Original Article
Growth factor independence 1 (Gfi1) regulates cell-fate decision of a bipotential granulocytic-monocytic precursor defined by expression of Gfi1 and CD48

Lothar Vassen[1], Ulrich Dührsen[2], Christian Kosan[1*], Hui Zeng[3], Tarik Möröy[1,4]

[1]Institut de recherches cliniques de Montréal IRCM, Montreal, QC, Canada; [2]Westdeutsches Tumorzentrum Klinik für Hämatologie Universitätsklinikum Essen Hufelandstrasse 55 45122 Essen, Germany; [3]Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China; [4]Département de Microbiologie et Immunologie, Université de Montréal, Montréal, QC, H2W1R7 Canada. *present address: Center for Molecular Biomedicine (CMB), Department of Biochemistry, Friedrich-Schiller-University Jena, Hans-Knöll-Str. 2, D-07745 Jena.

Abstract: The transcriptional repressor Gfi1 regulates the expression of genes important for survival, proliferation and differentiation of hematopoietic cells. Gfi1 deficient mice are severely neutropenic and accumulate ill-defined CD11b$^+$GR1^{int} myeloid cells. Here we show that Gfi1 expression levels determine mono- or granulocytic lineage choice in precursor cells. In addition, we identify CD48 as a cell surface marker which enables a better definition of monocytes and granulocytes in mouse bone marrow. Using the CD48/Gr1/Gfi1 marker combination we can show that the CD11b$^+$Gr1^{int} cells accumulating in Gfi1 deficient mice are monocytes and not granulocyte precursors. Expression of CD48, Gr1 and Gfi1 define different bone marrow subpopulations that are either committed to the granulocytic lineage, or bipotential precursors of granulocytes or monocytes. Finally, a comparison of genes differentially expressed between murine Gfi1 high granulocytic precursors and mature granulocytes with gene expression changes from human myeloblasts versus neutrophils show a strong resemblance of human and mouse differentiation pathways. This underlines the value of the markers CD48 and Gfi1 identified here to study human and murine granulo-monocytic differentiation.

Keywords: Gfi1, CD48, CD106, granulocyte, monocyte, myelopoiesis, neutropenia

Introduction

During evolution, multicellular organisms have developed very complex defense mechanisms to manage the detection of pathogen associated molecular patterns or damaged cells. The innate immune system is an immediately available first line of defense composed of two major components: The humoral innate immune system that encompasses the complement system, cytokines and lysozyme and the cellular innate immune system that engages a variety of cell types including mast cells, natural killer cells, eosinophils and basophils as well as phagocytotic dendritic cells, macrophages and neutrophils. Eosinophils, basophils and polymorphonuclear neutrophils are subsets of a larger cellular entity called granulocytes which represent the most abundant leukocyte population in humans. Neutrophil granulocytes can migrate to and rapidly accumulate at sites of infection. After activation they phagocyte and kill pathogens by a series of antimicrobial and proteolytic proteins released from intracellular granules. These granules are subdivided into specific, azurophil and gelatinase granules, which contain different sets of proteins and release their content to either phagosomes (specific and azurophil granules) or to the extracellular environment (specific and gelatinase granules). The proteins stored in neutrophil granules are produced in advance during granulocyte maturation and thus allow for an instant response to invading pathogens.

Failure of proper maturation of neutrophils can induce severe congenital neutropenia (SCN), a primary immunodeficiency with severe clinical

symptoms. The study of genetic defects associated with neutropenia has shed light on the mechanisms of neutrophil maturation. For instance, mutations in neutrophil elastase (Elane/Ela2) [1] and glucose-6-phosphatase-beta (G6PC3) [2, 3] cause endoplasmic reticulum stress and apoptosis of neutrophils. Human adenylate-kinase-2 (AK2) [4] and the HS-1-associated protein X (Hax1) [5] are mitochondrial proteins whose absence causes apoptosis of myeloid progenitor cells. Rare causes of SCN are gain of function mutations in the Wiskott-Aldrich syndrome protein (Was) [6] affecting actin polymerization and loss of function mutations in the only known transcription factor directly associated with SCN, growth-factor-independence-1 (Gfi1) [7, 8]. Mice lacking Gfi1 are severely neutropenic and accumulate a $CD11b^{hi}Gr1^{int}$ cell population [7-9] considered to contain arrested myeloid precursors and monocytes. Gfi1 is crucial for granulopoiesis but not for monopoiesis suggesting that Gfi1 exerts different functions in defined monocyte- and granulocyte precursors. However, how this is achieved remains to be elucidated. In addition, while elaborate protocols for the identification of almost all steps of granulopoiesis exist for human cells [10, 11], the separation of the different neutrophil maturation steps in mice by flow cytometry (FACS) has still to be developed [12]. Efforts to analyze the final steps of granulocyte maturation as well as mouse models of maturation defects or myeloid diseases would greatly benefit from the establishment of a more precise surface marker definition for neutrophil maturation.

Here we used Gfi1:GFP knock in reporter mice [13] to analyze Gfi1 expression of bone marrow derived $CD11b^{hi}Gr1^{lo}$ cells, which comprise monocytes and cells of the granulocytic differentiation pathway. We found that $CD11b^{hi}Gr1^{lo}$ cells contain Gfi1 high and Gfi1 low expressing subsets. This differential expression also indicated a functional separation since these two subpopulations were found to be primed to differentiate in response to GM-CSF into the granulocytic or monocytic lineage, respectively. Whole genome gene expression analysis indicated that both Gfi1 high and low subsets run different genetic programs that ensure their lineage potential. In the case of Gfi1 high cells this was consistent with terminal granulocyte maturation [14, 15]. In addition the analysis of the genetic program of Gfi1 high and low

$Cd11b^{hi}Gr1^{lo}$ cells showed that the surface markers CD48 and to a limited extend also CD106/Vcam1, can be used to follow granulocyte maturation and for the identification of bipotential granulocytic-monocytic precursor cells.

Materials and methods

Mice

The generation of GFP-Gfi1 knock in mice has been described previously [13]. Mice were housed under specific pathogen-free conditions. Institutional animal ethics committees reviewed animal experimentation protocols and certified animal technicians regularly observed the mice. All mice were backcrossed with C57/B6 mice for at least 10 generations. No phenotype or differences in number of cells was observed for $Gfi1^{KI/WT}$ mice.

Flow cytometry, cell sorting, microarray analysis and Q-PCR

Monocytic and granulocytic cell populations were analyzed by flow cytometry using an LSR and sorted using a FACSVantageDiVa (Becton Dickinson) or MoFlo (Beckman Coulter) from adult mouse bone marrow from heterozygous or homozygous EGFP:Gfi1 knock-in mice. Antibodies were purchased from BD Biosciences (CD11b-PerCP-Cy5.5, M1/70; Gr1-PE, RB6-8C5), BioLegend (CD48-APC, HM48-1) or Serotec (CD106-Alexa-Fluor 647, MVCAMA(429). TRIzol (INVITROGEN) or triReagent (MRC) was applied to isolate RNA from sorted cells. Quantitative RT-PCR was performed using PerfeCTA SYBR Green SuperMix, UNG, Low ROX (Quanta BioSciences, Gaithersburg, MD) in a 14-µl reaction volume containing 20pM of each primer and 1µl of cDNA generated using the RT2 First Strand Kit (SA-Biosciences, Hilden, Germany) according to the manufacturer's instructions. The expression of the gene of interest was calculated relative to the ß-actin mRNA levels as $2^{-\Delta ct}$ and error bars represent the standard deviations (+-SD). All analyses were done on at least two biological replicates and triplicate measurements. Primers used were as follows: Vcam1-left: tggtgaaatggaatctgaacc; Vcam1-right: cccagatggtggtttcctt; CD48-left: cgagttgaagataaccctgga; CD48-right: tcgacgcttcagtcttattgatt; Kit-left: gatctgctctgcgtcctgtt; Kit-right: cttgcagatggctga-

ga-cg; Myb-left: tgtcaacagagaacgagctga; Myb-right: gctgcaagtgtggttctgtg; Csf1r-left: cgagg-ga-gactccagctaca; Csf1r-right: gactggagaagc-cactg-tcc; Ccr2-left: acctgtaaatgccatgcaagt; Ccr2-right: tgtcttccatttcctttgatttg; ß-Actin-US: ctacaatgagctgcgtgtggc; ß-Actin-LS: caggtc-cagacgcaggatggc. For whole genome gene expression analysis, a total of 10 µg cRNA from sorted cells was hybridized on Affymetrix Mouse Genome 430A 2.0 arrays (GPL8321). After washing and staining, hybridized chips were scanned using the Affymetrix scanner M10, and data were analyzed with AltAnalyze [16] software using default analysis settings. Single samples were generated for each subgroup. Gene set enrichment analysis was performed using the GSEA software (www.broadinstitute.org/gsea) and hierarchical clustering analysis and scatter plots were generated using the Spotfire Decision Site software (www.spotfire.tibco.com).

Array data have been deposited in the public database Gene Expression Omnibus repository, accession: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE35970

Cell culture

Cytospins of sorted bone marrow derived cell populations were prepared either immediately or after culture in Dulbecco modification of Eagle medium (DMEM, Wisent) in the presence of 10% heat inactivated fetal calf serum (FCS, HyClone, Thermo Scientific) and either 15 ng/ml mouse recombinant GM-CSF, 15 ng/ml G-CSF or 15 ng/ml M-CSF (all Peprotech) for up to three days.

Results

Gfi1 expression levels separate two CD11b^{hi}-Gr1^{lo} subpopulations

Gfi1 is expressed in GMPs, which are very early bi-potential granulocytic-monocytic precursors, but is absent from monocytes and is not required for monocytic differentiation. In contrast, Gfi1 is expressed in granulocytes and is strictly required for granulocytic differentiation [7-9]. Therefore GMPs have to produce precursors for these lineages that either maintain or extinguish Gfi1 expression. To identify these precursors, we used Gfi1:EGFP knock-in (Gfi1^{KI}) mice, in which the expression of endogenous

Gfi1 can be monitored by measuring green fluorescence [13] (**Figure 1A**). Heterozygous Gfi1$^{KI/WT}$ mice are indistinguishable from wt mice, while homozygous GFP knock-in mice (Gfi1$^{KI/KI}$) are Gfi1 deficient (i.e. Gfi1^{null}) and show the typical Gfi1 null neutropenic phenotype [13]. The analysis of bone marrow cells from Gfi1$^{KI/WT}$ mice by flow cytometry using CD11b (Mac1) and Gr1 markers showed that CD11b^{hi}Gr1^{lo} cells contain two populations of cells that express higher or lower Gfi1 levels (**Figure 1B**, gates A1 and A2), while the mature CD11b^{hi}Gr1^{hi} granulocytes are uniformly Gfi1-high (**Figure 1B**, gate B). Gfi1$^{KI/KI}$ mice lack CD11b^{hi}Gr1^{hi} neutrophils and accumulate CD11b^{hi} Gr1^{int} cells in gate C (**Figure 1B**, upper right). The two CD11b^{hi} Gr1^{lo} subpopulations from Gfi1$^{KI/WT}$ mice (**Figure 1B**, gates A1 and A2) and the granulocytes (**Figure 1B**, gate B) also differ in size and granularity (**Figure 1B**, lower rightmost panel). We hypothesized that the subpopulations defined by gates A1 and A2 may contain the precursor cells, which have either monocytic or granulocytic potential.

