Stem Cell Reports
Resource
OPEN ACCESS

Transcriptome-wide Profiling and Posttranscriptional Analysis of Hematopoietic Stem/Progenitor Cell Differentiation toward Myeloid Commitment

Daniel Klimmeck,[1,2,3,4,5] Nina Cabezas-Wallscheid,[1,2,4,5] Alejandro Reyes,[3,5] Lisa von Paleske,[1] Simon Renders,[1,2] Jenny Hansson,[3] Jeroen Krijgsveld,[3] Wolfgang Huber,[3,6,*] and Andreas Trumpp[1,2,4,6,*]
[1]Division of Stem Cells and Cancer, Deutsches Krebsforschungszentrum (DKFZ), 69120 Heidelberg, Germany
[2]Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM gGmbH), 69120 Heidelberg, Germany
[3]European Molecular Biology Laboratory (EMBL), Genome Biology Unit, 69120 Heidelberg, Germany
[4]German Cancer Consortium (DKTK)
[5]Co-first author
[6]Co-senior author
*Correspondence: wolfgang.huber@embl.de (W.H.), a.trumpp@dkfz.de (A.T.)
http://dx.doi.org/10.1016/j.stemcr.2014.08.012

SUMMARY

Hematopoietic stem cells possess lifelong self-renewal activity and generate multipotent progenitors that differentiate into lineage-committed and subsequently mature cells. We present a comparative transcriptome analysis of ex vivo isolated mouse multipotent hematopoietic stem/progenitor cells ($Lin^{neg}SCA\text{-}1^{+}c\text{-}KIT^{+}$) and myeloid committed precursors ($Lin^{neg}SCA\text{-}1^{neg}c\text{-}KIT^{+}$). Our data display dynamic transcriptional networks and identify a stem/progenitor gene expression pattern that is characterized by cell adhesion and immune response components including kallikrein-related proteases. We identify 498 expressed lncRNAs, which are potential regulators of multipotency or lineage commitment. By integrating these transcriptome with our recently reported proteome data, we found evidence for posttranscriptional regulation of processes including metabolism and response to oxidative stress. Finally, our study identifies a high number of genes with transcript isoform regulation upon lineage commitment. This in-depth molecular analysis outlines the enormous complexity of expressed coding and noncoding RNAs and posttranscriptional regulation during the early differentiation steps of hematopoietic stem cells toward the myeloid lineage.

INTRODUCTION

In the adult hematopoietic system, short-lived mature cells are constantly lost and need to be replaced in order to maintain blood homeostasis (Murphy et al., 2005; Weissman and Shizuru, 2008). This essential task is fulfilled by hematopoietic stem cells (HSCs), which reside in the trabecular areas of the bone marrow (Purton and Scadden, 2007; Till and McCulloch, 1961; Wilson et al., 2009). HSCs possess the highest self-renewal capacity and produce multipotent progenitors (MPPs) with steadily decreasing self-renewal activity (Trumpp et al., 2010; Weissman and Shizuru, 2008). HSCs and MPPs (HSPC) are contained within a compartment immunophenotypically defined as negative for mature blood cell markers (Lin^{-}) and positive for stem cell markers SCA-1 and c-KIT ($LS^{+}K$; Weissman and Shizuru, 2008). HSPCs eventually commit to more mature lymphoid or myeloid progenitors with increasingly restricted self-renewal and differentiation potential (Graf and Enver, 2009). The myeloid committed progenitor subset (Lin^{-}, $SCA\text{-}1^{-}$ and $c\text{-}KIT^{+}$; [$LS^{-}K$]) comprises common myeloid progenitors (CMPs) as well as more specialized granulocyte-macrophage progenitors (GMPs) and megakaryocyte-erythroid progenitors (MEPs) (Akashi et al., 2000; Pronk et al., 2007), which differentiate toward mature effector cells.

Two crucial aspects of early hematopoiesis are multipotency and lineage commitment (Graf and Enver, 2009; Trumpp et al., 2010). Expression profiling of HSPCs by cDNA microarrays has elucidated important aspects of hematopoietic stem cell biology, including the relevance of the KIT^{-} and Wnt^{-} signaling pathways (Gazit et al., 2013; Kent et al., 2008; Luis et al., 2012; Seita and Weissman, 2010) for multipotency. Transcriptional control networks active in early hematopoiesis have been studied using single-gene expression analysis (Moignard et al., 2013), but their impact on protein levels and posttranscriptional gene expression regulation in HSPCs has not been described.

Recently, transcriptome profiling by next-generation sequencing (NGS; e.g., RNA sequencing [RNA-seq]) has significantly extended the possibilities to study gene expression (Ozsolak and Milos, 2011), which was also used to investigate young versus aged HSCs (Sun et al., 2014). It permits not only the analysis of differential mRNA expression of low abundant regulatory factors, but also the detection of alternative splicing events that can generate different protein isoforms and the identification of noncoding RNAs. Long noncoding RNAs (lncRNAs) (Mercer et al., 2009) are involved in the regulation of gene expression at various levels (Pauli et al., 2011;

Yoon et al., 2013) and can function as oncogenes or tumor-suppressor genes (Gutschner and Diederichs, 2012). Although efforts have been made to identify and elucidate the roles of lncRNAs in stem cells (Qureshi and Mehler, 2012; Uchida et al., 2012), little is known about the expression of lncRNAs or their functions in hematopoietic stem/progenitors (Paralkar and Weiss, 2013). Further, the advent of improved proteome techniques has enabled in-depth comparative analysis of RNA and protein signatures in diverse systems (Cox and Mann, 2007; Schwanhäusser et al., 2011; Vogel and Marcotte, 2012). Although attempts were made to correlate transcriptome and proteome signatures of hematopoietic immature cells (Spooncer et al., 2008), a comprehensive comparison is still lacking.

We performed a genome-wide RNA-seq analysis of primary multipotent and self-renewing hematopoietic stem/progenitors and myeloid committed precursors. We report robust and reproducible transcriptome data with more than 19,000 quantified genes including more than 1,300 noncoding RNA species. To address how gene expression is regulated in multipotency and commitment, we integrated our RNA-seq data with the recently reported proteome data set of the identical cell populations (Klimmeck et al., 2012). These data sets outline the dynamic expression changes that occur during the transition of stem/progenitors toward myeloid commitment.

RESULTS

Quantitative Transcriptomic Analysis of Hematopoietic Stem and Progenitor Cells

Whole-transcriptome analysis was performed to investigate differences in the gene expression profiles between multipotent hematopoietic stem progenitor cells (HSPCs; LS^+K) and myeloid committed cells (LS^-K) of the mouse bone marrow (Figure 1A). We fluorescence-activated cell sorting (FACS)-sorted 50,000 primary cells of each population in three independent biological experiments (Figure 1B and Figure S1 available online) and enriched for polyadenylated RNA. We generated paired-end libraries and sequenced more than 2×10^8 reads per sample (Figures S1 and S2). Quality-control metrics indicated that the data were reproducible and of high quality (Figures 1C, S1, and S2). We identified the expression of 19,824 genes (Table S1). We classified the quantified genes by RNA categories. As expected after poly(A)-RNA enrichment, the majority of transcripts were categorized as protein-coding genes (78%; 15,474; Figure 1D). In addition, we classified hits to 23 other noncoding RNA categories including pseudogenes (1,783) and lncRNAs (498).

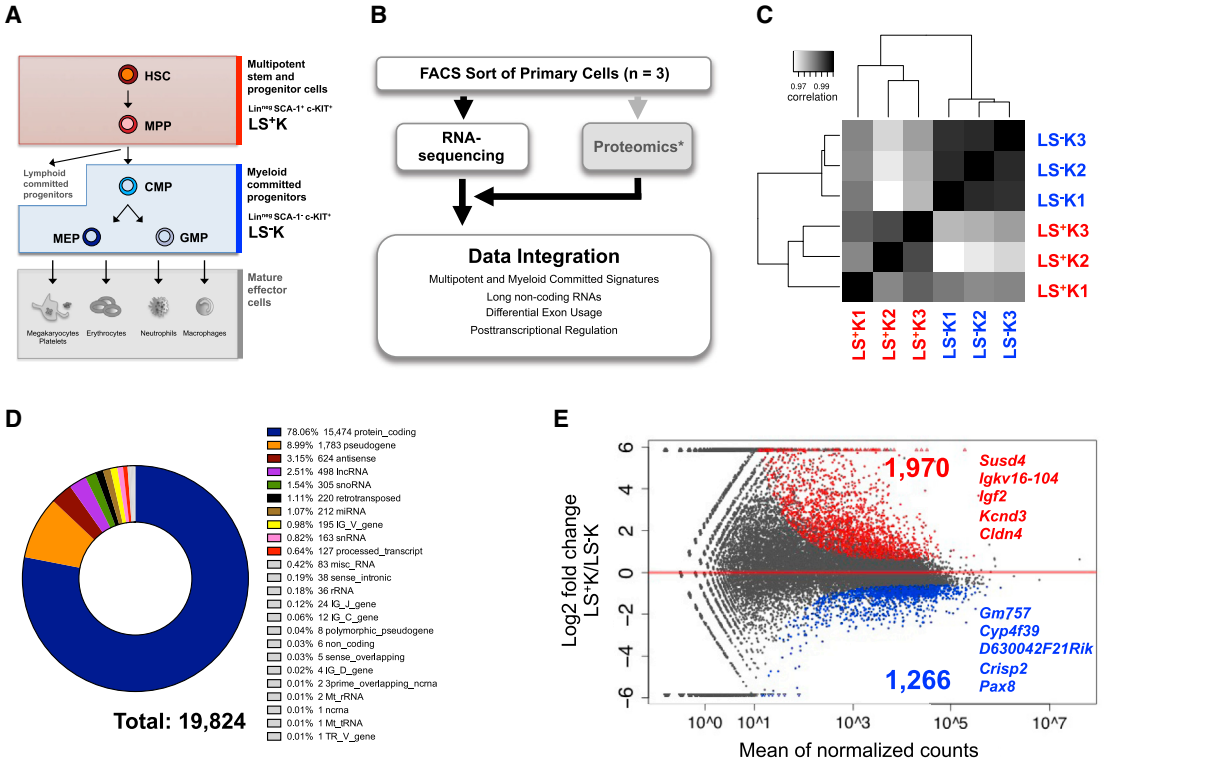
Differential Gene Expression Analysis Reveals Significant Transcriptional Divergence between HSPCs and Myeloid Committed Precursors

