blood®
Regular Article

PHAGOCYTES, GRANULOCYTES, AND MYELOPOIESIS

Tracing the evolutionary history of blood cells to the unicellular ancestor of animals

Yosuke Nagahata,[1,2] Kyoko Masuda,[1] Yuji Nishimura,[1] Tomokatsu Ikawa,[3] Shinpei Kawaoka,[4] Toshio Kitawaki,[2] Yasuhito Nannya,[5] Seishi Ogawa,[5] Hiroshi Suga,[6] Yutaka Satou,[7] Akifumi Takaori-Kondo,[2] and Hiroshi Kawamoto[1]

[1]Laboratory of Immunology, Institute for Life and Medical Sciences and [2]Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [3]Division of Immunology and Allergy, Research Institute for Biomedical Sciences, Tokyo University of Science, Chiba, Japan; [4]Inter-Organ Communication Research Team, Institute for Life and Medical Sciences and [5]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [6]Department of Life and Environmental Sciences, Prefectural University of Hiroshima, Shobara, Japan; and [7]Department of Zoology, Graduate School of Science, Kyoto University, Kyoto, Japan

KEY POINTS

- The initial blood cells emerged in the common ancestor of animals inheriting a phagocytic program from unicellular organisms.
- In murine hematopoiesis, CEBPα is commonly repressed by polycomb complexes to maintain nonphagocytic lineages.

Blood cells are thought to have emerged as phagocytes in the common ancestor of animals followed by the appearance of novel blood cell lineages such as thrombocytes, erythrocytes, and lymphocytes, during evolution. However, this speculation is not based on genetic evidence and it is still possible to argue that phagocytes in different species have different origins. It also remains to be clarified how the initial blood cells evolved; whether ancient animals have solely developed de novo programs for phagocytes or they have inherited a key program from ancestral unicellular organisms. Here, we traced the evolutionary history of blood cells, and cross-species comparison of gene expression profiles revealed that phagocytes in various animal species and *Capsaspora* (*C.*) *owczarzaki*, a unicellular organism, are transcriptionally similar to each other. We also found that both phagocytes and *C. owczarzaki* share a common phagocytic program, and that CEBPα is the sole transcription factor highly expressed in both phagocytes and *C. owczarzaki*. We further showed that the function of CEBPα to drive phagocyte program in nonphagocytic blood cells has been conserved in tunicate, sponge, and *C. owczarzaki*. We finally showed that, in murine hematopoiesis, repression of CEBPα to maintain nonphagocytic lineages is commonly achieved by polycomb complexes. These findings indicate that the initial blood cells emerged inheriting a unicellular organism program driven by CEBPα and that the program has also been seamlessly inherited in phagocytes of various animal species throughout evolution.

Introduction

Among various lineage blood cells, such as erythrocytes and lymphocytes, phagocytes including macrophages and neutrophils have been thought to represent the most evolutionarily ancient blood cells because phagocytes can be found in any animal including organisms that are morphologically very simple multicellular like the sponge,[1] whereas more lineage types can be seen in more complex animals.[2-5] It has thus been speculated that the evolutionary initial blood cells emerged as phagocytes in the common ancestor of animals, and that various nonphagocyte lineages have evolved from the primordial phagocytes during evolution. Concerning this issue, we have demonstrated that the potential to produce phagocytes is retained in the early progenitors primed for erythroid, T- and B-cell lineages in murine hematopoiesis.[6-10] Based on such findings, we have proposed that the retention of phagocyte potential in these lineage progenitors is a vestige of the phylogenic process, where each of these lineages has evolved from ancestral phagocytes.[2,11] The vestige has also been found in other vertebrates: thrombocytes, erythrocytes, and B cells in shark, bony fish, and frog have phagocytic potential.[12-14]

One thing to note here is that such speculation can be made provided that all phagocytes have the same origin during phylogeny. However, genetic evidence supporting this model has been insufficient, and we can still argue a possibility of convergent evolution: phagocytes in different animal species have different origins. Furthermore, it remains to be clarified how the initial blood cells evolved. We can argue 2 possible cases: the first is that ancient animals have solely developed de novo programs for phagocytes, and the second is that they inherited a key program from ancestral unicellular organisms.

To address this issue, we decided to clarify whether a common program has been shared in phagocytes of various animal species and whether the program is also shared with a unicellular organism. To this end, we compared gene expression

profiles in phagocytes and nonphagocytes of various animal species, and unicellular organisms.

Methods

Mice

Ert2Cre-Cdkn2a$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$, *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$-*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ and *LckCre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were generated and maintained in our animal facility. All mice were maintained in specific pathogen-free conditions in our animal facility. All experiments were performed in accordance with the guidelines of the Kyoto University animal experiment committee and approved by our institutional committee.

Tunicate

Ciona intestinalis (type A; also called *Ciona robusta*) adults were obtained from the National BioResource Project for *Ciona*.

Capsaspora

Capsaspora owczarzaki was maintained at 23°C in the ATCC 1034 medium as previously reported.[15]

Data and code availability

Public data of mouse in EMBL-EBI (supplemental Table 1, available on the *Blood* website) and data of mouse, tunicate, sponge, *C. owczarzaki*, *Salpingoeca rosetta*, and *Creolimax fragrantissima* in previous reports were analyzed[15-23]. RNA sequencing (RNA-seq) data of tunicate phagocytes and *Ring1a/b* knockout (KO) myeloid cells are available at DNA Data Bank of Japan database (DRA013007 and DRA014437).

Cross-species transcriptomic comparison

We identified homologs in *Mus musculus*, *C. intestinalis*, *Amphimedon queenslandica*, and *C. owczarzaki* using the OrthoFinder (supplemental Table 2).[24] Homolog groups commonly conserved across the 4 species were selected and used for cross-species comparison (supplemental Table 3). Cross-species analysis of 6 species adding *S. rosetta* and *C. fragrantissima* was also performed (supplemental Tables 4-5).

Transcription factors (TFs) and phagocytosis-related genes

For selecting TFs and phagocytosis/lysosome-related genes, we used the AmiGO2 database (http://amigo.geneontology.org/amigo) (supplemental Table 6).

Isolation of mouse progenitors

Single-cell suspensions of the thymus or bone marrow (BM) were prepared and progenitors were isolated by fluorescence-activated cell sorting. Gating strategies are shown in supplemental Figure 1.

CEBPα and Ring1B encoding vectors

Codon-optimized DNA sequences of CEBPα and Ring1B were synthesized using GeneArt (Thermo Fisher Scientific) (supplemental Table 7).

Retrovirus production and transduction

CEBPα- and Ring1B-encoding vectors were transfected into the Plat-E cells (CosmoBio) and supernatants were harvested. For transduction, purified progenitors were resuspended with the supernatant, and were centrifuged for 90 minutes at 1000×*g* at 32°C.

Phagocytosis assay

pHrodo-green zymosan or *Staphylococcus aureus* beads (Invitrogen) were added to each culture. One hour later, the medium was replaced with phosphate-buffered saline and phagocytosis was observed using a fluorescence microscope.

RNA extraction and real-time quantitative polymerase chain reaction

Total RNA was isolated using an RNeasy kit (Qiagen). Complementary DNA synthesis was performed using a SuperScript IV VILO Master Mix complementary DNA synthesis kit (Invitrogen). Real-time polymerase chain reaction was performed using PowerUp SYBR Green Master Mix (Applied Biosystems) and analyzed by StepOnePlus (Applied Biosystems).

RNA-seq of tunicate phagocytes and *Ring1a/b* KO myeloid cells

Libraries were prepared using SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara) and Nextera XT DNA Library Prep kit (Illumina) and sequenced on a NovaSeq 6000 (Illumina).

In vitro deletion of *Ring1b*

The isolated progenitors were cocultured with TSt4[25] or TSt4-DLL1[26] cells for 4 to 12 days, and *Ring1b* was deleted by 4-hydroxytamoxifen (4-OHT).

BM chimera mice

Hemolyzed whole BM cells (2 × 10^6 cells) were IV injected into sublethally irradiated (4 Gy) *Rag2*$^{-/-}$ mice. For long term observation, 1 × 10^6 BM cells were transplanted with 1 × 10^6 competitor cells.

Statistical analysis

Survival rates were estimated using Kaplan-Meier methods and compared using log-rank tests. Continuous and categorical variables were compared using 2-tailed *t* tests and Fisher exact test, respectively.

Further experimental details are provided in supplemental methods.

Results

Phagocytes of mouse, tunicate, and sponge are transcriptionally similar to a unicellular organism

We compared gene expression profiles of various lineage or stage cells among 4 species: mouse (*M. musculus*), tunicate (*C. intestinalis*), sponge (*A. queenslandica*), and *C. owczarzaki*, a unicellular organism (hereafter *Capsaspora*) (Figure 1A). Among invertebrates, we selected tunicate and sponge because tunicate belongs to chordates and is close to vertebrates, whereas sponge is the animal oldest and farthest from vertebrates.[27,28] Among unicellular organisms, *Capsaspora* was selected because it is phylogenetically close to animals, forming a clade termed Holozoa together with Metazoa (Figure 1A).[29-31] We first searched homologs conserved among the 4 species and 3237 homolog groups were identified; 5911 genes in mouse, 4031 genes in tunicate, 5443 genes in sponge, and 4096 genes

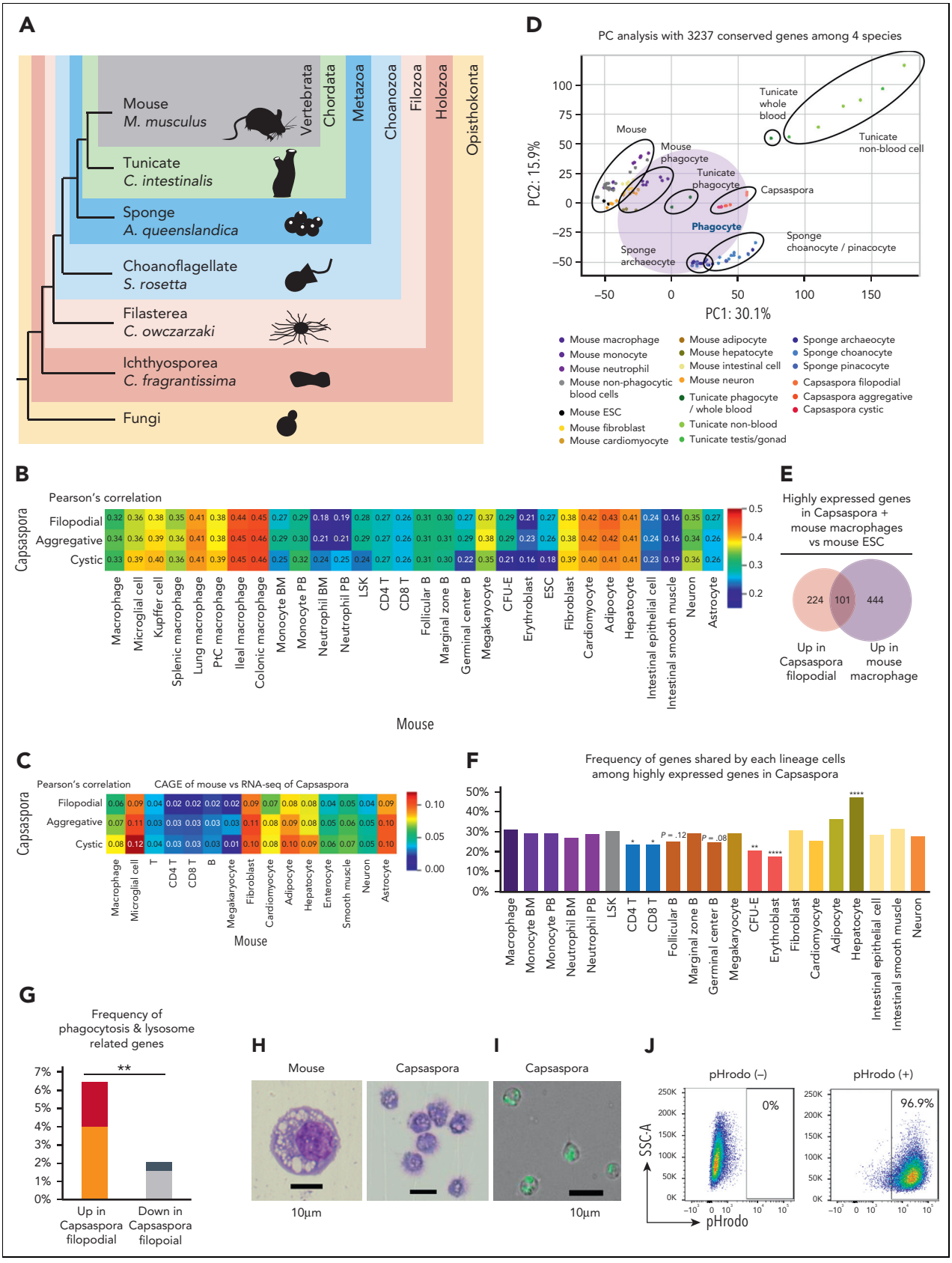


Figure 1.

in *Capsaspora* were assigned to the 3237 homolog groups. Then, gene expression profiles were compared based on the homolog groups (supplemental Figure 2A). As expected, mouse, tunicate, sponge, and *Capsaspora* were very different form each other (supplemental Figure 2B). Among blood cells, macrophages were more similar to *Capsaspora* than nonphagocytic cells were (Figure 1C-D). Macrophages were also more similar to *Capsaspora* than neutrophils, in line with the fact that neutrophils with multilobulated nuclei are unique to vertebrates.[32] In order to exclude batch effect between mouse data sets, comparison using a single data set of mouse cells with the cap analysis gene expression (CAGE) method was also performed (Figure 1C). In both the analysis with RNA-seq and CAGE data sets, macrophages, hepatocytes, fibroblasts, and adipocytes among mouse cells showed high similarity to Capsasapora (Figure 1B-C). Because hepatocytes, fibroblasts and adipocytes are known to have phagocytic potential,[33-35] macrophages and these 3 lineage cells can be categorized as phagocytes. In principle component (PC) analysis, phagocytes of mouse and tunicate, sponge archaeocytes, which are known to have phagocytic potential,[1] and *Capsaspora* showed similarity to each other (Figure 1D).

Next, we examined how frequently *Capsaspora* and various mouse cell lineages share highly expressed genes; number of genes expressed higher than ESCs were examined. *Capsaspora* and macrophages highly expressed 325 and 545 genes, respectively, and they shared 101 genes (Figure 1E). Macrophages shared more genes with *Capsaspora* than other blood cell lineages (Figure 1F and supplemental Figures 3-4). Hepatocytes also shared many genes with *Capsaspora* and shared more with macrophages among nonblood cells (supplemental Figures 3-4). Kyoto Encyclopedia of Genes and Genomes pathway analysis showed that lysosome-related genes were among genes shared by *Capsaspora*, macrophages, and hepatocytes (supplemental Figure 5). Gene ontology analysis using AmiGO2 database showed that 325 genes highly expressed in *Capsaspora* were more frequently phagocytosis/lysosome-related genes compared with the 2252 low expressed genes (Figure 1G). These data suggested that phagocytosis- and lysosome-related genes shape the similarity between *Capsaspora* and mouse phagocytes. In fact, *Capsaspora* cells showed mouse macrophage–like cytology with several vacuoles in the cytoplasm (Figure 1H) and robust phagocytic activity (Figure 1I-J). These data suggested that the transcriptional profile of phagocytes has been conserved from common ancestors of *Capsaspora* and animals.

Phagocytes and a unicellular organism share a CEBPα-driven phagocytic program

Next, we compared gene expression profiles of *Capsaspora* and mouse macrophages with mouse ESCs and the nonphagocytic blood cells, $Lin^{-}Sca1^{+}ckit^{+}$ cells, T cells, B cells, megakaryocytes, and erythroid cells. Eleven genes were highly expressed in both mouse macrophages and *Capsaspora* (Figure 2A and supplemental Figure 6A), and these 11 genes were lysosome related, suggesting that these genes contribute to phagocytosis in phagosome/lysosome pathway (Figure 2B and supplemental Figure 6B). Nine of the 11 genes were also highly expressed in hepatocytes (supplemental Figure 6A). Next, we attempted to reveal which TFs commonly play a key role in both *Capsaspora* and mouse phagocytes. We found that 62 TFs were conserved among the 4 species, and then we compared their expression levels. As with the comparison based on the 3237 conserved genes (Figure 1B-C), the comparison based on the 62 conserved TFs showed that mouse phagocytes were closer to *Capsaspora* than mouse nonphagocytes were (Figure 2C and supplemental Figure 7A-B). CEBPα was the sole TF highly expressed in both *Capsaspora* and mouse macrophages compared with mouse ESCs and nonphagocytic blood cells (Figure 2D-F and supplemental Figure 8A). Several regions of the CEBPα homologs, especially DNA binding bZIP domain, were conserved among the 4 species (supplemental Figure 9). Other TFs were also conserved among the 4 species (supplemental Figure 8B), and CEBPγ, another CEBP homolog, was also examined because we were not able to distinguish which was a functional CEBPα homolog in the phylogenetic tree (supplemental Figure 10). However, we found that expression levels of CEBPγ were not highly expressed in *Capsaspora* (supplemental Figure 8C). Expression levels of GATA1-6 homologs in *Capsaspora*, macrophages, and hepatocytes were lower than in megakaryocytes and erythroid cells (supplemental Figure 8D), and those of EBF1-4 were lower than B cells (supplemental Figure 8E). Relatively high expression levels of GATA and EBF families in *Capsaspora* and some mouse nonhematopoietic cells suggested that these TFs determine programs conserved among *Capsaspora* and mouse nonhematopoietic lineages.[36-38] Although PU.1 and IRF are important in murine myeloid cells,[39,40] their homologs were not detected in *Capsaspora* (supplemental Figure 8B). When gene expression levels were compared between the 3 stages of *Capsaspora*, CEBPα was expressed more in filopodial or cystic stages than in the aggregative stage (Figure 2F). Among the 11 genes highly expressed in mouse macrophages and *Capsaspora*, PLA2G15 was also expressed more in filopodial and cystic stages (Figure 2F and supplemental Figure 6A). PLA2G15 is a lysosomal protein and plays a role in host defense and efferocytosis by human phagocytes.[41,42] These data suggested that a CEBPα-driven phagocytic program including PLA2G15 expression has been conserved between a unicellular organism and vertebrates.

We also performed cross-species analysis adding a choanoflagellate (*S. rosetta*) and Ichthyosporea (*C. fragrantissima*). In

Figure 1. Phagocytes of mouse, tunicate, and sponge are transcriptionally similar to a unicellular organism. (A) Phylogenetic tree of mouse, tunicate, sponge, choanoflagellate, *Capsaspora*, Ichthyosporea, and fungi. (B-C) Heat map with Pearson correlation of various mouse cell lineages and *Capsaspora*. Gene expression profiles were compared among 3 stages of *Capsaspora* and 30 mouse lineages (B) or 15 lineages (C) based on 3237 conserved homologs. Transcriptome data examined by RNA-seq (B) or CAGE method (C) were analyzed. (D) PC analyses of various lineages or stages of 4 species: *Capsaspora*, sponge, tunicate, and mouse. Expression levels of 3237 conserved homologs were normalized and compared. (E) Venn diagrams with the number of highly expressed genes in *Capsaspora* filopodial stage or mouse macrophages compared with mouse ESCs. (F) Frequency of genes shared by various mouse cell lineages among 325 highly expressed genes in *Capsaspora* filopodial stage. Statistical significance of differences between macrophage and the other lineages are also shown. (G) Frequency of phagocytosis-related genes among 325 genes highly expressed in *Capsaspora* filopodial stage and 2252 genes low expressed in *Capsaspora* filopodial stage compared with mouse ESCs. Frequency of phagocytosis- and lysosome-related genes expressed higher in mouse macrophages than mouse ESCs are shown. Frequency of genes highly expressed in macrophages compared to mouse ESCs and nonphagocytic blood cells are shown in red and black, respectively. (H) Cytology of mouse phagocyte (left) and *Capsaspora* (right) was examined by Wright-Giemsa staining. (I-J) Phagocytic activity of *Capsaspora* was evaluated by engulfment of pHrodo-green beads (I), and frequency of phagocytic cells was evaluated by flow cytometry (J). Data are representative of 2 independent experiments. $*P < .05$, $**P < .01$, $****P < .0001$. ESCs, embryonic stem cells; PtC, peritoneal cavity.

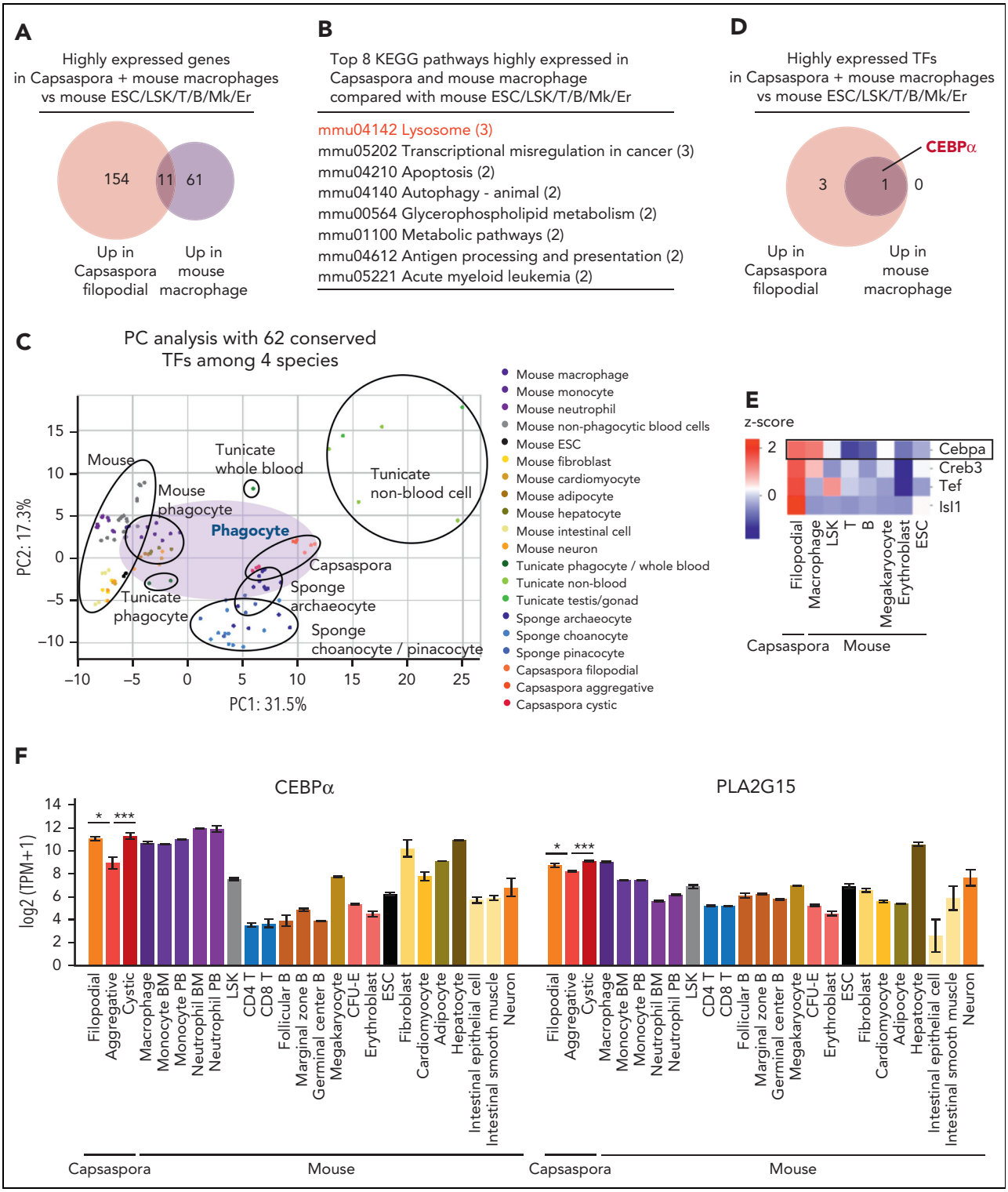


Figure 2. Phagocytes and a unicellular organism share a CEBPα-driven phagocytic program. (A,D) Venn diagrams with the number of highly expressed genes (A) and TFs (D) in *Capsaspora* or mouse macrophages compared with mouse ESCs and nonphagocytic blood cells. (B) Top 8 KEGG pathways involved in the 11 genes highly expressed in *Capsaspora* and mouse macrophages compared with mouse ESCs and nonphagocytic blood cells. (C) PC analyses of various lineages or stages of 4 species: *Capsaspora*, sponge, tunicate, and mouse. Expression levels of 62 conserved TFs were compared. (E) Heatmap of scaled expression levels (z score) of TFs in *Capsaspora*, mouse macrophages, mouse ESCs, and mouse nonphagocytic blood cells. Four TFs expressed higher in *Capsaspora* or mouse macrophages than in mouse ESCs and nonphagocytic blood cells were selected. Expression levels were scaled among the 8 cell groups. (F) Expression levels of CEBPα homologs and PLA2G15 homologs in *Capsaspora*, and various mouse cell lineages. Data are mean ± standard error of the mean. Statistical significance of differences between 3 stages of *Capsaspora* are shown, $^{*}P < .05$, $^{**}P < .01$. KEGG, Kyoto Encyclopedia of Genes and Genomes.

this analysis, phagocytes of various species also showed similarity to each other and to unicellular organisms (supplemental Figure 11A-C). In mouse cell lineages, macrophages and adipocytes showed high similarity to unicellular organisms (supplemental Figure 11B-C). *Hgd* was highly expressed in mouse macrophages, *Capsaspora*, and *C. fragrantissima* (supplemental Figure 11D). However, because both *S. rosetta* and *C. fragrantissima* lack CEBPα, no TFs highly expressed in all of mouse macrophages, *Capsaspora*, and *C. fragrantissima* were detected. Some important genes other than CEBPα may determine the similarity of these cells.

Tunicate and sponge phagocytes highly express CEBPα homologs

Next, we examined whether expression levels of CEBPα were different between phagocytes and nonphagocytic blood cells in sponge and tunicate. In sponges, we focused on archaeocytes, which behave like blood cells in that they circulate around the body cavity and have phagocytic potential.[1] Analysis of archaeocytes showed that CEBPα expression levels were positively correlated with those of phagocytosis-related genes and PLA2G15, but CEBPγ levels were not (Figure 3A).

