Comparative single cell transcriptomic analysis of the hematopoietic system between human and mouse

Shouguo Gao, PhD1, Zhijie Wu, MD, PhD1, Xingmin Feng, PhD1, Sachiko Kajigaya, PhD1, Neal S. Young, MD1

*Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD,*

Single cell RNA-seq has been extensively applied to the hematopoiesis of human and mouse, but the cross-species comparison of hematopoietic hierarchy is still not well studied. Here, through the 10X single cell platform and a canonical correlation analysis (CCA) computational strategy, we conducted comparative transcriptomic analysis of hematopoietic hierarchy in human and mouse. We found that the hematopoietic stem and progenitor cell (HSPC) compartment in the two species is composed of subpopulations characterized the same set of homologous genes, and the hematopoietic lineages and transcriptional profiling in hematopoiesis are well conserved between human and mouse, indicating an evolutionary similarity in their hematopoietic systems.

We constructed a single-cell resolution transcriptomic atlas of HSPCs in human and mouse, having a total number of 32,805 single cells. We only kept the orthologous genes of human and mouse collected in InParanoid. With known marker genes, we grouped human cells as hematopoietic stem cell (HSC), multilymphoid progenitor (MLP), granulocyte-monocyte progenitor (GMP), Pro-B cell (ProB), earliest thymic progenitor (ETP), and megakaryocytic-erythroid progenitor (MEP); and mouse cells as long-term HSCs (LTHSC), lymphoid multipotent progenitors (LMPP), and multipotent progenitor (MPP), GMP, MEP and common myeloid progenitors (CMP).

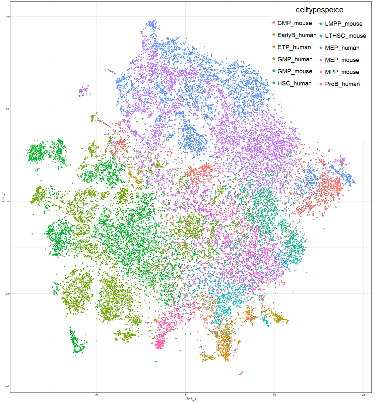
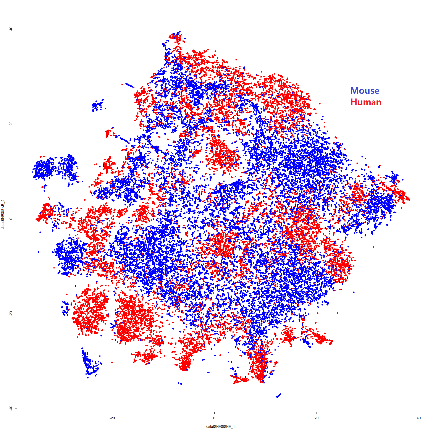
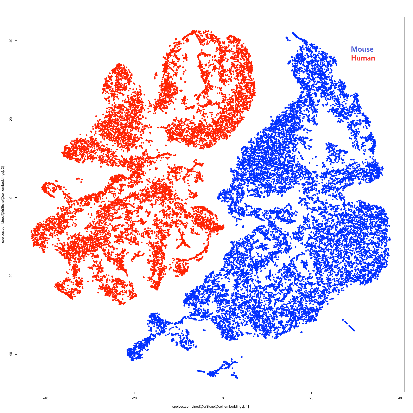
tSNE plots showed that cells were grouped by species, instead of cell types, due to species specificity and batch effects (Fig A, mouse and human cells were profiled at different time). The CCA algorithm is a multivariate statistical technique for detecting the statistical common factors among two digital gene expression (DGE) matrices, which vary from each other due to batch effects. After alignment with CCA, the cells of mouse and human were well mixed and separated by same cell type categories (Fig B-C). The cells were grouped into 17 subpopulations (15245 and 17560 cells, and 17 and 16 subpopulations from human and mouse, respectively) by computational analysis (Fig C). We checked the cluster specific genes and found that they were conserved between mouse and human and share same functional themes.

Furthermore, we explored the similarities and differences of HSPCs between two species. To obtain a detailed view on the cellular evolution from mouse to human in the HSPC system, we used mouse and human orthologous genes, and calculated an average of expression of cells in each population of human and mouse. After hierarchical clustering, a cluster dendrogram indicated that cell types were highly conserved between human and mouse (Fig D). For example, MEP and GMP of mouse and human shared a very similar transcriptome pattern. Human HSC was firstly clustered with mouse LTHSC and then with mouse MPP. Further, MPP and LMPP in mouse had similar transcriptomes, which was observed in human already.