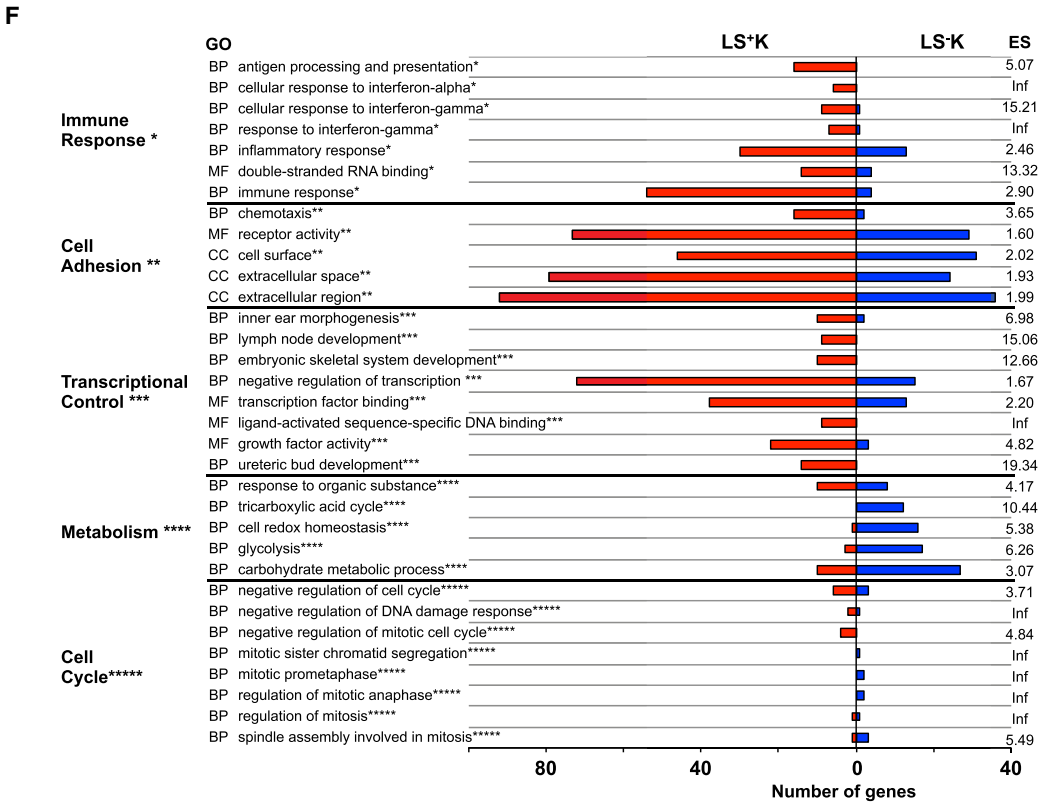
We found 3,236 genes to be differentially expressed (false discovery rate [FDR] = 0.05) between multipotent and myeloid committed cells, which indicated a high divergence in transcript levels (Figure 1E; Table S1). Of these, 1,970 genes were highly expressed in HSPCs and 1,266 were higher in LS^-K cells. As expected, transcripts encoding SCA-1, which was used to sort this population, as well as FLT3 and THY1 surface markers (Adolfsson et al., 2005; Weissman and Shizuru, 2008) were preferentially expressed in LS^+K cells (Table S1). In contrast, neutrophil serine proteases *Ctsg* and *Elane* were highly expressed in LS^-K cells (Korkmaz et al., 2008). We validated 24 differentially expressed genes by quantitative real-time PCR, which confirmed the robustness of the data for a wide range of gene expression transcript abundances (from 2 to 2,314,409 of the mean of sequenced fragments; Figure S2D).

To investigate the biological roles of our set of differentially expressed genes, we applied a gene ontology (GO) enrichment analysis, using a stratified approach that controls for expression strength biases (Figure 1F; Table S2). The GO terms overrepresented in LS^+K cells fell into three main categories: immune response (* in Figure 1F; e.g., antigen processing and presentation, inflammatory response), cell adhesion (**; chemotaxis), and transcriptional control (***; negative regulation of transcription). In contrast, metabolism (****; cell redox homeostasis, glycolysis) was significantly overrepresented in myeloid committed cells. Cell cycle (*****) showed a dual pattern with e.g., negative regulation of cell cycle enriched in LS^+K and spindle assembly involved in mitosis enriched in myeloid cells, reflecting the higher proliferative state in the committed progenitors. We determined genes that were uniquely detected in HSPCs but not in LS^-K and vice versa (Figure S3). Although four genes were exclusively detected in LS^-K myeloid progenitors (*Ivl*, *Gm14705*, *Gp6*, *Zfp819*), 69 genes of diverse RNA categories were only detected in LS^+K cells and categorized primarily to cell adhesion and immune system process (e.g., Kallikrein-related peptidases *Klk1*, *Klk9*, and *Klk10*; Figure S3), suggesting particular importance of these processes for the HSPC state. In conclusion, the differential expression pattern for key cellular processes demonstrates a distinct transcriptomic composition of multipotent and myeloid committed cells.

Cell-Cycle Activity Is Tightly Regulated in HSPCs and upon Myeloid Commitment

Next, we analyzed the protein-protein interaction network of the differentially expressed genes related to cell cycle (Figure 2A; ***** in Figure 1F). We found negative regulators of mitosis to be highly expressed in LS^+K cells (*Cdkn1b/p27*),





(legend on next page)

in line with previous work demonstrating regulation of cell-cycle activity in HSPCs (Zou et al., 2011; Tesio and Trumpp, 2011). In contrast, a highly interconnected group of proliferation-related genes were enriched in LS$^-$K (*Plk1*, *Rps6ka2*). We validated these findings by real-time PCR and extended this expression analysis to more refined myeloid committed progenitors (CMP, MEP, GMP; Figures 2B, S4A, and S4B). Interestingly, MEPs and GMPs showed higher expression of cell-cycle activators but lower expression of inhibitors compared to HSPCs and CMPs. To test whether these gene expression patterns were reflected functionally, we assessed cell-cycle stages by flow cytometry (Figures 2C and 2D). In line, HSPCs showed significantly more cells in G0 (36%) compared to overall myeloid committed precursors (17%). Within the LS$^-$K compartment, CMPs were cycling less (23%; G0) compared to the more committed MEP (13%) and GMP (11%) populations. Together, these results suggest that the low cell-cycle activity in HSPCs becomes stepwise activated upon myeloid commitment.

Global Analysis of Genes Involved in Transcriptional Control of Multipotency and Commitment

Next, we investigated the differentially expressed genes related to transcriptional control (*** in Figure 1F). Notably, the majority of these were enriched in HSPCs (Figure 3; 146 out of 177; 82%) and only few in myeloid progenitors (31 out of 177). Classification of these 177 genes according to their molecular function (Figure 3A; Table S3) revealed various functional categories including ligand-receptors (31) and transcription factors (TFs; 62) (Figure 3A). We then assigned these differential regulators to signaling pathways (Table S3). We found the expression of the Bmp-Smad-TGF-beta$^-$ (*Bmp4*, *Smad1*; Blank and Karlsson, 2011) and Wnt$^-$Notch$^-$ (*Wnt10b*, *Jag2*; Bigas and Espinosa, 2012; Luis et al., 2012) signaling pathways to be highly enriched in HSPCs. Notably, six members of the Forkhead box protein family (e.g., *Foxa3*, *Foxr1*) were enriched in HSPCs (2- to 32-fold), whereas *Foxh1* was lowly expressed (−8-fold). Next, we screened the expression of TF targets of Fox family members (Figure S4C). This analysis revealed genes possibly activated downstream of *Foxo1* (*Ccng2*, *Trib3*) in LS$^+$K and *Foxh1* (*Aldh1a1*, *Mixl1*) in LS$^-$K cells; it also highlights factors potentially repressed by *Foxh1* upon myeloid commitment (*Aldh1a2*, *Pitx2*).

Next, we investigated the expression of long noncoding RNAs (lncRNAs) in LS$^+$K and LS$^-$K cells (Figures 3B–3F). From the 489 ENSEMBL-annotated lncRNAs for which we identified expression in our data, 67 were differentially expressed between HSPCs and LS$^-$K (FDR = 0.05; 55 up/ 12 down; Table S1). Strikingly, only two out of the 67 differentially expressed lncRNA set were functionally annotated (*H19*, *Meg3*; (Venkatraman et al., 2013; Yoshimizu et al., 2008; Zhou et al., 2012)). In contrast, other lncRNAs were also exclusive or strongly differentially expressed in HSPCs (*A930001C03Rik*, *Gm12066*) or LS$^-$Ks (*Gm12708*), but they lacked characterization. One of the most abundant lncRNA was *Malat1* (>800,000 counts in LS$^-$K cells; Figure 3C; Table S1), which showed more than 2-fold higher expression in HSPCs. Increased expression of *Meg3* and *Malat1* in HSPCs was confirmed by independent real-time PCR (Figure 3D). In order to examine the expression of *Malat1* within the LS$^+$K compartment, we performed real-time PCR in refined HSC and MPP populations (Wilson et al., 2008). We found *Malat1* to be highly expressed in the most immature

Figure 1. Quantitative Transcriptomic Analysis of Hematopoietic Stem and Progenitor Cells

(A) Early hematopoiesis. Multipotent hematopoietic stem cells (HSCs) give rise to multipotent progenitors (MPPs), which commit either to myeloid specified progenitors (CMPs, common myeloid progenitors; GMP, granulocyte-macrophage progenitor; MEP, megakaryocyte/erythrocyte progenitor) or to lymphoid specified progenitors.

(B) Experimental design. Cell fractions were purified with fluorescence-activated cell sorting (FACS) using specific surface markers for Lineage (Lin), SCA-1, and c-KIT, which distinguish multipotent stem progenitor cells (LS$^+$K) from myeloid committed (LS$^-$K) cells. For analysis of posttranscriptional regulation, the transcriptome was integrated with our previously published proteome (Klimmeck et al., 2012) (asterisk).

(C) Clustering of biological replicates. Heatmap represents similarity of samples from black (highest) to light gray (smallest).

(D) Classification of the quantified genes by RNA categories. The pie chart legend indicates percentage and absolute number of genes for each type of RNA. LncRNA, long noncoding RNA; snoRNA, small nucleolar RNA; miRNA, microRNA; IG_V gene, immunoglobin V gene; snRNA, small nuclear RNA; miscRNA, miscellaneousRNA; rRNA, ribosomal RNA; IG_J gene, immunoglobulin J gene; IG_C gene, immunoglobulin C gene; Mt_tRNA, mitochondrial transfer RNA; TR_V gene, V gene. Genes were called as quantified genes if they had at least 20 read counts per cell type.

(E) Differential gene expression between multipotent and myeloid committed progenitors. The plot shows for each gene (indicated by dots) a measure of its average expression (x axis) versus the logarithm of the ratio between expression levels in LS$^+$K and LS$^-$K. Red and blue coloring of dots represents differentially expressed genes (FDR = 0.05). Red dots represent upregulated genes in LS$^+$K (1,970 genes). Blue dots depict upregulated genes in LS$^-$K (1,266). The five most abundant genes in each fraction are shown in italics.

(F) Biological processes enriched in the annotation of differentially expressed genes. Within each gene ontology (GO) term, the number of genes with higher expression in LS$^+$K (red) and LS$^-$K (blue) are indicated. Processes related to *immune response; **cell adhesion; ***transcriptional control, ****metabolism, or *****cell cycle are labeled. The odd ratios are calculated with respect to background. For complete list of differential GO terms, see Table S2.

