

RESEARCH ARTICLE

# Linking Hematopoietic Differentiation to Co-Expressed Sets of Pluripotency-Associated and Imprinted Genes and to Regulatory microRNA-Transcription Factor Motifs

Mohamed Hamed<sup>1,2</sup>, Johannes Trumm<sup>2</sup>, Christian Spaniol<sup>1</sup>, Riccha Sethi<sup>1</sup>, Mohammad R. Irhimeh<sup>3\*</sup>, Georg Fuellen<sup>2</sup>, Martina Paulsen<sup>4</sup>, Volkhard Helms<sup>1\*</sup>

**1** Center for Bioinformatics, Saarland University, Saarbrücken, Germany, **2** Institute for Biostatistics and Informatics in Medicine and Ageing Research, Rostock University Medical Center, Rostock, Germany, **3** Kolling Institute of Medical Research, Royal North Shore Hospital, The University of Sydney, St Leonards, NSW, Australia, **4** Department of Genetics, Saarland University, Saarbrücken, Germany

✉ Current address: CTTWA, Royal Perth Hospital and University of Western Australia, Perth, Australia  
\* [volkhard.helms@bioinformatik.uni-saarland.de](mailto:volkhard.helms@bioinformatik.uni-saarland.de)



OPEN ACCESS

**Citation:** Hamed M, Trumm J, Spaniol C, Sethi R, Irhimeh MR, Fuellen G, et al. (2017) Linking Hematopoietic Differentiation to Co-Expressed Sets of Pluripotency-Associated and Imprinted Genes and to Regulatory microRNA-Transcription Factor Motifs. PLoS ONE 12(1): e0166852. doi:10.1371/journal.pone.0166852

**Editor:** Bibekanand Mallick, National Institute of Technology Rourkela, INDIA

**Received:** May 3, 2016

**Accepted:** November 5, 2016

**Published:** January 4, 2017

**Copyright:** © 2017 Hamed et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Gene expression datasets can be found at GEO with the following Accession numbers and URL's: GSE6506: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6506>; GSE14833: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE14833>; GSE34723: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE34723>; GSE10246: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10246>. Annotation of miRNA-target genes and of transcription factors was done as described in

## Abstract

Maintenance of cell pluripotency, differentiation, and reprogramming are regulated by complex gene regulatory networks (GRNs) including monoallelically-expressed imprinted genes. Besides transcriptional control, epigenetic modifications and microRNAs contribute to cellular differentiation. As a model system for studying the capacity of cells to preserve their pluripotency state and the onset of differentiation and subsequent specialization, murine hematopoiesis was used and compared to embryonic stem cells (ESCs) as a control. Using published microarray data, the expression profiles of two sets of genes, pluripotent and imprinted, were compared to a third set of known hematopoietic genes. We found that more than half of the pluripotent and imprinted genes are clearly upregulated in ESCs but subsequently repressed during hematopoiesis. The remaining genes were either upregulated in hematopoietic progenitors or in differentiated blood cells. The three gene sets each consist of three similarly behaving gene groups with similar expression profiles in various lineages of the hematopoietic system as well as in ESCs. To explain this co-regulation behavior, we explored the transcriptional and post-transcriptional mechanisms of pluripotent and imprinted genes and their regulator/target miRNAs in six different hematopoietic lineages. Therewith, lineage-specific transcription factor (TF)-miRNA regulatory networks were generated and their topologies and functional impacts during hematopoiesis were analyzed. This led to the identification of TF-miRNA co-regulatory motifs, for which we validated the contribution to the cellular development of the corresponding lineage in terms of statistical significance and relevance to biological evidence. This analysis also identified key miRNAs and TFs/genes that might play important roles in the derived lineage networks. These molecular associations suggest new aspects of the cellular regulation of the onset of cellular differentiation and during hematopoiesis involving, on one hand, pluripotent genes that were previously not discussed in the context of hematopoiesis and, on the other hand, involve

Hamed, M., Spaniol, C., Nazarieh, M., Helms, V. (2015) Nucl. Acids Res. doi: [10.1093/nar/gkv418](https://doi.org/10.1093/nar/gkv418). Lists of imprinted genes were obtained from the 4 catalogs: 1) IGC database (<http://www.otago.ac.nz/IGC/>); 2) Geneimprint (<http://www.geneimprint.com/site/genes-by-species.Musmusculus>); 3) WAMIDEX (<https://atlas.genetics.kcl.ac.uk>) [17]; 4) MouseBookTM (<http://www.mousebook.org/catalog.php?catalog=imprinting>). Lists of pluripotency genes were taken from PluriNet, see Som et al, PLOS ONE 5: e15165. For data related inquiries, please contact the corresponding author.

**Funding:** Support was provided by Deutsche Forschungsgemeinschaft (SFB 1027), DAAD (graduate fellowship to MH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

genes that are related to genomic imprinting. These are new links between hematopoiesis and cellular differentiation and the important field of epigenetic modifications.

## Introduction

The maintenance of cellular pluripotency, the onset of differentiation as well as cellular differentiation into particular lineages appear to be controlled by tightly regulated gene regulatory networks (GRNs) that describe the interactions between transcription factors (TFs) and microRNAs and their target genes [1]. For example, Fuellen and co-workers have manually compiled from the original literature a dataset of murine genes termed the PluriNetwork that are involved in the regulation of the pluripotent state [2]. Besides transcriptional control, epigenetic modifications such as DNA methylation and histone marks are increasingly gaining attention with respect to cellular differentiation. One of the hallmarks of epigenetics is the phenomenon of genomic imprinting, which describes parent-of-origin monoallelic expression [3]. As the importance of epigenetic modes of gene regulation is particularly evident for imprinted genes, these genes serve as common model systems. Therefore, we are focusing here on the expression patterns and modes of regulation of the genes that have been shown to be monoallelically expressed in the mouse.

Hematopoiesis describes the differentiation of hematopoietic stem cells (HSCs) into lineage-committed progenitor cells. Recent transcriptomics studies have identified important parts of the regulatory networks that control maintenance of HSCs [4] and progenitors [1, 5, 6]. During hematopoiesis, little is known about the imprinting status of imprinted genes. As an exception to this, maternal imprinting at the H19-Igf2 locus was shown to maintain adult hematopoietic stem cell quiescence [7]. Additionally, several lines of evidence exist for the importance of imprinted genes during the transition from the stem cell stage to differential commitment as well as during particular cell lineages, namely hematopoiesis. For example, a network of 15 co-regulated imprinted genes involved in embryonic growth has been identified [8]. Ten of these genes were downregulated in terminally differentiated mouse cells compared to long-term repopulating HSCs [9]. In multipotent progenitor cells, 13 out of 15 imprinted genes were clearly downregulated compared to HSC whereas the two imprinted genes *Gnas* and *Gatm* were upregulated in MPP3 and MPP4 relative to MPP1 and HSC [5]. Recently, we identified 10 imprinted genes that are transcriptionally regulated by the hematopoiesis-related TF NFAT. We also found 9 imprinted genes that are targets of the *FOXO4* TF [10]. In CD34<sup>+</sup> cells, the imprinted maternally expressed gene *p57 (Cdkn1c)* was the only cyclin-dependent kinase inhibitor to be rapidly upregulated by TGF $\beta$ , a negative regulator of hematopoiesis [11]. Additionally, we found that promoter regions around the transcription start sites of *Mkrn3*, *Igf2*, and *Gnas* genes contain DNA motifs that match to annotated binding site motifs for the TFs *E2A* and *TCF3* [10]. The latter plays major roles in determining tissue-specific cell fate during embryogenesis such as early B-cell lymphopoiesis and germinal center B-cell development [12]. Several studies from the Li group indicated that the expression of certain imprinted genes changes in HSCs during differentiation from quiescent to multi-lineage progenitors [13]. However, the transcriptional and post-transcriptional activities as well as the interwoven regulatory mechanisms of imprinted genes and imprinting-related miRNAs (that are regulators/targets of imprinted genes) in the onset and further progression of cell differentiation (for the example of hematopoiesis) and the aspects of their involvement have not been addressed in detail before.

Identifying regulatory elements that are related to the differentiation of hematopoietic cells from the pluripotent stage into a lineage committed stage is of great importance to the fields of

stem cell biology and hematology. The role of imprinted and pluripotent genes in hematopoiesis as compared to embryonic stem cells (ESC) and their expression patterns was not previously analyzed or studied in a systemic manner. Therefore, we conducted this study to investigate the functional connections of sets of genes in hematopoiesis and differentiation. We constructed lineage-specific regulatory networks involving imprinted genes, pluripotent genes, related miRNAs, and central hubs that could act as lineage drivers. Our results were validated via various analytical approaches and their biological relevance was demonstrated in explaining the differentiation process of HSCs that is tightly regulated through a set of imprinted, pluripotent and hematopoiesis-specific genes, providing a better understanding of the regulatory mechanisms of cellular differentiation of blood cells.

## Materials and Methods

### Gene selection

Three mouse gene lists were prepared (imprinted, pluripotent, and hematopoietic genes). Our selection of imprinted genes was not done manually as in some of our previous work [10, 14, 15] as the manually curated lists contained a rather restricted number of genes. In this study, our selection was based on the overlap of several well-known online catalogs of imprinted genes. Imprinted genes were downloaded from four well-known catalogs [IGC database (<http://www.otago.ac.nz/IGC>) [16], Geneimprint (<http://www.geneimprint.com/site/genes-by-species>, Musmusculus), WAMIDEX (<https://atlas.genetics.kcl.ac.uk>) [17], and MouseBook™ (<http://www.mousebook.org/catalog.php?catalog=imprinting>)]. A single list of 120 genes (called henceforth candidate imprinted genes) was created from the four catalogs by including only genes that appeared in at least two catalogs and by filtering out genes that have conflicting or unknown imprinting status in the various catalogs (*i.e.* whether they are maternally or paternally expressed) (Table A in S1 File). As this is a computational study, we did not verify experimentally whether these genes are actually monoallelically expressed in the investigated cell lines or not.

The pluripotent gene list including 274 genes was obtained from the PluriNetWork [2], a hand curated pluripotency-controlling gene network in mouse with 574 regulatory interactions. To the best of our knowledge, no generally accepted GRN for the global hematopoiesis system has been established. In the absence of such a model, the 615 genes that are annotated in the Gene Ontology [18] for the GO term *hematopoietic or lymphoid organ development* (GO:0048534) were considered as ‘hematopoiesis genes’. Not all genes in the three gene lists were annotated in the Affymetrix array. Of the 120 imprinted genes only 86 were annotated (the rest were mostly non-coding RNAs, which are thus not considered), whereas only 2 out of 274 pluripotent genes and 53 out of the 615 hematopoietic genes were not annotated. The counts of overlapping genes are shown in Figure A in S1 File.

