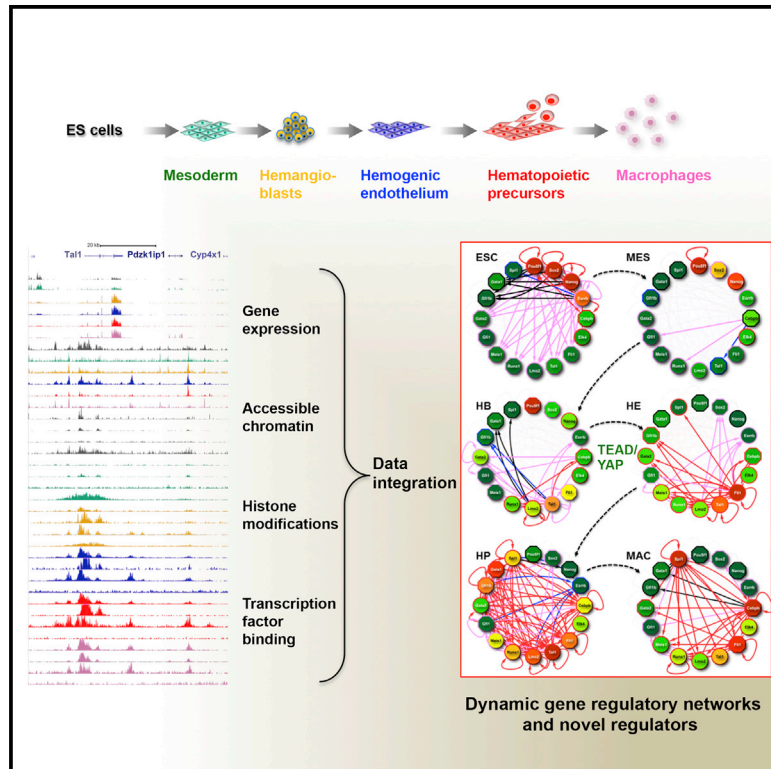


# Developmental Cell

## Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation

### Graphical Abstract



### Authors

Debbie K. Goode, Nadine Obier, M.S. Vijayabaskar, ..., Valerie Kouskoff, Berthold Göttgens, Constanze Bonifer

### Correspondence

c.bonifer@bham.ac.uk

### In Brief

Goode, Obier, Vijayabaskar et al. isolate cells at six different stages of hematopoietic differentiation, starting from embryonic stem cells, and perform a comprehensive multi-omics analysis of this developmental pathway. The data identify regulators of hematopoietic specification and highlight the minimum requirements for the reprogramming of non-blood cells to blood.

### Highlights

- Comprehensive genome-scale resource for studying embryonic blood cell specification
- Genome-scale definition of *cis* elements driving differential gene expression
- A gene regulatory network model for hematopoiesis aiding reprogramming experiments
- Analysis suggests a role for TEAD factors in hematopoietic specification

### Accession Numbers

GSE69101



# Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation

Debbie K. Goode,<sup>1,5</sup> Nadine Obier,<sup>2,5</sup> M.S. Vijayabaskar,<sup>4,5</sup> Michael Lie-A-Ling,<sup>3</sup> Andrew J. Lilly,<sup>3</sup> Rebecca Hannah,<sup>1</sup> Monika Lichtinger,<sup>2</sup> Kiran Batta,<sup>3</sup> Magdalena Florkowska,<sup>3</sup> Rahima Patel,<sup>3</sup> Mairi Challinor,<sup>3</sup> Kirstie Wallace,<sup>3</sup> Jane Gilmour,<sup>2</sup> Salam A. Assi,<sup>2</sup> Pierre Cauchy,<sup>2</sup> Maarten Hoogenkamp,<sup>2</sup> David R. Westhead,<sup>4,6</sup> Georges Lacaud,<sup>3,6</sup> Valerie Kouskoff,<sup>3,6</sup> Berthold Göttgens,<sup>1,6</sup> and Constanze Bonifer<sup>2,6,\*</sup>

<sup>1</sup>Department of Haematology, Cambridge Institute for Medical Research and Wellcome Trust and MRC Cambridge Stem Cell Institute, Cambridge CB2 0XY, UK

<sup>2</sup>Institute of Cancer and Genomic Sciences, College of Medicine and Dentistry, University of Birmingham, Birmingham B152TT, UK

<sup>3</sup>CRUK Manchester Institute, University of Manchester, Manchester M20 4BX, UK

<sup>4</sup>School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

<sup>5</sup>Co-first author

<sup>6</sup>Co-senior author

\*Correspondence: [c.bonifer@bham.ac.uk](mailto:c.bonifer@bham.ac.uk)

<http://dx.doi.org/10.1016/j.devcel.2016.01.024>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## SUMMARY

Metazoan development involves the successive activation and silencing of specific gene expression programs and is driven by tissue-specific transcription factors programming the chromatin landscape. To understand how this process executes an entire developmental pathway, we generated global gene expression, chromatin accessibility, histone modification, and transcription factor binding data from purified embryonic stem cell-derived cells representing six sequential stages of hematopoietic specification and differentiation. Our data reveal the nature of regulatory elements driving differential gene expression and inform how transcription factor binding impacts on promoter activity. We present a dynamic core regulatory network model for hematopoietic specification and demonstrate its utility for the design of reprogramming experiments. Functional studies motivated by our genome-wide data uncovered a stage-specific role for TEAD/YAP factors in mammalian hematopoietic specification. Our study presents a powerful resource for studying hematopoiesis and demonstrates how such data advance our understanding of mammalian development.

## INTRODUCTION

Cellular identities in multicellular organisms are defined by their individual gene expression programs and are established in a series of cell fate changes starting from pluripotent cells of the embryo. The information on the balanced and coordinated up- and downregulation of gene expression is encoded in our genome and is read by transcription factors (TFs), which interact with the epigenetic regulatory machinery to program the chromatin of lineage-specific genes into active and inactive states. To un-

derstand the mechanisms by which TFs establish and maintain specific transcriptional programs, it is essential to investigate developing biological systems, as illustrated by studies in non-vertebrate models (Van Nostrand and Kim, 2011; Zinzen et al., 2009).

Embryonic blood cells arise from early mesodermal cells via hemangioblast and hemogenic endothelial intermediates (Medvinsky et al., 2011). Studies of chromatin programming and gene expression during the generation of mature blood cells from hematopoietic stem cells were instrumental in defining the concept that development at the level of chromatin is a gradual and hierarchical process starting long before the overt transcriptional activation of lineage-specific genes (Bonifer et al., 2008; Hoogenkamp et al., 2009; Org et al., 2015; Wamstad et al., 2012; Wang et al., 2015). This notion is illustrated by the regulatory circuit essential for macrophage differentiation, the gene encoding TF PU.1 (*Spi1*), and its target, the *Csf1r* growth factor receptor gene (reviewed in Bonifer et al., 2008). Both are targets of RUNX1, but *Spi1* expression is induced prior to *Csf1r*. Early *Spi1* induction follows an initial enhancer priming event by TFs upstream of RUNX1 followed by upregulation via autoregulation (Leddin et al., 2011; Lichtinger et al., 2012), whereas subsequent full expression of *Csf1r* requires the concerted action of RUNX1, PU.1, and PU.1-induced factors (Krysinska et al., 2007; Lichtinger et al., 2012). This example illustrates the complexity of the molecular mechanisms underlying the establishment of cell-type-specific expression profiles. However, the global transcriptional control mechanisms underlying such dynamic progression events have remained largely obscure, because of a lack of comprehensive information on TF binding and the dynamic nature of the chromatin template with which they interact. We also know very little about how such transcriptional control mechanisms are interlinked with outside signaling.

The developmental hierarchies of early embryonic hematopoiesis are recapitulated in differentiating embryonic stem cells (ESCs) (Lancrin et al., 2010), which provide a tractable system capable of generating the cell numbers required for performing multiple genome-wide assays on the same samples. Recent studies have investigated the function of individual regulators at specific

developmental stages, such as early mesodermal patterning functions of the TF SCL/TAL1 and the RUNX1-controlled transition from hemogenic endothelium to hematopoietic progenitors (HPs) (Lancrin et al., 2012; Lichtinger et al., 2012; Lie-A-Ling et al., 2014; Liu et al., 2015; Tanaka et al., 2012). However, while a number of studies have examined individual cell fate transitions or investigated the differentiation of mature blood cells from hematopoietic stem cells (Garber et al., 2012; Lara-Astiaso et al., 2014; Tsankov et al., 2015), no study to date has reported an integrated genome-scale analysis of an entire developmental time course from early ESCs to fully defined blood cells.

In this study, we surveyed the global transcriptional journey from the ESC to the terminally differentiated state of macrophages via blood precursor cells by generating data for RNA sequencing (RNA-seq), DNase sequencing (DNA-seq), and chromatin immunoprecipitation sequencing (ChIP-seq) for histone marks and 16 different TFs across six sequential developmental stages. To facilitate access across the wider scientific community, we have integrated all genome-scale datasets into an online resource with advanced browse, search, and analysis capabilities. We have exploited our datasets to assemble a core regulatory network model that was able to inform the design of TF-mediated reprogramming strategies for the production of blood cells from fibroblasts. Furthermore, computational analysis of regulatory elements revealed the nature of TFs involved in stage-specific priming of distal elements, and informed functional validation experiments identifying TEAD/YAP interaction as a stage-specific regulator of early murine blood specification *in vitro* and *in vivo*. Finally, we identified TEAD target genes and their associated pathways, thus significantly enhancing our understanding of the signaling processes driving embryonic blood cell development.

## RESULTS

### Capturing a Complete Developmental Pathway using Genome-Scale Technologies

To study the specification of hematopoietic cells and their further differentiation, we employed mouse ESC *in vitro* differentiation to purify well-defined intermediate cell populations en route from pluripotent ESCs to adherent macrophages (Lancrin et al., 2009; Sroczynska et al., 2009), making use of a *Brachyury* GFP reporter (Fehling et al., 2003) and surface marker expression. Full details of this strategy are given in Figure S1A. In brief, pluripotent ESCs differentiate to mesoderm (MES) cells (Bry:GFP<sup>+</sup>/Flk1<sup>-</sup>), which then progress to the hemangioblast (HB) stage (Bry:GFP<sup>+</sup>/Flk1<sup>+</sup>) with smooth muscle, endothelial, and hematopoietic potential, followed by the hemogenic endothelium (HE) stage that has both endothelial and hematopoietic potential (CD41<sup>-</sup>/Tie2<sup>+</sup>/Kit<sup>+</sup>). HE cells then undergo the endothelial-hematopoietic transition (EHT) involving a shape change, after which they are fully committed to blood (CD41<sup>+</sup> cells). CD41<sup>+</sup> cells were further differentiated to generate CD11b<sup>+</sup> macrophages (MAC). From purified cells we determined global gene expression profiles by RNA-seq and mapped the full set of *cis*-regulatory elements at each developmental stage by global DNaseI hypersensitive site (DHS) mapping (DNaseI-seq). We used ChIP-seq to generate global maps of TF binding for key regulators across this entire developmental pathway as well as global patterns of H3K4me3, H3K9ac,

H3K27ac, and H3K27me3 histone modifications to investigate how TFs programmed the chromatin landscape. TFs were chosen according to the cell type in which they were expressed (Figure S1B), and all integrative analysis of ChIP and DHS data was focused on genomic regions found in at least two independent biological experiments (Table S1). Our datasets were complemented with published data for undifferentiated mouse ESCs (Chen et al., 2008; Whyte et al., 2013). The quality of this data resource is exemplified in a browser window depiction of sequence tags aligning to the *Tal1* locus (Figure 1A), which encodes a key regulator of early blood specification (Shivdasani et al., 1995; Wilson et al., 2009).