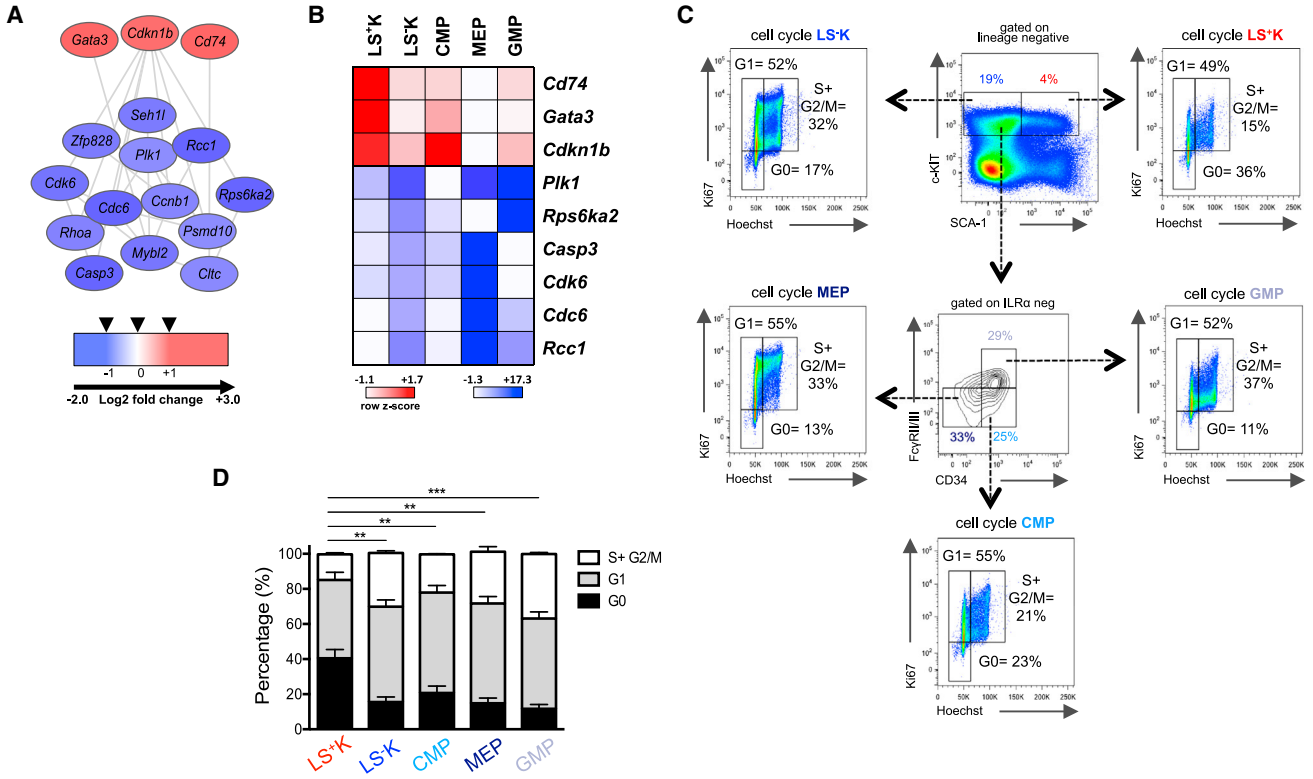


Figure 2. Cell-Cycle Stages of HSPCs and Myeloid Committed Precursors

(A) Protein-protein interaction analysis built based on GO terms related to cell cycle. Genes highly expressed in LS$^+$K cells are shown in red, and genes higher in LS$^-$K cells are shown in blue. Log2 fold change is represented.

(B) Heatmap of cell-cycle regulators. mRNA expression levels in LS$^+$K, LS$^-$K, CMPs, GMPs, and MEPs based on real-time PCR. *Z* scores were calculated per row to highlight expression differences per gene between populations. See Figures S4A and S4B.

(C and D) FACS analysis of cell-cycle stages of HSPCs and myeloid progenitors. (C) Shows representative gating scheme. (D) Quantification of three biological replicates. Error bars indicate ±SD of three independent biological replicates.

HSCs and MPP1 cells compared to later MPPs and LS$^-$Ks (Figure 3E). On the contrary, the high expression of *1700006J14Rik* in LS$^-$K was mostly due to its increased expression in MEPs, as shown by real-time PCR (Figure 3F). In addition to ENSEMBL-annotated lncRNAs, we predicted de novo transcripts using Cufflinks (Trapnell et al., 2010). In order to increase the confidence of the assemblies, we considered only multiexonic assemblies and we filtered these assemblies based on (1) their length (>200 bp), (2) proximity to annotated genes (>10 Mb), (3) coverage (>20 mean fragment counts across all the samples), and (4) lack of coding sequences (<100 consecutive codons). After these filters, we predicted 713 potential de novo lncRNAs, of which 149 were differentially expressed (FDR = 0.1; Table S4). This set represents a predictive starting point for further validation. Taken together, we provide a comprehensive expression landscape of transcriptional control factors and lncRNAs, which serves as a resource to investigate regulatory circuits controlling early adult hematopoiesis.

Regulation of Gene Expression in HSPCs and Myeloid Precursors

Next, we tested for differential usage of exons between LS$^+$K and LS$^-$K cells using DEXseq (Figures 4A and 4B; Anders et al., 2012). In total, 4,096 genes showed evidence of differential exon usage (http://www-huber.embl.de/DEULSK/testForDEU.html), out of which 755 were differentially expressed genes. Because we identified a transcriptional control cluster (Table S3) to be enriched in HSPCs, we further investigated exon usage differences within this regulatory network. We found 48 out of 177 genes showing differential exon usage (Figure 4A). Of these, we exemplify the MyoD family inhibitor, *Mdfi*, a regulator of Wnt signaling (Kusano and Raab-Traub, 2002). Its fifth exon was almost exclusively expressed in HSPCs, suggesting specific roles for different *Mdfi* isoforms in stem/progenitors (Figure 4B). Complementarily, we used MISO (Katz et al., 2010) to quantify annotated alternative splicing events and used DEXSeq to test for differences in these annotated splicing events between the LS$^+$K and LS$^-$K populations.

Using this approach, we found 214 differentially spliced events (FDR = 0.1; Table S5). These findings suggest transcript isoform regulation as an unexpectedly abundant regulatory mechanism at the transition from HSPCs to LS$^-$K.

In order to explore potential posttranscriptional regulation, we integrated these RNA-seq data with our previously reported proteome data set generated from the identical cell populations (Klimmeck et al., 2012) (Figures 4C–4F; Table S6). Ninety-eight percent of all quantified proteins were assigned to the respective gene identifier (4,919 out of 5,027; Figure 4C). Of these, 419 were found to be differentially expressed both at the RNA and protein level (FDR = 0.1). Overall, the correlation coefficient of the RNA and protein fold changes was 0.39 (Figure 4D). However, when we restricted the genes to those detected as differentially expressed on RNA and protein levels, the correlation coefficient increased substantially (R = 0.81; Figure 4D). We found 82% of the differentially expressed genes with consistent sign of fold changes in mRNA and protein levels (Figure 4E, 342 out of 419). In contrast, a group of 77 hits showed anticorrelation in their fold changes, with increased protein but decreased transcript levels in HSPCs compared to LS$^-$Ks (Figure 4E; Table 1). These anticorrelated genes were enriched for metabolic process and response to oxidative stress suggesting the involvement of posttranscriptional mechanisms in the regulation of these processes (Figure 4F). No significant anticorrelated hits with increased transcript but decreased protein level in HSPCs were found, suggesting less pronounced effects of paused translation or protein degradation during hematopoietic commitment.

Extended Self-Protective Signature Highlights Posttranscriptional Regulation

To extend our findings to differentially expressed genes related to immune stress response, a protein-protein interaction analysis was carried out (Figures 5 and S5; Table S7), integrating the GO biological process annotation of the 3,236 differentially expressed genes (* in Figure 1F) and our previously described proteome immune signature (Klimmeck et al., 2012). The resulting network covered a broad range of cellular compartments and gene classes including e.g., secreted inflammatory cytokines (*Cxcl2*, *Pf4*), plasma membrane receptors (*Cxcr1*, *Flt3*), and transcription regulators (*Mecom*, *Gata3*). Enrichment of self-protective processes in HSPCs highlighted by our proteome analysis (e.g., viral dsRNA sensors [*Ddx58*, *Oas3*]) was confirmed by the combined analysis on RNA and protein level (Figure S5), indicating transcriptional shutdown of these programs upon commitment. Conversely, all detected genes involved in stress response to unfolded protein at the endoplasmic reticulum (*Calr*, *Hsp90b*) showed consistent suppression at RNA and protein level in HSPCs (Figures 5 and S5). In contrast, many proteins involved in mitochondrial response to oxidative stress (*Nqo1*, *Sod2*) showed strong anticorrelation with decreased RNA but increased protein levels, suggesting a coordinated regulation of translation in HSPCs and whereas *Tfrc*, a modulator of antibacterial cellular iron homeostasis was increased on both RNA and protein level in myeloid LS$^-$K cells, its two ligands *Trf* and *Ltf* as well as iron transporter *Fth1* showed an anticorrelation. Finally, several factors involved in DNA repair (*Xrcc6*, *Xpc*) showed decreased RNA but increased protein expression in HSPCs, pointing to specific mechanisms leading to increased translation or protein stability (Figure 5). Taken together, by integrating transcriptome and proteome data, we achieved a global perspective on self-protective mechanisms in early hematopoiesis, which refines gene expression regulation to specific defense processes.

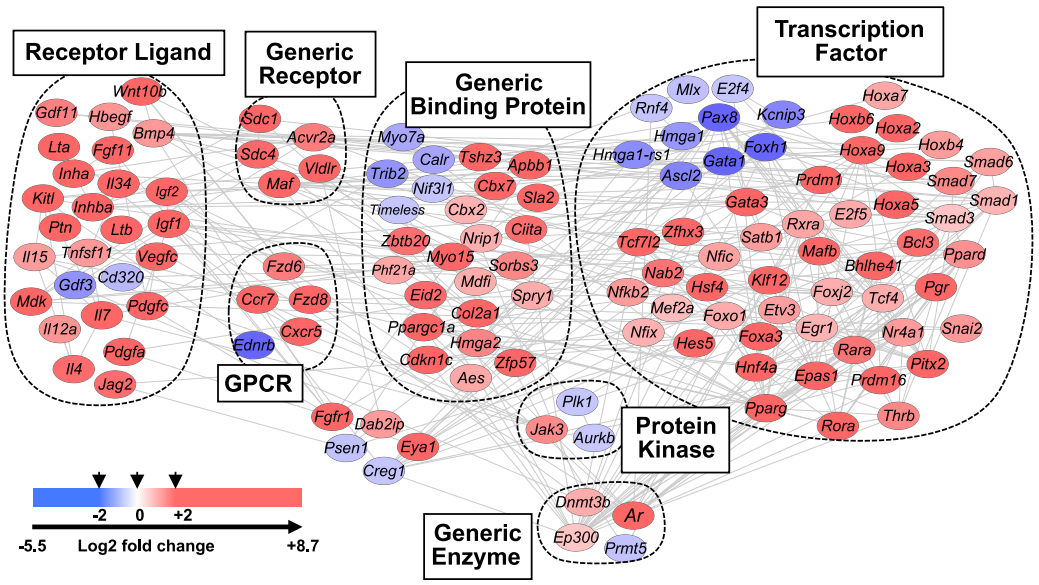
Energy Metabolism in HSPCs

Next, we characterized the regulation of gene expression in metabolic genes (Figure 6; Table S8). Strikingly, more than half (seven out of 13) of the compared glycolytic enzymes showed anticorrelation between protein and RNA differences (Figures 6A and 6B). Whereas at the first step of glycolysis the mRNAs of all three hexokinases *Hk1-3* were lowly expressed in HSPCs, only Hk1 showed higher expression at the protein level. At the final step of glycolysis, the pyruvate kinases Pkm1 and Pkm2 were highly expressed on protein levels but lowly expressed on RNA levels (Figures 6A and 6B). Furthermore, several enzymes involved in intracellular glycogen breakdown (*Pgm1* and *Pgm2*) showed an anticorrelation between RNA and protein, which points to posttranscriptional regulation of this alternative entry into glycolysis. Although we did not detect significant changes for most enzymes involved in the TCA cycle, these transcripts showed higher mRNA levels in LS$^-$K. In addition, our analysis revealed anticorrelation in a cluster of 46 enzymes that are likely to contribute to the energy-creating infrastructure of HSPCs (e.g., galactose [*Glb1*/Bgal], superoxide/NADPH [*Prdx6*], and acetyl CoA fatty acid [*Acads*] metabolism). In summary, whereas parts of energy metabolism are consistently regulated both at mRNA and protein level, many enzymes show anticorrelation pointing to posttranscriptional regulation in HSPCs and LS$^-$K.

