75 % for signatures that contained a
210 low number of genes. The following parameters were used in the final analysis: for porcine cDC2
211 and tDC signatures we used 75 %, for all other porcine DC subsets (cDC1, pDC, DC3) we used
212 50%. For murine pDC we used 75 %, for all other murine DC subsets we used 5%. For human tDC
213 we used 75%, for all other human DC subsets we used 50%.
214 The enrichment score distributions across cells were visualized using histograms, allowing us to
215 set a threshold for selecting enriched cells. UMAP representation was used for visualizing
216 enrichment score in scRNA-seq data, from non-enriched cells (blue) to highly enriched cells
217 (yellow) (**Supp. Fig. 2**).

218 **Integration of equine and porcine DC subsets using SATURN**

219 Given the availability of in-house generated scRNA-seq data of enriched DC from equine and
220 porcine blood, we applied SATURN (19) to obtain cell embeddings that couple "biological

221 properties" of protein language models (pLMs) with RNA expression. We used SATURN with its
222 default ESM-1b model (esm1b_t33_650M_UR50S) (32) following a three-step protocol: first, we
223 created the equine and porcine protein embeddings, taking the protein translations/conversion
224 of Ensembl genome (release 111) for both species. Second, we added layers for raw counts for
225 each scRNA-seq dataset, and the cell type labeling to be used for integration. Additionally, we
226 randomly subset the porcine dataset to use the same number of cells as in the equine dataset
227 (5.5 k cells). Third, the following arguments of SATURN were adjusted to account for the size of
228 the datasets: (i) "pretrain_batch_size=64" and "batch_size=64", being the closest value to our DC
229 population with lowest number of cells, (ii) " model_dim=64" and "hidden dim=64", following the
230 logic of the default arguments, we set the same number of features to be extracted for the
231 integrated object, and (iii) "scale_expression=True". After integration, we examined the integrated
232 data by PCA to assess the overlap of cells by species and cell subsets.

233 **Preparation of Figures**

234 Figures were prepared using FlowJoTM v10.10.0 (BD Life Sciences), R v4.3.3, Rstudio v2024.04.1,
235 scanpy v1.10.3, matplotlib v3.9.2, seaborn v0.13.2, decoupler v1.8.0, Adobe Illustrator, and
236 Inkscape v1.3.2 (https://www.inkscape.org).

237 RESULTS

238 **Phenotyping of equine DC and monocytes**

239 As dendritic cells constitute a minor population within PBMC, it was necessary to enrich them by
240 fluorescence-activated cell sorting (FACS) prior to scRNA-seq. We analyzed expression of Flt3,
241 CD172a, CD14, CADM1 and CD163 to develop a gating strategy with optimal DC enrichment and
242 exclusion of (most) monocytes.
243 Visualizing Flt3 vs. CD172a expression did not allow for a clear delineation of DC and monocytes
244 due to high autofluorescence of the latter (**Fig. 1A**). Therefore Flt3-enriched cells were further
245 sub-gated to visualize expression of CD14 vs. CADM1. Within the resulting five populations,
246 putative DC and monocyte subsets were defined. While cDC1 ($CD14^{low}CADM1^{high}$) and pDC
247 ($CD14^{neg}CADM1^{int}$) were negative for *CD163*, cDC2 ($CD14^{int}CADM1^{int}$) expressed low levels of
248 *CD163*. Expression of CD172a was lowest on pDC, heterogenous on cDC1, and high on cD2.
249 Putative classical monocytes (cM, $CD14^{int\text{-}high}CADM1^{low}$) expressed high levels of CD163 and
250 CD172a. Putative non-classical monocytes (ncM, $CD14^{high}CADM1^{neg\text{-}low}$) expressed low levels of
251 CD163, the highest level of CD172a, and were lower in forward scatter (FSC-A), as reported for
252 ncM in blood of cattle (33). Analysis of *CCR7* upregulation following Resiquimod stimulation of

equine PBMC, as previously performed for bovine PBMC (34), further supports the gating strategy to enrich DC, as gated monocytes (putative cM and ncM) failed to upregulate *CCR7* in contrast to DC subsets (data not shown).

Single-cell RNA-seq of DC enriched from equine PBMC

Dendritic cells were enriched from freshly isolated equine PBMC of three horses and scRNA-seq (3' gene expression, 10X Genomics) was performed immediately after FACS-enrichment. As shown in **Figure 1B**, scRNA-seq of enriched DC resulted in 7 clusters (resolution 0.45). One horse was excluded from further analyses due to the low number of cells (**Table 1**, **Supp. Fig. 1**). For the remaining two horses, a total of 5,581 cells passed quality filtering (**Supp. Fig. 3**). Cluster proportions are shown in **Supp. Fig 3.** The complete and original gating strategies for Horse 1 and Horse 2 are shown in **Supplementary Figure 4**.

By visualizing the expression of key genes, we identified cluster 1 as cDC1 (*XCR1*, *BATF3*, *IRF8*, *CLEC9A*, *CLNK*, *CADM1*), clusters 0, 3, 4, and 5 as cDC2 (*FCER1A*, *CD1C*, *CD1A2*, *CSF1R*, *SIRPA*), cluster 2 as pDC (*TCF4*, *BLNK*, *IRF7*, *SPIB*), and cluster 6 as hematopoietic progenitors (*CD34*) (**Fig. 2B-D**). Notably, *TCF4* was also expressed by a subset of cells clustering as cDC2. These *TCF4*-expressing cells co-expressed *FCER1A* and appeared to exclusively lack *CSF1R* expression. Differential gene expression analysis (DEA; **Fig. 2E**) revealed additional markers for cDC1 (e.g. *CPVL*, *CST3*, *LRRK2*, *ID2*), cDC2 (e.g. *S100A4*, *MNDA*, *GIMAP5*, *TYROBP, IFI16, IL32*) and pDC (e.g. *CTSC*, *UGCG*, *FYB1*, *CLECL1*) (**Fig. 2D**). Furthermore, hematopoietic progenitors (c6) were enriched for *MEIS1*, *DACH1* and *MSI2* transcripts (**Fig. 2B**), as reported for mice (35,36). Complete gene lists from DEA are given in **Supplementary File 1.**

Gene expression patterns for *Flt3, CD14, CD172a, CADM1,* and MHC-II in the cDC1, cDC2 and pDC clusters (**Fig. 2F**) were in line with surface protein expression detected in flow cytometry (**Fig. 1**) and with previous phenotyping performed by Ziegler *et al.* (6) (**Fig. 1A**).

Delineation of transitional DC (tDC) and putative DC3 from cDC2

Four distinct clusters were apparent within cDC2 (**Fig. 3A**). Two clusters (c0 and c3) appeared in fact to be comprised of cDC2, as judging from the visualization of *CD1E* and *CD1A2* expression in feature plots (**Fig. 3B**). One cluster (c4) was identified as tDC according to its expression of pDC-associated genes (*TCF4*, *SPIB1*, *BLNK*) alongside cDC2-associated genes (**Fig. 3C**). Notably, one cluster (c5) displayed a monocyte-like signature with high expression of pro-inflammatory genes (*S100A12*, *FCN1*) (**Fig. 3D**). Based on co-expression of *FLT3* and *CD163* / *CSF1R*, we annotated this cluster as putative DC3 (**Fig. 3E**). Over-representation analysis confirmed the pro-

285 inflammatory signature of the putative DC3 cluster and additionally revealed an enrichment of
286 complement-associated transcripts (**Fig. 3F**).
287 Differential gene expression analysis was performed to determine specific signatures of cDC2
288 (c0&c3), tDC and DC3, as well as signatures differentiating tDC or DC3 from cDC2 (c0&3) (**Fig.**
289 **3G**). Expression of MHC-II and CD1 family genes was enriched in cDC2 clusters (c0, c3). Among
290 the top genes enriched in tDC (c4) was *DPP4*, encoding the peptidase CD26 that is reported to
291 regulate chemokine activity and thereby migration during inflammation (37). Moreover, *MS4A3*
292 and *PLXNA4* transcription was found to be enriched in equine tDC, when compared to cDC2 and
293 DC3. Notably MS4A3 is reported to regulate the cell cycle and to be specifically expressed in
294 myeloid precursors (38) Plexin 4 (*PLXNA4*) is described as a co-receptor for class 3 semaphorins,
295 suggesting an important role for semaphorin signaling in tDC, such as the reported dependence
296 on semaphorin 3A signaling for DC transmigration into the lymphatic system (39). Heatmaps and
297 selected violin plots of top enriched genes in the main DC subsets are shown in **Figure 3G-H**.
298 Complete gene lists are given in **Supplementary File 1**.
299 Among top gene transcripts enriched in putative DC3 were *LAMP1*, *FCN1*, *LYZ* and *S100A12*, i.e.
300 genes commonly associated with pro-inflammatory monocytes.

301 **Delineation of cDC2 from monocytes**

302 Precise delineation of cDC2 and DC3 from monocytes is challenging across species. As our DC-
303 enriched dataset likely contains very few monocytes, we took advantage of a previously
304 published scRNA-seq dataset from equine PBMC (18) to integrate equine monocytes with our
305 dataset (**Supp. Fig. 5**).
306 To prepare our dataset for integration with the publicly available dataset from Patel et al., our
307 dataset was re-processed using Ensembl v95 with the End-Sequence Analysis Toolkit (ESAT) (24).
308 As described in Patel et al., ESAT processing was necessary with the previous Ensembl
309 annotation (Ensembl v95) to rescue reads mapping to poorly annotated 3' untranslated regions
310 (18). Also in our dataset, ESAT-processing significantly increased the number of detected genes
311 for this previous annotation (Ensembl v95) (**Supp. Fig.5**).
312 We identified 12 distinct clusters among the integrated dataset (**Fig. 4A**), comprising co-
313 clustering of DC and monocytes from both datasets, as indicated by the visualization of DC- and
314 monocyte-related genes (**Fig. 4B** and **4C**).
315 Monocytes from the Patel dataset clustered to the left (c1, c3, c4, c6, c7, c10), expressing high
316 levels of *CD14* and *LYZ* (**Fig. 4D**). Cells from our DC-enriched dataset clustered mostly to the right
317 of the UMAP plot (c0, c2, c5, c8, c9, c11, c12), alongside *FLT3*-expressing cells from the Patel

318 dataset. Notably, some cells from our dataset clustered with monocytes, but expressed *FLT3*
319 alongside *CD14* and *LYZ* (**Fig. 4E**), potentially indicating DC3 identity.

