PLOS BIOLOGY

RESEARCH ARTICLE

Discovery of genes required for body axis and limb formation by global identification of retinoic acid–regulated epigenetic marks

Marie Berenguer[1], Karolin F. Meyer[1], Jun Yin[2], Gregg Duester[1]*

1 Development, Aging, and Regeneration Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California, United States of America, **2** Bioinformatics Core Facility, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California, United States of America

* duester@sbpdiscovery.org

Citation: Berenguer M, Meyer KF, Yin J, Duester G (2020) Discovery of genes required for body axis and limb formation by global identification of retinoic acid–regulated epigenetic marks. PLoS Biol 18(5): e3000719. https://doi.org/10.1371/journal.pbio.3000719

Academic Editor: Yi-Hsien Su, Academia Sinica, TAIWAN

Received: January 14, 2020

Accepted: April 29, 2020

Published: May 18, 2020

Data Availability Statement: RNA-seq data have been deposited in GEO under accession number GSE131584. ChIP-seq data have been deposited in GEO under accession number GSE131624. All other relevant data are within the paper and its Supporting Information files; individual numerical values can be found in S1 Data.

Funding: Research reported in this publication was supported by National Institute of Arthritis and Musculoskeletal and Skin Diseases at the National Institutes of Health (https://www.nih.gov/) under

Abstract

Identification of target genes that mediate required functions downstream of transcription factors is hampered by the large number of genes whose expression changes when the factor is removed from a specific tissue and the numerous binding sites for the factor in the genome. Retinoic acid (RA) regulates transcription via RA receptors bound to RA response elements (RAREs) of which there are thousands in vertebrate genomes. Here, we combined chromatin immunoprecipitation sequencing (ChIP-seq) for epigenetic marks and RNA-seq on trunk tissue from wild-type and *Aldh1a2*-/- embryos lacking RA synthesis that exhibit body axis and forelimb defects. We identified a relatively small number of genes with altered expression when RA is missing that also have nearby RA-regulated deposition of histone H3 K27 acetylation (H3K27ac) (gene activation mark) or histone H3 K27 trimethylation (H3K27me3) (gene repression mark) associated with conserved RAREs, suggesting these genes function downstream of RA. RA-regulated epigenetic marks were identified near RA target genes already known to be required for body axis and limb formation, thus validating our approach; plus, many other candidate RA target genes were found. Nuclear receptor 2f1 (*Nr2f1*) and nuclear receptor 2f2 (*Nr2f2*) in addition to Meis homeobox 1 (*Meis1*) and Meis homeobox *2* (*Meis2)* gene family members were identified by our approach, and double knockouts of each family demonstrated previously unknown requirements for body axis and/or limb formation. A similar epigenetic approach can be used to determine the target genes for any transcriptional regulator for which a knockout is available.

Introduction

Retinoic acid (RA) is generated from retinol by the sequential activities of retinol dehydrogenase 10 (RDH10) [1] and aldehyde dehydrogenase 1A2 (ALDH1A2) [2,3]. Knockout studies of these enzymes revealed an essential role for RA in many early developmental programs, including those controlling hindbrain anteroposterior patterning, neuromesodermal progenitor (NMP) differentiation, spinal cord neurogenesis, somitogenesis, forelimb bud initiation,

award number R01AR067731 (GD). 100%/$2,145,500 of the total project costs were financed with Federal funding. 0%/$0 of the total costs were financed with non-Federal funding. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ALDH1A2, aldehyde dehydrogenase 1A2; AU, arbitrary units; ChIP-seq, chromatin immunoprecipitation sequencing; DR1, direct repeat separated by 1 bp; DR2, direct repeat separated by 2 bp; DR5, direct repeat separated by 5 bp; E, embryonic day; ES, embryonic stem; FGF, fibroblast growth factor; FPKM, fragment per kilobase of transcript per million mapped reads; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; HDAC1, histone deacetylase 1; IPA, Ingenuity Pathway Analysis; NCOA, nuclear receptor coactivator; NCOR, nuclear receptor corepressor; NMP, neuromesodermal progenitor; PRC2, polycomb repressive complex 2; qRT-PCR, quantitative reverse transcription-polymerase chain reaction; RA, retinoic acid; RAR, retinoic acid receptor; RARE, retinoic acid response element; RDH10, retinol dehydrogenase 10; RXR, retinoid X receptor; sgRNA, single-guide RNA; TAD, topologically associated domain; TALE, 3 amino acid loop extension; UCSC, University of California Santa Cruz.

and heart anteroposterior patterning [4,5]. RA functions as a ligand for nuclear RA receptors (RARs) that bind DNA sequences known as RA response elements (RAREs) as a heterodimer complex with retinoid X receptors (RXRs) [6]. Binding of RA to RAR alters the ability of RAREs to recruit nuclear receptor coactivators (NCOAs) that activate transcription or nuclear receptor corepressors (NCORs) that repress transcription [7]. Thus, RA functions are mediated by transcriptional activation or repression of key genes via RAREs.

Identification of RA-regulated genes that are required for development has been difficult, as loss or gain of RA activity alters the mRNA levels of thousands of genes in various cell lines or animals, perhaps most being indirect targets of RA or regulated posttranscriptionally. As RA target genes are dependent upon RAREs, identification of RAREs by RAR-binding studies, cell line transfection assays, and enhancer reporter transgenes in mouse or zebrafish have been used to identify RA target genes that may be required for development, but progress is slow, as each gene is analyzed separately [5]. Genomic RAR chromatin immunoprecipitation sequencing (ChIP-seq) studies on mouse embryoid bodies and F9 embryonal carcinoma cells reported approximately 14,000 potential RAREs in the mouse genome [8,9], but it is unclear how many of these RAREs are required to regulate genes in any specific tissue, and many may not function in any tissue at any stage of development. Only a few RAREs have been shown to result in gene expression and developmental defects when subjected to deletion analysis in mouse, i.e., a RARE enhancer that activates *Hoxa1* in the hindbrain [10], a RARE enhancer that activates *Cdx1* in the spinal cord [11], and a RARE that functions as a silencer to repress caudal *Fgf8* in the developing trunk [7]. In 1 additional case, a RARE described within intron 2 of *Tbx5* that was suggested to be required for activation of *Tbx5* in the forelimb field based on a mouse enhancer reporter transgene [12] was found to be unnecessary for *Tbx5* activation and forelimb budding when subjected to CRISPR deletion analysis, suggesting *Tbx5* is not an RA target gene [13]. Many DNA control elements (including RAREs) that exhibit appropriate tissue-specific expression in enhancer reporter transgene assays have been shown to not be required as an enhancer in vivo when deleted; this may be due to enhancer redundancy or because the control element is really not an enhancer but appeared to be when inserted as a transgene at a random location in the genome near a heterologous promoter [14]. Thus, additional methods are needed (preferably genome-wide) to locate functional RAREs in a particular tissue that can be used to identify new candidate RA target genes that are required for development.

Epigenetic studies have found that histone H3 K27 acetylation (H3K27ac) associates with gene activation and histone H3 K27 trimethylation (H3K27me3) associates with gene repression [15,16]. We suggest that genes possessing nearby H3K27ac and H3K27me3 marks that are altered by loss of RA may point to direct transcriptional targets of RA (either activated or repressed) that are excellent candidates for performing functions downstream of RA. Here, we performed genomic ChIP-seq (H3K27ac and H3K27me3) and RNA-seq studies on embryonic day (E)8.5 mouse embryonic trunks from wild-type and *Aldh1a2*-/- mouse embryos lacking RA synthesis to globally identify RA target genes for embryonic trunk. Candidate targets are defined as genes whose mRNA levels are decreased or increased by genetic loss of RA that also have nearby RA-regulated epigenetic marks associated with conserved RAREs, suggesting they have important downstream functions. Our approach was able to identify many previously reported RA target genes known to control embryonic trunk development (including all 3 known RA target genes from RARE knockout studies: *Hoxa1*, *Cdx1*, and *Fgf8*); plus, we identified numerous new candidate RA target genes that may control trunk development. CRISPR knockout studies on several of these new candidate RA target genes validated them as being required for body axis and/or limb formation. Our approach is generally applicable to determine tissue-specific target genes for any transcriptional regulator that has a knockout available.

