A network of epigenetic regulators guides developmental haematopoiesis in vivo

Hsuan-Ting Huang1,2,6, Katie L. Kathrein2,6, Abby Barton2, Zachary Gitlin3, Yue-Hua Huang4, Thomas P. Ward2, Oliver Hofmann3, Anthony Dibiase3, Anhua Song3, Svitlana Tyekucheva4,5, Winston Hide3,4, Yi Zhou2 and Leonard I. Zon1,2,3,7

The initiation of cellular programs is orchestrated by key transcription factors and chromatin regulators that activate or inhibit target gene expression. To generate a comprehensive view of how specific chromatin factors and their associated complexes play a major role in the establishment of haematopoietic cells in vivo, a large-scale reverse genetic screen was conducted targeting orthologues of 425 human chromatin factors in zebrafish. A set of chromatin regulators was identified that target different stages of primitive and definitive blood formation, including factors not previously implicated in haematopoiesis. We identified 15 factors that regulate development of primitive erythroid progenitors and 29 factors that regulate development of definitive haematopoietic stem and progenitor cells. These chromatin factors are associated with SWI/SNF and ISWI chromatin remodelling, SET1 methyltransferase, HDACNuRD deacetylase, and Polycomb repressive complexes. Our work provides a comprehensive view of how specific chromatin factors and their associated complexes play a major role in the establishment of haematopoietic cells in vivo.

Gene expression programs required for the maintenance and differentiation of cell types are tightly regulated by a network of transcription factors and associated chromatin-modifying factors to facilitate or suppress gene expression. Epigenetic information consists of chemical modifications to both cytosine bases in DNA and histone proteins that fold the DNA into nucleosomes, as well as the repositioning, dissociation and/or reconstitution of entire nucleosomes. Mouse knockout models and zebrafish mutants have been used to investigate the role of chromatin factors in vertebrate development, but most chromatin factors have yet to be characterized.

Haematopoiesis is guided by cell-specific transcriptional regulators and associated chromatin factors that function to establish all mature blood cells. A hallmark of the process is the establishment of haematopoietic stem cells (HSCs) and depends on the function of transcriptional regulators such as Runx1, Scl and Lmo2, and chromatin factors, Mll and Bmi1, for stem cell production, self-renewal and survival. Additional factors coordinate the specialization of HSCs into multilineage progenitors that generate differentiated cells of the peripheral blood lineages. Members of the Polycomb (PcG) family have previously been identified as regulators of haematopoiesis. Bmi1 functions to positively regulate HSC proliferation by limiting cell cycle regulator expression, and HSCs with Bmi1 deficiency show impaired self-renewal capacity. The Mi-2–NuRD complex regulates a set of HSC-specific genes that maintain the HSC pool in the bone marrow. De-repression of these genes in Mi-2beta-deficient HSCs exhausts the HSC pool. Several chromatin factors have been identified as leukaemic translocation partners, underscoring the importance they have in normal development. MLL is rearranged in most infant leukaemias with patients generally having poor clinical outcomes. Similarly, translocation of PRDM16, another SET family member, is associated with a poor prognosis. Both promote the development of normal HSCs and leukaemic stem cells.

To identify chromatin factors that function during developmental haematopoiesis, we have undertaken a large-scale in vivo reverse genetic morpholino-based screen targeting zebrafish orthologues of 425 human chromatin factors. The zebrafish provides a suitable platform for rapid screening to assay the function of chromatin factors in haematopoiesis owing to their high fecundity, rapid development, evolutionary conservation, and ease in generating genetic knockdowns. We have identified 44 factors that affect the development of primitive and definitive blood, including 28 factors that have yet to be associated with haematopoiesis. We have also characterized different developmental...
stages during which these factors function, from the induction of stem and progenitor cells to differentiation into erythroid cells. By incorporating protein interaction data, we predict the BAF–PBAF, ISWI, SET1, CBP–p300–HBO1–NuA4, HDAC–NuRD and PRC1–PRC2 complexes as required for blood development. Taken together, our screen provides a valuable resource for elucidating the in vivo network of chromatin regulators of haematopoietic development.

RESULTS
A screen for chromatin regulators of developmental haematopoiesis
To identify the chromatin remodelling factors that function in developmental haematopoiesis, we conducted a large-scale in vivo reverse genetic screen targeting chromatin factors (Fig. 1a). We designed antisense oligonucleotide morpholinos to knock down expression of 488 zebrafish orthologues of 425 human chromatin factors (Supplementary Table 1). Our gene list included most of the known human factors containing chromatin-binding, -modifying or -remodelling domains curated from several public databases: CREMOFAC, SMART domain by NRDB, CDD at NCBI, Pfam and ChromDB (refs 13–17). In zebrafish, 488 orthogonal genes were identified by a reciprocal BLAST search of the unique human protein sequences into the zebrafish genome. Only 26 human proteins lacked a zebrafish orthologue.

Morpholinos targeting each chromatin factor were injected into single-cell embryos at three concentrations. These doses typically give a range of phenotypes from a hypomorph to a near complete knockdown for most messenger RNA products, similar to an allelic series. In some cases, complete knockdown could not be achieved because of lower targeting efficiency or embryonic lethality. Post-injection, embryos were collected at specific time points, using both standard morphological features of the whole embryo and hours post-fertilization (hpf) to stage to minimize differences in embryonic development caused by the morpholino injection18. The embryos were then assayed for haematopoietic defects by whole-mount in situ hybridization (WISH).