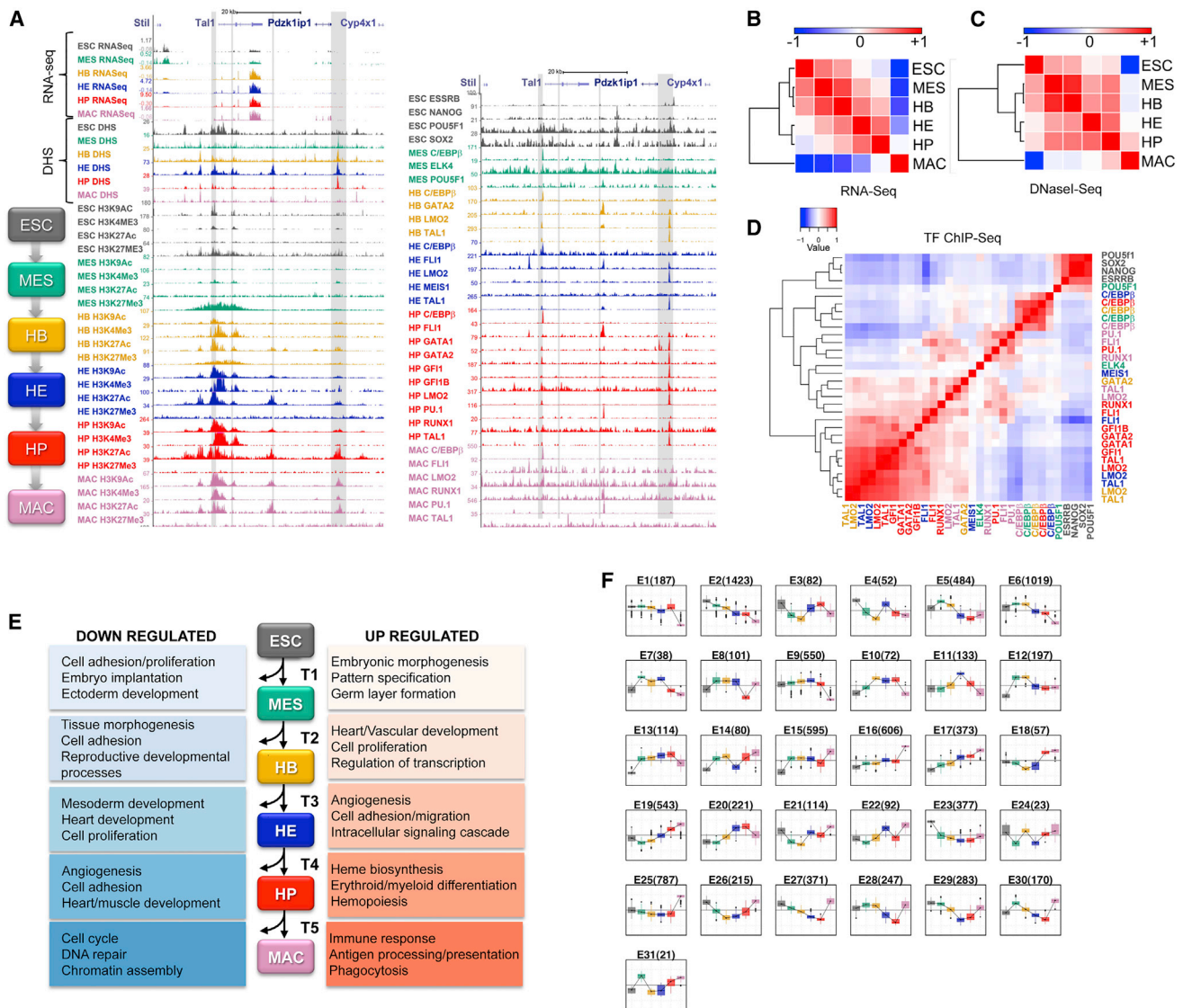
Initially, we used RNA-seq to investigate the dynamic changes of gene expression across the six differentiation stages and how these changes were reflected in the simultaneous changes in chromatin structure. To this end, we clustered RNA-seq (Figure 1B; Tables S2A and S2B) and DNaseI-seq data (Figure 1C). For both features, samples clustered in line with the known developmental progression, with an early cluster consisting of the ESC, and the more closely related MES and HB and a later cluster made up of HE and HP with the macrophage samples clustering separately. We then performed a similar analysis using TF binding data (Figure 1D). While cell-type-specific clustering of specific TF binding events were evident in ESCs and for certain TFs (e.g. FLI1) in HPs and MACs, others (such as C/EBP $\beta$ ) showed patterns predominantly driven by the identity of the factor rather than the tissue type (Figure 1D).

To facilitate inspection of individual genes and generate a resource for further data analysis, we developed a web interface to allow streamlined access for the wider scientific community: [http://www.haemopoiesis.leeds.ac.uk/data\\_analysis/](http://www.haemopoiesis.leeds.ac.uk/data_analysis/). The web portal provides access to both raw and processed data as well as user-driven analysis options. These include queries for specific genes and gene sets across our multi-omics datasets, as well as the visualization of all our data through a custom installation of the UCSC genome browser. In the following sections, we describe how our data can be explored to inform the functional validation of potential mechanisms.

### Identification of the Complete Set of Differentially Active *cis*-Regulatory Elements Driving Hematopoietic Specification

We next inspected the nature of genes changing expression at each cellular transition. 9,627 transcripts from 8,986 genes were dynamically expressed during the developmental time course (Figures S1C–S1E; Tables S2A and S2B). Expression changes between any two sequential developmental stages (transitions T1 to T5, Figure 1E) showed specific enrichment for functionality with the ensuing stage of development for upregulated genes (e.g. T4 shows enrichment for hematopoiesis), and alternative cell fates for downregulated genes (e.g. T4 angiogenesis, heart/muscle development; Table S2C).

To capture dynamic expression patterns across the entire developmental pathway and correlate such changes with alterations in chromatin structure and TF binding, we performed unsupervised/k-means clustering, which identified 31 major expression clusters E1 to E31 (Figure 1F and Table S3A) representing different gene ontology (GO) categories (Table S3B and Figure S4A). For example, E17–E19 represent clusters with



**Figure 1. Integrated Global Data over a Whole Developmental Pathway**

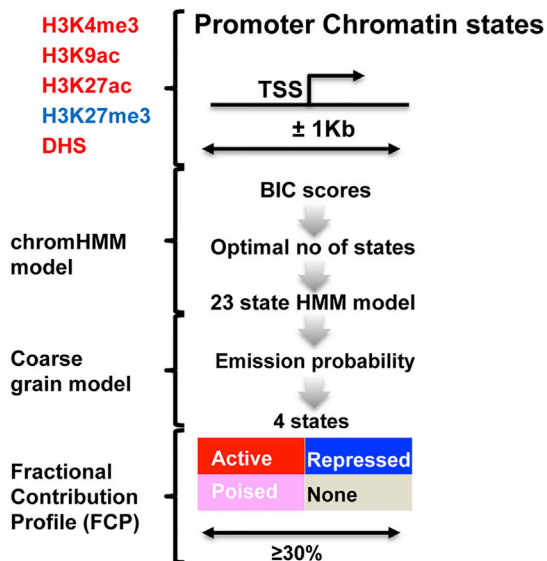
(A) UCSC browser screenshot depicting the *Ta11* locus aligning RNA-seq, DNase-seq, and ChIP-seq data from the six stages of development depicted in the left-hand flow chart. The stage-specific color scheme is used in all subsequent figures. Panels display ChIP-seq data for four histone modifications (left) and 16 different TFs (right) plus DHS data. The grayed-out regions indicate known regulatory regions: from left to right, promoters 1a and 1b, enhancers +19 and +40. (B–D) Hierarchical clustering of cell populations based on the normalized expression values of the genes (B), normalized correlation among the DHS sites (C), and correlation among the TF sites (D). The correlations were normalized between  $-1$  and  $+1$  to preserve the color scale. ESC, embryonic stem cell; HB, hemangioblast; HE, hemogenic endothelium; HP, hematopoietic progenitors, MES, mesoderm.

(E) Functional enrichment for genes that are differentially regulated during developmental transitions (T1–T5) in the progression of hematopoietic commitment. (F) The expression dynamics of the differentially expressed genes in the pathway given in (A) that are clustered into 31 patterns. The standardized expression values ( $z_{ij}$ ) of the differentially regulated genes in the developmental pathway (Figure S1E) were clustered into 31 expression patterns, and the plot shows the expression profiles of these patterns. The methodology is detailed in Supplemental Experimental Procedures.

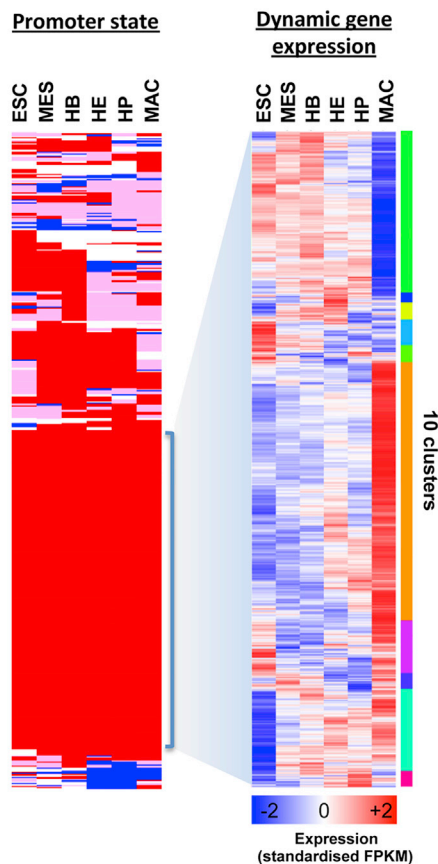
increased expression in macrophages, and all are enriched for functions relating to the immune response. Similarly, pattern E11 with upregulation in HE and downregulation in HP is enriched for functions relating to vasculogenesis and adhesion, whereas pattern E20 with upregulation toward HP is enriched for functions relating to hematopoiesis (Figure S4Ai–S4Aiv). Thus, our expression dataset defines distinct gene sets relevant for specific developmental transitions during early blood specification.