320 **Heterogeneity of cDC2**

321 Two clusters of cDC2 were evident in our dataset (**Fig. 5A**; cDC2.1 and cDC2.2) that may
322 resemble cDC2 subsets described for human and mouse. While cDC2.1 (c0) expressed higher
323 levels of *FCER1A*, cDC2.2 (c3) were clearly enriched in *CX3CR1* transcripts (**Fig. 5B**)
324 Notably, expression of CX3CR1 distinguishes murine T-bet cDC2 (cDC2B) from T-bet^{+} cDC2
325 (cDC2A) (40,41).
326 Accordingly, our cDC2.2 may correspond to murine cDC2B, which were described to have a more
327 pro-inflammatory profile and to be more capable of T-cell stimulation (41), as opposed to cDC2A
328 that appear to be involved in tissue repair.
329 The reported higher ability of murine cDC2A to stimulate naive T-cells, would be in line with higher
330 expression of MHC-II –related genes (*ELA-DRA*, *CD74*) on equine cDC2.2. (**Figure 5C**).
331 However, overall, equine cDC2.1 appeared more similar to pro-inflammatory monocytes, as
332 shown by their higher expression of *CCR2* and *GAPDH*. Moreover, expression of *TLR2* was higher
333 in CDC2.1, suggesting better recognition of bacterial lipoproteins by this subset. In line with high
334 *CX3CR1* expression in cDC2.2, gene expression for CD49d (*ITGA4*), acting as part of a coreceptor
335 for fractalkine (CX3CL1), was higher in this subset (39). Moreover, transcripts for the co-inhibitory
336 receptor CD31 (*PECAM1*) were enriched in CDC2.2, suggesting tolerogenic functions of cDC2.2
337 (43). Top 25 differentially expressed genes between cDC2 subclusters are given in **Figure 5D.** The
338 complete gene list is given in **Supplementary File 1.**
339 Taken together, gene transcription suggests pro-inflammatory classical monocyte-like functions
340 for cDC2.1, and T-cell modulatory functions reminiscent of nonclassical monocytes in cDC2.2.
341 This, however, would suggest that CX3CR1 expression on equine cDC2 subsets cannot be used
342 to differentiate subsets analogous to the cDC2 subsets described for mouse and human.
343

344 **Gene expression related to main DC functions**

345 Genes associated with DC functions, such as pattern recognition, migration and adhesion,
346 antigen presentation and T-cell modulation, as well as genes for other categories of interest, were
347 visualized for all seven analyzed DC subsets (**Fig. 6** and **Supp. Fig. 6**).
348 Among genes encoding pattern recognition receptors, *TLR4* transcripts were almost exclusively
349 detected in putative DC3, while *TLR9* transcripts were almost exclusively detected in pDC.
350 Additionally, both pDC and DC3 expressed the highest levels of *TLR7* (**Fig. 6B**). Plasmacytoid DC

351 expressed the highest levels of *CXCR3*, *SELL* (CD62L), *PECAM1* (CD31), *ITGAL* (CD11a), *ITGB7*,
352 and *ITGAE* (CD103), while cDC1 expressed the highest levels of *ITGA4* and *ITGB1* (**Fig. 6C**).
353 Expression of *ITGAX* (CD11c) was clearly enriched in cDC2 and putative DC3.
354 Among genes encoding co-stimulatory molecules (*CD80*, *CD83*, *CD86*), CD2.2 appeared to be
355 enriched in transcripts for *CD83* (**Fig. 6D**). Expression of CD86 was uniform across subsets, apart
356 from DC progenitors and pDC, where lower levels were detected. Expression of MHC-II-related
357 genes was highest in cDC1 and cDC2.2, while gene expression for CD1 molecules showed a more
358 complex pattern: cDC2, in particular cDC2.1, expressed the highest levels of *CD1A2*, *CD1C* and
359 *CD1E*, while transcripts for *CD1E1* were enriched in pDC. Notably, pDC also expressed
360 considerable levels of *CD1C*. As expected, *WDFY4* (involved in cross-presentation) showed the
361 highest expression in cDC1. Notably, the second highest level for *WDFY4* was detected in pDC.
362 Furthermore, plasmacytoid DC stood out by expressing the highest level of *IL10RB*, while DC3
363 showed very high expression of *LYZ* (**Supp. Fig. 6A**).
364 Overall, DC subsets expressed similar levels of glycolysis-related and other metabolic genes,
365 with the notable exception of the very high *SDHB* expression in pDC and the high *GAPDH*
366 expression in cDC2.1 (**Supp. Fig. 6**).

367 **Species comparisons support identification of equine tDC**

368 To explore the conservation of our DC populations across species, we first examined the relative
369 enrichment of DC signatures from human blood (11), murine spleen (31) and porcine blood (14)
370 in our equine scRNA-seq clusters by gene-set enrichment analysis (GSEA). Overall, pDC- and
371 DC3-derived signatures were specifically conserved between each compared species and horse
372 (**Fig. 7A-C**). Interestingly, porcine and human cDC2 signatures exhibited higher enrichment
373 scores in the equine cDC2.2 subcluster (c3), but were not enriched in the main cDC2.1
374 subcluster (c0) (**Fig. 7A** and **7C**). Moreover, murine gene signatures from cDC2 and tDC showed
375 the highest enrichment scores in clusters 0 and 3 (cDC2), and cluster 4 (tDC), respectively, thus
376 indicating the conservation of DC subset signatures between mouse and horse (**Fig. 7C**)
377 However, in comparison to pig and human gene sets, the murine cDC1-derived gene set showed
378 less specificity for equine cDC1 (c1), as shown by the high enrichment score in both cDC1 (c1)
379 and cDC2 (c0 and c3).
380 Next, we explored SATURN, a deep learning integration tool based on protein-language models
381 (19), to compare equine DC to our in-house generated dataset of porcine DC (14). Notably,
382 SATURN is described to bypass the limitation of one-to-one orthologs and the loss of information
383 that can be associated to species comparisons based on gene signatures. The selection of 6
384 principal components (PC), explaining 98.4 % of the variance in the dataset (**Supp. Fig. 7)**

resulted in co-clustering of corresponding DC subsets from the two species (**Fig. 8A**). More specifically, we found that cDC2 of horse and pig overlapped, while for the other DC subsets, cells from horse and pig came close and partially overlapped. Notably, the clustering of equine tDC with porcine tDC supports their classification. Surprisingly, with increasing number of PC used for the integration, cDC1 and pDC increasingly separated from their counterparts, while cDC2 of both species remained well integrated (**Supp. Fig. 7**).

To investigate the consequences of PC reduction in another dataset, we re-analyzed a dataset from a recently published benchmarking study (22). The authors used the default number of PC (i.e. 30) for integration between species and for some of their smallest datasets (e.g. human & pig, monkey & pig), SATURN performed relatively poorly in terms of integrating corresponding cell types. By reducing the number of PC in our re-analysis, we obtained good integration (overlapping cell types) also for these datasets **(Supp. Fig. 8)**.

To further explore the SATURN-integrated dataset, we performed differential expression analysis on the macrogenes. In SATURN, macrogenes are groups of genes that share similar protein embeddings. The top five macrogenes are visualized in **Figure 8B.** Two selected macrogenes with high enrichment for tDC ("1556" and "1043") were visualized in feature plots, showing enrichment for the majority of cells in the tDC cluster and for some cells in the cDC2 cluster (**Fig. 8C**). Weights given in **Figure 8D** define the importance of genes within macrogenes, but, importantly, do not reflect mean expression of that gene. Among the top five weighted genes for these two selected tDC-specific macrogenes, was pig_PLXNA4. Notably, this gene was detected as tDC-enriched in our DEG analysis (**Supp. File 1**), but was found to be expressed across DC subsets of pig and horse, as shown in **Figure 8D**. Complete lists of differentially expressed macrogenes across DC subsets are given in **Supplementary File 2**, alongside gene content for the top five macrogenes of all five DC subsets. To conclude, further exploration is necessary to determine whether macrogene modules can provide deeper insights into DC subsets across species.

DISCUSSION

By performing scRNA-seq on enriched Flt3^{+} DC from blood of two horses, we identified five DC subsets, including tDC and putative DC3 alongside cDC1, cDC2, pDC, and DC progenitors.

Equine cDC1 and pDC were detected as distinct clusters expressing key genes conserved across species (7). Equine cDC2, identified primarily through expression of *FCER1A*, formed the biggest cluster (approx. 50 % within DC) and could be separated into at least two subsets.

418 In a previous study, equine pDC were gated as CD14^{-}Flt3^{+}CADM1^{-}MHC-IIll^{ow}, as confirmed by
419 qPCR for conserved key genes (*PLAC8*, *RUNX2*, *TCF4*) (6). Our scRNA-seq dataset confirms this
420 gating strategy and the described phenotype, with pDC expressing relatively low levels of MHC-II
421 and presumably lacking expression of *CD172a, CADM1,* and *CD14*.
422 Putative cDC1 were described as Flt3^{+}CD4^{-}CD13^{+}CD14^{low}CD172a^{-}CADM1^{+}MHC-II^{high}, which is in
423 line with the high expression of *CADM1* and genes encoding MHC-II molecules in the cDC1
424 cluster of our scRNA-seq dataset. In the same study (6), putative cDC2 were described as
425 Flt3^{+}CD4^{-}CD13^{-}CD14^{low}CD172a^{+}CADM1^{low}MHC-II^{high}. As discussed below, our data (scRNA-seq
426 and flow cytometry) argue for adjusting this phenotype to CD14^{+} (6).
427
428 Both tDC and putative DC3 were found to cluster closely with cDC2. Identification of tDC was
429 based on expression of pDC-associated genes (*TCF4*, *SPIB*, *BLNK*) alongside cDC2-associated
430 genes like *FCER1A*. This transcriptomic signature in-between pDC and cDC2 is described across
431 species investigated so far, including human (11), mouse (13), and more recently pig (14), and
432 highlights their proposed origin from pro-pDC (9) and differentiation towards cDC2-like cells.
433 In terms of function, tDC are described to produce pro-inflammatory IL1-β during viral infection
434 in the mouse model (13), and to be recruited to human skin upon experimental inoculation with
435 UV-inactivated E. coli (44). Our data indicates expression of *TLR2* and *NLRP3* in equine tDC,
436 suggesting their capability to recognize gram-positive bacteria, rather than gram-negative
437 bacteria, and to produce active IL-1β via inflammasome activation (45).
438 Expression of MHC-II was lower in tDC than in cDC1 and cDC2, but was considerably higher in
439 tDC than in pDC, aligning with transcriptomic studies performed in all three species investigated
440 so far (13,14,31,44), and supporting their involvement in CD4 T-cell activation. Judging from the
441 level of *WDFY4* detection, tDC may also be capable of cross-presentation (46). In our study, tDC
442 were found to comprise 4 % and 6 % within DC for "Horse 1" and "Horse 2", respectively, which
443 is comparable to frequencies found in blood of humans and spleen of mice (11,13,40,44,47,48),
444 and in contrast to the high frequency (approx. 30 % detected within DC) that we recently reported
445 for blood of pigs (14). Future studies should address differences in frequency and gene
446 expression of tDC across tissues and species and should discuss possible implications for tDC
447 biology. A more general enrichment strategy, encompassing the whole mononuclear phagocyte
448 system, will be necessary for these comparisons to reduce gating bias in FACS, as discussed
449 further below.
450

