

39 **Abstract**

40

41 Runt-related transcription factor 1 (*Runx1*) can act as both an activator and a repressor.

42 Here we show that CRISPR-mediated deletion of *Runx1* in an embryonic kidney-derived

43 cell (mK4) results in large-scale genome-wide changes to chromatin accessibility and

44 gene expression. Open chromatin regions near down-regulated loci are enriched for Runx

45 sites, remain bound by Runx2, but lose chromatin accessibility and expression in *Runx1*

46 knockout cells. Unexpectedly, regions near upregulated genes are depleted of Runx sites

47 and are instead enriched for Zeb transcription factor binding sites. Re-expressing *Zeb2*

48 in *Runx1* knockout cells restores suppression. These data confirm that Runx1 activity is

49 uniquely needed to maintain open chromatin at many loci, and demonstrate that genome-

50 scale derepression is an indirect consequence of losing Runx1-dependent Zeb

51 expression.

52

53

54

55 **Introduction**

56 Transcription factors (TFs) play fundamental biological roles by controlling gene
57 expression, the first step in translating genomic DNA sequence into function. Mammalian
58 genomes encode over 1,000 TFs, which precisely control gene expression through
59 complex combinatorial interactions and transcriptional cascades (Lambert et al., 2018).
60 To achieve this precision, TFs use a wide range of mechanisms, including initiating the
61 activation or repression of gene expression through the recruitment of co-factors, initiating
62 new chromatin looping interactions between enhancers and target promoters, and
63 altering the chromatin landscape through the repositioning of nucleosomes (Lee and
64 Young, 2013). Achieving an understanding of the mechanisms underlying the control of
65 gene expression is thus an enduring and fundamental goal of molecular biology.

66
67 Runx/Runt TF family members recognize a characteristic TGTGGT DNA-binding motif,
68 and are present across all metazoans (Rennert et al., 2003). Members of the Runx family
69 play important roles in development and disease (Ito et al., 2015; Mevel et al., 2019),
70 notably during hematopoiesis (de Bruijn and Dzierzak, 2017; Seo and Taniuchi, 2020),
71 skin development (Glotzer et al., 2008; Hoi et al., 2010; Osorio et al., 2008), and
72 ossification (Komori, 2018; Mevel et al., 2019; Sierra et al., 2004; Zhang et al., 2008a).
73 Runx proteins can act as repressors, by recruiting the Groucho/TLE proteins via a C-
74 terminal tetrapeptide WRPY, or as activators, by heterodimerizing with Core binding
75 factor (CBF) β and recruiting cell context-specific activators. In several developmental
76 contexts, Runx proteins collaborate with Notch, at times facilitating Notch activity
77 (Giambra et al., 2012; Terriente-Felix et al., 2013), and at others acting downstream of

78 Notch (Kueh et al., 2016). The role that Runx1 plays in establishing chromatin
79 accessibility has been studied in some detail within the hematopoietic system (Lichtinger
80 et al., 2010), where it acts as a pioneer protein. However, how Runx proteins influence
81 gene regulatory networks in different cellular contexts remains to be elucidated.

82

83 Previously, we found that Runx binding sites were enriched near Notch-bound enhancers
84 in a diploid kidney metanephric mesenchymal cell line, mK4 (Hass et al., 2015). Two of
85 the three Runx orthologs, Runx1 and Runx2, are expressed in the kidney-derived mK4
86 cells (Valerius et al., 2002), facilitating detailed molecular comparison of Runx1 versus
87 Runx2 functions in regulating gene expression and their integration with multiple signaling
88 pathways. Such analyses are further aided in mK4 cells by the normal karyotype, the
89 ease of CRISPR-mediated genetic manipulation, and by short replication times, providing
90 sufficient material for a variety of genomic assays.

91

92 In this study, we show that Runx1 plays an important role in regulating chromatin
93 accessibility at many genomic loci in mK4 cells. In the absence of Runx1, Runx2 bound
94 most of the Runx1-bound chromatin but could not maintain Runx1-dependent
95 accessibility or gene expression. As Runx1 can repress expression of some genes, we
96 anticipated these to be re-expressed in Runx1KO cells; however, we were surprised to
97 discover that accessible chromatin near loci expressed only after Runx1 deletion were
98 instead enriched for Zeb sites and depleted of Runx sites, suggesting indirect involvement
99 of Runx1 in their regulation. Further investigation revealed that repression at multiple loci
100 throughout the genome is mediated by two Runx1-dependent targets, Zeb1 and Zeb2.

101 Restoring Zeb2-expression in Runx1KO cells restored repression of target genes. Thus,
102 the direct impact of Runx1 on chromatin in mK4 cells is mediated primarily through its
103 pioneer and transcriptional activator function, rather than through its repressor function.
104 Collectively, these data reveal an important role for Runx TFs in the maintenance of the
105 chromatin landscape and provide mechanistic insight into how Runx and Zeb TFs
106 interactively control gene expression in the kidney.

107

108 **Results**

109 **Generation and characterization of Runx1 knockout cells**

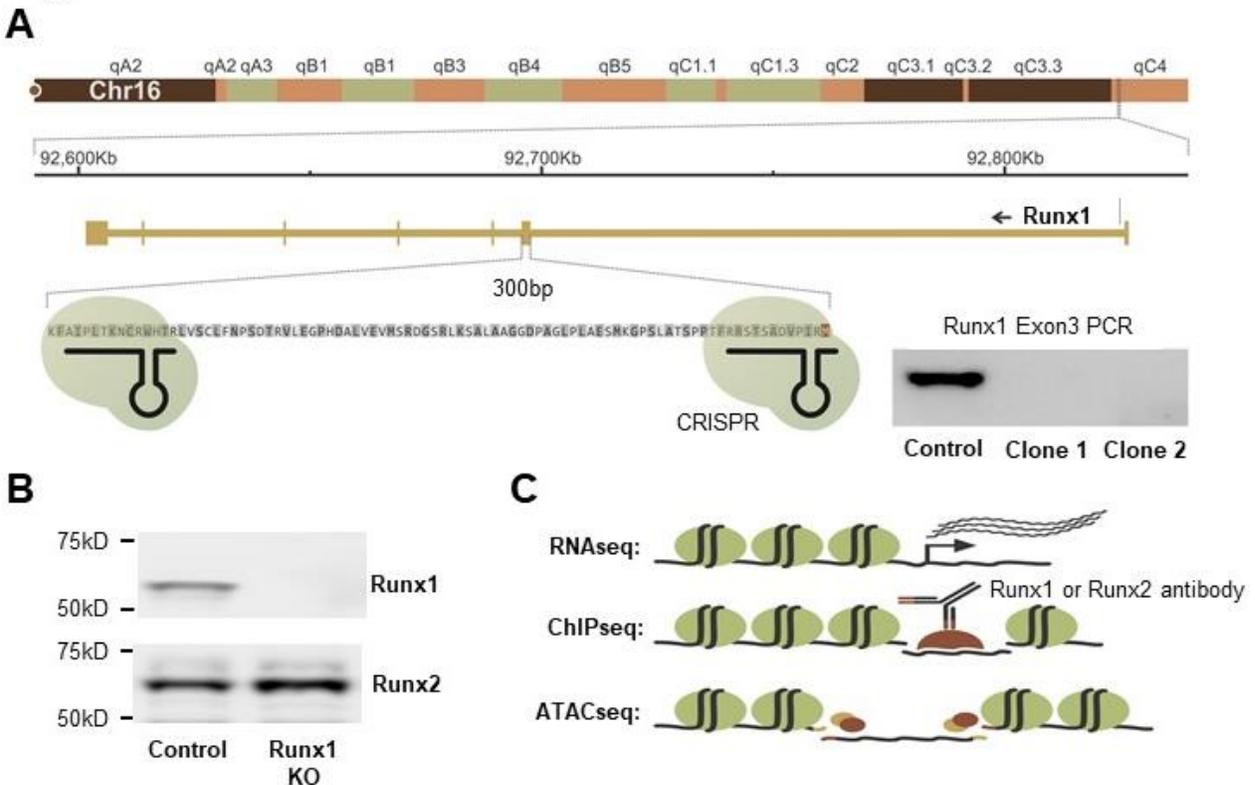
110 To generate cells lacking Runx1 activity, we targeted Runx1 with two gRNAs flanking
111 exon 3, which contains the start codon of the transcript expressed in mK4 cells and
112 encodes part of the Runt DNA binding domain (Figure 1A). Multiple clones grew after
113 selection with puromycin showing deletion of the targeted exon 3 (Figure 1A) by PCR
114 genotyping and loss of Runx1 protein as confirmed by Western blot (Figure 1B). The
115 expression of Runx2 in these cells remained unchanged (Figure 1B; Figure 1-
116 Supplemental Figure 1). Thus, any functional differences between Runx1KO cells and
117 the parental cell line (control) would indicate potential Runx1-specific roles that cannot be
118 compensated for by Runx2.

119

120 To determine the impact of Runx1-deficiency, we performed multiple genomic assays
121 comparing Runx1KO to control mK4 cells. Specifically, we analyzed gene expression
122 through RNA-seq, identified genomic locations bound by Runx1 and Runx2 through ChIP-
123 seq, and mapped chromatin architecture by ATAC-seq (Figure 1C). The RNA-seq, ChIP-

124 seq, and ATAC-seq experiments were all performed in biological triplicates to enable
125 statistical analyses for the identification of significant differences between control and
126 Runx1KO cells.

Figure 1:



127

128 **Figure 1: Generation and Characterization of Runx1KO Cells.** A) Diagram of the Runx1 exon
129 3 region targeted for deletion using CRISPR-Cas9 and confirmation of deletion by PCR. B)
130 Western blot showing that Runx1KO cells lack Runx1 protein but contain Runx2. C) Schematic
131 of genomic analyses utilized to characterize Runx1KO cells.

132

133 Runx1-deficiency induces dramatic changes in gene expression

134 RNA-seq analysis identified thousands of genes that were significantly altered in
135 Runx1KO cells relative to control cells (Figure 2A). The replicates were highly consistent
136 and revealed 1,705 upregulated and 1,182 down-regulated transcripts in the Runx1KO
137 cells (fold change > 2-fold, FDR < 0.05 across replicates) (Supplemental Table 1). GO

138 term analysis of the upregulated genes in Runx1KO cells identified enrichment for the
139 biological process of antigen processing and presentation (5.7 fold enriched, p-value
140 0.03, Supplemental Table 2), consistent with the critical role that Runx1 plays in the
141 immune system and with observations in human patients with Runx1 mutations (Awad et
142 al., 2018). Consistent with previous studies, Runx1KO cell downregulated genes were
143 enriched for the TGF-beta receptor signaling pathway (23.5 fold enriched, p-value 0.0012,
144 Supplemental Table 2) (Zhou et al., 2018). The widespread changes in gene expression
145 caused by Runx1-deficiency are consistent with previous observations of non-
146 redundancy with Runx2, as seen in other cellular contexts (Mevel et al., 2019). Thus,
147 Runx1 plays a critical role in controlling the transcriptome within mK4 cells in a manner
148 that cannot be compensated for by Runx2. We next examined if differences between
149 Runx1 and Runx2 effects on gene expression might be due to differences in DNA binding
150 preferences or differences in genomic binding locations.