We next investigated the correlation between expression kinetics and dynamic changes of chromatin at the gene promoters (Figures 2A and S2A) by using ChromHMM, which was reported as an automated computational system for annotating chromatin states (Ernst and Kellis, 2012). We modified this methodology to integrate both histone modifications and DNaseI accessibility data. The latter indicates regions of chromatin bound by TFs (Cockerill, 2011) and allows for the distinction between inactive

A



B Clustering of chromatin states



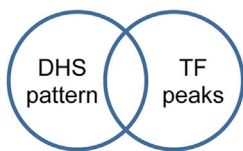
C

Dynamics of DHS occurrence

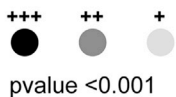
	ESC	MES	HB	HE	HP	MAC
1	1	0	0	0	0	0
1	1	1	0	0	0	0
1	1	1	1	0	0	0

Binary DHS pattern:  
presence/absence (1/0)

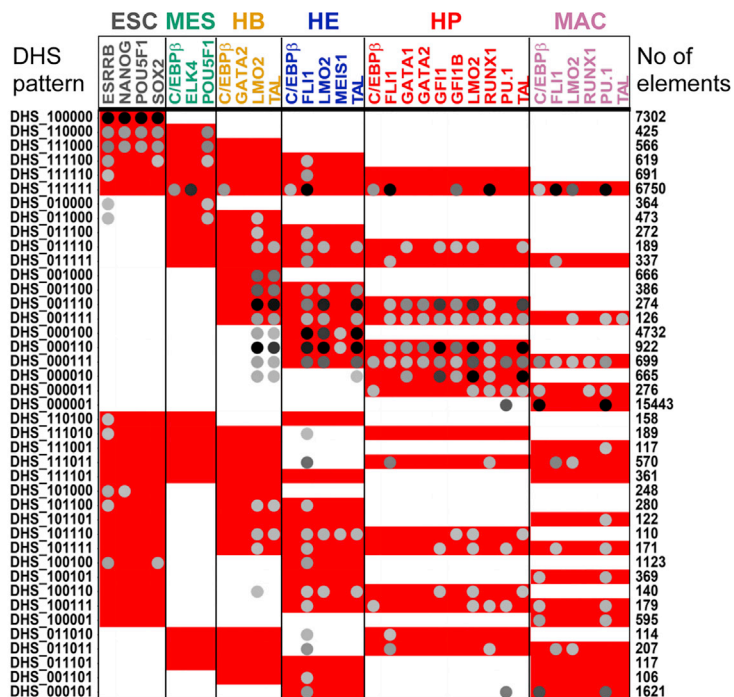
Overlap with TF peaks



Significance of overlap



D



(legend on next page)

chromatin regions (absence of DHS) and repressed/poised regions carrying H3K27me3. An initial number of 23 chromatin states (see Figure S2B) was further compressed, providing a simple four-state model of active (DHS marked with H3K4me3 and acetylated H3), repressed (marked with H3K27me3), poised (DHS marked with H3K27me3 but also acetylated H3 and/or marked with H3K4me3), or unmarked chromatin (Figure 2A). Examples (*Nanog*, *Runx1*) for dynamic alterations in promoter state are shown in Figure S2C, demonstrating that such changes occur gradually, both during the transition from the active to the inactive state and during gene activation. This behavior is also evident on a global scale with all differentially expressed genes (Figure S2D). It was proposed that poised promoters of key regulatory genes are held in this state until developmental cues shift the balance from poised to active or repressed states (as in the case of *Runx1*). The promoters of some genes highly expressed in macrophages indeed transit through a poised state, which in many cases is already evident in ESCs (Figure S2D). However, transitions from the unmarked or repressed state are more frequent (Figure S2D, last row at the bottom).

A direct correlation between promoter state and gene expression is not seen with all differentially expressed genes (Figure 2B). Promoters of around one-third of differentially regulated genes are persistently in the active state despite highly dynamic gene expression (Figure 2B, highlighted). When this gene set was investigated for GO term enrichment, we found that most of their functions were “housekeeping” roles pertinent to regulation of cell cycle, protein catabolism, transport, and localization (Figures 2B and S2D; Table S4).

To link gene expression with the chromatin state of distal *cis*-regulatory elements, we associated them with their nearest genes and correlated changes in their chromatin state with the 31 gene expression patterns across the differentiation pathway (Figures S3A and S3B depict the actual expression patterns as heatmaps). This comparison demonstrates a strong correlation between the dynamics of the chromatin state of distal elements and gene expression, indicating that most of these elements function as enhancers. We noted that the number of distal elements that displayed a poised or repressed chromatin state was small. These results add to the increasing evidence that cell-type-specific spatiotemporal expression patterns are largely driven by distal regulatory elements (Lara-Astiaso et al., 2014) and in addition demonstrate that such elements are in either the active or inactive chromatin state.

### Chromatin Dynamics and TF Binding Determines the Differential Activity of *cis*-Regulatory Elements

We next addressed the question of which TFs were responsible for the cell-stage-specific opening of chromatin. We therefore determined dynamic DHS patterns during the differentiation

time course and classified DHS patterns using a binary code with six digits (Figures 2C and 2D). We then performed pairwise comparisons between all our DHS patterns with each of the 32 TF ChIP-seq experiments (our own and publicly available data). Linking a set of regulatory genomic regions to annotated gene sets is sensitive to the varying sizes of the intergenic regions. We therefore used gene-set control analysis (GSCA) (Joshi et al., 2013), a tool designed to account for the differing sizes of such regions, to calculate pairwise correlation between TF ChIP-seq peak sets and expression gene sets, thus identifying all significant overlaps between TF binding events and DHS appearance. Figure 2D shows the most prevalent patterns of open chromatin over the six stages of development overlaid with the most significant TF binding events which in general, but not always, correlate with DHS presence. The most frequent DHS patterns are stage specific, over half of which involved DHS present only in macrophages (000001, 15,443 DHSs) and a quarter in ESC (100000, 7,302 DHSs). Notably, DHSs exclusively open in the HE (000100, 4,732 DHSs) are already primed by TF binding in HBs. The remaining patterns represent approximately 30% of all DHS whereby the majority of all patterns are continuous over at least two developmental stages. A common DHS pattern is 111111 (6,750 occurrences), the majority of which are CpG island promoters (Figure S4B) with a constitutively active chromatin state (Figure S4C). This class of DHS also contains the majority of binding events for C/EBP $\beta$  prior to the HP/MAC stages, suggesting a more widespread role of this transcription factor in development than previously thought.

Early binding of both LMO2 and TAL1 is highly significant in regulatory elements whose chromatin is first opened in HBs, HEs, or HPs, and include binding prior to the appearance of overt DHS sites, which is indicative of TF-mediated enhancer priming (see pattern DHS\_000110). PU.1 binding shows significant overlap with DHS patterns in HPs but is also found at sites that only become hypersensitive in MACs. This suggests that PU.1 can prime MAC-specific regulatory regions already in early multipotent progenitors, lending weight to the finding that it is capable of opening chromatin (Garber et al., 2012; Natoli et al., 2011; Barozzi et al., 2014; Heinz et al., 2010, 2015).

We next correlated the statistical significance of the dynamics of distal DHS patterns with dynamic gene expression patterns (Figure S3B). This again demonstrates that the dynamics of chromatin accessibility at distal sites correlates well with the dynamics of gene expression (Figure S3B).

### The Complex Interplay between Chromatin Dynamics, Gene Expression, and TF Binding Events

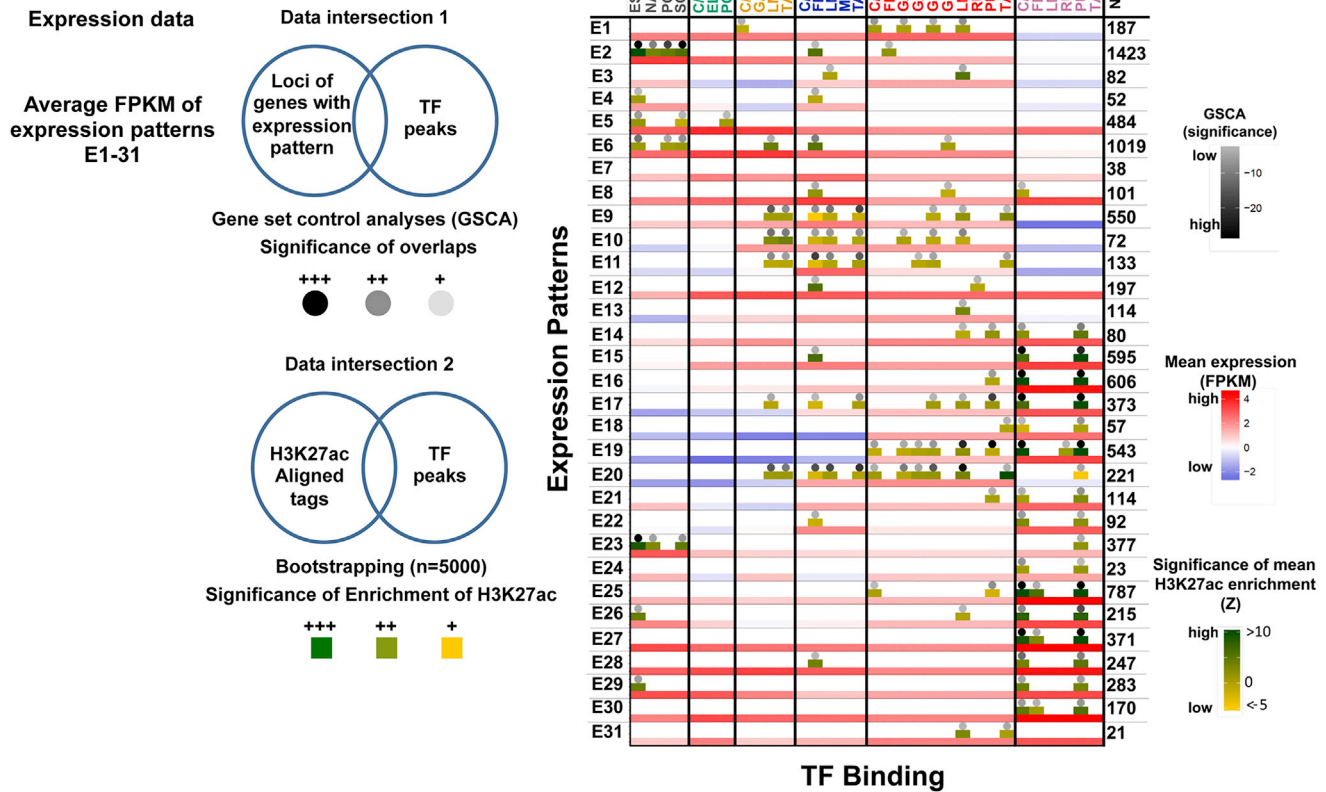
Our next analysis determined the combinatorial pattern of TF-DNA interactions driving target gene expression at key stages of blood development. We therefore interrogated the 31

#### Figure 2. Chromatin Programming during Progressive Lineage Commitment

(A) Schematic diagram of the method used to coarse grain the 23-state chromatin model to four potential chromatin states.

(B) Clustering of promoters (1Kb up or downstream of the transcription start site, TSS) based on their chromatin state patterns (left) and the clustering of the expression pattern of genes that are constitutively expressed (right).

(C and D) Integration of DHS pattern and TF binding. (C) Methodology of the integrative analysis of chromatin dynamics and TF binding events. (D) TF binding events and the p values denoting the significance of overlap are depicted as gray-scale density plots, shown as dots. Integration of DHS (rows) and TF binding (columns) patterns across the six stages with the population size of each DHS pattern given on the right-hand side. For significance calculations only DHS patterns with a population size >100 were considered.