Gfi1 expression defines two subsets of CD11b^{hi}Gr1^{lo} cells with different genetic programming

To test this hypothesis, we sorted the cells from gates A1, A2, B and the cells from Gfi1 deficient mice that were defined by the "C" gate (**Figure 1B**) and performed genome wide gene expression analyses using the Affymetrix microarray technology. Unsupervised hierarchical clustering of the expression data showed a very high similarity between CD11b^{hi} Gr1^{lo} Gfi1^{lo} cells (gate A1) and CD11b^{hi} Gr1^{lo} Gfi1$^{KI/KI}$ cells (gate C) and separated them from CD11b^{hi} Gr1^{lo} Gfi1^{hi} cells (gate A2) and CD11b^{hi} Gr1^{hi} Gfi1^{hi} granulocytes (gate B) (**Figure 2A**). As a measure for similarity, the correlation coefficients between A1 and the three other subsets were calculated and validated the clustering results (**Figure 2A**). CD11b^{hi}Gr1^{lo} Gfi1^{hi} and CD11b^{hi}Gr1^{lo} Gfi1^{lo} cells (gates A1 and A2) differ strongly in their gene expression pattern as is visualized in a scatterplot (**Figure 2B**, left panel). Genes that are typically highly expressed in granulocytes such as neutrophil elastase (Ela2/Elane) or myeloperoxidase (Mpo) [17] are also more highly expressed in Gfi1^{hi} (A2) cells than in Gfi1^{lo} cells (A1). In contrast, the colony stimulating factor 1 receptor (Csf1r/M-Csfr) gene is much higher expressed in Gfi1^{lo} (A1) cells than in Gfi1^{hi} cells
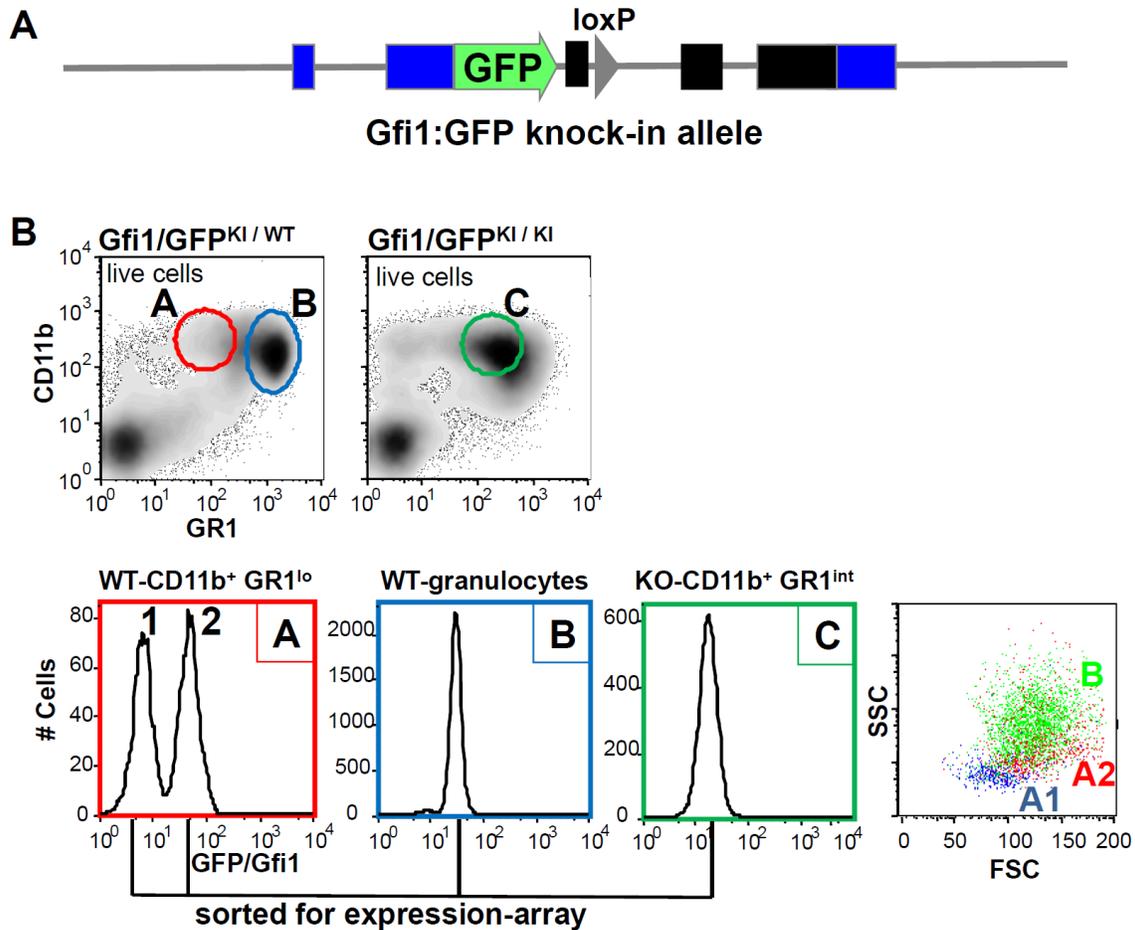
Figure 1. Gfi1 expression defines two subpopulations in the classical „monocytes" gate. A: The Gfi1:GFP knock-in allele replaces exon 3-5 of Gfi1 by GFP (A) in frame with the Gfi1 coding sequence and allows to follow Gfi1 expression by GFP-fluorescence analysis. B: Bone marrow cells from heterozygous and homozygous Gfi1:GFP knock-in mice were analyzed for Gr1, CD11b and GFP expression in the classical monocytes (A) (C for Gfi1-KI/KI mice) and granulocytes (B) gates by FACS. While granulocytes show a homogeneous, high Gfi1:GFP expression (B), monocytes subdivide into Gfi1-low (A1) and Gfi1-high (A2) populations which also differ in FSC/SSC (lower right panel). Gfi1 deficiency of homozygous Gfi1:GFP mice represents a phenotypical Gfi1-KO expanding CD11b$^+$Gr1^{int} cells (C). Cells from all 4 fractions were sorted for expression analysis on Affymetrix cDNA arrays. Array data were submitted to the NCBI GEO repository under ACC number GSE35970. (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE35970).

(A2) (**Figure 2B**, left panel). These data indicate that i) the level of Gfi1 expression indeed separates two different fractions of CD11b^{hi} Gr1^{lo} cells and ii) that Gfi1^{hi} cells (gate A2) seem to be more closely related to the granulocytic lineage (gate B) than A1 cells, which seem to be closely related to Gfi1-KO myeloid cells (**Figure 2B**, right panel and suppl Figure 1). If compared to the expression levels of A1 cells, 12.2% of all probesets in A2 differ more than 2-fold in expression level, while only 1.4% of the genes expressed in Gfi1-KO cells from gate C are more than 2-fold differentially expressed (**Figure 2B**, right panel). Interestingly, the hallmark neutro-

phil markers Ela2 and Mpo are more highly expressed in Gfi1-KO cells from gate C than in A1 monocytes (**Figure 2B**, right panel). Ela2 is a known Gfi1 target gene and therefore expected to be up-regulated in Gfi1-KO monocytes. Whether Mpo is also a true Gfi1 target as would be expected from this result, remains open.

Gene set enrichment analysis (GSEA) using the gene expression array data from populations A1 and A2 showed that CD11b^{hi}Gr1^{lo}Gfi1^{lo} cells (A1) up-regulate the expression of genes involved in the humoral immune response, surface chemokine receptors and members of the
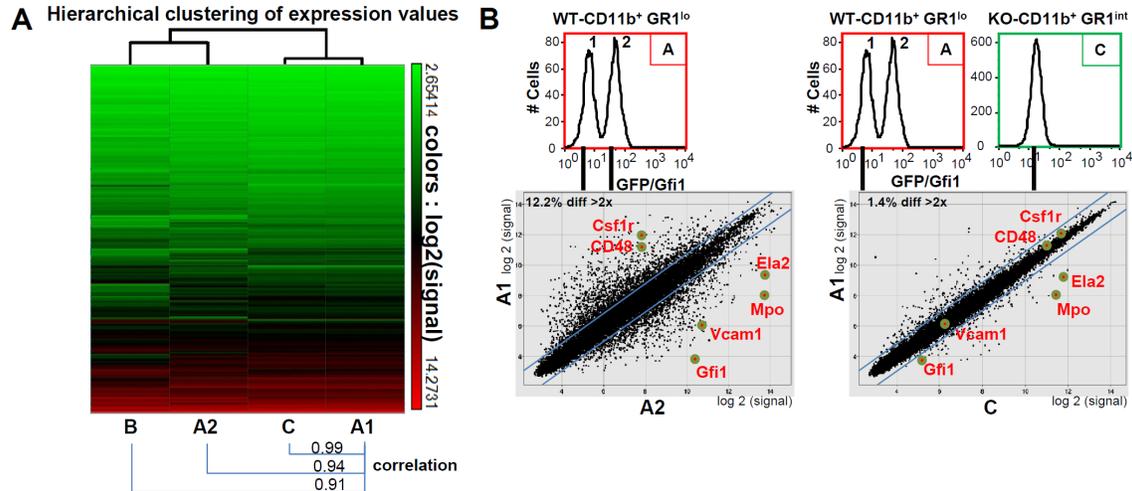
Figure 2. Gfi1-KO monocytes are highly similar to GFP/Gfi1-low CD11b⁺ Gr1ˡᵒ cells and differ from GFP/Gfi1-high CD11b⁺ Gr1ˡᵒ cells and granulocytes. A: Hierarchical clustering analysis of expression array data (log2 of normalized signal) from sorted cell populations. A1:Gfi1ˡᵒ CD11b⁺ Gr1ˡᵒ; A2:Gfi1ʰⁱ CD11b⁺ Gr1ˡᵒ; B:CD11b⁺ Gr1ʰⁱ (granulocytes); C:Gfi1KO CD11b⁺ Gr1ⁱⁿᵗ. The gene expression pattern of Gfi1KO CD11b⁺ Gr1ⁱⁿᵗ cells shows very high similarity to Gfi1ˡᵒ CD11b⁺ Gr1ˡᵒ monocytes, while Gfi1ʰⁱ CD11b⁺ Gr1ˡᵒ cells are more similar to granulocytes. As a measure of similarity between genesets, the correlation coefficient (correlation) between population A1 and C, A2 and B respectively was calculated. B: Scatter plot comparisons of gene expression levels (log-2 of normalized signal intensities) show a much higher similarity between Gfi1-KO CD11b⁺ Gr1ⁱⁿᵗ cells and GFP/Gfi1-low monocytes (right panel) than between Gfi1ʰⁱ CD11b⁺ Gr1ˡᵒ and Gfi1ˡᵒ CD11b⁺ Gr1ˡᵒ cells. Some marker genes identified in this study (Gfi1, Vcam1, CD48) and genes typically expressed in monocytes or neutrophils (Mpo, Ela2, Csf1r) are indicated in the scatter plots. Blue lines indicate the borderlines of 2-fold change of probeset intensities of the compared samples. The percentage of probesets with a more than 2-fold differential expression between compared samples is given in the upper left corner of each plot.