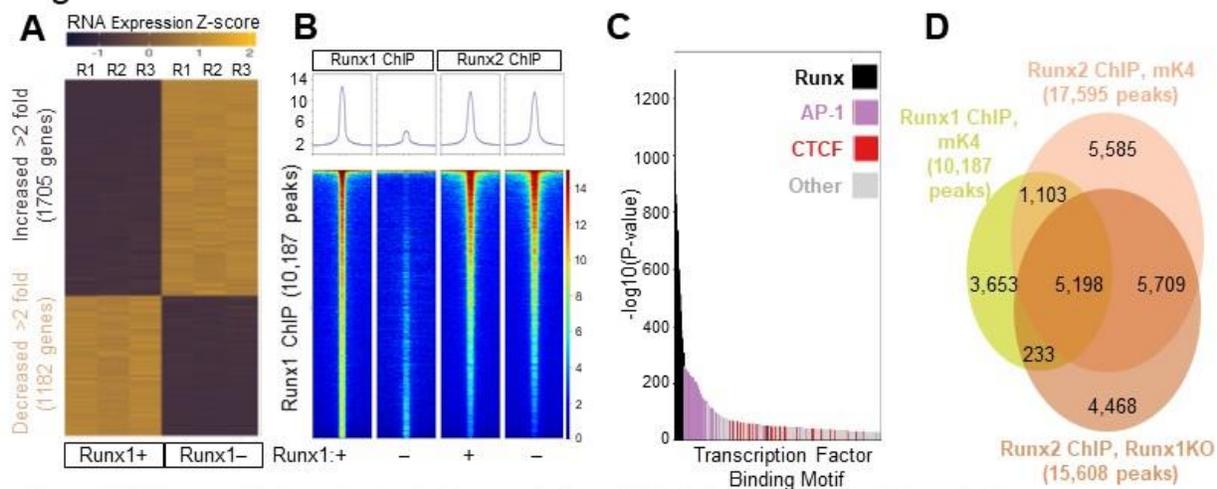
151

152 **Runx1 and Runx2 bind near genes that are down-regulated in Runx1KO cells**

153 To identify genomic loci occupied by Runx proteins, we performed Runx1 and Runx2
154 ChIP-seq. The specificity of the Runx1 antibody used for ChIP was determined by
155 performing the experiment in control and Runx1KO cells. ChIP-seq in control cells
156 produced 10,187 peaks that were highly reproducible between replicates but not present
157 in Runx1KO cells, as depicted in heatmaps of the Runx1 ChIP-seq reads mapped to
158 peaks in control and Runx1KO cells (Figure 2B; Figure 2-Supplemental Figure 1A).
159 HOMER transcription factor binding site motif enrichment analysis using mouse motifs
160 from the Cis-BP database (Lambert et al., 2019) identified Runx motifs as most highly

161 enriched in the control cell dataset (p-value 1×10^{-1298}) (Figure 2C and Supplemental
 162 Table 3). These data, combined with the limited number of peaks and relative lack of
 163 Runx motif enrichment in the Runx1KO cells, confirm the specificity of the antibody and
 164 that the identified genomic regions are bound by Runx1. The second most enriched class
 165 of motifs were for the AP-1 family. Runx1 genomic binding has previously been shown to
 166 be co-enriched with members of this family (Pencovich et al., 2011). Co-association with
 167 AP-1 suggested that many of the Runx1-bound regions are likely enhancers, given the
 168 known association of AP-1 sites with enhancers in most cell types (Andersson et al.,
 169 2014).

Figure 2:



170

171 **Figure 2: Widespread Transcriptional Changes in Runx1KO Cells Despite Runx2 Largely**
 172 **Occupying the Same Regions as Runx1.** A) Heatmap of RNA-seq triplicates showing 1,705
 173 upregulated and 1,182 downregulated genes (over 2 fold) in Runx1KO cells compared to control
 174 mK4 cells. B) Heatmap of ChIP-seq reads mapping to Runx1 peaks from Runx1 ChIP or Runx2
 175 ChIP in control versus Runx1KO cells. C) Graph displaying -log10 p-values of motif enrichment,
 176 revealing that Runx1 motifs are the most highly enriched motifs in the Runx1 ChIP peaks. D)
 177 Venn diagram showing strong overlap of Runx1 and Runx2 ChIP peaks.

178

179 To further determine whether these Runx1 bound regions were involved in transcriptional
 180 regulation, we assigned genes to the Runx1 ChIP peaks using the GREAT annotation

181 tool (McLean et al., 2010) and compared the genes near immunoprecipitated chromatin
182 to the genes that exhibiting expression changes of over 2-fold in Runx1KO cells
183 (Supplemental Table 1). This analysis showed that 45% (526/1182) of downregulated
184 genes in Runx1KO cells had a Runx1 ChIP-seq peak in their vicinity, a 1.85 fold
185 enrichment over what was expected by chance (hypergeometric p-value $4.60e \times 10^{-55}$).
186 In contrast, only 30% (511/1705) of upregulated genes had a nearby immunoprecipitated
187 peak (a 1.24 fold enrichment). These results are consistent with the loss of expression in
188 Runx1KO cells of genes predicted to be activated by Runx1 but are less consistent with
189 a model in which upregulated genes were repressed directly by Runx1.

190

191 The failure of Runx2 to regulate the same genes as Runx1, as reflected in the RNA-seq
192 data (Figure 2A), might reflect physical occlusion of Runx2 binding by Runx1.
193 Alternatively, certain regions of the genome might only be occupied by Runx1, and not
194 Runx2, due to differences in DNA binding preferences or protein interaction partners.
195 Finally, Runx1 and Runx2 might occupy the same loci, but Runx2 might have different
196 effects on gene expression compared to Runx1. To further investigate these possibilities,
197 we performed Runx2 ChIP-seq in both control and Runx1KO cells. The Runx2 ChIP
198 identified combined sets of 17,595 peaks present in control cells and 15,608 peaks
199 present in Runx1KO cells, with the Runx motif strongly enriched in both cell types (p-
200 value 1×10^{-2244} and 1×10^{-2363} , respectively (Supplemental Table 3)). Comparisons
201 between the chromatin bound by Runx1 and Runx2 revealed remarkable overlap of the
202 peaks in control cells, and retention of Runx2 ChIP signal at Runx1 peaks in Runx1KO
203 cells (Figures 2B, 2D, and Figure 2-Supplemental Figure 1A). Regulatory Element Locus

204 Intersection (RELI) analyses (Harley et al., 2018) confirmed the highly significant
205 agreement between the Runx1 (mk4), Runx2 (mk4), and Runx2 (Runx1 KO) ChIP-seq
206 datasets (control cell 194.46 fold enriched, p-value 2.0×10^{-219} ; Runx1KO cell 177.85 fold
207 enriched, p-value 2.32×10^{-219}) (Figure 2-Supplemental Figure 1B and Supplemental
208 Table 4). These results suggest that the regulatory regions near downregulated genes in
209 Runx1KO cells retain Runx2 binding, which evidently is not sufficient to drive their
210 expression. Further, the ChIP data indicate that Runx1 has transcriptional activator
211 function at a large subset of its target loci. Runx2 can also bind these loci, but this binding
212 is not sufficient to activate the expression of the associated genes

213

214 **Dramatic changes in chromatin accessibility drive expression changes in** 215 **Runx1KO cells**

216 The Runx ChIP data indicated that most regulatory regions remained accessible to Runx2
217 binding in Runx1KO cells (Figure 2B). To examine chromatin accessibility near down
218 regulated genes in Runx1KO cells, we next performed ATAC-seq experiments. As with
219 the RNA-seq and ChIP-seq data, all ATAC-seq replicates were highly reproducible
220 (Figure 3-Supplemental Figure 1A). The majority of open chromatin regions represented
221 by 37,481 ATAC-seq peaks displayed similar levels of reads between control and
222 Runx1KO cells, henceforth called Runx1-independent ATAC-seq peaks (Figure 3A,
223 intersect in Venn diagram, and heatmap in Figure 3B, left panel). Notably, we also
224 observed substantial and reproducible loss in chromatin accessibility after Runx1 was
225 deleted – 8,741 genomic loci had significantly lower accessibility in Runx1KO cells vs
226 control, which we denote as Runx1-dependent peaks (Figure 3A, left unique area in Venn

227 diagram and heatmap in Figure 3B, middle panel). Interestingly, a similar number of
228 regions (9,427) showed increased accessibility in Runx1KO cells vs control (Runx1KO-
229 induced; Figure 3A right unique area in Venn diagram and heatmap in Figure 3B, right
230 panel). We denote these sites as Runx1KO-induced.

231
232 These changes in chromatin accessibility could reflect Runx1 functioning as an activator
233 at some loci (i.e., by opening or maintaining the accessibility of Runx1-dependent sites)
234 and a repressor at other loci (i.e., by keeping Runx1KO-induced sites inaccessible). If
235 Runx1 is directly acting as both an activator and a repressor in this manner, then both
236 classes would be expected to be enriched for Runx1 motifs. To test this hypothesis, we
237 repeated the analyses described above and again found strong enrichment for AP-1
238 motifs (p-value 1×10^{-1601}), implicating Runx1-dependent ATAC-seq regions as likely
239 enhancers (Figure 3C and Supplemental Table 3). The 2nd most enriched motif class in
240 Runx1-dependent enhancers was Runx (p-value 1×10^{-388}) and as expected, the
241 dataset had highly significant overlap with our Runx1 ChIP data (RELI: 42.60 fold
242 enriched, p-value 8.6×10^{-215}) (Supplemental Table 4). Additionally, the Runx1-
243 dependent regions may be functionally important in multiple cellular contexts as the
244 RELI analysis also found significant enrichment for Runx1 ChIP sites in AML (10.73 fold
245 enriched, p-value 1.62×10^{-170}) and HPC-7 cells (9.73 fold enriched, p-value 2.18×10^{-146}) (Supplemental Table 4). Thus, Runx1 is likely playing an active role in keeping
247 these regions accessible in mK4 cells and potentially also in cancer cell types in which
248 Runx1 is known to play a critical

Figure 3:

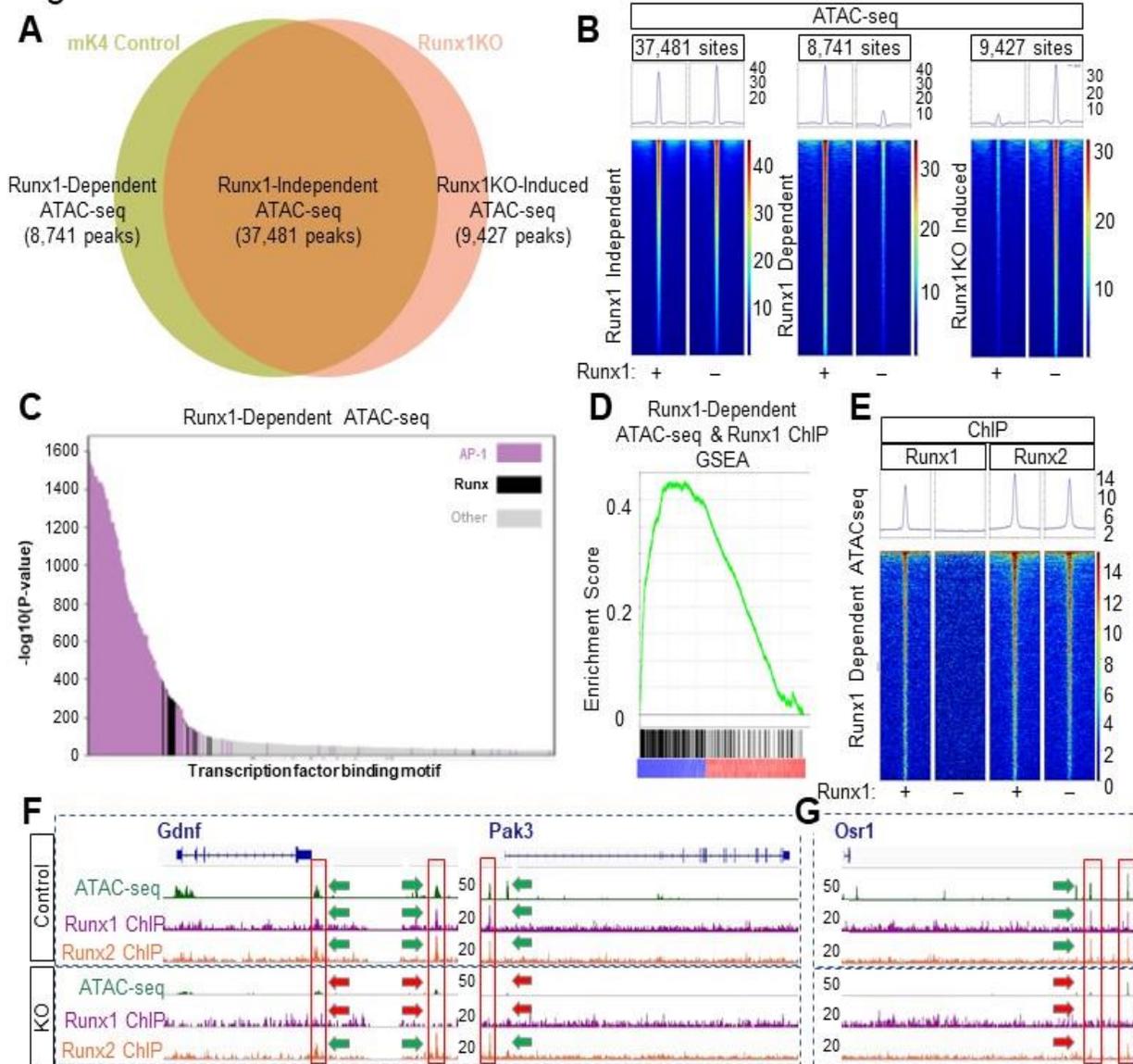


Figure 3: Runx1 Deletion Alters Chromatin Accessibility Despite the Presence of Runx2.

A) Venn diagram of ATAC-seq peaks in control and Runx1KO cells showing the number of regions open in both cell lines (Runx1-independent), regions open only in control cells (Runx1-dependent) and regions open only in Runx1KO cells (Runx1KO-induced). B) Heatmaps of the ATAC-seq reads mapping to Runx1-independent, Runx1-dependent, and Runx1KO-induced peaks in the control and Runx1KO cells. C) Graph of the $-\log_{10}$ p-values of motif enrichment, displaying that Runx1-dependent ATAC-seq peaks are strongly enriched for AP-1 and Runx motifs. D) Gene set enrichment analysis showing enrichment of transcriptionally down-regulated genes by Runx1-dependent ATAC-seq peaks that are bound by Runx1. E) Heatmap showing Runx1-Dependent ATAC-seq regions bound in Runx1 and Runx2 ChIP experiments. F) Genomic snapshots of *Gdnf* and *Pak3* genes that are downregulated in Runx1KO cells showing open chromatin regions present in control cells but not Runx1KO cells that are bound by Runx1 and Runx2 and retain Runx2 binding in the Runx1KO cells. G) Genomic snapshot of the Runx1KO downregulated gene *Osr1* that has genomic regions that lose chromatin accessibility in Runx1KO cells, which are bound by Runx1 and Runx2 in control cells, with reduced Runx2 binding in Runx1KO cells.

250 role. We next assigned Runx1-dependent ATAC-seq regions to nearby Runx1-
251 dependent transcripts using the GREAT annotation tool (McLean et al., 2010). Runx1-
252 dependent ATAC-seq regions were enriched 3.11 fold near down-regulated genes in
253 Runx1KO cells (hypergeometric p-value = 1.5×10^{-197}), indicating that the likely
254 enhancers identified by ChIP and ATAC-seq largely act by maintaining expression of
255 nearby genes.

256

257 Comparison of Runx1 ChIP peaks obtained in control cells to the three classes of ATAC-
258 seq peaks shown in Figures 3A and 3B revealed extensive Runx1 binding within both
259 Runx1-dependent peaks and peaks that remain open in the absence of Runx1 (Figure 3-
260 Supplemental Figure 1B), which do not require Runx1 to remain accessible. Interestingly,
261 regions that become inaccessible in Runx1KO cells (Runx1-dependent) that are bound
262 in the Runx1 ChIP show strong enrichment (4.03 fold, hypergeometric p-value $1.01e \times$
263 10^{-67}) for proximal genes whose expression decreases in Runx1KO cells (Figure 3D).
264 Thus, the combination of ATAC-seq and ChIP data helps to define a set of functional,
265 Runx1-dependent regulatory regions in the mK4 genome and supports the hypothesis
266 that Runx1 pioneer/activator function is maintaining accessible chromatin in mK4 cells.

267

268 The Runx1-dependent regions that close in the absence of Runx1 despite the presence
269 of Runx2 expression might do so because those specific regions are bound only by Runx1
270 and not by Runx2. To test this hypothesis, we compared the Runx2 ChIP-seq reads with
271 the three classes of ATAC-seq peaks shown in Figure 3A. As we observed for Runx1
272 ChIP, the Runx2 ChIP signal was present at both Runx1-independent and Runx1-

273 dependent ATAC-seq regions, and notably, they are clearly present in Runx1KO cells
274 (Figure 3-Supplemental Figure 1B). However, Runx2 ChIP signal was slightly decreased
275 at Runx1-dependent ATAC-seq sites in Runx1KO cells, likely due to decreased chromatin
276 accessibility (Figure 3E). For example, putative enhancers located near the Runx1KO
277 regulated genes *Gdnf* and *Pak3* are shown in Figure 3F (green arrows). Both *Gdnf* and
278 *Pak3* are members of signaling pathways previously reported to be regulated by Runx
279 (Chen et al., 2006; Ernsberger, 2008; Luo et al., 2007; Park et al., 2012; Rouillard et al.,
280 2016). In Runx1KO cells, they became inaccessible (Figure 3F, red arrows) while still
281 retaining Runx2 binding (Figure 3F, green arrow). Further examples are provided as
282 supplemental data to demonstrate that this pattern of lost ATAC-seq signal but retained
283 Runx2 binding is widespread near genes whose expression is reduced in Runx1KO cells
284 (Figure 3-Supplemental Figure 1D). These data suggest that while Runx2 can bind to
285 closed chromatin like Runx1 (Lichtinger et al., 2010), it cannot make the chromatin
286 accessible to other factors.

287

288 Other Runx1-dependent regions display greatly reduced Runx2 binding. For example,
289 two potential enhancers near *Osr1*, a reported Runx target gene (Stock et al., 2004), are
290 open and bound by both Runx1 and Runx2 in control cells (Figure 3G, green arrows), but
291 become inaccessible with limited binding of Runx2 in the Runx1KO cells (Figure 3G, red
292 arrows). We tested the hypothesis that sites that lose Runx2 binding in Runx1KO cells
293 may be enriched for downregulated genes by separating the Runx1-dependent ATAC-
294 seq regions bound by Runx1 into two groups: those sites that had an overlapping Runx2
295 ChIP peak in Runx1KO cells and those sites that did not (Figure 3-Supplemental Figure

296 1E). Enrichment analysis on these two classes, performed as above, revealed similar
297 strong enrichment for downregulated genes in the sites immunoprecipitated by Runx2 in
298 Runx1KO cells (4.12 fold enrichment, hypergeometric p-value $1.95e \times 10^{-61}$) as well as
299 the sites that are not immunoprecipitated by Runx2 in Runx1KO cells (3.88 fold
300 enrichment, hypergeometric p-value 1.51×10^{-13}). These results indicate that loss of both
301 Runx1 and Runx2 binding does not compromise expression more than loss of Runx1
302 alone. Collectively, the Runx1 and Runx2 ChIP data indicate that Runx1 binds to and
303 actively opens or maintains chromatin accessibility at a large number of loci, many of
304 which are associated with Runx1-responsive genes. Despite remaining bound to these
305 same regions in the absence of Runx1, Runx2 binding is unable to compensate for the
306 lack of Runx1.

307

308 **Runx1KO-induced chromatin regions are opened due to the loss of Zeb** 309 **transcriptional repressors**

310 Our analyses above suggest that Runx1KO-induced ATAC-seq regions require Runx1 to
311 remain inaccessible, but the lack of Runx1 binding to these regions is inconsistent with
312 direct repression by Runx1. To explore the possibility that particular Runx1-dependent
313 protein(s) are maintaining repression, we performed TF binding motif enrichment analysis
314 at these sites. Indeed, Runx motifs were absent from the top 1,500 enriched motifs
315 (Supplemental Table 3), consistent with an indirect mechanism whereby Runx1 acts
316 either by repressing a pioneer activator or by activating a repressor protein. Motif
317 enrichment analysis of Runx1KO-induced ATAC-seq peaks compared to Runx1-
318 dependent ATAC-seq sites revealed significant enrichment for the Zeb repressor motif

319 (Figure 4A; Figure 4-Supplemental Figure 1A). This is consistent with an indirect
320 mechanism in which Runx1 regulates Zeb expression, which in turn actively maintains
321 inaccessible chromatin architecture at multiple sites. Accordingly, we observed that the
322 levels of Zeb1 and Zeb2 mRNA were over 10 fold lower in Runx1KO cells (Figure 4B)
323 and confirmed this independently by real-time quantitative PCR (RT-qPCR) (Figure 4-
324 Supplemental Figure 1D). Using Western blot analysis, we found that the Zeb1 protein is
325 expressed in control cells but not in Runx1KO cells (Figure 4C). Unfortunately,
326 commercially available antibodies to Zeb2 failed to detect it in control cells. Next, we
327 asked if Runx1 was a direct regulator of Zeb1 expression by examining the Zeb1 locus
328 for Runx1 binding. An accessible regulatory region near Zeb1 was bound by Runx1 and
329 Runx2 in control cells (Figure 4D green arrows), but became inaccessible in Runx1KO
330 cells (Figure 4D, red arrows). Similarly, an accessible regulatory region near Zeb2, bound
331 by Runx1 and Runx2 in control cells, became inaccessible in Runx1KO cells (Figure 4E).
332 Additionally, the Zeb1 and Zeb2 TSS, open in control cells, were less accessible in
333 Runx1KO cells (Figure 4-Supplemental Figure 1B), mirroring the dramatic differences in
334 the expression levels of Zeb1 and Zeb2 between control and Runx1KO cells. These data
335 suggest that both Zeb1 and Zeb2 are regulated by Runx1-dependent enhancers in mK4
336 cells.