DISCUSSION

In this study, we integrated transcriptome and proteome data and characterized the inventory of HSPCs and their immediate myeloid committed progeny (LS$^-$K) (Figure S6). With almost 20,000 identified expressed genes including more than 3,000 significantly regulated ones, these data

A



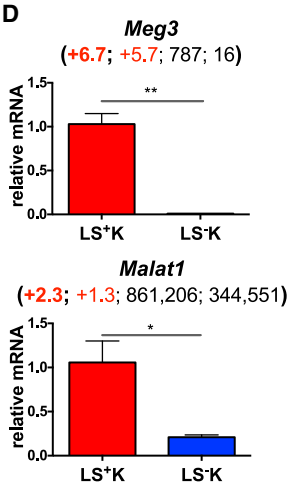
B **LncRNAs**

Gene symbol	Log2 fold change	p adj. value	Mean counts LS$^+$K	Mean counts LS$^-$K	Molecular function	Reference
H19	Inf	2.8 x 10^{-3}	70	0	Tumor suppressor/ oncogene	Yoshimizu 2008; Venkatraman 2013
A930001C03Rik	Inf	2.2 x 10^{-2}	32	0	unknown	
7SK	Inf	3.3 x 10^{-2}	30	0	unknown	
Gm12066	6.7	6.2 x 10^{-3}	112	1	unknown	
Meg3	5.7	1.1 x 10^{-5}	787	16	Tumor suppressor	Zhou 2012
Gm12575	5.6	1.2 x 10^{-4}	127	2	unknown	
Gm16070	5.1	8.5 x 10^{-13}	1,206	35	unknown	
D330050G23Rik	5.0	5.2 x 10^{-3}	91	3	unknown	
Gm13111	4.9	1.4 x 10^{-4}	155	5	unknown	
Gm13986	4.5	5.0 x 10^{-2}	415	18	unknown	
B430010I23Rik	-1.8	1.4 x 10^{-2}	86	292	unknown	
1190007F08Rik	-1.8	1.4 x 10^{-2}	1,822	6,277	unknown	
1700006J14Rik	-1.9	3.7 x 10^{-3}	1,514	5,443	unknown	
4933431E20Rik	-2.0	3.4 x 10^{-6}	425	1,645	unknown	
Gm13372	-2.1	1.7 x 10^{-2}	44	191	unknown	
Gm15290	-2.4	7.3 x 10^{-4}	58	302	unknown	
A730046J19Rik	-3.0	3.7 x 10^{-4}	25	201	unknown	
1700047F07Rik	-3.1	6.0 x 10^{-3}	12	99	unknown	
Gm12708	-3.2	1.9 x 10^{-2}	9	87	unknown	
Gm14705	-Inf	2.6 x 10^{-2}	0	29	unknown	

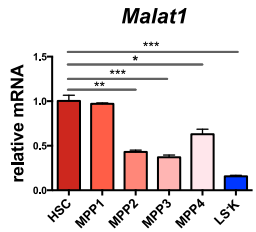
C

Malat1	1.3	1.7 x 10^{-1}	861,206	344,551	Oncogene	Gutschner 2013
AI662270	-0.8	5.3 x 10^{-2}	34,730	58,610	unknown	
4632427E13Rik	0.9	1.5 x 10^{-1}	21,752	11,576	unknown	
2810008D09Rik	0.9	1.1 x 10^{-2}	11,095	5,979	unknown	
Atp10d	1.4	3.9 x 10^{-2}	9,040	3,558	unknown	
1190007F08Rik	-1.8	1.4 x 10^{-2}	1,822	6,277	unknown	
2610307P16Rik	2.2	3.8 x 10^{-2}	6,545	1,468	unknown	
AI480526	1.1	6.0 x 10^{-2}	5,142	2,364	unknown	
2610035D17Rik	0.7	9.7 x 10^{-2}	4,329	2,724	unknown	
1700006J14Rik	1.3	3.7 x 10^{-3}	1,514	5,443	unknown	

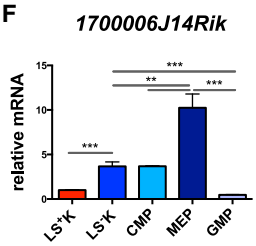
D



E



F



(legend on next page)

represent a deep analysis comparing these two cell populations that can be accessed as an interactive online resource (http://vega.embl.de/LSK).

The specific cell adhesion repertoire exclusively detected in HSPCs strongly supports the concept of a distinct autonomous microenvironment of HSPCs in the bone marrow (Hanoun and Frenette, 2013; Wilson and Trumpp, 2006), which probably serves important protective functions, among others. In line with this notion, secreted kallikrein-related proteases exert pleiotropic roles in a broad range of physiological processes including inflammatory response (Sotiropoulou et al., 2009) and thus might contribute to HSPC specific defense signaling via proteolytic events e.g., via TGF-beta- or Protease-activated receptor (PAR)-GPCR signaling. The immune response network presented in this study significantly extends our knowledge of distinct self-protective mechanisms elevated in HSPCs already at steady state and is in agreement with the recently observed reversible activation of HSCs under conditions of stress (Essers et al., 2009; Trumpp et al., 2010).

Maintenance of HSC quiescence under homeostatic conditions is tightly linked to mitochondrial oxidative response (Takubo et al., 2013; Yu et al., 2013). In particular, response to oxidative stress is known to be regulated by a variety of mechanisms including targeting of RNA-binding proteins, RNA half-life, and translation efficiency (Vogel et al., 2011). The comparison of the transcriptome and proteome data sets demonstrates an anticorrelation for central superoxide enzymes like *Gpx1* and *Sod2* and provides evidence for posttranscriptional regulation of ROS response and other defense mechanisms like iron metabolism in early hematopoiesis. Notably, expression of iron-response proteins (IRPs) like FTH1 or ACO2 is tightly regulated via iron-responsive elements (IREs) (Hentze et al., 2010). In addition, the expression of regulatory hormone Hepcidin, which is crucial for cellular iron homeostasis, has been shown to be strongly dependent on Bmp—(Hemojuvelin [HJV])-Smad signaling (Andriopoulos et al., 2009). Thus, the BMP-Smad signaling axis, of which we found its components to be highly expressed in HSPCs, might as well be involved in the regulation of iron homeostasis in multipotent cells.

Although most differentially expressed genes showed consistent expression on RNA and protein level and therefore may not require posttranscriptional regulation, our correlation analysis uncovered isozyme specific regulation for essential glycolytic enzymes (hexokinases, pyruvate kinases) on either RNA or protein level, which argues for blocked translation or increased protein turnover. The higher transcript levels of TCA cycle enzymes potentially boost oxidative phosphorylation after commitment. Notably, our findings highlight additional regulation of carboxylic acid metabolism, which allows novel perspectives on the coordination of metabolism along early hematopoiesis and suggests that energy metabolism as well as the response to immune mediated stress are modulated by different gene-regulatory mechanisms affecting differential and process-specific transcription, translation, mRNA, or protein turnover. The immune response signature described here at steady state should be complemented by a global analysis of induced stress response of HSPCs.

Our study identifies TFs and lncRNAs as candidates controlling multipotency and/or commitment. Indeed, loss of TFs including Foxo3a, one of the differentially abundant genes in our study, has been demonstrated to severely affect HSPC integrity in functional KO mouse studies (Rossi et al., 2012). Because some lncRNAs are known to facilitate expression of transcriptional regulators during development (Pauli et al., 2011), the differentially expressed

Figure 3. Differential Expression of Genes Involved in Transcriptional Control and Landscape of lncRNA in Multipotent and Myeloid Committed Progenitors

(A) Protein-protein interaction analysis built based on GO-terms related to transcriptional control. Each gene was functionally classified according to GeneGo MetaCore annotation and manual curation. Genes highly expressed in $LS^{+}K$ cells are shown in red, and genes higher expressed in $LS^{-}K$ cells are shown in blue. The color bar indicates the color code for Log2 fold change. For the complete list of differentially expressed genes involved in transcriptional control, see Table S3.

(B–E) Differential landscape of lncRNAs in multipotent and myeloid committed progenitors.

(B) Differentially abundant lncRNAs. Of the 67 lncRNAs with significant change in abundance, the table shows ten examples with changes in each direction. For the complete list, see Table S1. Inf, infinite number, meaning read count of one population is zero.

(C) Top abundant lncRNAs. Ten examples for highly abundant lncRNAs sorted by mean counts.

(D) Confirmation of differentially expressed lncRNAs by real-time PCR. mRNA expression levels of lncRNAs *Meg3* and *Malat1* in $LS^{+}K$ (red) and $LS^{-}K$ (blue) samples were quantified by real-time PCR. Numbers represent Log2 fold changes from real-time PCR (left, bold) and RNA-seq (right) data for $LS^{+}K$ compared to $LS^{-}K$. For all real-time PCR data, mean values means ± SD of three independent biological replicates with three technical replicates each are shown.

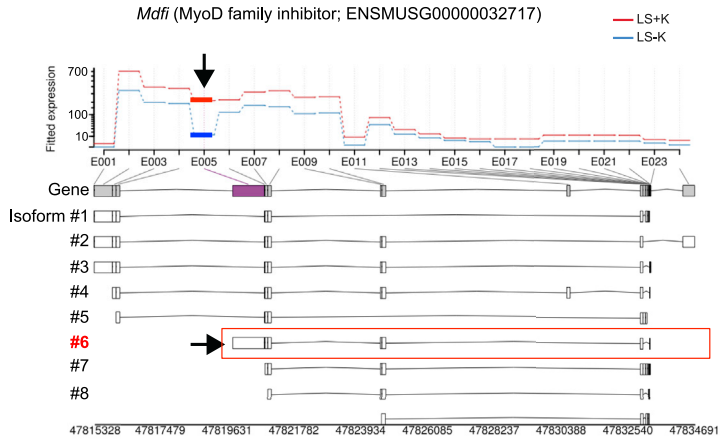
(E) Expression analysis of lncRNA *Malat1* in HSCs and MPPs. mRNA expression levels of *Malat1* in HSC, MPP1, MPP2, MPP3, MPP4 (Wilson et al., 2008), and $LS^{-}K$ cells were analyzed by real-time PCR.

(F) Expression analysis of lncRNA *1700006J14Rik* in myeloid committed precursors. RNA expression levels $LS^{+}K$, $LS^{-}K$, CMPs, GMPs, and MEPs based on real-time PCR. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.001$ (two-sided Student's t test).

