Main manuscript for

CHD7 and Runx1 interaction provides a braking mechanism for hematopoietic

differentiation

Jingmei Hsu[1,2,14†], Hsuan-Ting Huang[3,14††], Chung-Tsai Lee[2] , Avik Choudhuri[3], Nicola K.

Wilson[4], Brian J. Abraham[5], Victoria Moignard[4], Iwo Kucinski[4], Shuqian Yu[2], R. Katherine

Hyde[6†††], Joanna Tober[2], Xiongwei Cai[2], Yan Li[2], Yalin Guo[7], Song Yang[3], Michael Superdock[3],

Eirini Trompouki[3††††], Fernando J. Calero-Nieto[4], Alireza Ghamari[8], Jing Jiang[9], Peng Gao[10],

Long Gao[10], Vy Nguyen[3], Anne L. Robertson[3], Ellen M. Durand[3], Katie L. Kathrein[3], Iannis

Aifantis[11], Scott A. Gerber[12], Wei Tong[9], Kai Tan[10,13], Alan B. Cantor[8], Yi Zhou[3], P. Paul Liu[6],

Richard A. Young[5], Berthold Göttgens[4], Nancy A. Speck[2*], and Leonard I. Zon[3*]

Affiliations:

[1] Division of Hematology/Oncology, Department of Medicine, Hospital of the University of
Pennsylvania, Philadelphia, Pennsylvania 19104, USA

[2] Abramson Family Cancer Research Institute and Department of Cell and Developmental
Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

[3] Stem Cell Program and Division of Pediatric Hematology/Oncology, Boston Children's Hospital
and Dana Farber Cancer Institute, Howard Hughes Medical Institute, Harvard Stem Cell
Institute, Harvard University, Boston, Massachusetts 02115, USA

[4] Cambridge Institute for Medical Research, Department of Haematology, Wellcome Trust/MRC
Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom

[5] Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute
of Technology, Cambridge, Massachusetts 02142, USA

[6] National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland
20892, USA

[7] Department of Microbiology and Immunology, Geisel School of Medicine, Lebanon, New
Hampshire 03756, USA

[8]Division of Hematology/Oncology, Boston Children's Hospital and Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA

[9]Division of Hematology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA

[10]Division of Oncology and Center for Childhood Cancer Research, Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[11]Department of Pathology and Laura and Isaac Perlmutter Cancer Center, New York University School of Medicine, New York, New York, 10016, USA

[12]Department of Genetics, Geisel School of Medicine, Lebanon, New Hampshire 03756, USA

[13]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[†]Present address: BMT program, Division of Hematology/Oncology, Department of Medicine, Weill Cornell Medical Center, New York, NY, USA

[††]Present address: Department of Human Genetics, University of Miami Miller School of Medicine, Miami, Florida 33136, USA

[†††]Present address: Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, Omaha, Nebraska 68198, USA

[††††]Present address: Department of Cellular and Molecular Immunology, Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany

[14]These authors contributed equally to this work: Jingmei Hsu and Hsuan-Ting Huang

*Corresponding authors: zon@enders.tch.harvard.edu (L.I.Z.); nancyas@upenn.edu (N.A.S.).

Nancy A. Speck
https://orcid.org/0000-0002-1893-582X

Leonard Zon
https://orcid.org/0000-0003-0860-926X

Classification
Biological Sciences, Developmental Biology, Genetics

Keywords (3 to 5 words)
Hematopoiesis, RUNX1, CHD7

Author Contributions

J.H., X.C., J.T, Y.L., J.J., and S.Y. performed the mouse experiments. H-T.H. and V.N. performed zebrafish experiments. J.H. and C-T.L. performed biochemical experiments. W.T., J.J, J.H. and X.C. performed the limiting dilution transplant experiments. R.K.H. performed the *CBFB-MYH11* mouse experiments. N.K.W. and V.M. performed ChIP-seq in the inducible Cbfb-Myh11 HPC cell line generated by F.J.C-N. A.C., A.G., and I.K. generated CRISPR knockout cells and performed experiments. Y.G. and S.A.G. performed the mass spectrometry. J.H., H-T.H., A.C., N.K.W., B.J.A., V.M., I.K., S.Y., M.S., A.L.R., and L.G. performed bioinformatic analyses. E.T. generated the *Tg(runx1:hsp70,cmlc:dsRed)* zebrafish line. K.L.K. and P.G. performed ChIP-seq. E.M.D. performed microarray. P.P.L. participated in the *CBFB-MYH11* experiments. I.A., K.T., A.B.C., Y.Z., R.A.Y., and B.G. participated in bioinformatic analysis. J.H., H-T.H., N.A.S., and L.I.Z. designed and interpreted experiments and wrote the manuscript.

This PDF file includes:
Main Text
Figures 1 to 5
Figures S1 to S5

1 **Abstract**

2 Hematopoietic stem and progenitor cell (HSPC) formation and lineage differentiation involve

3 gene expression programs orchestrated by transcription factors and epigenetic regulators.

4 Genetic disruption of the chromatin remodeler chromodomain-helicase-DNA-binding protein 7

5 (CHD7) expanded phenotypic HSPCs, erythroid, and myeloid lineages in zebrafish and mouse

6 embryos. CHD7 acts to suppress hematopoietic differentiation. Binding motifs for RUNX and

7 other hematopoietic transcription factors are enriched at sites occupied by CHD7 and

8 decreased RUNX1 occupancy correlated with loss of CHD7 localization. CHD7 physically

9 interacts with RUNX1 and suppresses RUNX1-induced expansion of HSPCs during

10 development through modulation of RUNX1 activity. Consequently, the RUNX1:CHD7 axis

11 provides proper timing and function of HSPCs as they emerge during hematopoietic

12 development or mature in adults, representing a distinct and evolutionarily conserved control

13 mechanism to ensure accurate hematopoietic lineage differentiation.

14

15 **Significance**

16 Hematopoiesis involves the control of gene expression that regulates the processes of

17 proliferation and differentiation. We found that the chromatin remodeler CHD7 controls the

18 differentiation process. Knockdown or knockout of CHD7 leads to enhanced hematopoietic

19 differentiation in zebrafish and mice, suggesting that CHD7 acts as a brake on gene expression

20 associated with terminally differentiated blood cells.

21

22

23 **Introduction**

24   Hematopoiesis is established in three waves at different anatomic sites in all vertebrate

25 embryos. The first, primitive wave takes place in the yolk sac and generates primitive

26 erythrocytes, macrophages and megakaryocytes. In the second, definitive wave, hematopoietic

27 progenitor cells (HPCs) differentiate from hemogenic endothelium in the yolk sac and dorsal

28 aorta (1), and in the placenta, vitelline, and umbilical arteries in mice and humans (2). The third

29 wave, also derived from hemogenic endothelium in the dorsal aorta, vitelline and umbilical

30 arteries, produces hematopoietic stem cells (HSCs).

31   A large-scale *in vivo* reverse genetic screen targeting zebrafish homologs of 425 human

32 chromatin factors with antisense oligonucleotide morpholinos to identify genes controlling

33 embryonic hematopoietic stem and progenitor cell (HSPC) formation uncovered chromodomain-

34 helicase-DNA-binding protein 7 (Chd7) as the only factor that increased the expression of both

35 primitive and definitive hematopoietic genes, including *runx1*, when inhibited (3). The CHD class

36 of ATP-dependent chromatin remodeling enzymes alters nucleosome structure and has been

37 implicated in the maintenance of mouse embryonic stem cells, mammalian development, DNA

38 damage response, and transcription regulation (4). Autosomal dominant *CHD7* mutations cause

39 the inherited CHARGE and Kallmann syndromes (5). Mutations and copy number variations of

40 *CHD7* have been found in hematologic and other cancers (6).

41   RUNX1 is a master transcription factor absolutely required for hemogenic endothelial

42 specification and the endothelial to hematopoietic cell transition in zebrafish and mice (7-12).

43 Loss of RUNX1 in adult HSCs results in HPC and myeloid lineage expansion and lymphoid

44 lineage depletion (13). Here, we show that CHD7 genetically interacts with RUNX1 during

45 hematopoietic ontogeny and adult hematopoiesis, and disruption of CHD7 leads to enhanced

46 hematopoietic differentiation.

47

48

49

50

Results

CHD7 negatively regulates hematopoietic development.