We also examined tunicate blood cells and their expression of CEBP homologs and phagocytosis-related genes. CEBPα and phagocytosis-related genes were highly expressed in the blood cells, especially in phagocytes, but CEBPγ was not (Figure 3B). In order to investigate whether CEBPα is differently expressed among various blood lineage cells in tunicate, we collected blood cells from tunicates (Figure 3C). The blood cells were then sorted into 4 fractions based on characteristics of (1) small size (hemoblasts), (2) autofluorescence (morula cells), (3) fluorescence of engulfed beads (phagocytes), and (4) negative for these features (other blood cells) (Figure 3D). We found that the expression levels of CEBPα and PLA2G15 were remarkably higher in phagocytes compared to other lineages of blood cells, whereas the expression level of CEBPγ was not or only slightly (Figure 3E). These data may indicate that, in both sponge and tunicate, CEBPα commonly exert a phagocyte program.

Function of CEBPα to drive the phagocyte program has been conserved from a unicellular organism

Next, we asked whether CEBPα of the tunicate, sponge, and *Capsaspora* has a function similar to mouse CEBPα, the enforced expression of which has been shown to convert T and B cells into phagocytes.[43-46] First, mouse pro-B cells were transduced with CEBPα of mouse, tunicates, sponge, or *Capsaspora* (Figure 4A). CEBPα of tunicate and sponge, as well as mouse CEBPα, converted these B progenitors into cells that express CD11b, whereas the CEBPα of *Capsaspora*, and CEBPγ of sponge and *Capsaspora*, did not (Figure 4B and supplemental Figure 12A). The majority of the CD11b$^+$ cells induced by either tunicate or sponge CEBPα looked like macrophages and showed efficient phagocytic activity (Figure 4C-D). D-J–rearranged *IgH* genes were present in the generated CD11b$^+$ cells (supplemental Figure 12B), indicating that they were derived from pro-B cells. In order to clarify whether CEBPα of *Capsaspora* has the potential to drive the phagocyte program, we further examined other lineage progenitors. MkPs, ErPs, and DN3 T-cell progenitors were examined. CEBPα of *Capsaspora* as well as that of mouse, tunicate, and sponge converted MkPs into CD11b$^+$ phagocytes, whereas CEBPγ of *Capsaspora* did not (Figure 4E-G and supplemental Figure 12C), indicating that *Capsaspora* CEBPα has the potential to drive the phagocytic program. We also found that CEBPα of mouse, tunicate, sponge, and *Capsaspora* converted ErPs into CD11b$^+$ cells (Figure 4H and supplemental Figure 12D). DN3 T-cell progenitors were converted into CD11b$^+$ cells by mouse and sponge CEBPα but not by the tunicate and *Capsaspora* homologs (Figure 4I and supplemental Figure 12E).

We then examined how functionally similar the CEBPα homologs were. CEBPα is known to play roles in the differentiation of mouse neutrophils, and indeed, mouse CEBPα converted pro-B cells into neutrophil-like cells with ring-shaped or multilobulated nuclei, whereas CEBPα of tunicate and sponge hardly did so (Figure 4J-K). The expression levels of various genes were also compared between pro-B cells transduced with the mouse or sponge CEBPα, which converted pro-B cells into phagocytes to a similar extent (Figure 4B). To examine the direct consequence of *Cebpa* gene expression, we collected the cells on day 2, when they had not yet begun to express CD11b (supplemental Figure 12F-G). Sponge CEBPα upregulated phagocyte-associated genes to the same extent as mouse CEBPα, but mouse CEBPα was superior to sponge CEBPα in inducing expression of neutrophil-associated genes and in repressing B cell–associated genes (Figure 4L and supplemental Figure 12H).

Polycomb-mediated suppression of CEBPα is required for maintenance of various hematopoietic lineages in mouse

In mouse blood cells, CEBPα functions as master regulators of phagocytes, or myeloid cells in other words, having the potential to convert nonphagocytic lineage progenitors into myeloid cells,[43-50] implying that CEBPα must be strictly repressed for maintenance of nonphagocytic lineages. We attempted to reveal how CEBPα is repressed in nonphagocytic lineage cells, and hypothesized that the polycomb complex, one of major epigenetic repressors,[51] plays a role in suppression of the phagocyte program. We focused on Ring1A and B, which are catalytic components of polycomb complexes.[52] Expression levels of Ring1B were higher in nonphagocytic lineages than in myeloid cells reciprocally to those of CEBPα (supplemental Figure 13A-B). By analyzing published data, the *Cebpa* locus encoding CEBPα was found to be heavily marked with H3K27me3 in DN3 cells, pro-B cells, ErPs, and MkPs, but not in myeloid cells (supplemental Figure 13C). In contrast, *Spi1* locus encoding PU.1 was found not to be marked with H3K27me3 (supplemental Figure 13D). We also observed Ring1B binding at the *Cebpa* locus (supplemental Figure 13E). In order to confirm that CEBPα is suppressed by polycomb, we deleted *Ring1b* by using 4-OHT in progenitors of each lineage from *Ert2Cre -Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice (supplemental Figure 13F). In this experiment, we used *Cdkn2a*$^{-/-}$ background mice because *Ring1a/b* KO may cause derepression of Cdkn2a, leading to apoptosis of *Ring1a/b*-deleted cells.[53] Upon deletion of *Ring1b*, expression levels of CEBPα were remarkably elevated within a few days in all

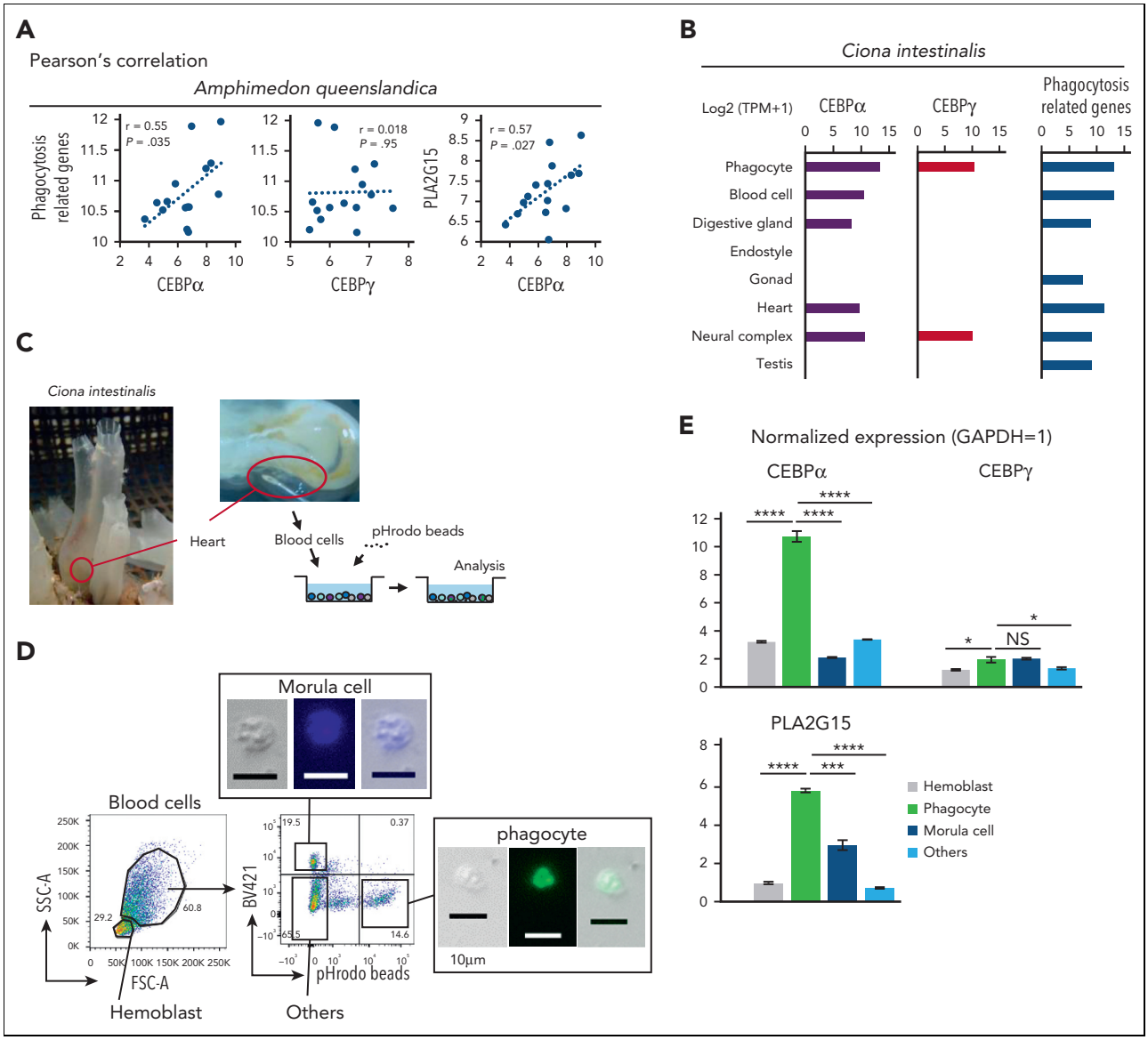


Figure 3. Tunicate and sponge phagocytes highly express CEBPα homologs. (A) Scatter plots of sponge archaeocytes with $\log_2$ (TPM + 1) values. The x-axes indicate sponge CEBPα and CEBPγ. The y-axes indicate total expression levels of phagocytosis-related genes and PLA2G15 (CEBPα homologs were excluded from phagocytosis-related genes in these analyses). (B) Expression levels with $\log_2$ (TPM + 1) values of CEBPα, CEBPγ, and phagocytosis-related genes in tunicate. Transcriptome data of phagocytes were examined by RNA-seq, and data of the other lineages were based on expressed sequence tag counts obtained from the Ghost Database (http://ghost.zool.kyoto-u.ac.jp/cgi-bin/gb2/gbrowse/kh/). (C) Blood cells of tunicate was aspirated by cardiac puncture. Collected blood cells were incubated with pHrodo beads and analyzed by flow cytometry. (D) Blood cells of tunicate were analyzed by flow cytometry based on their size, autofluorescence, and fluorescence of engulfed beads. (E) Normalized expression levels (*Gapdh* = 1) of CEBPα, CEBPγ, and PLA2G15 in various lineage blood cells of tunicate were evaluated by RT-qPCR. Data are mean ± standard error of the mean. $*P < .05$, $***P < .001$, $****P < .0001$.

lineages (supplemental Figure 13G). These data indicate that polycomb complexes commonly suppress CEBPα in various nonphagocytic lineages.

Next, we examined whether polycomb-mediated CEBPα suppression is physiologically important. We made BM chimera mice by transplantation of BM cells from *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$-*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice into sublethally irradiated *Rag2*$^{-/-}$ mice. Six weeks after transplantation, *Ring1b* was deleted by administration of tamoxifen, and mice were analyzed 2 weeks later (Figure 5A). The number of thymocytes, double-positive cells, DN cells, and DN3 cells in the green fluorescent protein–positive (GFP^+) fraction was decreased, whereas that of DN1 cells was increased in the *Ring1a/b* KO BM chimera mice (Figure 5B,E and supplemental Figure 14A-B,E). We also found a decrease in the number of pro-B cells and an increase of the number of B-1 progenitors, defined as $CD19^+B220^-$ cells (Figure 5C,F and supplemental Figure 14F). $Lin^-Sca1^+ckit^+$ cells including hematopoietic stem cells were decreased, whereas $Lin^-Sca1^-ckit^+$ cells were increased (supplemental Figure 14C,G). The proportion of ErPs and MkPs was decreased, whereas common myeloid progenitors were increased, and megakaryocyte-erythroid progenitors were intact (Figure 5D,G and supplemental Figure 14D,H). Because hematopoiesis of the BM chimera mice was severely impaired, they died within a few months (Figure 5H).

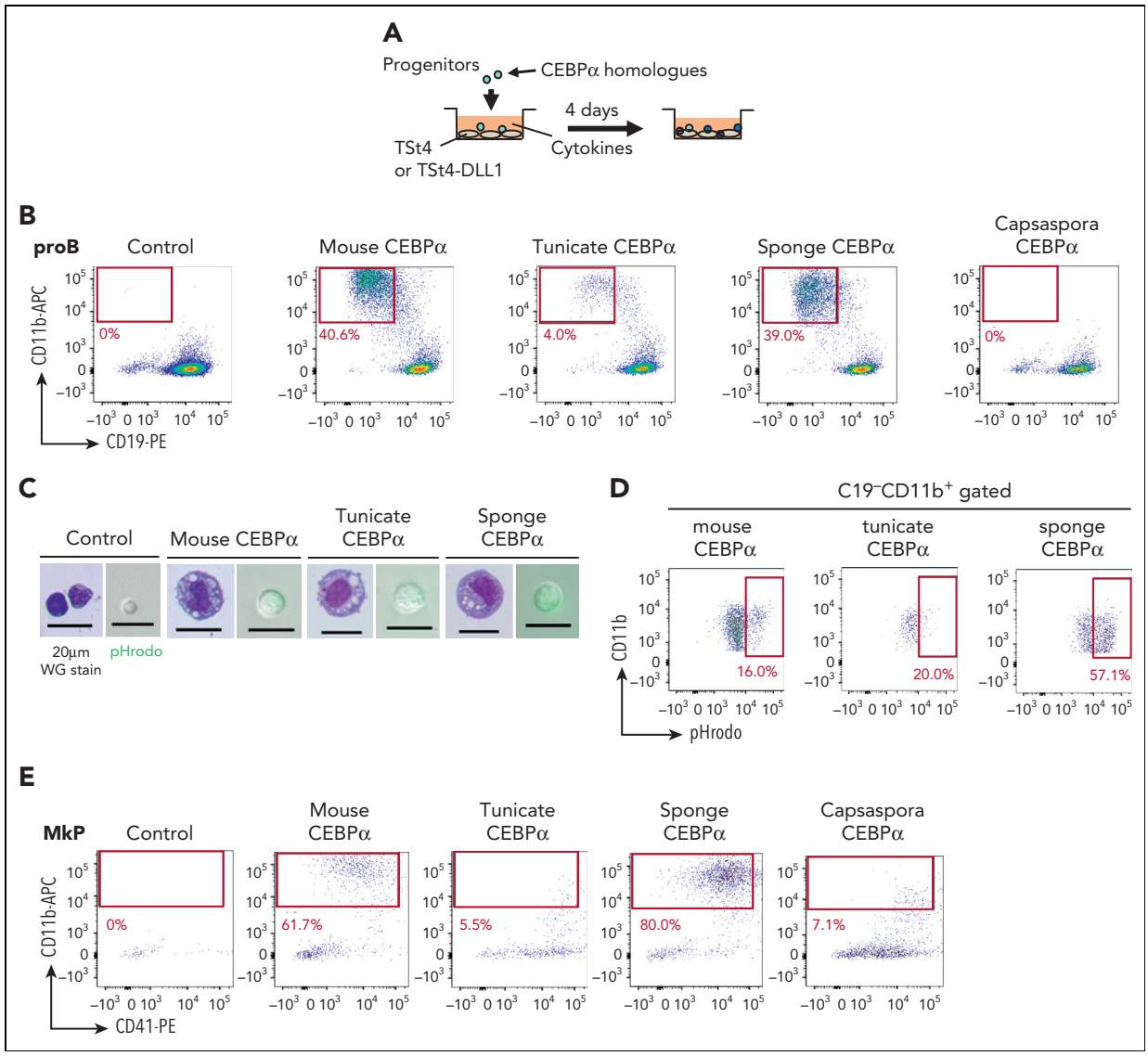


Figure 4. Function of CEBPα to drive the phagocyte program has been conserved from a unicellular organism. (A) Mouse CEBPα and its homologs from tunicate, sponge, and *Capsaspora* were transduced into pro-B cells, which were analyzed by flow cytometry 4 days later. (B,E,H-I) pro-B cells (B), MkPs (E), ErPs (H) and DN3 cells (I) were transduced with mouse, tunicate, sponge, or *Capsaspora* CEBPα and then examined by flow cytometry for the indicated lineage markers. Data are representative of 2 to 4 independent experiments. (C,F) The CD11b$^+$ cells generated by transduction with various CEBPα homologs into pro-B cells (C) and MkPs (F) were sorted and their cytology was examined by Wright-Giemsa staining (left). Their phagocytic activity was evaluated by engulfment of pHrodo-green beads (right). (D,G) Phagocytic activities of the generated CD11b$^+$ cells from pro-B cells (D) and MkPs (G) was evaluated by flow cytometry. (J) Wright-Giemsa stain of neutrophil-like cells with ring-shaped or multilobulated nuclei generated by transduction with mouse CEBPα into pro-B cells. (K) Frequency of cell types evaluated by cytology with Wright-Giemsa staining. Cells (n = 100) transduced with mouse, tunicate, or sponge CEBPα were examined. (L) Relative expression of neutrophil-associated genes in pro-B cells 2 days after CEBPα transduction. Relative expression levels (day 0 = 1) with $2^{-\Delta\Delta CT}$ values normalized with β-actin were shown. Data are mean ± standard error of the mean of 3 replicates. $^{**}P < .01$, $^{***}P < .001$. DN3, double-negative 3; ErPs, erythroid progenitors; MkPs, megakaryocyte progenitors.

To evaluate the long-term effect of *Ring1a/b* KO in blood cells, we performed transplantation of *Ring1a/b* KO BM cells with competitor BM cells, which should contribute normal hematopoiesis (Figure 5I). Eight weeks after deletion of *Ring1b*, almost all GFP$^+$ *Ring1a/b* KO cells became CD11b$^+$ myeloid cells (Figure 5J-K). Furthermore, the BM of *Ring1a/b* KO mice was occupied with myeloid cells and exhibited an anemic appearance, and the mice died within 3 months (Figure 5L and supplemental Figure 15A-D). These GFP$^+$ *Ring1a/b* KO myeloid cells expressed CD34 and looked like immature blasts (supplemental Figure 15E-F). Various lineage progenitors of thymocytes and BM cells, including competitor cells, were decreased, indicating that *Ring1a/b* KO myeloid cells were transformed into leukemic blasts and disturbed normal hematopoiesis (supplemental Figure 15G-L). We then examined whether sole *Ring1a/b* KO without *Cdkn2a* KO causes leukemia. We found that mice with *Cdkn2a*$^{+/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/\Delta}$ cells did not develop leukemia, and GFP$^+$ cells disappeared (Figure 5J and supplemental Figure 15D). This result suggested that the KO of *Ring1a/b*, leaving *Cdkn2a*$^{+/-}$ still present, led to an overexpression of Cdkn2a, resulting in apoptosis of KO cells as previously reported in the T cell–specific KO case.[53]

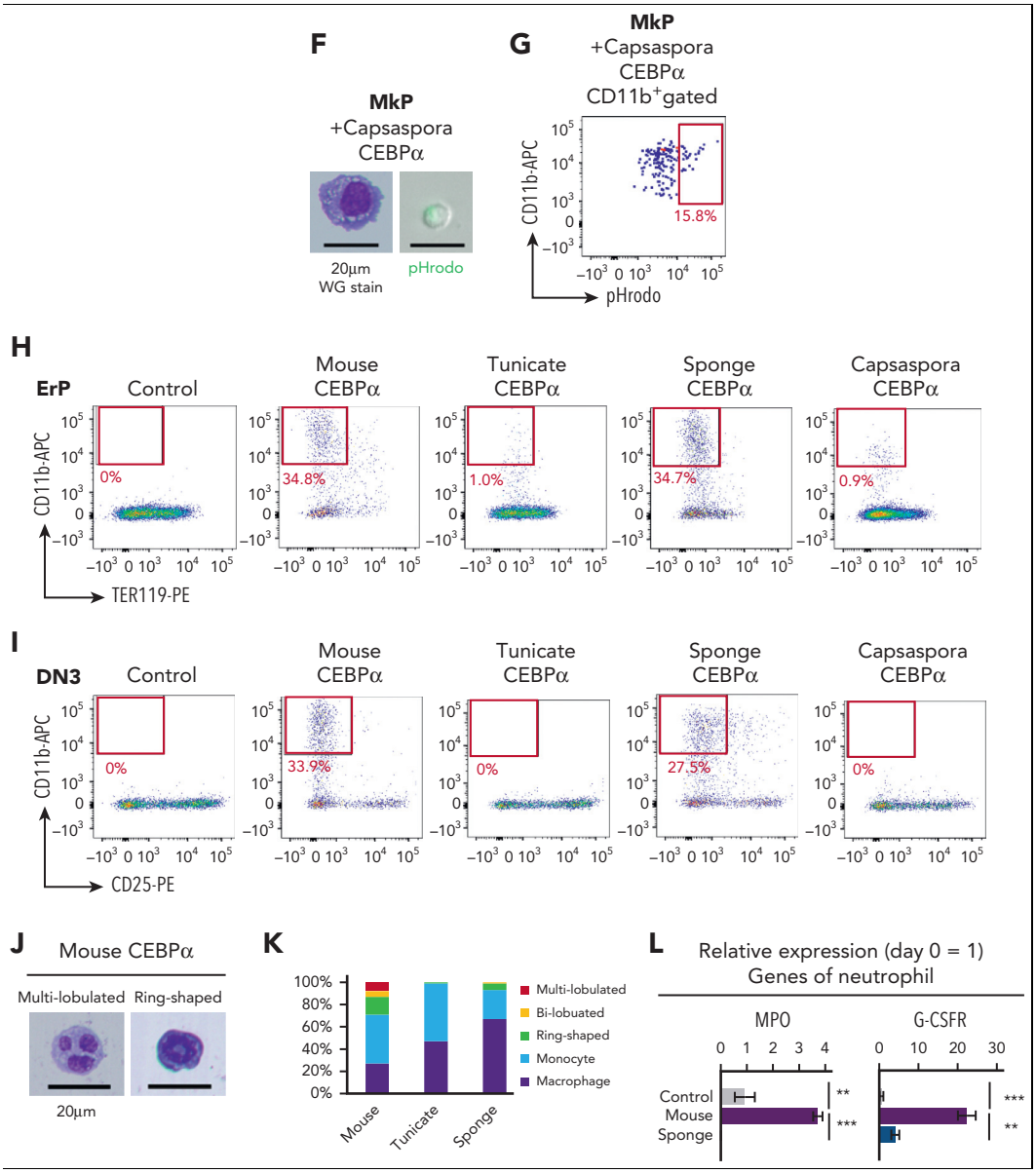


Figure 4 (continued)

Various lineage progenitors were reverted into the primordial lineage of phagocytes by *Ring1a/b KO*

In BM chimera mice, we showed that the number of various lineage progenitors was decreased, whereas that of myeloid cells was increased (Figure 5). Next, we tested whether cell fate conversion from each of the lineage progenitors into myeloid cells had occurred. First, we found that the myeloid cells from the *Ring1a/b* KO mice carried rearranged *IgH* genes but those of control mice showed no rearrangements (supplemental Figure 16A). Among 8 *Ring1a/b* KO BM chimera mice examined, 5 carried *IgH*-rearranged myeloid cells. These data indicate that B cells were converted into myeloid cells in vivo. In order to examine whether various lineage progenitors are converted into myeloid cells by *Ring1a/b* KO, DN3 cells, pro-B cells, ErPs, and MkPs of *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were cultured with or without 4-OHT (Figure 6A). Because these progenitors had already been determined to their respective lineages, control cells maintained their lineage identity (Figure 6B). In contrast, by deletion of *Ring1b*, these progenitors gave rise to CD11b^{+} macrophage-like cells (Figure 6B-C). DN3- and pro-B–derived myeloid cells harbored V-DJ–rearranged TCR genes and *IgH* genes, respectively, confirming that they had originated from T and B lineage progenitors (supplemental Figure 16B-C). We also observed lineage conversion from pro–B cells into myeloid cells via B-1 stage in vitro (supplemental Figure 16D-E), which was consistent with the increase in number of B-1 cells in BM chimera mice (Figure 5C and supplemental Figure 14F). We previously reported that *Ring1a/b* KO by *LckCre* converted T cells into B cells.[53] We again analyzed *LckCre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice and found that *Ring1a/b* KO DN3 cells expressed CD19 but some of them were B-1 phenotype lacking B220 expression