### Microarray analysis

Gene expression microarray data of three hematopoietic datasets (accession IDs GSE6506 [19], GSE14833 [20], GSE34723 [21]) and one non-hematopoietic dataset (control) (GSE10246 [22]) that also contains ESC samples, all based on the Affymetrix GeneChip Mouse Genome 430 2.0 Array, were downloaded from the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) [23]. Data normalization, model-based expression measurements, and annotation of the imprinted and pluripotent genes to their corresponding probes in the four datasets were done using the GC-RMA method and mouse 4302.db packages, respectively, by using the Bio-conductor software [24] within the statistical programming language R [25].

A gene expression similarity score was calculated to measure how similar the normalized expression of an individual gene present on the chip is to the distribution of normalized

expression values for the sets of pluripotent and hematopoietic genes, separately across the four datasets. Let us consider, for example, the similarity to pluripotent genes. In that case, the binned expression value of a gene  $g_i$  in each cell sample  $s_j$  was weighted by the number  $PD_{s_j}$  of pluripotent genes having expression values in the same bin in sample  $s_j$ . To reduce statistical noise, the GC-RMA normalized expression data was fitted by Gaussian curves. The respective values were then used as  $PD_{s_j}$  values.

This product was summed over all samples to give a representative similarity score for each gene. A detailed illustration is shown in the supplementary Figure B in [S1 File](#).

$$SimScore(g_i) = \sum_{j=1}^{cell\ samples} PD_{s_j} \times g_{i,s_j} \quad \text{with } i \in [1, \text{number of imprinted genes}] \\ \text{and } j \in [1, \text{all cell samples per dataset}]$$

Next, we separated the similarity scores of imprinted genes and non-imprinted genes and examined with the Mann-Whitney U-test whether imprinted genes have a higher gene expression similarity to pluripotent and hematopoietic genes than the background of all other genes. Additionally, we defined top scored genes as the highest 10% of the ranked genes and then applied the hyper-geometric test to investigate the significance of having imprinted genes among the defined top scored genes. We also tested the null distribution hypothesis by randomly shuffling the expression values of the imprinted genes and recalculating the similarity scores. The procedure was repeated 1000 times and the p-value was computed based on the number of random times in which the similarity scores were higher than the real score without shuffling.

For lineage specificity, six isometric lineages (three lymphoid and three myeloid) were constructed from the four expression datasets by following the hematopoietic differentiation model in Seita *et al.*, [21], (Table 1). We looked at three main hematopoiesis developmental stages: early progenitors (LTHSC and STHSC), intermediate progenitors (LMPP and CLP), and terminally differentiated blood cells (MKE and GM). Then we used a conservative

**Table 1. Selected hematopoietic lineages and their stages of sequential cell development.**

Lineage	Sequential Cell Development
<b>B-cell</b>	$HSC \rightarrow MPPa \rightarrow MPPb \rightarrow GMLPa \rightarrow GMLPb \rightarrow CLP \rightarrow BLP \rightarrow PREPROB \rightarrow FrB \rightarrow FrC \rightarrow FrD \rightarrow FrE$
<b>NK-cell</b>	$HSC \rightarrow MPPa \rightarrow MPPb \rightarrow GMLPa \rightarrow GMLPb \rightarrow CLP \rightarrow iNK \rightarrow mNK$
<b>T-cell</b>	$HSC \rightarrow MPPa \rightarrow MPPb \rightarrow GMLPa \rightarrow GMLPb \rightarrow CLP \rightarrow DN1 \rightarrow DN2 \rightarrow DN3a \rightarrow DN3b \rightarrow DN4 \rightarrow DPCD69^+ \rightarrow DPCD69^+ \rightarrow CD4^+ CD69^+ \rightarrow CD4^+ CD69^+$
<b>Erythrocytes</b>	$HSC \rightarrow MPPa \rightarrow pMEP \rightarrow MEP \rightarrow pCFU-E$
<b>Monocytes</b>	$HSC \rightarrow MPPa \rightarrow sCMP \rightarrow pGMPa \rightarrow pGMPb \rightarrow GMP \rightarrow Mono$
<b>Granulocytes</b>	$HSC \rightarrow MPPa \rightarrow sCMP \rightarrow pGMPa \rightarrow pGMPb \rightarrow GMP \rightarrow Gra$

*HSC*, Hematopoietic Stem Cells; *MPPa*, Multipotent Progenitor state A; *MPPb*, Multipotent Progenitor state B; *GMLPa*, Granulocyte Macrophage Lymphoid Progenitor state A; *GMLPb*, Granulocyte-Macrophage-Lymphoid Progenitor state B; *CLP*, Common Lymphoid Progenitor; *BLP*, Earliest B Lymphoid Progenitor; *PREPROB*, Precursor of B-cells Progenitor; *FrB*, Fraction B B-cell; *FrC*, Fraction C B-cell; *FrD*, Fraction D B-cell; *FrE*, Fraction E B-cell; *iNK*, intermediate Natural Killer Cell; *mNK*, mature Natural Killer Cell; *DN1*, Double Negative T-cell 1; *DN2*, Double Negative T-cell 2; *DN3a*, Double Negative T-cell 3a; *DN3b*, Double Negative T-cell 3b; *DN4*, Double Negative T-cell 4; *DPCD69<sup>+</sup>*, Double Positive CD69<sup>+</sup> T-cell; *DPCD69<sup>+</sup>*, Double Positive CD69<sup>+</sup> T-cell; *CD4<sup>+</sup>CD69<sup>+</sup>*, CD4<sup>+</sup> CD69<sup>+</sup> T-cell; *CD4<sup>+</sup>CD69<sup>+</sup>*, CD4<sup>+</sup> CD69<sup>+</sup> T-cell; *MEP*, Megakaryocyte/ Erythrocyte Progenitor; *pMEP*, pre of *MEP*; *pCFU-E*, pre of *CFU-E*; *sCMP*, Strict Common Myeloid Progenitor; *GMP*, Granulocyte-Macrophage-Progenitor; *pGMPa*, pre of GMP state A; *pGMPb*, pre of GMP state B; *Mono*, Monocytes; *Gra*, Granulocytes.

doi:10.1371/journal.pone.0166852.t001

differential expression approach based on moderated t-test to encode the differences between the three stages. P-values were adjusted using the Benjamini-Hochberg procedure [26] to limit the false discovery rate to 5%. Lineages that are constructed from dataset GSE6506 and contain only two stages (early progenitors and terminally differentiated cells) were analyzed by setting an on/off expression threshold (similar to Chambers *et al.*, [19]) to identify uniquely expressed genes in each stage of the cell development of each lineage. Finally, a gene was confirmed as differentially expressed gene if it appeared in the same lineage in at least two different datasets.

## miRNA Enrichment

We predicted the miRNAs associated with the list of differentially expressed mRNAs of each lineage by determining the set of miRNAs whose target genes as well as regulator TFs are significantly enriched within each differentially expressed gene set. For this, we used the hypergeometric distribution function followed by the Benjamini-Hochberg adjustment with a cutoff value of 0.001. The TF → miRNA interactions and miRNA → target gene interactions were compiled from the regulatory databases listed in [27].

## Construction of TF-miRNA regulatory networks and motif modules

The integrated association between the lineage markers and the enriched miRNAs were compiled from the regulatory databases of the murine TFmiR webserver [27]. All molecular interactions that are supported either by experimental and/or by predicted evidences were considered in this analysis. Next, 3-node FFL motifs (TF-miRNA co-regulatory networks consisting of a miRNA, a TF, and a joint target gene) were characterized using the computational procedure described in [27]. To identify critical network players, we computed the degree centrality measure (number of links incident upon a node) for each lineage network using the R package igraph [28] and highlighted the top 10% highest centrality nodes of the TFs/genes and miRNAs. The networks and co-regulatory motifs were visualized with Cytoscape V3.3.0 [29].

## Network validation and assessment of key player nodes

**1- Statistical validation: significance of the detected motifs.** To evaluate the significance of each co-regulatory motif, we compared how often it appears in the lineage regulatory network against the number of times it appears in randomized versions of these networks with preserved node degrees. The randomization procedure is explained in detail in [27]. The random networks were constructed  $N_r = 100$  times and compared to the real network. The p-value was calculated as

$$p - value = \frac{N_h}{N_r}$$

where  $N_h$  is the number of times that a certain motif type is identified in a randomized network at equal or higher number than in the lineage network. Only motifs having p-value  $< 0.05$  were considered for further analysis.

**2- Semantic validation: functional homogeneity within the motif nodes.** In order to assess the biological relevance of the identified TF-miRNA co-regulatory motifs and to better understand their functional roles, we analyzed the GO semantic similarity for all pairs of target genes regulated either by the TF or by the miRNA of each lineage-specific TF-miRNA motif. The GoSemSim R package [30] was used to calculate the semantic similarity scores according to the Gene Ontology (GO) annotations. Statistical significance was determined by randomly selecting the same number of co-regulated genes (genes targeted by either the TFs or the miRNA) from all Entrez genes with GO annotations, and computing their similarity

scores. The permutation procedure was repeated 100 times. Then, we utilized the Kolmogorov-Smirnov test to check whether the functional similarity scores of all gene pairs composing a regulatory motif are significantly higher than that of randomly selected pairs.

**3- Enrichment Analysis for genes and miRNAs.** The functional annotation tool in DAVID [31] was used to identify significantly enriched functional categories in the gene sets. For this, we determined which GO categories were annotated to at least 2 genes and are statistically overrepresented in the co-expressed genes against the full mouse genome (control) as previously shown in [10]. Then, enrichment was evaluated through the Fisher Exact test using p-value threshold of 0.05. Enrichment analysis of the miRNAs sets was performed using the TAM online tool [32] as previously done in [33].

**4- Literature support.** We used the PubMed database to search for evidence-based hematopoietic lineage commitment and cellular differentiation. Network nodes (TFs/genes and miRNAs) with such evidence supported by at least one published article are highlighted in the network visualization.

## Results and Discussion

In this study, we re-analyzed published gene expression microarray data deposited in GEO [23] [three hematopoietic datasets (accession IDs GSE6506 [19], GSE14833 [20], GSE34723 [21]) and one data set that is not specific to hematopoiesis (control) (GSE10246 [22])]. This was done in order to identify regulatory elements and their transcriptional and post-transcriptional interactions that are related to the transition of cells from the pluripotent stem cell stage into the lineage-committed stage. As explained in the methods section, we established three lists of imprinted, pluripotent, and hematopoietic genes. It is important to note that some genes in the pluripotent list might be directly involved in maintaining the pluripotency of embryonic stem (ES) cells and induced pluripotent stem cells (iPS), whereas some other genes might have indirect and more general functionalities, such as cell cycle regulation etc. From these lists, 86 imprinted, 272 pluripotent, and 562 hematopoietic genes are annotated on the microarray.