We conducted two screens simultaneously for primitive and definitive blood formation. For the primitive screen, developing erythrocytes in the posterior mesoderm of the embryo were assayed by β-globin e3 expression at the 16 somite stage, or 17 hpf (Fig. 1b; ref. 19). For the definitive screen, the establishment of haematopoietic stem and progenitor cells (HSPCs) in the aorta, gonad, mesonephros region (AGM) was detected with c-myb and runx1 expression at 36 hpf (Fig. 1c; ref. 20).

To establish the level of morpholino efficacy, the 21 splice-blocking morpholinos targeting the chromodomain (CHD) gene family were assayed for splicing activity by polymerase chain reaction with reverse transcription (RT–PCR). Of these 21 morpholinos, 10 did not result in any haematopoietic defect. For these 10, one gene could not be evaluated because no PCR product was detected. Only one of the nine remaining morpholinos did not show altered splicing activity, resulting in an estimated false negative rate (FNR) of 11% for the screen (Supplementary Fig. 1a,b). To expand on this limited approach, we verified the splicing activity of an additional 48 splice-blocking morpholinos that scored negative in both primitive and definitive screens. In total, 51 of 57 morpholinos caused altered splicing, resulting in the same estimated false negative rate of 11%. Furthermore, the knockdown efficiencies were comparable to those that gave a haematopoietic defect (Supplementary Fig. 5).

Classification of screen results
Gene expression phenotypes observed in morpholino-injected embryos, or morphants, were classified into one of three main categories: no change, decrease or increase. Owing to the range of decreased staining from subtle to complete absence of staining, we subdivided the decrease category into mild, intermediate or strong (Fig. 2a,b and Supplementary Table 2). Any morphant showing changes in β-globin e3 or c-myb–runx1 expression was considered a screen hit. Morphants with developmental abnormalities were listed separately (Supplementary Fig. 2a,b and Table 2). As morphologically normal morphants with the strongest decrease or increase in blood formation represented genes that were likely to be specific to blood development, these 26 primitive and 47 definitive factors were selected for further characterization.
**Figure 2** Classification of screen results. (a) Summary of primitive screen WISH results. Developing β-globin e3+ erythroid cells are found as two bilateral stripes in the posterior of the embryo, as shown by the red highlights in the schematic of a 16 somite stage (ss) embryo. Knockdown of the different chromatin factors resulted in no change, decrease, or increase of β-globin e3 expression. WT, wild type. (b) Summary of definitive screen WISH results. Induction of c-myb+ and runx1+ HSPCs occurs in the AGM region highlighted in red in the schematic of a 36 hpf embryo. Knockdown of the different chromatin factors screened resulted in no change, decrease or increase of c-myb and runx1 expression. Representative WISH results are shown for each phenotypic category with additional categories continued in Supplementary Fig. 2. Scale bars, 100 μm for low-magnification and 25 μm for high-magnification images.

**Chromatin factors regulate primitive blood development from the mesoderm**

Of the 26 morphants with altered primitive erythropoiesis, knockdown of 13 chromatin factors reduced β-globin e3 expression and 13 factors increased β-globin e3 expression at the 16 somite stage. To confirm these phenotypes, we re-screened for β-globin e3 expression. Of the 26 genes, 16 were verified; 6 from the reduced group and 10 from the increased group (Supplementary Table 3). Given that morpholinos can fail to inhibit their intended targets, splicing activity was confirmed by RT-PCR for the ten splice-blocking morpholinos used, and a second, nonoverlapping morpholino was tested to verify the initial screen result for all factors. Of the 16 genes, 15 were validated in this manner and characterized further (Fig. 3a,b and Supplementary Fig. 5). To show that the morpholinos did not just affect globin expression, one gene from each decrease category was re-evaluated with a second erythroid marker, band3. The decrease in band3 expression was consistent with the decrease in β-globin e3 expression and was not rescued by p53 loss (Fig. 3c), suggesting minimal, if any, morpholino toxicity. Overall, these additional tests provide further support for the validity of the screen results.

To examine early hematopoietic defects during the formation of mesodermal precursors and erythroid progenitors, we evaluated scl and gata1 expression at the 10–12 somite stage (14 hpf), respectively (Fig. 4). Of the 15 morphants tested, 12 showed changes in scl and gata1 expression consistent with the changes in β-globin e3, suggesting that most of the factors categorized in the strong decrease or increase
Chromatin factors regulate the induction of HSPCs

As in the primitive screen, we rescreened the 47 definitive genes (41 with a strong decrease and 6 with a increase in runx1–c-myb expression) and confirmed 31 genes: 26 morphants recapitulated the initial strong decrease in c-myb and runx1 expression in the AGM at 36 hpf, and 4 were verified for increased c-myb and runx1 expression (Supplementary Table 3). For morphants in the decrease category, expression of both markers was nearly abolished in the AGM. Morpholinos injected in a p53<sup>-/-</sup> background did not rescue the loss of expression of AGM markers, indicating that these phenotypes were not the result of morpholino toxicity (Fig. 3d). To verify the efficacy of the morpholinos, splicing activity was assessed and confirmed for all 21 splice-blocking morpholinos used. Of the 31 genes, 29 were validated with a second, non-overlapping morpholino and selected for further characterization (Fig. 3a,b and Supplementary Fig. 5). Collectively, these data provide additional verification of our definitive screen data.