Results

Comparison of RNA-seq and H3K27ac/H3K27me3 ChIP-seq for *Aldh1a2*-/- trunk tissue

Embryonic trunks were obtained from E8.5 embryos dissected to remove the head (including the pharyngeal region and anterior hindbrain), heart, and caudal tissue below the most recently formed somite as previously described [17]. We performed RNA-seq analysis comparing E8.5 trunk tissue from wild-type embryos and *Aldh1a2*-/- embryos that lack the ability to produce RA [3]. This analysis identified 4,298 genes whose mRNA levels in trunk tissue are significantly decreased or increased when RA is absent (fragments per kilobase of transcript per million mapped reads [FPKMs] > 0.5; a cutoff of log2 <-0.85 or >0.85 was employed to include *Sox2* known to be activated by RA; data available at GEO under accession number GSE131584).

We performed ChIP-seq analysis for H3K27ac and H3K27me3 epigenetic marks comparing E8.5 trunk tissue from wild-type and *Aldh1a2*-/- embryos isolated as described previously for RNA-seq. This analysis identified 314 RA-regulated ChIP-seq peaks for H3K27ac located within or near 214 genes (i.e., the genes with the nearest annotated promoters) using a log2 cutoff of <0.51 or >0.51 to include an RA-regulated peak near *Sox2* known to be activated by RA [18,19]. We identified 262 RA-regulated peaks for H3K27me3 located within or near 141 genes (i.e., the genes with nearest annotated promoters) using a log2 cutoff of <-0.47 or >0.47 to include an RA-regulated peak near *Fst* known to be repressed by RA [20]; all ChIP-seq data are available at GEO under accession number GSE131624. Thus, we found a much smaller number of RA-regulated ChIP-seq peaks for H3K27ac/H3K27me3 compared with the very large number of genes found to have altered mRNA levels with RNA-seq.

In order to identify genes that are good candidates for being transcriptionally activated or repressed by RA (RA target genes), we compared our ChIP-seq and RNA-seq results to identify RA-regulated ChIP-seq peaks in which nearby genes have significant changes in expression in wild type versus *Aldh1a2*-/- based on RNA-seq. We found 73 RA-regulated peaks for H3K27ac near 63 genes with significant changes in expression when RA is lost (S1 Table), plus 46 RA-regulated peaks for H3K27me3 near 41 genes with significant changes in expression when RA is lost (S2 Table). As some genes have more than 1 nearby RA-regulated peak for H3K27ac or H3K27me3, and some genes have nearby RA-regulated peaks for both H3K27ac and H3K27me3 (*Rarb*, *Dhrs3*, *Fgf8*, *Cdx2*, *Fst*, *Meis* homeobox 1 [*Meis1*], *Meis* homeobox 2 [*Meis2*], nuclear receptor 2f2 [*Nr2f2*], *Foxp4*, *Ptprs*, and *Zfhx4*), a total of 93 RA-regulated genes have nearby RA-regulated peaks for H3K27ac and/or H3K27me3 when RA is lost, thus identifying them as candidate RA target genes for trunk development (S1 and S2 Tables, S1A Fig).

Among the 93 candidate RA target genes for trunk development identified with our approach are many examples of genes previously reported to be regulated by RA in the trunk based on studies of *Aldh1a2*-/- embryos [5,21] or RA-treated NMPs [20]; this includes *Hoxa1*, *Cdx1*, *Rarb*, *Crabp2*, *Sox2*, *Dhrs3*, and *Pax6*, whose expression is increased by RA, plus *Fgf8*, *Cdx2*, and *Fst*, whose expression is decreased by RA (Table 1). H3K27ac peaks near *Cdx1*, *Rarb*, *Crabp2*, *Sox2*, *Dhrs3*, and *Pax6* are reduced in *Aldh1a2*-/- trunk, consistent with these being RA-activated genes, whereas H3K27ac peaks near *Fgf8*, *Cdx2*, and *Fst* are increased in *Aldh1a2*-/-, consistent with these being genes repressed by RA. Conversely, H3K27me3 peaks near *Fgf8*, *Cdx2*, and *Fst* are decreased in *Aldh1a2*-/-, whereas H3K27me3 peaks near *Rarb*, *Hoxa1*, and *Dhrs3* are increased in *Aldh1a2*-/-, consistent with the former being genes repressed by RA and the latter being genes activated by RA (Table 1). In addition to these 10 well-established RA target genes that are required for trunk development, we identified 83

Table 1. Comparison of ChIP-seq and RNA-seq for *Aldh1a2*-/- versus WT E8.5 trunk tissue showing identification of previously known RA target genes needed for trunk development plus some of the additional candidate RA target genes we identified here.

H3K27ac ChIP-seq versus RNA-seq				
H3K27ac ChIP-seq RA-regulated peak for *Aldh1a2* KO versus WT (mm10)	**log2 fold change: H3K27ac ChIP-seq for *Aldh1a2* KO versus WT**	**RARE: based on Homer TFBS analysis**	**Nearby gene with altered expression in *Aldh1a2* KO**	**log2 fold change for nearby gene: RNA-seq for *Aldh1a2* KO versus WT**
chr13:78197222–78204291	−1.23	DR1	Nr2f1	−2.02
chr4:145033496–145035860	−0.65	DR5	Dhrs3*	−1.11
chr14:16571405–16576397	−0.63	DR5	Rarb*	−1.64
chr11:18962656–18965461	−0.61	DR5, DR1	Meis1	−2.64
chr2:105689278–105690982	−0.58	-	Pax6*	−3.02
chr3:87956774–87961235	−0.58	DR2, DR1	Crabp2	−2.82
chr2:116019003–116024272	−0.58	DR2	Meis2	−1.10
chr7:70348715–70369942	−0.57	DR1	Nr2f2	−2.32
chr11:18956989–18958835	−0.57	DR5	Meis1	−2.64
chr11:19012000–19025444	−0.54	DR1	Meis1	−2.64
chr3:34678267–34680699	−0.54	DR2	Sox2*	−0.86
chr18:61033064–61036494	−0.52	DR2, DR1	Cdx1*	−2.00
chr3:34647848–34655776	−0.51	-	Sox2*	−0.86
chr19:45733505–45735997	0.53	DR1	Fgf8*	5.24
chr13:114456392–114460659	0.72	DR2	Fst*	1.15
chr5:147298587–147311126	0.73	DR2	Cdx2*	1.98
H3K27me3 ChIP-seq versus RNA-seq				
H3K27me3 ChIP-seq RA-regulated peak for *Aldh1a2* KO versus WT (mm10)	**log2 fold change: H3K27me3 ChIP-seq for *Aldh1a2* KO versus WT**	**RARE: based on Homer TFBS analysis**	**Nearby gene with altered expression in *Aldh1a2* KO**	**log2 fold change for nearby gene: RNA-seq for *Aldh1a2* KO versus WT**
chr18:38598986–38601292	−1.20	-	Spry4	3.43
chr5:147297983–147318733	−0.63	DR2	Cdx2*	1.98
chr19:45735049–45746658	−0.49	DR2	Fgf8*	5.24
chr13:114456076–114460873	−0.47	DR2	Fst*	1.15
chr4:144893360–144895562	0.59	-	Dhrs3*	−1.11
chr2:116072251–116077455	0.61	DR5	Meis2	−1.10
chr7:70356085–70361002	0.63	DR1	Nr2f2	−2.32
chr6:52156115–52158253	0.73	DR5, DR2	Hoxa1*	−5.43
chr11:19015536–19017169	0.78	DR1	Meis1	−2.64
chr11:19007512–19012358	0.87	DR2	Meis1	−2.64
chr14:16574377–16578138	1.02	DR5, DR1	Rarb*	−1.64

*Previously known to be required for trunk development (body axis, somites, or spinal cord). ChIP-seq values for RA-regulated peaks between *Aldh1a2*-/- (KO) and WT for H3K7ac (log2 <−0.51 or >0.51) and H3K27me3 (log2 <−0.47 or >0.47) with BHP <0.05; a cutoff of log2 <−0.51 or >0.51 for H3K27ac was employed to include an RA-regulated peak near *Sox2* known to be activated by RA; a cutoff of log2 <−0.47 or >0.47 was employed for H3K27me3 to include an RA-regulated peak near *Fst* known to be repressed by RA. RNA-seq values are log2 <−0.85 or >0.85 for differentially expressed genes with FPKM values (KO and WT) >0.5; a cutoff of log2 <−0.85 or >0.85 was employed to include the known RA target gene *Sox2*. Also see related data describing all known and new candidate RA target genes (S1, S2 and S3 Tables).