The two cDC2 clusters (termed here cDC2.1 and cDC2.2), identified in our dataset at first glance resembled CX3CR1-defined cDC2 subsets described in human and mouse. However, upon closer inspection, they did not appear to fit into the reported dichotomy of pro-inflammatory CX3CR1^{+} cDC2 and pro-resolving CX3CR1^{-} cDC2. Rather, equine cDC2.1 (lacking *CX3CR1* expression while expressing *CCR2*) transcriptomically resembled classical monocytes, and cDC2.2 (expressing *CX3CR1* while lacking *CCR2* expression) resembled non-classical monoctyes. Future scRNA-seq technologies with improved sensitivity, combined with protein-level and epigenomic analyses may provide more insights in this regard. However, cDC2 heterogeneity, presumably shaped by ontogeny (49) as well as environmental plasticity (50), will likely remain challenging to dissect – across species and tissues (9).

A subset of inflammatory cDC2 described in humans (51), was later discovered to comprise a separate DC lineage: dendritic cell type 3 (DC3), characterized by their co-expression of dendritic and monocytic markers (12,17).

In our dataset, putative DC3 (*FLT3*, *CD163*) were found to cluster with cDC2. Notably, in contrast to other investigated species (mouse, human, pig), *CD14* could not be used to differentiate DC3 from cDC2, as equine cDC2 were almost uniformly expressing *CD14*.

Across species, expression of CD14 is used to identify monocytes and to differentiate them from DC (52–54). With the discovery of *CD14*-expressing DC3 in human and mouse (10,12,14–16), this marker has been put into question. In pig, *CD14* expression appears to be limited to DC3 and monocytes (14), however, our scRNA-seq data and additional phenotyping data now clearly indicate that equine cDC2 express considerable levels of *CD14*. For the present study, we were able to exclude most monocytes by stringent gating based on CADM1 expression. While this enhanced enrichment of DC proofed beneficial to detect rare DC subsets like progenitors, tDC and DC3, this gating strategy may also have excluded some cDC2 and DC3 from our analysis, as continuous CADM1 expression on CD14^{+} cells did not allow for a clear separation of dendritic and monocytic cells.

This phenotypic overlap at the cDC2/DC3/monocyte interface is encountered across species and can likely not be resolved by additional phenotypic markers (51), especially under inflammatory conditions (9).

Sequencing of enriched DC alongside total mononuclear phagocytes should provide a more complete picture of DC heterogeneity, although a clear separation of DC and monocytes will likely remain challenging, as experienced with our recently published DC atlas for blood of pigs (14).

485 To assess conservation of DC signatures across species, we first applied GSEA. One limitation of
486 this analysis is that it relies on annotated orthologs in compared species. Moreover, signatures
487 for GSEA are derived from DEG analysis. In these DC subset comparisons, signatures of cDC2
488 are known to be composed of relatively few genes. Given that this signature is further restricted
489 by the lack of annotated orthologs, this likely explains our findings. Conversion of gene names
490 across compared species, as performed for GSEA on our porcine DC dataset (14), may improve
491 the outcome of these analyses.
492 Cross-species mapping of tDC signatures by GSEA confirmed our annotation, although tDC
493 signatures derived from a porcine (14) and a human dataset (11) primarily mapped onto equine
494 pDC. Given the expression of many pDC-associated genes in tDC, this finding could be expected.
495 In addition to GSEA, we explored SATURN (19), a deep learning tool based on protein-language
496 models (19) that is reported to outperform other tools for species comparisons (22).
497 For each species or dataset to be integrated, SATURN converts transcriptomic data to protein
498 sequences and then couples them to a protein language model to extract the protein
499 embeddings. Based on these embeddings, it learns the representation (integrated object across
500 species) of the transcriptomic datasets. As this tool is based on protein sequences in a joint
501 embedding, it may overcome limitations from poorly annotated reference genomes. We used
502 SATURN to integrate the scRNA-seq data of equine DC with our recently published scRNA-seq
503 data of porcine DC (14). DC subsets co-clustered as expected, with equine tDC clearly clustering
504 together with porcine tDC.
505 Here, SATURN was tried for the first time on closely related cell types (DC subtypes). Moreover,
506 our dataset was considerably smaller than the datasets used in the original SATURN paper.
507 Expected results (co-clustering of DC subsets across species) were obtained, when parameters
508 (like the number of used PCs) were adjusted. Rather unexpected was the clustering observed
509 upon inclusion of a higher number of PCs (explaining close to 100 % of the variation), with cDC2
510 of horse and pig remaining co-clustered and cDC1 and pDC increasingly separating from their
511 counterparts. In fact, cDC2 are known to be very heterogeneous, and cross-species analyses
512 based on gene signatures usually give conflicting results with cDC2. On the other hand, cDC1
513 and pDC are relatively well-defined and homogeneous cell types with specific core signatures
514 conserved across species (52,53) and we would have expected them to cluster together
515 independently of included PCs.
516 As the top 5 selected PCs explained at least 90 % of variation in all three datasets analyzed, it
517 appears justified to restrict the number of PCs and one could argue that cell-type information is
518 contained in these first few PCs, whereas species-specific information and noise is contained in

later PCs, explaining much less of the observed variance. Additional analyses on different datasets (e.g. small vs. large, expected low vs. high technical batch effects) are necessary to better understand the driving factors behind this deep-learning tool. Furthermore, it remains to be determined if SATURN is useful to extract additional information from macrogenes – for example novel conserved DC markers, with potential insights into conserved functions of DC subsets.

Others have characterized equine immune cells by scRNA-seq of PBMC (18), BALF (bronchoalveolar lavage fluid) (55), synovial fluid (56), and endometrium (57), but detected only low frequencies of DC. Through DC enrichment prior to scRNA-seq, our study provides the first in-depth analysis of equine DC heterogeneity.

With the present characterization of transitional DC in equine blood, we pave the way for identifying and investigating this DC subset in tissues like the skin, as their recruitment to skin and their adoption of a pro-inflammatory profile has been reported for humans (44).

Moreover, it remains to be determined if equine tDC have a role in driving inflammation during viral infections, as suggested based on murine studies (13). With more and more species contributing to a universal DC atlas, comparative analyses promise to considerably improve our understanding of this complex DC network.

AUTHOR CONTRIBUTIONS

MB: bioinformatic data analysis (scVI, SATURN), preparation of figures, manuscript editing

AB: manuscript writing and editing

SO: bioinformatic data analysis

IC: laboratory work

SJ: laboratory work

AZ: data analysis

FB: bioinformatic data analysis

AS: manuscript editing

EM: study design, manuscript editing

SCT: study design, supervision of laboratory work, data analysis, figure design and preparation, manuscript writing

ACKNOWLEDGEMENTS

We thank Roosheel Patel and Brad Rosenberg (Department of Microbiology, Icahn School of Medicine at Mount Sinai, USA) for ESAT processing of our dataset. Moreover, we thank Stefan

Müller (FCCS, University of Bern, Switzerland) for cell sorting, the team of the NGS platform (NGSP, University of Bern, Switzerland) for single-cell RNA sequencing, and the Interfaculty Bioinformatics Unit (IBU, University of Bern, Switzerland) for access to their compute cluster.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

557 **REFERENCES**

558 1. Eisenbarth SC. Dendritic cell subsets in T cell programming: location dictates function. Nat
559 Rev Immunol. 2019 Feb;19(2):89–103.

560 2. Collin M, Bigley V. Human dendritic cell subsets: an update. Immunology. 2018;154(1):3–20.

561 3. Reynolds G, Tirard A, Villani AC. Plasmacytoid dendritic cells: Welcome back to the DC fold.
562 Immunity. 2022 Mar 8;55(3):380–2.

563 4. Ziegler-Heitbrock L, Ohteki T, Ginhoux F, Shortman K, Spits H. Reply to 'Reclassification of
564 plasmacytoid dendritic cells as innate lymphocytes is premature.' Nat Rev Immunol. 2023
565 May;23(5):338–9.

566 5. Reizis B. Plasmacytoid Dendritic Cells: Development, Regulation, and Function. Immunity.
567 2019 Jan 15;50(1):37–50.

568 6. Ziegler A, Marti E, Summerfield A, Baumann A. Identification and characterization of equine
569 blood plasmacytoid dendritic cells. Dev Comp Immunol. 2016 Dec;65:352–7.

570 7. Summerfield A, Auray G, Ricklin M. Comparative dendritic cell biology of veterinary mammals.
571 Annual review of animal biosciences. 2015;3:533–57.

572 8. Entrican G, Lunney JK, Wattegedera SR, Mwangi W, Hope JC, Hammond JA. The Veterinary
573 Immunological Toolbox: Past, Present, and Future. Front Immunol. 2020 Jul 28;11:1651

574 9. Tussiwand R, Rodrigues PF. Where's Waldo: Identifying DCs within Mononuclear Phagocytes
575 during Inflammation. Immunity. 2020 Jun 16;52(6):892–4.