337
338 In support of the hypothesis that the opening of the chromatin in Runx1KO cells is due to
339 the loss of Zeb repressors, we examined a known Zeb1 repressed target, *Pard6b*, that
340 was upregulated in Runx1KO cells nearly 10-fold based on RNA-seq analysis (Figure 4-
341 Supplemental Data 1C) and RT-qPCR (Figure 4-Supplemental Figure 1D). A predicted
342 Zeb binding site upstream of *Pard6b* is only accessible in Runx1KO cells (Figure 4F).

Figure 4:

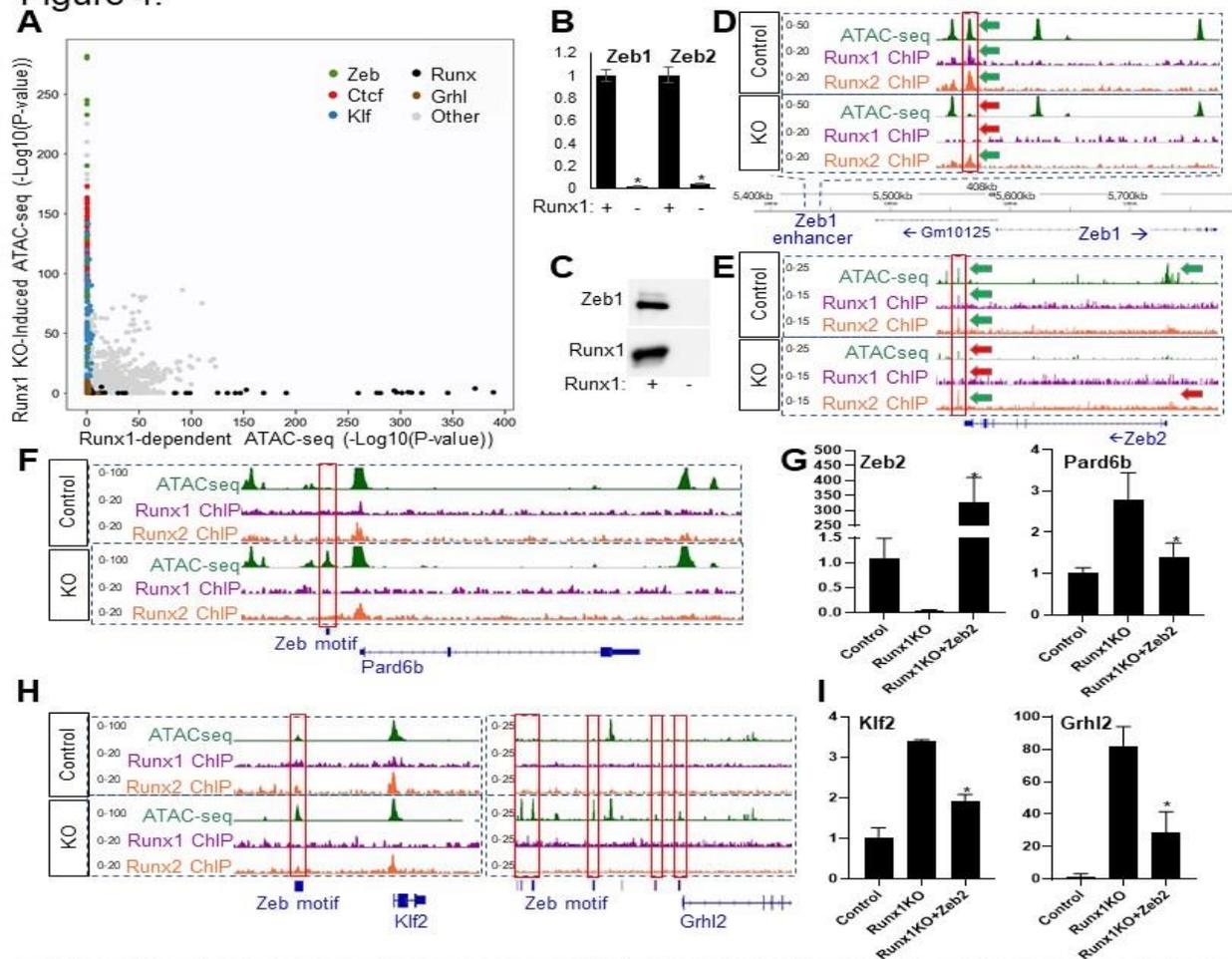


Figure 4: Runx1KO Cells Lack Zeb Repressors, Leading to the Opening of Chromatin. A) Graph displaying p-values of transcription factor motif enrichment in Runx1-dependent versus Runx1-induced ATAC-seq, revealing that Zeb motifs are specifically enriched in Runx1KO-Induced ATAC-seq peaks. Additionally, Ctfc, Klf, and Grhl motifs are enriched in the Runx1KO-induced ATAC-seq peaks, while Runx motifs are enriched in the Runx1-dependent ATAC-seq. Note that this graph has had AP-1 motif enrichment results removed in order to focus on other motif enrichment levels (see Figure 4-Supplemental Figure 1A for all transcription factor motifs). B) RT-qPCR showing that Runx1KO cells lose expression of Zeb1 and Zeb2. C) Western blot showing the absence of the Zeb1 protein in Runx1KO cells. D) Genomic snapshot showing a chromatin region near Zeb1 that is bound by Runx1 and Runx2 and loses chromatin accessibility in Runx1KO cells. E) Genomic snapshot of the Zeb2 locus showing a downstream potential enhancer bound by Runx1 and Runx2 that has decreased chromatin accessibility along with a loss of expression in Runx1KO cells. F) Genomic snapshot of the Zeb target gene Pard6b locus showing a promoter region containing a predicted Zeb binding site that is specifically open in Runx1KO cells. G) RT-qPCR showing that Zeb2 transient transfection of Runx1KO cells induces repression of Pard6b expression down to levels similar to those in control cells after 1 day of selection for transfected cells followed by 2 days of growth in media. H) Genomic snapshots of the Klf2 and Grhl2 loci displaying ATAC-seq regions that are specifically open in Runx1KO cells (outlined in red) that contain predicted Zeb binding sites. I) RT-qPCR confirmation of the upregulation of Klf2 and Grhl2 in Runx1KO cells, which is suppressed by transient transfection of Zeb2, as shown for RT-qPCR of Pard6b (panel G). The * denotes p < 0.05 in Student's t-test.

344 Thus, we hypothesized that a subset of the genes upregulated in the Runx1KO cells may
345 be the result of losing Runx1-dependent Zeb 1 and Zeb2 expression, and the subsequent
346 derepression of Zeb targets.

347
348 To test whether loss of Zeb expression is responsible for the widespread changes in
349 chromatin accessibility and subsequent gains in gene expression seen in Runx1KO cells,
350 we transiently transfected Runx1KO cells with an expression vector driving Zeb2 mRNA
351 levels 330 fold over baseline as determined by RT-qPCR (Figure 4G, left - recall that
352 there is no usable antibody to Zeb2). Next, we tested the expression of *Pard6b* in Zeb2-
353 transfected Runx1KO cells and found that its expression level was significantly reduced
354 (Figure 4G, right). This supports the hypothesis that many of the upregulated genes in
355 Runx1KO cells may be indirectly affected through the loss of Zeb-mediated repression.

356
357 Not all of the open chromatin gained in Runx1KO cells contains Zeb sites. Both Klf and
358 Grhl motifs are significantly enriched in the small subset of Runx1KO-induced ATAC-seq
359 fragments that lack Zeb motifs (Figure 4A and Supplemental Table 3). Genomic
360 snapshots of the *Klf2* and *Grhl2* loci reveal Zeb motif-containing chromatin regions that
361 become accessible in Runx1KO cells (Figure 4H), concomitant with increased *Klf2* and
362 *Grhl2* expression, suggesting that Klf2 and Grhl2 are suppressed by Zeb in mK4 cells. As
363 we had observed for *Pard6b*, transfection of Zeb2 in the Runx1KO cells led to significant
364 downregulation of both *Klf2* and *Grhl2* expression by RT-qPCR (Figure 4I), which
365 supports the interpretation that these genes are normally repressed by Zeb proteins in
366 mK4 cells. The inhibition of Grhl2 expression by Zeb proteins has been shown to play a
367 critical role in reciprocal negative feedback loops between these pathways during the

368 epithelial-mesenchymal transition (Cieply et al., 2013). Combined, these data suggest
369 that upregulation of Klf2 and Grhl2 have functional consequences via opening of
370 chromatin at their bound targets, and support the hypothesis that the upregulation of
371 genes in Runx1KO cells is due in large part to the loss of Zeb repressors, which cascade
372 due to derepression of additional transcriptional activators. Thus, these results reveal how
373 the loss of a single transcription factor can create ripple effects perturbing the entire
374 transcriptional network.

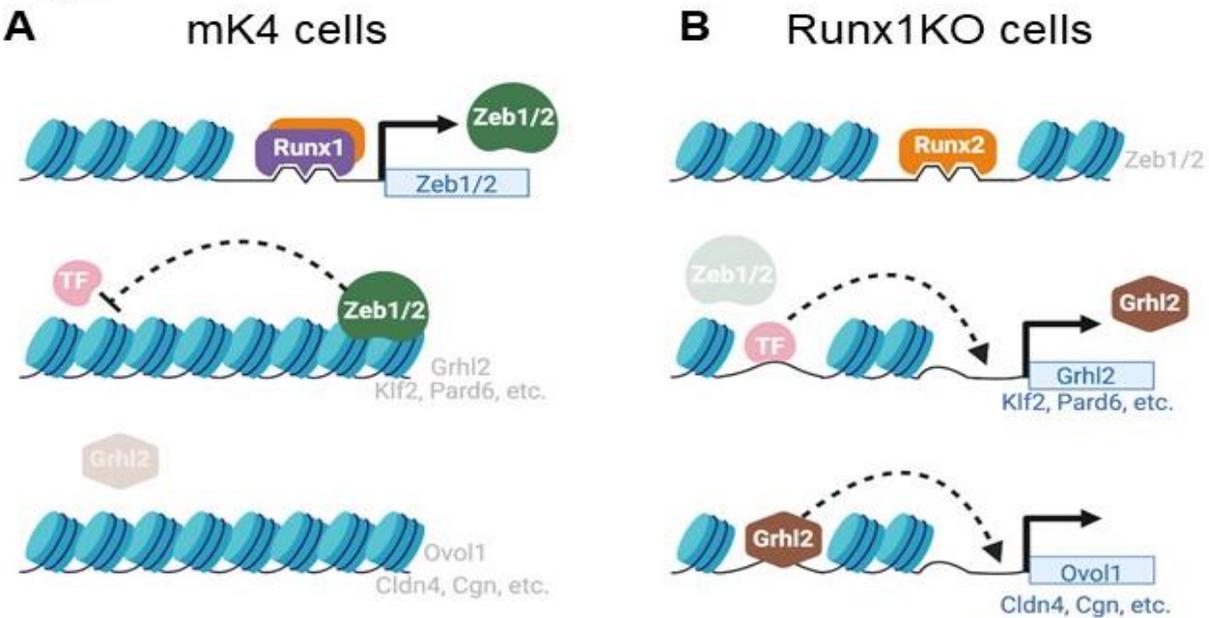
375

376 **Discussion**

377 We report herein that Runx1 regulates the chromatin landscape at multiple loci to broadly
378 impact transcription in a mouse kidney cell line. Even though Runx1 has been reported
379 to be both a transcriptional activator and a repressor (Brettingham-Moore et al., 2015;
380 Mevel et al., 2019; Seo et al., 2012), our analyses suggest that Runx1 functions primarily
381 as an activator in mK4 cells despite the expression of Gro/TLE1-3 (Supplemental Table
382 1). Runx1 does so by maintaining chromatin accessibility at many loci, including near loci
383 encoding the repressors Zeb1 and Zeb2. The Zeb proteins in turn act to maintain
384 inaccessible chromatin at many loci, including several other transcription factors (e.g.,
385 Klf2 and Grhl2), resulting in the cascading repression of further downstream indirect
386 targets (e.g., Ovol1). Runx1 deficiency leads to loss of accessibility at regulatory regions
387 of downregulated direct target genes and subsequently to gain in accessibility and
388 upregulation of actively repressed genes, including additional transcriptional activators
389 such as Grhl2 and Klf2 (Figure 5), leading ultimately to widespread genome-wide
390 changes to chromatin accessibility and transcription.