Abbreviations: *Aldh1a2*, aldehyde dehydrogenase 1A2; BHP, Benjamini and Hochberg method–corrected *p*-value; ChIP-seq, chromatin immunoprecipitation sequencing; DR1 or DR2 or DR5, direct repeat with 1 or 2 or 5 bp between each repeat; FPKM, fragment per kilobase of transcript per million mapped reads; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; KO, knockout; Meis1, Meis homeobox 1; Meis 2, Meis homeobox 2; Nr2f1, nuclear receptor 2f1; Nr2f2, nuclear receptor 2f2; RA, retinoic acid; RARE, retinoic acid response element; TFBS, transcription factor binding site; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.t001

additional genes that our findings indicate are candidate RA target genes for trunk, including *Nr2f1*, *Nr2f2*, *Meis1*, *Meis2*, and *Spry4* that were further examined here (Table 1); differential expression of these genes in E8.5 wild-type versus *Aldh1a2*-/- trunk was validated by quantitative reverse transcription-polymerase chain reaction (qRT-PCR) (S2 Fig). As our approach identified many known trunk RA target genes, it is a reliable approach for identifying new candidate RA target genes required for trunk development.

Identification of RAREs associated with RA-regulated deposition of H3K27ac or H3K27me epigenetic marks

As RA target genes need to be associated with a RARE, the DNA sequences within the RA-regulated H3K27ac/H3K27me3 ChIP-seq peaks we found near our list of 93 RA-regulated genes were searched for RARE sequences using the Homer transcription factor binding site program for the mm10 genome; we searched for 3 types of RAREs including those with a 6-bp direct repeat separated by either 5 bp (DR5), 2 bp (DR2), or 1 bp (DR1) [5], and the presence or absence of RAREs is summarized (S1 and S2 Tables). We found that 46 of these 93 genes contained at least 1 RARE in their nearby RA-regulated H3K27ac and/or H3K27me3 ChIP-seq peaks, thus narrowing down our list of candidate RA target genes to 49% of the genes originally identified. Our approach identified the 3 RAREs previously shown to have required functions during trunk development in vivo by knockout studies (RAREs for *Hoxa1*, *Cdx1*, *Fgf8*) plus several RAREs associated with known RA-regulated genes in the E8.5 trunk from *Aldh1a2*-/- studies (*Rarb*, *Crabp2*, *Sox2*, *Dhrs3*, *Cdx2*, *Fst*), thus validating our approach for identifying RA-regulated genes required for trunk development. The sequences of all the RAREs near these 46 RA target genes here are summarized; included are 65 RAREs near 34 RA-activated genes (we refer to these as RARE enhancers associated with increased H3K27ac and/or decreased H3K27me3 in the presence of RA) and 20 RAREs near 12 RA-repressed genes (we refer to these as RARE silencers associated with increased H3K27me3 and/or decreased H3K27ac in the presence of RA) (S3 Table).

The results here provide evidence that many of the RA-regulated H3K27ac and H3K27me3 marks are associated with regulation of the nearest genes. However, it is possible that some H3K27ac and H3K27me3 RA-regulated peaks may be related to RA-regulated genes located further away in the same topologically associated domain (TAD). In order to address this issue, we assigned each RA-regulated H3K27ac and H3K27me3 peak to a TAD using the 3D Genome Browser (http://promoter.bx.psu.edu/hi-c/view.php); TAD analysis has not been performed on mouse E8.5 trunk tissue, but as TAD domains are similar between different tissues [22], we used the TAD database for mouse embryonic stem (ES) cells, which is the closest biologically relevant database in the 3D Genome Browser. Then, the genes in each TAD containing an RA-regulated peak were searched in our RNA-seq database to identify genes whose mRNA levels are decreased or increased when RA is lost, and if at least 1 gene was found, we determined whether a RARE is present in the ChIP-seq peak. This analysis resulted in the identification of 82 additional RARE enhancers near RA-activated genes and 40 additional RARE silencers near RA-repressed genes, where the gene is not the gene nearest to the RARE in the TAD; in some cases, more than 1 RA-regulated gene was identified in a TAD (S3 Table).

Up to now, *Fgf8* represents the only example of a gene that is directly repressed by RA at the transcriptional level as shown by developmental defects upon knockout of the RARE at −4.1 kb and by the ability of this RARE to stimulate binding of NCOR and polycomb repressive complex 2 (PRC2) plus deposition of H3K27me3 in an RA-dependent manner [7,17]. Here, in addition to *Fgf8*, we found many more candidates for genes repressed by RA in the trunk based on identification of nearby RARE silencers (S3 Table).