576 10. Bourdely P, Anselmi G, Vaivode K, Ramos RN, Missolo-Koussou Y, Hidalgo S, et al.
577 Transcriptional and Functional Analysis of CD1c+ Human Dendritic Cells Identifies a CD163+
578 Subset Priming CD8+CD103+ T Cells. Immunity. 2020 Aug;53(2):335-352.e8.

579 11. Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. Single-cell RNA-seq
580 reveals new types of human blood dendritic cells, monocytes, and progenitors. Science. 2017
581 Apr 21;356(6335):eaah4573.

582 12. Liu Z, Wang H, Li Z, Dress RJ, Zhu Y, Zhang S, et al. Dendritic cell type 3 arises from Ly6C+
583 monocyte-dendritic cell progenitors. Immunity. 2023 Aug 8;56(8):1761-1777.e6.

584 13. Sulczewski FB, Maqueda-Alfaro RA, Alcántara-Hernández M, Perez OA, Saravanan S, Yun TJ,
585 et al. Transitional dendritic cells are distinct from conventional DC2 precursors and mediate
586 proinflammatory antiviral responses. Nat Immunol. 2023 Aug;24(8):1265–80.

587 14. Baillou A, Auray G, Huber A, Botos M, Brito F, Summerfield A, et al. Unraveling dendritic-cell
588 diversity in blood of pigs: welcome tDC and DC3. bioRxiv; 2025.

589 15. Amon L, Seichter A, Vurnek D, Heger L, Lächele L, Tochoedo NR, et al. Clec12A, CD301b,
590 and FcγRIIB/III define the heterogeneity of murine DC2s and DC3s. Cell Reports. 2024
591 Mar;43(3):113949.

592 16. Villar J, Segura E. The More, the Merrier: DC3s Join the Human Dendritic Cell Family.
593 Immunity. 2020 Aug;53(2):233–5.

594 17. Cytlak U, Resteu A, Pagan S, Green K, Milne P, Maisuria S, et al. Differential IRF8 Transcription
595 Factor Requirement Defines Two Pathways of Dendritic Cell Development in Humans.
596 Immunity. 2020 Aug 18;53(2):353-370.e8.

597 18. Patel RS, Tomlinson JE, Divers TJ, Van De Walle GR, Rosenberg BR. Single-cell resolution
598 landscape of equine peripheral blood mononuclear cells reveals diverse cell types including
599 T-bet+ B cells. BMC Biol. 2021 Dec;19(1):13.

600 19. Rosen Y, Brbić M, Roohani Y, Swanson K, Li Z, Leskovec J. Toward universal cell embeddings:
601 integrating single-cell RNA-seq datasets across species with SATURN. Nat Methods. 2024 Feb
602 16;1–9.

603 20. Song Y, Miao Z, Brazma A, Papatheodorou I. Benchmarking strategies for cross-species
604 integration of single-cell RNA sequencing data. Nat Commun. 2023 Oct 14;14(1):6495.

605 21. Luecken MD, Büttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al.
606 Benchmarking atlas-level data integration in single-cell genomics. Nat Methods. 2022
607 Jan;19(1):41–50.

608 22. Zhong H, Han W, Gomez-Cabrero D, Tegner J, Gao X, Cui G, et al. Benchmarking cross-
609 species single-cell RNA-seq data integration methods: towards a cell type tree of life. Nucleic
610 Acids Research. 2025 Jan 13;53(1):gkae1316.

611 23. Guzylack-Piriou L, Alves MP, McCullough KC, Summerfield A. Porcine Flt3 ligand and its
612 receptor: generation of dendritic cells and identification of a new marker for porcine dendritic
613 cells. Developmental and comparative immunology. 2010 Apr;34(4):455–64.

614 24. Derr A, Yang C, Zilionis R, Sergushichev A, Blodgett DM, Redick S, et al. End Sequence
615 Analysis Toolkit (ESAT) expands the extractable information from single-cell RNA-seq data.
616 Genome Res. 2016 Oct 1;26(10):1397–410.

617 25. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis.
618 Genome Biology. 2018 Feb 6;19(1):15.

619 26. Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, et al. Best practices for
620 single-cell analysis across modalities. Nat Rev Genet. 2023 Aug;24(8):550–72.

621 27. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in
622 Single-Cell Transcriptomic Data. Cell Systems. 2019 Apr 24;8(4):281-291.e9.

623 28. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell
624 transcriptomics. Nat Methods. 2018 Dec;15(12):1053–8.

625 29. Badia-i-Mompel P, Vélez Santiago J, Braunger J, Geiss C, Dimitrov D, Müller-Dott S, et al.
626 decoupleR: ensemble of computational methods to infer biological activities from omics data.
627 Bioinformatics Advances. 2022 Jan 1;2(1):vbac016.

628 30. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular
629 Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015 Dec 23;1(6):417–
630 25.

631 31. Valente M, Collinet N, Vu Manh TP, Popoff D, Rahmani K, Naciri K, et al. Novel mouse models
632 based on intersectional genetics to identify and characterize plasmacytoid dendritic cells. Nat
633 Immunol. 2023 Apr;24(4):714–28.

634 32. Rives A, Meier J, Sercu T, Goyal S, Lin Z, Liu J, et al. Biological structure and function emerge
635 from scaling unsupervised learning to 250 million protein sequences. Proceedings of the
636 National Academy of Sciences. 2021 Apr 13;118(15):e2016239118.

637 33. Talker SC, Barut GT, Lischer HEL, Rufener R, Bruggmann R, Summerfield A. Monocyte biology
638 conserved across species: Functional insights from cattle. Front Immunol. 2022 Jul 29;13.

639 34. Barut GT, Lischer HEL, Bruggmann R, Summerfield A, Talker SC. Transcriptomic profiling of
640 bovine blood dendritic cells and monocytes following TLR stimulation. European Journal of
641 Immunology. 2020;50(11):1691–711.

642 35. Rodrigues PF, Trsan T, Cvijetic G, Khantakova D, Panda SK, Liu Z, et al. Progenitors of distinct
643 lineages shape the diversity of mature type 2 conventional dendritic cells. Immunity. 2024 Jul
644 9;57(7):1567-1585.e5.

645 36. Kharas MG, Lengner CJ, Al-Shahrour F, Bullinger L, Ball B, Zaidi S, et al. Musashi-2 regulates
646 normal hematopoiesis and promotes aggressive myeloid leukemia. Nat Med. 2010
647 Aug;16(8):903–8.

648 37. Metzemaekers M, Van Damme J, Mortier A, Proost P. Regulation of Chemokine Activity - A
649 Focus on the Role of Dipeptidyl Peptidase IV/CD26. Front Immunol. 2016;7:483.

650 38. Silva-Gomes R, Mapelli SN, Boutet MA, Mattiola I, Sironi M, Grizzi F, et al. Differential
651 expression and regulation of MS4A family members in myeloid cells in physiological and
652 pathological conditions. J Leukoc Biol. 2022 Apr;111(4):817–36.

653 39. Takamatsu H, Takegahara N, Nakagawa Y, Tomura M, Taniguchi M, Friedel RH, et al.
654 Semaphorins guide the entry of dendritic cells into the lymphatics by activating myosin II. Nat
655 Immunol. 2010 Jul;11(7):594–600.

656 40. Rodrigues PF, Trsan T, Cvijetic G, Khantakova D, Panda SK, Liu Z, et al. Progenitors of distinct
657 lineages shape the diversity of mature type 2 conventional dendritic cells. Immunity. 2024 Jul
658 9;57(7):1567-1585.e5.

659 41. Brown CC, Gudjonson H, Pritykin Y, Deep D, Lavallée VP, Mendoza A, et al. Transcriptional
660 Basis of Mouse and Human Dendritic Cell Heterogeneity. Cell. 2019 Oct 31;179(4):846-
661 863.e24.

662 42. Fujita M, Takada YK, Takada Y. Integrins αvβ3 and α4β1 Act as Coreceptors for Fractalkine,
663 and the Integrin-Binding Defective Mutant of Fractalkine Is an Antagonist of CX3CR1. The
664 Journal of Immunology. 2012 Dec 15;189(12):5809–19.

665 43. Clement M, Fornasa G, Guedj K, Ben Mkaddem S, Gaston AT, Khallou-Laschet J, et al. CD31
666 is a key coinhibitory receptor in the development of immunogenic dendritic cells. Proceedings
667 of the National Academy of Sciences. 2014 Mar 25;111(12):E1101–10.

668 44. Lubin R, Patel AA, Mackerodt J, Zhang Y, Gvili R, Mulder K, et al. The lifespan and kinetics of
669 human dendritic cell subsets and their precursors in health and inflammation. Journal of
670 Experimental Medicine. 2024 Nov 4;221(11):e20220867.

671 45. Swanson KV, Deng M, Ting JPY. The NLRP3 inflammasome: molecular activation and
672 regulation to therapeutics. Nat Rev Immunol. 2019 Aug;19(8):477–89.

673 46. Theisen DJ, Davidson JT, Briseño CG, Gargaro M, Lauron EJ, Wang Q, et al. WDFY4 is required
674 for cross-presentation in response to viral and tumor antigens. Science. 2018 Nov
675 9;362(6415):694–9.

676 47. Alcántara-Hernández M, Leylek R, Wagar LE, Engleman EG, Keler T, Marinkovich MP, et al.
677 High-Dimensional Phenotypic Mapping of Human Dendritic Cells Reveals Interindividual
678 Variation and Tissue Specialization. Immunity. 2017 Dec;47(6):1037-1050.e6.

679 48. Leylek R, Alcántara-Hernández M, Lanzar Z, Lüdtke A, Perez OA, Reizis B, et al. Integrated
680 Cross-Species Analysis Identifies a Conserved Transitional Dendritic Cell Population. Cell
681 Reports. 2019 Dec;29(11):3736-3750.e8.

682 49. Minutti CM, Piot C, Pereira da Costa M, Chakravarty P, Rogers N, Huerga Encabo H, et al.
683 Distinct ontogenetic lineages dictate cDC2 heterogeneity. Nat Immunol. 2024 Mar;25(3):448–
684 61.