391

Figure 5:



392

393 **Figure 5: Model of Transcription Factor Network Perturbation by Runx1-Deficiency.** A) In
394 control mK4 cells, Runx1 induces the transcriptional repressors Zeb1 and Zeb2 that inhibit other
395 transcriptional activators such as Grhl2, resulting in inhibition of downstream Grhl2 target genes.
396 B) Runx1KO cells lose expression of Zeb1 and Zeb2, which derepresses their targets including
397 Grhl2 and Klf2, which in turn leads to upregulation of their downstream targets such as Ovo1,
398 Cldn4, and Cgn. Figure was created using BioRender.com.

399

400 Interestingly, these broad effects on chromatin accessibility and transcription occur
401 despite the presence of Runx2 in these cells. Runx2 remains bound to most of the sites
402 immunoprecipitated by Runx1 but is unable to compensate for Runx1 loss. Despite the
403 conserved ability of both proteins to bind inaccessible chromatin, our results are
404 consistent with Runx1, but not Runx2, acting as a pioneer factor (Lichtinger et al., 2010).
405 The molecular mechanism underlying the specific ability of Runx1 to maintain chromatin
406 accessibility near its target genes remains under investigation. In addition to facilitating
407 transcription by enhancing accessibility, TFs can regulate transcription by recruiting

408 additional factors or by contributing activity to preassembled complexes. In mK4 cells,
409 Runx1 appears to largely act by making chromatin accessible to other TFs and by
410 inducing the expression of repressors to prevent a myriad of other sites from responding
411 to their regulators. The enrichment for down-regulated genes near Runx1-dependent
412 ATAC-seq is greater than that observed for Runx1 ChIP, which suggests that impact on
413 accessibility spreads across chromatin regions beyond the sites directly bound by Runx1.
414 This raises the possibility that Runx1 may play an additional role in facilitating the loading
415 of other TFs through the opening or maintaining of enhancer accessibility, to allow full
416 transcriptional activation of its target genes, as has been described for other transcription
417 factors (van Bakel, 2011). The integration of signaling pathways through control of
418 chromatin accessibility in co-occupied regions, even in the absence of direct interactions,
419 has been shown to be an important regulatory mechanism for other TFs, such as SOX2
420 and OCT4 (Friman et al., 2019). It will be interesting to determine whether Runx1 plays a
421 generalized master-regulator role controlling the activity of other signaling pathways
422 through the modulation of chromatin accessibility.

423

424 The finding that gains in chromatin accessibility in Runx1KO cells largely reflects a loss
425 of Zeb repressor protein activity suggests that Runx1 “repressor” functions may be
426 executed by Zeb1 in many cells and tissues, where Zeb expression requires Runx1. This
427 may reflect an underappreciated and widespread collaboration between these proteins.
428 Accordingly, examination of public functional genomics data revealed that the Runx1-
429 bound Zeb1 enhancer identified in mK4 cells is accessible (DNase hypersensitive) in
430 other tissues including several hemopoietic cell lines (Supplemental Table 4) and is also

431 bound by Runx1 in AML. Further, gaining Zeb expression could be relevant not only to
432 the role of Runx1 in AML but might also contribute to solid organ malignancies, where
433 Zeb proteins are critical inducers of the epithelial-mesenchymal transition (EMT). In
434 agreement with this hypothesis, co-expression of both Runx1 and Zeb2 in circulating
435 tumor cells has been shown to significantly correlate with cancer reoccurrence (Alonso-
436 Alconada et al., 2014). Additionally, Runx1 has been shown to be critical for TGFbeta
437 induced EMT during renal fibrosis both in HK-2 cells and *in vivo* (Zhou et al., 2018). The
438 interplay between the Runx and Zeb calls for a reexamination of whether TGFbeta
439 induction of EMT through increased expression of Zeb proteins (Xu et al., 2009) is
440 mediated through Runx1. Interestingly, the cell-specific dependence of Zeb expression
441 on Runx1 may explain why Runx1 has been suggested to either promote or suppress
442 EMT in different cell lines (Zhou et al., 2018). For example, Runx1 has been shown to
443 inhibit expression of Zeb1 in breast cancer cells, thereby suppressing EMT (Hong et al.,
444 2018; Hong et al., 2017; Li et al., 2019). This raises the possibility that a switch between
445 positive and negative regulation of Zeb proteins by Runx1 could underlie the contrasting
446 reports on its role in promoting or inhibiting EMT. This indirect mechanism by which Runx1
447 mediates repression through the upregulation of other proteins may be widespread and
448 include additional repressors and activators that could be uncovered by the enrichment
449 for their motifs revealed by experiments such as ATAC-seq in different cellular contexts.
450 Collectively, our data indicate that loss of Runx1 produces widespread genomic and
451 transcriptional changes through a cascade of direct and indirect sequelae involving
452 multiple transcriptional repressors and activators, and reveal key members of this
453 complex network of interacting TFs.

454

455 **Materials and Methods**

456 **Tissue culture:** The mK4 and Runx1KO cells were grown in DMEM supplemented with
457 10% FBS, L-glutamine, penicillin/streptomycin, and sodium pyruvate. The cells were
458 transfected using Lipofectamine 2000 (Invitrogen) following the manufacturer's directions.

459

460 **Generation of Runx1KO cells:** We used the mK4 cell line as our control cell line, as
461 described in (Valerius et al., 2002). From the control cell line we generated a sub-line that
462 does not express Runx1 through the use of guide RNAs (gRNAs) in the px458 and px459
463 CRISPR/Cas9 vectors to delete the third exon of Runx1 that contains the start codon
464 utilized in mK4 cells. The targeting sequences were generated with the method and tools
465 described in (Haeussler et al., 2016). These cells were then transfected with
466 Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. The cells
467 underwent selection with puromycin for two days. Subsequent clones were picked
468 approximately a week later using cloning disks and the clones were screened for exon 3
469 deletion by PCR and for loss of Runx1 protein expression by Western blot.

470

471 **Western blot:** Confluent control cells or Runx1KO mK4 cells were collected in 100 ul of
472 RIPA-DOC with protease inhibitors plus 100 ul of 2X sample buffer. Protein samples were
473 run on 7% polyacrylamide gels and then transferred to PVDF. Indicated antibodies were
474 applied at 1:1000 dilutions overnight at 4 degrees and then secondary antibodies were
475 used at 1:5000 at room temperature for 1hr. The Western blot signal was detected using

476 Thermofisher Supersignal Femto ECL reagent using a Bio Rad Chemidoc MP Imaging
477 System.

478

479 **RNA-seq:** mK4 control and Runx1KO cells were cultured in triplicate in standard mK4
480 media (DMEM plus 10% FBS, 2% L-glutamine, 1% Pen/Strep, and 1% Sodium Pyruvate)
481 on 12 well plates until nearly confluent. Cells were removed from the plate with trypsin
482 that was subsequently inactivated using mK4 conditioned media to prevent a feeding
483 effect from fresh media activating signaling pathways. RNA was collected using
484 Invitrogen's Purelink RNA Mini kit according to the manufacturer's directions. RNA-seq
485 on polyA isolated RNA was performed by the CCHMC sequencing core to produce over
486 20 million reads per sample.

487

488 **RT-qPCR:** Biological triplicate samples of RNA were converted to cDNA using
489 Superscript II Reverse Transcriptase from Invitrogen following the company's protocol.
490 The cDNA was diluted to 40 ng/ μ l, and 5 μ l of each sample was added to each RT-qPCR
491 reaction that were amplified using iTaq Universal SYBR Green Supermix from Bio-Rad
492 and read on a StepOnePlus Real-Time PCR System from Applied Biosystems. Gene
493 expression levels were normalized to Gapdh and changes were determined relative to
494 control cells, with significance calculated using Student's t-test.

495

496 **ATAC-seq:** ATAC-seq experiments were carried out in the control mK4 cells and
497 Runx1KO cells in triplicate. Experiments were performed following the protocol laid out
498 by the Kaestner Lab (Ackermann et al., 2016). The Tn5 used in the experiment was

499 prepared using the method outlined in (Buenrostro et al., 2013). The purification of the
500 library prep was done in accordance with (Corces et al., 2017).

501 **ChIP-seq:** Control and Runx1KO mK4 cells were grown on 10cm plates in triplicate until
502 nearly confluent and removed from the plate using trypsin that was inactivated with
503 conditioned media. Individual cells were counted and 10^6 cells were used to make the
504 ChIP lysates. Cells were incubated in crosslinking solution (1% formaldehyde, 5 mM
505 HEPES [pH 8.0], 10 mM sodium chloride, 0.1 mM EDTA, and 0.05 mM EGTA in RPMI
506 culture medium with 10% FBS) and placed on a tube rotator at room temperature for 10
507 min. To stop the crosslinking, glycine was added to a final concentration of 0.125 M and
508 tubes were placed back on the rotator at room temperature for 5 min. Cells were washed
509 twice with ice-cold PBS, resuspended in lysis buffer 1 (50 mM HEPES [pH 8.0], 140 mM
510 NaCl, 1 mM EDTA, 10% glycerol, 0.25% Triton X-100, and 0.5% NP-40), and placed on
511 a tube rotator at 4C for 10 minutes. Nuclei were harvested after centrifugation at 10,000g
512 for 5 min, resuspended in lysis buffer 2 (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 200 mM
513 NaCl, and 0.5 mM EGTA), and placed on a tube rotator at room temperature for 10 min.
514 Nuclei were collected again by centrifugation at 10,000g for 5 minutes. Protease and
515 phosphatase inhibitors were added to both lysis buffers. Nuclei were then resuspended
516 in the sonication buffer (10 mM Tris [pH 8.0], 1 mM EDTA, and 0.1% SDS). A S220
517 focused ultrasonicator (COVARIS) was used to shear chromatin (150- to 500-bp
518 fragments) with 10% duty cycle, 175 peak power, and 200 bursts per cycle for 7 min. A
519 portion of the sonicated chromatin was run on an agarose gel to verify fragment sizes.
520 Sheared chromatin was precleared with 20 μ l Dynabeads Protein A (Life Technologies)
521 at 4 °C for 1 hr.