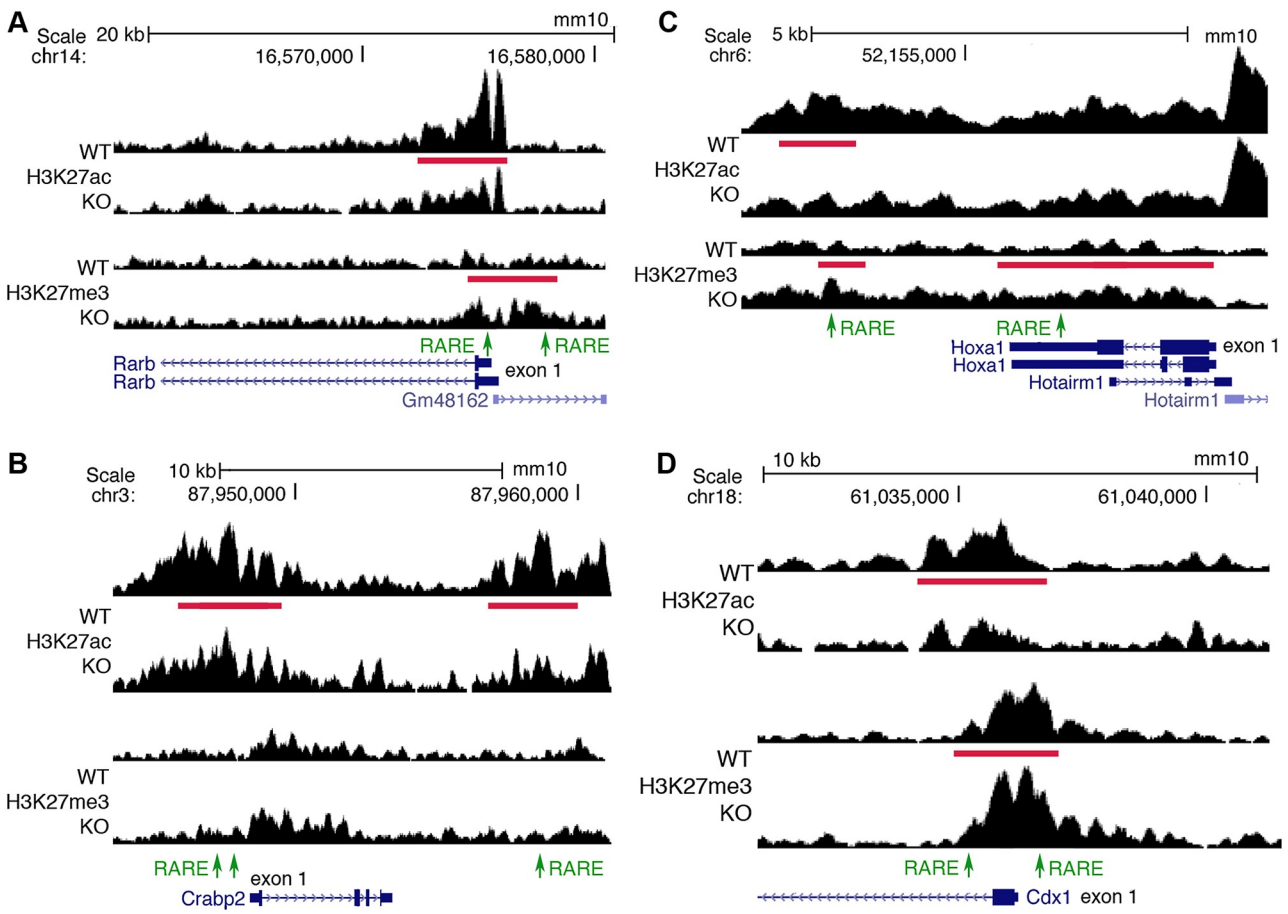

Fig 1. ChIP-seq findings for *Rarb*, *Crabp2*, *Hoxa1*, and *Cdx1* showing that RA-regulated peaks for H3K27ac and H3K7me3 are located near known RARE enhancers. (A) Shown for *Rarb* are RA-regulated ChIP-seq peaks for H3K27ac and H3K27me3 (red bars) when RA is lost in E8.5 trunk comparing WT versus *Aldh1a2*-/- (KO) as well as RAREs (green). A RARE in the 5′-untranslated region is known to function as an RA-dependent enhancer in mouse transgene studies [23]; here, H3K27ac is decreased and H3K27me3 increased near the native RARE when RA is lost in trunk tissue, supporting its function as a RARE enhancer in vivo. We also found a RARE in the 5′-noncoding region of *Rarb* within an H3K27me3 ChIP-seq peak that is increased when RA is lost. (B) RA-regulated peaks for H3K27ac and RAREs are shown for *Crabp2*. The 2 RAREs in the 5′-noncoding region were previously shown to function as RA-dependent enhancers in cell line studies [24]. Our epigenetic studies also identified another RARE enhancer in the 3′-noncoding region. (C) RA-regulated peaks for H3K27ac and/or H3K27me3 and RAREs are shown for *Hoxa1*. KO studies in mouse embryos have shown that the RARE in the 3′-noncoding region is essential for hindbrain *Hoxa1* expression and development [10]. (D) RA-regulated peaks for H3K27ac and H3K27me3 and RAREs are shown for *Cdx1*. KO studies in mouse embryos have shown that the RARE in the 5′-noncoding region is essential for *Cdx1* expression and body axis development [11]. RA-regulated peaks in the genome browser view shown here and elsewhere are for 1 replicate, with the other replicate showing a similar result. ChIP-seq, chromatin immunoprecipitation sequencing; E, embryonic day; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; KO, knockout; RA, retinoic acid; RARE, RA response element; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.g001

Analysis of known RA target genes for trunk validates our approach for finding new targets

The RA-regulated H3K27ac and/or H3K27me3 peaks we identified near *Rarb*, *Crabp2*, *Hoxa1*, and *Cdx1* all overlap previously reported RAREs for these genes (Fig 1). In the case of *Rarb*, the DR5 RARE in the 5′-untranslated region [23] overlaps RA-regulated peaks for both H3K27ac and H3K27me3, suggesting that this RARE in the presence of RA stimulates deposition of H3K27ac and removal of H3K27me3 during activation of *Rarb*; we identified a DR1 RARE in the 5′-noncoding region of *Rarb* within another RA-regulated H3K27me3 ChIP-seq peak (Fig 1A). For *Crabp2*, 2 closely spaced RAREs previously reported in the 5′-noncoding

region [24] associate with RA-regulated peaks for H3K27ac; plus, another RARE we identified in the 3′-noncoding region associates with changes in H3K27ac (Fig 1B). For *Hoxa1*, the RARE located in the 3′-noncoding region is associated with RA-regulated peaks for both H3K27ac and H3K27me3; plus, another RARE we identified in the 3′-untranslated region is associated with RA-regulated peaks for H3K27me3 (Fig 1C); importantly, knockout studies on the *Hoxa1* RARE in the 3′-noncoding region demonstrated that it is required in vivo for *Hoxa1* expression and normal development [10]. For *Cdx1*, 2 RAREs have been reported, 1 in the 5′-noncoding region that was shown by knockout studies to be required for *Cdx1* expression and body axis development [11] plus another RARE in intron 1 [25]. Both of these *Cdx1* RAREs are overlapped by RA-regulated peaks for both H3K27ac and H3K27me3 (Fig 1D). These findings demonstrate that our approach can identify genes that are already known to be transcriptionally activated by RA via a RARE and required for development.

Identification of RA-regulated epigenetic marks and RAREs near RA-regulated genes known to control NMPs

Ingenuity Pathway Analysis (IPA) of our list of 93 RA target genes shows enrichment for the pathway "development of body trunk", including *Sox2*, *Cdx2*, and *Fgf8* known to be required for NMP function during trunk development (S1B Fig). NMPs are bipotential progenitor cells in the caudal region coexpressing *Sox2* and *T/Bra* that undergo balanced differentiation to either spinal cord neuroectoderm or presomitic mesoderm to generate the postcranial body axis [26–33]. NMPs are first observed in mouse embryos at about E8.0 near the node and caudal lateral epiblast lying on each side of the primitive streak [34–36]. Caudal Wnt and fibroblast growth factor (FGF) signals are required to establish and maintain NMPs [34,36–41]. *Cdx2* is required for establishment of NMPs [33]. During development, RA is first produced at E7.5 in presomitic mesoderm expressing *Aldh1a2* to generate an anteroposterior gradient of RA with high activity in the trunk and low activity caudally [5]. Loss of RA does not prevent establishment or maintenance of NMPs but does result in unbalanced differentiation of NMPs, with decreased caudal *Sox2* expression and decreased appearance of neural progenitors, plus increased caudal *Fgf8* expression and increased appearance of mesodermal progenitors and small somites because of the encroachment of caudal *Fgf8* expression into the trunk where it reduces epithelial condensation of presomitic mesoderm needed to form somites [19,20,42,43]. *Cdx2* expression is increased when RA is lost in *Aldh1a2*-/- embryos [44].

Here, when RA is lost, we observed RA-regulated H3K27ac and/or H3K27me3 peaks near several genes required for NMP function that show decreased (*Sox2*) or increased (*Fgf8* and *Cdx2*) expression (Fig 2A–2C). Most of these RA-regulated peaks contain RAREs, providing evidence that *Sox2*, *Fgf8*, and *Cdx2* are direct RA target genes (S3 Table). For *Sox2*, we observed 2 RA-regulated H3K27ac ChIP-seq peaks, but only the one in the 3′-noncoding region was found to have a RARE (Fig 2A). In the case of *Fgf8*, previous studies reporting knockout of the RARE located in the 5′-noncoding region at −4.1 kb resulted in increased caudal *Fgf8* expression and a small somite phenotype (although the defect is not as severe as for *Aldh1a2*-/- embryos), demonstrating that this RARE functions in vivo as a silencer by RA-dependent recruitment of NCORs [7]; RARE redundancy may explain the milder phenotype, as our approach suggests that *Fgf8* has 2 additional candidate RARE silencers (Fig 2B). RARE redundancy may be common, as we observe that *Cdx2* has 2 candidate RARE silencers (Fig 2C), and our overall analysis shows that many genes have more than 1 nearby RARE (S3 Table). These findings provide evidence that RA controls NMP differentiation directly at the transcriptional level by activating *Sox2* and repressing *Fgf8* and *Cdx2* as progenitor cells progress from a caudal to a trunk location.