**Figure 3. Integration of Chromatin Dynamics, TF Binding Events, and Gene Expression during Hematopoietic Specification**

(Left) Flow diagram of data integration. The average expression values ( $\log_{10}(\text{FPKM})$ ) of genes in expression patterns E1 to E31 were calculated for each developmental stage. The significance ( $p < 0.0001$ ) of the overlap between the genes in each expression pattern and a given TF ChIP-seq peak set was obtained using gene-set control analysis. Z scores were obtained from the mean enrichment of H3K27ac in TF binding sites at these loci using bootstrapping. (Right) Average expression levels for each pattern (rows) are shown as a red-blue heatmap (see key), with columns for each cell type labeled at the top. The columns are further divided into TF ChIP-seq experiments, and the significant overlap between TF binding events and gene sets belonging to each expression pattern are depicted by gray-scale density plots shown as dots (see key). Significant overlap of these binding events and H3K27ac sites are also shown as a density plot depicted by yellow-green boxes (see key).

expression clusters (Figure 1F) to ascertain (1) whether expression patterns correlated with enriched binding of any of our examined 32 TF datasets to these genes, and (2) how such binding events correlated with histone H3K27 acetylation at this position. For visual inspection, the TF binding and histone acetylation data were then overlaid onto a heatmap summarizing gene expression for patterns E1–E31 (Figure 3). This analysis shows the overall correlation between dynamic transcription factor binding, histone acetylation, and gene expression. The genes expressed in patterns E17–E20 are associated with increased gene expression during hematopoiesis, all showing early low-level induction prior to high-level expression (Table S3). This induction is associated with significant binding of hematopoietic regulators, but not MEIS1. Highly significant early binding of LMO2/TAL1 in HB and FLI1/LMO2/TAL1 in HE occurs in genes expressed in patterns E9–E11. All three patterns are associated with binding of the repressor GF11 in HP and with the repression of gene expression in macrophages. Patterns E9 and E11 involve upregulation of genes in the major HB–HE transition but then downregulation in HPs. Both sets of genes are enriched for func-

tions relating to vasculogenesis, heart development, and cell adhesion (Figure S4A). Our results therefore highlight GF11 as a candidate regulator involved in downregulating genes involved in non-hematopoietic cell fates following the HE to HP transition. This is consistent with data in the mouse that demonstrate a failure of EHT in the combined absence of GF11 and GF1B in addition to the continued expression of endothelial genes (Lancrin et al., 2012; Lie-A-Ling et al., 2014; Thambyrajah et al., 2016). In summary, our analysis provides a highly informative integrated view of the dynamic relationships between gene expression, chromatin state, and TF binding.

#### A Dynamic Core Gene Regulatory Network Driving Blood Development

To uncover the hierarchy of transcription factors driving blood specification forward, we generated gene regulatory network (GRN) representations connecting all 16 TFs analyzed by ChIP-seq, with separate representations for all six stages of development. To visualize different features, we illustrated multiple different data types within a single GRN representation at each

locus (Figure 4). Annotation for each of the six sequential developmental stages provided effective representation of the dynamics of cellular states, highlighted the chromatin features of the promoter of each gene locus, and indicated how interactions between a core set of key regulators drives developmental progression and terminal differentiation.

In ESCs all four pluripotency TFs participate in a highly connected core network circuit and already at this stage bind loci for hematopoietic TFs, including *Cebpb*, *Elk4*, *Gata2*, *Lmo2*, *Meis1*, *Runx1*, and *Tal1*, which display open or poised chromatin at their promoters, but also bind *Gfi1b*, *Gata1*, and *Spi1*, whose promoters are organized in closed-unmarked/repressed chromatin. As early as the HB stage, several hematopoietic regulator genes are upregulated, including *Tal1* and *Lmo2*, which exhibit autoregulation and co-regulate multiple genes. These include *Fli1* and *Meis1*, both of which are upregulated upon differentiation into HE. This stage is characterized by the involvement of LMO2, TAL1, and FLI1 (and in some cases MEIS1) in co-regulating genes for multiple hematopoietic TFs, revealing a densely connected GRN composed of potential feedback loops, which is likely to set the stage for the next step of hematopoietic commitment.

The HP stage shows highest expression for many of the key hematopoietic TFs, with binding events being complex and combinatorial. All ten TFs tested at this stage bind to *Gfi1*, *Gfi1b*, and *Runx1*, and nine out of ten bind to *Cebpb* and *Tal1* (the exceptions being RUNX1 and GFI1B, respectively). LMO2, TAL1, and to some extent FLI1 continue to co-bind and at this stage all target genes are shared with GFI1, consistent with the results shown in Figure 1D. FLI1 no longer binds to *Lmo2* or *Meis1*, both of which are strongly upregulated. While *Gata1* is upregulated by a combination of GFI1/LMO2/TAL1, *Gata2* is bound by C/EBP $\beta$ , TAL1, LMO2, FLI1, GATA1, GFI1, and GFI1B and is downregulated, uncovering a potential feedback mechanism regulating this TF within the network.

In macrophages, part of the HP-specific network is deconstructed with the further downregulation of early hematopoietic regulator genes such as *Gata2* and *Tal1* and a strong increase in the expression of PU.1 and C/EBP $\beta$ , which dominate global binding patterns. Within the GRN both TFs already share many target genes in HPs, including all genes encoding for experimental TFs (14 loci in total). Nine of these loci (*Cebpb*, *Elk4*, *Fli1*, *Gata2*, *Lmo2*, *Meis1*, *Runx1*, *Spi1*, *Tal1*) are also bound by FLI1 at this stage where *Lmo2* and *Tal1* are downregulated. Taken together, our datasets provide deep insights into the regulatory processes that control the dynamic rewiring of network connections during blood cell specification and differentiation. In the remaining part of this article we provide examples of how these data can be used to gain insights into the regulation of hematopoietic specification.

### Hierarchy Matters: TAL1/LMO2, but Not FLI1/GATA2, Can Reprogram Fibroblasts into Hematopoietic Cells

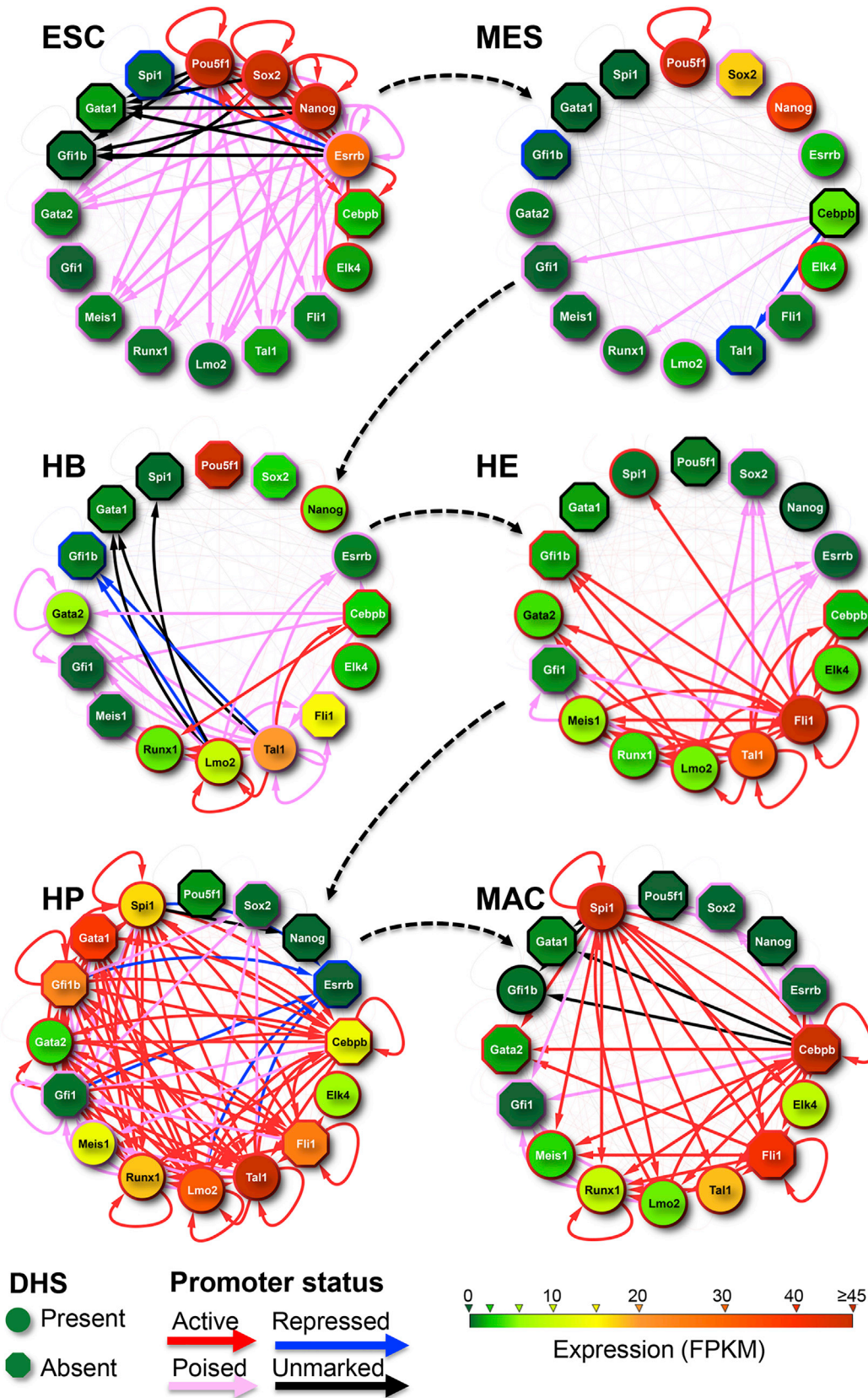
A number of recent publications reported a variety of TF combinations capable of generating blood cells via the reprogramming route (Batta et al., 2014; Elcheva et al., 2014; Pereira et al., 2013; Riddell et al., 2014; Sandler et al., 2014). We reasoned that the most likely factors capable of activating such a program would be those that (1) are expressed first during blood specification

and (2) bind to a large number of genes required for blood cell development. Four tested factors fulfill these criteria, GATA2, TAL1, LMO2, and FLI1, with all of their respective genes being activated at the hemangioblast stage (Figure S1B). Figure S5A shows an extended transcriptional network highlighting TFs that have been used for reprogramming experiments demonstrating that most binding events within the hematopoietic transcriptional network involve TAL1/LMO2 which interact with one another, autoregulate, and bind *Fli1* and *Gata2*. Moreover, these factors synergize in driving hematopoietic development in zebrafish (Patterson et al., 2007). We therefore tested the hypothesis that TAL1/LMO2 overexpression would be sufficient to activate the hematopoietic developmental program in mouse embryonic fibroblasts (MEFs). To this end we transduced wild-type MEFs or MEFs carrying a doxycycline-inducible allele of *Tal1* with different combinations of expression vectors for the four factors (Figure 5A) and ensured that each construct was efficiently overexpressed (Figure S5B). We then scored the number of hematopoietic colonies (Figure 5B) and measured the activation of a blood-cell-specific gene expression program using RNA-seq (Figure 5C). These experiments show that (1) reprogramming generates cells with a gene expression profile that is highly correlated with that of HPs, (2) TAL1 and LMO2 are sufficient for reprogramming, (3) both are also necessary even in the presence of GATA2 and FLI1, and (4) GATA2 and FLI1 alone cannot reprogram efficiently even in the presence of either TAL1 or LMO2. Figure 5D shows that at least 13 important hematopoietic regulator genes are bound by the LMO2/TAL1 complex during the HB-HE transition, far exceeding those by any of the other tested factors. In addition, in HP, TAL1 and LMO2 cooperate to upregulate a battery of genes encoding downstream factors whose expression is upregulated, such as *Runx1*, *Gata1*, and *Spi1* (*Pu.1*). These experiments demonstrate that the integrated analysis of time course TF binding, chromatin accessibility/modification, and expression data enables to highlight those factors that are on top of the hierarchy of tissue specification and are involved in the activation of the majority of genes governing lineage-specific gene expression programs. We believe that this principle will be applicable in multiple developmental settings.