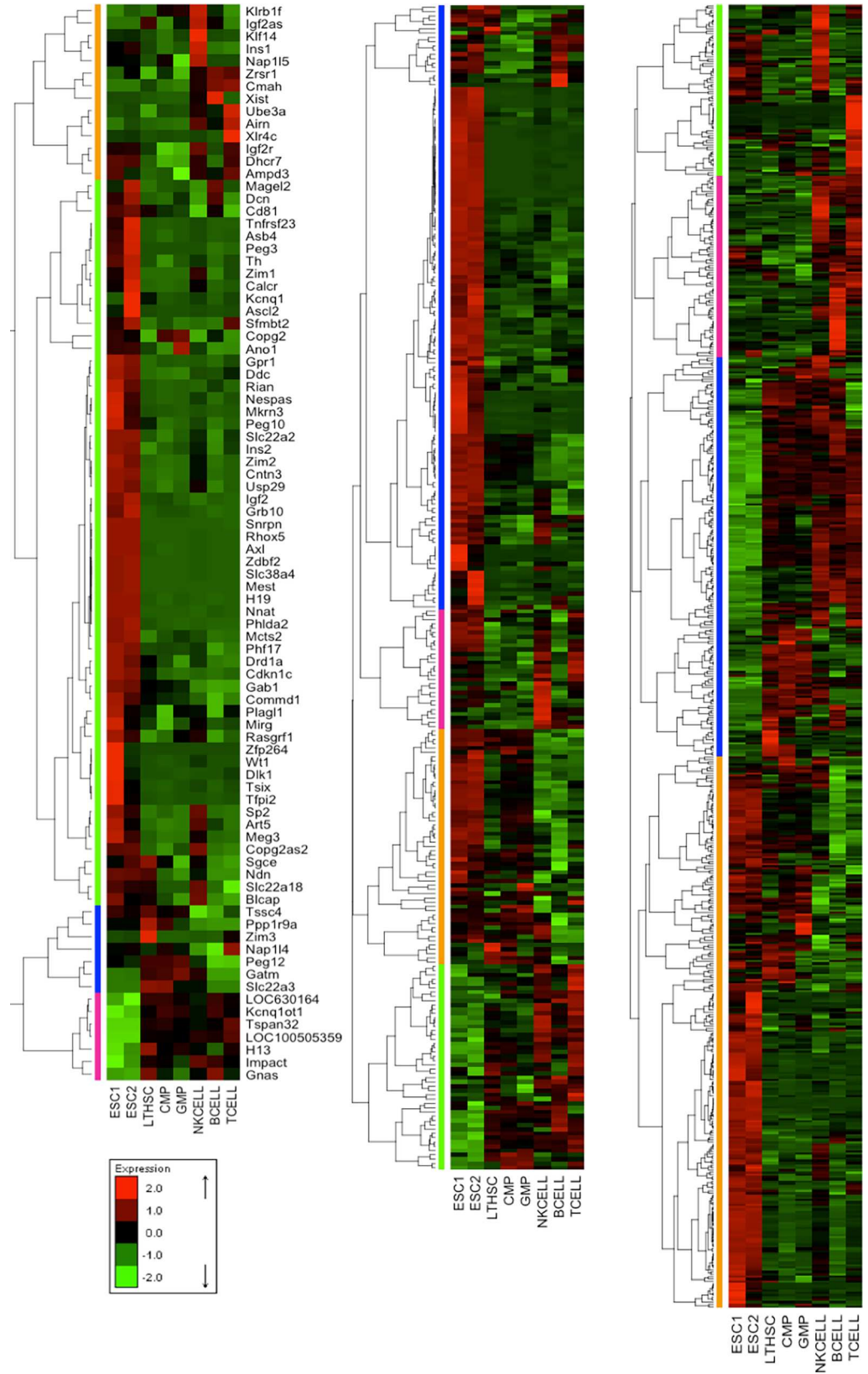
We start with a compact overview of the different stages of our computational analysis. First, the expression patterns of the imprinted and pluripotent genes during different stages of the hematopoietic lineages were compared to the global expression patterns of hematopoietic genes. We identified marker genes that are exclusively differentially expressed in each of the six main hematopoietic lineages and we postulated their functional roles by scrutinizing the mammalian phenotypes associated with hematopoietic abnormalities. Then, we predicted for each lineage marker gene set the miRNAs with significantly enriched target genes and significantly enriched regulatory TFs of the marker genes. Using regulatory data sources, we constructed lineage-specific 'TF-miRNA-mediated regulatory networks'. In these networks, we identified 3-node Feed Forward Loop (FFL) motifs consisting of imprinted/pluripotent TFs and co-targeted genes as well as imprinting/pluripotency-associated miRNAs. Finally, we validated/tested the contribution of these TF-miRNA co-regulatory motifs to the transcriptional activities of the corresponding hematopoietic lineage in terms of statistical significance and their association with biological evidence. It turned out that many miRNAs and TFs/genes of the identified lineage-specific co-regulatory motifs were implicated by previous work in the sequential cell development of either the related hematopoietic lineage or hematopoiesis in general. Our results demonstrate that the constructed lineage-specific mediated regulatory networks (comprising imprinting- and pluripotency-associated TFs and miRNAs) contain valuable new information for researchers that may aid in identifying critical TFs, miRNAs, and their targets for further experimental design, and enhance the understanding of the regulatory mechanisms of hematopoietic cellular differentiation.

## Imprinted genes show similar expression patterns to pluripotent and hematopoietic genes

To provide an overview, [Fig 1](#) shows clustered normalized expression profiles for two ESC lines, three progenitor cell lines (Long Term HSC: LTHSC, Common Myeloid Progenitor: CMP, and Granulocyte-Macrophage-Progenitor: GMP), and three terminally differentiated cell lines (NK-cells, B-cells, T-cells). Clustering as well as visual inspection revealed three main classes of expression patterns, which are shared by most imprinted, pluripotent and hematopoietic genes. The first class (1<sup>st</sup>) contains genes that have high expression levels in ESC and have gradually decreasing expression levels during the two stages of hematopoiesis (early and intermediate progenitors and terminally differentiated blood cells). More than half of the imprinted genes ([Fig 1](#), left panel, green) and of the pluripotent genes ([Fig 1](#), middle panel, blue) belong to this class (1<sup>st</sup>). Also, about one third of the hematopoiesis genes ([Fig 1](#), right panel, orange) belong to the same 1<sup>st</sup> class. Genes of the second class (2<sup>nd</sup>) are characterized by overexpression in the early and intermediate progenitors (more specifically in Common Lymphoid Progenitor: *CLP*) and have relatively low expression levels in both ESC and terminally differentiated cells. The third class (3<sup>rd</sup>) includes genes that are predominantly upregulated in matured blood cells. Interestingly, the 2<sup>nd</sup> and 3<sup>rd</sup> classes contain genes from all three gene sets. On the basis of gene ontology (GO) annotation, we investigated the functional similarity of the three genes sets among each other and with respect to randomly selected genes (Figure C in [S1 File](#)). This analysis revealed that pluripotent genes and hematopoietic genes share the highest similarity of GO annotation (about 0.75). This is quite expected since the genes from both sets are involved in regulating cell fate. Comparing the functional similarity of the pluripotent genes of the 1<sup>st</sup> class to the hematopoiesis genes of the 1<sup>st</sup> class resulted in slightly higher numbers than measuring the functional similarity between mixed classes. In comparison, the average functional similarity of imprinted genes with pluripotent genes or with hematopoiesis genes was lower (about 0.6), but still clearly higher than that with randomly selected genes (about 0.3).

To quantify this visual impression, we ranked all genes according to their gene expression similarity score (see [methods](#) section) across all considered hematopoietic samples. This score measures how similar the expression of an individual gene is to the distribution of expression values for the sets of pluripotent and hematopoietic genes. Notably, all p-values for the three hematopoietic datasets (that encompass differentiation and cell development data only) were significant (between 0.001 and 0.01) ([Table 2](#)). Moreover, a large portion of imprinted genes was found to belong to the highest 10% of the ranked genes in GSE6506 and GSE34723 datasets (66% and 59%, respectively, [Table 2](#)). In contrast, no significant difference was found between the ranking of imprinted genes and the background of all genes of the control (GSE10246; largely non-hematopoietic) and the number of top ranked imprinted genes was lowest here. Many imprinted genes shared very similar expression patterns with the pluripotent and hematopoietic gene sets ([Fig 1](#), [Table 2](#)) similar to observations made for a smaller set of imprinted genes in murine HSCs [9]. This similarity was most pronounced in 1<sup>st</sup> class genes that are overexpressed in ESCs. However, the functional similarity of imprinted genes, pluripotent, and hematopoietic genes does not quite reach the level that is seen for the similarity between PluriNetWork and hematopoiesis genes. We attribute the observation that particular imprinted genes show a high variability of expression among the various stages of differentiation to the different roles played by these genes in the cell.

Comparing the expression levels of known pluripotent with hematopoietic genes showed that the compiled PluriNetWork [2] contains not only the GRN that keeps cells in the pluripotent state but appears also to be related to the regulation of the onset of cellular differentiation





**Fig 1. Heatmaps showing transient changes in expression profiles.** Different groups of ESC and hematopoietic cells (e.g stem cells, intermediate progenitors, and terminally differentiated blood cells) from the GSE10246 dataset for (left panel) imprinted genes, (middle panel) pluripotent genes and (right panel) hematopoietic genes were compared. Green spots represent downregulated genes, and red spots represent upregulated genes. The order of genes is obtained by hierarchical clustering, which shows three similar pattern classes between imprinted, pluripotent and hematopoietic genes.

doi:10.1371/journal.pone.0166852.g001

such as hematopoiesis. In fact, the GO terms *hematopoietic or lymphoid organ development*, *haemopoiesis*, *myeloid cell differentiation*, *leukocyte differentiation* are annotated to 44, 39, 27, and 23 pluripotent genes, respectively, suggesting that a significant portion of the pluripotent genes is indeed involved in hematopoiesis regulation (Table C in S1 File). Moreover, Figs 1 and 2 demonstrate convincingly that the full set of pluripotent genes displayed pronounced variations during different stages of hematopoiesis and in individual cell lineages as well. These findings agree with previous studies that discussed the role of pluripotent genes in determining cell fate and controlling differentiation [34].

### All three gene sets contribute to hematopoietic lineage specificity

With a focus on the differentiation along particular lineages, we subjected the selected microarray datasets to differential expression analysis. For this, we divided cell samples into three classes [early progenitors (e.g HSC and MPP), intermediate progenitors (e.g GMLP and CLP), and terminally differentiated blood cells (e.g Monocytes and NK-cells)] so that this analysis was now based on far more cell types than the global analysis in Fig 1. Table B in S1 File lists those imprinted, pluripotent, and hematopoietic genes having statistically significantly different expression patterns during the sequential stages of specific blood cell lineages.