Jak/Stat signaling pathway (**Figure 3A**). In contrast CD11bʰⁱGr1ˡᵒGfi1ʰⁱ cells (A2) increase the expression of genes associated with cell cycle progression (**Figure 3A** and <u>suppl Figure 2</u>). Hoechst staining showed that a higher proportion of CD11bʰⁱGr1ˡᵒGfi1ʰⁱ cells (A2) are in S/G2/M phase than of CD11bʰⁱGr1ˡᵒGfi1ˡᵒ cells (A1) (**Figure 3B**), suggesting that CD11bʰⁱGr1ˡᵒGfi1ʰⁱ (A2) cells are cycling, whereas CD11bʰⁱGr1ˡᵒGfi1ˡᵒ (A1) cells are arrested in G1.

To define the factors involved in the different transcriptional programming of CD11bʰⁱGr-1ˡᵒGfi1ˡᵒ (A1) compared to CD11bʰⁱGr1ˡᵒGfi1ʰⁱ cells (A2) we searched for transcription factor binding sites that were enriched in the promoter regions of differentially expressed genes by GSEA. The promoters of genes more highly expressed in CD11bʰⁱGr1ˡᵒGfi1ˡᵒ cells (A1) were found to contain binding sites in their promoters for Stat5/Stat3, which are important effectors of signaling through cytokine receptors such as GM-CSF [18]. Genes with higher expression in the CD11bʰⁱGr1ˡᵒGfi1ʰⁱ fraction

(A2) contain binding sites in their promoters for E2F, which is a master regulator of cell cycle progression (**Figure 3B**). These findings suggest that Gfi1 expression distinguishes two subpopulations in CD11bʰⁱGr1ˡᵒ cells: one subset, which is Gfi1ˡᵒ, is receiving cytokine signals and is quiescent (A1 cells) and the other subset, which is Gfi1ʰⁱ, is proliferating (A2 cells).

Gfi1 expression defines the lineage potential of CD11bʰⁱGr1ˡᵒ cells for monocytic or granulocytic differentiation

To test the lineage potential of CD11bʰⁱGr1ˡᵒGfi1ˡᵒ (A1) and CD11bʰⁱGfi1ʰⁱ (A2) cells, we FACS-sorted and stained them for a morphological analysis either directly or after 3 days of culture in the presence of GM-CSF, which can initiate both granulocytic and monocytic differentiation (**Figure 4**). CD11bʰⁱGr1ˡᵒGfi1ˡᵒ cells (A1 gate) were almost all monocytes and developed into macrophages in the presence of GM-CSF (**Figure 4A**, **4B**). In contrast, CD11bʰⁱGr1ˡᵒGfi1ʰⁱ cells (A2 gate) resembled myelocytes and metamyelocytes and developed into mature

A in *silico* / in *vivo*

GO-pathways

enriched in A1	p-val
RECEPTOR_COMPLEX	<0.01
HUMORAL_IMMUNE_RESPONSE	<0.01
RESPONSE_TO_VIRUS	<0.01
IMMUNE_RESPONSE	<0.01
CHEMOKINE_RECEPTOR_BINDING	<0.01
JAK_STAT_CASCADE	<0.01

enriched in A2	p-val
CELL_CYCLE_PROCESS	<0.01
CELL_CYCLE_PHASE	<0.01
M_PHASE	<0.01
MITOTIC_CELL_CYCLE	<0.01
DNA_REPLICATION	<0.01
CELL_CYCLE_GO_0007049	<0.01

B promoter elements

enriched in A1	p-val	enriched in A2	p-val
STAT5B_01	<0.01	E2F_Q6_01	<0.01
STAT3_01	0.01	E2F_Q4_01	<0.01
ISRE_01	<0.01	E2F_01	<0.01
TATCTGG,MIR-488	0.01	E2F_03	<0.01
HNF1_C	<0.01	E2F1_Q3	<0.01
ATCTTGC,MIR-31	0.02	SGCGSSAAA_V$E2F1DP2	<0.01
GGTGTGT,MIR-329	<0.01	E2F1_Q6	<0.01
AACATTC,MIR-409-3P	0.01	E2F1_Q6_01	<0.01
SREBP1_02	0.01	E2F_Q3_01	<0.01
STAT5A_01	<0.01		

Figure 3. Geneset enrichment analysis (GSEA) shows physiological differences between Gfi1^{hi} CD11b$^+$ Gr1^{lo} and Gfi1^{lo}CD11b$^+$Gr1^{lo} cells in differentiation and cell cycle. A: GSEA of array data comparing Gfi1^{lo} CD11b$^+$Gr1^{lo} (A1) and Gfi1^{hi} CD11b$^+$Gr1^{lo} (A2) cells (in *silico*) revealed an enrichment in GO-pathways associated with humoral immune response and other typical mature monocyte functions and signaling cascades in Gfi1-low cells, while cell cycle associated gene sets are down in Gfi1^{lo} cells compared to Gfi1^{hi} cells. B: In *vivo* analysis of cell cycling using HOECHST staining in FACS analysis proved that Gfi1^{hi} CD11b$^+$ Gr1^{lo} cells cycle more than Gfi1^{lo} cells. Variance analysis of three independent experiments is indicated for S/G2/M phase cells. C: Analysis of transcription factor binding sites enriched in the promoters of genes differentially expressed in Gfi1^{lo} (A1) and Gfi1^{hi} CD11b$^+$ Gr1^{lo} (A2) cells defines genesets with higher expression in Gfi1^{lo} cells for primary immune response regulated by STAT transcription factors (left panel). Cell cycle regulators enriched in Gfi1^{hi} cells are targets of E2F transcription factors (right panel).

neutrophils (73%) in the presence of GM-CSF (**Figure 4A**, **4B**). This indicated that the

CD11b^{hi}Gr1^{lo} population falls into two populations with the potential to either differentiate
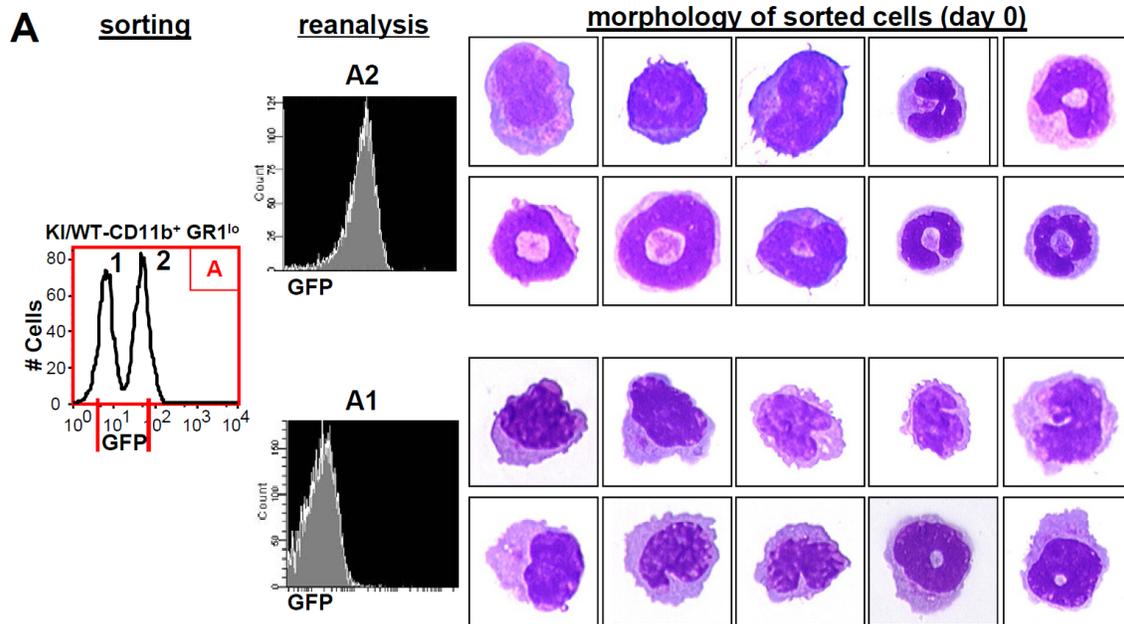
Figure 4. Gfi1^{lo} CD11b$^+$ Gr1^{lo} cells are monocytes, Gfi1^{hi} CD11b$^+$ Gr1^{lo} cells differentiate exclusively into granulocytes. A: CD11b$^+$ Gr1^{lo} bone marrow cells from Gfi1:GFP knock-in mice were separated for GFP-low (A1) or GFP-high (A2) expression by FACS and their morphology was either analyzed directly by May–Grünwald–Giemsa-staining or after three days of culture in the presence of GM-CSF to induce granulocytic / monocytic maturation. B: Phenotypic characterization of sorted cells directly or after 3 days of culture in the presence of GM-CSF (15ng/ml) identifies almost all GFP/Gfi1^{lo} (A1) cells as monocytes, differentiating into macrophages during culture. GFP/Gfi1^{hi} (A2) cells are immature myelocytes or metamyelocytes differentiating mainly into granulocytes.

into monocytes (A1 cells) or into granulocytes (A2 cells) and that the expression level of Gfi1 in these two subsets determines this lineage choice.