### Identification of Factors Driving Key Stages of Blood Specification: A Role for TEAD/YAP

We next used our dataset to identify transcriptional regulators of blood cell specification. We reasoned that cell-stage-specific regulators would leave their mark in the epigenome by occupying their respective binding motifs within cell-stage-specific DHS. To capture all relevant regulatory regions at the genome-wide scale, we used our chromatin accessibility data to perform a pairwise comparison of distal DHS from one cell type with all others as outlined in Figure 6A. For each set of DHSs unique to a given cell population, we determined relative enrichment for sequence motifs and performed a clustering analysis against cognate motifs of TFs expressed in these cells. Our analyses recovered the known role of specific factors in the relevant cell types. In ESCs the pluripotency factor motifs form a distinct cluster, whereas the RUNX motif is predominantly enriched at the HP stage (Figure 6B, blue arrow), where this TF is critically required (Chen et al., 2009; Lancrin et al., 2009). These results were





(legend on next page)

confirmed by analyzing enriched motifs within the TF ChIP-seq peaks (Figure S6A). Besides confirming the presence of the motifs for the assayed factors, we also discovered a significant cell-stage-specific enrichment of co-localizing motifs. An example is the significant enrichment of GATA motifs near the TAL1/LMO2 complex over several developmental stages, confirming the important role of this factor in forming a complex regulating hematopoietic genes (Wadman et al., 1997; Wilson et al., 2010). We also observed a strong enrichment of TCF7L1/2 motifs in TAL1/LMO2 peaks at the HB stage, which are also predominant in the cell-specific DHS at this stage. TCF/LEF are mediators of Wnt signaling and have been shown to regulate hematopoietic specification (Sturgeon et al., 2014), raising the possibility that at this developmental stage elements binding the TAL1/LMO2 complex are WNT-signaling responsive.

The same analysis also uncovered enriched motifs for factors not yet linked with mammalian hematopoietic development, such as a significant enrichment for TEAD binding motifs early in hematopoietic development, specifically at the HB stage (Figure 6B, red arrow). TEAD motifs significantly co-localized with peaks for LMO2 and TAL1 in HBs and with TAL1/LMO2/FLI1 in the HE, but not in HPs (Figure 6C). TEAD transcriptional activity is controlled by the Hippo signaling network, which has emerged as a highly conserved pathway controlling cell proliferation, cell shape, organ size, and cell fate decisions in several differentiation pathways, including hematopoiesis in *Drosophila* (Dong et al., 2007; Ferguson and Martinez-Agosto, 2014; Milton et al., 2014). In mammals, when the Hippo pathway is activated, the MST and LATS kinases phosphorylate the transcriptional co-regulator YAP, which is then sequestered in the cytoplasm and consequently cannot form a complex with its nuclear DNA binding partner TEAD. When Hippo signaling is inactive, YAP interacts with TEAD factors in the nucleus to positively or negatively regulate Hippo signaling-responsive target genes (Yu and Guan, 2013).

To test whether TEAD factors are involved in regulating mammalian hematopoietic specification, we first looked for the presence of TEAD and YAP and any change in nuclear localization of YAP in mouse embryos. Figure 7A shows cross sections of developing blood islands from the yolk sac of embryonic day 7.5 (E7.5) mouse embryos stained with antibodies against TIE2 to identify endothelial cells, TEAD (upper panel) and YAP (lower panel). The images demonstrate a nuclear localization of YAP prior to the EHT but a cytoplasmic localization in hematopoietic cells. Staining of hematopoietic clusters emerging from the dorsal aorta from E10.5 mouse embryos (Figures S7A and S7B) also shows the absence of nuclear YAP in committed hematopoietic cells. The same is true for in vitro differentiated cells, where YAP is predominantly localized in the nucleus at the HB stage (Figures S6B and S6C) and then is localized in the cytoplasm in CD41<sup>+</sup> HP cells (Figure S6D), indicating a precise temporal regulation of TEAD activity during hematopoietic specification.

To test whether the interaction of YAP with TEAD factors is important for hematopoietic differentiation, we performed both in vitro and ex vivo experiments using verteporfin, which specifically inhibits TEAD-YAP complex formation, thus mimicking the Hippo pathway activation (Liu-Chittenden et al., 2012). Culturing ESC-derived embryoid bodies with the inhibitor blocked the emergence of CD41<sup>+</sup> HPs, but only when administered at early time points prior to day 5 of embryoid body (EB) culture when the emergence of blood cells occurs (Figures 7B and 7C). During normal HB development from mesodermal cells, the expression of *Fgf5* and *Bry* are downregulated concomitant with the upregulation of *Flk1* expression (Fehling et al., 2003). This process was abolished after treatment of isolated Bry<sup>+</sup> ME cells with verteporfin (Figure S6E). We also explanted FLK1<sup>+</sup>/CD41<sup>-</sup> cells from E7.5 embryos and differentiated them into CD41<sup>+</sup> hematopoietic precursor cells on stromal cultures in the presence or absence of verteporfin (Figure S6F), and compared their response to cultured committed CD45<sup>+</sup> hematopoietic cells from E10.5 embryos (Figure S6G). Again, the addition of the inhibitor inhibited blood cell emergence and survival prior to, but not after hematopoietic specification. Moreover, genes encoding YAP and all TEAD factors were downregulated from the HP stage onward with expression essentially absent in macrophages (Table S2A). Together, these data suggest that (1) TEAD/YAP interaction is required at early stages of hematopoietic commitment and (2) YAP localizes outside of the nucleus after the EHT, suggesting that Hippo signaling is switched on in these cells.

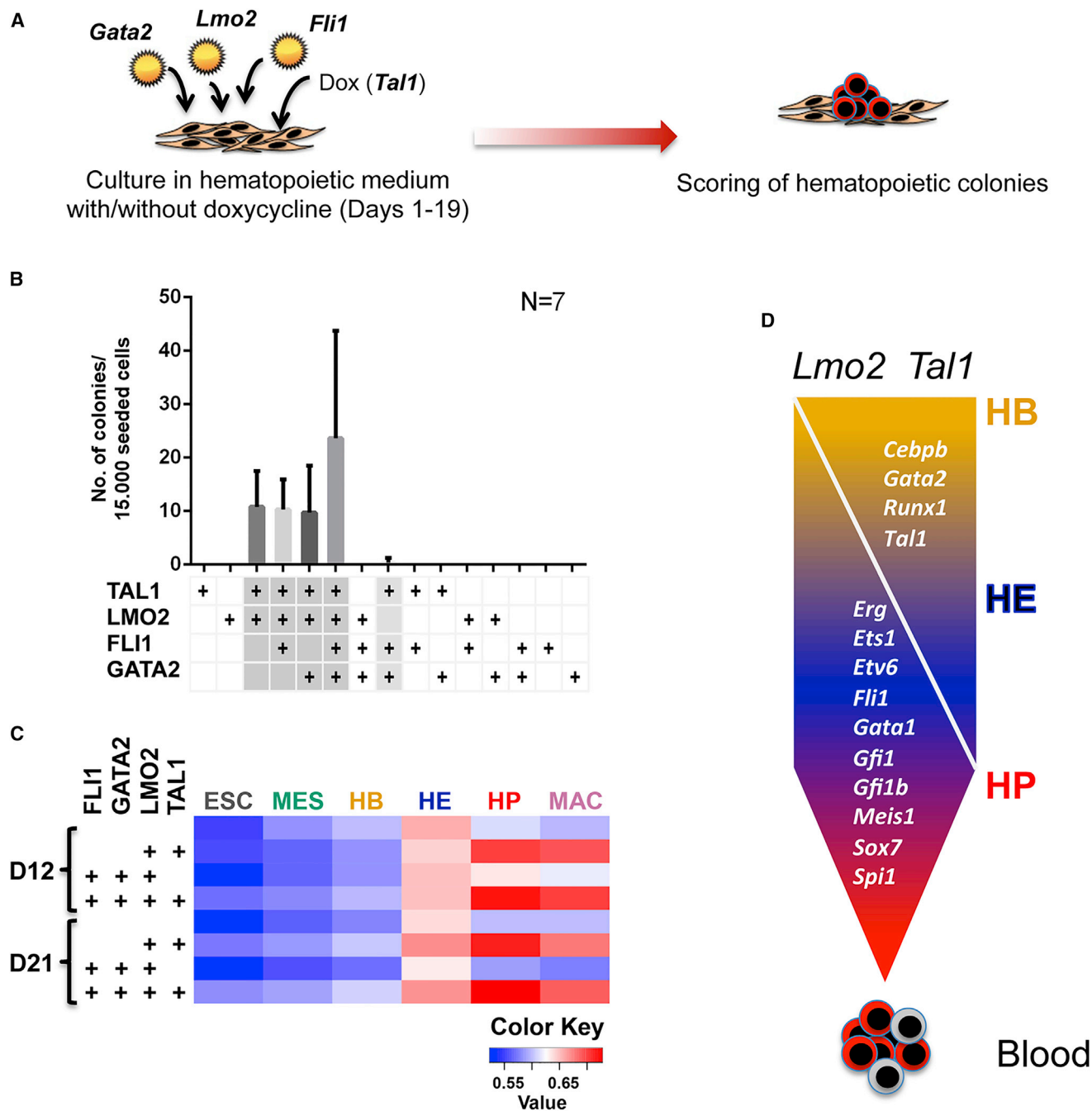
#### Identification of TEAD Target Genes

Having established an important role of TEAD and YAP interaction in hematopoietic specification, we performed ChIP-seq for TEAD4 from in vitro differentiated Flk1<sup>+</sup> heman-gioblast cells to map its target genes. We mapped 5,234 TEAD4 binding regions (with manual validations shown in Figure S7C), including the *Tal1* locus (Figure 7D). Two-thirds of the binding sites occur in distal regions. Approximately 30% of all TEAD peaks (1,563) are in DHS that are active throughout all stages of differentiation (DHS pattern 111111, Figure 3, left panel, and Figures S4B and S4C). Most of these sites (1,342) are promoters, indicating that TEAD may fine-tune the expression of CG island promoters that predominate in such sites. TEAD binding sites in distal peaks are found predominantly in stage-specific DHS. A significant number of TEAD4 binding sites overlap with LMO2 and TAL1 binding sites in HB (Figure S7D), which is in concordance with the enrichment for TEAD motifs in early LMO2 and TAL1 binding events (Figures 6B and 6C).

Analysis of pathways enriched in TEAD4 bound genes (Table S8) revealed focal adhesion and Rap1 signaling as well as Wnt and transforming growth factor  $\beta$  signaling to be the top-scoring

#### Figure 4. Dynamic Gene Regulatory Network Driving Hematopoietic Specification

For each developmental stage the 16 TFs used in ChIP-seq experiments are shown as nodes in a GRN. The color of each node corresponds to the level of gene expression (see key). The chromatin accessibility at each promoter is shown as open/circular (DHS presence) or closed/octagonal (DHS absence), and the border color of the node corresponds to the promoter state according to the coarse-grain four-state model mentioned in Figure 2 (see key). Arrows indicate binding events of a TF (source) at loci encoding all TFs (target). The arrow color relates to the promoter state of the target TF encoding gene. No emanating arrow from a node indicates absence of ChIP data. For information about which ChIP experiments were conducted in which cell type, see Figure S1B. Note that we did not include C/EBP $\beta$  binding in the ES cell GRN, since the publicly available datasets did not contain the respective data.