522 Immunoprecipitation of Runx-chromatin complexes was performed with an SX-8X IP-
523 STAR compact automated system (Diagenode). Beads conjugated to antibodies against
524 Runx1 (Rabbit mAb #8529, Cell Signaling) or Runx2 (Rabbit mAb #8486, Cell Signaling)
525 were incubated with precleared chromatin at 4°C for 8 hours. The beads were then
526 washed sequentially with wash buffer 1 (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM
527 EDTA, 0.1% SDS, 0.1% NaDOC, and 1% Triton X-100), wash buffer 2 (50 mM Tris-HCl
528 [pH 7.5], 400 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.1% NaDOC, and 1% Triton X-100),
529 wash buffer 3 (2 mM EDTA, 50 mM Tris-HCl [pH 7.5] and 0.2% Sarkosyl Sodium Salt),
530 and wash buffer 4 (10 mM Tris-HCl [pH 7.5], 1 mM EDTA, and 0.2% Triton X-100). Finally,
531 the beads were resuspended in 10 mM Tris-HCl (pH 7.5) and used to prepare libraries
532 via ChIPmentation (Schmidl et al., 2015).

533

534 **Processing of functional genomics data:** RNA-seq, ATAC-seq, and ChIP-seq reads
535 (in FASTQ format) were first subjected to quality control using FastQC (v0.11.7)
536 (parameter settings: --extract -o output_fastqc -f R1.fastq.gz) (Kalita et al., 2018). Adapter
537 sequences were removed using Trim Galore (v0.4.2) (parameter settings: -o folder --
538 path_to_cutadapt cutadapt --paired R1.fastq.gz R2.fastq.gz) (Goodwin et al., 2016), a
539 wrapper script that runs cutadapt (v1.9.1) (Bentley et al., 2008) to remove adapter
540 sequences from the reads. The quality-controlled reads were aligned to the reference
541 mouse genome version NCBI37/mm9 using STAR v2.6.1e (Dobin et al., 2013). Duplicate
542 reads were removed using the program sambamba v0.6.8 (parameter settings: -q
543 markdup -r -t 8 trimmed.bam trimmed_dedup.bam) (Tarasov et al., 2015). Gene

544 annotations for RNA-seq analysis were downloaded from the UCSC Table Browser
545 (Karolchik et al., 2004) for the NCBI37/mm9 genome in GTF format.

546

547 ATAC-seq and ChIP-seq data were processed using the following steps. Peaks were
548 called using MACS2 v2.1.2 (parameter settings: callpeak -g mm -q 0.01 --broad -t
549 trimmed_dedup.bam -f BAM -n trimmed_dedup_peaks) (Zhang et al., 2008b). Specific
550 ChIP peaks were identified by MACS2 peak calling on the combined replicate reads and
551 removing peaks that overlapped with non-specific background peaks called in the Runx1
552 ChIP in the Runx1KO cells. Peaks shared across experiments (i.e., peaks shared
553 between replicates or shared between treatments/conditions) were identified as peaks
554 with 50% or greater overlap, using BEDtools v2.27.0 (Quinlan and Hall, 2010). The final
555 peak sets for each condition were obtained by requiring peaks to be present in at least
556 two out of the three biological replicates. When comparing across treatments or
557 conditions, peak overlap between any of the three replicates in either treatment/conditions
558 was considered a shared peak between the treatments/conditions. Final peaks, originally
559 in BED format, were converted to Gene Transfer Format (GTF) format to enable fast
560 counting of reads under the peaks using the program featureCounts v1.6.2 (Rsubread
561 package) (parameter settings: featureCounts --ignoreDup -M -t peak -s 0 -O -T 4 -a
562 common_peaks.gtf -o output_counts.txt trimmed_dedup.bam). The resulting matrix of
563 raw counts was normalized for all experiment types to transcripts per million values
564 (TPMs). TF binding site motif enrichment analysis was performed using the HOMER
565 software package (Heinz et al., 2010), which was modified to use a log base 2 scoring

566 system and the set of mouse motifs contained in build 2.0 of the Cis-BP database
567 (Lambert et al., 2019).

568

569 **Acknowledgements:** Dr. Eric Brunskill, Hope Rowden, Carmy Forney, Kevin Ernst, and
570 Dr. Xiaoting Chen for critical discussions, technical assistance, organization, and/or
571 computational help in the production of this manuscript. This work was made possible
572 through funding from the following sources: R01 NS099068 to R.K. and M.T.W. and
573 Cincinnati Children's Hospital "Trustee Award", "Center for Pediatric Genomics Award"
574 and "CCRF Endowed Scholar Award" to M.T.W.

575

576 **Competing interests:** The authors' have no competing interests to declare.

577

578 **Resource availability:** All the sequencing data has been deposited at the Gene
579 Expression Omnibus (GEO) and can be located under accession number GSE158093.
580 Additionally, a genome browser session with the data loaded will be made publicly available upon
581 publication of this manuscript. All the cell lines and plasmids used in this manuscript will be
582 distributed upon request.

583

584 **References**

585

586 Ackermann, A.M., Wang, Z., Schug, J., Naji, A., and Kaestner, K.H. (2016). Integration of
587 ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol*
588 *Metab* 5, 233-244.

589 Alonso-Alconada, L., Muinelo-Romay, L., Madisoo, K., Diaz-Lopez, A., Krakstad, C.,
590 Trovik, J., Wik, E., Hapangama, D., Coenegrachts, L., Cano, A., *et al.* (2014). Molecular
591 profiling of circulating tumor cells links plasticity to the metastatic process in endometrial
592 cancer. *Mol Cancer* 13, 223.

- 593 Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen,
594 Y., Zhao, X., Schmidl, C., Suzuki, T., *et al.* (2014). An atlas of active enhancers across
595 human cell types and tissues. *Nature* 507, 455-461.
- 596 Awad, S., Kankainen, M., Dufva, O., Heckman, C.A., Porkka, K., and Mustjoki, S. (2018).
597 RUNX1 Mutations Identify an Entity of Blast Phase Chronic Myeloid Leukemia (BP-CML)
598 Patients with Distinct Phenotype, Transcriptional Profile and Drug Vulnerabilities. *Blood*
599 132, 4257-4257.
- 600 Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G.,
601 Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., *et al.* (2008). Accurate whole human
602 genome sequencing using reversible terminator chemistry. *Nature* 456, 53-59.
- 603 Brettingham-Moore, K.H., Taberlay, P.C., and Holloway, A.F. (2015). Interplay between
604 Transcription Factors and the Epigenome: Insight from the Role of RUNX1 in Leukemia.
605 *Front Immunol* 6, 499.
- 606 Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013).
607 Transposition of native chromatin for fast and sensitive epigenomic profiling of open
608 chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10, 1213-1218.
- 609 Chen, C.L., Broom, D.C., Liu, Y., de Nooij, J.C., Li, Z., Cen, C., Samad, O.A., Jessell,
610 T.M., Woolf, C.J., and Ma, Q. (2006). Runx1 determines nociceptive sensory neuron
611 phenotype and is required for thermal and neuropathic pain. *Neuron* 49, 365-377.
- 612 Cieply, B., Farris, J., Denvir, J., Ford, H.L., and Frisch, S.M. (2013). Epithelial-
613 mesenchymal transition and tumor suppression are controlled by a reciprocal feedback
614 loop between ZEB1 and Grainyhead-like-2. *Cancer Res* 73, 6299-6309.
- 615 Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A.,
616 Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., *et al.* (2017). An improved
617 ATAC-seq protocol reduces background and enables interrogation of frozen tissues.
618 *Nature methods* 14, 959-962.
- 619 de Bruijn, M., and Dzierzak, E. (2017). Runx transcription factors in the development and
620 function of the definitive hematopoietic system. *Blood* 129, 2061-2069.
- 621 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,
622 Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner.
623 *Bioinformatics* 29, 15-21.
- 624 Ernsberger, U. (2008). The role of GDNF family ligand signalling in the differentiation of
625 sympathetic and dorsal root ganglion neurons. *Cell Tissue Res* 333, 353-371.

- 626 Friman, E.T., Deluz, C., Meireles-Filho, A.C., Govindan, S., Gardeux, V., Deplancke, B.,
627 and Suter, D.M. (2019). Dynamic regulation of chromatin accessibility by pluripotency
628 transcription factors across the cell cycle. *eLife* 8.
- 629 Giambra, V., Jenkins, C.R., Wang, H., Lam, S.H., Shevchuk, O.O., Nemirovsky, O., Wai,
630 C., Gusscott, S., Chiang, M.Y., Aster, J.C., *et al.* (2012). NOTCH1 promotes T cell
631 leukemia-initiating activity by RUNX-mediated regulation of PKC-theta and reactive
632 oxygen species. *Nat Med* 18, 1693-1698.
- 633 Glotzer, D.J., Zelzer, E., and Olsen, B.R. (2008). Impaired skin and hair follicle
634 development in Runx2 deficient mice. *Dev Biol* 315, 459-473.
- 635 Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: ten years
636 of next-generation sequencing technologies. *Nat Rev Genet* 17, 333-351.
- 637 Haeussler, M., Schonig, K., Eckert, H., Eschstruth, A., Mianne, J., Renaud, J.B.,
638 Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., *et al.* (2016). Evaluation
639 of off-target and on-target scoring algorithms and integration into the guide RNA selection
640 tool CRISPOR. *Genome Biol* 17, 148.
- 641 Harley, J.B., Chen, X., Pujato, M., Miller, D., Maddox, A., Forney, C., Magnusen, A.F.,
642 Lynch, A., Chetal, K., Yukawa, M., *et al.* (2018). Transcription factors operate across
643 disease loci, with EBNA2 implicated in autoimmunity. *Nature genetics* 50, 699-707.
- 644 Hass, M.R., Liow, H.H., Chen, X., Sharma, A., Inoue, Y.U., Inoue, T., Reeb, A., Martens,
645 A., Fulbright, M., Raju, S., *et al.* (2015). SpDamID: Marking DNA Bound by Protein
646 Complexes Identifies Notch-Dimer Responsive Enhancers. *Mol Cell* 59, 685-697.
- 647 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre,
648 C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining
649 transcription factors prime cis-regulatory elements required for macrophage and B cell
650 identities. *Mol Cell* 38, 576-589.
- 651 Hoi, C.S., Lee, S.E., Lu, S.Y., McDermitt, D.J., Osorio, K.M., Piskun, C.M., Peters, R.M.,
652 Paus, R., and Tumber, T. (2010). Runx1 directly promotes proliferation of hair follicle stem
653 cells and epithelial tumor formation in mouse skin. *Mol Cell Biol* 30, 2518-2536.
- 654 Hong, D., Fritz, A.J., Finstad, K.H., Fitzgerald, M.P., Weinheimer, A., Viens, A.L.,
655 Ramsey, J., Stein, J.L., Lian, J.B., and Stein, G.S. (2018). Suppression of Breast Cancer
656 Stem Cells and Tumor Growth by the RUNX1 Transcription Factor. *Mol Cancer Res* 16,
657 1952-1964.