Morpholino (MO) knockdown of *chd7* in zebrafish embryos (*chd7* morphants) increased the expression of primitive erythroid-specific *β-globin e3* (*hbbe3)* at 17 hours post-fertilization (hpf), and definitive hematopoietic markers *myb* and *runx1* in the dorsal aorta (DA) at 36 hpf (SI Appendix, Fig. S1A-D). Markers of hematopoietic mesoderm at 10-12 hpf (*tal1* and *lmo2*) were elevated in *chd7* morphants, while earlier expression of the pan-mesodermal markers *ntla* and *eve1,* and the early hematopoietic marker *drl* at 6 hpf were normal (Fig. 1A and SI Appendix, Fig. S2A). Markers of primitive erythroid (*gata1, gata2, hbbe3*) and myeloid cells (*mpx* and *spi1b)* were increased in the posterior lateral mesoderm, and expression of definitive myeloid (*myb*) and erythroid (*hbae1, hbbe1*) markers were elevated in the caudal hematopoietic tissue (CHT), a site of HSPC colonization (Fig. 1A,B and SI Appendix, Fig. S2A). Expression of endothelial *kdrl* was normal, and analysis of *Tg(kdrl:EGFP) chd7* morphants showed no increase in GFP[+] endothelial cells (SI Appendix, Fig. S1E,F). Hence, *chd7* regulates both primitive and definitive hematopoietic lineage gene expression in the zebrafish embryo. The increased expression of hematopoietic genes correlated with enhanced hematopoiesis. *chd7 Tg(myb:EGFP)* morphants had 1.6-fold more myb:EGFP[+] cells in the DA and 2.8-fold more in the posterior tail region than control embryos (Fig. 1C). Similar increases were observed in two other hematopoietic transgenic lines, *Tg(lmo2:DsRed)* and *Tg(itga2b:EGFP)* (SI Appendix, Fig. S1G), and in an independent study (14). We confirmed a reduction of *rag1* expression in the thymus indicative of a decrease in T lymphocyte progenitors (SI Appendix, Fig. S2A), which was attributed in a previous study to severely impaired thymus organogenesis (14). We conclude that *chd7* negatively regulates HSPC formation in the zebrafish embryo.

To determine if the function of CHD7 in hematopoiesis is conserved in the mouse, we measured the number of phenotypic HSPCs in the aorta-gonad-mesonephros (AGM) region

of *Chd7* mutant embryos. HSPCs in mouse embryos briefly accumulate as clusters of Runx1[+]CD31[+]Kit[+] cells attached to luminal endothelial cells in the major caudal arteries, peaking in number at embryonic day (E) 10.5 (15). Germline deletion of CHD7 caused a developmental delay by E10.5 and lethality by E11.5 (16), preventing accurate assessment of AGM hematopoiesis in null embryos. Therefore, we enumerated Runx1[+]CD31[+]Kit[+] hematopoietic cluster cells in *Chd7*[+/-] embryos, which are viable, and in embryos with *Chd7*[f] alleles deleted by Cre driven by vascular endothelial cadherin (*Cdh5*) regulatory sequences (*Cdh5*-Cre) (SI Appendix, Fig. S3A). There was a significant increase in Runx1[+]CD31[+]Kit[+] cells in the AGM region of *Chd7*[f/f]*;Cdh5-Cre* embryos and a trend towards increased numbers in *Chd7*[+/-] embryos (Fig. 1D,E). The numbers of definitive erythroid and myeloid progenitors in the yolk sac of mutant embryos were also increased (Fig. 1F), but lymphoid progenitor numbers in the AGM, umbilical and vitelline arteries (A+U+V) were lower (Fig. 1G). The number of primitive erythroid progenitors in the yolk sac was not altered (Fig. 1H). We conclude that CHD7 restrains the numbers of definitive erythroid and myeloid progenitors differentiating from the yolk sac and phenotypic HSPCs in the AGM region in both zebrafish and mouse embryos.

CHD7 functions to regulate myeloid output from HSCs.

We evaluated the hematopoietic function of CHD7 in the mouse by excising *Chd7*[f] alleles with *Vav1-Cre* (SI Appendix, Fig. S3A,B). CHD7 is most highly expressed in long term repopulating HSCs (LT-HSCs) (SI Appendix, Fig. S3C), but deletion of CHD7 did not significantly affect the percentages of phenotypic LT-HSCs (CD48[-]CD150[+] lineage[-]Sca1[+]Kit[+] (LSK)) or restricted hematopoietic progenitors (HPC-1, CD48[+]CD150[-]LSK), although there was a significant increase in CD48[-]CD150[-] LSK cells, which contain several populations of multipotent progenitors (MPPs) (17) (Fig. 2A). The frequency of functional CHD7 deficient LT-HSCs in both whole BM and in purified CD48[-]CD150[+]LSK cells, determined by limiting dilution

103   transplants, was increased >2-fold when donor contribution to Mac1$^+$ peripheral blood (PB) cells

104   was scored (Fig. 2B,C). In contrast, no significant increase in LT-HSC frequency was detected

105   when donor contribution to CD48$^-$CD150$^+$LSK cells in the recipient BM was scored (SI

106   Appendix, Fig. S3D), suggesting that CHD7 does not affect the frequency of LT-HSCs, but

107   constrains the myeloid lineage output of LT-HSCs and potentially other downstream

108   progenitors. Serial transplantation of BM cells revealed no differences between control and

109   CHD7 deficient HSCs by the fourth transplant, thus CHD7 deficient LT-HSCs had normal self-

110   renewal capacity (SI Appendix, Fig. S3E,F). There were no differences in proliferation,

111   quiescence, or apoptosis between CHD7 deficient and control LT-HSCs, MPPs, or HPC-1s in

112   adult mice (SI Appendix, Fig. S3G-I). The frequencies of phenotypic LT-HSCs and HPC-1s were

113   not differentially affected by stress induced by 5-fluorouracil injection (SI Appendix, Fig. S3J).

114   We conclude that CHD7 loss does not negatively impact most functional properties of LT-HSCs.

115

116   Gene expression profiling did, however, reveal differences between control and CHD7

117   deficient LT-HSCs. Ingenuity Pathway Analysis determined that genes upregulated in CHD7

118   deficient LT-HSCs were significantly associated with hematopoietic system development and

119   function, immune cell trafficking, cell to cell signaling and interaction, and cellular movement

120   (Fig. 2D and Supplementary Dataset 1,2). Genes representative of each blood lineage,

121   including erythroid (*Hbb-b1*, *glycophorins*), myeloid (*Mpo, Lyz1*, *Alox5*), and lymphoid (*Thy1*)

122   were upregulated in CHD7 deficient LT-HSCs (Fig. 2E, Supplementary Dataset 1), suggesting

123   that CHD7 deficiency results in LT-HSCs that are more primed for multilineage differentiation.

124   Additional evidence that CHD7 constrains myeloid lineage differentiation include an elevated

125   frequency of differentiated Gr1$^+$Mac1$^+$ cells in the liver of E14.5 *Chd7$^{+/-}$* fetuses (SI Appendix,

126   Fig. S4A) and in adult BM and spleen of *Chd7$^{f/f}$;Vav1-Cre* mice (SI Appendix, Fig. S4B), and

127   increased numbers of granulocyte/monocyte progenitors in culture (SI Appendix, Fig. S4C). We

128 conclude that CHD7 deficient LT-HSCs are more primed for differentiation, particularly of

129 myeloid lineage cells.