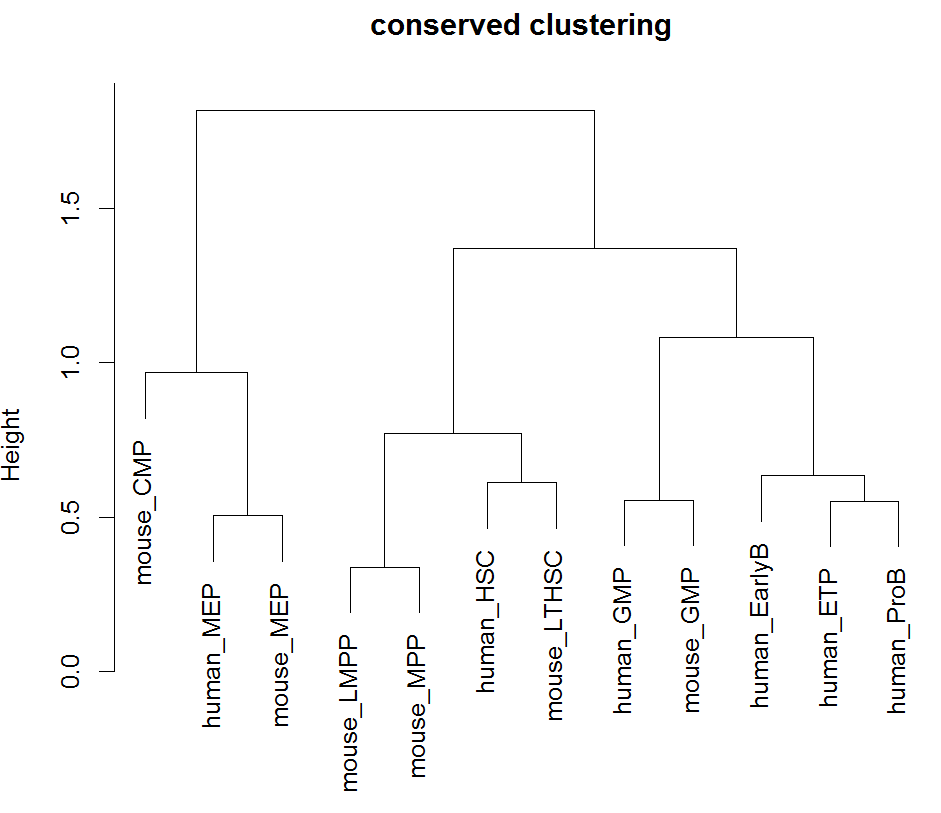
We used monocle to examine the differentiation trajectory of hematopoiesis in human and mouse, and we defined the HSC and LTHSC as roots so that they located at starting points of the differentiation hierarchy. When applied monocle to the human and mouse cells profiled, with a graph, an intuitive representation of HSPC differentiation emerged. In both mouse and human, three branches came out from HSC and LTHSC. We checked the cell types and found that three branches were Erythroid/megakaryocytic, Myeloid, and Lymphoid (Fig E-F). We also examined expression level changes of individual genes during trajectory. Such as, in both mouse and human, Gata1 and CD79A expression levels increased along Erythroid branch and lymphoid branch, respectively, and Procr expression decreased with differentiation.