Previous work has shown that HSPCs emerge from the haemogenic endothelium of the dorsal aorta and that proper vessel development and establishment of artery identity is necessary for AGM stem cell induction<sup>22,23</sup>. The expression of the vascular marker flk1 and arterial identity marker ephrinB2 were analysed for the 29 verified genes. Most (20 of 29) showed normal flk1 and ephrinB2 expression. Although these
Chromatin factors regulate distinct steps of primitive erythroid development. Fifteen primitive genes were screened for changes in \( \text{sc} \)l expression marking haematopoietic mesodermal precursors (yellow highlight) and \( \text{g} \)ata1 expression marking erythroid progenitor formation (orange highlight) illustrated in the top panel (flatmount view). \( \beta \)-globin e3 staining was repeated to verify primary screen results (posterior view). A \( \text{myoD} \) probe for labelling somites was included as a staging marker. Blue downward arrows represent reduced expression. Double blue arrows indicate near absence of expression. Magenta upward arrows represent increased expression. The hyphen indicates no change in gene expression. Square brackets indicate the thickness of the stripes. Scale bars, 100 \( \mu \)m.

### Results

- The remaining ten chromatin factors were found to function at earlier stages of vessel specification on the basis of the presence of vascular defects. Four of the factors, \( \text{hdac}4 \), \( \text{mbd3b} \), \( \text{phf21a} \) and \( \text{svr39h1} \), showed normal \( \text{flk1} \) levels but reduced \( \text{ephrinB2} \) expression, suggesting that they function in the establishment of aorta identity upstream of HSPC formation. Finally, 5 morphants, \( \text{mier}1 \), \( \text{jhdmlbb} \), \( \text{jmjd1ca} \), \( \text{ing4} \) and \( \text{rbbp3} \), lost both intersomitic \( \text{flk1} \) and arterial \( \text{ephrinB2} \) expression; hence, the loss of HSPCs in these morphants is probably due to the absence of haemogenic endothelium. Overall, our definitive screen uncovered chromatin regulators involved in the development of HSPCs from the AGM and during vascular development.

### Genes associated with mild to moderate knockdown phenotypes are important regulators of haematopoietic development

In addition to chromatin factors with strong decrease or increase phenotypes, other factors with moderate to mild phenotypes also function in haematopoietic development (Supplementary Table 3). We characterized seven genes from the primitive screen that showed only an intermediate reduction in \( \beta \)-globin e3 expression on reinjection. Of these seven morphants, six had normal \( \text{sc} \)l but reduced \( \text{g} \)ata1 expression, suggesting that they probably function at the erythroid specification stage. Two of these factors, \( \text{CHD4} \) and \( \text{CBX8} \), associate with the FOG1–\( \text{GATA1} \) transcriptional complex and the TIF1–\( \gamma \) elongation complex, respectively; both complexes are key regulators of erythroid development. Finally, \( \text{k} \)at5 showed a decrease in \( \beta \)-globin e3 without loss of \( \text{sc} \)l and \( \text{gata1} \) expression, indicating that it plays a role in erythroid cell differentiation. In comparison with the 15 genes with the strongest phenotypes, those with more mild phenotypes probably function at later stages of erythropoiesis after induction of the \( \text{sc} \)l+ mesoderm.

Similarly, for genes with moderate phenotypes from the definitive screen, we did not observe any defects in \( \text{flk1} \) and \( \text{ephrinB2} \) expression.
Chromatin factors regulate distinct steps of definitive HSPC development. 29 definitive genes were screened for changes in flk1 expression marking the vessels (tan highlight) and ephrinB2 expression marking the artery endothelium (purple highlight) illustrated in the top panel. c-myb and runx1 staining was repeated to verify primary screen results. Blue downward arrows represent reduced expression. The magenta upward arrow represents increased expression. The black dashes indicate no change in gene expression. Scale bar, 25 μm.

To identify all possible chromatin complexes represented by our screen data sets, we generated a human protein–protein interaction map of the 425 chromatin factors screened and mapped the factors found in our screen (102 primitive and 116 definitive factors) onto the network. We identified the same complexes as we did in the previous analysis with the addition of a ubiquitylation complex (Fig. 7 and Supplementary Fig. 3).

As chromatin factors associated with the same complex probably share target binding sites, we analysed 34 published ChIP-seq (chromatin immunoprecipitation followed by sequencing) data sets in K562 erythroleukaemia cells of chromatin factors in our screen27. Two major complexes were found. The first group includes SIN3A, CHD4, HDAC1, TAF1 and JARID1C associated with the HDAC–NuRD complex, and the second group includes RNF2, SUZ12, CBX2 and...
**Figure 6** Identification of chromatin-modifying complexes using protein interaction data for the 44 validated primitive and definitive genes. Protein complexes associated with the 44 chromatin factors from the decrease (strong) and increase categories were identified, and additional genes from the screen present in these complexes were included (Uniprot). Results for SWI/SNF, ISWI, SET1 and CBP-p300 are illustrated with additional complexes shown in Supplementary Fig. 3. Filled in circles represent chromatin factors that were identified as screen hits with their respective phenotypic classification denoted by different colours. Dotted lines indicate alternative complex associations. Scale bar, 100 μm.