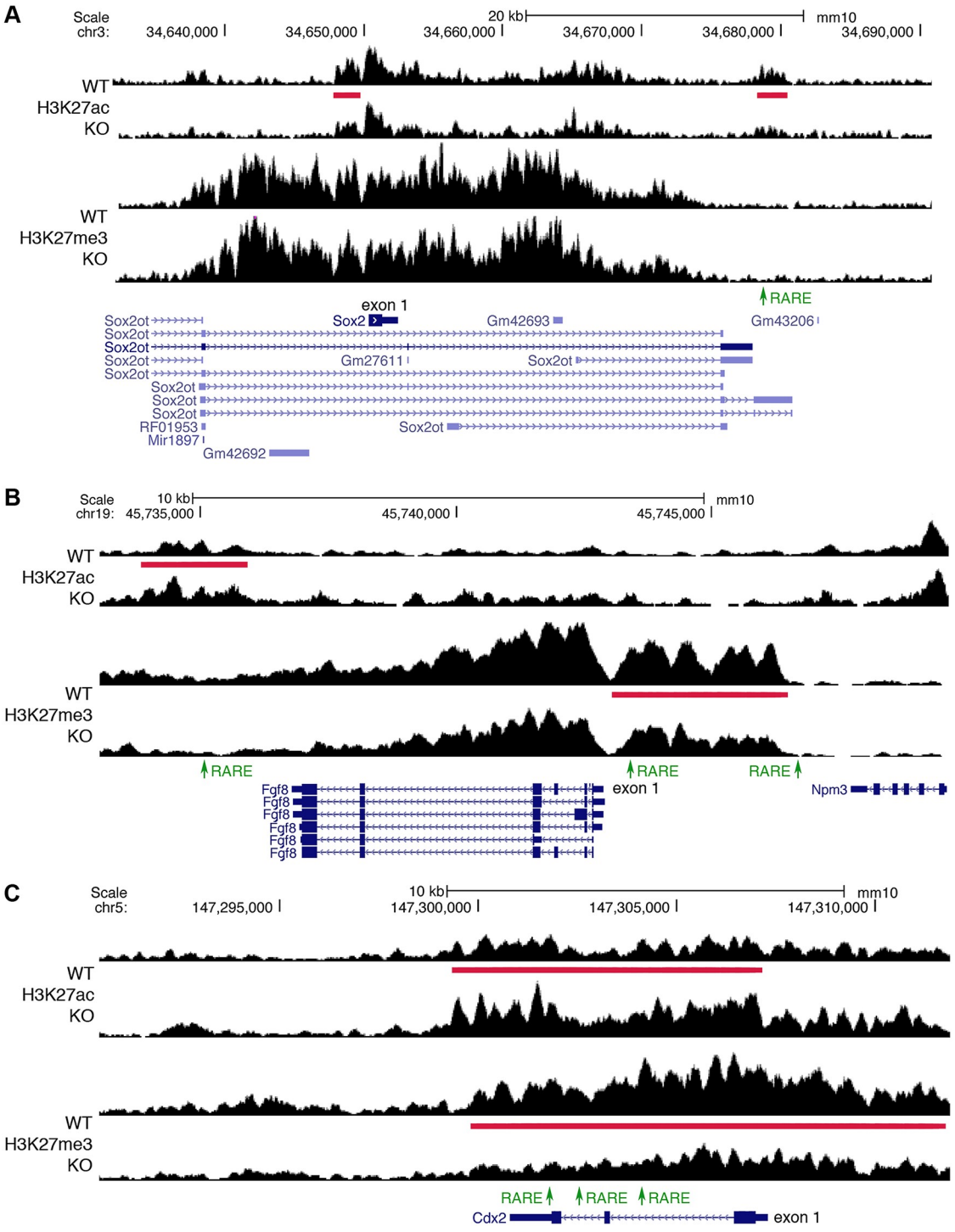

Fig 2. ChIP-seq findings identify RAREs near genes required for NMP function. (A) Two RA-regulated ChIP-seq peaks for H3K27ac (red bars) near *Sox2* are shown for trunk tissue from E8.5 WT versus *Aldh1a2*-/- (KO). A RARE (green) was found in the 3′-noncoding peak (but not the 5′-noncoding peak), suggesting it may function as a RARE enhancer as the H3K27ac peak is decreased when RA is lost. (B) Shown are RA-regulated ChIP-seq peaks for H3K27me3 and H3K27ac near *Fgf8*. In the 5′-noncoding region of *Fgf8*, we found 2 RAREs on either end of the peak for H3K27me3 (repressive mark) that is decreased in KO, indicating they are candidate RARE silencers; the RARE furthest upstream in the 5′-noncoding region at −4.1 kb was shown by knockout studies to function as an RA-dependent RARE silencer required for caudal *Fgf8* repression and somitogenesis [7]. We also found another RARE in the 3′-noncoding region of *Fgf8* that is another candidate for a RARE silencer, as it is contained within an RA-regulated peak for H3K27ac (activating mark) that is increased when RA is lost. (C) *Cdx2* has a peak for H3K27ac that is increased

and an overlapping peak for H3K27me3 that is decreased, along with 3 RAREs included within both peaks, indicating that all these RAREs are candidates for RARE silencers. ChIP-seq, chromatin immunoprecipitation sequencing; E, embryonic day; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; KO,; NMP, neuromesodermal progenitor; RA, retinoic acid; RARE, RA response element; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.g002

Evidence for genes regulated indirectly by RA at the transcriptional level

Our studies show that many genes that are down-regulated or up-regulated following loss of RA are associated with RA-regulated peaks for H3K27ac or H3K27me3 (either nearby or in the same TAD) that do not contain RAREs (S1 and S2 Tables). Such genes may be indirectly activated or repressed by RA at the transcriptional level. In the case of *Pax6*, our results indicate that RA stimulates H3K27ac deposition in *Pax6* introns 2 and 6 that do not contain RAREs, with no other RA-regulated peaks in the same TAD (Fig 3A). Previous studies identified an enhancer in *Pax6* intron 6 containing a SOXB1 binding site that is important for activation in the spinal cord [45]. Activation of *Pax6* in the spinal cord requires CDX proteins in the posterior-most neural tube, and CDX binding sites have been identified in *Pax6* intron 2 [46]; in addition to expression in the caudal progenitor zone, mouse *Cdx1* is expressed in the posterior neural plate where *Pax6* is activated, and this expression domain requires RA [44]. Activation of *Pax6* requires that caudal FGF signaling be down-regulated [43]. Thus, although it is possible that our H3K27ac/H3K27me3 studies failed to identify an unknown RARE near *Pax6*, our findings suggest that the RA requirement for *Pax6* activation may operate through several indirect mechanisms because of the ability of RA to activate *Sox2* and *Cdx1* and repress *Fgf8* (Figs 1 and 2).

We observed that *Spry4* (shown here to be down-regulated by RA) does not have a RARE associated with its RA-regulated ChIP-seq peak for H3K27me3; no other RA-regulated peaks were found in its TAD (Fig 3B). Many of the RA-regulated ChIP-seq peaks observed with our approach that do not contain RAREs may be indirect RA-regulated peaks that contain DNA binding sites for transcription factors other than RARs whose expression or activity is altered by loss of RA, thus resulting in changes for H3K27ac/H3K27me3 marks that are caused by the other transcription factors.

Conservation of RAREs identified with our approach identifies candidate RA target genes

Although it possible that some RAREs that are conserved only in mammals perform mammal-specific functions, RAREs that are conserved from mammals to birds or lower may play fundamental roles in regulation of target genes. The candidate RARE enhancers and RARE silencers we identified here that are associated with RA-regulated epigenetic marks were searched for evolutionary conservation using the University of California Santa Cruz (UCSC) genome browser. Among the RAREs in which the nearest gene is RA-regulated, we found 6 RAREs that are conserved from mouse to zebrafish, 11 conserved to frog (*Xenopus tropicalis*), 18 conserved to reptile (lizard, painted turtle), 20 conserved to bird (chicken, turkey), 39 conserved to human, 65 conserved to rodent (rat), and 20 that are not conserved with rat (S3 Table). The large number of RAREs (i.e., 20) conserved beyond mammals to bird, lizard, frog, or fish demonstrate that our approach is able to identify highly conserved RAREs that point to excellent candidate genes required for development. Among the additional RAREs we found located farther away in the TAD from an RA-regulated gene, we identified only 4 more RAREs conserved beyond mammals to bird, lizard, frog, or fish, thus bringing the total to 24 highly conserved RAREs (S3 Table). Thus, most of the highly conserved RAREs we identified are located

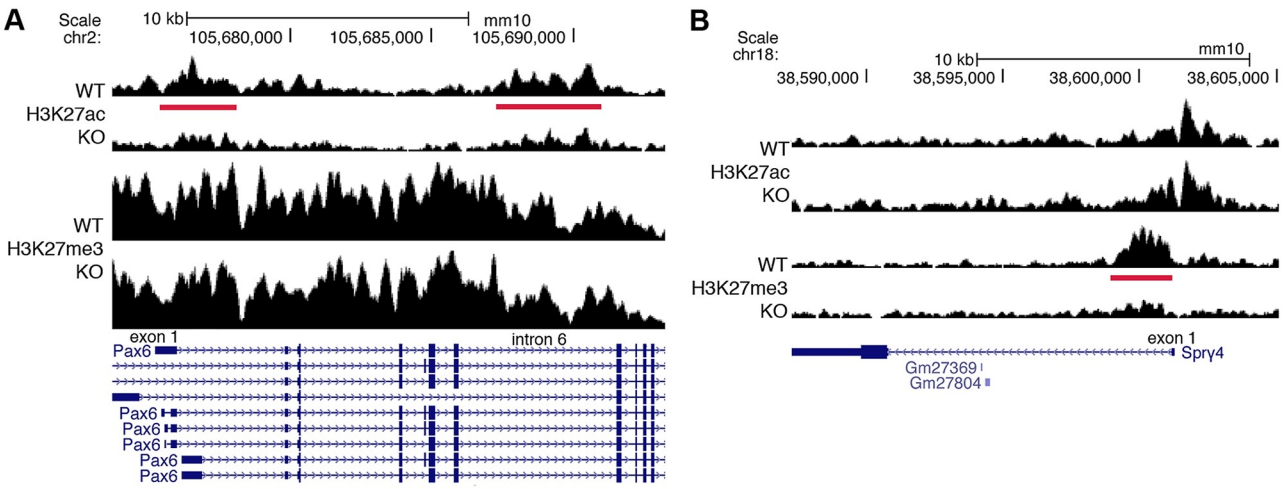

Fig 3. ChIP-seq findings for *Pax6* and *Spry4* that lack RARE enhancers or silencers. These genes are good candidates for being indirect transcriptional targets of RA as their RA-regulated ChIP-seq peaks do not contain RAREs. (A) *Pax6* has 2 RA-regulated peaks (red bars) for H3K27ac (decreased) when RA is lost in E8.5 trunk tissue from *Aldh1a2*-/- (KO) compared with WT; these RA-regulated peaks do not contain RAREs, suggesting that transcription of *Pax6* is indirectly activated by RA. (B) *Spry4* has an RA-regulated peak for H3K27me3 (decreased) when RA is lost with no associated RARE, suggesting that transcription of *Spry4* is indirectly repressed by RA. ChIP-seq, chromatin immunoprecipitation sequencing; E, embryonic day; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; KO, knockout; RA, retinoic acid; RARE, RA response element; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.g003

very close to an RA-regulated gene rather than further distant in the TAD. In addition, all these highly conserved RAREs are either identical to the RARE consensus or have only 1 mismatch. Here we summarize the 24 most highly conserved RAREs that point to 38 RA-regulated genes that may be required for development (Table 2).