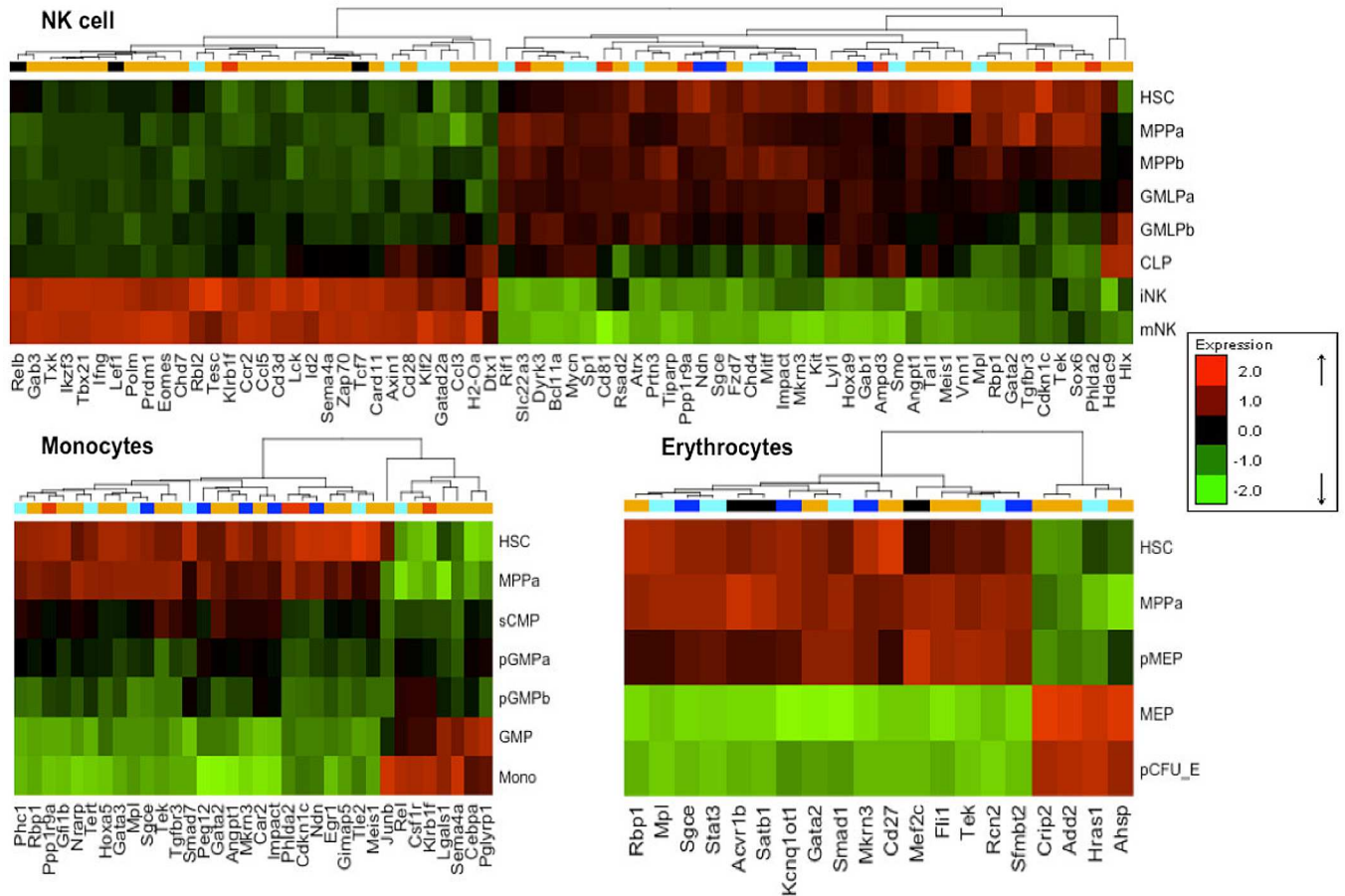
Lineage-specific differentially expressed genes (here termed marker genes) found in the three gene sets and the related expression heatmaps for NK-cells, monocytes, and erythrocytes are shown in Fig 2 (heatmaps for the other 3 lineages are shown in Figure D in S1 File). The number of significant lineage markers varies between 23 genes (in granulocytes) and 193 genes (in B-cells). Only the three genes *Rbp1*, *Sgce* and *Mkrn3* (the last two are imprinted

**Table 2. Statistical comparison of gene expression similarity (gene similarity scores).**

Dataset	Compared genes to background	Mann-Whitney U-Test	Permutation Test	Top Scored Imprinted Genes	Hyper-geometric Test
GSE6506 [19]	Pluripotent	0.006	0.019	55	0.006
	Hematopoietic	0.044	0.032	57	0.010
GSE34723 [21]	Pluripotent	0.004	0.022	50	0.004
	Hematopoietic	0.003	0.006	51	0.009
GSE14833 [20]	Pluripotent	0.003	0.014	18	0.195 *
	Hematopoietic	0.006	0.072 *	24	0.214 *
GSE10246 [22] (Control)	Pluripotent	0.106	0.267	11	0.784
	Hematopoietic	0.101	0.089	14	0.700

First, the Mann-Whitney U-test was used to test whether imprinted genes have a higher gene expression similarity to pluripotent and hematopoietic genes than to the background of all other genes (non-imprinted genes). Then the top 10% scoring genes were tested using hyper-geometric test to find out the significance of having imprinted genes among them (rightmost column). Secondly, also a permutation test was performed to test the null distribution hypothesis by randomly shuffling the expression values of the imprinted genes and recalculating the similarity scores. The procedure was repeated 1000 times. The p-value was computed based on the number of random times where the similarity scores were higher than the real score without shuffling. (\*) Among the three hematopoietic datasets, only the p-value of the hyper-geometric test for GSE14833 and the respective permutation test does not meet the significance threshold of 0.05.

doi:10.1371/journal.pone.0166852.t002



**Fig 2. Heatmaps of differentially expressed imprinted genes.** The order of genes is obtained by hierarchical clustering of three blood lineages (NK-cells, Monocytes, and Erythrocytes) based on the GSE34723 dataset. Gene clustering color coding is (blue) for paternally expressed genes, (red) for maternally expressed, (cyan) for pluripotent genes, and (orange) for hematopoietic genes. Shared genes between the pluripotent and hematopoietic gene sets are marked in black. Green spots represent downregulated genes, and red spots represent upregulated genes. The clustering reveals that for every lineage, there exist imprinted as well pluripotent and hematopoietic genes showing similar expression changes during cell development. The other three lineages (B-cells, T-cells, and granulocytes) are shown in the supplementary Figure D in [S1 File](#).

doi:10.1371/journal.pone.0166852.g002

genes) are shared by all myeloid branches (erythrocytes, monocytes, and granulocytes) (Table B in [S1 File](#)). A recent study found that the expression of *Rbp1* was affected significantly by the medical indication of acute myeloid leukemia (AML) [35]. Additionally, we identified 16 marker genes (e.g. *Lgals1*, *Gimap5*, *Pml*, and *Hoxa5*) that are exclusively differentially expressed in myeloid lineages (not in lymphoid). These 16 genes are annotated for terms such as GO:0002317 "plasma cell differentiation", GO:0043011 "myeloid dendritic cell differentiation", GO:0030099 "myeloid cell differentiation", and GO:0045639 "positive regulation of myeloid cell differentiation", respectively. Along the same lines, the lymphoid markers contain 30 genes shared by all lymphoid peers (B-cell, T-cell, and NK-cell) and 226 genes that were only detected for individual lymphoid lineages (not myeloid) such as *Tcf7*, *Lef1*, and *Rel*. The latter gene plays a role in differentiation and lymphopoiesis [36]. A complete list of the identified genes is given in Table B in [S1 File](#).

Different cell types showed pronounced differences in their gene expression profiles: most prominent was the high number of differentially expressed genes in B-cells and T-cells with a major contribution of pluripotent genes and to some extent also imprinted genes. This expression profile is interesting as a substantial portion of B-cells and T-cells serve as memory cells

that can be induced by secondary infections to undergo further cell divisions [37, 38]. NK-cells that have recently been shown to have some potential for further cell divisions [39] tend to have more differentially expressed genes compared to the differentiated myeloids. Many of the lineage markers identified in this work are concordant with the findings of previous work. Generally, the ten lineage markers (*Cdkn1c*, *Ndn*, *Gatm*, *Phlda2*, *Air*, *Igf2r*, *Slc22a3*, *H13*, *Sfmbt2*, and *Peg12*) that participate in most lineages were demonstrated to be differentially expressed in the early onset of the hematopoietic process [7, 40]. More specifically, the identified erythrocytes lineage markers *Fli1*, *Mpl*, and *Gata2* were previously found to determine the erythrocytes signature [5, 41–43].

Separate labeling of maternally and paternally expressed genes did not reveal a clear-cut separation, which is consistent with previous findings [10]. Nevertheless, only paternally expressed genes were differentially expressed in the erythrocyte lineage (Fig 2). In contrast, the imprinted genes that were overexpressed during late stages of hematopoiesis tend to be maternally expressed (e.g. *Cmah* and *Nap1l4* in B- and T-cells, *Klrb1f* in monocytes and NK-cells, *Th* and *Igf2r* in T-cells) rather than paternally expressed (*Sp2*, *Mcts2*, and *Ddc* only in B-cells and T-cells). Three imprinted genes (*Ndn*, *Peg3*, and *Peg12*) that were annotated by Chambers and colleagues [19] as HSC specific genes were identified here as marker genes for differentiated lineages. Consistent with the findings of Chambers *et al.*, they are highly expressed in HSCs and downregulated in differentiated states.

The postulated functional role of the identified lineage markers during hematopoiesis was backed up by inspecting the mammalian phenotypes associated with hematopoiesis abnormalities using the MGI database [44], (Table B in S1 File). Apparently, deleting the lineage-specific marker genes leads to deficiencies in either functionalities or differentiation of a specific lineage, validating the used approach in identifying the lineage markers. An example from the B-cell lineage is the knockout of the imprinted gene *CD81*. This is reported to cause *abnormal B-cell morphology* (MGI ID: MP:0004939), *decreased B-1 B-cell number* (MP:0004978), and *instability in B-cell proliferation* (MP:0005154, MP:0005093). More generally, the knockout of the imprinted gene *Cdkn1c* leads to *decreasing hematopoietic stem cell number* (MP:0004810) and *abnormal hematopoietic stem cell physiology* (MP:0010763). From the set of pluripotent genes, gene knockout of *Relb* exhibits also several abnormalities such as *decreased B-cell number* (MP:0005017), *decreased B-cell proliferation* (MP:0005093), *absent lymph nodes* (MP:0008024), *decreased pre-B-cell number* (MP:0008209), and *extra-medullary hematopoiesis* (MP:0000240).

## Construction of TF-miRNA mediated regulatory network and hub genes

Subsequently, we identified the miRNAs associated with each marker gene set and constructed for each lineage a GRN network that represents the combinatorial regulatory mechanisms of miRNAs and TFs (see [Material and Methods](#)) to the extent these are reflected in the regulatory databases that this work was based on. Each such lineage network contains three types of nodes, namely miRNAs, TFs, and target genes. In order to assess the contribution of each node type in the lineage-specific GRN networks to their stability and robustness, we computed the node degree centrality parameters and ranked the nodes according to their degrees.