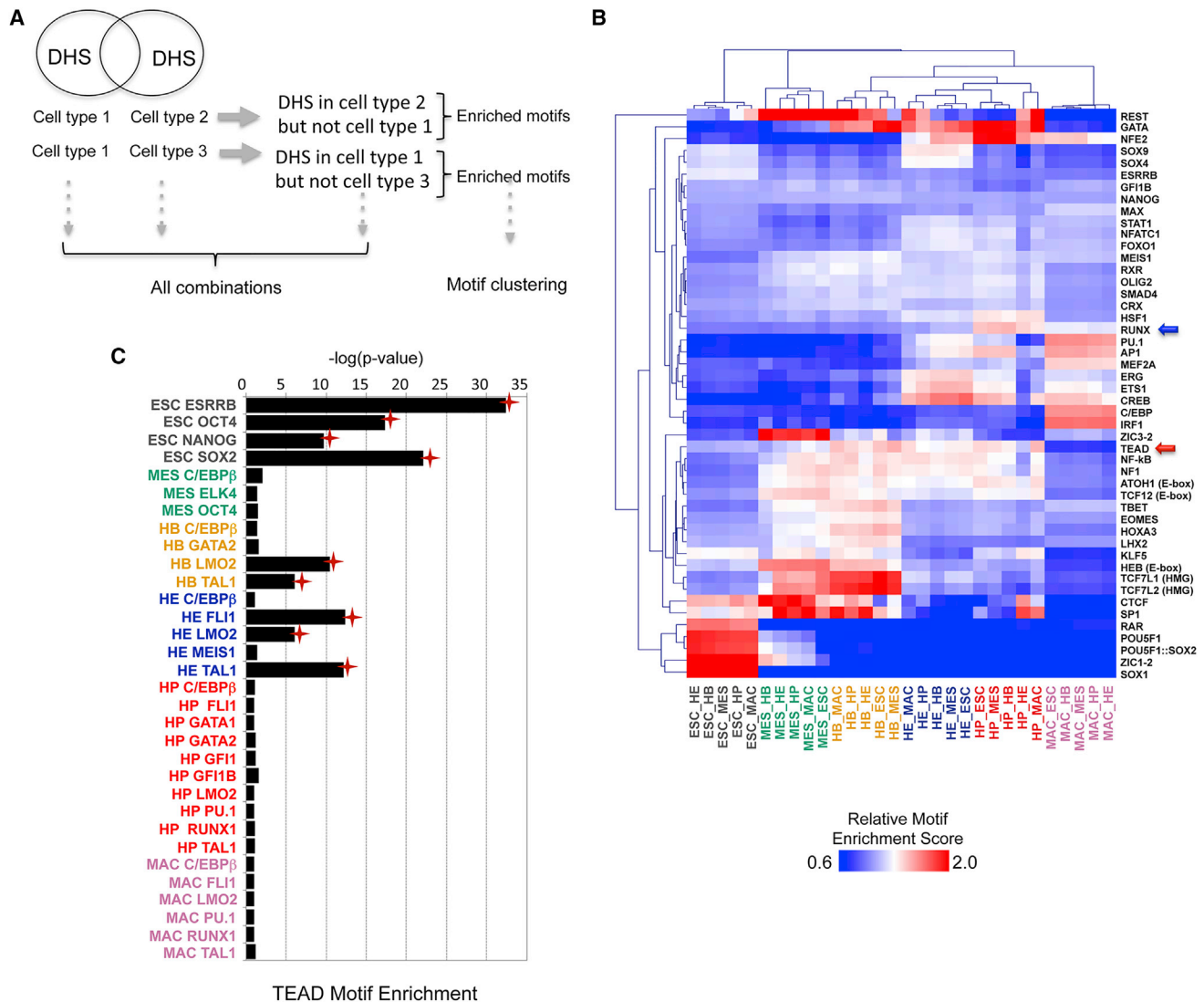
**Figure 5. Gene Regulatory Network Analysis Is Informative for Reprogramming Success**

(A) Schematic of reprogramming experiments.

(B) Bar chart showing the number of hematopoietic colonies generated from fibroblast cells after overexpression of different combinations of TAL1, LMO2, FLI1 and GATA2 as indicated in the table below. Grayed areas in this table highlight successful production of hematopoietic colonies. The number of hematopoietic colonies generated from fibroblast cells after overexpression of *Tal1*, *Lmo2*, *Fli1*, *Gata2*, and the indicated combinations of these four TF encoding genes. Data presented are mean  $\pm$  SEM of individual experiments ( $n = 7$ ).

(C) Correlation coefficient analyses of gene expression profiles generated by RNA-seq from hematopoietic cells generated by reprogramming from MEFs at day 12 and day 21 (D12 and D21) of the experiment with gene expression patterns generated from in vitro differentiated cells. The heatmap shows the correlation between expression data from each stage (columns) and each experiment (rows). All red/pink rectangles show correlation values that are significantly different from all blue rectangles (Z transform test,  $p < 0.01$ ).

(D) Model depicting the hierarchy of transcriptional regulation during hematopoietic differentiation. Target TF genes are listed in the arrow, with those that are upregulated after binding events listed on the left and those that are downregulated listed on the right.



**Figure 6. Chromatin and Transcription Factor Binding Dynamics and Identification of Regulators**

(A) Schematic representation of the methodology for pairwise motif clustering at distal DHS (for used motifs see Table S6).

(B) Relative motif enrichment (RE) scores for the motifs in DHS unique to a given cell type compared with each of the other five cell types were clustered. RUNX and TEAD motifs are indicated by blue and red arrows, respectively. The significance of the RE scores were computed using the bootstrapping method (Table S7).

(C) Significance of co-localization of TEAD motifs in TF ChIP peaks. Red stars indicate significant co-localization.

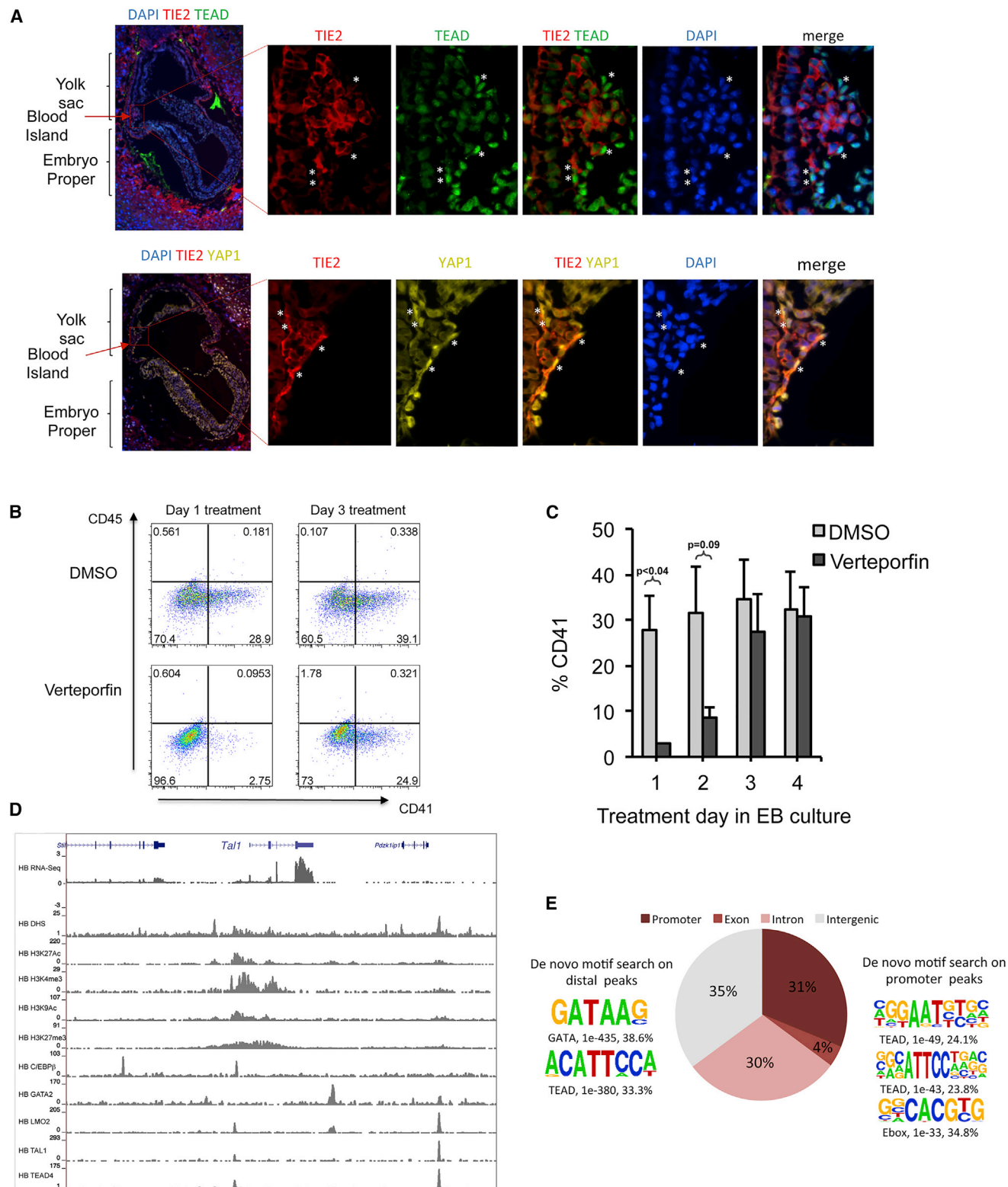
pathways, all of which are known to influence hematopoietic and endothelial specification. In addition, a number of genes important for hematopoiesis are bound by TEAD, including the gene for the hematopoietic master regulator RUNX1 as well as *Kit*, which encodes a growth factor receptor crucial for the growth of hematopoietic precursor cells. In summary, our data indicate that TEAD factors target genes that regulate hematopoietic specification during the critical period leading to the EHT.

## DISCUSSION

### A Gene Regulatory Network Model for Hematopoietic Specification in the Embryo

The comprehensive experimental and computational studies of embryonic blood cell development reported here elucidate in

fine detail how the interplay between cell-stage-specific TFs and the chromatin landscape drives differential gene expression during ontogeny. As indicated by the presence of a DHS, the majority of active cell-type-specific distal *cis*-regulatory elements correlate with significant binding of the measured TFs. Our data show a high complexity of developmental stage-specific TF assembly, with some factors binding first to be joined by others later in development, and with chromatin modifications following suit. Concurrent with the long-standing concept of developmental priming of distal *cis*-regulatory elements, a number of genes are bound by TFs in progenitor stages prior to high-level expression later in development. Recent experiments have shown that the early binding of TFs to distal elements of lineage-specific genes is required for the correct timing of gene activation and the repression of alternative fates (Lichtinger



**Figure 7. A Role for TEAD Factors in Early Hematopoietic Specification**

(A) TEAD and YAP localize to the nucleus of a subset of TIE2<sup>+</sup> endothelium within yolk sac blood island of E7.5 embryos. E7.5 embryo sections were stained as indicated. Asterisks mark TIE2<sup>+</sup> cells on the outer edge of the developing blood island, which show nuclear localization for both TEAD and YAP. Cells within the blood island are maturing primitive erythrocytes and do not show nuclear localization for either TEAD or YAP.