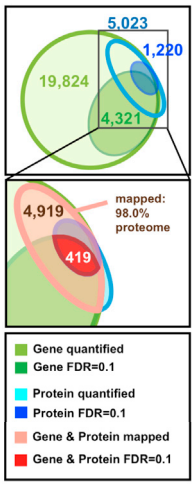
A

gene symbol	chr.	start	end	total exons	exon changes
Aes	10	81559494	81566362	7	5
Arap1	7	101348067	101412586	74	19
Calr	8	84841850	84846934	17	2
Cbx2	11	119022962	119031270	7	1
Cbx6	15	79823896	79834688	9	2
Ccr7	11	99144196	99155077	3	1
Cdkn1c	7	143458350	143461050	5	1
Creg1	1	165763746	165775308	7	1
Dab2ip	2	35558266	35730994	41	1
E2f4	8	105297663	105305369	10	7
Ep300	15	81586214	81652077	31	2
Epas1	17	86753907	86833410	16	1
Fgfr1	8	25513654	25575714	52	1
Foxp4	17	47867133	47924645	34	1
Gpi1	7	34202122	34230289	24	9
Hdgfrp3	7	81881251	81934473	10	3
Hmga1	17	27556620	27563674	18	1
Hmga2	10	120361275	120476469	8	2
Hnf4a	2	163506808	163572910	17	1
Hoxa7	6	52214491	52221854	13	1
Hoxa9	6	52223100	52231089	9	1
Igf1	10	87858265	87937042	20	1
Jak3	8	71676296	71688558	33	10
Kcnip3	2	127456498	127522094	14	1
Kitl	10	100015630	100100413	15	2
Mdfi*	**17**	**47815328**	**47834691**	**24**	**1**
Myo15	11	60469339	60528369	70	1
Myo7a	7	98051060	98119524	72	2
Nfic	10	81396186	81431005	21	3

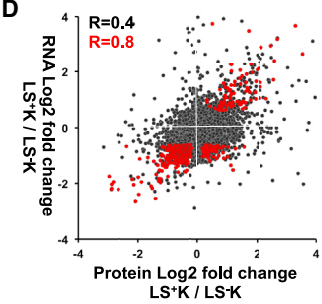
B



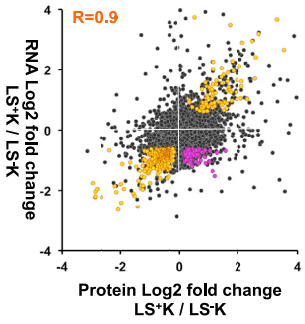
C



D



E



F

GO Id	Term	Count	p-value
GO:19752	Carboxylic Acid Metabolism Process	33/517	4.94E-17
	LDHB|ALDH18A1|NARS|ASS1|CTPS|PRKAG1|PGD|PRKAG2|ACOT2|ASNS|ACSF2|MTHFD1L|IARS|MTHFD1|ACSS1|ACSL1|IDH1|PRKAA1|HSD17B4|MARS|ACSL5|CHDH|P4HB|SHMT2|GATM|NCF1|ACLY|PADI4|GMPS|FRRS1|PKLR|LTA4H|PSAT1		
GO:6950	Response to Stress	38/1244	3.79E-07
	F13A1|PRKAG1|G6PDX|PRKAG2|ACOT2|PDIA5|AQP1|MANF|CASP3|TRP53BP1|NPM1|CAMK2D|IDH1|HSPA5|ERO1L|EIF2B3|EIF2B4|HSPA9|TRIP13|DNAJA2|H2-K1|PTPRC|NCF1|ELANE|EIF2B1|LSP1|DDX58|HYOU1|PNKP|ERP44|HSP90B1|PYGM|TIMELESS|EIF2S1|PKLR|MPO|HSPD1|CTSG		
GO:2376	Immune System Process	18/700	5.95E-03
	H2-K1|PTPRC|IL16|FLT3|NCF1|ELANE|G6PDX|DRG1|CDK6|IGHM|MEIS1|DDX58|CASP3|PLCG1|PLCG2|MPO|HSPD1|CTSG		
GO:8152	Metabolic Process	50/6470	1.05E-15
	IMPA1|KCNAB2|SAE1|GLB1|WARS|GPX1|GOT1|DKC1|SPR|NQO1|PDXK|ACO2|ACO1|ACADS|SUCLG2|SUCLG1|PTBP1|QDPR|LYPLA1|SIGMAR1|LAP3|PPP1CA|PRDX6|PGM1|CTSD|ADSL|6530401N04RIK|NHP2|ALDH9A1|TALDO1|SORD|HDGF|HK1|WBP11|PFAS|HNRNPM|TPI1|HNRNPK|AKR1A1|ETFDH|TCEA1|ACTL6A|GLO1|AGK|APEX1|SOD2|APRT|SLC25A12|SLC25A13|RBM14		
GO:5975	Carbohydrate Metabolic Process	11/433	3.78E-07
	TPI1|PPP1CA|SORD|TALDO1|GOT1|IMPA1|AKR1A1|PGM1|HK1|GLO1|GLB1		
GO:6979	Response to Oxidative Stress	6/108	1.02E-05
	GPX1|PRDX6|ETFDH|NQO1|APEX1|SOD2		

Figure 4. Gene Expression Regulation in Multipotency and Commitment

(A and B) Analysis of transcript isoform regulation.

(A) Differential exon usage in genes involved in transcriptional control. Forty-eight out of 177 genes exhibited exon usage changes between LS$^+$K and LS$^-$K cells in one to 12 exons (top 30 genes in alphabetical order are shown; for full annotation, see Table S3).

(B) Example: TF (*Mdfi*) shows differential exon usage. The data are consistent with higher levels of isoform #6 of *Mdfi* in LS$^+$K (red) compared to LS$^-$K (blue).

(C–F) Analysis of posttranscriptional regulation. (C) Overview of integration of transcriptome and proteome data. Transcriptome RNA-seq data (19,824 genes quantified), and proteome data (5,027 proteins quantified) were mapped to each other based on Uniprot IDs using BioMart. Ninety-eight percent of quantified proteins could be mapped (4,919 out of 5,027 total). (D) Overall correlation. Correlation of RNA to protein expression ratios (Log2 fold change) for all 4,919 genes. Four hundred nineteen hits were significant on both RNA and protein level (FDR = 0.1), shown in red. R, Pearson correlation coefficient. (E) Correlation of 419 hits significant on RNA and protein level (FDR = 0.1). (F) GO enrichment analysis of 419 hits significant on RNA and protein level (FDR = 0.1). GO enrichment analysis was carried out on 342 hits, which were consistently up- or downregulated (yellow), and 77 hits, which were anticorrelated with increased protein but decreased RNA ratios (pink).

Table 1. List of 77 Anticorrelating Hits between RNA and Protein Level

Uniprot ID	Ensembl Transcript ID	Gene Name	Protein Name	RNA-seq Log2 Fold Change LS^+K/LS^-K	RNA-seq adj. p Value	Proteome Log2 Fold Change LS^+K/LS^-K	Proteome adj. p Value	Differential Exon Usage
Q8BHA3	ENSMUSG00000020956	*6530401 N04Rik*	D-tyrosyl-tRNA(Tyr) deacylase 2	−0.70	0.0692	0.65	0.0901	no
Q07417	ENSMUSG00000029545	*Acads*	acyl-coenzyme A dehydrogenase, short chain, isoform CRA_a	−0.79	0.0350	0.53	0.0035	no
P28271	ENSMUSG00000028405	*Aco1*	aconitase 1	−0.98	0.0041	0.39	0.0599	no
Q99KI0	ENSMUSG00000022477	*Aco2*	aconitate hydratase, mitochondrial	−0.93	0.0058	0.26	0.0362	yes
Q9Z2N8	ENSMUSG00000027671	*Actl6a*	actin-like protein 6A;23 kDa protein	−0.75	0.0364	0.28	0.0888	no
P54822	ENSMUSG00000022407	*Adsl*	adenylosuccinate lyase	−0.68	0.0630	0.25	0.0350	no
Q9ESW4	ENSMUSG00000029916	*Agk*	acylglycerol kinase, mitochondrial	−0.71	0.0597	0.45	0.0270	yes
P50247	ENSMUSG00000048087	*Ahcy*	adenosylhomocysteinase	−1.20	0.0004	0.43	0.0050	no
Q9JII6	ENSMUSG00000028692	*Akr1a1*	alcohol dehydrogenase [$NADP^+$]	−0.88	0.0101	0.39	0.0423	yes
Q9JLJ2	ENSMUSG00000026687	*Aldh9a1*	aldehyde dehydrogenase 9A1	−1.05	0.0015	0.20	0.0548	no
P28352	ENSMUSG00000035960	*Apex1*	DNA-(apurinic or apyrimidinic site) lyase	−0.64	0.0997	0.67	0.0290	no
Q564P4	ENSMUSG00000006589	*Aprt*	adenine phosphoribosyltransferase	−0.85	0.0137	0.55	0.0871	no
P84078	ENSMUSG00000048076	*Arf1*	ADP-ribosylation factor 1	−1.02	0.0020	0.28	0.0798	yes
Q99PT1	ENSMUSG00000025132	*Arhgdia*	Rho GDP-dissociation inhibitor 1	−0.73	0.0391	0.18	0.0737	yes
P47754	ENSMUSG00000015733	*Capza2*	F-actin-capping protein subunit alpha-2	−0.69	0.0915	0.30	0.0708	yes
Q32P00	ENSMUSG00000057886	*Cbx3*	chromobox homolog 3	−0.81	0.0654	0.35	0.0082	no
Q91WS0	ENSMUSG00000037710	*Cisd1*	CDGSH iron sulfur domain-containing protein 1	−0.63	0.0942	0.43	0.0235	no
P61202	ENSMUSG00000027206	*Cops2*	COP9 signalosome complex subunit 2	−0.77	0.0316	0.28	0.0484	no
Q8C243	ENSMUSG00000007891	*Ctsd*	cathepsin D	−0.63	0.0903	0.27	0.0281	yes
Q91VR5	ENSMUSG00000037149	*Ddx1*	ATP-dependent RNA helicase DDX1	−0.87	0.0115	0.22	0.0858	yes
Q9ESX5	ENSMUSG00000031403	*Dkc1*	dyskeratosis congenita 1	−0.71	0.0492	0.27	0.0288	yes
P17182	ENSMUSG00000059040	*Eno1*	alpha-enolase	−1.04	0.0106	0.18	0.0798	no
Q921G7	ENSMUSG00000027809	*Etfdh*	electron transfer flavoprotein-ubiquinone oxidoreductase, mitochondrial	−0.77	0.0325	0.39	0.0206	no
Q78JE5	ENSMUSG00000032309	*Fbxo22*	F-box only protein 22	−0.73	0.0494	0.46	0.0539	no
P23780	ENSMUSG00000045594	*Glb1*	β-galactosidase	−0.69	0.0633	0.51	0.0447	no
Q9CPU0	ENSMUSG00000024026	*Glo1*	lactoylglutathione lyase	−0.74	0.0424	0.80	0.0101	no
Q3UJH8	ENSMUSG00000025190	*Got1*	aspartate aminotransferase 1	−0.73	0.0524	0.52	0.0167	no
P11352	ENSMUSG00000063856	*Gpx1*	glutathione peroxidase 1	−1.09	0.0216	0.37	0.0167	no

(*Continued on next page*)