CD48 as a marker for granulo-monocytic maturation

To identify a surface-marker protein that similar to Gfi1 could define the potential of precursors to differentiate into monocytes or granulocytes, we further analyzed the gene expression data of A1 and A2 cells. We found that Vcam1 clustered with genes that are highly expressed in A2 cells, but are lowly expressed in A1 cells and

in the B and C fraction (**Figure 5A**, heat map, left). In addition, FACS analysis showed that A2 cells express high levels of Vcam1, but also indicated that Vcam1 didn't separate A1 and A2 cells as well as Gfi1 expression (**Figure 5A**, right panel), In contrast, CD48, which is expressed on a wide range of lymphoid and myeloid cells [19] was found to be expressed at lower levels in populations B and A2, where Gfi1 expression is high, and was high in populations C and A1 where Gfi1 expression is absent or low (**Figure 5B**, heat map, left). In addition, using flow cytometry, two distinct populations could be defined using CD48 and Gr1, which were GFP low (i.e. had low levels of Gfi1) or were
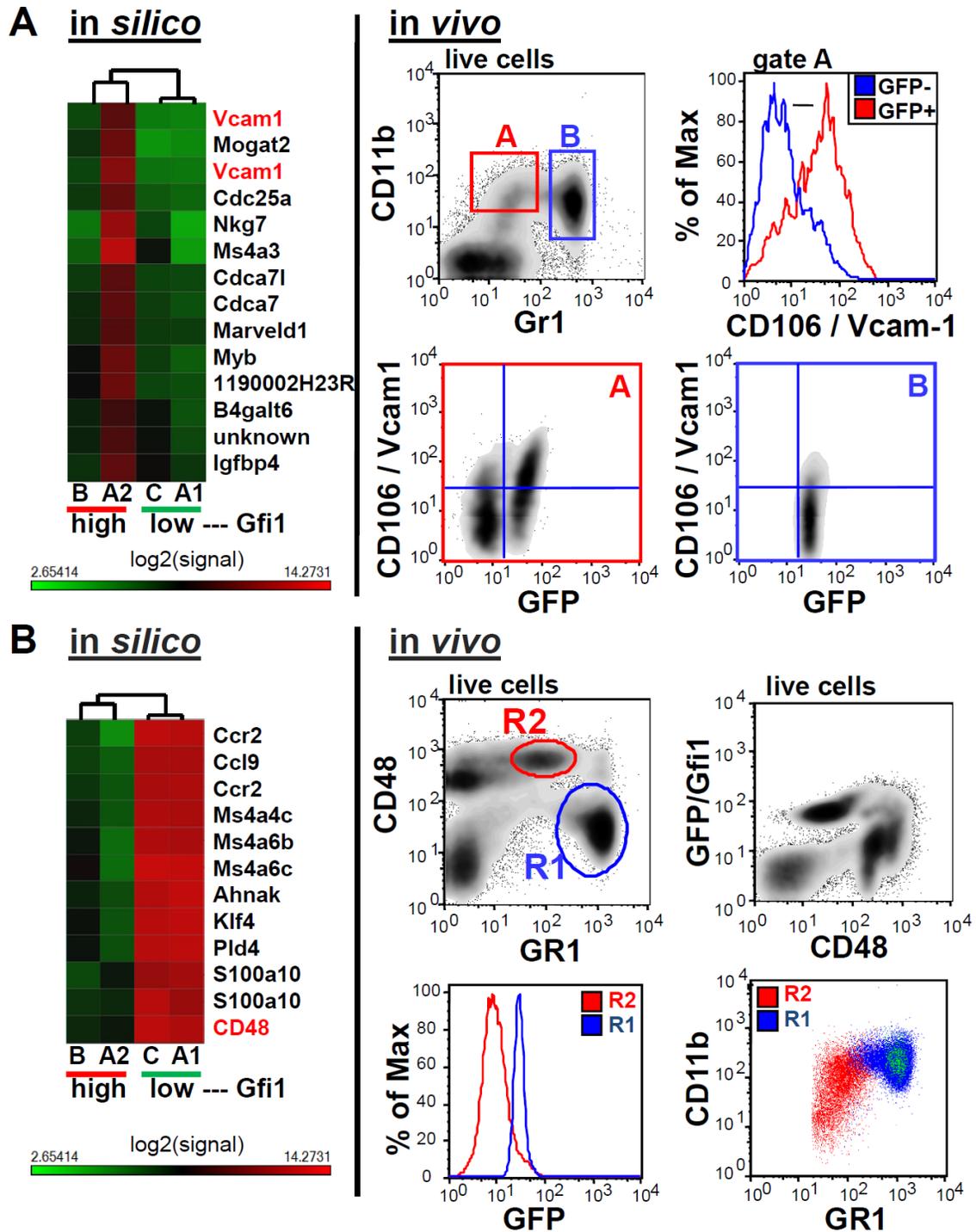
Figure 5. Hierarchical clustering analysis predicts suitable markers for monocyte / granulocyte development. A: Hierarchical cluster analysis of expression values for all four sorted cell populations was performed. Vcam1/CD106 expression is correlated to GFP:Gfi1 expression *in silico* in CD11b^hiGr1^lo cells (A1, A2), but not in granulocytes (B). This could be verified *in vivo* using FACS analysis (right panel), although the correlation is not perfect in monocytes. B: CD48 expression is inversely correlated to GFP/Gfi1 expression *in silico* (left panel) and CD48 separates Gfi1 positive (R1 granulocytes) and negative (R2, monocytes) CD11b^+ Gr1^lo fractions *in vivo* (right panel). When Gfi1 expression is turned on, CD48 goes down (upper right plot). CD48 can replace CD11b for FACS analysis of bone marrow derived myeloid cells resulting in a higher resolution analysis of granulocytic development. FACS analyses shown are representative for at least four independent experiments.
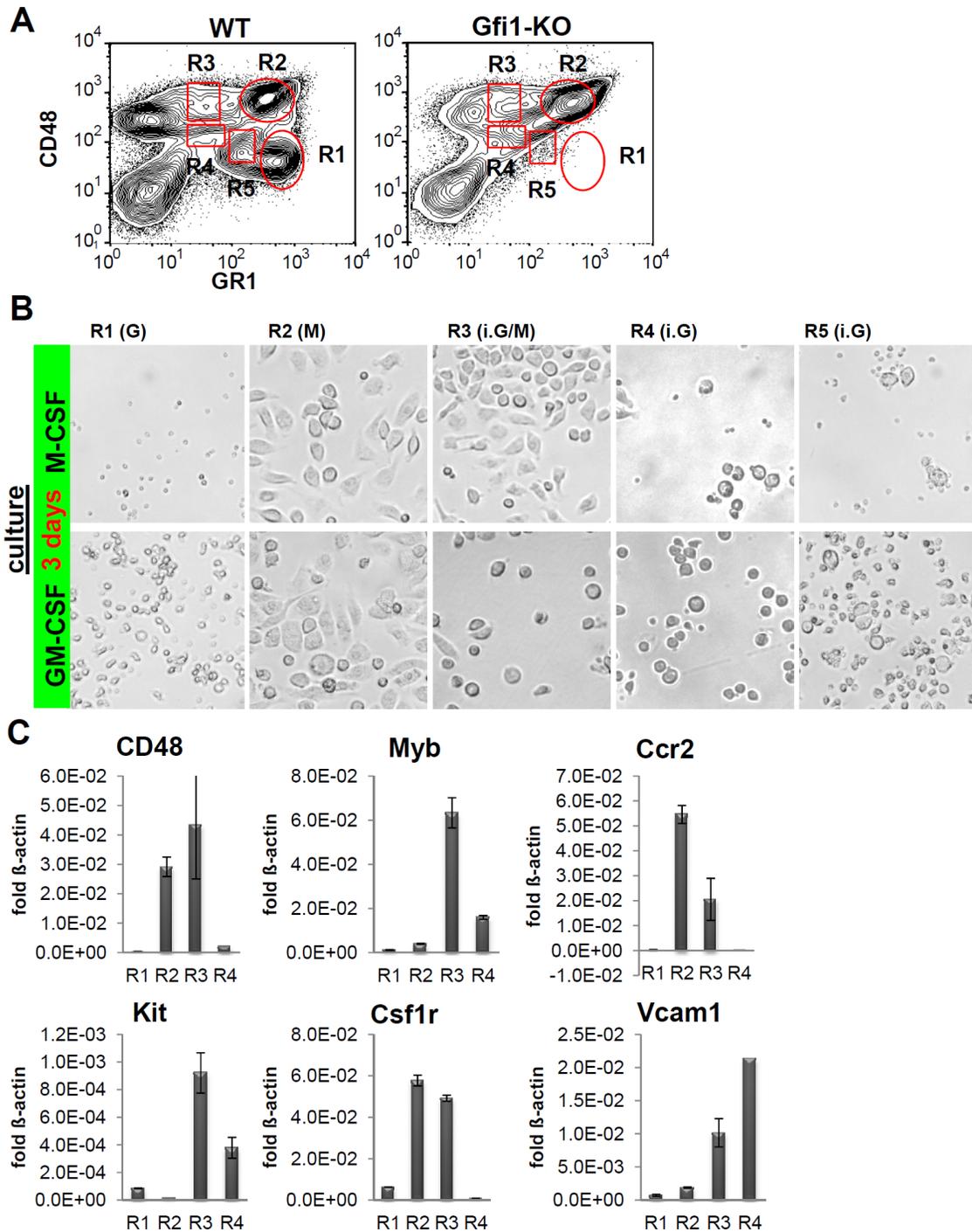
Figure 6. Gfi1 regulates cell fate decision of the bipotential monocyte / granulocyte precursor. A: FACS analysis of bone marrow cells from wt or Gfi1-KO mice using CD48 in conjunction with Gr1. The complete CD48^{lo} Gr1l^{o-hi} cell population is missing in the Gfi1-KO. B: Cells from the gates R1-R4 were sorted and cultured for 3 days in the presence of M-CSF or GM-CSF. Monocytes maturing to adherent cells (macrophages), are clearly separable from immature granulocytes incapable to produce adherent cells. C: Cells from the gates R1-R4 were sorted and total RNA was subjected to quantitative RT-PCR for the genes indicated. CD48 was analyzed as a control for sorting. Chemokine (C-C-motif) receptor 2 (CCR2) and colony stimulating factor 1 receptor (Csf1r/M-Csfr) with higher expression levels in CD48-high cells are typically highly expressed in monocytes/macrophages. The myeloblastosis oncogene (Myb) and stem cell factor receptor Kit, more highly expressed in Gr1-low cells,. Vascular cell adhesion molecule 1 (Vcam1) shows highest expression in R4, which corresponds well to Gfi1^{hi}Cd11b^{+}Gr1^{lo} (A2) cells analyzed in our arrays (see Figure 5).

A   correlation = 0.74   p = 4.3 E-8

B

low in immature cells

Category	Term	PValue
BP	immune response	1.96E-16
BP	response to wounding	8.79E-10
BP	defense response	3.02E-08
BP	cell activation	5.34E-08
BP	regulation of cytokine production	2.46E-07
BP	inflammatory response	4.45E-07
CC	plasma membrane	1.01E-06
CC	vesicle	0.00138973
CC	anchored to membrane	0.00169814
CC	cytosol	0.00245792
CC	cytoplasmic vesicle	0.00295908
CC	membrane raft	0.0052855
MF	cytokine binding	5.29E-06
MF	cytokine receptor activity	1.42E-05
MF	transcription factor activity	0.0021566
MF	cytoskeletal protein binding	0.00305239
MF	lipopolysaccharide binding	0.00560374
MF	enzyme binding	0.00654102

high in immature cells

Category	GO-Term	PValue
BP	cell cycle phase	6.30E-35
BP	cell cycle	4.42E-34
BP	DNA metabolic process	1.31E-33
BP	M phase	4.07E-33
BP	mitotic cell cycle	1.15E-32
BP	cell cycle process	8.58E-32
CC	ribonucleoprotein complex	4.40E-39
CC	non-membrane-bounded organelle	1.06E-36
CC	intracellular non-membrane-bounded organelle	1.06E-36
CC	membrane-enclosed lumen	9.43E-32
CC	intracellular organelle lumen	2.24E-30
CC	organelle lumen	2.92E-30
MF	nucleotide binding	5.24E-26
MF	ATP binding	4.84E-21
MF	adenyl ribonucleotide binding	1.37E-20
MF	adenyl nucleotide binding	5.16E-20
MF	nucleoside binding	7.00E-20
MF	RNA binding	9.05E-20

Figure 7. Murine and human granulocyte (neutrophil) maturation involves a highly overlapping gene set. A: Scatter plot comparing the genes whose expression changes at least 2-fold during the maturation of mouse and human granulocytes. Y-axis is showing the log2 of fold changes of mouse bone marrow derived Gfi1[hi] CD11b[+] Gr1[lo] cells versus granulocytes, the X-axis shows the log2 of fold changes of human myeloblasts compared to neutrophils (reanalysis of data taken from the NCBI GEO database GSE12837, [15]). Human/mouse probeset correlation was based on the "homologene database". Numbers in the corners indicate the number of genes showing equal (lower left and upper right) or opposite (upper left and lower right) regulation. Over 12 fold more genes are regulated the same way than the opposite way. A positive correlation coefficient of 0.74 was determined and students T-test revealed a p-value of 4.3 E-08. This strongly supports the finding, that Gfi1[hi] CD11b[+] Gr1[lo] cells are immature granulocytes. B: Functional annotation of genes with either lower (left panel) or higher (right panel) expression in immature compared to mature mouse- and human granulocytes. Probesets that were differentially expressed the same way in mouse and human were subjected to functional annotation using the DAVID online tool and enriched GO-categories were selected. The 6 most significantly enriched terms of each of the three GO-categories, "Biological Process" BP FAT (BP), "Cellular Component" CC FAT (CC) and "Molecular Function" MF FAT (MF) are listed for both analyses.