130

131 **CHD7 cooperates with the RUNX1 transcription factor.**

132 To determine how CHD7 regulates hematopoietic genes, we performed chromatin

133 immunoprecipitation followed by sequencing (ChIP-seq) to identify CHD7 occupied loci in the

134 murine 416B HPC line (18) (Supplementary Dataset 3). CHD7-occupied gene regions are

135 DNaseI hypersensitive and co-localize with multiple hematopoietic transcription factors (19)

136 (Fig. 3A-D). Ets, Runx, and Gata motifs were enriched at CHD7 bound sites, suggesting that

137 CHD7 functions in part through key hematopoietic transcription factors to regulate

138 hematopoiesis (Fig. 3E). Given that RUNX1 was shown to interact with CHD7 (20), we

139 performed ChIP-seq to determine if RUNX1 influenced CHD7 occupancy. We were unable to

140 generate a RUNX1 deficient HPC line, so instead used a Tet-inducible, neomorphic *CBFB-*

141 *MYH11* fusion gene to block RUNX1 activity (SI Appendix, Fig. S5A). CBFβ-SMMHC (encoded

142 by *CBFB-MYH11*) blocks RUNX1 activity, in part, by sequestering it off the DNA (21). Upon

143 induction of *CBFB-MYH11* expression (+Dox), we identified peaks with >4-fold decreases in

144 RUNX1 binding (Fig. 3F and SI Appendix, Fig. S5B). Of the 3,036 peaks that lost RUNX1

145 binding, 1,043 (34.4%) showed >2-fold and 270 (8.9%) >4-fold loss in CHD7 binding (Fig. 3F

146 ii,iii), as illustrated for the *Evi5* gene (Fig. 3G). In contrast, of the 10,045 RUNX1 peaks that did

147 not decrease upon *CBFB-MYH11* expression, only 781 (7.8%) showed >2-fold and 65 (0.6%)

148 >4-fold CHD7 binding loss (Fig. 3F ii,iv). These results are reproducible and indicate that CHD7

149 is selectively lost from regions of the genome where RUNX1 binding is attenuated

150 (Supplementary Dataset 4). This observation is supported by short-term analysis of

151 CRISPR/Cas9 knockouts for CHD7 and RUNX1 in which initial gene expression changes when

152 either gene is perturbed are positively correlated (SI Appendix, Fig. S5C and Supplementary

153 Dataset 5), and consistent with previous data showing that ~30% of genes that are differentially

154   expressed in *CBFB-MYH11*-expressing cells upon CHD7 loss are direct RUNX1 targets (20).

155   Consequently, CHD7 and RUNX1 function together on target genes important for

156   hematopoiesis in HPCs at the genomic level.

157

158   **CHD7 physically interacts with RUNX1**.

159   We independently identified CHD7 in an unbiased mass spectrometry screen for

160   proteins that interact with RUNX1 and its non-DNA binding partner CBFβ (Fig. 4A and Table

161   S1). We and others confirmed the interaction of endogenous CHD7 with RUNX1 and CBFβ by

162   co-immunoprecipitation in a cell line that expresses high levels of all three proteins (20) (Fig.

163   4B). CHD7 binds the activation domain of RUNX1, as RUNX1 proteins with deletions impinging

164   on the activation domain do not immunoprecipitate CHD7 (Fig. 4C-E). Thus, the RUNX1

165   activation domain, which is essential for RUNX1 function (22), mediates the interaction with

166   CHD7.

167

168   We functionally mapped domains in CHD7 responsible for its ability to constrain definitive

169   hematopoiesis. Overexpression of human CHD7 (hCHD7) mRNA (23) suppresses HPCs in the

170   CHT of zebrafish embryos (Fig. 4F). We injected five different hCHD7 truncation mutants (24)

171   into zebrafish embryos to determine which mutant (mut) proteins could suppress *myb/runx1*

172   expression. Only hCHD7 mut 5 lacking the N-terminal portion of the ATPase/helicase domain

173   failed to suppress *myb/runx1* expression, whereas deletion of the chromodomains (mut 6) or the

174   SLIDE/SANT/BRK domains (mut 2-4) had no effect (Fig. 4G,H, SI Appendix, Fig. S2B). The

175   catalytically dead mutant (23) hCHD7^{K999R} also failed to suppress *myb* expression in the CHT

176   (Fig. 4F, Appendix, Fig. S2B). The ability of CHD7 to suppress hematopoiesis requires its

177   ATPase/helicase activity.

178

CHD7 genetically interacts with RUNX1 to regulate hematopoiesis.

We tested if CHD7 and RUNX1 genetically interact. We observed an increased percentage of Gr1$^+$Mac1$^+$ cells in the spleen of *Runx1$^{+/-}$;Chd7$^{+/-}$* compared to *Chd7$^{+/-}$* mice, suggesting that CHD7 represses adult myelopoiesis in collaboration with RUNX1 (Fig. 5A). We also identified a genetic interaction during embryonic hematopoiesis. In the absence of RUNX1, primitive erythropoiesis in the murine yolk sac is delayed, as evidenced by an increase in immature Ter119^{lo}Kit$^-$ cells and a commensurate decrease in mature Ter119^{hi}Kit$^-$ cells (Fig. 5B,C). This delay is more pronounced in embryos heterozygous for the neomorphic *Cbfb-MYH11* allele (*Cbfb$^{+/M}$*) (Fig. 5C). The dominant negative effect of the *Cbfb^M* allele requires RUNX1, as it is suppressed by RUNX1 deficiency (*Cbfb$^{+/M}$;Runx1$^{-/-}$*) (25) (Fig. 5B). Deletion of either one or both *Chd7* alleles (*Chd7$^{+/-}$* or *Chd7$^{-/-}$*) did not alter the percentages of Ter119^{lo}Kit$^-$ and Ter119^{hi}Kit$^-$ cells, but in *Cbfb$^{+/M}$* embryos, loss of *Chd7* partially restored primitive erythrocyte differentiation (Fig. 5B,C). Since either loss of *Runx1* or *Chd7* can suppress the *Cbfb^M* allele, they both function in the same genetic pathway and are required for the neomorphic activity of the *CBFB-MYH11* fusion gene. This is consistent with a previous study showing loss of CHD7 delays leukemogenesis caused by *CBFB-MYH11* (20).

We also examined the functional relationship between *chd7* and *runx1* during embryonic hematopoiesis in the zebrafish. Definitive hematopoiesis in the DA of zebrafish embryos is impaired by a *runx1^{w84x}* mutation that truncates the Runx1 protein (26). Knockdown of *chd7* in *runx1$^{w84x/w84x}$* embryos did not restore *myb* expression in the DA, as *runx1* is necessary for HSPC development, *myb* expression, and CHD7 activity (Fig. 5D). Therefore, we tested the effect of *runx1* overexpression on CHD7 activity. Ectopic expression of *hCHD7* mRNA decreases *myb* expression in the CHT, while overexpression of *runx1* from an inducible heat shock promoter increases *myb* mRNA levels as expected (Fig. 5E). When *hCHD7* mRNA was

204 expressed in the context of heat shock induced *runx1,* the ability of *runx1* to increase *myb*

205 expression was suppressed (Fig. 5E). Altogether, the data indicate that *chd7* functions to inhibit

206 *runx1* activity during embryonic hematopoiesis.

207

208

209

210 **Discussion**

211 We uncovered a novel mechanism in hematopoietic differentiation in which the activity of

212 the key transcription factor RUNX1 is regulated by the chromatin-remodeling enzyme CHD7.

213 Expression of CHD7 is not blood specific; nonetheless it functions to negatively regulate several

214 stages of embryonic and adult hematopoiesis that are well conserved from zebrafish to mouse.

215 Loss of CHD7 enhances HSPC formation and lineage differentiation, particularly of the myeloid

216 lineage. In adult mice, CHD7 deficient LT-HSCs have a gene expression profile suggestive of a

217 primed state that promotes multilineage differentiation, indicating that the enhanced myeloid cell

218 contribution originates from the LT-HSC. Genomic analysis of CHD7 occupancy suggests that it

219 can cooperate with multiple hematopoietic transcription factor genes. Here, we demonstrate that

220 CHD7 genetically and physically interacts with RUNX1 through the RUNX1 transactivation

221 domain, and that RUNX1 recruits CHD7 to a subset of its target genes. The outcome of the

222 interaction between RUNX1 and CHD7 depends on the developmental context. RUNX1 and

223 CHD7 have opposing effects during HSPC formation from hemogenic endothelium, with RUNX1

224 promoting and CHD7 dampening the process. In the adult, however, RUNX1 and CHD7 both

225 function to restrict myeloid lineage output from LT-HSCs (27, 28). This may have more to do

226 with RUNX1's opposing functions in determining myeloid progenitor numbers in the embryo

227 versus the adult, as we show here that in multiple contexts, in zebrafish embryos, mouse

228 embryos, and adult mice, CHD7 appears to constrain the differentiation of myeloid lineage cells.

229   Distinct functions and mechanisms for CHD7 have been described in different cell types

230   and developmental stages. In murine ES cells, CHD7 is associated with active chromatin and

231   actively transcribed genes, yet its dominant activity is to suppress gene expression (29). In

232   contrast, CHD7 primarily augments gene expression in pre-leukemic HSPCs expressing the

233   CBFB-MYH11 protein (20). CHD7's function in embryonic and adult hematopoiesis is similar to

234   what has been observed in ES cells, i.e. that it primarily dampens the expression of actively

235   transcribed genes.