CBX8 from the Polycomb complexes. We ranked triplet combinations of these factors together with all other groups of three factors on the basis of the percentage of overlap of target genes. The HDAC–NuRD and PRC1–2 complex combinations predicted from our screen, including those that have been shown to interact biochemically, fell within the top 20% of all possible combinations of three factors (Supplementary Fig. 4a,c). After excluding a subset of ten factors that have large target gene lists (>8,000 target genes), which could skew the distribution, this filtered analysis resulted in the predicted interactions falling within the top 5% of all combinations (Supplementary Fig. 4b,c). Both predicted interactions are significantly enriched in the upper tail of the distribution in the two analyses, therefore suggesting that our screen has identified chromatin factors that function in distinct complexes to regulate haematopoietic development.

**Genetic interaction of predicted chromatin complex subunits**

On the basis of the complexes identified, we expected that chromatin factor subunits that scored positive in our screen would interact genetically. We tested the interaction by combinatorial knockdown of ISWI subunits required for primitive erythropoiesis, smarca1, chrac1 and rsf1b (Fig. 8a; refs 28, 29). As described in the previous section, loss of these factors individually resulted in decreased gata1 and...
**DISCUSSION**

Haematopoietic stem cells undergo proliferation and differentiation under the control of cell-specific transcription factors whose function is facilitated by chromatin factors. These factors establish an epigenetic landscape that controls self-renewal and provides lineage priming, driving differentiation. To better understand the epigenetic regulation of haematopoiesis, we undertook the first reverse genetic approach to define the function of chromatin factors in the zebrafish. A library of zebrafish genes orthologous to 425 human chromatin factors were identified, containing canonical ‘readers’, ‘writers’ and ‘erasers’ of chromatin and other, less characterized, families.

In this study, we characterized a cohort of 15 chromatin factors that regulate primitive haematopoiesis and 29 that regulate definitive haematopoiesis, including both known and previously unidentified factors. On the basis of our validation work, the data suggest that these factors function at the level of erythroid and HSPC specification. Our analysis of several blood-specific markers at several distinct time points has been used to describe differentiation defects in many zebrafish blood mutants, where the absence of marker expression indicates a loss of haematopoiesis\(^32,33\). However, a morphant with delayed haematopoiesis would phenotypically resemble a morphant with a loss of haematopoiesis, and both would be classified as hits in the ‘decrease’ category. In addition, it is possible that an accelerated emergence of blood cells could confound our analysis except when it results in an overall increase in blood production. Regardless of whether blood development is selectively delayed or accelerated, our results ultimately show that expression of the blood markers is altered, whether directly or indirectly affecting any number of pathways, such as metabolism, transcriptional elongation, and cell cycle regulation. Further work will be required to determine the mechanism of action on haematopoiesis and whether these phenotypes are cell autonomous.
Figure 8 Genetic interaction of ISWI chromatin factors by combinatorial knockdown. (a) gata1 and β-globin e3 expression levels were examined by injecting single and combined suboptimal doses of morpholinos against smarca1, chrac1 and rsf1b. The combined dose of 7.5 ng includes 4 ng of smarca1, 0.8 ng of chrac1, and 2.7 ng of rsf1b. (b) Individual or combined knockdown of SWI/SNF subunits baf57, baf170 and brg1 did not alter β-globin e3 expression. These factors were not identified as hits in the screen. The combined dose of 12 ng includes 4 ng of each of the three morpholinos. Morpholino doses are indicated as nanograms (ng). Scale bar, 100 μm.

Disruption of both positive and negative regulators of chromatin frequently resulted in the same phenotype in our screen. For example, knockdown of p300, which acetylates histones, and hdac6, which deacetylates histones, each resulted in loss of c-myb+ and runx1+ cells in the AGM. Although they probably serve opposing roles in regulation of their respective target genes, their functions are both required for proper HSC specification. These data are in concordance with proteomics data showing transcription factors, such as GATA1, recruiting both positive and negative regulators to activate and repress target genes, respectively.

Although members of the same chromatin family could compensate for each other, individual knockdown of many of these factors still resulted in a haematopoietic phenotype, suggesting non-redundant functions among related chromatin factors. Two factors that were identified from our screen, p300 and crebbp, share similar functions but showed opposing phenotypes. p300 and CBP are homologous proteins that share a bromodomain and histone acetyltransferase domain. Mouse knockouts of p300 and Cbp exhibit similar phenotypes. Despite their overlapping roles, evidence of differential regulation has been accumulating. In HSCs, Cbp plays an important role in HSC self-renewal whereas p300 regulates HSC differentiation. Recent ChIP-seq results identified distinct binding sites between the two factors. Consequently, the effect that chromatin factors have in vivo cannot be predicted solely on the basis of their domain function. Future in vivo studies will be important for our understanding of chromatin regulation and gene expression.

In a recent study investigating histone modifications on differentiating erythroid cells in mouse fetal liver, five histone marks were induced during this transition, H3K4me2, H3K4me3, H3K9Ac, H4K16Ac and H3K79me2. Consistent with these findings, our strongest primitive hits are composed of chromatin factors involved in methylation and acetylation of histones, including H3K4 methylation. Although similar work characterizing changes in histone marks during various stages of definitive HSC formation has not been performed, we predict that they will include histone modifications such as methylation of H3K4, H3K9 and H3K36 on the basis of the chromatin factors identified in our screen.