Table 1. *Continued*

Uniprot ID	Ensembl Transcript ID	Gene Name	Protein Name	RNA-seq Log2 Fold Change $LS^{+}K/LS^{-}K$	RNA-seq adj. p Value	Proteome Log2 Fold Change $LS^{+}K/LS^{-}K$	Proteome adj. p Value	Differential Exon Usage
P51859	ENSMUSG00000004897	*Hdgf*	hepatoma-derived growth factor	−0.67	0.0684	0.27	0.0334	yes
Q9ERZ0	ENSMUSG00000028332	*Hemgn*	hemogen	−1.06	0.0580	1.01	0.0582	no
Q3UE51	ENSMUSG00000037012	*Hk1*	hexokinase 1	−0.69	0.0520	1.03	0.0004	yes
P17095	ENSMUSG00000046711	*Hmga1*	high mobility group protein A1	−1.36	0.0000	1.12	0.0562	yes
O54879	ENSMUSG00000015217	*Hmgb3*	high mobility group protein B3	−0.83	0.0171	0.42	0.0202	no
B2M1R6	ENSMUSG00000021546	*Hnrnpk*	heterogeneous nuclear ribonucleoprotein K	−0.81	0.0185	0.23	0.0494	yes
Q9D0E1	ENSMUSG00000059208	*Hnrnpm*	heterogeneous nuclear ribonucleoprotein M	−0.79	0.0235	0.19	0.0809	yes
Q3TME6	ENSMUSG00000027531	*Impa1*	inositol (myo)-1(or 4)-monophosphatase 1	−0.86	0.0124	0.65	0.0059	no
P62482	ENSMUSG00000028931	*Kcnab2*	voltage-gated potassium channel subunit beta-2	−0.68	0.0602	0.36	0.0782	yes
Q9CPY7	ENSMUSG00000039682	*Lap3*	cytosol aminopeptidase	−0.80	0.0225	0.24	0.0495	yes
P48678	ENSMUSG00000028063	*Lmna*	ILamin-A/C	−1.52	0.0051	1.19	0.0106	no
P97823	ENSMUSG00000025903	*Lypla1*	acyl-protein thioesterase 1	−0.84	0.0152	0.48	0.0220	yes
Q9CRB2	ENSMUSG00000001056	*Nhp2*	H/ACA ribonucleoprotein complex subunit 2	−0.79	0.0615	0.29	0.0862	no
Q9D0T1	ENSMUSG00000063543	*Nhp2l1*	NHP2-like protein 1	−1.01	0.0181	0.31	0.0283	no
Q64669	ENSMUSG00000003849	*Nqo1*	NAD(P)H dehydrogenase [quinone] 1	−1.05	0.0302	1.02	0.0430	no
Q63850	ENSMUSG00000043858	*Nup62*	nuclear pore glycoprotein p62	−0.88	0.0100	0.36	0.0560	no
Q8K183	ENSMUSG00000032788	*Pdxk*	pyridoxal kinase	−0.73	0.0432	0.50	0.0141	yes
P70296	ENSMUSG00000032959	*Pebp1*	phosphatidylethanolamine-binding protein 1	−0.76	0.0321	0.57	0.0027	no
Q5SUR0	ENSMUSG00000020899	*Pfas*	phosphoribosylformylglycinamidine synthase	−0.79	0.0240	0.19	0.0780	yes
Q9D0F9	ENSMUSG00000029171	*Pgm1*	phosphoglucomutase-1	−1.11	0.0007	0.89	0.0005	yes
Q7TSV4	ENSMUSG00000029171	*Pgm2*	phosphoglucomutase-2	−1.11	0.0007	0.64	0.0095	yes
P52480	ENSMUSG00000032294	*Pkm*	isoform M2 of Pyruvate kinase isozymes M1/M2	−0.76	0.0340	0.30	0.0317	yes
P52480	ENSMUSG00000032294	*Pkm*	isoform M1 of Pyruvate kinase isozymes M1/M2	−0.76	0.0340	1.27	0.0447	yes
P62137	ENSMUSG00000040385	*Ppp1ca*	serine/threonine-protein phosphatase PP1-alpha	−0.78	0.0263	0.18	0.0886	no
P58389	ENSMUSG00000039515	*Ppp2r4*	serine/threonine-protein phosphatase 2A regulatory subunit B	−0.71	0.0482	0.36	0.0218	yes
Q6A0D0	ENSMUSG00000026701	*Prdx6*	peroxiredoxin-6	−1.05	0.0012	0.55	0.0050	yes

(*Continued on next page*)

Table 1. ***Continued***

Uniprot ID	Ensembl Transcript ID	Gene Name	Protein Name	RNA-seq Log2 Fold Change LS^+K/LS^-K	RNA-seq adj. p Value	Proteome Log2 Fold Change LS^+K/LS^-K	Proteome adj. p Value	Differential Exon Usage
Q3U5I2	ENSMUSG00000006498	*Ptbp1*	polypyrimidine tract binding protein 1	−1.02	0.0019	0.26	0.0338	yes
Q8BVI4	ENSMUSG00000015806	*Qdpr*	dihydropteridine reductase	−0.63	0.0973	0.34	0.0087	no
Q8C2Q3	ENSMUSG00000006456	*Rbm14*	RNA-binding protein 14	−0.82	0.0264	0.64	0.0081	yes
Q9R1T2	ENSMUSG00000052833	*Sae1*	SUMO-activating enzyme subunit 1	−0.74	0.0381	0.18	0.0875	yes
Q9D154	ENSMUSG00000044734	*Serpinb1a*	leukocyte elastase inhibitor A	−0.69	0.0601	1.56	0.0029	no
Q9JJU8	ENSMUSG00000031246	*Sh3bgrl*	SH3 domain-binding glutamic acid-rich-like protein	−0.71	0.0494	0.27	0.0616	no
O55242	ENSMUSG00000036078	*Sigmar1*	sigma 1-type opioid receptor	−0.68	0.0663	0.34	0.0421	no
Q8BH59	ENSMUSG00000027010	*Slc25a12*	calcium-binding mitochondrial carrier protein Aralar1	−1.01	0.0027	0.72	0.0082	yes
Q9QXX4	ENSMUSG00000015112	*Slc25a13*	calcium-binding mitochondrial carrier protein Aralar2	−0.79	0.0334	0.42	0.0360	no
P09671	ENSMUSG00000006818	*Sod2*	superoxide dismutase [Mn], mitochondrial	−0.81	0.0203	0.60	0.0068	no
Q64442	ENSMUSG00000027227	*Sord*	sorbitol dehydrogenase	−0.68	0.0629	0.63	0.0108	yes
Q64105	ENSMUSG00000033735	*Spr*	sepiapterin reductase	−0.67	0.0762	0.87	0.0068	yes
Q9WUM5	ENSMUSG00000052738	*Suclg1*	succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial	−0.64	0.0855	0.33	0.0109	no
Q9Z2I8	ENSMUSG00000061838	*Suclg2*	succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial	−0.69	0.0617	0.57	0.0140	no
Q93092	ENSMUSG00000025503	*Taldo1*	transaldolase	−0.74	0.0386	0.20	0.0583	no
P10711	ENSMUSG00000033813	*Tcea1*	transcription elongation factor A protein 1	−0.74	0.0423	0.26	0.0562	yes
P17751	ENSMUSG00000023456	*Tpi1*	triosephosphate isomerase 1	−0.90	0.0078	0.86	0.0008	no
Q9QZE7	ENSMUSG00000056820	*Tsnax*	translin-associated protein X	−0.74	0.0408	0.76	0.0539	no
Q64727	ENSMUSG00000021823	*Vcl*	vinculin	−0.90	0.0178	0.28	0.0548	no
Q60932	ENSMUSG00000020402	*Vdac1*	voltage-dependent anion-selective channel protein 1	−1.00	0.0023	0.23	0.0698	no
P32921	ENSMUSG00000021266	*Wars*	tryptophanyl-tRNA synthetase, cytoplasmic	−0.66	0.0708	0.31	0.0430	yes
Q923D5	ENSMUSG00000030216	*Wbp11*	WW domain-binding protein 11	−0.81	0.0203	0.21	0.0895	no
P62259	ENSMUSG00000020849	*Ywhae*	14-3-3 protein epsilon	−0.79	0.0238	0.34	0.0086	no

FDR = 0.1 for RNA and protein.

lncRNAs identified might exert similar roles in HSPCs, leading to stabilization of multipotency. *Malat1*, which is upregulated in several human tumors and plays an important role in extravasation in lung cancer metastasis (Gutschner et al., 2013), might be involved in promoting stem cell motility and HSC-niche interactions in the bone marrow. In addition, because the vast majority of differentially expressed lncRNAs lack molecular and biological

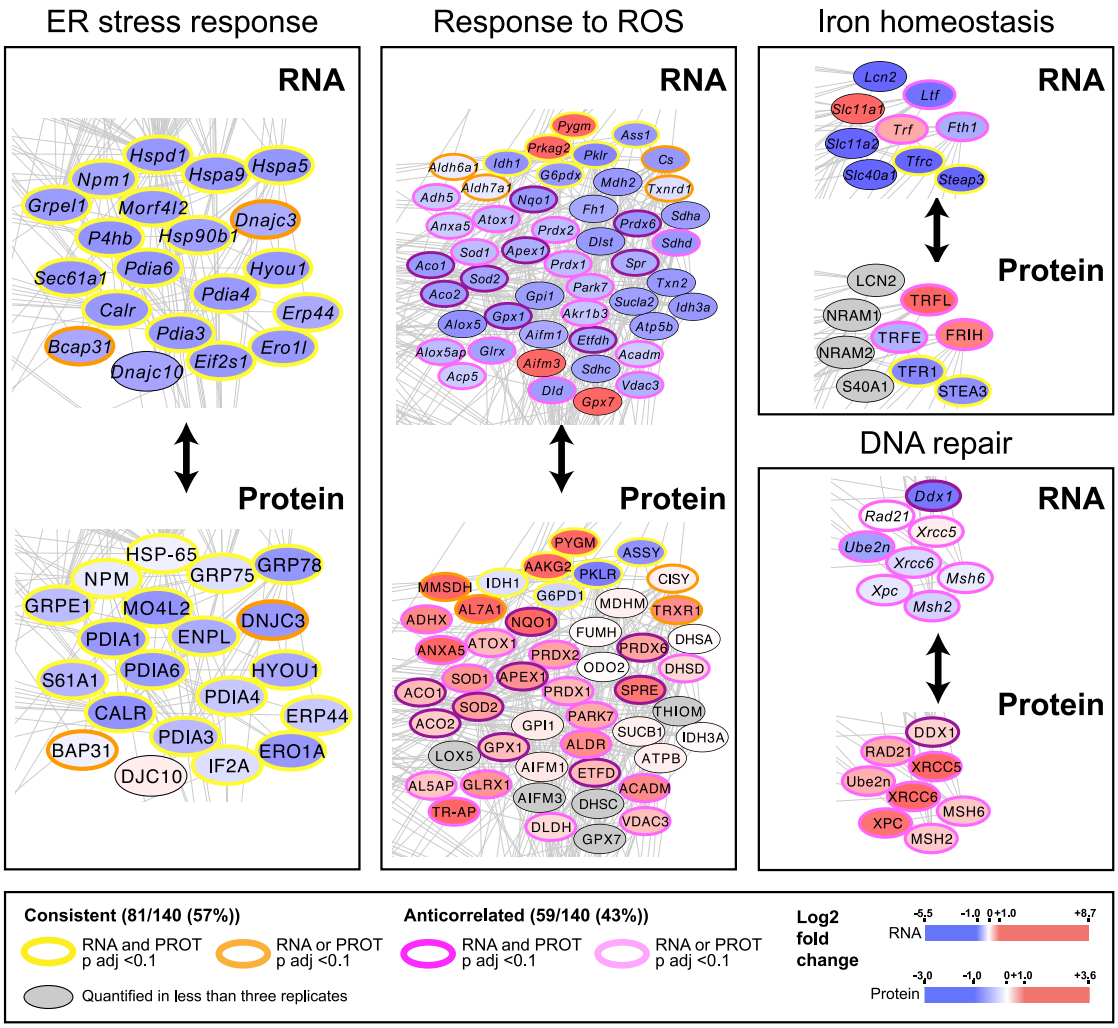


Figure 5. Gene Expression Regulation of Self-Protective Immune Response in Multipotency and Commitment

Protein-protein interaction visualization of significantly overrepresented GO processes related to immune response (see * in Figure 1F). For full list of involved differentially expressed genes, see Table S7. For high-resolution image of the interaction network, see Figure S5. Protein expression ratios were integrated for 138 out of 480 nodes total. Expression levels are displayed in color code (red, enriched in multipotent $LS^{+}K$; blue, enriched in $LS^{-}K$).

characterization, they are now attractive candidates for further exploration in vivo. In summary, in this study we provide the global landscape gene expression of protein-coding and noncoding transcripts during early hematopoiesis. It represents a comprehensive resource for the stem cell field and will serve as a valuable resource for functional exploration of self-renewal, multipotency, and lineage determination.

EXPERIMENTAL PROCEDURES

Animals

Eight- to 12-week-old female C57BL/6 mice purchased from Harlan Laboratories were used throughout the study. All mice were maintained in the animal facility at DKFZ, under specific pathogen-free (SPF) conditions and kept in individually ventilated cages (IVCs). Animal procedures were performed according to protocols approved by the German authorities, Regierungspräsidium Karlsruhe (Nr. Z110/02, DKFZ #261).