236

237   Mutational analysis showed that the catalytic activity of the ATPase/helicase domain is

238   required for CHD7 to suppress *myb* expression in zebrafish embryos. Other domains such as

239   the chromodomains and SANT domain, which are required for CHD7's ATPase/helicase activity

240   *in vitro* and are mutated in CHARGE syndrome (24), were not required to repress embryonic

241   hematopoiesis. A possible explanation for this paradox is overexpression of CHD7 lacking

242   chromodomains or the SANT domain may bypass the stricter requirement for these domains in

243   biochemical assays using purified proteins, or *in vivo* where CHD7 is present at physiological

244   levels. ATP binding may be necessary for the ATPase/helicase domain to interact with other

245   proteins. An example of this is the RNA helicase UAP56, which must bind ATP to interact with

246   U2AF in spliceosome assembly (30). Five proteins detected in our CHD7 pulldown are known to

247   be part of, or interact with, the Mi-2/NuRD repressive complex, thus loss of ATP binding by

248   CHD7 could impair the recruitment of repressors at enhancers or gene bodies to regulate gene

249   transcription, leading to increased expression of RUNX1 and other hematopoietic transcription

250   factor targets. A third possibility is there may be multiple ways by which CHD7 is recruited to

251   chromatin and catalysis is activated; in developmental hematopoiesis proteins such as RUNX1

252   may substitute for the chromodomains and SANT domain to recruit CHD7 to chromatin and

253   stimulate its enzymatic activity. Future work elucidating the molecular mechanism of how CHD7

254   navigates the chromatin to regulate transcription will provide additional insights into how

255 epigenetic regulators function with transcription factors to promote appropriate lineage

256 differentiation.

257

258

259 **Materials and Methods**

260 **Animal models.**

261 Zebrafish (*Danio rerio*) Tüebingen strain were bred and maintained according to

262 institutional animal care and use committee guidelines at Boston Children's Hospital.

263 Morpholino sequences were: *chd7 exon 3*: ACTCGTTTATACTCTACACGTACCT; *chd7 exon 4*:

264 TTACAAGCAAGTTTACCTGAAC ACC (Gene Tools, LLC). *Chd7* morpholinos were

265 resuspended in nuclease-free water, and equal amounts of each morpholino were combined

266 (12-15ng) for microinjection at the single cell stage. Standard control morpholino from Gene

267 Tools was used. Fish lines *Tg(myb:EGFP), Tg(lmo2:dsRed), Tg(itga2b:EGFP, Tg(kdrl:RFP),*

268 *Tg(kdrl:EGFP), Tg(hsp70:runx1;cmlc:dsRed),* and *runx1$^{w84x/+}$* were previously described *(26, 31-*

269 *36).*

270

271 Mouse 129S1/Svlmj and C57BL6/J strains were bred and maintained according to

272 institutional animal care and use committee guidelines at University of Pennsylvania.

273 Generation of the murine *Chd7* targeting vector and electroporation into C57BL/6J ES cells was

274 performed by InGenious Targeting Laboratory, Ronkonkoma, NY (details available upon

275 request). Chimeric mice were mated to 129S1/Svlmj x C57BL6/J F1 mice, and progeny

276 backcrossed to C57BL/6J mice for 6 generations. Primers for *Chd7* genotyping were JMH61:

277 AAAATGTGGATCTCTCTCCAAACT, JMH65:TTATTTTCTTGAGACAAGGCCTCAC, JMH66:

278 GGTAACACAC TCCTTTAAACCCAGA. *Vav1*-Cre mice were provided by Thomas Graf (37).

279   *Runx1$^{+/-}$* (*Runx1^{tm1Spe}*), *Cbfb$^{+/M}$* (*Cbfb^{tm1Ihc}*), *Cdh5-Cre,* and *β-actin-Cre* mice were described

280   previously (28, 38-40). Both male and female animals were used for experiments.

281

282   **Embryo staining and microscopy.**

283   Whole mount *in situ* hybridization on zebrafish embryos fixed in 4% paraformaldehyde

284   was performed as described previously (41) with a minimum of two independent replicates

285   performed for each staining. Ratios represent number of embryos with indicated phenotype/total

286   number scored. Stained embryos were imaged using a Nikon stereoscope with a Nikon Coolpix

287   4500 camera or Zeiss camera. Embryos mounted in glycerol were imaged on a Nikon E600

288   compound microscope. Confocal imaging was performed on a Zeiss spinning disk confocal

289   microscope using Volocity (PerkinElmer) or ZEN (ZEISS) software for image acquisition.

290

291   Whole mount immunostaining of mouse embryos were processed as described (15).

292   Primary antibodies used were rabbit anti-mouse CD117 (Thermo Fisher Scientific Cat# 14-1171-

293   82,, RRID:AB_467433), rat anti-mouse CD31 (BD Biosciences Cat# 550274, RRID:AB_393571), and

294   rabbit anti-human/mouse RUNX1 (Abcam Cat# 2593-1, RRID:AB_1580795). Secondary antibodies

295   were purchased from Invitrogen: goat anti-rat Alexa Fluor 647 (Thermo Fisher Scientific Cat# A-

296   21247, RRID:AB_141778), goat-anti rat Alexa Fluor 555 (**Thermo Fisher Scientific Cat# A-21434,**

297   **RRID:AB_2535855**), and goat anti-rabbit Alexa Fluor 488 (**Thermo Fisher Scientific Cat# A-11006,**

298   **RRID:AB_2534074**). Images were collected on a Zeiss LSM 710 confocal microscope equipped

299   with 488-, 543- and 633-nm wavelengths and a 20x immersion objective (Plan-Apochromat

300   25x/0.8 NA). Data were acquired using Zeiss ZEN 2011 and processed using FIJI software (42),

301   LOCI Bio-Formats Importer (http://dev.loci.wisc.edu/fiji/) and the cell counter plugin (version 29

302   February 2008, Kurt De Vos, http://rsb.info.nih.gov/ij/plugins/cell-counter.html). To count

303   hematopoietic clusters, 2 to 3 micron thick Z-sections were collected.

304

Transcription Assays.

Limiting dilution transplantations were performed using total bone marrow cells and sorted $CD48^-CD150^+Lin^{neg}Sca1^+Kit^+$ cells from $Chd7^{f/f}$ or $Chd7^{f/f};Vav1-Cre$ mice (C57BL/6J), along with $2x10^5$ total bone marrow competitor cells (129S1/SvlmJ x B6.SJL-$Ptprc^aPepC^b$/BoYJ F1) into 9 Gy lethally irradiated recipients (B6.SJL-$Ptprca\ Pepcb$/BoyJ (B6-LY5.2/Cr)). To purify $CD48^-CD150^+Lin^{neg}Sca1^+Kit^+$ donor cells, bone marrow cells were lineage depleted using biotinylated antibodies to CD3, CD5, CD19, B220, Gr1, Mac1 and anti-biotin MACS beads on a MACS column (Miltenyi Biotec), then stained with lineage antibodies (B220, Gr1, Mac1, CD3e, Ter119; eFlour450), Sca1 PerCP/Cy5.5, Kit APC-eFlour780, CD48 APC, CD150 PE-Cy7 and sorted on a BD FACSAria.

For serial transplantation, 2×10^6 total BM donor cells were transplanted into lethally irradiated B6-LY5.2/Cr recipient mice without competitors as described above. At 4 months, 2×10^6 BM cells from each primary recipient were transplanted into secondary B6-LY5.2/Cr recipients. The transplant was carried out until tertiary recipients were transplanted into fourth B6-LY5.2/Cr recipients.

Positive engraftment was scored three different ways: 1) ≥1% of all cell lineages in PB were donor-derived at 4 months post-transplantation; 2) ≥1% $Mac1^+$ cells in PB were donor-derived at 4 months post-transplantation; 3) ≥1% $CD48^-CD150^+$LSK cells in BM were donor-derived. ELDA (43) R statistical software was used to assess differences between paired sets of limiting dilution analyses to give the estimate of functional LT-HSC numbers.

To assess response to stress, 150 mg/kg 5-fluorouracil (FU) (Sigma) was injected intravenously and HSCs were analyzed 7 days later by flow cytometry (27).

331 **Progenitor Assays.**

332   Myeloid methylcellulose assay using bone marrow and fetal cells was described

333 previously (44). Embryonic lymphoid T and B progenitors assays using OP9 and OP9-DL1

334 stromal cells were performed as described (45, 46). L-Calc (Stem Cell Technologies) was used

335 to calculate the progenitor frequencies.