685 50. León B. Type 2 conventional dendritic cell functional heterogeneity: ontogenically
686 committed or environmentally plastic? Trends in Immunology. 2025 Feb;46(2):104–20.

687 51. Dutertre CA, Becht E, Irac SE, Khalilnezhad A, Narang V, Khalilnezhad S, et al. Single-Cell
688 Analysis of Human Mononuclear Phagocytes Reveals Subset-Defining Markers and Identifies
689 Circulating Inflammatory Dendritic Cells. Immunity. 2019 Sep 17;51(3):573-589.e8.

690 52. Auray G, Keller I, Python S, Gerber M, Bruggmann R, Ruggli N, et al. Characterization and
691 Transcriptomic Analysis of Porcine Blood Conventional and Plasmacytoid Dendritic Cells
692 Reveals Striking Species-Specific Differences. The Journal of Immunology. 2016 Dec
693 15;197(12):4791–806.

694 53. Talker SC, Baumann A, Bruggmann R. Precise Delineation and Transcriptional
695 Characterization of Bovine Blood Dendritic-Cell and Monocyte Subsets. Front Immunol. 2018
696 Oct 30;9.

54. Collin M, Bigley V. Human dendritic cell subsets: an update. Immunology. 2018;154(1):3–20.

55. Sage SE, Leeb T, Jagannathan V, Gerber V. Single-cell profiling of bronchoalveolar cells reveals a Th17 signature in neutrophilic severe equine asthma. Immunology. n/a(n/a).

56. Ammons DT, Chow L, Goodrich L, Bass L, Larson B, Williams ZJ, et al. Characterization of the single cell landscape in normal and osteoarthritic equine joints. Ann Transl Med. 2024 Oct;12(5):88–88.

57. Jaworska J, Tobolski D, Salem SE, Kahler A, Wocławek-Potocka I, De Mestre AM. Single-cell atlas of the pregnant equine endometrium before and after implantation. Biology of Reproduction. 2025 Jan 5;ioaf004.

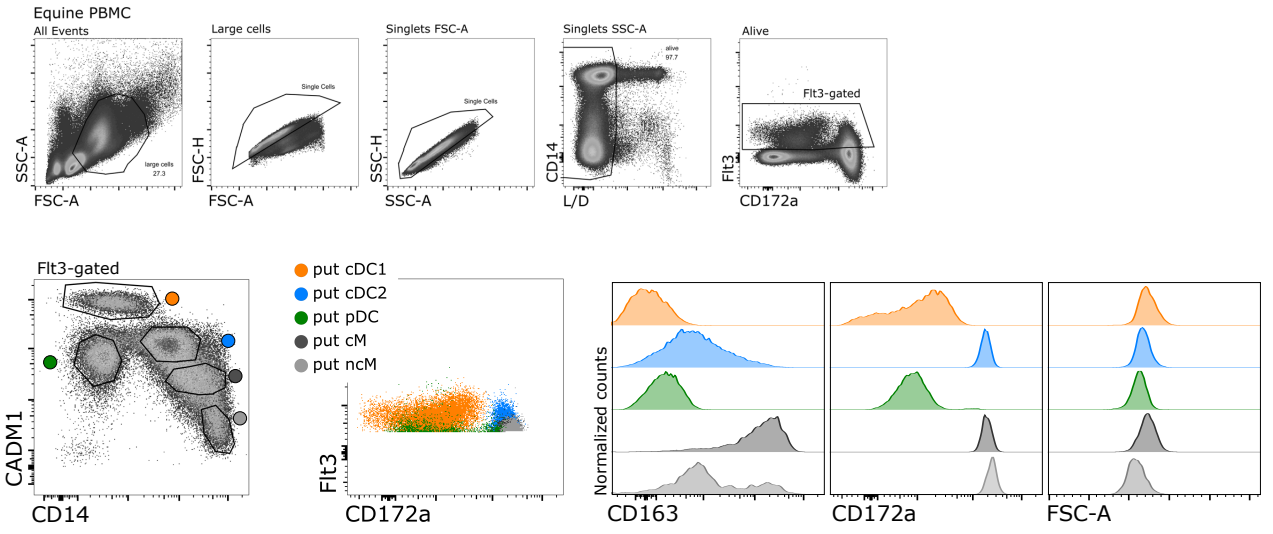
58. Sage SE, Leeb T, Jagannathan V, Gerber V. Single-cell profiling of bronchoalveolar cells reveals a TH17 signature in neutrophilic severe equine asthma. Immunology. 2024 Apr;171(4):549–65.

59. Sage SE, Nicholson P, Peters LM, Leeb T, Jagannathan V, Gerber V. Single-cell gene expression analysis of cryopreserved equine bronchoalveolar cells. Front Immunol. 2022 Aug 23;13:929922.

60. Riihimäki M, Fegraeus K, Nordlund J, Waern I, Wernersson S, Akula S, et al. Single-cell transcriptomics delineates the immune cell landscape in equine lower airways and reveals upregulation of FKBP5 in horses with asthma. Sci Rep. 2023 Sep 27;13(1):16261.

Figure 1

A Phenotyping to establish enrichment strategy for equine DC



B Enrichment of DC by FACS for scRNA-seq

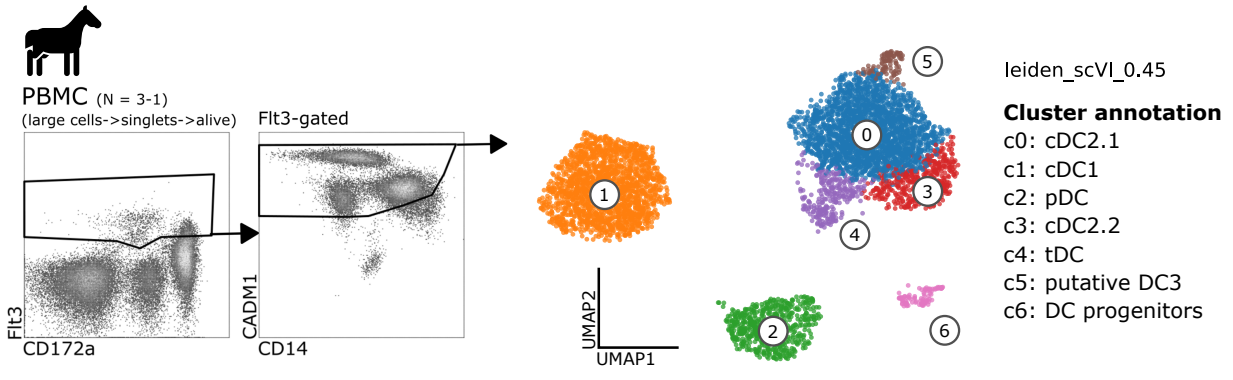


Figure 1: Phenotyping of equine DC and monocytes. **(A)** Gating strategy for phenotyping of putative DC and monocyte subsets. Large cells within PBMC were gated, doublets and dead cells excluded, and a first gating was set to include most DC as well as monocytes falling in the Flt3 gate. Putative DC and monocyte subsets were defined based on expression of CD14 and CADM1. Overlayed dotplot shows localization of subsets in CD172a vs. Flt3 plot. For each subset, expression levels of CD163 and CD172a were visualized in histograms, alongside forward scatter intensity (FSC-A). **(B)** Enrichment strategy for scRNA-seq of DC, and annotation of scRNA-seq clusters.

Figure 2

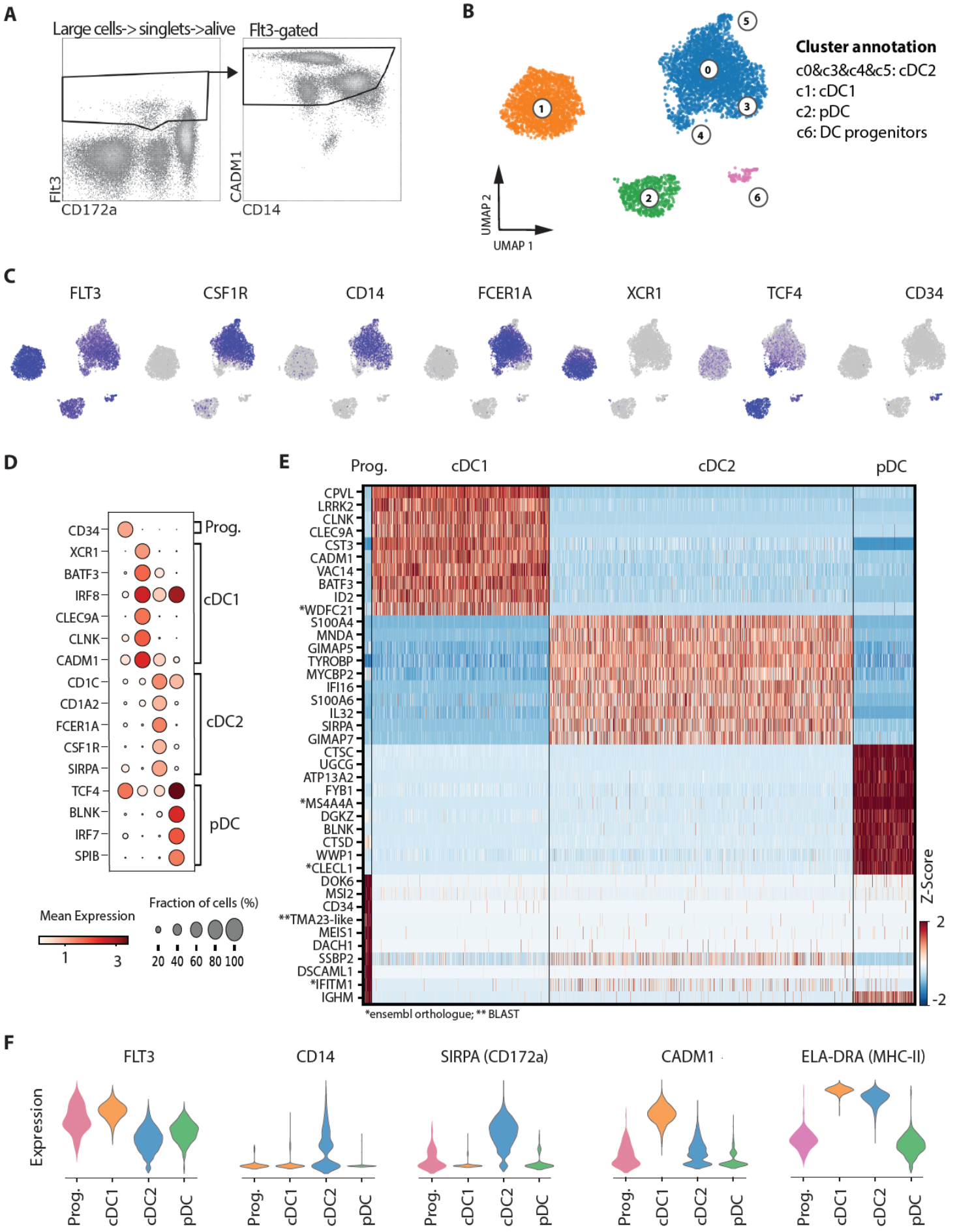


Figure 2. Identification of main DC subsets in blood of horse. (A) Enrichment of DC for scRNA-seq. **(B)** Annotation of traditional DC subsets (cDC1, CDC2, pDC) and DC progenitors. **(C)** Feature plots of key markers used for subset identification. **(D)** Dotplot with additional DC subset-specific genes. **(E)** Top 20 genes obtained from differential expression analysis. Complete gene lists are given in **Supplementary File 1**. **(F)** Violin plots show gene expression for phenotypic markers previously analyzed by flow cytometry.