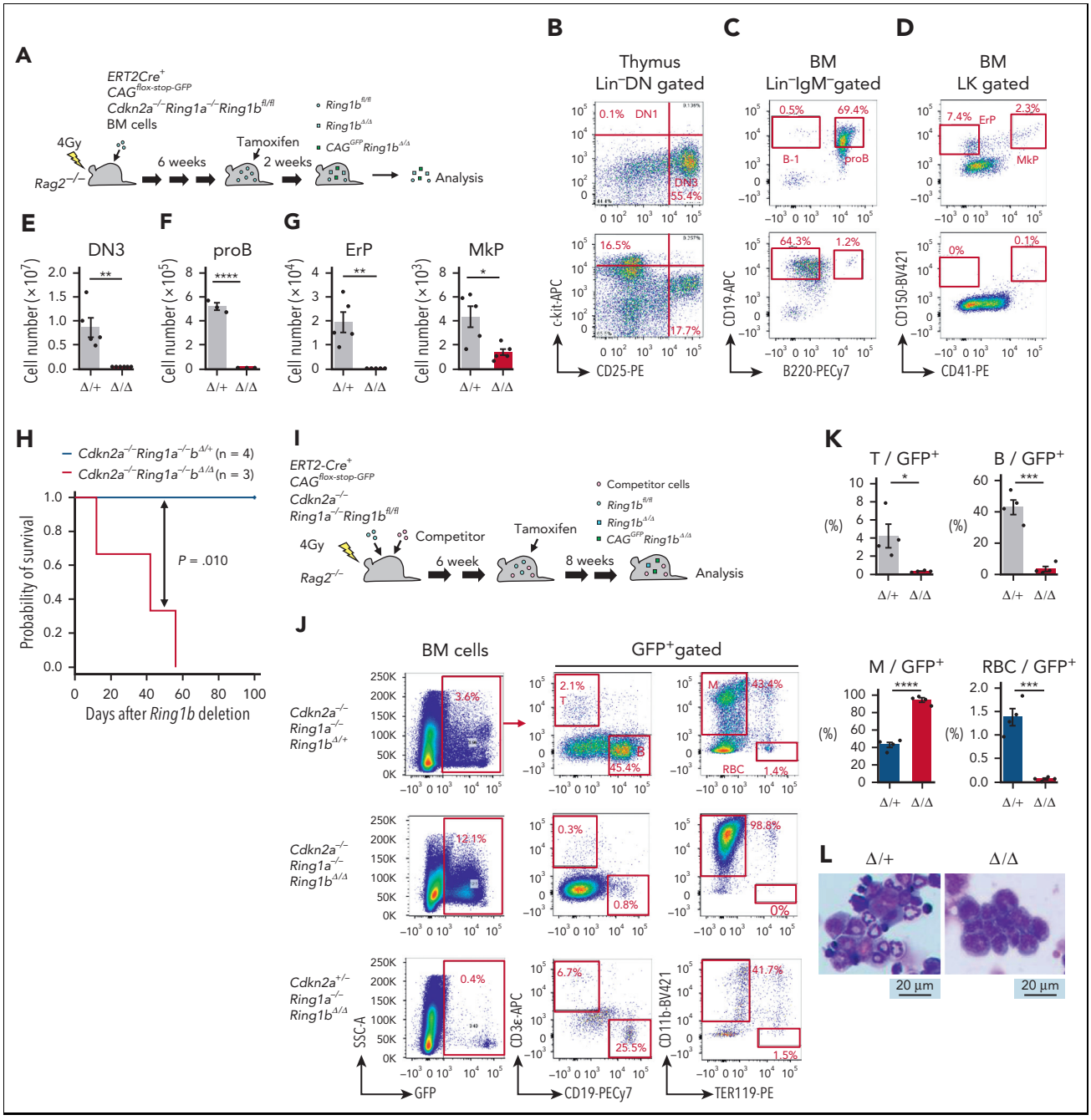


Figure 5. Polycomb-mediated suppression of CEBPα is required for maintenance of various hematopoietic lineages in mouse. (A,I) Experimental procedure for conditional inactivation of polycomb function. BM cells of *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$*-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice or *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$*-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/+}$ mice were transplanted without (A) or with (I) competitor cells by IV injection into sublethally irradiated *Rag2*$^{-/-}$ mice. Six weeks later, the transplanted mice were administrated tamoxifen intraperitoneally to delete *Ring1b* in blood cells. Two (A) or 8 (I) weeks after *Ring1b* deletion, mice were sacrificed and analyzed. (B-D) Flow cytometric profiles of GFP$^+$ thymocytes (B) and GFP$^+$ BM cells (C-D). Upper and lower panels show data of control (Δ/+; *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{4/+}$, n = 5 in panels B and D, and n = 3 in panel C) and *Ring1a/b* KO (Δ/Δ; *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{Δ/Δ}$, n = 6 in panel B, n = 3 in panel C, and n = 5 in panel D) mice, respectively. (E-G) Number of GFP$^+$ DN3 cells (E), pro-B cells (F), and ErPs and MkPs (G) of control (black) and *Ring1a/b* KO (red) mice. (H) Survival curve with Kaplan-Meier plots after BM transplantation to sublethally irradiated *Rag2*$^{-/-}$ mice. Blue and red lines show survival curve of control (*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{Δ/+}$, n = 4) and *Ring1a/b* KO (*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{Δ/Δ}$, n = 3) mice, respectively. Statistical significance of differences between the survival rates were calculated with the log-rank test. (J) Flow cytometric profiles of whole BM cells of control (n = 4), *Ring1a/b* KO in *Cdkn2a*$^{-/-}$ background (n = 4), and *Ring1a/b* KO in *Cdkn2a*$^{+/-}$ background (n = 3) mice with competitor cells. (K) Percentage of myeloid cells, RBCs, T cells, and B cells among GFP$^+$ BM cells of control (blue) and *Ring1a/b* KO (red) mice with competitor cells. (L) Wright-Giemsa stain of BM smears obtained from control and *Ring1a/b* KO mice with competitor cells. Data are mean ± standard error of the mean. *P < .05, **P < .01, ***P < .001, ****P < .0001.

(supplemental Figure 16F). In addition, DN3 cells of *LckCre* mice were converted into myeloid cells via B lineage cells carrying rearranged *IgH* and *Tcrb* genes (supplemental Figure 16G-J). Although *Ring1b* deletion converted non-phagocytic lineage cells into phagocytes, Ring1B overexpression did not convert phagocytes into non-phagocytic lineage cells (supplemental Figure 17A-E), indicating that polycomb complexes play a role in the maintenance of nonphagocytic lineages but not in induction of nonphagocytic lineages.

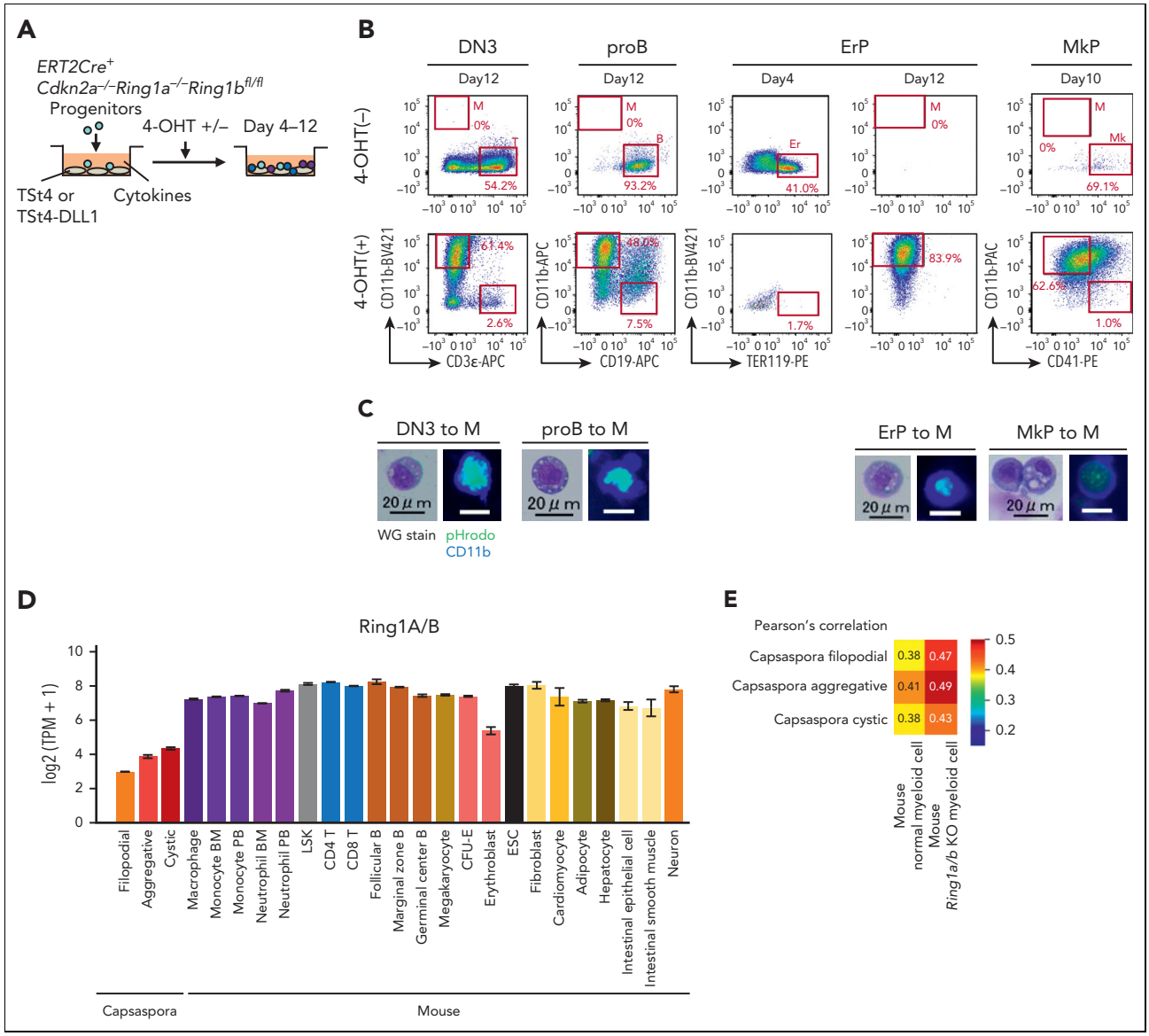


Figure 6. Various lineage progenitors were reverted into the primordial lineage of phagocytes by *Ring1a/b* KO. (A) DN3 cells, pro–B cells, ErPs, and MkPs isolated from *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were cocultured with TSt4 or TSt4-DLL1 cells for 4 to 12 days with or without 4-OHT in the presence of 10 ng/mL of stem cell factor, Flt3-L, interleukin 1α (IL-1α), IL-3, IL-7, tumor necrosis factor α, and granulocyte-macrophage colony-stimulating factor. For ErPs and MkPs, 2 U/mL of erythropoietin and 50 ng/mL of thrombopoietin were added, respectively. (B) Flow cytometric profiles of the cultured cells. Data are representative of 3 independent experiments. (C) Cytology of the generated CD11b^{+} cells was examined by Wright-Giemsa staining (left), and their phagocytic activity was evaluated by pHrodo-green beads with CD11b-BV421 staining (right). (D) Expression levels of Ring1A/B homologs in *Capsaspora* and various mouse cell lineages. (E) Heat map with Pearson correlation of mouse normal and *Ring1a/b* KO myeloid cells and *Capsaspora*.

Lastly, we found that expression levels of Ring1A/B homologs were low in *Capsaspora* (Figure 6D), and *Ring1a/b* KO myeloid cells were more similar with *Capsaspora* than normal myeloid cells (Figure 6E). These data suggested that *Ring1a/b* KO reverted mouse cells toward a primordial status close to *Capsaspora*, and that Ring1A/B has played a role in acquiring new lineages in evolution.

Discussion

Animals evolved from unicellular organisms,[29,30,54-57] and *Capsaspora*, which is known to exhibit typical filopodial features, is phylogenetically close to animals.[15,20,58-61] The present study enabled us to envisage that the phenotype of *Capsaspora* represents the origin of phagocytes in animals. We showed that *Capsaspora* has phagocytic potential and exhibit gene expression profiles similar to phagocytes of animals characterized by high CEBPα expression. Furthermore, we showed that CEBPα homologs converted murine nonphagocyte progenitors into phagocytes.

Here, we propose the following scenario in the evolutionary history of blood cells: when a unicellular ancestor came to form a multicellular organism, a body cavity structure surrounded by epithelium would have formed. In such a situation it would have been advantageous if the organism had an ancestral type of cell

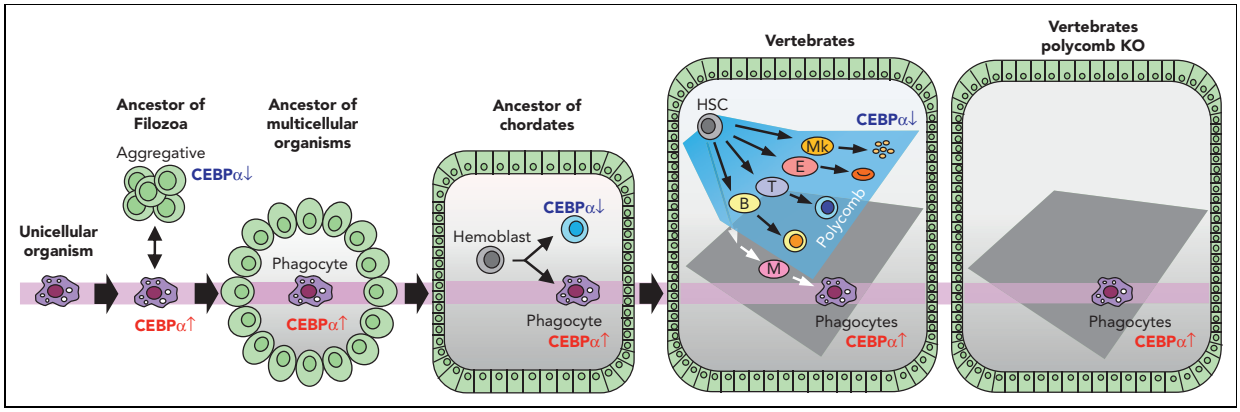


Figure 7. Schematic illustration for the evolution of blood cells. A component of the unicellular organism phenotype has been seamlessly inherited as phagocytes in multicellular animals. Vertebrates acquired various lineage blood cells by suppressing CEBPα using polycomb complexes. When polycomb function was impaired, hematopoiesis was reverted into a primitive one with phagocytes alone.

in the cavity that was able to patrol the cavity to eliminate pathogens and dead cells by phagocytosis. Thus, the multicellular organism should have survived after succeeding in holding such cells by inheriting the ancestral program for phagocytic characteristics driven by CEBPα, bringing about the birth of the initial blood cells (Figure 7).

Thereafter, megakaryocyte, erythroid, T-cell, and B-cell lineages were generated during the evolution of animals. An early study reported that the sea urchin has blood cells with clotting function,[62] so it is probable that the megakaryocyte lineage had been segregated at an earlier stage than echinoderms in the branch of Deuterostomia. An early branch of the megakaryocyte lineage in the hematopoietic differentiation pathway[63] should reflect its evolutionary early segregation. In chordates, at the level of protochordates, blood cells are segregated into several lineages,[5,64] and, in accordance with this finding, we showed that CEBPα is specifically expressed in the phagocytic blood cells. In the evolutionary history of vertebrates, before branching into jawless and jawed fish, the erythroid and lymphoid lineages should have arisen, because both jawless and jawed fish have these 2 cell types.[65-67] In vertebrate hematopoiesis, CEBPα is specifically expressed in phagocytes, and it is now clear, based on this study, that repression of CEBPα to maintain nonphagocytic lineages is commonly achieved by polycomb complexes in vertebrates (Figure 7). The findings that Ring1a/b KO leads to leukemogenesis in absence of Cdkn2a further suggest that Cdkn2a has been employed for secure hematopoiesis, so that dysfunction of the polycomb complex results in apoptosis (Figure 5J and supplemental Figure 15D).

In vertebrate hematopoiesis, phagocytic blood lineages and CEBPα has also been diverged. It is known that quadruplication of the genome took place in an ancestor of vertebrates after segregation from tunicates,[68,69] and vertebrates have quadruple CEBPα genes: CEBPα, CEBPβ, CEBPδ, and CEBPε. Such quadruplication of CEBPα has enabled vertebrates to acquire various phagocytic blood cells; for example, CEBPδ and CEBPε are important in granulocyte.[46,70,71] Homologs of other TFs essential to myeloid cells in vertebrates, such as PU.1 and IRF, were not found in *Capsaspora* (supplemental Figure 8B). It is probable that these genes have emerged after multicellular organisms had evolved from unicellular organisms and have enabled vertebrates to acquire another phagocytic blood cells, for example, dendritic cells.

We further argue whether findings in this study shows some implications regarding multicellularization in ancestral unicellular organisms. Phagocytosis itself is common among some unicellular eukaryotes,[72,73] but CEBP homologs has been found only in Filozoa.[60] Acquisition of CEBPα in ancestral Filozoan organisms, together with cis-regulatory system,[61] should have enabled them to regulate a phagocytic program. Lower expression of CEBPα homolog and higher expression of Ring1A/B homologs in the aggregative stage of *Capsaspora* than the filopodial stage (Figures 2F and 6D) suggested that polycomb complexes have played a role in repressing CEBPα and a phagocytic program in ancestral Filozoa. It is tempting to speculate that polycomb-mediated CEBPα repression has contributed to aggregation and multicellularization.

Of note was that hepatocytes, fibroblasts, and adipocytes, in which CEBPα is also known to be expressed, showed similarity to *Capsaspora*. Because these cells are known to have phagocytic potential, it is likely that these cells also inherited *Capsaspora* program driven by CEBPα. Further study is required to unveil whether such programs have been seamlessly maintained in the evolutionary history of these cells. Another unsolved issue is the evolutionary history of Protostomia blood cells. It remains to be clarified whether they have seamlessly inherited the CEBPα-driven program or have inherited an alternative program driven by different TFs.

Overall, this study has provided insight into the origin of blood cells in the animal kingdom, where the primary phagocytes in the ancestor of animals arose by activating the CEBPα-driven phagocytic program inherited from a unicellular organism, and has clarified the molecular mechanism by which the phagocytic program is suppressed to maintain nonphagocytic lineage cells in vertebrate hematopoiesis, that is, polycomb-mediated epigenetic suppression of CEBPα.

Acknowledgments

The authors thank Shimon Sakaguchi (Osaka University) for kindly providing *Rag2*$^{-/-}$ mice; Haruhiko Koseki (RIKEN) and Miguel Vidal (Centro de Investigaciones Biologicas) for kindly providing *Ert2Cre* mice and *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice; Jun-ichi Miyazaki (Osaka University) for kindly providing *CAG*$^{flox\text{-}stop\text{-}GFP}$ mice; Ellen V. Rothenberg (Caltech) and Hiroyuki Hosokawa (Tokai University) for kindly providing the pMXs-IRES-hNGFR vector; and Peter Burrows (University of Alabama at Birmingham) for critical reading of the manuscript.

This work was supported by funds from Japan Society for the Promotion of Science KAKENHI, Grant-in-Aid for Scientific Research (B) (JP15H04743), and Grant-in-Aid for Scientific Research on Innovative Areas (JP 19H05747). LiMe Office of Director's Research Grants 2022 (No. 6) also supported this work.

Authorship

Contribution: Y. Nagahata and H.K. conceived and designed the project; Y. Nagahata, K.M., T.I., Y. Nishimura, and S.K. designed and optimized experimental methodologies using mice, Y. Nagahata and Y.S. did so using tunicates, and Y. Nagahata and H.S. did so using *Capsaspora*; Y. Nagahata, H.S., and Y.S. performed experiments; Y. Nagahata, H.S., and Y.S. analyzed the data; T.K., Y. Nannya, S.O., and A.T.-K gave advice in performing the experiments; and Y. Nagahata., K.M., and H.K. wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: Y.N., 0000-0002-3157-4168; S.O., 0000-0002-7778-5374; H.S., 0000-0003-1795-2174; Y.S., 0000-0001-5193-0708; A.T.-K., 0000-0001-7678-4284; H.K., 0000-0002-7670-9900.

Correspondence: Hiroshi Kawamoto, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan; email: kawamoto@infront.kyoto-u.ac.jp.

Footnotes

Submitted 11 March 2022; accepted 26 August 2022; prepublished online on *Blood* First Edition 16 September 2022. https://doi.org/10.1182/blood.2022016286.

Expression levels of homologs were found in supplemental Tables 3 and 5. Other sources were also available with the online version of this article.

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

REFERENCES

1. Mukherjee S, Ray M, Ray S. Phagocytic efficiency and cytotoxic responses of Indian freshwater sponge (Eunapius carteri) cells isolated by density gradient centrifugation and flow cytometry: a morphofunctional analysis. *Zoology (Jena)*. 2015;118(1):8-18.
2. Kawamoto H, Ikawa T, Masuda K, Wada H, Katsura Y. A map for lineage restriction of progenitors during hematopoiesis: the essence of the myeloid-based model. *Immunol Rev*. 2010;238(1):23-36.
3. Cooper MD, Alder MN. The evolution of adaptive immune systems. *Cell*. 2006;124(4):815-822.
4. Boehm T. Evolution of vertebrate immunity. *Curr Biol*. 2012;22(17):R722-732.
5. Rosental B, Kowarsky M, Seita J, et al. Complex mammalian-like haematopoietic system found in a colonial chordate. *Nature*. 2018;564(7736):425-429.
6. Katsura Y, Kawamoto H. Stepwise lineage restriction of progenitors in lympho-myelopoiesis. *Int Rev Immunol*. 2001;20(1):1-20.
7. Kawamoto H, Ohmura K, Katsura Y. Direct evidence for the commitment of hematopoietic stem cells to T, B and myeloid lineages in murine fetal liver. *Int Immunol*. 1997;9(7):1011-1019.
8. Lu M, Kawamoto H, Katsube Y, Ikawa T, Katsura Y. The common myelolymphoid progenitor: a key intermediate stage in hemopoiesis generating T and B cells. *J Immunol*. 2002;169(7):3519-3525.
9. Masuda K, Kakugawa K, Nakayama T, Minato N, Katsura Y, Kawamoto H. T cell lineage determination precedes the initiation of TCR gene rearrangement. *J Immunol*. 2007;179(6):3699-3706.
10. Wada H, Masuda K, Satoh R, et al. Adult T-cell progenitors retain myeloid potential. *Nature*. 2008;452(7188):768-772.
11. Kawamoto H. A close developmental relationship between the lymphoid and myeloid lineages. *Trends Immunol*. 2006;27(4):169-175.
12. Li J, Barreda DR, Zhang YA, et al. B lymphocytes from early vertebrates have potent phagocytic and microbicidal abilities. *Nat Immunol*. 2006;7(10):1116-1124.
13. Nagasawa T, Nakayasu C, Rieger AM, Barreda DR, Somamoto T, Nakao M. Phagocytosis by thrombocytes is a conserved innate immune mechanism in lower vertebrates. *Front Immunol*. 2014;5:445.
14. Stokes EE, Firkin BG. Studies of the peripheral blood of the Port Jackson shark (Heterodontus portusjacksoni) with particular reference to the thrombocyte. *Br J Haematol*. 1971;20(4):427-435.
15. Sebe-Pedros A, Irimia M, Del Campo J, et al. Regulated aggregative multicellularity in a close unicellular relative of metazoa. *Elife*. 2013;2:e01287.
16. Consortium F, the RP, Clst, et al. A promoter-level mammalian expression atlas. *Nature*. 2014;507(7493):462-470.
17. Choi J, Baldwin TM, Wong M, et al. Haemopedia RNA-seq: a database of gene expression during haematopoiesis in mice and humans. *Nucleic Acids Res*. 2019;47(D1):D780-D785.
18. Imai KS, Hino K, Yagi K, Satoh N, Satou Y. Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. *Development*. 2004;131(16):4047-4058.
19. Satou Y, Kawashima T, Shoguchi E, Nakayama A, Satoh N. An integrated database of the ascidian, Ciona intestinalis: towards functional genomics. *Zoolog Sci*. 2005;22(8):837-843.
20. Suga H, Chen Z, de Mendoza A, et al. The Capsaspora genome reveals a complex unicellular prehistory of animals. *Nat Commun*. 2013;4:2325.
21. Sogabe S, Hatleberg WL, Kocot KM, et al. Pluripotency and the origin of animal multicellularity. *Nature*. 2019;570(7762):519-522.
22. de Mendoza A, Suga H, Permanyer J, Irimia M, Ruiz-Trillo I. Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *Elife*. 2015;4:e08904.
23. Fairclough SR, Chen Z, Kramer E, et al. Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate Salpingoeca rosetta. *Genome Biol*. 2013;14(2):R15.
24. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):238.
25. Ohmura K, Kawamoto H, Fujimoto S, Ozaki S, Nakao K, Katsura Y. Emergence of T, B, and myeloid lineage-committed as well as multipotent hemopoietic progenitors in the aorta-gonad-mesonephros region of day 10 fetuses of the mouse. *J Immunol*. 1999;163(9):4788-4795.
26. Masuda K, Kubagawa H, Ikawa T, et al. Prethymic T-cell development defined by the expression of paired immunoglobulin-like receptors. *EMBO J*. 2005;24(23):4052-4060.

27. Turner EC. Possible poriferan body fossils in early Neoproterozoic microbial reefs. *Nature*. 2021;596(7870):87-91.

28. Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*. 2011;334(6059):1091-1097.

29. Shalchian-Tabrizi K, Minge MA, Espelund M, et al. Multigene phylogeny of choanozoa and the origin of animals. *PLoS One*. 2008;3(5):e2098.

30. Torruella G, Derelle R, Paps J, et al. Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. *Mol Biol Evol*. 2012;29(2):531-544.

31. Torruella G, de Mendoza A, Grau-Bove X, et al. Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Curr Biol*. 2015;25(18):2404-2410.

32. Hartenstein V. Blood cells and blood cell development in the animal kingdom. *Annu Rev Cell Dev Biol*. 2006;22:677-712.

33. Soji T, Murata Y, Ohira A, Nishizono H, Tanaka M, Herbert DC. Evidence that hepatocytes can phagocytize exogenous substances. *Anat Rec*. 1992;233(4):543-546.

34. Villena JA, Cousin B, Penicaud L, Casteilla L. Adipose tissues display differential phagocytic and microbicidal activities depending on their localization. *Int J Obes Relat Metab Disord*. 2001;25(9):1275-1280.

35. Romana-Souza B, Chen L, Leonardo TR, Chen Z, DiPietro LA. Dermal fibroblast phagocytosis of apoptotic cells: a novel pathway for wound resolution. *FASEB J*. 2021;35(4):e21443.

36. Molkentin JD, Kalvakolanu DV, Markham BE. Transcription factor GATA-4 regulates cardiac muscle-specific expression of the alpha-myosin heavy-chain gene. *Mol Cell Biol*. 1994;14(7):4947-4957.

37. Milatovich A, Qiu RG, Grosschedl R, Francke U. Gene for a tissue-specific transcriptional activator (EBF or Olf-1), expressed in early B lymphocytes, adipocytes, and olfactory neurons, is located on human chromosome 5, band q34, and proximal mouse chromosome 11. *Mamm Genome*. 1994;5(4):211-215.

38. Liu X, Rowan SC, Liang J, et al. Categorization of lung mesenchymal cells in development and fibrosis. *iScience*. 2021;24(6):102551.

39. Nerlov C, Graf T. PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes Dev*. 1998;12(15):2403-2412.

40. Tsujimura H, Tamura T, Gongora C, et al. ICSBP/IRF-8 retrovirus transduction rescues dendritic cell development in vitro. *Blood*. 2003;101(3):961-969.

41. Shayman JA, Tesmer JJG. Lysosomal phospholipase A2. *Biochim Biophys Acta Mol Cell Biol Lipids*. 2019;1864(6):932-940.

42. Mota AC, Dominguez M, Weigert A, Snodgrass RG, Namgaladze D, Brune B. Lysosome-dependent LXR and PPARdelta activation upon efferocytosis in human macrophages. *Front Immunol*. 2021;12:637778.