FACS Staining, Sorting, and Cell-Cycle Analysis

Bone marrow was isolated from hind legs (femur, tibia), hips (ilium), and backbone (vertebra). Muscle, connective tissue, and spinal cord were removed; bones were crushed in RPMI/2% FBS (GIBCO) using mortar and pestle. Single-cell suspensions were made by flushing through a 40 μm filter mesh. Cell numbers and viability were determined using a ViCell Counter (Beckman Coulter). To deplete lineage-positive cells, total bone marrow was stained 30 min with a combination of monoclonal rat

antibodies directed against mature cell specific lineage markers (for detailed specifications, see Supplemental Experimental Procedures). Labeled cells were incubated for 20 min with polyclonal sheep anti-rat immunoglobulin-G-coated magnetic Dynabeads (Invitrogen) at a ratio 2:1 beads to cell and depleted using a magnet, enriching for the lineage-negative (Lin^{neg}) cell fraction. Beads were washed twice with RPMI/2% FBS to harvest residual cell fractions. Centrifugation steps were carried out at 1,500 rpm and 4°C for 5 min (5810r, Eppendorf). To specify multipotent and myeloid progenitor fractions, as well as refined HSC and MPP populations, the Lin^{neg} fraction was stained 30 min using rat monoclonal fluorochrome-coupled antibodies (see Supplemental Experimental Procedures). All antibodies were titrated prior to use. Cell sorting was performed on a FACS Aria I or II (Becton Dickinson) at the DKFZ Flow Cytometry Service Unit, using the following sort parameters: 70 μm nozzle; 15,000 evt/s; 70 psi. $LS^{+}K$ cells and $LS^{-}K$ cells were obtained by sorting $Lin^{neg}SCA\text{-}1^{+}c\text{-}KIT^{+}$ and $Lin^{neg}SCA\text{-}1^{+}c\text{-}KIT^{-}$, respectively, in biological triplicate (Figures S1A and S1B). Sorted cells were collected into ice-cold RNA lysis buffer (ARCTURUS PicoPure RNA Isolation Kit [Life Technologies, Invitrogen]) and stored at −80°C until further usage. Of note, to determine sample's purity we resorted a fraction of each population (Figure S1B; purity greater than 95%). Cell-cycle analysis together with six-color surface staining to define myeloid subsets was performed as previously described (Wilson et al., 2004) using Ki67-FITC (BD Biosciences) and Hoechst 33342 (Molecular Probes).

Total RNA Isolation and RNA-Seq

Total RNA isolation was performed from the indicated populations using ARCTURUS PicoPure RNA Isolation Kit (Life Technologies, Invitrogen) according to the manufacturer's instructions. DNase treatment was performed using RNase-free DNase Set (QIAGEN). Total RNA was used for quality controls and for normalization of starting material (Figure S2). cDNA-libraries were generated with 10 ng of total RNA using the SMARTer Ultra Low RNA Kit for Illumina Sequencing (Clontech) according to the manufacturer's indications. Of note, 12 cycles were used for the amplification of cDNA, respectively. Paired-end adaptors were applied to each population. Sequencing was performed with the HiSeq2000 device (Illumina) and one sample per lane. Quality controls before and after sequencing and schematic overview of sampling workflow are shown in Figures S1 and S2.

Processing of RNA-Seq Data

RNA-seq reads were aligned to the reference genome of *Mus musculus* GRcm38 from ENSEMBL release 69 (Flicek et al., 2013) using GSNAP version 2012-07-20. The alignment quality statistics were computed using scripts based on the HTSeq Python library (Anders et al., 2014) and R/Bioconductor (Gentleman et al., 2004). The results from these statistics are available in Figure S2. Only uniquely aligned reads unambiguously assigned to annotated exons were considered for differential expression analysis. The differential expression analysis was done using DESeq (Anders and Huber, 2010), and the differential exon usage analysis (http://www-huber.embl.de/DEULSK/testForDEU.html) was done using DEXSeq (Anders et al., 2012). In RNA-seq count data, the power to detect differentially expressed genes varies widely through the dynamic range. Therefore, to avoid associated biases in gene set enrichment analysis, we generated background set of genes whose expression distribution approximated the expression distribution of our set of differentially expressed genes (Ho et al., 2007). Genes were called as quantified genes if they had at least 20 read counts per cell type. We tested differentially expressed genes from the RNA-seq data for overrepresented GO categories using Fisher's exact test. The R/Bioconductor scripts used for this study are available as an online resource (http://www-huber.embl.de/DEULSK/supplementary_file.pdf).

Gene Expression Analysis by Real-Time PCR

For real-time PCR, total RNA was isolated as described above and reverse-transcribed using Superscript III reverse transcriptase (Life Technologies, Invitrogen). The PCR was performed using the Fast SYBR Green Master Mix (Applied Biosystems) on a ViiA 7 Real-Time PCR System (Applied Biosystems) according to the manufacturer's instructions. Primers were designed using the Universal Probe Library Assay Design Center (Roche). For primers used, see Supplemental Experimental Procedures. The analysis of amplification curves was carried out using the ViiA 7 Software v1.1 (Applied Biosystems).

Bioinformatic Analysis

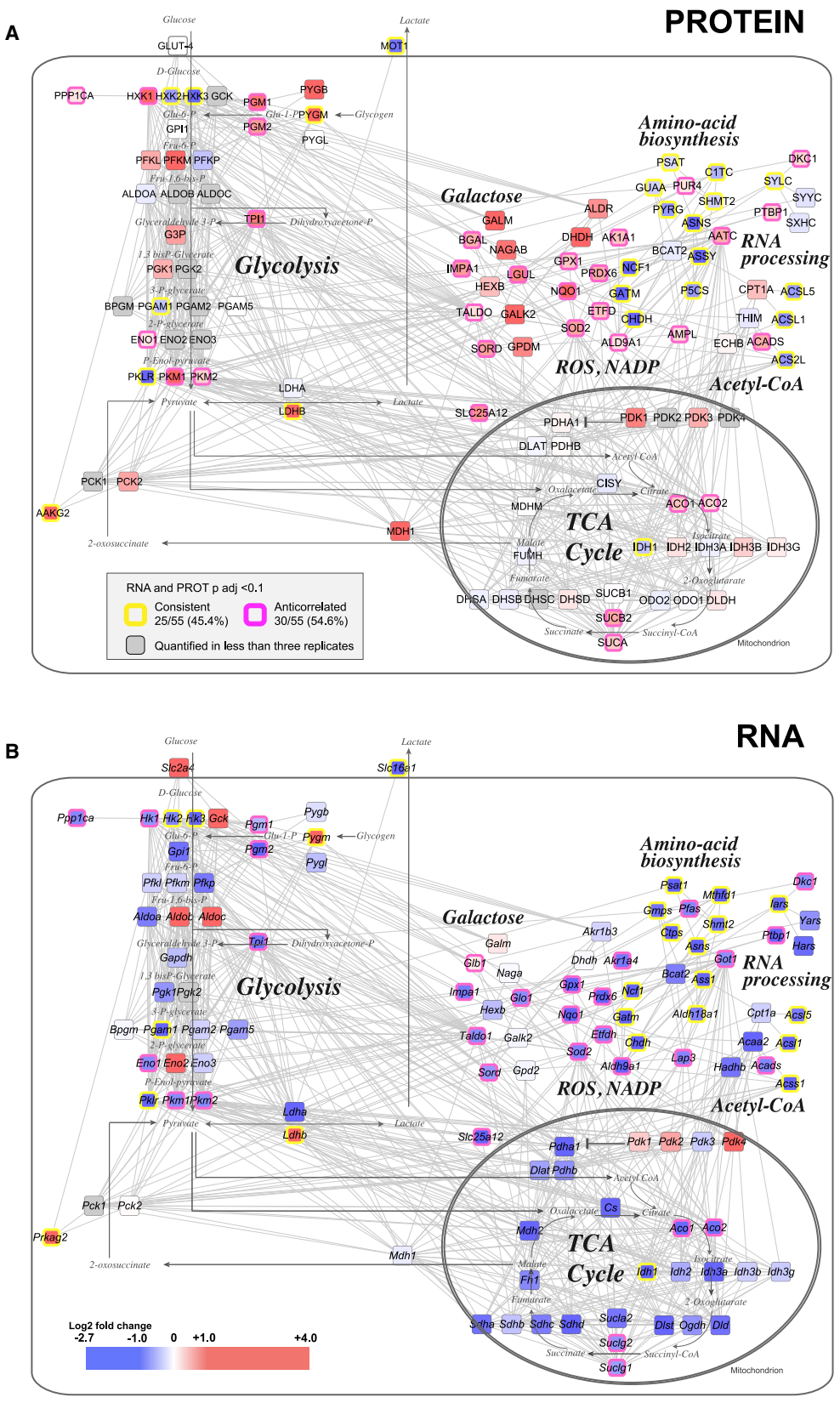
For Integration of transcriptomics and proteomics, the ENSEMBL identifiers were associated to their respective UniProt identifiers using BioMart (Durinck et al., 2005). For the purpose of this analysis, RNAs and proteins were considered differentially expressed if the adjusted p value <0.1. We tested for GO enriched categories using Fishers' exact test, the functional annotation tool of MetaCore (GeneGo; Nikolsky et al., 2005), protein classification by Panther (Mi et al., 2007), and the BiNGO plugin (Maere et al., 2005) for Cytoscape (Shannon et al., 2003). To investigate interactions between differentially expressed proteins, networks were constructed using STRING (Szklarczyk et al., 2011). Interaction networks were visualized using Cytoscape; nodes were arranged after manual curation according to Gene Ontology (UniProt; (Jain et al., 2009), extensive literature search, and STRING interaction scores. To identify possible transcription factor targets, the Transcription Factor Encyclopedia (Yusuf et al., 2012) was used as a resource.

ACCESSION NUMBERS

The RNA-seq data sets from $LS^{+}K$ and $LS^{-}K$ cells can be accessed through ArrayExpress (https://www.ebi.ac.uk/arrayexpress; accession number E-MTAB-1963). An interactive online resource that also integrates the previous proteome data (Klimmeck et al., 2012) can be found under the following link: http://vega.embl.de/LSK.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2014.08.012.



(legend on next page)

AUTHOR CONTRIBUTIONS

D.K., N.C.-W., and A.T. designed and coordinated the study. A.T. and W.H. (bioinformatics) designed and supervised the experiments and interpreted the data. N.C.-W., A.R., and D.K performed the RNA-seq experiments, bioinformatic analysis, and data interpretation. A.R. built the interactive online data resource. D.K. and N.C.-W. coordinated the animal experiments and performed FACS. L.v.P., S.R., and D.K. performed real-time PCR. D.K., N.C.-W., A.R., and J.H. together with J.K., W.H., and A.T. wrote the manuscript.

ACKNOWLEDGMENTS

We thank S. Wolf, S. Schmidt, U. Ernst, A. Hotz-Wagenblatt, and C. Previti from the Genomics Proteomics Core Facility and A. Atzberger and S. Schmitt from the Flow Cytometry Core Facility at DKFZ for their expert assistance. This work was supported by the BioRN Spitzencluster "Molecular and Cell Based Medicine" supported by the German Bundesministerium für Bildung und Forschung, the Sonderforschungsbereich SFB873 funded by the Deutsche Forschungsgemeinschaft, and the Dietmar Hopp Foundation (A.T.). W.H and A.R. acknowledge funding from the European Community's Seventh Framework Programme (FP7/2007-2013), project RADIANT (grant agreement no. 305626).