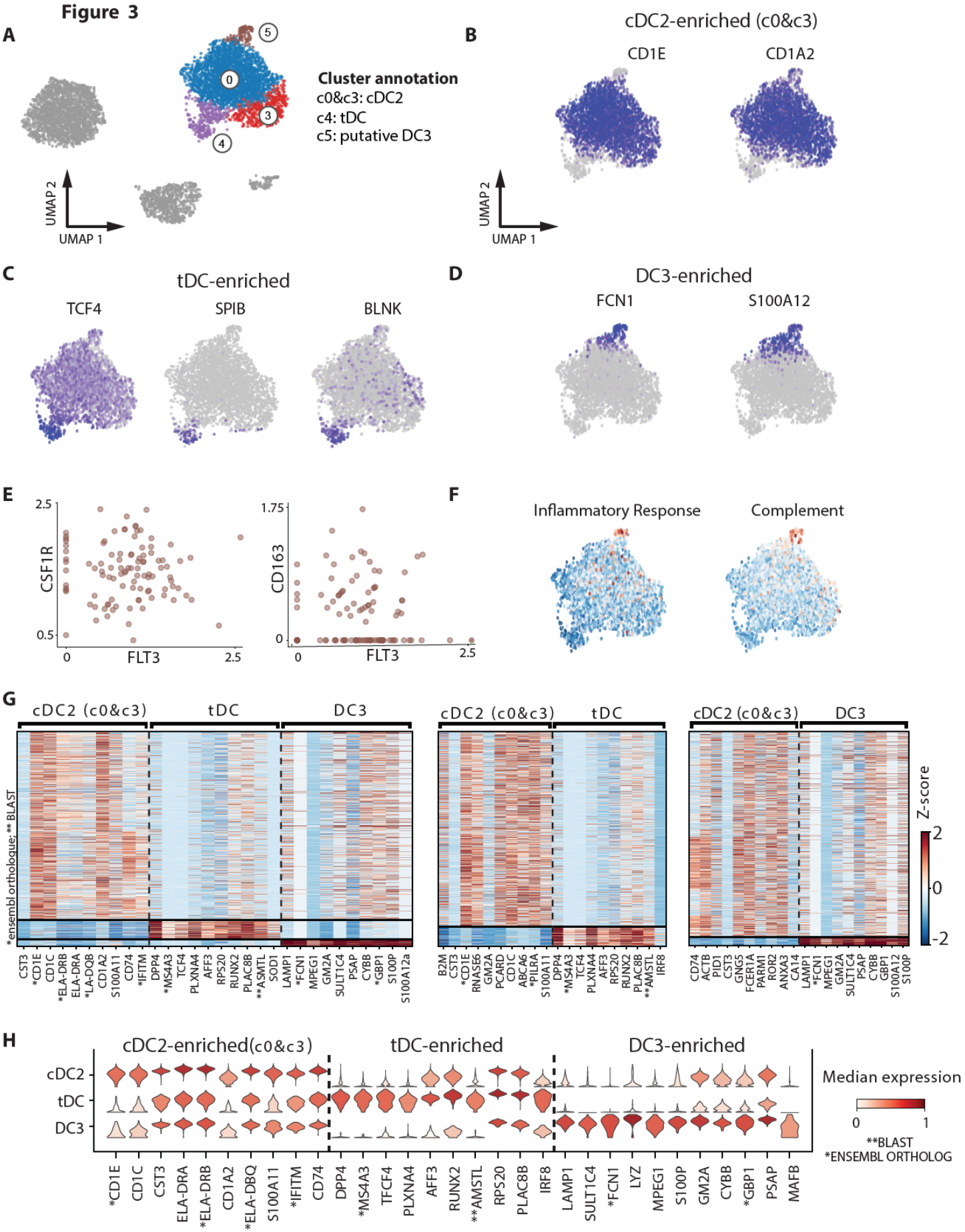


Figure 3: Delineation of tDC and putative DC3 from cDC2. (A) Cluster annotation. **(B)** Annotation of c0 and c3 as CDC2 based on expression of CD1 genes. **(C)** Annotation of c4 as tDC based on expression of pDC-related genes. **(D-F)** Annotation of c5 as DC3, based on pro-inflammatory profile and co-expression of FLT3 and monocyte markers on the single-cell level (scatter plot). **(G)** Heatmaps showing top 10 differentially expressed genes (DEG) between cDC2, DC3 and tDC (left), cDC2 and tDC (middle), and cDC2 and DC3 (right). **(H)** Violin plots show selected genes from DEG analysis enriched in either cDC2, tDC, or DC3. Complete gene lists are given in **Supplementary File 1.**

Figure 4

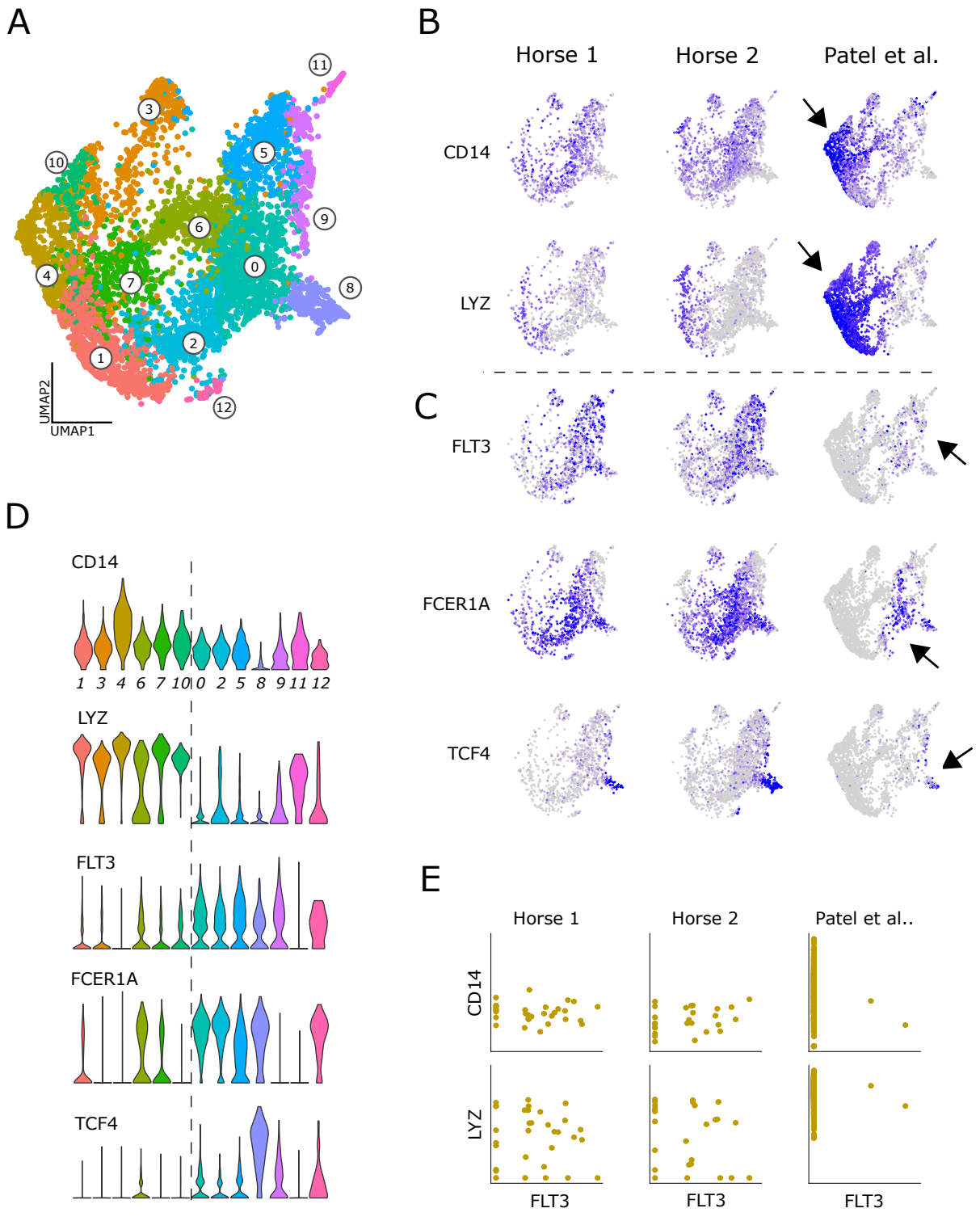


Figure 4: Delineation of cDC2 from monocytes. DC and monocytes from a previously published scRNA-seq dataset (equine PBMC, Patel et al.) were re-processed, subsetted and integrated with our re-processed DC-enriched dataset **(Suppl. Fig. 5)** Our dataset was re-processed with Ensembl v95 and the ESAT tool to match the dataset available from Patel et al. **(A)** Clustering of integrated dataset. (B-C) Feature plots show expression of monocyte-associated genes **(B)** and DC-associated genes **(C)** for "Horse 1" and "Horse 2" from our dataset and for the seven horses analyzed by Patel et al. **(D)** Violin plots visualize gene expression for key monocyte/ DC markers on the cluster level. **(E)** Scatter plots show cells in cluster 4 co-expressing FLT3 and monocytic markers (CD14, LYZ).