As RAREs need to be bound by an RAR in order to function, we examined previously reported RAR ChIP-seq databases for mouse embryoid bodies [8] and mouse F9 embryonal carcinoma cells [9] to determine whether the highly conserved RAREs we identified are included in RAR-binding regions. We found that 19 of our 24 highly conserved RAREs are included in the RAR ChIP-seq peaks from at least 1 of those studies (S4 Table).

Our list of best candidate RA target genes (Table 2) includes several for which knockout studies have already demonstrated required functions during trunk development, i.e., in RA signaling (*Rarb*, *Dhrs3*), body axis formation (*Hoxa1*, *Hoxa4*, *Hoxa9*, *Sox2*, *Fgf8*, *Pbx1*, *Tshz1*, *Zbtb16*), and foregut formation (*Foxp4*); mouse knockout data are summarized by Mouse Genome Informatics (http://www.informatics.jax.org). This list also includes many genes for which knockout studies have either not been performed or knockouts resulted in no reported early developmental defects. This list of genes thus contains excellent new candidates that can be tested for function during trunk development by generating knockouts or double knockouts in the case of gene families.

Nr2f and *Meis* gene families have nearby RA-regulated epigenetic marks associated with highly conserved RARE enhancers

We identified 2 gene families (*Nr2f* and *Meis*) in which 2 family members have decreased expression when RA is lost and nearby RA-regulated peaks for H3K27ac or H3K27me3 containing RAREs. *Nr2f1* and *Nr2f2* encode orphan nuclear receptors NR2F1 and NR2F2 (previously known as COUP-TFI and COUP-TFII, respectively) that regulate transcription, although they have not been found to have endogenous ligands that control their activity [47].

Table 2. DNA sequences of highly conserved RAREs located in RA-regulated ChIP-seq peaks for H3K27ac or H3K27me3 near all RA-regulated genes in same TAD.

RARE Motifs (Homer): DR5 = RAR:RXR(NR),DR5; DR2 = Reverb(NR),DR2; DR1 = TR4(NR),DR1											
Nearest Gene with Decreased or Increased Expression in *Aldh1a2* KO	**Other Genes in Same TAD with Decreased or Increased Expression in *Aldh1a2* KO**	**RARE DNA Sequence 5′-3′ Overall Consensus: AGGTCA N5,N2 AGGTCA G T or N1 G T**	**Type of RARE**	**Conserved**						**Genomic Coordinates (mm10)**
				rodent	**human**	**bird**	**reptile**	**frog**	**fish**	
RARE Enhancers (RA stimulates gain of H3K27ac and/or loss of H3K27me3 near RARE and activates gene in same TAD)										
C1d	none	GGGTCA G GGGTTA	DR1	x	x	x	x	x	x	chr11:18748180–18748192
Clstn1	Lzic, Nmnat1, Kif1b	GGGTCA GA AGGTCA	DR2	x	x	x	x	x		chr4:149907094–149907107
Dach1	None	AGTTCA CACAA AGTTCA	DR5	x	x	x	x	x	x	chr14:98035388–98035404
Dhrs3	None	GGGTCA TTCCA AGTTCA	DR5	x	x	x	x	x		chr4:145034810–145034826
		GGTTCA TCGGG AGGGCA	DR5	x	x	x	x	x		chr4:145034847–145034863
Foxp4	None	GGGTGA C AGGTCA	DR1	x	x	x	x			chr17:47898625–47898637
Hoxa1	Hoxa4, Hoxa9, Skap2	GGTTCA CCGAA AGTTCA	DR5	x	x	x	x	x		chr6:52153426–52153442
		GGTTCA AGAAG AGTTCA	DR5	x	x	x	x	x	x	chr6:52175533–52175549
Meis1	none	AGGCCA CTGAG AGGTCA	DR5	x	x	x	x	x		chr11:18963875–18963891
Meis2	Dph6	AGGTCA AAAAC AGTTCA	DR5	x	x	x	x			chr2:116071242–116071258
Nr2f1	none	GTGTCA A AGTTCA	DR1	x	x	x	x	x	x	chr13:78200425–78200437
Nr2f2	none	GTGTCA A AGTTCA	DR1	x	x	x	x	x	x	chr7:70361772–70361784
Pbx1	Lmx1a	GGGTCG CT GGGTCA	DR2	x	x	x	x			chr1:169238844–169238857
Rarb	none	GGTTCA CCGAA AGTTCA	DR5	x	x	x	x			chr14:16575513–16575529
Sox2	none	GGGTCA GG AGGTCA	DR2	x	x	x	x	x	x	chr3:34679067–34679080
Tshz1	none	GGGTCA TTCAT AGTTCA	DR5	x	x	x	x			chr18:84073476–84073492
		AGGTCA GG AGGTGA	DR2	x	x	x	x			chr18:83839858–83839871
		GGGTGA ACTCA GGTTCA	DR5	x	x	x	x			chr18:83839869–83839885
Zbtb16	none	GGGTCA CA GGGTCA	DR2	x	x	x	x	-	x	chr9:48694721–48694734
		GGGTCA G GGGTTA	DR1	x	x	x	x			chr9:48695827–48695839
Zfhx4	Pex2	GGGTCA GCCTG AGGTCA	DR5	x	x	x	x	x	x	chr3:5388103–5388119
Zfp386	none	GAGTCA A AGGTCA	DR1	x	-	x	x			chr12:117352086–117352098
Zfp638	none	GGTTCA GCCAA AGGTGA	DR5	x	x	x	x	x		chr6:84976840–84976856
RARE Silencers (RA stimulates gain of H3K27me3 and/or loss of H3K27ac near RARE and represses gene in same TAD)										
Fgf8	Poll, Btrc, Mrpl43, Chuk, Sema4g, Dnmbp, Erlin1, Entpd7, Got1, Slc25a28	GGGTCA GC AGTTCA	DR2	x	x	x				chr19:45747043–45747056

RAREs shown here are conserved from mouse to bird, reptile, frog, or fish. RAREs contain no more than 1 mismatch to Homer consensus DR5, DR2, or DR1 RARE motifs shown here.

Abbreviations: *Aldh1a2*, aldehyde dehydrogenase 1A2; ChIP-seq, chromatin immunoprecipitation sequencing; DR, direct repeat; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; KO, knockout; RA, retinoic acid; RARE, RA response element; TAD, topologically associated domain

https://doi.org/10.1371/journal.pbio.3000719.t002

Meis1 and *Meis2* encode transcription factors belonging to the three amino acid loop extension (TALE) family of homeodomain-containing proteins [48,49].

Previous studies suggested that *Nr2f* genes are activated by RA in *Ciona*, zebrafish, and mouse cell lines [50–53]. Here, *Nr2f1* and *Nr2f2* were both found to have a single RARE in the 5′-noncoding region close to exon 1 that is overlapped by or close to the edge of RA-regulated H3K27ac and H3K27me3 peaks (Fig 4A and 4B). Recent studies in zebrafish identified RAREs in similar locations in the *nr2f1a* and *nr2f2* genes [52], and this conservation to mouse was detected by our analysis (Table 2).