43. Xie H, Ye M, Feng R, Graf T. Stepwise reprogramming of B cells into macrophages. *Cell*. 2004;117(5):663-676.

44. Laiosa CV, Stadtfeld M, Xie H, de Andres-Aguayo L, Graf T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. *Immunity*. 2006;25(5):731-744.

45. Collombet S, van Oevelen C, Sardina Ortega JL, et al. Logical modeling of lymphoid and myeloid cell specification and transdifferentiation. *Proc Natl Acad Sci U S A*. 2017;114(23):5792-5799.

46. Cirovic B, Schonheit J, Kowenz-Leutz E, et al. C/EBP-induced transdifferentiation reveals granulocyte-macrophage precursor-like plasticity of B cells. *Stem Cell Rep*. 2017;8(2):346-359.

47. Suh HC, Gooya J, Renn K, Friedman AD, Johnson PF, Keller JR. C/EBPalpha determines hematopoietic cell fate in multipotential progenitor cells by inhibiting erythroid differentiation and inducing myeloid differentiation. *Blood*. 2006;107(11):4308-4316.

48. Zhang DE, Zhang P, Wang ND, Hetherington CJ, Darlington GJ, Tenen DG. Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proc Natl Acad Sci U S A*. 1997;94(2):569-574.

49. Zhang P, Iwasaki-Arai J, Iwasaki H, et al. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity*. 2004;21(6):853-863.

50. Giladi A, Paul F, Herzog Y, et al. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat Cell Biol*. 2018;20(7):836-846.

51. Piunti A, Shilatifard A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science*. 2016;352(6290):aad9780.

52. Wang H, Wang L, Erdjument-Bromage H, et al. Role of histone H2A ubiquitination in Polycomb silencing. *Nature*. 2004;431(7010):873-878.

53. Ikawa T, Masuda K, Endo TA, et al. Conversion of T cells to B cells by inactivation of polycomb-mediated epigenetic suppression of the B-lineage program. *Genes Dev*. 2016;30(22):2475-2485.

54. Sebe-Pedros A, Degnan BM, Ruiz-Trillo I. The origin of Metazoa: a unicellular perspective. *Nat Rev Genet*. 2017;18(8):498-512.

55. Cavalier-Smith T, Chao EE. Phylogeny of choanozoa, apusozoa, and other protozoa and early eukaryote megaevolution. *J Mol Evol*. 2003;56(5):540-563.

56. Steenkamp ET, Wright J, Baldauf SL. The protistan origins of animals and fungi. *Mol Biol Evol*. 2006;23(1):93-106.

57. Ros-Rocher N, Perez-Posada A, Leger MM, Ruiz-Trillo I. The origin of animals: an ancestral reconstruction of the unicellular-to-multicellular transition. *Open Biol*. 2021;11(2):200359.

58. Stibbs HH, Owczarzak A, Bayne CJ, DeWan P. Schistosome sporocyst-killing Amoebae isolated from Biomphalaria glabrata. *J Invertebr Pathol*. 1979;33(2):159-170.

59. Owczarzak A, Stibbs HH, Bayne CJ. The destruction of Schistosoma mansoni mother sporocysts in vitro by amoebae isolated from Biomphalaria glabrata: an ultrastructural study. *J Invertebr Pathol*. 1980;35(1):26-33.

60. Sebe-Pedros A, de Mendoza A, Lang BF, Degnan BM, Ruiz-Trillo I. Unexpected repertoire of metazoan transcription factors in the unicellular holozoan Capsaspora owczarzaki. *Mol Biol Evol*. 2011;28(3):1241-1254.

61. Sebe-Pedros A, Ballare C, Parra-Acero H, et al. The dynamic regulatory genome of Capsaspora and the origin of animal multicellularity. *Cell*. 2016;165(5):1224-1237.

62. Kindred JE. Phagocytosis and clotting in the perivisceral fluid of Arbacia. *Biol Bull*. 1921;41(3):144-152.

63. Yamamoto R, Morita Y, Ooehara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*. 2013;154(5):1112-1126.

64. Parrinello D, Parisi M, Parrinello N, Cammarata M. Ciona robusta hemocyte populational dynamics and PO-dependent cytotoxic activity. *Dev Comp Immunol*. 2020;103:103519.

65. Hagerstrand H, Danieluk M, Bobrowska-Hagerstrand M, et al. The lamprey (Lampetra fluviatilis) erythrocyte; morphology, ultrastructure, major plasma membrane proteins and phospholipids, and cytoskeletal organization. *Mol Membr Biol*. 1999;16(2):195-204.

66. Pancer Z, Amemiya CT, Ehrhardt GR, Ceitlin J, Gartland GL, Cooper MD. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature*. 2004;430(6996):174-180.

67. Bajoghli B, Guo P, Aghaallaei N, et al. A thymus candidate in lampreys. *Nature*. 2011;470(7332):90-94.

68. Ohno S. *Evolution by Gene Duplication*. Springer; 1970.

69. Furlong RF, Holland PW. Were vertebrates octoploid? *Philos Trans R Soc Lond B Biol Sci*. 2002;357(1420):531-544.

70. Scott LM, Civin CI, Rorth P, Friedman AD. A novel temporal expression pattern of three C/EBP family members in differentiating myelomonocytic cells. *Blood*. 1992;80(7):1725-1735.

71. Yamanaka R, Barlow C, Lekstrom-Himes J, et al. Impaired granulopoiesis, myelodysplasia, and early lethality in CCAAT/enhancer binding protein epsilon-deficient mice. *Proc Natl Acad Sci U S A*. 1997;94(24):13187-13192.

72. Cavalier-Smith T. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int J Syst Evol Microbiol*. 2002;52(Pt 2):297-354.

73. Yutin N, Wolf MY, Wolf YI, Koonin EV. The origins of phagocytosis and eukaryogenesis. *Biol Direct*. 2009;4:9.

Supplemental information

Supplemental Methods

Materials & Correspondence

Correspondence and requests for materials should be addressed to Hiroshi Kawamoto (kawamoto@infront.kyoto-u.ac.jp)

Data and code availability

For ChIP-seq analyses, published data sets were used: GSE143682[74] for myeloid cells, GSE31233[75] for DN3 cells, GSE48534[76] for proB cells, and GSE36028 by ENCODE project[77] for ErPs and MkPs.

For comparison of gene expression profiles among different species, public data of mouse in EMBL-EBI (https://www.ebi.ac.uk/gxa/experiments?species=mus%20musculus) (supplemental Table 1) and data of mouse, tunicate, sponge, choanoflagellate, Capsaspora, and Creolimax in previous reports were analyzed[15-23]. RNA sequencing data of tunicate phagocytes and *Ring1a/b* KO myeloid cells were available at DNA Data Bank of Japan (DDBJ) database (DRA013007 and DRA014437).

Mice

6–15-week-old mice (*Mus musculus*) were used. For the transplantation experiments, 6–8-week-old mice were used. Both male and female mice were used, as we have not observed any differences associated with sex; mice were randomly assigned to the different groups in the experiments. C57BL/6J (B6) mice were purchased from Japan SLC, Inc. *Rag2*$^{-/-}$ mice were kindly provided by Shimon Sakaguchi (Osaka University). *Ert2Cre* mice and *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were kindly provided by Haruhiko

Koseki (RIKEN, Japan) and Miguel Vidal (Centro de Investigaciones Biologicas, Spain)[78]. *CAG*$^{flox\text{-}stop\text{-}GFP}$ mice were kindly provided by Jun-ichi Miyazaki (Osaka University, Japan)[79]. *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$-*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$, and *LckCre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were generated and maintained in our animal facility. All mice were maintained in SPF conditions in our animal facility. All experiments were performed in accordance with the guidelines of the Kyoto University Animal Experiment Committee and approved by our institutional committee.

Cross-species transcriptomic comparison

For the cross-species comparison, we first identified homologs in the proteomes of *M. musculus*, *C. intestinalis*, *A. queenslandica*, and *C. owczarzaki* using the OrthoFinder[24]. Amino acid sequences were obtained from the Ensembl database and previous reports[15,20,80]. Instead of default algorithm of Diamond, we used MMseqs2 for identifying homologs. CAOG_02294 and CAOG_03823 were manually added to the orthogroup of CEBPα and that of CEBPγ, respectively, based on the findings of this study (Figure 4; supplemental Figure 10). We distinguished CEBPα/β/δ/ε homologs and CEBPγ homologs based on phylogenetic analysis and their function of inducing phagocytes (Figure 4; supplemental Figure 10). CAOG_00371 was added to the orthogroup of Ring1A/B based on result from BLASTP search. ENSCING00000009562 was also added to the orthogroup of Ring1A/B and that of exosome component protein based on the ghost database. Orthogroups commonly conserved across the four species were selected and used for cross-species comparison (supplemental Table 3). If one orthogroup contained two or more homologs in one species, the total amount of TPM values of the homologs

was used for TPM value of the orthogroup, and the gene name was represented with the gene that had the highest TPM values. We then normalized TPM values only for the subset of conserved orthogroups, transformed them to $\log_2$ (TPM +1). Clustering was performed by the UPGMA method and Euclidean metric. The PC analyses of the expression data was performed using scikit-learn. *C. intestinalis* TPM values were calculated from RNA sequencing data (DRA 013007) and EST counts (http://ghost.zool.kyoto-u.ac.jp/datas/seqN.zip, http://ghost.zool.kyoto-u.ac.jp/datas/seqC.zip) mapped on the genome of Ensembl database (ftp.ensembl.org/pub/release-104/fasta/ciona_intestinalis/cdna/Ciona_intestinalis.KH.cdna.all.fa.gz).

Differentially expressed genes

In comparison of gene expression levels between Capsaspora, mouse phagocytes, and mouse non-phagocytes, mean values of each cell type were compared: mean values of the three stages (filopodial, aggregative, and cystic) for Capsaspora, mean values of three types of phagocytes (macrophages, monocytes, and neutrophils) for mouse phagocytes, and mean values of non-phagocytic blood cells and non-blood cells for mouse non-phagocytes. Genes which were >2 fold highly expressed in Capsaspora or phagocytes than non-phagocytes with $p < 0.05$ were selected. In comparison between the three stages of Capsaspora, genes which were >1 fold highly expressed in both filopodial and cystic stages than in aggregative stages with $p < 0.05$ were selected.

TFs and phagocytosis related genes

For selecting TFs and phagocytosis/lysosome related genes, we used the AmiGO2 database (http://amigo.geneontology.org/amigo). TFs candidates were searched with the words “transcription factor” in *M. musculus*, and 1712 genes were identified. Among

them 485 non-TF genes such as polycomb proteins were manually removed and the remaining 1227 genes were selected as TFs. For phagocytosis and lysosome related genes, candidates were searched with the word “phagocytosis” and “lysosome” in *M. musculus*, and 369 and 555 genes were identified, respectively. Among them, only genes which were >2 fold more highly expressed in phagocytes than in non-phagocytes with $p < 0.05$ were selected as phagocytosis-related genes. For cross-species comparison, 11 commonly conserved phagocytosis related genes and 62 conserved TFs across the four species were used (supplemental Table 6).

KEGG pathway analysis

KEGG pathway analyses were performed for selected genes and their mouse paralogs using KEGG Mapper (https://www.genome.jp/kegg/mapper/).

Phylogenetic analysis of CEBP and other bZIP homologs

We searched the UniProt (https://www.uniprot.org) database and the whole genome databases of *Nematostella vectensis*[81], *Trichoplax adhaerens*[82], *Amphimedon queenslandica*[83], *Monosiga brevicollis*[84], and *Capsaspora owczarzaki*[20] for CEBP and other closely-related gene families. The collected amino acid sequences were aligned manually. The 56 amino acids sites of the alignment, which correspond to the sequence from Met227 to Phe262 of mouse CEBPβ, were used for the phylogenetic analysis. Four CEBP genes of mouse and two CEBP-related genes of *Ciona intestinalis* were added to the alignment manually. We inferred the Maximum-likelihood (ML) tree using RAxML 7.3.4[85] with the LG+Γ substitution model, by choosing the topology with the best likelihood out of 30 independent calculations using the default hill-climbing algorithm. A bootstrap test with 100 replicates was performed by the rapid bootstrap algorithm

implemented in the program.

Amino acid sequences of CEBPα homologs were aligned and compared with MEGA11 software and ClustalW method[86].

Antibodies

The following antibodies were purchased from BioLegend: PE-CD4 (RM4-5; 100512), PECy7-Sca1 (D7; 108114), APC-CD3ε (145-2C11; 100312), APC-CD11b (M1/70; 101212), APC-CD19 (1D3; 152410), BV421-CD11b (M1/70; 101236), BV421-CD45 (30-F11; 103134), BV421-CD150 (TC15-12F12.2; 115925), biotin-CD4 (RM4-5; 100508), biotin-CD8α (53-6.7; 100704), biotin-CD11b (M1/70; 101204), biotin-CD11c (N418; 117304), biotin-TER119 (TER119; 116204).

The following antibodies were purchased from eBioscience: PE-CD25 (PC61.5; 12-0251-83), PE-CD41 (eBioMWReg30; 12-0411-81), PE-IgM (II/41; 12-5790-82), PE-TER119 (TER119; 12-5921-81), PE-F4/80 (BM8; 12-4801-80), eFluor660-CD34 (RAM34; 50-0341-80), APCeFluor780-ckit (ACK2; 47-1172-82), biotin-IL7R (A7R34; 13-1271-85), biotin-NK1.1 (PK136; 13-5941-85).

The following antibodies were purchased from BD Biosciences: FITC-CD3ε (1452C11; 553062), FITC-streptavidin (554060), PE-FcR (2.4G2; 553145), PECy7-CD19 (1D3; 552854), PECy7-B220 (RA3-6B2; 561881), APC-CD8α (53-6.7; 553035), APC-ckit (2B8; 553356), APC-Gr1 (RB6-8C5; 553129), APCCy7-streptavidin (554063), BV421-human CD271 (NGFR) (C40-1457; 562562), V450-strepavidin (560797), biotin-CD3ε (145-2C11; 553060), biotin-CD19 (1D3; 553784), biotin-FcR (2.4G2; 553143).

The following antibodies were purchased from TONBO Biosciences: PECy7-CD11c (N418; 60-0114-U100).

Growth factors

Recombinant murine SCF, Flt3-Ligand (L), IL-1α, IL3, IL-7, TNFα, M-CSF GM-CSF, TPO, and human EPO were purchased from Peprotech.

Isolation of progenitors

Single-cell suspensions of the thymus or femoral and tibial BM from *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice, *LckCre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$, or wild type mice were prepared. DN3 cells were isolated from thymocytes and proB cells, ErPs, and MkPs were isolated from BM cells. For ChIP experiments, ErPs and MkPs were isolated from E15.5 fetal livers. DN3 cells were sorted as Lin$^-$ckit$^-$CD25$^+$ cells after staining with FITC-streptavidin/biotin-Lin (CD3ε, CD4, CD8α, CD11b, CD11c, CD19, NK1.1, and TER119), PE-CD25, and APC-ckit antibodies. ProB cells were sorted as Lin$^-$IgM$^-$B220$^+$CD19$^+$ cells after staining with FITC-streptavidin/biotin-Lin (CD3ε, CD11b, CD11c, NK1.1, and TER119), PE-IgM, PECy7-B220 and APC-CD19 antibodies. ErPs and MkPs were sorted as Lin$^-$Sca1$^-$ckit$^+$CD150$^+$CD41$^-$ and Lin$^-$Sca1$^-$ckit$^+$CD150$^+$CD41$^+$ cells, respectively, after staining with FITC-streptavidin/biotin-Lin (CD3ε, CD11b, CD11c, NK1.1, TER119, IL-7R, and FcR), PE-CD41, PECy7-Sca1, APC-ckit, and BV421-CD150 antibodies. Macrophages were sorted as Lin$^-$CD11c$^-$CD11b$^+$Gr1^{low}SSC-A^{low}F4/80$^+$ cells, after staining with FITC-streptavidin/biotin-Lin (CD3ε, CD19, NK1.1, TER119), PE-F4/80, PECy7-CD11c, APC-Gr1, and BV421-CD11b antibodies. Gating strategies are shown in supplemental Figure 1.

CEBPα and Ring1B encoding vectors

Original CEBPα and Ring1B sequences were obtained from the NCBI database and

codon optimized genes were synthesized using GeneArt (Thermo Fisher Scientific) (supplemental Table 7). The pMXs-IRES-hNGFR vector was kindly gifted by Ellen V. Rothenberg (Caltech) and Hiroyuki Hosokawa (Tokai University, Japan)[87], and synthesized CEBPα homologs and Ring1B were inserted into the vector.

Retrovirus production and transduction

CEBPα and Ring1B encoding vectors were transfected into the Plat-E cells (CosmoBio) using FuGENE HD (Promega) and supernatants were harvested 2 and 3 days later. For transduction, purified progenitors were resuspended with the supernatant at 1 - 3 x 10^5 cells/ml for proB cells, DN3 cells, and macrophages, 2 x 10^4 cells/ml for ErPs and MkPs with 8 μg/ml of polybrene (Sigma) and 10 ng/ml each of IL-3, IL-7, Flt-3, SCF, and M-CSF. The cells were centrifuged for 90 min at 1,000 g at 32℃. Then, the supernatant was removed and the cells were resuspended in RPMI1640 supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 2 mg/mL sodium bicarbonate, 0.1 mM nonessential amino acid solution, 5×10^{-5} M 2-ME, 100 mg/mL streptomycin, 100 U/mL penicillin, and 10 ng/ml each of IL-3, IL-7, Flt-3, SCF, and M-CSF. For ErPs and MkPs 2 U/ml of EPO and 50 ng/ml of TPO were added instead of IL-7. The transduced progenitors were co-cultured with TSt4 for proB cells and ErPs, and with TSt4-DLL1 for DN3 cells and MkPs. For evaluation of phagocytic activity of MkP-derived cells induced by transduction of Capsaspora CEBPα (CAOG_02294), MkPs were cultured for 8 days, and 10 ng/ml each of IL-1α, IL-7, TNFα, and GM-CSF were added.

Wright-Giemsa stain

The isolated cells were attached to microscope slide glass by Cytospin 4 (Thermo Fisher

Scientific) for 5 min at 1500 rpm. The slides were stained and fixed with Wright solution for 5 min, washed with 1/150 M phosphate buffer twice, and stained with diluted Giemsa solution for 5 minutes.

Phagocytosis assay

To evaluate phagocytic activity, pHrodo-green zymosan beads (Invitrogen) were added to each culture and incubated. One hour later, medium containing beads and antibody were replaced with PBS and phagocytosis was observed using a BIOREVO BZ-9000 fluorescence microscope (KEYENCE). For evaluation of phagocytic activity of Capsaspora, pHrodo-green *S. aureus* beads (Invitrogen) and BIOREVO BZ-X810 fluorescence microscope (KEYENCE) were used.

RNA extraction and RT–qPCR

Total RNA was isolated using an RNeasy kit (Qiagen). cDNA synthesis was performed using a SuperScript IV VILO Master Mix cDNA synthesis kit (Invitrogen) following the manufacturer's protocol. Realtime PCR was performed using PowerUp SYBR Green Master Mix (Applied Biosystems) and analyzed by StepOnePlus (Applied Biosystems). The reactions were performed as follows: 10min at 95°C followed by 40 cycles of 15 sec at 95°C and 60 sec at 60°C. The primer sequences used are shown in supplemental Table 8. Primers sequences for M-CSFR and MPO can be found in a previous report[43].

Blood cells of tunicates

Blood of *C. intestinalis* was aspirated by cardiac puncture with a syringe in the presence of artificial sea water (Tomita Pharmaceutical). Blood cells were collected from pooled blood of several dozen tunicates by centrifugation (800 g, 5 minutes, 4℃). In order to distinguish phagocytes, blood cells were incubated with pHrodo-green zymosan beads in

184 the artificial sea water for 1 hour at room temperature. Isolation of each lineage of blood
185 cells was performed by FACS. Hemoblasts were identified as small size cells, and large
186 size cells were subdivided into 3 fractions: phagocytes as pHrodo-green positive cells,
187 granulocytes as autofluorescent positive cells detected by a violet laser (405 nm) and
188 V450 channel (450 nm), and other cells as pHrodo-green negative, auto-fluorescent
189 negative cells. Morphology of the sorted cells from each fraction were observed by
190 microscopy (BIOREVO BZ-9000, KEYENCE). RNA extraction and cDNA synthesis
191 were performed as with mouse blood cells, followed by qPCR with primers for *C.*
192 *intestinalis* (supplemental Table 8), designed according to the ghost database, NCBI
193 database and a previous report[88].

194 **ChIP-qPCR analysis**

195 ChIP assays were performed as previously described[89]. DN3 cells and proB cell were
196 isolated from thymus and BM of wild-type mice, respectively. ErPs and MkPs were
197 isolated from fetal livers. 5×10^4 - 5×10^6 cells were fixed with 1% formaldehyde and
198 lysed in cell lysis buffer (10 mM Tris pH 8.0, 10 mM, 0.2 % NP-40) on ice for 10 min.
199 After centrifugation, the cell lysis buffer was removed and nuclei lysis buffer (50 mM
200 Tris pH 8.0, 10 mM EDTA, 1 % SDS) was added and mixed by pipetting. An equivalent
201 volume of IP dilution buffer (20 mM Tris pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.01 %
202 SDS, 1 % Triton X-100) was added to the lysates and samples were sonicated. 2×10^4 –
203 2×10^6 cells were used for each immunoprecipitation. Anti-Ring1B antibody (Active
204 Motif, 39663) and control IgG (Thermo Fisher Scientific, 10400C) were used.
205 Immunocomplexes were bound to Dynabeads Protein G (Thermo Fisher Scientific,
206 DB10003) and washed six times. The immune complexes were then eluted by incubating

the beads at 65°C for 6 hours in elution buffer (0.1 M sodium bicarbonate, 1 % SDS, 0.3 M NaCl, and 10 μg/ml RNase). Eluted DNA was digested with proteinase K. DNA was purified using a DNeasy kit (Qiagen). Immunoprecipitated and input DNA were quantified by real-time PCR with the primers shown in supplemental Table 8.

ChIP-seq analysis

For ChIP-seq analysis of H3K27me3, published data sets were downloaded as FASTQ files: GSE143682 for myeloid cells, GSE31233 for DN3 cells, and GSE48534 for proB cells. Then, they were changed to BAM format and mapped as bamCoverage data on the UCSC genome browser (https://genome.ucsc.edu/) using Galaxy (https://usegalaxy.org/). For ErPs and MkPs, uploaded data in the UCSG genome browser (GSE36028) were used.

***In vitro* deletion of *Ring1b* in sorted progenitors for RT-qPCR**

The purified progenitors from *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were incubated on TSt-4 (for proB cells) [25] or on TSt-4/DLL1 (for DN3 cells)[26], or without feeder cells (for ErPs and MkPs) for 1-3 days in RPMI 1640 medium (GIBCO-BRL) supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 2 mg/mL sodium bicarbonate, 0.1 mM nonessential amino acid solution, 5×10^{-5} M 2-ME, 100 mg/mL streptomycin, 100 U/mL penicillin, 0.1μM 4-OHT, 10 ng/mL of SCF, Flt3-L IL-1α, IL-3, IL-7, TNFα, and GM-CSF. For ErPs, 10 ng/mL of M-CSF and 2 U/mL of EPO were added and for MkPs, 10 ng/mL of M-CSF and 50 ng/mL of TPO were added. To purify the cultured proB cells and DN3 cells after co-culture with feeder cells, harvested cells were stained with anti-CD45 antibody, and CD45^{+} cells were sorted by FACS Aria (BD). Cultured ErPs and MkPs were harvested without sorting. Collected cells were used for RT-qPCR analysis.

Bone marrow chimera mice

Hemolyzed whole bone marrow cells (2 × 10^6 cells) from *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$-*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were intravenously injected into the tail veins of sublethally irradiated (4 Gy) *Rag2*$^{-/-}$ mice. *Ert2Cre-CAG*$^{flox\text{-}stop\text{-}GFP}$-*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/+}$ mice were used as a control. Mice carrying the *CAG*$^{flox\text{-}stop\text{-}GFP}$ construct were used so that we could gate on cells in which the Cre-lox system had worked, based on the expression of GFP. Six weeks after bone marrow transplantation, *Ring1b* of blood cells was deleted by intraperitoneal injection with 75 μl of tamoxifen dissolved in corn oil at a concentration of 20 mg/ml for 3 times every other day. Mice were analyzed 2 weeks after the first injection of tamoxifen. For long term observation, 1 × 10^6 bone marrow cells were transplanted with 1 × 10^6 competitor cells from *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/+}$ mice or *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice, and mice were analyzed 8 weeks after tamoxifen injection.

For RNA-seq of Ring1*a/b* KO myeloid cells, CD19$^-$CD11b$^+$CD34$^+$c-kit$^+$ BM myeloid bast cells were sorted and expanded by culture in RPMI 1640 medium (GIBCO-BRL) supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 2 mg/mL sodium bicarbonate, 0.1 mM nonessential amino acid solution, 5 × 10^{-5} M 2-ME, 100 mg/mL streptomycin, 100 U/mL penicillin, 2 ng/mL of SCF, Flt3-L, IL-3, and M-CSF. CD19$^-$CD11b$^+$ BM myeloid cells were isolated from *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/+}$ mice or *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice as normal myeloid cells.

PCR analysis of *Igh* and *Tcrb* gene rearrangement

The analysis of *Igh* and *Tcrb* gene rearrangement was performed as described previously[90]. Genomic DNA was prepared from myeloid cells generated from T and B progenitors *in*

vitro using the directPCR (Viagen Biotech). The reaction volume was 20 μL, containing 8 μL of genomic DNA (approximately equivalent to 100 cells for *in vitro* experiments), 4 pmol of each primer, and 10 μL of 2× Quick Taq (TOYOBO). The PCR reactions were performed as follows with the primers in supplemental Table 9: 2 min at 94°C followed by 40 cycles of 30 sec at 94°C, 30 sec at 60°C, 1.5 min at 68°C, and finally 5 min at 68°C. Amplified DNA products were analyzed on an agarose gel mixed with RedSafe (iNtRON). For myeloid cells derived from BM chimera mice and control mice, 2× TaKaRa Taq HS Low DNA (Takara) was used and the PCR reactions were performed as follows: 2 min at 94°C followed by 40 cycles of 5 sec at 94°C, 1 sec at 60°C, 1 min at 68°C, and finally 5 min at 68°C. Annealing time was modified to 5 sec for *Igh* gene rearrangement.