The most connected nodes (miRNAs and TFs and target genes) are candidates for central hubs (key network players) that could possibly drive the cellular switch towards the related hematopoietic lineage commitment and thus could act as potential master regulator. In Table 3, we list the top 10% nodes with highest degree for each lineage. For instance, we identified 5 hub TFs/genes (*Rel*, *Ppp1r9a*, *Smad7*, *Ndn*, and *Cdkn1c*) and 10 hub miRNAs (*mir-92a-3p*, *mir-377-3p*, *mir-17-5p*, *mir-96-5p*, *mir-875-3p*, *mir-875-5p*, *mir-721*, *mir-330-3p*, *mir-20a-5p*, and *mir-19a-3p*) for the monocyte lineage (Figure E in S1 File). Among them, genes *Rel*,

**Table 3. The identified central hub nodes (TFs/miRNAs) and their enrichment analysis.**

Lineage	Hub TFs/genes	Related enriched Go Terms	Hub miRNAs	Related enriched Functional terms and diseases (p-value)
<b>B-cell</b>	Pou2f1, Zfx, Creb1, Myc, Klf4, Mycn, Rel, Rela, Dnmt3a, Satb2, Ppp1r9a, Mtf2, Phf17, Ccnd1, Rbbp7, Pbrm1, Smad2, Sgk1, Smarca4, Acvr1, Smad4, Mitf, Dnmt1, Smarca2, Hdac2, Terf2, Il6st, Smarcd1, Ctbp2, Smarcc1, Mapk1, H3f3a, Mta2, Ewsr1, Fgf4, Ndn, Dhx9, Rcn2, Hdac1, Gab1, Acvr1b, Irs1, Smad1, Lef1, Chd4, Zfp143, Tcf3, Klf2, Arid3b, Bmpr2, Ehmt2, Mkrn3, Fgfr1, Sfmbt2, Gatm, Sumo1, Cdkn1c, Relb, Sp2, Cd81, Stk40, Carm1, Cdk2, Hira, Axl, Ssrp1, Smarca5, Parp1	<ul style="list-style-type: none"> <li>•Positive regulation of B-cell proliferation (0.009)</li> <li>•Embryonic pattern specification (0.01)</li> <li>•Cell fate commitment (0.0046)</li> <li>•Cell differentiation (1.6e-7)</li> <li>•Embryonic morphogenesis (6.7e-8)</li> <li>•Cell-cell signaling involved in cell fate specification (0.011)</li> </ul>	mmu-mir-19a-3p, mmu-mir-19b-3p, mmu-mir-92a-3p, mmu-mir-495-3p, mmu-mir-17-5p, mmu-mir-200c-3p, mmu-mir-9-5p, mmu-mir-302b-3p, mmu-mir-302a-3p, mmu-mir-20a-5p, mmu-mir-381-3p, mmu-mir-302d-3p, mmu-mir-137-3p, mmu-mir-124-3p, mmu-mir-497-5p, mmu-mir-32-5p, mmu-mir-200b-3p, mmu-mir-195a-5p, mmu-let-7i-5p, mmu-mir-875-3p, mmu-mir-694, mmu-mir-340-5p, mmu-mir-338-5p, mmu-mir-30e-5p, mmu-mir-30b-5p	<ul style="list-style-type: none"> <li>•Embryonic stem cell regulation (0.0047)</li> <li>•Cell fate determination (0.00039)</li> <li>•Cell proliferation (Hwang et al BJC 2006) (5.2e-7)</li> <li>•Hematopoiesis (0.0071)</li> <li>•Onco-miRNA (0.007)</li> <li>Leukemia, B-Cell (0.000034)</li> <li>•Leukemia, Myeloid (0.000004)</li> <li>•Lymphoma, B-Cell (0.000106)</li> <li>•Lymphoma, Large-Cell, Anaplastic (0.0000456)</li> </ul>
<b>T-cell</b>	Ctcf, Creb1, Mycn, Trp53, Foxd3, Atf2, Mef2c, Ppp1r9a, Ccnd1, Il6st, Hif1a, Mitf, Terf2, Smad2, Kdm6b, Igf2r, Ndn, Socs1, Ctbp2, Gab1, Kdm6a, Acvr1b, Gatm, Smad1, Ewsr1, Rcn2, Arid3b, Sfmbt2, Smurf1, Dhx9, Smad3, Mta2, Mkrn3, Cdkn1c, Bmpr2, Stk40, Sp2, Klf2, Lef1, Ehmt2, H13, Spp1, Notch1, Cdk2, Ocln, Satb1, Zfp219, Tcf7, Pim1, Hdac1, Cd81, Fgfr1, Hira, Grb2, Impact, Tcf3, Med12, Axin1	<ul style="list-style-type: none"> <li>•Fat cell differentiation (0.023)</li> <li>•Embryonic pattern specification (0.008)</li> <li>•Immune system development (0.005)</li> <li>•Leukocyte differentiation (0.017)</li> <li>•hemopoiesis (0.016)</li> <li>•B-cell lineage commitment (0.019)</li> <li>•hemopoietic or lymphoid organ development (0.023)</li> </ul>	mmu-mir-381-3p, mmu-mir-19b-3p, mmu-mir-19a-3p, mmu-mir-20a-5p, mmu-mir-17-5p, mmu-mir-30e-5p, mmu-mir-302d-3p, mmu-mir-302b-3p, mmu-mir-30c-5p, mmu-mir-302a-3p, mmu-mir-291b-3p, mmu-mir-20b-5p, mmu-mir-124-3p, mmu-mir-106b-5p, mmu-mir-106a-5p, mmu-mir-96-5p, mmu-mir-93-5p, mmu-mir-92a-3p, mmu-mir-32-5p, mmu-mir-30b-5p, mmu-mir-30a-5p, mmu-mir-291a-3p, mmu-mir-497-5p	<ul style="list-style-type: none"> <li>•Angiogenesis (0.001768)</li> <li>•Embryonic stem cell regulation (0.00071201)</li> <li>•cell fate determination (0.017002)</li> <li>•Hematopoiesis (0.005842)</li> <li>•Immune system(Xiao's et al Cell 2009) (2.47e-7)</li> <li>•Adipocyte differentiation (0.003)</li> <li>•Hematologic Neoplasms (0.00000138)</li> <li>•Leukemia-Lymphoma, Adult T-Cell (4.1e-7)</li> <li>•Lymphoma, T-Cell (0.00000138)</li> </ul>
<b>NK-cell</b>	Mycn, Sp1, Ppp1r9a, Atrx, Ndn, Rbl2, Mitf, Gab1, Klf2, Cd81, Chd4, Cdkn1c, Tcf7, Mkrn3, Impact, Relb, Gatad2a, Lef1	<ul style="list-style-type: none"> <li>•hemopoiesis (0.035)</li> <li>•myeloid leukocyte differentiation (0.039)</li> <li>•immune system development (0.047)</li> <li>•myeloid cell differentiation (0.0057)</li> </ul>	mmu-mir-92a-3p, mmu-mir-381-3p, mmu-mir-25-3p, mmu-mir-20a-5p, mmu-mir-200c-3p, mmu-mir-17-5p, mmu-mir-92b-3p, mmu-mir-367-3p, mmu-mir-363-3p, mmu-mir-32-5p, mmu-mir-200b-3p	<ul style="list-style-type: none"> <li>•Akt pathway (0.0036)</li> <li>•Angiogenesis (0.01)</li> <li>•Hematopoiesis (0.02)</li> <li>•Immune system (0.000009)</li> <li>•Hormones regulation (0.004)</li> <li>•Cell proliferation (0.0001)</li> <li>•Leukemia, Myeloid (0.0003)</li> <li>•Lymphoma, Large-Cell, Anaplastic (0.00033)</li> <li>•Hematologic Neoplasms (0.0002)</li> </ul>
<b>Monocyte</b>	Rel, Ppp1r9a, Smad7, Ndn, Cdkn1c	<ul style="list-style-type: none"> <li>•cell development (0.01) developmental process (0.0245)</li> <li>•neuron differentiation (0.004)</li> <li>•regulation of metabolic process (0.029)</li> </ul>	mmu-mir-92a-3p, mmu-mir-377-3p, mmu-mir-17-5p, mmu-mir-96-5p, mmu-mir-875-5p, mmu-mir-875-3p, mmu-mir-721, mmu-mir-330-3p, mmu-mir-20a-5p, mmu-mir-19a-3p	<ul style="list-style-type: none"> <li>•Granulopoiesis (0.013)</li> <li>•Cell proliferation (2.9e-7)</li> <li>•Hematopoiesis (0.00052)</li> <li>•Angiogenesis (0.000186)</li> <li>•Leukemia, Myeloid, Acute (0.00131697)</li> <li>•Myeloproliferative Disorders (0.0018)</li> <li>•Hematologic Neoplasms (6.8e-7)</li> </ul>

(Continued)

Table 3. (Continued)