Received: April 27, 2014
Revised: August 20, 2014
Accepted: August 21, 2014
Published: September 25, 2014

REFERENCES

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell *121*, 295–306.

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193–197.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. Genome Res. *22*, 2008–2017.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq–a Python framework to work with high-throughput sequencing data. bioRxiv. http://dx.doi.org/10.1101/002824.

Andriopoulos, B., Jr., Corradini, E., Xia, Y., Faasse, S.A., Chen, S., Grgurevic, L., Knutson, M.D., Pietrangelo, A., Vukicevic, S., Lin, H.Y., and Babitt, J.L. (2009). BMP6 is a key endogenous regulator of hepcidin expression and iron metabolism. Nat. Genet. *41*, 482–487.

Bigas, A., and Espinosa, L. (2012). Hematopoietic stem cells: to be or Notch to be. Blood *119*, 3226–3235.

Blank, U., and Karlsson, S. (2011). The role of Smad signaling in hematopoiesis and translational hematology. Leukemia *25*, 1379–1388.

Cox, J., and Mann, M. (2007). Is proteomics the new genomics? Cell *130*, 395–398.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439–3440.

Essers, M.A., Offner, S., Blanco-Bose, W.E., Waibler, Z., Kalinke, U., Duchosal, M.A., and Trumpp, A. (2009). IFNalpha activates dormant haematopoietic stem cells in vivo. Nature *458*, 904–908.

Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., et al. (2013). Ensembl 2013. Nucleic Acids Res. *41*, D48–D55.

Gazit, R., Garrison, B.S., Rao, T.N., Shay, T., Costello, J., Ericson, J., Kim, F., Collins, J.J., Regev, A., Wagers, A.J., and Rossi, D.J.; Immunological Genome Project Consortium (2013). Transcriptome analysis identifies regulators of hematopoietic stem and progenitor cells. Stem Cell Rev. *1*, 266–280.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

Graf, T., and Enver, T. (2009). Forcing cells to change lineages. Nature *462*, 587–594.

Gutschner, T., and Diederichs, S. (2012). The hallmarks of cancer: a long non-coding RNA point of view. RNA Biol. *9*, 703–719.

Gutschner, T., Hämmerle, M., Eissmann, M., Hsu, J., Kim, Y., Hung, G., Revenko, A., Arun, G., Stentrup, M., Gross, M., et al. (2013). The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res. *73*, 1180–1189.

Hanoun, M., and Frenette, P.S. (2013). This niche is a maze; an amazing niche. Cell Stem Cell *12*, 391–392.

Hentze, M.W., Muckenthaler, M.U., Galy, B., and Camaschella, C. (2010). Two to tango: regulation of Mammalian iron metabolism. Cell *142*, 24–38.

Ho, D.E., Imai, K., King, G., and Stuart, E.A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit. Anal. *15*, 199.

Figure 6. Gene Expression Regulation of Energy Metabolism

Protein-protein interaction analysis of significantly overrepresented GO processes related to metabolism (**** in Figure 1F). This network was imposed on a pathway map based on GeneGo pathway maps of glycolysis and gluconeogenesis. Expression levels are displayed in color code (red, enriched in multipotent $LS^{+}K$; blue, enriched in $LS^{-}K$). For full list of involved differential hits, see Table S8.

(A) Protein expression levels.

(B) Gene expression levels. Genes with p adj <0.1 for both RNA and protein are shown with yellow (consistent changes) or pink boxes (anticorrelating changes). Gray squares indicate quantification in less than three replicates.

Jain, E., Bairoch, A., Duvaud, S., Phan, I., Redaschi, N., Suzek, B.E., Martin, M.J., McGarvey, P., and Gasteiger, E. (2009). Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinformatics *10*, 136.

Katz, Y., Wang, E.T., Airoldi, E.M., and Burge, C.B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat. Methods *7*, 1009–1015.

Kent, D., Copley, M., Benz, C., Dykstra, B., Bowie, M., and Eaves, C. (2008). Regulation of hematopoietic stem cells by the steel factor/KIT signaling pathway. Clin. Cancer Res. *14*, 1926–1930.

Klimmeck, D., Hansson, J., Raffel, S., Vakhrushev, S.Y., Trumpp, A., and Krijgsveld, J. (2012). Proteomic cornerstones of hematopoietic stem cell differentiation: distinct signatures of multipotent progenitors and myeloid committed cells. Mol. Cell. Proteomics *11*, 286–302.

Korkmaz, B., Moreau, T., and Gauthier, F. (2008). Neutrophil elastase, proteinase 3 and cathepsin G: physicochemical properties, activity and physiopathological functions. Biochimie *90*, 227–242.

Kusano, S., and Raab-Traub, N. (2002). I-mfa domain proteins interact with Axin and affect its regulation of the Wnt and c-Jun N-terminal kinase signaling pathways. Mol. Cell. Biol. *22*, 6393–6405.

Luis, T.C., Ichii, M., Brugman, M.H., Kincade, P., and Staal, F.J. (2012). Wnt signaling strength regulates normal hematopoiesis and its deregulation is involved in leukemia development. Leukemia *26*, 414–421.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics *21*, 3448–3449.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nat. Rev. Genet. *10*, 155–159.

Mi, H., Guo, N., Kejariwal, A., and Thomas, P.D. (2007). PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. Nucleic Acids Res. *35*, D247–D252.

Moignard, V., Macaulay, I.C., Swiers, G., Buettner, F., Schütte, J., Calero-Nieto, F.J., Kinston, S., Joshi, A., Hannah, R., Theis, F.J., et al. (2013). Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. Nat. Cell Biol. *15*, 363–372.

Murphy, M.J., Wilson, A., and Trumpp, A. (2005). More than just proliferation: Myc function in stem cells. Trends Cell Biol. *15*, 128–137.

Nikolsky, Y., Ekins, S., Nikolskaya, T., and Bugrim, A. (2005). A novel method for generation of signature networks as biomarkers from complex high throughput data. Toxicol. Lett. *158*, 20–29.

Ozsolak, F., and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. Nat. Rev. Genet. *12*, 87–98.

Paralkar, V.R., and Weiss, M.J. (2013). Long noncoding RNAs in biology and hematopoiesis. Blood *121*, 4842–4846.

Pauli, A., Rinn, J.L., and Schier, A.F. (2011). Non-coding RNAs as regulators of embryogenesis. Nat. Rev. Genet. *12*, 136–149.

Pronk, C.J., Rossi, D.J., Månsson, R., Attema, J.L., Norddahl, G.L., Chan, C.K., Sigvardsson, M., Weissman, I.L., and Bryder, D. (2007). Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. Cell Stem Cell *1*, 428–442.

Purton, L.E., and Scadden, D.T. (2007). Limiting factors in murine hematopoietic stem cell assays. Cell Stem Cell *1*, 263–270.

Qureshi, I.A., and Mehler, M.F. (2012). Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. Nat. Rev. Neurosci. *13*, 528–541.

Rossi, L., Lin, K.K., Boles, N.C., Yang, L., King, K.Y., Jeong, M., Mayle, A., and Goodell, M.A. (2012). Less is more: unveiling the functional core of hematopoietic stem cells through knockout mice. Cell Stem Cell *11*, 302–317.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. Wiley Interdiscip. Rev. Syst. Biol. Med. *2*, 640–653.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Sotiropoulou, G., Pampalakis, G., and Diamandis, E.P. (2009). Functional roles of human kallikrein-related peptidases. J. Biol. Chem. *284*, 32989–32994.

Spooncer, E., Brouard, N., Nilsson, S.K., Williams, B., Liu, M.C., Unwin, R.D., Blinco, D., Jaworska, E., Simmons, P.J., and Whetton, A.D. (2008). Developmental fate determination and marker discovery in hematopoietic stem cell biology using proteomic fingerprinting. Mol. Cell. Proteomics *7*, 573–581.

Sun, D., Luo, M., Jeong, M., Rodriguez, B., Xia, Z., Hannah, R., Wang, H., Le, T., Faull, K.F., Chen, R., et al. (2014). Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. Cell Stem Cell *14*, 673–688.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., et al. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. *39*, D561–D568.

Takubo, K., Nagamatsu, G., Kobayashi, C.I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., Goda, N., Rahimi, Y., Johnson, R.S., Soga, T., et al. (2013). Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. Cell Stem Cell *12*, 49–61.

Tesio, M., and Trumpp, A. (2011). Breaking the cell cycle of HSCs by p57 and friends. Cell Stem Cell *9*, 187–192.

Till, J.E., and McCulloch, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. *14*, 213–222.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Trumpp, A., Essers, M., and Wilson, A. (2010). Awakening dormant haematopoietic stem cells. Nat. Rev. Immunol. *10*, 201–209.

Uchida, S., Gellert, P., and Braun, T. (2012). Deeply dissecting stemness: making sense to non-coding RNAs in stem cells. Stem Cell Rev. *8*, 78–86.

Venkatraman, A., He, X.C., Thorvaldsen, J.L., Sugimura, R., Perry, J.M., Tao, F., Zhao, M., Christenson, M.K., Sanchez, R., Yu, J.Y., et al. (2013). Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. Nature *500*, 345–349.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. *13*, 227–232.

Vogel, C., Silva, G.M., and Marcotte, E.M. (2011). Protein expression regulation under oxidative stress. Mol. Cell. Proteomics *10*, 009217.

Weissman, I.L., and Shizuru, J.A. (2008). The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. Blood *112*, 3543–3553.

Wilson, A., and Trumpp, A. (2006). Bone-marrow haematopoietic-stem-cell niches. Nat. Rev. Immunol. *6*, 93–106.

Wilson, A., Murphy, M.J., Oskarsson, T., Kaloulis, K., Bettess, M.D., Oser, G.M., Pasche, A.C., Knabenhans, C., Macdonald, H.R., and Trumpp, A. (2004). c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. Genes Dev. *18*, 2747–2763.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R.C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C.F., Eshkind, L., Bockamp, E., et al. (2008). Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. Cell *135*, 1118–1129.

Wilson, A., Laurenti, E., and Trumpp, A. (2009). Balancing dormant and self-renewing hematopoietic stem cells. Curr. Opin. Genet. Dev. *19*, 461–468.

Yoon, J.H., Abdelmohsen, K., and Gorospe, M. (2013). Posttranscriptional gene regulation by long noncoding RNA. J. Mol. Biol. *425*, 3723–3730.

Yoshimizu, T., Miroglio, A., Ripoche, M.A., Gabory, A., Vernucci, M., Riccio, A., Colnot, S., Godard, C., Terris, B., Jammes, H., and Dandolo, L. (2008). The H19 locus acts in vivo as a tumor suppressor. Proc. Natl. Acad. Sci. USA *105*, 12417–12422.

Yu, W.M., Liu, X., Shen, J., Jovanovic, O., Pohl, E.E., Gerson, S.L., Finkel, T., Broxmeyer, H.E., and Qu, C.K. (2013). Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. Cell Stem Cell *12*, 62–74.

Yusuf, D., Butland, S.L., Swanson, M.I., Bolotin, E., Ticoll, A., Cheung, W.A., Zhang, X.Y., Dickman, C.T., Fulton, D.L., Lim, J.S., et al. (2012). The transcription factor encyclopedia. Genome Biol. *13*, R24.

Zhou, Y., Zhang, X., and Klibanski, A. (2012). MEG3 noncoding RNA: a tumor suppressor. J. Mol. Endocrinol. *48*, R45–R53.

Zou, P., Yoshihara, H., Hosokawa, K., Tai, I., Shinmyozu, K., Tsukahara, F., Maru, Y., Nakayama, K., Nakayama, K.I., and Suda, T. (2011). p57(Kip2) and p27(Kip1) cooperate to maintain hematopoietic stem cell quiescence through interactions with Hsc70. Cell Stem Cell *9*, 247–261.