In vitro conversion of progenitors into phagocytes

The isolated progenitors were co-cultured with TSt4 (for proB cells and ErPs) or TSt4-DLL1 (for DN3 cells and MkPs) cells for 4-12 days in RPMI1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 2 mg/mL sodium bicarbonate, 0.1 mM nonessential amino acid solution, 5×10^{-5} M 2-ME, 100 mg/mL streptomycin, 100 U/mL penicillin, 10 ng/mL of SCF, Flt3-L IL-1α, IL-3, IL-7, TNFα, and GM-CSF. For ErPs and MkPs, 2 U/mL of EPO and 50 ng/mL of TPO were added, respectively. 0.1 μM 4-OHT was added to delete *Ring1b* in *in vitro*.

Statistical analysis

Survival rates were estimated using Kaplan-Meier methods and compared using Log-rank tests. Continuous and categorical variables were compared using 2-tailed t tests and Fisher's exact test, respectively. Statistical analyses for survival rates were performed using EZR 1.40 (Jichi Medical University, Japan)[91]. Statistical analyses for correlation of

276 gene expression levels were performed using Anaconda 4.10.1 and Jupyter Notebook
277 6.3.0. Bar-dot graphs were drawn using bioinfokit[92].
278

Supplemental References

74. Cabal-Hierro L, van Galen P, Prado MA, et al. Chromatin accessibility promotes hematopoietic and leukemia stem cell activity. *Nat Commun*. 2020;11(1):1406.

75. Zhang JA, Mortazavi A, Williams BA, Wold BJ, Rothenberg EV. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell*. 2012;149(2):467-482.

76. Lane AA, Chapuy B, Lin CY, et al. Triplication of a 21q22 region contributes to B cell transformation through HMGN1 overexpression and loss of histone H3 Lys27 trimethylation. *Nat Genet*. 2014;46(6):618-623.

77. Yue F, Cheng Y, Breschi A, et al. A comparative encyclopedia of DNA elements in the mouse genome. *Nature*. 2014;515(7527):355-364.

78. Roman-Trufero M, Mendez-Gomez HR, Perez C, et al. Maintenance of undifferentiated state and self-renewal of embryonic neural stem cells by Polycomb protein Ring1B. *Stem Cells*. 2009;27(7):1559-1570.

79. Kawamoto S, Niwa H, Tashiro F, et al. A novel reporter mouse strain that expresses enhanced green fluorescent protein upon Cre-mediated recombination. *FEBS Lett*. 2000;470(3):263-268.

80. Sebe-Pedros A, Chomsky E, Pang K, et al. Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nat Ecol Evol*. 2018;2(7):1176-1188.

81. Putnam NH, Srivastava M, Hellsten U, et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science*. 2007;317(5834):86-94.

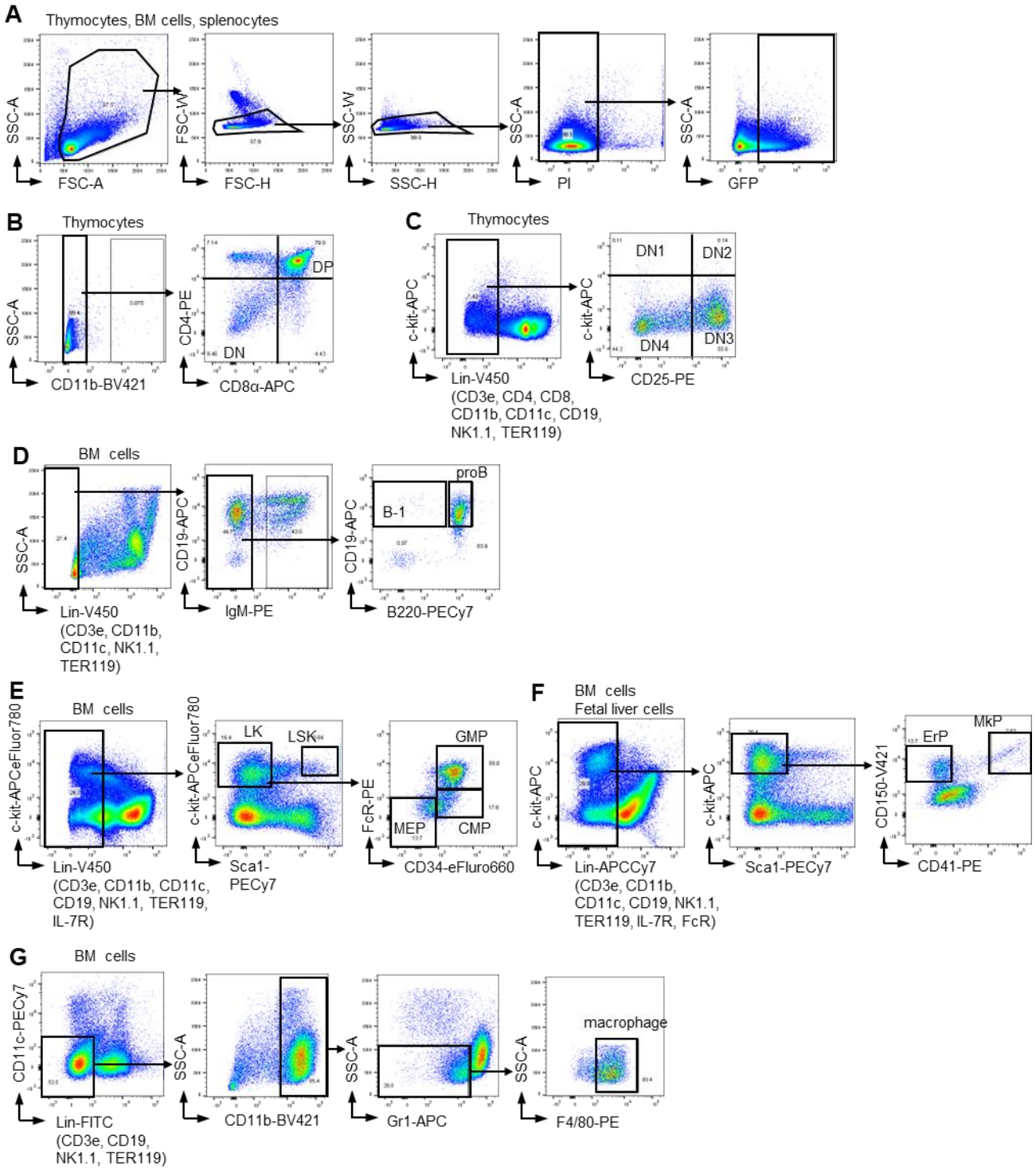
82. Srivastava M, Begovic E, Chapman J, et al. The Trichoplax genome and the

302 nature of placozoans. *Nature*. 2008;454(7207):955-960.
303 83. Srivastava M, Simakov O, Chapman J, et al. The Amphimedon queenslandica
304 genome and the evolution of animal complexity. *Nature*. 2010;466(7307):720-726.
305 84. King N, Westbrook MJ, Young SL, et al. The genome of the choanoflagellate
306 Monosiga brevicollis and the origin of metazoans. *Nature*. 2008;451(7180):783-788.
307 85. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic
308 analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22(21):2688-
309 2690.
310 86. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics
311 Analysis Version 11. *Mol Biol Evol*. 2021;38(7):3022-3027.
312 87. Hosokawa H, Ungerback J, Wang X, et al. Transcription Factor PU.1 Represses
313 and Activates Gene Expression in Early T Cells by Redirecting Partner Transcription
314 Factor Binding. *Immunity*. 2018;48(6):1119-1134 e1117.
315 88. Hamada M, Goricki S, Byerly MS, Satoh N, Jeffery WR. Evolution of the
316 chordate regeneration blastema: Differential gene expression and conserved role of notch
317 signaling during siphon regeneration in the ascidian Ciona. *Dev Biol*. 2015;405(2):304-
318 315.
319 89. Ikawa T, Kawamoto H, Goldrath AW, Murre C. E proteins and Notch signaling
320 cooperate to promote T cell lineage specification and commitment. *J Exp Med*.
321 2006;203(5):1329-1342.
322 90. Ikawa T, Kawamoto H, Wright LY, Murre C. Long-term cultured E2A-deficient
323 hematopoietic progenitor cells are pluripotent. *Immunity*. 2004;20(3):349-360.
324 91. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for

medical statistics. *Bone Marrow Transplant*. 2013;48(3):452-458.

92. Bedre R. reneshbedre/bioinfokit: Bioinformatics data analysis and visualization toolkit. Zenodo.; 2020.

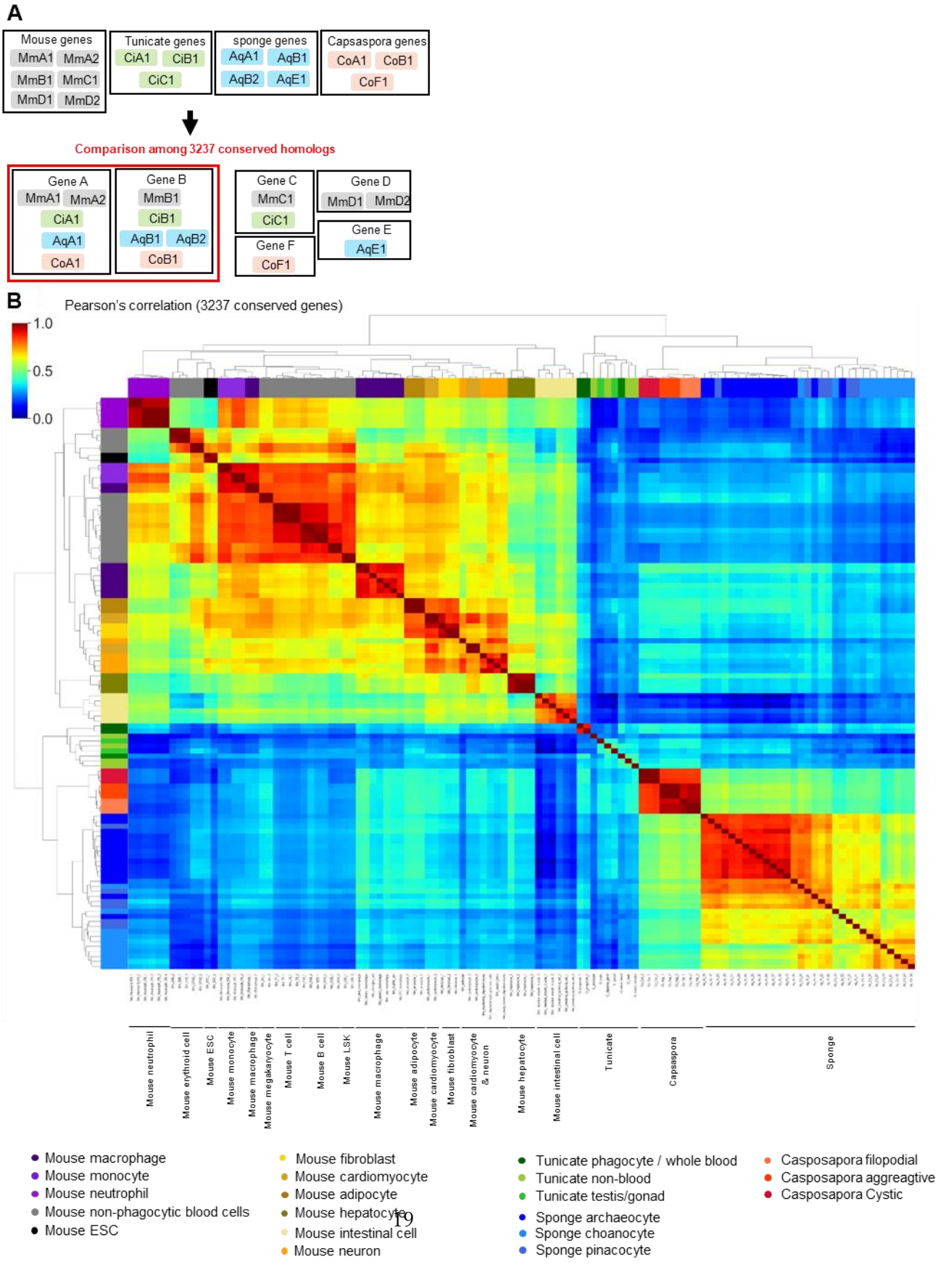
Supplemental Figure 1



Supplemental Figure 1. Gating strategy for thymocytes and BM cells.

(A) Live singlets were gated, followed by the further strategy below to identify various cell types. When BM chimera mice were examined, GFP^+ cells were gated and analyzed. (B-F) Further gating strategy for DN and DP cells (B), for DN1–4 cells (C), for proB cells and B-1 cells (D), for LSK cells, LK cells, CMPs, GMPs, and MEPs (E), for ErPs and MkPs(F), and for macrophages (G).

Supplemental Figure 2



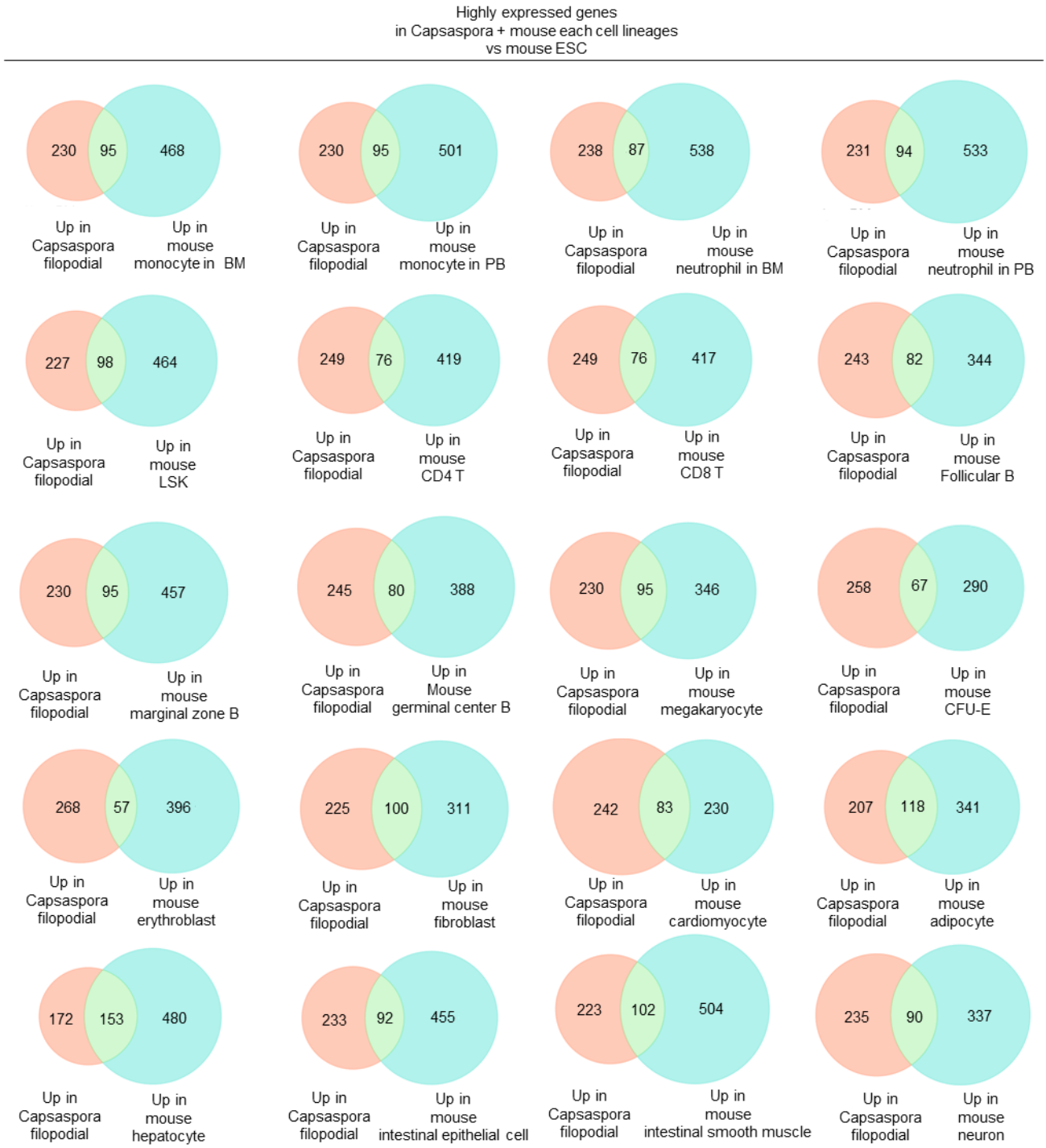
338

Supplemental Figure 2. Cross-species comparison of gene expression profiles.

(A) Conserved genes among mouse, tunicate, sponge, and Capsaspora were searched with OrthoFinder[24]. In brief, all protein coding genes of the four species were classified into homolog groups, and expression levels of 3237 homologs conserved among the four species were compared.

(B) Clustermap with Pearson correlation of Capsaspora, sponge, tunicate and mouse. Analysis based on 3237 conserved homologs is shown.

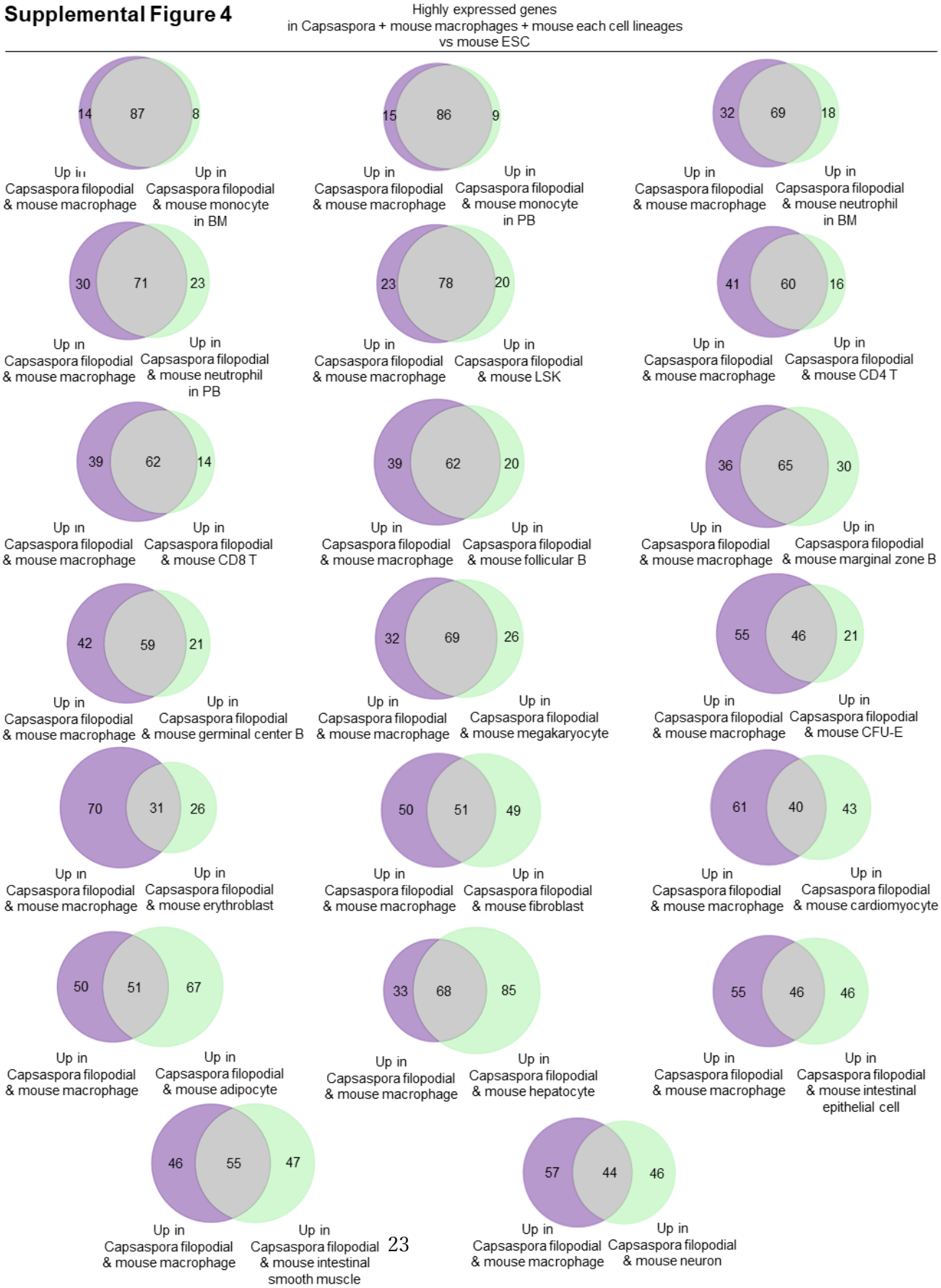
Supplemental Figure 3



Supplemental Figure 3. Number of highly expressed genes in Capsaspora and mouse various cell lineages.

Venn diagrams with the numbers of highly expressed genes in Capsaspora filopodial stage or mouse various cell lineages compared with mouse ESCs. BM, bone marrow; PB, peripheral blood.

Supplemental Figure 4



355

Supplemental Figure 4. Number of highly expressed genes in Capsaspora, mouse macrophages, and mouse various cell lineages.

Venn diagrams with the numbers of highly expressed genes in Capsaspora filopodial stage and mouse macrophages or Capsaspora filopodial stage and mouse various cell lineages compared with mouse ESCs.

Supplemental Figure 5

A KEGG pathway highly expressed in Capsaspora and mouse macrophage compared with mouse ESC

mmu01100 Metabolic pathways (46)
mmu04120 Ubiquitin mediated proteolysis (13)
mmu00982 Drug metabolism - cytochrome P450 (13)
mmu04142 Lysosome (11)
mmu00430 Taurine and hypotaurine metabolism (9)
mmu04140 Autophagy - animal (8)
mmu04024 cAMP signaling pathway (7)
mmu05010 Alzheimer disease (7)
mmu04925 Aldosterone synthesis and secretion (7)
mmu04022 cGMP-PKG signaling pathway (6)
mmu05171 Coronavirus disease - COVID-19 (6)
mmu04141 Protein processing in endoplasmic reticulum (6)
mmu04972 Pancreatic secretion (6)
mmu05417 Lipid and atherosclerosis (6)
mmu04810 Regulation of actin cytoskeleton (6)
mmu04144 Endocytosis (6)
mmu00340 Histidine metabolism (6)
mmu00600 Sphingolipid metabolism (6)

B KEGG pathway highly expressed in Capsaspora and mouse hepatocyte compared with mouse ESC

mmu01100 Metabolic pathways (80)
mmu00982 Drug metabolism (13)
mmu00280 Valine, leucine and isoleucine degradation (12)
mmu00410 beta-Alanine metabolism (12)
mmu00380 Tryptophan metabolism (11)
mmu04142 Lysosome (11)
mmu04714 Thermogenesis (10)
mmu00430 Taurine and hypotaurine metabolism (10)
mmu01200 Carbon metabolism (10)
mmu00270 Cysteine and methionine metabolism (9)
mmu00340 Histidine metabolism (9)
mmu05022 Pathways of neurodegeneration - multiple diseases (9)
mmu05020 Prion disease (8)
mmu00260 Glycine, serine and threonine metabolism (8)
mmu00360 Phenylalanine metabolism (8)
mmu04928 Parathyroid hormone synthesis, secretion and action (8)
mmu00350 Tyrosine metabolism (8)
mmu04911 Insulin secretion (8)
mmu04925 Aldosterone synthesis and secretion (8)
mmu04140 Autophagy - animal (8)

C KEGG pathway highly expressed in Capsaspora, mouse macrophage, and mouse hepatocyte compared with mouse ESC

mmu01100 Metabolic pathways (37)
mmu00982 Drug metabolism - cytochrome P450 (13)
mmu04142 Lysosome (10)
mmu00430 Taurine and hypotaurine metabolism (9)
mmu04140 Autophagy - animal (8)
mmu05010 Alzheimer disease (6)
mmu00600 Sphingolipid metabolism (6)
mmu00340 Histidine metabolism (6)
mmu00350 Tyrosine metabolism (5)
mmu04071 Sphingolipid signaling pathway (5)
mmu04621 NOD-like receptor signaling pathway (5)
mmu01240 Biosynthesis of cofactors (5)
mmu00010 Glycolysis / Gluconeogenesis (5)
mmu00410 beta-Alanine metabolism (5)
mmu04130 SNARE interactions in vesicular transport (5)
mmu00980 Metabolism of xenobiotics by cytochrome P450 (4)
mmu05417 Lipid and atherosclerosis (4)
mmu00360 Phenylalanine metabolism (4)
mmu05415 Diabetic cardiomyopathy (4)
mmu05152 Tuberculosis (4)
mmu04145 Phagosome (4)
mmu04136 Autophagy (4)

Supplemental Figure 5. KEGG pathway analysis.

KEGG pathways involved in genes highly expressed in Capsaspora filopodial stage and mouse macrophages (A), Capsaspora filopodial stage and mouse hepatocytes (B), and Capsaspora filopodial stage, mouse macrophages, and mouse hepatocytes (C).