By examining our screen results, we identified relevant chromatin-modifying complexes for blood development including BAF–PBAF, ISWI, HDAC–NuRD, NuA4–p300–CBP–HBO1, SET1 and PRC1–2 complexes. Hypotheses regarding the subunit composition of the chromatin factor complexes can be generated using our data set. One of
the most striking results from our primary screen was the knockdown of smarca1, which abrogates scl, gata1 and β-globin e3 expression in the embryo. chrac1 and rsf1b, other components that form the ISWI complex, were also identified in the screen. In mammalian data, these factors form a complex with another family member, smarca5. Our data suggest that ISWI chromatin remodelling is important for primitive and definitive haematopoiesis and that the complex contains smarca1 (not smarca5), chrac1 and rsf1b. In addition, by comparing chromatin occupancy of complex members, we observed higher proportions of bound genes among factors predicted to be in the same complex. Taken together, these data suggest that our screen has identified chromatin factors that function in distinct complexes to regulate haematopoiesis.

Overall, we have identified a set of genes involved in the regulation of developmental haematopoiesis, including primitive erythropoiesis and definitive HSPC specification, and provide a resource for the identification and characterization of previously unidentified regulators. Studies focusing on the interactions between haematopoietic transcription factors and our chromatin factors will provide a more complete transcriptional network of gene regulation in blood development. In combination with other genetic and biochemical studies, our screen helps to unravel the epigenetic code that establishes the programs of gene expression for self-renewal and differentiation in haematopoietic cells.

METHODS

Methods and any associated references are available in the online version of the paper.

ACKNOWLEDGEMENTS

We thank O. Tamplin, T. V. Bowman, P. Cahan and C. K. Kaufman for helpful discussions. The work was supported by NIH NIDDK SR01DK053298-15, NIH NHLBI 5R01HL084801-21, NIH NIDDK 5P30 DK49216-19, NIH NIDDK DK53298-15, NIH NIDDK R43 DK092760-02, HHMI (to L.I.Z.), NIH NHLBI T32 HL066987-09 and NIH NIDDK 1F32DK089876-01 (to K.L.K.).

REFERENCES

1. Li, B., Carey, M. & Workman, J. The role of chromatin during transcription. Nature 463, 474–484 (2010).
2. Cunliffe, V. Histone deacteylase 1 is required to repress Notch target gene expression during zebrafish neurogenesis and to maintain the production of motoneurones in response to hedgehog signalling. Development 131, 2983–2995 (2004).
3. Gregg, R., Miller, G., Badou, J., Dowling, J. & Link, B. Positional clonalization of the young mutation identifies an essential role for the Brahma chromatin remodelling complex in mediating retinal cell differentiation. Proc. Natl Acad. Sci. USA 100, 6535–6540 (2003).
4. Miller, C., Maves, L. & Kimmel, C. maz regulates Hox expression and pharyngeal segmental identity in zebrafish. Development 131, 2443–2461 (2004).

6. Van der Velden, U. Y., Wang, L., van Lohuizen, M. & Haramis, A. P. The Polycomb group protein Ring1b is essential for pectoral fin development. Development 139, 2210–2220 (2012).
7. Orkin, S. & Zon, L. Hematopoiesis: an evolving paradigm for stem cell biology. Cell 132, 631–644 (2008).
8. Takahara, Y. Role of Polycomb-group genes in sustaining activities of normal and malignant stem cells. Int. J. Hematol. 87, 25–34 (2008).
9. Yoshida, T. et al. The role of the chromatin remodeler Mi-2beta in haematopoietic cell self-renewal and multilineage differentiation. Genes Dev. 22, 1174–1189 (2008).
10. Krivtsov, A. & Armstrong, S. MLL translocations, histone modifications and leukemia stem-cell development. Nat. Rev. Cancer 7, 823–833 (2007).
11. Duhoux, F. P. et al. PRDM16 (1p36) translocations define a distinct entity of myeloid malignancies with poor prognoses but may also occur in lymphoid malignancies. Br. J. Haematol. 156, 76–88 (2012).
12. Chuiakov, S., Levi, B., Smith, M. & Morrison, S. Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. Nature Cell Biol. 12, 1009–1016 (2010).
13. Shipra, A., Chetan, K. & Rao, M. R. CREMOFAC-a database of chromatin remodeling factors. Bioinformatics 22, 2940–2944 (2006).
14. Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl Acad. Sci. USA 95, 5857–5864 (1998).
15. Marchler-Bauer, A. et al. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39, D225–D229 (2011).
16. Punta, M. et al. The Pfam protein families database. Nucleic Acids Res. 40, D319–D326 (2012).
17. Gendler, K., Paulsen, T. & Napoli, C. ChromDB: the chromatin database. Nucleic Acids Res. 36, D298–D302 (2008).
18. Kimmel, C., Ballard, W., Kimmel, S., Ullmann, B. & Schilling, T. Stages of embryonic development of the zebrash. Dev. Dyn. 203, 253–310 (1995).
19. Brownlie, A. et al. Characterization of embryonic globin genes of the zebrafish. Dev. Biol. 255, 48–61 (2003).
20. North, T. et al. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. Nature 447, 1007–1011 (2007).
21. Liao, E. et al. SCL/Tal-1 transcription factor acts downstream of cdc2 to specify haematopoietic and vascular progenitors in zebrafish. Genes Dev. 12, 621–626 (1998).
22. Kissa, K. & Herborn, P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature 464, 112–115 (2010).
23. Burns, C. et al. A genetic screen in zebrafish defines a hierarchical network of pathways required for haematopoietic stem cell emergence. Blood 113, 5776–5782 (2009).
24. Hong, W. et al. The NuRD repressor complex to mediate transcriptional repression by GATA-1. EMBO J. 24, 2367–2445 (2005).
25. Bai, X. et al. TIF1gamma controls erythroid cell fate by regulating transcription elongation. Cell 142, 133–143 (2010).
26. Apweiler, R. et al. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. 40, D71–D75 (2012).
27. Peter, B. B. A User’s Guide to the Encyclopedia of DNA Elements (ENCODE). PLoS Biol. 9, e1001046 (2011).
28. LeJoy, F., Orphandies, G., Lane, W. S. & Reinberg, D. Requirement of RNF and FACT for transcription of chromatin templates in vitro. Science 282, 1900–1904 (1998).
29. Varga-Wesz, P. & Becker, P. Chromatin-remodeling factors: machines that regulate? Curr. Opin. Cell Biol. 10, 346–399 (1998).
30. Wang, W. et al. Architectural DNA binding by a high-mobility-group/kinesin-like subunit in mammalian SWI/SNF-related complexes. Proc. Natl Acad. Sci. USA 95, 492–498 (1998).
31. Wang, W. et al. Diversity and specialized of mammalian SWI/SNF complexes. Genes Dev. 10, 2117–2130 (1996).
32. Ransoms, D. et al. Characterization of zebrafish mutants with defects in embryonic hematopoiesis. Development 123, 311–319 (1996).
33. Thompson, M. et al. The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. Dev. Biol. 197, 248–269 (1998).
34. Rodriguez, P. et al. GATA-1 forms distinct activating and repressive complexes in erythroid cells. EMBO J. 24, 2354–2420 (2005).
35. Ogyuzko, V., Schiltz, R., Russanov, V., Howard, B. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. EMBO J. 15, 2533–2540 (1996).
36. Yao, T. et al. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. Cell 93, 361–433 (1998).
37. Weidinger, M. et al. The role of chromosome separation in embryonic development. Proc. Natl Acad. Sci. USA 99, 14789–14803 (2002).
38. Ramos, Y. et al. Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. Nucleic Acids Res. 38, 5396–5404 (2010).
39. Wong, P. et al. Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. Blood 118, e128–e138 (2011).
40. Turner, B. et al. iRefWeb: interactive analysis of consolidated protein interaction data and their supporting evidence. Database 2010, baq023 (2010).
METHODS