- 658 Hong, D., Messier, T.L., Tye, C.E., Dobson, J.R., Fritz, A.J., Sikora, K.R., Browne, G.,
659 Stein, J.L., Lian, J.B., and Stein, G.S. (2017). Runx1 stabilizes the mammary epithelial
660 cell phenotype and prevents epithelial to mesenchymal transition. *Oncotarget* **8**, 17610-
661 17627.
- 662 Ito, Y., Bae, S.C., and Chuang, L.S. (2015). The RUNX family: developmental regulators
663 in cancer. *Nat Rev Cancer* **15**, 81-95.
- 664 Kalita, C.A., Moyerbrailean, G.A., Brown, C., Wen, X., Luca, F., and Pique-Regi, R.
665 (2018). QuASAR-MPRA: accurate allele-specific analysis for massively parallel reporter
666 assays. *Bioinformatics* **34**, 787-794.
- 667 Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and
668 Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* **32**,
669 D493-496.
- 670 Komori, T. (2018). Runx2, an inducer of osteoblast and chondrocyte differentiation.
671 *Histochem Cell Biol* **149**, 313-323.
- 672 Kueh, H.Y., Yui, M.A., Ng, K.K., Pease, S.S., Zhang, J.A., Damle, S.S., Freedman, G.,
673 Siu, S., Bernstein, I.D., Elowitz, M.B., *et al.* (2016). Asynchronous combinatorial action of
674 four regulatory factors activates Bcl11b for T cell commitment. *Nat Immunol* **17**, 956-965.
- 675 Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale,
676 J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. *Cell* **175**,
677 598-599.
- 678 Lambert, S.A., Yang, A.W.H., Sasse, A., Cowley, G., Albu, M., Caddick, M.X., Morris,
679 Q.D., Weirauch, M.T., and Hughes, T.R. (2019). Similarity regression predicts evolution
680 of transcription factor sequence specificity. *Nature genetics* **51**, 981-989.
- 681 Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in
682 disease. *Cell* **152**, 1237-1251.
- 683 Li, Q., Lai, Q., He, C., Fang, Y., Yan, Q., Zhang, Y., Wang, X., Gu, C., Wang, Y., Ye, L.,
684 *et al.* (2019). RUNX1 promotes tumour metastasis by activating the Wnt/beta-catenin
685 signalling pathway and EMT in colorectal cancer. *J Exp Clin Cancer Res* **38**, 334.
- 686 Lichtinger, M., Hoogenkamp, M., Krysinska, H., Ingram, R., and Bonifer, C. (2010).
687 Chromatin regulation by RUNX1. *Blood Cells Mol Dis* **44**, 287-290.
- 688 Luo, W., Wickramasinghe, S.R., Savitt, J.M., Griffin, J.W., Dawson, T.M., and Ginty, D.D.
689 (2007). A hierarchical NGF signaling cascade controls Ret-dependent and Ret-

- 690 independent events during development of nonpeptidergic DRG neurons. *Neuron* *54*,
691 739-754.
- 692 McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M.,
693 and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory
694 regions. *Nat Biotechnol* *28*, 495-501.
- 695 Mevel, R., Draper, J.E., Lie, A.L.M., Kouskoff, V., and Lacaud, G. (2019). RUNX
696 transcription factors: orchestrators of development. *Development* (Cambridge, England)
697 *146*.
- 698 Osorio, K.M., Lee, S.E., McDermitt, D.J., Waghmare, S.K., Zhang, Y.V., Woo, H.N., and
699 Tumber, T. (2008). Runx1 modulates developmental, but not injury-driven, hair follicle
700 stem cell activation. *Development* *135*, 1059-1068.
- 701 Park, B.Y., Hong, C.S., Weaver, J.R., Rosocha, E.M., and Saint-Jeannet, J.P. (2012).
702 Xaml1/Runx1 is required for the specification of Rohon-Beard sensory neurons in
703 *Xenopus*. *Developmental biology* *362*, 65-75.
- 704 Pencovich, N., Jaschek, R., Tanay, A., and Groner, Y. (2011). Dynamic combinatorial
705 interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation
706 in cell line models. *Blood* *117*, e1-14.
- 707 Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing
708 genomic features. *Bioinformatics* *26*, 841-842.
- 709 Rennert, J., Coffman, J.A., Mushegian, A.R., and Robertson, A.J. (2003). The evolution
710 of Runx genes I. A comparative study of sequences from phylogenetically diverse model
711 organisms. *BMC Evol Biol* *3*, 4.
- 712 Rouillard, A.D., Gundersen, G.W., Fernandez, N.F., Wang, Z., Monteiro, C.D.,
713 McDermott, M.G., and Ma'ayan, A. (2016). The harmonizome: a collection of processed
714 datasets gathered to serve and mine knowledge about genes and proteins. *Database*
715 (Oxford) *2016*.
- 716 Schmidl, C., Rendeiro, A.F., Sheffield, N.C., and Bock, C. (2015). ChIPmentation: fast,
717 robust, low-input ChIP-seq for histones and transcription factors. *Nature methods* *12*,
718 963-965.
- 719 Seo, W., Tanaka, H., Miyamoto, C., Levanon, D., Groner, Y., and Taniuchi, I. (2012).
720 Roles of VWRPY motif-mediated gene repression by Runx proteins during T-cell
721 development. *Immunol Cell Biol* *90*, 827-830.

- 722 Seo, W., and Taniuchi, I. (2020). The Roles of RUNX Family Proteins in Development of
723 Immune Cells. *Mol Cells* 43, 107-113.
- 724 Sierra, O.L., Cheng, S.L., Loewy, A.P., Charlton-Kachigian, N., and Towler, D.A. (2004).
725 MINT, the Msx2 interacting nuclear matrix target, enhances Runx2-dependent activation
726 of the osteocalcin fibroblast growth factor response element. *J Biol Chem* 279, 32913-
727 32923.
- 728 Stock, M., Schafer, H., Fliegau, M., and Otto, F. (2004). Identification of novel genes of
729 the bone-specific transcription factor Runx2. *J Bone Miner Res* 19, 959-972.
- 730 Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast
731 processing of NGS alignment formats. *Bioinformatics* 31, 2032-2034.
- 732 Terriente-Felix, A., Li, J., Collins, S., Mulligan, A., Reekie, I., Bernard, F., Krejci, A., and
733 Bray, S. (2013). Notch cooperates with Lozenge/Runx to lock haemocytes into a
734 differentiation programme. *Development* 140, 926-937.
- 735 Valerius, M.T., Patterson, L.T., Witte, D.P., and Potter, S.S. (2002). Microarray analysis
736 of novel cell lines representing two stages of metanephric mesenchyme differentiation.
737 *Mech Dev* 110, 151-164.
- 738 van Bakel, H. (2011). Interactions of transcription factors with chromatin. *Sub-cellular*
739 *biochemistry* 52, 223-259.
- 740 Xu, J., Lamouille, S., and Derynck, R. (2009). TGF-beta-induced epithelial to
741 mesenchymal transition. *Cell Res* 19, 156-172.
- 742 Zhang, Y., Hassan, M.Q., Li, Z.Y., Stein, J.L., Lian, J.B., van Wijnen, A.J., and Stein, G.S.
743 (2008a). Intricate gene regulatory networks of helix-loop-helix (HLH) proteins support
744 regulation of bone-tissue related genes during osteoblast differentiation. *J Cell Biochem*
745 105, 487-496.
- 746 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum,
747 C., Myers, R.M., Brown, M., Li, W., *et al.* (2008b). Model-based analysis of ChIP-Seq
748 (MACS). *Genome Biol* 9, R137.
- 749 Zhou, T., Luo, M., Cai, W., Zhou, S., Feng, D., Xu, C., and Wang, H. (2018). Runt-Related
750 Transcription Factor 1 (RUNX1) Promotes TGF-beta-Induced Renal Tubular Epithelial-
751 to-Mesenchymal Transition (EMT) and Renal Fibrosis through the PI3K Subunit
752 p110delta. *EBioMedicine* 31, 217-225.
- 753

Figure 1-Supplemental Figure 1

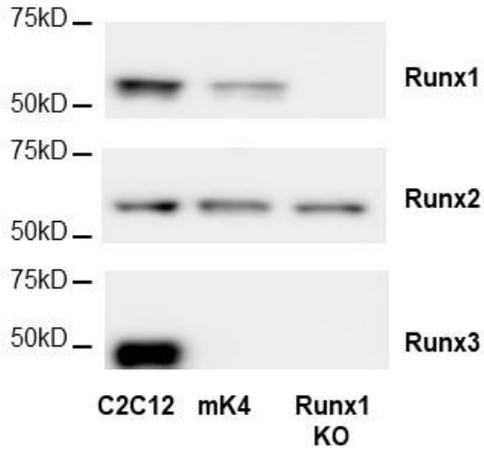


Figure 1-Supplemental Figure 1. Western blot for Runx1, Runx2, and Runx3 in C2C12, mK4 and mK4-Runx1KO cells showing that all three Runx proteins are expressed in C2C12, Runx1 and Runx2 in mK4 cells and only Runx2 in the mK4-Runx1KO cells.

Figure 2-Supplemental Figure 1

Figure 2-Supplemental Figure 1. A) Heatmaps of Runx1 and Runx2 ChIP reads mapped to Runx1 ChIP peaks showing the reproducibility of the ChIP replicates. B) Bar graphs of transcription factor motif enrichment $-\log_{10}$ p-values in the Runx1 ChIP in mK4 cells and Runx2 ChIP in mK4 and Runx1KO cells that confirm the strongest enrichment of Runx motifs.