Supplemental Figure 6

A Eleven genes upregulated in Capsaspora and mouse macrophage

	Log2(fold change)																			
Gene Name	Co Filo /ESC	Co_Filo /LSK	Co_Filo /T4	Co_Filo /T8	Co_Filo /FoB	Co_Filo /MZB	Co_Filo /GCB	Co_Filo /Mk	Co_Filo /CFUE	Co_Filo /ErBl	Mac /ESC	Mac /LSK	Mac /T4	Mac /T8	Mac /FoB	Mac /MZB	Mac /GCB	Mac /Mk	Mac /CFUE	Mac /ErBl
Ctsl	6.1	7.2	9.0	8.4	9.7	10.0	11.5	7.1	7.4	10.8	2.3	3.4	5.2	4.5	5.8	6.1	7.6	3.2	3.5	6.9
Pla2g15	**1.8**	**1.8**	**3.5**	**3.5**	**2.6**	**2.5**	**2.9**	**1.7**	**3.5**	**4.2**	**2.1**	**2.2**	**3.8**	**3.9**	**2.9**	**2.8**	**3.3**	**2.1**	**3.8**	**4.5**
Cebpa	**4.8**	**3.5**	**7.5**	**7.4**	**7.1**	**6.2**	**7.2**	**3.3**	**5.7**	**6.5**	**4.5**	**3.2**	**7.2**	**7.0**	**6.8**	**5.8**	**6.8**	**2.9**	**5.4**	**6.2**
Dpep1	3.0	5.1	5.7	5.1	5.5	5.3	5.6	2.8	5.4	5.5	7.2	9.4	9.9	9.3	9.8	9.6	9.9	7.0	9.7	9.7
Ctsb	7.0	6.1	6.1	6.6	6.2	5.4	6.7	6.6	5.9	3.2	5.5	4.6	4.6	5.0	4.7	3.9	5.2	5.1	4.4	1.6
Mfsd7a	3.2	1.8	4.5	4.5	4.7	4.4	4.6	1.1	4.1	4.5	5.9	4.5	7.2	7.2	7.4	7.1	7.3	3.8	6.8	7.2
Abhd3	3.1	4.7	5.4	5.5	5.1	5.4	5.5	5.0	5.5	5.5	2.4	4.0	4.7	4.8	4.4	4.7	4.8	4.3	4.7	4.8
Agmo	8.1	6.3	8.1	7.9	7.7	7.9	8.0	7.7	7.9	7.8	5.7	4.0	5.7	5.6	5.4	5.6	5.6	5.4	5.5	5.4
Hal	2.7	4.2	3.4	3.3	4.4	3.8	4.4	4.3	4.4	4.4	6.8	8.3	7.5	7.4	8.5	7.9	8.5	8.4	8.5	8.5
Gas7	2.2	4.2	4.5	2.9	4.1	4.2	4.3	3.3	4.3	4.5	5.0	7.1	7.3	5.7	6.9	7.0	7.1	6.1	7.2	7.3
Hgd	6.1	6.6	6.2	6.5	6.5	5.9	6.6	6.5	6.3	6.5	5.3	5.9	5.4	5.8	5.7	5.1	5.8	5.8	5.5	5.8

Gene Name	Up in fibroblast vs ESC	Up in cardiomyocyte vs ESC	Up in adipocyte vs ESC	Up in hepatocyte vs ESC	Up in intestinal epithelial cell vs ESC	Up in intestinal smooth muscle vs ESC	Up in neuron vs ESC
Ctsl	○		○	○			
Pla2g15				○			
Cebpa	○		○	○			
Dpep1	○		○		○	○	
Ctsb	○	○	○	○	○	○	○
Mfsd7a				○		○	
Abhd3				○			○
Agmo		○	○	○	○	○	○
Hal				○			
Gas7	○	○	○		○	○	○
Hgd				○			

B

Orthogroup	Ensembl Gene ID	Gene Name	Full Gene Name	Function
OG0000348	ENSMUSG00000021477	Ctsl	cathepsin L	peptidase C1 (papain) family of cysteine proteases
	ENSMUSG00000042243	BC051665	cDNA sequence BC051665	collagen binding activity; fibronectin binding activity; and serpin family protein binding activity
	ENSMUSG00000056728	Ctsll3	cathepsin L-like 3	collagen binding activity; fibronectin binding activity; and serpin family protein binding activity
OG0000986	ENSMUSG00000031903	Pla2g15	phospholipase A2, group XV	acylglycerol O-acyltransferase activity; calcium-independent phospholipase A2 activity; and phospholipase A1 activity
	ENSMUSG00000035237	Lcat	lecithin cholesterol acyltransferase	phosphatidylcholine-sterol O-acyltransferase activity and platelet-activating factor acetyltransferase activity
OG0001223a	ENSMUSG00000034957	Cebpa	CCAAT/enhancer binding protein (C/EBP), alpha	transcription factor
	ENSMUSG00000052435	Cebpe	CCAAT/enhancer binding protein (C/EBP), epsilon	transcription factor
	ENSMUSG00000056501	Cebpb	CCAAT/enhancer binding protein (C/EBP), beta	transcription factor
	ENSMUSG00000071637	Cebpd	CCAAT/enhancer binding protein (C/EBP), delta	transcription factor
OG0001444	ENSMUSG00000019278	Dpep1	dipeptidase 1	beta-lactamase activity and dipeptidase activity
	ENSMUSG00000031898	Dpep3	dipeptidase 3	dipeptidase activity
	ENSMUSG00000053687	Dpep2	dipeptidase 2	dipeptidase activity
	ENSMUSG00000115067	Dpep2	dipeptidase 2	dipeptidase activity
	ENSMUSG00000115768	Dpep2nb	NA	NA
OG0001871	ENSMUSG00000021939	Ctsb	cathepsin B	peptidase C1 family
OG0001920	ENSMUSG00000029490	Mfsd7a	major facilitator superfamily domain containing 7A	transmembrane transporter activity
OG0002586	ENSMUSG00000002475	Abhd3	abhydrolase domain containing 3	carboxylic ester hydrolase activity
OG0002726	ENSMUSG00000050103	Agmo	alkylglycerol monooxygenase	glyceryl-ether monooxygenase activity and iron ion binding activity
OG0003556	ENSMUSG00000020017	Hal	histidine ammonia lyase	histidase protein family
OG0003780	ENSMUSG00000033066	Gas7	growth arrest specific 7	actin filament binding activity
OG0005031	ENSMUSG00000022821	Hgd	homogentisate 1, 2-dioxygenase	homogentisate 1,2-dioxygenase activity and identical protein binding activity

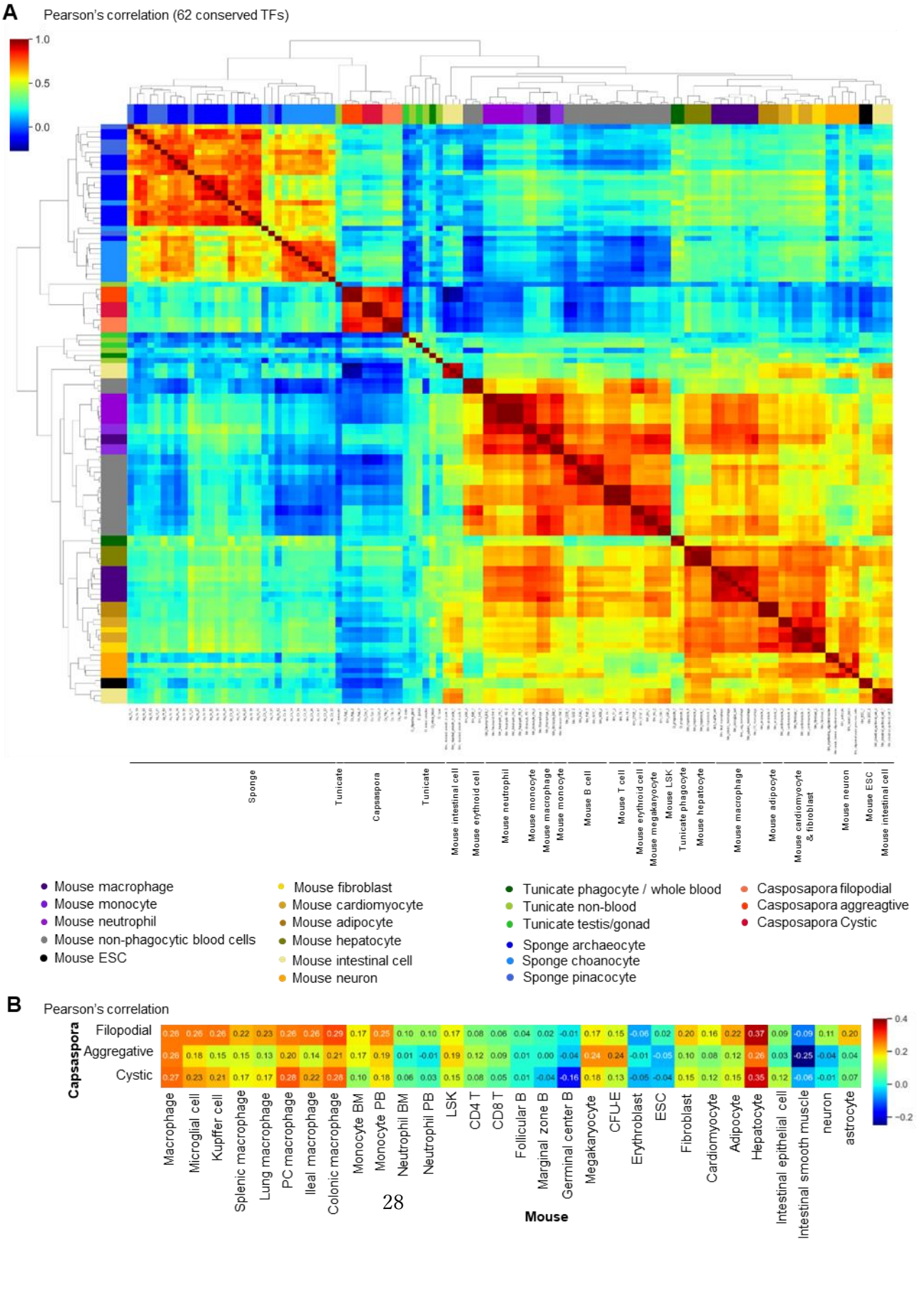
368

Supplemental Figure 6. Genes highly expressed in Capsaspora and mouse macrophages.

(A) List of the 11 genes highly expressed in both Capsaspora filopodial stage and mouse macrophages compared with mouse ESCs and non-phagocytic blood cells. Log2 fold-change values were shown in upper table, and genes shared by mouse other cell lineages were indicated with circle in lower table. Genes highly expressed in filopodial and cystic stages of Capsaspora are shown in bold font and the *Cebpa* in red.

(B) List of the 11 genes and their mouse paralogs with predicted function in NCBI database (https://www.ncbi.nlm.nih.gov/gene).

Supplemental Figure 7



379

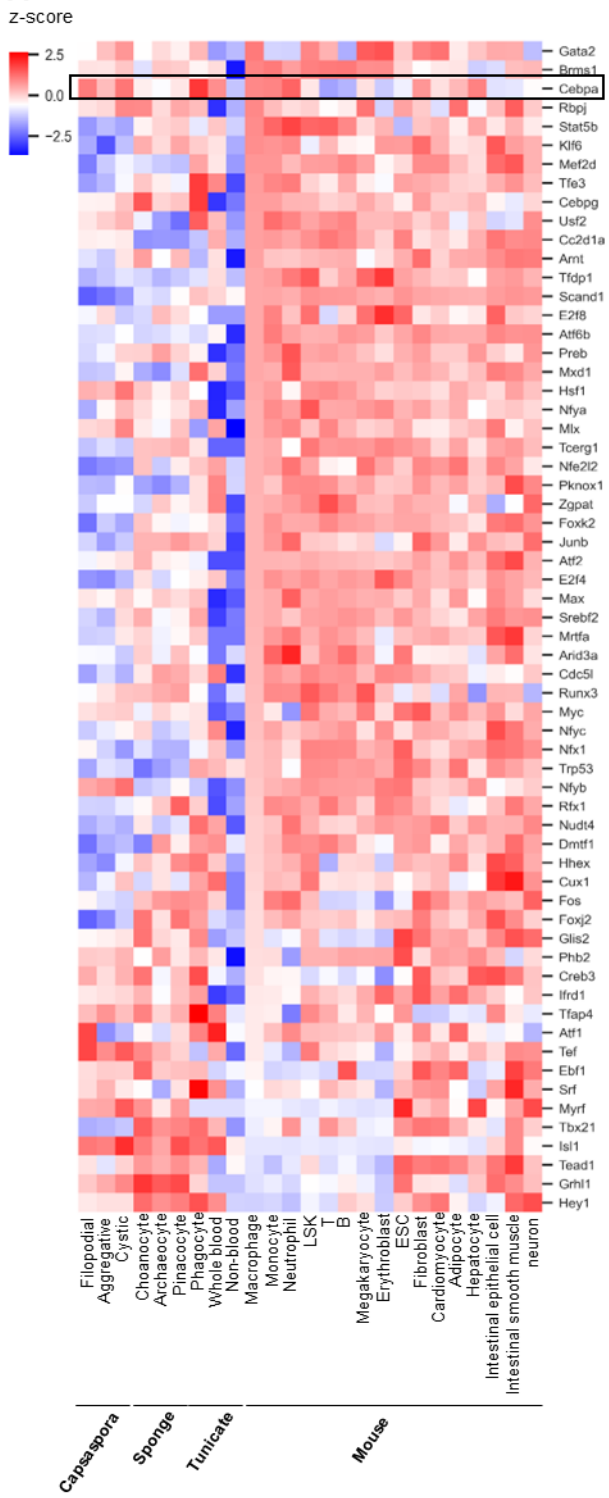
Supplemental Figure 7. Cross-species comparison of gene expression profiles based on TFs.

(A) Clustermap with Pearson correlation of Capsaspora, sponge, tunicate and mouse. Analysis based on 62 conserved TFs is shown.

(B) Heatmap with Pearson correlation of various mouse cell lineages and Capsaspora. Gene expression profiles were compared among 3 stages of Capsaspora and 30 mouse lineages based on 62 conserved TFs.

Supplemental Figure 8

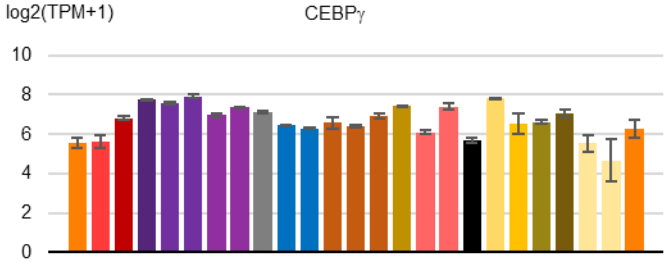
A



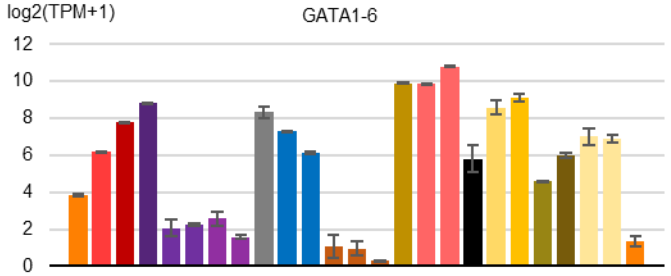
B

	mouse	tunicate	sponge	Capsaspora
CEBPα, β, δ, ε	4	1	1	1
CEBPγ	1	1	1	1
SPI1(PU.1), SPIB, SPIC	3	1	0	0
IRF1,2	2	0	0	0
IRF3-8	6	1	1	0
GATA1-6	6	2	1	1
TCF7	4	1	1	0
BCL11a/b	2	1	0	0
IKZF1-4	4	1	0	0
PAX2,5,6,8	4	1	1	0
EBF1-4	4	1	1	1

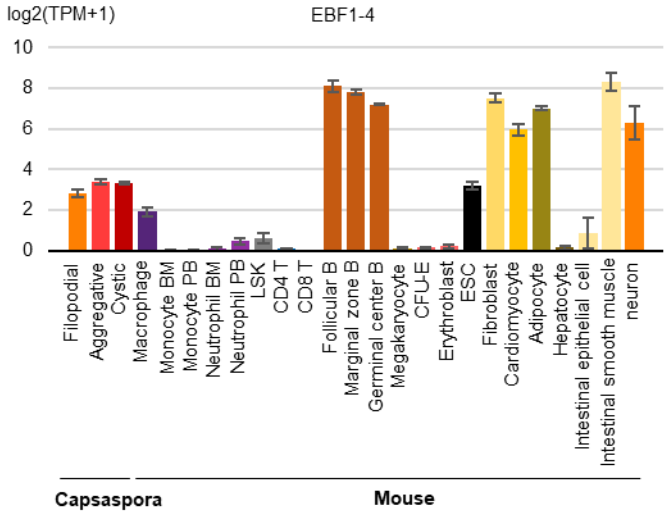
C



D



E



Supplemental Figure 8. Expression levels of conserved TFs.

(A) Heatmap of scaled expression levels (z-score) of 62 conserved TFs in Capsaspora, mouse, tunicate, and sponge cells. Expression levels were scaled with the 25 stages or lineages of four species.

(B) Conservation of TFs among four species. Eleven TFs were selected.

(C-E) Expression levels of CEBPγ (C), GATA1-6 (D), and EBF1-4 (E) homologs in Capsaspora and mouse various cell linages.

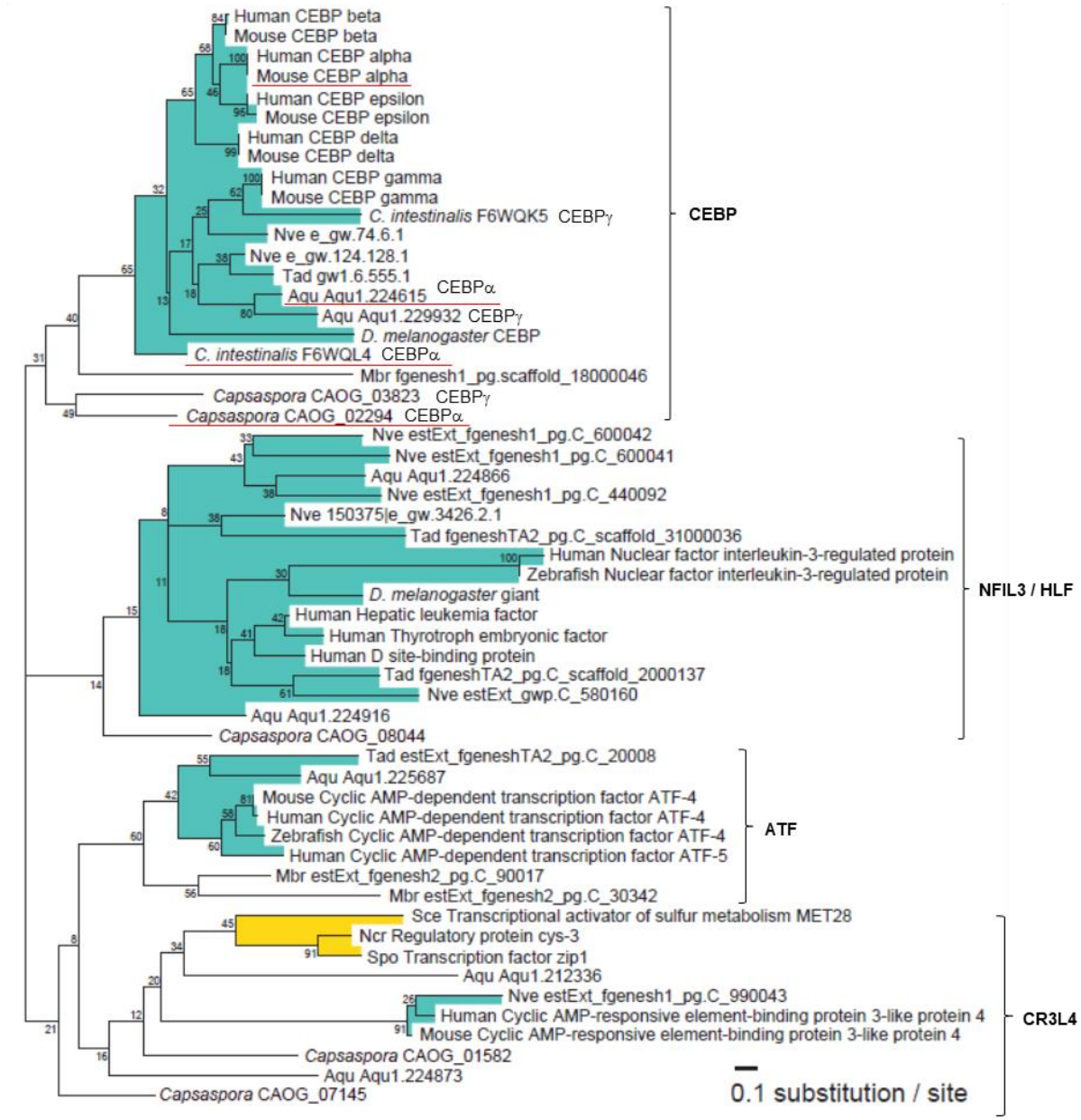
Supplemental Figure 9

398

Supplemental Figure 9. Multiple alignment of amino acid sequences of CEBP homologs.

Multiple alignment of amino acid sequences of CEBPα homologs. Transcription activation domain (TAD) and bZIP domain were indicated.

Supplemental Figure 10



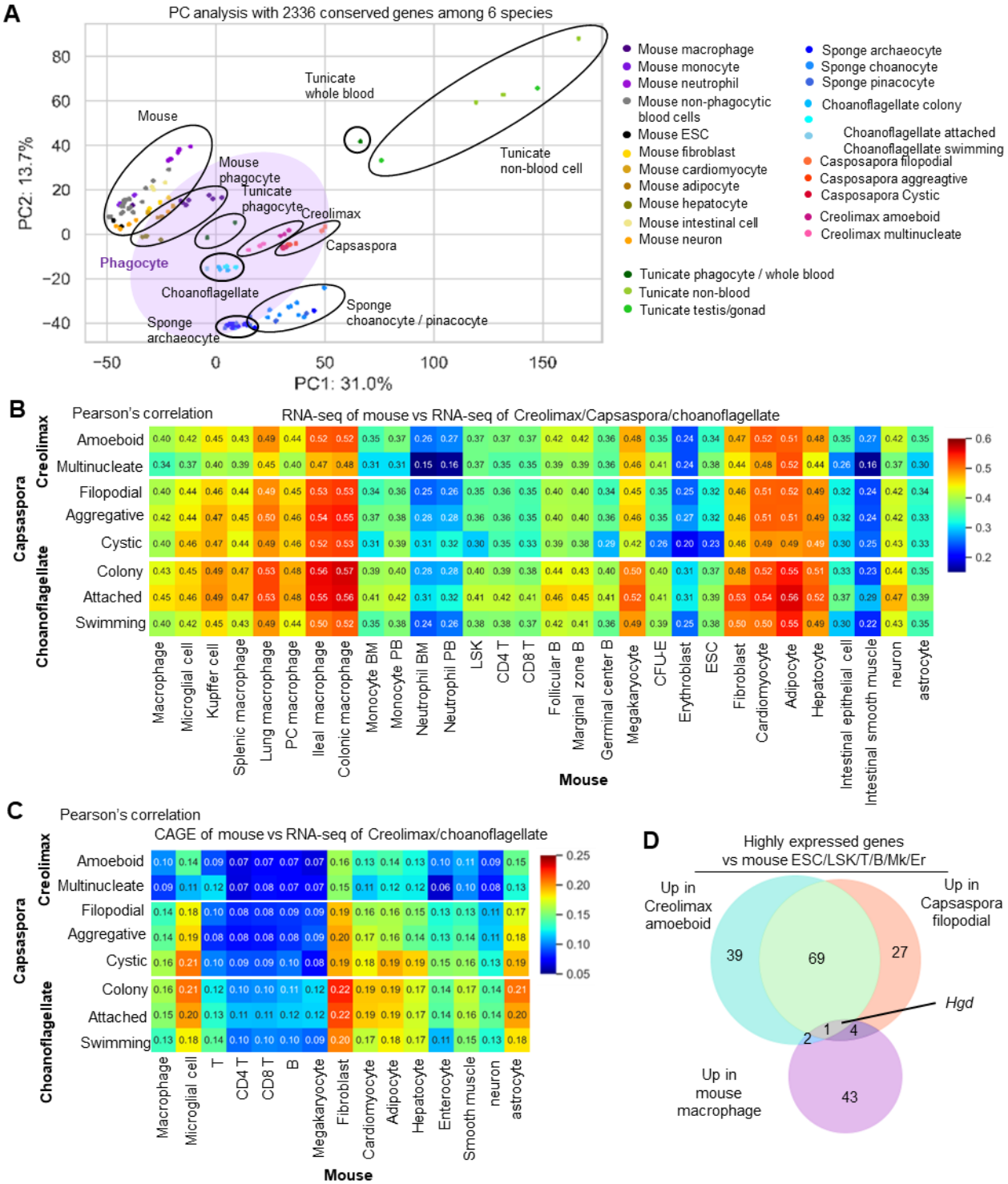
404

405

Supplemental Figure 10. CEBP homologs of various species.

Maximum likelihood tree of CEBP family and other related families inferred from bZIP domain sequences. Numbers at the branches represent the bootstrap values. Gene cluster of animals and fungi are shaded blue and yellow, respectively. Nve, *Nematostella vectensis*; Tad, *Trichoplax adhaerens*; Aqu, *Amphimedon queenslandica*; Mbr, *Monosiga brevicollis*; CEBP, CCAAT enhancer binding protein; NFIL3, Nuclear factor, interleukin 3 regulated; HLF, Hepatic leukemia factor; ATF, activating transcription factor; CR3L4, CAMP responsive element binding protein 3 like 4.