(legend continued on next page)

et al., 2012; Org et al., 2015). Early transcription factor binding is therefore required to express genes at the correct developmental stage and at the appropriate level.

### Recapitulation of Early Developmental Processes during Reprogramming

The activation of a hematopoietic gene expression program in unrelated cells by external factors requires the alteration of the chromatin and transcriptional landscape and the activation of hematopoietic genes. Our dynamic GRN model provided essential clues about which factors (TAL1 and LMO2) were the most likely to succeed in this task and which ones would not. However, the actual molecular mechanism by which this occurs requires further investigation. E-Box motifs are widespread in the genome and are used by a variety of factors; we therefore suggest that the interaction with the bridging molecule LMO2 (Wadman et al., 1997) is crucial in bringing in additional factors. LMO2 forms part of a complex consisting of various E-box binding proteins as well as GATA, ETS, or RUNX1, and recruits LDB1 (Mylyona et al., 2013). The latter mediates interactions between enhancer and promoter elements (Deng et al., 2012). The ETS factor ETV2 (ER71) is expressed in hemangioblasts and has been shown to be absolutely required for hematopoietic specification (Liu et al., 2015). This factor, together with GATA2 and TAL1, is capable of inducing enhanced hemangioblast formation during ES cell differentiation (Liu et al., 2013). However, TAL1 expression could rescue the hematopoietic defect in ETV2<sup>-/-</sup> cells, but FLI1 and GATA2 could not (Wareing et al., 2012), and when included in preliminary reprogramming experiments, ETV2 was incapable of generating hematopoietic colonies, on its own or in combination with either LMO2 or TAL1 (data not shown). These findings highlight the central role of the TAL1/LMO2 complex in driving a blood-cell-specific gene expression program and indicate that factors too far upstream in the hematopoietic specification hierarchy cannot substitute for hematopoietic factors.

### TEAD Factors Regulate Hematopoietic Specification in Mice

Our study provided an example of how our data resource can be used to advance our understanding of hematopoietic specification, in this case by using the differential analysis of enriched motifs in DHS. This type of analysis confirmed the role of known factors (such as RUNX1) but also identified a number of potential cell-stage-specific regulators of hematopoietic specification, leading to the identification of motifs for the Hippo-regulated TEAD TFs enriched specifically in MES, HB, and HE. Immunostaining analyses confirmed the presence of YAP in the nucleus of cells at stages prior to the EHT in both ESCs and embryo-derived cells, but not thereafter. Our functional studies show that the interaction of TEAD factors and the YAP co-factor in the nucleus at this early stage is strictly

required for the formation of hematopoietic precursor cells from both sources.

How could TEAD factors impact on hematopoietic specification in the embryo? The analysis of their target genes in this study suggests that the TEAD/YAP complex may integrate the response to different types of signaling via interaction with genes encoding numerous signaling molecules whose balance controls hematopoietic specification such as bone morphogenetic proteins, NODAL and WNT, and fibroblast growth factor receptors (Pouget et al., 2014; Simoes et al., 2011). Moreover, TEAD/YAP interacts with genes of the Rap1 pathway, which controls cell adhesion dynamics and is essential for embryonic, but not adult hematopoiesis (Satyanarayana et al., 2010). Several recent studies have shown that the TEAD/YAP complex cooperates with tissue-specific TFs to either activate or repress gene expression. In human ESCs, TEAD/YAP interact with OCT4 to maintain the expression of pluripotency genes and represses mesendodermal genes, whereby repression can be overcome by the activation of BMP and Wnt signaling (Beyer et al., 2013; Estaras et al., 2015). The footprint of this interaction in the epigenome can be seen in our data as a co-localization of TEAD motifs with binding motifs for pluripotency factors in ESCs (Figures 6B and S6A). Our data show that in developing hematopoietic cells TEAD peaks co-localize with TAL1/LMO2 complexes, and SCL/TAL1 motif co-association continues to be significant up to the HE stage, but, again, not thereafter. This finding suggests that the factor complexes on such *cis*-regulatory elements may respond to Hippo signaling before, but not after the EHT. We have previously shown that genes associated with TAL1 and FLI1 binding in the HE are highly enriched for genes regulating cell shape and focal adhesion. During the EHT RUNX1 is strongly upregulated and relocates TAL1 and FLI1 to new binding sites after the EHT (Lichtinger et al., 2012), thus explaining the absence of co-localizing TEAD motifs in HPs. The expression of all TEADs as well as YAP is downregulated during terminal differentiation of HPs (Table S1A), thus uncoupling *cis*-element activity from TEAD-mediated signaling processes. Our analysis suggests the presence of a dynamic GRN of tissue-specific and lineage-specifying TFs that is intricately connected with signaling-responsive TFs such as TEAD or TCF7L1/2. Our data resource will enable numerous further functional and computational studies to examine the role of these different factors and thus gain insights into the molecular control mechanisms that underpin crucial steps in blood specification.

### EXPERIMENTAL PROCEDURES

A detailed description of all experiments can be found in [Supplemental Experimental Procedures](#). All animal work was performed under regulation in accordance with the United Kingdom Animal Scientific Procedures Act (ASPA) 1986. Animal experiments were approved by the Animal Welfare and Ethics Review Body (AWERB) of the Cancer Research UK Manchester Institute.

(B and C) TEAD activity is required during the early phase of hematopoietic commitment. The TEAD-YAP inhibitor verteporfin (9.6  $\mu$ M) or DMSO vehicle was added on day 1, 2, 3, or 4 of EB culture. Day-1 EB corresponds to mesoderm (MES) commitment, day 2-3 to hemangioblast (HB), commitment and day 4-5 to HE and HP specification. The frequency of CD41<sup>+</sup> hematopoietic cells was determined on day 7. (B) Representative FACS plots. (C) Quantification of the percentage of CD41<sup>+</sup> cells from  $n \geq 3$  independent experiments. Data presented are mean  $\pm$  SEM, paired t test.

(D) Genome browser screenshot showing TEAD4 binding to the *Tal1* locus in HB together with other chromatin and binding features.

(E) Genomic distribution of TEAD4 peaks together with TF binding motifs enriched in distal (left) and proximal (right) peaks.

### Isolation of Cell Populations

A mouse ES cell line carrying a brachyury (Bry) GFP<sup>+</sup> reporter gene was cultured as described by Sroczynska et al. (2009). GFP and cell surface marker staining were used to identify each cell population by cell sorting. MES cells were (Bry<sup>+</sup>Flk1<sup>-</sup>), HB cells were (Bry<sup>+</sup>/Flk1<sup>+</sup>), HE cells were Tie2<sup>+</sup>/cKit<sup>+</sup>/CD41<sup>-</sup>, and HPs were CD41<sup>+</sup>. Macrophages were isolated by differentiation of CD41<sup>+</sup> cells to CD11b-expressing cells.

### Chromatin Immunoprecipitation followed by Next-Generation Sequencing

For each stage of differentiation cells were sorted, crosslinked, and stored either as frozen cells (histone modification ChIP) or nuclei (TF ChIP) for subsequent ChIP assays, performed as described previously (Lichtinger et al., 2012; Wilson et al., 2009). A full list of antibody sources is given in Supplemental Experimental Procedures.

DNaseI and ChIP samples were amplified and sequenced according to the manufacturer's instructions. Some samples (MES, HB and HP H3K4me3, HP H3K27me3, and HP H3K27ac) were processed using an ABI SOLiD 4 sequencer, and subsequently either the Illumina 2G Genome Analyzer or the Hi-Seq 2000 were used.

### RNA Sequencing

RNA preparation and sequencing were performed as described previously (Lie-A-Ling et al., 2014).

### DNaseI Sequencing

One to three million freshly sorted cells were digested with DNaseI enzyme as described in detail previously (Ptasinska et al., 2014) and size selected for 50- to 300-bp fragments.

### Reprogramming Experiments

E14.5 murine wild-type fibroblasts (MEFs) or from iTal1-2A-GFP transgenic mice carrying rtTA and TRE-TAL1 cassettes allowing inducible expression of TAL1 upon addition of doxycycline were prepared as described previously (Sroczynska et al., 2009). Reprogramming experiments were carried out with either wild-type (n = 5) or iTal1-2A-GFP (n = 2) MEFs as described by Batta et al. (2014).

### Inhibition of TEAD/YAP Interaction In Vitro and Ex Vivo

Verteporfin was added to day-1, -2, -3, or -4 EB cultures at a final concentration of 9.6 μM, and EB-derived cells were stained with CD41-PE and CD45-brilliant violet 421 (Becton Dickinson) and analyzed by flow cytometry. BRY<sup>+</sup>/FLK1<sup>-</sup> mesoderm-enriched population was cultured for 48 hr with or without verteporfin. Expression of *Fgf5*, *Bry*, and *Flk1* was measured by RT-PCR. FLK1<sup>+</sup>/CD41<sup>-</sup> HE-enriched cells were sorted from E7.5 embryos by fluorescence-activated cell sorting (FACS) and cultured for 4 days on irradiated OP9 in HE media, and hematopoietic differentiation of HE cells was measured by staining with a CD41-PE antibody. CD45<sup>+</sup> committed hematopoietic cells were sorted from E10.5 embryos and cultured on OP9 with verteporfin or DMSO. Growth and survival of hematopoietic cells was assessed by FACS after staining with a CD45-fluorescein isothiocyanate antibody.

### Immunostaining of Embryos

E7.5 embryos were fixed in 4% paraformaldehyde, mounted, and stained with rabbit Pan-TEAD (D3F7L, New England Biolabs) (1:100), rabbit YAP (D8HIX XP, New England Biolabs) (1:100), or purified anti-mouse Tie2 (Tek-CD202B, 14-5987-85, eBioscience) (1:100). Secondary antibodies were Alexa Fluor 488 goat anti-rat immunoglobulin G (IgG) (A11006, Life Technologies) and Alexa Fluor 647 F(ab')<sub>2</sub> fragment of goat anti-rabbit IgG (H + L) (A21246, Life Technologies).

### Data Analysis

A detailed description of all bioinformatics methods can be found in Supplemental Experimental Procedures.

### ACCESSION NUMBERS

All data and associated analyses are provided in our web server <http://www.bioinformatics.leeds.ac.uk/labpages/hematopoiesis/>. The tag densities for all our experiments have been integrated into the genome browser in <http://www.haemopoiesis.leeds.ac.uk>. Data can also be found at: <http://codex.stemcells.cam.ac.uk/>. The raw fastq, aligned bam, peak, and tag density files are available from the NCBI Gene Expression Omnibus portal (GEO: GSE69101).

haemopoiesis.leeds.ac.uk. Data can also be found at: <http://codex.stemcells.cam.ac.uk/>. The raw fastq, aligned bam, peak, and tag density files are available from the NCBI Gene Expression Omnibus portal (GEO: GSE69101).

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and eight tables and can be found with this article online at <http://dx.doi.org/10.1016/j.devcel.2016.01.024>.

### AUTHOR CONTRIBUTIONS

D.K.G., N.O., M.L., M.L.A.L., A.J.L., M.F., K.B., M.C., K.W., J.G., and M.H. performed experiments; M.S.V., H.R., S.A.A., and P.C. analyzed data. The senior authors (D.R.W., G.L., V.K., B.G., and C.B.) jointly obtained funding, conceived, and directed the study, and, together with D.K.G., wrote the paper. Within the consortium, D.R.W. directed data analysis/integration and creation of the website; G.L. and V.K. oversaw cell differentiation and purification, RNA-seq data generation, reprogramming and the in vivo and inhibitor studies; B.G. oversaw the transcription factor ChIP data generation and further data analysis; and C.B. was the project coordinator and oversaw chromatin data generation as well as further data analysis and integration.