Screen database. All screen data can be accessed at http://zfbmaps.tch.harvard.edu/pfam, supported by Chrome, Firefox and Safari browsers. Information about experimental design and methods are included in the database. All images of WISH results are viewable, including all primary screen data as well as any validation data. The database also provides a collection of morpholino sequences that result in abnormal embryonic phenotypes in morphants, which can be deduced on the basis of the morphological defects or the abnormal staining pattern of the blood markers in the embryo. Live images for select morphants with interesting phenotypes can also be found.

Zebrafish maintenance. Zebrafish (Danio rerio) Tübingen strain were bred and maintained according to Animal Research guidelines of the Institutional Animal Care and Use Committee at Boston Children’s Hospital41. The p53+/− fish used for this study were previously described42. Zebrafish were bred and embryos were collected for microinjection within 45 min post fertilization. Post microinjection, embryos were raised in E3 in 100 mm Petri dishes at a density of fewer than 100 embryos per dish. All larvae were incubated at 25°C and were staged according to standard morphological features for the 17 and 36 h post-fertilization stages rather than by strict hours post fertilization19.

Zebrafish BLAST search. Orthologous genes in zebrafish were identified by reciprocal BLAST search. Briefly, BLAST was performed using the unique human protein sequences to search in the zebrafish non-redundant database and genomic sequences when necessary. Zebrafish orthologues were then used to BLAST search back into the human genome to verify the orthologue identification. Owing to the occurrence of a whole genome duplication event in zebrafish evolution, there are several human genes that have two or more orthologous sequences that can be found in the zebrafish genome. For these genes, we have designated with an a and b suffix according to standard zebrafish nomenclature guidelines. Of the 425 human proteins, only 26 lacked zebrafish orthologues. Note the search was initiated in 2006.

Morpholino design and synthesis. For each zebrafish orthologue, the sequences were analysed to identify the most suitable target sites for morpholino knockdown. Translation blocking, or ATG, morpholinos were designed if the transcriptional start was well annotated. Target regions fall within a 25 base pair region surrounding the ATG site. For genes where the ATG is less well characterized, splice-blocking morpholinos were designed. For these genes, target exons were prioritized on the basis of the proximity to the 5’ end of the gene to produce the shortest possible mRNA product. Exons with a base pair number that is not a multiple of three were selected, when possible, to increase the likelihood of a frameshift mutation after exon excision. All splicing morpholinos targeted exons before or within the chromatin domain of interest. Once the target site was selected, DNA sequences were submitted for morpholino oligonucleotide design and synthesis (Gene-Tools).

To prepare morpholinos for injection, lyophilized morpholinos were dissolved in nuclease-free water to a stock concentration of 2 mM and stored at room temperature. Before microinjection, morpholinos were diluted to the appropriate concentration and heated for 5 min at 65°C to minimize secondary structures. For combined morpholino injections, a single aliquot containing all three morpholinos at the appropriate concentration was prepared. The suboptimal dose for each morpholino in the combination was determined by the dose response of the individual morpholino.