Meis1 and *Meis2* were previously shown to be up-regulated by RA in chick limbs treated with RA [54], and loss of RA in *Aldh1a2*-/- embryos results in reduced expression of *Meis2* in paraxial mesoderm [55]. Other studies have shown that *Meis1* and *Meis2* are activated by RA in embryonic stem cells and other cell lines, and RAREs were identified in their 5′-noncoding regions [56,57]. Here, *Meis1* was found to have 4 RAREs in introns 1, 6, and 7 that are overlapped by RA-regulated peaks for H3K27ac and/or H3K27me3; plus, we identified the previously reported RARE in the 5′-noncoding region that is located at the edge of a small RA-regulated H3K27ac peak (Fig 4C). *Meis2* was found to have 2 RAREs that are overlapped by RA-regulated peaks for H3K27ac and/or H3K27me3, 1 in the 5′-noncoding region (previously identified) and another in intron 7 (Fig 4D). Our analysis shows that *Meis1* and *Meis2* each have a highly conserved DR5 RARE enhancer (Table 2). Together, these studies identify *Nf2f1*, *Nr2f2*, *Meis1*, and *Meis2* as candidate RA target genes in the developing trunk.

Nr2f1 and *Nr2f2* function redundantly to control body axis formation

In order to be a biologically important direct RA target gene, the gene must be controlled by a RARE and must also perform a function downstream of RA during trunk development, which can be determined by gene knockout studies. Here, we sought to validate our approach by performing knockout studies on some of the candidate RA target genes, particularly those that have nearby highly conserved RAREs. One could also undertake deletion studies of the RAREs, but this is only relevant after a knockout of the gene itself shows a defect. In addition, as genes are often controlled by redundant enhancers (which we observed here for many genes that have 2 or more RAREs associated with RA-regulated epigenetic marks; S3 Table), studies in which predicted enhancers are deleted often have no effect on development [13,14,58–60]; this includes knockout studies we performed on a RARE that was predicted by enhancer transgene studies to be needed for *Tbx5* expression in forelimb bud that had no effect on *Tbx5* or development [13]. Next, we describe gene knockout studies on candidate RA target genes with nearby highly conserved RAREs to determine whether these genes have a required function in trunk development.

Nr2f1 and *Nr2f2* were selected for gene knockout because they both have nearby candidate RARE enhancers (identified by our H3K27ac/H3K27me3 ChIP-seq analysis) that are conserved from mouse to zebrafish (Table 2). *Nr2f1* (formerly known as COUP-TFI) and *Nr2f2* (formerly known as COUP-TFII) are both expressed at E8.5 in somites and presomitic mesoderm but not the spinal cord, suggesting they may function in mesoderm formation during body axis formation [61,62]. Here, in situ hybridization analysis shows that *Nr2f1* and *Nr2f2* have reduced expression in the trunk of E8.5 *Aldh1a2*-/- embryos compared with wild type (S3 Fig).

The *Nr2f1* knockout is lethal at birth with brain defects, but no somite, spinal cord, or body axis defects are observed [63]. The *Nr2f2* knockout is lethal at E10.5 with defects in heart development but not body axis formation [64]. As redundancy may have masked a body axis defect, we generated *Nr2f1/Nr2f2* double mutants. As it would be quite time-consuming and

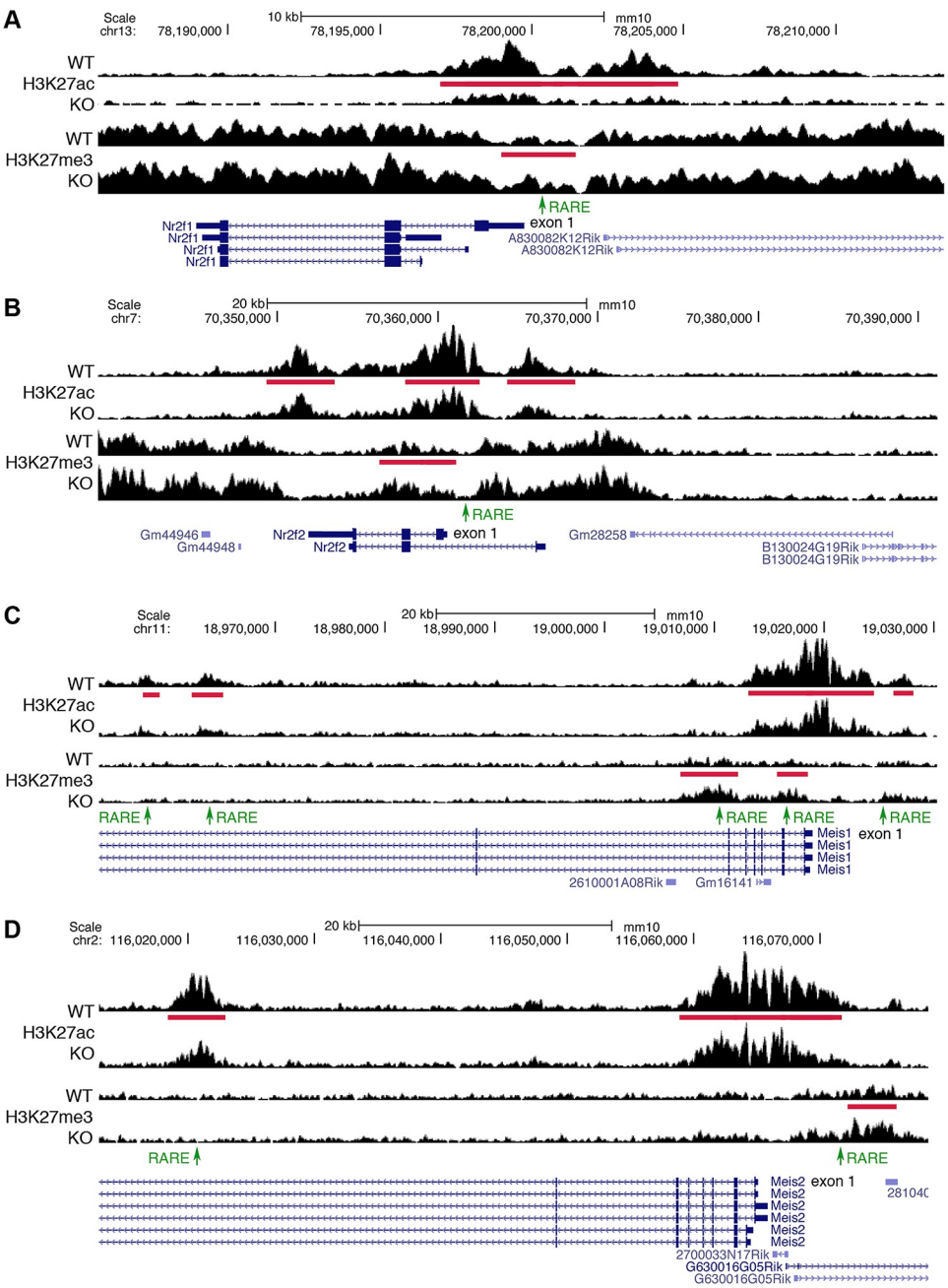

Fig 4. ChIP-seq findings for *Nr2f1*, *Nr2f2*, *Meis1*, and *Meis2* identify RARE enhancers in gene families. (A-B) *Nr2f1* and *Nr2f2* have RA-regulated peaks (red bars) for both H3K27ac (decreased) and H3K27me3 (increased) when RA is lost in E8.5 trunk from *Aldh1a2*-/- (KO) compared with WT. Each family member has one RARE (green) contained within these RA-regulated peaks that are candidates for RARE enhancers. (C-D) *Meis1* and *Meis2* have RA-regulated peaks for both H3K27ac (all decreased) and H3K27me3 (all increased) when RA is lost, along with associated RAREs for each peak that are candidates for RARE enhancers. ChIP-seq, chromatin immunoprecipitation sequencing; E, embryonic day; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; KO, knockout; *Meis1*, *Meis* homeobox 1; *Meis2*, *Meis* homeobox 2; *Nr2f1*, nuclear receptor 2f1; *Nr2f2*, nuclear receptor 2f2; RA, retinoic acid; RARE, RA response element; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.g004