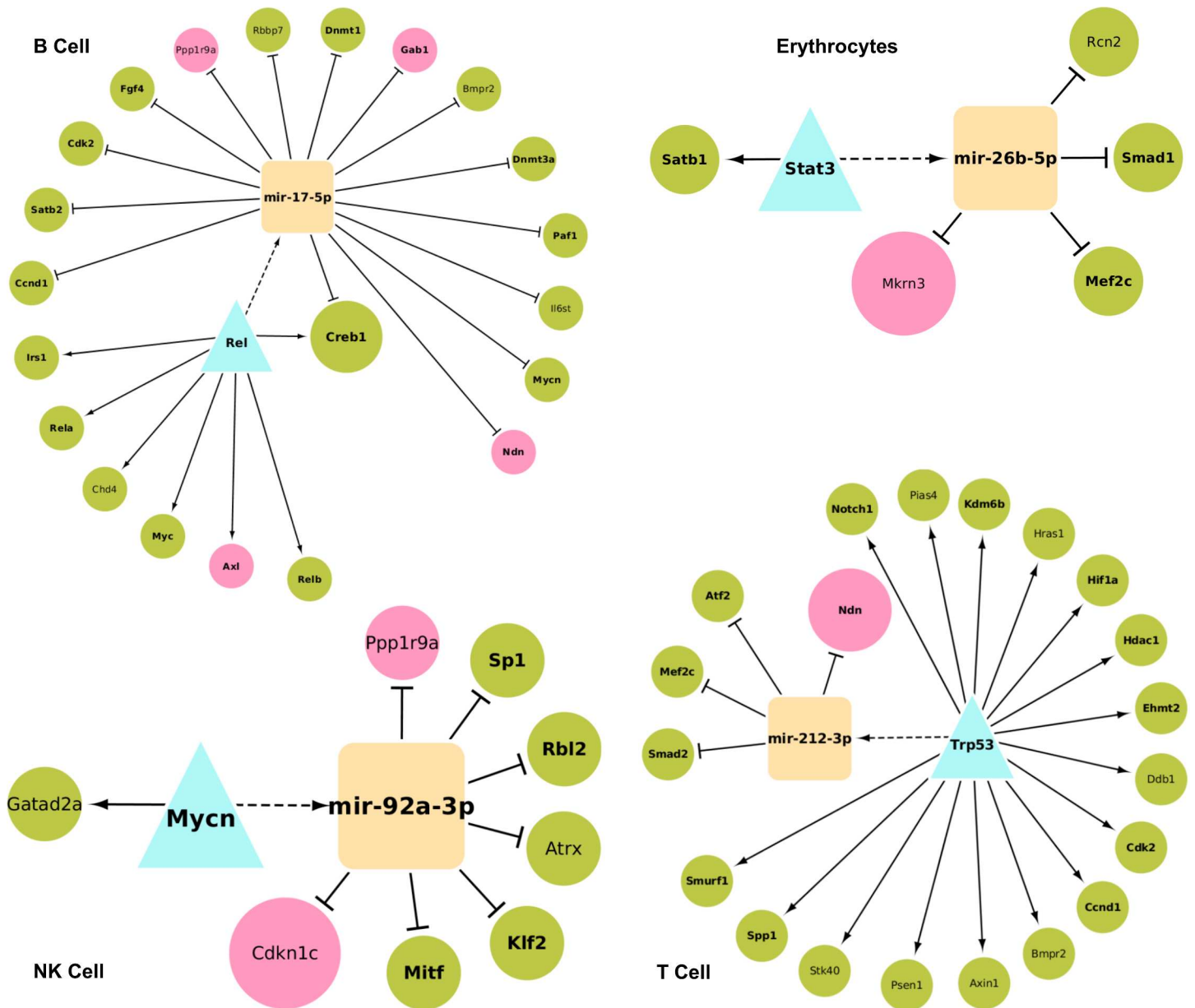
Lineage	Hub TFs/genes	Related enriched Go Terms	Hub miRNAs	Related enriched Functional terms and diseases (p-value)
Granulocyte	Ppp1r9a, Ndn, Mkrn3	<ul style="list-style-type: none"> <li>•Cell projection organization (0.044)</li> <li>•Neuron development (0.04)</li> </ul>	—	—
Erythrocyte	Stat3, Mef2c, Rcn2, Smad1, Sfmbt2, Acvr1b, Mkrn3, Satb1, Hras1	<ul style="list-style-type: none"> <li>-Cell differentiation (0.033)</li> <li>-Positive regulation of macromolecule biosynthetic process (0.0257)</li> <li>-Cellular macromolecule metabolic process (0.037)</li> <li>-Organ morphogenesis (0.031)</li> </ul>	mmu-mir-26b-5p, mmu-mir-495-3p, mmu-mir-26a-5p, mmu-mir-93-5p, mmu-mir-467e-3p, mmu-mir-466a-3p, mmu-mir-329-3p, mmu-mir-291b-5p, mmu-mir-291a-5p, mmu-mir-223-3p, mmu-mir-219-5p, mmu-mir-199a-3p, mmu-mir-183-5p, mmu-mir-135a-5p, mmu-mir-106b-5p, mmu-let-7c-5p, mmu-let-7b-5p, mmu-mir-96-5p, mmu-mir-92b-3p, mmu-mir-9-5p, mmu-mir-880-3p, mmu-mir-878-3p, mmu-mir-762	<ul style="list-style-type: none"> <li>•Embryonic stem cell (hESC) regulation (0.009)</li> <li>•Cell division (0.013)</li> <li>•Cell fate determination (0.00005)</li> <li>•immune system (0.015)</li> <li>•miRNA tumor suppressor (0.02)</li> <li>•Anti-cell proliferation (0.043)</li> <li>•Leukemia, Myeloid, Acute (0.027)</li> <li>•melanoma (0.00807)</li> <li>•Lymphoma, Primary Effusion (0.043)</li> </ul>

doi:10.1371/journal.pone.0166852.t003

*Smad7*, and *Cdkn1c* had been reported to be essential TFs in cell fate determination of hematopoietic progenitors of monocytes [45–48] and the absence of the *Rel* gene leads to multiple hematopoietic cell defects and apoptosis during monocyte differentiation [45]. Moreover, we assessed the biological role of the identified hub nodes by linking them to functional and disease annotations via overrepresentation analysis (ORA). We identified the significant functional terms and associated diseases that were enriched in the group of identified hub miRNAs of each lineage network, see Table 3. Strikingly, in 5 out of 6 lineages, the identified hub miRNAs were enriched with hematopoiesis-related biological processes such as *angiogenesis* (p-value < 0.0017), *embryonic stem cell regulation* (p-value < 0.00071), *cell fate determination* (p-value < 0.017002), *immune system* (p-value < 2.47e-7) as well as with hematopoietic lineage-specific diseases such as leukemia T-Cell (p-value < 4.1e-7), and Lymphoma B-Cell (p-value < 1.06e-4). The networks constructed for monocytes and granulocytes and their hub genes/miRNAs are visualized in Figures E and F in S1 File, respectively.

### Identification of TF-miRNA co-regulatory motifs

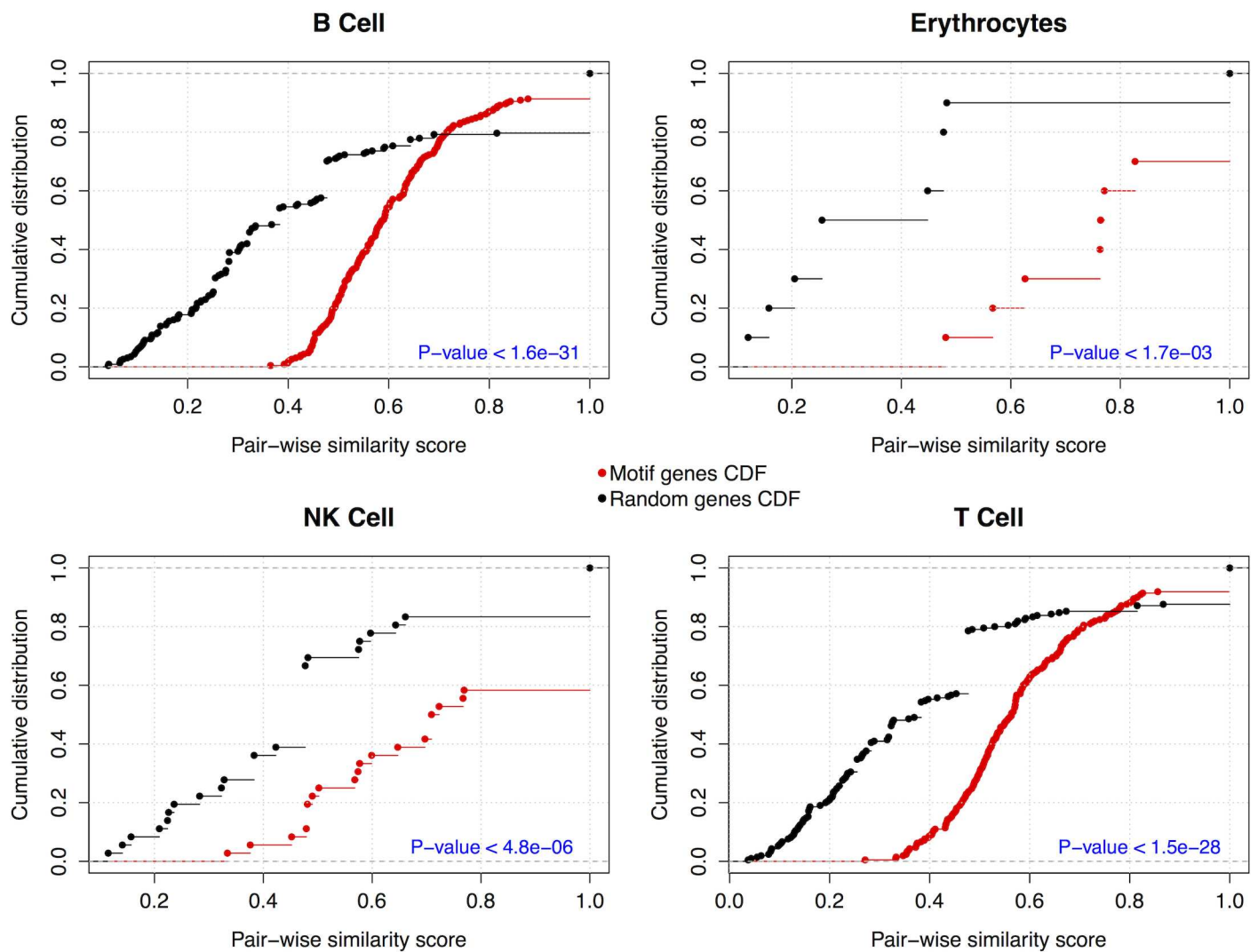
Transcriptional regulation networks often contain Feed Forward Loop (FFL) motifs as functional interconnection patterns [49] that govern many aspects of normal cell functions and diseases [50, 51]. Here, we checked for the presence of 3-node co-regulatory FFL motifs involving the lineage marker TFs/genes and miRNAs in the GRNs of individual hematopoietic lineages (see Materials and Methods). We found that such 3-node motifs were significantly enriched only in the B-cell, T-cell, erythrocyte, and NK-cell lineage networks. The statistical significance of the motifs was tested by comparing their count in the real network to their counts in randomized variants of these networks preserving the same node degrees. In the B-cell, T-cell, erythrocyte, and NK-cell lineages 1727, 1182, 11, and 60 statistically significant FFL motifs were detected, respectively. Fig 3 illustrates the most significant motif detected in each lineage network. Interestingly, these motifs involved both imprinted and pluripotent genes as either a TF or a target gene suggesting the collaborative role and their participation in cellular differentiation and development in the corresponding lineages. A good example for that is the detected



**Fig 3. The most significant TF-miRNA co-regulatory motifs for 4 hematopoietic lineages.** TFs are represented by a turquoise triangle whereas miRNAs are shown as orange squares. Green circles denote the pluripotent genes whereas imprinted genes are colored in pink. Bold label nodes are supported by literature evidence.

doi:10.1371/journal.pone.0166852.g003

B-cell motif that comprises the miRNA mir-17-5p, the imprinted genes *Axl*, *Ndn*, *Gab1*, *Ppp1r9a* and the pluripotent genes *Rel*, *Ccnd1*, *Satb2*, *Creb1*, *Cdk2*, *Fgf4*, *Rbbp7*, *Dnmt1*, *Bmpr2*, *Dnmt3a*, *Paf1*, *Il6st*, *Mycn*, *Myc*, *Chd4*, *Rela*, *Irs1*, *Relb*. Noteworthy is the TF *Rel* gene that was characterized as a hub gene of the B-cell lineage (Table 3) and was reported to play a critical role in cell fate determination of multiple hematopoietic progenitors including B-cell [45]. Also, the TF *Myc* belongs to the four known Yamanaka factors that play a significant role in cell reprogramming [52] and were shown to be sufficient for reprogramming differentiated cells into iPS [53]. *Myc* is believed to regulate the expression of 15% of all human genes [54] and plays an important role in B-cell proliferation [55]. The imprinted gene *Gab1* was found



**Fig 4. Functional homogeneity of the identified co-regulatory motifs.** Cumulative distribution of GO functional semantic scores of gene pairs of co-regulated genes in the examined motifs (red) versus randomly selected genes (black). The p-value was calculated using the Kolmogorov-Smirnov test.

doi:10.1371/journal.pone.0166852.g004

to recruit cytosolic signaling proteins to cellular membranes where they can act on membrane-bound targets during the activation of multiple B-cell antigen receptor signaling pathways [56].