336

337 **RNA isolation, microarray processing, and analysis.**

338   Mouse genome 1.0 arrays were used to perform microarray analysis of $Chd7^{f/f}$ and

339 $Chd7^{f/f};Vav1\text{-}Cre$ CD48$^-$CD150$^+$LSK cells. Following lineage depletion, CD48$^-$

340 CD150$^+$Lin^{neg}Sca1$^+$Kit$^+$ cells were sorted directly into TRIzol LS (Ambion, 10296028). cDNAs

341 were generated using the Nugen WT-Ovation Pico system with Exon Module 3 (Nugen, 3300

342 and 2000-12) and were hybridized to Affymetrix Mouse 430v.2 chip (Affymetrix Mouse Gene

343 2.0ST). Intensity CEL files were normalized with the Robust Multichip Average (RMA) algorithm.

344 Expression value fold change cutoff was set at 2-fold and p-value set as 0.05 for Ingenuity

345 Pathway Analysis (IPA) using default parameters.

346

347 MouseGene 2.0 ST arrays were used to perform microarray analysis of LT-HSCs (CD48$^-$

348 CD150$^+$CD34$^-$LSK) ST-HSCs (CD48$^-$CD150$^+$CD34$^+$LSK), and MPPs (CD48$^+$CD150$^-$CD34$^+$LSK)

349 populations from 8 week old C57BL6J mice for Chd7 expression. Differentially expressed genes

350 were assessed as those with at least a log fold expression change of 1 and an FDR based

351 adjusted p-value of <0.1.

352

353   CHD7 deletion mutants cloned in pcDNA3.1 vector (24) were used to synthesize mRNA

354 for microinjection in zebrafish embryos. Constructs were linearized with AvrII (NEB), then

355 purified by ethanol precipitation after phenol:chloroform extraction for mRNA synthesis using

356 mMessage mMachine T7 Ultra kit (Ambion).

17

357

358

359

360   **Flow cytometry.**

361   Flow cytometry was performed on BD LSRII or FACSAria, and data were analyzed with

362   Flowjo (Tree Star Inc.). For analysis of *Tg(kdrl:EGFP)* embryos, embryos were manually

363   dissociated in PBS containing Liberase (Sigma-Aldrich), then washed and filtered in 0.9X

364   PBS/2% FBS.

365

366   Monoclonal antibodies used for analysis of adult and embryonic blood cells in mice were

367   as described above (27) with the following additions/exceptions: CD19 (**BD Biosciences Cat#**

368   **561738, RRID:AB_10893995**), CD71 (**BD Biosciences Cat# 553266, RRID:AB_394743**). FACS

369   staining of the embryonic peripheral blood was performed as described previously (47, 48). For

370   fetal liver HSPC analysis, lineage antibodies included CD3, CD5, CD19, B220 and Gr1, but not

371   Mac1.

372

373   **Cell cycle, proliferation, and apoptosis analysis.**

374   Mouse BrdU assays were performed as described previously (27). Antibodies to Ki-67

375   (**BioLegend Cat# 652405, RRID:AB_2561929**) and Annexin V (**Thermo Fisher Scientific Cat# 88-**

376   **8005-72, RRID:AB_2575162**) were used to analyze proliferation and apoptosis.

377

378   **RT- and qPCR.**

379   Pools of 20 to 50 embryos were homogenized in TRIzol, and RNA extracts were prepared

380   according to the manufacturer's protocol (Life Technologies). Genomic DNA was removed with

381   TURBO DNA-free kit (Ambion), followed by cDNA synthesis using SuperScript III First Strand

382   synthesis kit according to the manufacturer's protocol (Life Technologies). Primers used for RT-

18

383 PCR were: *chd7 ex2* forward: 5'-GGGCACCTACTCACCAATCA-3', *chd7 ex4* reverse: 5'-

384 GCCTCTTTCTTGGTGCTGTT-3', *chd7 ex3* forward: 5'-TCCCAAGACACCCAAAGAAC-3', *chd7*

385 ex5 reverse: 5'-GCCTCTTTCTTGGTGCTGTT-3', *ef1α* forward: 5'-ATCTACAAATGC

386 GGTGGAAT-3', *ef1α* reverse: 5'-ATACCAGCCTCAAACTCACC-3'. qPCR was performed using

387 SsoFast EvaGreen Supermix on a BioRad C1000 CFX-384 real-time PCR machine. qPCR

388 primers used were: *myb* forward: 5'-CCGACAGAAGCCGGATGA-3', *myb* reverse: 5'-

389 TGGCACTTCGCCTCAACTG-3', *runx1* forward: 5'-CGTCTTCACAAACCCTCCTCAA-3', and

390 *runx1* reverse: 5'-GCTTTACTGCTTCATCCGGCT-3'. $^{\Delta}$Ct values were normalized to *ef1α*.

391

392 **Western blot and LC-MS/MS protein identification.**

393 Nuclear extracts from murine T-ALL cell line 720 (49) were incubated with protein A Dynabeads

394 coupled with anti-CHD7 (Abcam Cat# ab31824, RRID:AB_869129), RUNX1 (Millipore Cat# PC284,

395 RRID:AB_2254229), CBFβ (β141.1, Speck lab), Flag (Sigma-Aldrich Cat# A2220,

396 RRID:AB_10063035) or CHD4 (Abcam Cat# ab72418, RRID:AB_1268107) antibodies overnight at

397 4°C or at room temperature for 1 hour.

398

399 For LC-MS/MS, FLAG-tagged CBFβ was expressed in and immunoprecipitated from extracts of

400 the 720 T-ALL cell line (49). The CBFβ immunoprecipitates will contain RUNX subunits and

401 other RUNX-CBFβ interacting proteins. As a control, a FLAG-tagged CBFβ containing two

402 amino acid substitutions that decrease RUNX1 binding (50) was immunoprecipitated as a

403 negative control to subtract out proteins that bound to FLAG-CBFβ in the absence of RUNX.

404 Identification of protein-protein interactors was performed by excision of regions of SDS-PAGE-

405 separated immunoprecipitates, in-gel digestion and peptide shotgun sequencing as described

406 (51). Briefly, gel regions were destained to clarity, dehydrated, digested with trypsin, extracted

407 and analyzed by LC-MS/MS on an LTQ Orbitrap mass spectrometer. The resulting MS/MS

19

408   spectra were data searched using the SEQUEST algorithm (52) against a mouse proteome

409   database (International Protein Index; IPI) and curated to a <1% false discovery rate (FDR)

410   using the target-decoy strategy (53). Proteins were required to have a minimum of 3 peptides to

411   be declared as a positive identification.

412

413   **ChIP-sequencing.**

414   ChIP-seq and DNase I hypersensitivity mapping in 416B cells was performed as described

415   previously (19, 54). *De novo* motif analysis was performed using HOMER (55).

416

417   **RNA-sequencing.**

418   Transcriptional profiles of 416B cells expressing sgRNAs were analyzed using a protocol based

419   on the Smart-Seq2 system (56) with the following modifications: pools of 75 cells expressing

420   each sgRNA were sorted into 96-well plates containing lysis buffer, and cDNA was amplified

421   with 13 PCR cycles. Each experimental condition was analyzed in quadruplicates. Illumina

422   sequencing was perfomed on a HiSeq 4000 instrument. Reads were aligned against mouse

423   reference genome GRCm38/mm10 using GSNAP, and those overlapping exons (ENSEMBL

424   m38.81) were counted using HTSeq (57). Differential gene expression was performed using the

425   DESeq2 package (58).

426

427   **Cell lines.**

428   To generate CBFB-MYH11 expressing cells, the mouse myeloid progenitor cell line 416B was

429   co-transfected with: 1) a plasmid containing the tetracycline transcription silencer (tTS), the

430   tetracycline transactivator (rtTA) and blasticidin resistance under the control of a Ef1α promoter

431   and 2) a plasmid containing the entire Cbfb-Myh11 type A cDNA in frame with an F2A element

432   and mCherry protein under the control of a tetracycline responsive element. As control, cells

433   were alternatively transfected with a plasmid lacking the Cbfb-Myh11 cDNA. Transposase

20

434   PL623 (59) (kindly donated by Pentao Liu, Sanger Institute, Cambridge) was also transiently

435   expressed to promote simultaneous stable integration of the constructs. Plasmids were

436   transfected into 416B cells by electroporation using a BioRad electroporator (220V, 900 µF).