Figure 5

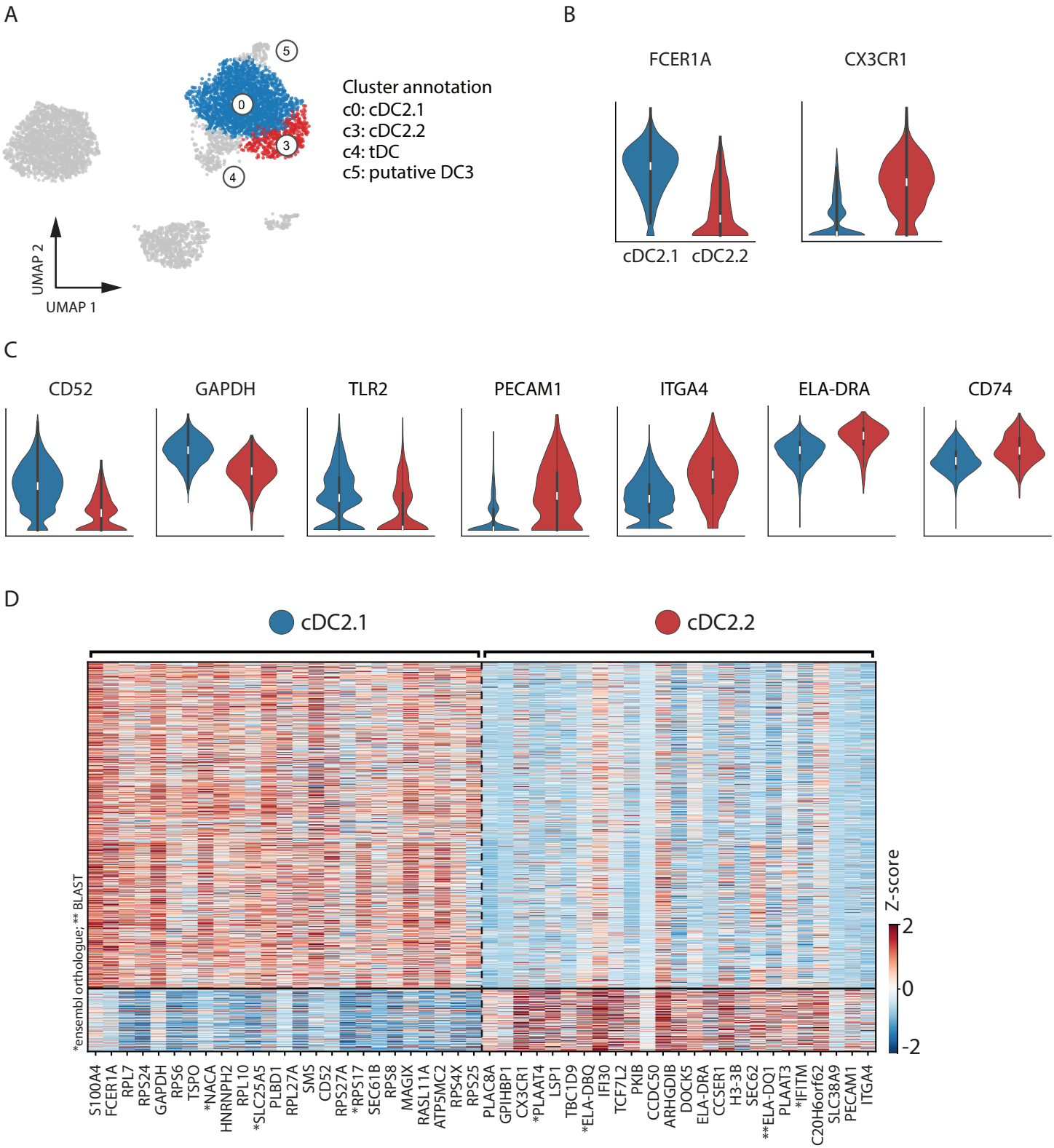
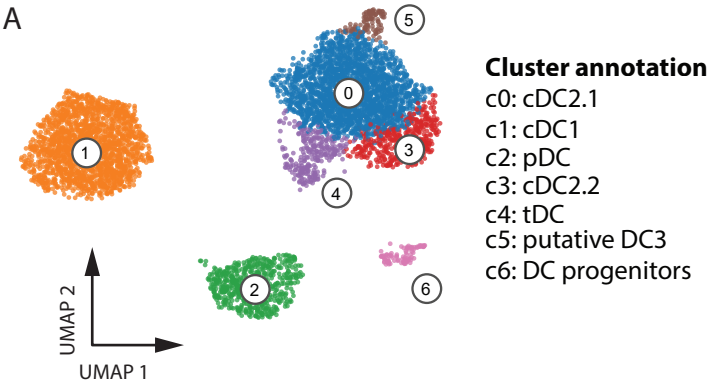


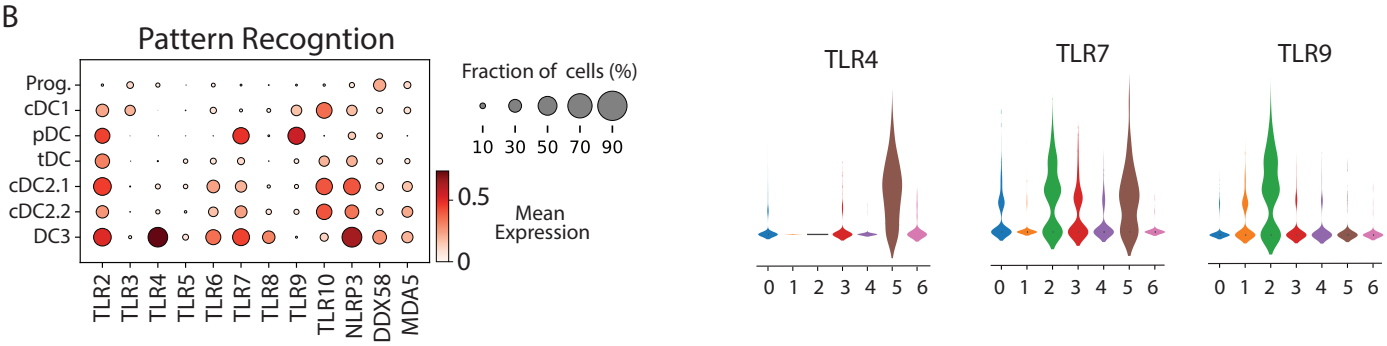
Figure 5. Heterogeneity of cDC2. (A) Cluster annotation. **(B)** Violin plots show gene expression of FCER1A and CX3CR1 in cDC2 subsets. **(D)** Selected violin plots based on DEG analysis. **(C)** Heatmap shows top 25 differentially expressed genes (DEG) between cDC2.1 and cDC2.2. Complete gene lists are given in **Supplementary File 1**.

Figure 6

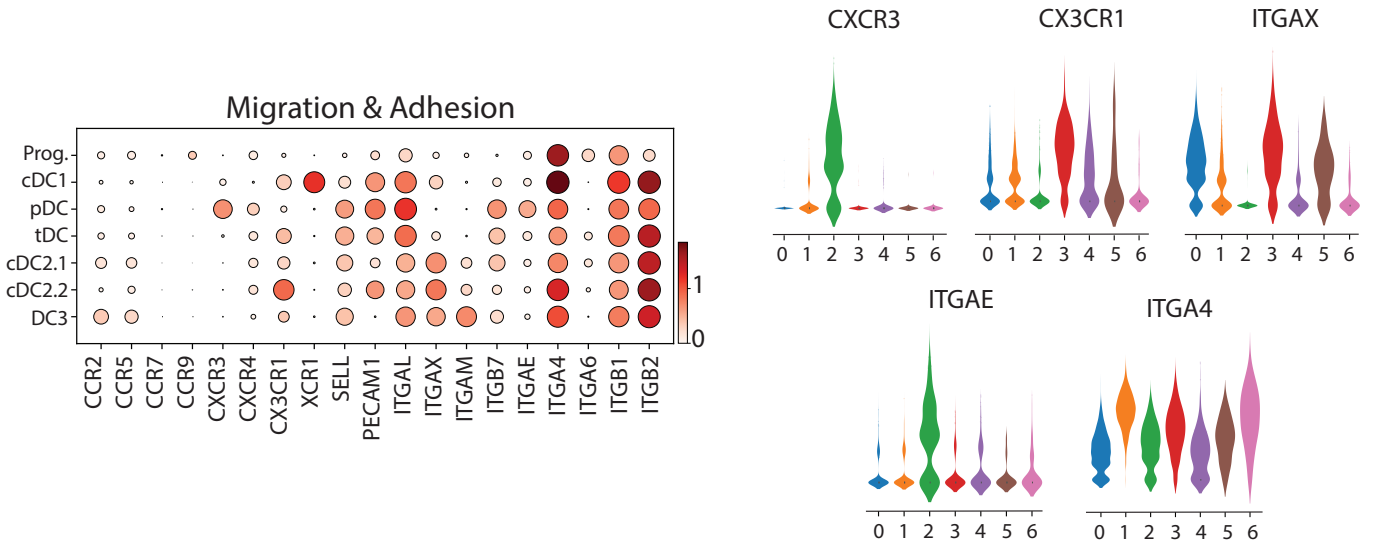
A



B



C



D

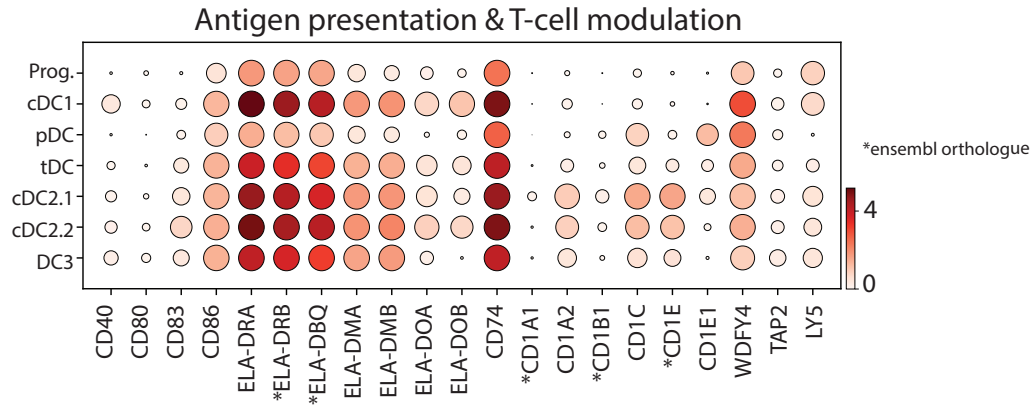


Figure 6. Gene expression related to main DC functions. (A) Cluster annotation. **(B-D)** Dotplots and selected violin plots visualize gene expression across all annotated clusters related to pattern recognition **(B)**, migration & adhesion **(C)**, and antigen presentation & T-cell modulation **(D)**. Gene expression related to inflammation and metabolism is shown in **Supplementary Figure 5**.