Microinjection. Microinjection of morpholinos was performed at the 1–2 cell stage. Injection volumes were measured using a stage micrometer to a diameter of ~91 µm (0.4 nl) and injected either once, twice or three times into the same embryo to achieve a 3 dose range. All morpholinos were injected into the zebrafish yolk. For ATG morpholinos, the dose range used was 2, 4 and 6 ng and for splice-blocking morpholinos, the dose range used was 4, 8 and 12 ng (ref. 43.44). The higher dose range used for splice-blocking morpholinos is due to less efficient targeting compared with ATG morpholinos. For morpholinos that induced severe morphological defects at this initial dose range, appropriate dilutions were made and the morpholino was rescreened at a lower dose range.

Whole-mount in situ hybridization and imaging. Zebrafish embryos were dechorionated by hand or using promace and fixed in 4% paraformaldehyde at 4°C. Embryos 24 hpf and older were bleached to remove pigmentation, then dehydrated in methanol or ethanol at 20°C. Complementary DNA probes were synthesized from restriction enzyme linearized plasmid and transcribed with T7, SP6 or T3 polymerase in the presence of RNA DIG labelling mix to generate antisense probes (Roche). Newly synthesized probes were purified using RNeasy mini spin columns according to manufacturer’s protocol (Qiagen). Probes used were c-myb, runx1, gata1, scl, β-globin, flk1, band3 and ephrinB2 (refs 19–21,23).

In situ hybridization steps were performed as described previously with the BioRad HT1 in situ robot (Intaviv)43. Briefly, embryos stored in methanol were transferred to standard 100 µm mesh size plates and placed on the robot with 2:1 methanol/PBST (1 x PBS + 0.1% Tween-20) in the tray. Proteinase K treatment time for 36 hpf embryos was 5 min. For 17 hpf and 24 hpf embryos, proteinase K treatment was performed off the robot for 30 s and 2.5 min, respectively. The protocol for day 1 can be found in Supplementary Table 4, Day 1.

Once the pre-hybridization step was completed, the samples were transferred from the robot to an airtight plastic container and incubated overnight in a 70°C hybridization oven. Following the probe incubation step, the samples were transferred back to the robot with hygiene solution in the tray. The protocol for day 2 can be found in Supplementary Table 4, Day 2.

After program completion, trays were transferred to airtight plastic containers for staining. Upon completion of staining, embryos were scored manually. The in situ results for each morpholino injection were determined on the basis of a minimum of 65% of the embryos exhibiting the same phenotype. No statistical method was used to predetermine sample size, but at least 20 embryos per group were assayed in most cases. Representative embryos were imaged using a Nikon stereoscope with a Nikon Coolpix 4500 camera or Zeiss Axiocam camera. Flat-mounted embryos in glycerol were imaged on a Nikon E600 compound microscope. All photos were taken at the same magnification.

RT–PCR. Embryos were collected in Trizol (Invitrogen) and homogenized using a homogenizer. RNA was extracted according to the manufacturer’s protocol and subsequently treated with Turbo DNA-free (Ambion). For reverse transcription reactions, 1 μg of total RNA was used. cDNA was synthesized using the Superscript III (Invitrogen) according to the manufacturer’s instructions. All gene-specific primers used are listed in Supplementary Table 5, and dTMS was used as the control gene. General PCR conditions used were: 95°C for 10 min, 30 cycles of 95°C for 10 s, 58°C for 30 s and 72°C for 1 min.

Human protein interaction network. Protein interactions for the chromatin factors screened were retrieved from the IRef database and visualized using Cytoscape46.48. To increase the connectivity of the network, all additional genes that interacted with more than three nodes were added. Highly interconnected modules involving screen hits were retrieved using NeMo and AllegroMCODE (AllegroVIVA) plugins and refined by manual curation based on complex annotations found in Uniprot49–52.

ChIP-seq data analysis. Putative target genes from each ChIP-seq data set for the 34 chromatin factors screened: CREBBP, p300, HDAC6, HDAC1, SETDB1, JARID1C, CHD7, RNFL, EZH2, SUZ12, CBX8, CBX2, PHF8, RBPP5, LSD1, CHD1, JARID1B, SIRT6, NSD2, HDAC2, PCAF (GEO Accession number GSE32509), ARID3A (GEO accession number GSM955356), HMGN3 (GEO accession number GSM119410), CBX3 (GEO accession number GSM1003568), HDAC8 (GEO accession number GSE1363), CHD4 (GEO accession number GSM1003510), SIN3A (GEO accession number GSM803525), CTCF (GEO accession number GSM733719), SNF5 (GEO accession number GSM955364), CHD2 (GEO accession number GSM953902), SMARCA4 (GEO accession number GSM953635), TAF1 (GEO accession number GSM803431), TRIM28 (GEO accession number GSM103900) and UBTF (GEO accession number GSM103901) were extracted from the ENCODE Consortium and imported into Genomic Regions Enrichment of Annotations Tool (GREAT). All gene sets were analysed using default settings to generate target gene lists26,47. HDAC1, TAF1, PHF8, RBPP5, LSD1, CHD1, CTCF, CBX3, ARID3A and HMGN3 have >8,000 target genes and, thus, a large amount of redundancy. These gene lists were excluded from the filtered analysis to eliminate any skewing that might be caused by these large gene sets. The percentage of overlap of target genes was calculated for all possible combinations of 3 factors and ranked from highest to lowest. Percentage overlap was calculated as the proportion of intersecting target genes over the total number of genes for each 3 factor combination.