437   After 24 hours, cells were selected in 1 µg/ml of blasticidin (InvivoGen). After 14 days, mCherry-

438   negative single cells that did not stain with DAPI (Sigma) were sorted into 96-well plates using a

439   BD Influx sorter and cultured for typically 2 weeks. Clonal cultures were then tested for induction

440   and expression levels of Cbfb-Myh11 using 1µg/ml of doxycycline. Induction of Cbfb-Myh11 for

441   ChIP experiments, performed in replicate, was confirmed by flow cytometry for mCherry

442   expression on a BD Fortessa.

443

444   416B cells expressing Cas9 protein were obtained by transduction with pKLV2-EF1a-Cas9Bsd-

445   W lentivirus (Addgene #68343) and selected with 10 µg/ml of blasticidin (InvivoGen). Three

446   different *Chd7* and one *Runx1* sgRNAs were cloned into the Perturb-seq GBC library backbone

447   (Addgene #85968), and lentivirally transduced BFP$^+$7AAD$^-$ cells were sorted after 4 days by flow

448   cytometry for RNA-seq or genotyping. Efficiency of CRISPR editing was confirmed by high-

449   throughput sequencing of genomic DNA from 2×10^5 cells. Targeting sgRNAs sequences were:

450   Chd7 sgRNA1 – AGACGCCAATCCGTTCCCG, Chd7 sgRNA2 –

451   TGGTACCTGAACGGCCCGG, Chd7 sgRNA3 – GACATGCCCATAAACGAACG, Runx1

452   sgRNA1 – GCGCACTAGCTCGCCAGGG.

453

454   All cell lines were tested negative for mycoplasma contamination.

455

456   **Statistics.**

457   Quantitative data is shown as mean ± s.d. with *P* values calculated using unpaired two-tailed

458   Student's *t*-test or analysis of variance (ANOVA). No statistical method was used to

459 predetermine sample size. The experiments were not randomized. The investigators were not

460 blinded to allocation during experiments and outcome assessment.

461

462 **Data Availability.**

463 Genomic and microarray data are deposited in Gene Expression Omnibus (GEO) under the

464 following accession numbers: GSE84136, GSE83956.

465

483

References

1. Orkin SH & Zon LI (2008) Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 132(4):631-644.
2. Dzierzak E & Speck NA (2008) Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat Immunol* 9(2):129-136.
3. Huang HT*, et al.* (2013) A network of epigenetic regulators guides developmental haematopoiesis in vivo. *Nat Cell Biol* 15(12):1516-1525.
4. Sims JK & Wade PA (2011) SnapShot: Chromatin remodeling: CHD. *Cell* 144(4):626-626 e621.
5. Layman WS, Hurd EA, & Martin DM (2010) Chromodomain proteins in development: lessons from CHARGE syndrome. *Clin Genet* 78(1):11-20.
6. Forbes SA*, et al.* (2011) COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 39(Database issue):D945-950.
7. Tober J, Maijenburg MW, & Speck NA (2016) Taking the Leap: Runx1 in the Formation of Blood from Endothelium. *Curr Top Dev Biol* 118:113-162.
8. Bertrand JY*, et al.* (2010) Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* 464(7285):108-111.
9. Kissa K & Herbomel P (2010) Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* 464(7285):112-115.
10. Boisset JC*, et al.* (2010) In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* 464(7285):116-120.
11. Eliades A*, et al.* (2016) The Hemogenic Competence of Endothelial Progenitors Is Restricted by Runx1 Silencing during Embryonic Development. *Cell reports* 15(10):2185-2199.
12. Yzaguirre AD, Howell ED, Li Y, Liu Z, & Speck NA (2018) Runx1 is sufficient for blood cell formation from non-hemogenic endothelial cells in vivo only during early embryogenesis. *Development* 145(2).
13. Link KA, Chou FS, & Mulloy JC (2010) Core binding factor at the crossroads: determining the fate of the HSC. *J Cell Physiol* 222(1):50-56.
14. Liu ZZ*, et al.* (2018) Chd7 Is Critical for Early T-Cell Development and Thymus Organogenesis in Zebrafish. *Am J Pathol* 188(4):1043-1058.
15. Yokomizo T*, et al.* (2012) Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nat Protoc* 7(3):421-431.
16. Hurd EA*, et al.* (2007) Loss of Chd7 function in gene-trapped reporter mice is embryonic lethal and associated with severe defects in multiple developing tissues. *Mamm Genome* 18(2):94-104.
17. Oguro H, Ding L, & Morrison SJ (2013) SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* 13(1):102-116.
18. Dexter TM, Allen TD, Scott D, & Teich NM (1979) Isolation and characterisation of a bipotential haematopoietic cell line. *Nature* 277(5696):471-474.
19. Schutte J*, et al.* (2016) An experimentally validated network of nine haematopoietic transcription factors reveals mechanisms of cell state stability. *Elife* 5:e11469.
20. Zhen T*, et al.* (2017) Chd7 deficiency delays leukemogenesis in mice induced by Cbfb-MYH11. *Blood* 130(22):2431-2442.
21. Shigesada K, van de Sluis B, & Liu PP (2004) Mechanism of leukemogenesis by the inv(16) chimeric gene CBFB/PEBP2B-MHY11. *Oncogene* 23(24):4297-4307.
22. Okuda T*, et al.* (2000) Biological characteristics of the leukemia-associated transcriptional factor AML1 disclosed by hematopoietic rescue of AML1-deficient embryonic stem cells by using a knock-in strategy. *Mol Cell Biol* 20(1):319-328.

23. Bajpai R*, et al.* (2010) CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature* 463(7283):958-962.

24. Bouazoune K & Kingston RE (2012) Chromatin remodeling by the CHD7 protein is impaired by mutations that cause human developmental disorders. *Proc Natl Acad Sci U S A* 109(47):19238-19243.

25. Hyde RK, Zhao L, Alemu L, & Liu PP (2015) Runx1 is required for hematopoietic defects and leukemogenesis in Cbfb-MYH11 knock-in mice. *Leukemia* 29(8):1771-1778.

26. Sood R*, et al.* (2010) Development of multilineage adult hematopoiesis in the zebrafish with a runx1 truncation mutation. *Blood* 115(14):2806-2809.

27. Cai X*, et al.* (2011) Runx1 loss minimally impacts long-term hematopoietic stem cells. *PLoS One* 6(12):e28430.

28. Wang Q*, et al.* (1996) Disruption of the *Cbfa2* gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc. Natl. Acad. Sci. USA* 93:3444-3449.

29. Schnetz MP*, et al.* (2010) CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet* 6(7):e1001023.

30. Shen H*, et al.* (2008) Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. *Genes Dev* 22(13):1796-1803.

31. North TE*, et al.* (2007) Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature* 447(7147):1007-1011.

32. Zhu H*, et al.* (2005) Regulation of the lmo2 promoter during hematopoietic and vascular development in zebrafish. *Dev Biol* 281(2):256-269.

33. Traver D*, et al.* (2003) Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol* 4(12):1238-1246.

34. Huang H, Zhang B, Hartenstein PA, Chen JN, & Lin S (2005) NXT2 is required for embryonic heart development in zebrafish. *BMC Dev Biol* 5:7.

35. Choi J*, et al.* (2007) FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Dev Biol* 304(2):735-744.

36. Jing L*, et al.* (2015) Adenosine signaling promotes hematopoietic stem and progenitor cell emergence. *J Exp Med* 212(5):649-663.

37. Stadtfeld M & Graf T (2005) Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development* 132(1):203-213.

38. Landrette SF*, et al.* (2005) Plag1 and Plagl2 are oncogenes that induce acute myeloid leukemia in cooperation with Cbfb-MYH11. *Blood* 105(7):2900-2907.

39. Chen MJ, Yokomizo T, Zeigler BM, Dzierzak E, & Speck NA (2009) Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* 457(7231):889-891.

40. Lewandoski M, Meyers EN, & Martin GR (1997) Analysis of Fgf8 gene function in vertebrate development. *Cold Spring Harb Symp Quant Biol* 62:159-168.

41. Thisse C & Thisse B (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc* 3(1):59-69.

42. Schindelin J*, et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676-682.

43. Hu Y & Smyth GK (2009) ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *Journal of immunological methods* 347(1-2):70-78.