Gfi1^{hi} (**Figure 5B** right panel, gates R1 and R2). This suggested that cells down-regulate CD48 as soon as they upregulate Gfi1. These populations (R1 and R2) although clearly defined by CD48 and Gr1 expression, could not be clearly separated by the CD11b and Gr1 markers (**Figure 5B**, lower right panel) indicating that the use of CD48 and Gr1 as markers allows a better resolution of myeloid precursors than the use of CD11b and Gr1 markers (**Figure 5B**).

Gfi1 defines cell fate at the state of the bipotential granulocytic/monocytic precursor

To further validate the CD48/Gr1/Gfi1 marker combination, we isolated bone marrow from Gfi1:GFP reporter mice and analyzed cells for GFP expression falling into the previously defined gates R1 and R2 or the newly defined gates R3, R4 and R5 (**Figure 6A** and suppl Figure 3). R1, R2 and R5 cells were clearly identified as homogenous populations with either low (R2) or high (R1, R5) Gfi1 expression, whereas cells from gates R3 and R4 showed a broader spectrum of Gfi1 expression levels (suppl Figure 3). In Gfi1 deficient mice that lack granulocytes, cells falling onto gate R1 (CD48^{low}/Gr1^{high}) are almost completely absent from the bone marrow (**Figure 6A**), suggesting that R1 cells are granulocytes. This analysis also suggests that the accumulating cells seen in Gfi1 deficient mice are monocytes, since they fall into gate R2 in this type of analysis (**Figure 6A**).

In addition, cells from the Gr1^{hi} CD48^{hi} monocytes-gate accumulating in the Gfi1KO mice show a homogeneous GFP expression (suppl Figure 4), which would not be expected from a heterogeneous population with differential Gfi1-promoter activity. R4 and R5 populations were absent or strongly reduced in Gfi1 deficient mice (**Figure 6A**), suggesting that these subsets could be granulocytic precursors.

To test the differentiation potential of these subsets, we sorted bone marrow cells falling into the R1, R2, R3, and R4 gates and cultured them for three days in the presence of either M-CSF or GM-CSF (**Figure 6B**). As expected, Gr1^{+}CD48^{lo} cells (R1) could not grow in the presence of M-CSF, nor gave rise to adherent cells in culture, but grew in the presence of GM-CSF (**Figure 6B**, R1), indicating again that these cells are granulocytes. Almost all the cells from the Gr1^{int}CD48^{hi} gate (R2) gave rise

to adherent macrophages in the presence of M-CSF as well as GM-CSF after 3 days in culture, confirming that these cells represent indeed monocytes (**Figure 6B**, R2). In contrast, Gr1^{vlo}CD48^{hi} cells (R3) proliferated well in the presence of M-CSF and GM-CSF and after 3 days in culture gave rise to adherent cells efficiently in the presence of M-CSF. In the presence of GM-CSF adherent cells and non-adherent cells grew out (**Figure 6B**, R3). Gr1^{vlo}CD48^{lo} cells (R4) proliferated poorly in the presence of M-CSF and only very few adherent macrophages were seen after 3 days of culture. In the presence of GM-CSF these cells proliferated, but could not give rise to adherent cells even after 3 days of culture (**Figure 6B**, R4). These data suggest that R3 cells that express Gfi1 at intermediate levels and are present in the Gfi1KO mice are either bipotential granulomonocytic precursors or contain cells that are able to differentiate into either lineage. Also, these findings indicate that R4 cells, which are lost in Gfi1KO mice, are already committed to granulocytic differentiation.

Next, we analyzed the sorted bone marrow cells falling into the R1-R4 gates for the expression of marker genes for mature monocytes and the maturation of myeloid cells by Q-PCR (**Figure 6C**). We found that those genes that are typically expressed in monocytes, such as chemokine (C-C motif) receptor 2 (CCR2) and colony stimulating factor 1 receptor (Csf1r/M-Csfr) were also most highly expressed in Gr1^{int}CD48^{hi} cells (gate R2) indicating again that these cells are monocytes. Gr1^{vlo}CD48^{hi} (gate R3) cells show the highest expression of myeloblastosis viral oncogene homolog (Myb) and stem cell factor receptor Kit, indicating their immature state. Myb and Kit expression drops in Gr1^{vlo}CD48^{lo} (R4) cells suggesting that they are committed for neutrophil differentiation. Myb and Kit are almost absent in R2 and R1 cells, which is in agreement that these cells represent more mature monocytes (R2) and granulocytes (R1), respectively. Finally, the expression of Csf1r and Myb also distinguishes cells from gates R3 and R4, indicating that R3 cells but not R4 cells have monocytic potential (**Figure 6C**).

Human myeloblast versus neutrophil differentiation

The key features of human granulo-monocytic differentiation are also found in the murine sys-

tem. Hence mouse models can mimic to a certain extent specific human conditions in particular neutropenia and myelodysplastic syndromes (for review see [20, 21]), which are often associated with acute myeloid leukemia (AML) [22]. To test whether this is also true on a molecular basis, we compared our mouse expression data to data that were available for human cells. We observed that the changes in gene expression from $CD11b^{hi}Gr1^{lo}Gfi1^{hi}$ cells (gate A2 in **Figure 1**) to $CD11b^{hi}Gr1^{hi}Gfi1^{hi}$ granulocytes (gate B in **Figure 1**) correlated well with changes in gene expression that occur during differentiation from human myeloblasts to neutrophils (http://www.ncbi.nlm.nih. gov/gds/?term=GSE12837), [15] (**Figure 7**). Genes that were differentially expressed between immature and mature granulocytes in our murine dataset and in the human datasets as well were analyzed for enriched GO-categories using the DAVID online-tool [23, 24]. As in the murine system (**Figure 3**), immature cells show a higher expression of cell cycle associated genes (**Figure 7B**, right panel) and a lower expression of immune response genes necessary for normal granulocyte function (**Figure 7B**, left panel) suggesting that Gfi1 or CD48 expression are useful markers for neutrophil maturation and are relevant to study myeloid differentiation or myeloproliferative diseases in mouse models mimicking human conditions.

Discussion

In this study we report that the expression level of the transcription factor Gfi1 correlates with the potential of myeloid precursors to differentiate into either monocytes or granulocytes. Our data provide evidence that a subset of cells defined by expression of the markers Gr1 and CD11b ($Gr1^{lo}CD11b^{hi}$) contains two subpopulations that express high or low levels of Gfi1. Gfi1 low expressing cells are restricted to differentiate into monocytes and, vice versa, Gfi1 high expressing cells are determined to differentiate into granulocytes. This role of Gfi1 is supported by in vitro differentiation experiments and also by previous reports that indicate that Gfi1 deficiency leads to a lack of granulocytes and a surplus or accumulation of monocytes [7]. Moreover, we show that gene expression patterns of Gfi1 high and low expressing cellular subsets support this hypothesis. Gfi1 low expressing cells and Gfi1-KO monocytes show a

highly similar expression pattern and Gfi1 high expressing cells share more similarity with granulocytes with regard to their gene expression pattern.

Using the information that Gfi1 expression predicts the differentiation potential of precursors for either the monocytic or granulocytic lineage, we were able to identify CD48 as a new marker for the definition of monocytes, granulocytes and two novel precursor populations that are either bipotential and can give rise to monocytes and granulocytes or are already committed to the granulocytic lineage. A number of experimental findings corroborate this hypothesis. First, gene expression array analysis indicates that the expression of Gfi1 and CD48 is inversely correlated and that cells that express low levels of Gfi1 and are prone to develop into monocytes are expressing high levels of CD48. Conversely those cells that we identified to express high levels of Gfi1 and are able to differentiate into granulocytes express lower levels of CD48. Second, the analysis of bone marrow cells from Gfi1 deficient mice that are known to lack granulocytes and to accumulate monocytes lacked $Gr-1^{hi}CD48^{lo}$ cells and accumulated the $Gr1^{hi}CD48^{hi}$ population. Third, $Gr1^{lo}CD48^{lo}$ cells do neither express CCR2, nor Csf1r, typically expressed in monocytes, but show high expression levels of Myb and Kit, markers for immature cells, although lower than in $Gr1^{lo}CD48^{hi}$ bipotential cells. We conclude from these data that the marker CD11b can be replaced by CD48 to more clearly define monocytes and granulocytes and their precursors. Our data also suggest that this new marker combination may not only be useful in the analysis of murine myeloid differentiation but also for the analysis of mouse models for myeloproliferative diseases or myeloid leukemia.

Using the marker combination of CD48 and Gr1, we were able to define two precursor populations. One population, which is defined as $Gr1^{lo/int}CD48^{hi}$ and contains Gfi1 high and low expressing cells, is a bipotential precursor population for granulocytic and monocytic differentiation (termed here R3). This population can give rise to adherent cells and non-adherent cells in the presence of M-CSF or GM-CSF. In addition, we have defined another precursor population $Gr1^{lo/int}CD48^{lo}$ (Gfi1-high, termed

here R4) that has lost the potential to differentiate into monocytes and only maintains potential for granulocytic lineage, since under culture conditions where only M-CSF is present these cells do not divide and die, but give rise to non-adherent cells when GM-CSF is present. Our findings suggest that the expression of CD48, GR1 and Gfi1 can be used to define precursor populations for monocytic and granulocytic differentiation and that low CD48 and high Gfi1 levels correlate with commitment to the granulocytic lineage.