Figure 7

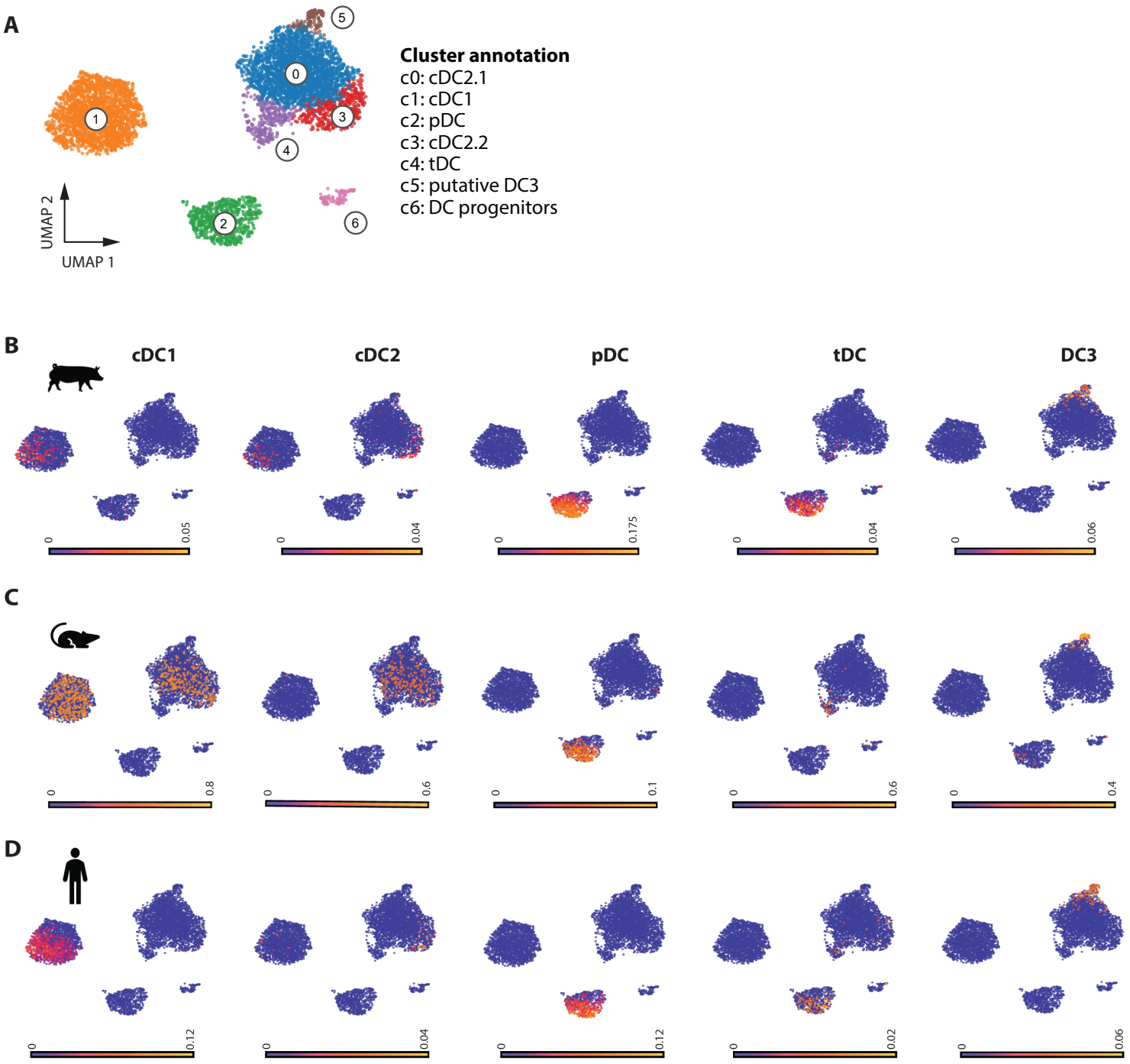
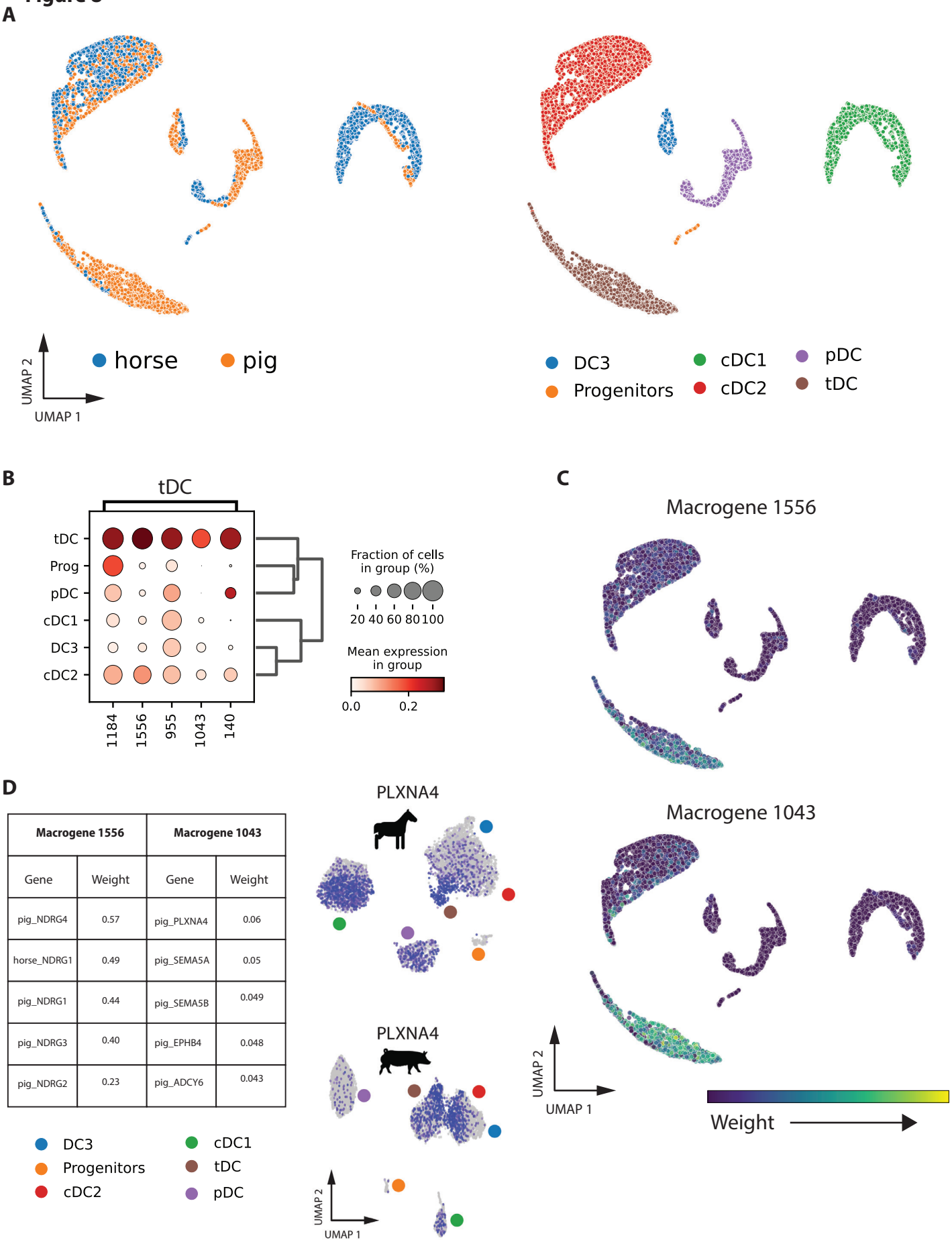


Figure 7: Cross-species comparison of DC-subset specific gene signatures by gene set enrichment analysis. (A) Cluster annotation of equine dataset. Enrichment of DC subset transcriptomic signatures from porcine blood (bulk; from Baillou et al.) **(B)**, from murine spleen (FB5P-seq from Valente et al.) **(C)**, and from human blood (sc SMART-seq; from Villani et al.) **(D)** in the equine single-cell RNA-seq clusters, represented by UMAP plots. Relative enrichment scores range from low (blue) to high (yellow). Porcine "DC3" signature is derived from sorted CD14^{+} monocytes.

Figure 8



Macrogene 1556		Macrogene 1043	
Gene	Weight	Gene	Weight
pig_NDRG4	0.57	pig_PLXNA4	0.06
horse_NDRG1	0.49	pig_SEMA5A	0.05
pig_NDRG1	0.44	pig_SEMA5B	0.049
pig_NDRG3	0.40	pig_EPHB4	0.048
pig_NDRG2	0.23	pig_ADCY6	0.043

Figure 8. Integration of equine and porcine scRNA-seq datasets with SATURN. (A) UMAP plots colored by species and cell types. **(B)** Dot plot of the most significant differentially expressed macrogenes for tDC. **(C)** UMAP plot representing the cells with the highest expression of selected macrogenes. **(D)** Genes of porcine and equine origin contained in macrogenes and visualization of PLXNA4 as top-weighted gene within tDC-enriched macrogene. **Supplementary Figure 9** shows key genes visualized in feature plots for the equine and porcine datasets.