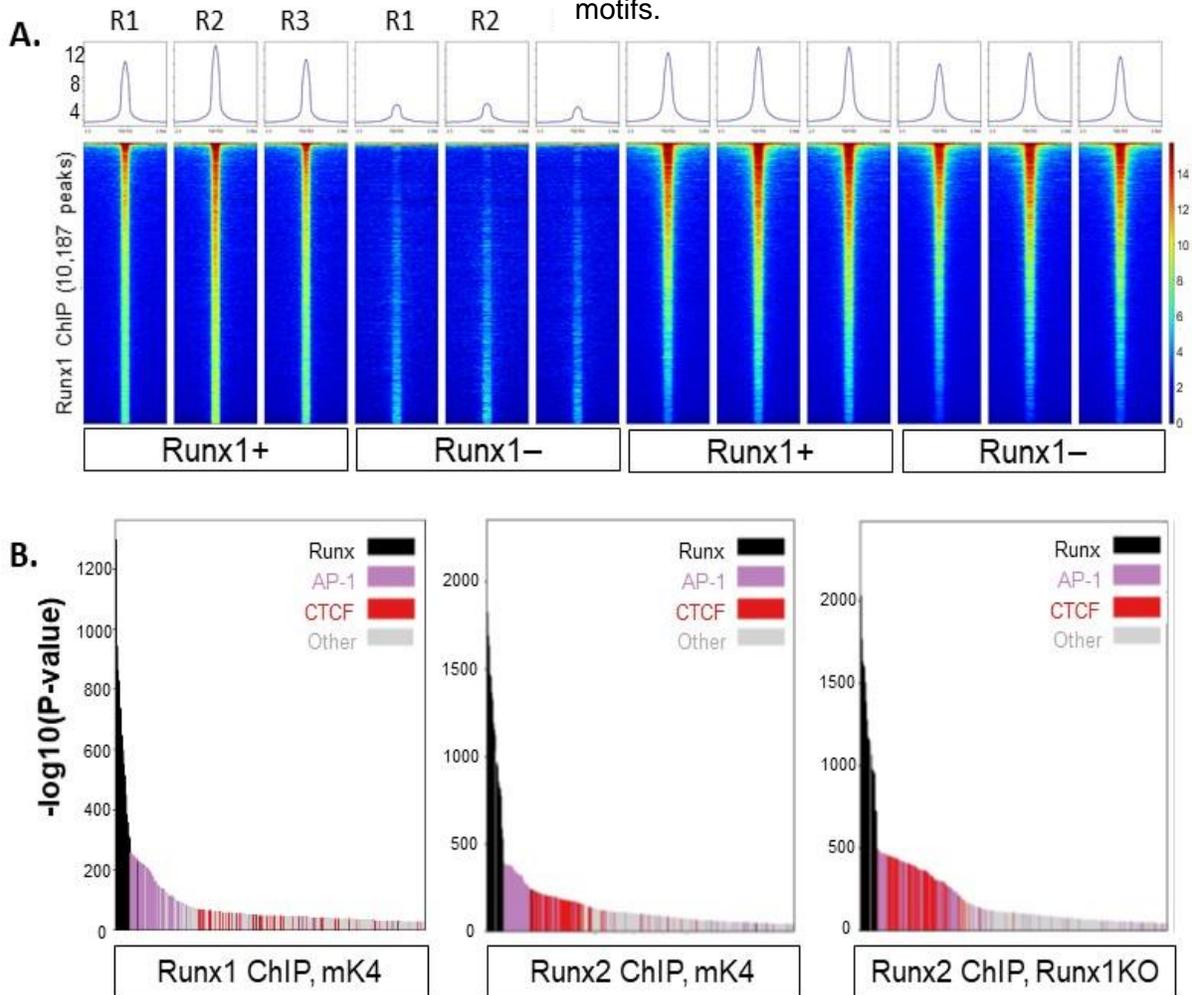
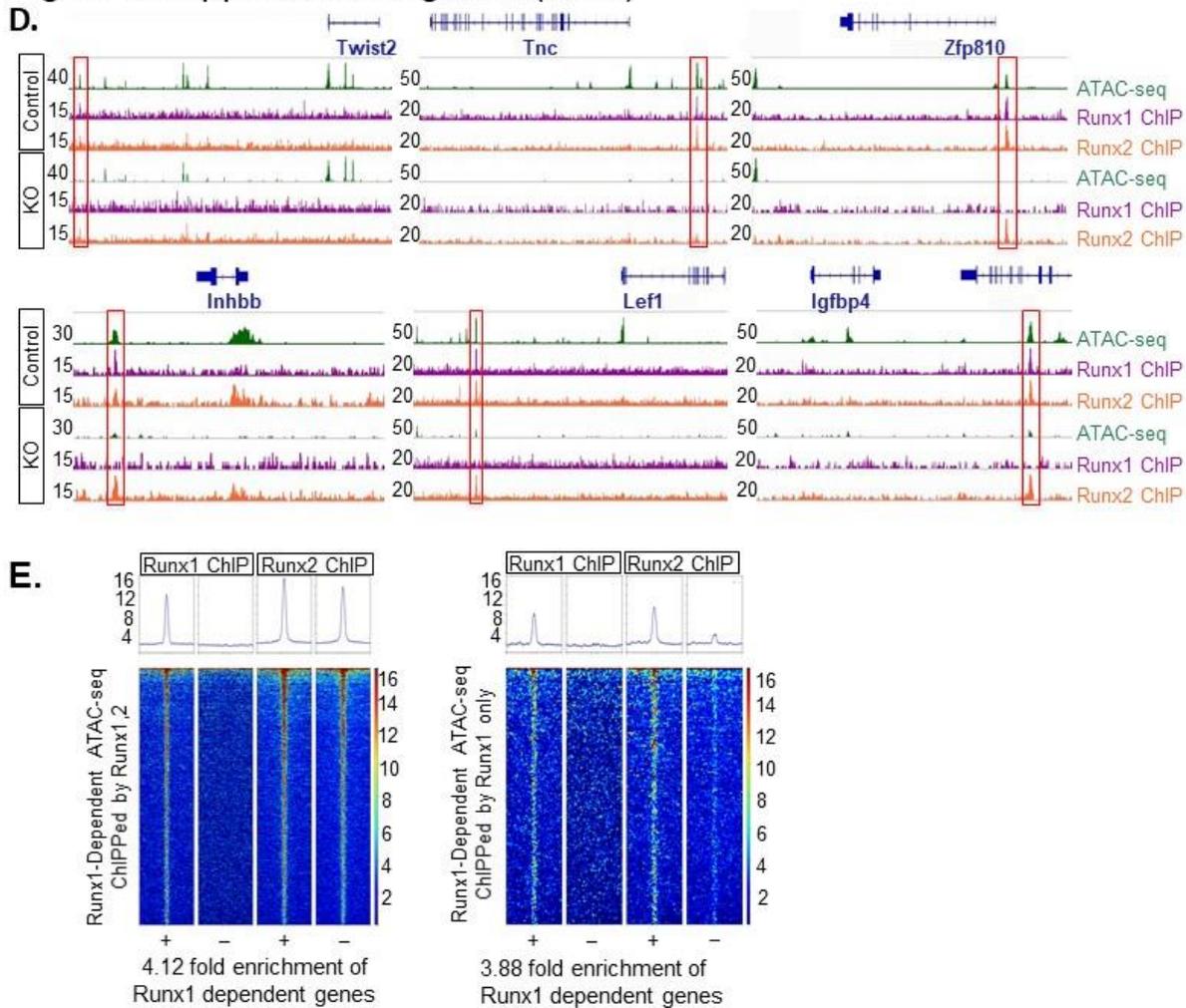


Figure 3-Supplemental Figure 1 (cont.)

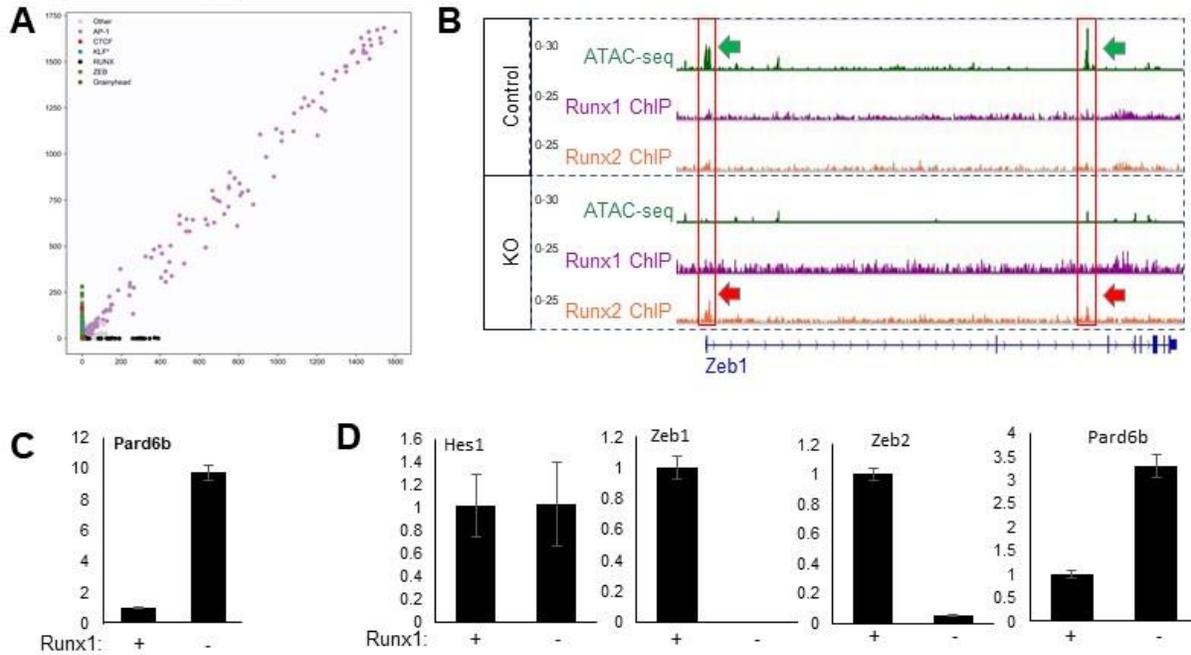


755

756

Figure 3-Supplemental Figure 1. A) Heatmap of Z-scores of individual ATAC-seq samples from mK4 or Runx1KO cells in Runx1-dependent or Runx1KO-induced ATAC-seq peaks showing the reproducibility of the ATAC-seq signal in the replicates and the differences between the mK4 and Runx1KO cells. B) Heatmaps of Runx1 or Runx2 ChIP showing Runx1 binding to both the Runx1-independent and Runx1-dependent ATAC-seq peaks but very little binding to the Runx1KO-induced ATAC-seq peaks. C) Bar graphs of $-\log_{10}$ p-values of motif enrichment in the 3 different classes of ATAC-seq peaks. Runx1-independent ATAC-seq peaks are enriched for AP-1 and Ctfc motifs, Runx1-dependent ATAC-seq enrich for AP-1 and Runx motifs, and Runx1KO-Induced ATAC-seq display enrichment of AP-1, Zeb, and Ctfc motifs. D) Genomic snapshots of Runx1 and Runx2 ChIP and ATAC-seq around the Runx1KO downregulated genes *Twist2*, *Tnc*, *Zfp810*, *Inhbb*, *Lef1*, and *Igfbp4* that shows Runx1 and Runx2 binding to ATAC-seq peaks that are lost in Runx1KO cells. E) Heatmaps Runx1 or Runx2 ChIP on Runx1-dependent ATAC-seq peaks split into those sites that retain Runx2 binding in Runx1KO cells or sites where Runx2 binding is lost. The Runx1-dependent ATAC-seq that lose Runx2 binding in Runx1KO cells fail to enrich for genes that are down-regulated in Runx1KO cells more than the Runx1-dependent ATAC-seq that retain binding of Runx2.

Figure 4-Supplemental Figure 1



757
758

759

760

761

Figure 4-Supplemental Figure 1. A) Plot showing transcription factor motif enrichment (-log₁₀ p-value) in Runx1-dependent versus Runx1KO-induced ATAC-seq peaks that shows AP-1 motifs are strongly enriched in both but Zeb, Ctf, Klf, and Grainyhead motifs are specific to Runx1KO-induced ATAC-seq peaks while Runx1 motifs are enriched only in Runx1-dependent ATAC-seq peaks. B) Genomic snapshot of Zeb1 that shows the loss of ATAC-seq signal at the TSS that is lost in Runx1KO cells, which is consistent with the loss of expression in these cells. C) Graph of normalized RNA-seq reads showing the upregulation of Pard6b in Runx1KO cells as expected for a Zeb repressed gene. D) RT-qPCR analysis confirming the dramatic downregulation of both Zeb1 and Zeb2 and upregulation of Pard6b in Runx1KO cells, but lack of difference for the unrelated gene Hes1.