Another example is the erythrocytes motif that involves the pluripotent TF *Stat3* (which was previously reported to be essential in the regulation of erythropoiesis and erythrocytes development [57]) and the miRNA mir-26b-5p (which exhibits moderate changes during erythroid differentiation [58]). The TF regulates the miRNA which in turns regulates the imprinted gene *Mkrn3* whose expression level was significantly increased in several hematopoietic lineages with distinct patterns [59].

Furthermore, we evaluated the biological evidence of these motifs to better understand their functional roles in the development of lineage cell states. For this, we computed the functional similarity scores between co-regulated gene pairs as a measure of their functional homogeneity and integrity. The distribution of the resulting similarity scores was compared to the

similarity score distribution of randomly selected gene pairs (p-values < 0.00173, Kolmogorov-Smirnov test), see Fig 4. As the co-regulated genes have significantly more similar cellular functions than randomly selected genes, these FFL motifs could provide novel insights into TF-miRNA networks during lineage commitment and specification and suggest a cooperative functional role between the TFs/genes and potential miRNA partners. Especially when combined with experimental validation, these co-regulatory motifs as well as the identified hub TFs/genes and miRNAs could expand our understanding of prospective lineage-specific genomic drivers and therefore could help to better understand the cell fate determination of hematopoietic progenitor cells.

In summary, we unraveled specific molecular regulatory interactions involving imprinted genes that explain how these genes contribute to cellular differentiation processes and to different developmental stages of hematopoiesis. The presented analysis suggests new aspects of the cellular regulation of the onset of cellular differentiation and during hematopoiesis. These involve, on the one hand, pluripotent genes that were previously not discussed in the context of hematopoiesis and, on the other hand, involve imprinted genes that are related to genomic imprinting.

### Supporting Information

**S1 File. Figure A. Venn diagram of the 3 gene sets involved in the analysis.** Imprinted, pluripotent, and hematopoietic genes. **Figure B. A sketch diagram showing the calculation of the similarity score between the two observation groups:** 1- the similarity scores between imprinted genes and (pluripotent or hematopoietic) genes and 2- the similarity scores between non-imprinted genes and (pluripotent or hematopoietic) genes as a background. The expression values of the imp/non-imp genes are weighted by the Gaussian fitting of the expression values of the corresponding pluripotent/ hematopoietic sample. **Figure C. The functional similarity between genes belonging to the first class (c1) of each gene set in Fig 1 in comparison to 2 background sets.** In bg1, class 1 of the first gene set is compared with "not class 1" of the other gene set. In bg2, class 1 of the first gene set is compared with a randomly selected set of genes from the other gene set that has the same size like class 1. Therefore the following labels are used in the plot:

imp_hema:	"imp(c1) vs hema(c1)",
i_h(bg1):	"imp(c1) vs hema (notc1)",
i_h(bg2):	"imp(c1) vs random genes of size hema c1"
imp_pluri:	"imp(c1) vs pluri(c1)",
i_p(bg1):	"imp(c1) vs pluri (notc1)",
i_p(bg2):	"imp(c1) vs random genes of size pluri c1",
plur_hema:	"pluri(c1) vs hema(c1)",
p_h(bg1):	"pluri(c1) vs hema (notc1)",
p_h(bg2):	"pluri(c1) vs random genes of size hema c1"

doi:10.1371/journal.pone.0166852.t004

**Figure D. Heatmaps of differentially expressed imprinted genes (paternally expressed are in blue and maternally expressed are in red), pluripotency genes (cyan), and hematopoietic genes (orange) along three blood lineages (B cells, T cells, and granulocytes) based on GSE34723 dataset.** The other three lineages are shown in Fig 2. Shared genes between pluripotency and hematopoietic gene sets are marked in black. Green spots represent down-regulated



genes, and red spots represent up-regulated genes. The clustering reveals that for every developmental line, there exist imprinted as well pluripotency and hematopoietic genes showing similar expression changes during development. **Figure E. A TF-miRNA network for the monocyte lineage.** Green color denotes the pluripotent genes whereas imprinted genes are colored in pink. miRNAs are colored in orange. Darker nodes are the identified putative driver TFs/miRNAs. Bold label nodes are nodes, which are supported by literature evidence. **Figure F. A TF-miRNA network for the granulocyte lineage.** Green color denotes the pluripotent genes whereas imprinted genes are colored in pink. miRNAs are colored in orange. Darker nodes are the identified putative driver TFs/miRNAs. Bold label nodes are nodes, which are supported by literature evidence. **Table A in S1 File. List of the 86 imprinted genes that are annotated in the MS4302.0 microarray chip. Table B in S1 File. List of lineage-specific imprinted, pluripotent, and hematopoietic genes.** in the investigated blood lineages and the associated mammalian phenotypes due to gene knock outs according to the MGI database. **Table C in S1 File. Enrichment of Pluripotency genes with Hematopoietic functional terms.**  
(DOCX)

## Acknowledgments

We thank Dr. Margaret Goodell (Baylor College of Medicine, Texas, USA) for valuable comments on an earlier version of the manuscript.

## Author Contributions

**Conceptualization:** MH VH.

**Data curation:** MH JT CS.

**Formal analysis:** MH RS.

**Funding acquisition:** VH MH.

**Investigation:** MH RS.

**Methodology:** MH JT CS.

**Project administration:** MH VH.

**Resources:** MH VH.

**Software:** MH JT CS.

**Supervision:** VH MP.

**Validation:** MH.

**Visualization:** MH JT CS RS.

**Writing – original draft:** MH VH.

**Writing – review & editing:** MH MRI GF MP VH.

## References

1. Moignard V, Macaulay IC, Swiers G, Buettner F, Schütte J, Calero-Nieto FJ, et al. Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nature cell biology*. 2013; 15(4):363–72. doi: [10.1038/ncb2709](https://doi.org/10.1038/ncb2709) PMID: [23524953](https://pubmed.ncbi.nlm.nih.gov/23524953/)

2. Som A, Harder C, Greber B, Siatkowski M, Paudel Y, Warsaw G, et al. The PluriNetWork: An electronic representation of the network underlying pluripotency in mouse, and its applications. *PLoS one*. 2010; 5(12):e15165. doi: [10.1371/journal.pone.0015165](https://doi.org/10.1371/journal.pone.0015165) PMID: [21179244](https://pubmed.ncbi.nlm.nih.gov/21179244/)
3. Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics*. 2001; 2(1):21–32. doi: [10.1038/35047554](https://doi.org/10.1038/35047554) PMID: [11253064](https://pubmed.ncbi.nlm.nih.gov/11253064/)
4. Park S- J, Umemoto T, Saito-Adachi M, Shiratsuchi Y, Yamato M, Nakai K. Computational Promoter Modeling Identifies the Modes of Transcriptional Regulation in Hematopoietic Stem Cells. *PLoS one*. 2014; 9(4):e93853. doi: [10.1371/journal.pone.0093853](https://doi.org/10.1371/journal.pone.0093853) PMID: [24710559](https://pubmed.ncbi.nlm.nih.gov/24710559/)
5. Cabezas-Wallscheid N, Klimmeck D, Hansson J, Lipka DB, Reyes A, Wang Q, et al. Identification of Regulatory Networks in HSCs and Their Immediate Progeny via Integrated Proteome, Transcriptome, and DNA Methylation Analysis. *Cell stem cell*. 2014; 15(4):507–22. doi: [10.1016/j.stem.2014.07.005](https://doi.org/10.1016/j.stem.2014.07.005) PMID: [25158935](https://pubmed.ncbi.nlm.nih.gov/25158935/)
6. Klimmeck fD, Cabezas-Wallscheid N, Reyes A, von Paleske L, Renders S, Hansson J, et al. Transcriptome-wide Profiling and Posttranscriptional Analysis of Hematopoietic Stem/Progenitor Cell Differentiation toward Myeloid Commitment. *Stem cell reports*. 2014; 3(5):858–75. doi: [10.1016/j.stemcr.2014.08.012](https://doi.org/10.1016/j.stemcr.2014.08.012) PMID: [25418729](https://pubmed.ncbi.nlm.nih.gov/25418729/)
7. Venkatraman A, He XC, Thorvaldsen JL, Sugimura R, Perry JM, Tao F, et al. Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature*. 2013.
8. Varrault A, Gueydan C, Delalbre A, Bellmann A, Houssami S, Aknin C, et al. *Zac1* regulates an imprinted gene network critically involved in the control of embryonic growth. *Developmental cell*. 2006; 11(5):711–22. doi: [10.1016/j.devcel.2006.09.003](https://doi.org/10.1016/j.devcel.2006.09.003) PMID: [17084362](https://pubmed.ncbi.nlm.nih.gov/17084362/)
9. Berg JS, Lin KK, Sonnet C, Boles NC, Weksberg DC, Nguyen H, et al. Imprinted Genes That Regulate Early Mammalian Growth Are Coexpressed in Somatic Stem Cells. *PLoS one*. 2011; 6(10):e26410. doi: [10.1371/journal.pone.0026410](https://doi.org/10.1371/journal.pone.0026410) PMID: [22039481](https://pubmed.ncbi.nlm.nih.gov/22039481/)
10. Hamed M, Ismael S, Paulsen M, Helms V. Cellular Functions of Genetically Imprinted Genes in Human and Mouse as Annotated in the Gene Ontology. *PLoS one*. 2012; 7(11):e50285. doi: [10.1371/journal.pone.0050285](https://doi.org/10.1371/journal.pone.0050285) PMID: [23226257](https://pubmed.ncbi.nlm.nih.gov/23226257/)
11. Scandura JM, Boccuni P, Massagué J, Nimer SD. Transforming growth factor  $\beta$ -induced cell cycle arrest of human hematopoietic cells requires p57KIP2 up-regulation. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(42):15231–6. doi: [10.1073/pnas.0406771101](https://doi.org/10.1073/pnas.0406771101) PMID: [15477587](https://pubmed.ncbi.nlm.nih.gov/15477587/)
12. Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C, Malin S, et al. Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity*. 2008; 28(6):751–62. doi: [10.1016/j.immuni.2008.04.014](https://doi.org/10.1016/j.immuni.2008.04.014) PMID: [18538592](https://pubmed.ncbi.nlm.nih.gov/18538592/)
13. Venkatraman A, He XC, Thorvaldsen JL, Sugimura R, Perry JM, Tao F, et al. Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature*. 2013.
14. Hutter B, Bieg M, Helms V, Paulsen M. Imprinted genes show unique patterns of sequence conservation. *BMC genomics*. 2010; 11(1):649.
15. Hutter B, Bieg M, Helms V, Paulsen M. Divergence of imprinted genes during mammalian evolution. *BMC evolutionary biology*. 2010; 10(1):116.
16. Morison IM, Paton CJ, Cleverley SD. The imprinted gene and parent-of-origin effect database. *Nucleic Acids Research*. 2001; 29(1):275–6. PMID: [11125110](https://pubmed.ncbi.nlm.nih.gov/11125110/)
17. Schulz R, Woodfine K, Menhenniott TR, Bourc'his D, Bestor T, Oakey RJ. WAMIDEX: a web atlas of murine genomic imprinting and differential expression. *Epigenetics: official journal of the DNA Methylation Society*. 2008; 3(2):89.
18. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nature genetics*. 2000; 25(1):25–9. doi: [10.1038/75556](https://doi.org/10.1038/75556) PMID: [10802651](https://pubmed.ncbi.nlm.nih.gov/10802651/)
19. Chambers SM, Boles NC, Lin KYK, Tierney MP, Bowman TV, Bradfute SB, et al. Hematopoietic fingerprints: an expression database of stem cells and their progeny. *Cell Stem Cell*. 2007; 1(5):578–91. doi: [10.1016/j.stem.2007.10.003](https://doi.org/10.1016/j.stem.2007.10.003) PMID: [18371395](https://pubmed.ncbi.nlm.nih.gov/18371395/)
20. Di Tullio A, Manh TPV, Schubert A, Månsson R, Graf T. CCAAT/enhancer binding protein  $\alpha$  (C/EBP $\alpha$ )-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proceedings of the National Academy of Sciences*. 2011; 108(41):17016–21.
21. Seita J, Sahoo D, Rossi DJ, Bhattacharya D, Serwold T, Inlay MA, et al. Gene Expression Commons: An Open Platform for Absolute Gene Expression Profiling. *PLoS one*. 2012; 7(7):e40321. doi: [10.1371/journal.pone.0040321](https://doi.org/10.1371/journal.pone.0040321) PMID: [22815738](https://pubmed.ncbi.nlm.nih.gov/22815738/)
22. Lattin JE, Schroder K, Su AI, Walker JR, Zhang J, Wiltshire T, et al. Expression analysis of G Protein-Coupled Receptors in mouse macrophages. *Immunome research*. 2008; 4(1):5.

23. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*. 2002; 30(1):207–10. PMID: [11752295](#)
24. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*. 2004; 5(10):R80. doi: [10.1186/gb-2004-5-10-r80](#) PMID: [15461798](#)
25. Ihaka R, Gentleman R. R: A language for data analysis and graphics. *Journal of computational and graphical statistics*. 1996:299–314.
26. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995:289–300.
27. Hamed M, Spaniol C, Nazarieh M, Helms V. TFmiR: a web server for constructing and analyzing disease-specific transcription factor and miRNA co-regulatory networks. *Nucleic acids research*. 2015: gkv418.
28. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal Complex Systems*. 2006; 1695:1695.
29. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*. 2011; 27(3):431–2. doi: [10.1093/bioinformatics/btq675](#) PMID: [21149340](#)
30. Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics*. 2010; 26(7):976–8. doi: [10.1093/bioinformatics/btq064](#) PMID: [20179076](#)
31. Da Wei Huang BTS, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2008; 4(1):44–57.
32. Lu M, Shi B, Wang J, Cao Q, Cui Q. TAM: a method for enrichment and depletion analysis of a microRNA category in a list of microRNAs. *BMC bioinformatics*. 2010; 11(1):419.
33. Hamed M, Spaniol C, Zapp A, Helms V. Integrative network-based approach identifies key genetic elements in breast invasive carcinoma. *BMC genomics*. 2015; 16(Suppl 5):S2.
34. Göke J, Jung M, Behrens S, Chavez L, O’Keeffe S, Timmermann B, et al. Combinatorial binding in human and mouse embryonic stem cells identifies conserved enhancers active in early embryonic development. *PLoS computational biology*. 2011; 7(12):e1002304. doi: [10.1371/journal.pcbi.1002304](#) PMID: [22215994](#)
35. Mougeot JL, Bahrani-Mougeot FK, Lockhart PB, Brennan MT. Microarray analyses of oral punch biopsies from acute myeloid leukemia (AML) patients treated with chemotherapy. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2011; 112(4):446–52. doi: [10.1016/j.tripleo.2011.05.009](#) PMID: [21862359](#)
36. Grumont R, Gerondakis S. The murine c-rel proto-oncogene encodes two mRNAs the expression of which is modulated by lymphoid stimuli. *Oncogene research*. 1990; 5(4):245. PMID: [2204017](#)
37. Hu H, Huston G, Duso D, Lepak N, Roman E, Swain SL. CD4+ T cell effectors can become memory cells with high efficiency and without further division. *Nature immunology*. 2001; 2(8):705–10. doi: [10.1038/90643](#) PMID: [11477406](#)
38. Janssen EM, Lemmens EE, Wolfe T, Christen U, von Herrath MG, Schoenberger SP. Cd4&plus; t cells are required for secondary expansion and memory in cd8&plus; t lymphocytes. *Nature*. 2003; 421(6925):852–6. doi: [10.1038/nature01441](#) PMID: [12594515](#)
39. Sun JC, Beilke JN, Lanier LL. Adaptive immune features of natural killer cells. *Nature*. 2009; 457(7229):557–61. doi: [10.1038/nature07665](#) PMID: [19136945](#)
40. Haug JS, He XC, Grindley JC, Wunderlich JP, Gaudenz K, Ross JT, et al. N-cadherin expression level distinguishes reserved versus primed states of hematopoietic stem cells. *Cell stem cell*. 2008; 2(4):367–79. doi: [10.1016/j.stem.2008.01.017](#) PMID: [18397756](#)
41. Månsson R, Hultquist A, Luc S, Yang L, Anderson K, Kharazi S, et al. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity*. 2007; 26(4):407–19. doi: [10.1016/j.immuni.2007.02.013](#) PMID: [17433729](#)
42. Gekas C, Graf T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood*. 2013; 121(22):4463–72. doi: [10.1182/blood-2012-09-457929](#) PMID: [23564910](#)
43. Sanjuan-Pla A, Macaulay IC, Jensen CT, Woll PS, Luis TC, Mead A, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013; 502(7470):232–6. doi: [10.1038/nature12495](#) PMID: [23934107](#)
44. Bult C, Blake J, Kadin J, Eppig J, Ringwald M, Richardson J. P4-S The Mouse Genome Informatics Database: An Integrated Resource for Mouse Genetics and Genomics. *Journal of Biomolecular Techniques: JBT*. 2007; 18(1):2.

45. Grossmann M, Metcalf D, Merryfull J, Beg A, Baltimore D, Gerondakis S. The combined absence of the transcription factors Rel and RelA leads to multiple hemopoietic cell defects. *Proceedings of the National Academy of Sciences*. 1999; 96(21):11848–53.
46. Kitamura K, Aota S-i, Sakamoto R, Yoshikawa S-I, Okazaki K. Smad7 selectively interferes with different pathways of activin signaling and inhibits erythroid leukemia cell differentiation. *Blood*. 2000; 95(11):3371–9. PMID: [10828018](#)
47. Lee EO, Kang JL, Chong YH. The Amyloid- $\beta$  Peptide Suppresses Transforming Growth Factor- $\beta$ 1-induced Matrix Metalloproteinase-2 Production via Smad7 Expression in Human Monocytic THP-1 Cells. *Journal of Biological Chemistry*. 2005; 280(9):7845–53. doi: [10.1074/jbc.M409101200](#) PMID: [15632190](#)
48. Frankenberger M, Hofer TP, Marei A, Dayyani F, Schewe S, Strasser C, et al. Transcript profiling of CD16-positive monocytes reveals a unique molecular fingerprint. *European journal of immunology*. 2012; 42(4):957–74. doi: [10.1002/eji.201141907](#) PMID: [22531920](#)
49. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature genetics*. 2002; 31(1):64–8. doi: [10.1038/ng881](#) PMID: [11967538](#)
50. He L, He X, Lim LP, De Stanchina E, Xuan Z, Liang Y, et al. A microRNA component of the p53 tumour suppressor network. *Nature*. 2007; 447(7148):1130–4. doi: [10.1038/nature05939](#) PMID: [17554337](#)
51. Li X, Cassidy JJ, Reinke CA, Fischboeck S, Carthew RW. A microRNA imparts robustness against environmental fluctuation during development. *Cell*. 2009; 137(2):273–82. doi: [10.1016/j.cell.2009.01.058](#) PMID: [19379693](#)
52. Yamanaka S. Induction of pluripotent stem cells from mouse fibroblasts by four transcription factors. *Cell proliferation*. 2008; 41:51–6. doi: [10.1111/j.1365-2184.2008.00493.x](#) PMID: [18181945](#)
53. Takahashi K, Okita K, Nakagawa M, Yamanaka S. Induction of pluripotent stem cells from fibroblast cultures. *Nature protocols*. 2007; 2(12):3081–9. doi: [10.1038/nprot.2007.418](#) PMID: [18079707](#)
54. Gearhart J, Pashos EE, Prasad MK. Pluripotency redux—advances in stem-cell research. *The New England journal of medicine*. 2007; 357(15):1469–72. doi: [10.1056/NEJMp078126](#) PMID: [17928593](#)
55. de Alboran IM, O'Hagan RC, Gartner F, Malynn B, Davidson L, Rickert R, et al. Analysis of C-MYC function in normal cells via conditional gene-targeted mutation. *Immunity*. 2001; 14(1):45–55. PMID: [11163229](#)
56. Ingham RJ, Holgado-Madruga M, Siu C, Wong AJ, Gold MR. The Gab1 protein is a docking site for multiple proteins involved in signaling by the B cell antigen receptor. *Journal of Biological Chemistry*. 1998; 273(46):30630–7. PMID: [9804835](#)
57. Watowich SS, Mikami A, Busche RA, Xie X, Pharr PN, Longmore GD. Erythropoietin receptors that signal through Stat5 or Stat3 support fetal liver and adult erythropoiesis: lack of specificity of stat signals during red blood cell development. *Journal of Interferon & Cytokine Research*. 2000; 20(12):1065–70.
58. Bruchova H, Yoon D, Agarwal AM, Mendell J, Prchal JT. Regulated expression of microRNAs in normal and polycythemia vera erythropoiesis. *Experimental hematology*. 2007; 35(11):1657–67. doi: [10.1016/j.exphem.2007.08.021](#) PMID: [17976518](#)
59. Liang J, Song W, Tromp G, Kolattukudy PE, Fu M. Genome-wide survey and expression profiling of CCCH-zinc finger family reveals a functional module in macrophage activation. *PLoS One*. 2008; 3(8): e2880. doi: [10.1371/journal.pone.0002880](#) PMID: [18682727](#)