Supplemental Figure 11



415

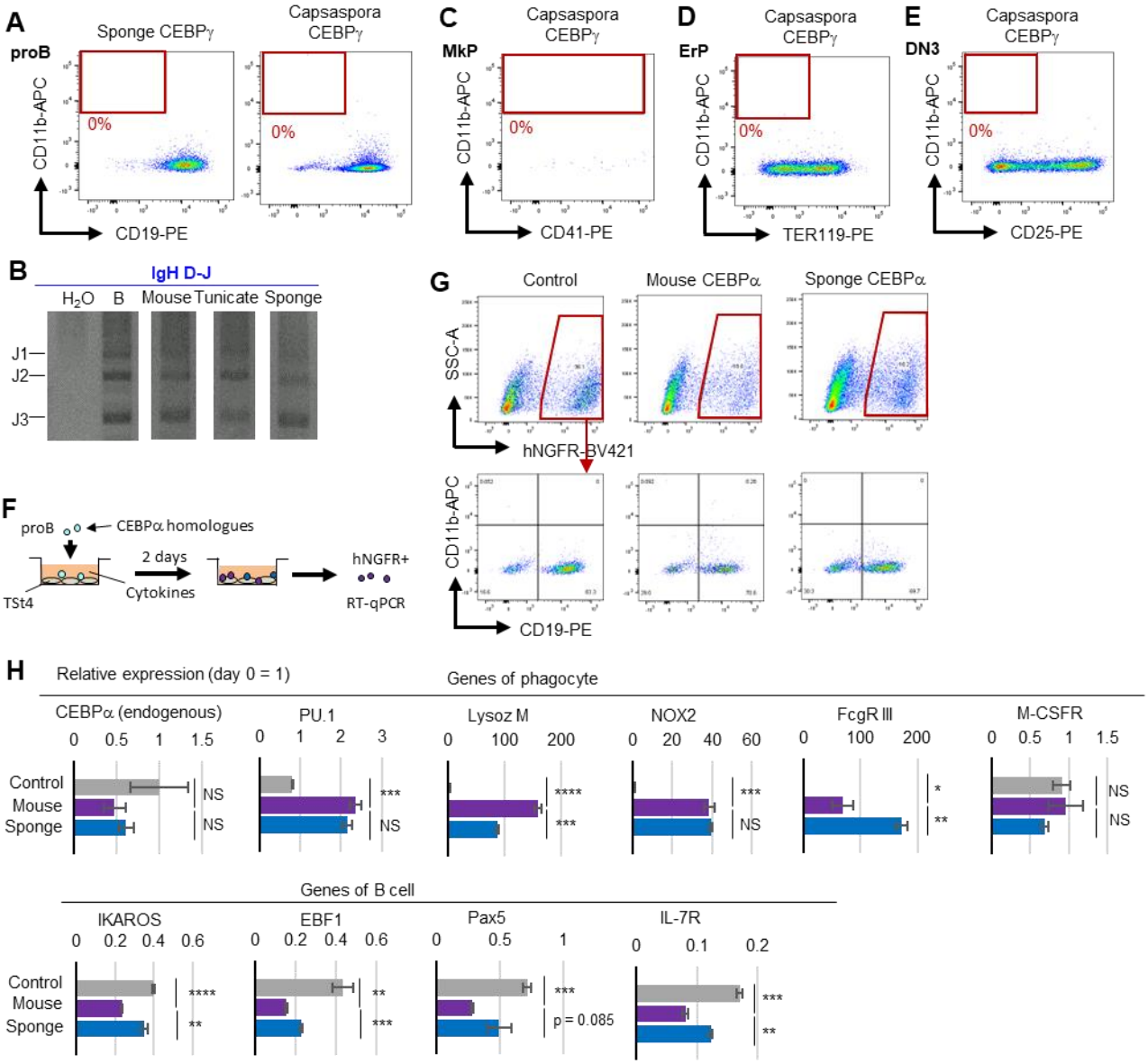
416

Supplemental Figure 11. Cross species analysis with 6 species.

(A-C) Gene expression profiles were compared based on 2336 homologs conserved among mouse, tunicate, sponge, choanoflagellate, Capsaspora, and Creolimax. PC analysis (A) and Heatmap with Pearson correlation of various mouse cell lineages, choanoflagellate, Capsaspora and Creolimax (B-C) were shown. Transcriptome data examined by RNA-seq (B) or CAGE method (C) were analyzed.

(D) Venn diagrams with the numbers of highly expressed genes in Creolimax amoeboid stage, Capsaspora filopodial stage, or mouse macrophages compared with mouse ESCs and non-phagocytic blood cells.

Supplemental Figure 12



Supplemental Figure 12. Comparison of function of CEBPα homologs.

(A) Flow cytometric profiles of cultured proB cells overexpressing CEBPγ of sponge and Capsaspora. Data are representative of 3 independent experiments.

(B) The generated $CD11b^+$ cells transduced with various CEBPα homologs were purified and the rearrangement status of their IgH genes was examined by PCR.

(C-E) MkPs (C), ErPs (D), and DN3 cells (E) were transduced with Capsaspora CEBPγ. Data are representative of 2–3 independent experiments.

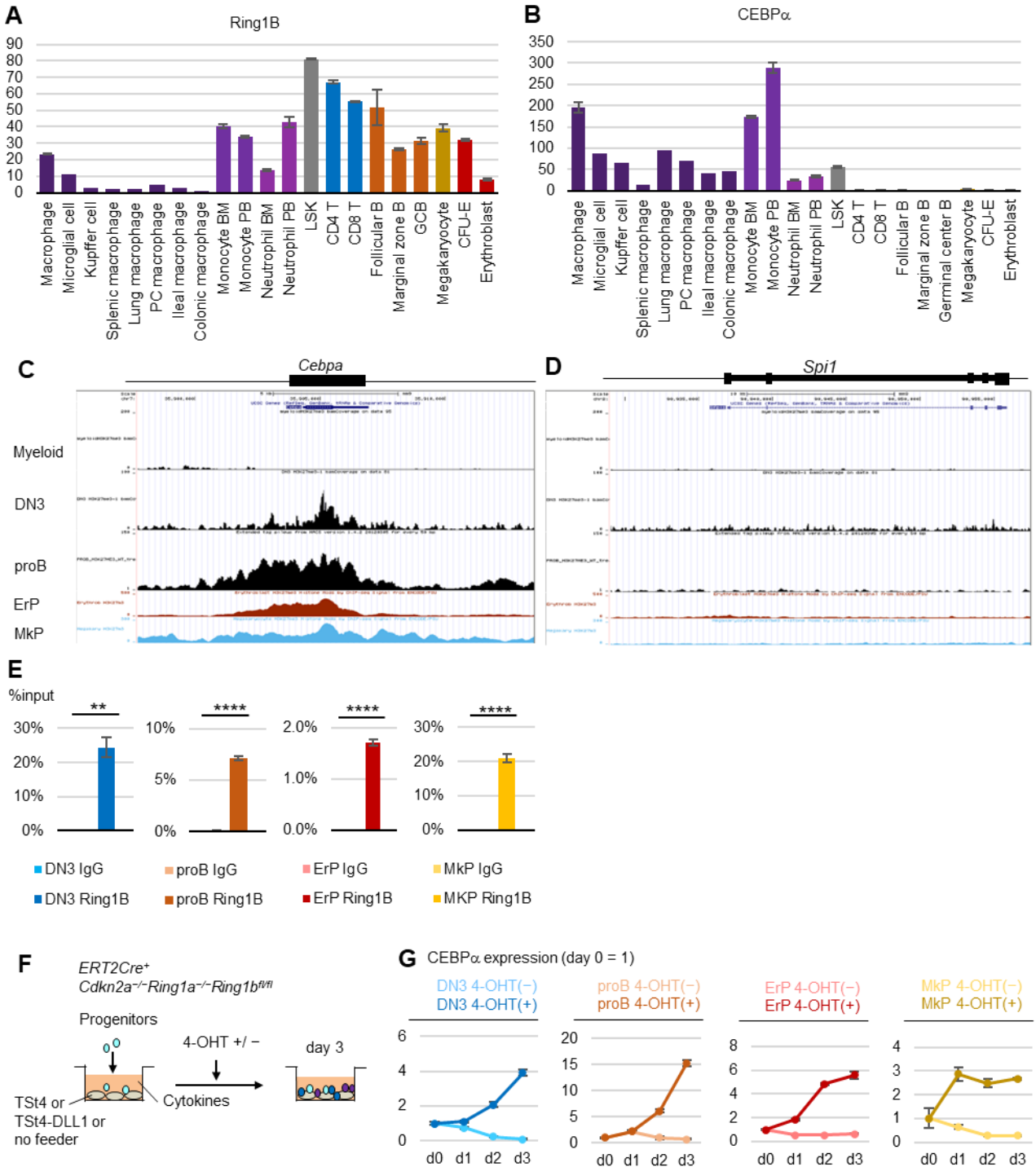
(F) Mouse CEBPα or its sponge homolog were transduced into proB cells. After 2 days, $hNGFR^+$ cells were analyzed by RT-qPCR.

(G) Flow cytometric profiles of proB cells transduced with CEBPα homologs after 2 days in culture are shown. $hNGFR^+$ cells were sorted and their gene expression patterns were examined by RT-qPCR. Data are representative of 3 independent experiments.

(H) Relative expression of lineage genes in proB cells 2 days after CEBPα transduction. Data are mean ± SEM. Relative expression levels (day 0 =1) with $2^{-\Delta\Delta CT}$ values normalized with β-actin were shown.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

Supplemental Figure 13



Supplemental Figure 13. Polycomb suppresses CEBPα in various lineage hematopoietic progenitors.

(A-B) Expression levels of Ring1B (A) and CEBPα (B) in various mouse hematopoietic lineages. TPM values were obtained from EMBL-EBI database and a previous report[17].

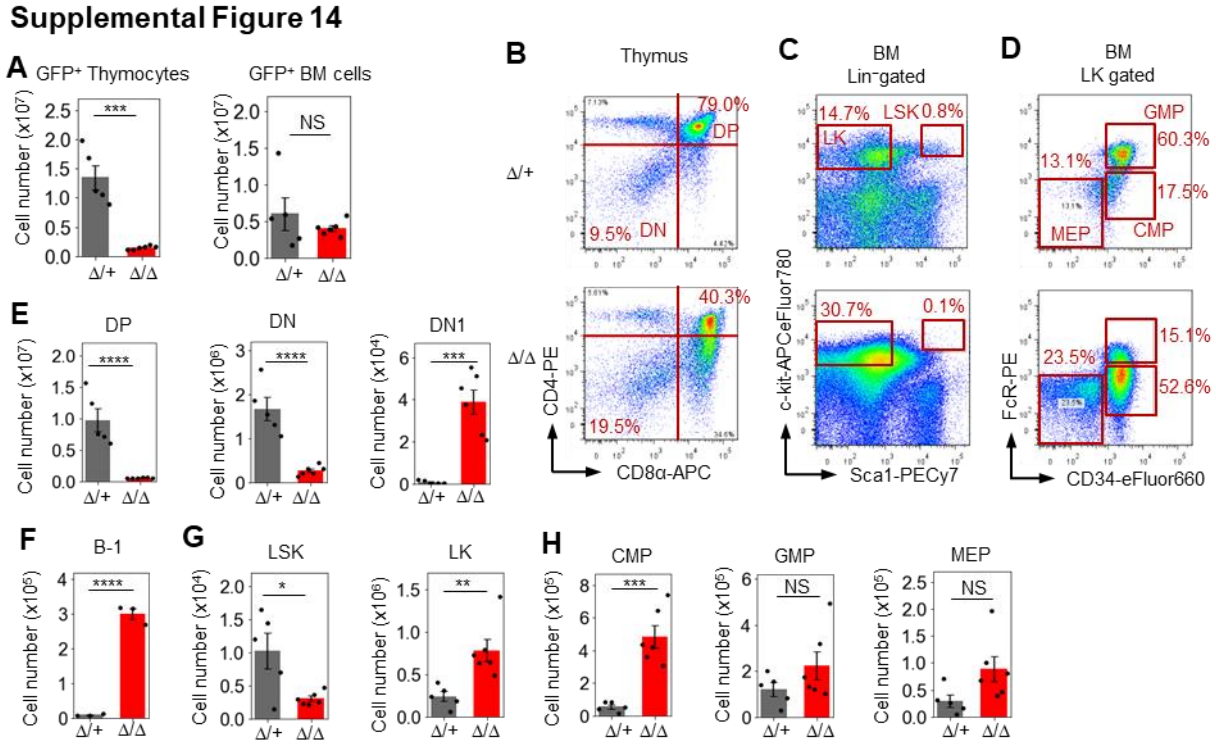
(C-D) H3K27me3 at *Cebpa* (C) or *spi1* (D) loci in each indicated lineage. FASTQ data were obtained from published data sets and mapped using the UCSC genome browser.

(E) Ring1B binding at *Cebpa* genes in DN3 cells, proB cells, ErPs, and MkPs was evaluated by ChIP-qPCR. IgG was used as a negative control. Data are representative of 3 independent experiments and are mean ± SEM.

(F) DN3 cells, proB cells, ErPs, and MkPs were isolated from *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice and cultured for 3 days in the presence or absence of 4-OHT. DN3 cells and proB cells were co-cultured with TSt4-DLL1 and TSt4, respectively. ErPs and MkPs were cultured without feeder cells.

(G) Expression levels of CEBPα (*Cebpa*) in each type of progenitor cell cultured in the presence or absence of 4-OHT were evaluated. Gene expression data were calculated by ΔΔCt methods and β-Actin (*Actb*) was used as internal control. Expression levels were normalized to gene expression at day 0. Data are representative of 3 independent experiments and are mean ± SEM.

** $p<0.01$, **** $p<0.0001$



Supplemental Figure 14. Polycomb mediated suppression of CEBPα is required for maintenance of various hematopoietic lineages.

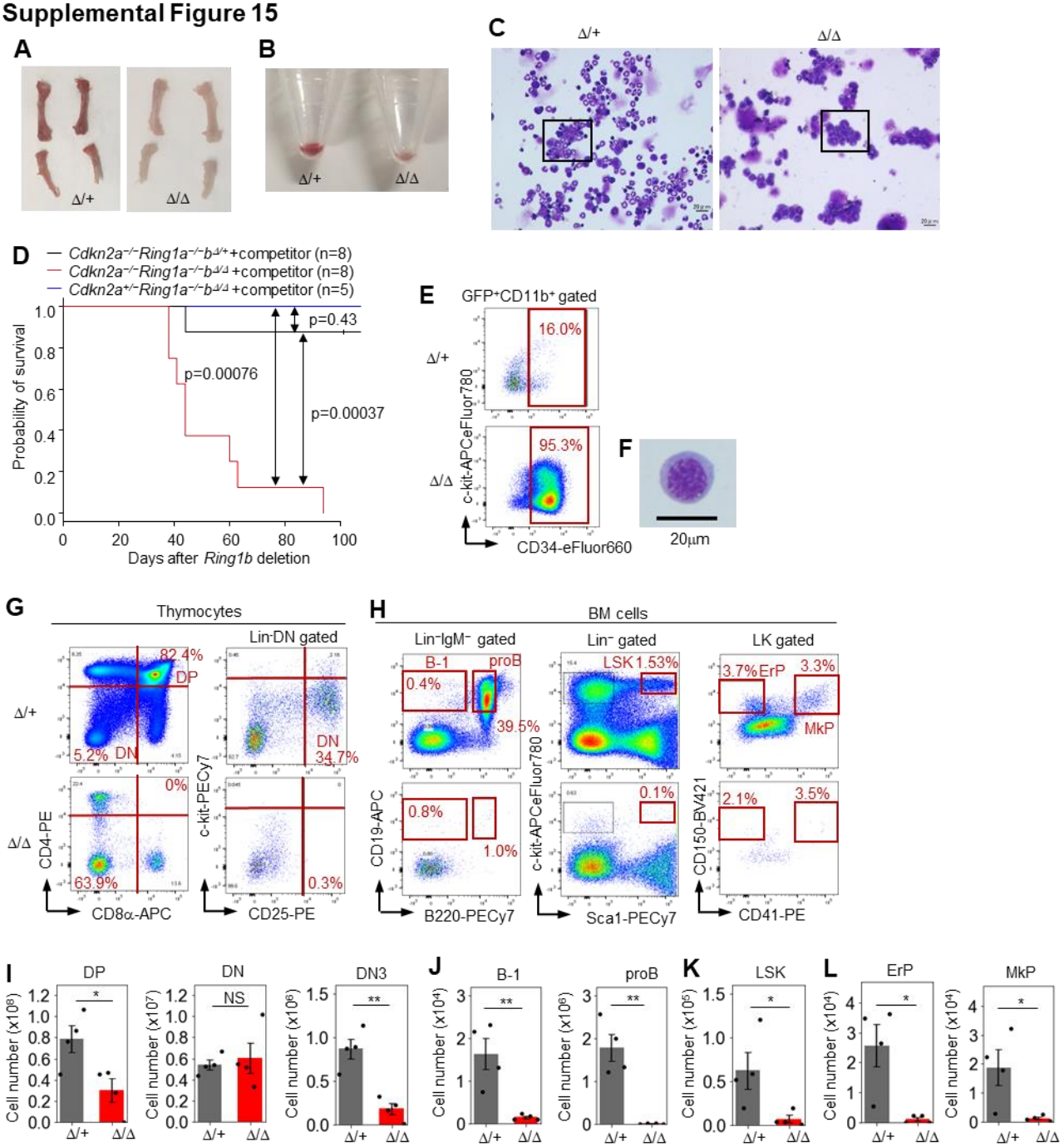
(A) The number of GFP$^+$ thymocytes and GFP$^+$ BM cells in the BM chimera mice. n = 5 in control mice and n = 6 in KO mice.

(B-D) Flow cytometric profiles of GFP$^+$ thymocytes (B) and GFP$^+$ BM cells (C, D). Upper panels show data of control mice (Δ/+, n = 5) and lower panels show data of *Ring1a/b* KO mice (Δ/Δ, n = 6)

(E-H) Number of GFP$^+$ DP cell, DN cells, and DN1 cells (E), B-1 progenitors (F), LSK cells and LK cells (G), and CMPs, GMPs, and MEPs (H). The black bars and red bars indicate data of control and *Ring1a/b* KO mice, respectively. Data are mean ± SEM.

* p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

Supplemental Figure 15



Supplemental Figure 15. *Ring1a/b* KO BM was occupied with phagocytes.

(A-B) Femoral and tibial bones (A) and pellets of BM cells (B) obtained from control (Δ/+, *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/+}$, n = 4) and *Ring1a/b* KO (Δ/Δ, *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/+}$, n = 4) mice with competitor cells.

(C) Wright-Giemsa stain of BM smears obtained from control and *Ring1a/b* KO mice. The boxed areas are shown at higher magnification in Figure 5L.

(D) Survival curve with Kaplan-Meier plots of BM transplantation into sublethally irradiated *Rag2*$^{-/-}$ mice with competitor cells. Black and red lines show survival curve of control (*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/+}$, n = 8), *Ring1a/b* KO in *Cdkn2a*$^{-/-}$ background (*Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/\Delta}$, n = 8) mice, and *Ring1a/b* KO in *Cdkn2a*$^{+/-}$ background (*Cdkn2a*$^{+/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/\Delta}$, n = 5) mice, respectively. Statistical significance of differences between the survival rates was calculated with the Log-rank test.

(E) Flow cytometric profiles of whole GFP^{+}CD11b^{+} BM cells of *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/+}$ (n = 3) and *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/\Delta}$ (n = 3) mice with competitor cells.

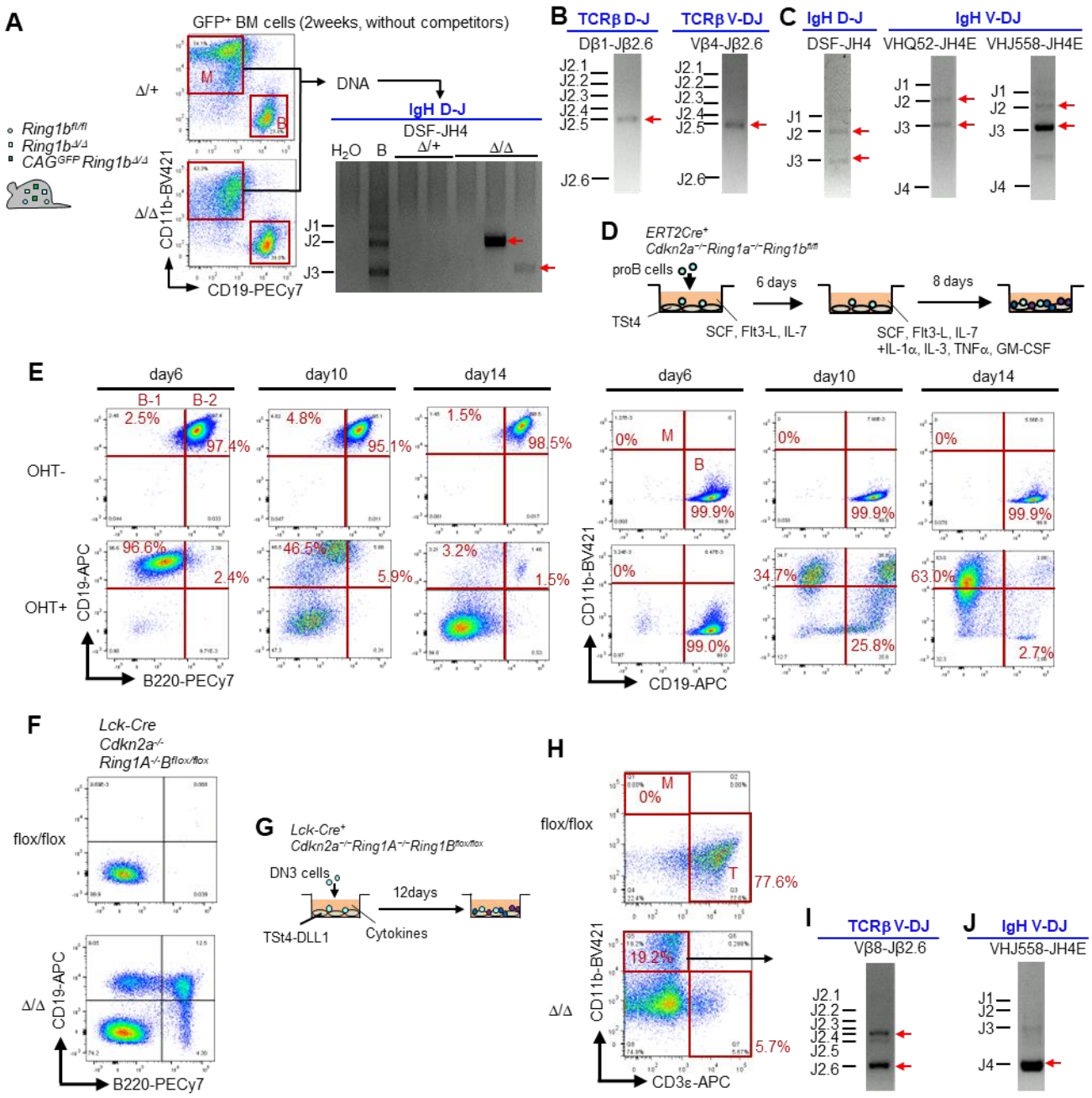
(F) GFP^{+}CD11b^{+}CD34^{+} cells from *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/\Delta}$ mice were sorted and their cytology and was evaluated by Wright-Giemsa staining.

(G-H) Flow cytometric profiles of whole thymocytes (G) and whole BM cells (H) of *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/+}$ (n = 4) and *Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{\Delta/\Delta}$ (n = 4) mice.

(I-L) Number of DP, DN, and DN3 cells (I), B-1 progenitors and proB cells (J), LSK cells (K), and ErPs and MkPs (L), in BM chimera mice with competitor cells. Data are mean ± SEM.

* $p<0.05$, ** $p<0.01$

Supplemental Figure 16



Supplemental Figure 16. *In vitro* deletion of *Ring1a/b* simulated *Ring1a/b* KO BM chimera mice.

(A) $GFP^+CD19^-CD11b^+$ myeloid cells were isolated from BMs of BM chimera mice, their genomic DNA was extracted, and the rearrangement status of IgH genes was examined by PCR. Data are representative of 5 control mice and 8 KO mice.

(B-C) For DN3- (B) and proB-derived (C) myeloid cells, D-J and V-DJ rearrangement status of TCR (B) and IgH genes (C) were examined by PCR, respectively.

(D) proB cells isolated from *Ert2Cre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were co-cultured with TSt4 cells for 4 days with or without 4-OHT in the presence of 10 ng/ml of SCF, Flt3-L, and IL-7, followed by culture for 8 days in the presence of 10 ng/ml of SCF, Flt3-L, IL-1α, IL-3, IL-7, TNFα, and GM-CSF.

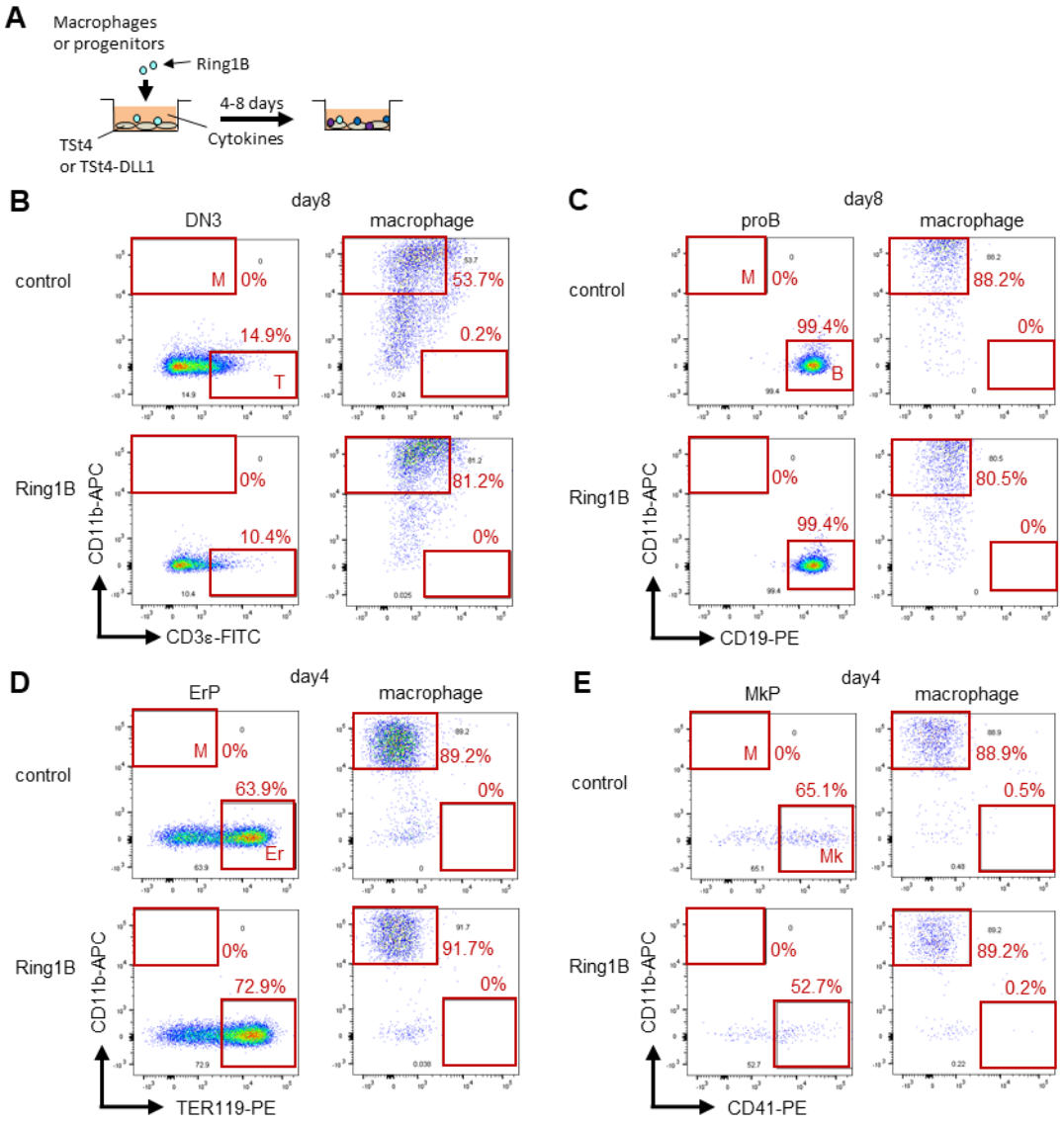
(E) Flow cytometric profiles of the cultured proB cells. Data are representative of 2 independent experiments.