In humans, several mutations in proteins with different functions have been described that cause neutropenia (for review see [25, 26]), but Gfi1 is the only transcription factor known to cause neutropenia in man and mice when mutated [8] or knocked out by gene targeting [7]. Although Gfi1KO mice are severely neutropenic, it remained unclear at which state of development granulopoiesis is affected by loss of Gfi1 and what are the important factors regulated by Gfi1. The classical markers for monocytes/granulocytes CD11b and Gr1 do not distinguish (immature-) monocytes from immature granulocytes and for Gfi1 knock-out mice, they only indicated clearly the lack of mature neutrophils. Our finding that that the CD11b^{hi} Gr1^{lo} population is heterogeneous for Gfi1 expression and that on the transcriptional level CD11b^{hi}Gr1^{lo}Gfi1^{lo} cells show very few differences to CD11b^{hi}Gr1^{int}Gfi1$^{-/-}$ cells, while CD11b^{hi}Gr1^{lo}Gfi1^{hi} cells are more closely related to granulocytes indicated that that the CD11b^{hi}Gr1^{int} cells accumulating in Gfi1$^{-/-}$ mice are indeed monocytes. This is supported by our observations that all CD11b^{hi}Gr1^{lo}Gfi1^{lo} cells morphologically resemble monocytes, show a homogeneous low activity of the Gfi1-promoter and give rise to macrophages induced by GM-CSF in culture. The results of a hierarchical clustering of gene expression profiles also support this conclusion. Finally, CD11b^{hi}Gr1^{lo}Gfi1^{hi} cells morphologically are (meta-)myelocytes and can only give rise to neutrophils, but not mononcytes and therefore have all the features of immature granulocytes.

In human neutropenia in the context of autosomal dominant mutations in Gfi1, a remarkable accumulation of myeloid progenitors is observed [27] very similar to the situation observed in Gfi1 deficient mice. It is thus conceivable that this also represents and accumulation of monocytes as in Gfi1 deficient mice. Our observation that the changes in the transcriptional program between CD11b^{hi}Gr1^{lo}Gfi1^{hi} cells (termed here A2) and granulocytes largely reflect the changes of the transcriptional program of human myeloblasts differentiating to neutrophils supports these conclusions and indicates that our findings reported here are also relevant to human granulo-monocytic differentiation.

Contributions

L.V., U.D., C.K. and H.Z. did the experiments. L.V. did the analysis of the array data. L.V., H.Z. and T.M. have written the manuscript. T.M. has initiated the project, oversaw the interpretation and presentation of the data and provided the funding.

Disclosure statement

The authors have nothing to disclose.

Address correspondence to: Dr. Tarik Möröy, Institut de recherches cliniques de Montréal (IRCM), 110 des Pins West Avenue, Montréal, Québec, H2W 1R7, Canada. Tel: 1 (514) 9875501; Fax: 1 (514) 9875679; E-mail: Tarik.Moroy@ircm.qc.ca; Or: Dr. Hui Zeng, Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China. E-mail: zenghui@ccmu.edu.cn

References

[1]   Horwitz M, Benson KF, Person RE, Aprikyan AG and Dale DC. Mutations in ELA2, encoding neutrophil elastase, define a 21-day biological clock in cyclic haematopoiesis. Nat Genet 1999; 23: 433-436.

[2] Boztug K, Appaswamy G, Ashikov A, Schaffer AA, Salzer U, Diestelhorst J, Germeshausen M, Brandes G, Lee-Gossler J, Noyan F, Gatzke AK, Minkov M, Greil J, Kratz C, Petropoulou T, Pellier I, Bellanne-Chantelot C, Rezaei N, Monkemoller K, Irani-Hakimeh N, Bakker H, Gerardy-Schahn R, Zeidler C, Grimbacher B, Welte K and Klein C. A syndrome with congenital neutropenia and mutations in G6PC3. N Engl J Med 2009; 360: 32-43.

[3] Cheung YY, Kim SY, Yiu WH, Pan CJ, Jun HS, Ruef RA, Lee EJ, Westphal H, Mansfield BC and Chou JY. Impaired neutrophil activity and increased susceptibility to bacterial infection in mice lacking glucose-6-phosphatase-beta. J Clin Invest 2007; 117: 784-793.

[4] Lagresle-Peyrou C, Six EM, Picard C, Rieux-Laucat F, Michel V, Ditadi A, Demerens-de Chappedelaine C, Morillon E, Valensi F, Simon-Stoos KL, Mullikin JC, Noroski LM, Besse C, Wulffraat NM, Ferster A, Abecasis MM, Calvo F, Petit C, Candotti F, Abel L, Fischer A and Cavazzana-Calvo M. Human adenylate kinase 2 deficiency causes a profound hematopoietic defect associated with sensorineural deafness. Nat Genet 2009; 41: 106-111.

[5] Klein C, Grudzien M, Appaswamy G, Germeshausen M, Sandrock I, Schaffer AA, Rathinam C, Boztug K, Schwinzer B, Rezaei N, Bohn G, Melin M, Carlsson G, Fadeel B, Dahl N, Palmblad J, Henter JI, Zeidler C, Grimbacher B and Welte K. HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). Nat Genet 2007; 39: 86-92.

[6] Devriendt K, Kim AS, Mathijs G, Frints SG, Schwartz M, Van Den Oord JJ, Verhoef GE, Boogaerts MA, Fryns JP, You D, Rosen MK and Vandenberghe P. Constitutively activating mutation in WASP causes X-linked severe congenital neutropenia. Nat Genet 2001; 27: 313-317.

[7] Karsunky H, Zeng H, Schmidt T, Zevnik B, Kluge R, Schmid KW, Duhrsen U and Moroy T. Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. Nat Genet 2002; 30: 295-300.

[8] Person RE, Li FQ, Duan Z, Benson KF, Wechsler J, Papadaki HA, Eliopoulos G, Kaufman C, Bertolone SJ, Nakamoto B, Papayannopoulou T, Grimes HL and Horwitz M. Mutations in proto-oncogene GFI1 cause human neutropenia and target ELA2. Nat Genet 2003; 34: 308-312.

[9] Hock H, Hamblen MJ, Rooke HM, Traver D, Bronson RT, Cameron S and Orkin SH. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. Immunity 2003; 18: 109-120.

[10] Terstappen LW, Safford M and Loken MR. Flow cytometric analysis of human bone marrow. III.

[11] Elghetany MT, Ge Y, Patel J, Martinez J and Uhrova H. Flow cytometric study of neutrophilic granulopoiesis in normal bone marrow using an expanded panel of antibodies: correlation with morphologic assessments. J Clin Lab Anal 2004; 18: 36-41.

[12] Lagasse E and Weissman IL. Flow cytometric identification of murine neutrophils and monocytes. J Immunol Methods 1996; 197: 139-150.

[13] Yucel R, Kosan C, Heyd F and Moroy T. Gfi1:green fluorescent protein knock-in mutant reveals differential expression and autoregulation of the growth factor independence 1 (Gfi1) gene during lymphocyte development. J Biol Chem 2004; 279: 40906-40917.

[14] Theilgaard-Monch K, Jacobsen LC, Borup R, Rasmussen T, Bjerregaard MD, Nielsen FC, Cowland JB and Borregaard N. The transcriptional program of terminal granulocytic differentiation. Blood 2005; 105: 1785-1796.

[15] Coppe A, Ferrari F, Bisognin A, Danieli GA, Ferrari S, Bicciato S and Bortoluzzi S. Motif discovery in promoters of genes co-localized and co-expressed during myeloid cells differentiation. Nucleic Acids Res 2009; 37: 533-549.

[16] Emig D, Salomonis N, Baumbach J, Lengauer T, Conklin BR and Albrecht M. AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. Nucleic Acids Res 2010; 38: W755-762.

[17] Friedman AD. Transcriptional regulation of granulocyte and monocyte development. Oncogene 2002; 21: 3377-3390.

[18] Mui AL, Wakao H, O'Farrell AM, Harada N and Miyajima A. Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin-5 transduce signals through two STAT5 homologs. EMBO J 1995; 14: 1166-1175.

[19] Yokoyama S, Staunton D, Fisher R, Amiot M, Fortin JJ and Thorley-Lawson DA. Expression of the Blast-1 activation/adhesion molecule and its identification as CD48. J Immunol 1991; 146: 2192-2200.

[20] Wegrzyn J, Lam JC and Karsan A. Mouse models of myelodysplastic syndromes. Leuk Res 2011; 35: 853-862.

[21] Beachy SH and Aplan PD. Mouse models of myelodysplastic syndromes. Hematol Oncol Clin North Am 2010; 24: 361-375.

[22] Fortier JM and Graubert TA. Murine models of human acute myeloid leukemia. Cancer Treat Res 2010; 145: 183-196.

[23] Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.

Neutrophil maturation. Leukemia 1990; 4: 657-663.

[24] Huang da W, Sherman BT and Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009; 37: 1-13.

[25] Boztug K and Klein C. Genetic etiologies of severe congenital neutropenia. Curr Opin Pediatr 2011; 23: 21-26.

[26] Vandenberghe P and Beel K. Severe congenital neutropenia, a genetically heterogeneous disease group with an increased risk of AML/MDS. Pediatr Rep 2011; 3 Suppl 2: e9.

[27] Hock H, Hamblen MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y and Orkin SH. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature 2004; 431: 1002-1007.

Suppl Figure 1. Gfi1^{hi} CD11b⁺ Gr1^{lo} cells are more similar to granulocytes than Gfi1^{lo} CD11b⁺ Gr1^{lo} cells. A: Scatter plot comparisons of gene expression levels (log-2 of normalized signal intensities) between CD11b⁺Gr1ʰⁱ granulocytes (fraction B) and GFP/Gfi1ʰⁱCD11b⁺Gr1ˡᵒ (fraction A2) cells. B: As in (A) comparing CD11b⁺Gr1ʰⁱ granulocytes (fraction B) and GFP/Gfi1ˡᵒCD11b⁺Gr1ˡᵒ (fraction A1) cells. C: As in (A) comparing GFP/Gfi1ʰⁱCD11b⁺Gr1ˡᵒ (fraction A2) and Gfi1ᴷᴼCD11b⁺Gr1ⁱⁿᵗ (fraction C) cells. Blue lines indicate the borderlines of 2-fold change of probeset intensities of the compared samples. The percentage of probesets with a more than 2-fold differential expression between compared samples is given in the upper left corner of each plot. Some marker genes identified in this study (Gfi1, Vcam1, CD48) and genes typically expressed in monocytes / neutrophils (Mpo, Ela2, Csf1r) are indicated in the scatter plots.