44. Tober J, Yzaguirre AD, Piwarzyk E, & Speck NA (2013) Distinct temporal requirements for Runx1 in hematopoietic progenitors and stem cells. *Development* 140(18):3765-3776.

45. Li Y, *et al.* (2014) Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev* 28(23):2597-2612.
46. Mohtashami M, Zarin P, & Zuniga-Pflucker JC (2016) Induction of T Cell Development In Vitro by Delta-Like (Dll)-Expressing Stromal Cells. *Methods Mol Biol* 1323:159-167.
47. Hyde RK, *et al.* (2010) Cbfb/Runx1 repression-independent blockage of differentiation and accumulation of Csf2rb-expressing cells by Cbfb-MYH11. *Blood* 115(7):1433-1443.
48. Kamikubo Y, *et al.* (2013) The C-terminus of CBFbeta-SMMHC is required to induce embryonic hematopoietic defects and leukemogenesis. *Blood* 121(4):638-642.
49. O'Neil J, *et al.* (2006) Activating Notch1 mutations in mouse models of T-ALL. *Blood* 107(2):781-785.
50. Tang Y-Y, *et al.* (2000) Energetic and functional contribution of residues in the core binding factor b (CBFb) subunit to heterodimerization with CBFa. *The Journal of biological chemistry* 275:39579-39588.
51. Kettenbach AN, *et al.* (2011) Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Sci Signal* 4(179):rs5.
52. Faherty BK & Gerber SA (2010) MacroSEQUEST: efficient candidate-centric searching and high-resolution correlation analysis for large-scale proteomics data sets. *Anal Chem* 82(16):6821-6829.
53. Elias JE & Gygi SP (2010) Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol Biol* 604:55-71.
54. Wilson NK, *et al.* (2016) Integrated genome-scale analysis of the transcriptional regulatory landscape in a blood stem/progenitor cell model. *Blood* 127(13):e12-23.
55. Heinz S, *et al.* (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38(4):576-589.
56. Picelli S, *et al.* (2014) Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* 9(1):171-181.
57. Anders S, Pyl PT, & Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2):166-169.
58. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
59. Wang W, *et al.* (2011) Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. *Proc Natl Acad Sci U S A* 108(45):18283-18288.

Figure 1. *Chd7* negatively regulates embryonic hematopoiesis. **A,** *Chd7* knockdown increases expression of hematopoietic mesodermal precursor, primitive erythroid and myeloid but not early mesoderm genes. Representative embryos for whole mount *in situ* hybridization are shown, with additional genes shown in Appendix, Fig. S2. Regions of blood development are highlighted in red in the embryo schematic. Red arrows and arrowheads, increase. Grey arrows, no change. Scale bars = 50 µm. Replicates: 2. **B,** *Chd7* knockdown increases expression of definitive HSPC and definitive myeloid and erythroid genes. Same descriptions as in panel a. **C,** *Chd7* knockdown in *Tg(myb:EGFP)* embryos increases EGFP$^+$ cells in the dorsal aorta (DA) and tail region (left), quantified in graph (right) (n=53-55). Representative embryos shown from 3 independent replicates. **D**, *Chd7* deletion in mice increases Runx1$^+$CD31$^+$Kit$^+$ hematopoietic clusters detected by confocal imaging of E10.5 *Chd7$^{+/+}$*, *Chd7$^{+/-}$*, and *Chd7$^{f/f}$; Cdh5-Cre* AGM regions. Representative clusters shown. i, one somite pair (sp) area; ii, individual cluster. **E**, Quantification of data from panel D (n=7-13). One-way ANOVA, Dunnett's multiple comparison test; # = comparator. **F,** Increased number of burst forming unit–erythroid (BFU-E) and granulocyte/monocyte progenitors (CFU-GM) in E10.5 *Chd7$^{+/-}$* yolk sacs (n=8-14). GEMM, granulocyte/erythrocyte/monocyte/megakaryocyte progenitors. **G,** Reduced number of lymphoid progenitors in E10.5 *Chd7$^{+/-}$* embryos (n=10-12). A+U+V =AGM, umbilical and vitelline arteries. **H**, Left: the number of erythroid progenitors (EryP) in the yolk sac of *Chd7$^{f/f}$;Cdh5-Cre* embryos is not altered (n=14-15). Right: Both *Chd7^f* alleles were deleted in 65% of the colonies, and one allele was deleted in 27% of the colonies, thus *Cdh5*-Cre was active in the majority of EryP or their precursors (n=colonies from 6-8 yolk sacs). All graphs show mean ± s.d., unpaired two-tailed *t*-test unless otherwise specified.

Figure 2. CHD7 regulates hematopoiesis cell autonomously. **A,** *Chd7* deficiency does not affect phenotypic LT-HSCs. Flow cytometry of LT-HSCs (CD48$^-$CD150$^+$), MPPs (CD48$^-$CD150$^-$), and HPC-1s (CD48$^+$CD150$^-$) from Lin^{neg}Sca1$^+$Kit$^+$ (LSK) bone marrow populations (left),

quantified in bar graph (right) (n=6-7). Mean ± s.d., unpaired two-tailed *t*-test. **B,** Schematic

diagram of mouse limiting dilution transplantation experiments. **C,** The frequency of functional

CHD7 deficient LT-HSCs was increased 2-fold in whole BM (left) and in purified CD48$^-$

CD150$^+$LSK cells (right) when ≥1% donor contribution to Mac1$^+$ PB was scored at 4 months.

LT-HSC frequency was calculated by ELDA (n=7-14 recipients per dose). **D,** Loss of CHD7

increases multilineage hematopoiesis. The top functional categories for genes upregulated in

CHD7 deficient mouse LT-HSCs were enriched for hematopoietic related functions by IPA.

Clusters of individual functional gene annotations (squares) belonging to each category are

labeled numerically. Replicates: 4. **E,** Heatmap of representative lineage specific genes

upregulated in CHD7 deficient CD48$^-$CD150$^+$LSK cells.

Figure 3. CHD7 cooperates with hematopoietic transcription factors to regulate

hematopoiesis. **A,** CHD7 binding distribution in the murine 416B HPC cell line by ChIP-seq.

Replicates: 2. **B,** Gene track of CHD7 binding overlaps with DNaseI hypersensitive sites at the

Tal1 gene. **C,** Overlap of CHD7 binding and DNaseI hypersensitive sites. **D,** Overlap of CHD7

binding and hematopoietic transcription factors. **E,** CHD7 binding sites are highly enriched for

Ets, Runx, and Gata motifs by HOMER motif analysis. **F,** CHD7 binding is selectively lost from

genomic regions where RUNX1 binding is attenuated by CBFB-MYH11. CBFB-MYH11

expression was induced in myeloid progenitor cells by doxycycline (Dox). RUNX1 occupancy

in a i) control clone and ii) *CBFB-MYH11* expressing clone. Loss of CHD7 occupancy is iii)

higher in regions of >4-fold RUNX1 occupancy loss and iv) minimally changed in regions of

<2-fold RUNX1 occupancy loss. Black dotted line, no change (n.c.). Grey line, 2-fold change.

Red line, 4-fold change. Replicate experiment shown in Appendix, Fig. S4B. **G,** Gene tracks

showing loss of RUNX1 and CHD7 binding to *Evi5* (red arrows) in Dox induced *CBFB-MYH11*

expressing cells.

Figure 4. CHD7 interacts with Runx1 and restrains RUNX1 activity. **A,** Scheme for identifying RUNX1-CBFβ interacting proteins in a murine T-ALL cell line. FLAG-tagged CBFβ containing two amino acid substitutions (red stars) that decrease RUNX1 binding was used as a negative control. **B,** CHD7 co-immunoprecipitates RUNX1-CBFβ but not CHD4 in murine T-ALL cells. I, input; S, depleted supernatant following immunoprecipitation; IP, immunoprecipitate. **C,** Deletions impinging on the RUNX1 activation domain decrease the interaction between RUNX1 and CHD7. CHD7 was immunoprecipitated, and Western blots probed with antibodies to CHD7 or FLAG. F-RUNX1, FLAG-RUNX1; Δ deleted amino acids; F, vector expressing FLAG alone. Arrows indicate CHD7 (top panel) or full length and internally deleted RUNX1 proteins (bottom panel). **D,** C-terminal RUNX1 deletions. **E,** Summary of RUNX1 mapping experiments. RD, DNA and CBFβ-binding Runt domain; AD, transactivation domain; ID, inhibitory domain. **F,** Expression of *hCHD7* but not the catalytically dead mutant *hCHD7^{K999R}* in zebrafish embryos reduces *myb* expression in the CHT by whole mount *in situ* hybridization. Representative embryos shown. Blue arrows, decrease. Grey arrows, no change. Scale bars = 50 μm. Replicates: 2. **G,** Mutation mapping of *hCHD7* domains show the ATPase/helicase domain is required to suppress *myb* and *runx1* expression in the CHT. Same descriptions as in panel F. **H,** Summary of hCHD7 mapping experiments. FL, full length. CD, chromodomain. HD, ATPase/helicase domain. SL/SD/BD, SLIDE/SANT/BRK domains. y, yes; n, no. Quantification of results from panels F, G are in Appendix, Fig. S2B.