(F) Flow cytometric profiles of DN3 cells in thymi of control (n = 3) and *LckCre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ (n = 3) mice.

(G) DN3 cells isolated from *LckCre-Cdkn2a*$^{-/-}$*Ring1a*$^{-/-}$*Ring1b*$^{fl/fl}$ mice were co-cultured with TSt4-DLL1 cells for 12 days in the presence of 10 ng/ml of SCF, Flt3-L, IL-1α, IL-3, IL-7, TNFα, and GM-CSF.

(H-J) Flow cytometric profiles of the cultured DN3 cells. Data are representative of 3 independent experiments (H). For DN3 derived myeloid cells, V-DJ rearrangement status of TCRβ (I) and IgH genes (J) were examined by PCR, respectively.

Supplemental Figure 17



Supplemental Figure 17. Ring1B over expression did not convert macrophages into non-phagocytic lineage cells.

(A) Macrophages or each lineage progenitors isolated from wild type mice were co-cultured with TSt4 or TSt4-DLL1 cells for 4 - 8 days in the presence of 10 ng/ml of SCF, Flt3-L, IL-3, and M-CSF. In culture medium for T and B cells, erythroid cells, and megakaryocytes, 10ng/ml of IL-7, 2 U/ml of EPO, and 50 ng/ml of TPO were added, respectively.

(B-E) Flow cytometric profiles of the cultured cells. Data are representative of 3 independent experiments.

Supplementarl Table 1. Sources of RNA-seq data of various mouse lineage cells.		
sample name in analysis	source	original sample name
RNA-seq		
Mm_microglial_cell	E-GEOD-63340	brain microglial cell
Mm_Kupffer_cell	E-GEOD-63340	Kupffer cell
Mm_splenic_macrophage	E-GEOD-63340	splenic red pulp macrophage
Mm_lung_macrophage	E-GEOD-63340	lung macrophage
Mm_PC_macrophage	E-GEOD-63340	peritoneal cavity macrophage
Mm_ileal_macrophage	E-GEOD-63340	ileal macrophage
Mm_colonic_macrophage	E-GEOD-63340	colonic macrophage
Mm_macrophage	GSE116177	Mac
Mm_monocyte_BM	GSE116177	MonoBM
Mm_monocyte_PB	GSE116177	MonoPB
Mm_neutrophil_BM	GSE116177	NeutBM
Mm_neutrophil_PB	GSE116177	NeutPB
Mm_LSK	GSE116177	LSK
Mm_T4	GSE116177	CD4T
Mm_T8	GSE116177	CD8T
Mm_FoB	GSE116177	FoB
Mm_MZB	GSE116177	MZB
Mm_GCB	GSE116177	GCB
Mm_Mk	GSE116177	MegTPO
Mm_CFUE	GSE116177	CFUE
Mm_ErBl	GSE116177	ErBlPB
Mm_ESC_1	E-GEOD-27843	embryonic stem cell
Mm_ESC_2	E-GEOD-47948	embryonic stem cell
Mm_fibroblast	E-GEOD-49906	fibroblast
Mm_cardiomyocyte	E-GEOD-49906	cardiomyocyte
Mm_adipocyte	E-MTAB-6502	preadipocyte, 2 month
Mm_hepatocyte	E-GEOD-55552	hepatocyte
Mm_intestinal_epithelial_cell	E-MTAB-9180	epithelium of crypt Lieberkuhn
Mm_intestinal_smooth_muscle	E-MTAB-9180	intestine smooth muscle
Mm_neuron_brain	E-GEOD-52564	neuron
Mm_astrocyte	E-GEOD-52564	astrocyte
Mm_oligodendrocyte_precursor_cells	E-GEOD-52564	oligodendrocyte precursor cells
Mm_newly_formed_oligodendrocytes	E-GEOD-52564	newly formed oligodendrocytes
Mm_myelinating_oligodendrocytes	E-GEOD-52564	myelinating oligodendrocytes
CAGE		
Mm_macrophage	E-MTAB-3578	macrophage
Mm_microglial_cell	E-MTAB-3578	microglial cell
Mm_T_cell	E-MTAB-3578	T cell
Mm_CD4T	E-MTAB-3578	t cell, CD4+
Mm_CD8T	E-MTAB-3578	t cell, CD8+
Mm_B_cell	E-MTAB-3578	b cell, CD19+
Mm_megakaryocyte	E-MTAB-3578	megakaryocyte
Mm_fibroblast	E-MTAB-3578	fibroblast
Mm_cardiomyocyte	E-MTAB-3578	cardiac muscle cell
Mm_adipocyte	E-MTAB-3578	adipocyte
Mm_hepatocyte	E-MTAB-3578	hepatocyte
Mm_enterocyte	E-MTAB-3578	enterocyte
Mm_smooth_muscle_cell	E-MTAB-3578	smooth muscle cell
Mm_neuron	E-MTAB-3578	neuron
Mm_astrocyte	E-MTAB-3578	astrocyte

Supplemental Table 1. Sources of RNA-seq data of various mouse lineage cells.

Supplemental Table 2. Orthogroups of protein coding genes among mouse, tunicate, sponge, and Capsaspora.

Supplemental Table 3. Normalized TPM values of 3237 conserved homolog groups among 4 species.

Supplemental Table 4. Orthogroups of protein coding genes among mouse, tunicate, sponge, choanoflagellate, Capsaspora, and Creolimax.

Supplemental Table 5. Normalized TPM values of 2336 conserved homolog groups among 6 species.

Supplemental Table 6. List of transcription factors and phagocytosis related genes.

Supplemental Table 7. Sequences of synthesized CEBPα homologs and Ring1B.	
CEBPα homologs & Ring1B	sequence
Mus musculus CEBPα	GCCACCATGGAAAGCGCCGATTTCTACGAGGTGGAACCCAGACCTCCTATGAGCAGCCATCTGCAGTCTCCACCTCACGCTCCTAGCAA CGCTGCTTTCGGCTTTCCTAGAGGCGCTGGACCTGCTCCTCCTCCTGCACCTCCTGCTGCTCCAGAACCTCTTGGCGGCATCTGCGAGCAC GAGACATCCATCGACATCAGCGCCTACATCGACCCTGCCGCCTTCAACGATGAGTTCCTGGCCGATCTGTTCCAGCACAGCAGACAGCA AGAGAAGGCCAAGGCTGCTGCTGGACCAGCTGGCGGAGGCGGCGATTTTGATTATCCTGGTGCTCCTGCTGGCCCTGGCGGAGCTGTTA TGTCTGCTGGTGCACATGGACCTCCACCTGGCTACGGATGTGCCGCTGCTGGATATCTGGACGGCAGACTGGAACCCCTGTACGAGAGA GTTGGCGCTCCTGCTCTTAGACCCCTGGTCATCAAGCAAGAGCCCAGAGAAGAGGACGAGGCCAAGCAGCTTGCTCTGGCTGGACTGTT CCCTTACCAGCCTCCTCCGCCTCCACCACCTCCTCATCCTCATGCATCTCCAGCTCACCTGGCCGCTCCTCACCTCCAGTTTCAGATCGCTC ACTGTGGCCAGACCACCATGCACCTTCAGCCTGGACATCCCACACCTCCACCAACACCTGTGCCTTCTCCACATGCAGCACCAGCTCTTG GAGCAGCTGGACTTCCTGGACCTGGCTCTGCTCTGAAAGGACTGGCTGGCGCTCACCCTGATCTGAGAACAGGTGGTGGCGGTGGTGG ATCTGGTGCTGGCGCTGGAAAGGCCAAGAAAAGCGTGGACAAGAACAGCAACGAGTACAGAGTGCGGAGAGAGCGGAACAATATCGC CGTGCGCAAGAGCAGAGACAAGGCTAAGCAGAGAAACGTGGAAACCCAGCAGAAGGTCCTGGAACTGACCAGCGACAACGACAGACT GAGAAAGAGAGTGGAACAGCTGAGCAGAGAGCTGGACACCCTGAGAGGCATCTTCAGACAGCTGCCTGAGAGCAGCCTGGTCAAGGC CATGGGCAACTGTGCTTGA
Ciona intestinalis CEBPα	GCCACCATGGCCTCTAGAAGCAACCCCAACCTGAACTCTGGCGCTGGCGCCCACAAGAACTACATGAGCGCCGACAACACCCTGGACT TCACCGACCTGTACCAGAACGAGGAAAGCTTCGACCTGAGCTGTTACATGAGCACAAGCCCTAACCTGGAACTGTGCAACGACGACCTG TTCGAGAACCTGTTTCCTCCAACAAGCCCCAGCCTGCCTGCCGACCTGAAGAAAGAGAAGTTCGAGGACTTCGAGGAAGAACTGGACAA CTGGTACTACACCACCACCAGCGAGCCTCCAAGCAACTGTGCCAGCGCCTGTTCTAGCGCTGCTCCTTCTCCTGTGCCTTACAACTTCTAC AACACCCCTCCTGTGACACCCCAGAGGCCTCAGTCTCCTCACCATCTGCAGTTCGACGGCCTGACCACACCTATGAGCAGCTCTAGATTC GCCGCCGTGATGGCCGAGGATAGACAGGTTCTGCAGCAGAGACAGGTGCCCCAGAGAATCAACGAGAGCAACATGCCCATCAGCACC CCTATCATCGTGAAGCCCGATCCTGAGGAACAGATGTACCTGACCGAGCAGAGCTTCGCCAACATGAAGGACGAGCCCGTGACCGCTA ACCACGTGCCAACATCTATCGCCAGCCAGCCTGACATCGACACCAAGACCATGAAGAGGCCCGCTCCTATCGACTACAGCAGCAGCAA AAAGACCAAGATCCTGATCAAGGGCAGCAAAGAGTACGTCCAGAAGCGCGAGAGAAACAACGTGGCCGTGCGGAGAAGCAGAGATAA GGCCAAGAGAAAGGCCGCCGAGACACAAGTGAAGGTGGACCAGCTGCAGAATGAGAACCTGAAGCTGCACGAGAAGGTGGCCGAGC TGACCCACGAACTGACAACCCTGAAGAACCTGCTGAAGGCCCTGCCTCACGCTCAGTCTTAA
Amphimedon queenslandica CEBPα	GCCACCATGGATGGAAACAACAGCCTGTACGACCTGTACCTGGCTGCCGGCAAGGGCTTCGTGGTCACAAGCAGCAGCTTCGACACAA ACCCCGTGTACCTGAGGCCTCCTGTGGGCCCTTCTACAATCAACCCTCCTCTGCTGTACCACAACGAGGAAAGCGTGGACTTCAGCGACC TGGTGGAAGCCTGCGAGAAGAACATCACCCCTCCAGCTCCTATCAGCGGCAGCGACTGTAGCTCTGCCTGCTGCTTCTTCGAGGAAAAG CCCTTTCCATTCGACAAGATCCTGCCTATGAGCAGCCCCATGTACAACGAAGAGGCTAGCCCCTACTGCGTGCCACCTCCTCCTCCACCT CATCCACCTAGCACAAGCCACGACTTCGTGCCCCACGAGAACCCTTCTCTGCCTCTGGATGGCATCCTGACCAGCACACCCCACTTCAGC CTGCCTCATCCTACACCTAATGTGCCTGCCGCCACAAGCGTGGCCAGCACAACATCTAGACAGCAGCCCAGACAGAGCAGACTGACCAC CAAGCAGGCCAAGAAGATCTACAACAAGAACAGCGAGGAATACCGGGACAAGCGCGAGAGAAACAACGTGGCCGTGCGGAAGAGCA GAAACAAGATGAAGATGAGGGCCCAAGAGACAGAGCGGAGAGTGCACGAGCTGGAAGAGGAAAACACAGCCCTGCAGAACCAGGTG TCCCTGCTGCTGAAAGAACTGAAGGTGCTGAAGGGCCTGCTGAGCAGCGCTGATGGAATCCAGCAGCCTCAAGTGACCAGCAACATCG ACCCCGTGACCTCCAGATGCAGCTGA
Amphimedon queenslandica CEBPγ	GCCACCATGGCCGATAAGGACAGCCTGTACAGCATCGAGCAGTACCTGGGCCTCGCCGCTTCTGCCAAAGCCACATCTGGCAATCTGTA CCCCTTCGGCCAGGCTCCTCCTAGCTCTGTGCATCTGGGCTTTCAGCTGGAAGAGGAAAGCGTGGACTTCAGCGAGCTGGTCAAGGCCT GCGAGCAGAAGGACTACTACAGCCAGACCACCAACGACTACTACTCCAGCGGCTACTGGTATCAGCAGAAACCTCCTGGCATGCCCATC AGAAGCGAGGAAGAGACAAGCACCCTCGTCGGCGTGGGCACAACAAGCAATCCTCACGACGTGACCAACGTGTCCGCCAATCCTACAC TGCCTCCAGTGGCCAGCTCTCTGTCTGCTCCTTCTCCACCTCCTACCGTGCAGGATACAGCCGCCAGAACCAGACAGACCACAAAGAAG CAGAAAATCAACGACAAGTGCACCGACGACTACAAGGACAAGAGACACCGGAACAATATCGCCGTGCGCAAGAGCAGAAGCAAGTTC AGAAAGAGAGTGCTGGAAACCGAGAAGCGGGTGCAAGAACTGGAAGAGAACAACGCCAAGCTGAAGAACTACGTGGCCCTGCTGCAG AAAGAACTGGCCGTGCTGAAGGGCCTGTTCAGCTCTGCCTCTAACGCTAGCGGCGACGAGAGATACAGCTGA
Capsaspora owczarzaki CEBPα (CAOG_02294)	GCCACCATGGACCTGCCTTGTAGACAGCAGCAGAAGCAGCAACAGCAACAACAGTGCGCCTGCCTGCTGTCTGTCAGAGGCGGAAGAG GCTCTATCCTGAACGCCTGGCACAGCGACAGCTTCATCGGATCTGCTGCCCCTTTCGGCACCTTCGGCTCTACACCTTTCAGCCTGGCCAC AGCCGCTTCTCAGAGAGCCGCATCTATCCTGCCTCCAAGCTTCGACTCCGTGTCTAGCCCTCTGAGCACACTGAGCAGCCCACCTTCTAC ACTGGACGCCTGTGACGACATCCCCGCTCCTACACCTGTGGCCGCTTTTTACCTGCACCAGCTGCTGGGCAGCAGCAGCAAAGAAAGAG ATGTGGACAGACTGCTGCATCACCCCGACATGCTGTCTTACCTGGCTGGCCATTCTGCCGCTCCATTCAGAAGGGCCTCTACCGCCGCCA CAAAGAGGCCTAGTTCTGAGGCTCTGGCTAGCCCTCCTGCACAGAGGCCAAGACTGGATGACGGCTCTCACTCTTCTGCCGCAGCTGCT GCTCCTCCTAGCTCTGCTGTGGATTGGTTCGAGCACGCCAGCCTGAGAGACATCCTGAATGACGAGAACCTGGTGCTGCACCTGGATGC CCCTAGCCTGTCTATGCAGCAGCTCGATCCTAGCCTGGCCTCTCCACTGCCTGTTCAGACACAGCCTGTGGCTGTCAGACTGGACGTGCC ATCTTCTCCACCTGTGCTGGCTAGCATCCCTCTGGTGCCTAGCGCTACAGACCTGGTCACACTGGCCGACGATCACCATCTGCTGCTCGCC GACCATCCTAGAAGCCTGTCTGTGGAAGGCCTGCTGCTGAAGTCTAGCCCACTGGTGTCCAGCCACGACCACAATCTGGCTGTTGCTCCA CTGCTGCTTGCCGGCGATTCACTGCCAGCTGGATCTTGGAGTCTGCCCGTGTCCAACAGCCACTCCGACACATCTCCAAGCAGAGCCCTG CAGAGCCTGCTGGACGACTCTTCTGTGGAACACGCCCTGAACAGCAGCTCCGTGGCTAACGCTCACTTCCCTGATTTCGCCGGCCTGTTC GAGTTCGACGATCTGAGCCCTTTCCAGCCTCCTGCCGCTTTCTTCAGCGAGCAGTCTTTCAGCCCCGCTTCTCCACTGAAAGAGAGCACCC CTCAGCCTTCTCCATCTGCCGCTGCAGCTACTCCAGCTGCTCCTGCAGCTCCTGTGGTGGAACAGAGCCCTATCAGCAAGCCTAGAGCCG CCAGCTCCAGATCCAGAGCCACAACACAAAGAGCCGCCGCTAGCTCCAACAAGCCTGTGAAGGCTTCTGCCGTGGCTGCAGTGGCTGC TGAAAAAGCCGCTGCTGCCGCACCTTCTGCCAGCTCTGATTCTGATGGCGCCCAAGAACTGGACGCTGGCGATCTGCCTCAGAGAAAGC TGAGCAGAAGAGAGCGGAACAATATCGCCGTGCGGAGATGCAGAGACAAGAACAGAGAGAAGTCCCTGGCCGCCAAGAGCCAGTGC GAAACAGTGGCTCAAGAGAACGCTAACCTGAGAGTGCGCATCCACAGCCTGGAACAAGAGGTGTCCTACCTGAAGTCCATGCTGCTGA GCCAAGTGATCCCCGGCATGCCTCTGCCAAGCGCCTCTAGATGA
Capsaspora owczarzaki CEBPγ (CAOG_03823)	GCCACCATGACAATCGAGCCTAGCTTCGCCAGCCAGATCCTGAACTTCCCTTGCACAGCCCCTATCAGCGCTATCGCCGATGACACCTCT GCCATGATGCTGATGCTGGAACCCAGCTTCGTGCCCCAGTTCAGAGGCGACCTGAACGACTTCCCTAGCAGCGACATGCTGGAATTCGA CGCCGCTCTGCTGGCTTTCGAAGCTGCCGCTGCTTCTGAAGGCCTGGCCAGAGTGTGTAAACAGAGCCCTTGTGCCAGCGACGCCAGCA GCGTTAGCGATTCTGCTGTGTCTGACGCCGCCGACAGAGCCTCTTTTGCTTCTGAGCCTGCCTGGCTGAGCCAGACCTTCGATTTCAGCCT GCACAGCAACGCCGACGTGAAGAAGAGATCTGGCAGAGCCCCTACACACCACACCAACACAAAGAACAACATGAGCTTCATCGACGAG AGCGTGTCCCCAAGAAAGCGGAGAGCTGAGGAAGCCCTGAGGGACCCTAGATCTTCTGGCGCCTCTCAGTCTCTGAGCGGCAGCCCTA GAGTGTCCCCTGTTGCTTCTCCTGCCGTGCACCTGAGCACCGACAGACAACTTGAAGCCGGCGAGACACTGGCCGACTACCTGGACAAG AGAAAGAAAAACAACGACGCCGTCAAGAAGTGCAGAGCCCGGAAGAGAATGGCCGTGGTGGCTACCGAAGAGGAATGCCAGAGACT GTCTGGCGAGAACGCCAGCCTGAGAGACAGAGTGGGATCTCTGGAAGCCGAGGTGGCCTATCTGAAGAACCTGCTGATCTCTGCCACA GCCGTGGCCGCTAATCTGGCTTAA
Mus musculus Ring1B	GCCACCATGTCTCAGGCTGTGCAGACAAACGGCACCCAGCCTCTGAGCAAGACATGGGAGCTGAGCCTGTACGAGCTGCAGAGAACAC CCCAAGAGGCCATCACAGACGGCCTGGAAATCGTGGTGTCCCCTAGAAGCCTGCACAGCGAGCTGATGTGCCCTATCTGCCTGGACATG CTGAAGAACACCATGACCACCAAAGAGTGCCTGCACCGGTTCTGCGCCGACTGCATCATCACAGCTCTGAGAAGCGGCAACAAAGAGT GTCCTACCTGCAGAAAGAAACTGGTGTCCAAAAGATCCCTGCGGCCTGATCCTAACTTCGACGCCCTGATCAGCAAGATCTACCCCAGCA GGGATGAGTACGAGGCCCACCAAGAAAGAGTGCTGGCCAGAATCAACAAGCACAACAACCAGCAGGCCCTGAGCCACAGCATCGAGG AAGGACTGAAGATCCAGGCCATGAACAGACTGCAGAGAGGCAAGAAGCAGCAGATCGAGAACGGCTCTGGCGCCGAGGATAACGGC GACAGCTCTCACTGTAGCAACGCCAGCACACACAGCAATCAAGAGGCTGGCCCTAGCAACAAGAGGACCAAGACAAGCGACGACTCTG GCCTGGAACTGGACAACAACAACGCCGCCGTGGCCATCGATCCTGTGATGGATGGCGCTAGCGAGATCGAGCTGGTGTTCAGACCCCA TCCTACACTGATGGAAAAGGACGACAGCGCCCAGACACGGTACATCAAGACATCTGGCAACGCCACAGTGGACCACCTGAGCAAGTAC CTGGCTGTGCGACTGGCCCTGGAAGAACTGAGAAGCAAGGGCGAGAGCAACCAGATGAACCTGGACACCGCTTCCGAGAAGCAGTAC ACCATCTATATCGCCACCGCCAGCGGCCAGTTCACAGTGCTGAACGGCAGCTTCAGTCTGGAACTCGTGTCCGAGAAGTACTGGAAAGT GAACAAGCCCATGGAACTGTACTACGCCCCTACCAAAGAACACAAGTGA

555

556 **Supplemental Table 7.** Sequences of synthesized CEBPα homologs and mouse Ring1B.

Supplemental Table 8. Primer sequences for RT-qPCR and ChIP-qPCR.		
	Forward	Reverse
RT-qCPR of *M. musculus*		
B-actin	TCCTGTGGCATCCATGAAACT	GAAGCACTTGCGGTGCACGAT
Cebpa	CAAGAACAGCAACGAGTACC	GGTDATTGTCACTGGTCAAC
Sfpi1	ACTTCCCTGAGAACCACTTC	TGTAGGAAACCTGGTGAGA
M-CSFR	GCGATGTGTGAGCAATGGCAGT	AGACCGTTTTGCGTAAGACCTG
Lysoz M	GTCAGCCTGGCCGACTGGGTG	GACTCCGCAGTTCCGAATATA
Nox2	CTTTCTCAGGGGTTCCAGTG	TGCAGTGCTATCATCCAAGC
MPO	ATGCAGTGGGGACAGTTTCTG	GTCGTTGTAGGATCGGTACTG
Ebf1	TGGGTTACAGGTCATATTCG	GAACTGCTTGGACTTGTACG
Pax5	CATTCGGACAAAAGTACAGC	GATGCCACTGATGGAGTATG
IL-7R	CAATGAGTGCCCTACCTGAAACT	GGTTATACACAGCAGGGATTCAAA
IKAROS	GAGCTGGACTTGTGTCTTTG	TTTCACACCTTGGTGAACTC
ChIP-qPCR of *M. musculus*		
Cebpa promoter	GAGAAGGCGGGCTCTAAGAC	CGGGGACCGCTTTTATAGAG
RT-qPCR of *C. intestinalis*		
GAPDH	GGCTAAAGCTGTCGGCAAAG	TGGCTTCTCCAAACGAACTG
C/EBPα	TCTTTTCCCACCAACCTCAC	AATTTGACGGTGGCTCACTC
C/EBPγ	TTTATGGACCATGCCAAAGG	CCTGGTTTGTCGGTCAGTTT
PLA2G15	TGGGTTCACAGCTAGAAGCA	AACGGCAAAAGTCCATCAAG

Supplemental Table 8. Primer sequences for RT-qPCR and ChIP-qPCR.

Supplemental Table 9. Primer sequences for TCRβ and IgH rearrangement.	
IgH primer (D-J)	
DSF	GGGATCCTTGTGAAGGGATCACTACTGTG
JH4	AAAGACCTGCAGAGGCCATTCTTACC
IgH primer (V-DJ)	
VH7183	CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC
VHQ52	GCGAAGCTTCTCACAGAGCCTGTCCATCAC
VHg3.8	CAAGGGACGGTTTGCCTTCTCTTTGGAA
VH3609	KCYYTGAAGAGCCRRCTCACAATCTCC
VHJ558	GCGAAGCTTARGCCTGGGRCTTCAGTGAAG
JH4E	AGGCTCTGAGATCCCTAGACAG
TCRβ primer (D-J)	
Dβ1	TTATCTGGTGGTTTCTTCCAGC
Dβ2	GCACCTGTGGGGAAGAAACT
Jβ1.6	GGTAGAAAGGTAGAGGGTTCCAGA
Jβ2.6	TGAGAGCTGTCTCCTACTATCGATT
TCRβ primer (V-DJ)	
Vβ2	GGGTCACTGATACGGAGCTG
Vβ4	GGACAATCAGACTGCCTCAAGT
Vβ5.1	GTCCAACAGTTTGATGACTATCAC
Vβ8	GATGACATCATCAGGTTTTGTC
Vβ14	CTTCTACCTCTGTGCCTGGAGT
Jβ2.6	TGAGAGCTGTCTCCTACTATCGATT

Supplemental Table 9. Primer sequences for TCRβ and IgH rearrangement.