1

Gfi1 regulates granulocytic cell fate
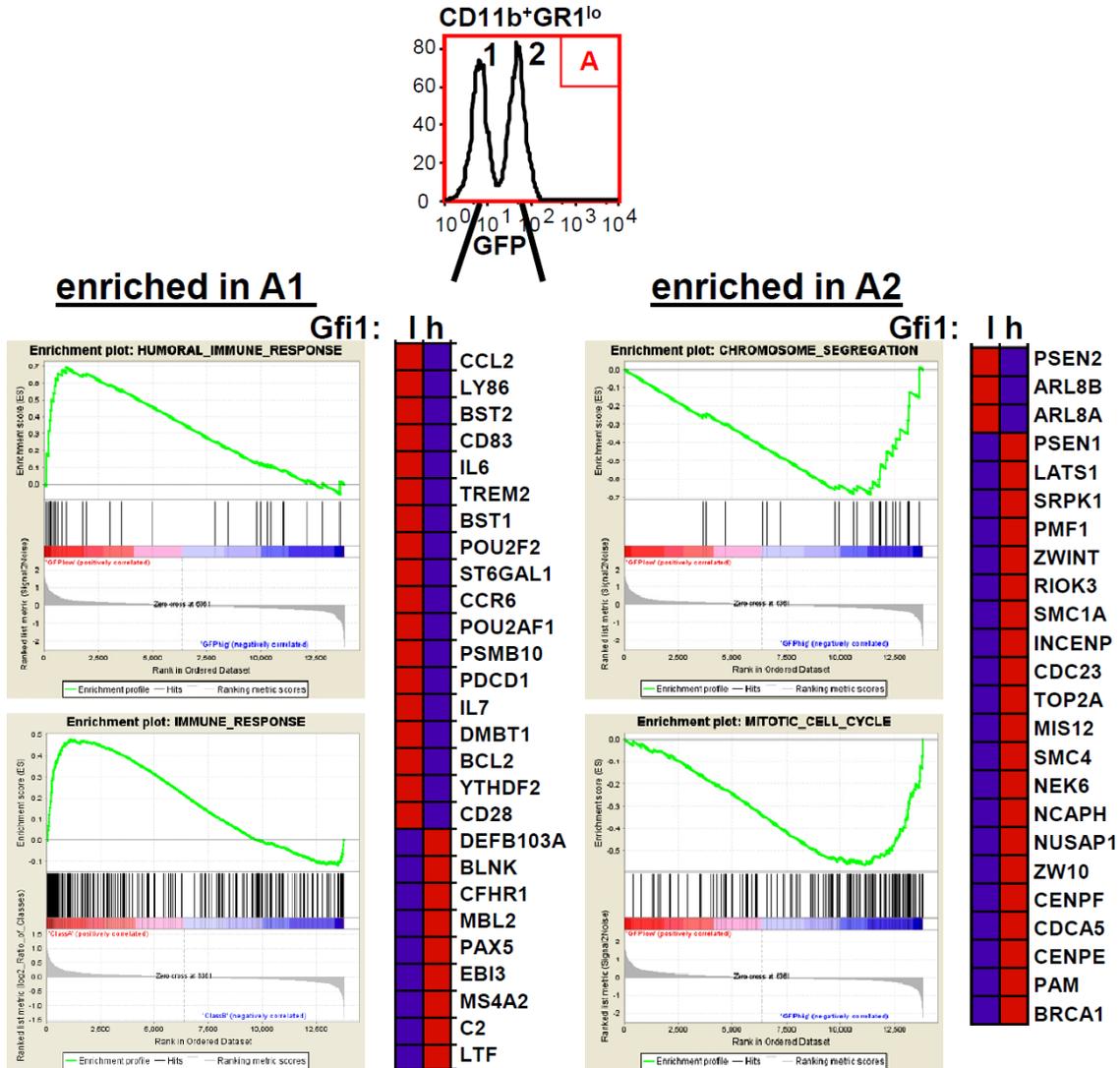
Suppl Figure 2. CD11b$^+$ Gr1^{lo} cell subsets separated by Gfi1 expression differ in primary immune response and cell cycle. GSEA analysis (enrichment plots) of gene-expression comparing Gfi1^{lo} (A1) and Gfi^{hi} (A2) CD11b$^+$ Gr1^{lo} bone marrow cells defines primary immune response genes overrepresented in Gfi1^{lo} cells (left panel). Cell cycle regulators are enriched in Gfi1^{hi} cells. Genes indicated with higher (red) or lower (blue) expression in A1 corresponding to "humoral immune response" (left panel) or "chromosome segregation" (right panel) are given.

A
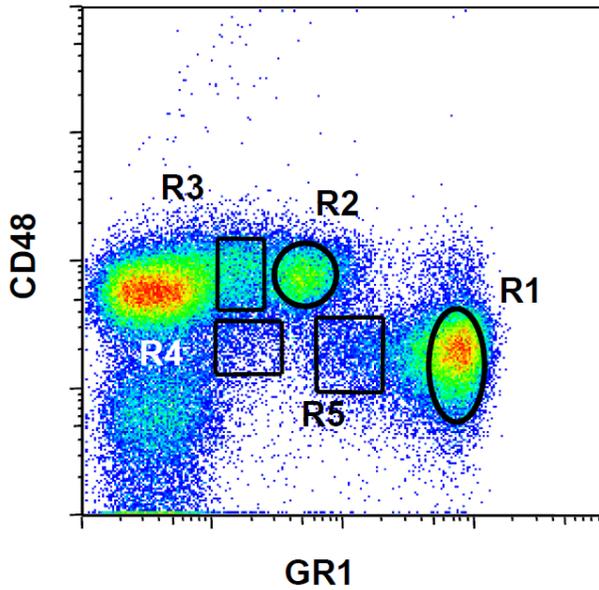
B
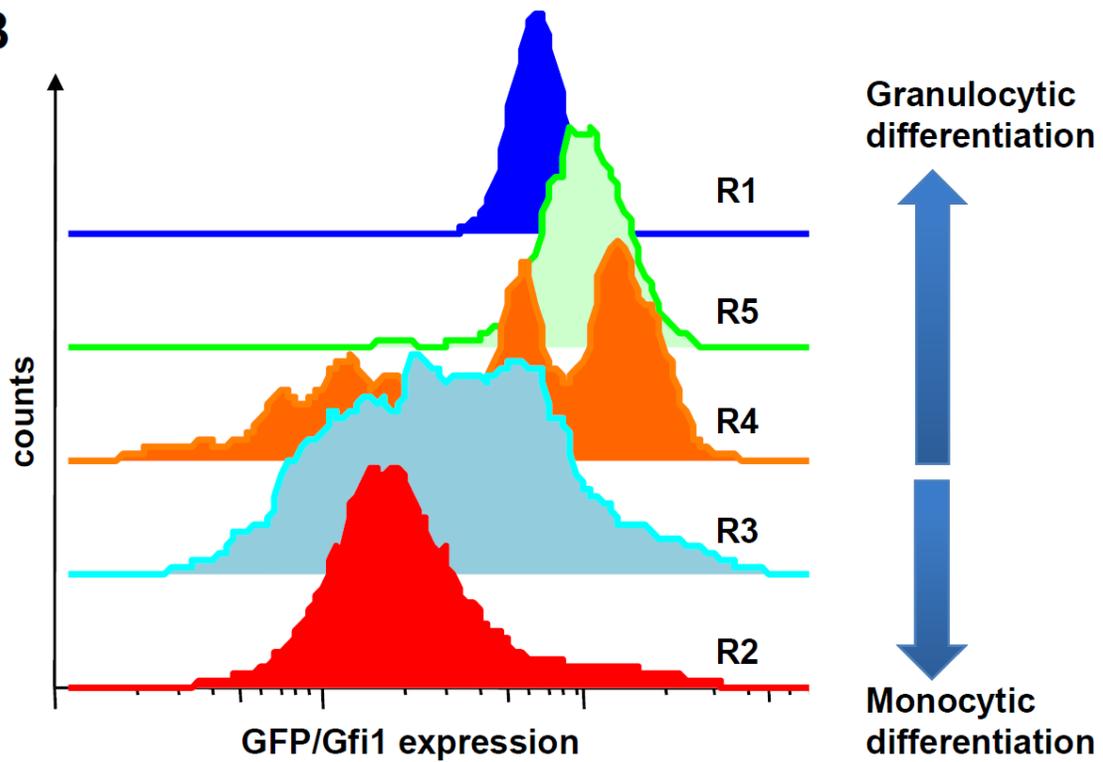
Suppl Figure 3. CD48 and Gr1 define cell subsets with differential Gfi1 expression. A: Bone marrow cells from heterozygous Gfi1:GFP knock-in mice were sorted for the surface markers Gr1 and CD48 by flow cytometry as indicated. B: Cells from the gates R1-R5 were analyzed for Gfi1:GFP expression levels. FACS plots are representative for at least three biological replicates.
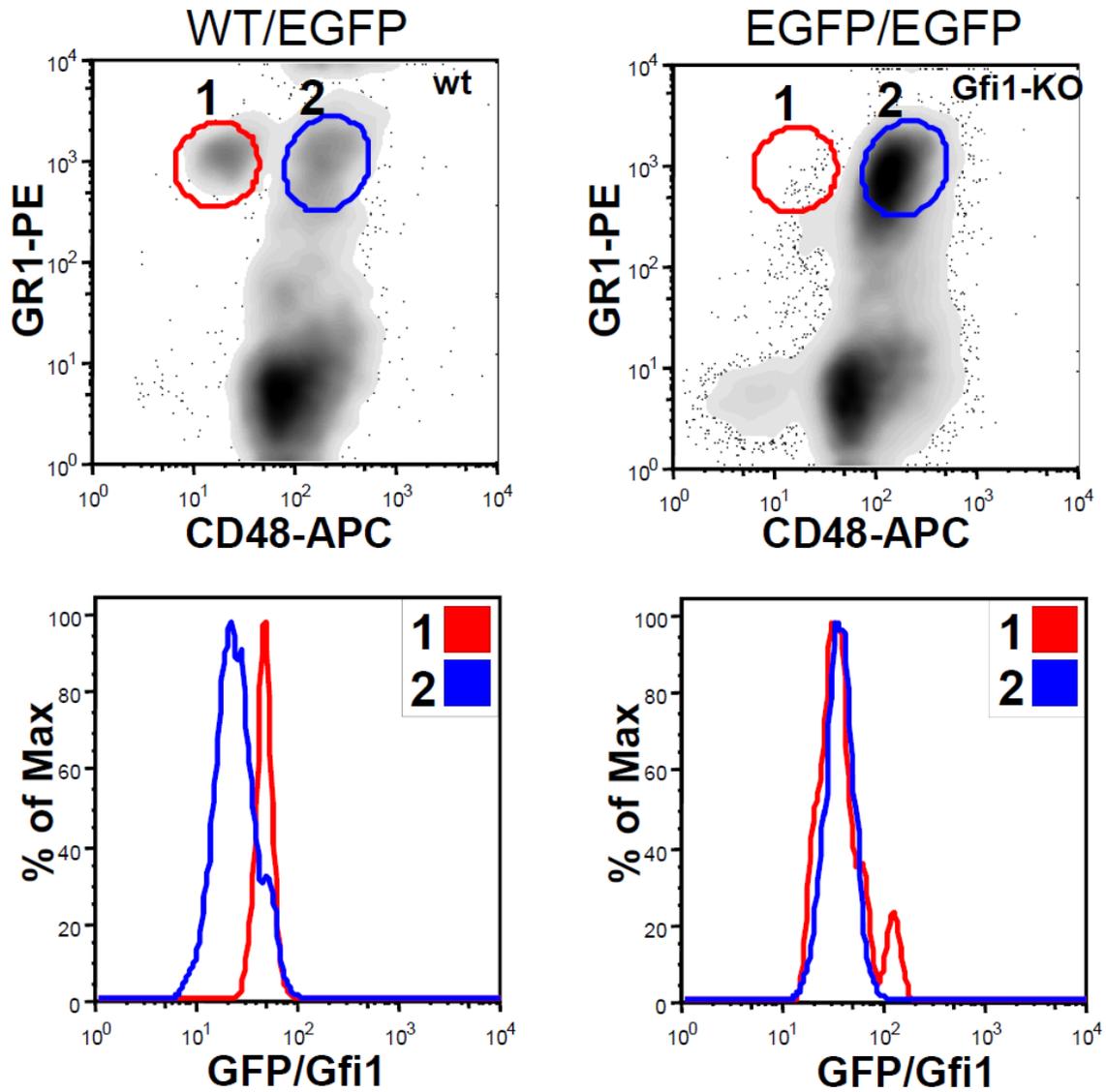
Suppl Figure 4. CD48hi Gr1hi CD11b positive cells accumulating in Gfi1KO-mice show a homogeneous Gfi1-promoter activity on monocyte level. Bone marrow cells from heterozygous (left) and homozygous (right) Gfi1:GFP knock-in mice were analyzed by flow cytometry. CD11b positive live cells were analyzed for the surface markers Gr1 and CD48 and gated for granulocytes (1) and monocytes (2) as indicated. Gated cells were then analyzed for GFP/Gfi1 expression revealing a homogeneous GFP expression for the Gfi1-KO cells in the monocytes gate (2).
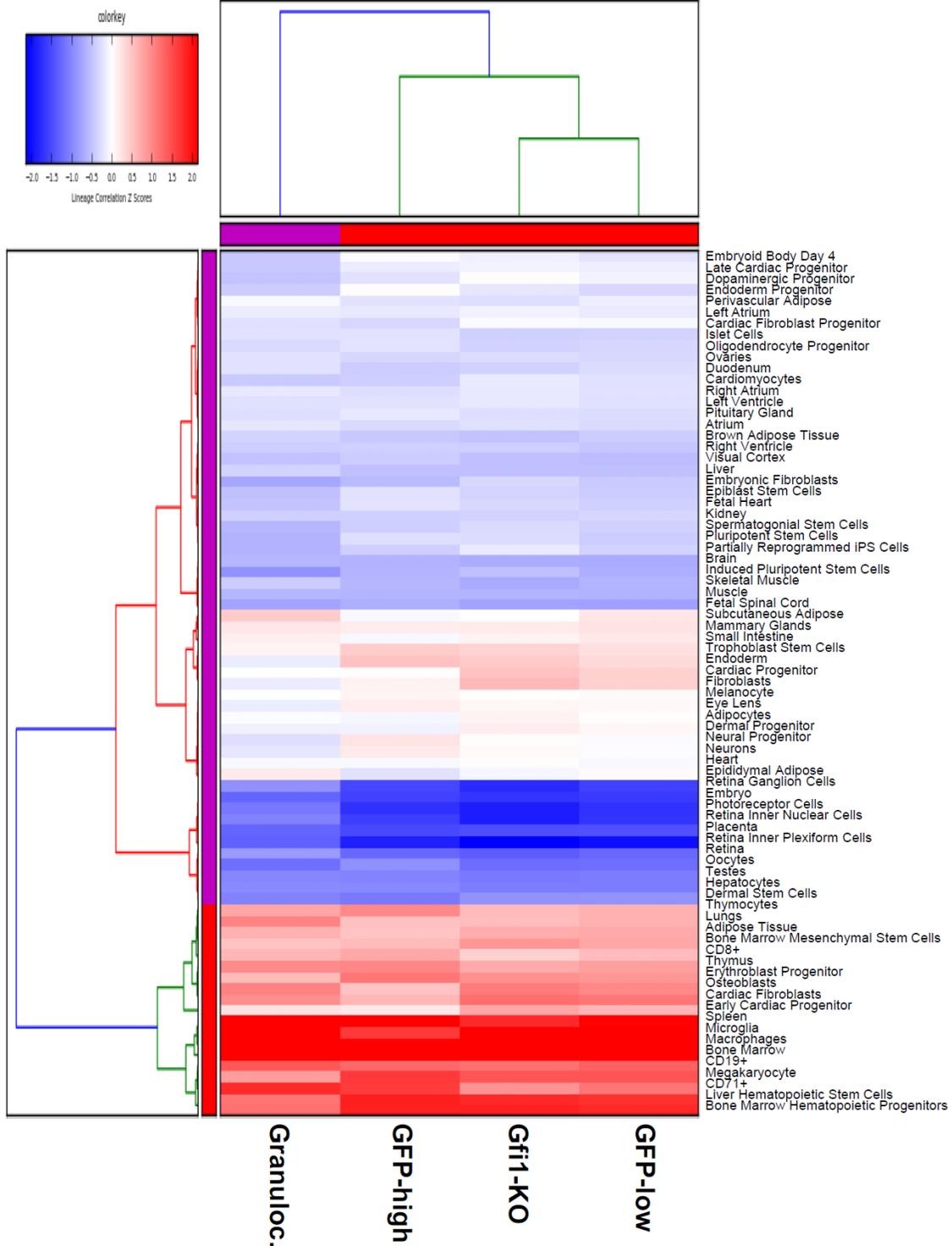
Suppl Figure 5. Gfi1lo CD11b+ Gr1lo cells are most closely related to Gfi1-KO monocytes in expression of lineage correlated protein coding RNA. Array data of the four sorted cellular subsets indicated were analyzed by clustering of lineage correlations of protein coding sequences using the AltAnalyze software tool (http://www.altanalyze.org). Expression levels of marker proteins for the cell-lineages indicated to the right were clustered and similarities are shown on top.