### ACKNOWLEDGMENTS

This work was funded by a Longer Larger (LoLa) consortium grant from the Biotechnology and Biological Sciences Research Council, UK, to the senior authors and the corresponding author, computing infrastructure grants from the Wellcome Trust and National Institute for Health Research to B.G., grants from Cancer Research UK to G.L. and V.K., and funding from the Bloodwise charity to C.B.

Received: May 28, 2015

Revised: December 4, 2015

Accepted: January 26, 2016

Published: February 25, 2016

### REFERENCES

- Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., and Natoli, G. (2014). Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol. Cell* 54, 844–857.
- Batta, K., Florkowska, M., Kouskoff, V., and Lacaud, G. (2014). Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *Cell Rep.* 9, 1871–1884.
- Beyer, T.A., Weiss, A., Khomchuk, Y., Huang, K., Ogunjimi, A.A., Varelas, X., and Wrana, J.L. (2013). Switch enhancers interpret TGF-beta and Hippo signaling to control cell fate in human embryonic stem cells. *Cell Rep.* 5, 1611–1624.
- Bonifer, C., Hoogenkamp, M., Krysinska, H., and Tagoh, H. (2008). How transcription factors program chromatin—lessons from studies of the regulation of myeloid-specific genes. *Semin. Immunol.* 20, 257–263.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117.
- Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* 457, 887–891.
- Cockerill, P.N. (2011). Structure and function of active chromatin and DNase I hypersensitive sites. *FEBS J.* 278, 2182–2210.
- Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149, 1233–1244.

- Dong, J., Feldmann, G., Huang, J., Wu, S., Zhang, N., Comerford, S.A., Gayyed, M.F., Anders, R.A., Maitra, A., and Pan, D. (2007). Elucidation of a universal size-control mechanism in *Drosophila* and mammals. *Cell* **130**, 1120–1133.
- Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.-H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., et al. (2014). Direct induction of haematopoietic programs in human pluripotent stem cells by transcriptional regulators. *Nat. Commun.* **5**, 4372.
- Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216.
- Estaras, C., Benner, C., and Jones, K.A. (2015). SMADs and YAP compete to control elongation of beta-catenin:LEF-1-recruited RNAPII during hESC differentiation. *Mol. Cell* **58**, 780–793.
- Fehling, H.J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., and Kouskoff, V. (2003). Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* **130**, 4217–4227.
- Ferguson, G.B., and Martinez-Agosto, J.A. (2014). Yorkie and Scalloped signaling regulates Notch-dependent lineage specification during *Drosophila* hematopoiesis. *Curr. Biol.* **24**, 2665–2672.
- Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., et al. (2012). A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* **47**, 810–822.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589.
- Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154.
- Hoogenkamp, M., Lichtinger, M., Krysinska, H., Lancrin, C., Clarke, D., Williamson, A., Mazzarella, L., Ingram, R., Jorgensen, H., Fisher, A., et al. (2009). Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. *Blood* **114**, 299–309.
- Joshi, A., Hannah, R., Diamanti, E., and Gottgens, B. (2013). Gene set control analysis predicts hematopoietic control mechanisms from genome-wide transcription factor binding data. *Exp. Hematol.* **41**, 354–366.e314.
- Krysinska, H., Hoogenkamp, M., Ingram, R., Wilson, N., Tagoh, H., Laslo, P., Singh, H., and Bonifer, C. (2007). A two-step, PU.1-dependent mechanism for developmentally regulated chromatin remodeling and transcription of the *c-fms* gene. *Mol. Cell Biol.* **27**, 878–887.
- Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* **457**, 892–895.
- Lancrin, C., Sroczynska, P., Serrano, A.G., Gandillet, A., Ferreras, C., Kouskoff, V., and Lacaud, G. (2010). Blood cell generation from the hemangioblast. *J. Mol. Med. (Berl)* **88**, 167–172.
- Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N.K., Moroy, T., Bonifer, C., et al. (2012). GF11 and GF11B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood* **120**, 314–322.
- Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretzky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943–949.
- Leaddin, M., Perrod, C., Hoogenkamp, M., Ghani, S., Assi, S., Heinz, S., Wilson, N.K., Follows, G., Schonheit, J., Vockentanz, L., et al. (2011). Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. *Blood* **117**, 2827–2838.
- Lichtinger, M., Ingram, R., Hannah, R., Muller, D., Clarke, D., Assi, S.A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M.S., Wu, M., et al. (2012). RUNX1 re-shapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J.* **31**, 4318–4333.
- Lie-A-Ling, M., Marinopoulou, E., Li, Y., Patel, R., Stefanska, M., Bonifer, C., Miller, C., Kouskoff, V., and Lacaud, G. (2014). RUNX1 positively regulates a cell adhesion and migration program in murine hemogenic endothelium prior to blood emergence. *Blood* **124**, e11–20.
- Liu, F., Bhang, S.H., Arentson, E., Sawada, A., Kim, C.K., Kang, I., Yu, J., Sakurai, N., Kim, S.H., Yoo, J.J.W., et al. (2013). Enhanced hemangioblast generation and improved vascular repair and regeneration from embryonic stem cells by defined transcription factors. *Stem Cell Rep.* **1**, 166–182.
- Liu, F., Li, D., Yu, Y.Y., Kang, I., Cha, M.J., Kim, J.Y., Park, C., Watson, D.K., Wang, T., and Choi, K. (2015). Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. *EMBO Rep.* **16**, 654–669.
- Liu-Chittenden, Y., Huang, B., Shim, J.S., Chen, Q., Lee, S.-J., Anders, R.A., Liu, J.O., and Pan, D. (2012). Genetic and pharmacological disruption of the TEAD-YAP complex suppresses the oncogenic activity of YAP. *Genes Dev.* **26**, 1300–1305.
- Medvinsky, A., Rytsov, S., and Taoudi, S. (2011). Embryonic origin of the adult hematopoietic system: advances and questions. *Development* **138**, 1017–1031.
- Milton, C.C., Grusche, F.A., Degoutin, J.L., Yu, E., Dai, Q., Lai, E.C., and Harvey, K.F. (2014). The Hippo pathway regulates hematopoiesis in *Drosophila melanogaster*. *Curr. Biol.* **24**, 2673–2680.
- Mylona, A., Andrieu-Soler, C., Thongjuea, S., Martella, A., Soler, E., Jorna, R., Hou, J., Kockx, C., van Ijcken, W., Lenhard, B., et al. (2013). Genome-wide analysis shows that Ldb1 controls essential hematopoietic genes/pathways in mouse early development and reveals novel players in hematopoiesis. *Blood* **121**, 2902–2913.
- Natoli, G., Ghisletti, S., and Barozzi, I. (2011). The genomic landscapes of inflammation. *Genes Dev.* **25**, 101–106.
- Org, T., Duan, D., Ferrari, R., Montel-Hagen, A., Van Handel, B., Kerenyi, M.A., Sasidharan, R., Rubbi, L., Fujiwara, Y., Pellegrini, M., et al. (2015). Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. *EMBO J.* **34**, 759–777.
- Patterson, L.J., Gering, M., Eckfeldt, C.E., Green, A.R., Verfaillie, C.M., Ekker, S.C., and Patient, R. (2007). The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish. *Blood* **109**, 2389–2398.
- Pereira, C.-F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C.E., Clark, N.R., Nomura-Kitabayashi, A., Kovacic, J.C., Ma'ayan, A., et al. (2013). Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* **13**, 205–218.
- Pouget, C., Peterkin, T., Simoes, F.C., Lee, Y., Traver, D., and Patient, R. (2014). FGF signalling restricts haematopoietic stem cell specification via modulation of the BMP pathway. *Nat. Commun.* **5**, 5588.
- Ptasinska, A., Assi, S.A., Martinez-Soria, N., Imperato, M.R., Piper, J., Cauchy, P., Pickin, A., James, S.R., Hoogenkamp, M., Williamson, D., et al. (2014). Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep.* **8**, 1974–1988.
- Riddell, J., Gazit, R., Garrison, B.S., Guo, G., Saadatpour, A., Mandal, P.K., Ebina, W., Volchkov, P., Yuan, G.-C., Orkin, S.H., et al. (2014). Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell* **157**, 549–564.
- Sandler, V.M., Lis, R., Liu, Y., Kedem, A., James, D., Elemento, O., Butler, J.M., Scandura, J.M., and Rafii, S. (2014). Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature* **511**, 312–318.
- Satyanarayana, A., Gudmundsson, K.O., Chen, X., Coppola, V., Tessarollo, L., Keller, J.R., and Hou, S.X. (2010). RapGEF2 is essential for embryonic hematopoiesis but dispensable for adult hematopoiesis. *Blood* **116**, 2921–2931.



- Shivdasani, R.A., Mayer, E.L., and Orkin, S.H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**, 432–434.
- Simoes, F.C., Peterkin, T., and Patient, R. (2011). Fgf differentially controls cross-antagonism between cardiac and haemangioblast regulators. *Development* **138**, 3235–3245.
- Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V., and Lacaud, G. (2009). In vitro differentiation of mouse embryonic stem cells as a model of early hematopoietic development. *Methods Mol. Biol.* **538**, 317–334.
- Sturgeon, C.M., Ditadi, A., Awong, G., Kennedy, M., and Keller, G. (2014). Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 554–561.
- Tanaka, Y., Joshi, A., Wilson, N.K., Kinston, S., Nishikawa, S., and Gottgens, B. (2012). The transcriptional programme controlled by Runx1 during early embryonic blood development. *Dev. Biol.* **366**, 404–419.
- Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Möröy, T., et al. (2016). GF11 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nat. Cell Biol.* **18**, 21–32.
- Tsankov, A.M., Gu, H., Akopian, V., Ziller, M.J., Donaghey, J., Amit, I., Gnirke, A., and Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518**, 344–349.
- Van Nostrand, E.L., and Kim, S.K. (2011). Seeing elegance in gene regulatory networks of the worm. *Curr. Opin. Genet. Dev.* **21**, 776–786.
- Wadman, I.A., Osada, H., Grutz, G.G., Agulnick, A.D., Westphal, H., Forster, A., and Rabbitts, T.H. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.* **16**, 3145–3157.
- Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* **151**, 206–220.
- Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N.A., Muth, K., Palmer, J., Qiu, Y., Wang, J., et al. (2015). Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16**, 386–399.
- Wareing, S., Mazan, A., Pearson, S., Gottgens, B., Lacaud, G., and Kouskoff, V. (2012). The Flk1-Cre-mediated deletion of ETV2 defines its narrow temporal requirement during embryonic hematopoietic development. *Stem Cells* **30**, 1521–1531.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319.
- Wilson, N.K., Miranda-Saavedra, D., Kinston, S., Bonadies, N., Foster, S.D., Calero-Nieto, F., Dawson, M.A., Donaldson, I.J., Dumon, S., Frampton, J., et al. (2009). The transcriptional program controlled by the stem cell leukemia gene *Scf/Tal1* during early embryonic hematopoietic development. *Blood* **113**, 5456–5465.
- Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schutte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532–544.
- Yu, F.-X., and Guan, K.-L. (2013). The Hippo pathway: regulators and regulations. *Genes Dev.* **27**, 355–371.
- Zinzen, R.P., Girardot, C., Gagneur, J., Braun, M., and Furlong, E.E. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* **462**, 65–70.