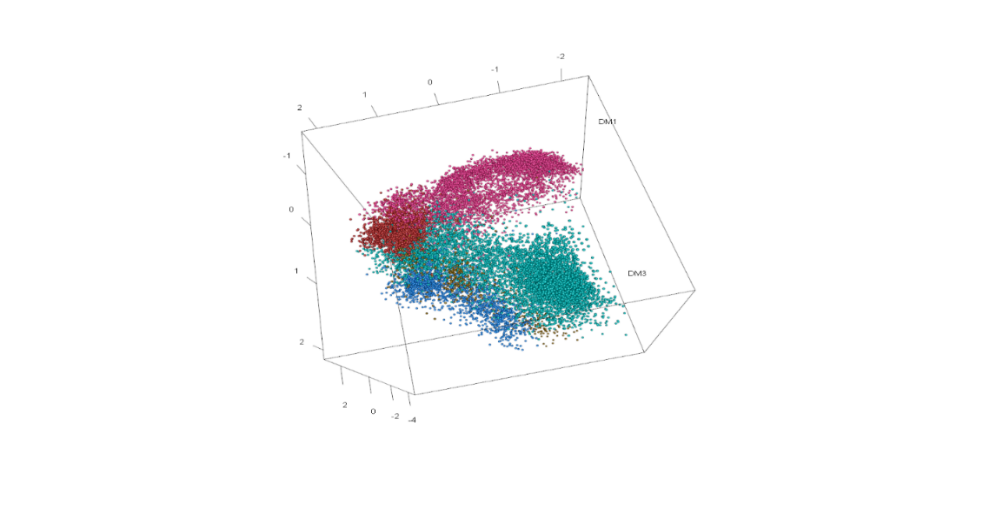
To understand the species conservation of cell populations of hematopoiesis in human and mouse, single-cell transcriptomes of human cells were compared with those of mouse cells using scmap. This method allowed comparison of multiple datasets by projecting a cell onto a reference dataset and inferring cellular identity with the cell types defined in the reference. First, most of human MEP cells (85%) were mapped to mouse MEP cell types based on transcriptional similarity, suggesting their functional similarity and species conservation. Further, 45% and 24% human HSC cells were mapped to mouse LTHSC and MPP cell types, respectively, indicating the similarity of MPP and HSC (Fig G-H). Cell type mapping results of other cell types were all as expected. All these confirmed the conservation of hematopoiesis between human and mouse. We did not identify the MPP population in human because it cannot be separated from HSC with transcriptome.

Similar results were obtained when we compared the hematopoietic transcriptomes using the datasets from GSE81682 (mouse) and The Human Cell Atlas (human) with same analytical strategies.



A B C





**Erythroid/megakaryocytic**

**Myeloid**

**Lymphoid**

E



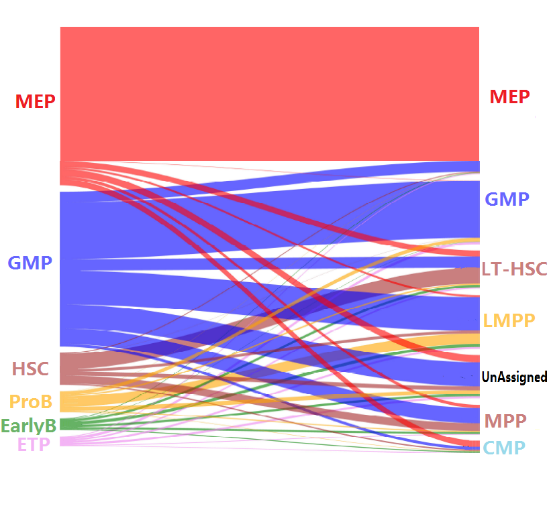
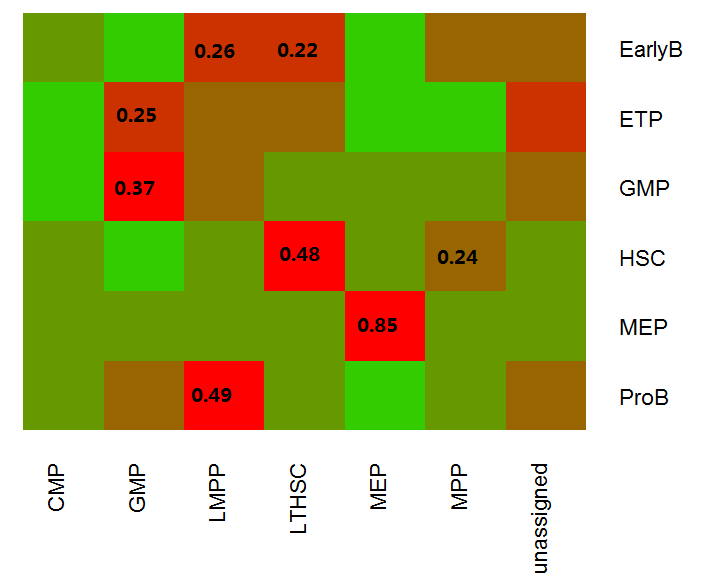
**Erythroid/megakaryocytic**

**Myeloid**

**Lymphoid**

D

F

G H