Figure 5. CHD7 interacts genetically with RUNX1 to regulate hematopoiesis. **A,** *Chd7* and *Runx1* interact genetically to repress myelopoiesis in the spleen of adult mice by flow cytometric analysis (n=4). **B,** Restoration of primitive erythrocyte maturation with *Chd7* deletion in *Cbfb^{M}* embryos with peripheral blood analysis by flow cytometry (n=6-42). Representative plots shown. Simplified genotypes are: *Chd7$^{+/-}$ = Chd7$^{+/f}$;β-actin-Cre, Chd7$^{-/-}$ = Chd7$^{f/f}$;β-actin-Cre, Cbfb$^{+/M}$ =*

Cbfb^{+/MYH11};*β-actin-Cre.* **C,** *Chd7* or *Runx1* deletion partially restores normal maturation of primitive erythrocytes in E10.5 embryos expressing the dominant neomorphic *Cbfb-MYH11* allele (*Cbfb*^{M}). All values were significantly different as compared to *Cbfb*$^{+/M}$. ANOVA and Dunnett's multiple comparison test. **D,** Expansion of myb$^+$ HSPCs in *chd7* morphant (MO) embryos is suppressed in *runx1*^{w84x} mutants expressing truncated Runx1. Whole mount *in situ* hybridization of representative embryos shown, with phenotypic results quantified in bar graph (right). *p < 0.01 by Chi-square test. Red arrows, increase. Blue arrows, decrease. Scale bar = 50 μm. Replicates: 2. **E,** Overexpression of *hCHD7* suppresses the expansion of myb$^+$ HSPCs caused by heat-shock induced *runx1* overexpression. Same descriptions as in panel D.

A

mesoderm formation (6 hpf) → mesoderm induction (12 hpf) → erythroid lineage (17 hpf) → myeloid lineage (24 hpf)

Control / chd7 MO

	ntla	tal1	gata2	mpx
Control	n=19	n=25	n=20	n=13
chd7 MO	n=26/26	n=16/19	n=17/23	n=11/11

ntla ↔ tal1 ↑ gata2 ↑ mpx ↑
eve1 ↔ lmo2 ↑ gata1 ↑ spi1b ↑
drl ↔ hbbe3 ↑

B

DA induction (36 hpf) / CHT hematopoiesis (2 - 5 dpf)

	runx1	hbae1 (5 dpf)
Control	n=25	n=10
chd7 MO	n=25/28	n=14/14

runx1 ↑ myb ↑
myb ↑ hbae1 ↑
ikzf1 ↑ hbbe1 ↑

C

Tg(myb:EGFP) 36 hpf

Control MO — DA, Tail
chd7 MO

Number of EGFP+ cells

○ Control MO   □ chd7 MO

DA 0.006   Tail 0.003

D

1 sp E10.5

Chd7+/+
i 1 sp
ii CD31 / Runx1 / Kit / merge

Chd7+/−
i 1 sp
ii CD31 / Runx1 / Kit / merge

E

☐ +/+   ☐ Chd7+/−   ■ Chd7f/f;Cdh5-Cre

Number of CD31+Runx1+Kit+ cells per sp

≤ 0.002   ***   #

F

○ Chd7f/f   ☐ Chd7f/f;Cdh5-Cre

Number of progenitors per yolk sac

GM 0.05   GEMM   BFU-E 0.02

G

Number of lymphoid progenitors / A+U+V

☐ T   ■ B

≤ 0.02

+/+   Chd7+/−

H

○ Chd7f/f   ☐ Chd7f/f;Cdh5-Cre

EryP progenitors

Chd7f/f;Cdh5-Cre

% deleted alleles

■ f/f   ☐ f/Δ   ☐ Δ/Δ

A

Chd7^f/f

Chd7^f/f
Vav1-Cre

4.4 — Kit (APC-Cy7) vs Lin (eFlour450)
11.6 LSK — Kit (APC-Cy7) vs Sca-1 (PerCp/Cy5.5)
HPC-1 73.7, MPP 6.7, LT-HSC 4.3 — CD48 (FITC) vs CD150 (PE-Cy7)

4.3
7.5
62.1, 16.8, 5.9

○ Chd7^f/f
□ Chd7^f/f;Vav1-Cre

% LSK cells

0.005 0.03

150⁺48⁻   150⁻48⁻   150⁻48⁺

B

Chd7^f/f or Chd7^f/f;Vav1-Cre

Donor (CD45.2) → Recipient (CD45.1)

Competitor (CD45.1/CD45.2) →

Donor contribution to:
CD48⁻CD150⁺LSK cells
Mac1⁺ PB cells

C

Contribution to Mac1⁺ PB cells

● Chd7^f/f   ● Chd7^f/f;Vav1-Cre

% Negative recipients vs Number BM cells
1:66,787
1:18,242
$P \leq 0.0025$

% Negative recipients vs Number CD48⁻CD150⁺LSK cells
1:75.1
1:20.7
$P \leq 0.032$

D

Chd7^f/f; Vav1-Cre CD48⁻CD150⁺LSK cells

P value: -3.6e⁻⁶ to 3.3e⁻³

1. Hematological system development and function
2. Inflammatory response
3. Cell death and survival
4. Tissue morphology
5. Cellular growth and proliferation
6. Organismal injury and abnormalities
7. Immune cell trafficking
8. Cell to cell signaling and interaction
9. Cellular development
10. Hematological disease
11. Cell function and maintenance
12. Cancer
13. Cell cycle
14. Cell movement
15. Tissue development
16. Hematopoiesis
17. Hereditary disorder

E

Z-score: -2 -1 0 1 2

Chd7^f/f   Chd7^f/f;
Vav1-Cre

Anxa3, Gypa] Erythroid
Ctsg
Cd33
C3
Elane
Gp49a   } Neutrophil
Lyz1
Mmp8
Mpo
Ngp
Alox5
Fcer1g] Eosinophil
Cd53
Lilrb4 } Lymphoid
Thy1

A FLAG-CBFβ FLAG-CBFβ (G61A/N104A)

Immunoprecipitation
Shotgun mass spec sequencing
Compare lists

B IP: CHD7   IgG

C IP: CHD7
F-RUNX1: Δ190-330 Δ190-340 Δ190-350 Δ190-360 F-Runx1 F

D IP: CHD7
F-RUNX1: 1-451 1-214 1-302 1-375 1-387

E

RUNX1		CHD7 binding
1-451 F RD AD ID 451		+++
Δ190-330 F RD 190 330 ID		++
Δ190-340 F RD 190 340 ID		++
Δ190-350 F RD 190 350 ID		+
Δ190-360 F RD 190 360 ID		+/−
1-387 F RD AD ID 387		+++
1-375 F RD AD 375		+++
1-302 F RD 302		−
1-214 F RD 214		−

F

Control	hCHD7 mRNA	hCHD7^{K999R} mRNA
n=6	n=9/11 ↓	n=9/9 ↔

probe: myb

CHT 4-5 dpf

G

Control probe: myb/runx1   n=23

FL n=15/22 ↓	mut 4 n=16/24 ↓
mut 2 n=11/18 ↓	mut 5 n=17/21 ↔
mut 3 n=14/25 ↓	mut 6 n=13/16 ↓

H

CHD7		myb/runx1 suppression
FL 1-2997 CD HD SL SD BD		y
mut 2 1-1899 CD HD		y
mut 3 1-1547 CD HD		y
mut 4 787-1899 CD HD		y
mut 5 1285-2138 SL SD		n
mut 6 936-2997 HD SL SD BD		y

A Spleen

B

C

D *runx1^w84x* probe: *myb*

E *Tg(hsp70:runx1)* probe: *myb*