Array quality control

A

Biological Sample Expression Distribution - exp.Granuloc-GfiKO

Legend:
- (A) Gfi1-KI_GFPhigh-CHR3_430A2
- (B) Gfi1-KI_GFPlow-CHR3_430A2
- (C) Gfi1-KI_Mac1&Gr1_high-CHR3_430A2
- (D) Gfi1_ko_mono-CHR3_430A2

Y-axis: Number of Observations (0, 100, 200, 300, 400, 500, 600)

X-axis: Log2 Expression (x10-1 Precision) (2, 4, 6, 8, 10, 12, 14, 16)

B

Average Raw Intensity Signal - rma.report

Y-axis: Signal Intensity (0, 200, 400, 600, 800, 1000, 1200)

X-axis categories: Granuloc., GFP-low, GFP-high, Gfi1-KO

Suppl Figure 6. Quality control of array data. Biological sample distribution (A) and average raw intensity signal (RMA) analysis of the array data of sorted cells used in this publication prove the quality of the data acquired. Quality control was done using the AltAnalyze software tool (http://www.altanalyze.org).

Suppl Table 1. Gfi1^{lo} CD11b^{+} Gr1^{lo} cells differ from Gfi1KO monocytes mainly in activation response genes. Geneset enrichment analysis (GSEA) of genes that differ between Gfi1^{lo} CD11b^{+} Gr1^{lo} cells (monocytes) and cells from the "aberrant" Gfi1-KO monocytes fraction at least two fold in expression level, with a raw-p-value of 0.05 or smaller. Enriched genesets that fall into the category "biological process" were sorted for their z-score calculated by the AltAnalyze software toolbox

Ontology Name	# Chang.	# Measured	# in Ontology	% Changed	% Pres.	Z- Score	FisherExactP	gene symbols
cellular response to interferon-beta	8	18	20	44.444	90.0	17.2	0.00	Ifi203,Ifi204,Ifi205,Ifit1,Ifit3,Igtp,Pyhin1,Tgtp1,Tgtp2
cellular response to interferon-alpha	4	5	6	80.000	83.3	16.5	0.00	I830012O16Rik,Ifit1,Ifit2,Ifit3,Oas1g
antigen processing and presentation of peptide antigen	8	33	36	24.242	91.6	12.4	0.00	Cd74,Ctse,H2-Aa,H2-Ab1,H2-Eb1,H2-Q4,H2-Q6,H2-Q7
response to virus	16	123	160	13.008	76.8	12.3	0.00	Bst2,Cd40,Cxcl10,Ifi27l2a,Ifih1,Ifit1,Ifit2,Irf7,Isg15,Isg20,Itgax,Mx1,Oas1g,Oas3,Rsad2,Tgtp1,Tgtp2
positive regulation of cholesterol storage	3	6	6	50.000	100.0	11.2	0.00	Cd36,Lpl,Msr1
immune response	23	349	434	6.590	80.4	9.6	0.00	Axl,Bst2,Cd74,Cxcl10,Dhx58,H2-Aa,H2-Ab1,H2-Eb1,H2-Q4,H2Q6,H2Q7,Ifih1,Igj,Irf7,Irgm1,Mx1,Oas1a,Oas1g,Oas2,Oas3,Oasl1,Oasl2,Pglyrp1,Tgtp1,Tgtp2
positive regulation of macrophage derived foam cell differentiation	3	8	8	37.500	100.0	9.6	0.00	Cd36,Lpl,Msr1
response to stilbenoid	3	13	14	23.077	92.8	7.4	0.00	Cd36,Ifit3,Usp18
negative regulation of viral reproduction	3	15	16	20.000	93.7	6.8	0.00	Ifit1,Oas1a,Oas1g,Oas3
response to interferon-gamma	5	39	47	12.821	82.9	6.8	0.00	H2-Aa,H2-Ab1,Irgm2,Snca,Stat1
defense response to Gram-positive bacterium	4	39	48	10.256	81.2	5.3	0.00	Camp,Cd36,Klrk1,Pglyrp1
negative regulation of multi-organism process	3	24	26	12.500	92.3	5.2	0.00	Camp,Cd36,Ifit1
regulation of angiogenesis	7	118	132	5.932	89.3	4.8	0.00	Camp,Cd36,Cxcl10,Ets1,Flt1,Mmp9,Stat1
regulation of immune system process	18	550	636	3.273	86.4	4.7	0.00	Axl,Bcl2a1d,Cd274,Cd36,Cd40,Cd74,Cd83,Cxcl10,Dpp4,Ets1,H2-Aa,H2-Ab1,Hspa1b,Ifit1,Klrk1,Orm1,Pglyrp1,Snca
response to mechanical stimulus	7	127	135	5.512	94.0	4.6	0.00	Cd36,Cd40,Cxcl10,Ets1,Mmp9,Pkd2,Stat1
response to lipopolysaccharide	9	196	214	4.592	91.5	4.5	0.00	Ace,Cd36,Cmpk2,Cxcl10,Klrk1,Mmp9,S100a9,Snca,Stat1
neutral lipid metabolic process	4	51	61	7.843	83.6	4.4	0.00	Cd36,Lpin1,Lpl,Snca
RNA catabolic process	4	54	82	7.407	65.8	4.3	0.00	Hspa1a,Hspa1b,Isg20,Oas2
cellular response to oxidative stress	5	82	87	6.098	94.2	4.2	0.00	Axl,Cd36,Ets1,Hba-a1,Hba-a2,Snca
response to heat	4	62	73	6.452	84.9	3.9	0.00	Cxcl10,Hspa1a,Hspa1b,Mmp9

GSEA: GFP-lo vs Gfi1KO-Monocytes; upregulated fold 2, GO-term: biological process.