41. Lawrence, C., Adatto, I., Best, J., James, A. & Maloney, K. Generation time of zebrafish (Danio rerio) and medakas (Oryzias latipes) housed in the same aquarium facility. Lab Anim. 158165 (2012).
42. Berghmans, S. p. et al. 61 Suppl 1, 407–412 (2005).
43. Rai, K. et al. Zebrafish Dnmt1 and Suv39h1 regulate organ-specific terminal differentiation during development. Mol. Cell Biol. 26, 7077–7085 (2006).
44. Shimoda, N., Yamaokshi, K., Miyake, A. & Takeda, H. Identification of a gene required for de novo DNA methylation of the zebrafish no tail gene. Dev. Dyn. 233, 1509–1525 (2005).
45. Thiese, C. & Thiese, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nat. Protoc. 3, 59–128 (2008).
46. Libuti, C. T. et al. Cytoscape Web: an interactive web-based network browser. Bioinformatics 26, 2347–2348 (2010).
47. Rivera, C. G., Vakil, R. & Bader, J. S. NeMo: network module identification in cytoscape. BMC Bioinformatics 11 Suppl 1, S61 (2010).
48. McLean, C. et al. GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28, 495–501 (2010).

© 2013 Macmillan Publishers Limited. All rights reserved.

DOI: 10.1038/nclb2870
Supplementary Figure 1 Efficacy of chromodomain family splice blocking morpholinos. (a) Splicing activity for splice blocking morpholinos targeting chd genes that result in hematopoietic phenotype. (b) Splicing activity for splice blocking morpholinos targeting chd genes that do not result in a hematopoietic phenotype. chd6 had no expression detected. Filled in arrowheads indicate wild-type bands. Empty arrowheads indicate splice variants.
Supplementary Figure 2 Screen hits with morphological defects. (a) Summary of primitive screen WISH results with morphological defects as determined by the abnormal patterning of the bilateral stripes of β-globin e3+ erythroid cells. (b) Summary of definitive screen WISH results with morphological defects as determined by stunted embryonic growth in the tail. Representative WISH results are shown for each phenotypic category. “n” is the number of chromatin factors with the indicated phenotype. Blue downward arrows represent reduced marker expression. One arrow indicates a mild change, two arrows an intermediate change, and three arrows a strong change. Scale bars: 100μm for low magnification and 25 μm for high magnification.
Supplementary Figure 3 Protein interaction modules containing chromatin factors identified from primitive and definitive screens. Subnetworks isolated from Fig. 6 are shown below including first neighboring interacting nodes. Each node annotated to a specific chromatin complex corresponding to those in Fig. 6 is indicated by uniquely colored borders. The entire complex is demarcated by dotted lines in the same color. Screen classifications for each chromatin factor are indicated by the color of each node.
Supplementary Figure 4 Chromatin factor hits from predicted protein complexes have overlapping target genes in K562 erythroleukemia cells. Distribution plots of percent overlap of target genes for chromatin factor ChIP-seq datasets. Tickmarks along the x-axis represent each chromatin factor triplet combination. PRC1/PRC2 combinations are shown in blue, HDAC/NuRD combinations are shown in green. All others are shown in grey. (a) Distribution plot of triplets from 24 chromatin factor datasets excluding chromatin factors that bind to >8,000 genes (filtered). NuRD triplets containing HDAC1 and TAF1 were ignored PRC1/PRC2 and HDAC/NuRD combinations are significantly overrepresented in the top 5% of the distribution (Fisher’s exact test; p = 6.1x10^-6 and 0.05*, respectively). (b) Distribution plot of triplets from 34 chromatin factor datasets (all). PPRC1/PRC2 and HDAC/NuRD combinations are significantly overrepresented in the top 20% of the distribution (Fisher’s exact test; p = 0.002 and 1x10^-7, respectively). (c) Chart listing combinations for PRC1/PRC2 and HDAC/NuRD complexes. % overlap is the proportion of intersecting target genes over the total target genes for the 3 factors. Rank is ordered from highest to lowest % overlap. *Only one NuRD triplet was tested for this distribution based on the filtering criterion.
Supplementary Figure 5 Uncropped gel images from RT-PCRs.
Supplementary Table 1  Morpholino screen target gene list. Genes are organized by chromatin associated families. Human gene name, human RefSeq ID, zebrafish gene name, zebrafish RefSeq ID, chromosome location, type of morpholino (ATG or splicing target exon), and morpholino sequences are listed.

Supplementary Table 2  Summary of screen results. Each chromatin factor was classified based on marker expression levels following in situ hybridization. Results for the two screens (Primitive and Definitive) are listed separately. The classes of expression levels are no phenotype, no phenotype (with morphology), mild, mild (with morphology), intermediate, intermediate (with morphology), strong, and strong (with morphology).

Supplementary Table 3  Summary of validation results. Each chromatin factor was classified based on marker expression levels following in situ hybridization. Results for the two screens (Primitive and Definitive) are listed separately. The classes of expression levels are normal (N), low, decrease (L), and high, increase (H).

Supplementary Table 4  Whole mount in situ robot protocol. Step-wise details of the in situ program for the Biolane HTI in situ robot.

Supplementary Table 5  RT-PCR primer sequences. Primers used for detecting splicing activity in splice blocking morpholinos. Primers were designed using the Primer3 tool.