expensive to obtain (if possible) the previously described *Nr2f1* and *Nr2f2* single-knockout mouse lines, then generate a double heterozygote mouse line, and then generate double homozygote embryos at a ratio of 1:16, we employed CRISPR/Cas9 gene editing to examine F0 embryos as we previously described for *Ncor1/Ncor2* double mutants [7]. Fertilized mouse oocytes were injected with single-guide RNAs (sgRNAs) designed to generate frameshift knockout deletions in the second exons of both *Nr2f1* and *Nr2f2*. After dissecting F0 embryos at E9.0, we obtained *Nr2f1/Nr2f2* double knockouts that exhibited a body axis growth defect, more similar in size to that of wild-type E8.25 embryos (Fig 5). Genotyping showed that embryos carrying 1 or 2 knockout alleles were normal in size compared with E9.0 wild type (Fig 5A), whereas embryos carrying either 3 or 4 knockout alleles exhibited a defect in body axis extension and are similar in size to E8.25 wild type; $n = 7$ (Fig 5B and 5C). Staining for *Uncx* somite expression demonstrated that embryos with 1–2 knockout alleles all have a normal number of somites with normal size (Fig 5A), whereas embryos with 3–4 knockout alleles all have fewer somites that are smaller in size; embryos with 3 knockout alleles (either *Nr2f1*-het/*Nr2f2*-hom or *Nr2f1*-hom/*Nr2f2*-het) or 4 knockout alleles (*Nr2f1*-hom/*Nr2f2*-hom) have a similar small somite defect (Fig 5B and 5C). As E9.0 *Nr2f1/Nr2f2* mutants carrying 3–4 knockout alleles are more similar in size to E8.25 wild type, in order to estimate somite size along the anteroposterior axis, we compared them to *Uncx*-stained E8.25 wild-type embryos (Fig 5D), thus revealing that the E9.0 mutants have somites about 57% the size of somites in E8.25 wild-type embryos, showing they have a specific defect in trunk development rather than a global body growth defect (Fig 5E).

Overall, our findings show that loss of 3 or 4 alleles of *Nr2f1* and *Nr2f2* hinders body axis formation and results in smaller somites, thus validating our approach for finding new genes regulated by RA that are required for body axis formation. *Nr2f1/Nr2f2* double mutants exhibit a small somite phenotype (57% of normal) that is similar to RA-deficient *Aldh1a2*-/- embryos (50% of normal); however, they cease body axis extension sooner, as E9.0 *Nr2f1/Nr2f2* double mutants have fewer than 10 somites (Fig 5B and 5C), whereas most E9.0 *Aldh1a2*-/- embryos reach 15 somites [65]. Thus, the more severe growth defect we observe for *Nr2f1/Nr2f2* double mutants may be caused at least in part by a cardiac defect [64]. However, as our RNA-seq and ChIP-seq studies were performed on trunk tissue in which the heart was excluded, our findings show that RA-regulated epigenetic marks are observed near *Nr2f1/Nr2f1* genes located outside the heart in the trunk. Also, our results show that double mutants have smaller somites than wild-type embryos of a comparable size, revealing a specific effect on somitogenesis (body axis extension) rather than just an overall effect on embryonic growth. In the future, more detailed studies of *Nr2f1/Nr2f2* double mutants can be performed to determine how these genes control body axis extension. In addition, future studies can be performed to determine how RARE enhancers function along with other factors to control *Nr2f1* and *Nr2f2* expression during body axis formation.

Meis1 and *Meis2* function redundantly to control both body axis and limb formation

Meis1 and *Meis2* were selected for gene knockout, as *Meis1* has a nearby candidate RARE enhancer conserved from mouse to frog, and *Meis2* has a nearby candidate RARE enhancer conserved from mouse to bird (Table 2). *Meis1* and *Meis2* are both expressed throughout the trunk and in the proximal regions of limb buds [54]. Here, in situ hybridization analysis shows that *Meis1* and *Meis2* have reduced expression in the trunk of E8.5 *Aldh1a2*-/- embryos compared with wild type (S3 Fig).

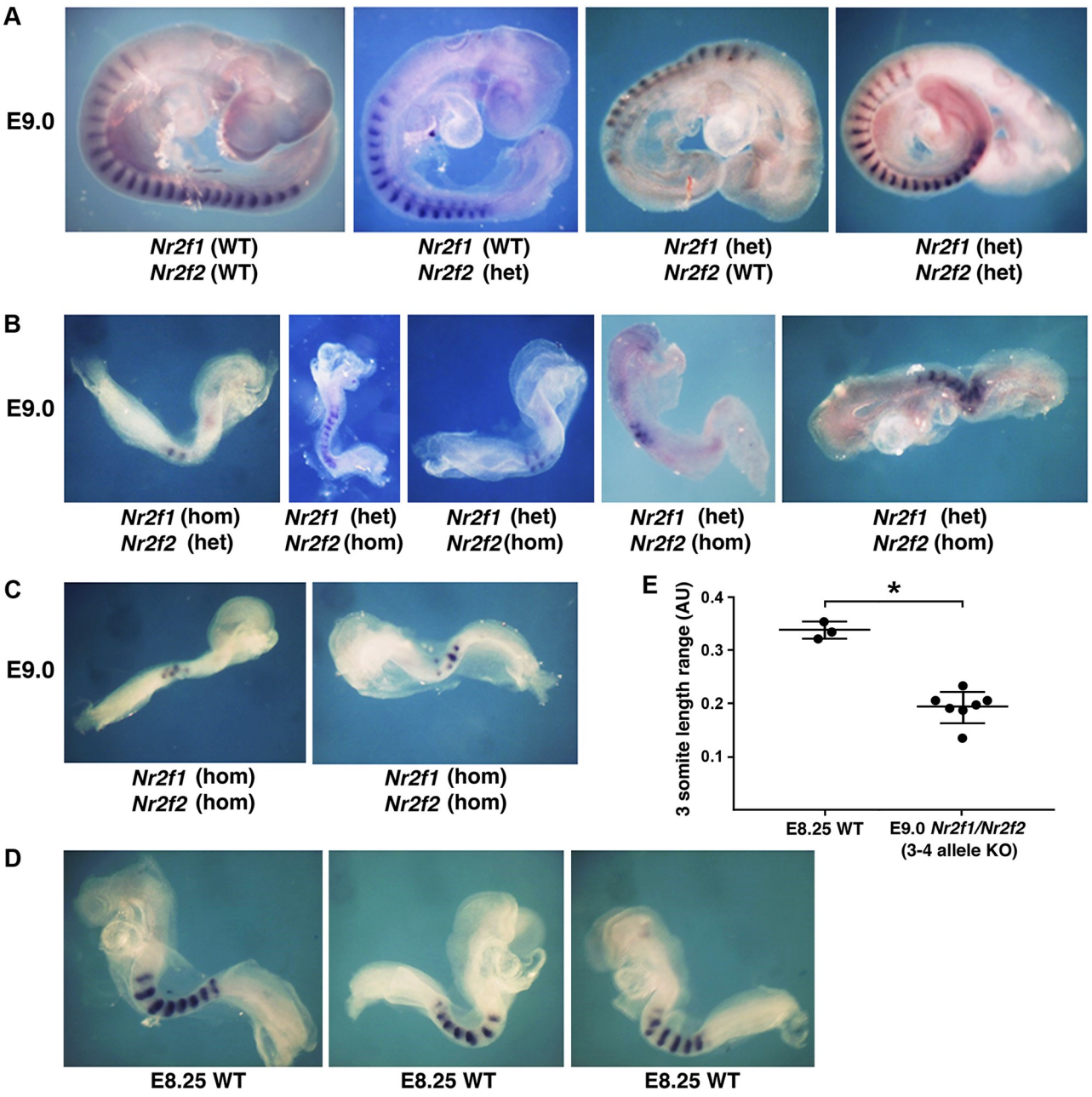

Fig 5. *Nr2f1/Nr2f2* double mutants exhibit defects in body axis formation. (A) Embryos dissected at E9.0 carrying 0–2 knockout alleles for *Nr2f1* or *Nr2f2* have normal somites and body axis formation based on expression of the somite marker *Uncx*. (B-C) Embryos dissected at E9.0 and stained for *Uncx* that carry 3 or 4 knockout alleles for *Nr2f1* or *Nr2f2* exhibit small somites and reduced body axis growth resembling the size of embryos at E8.25. (D) WT E8.25 embryos stained for *Uncx* expression. (E) Comparison of somite size along the anteroposterior axis between E8.25 WT and E9.0 *Nr2f1/Nr2f2* knockout embryos (3–4 knockout alleles); $^{*}p < 0.05$, data expressed as mean ± SD, 1-way ANOVA (nonparametric test); WT, $n = 3$ biological replicates; *Nr2f1/Nr2f2* 3–4 allele double knockout, $n = 7$ biological replicates. Data associated with this figure can be found in S1 Data. AU, arbitrary units; E, embryonic day; het, heterozygous; hom, homozygous; *Nr2f1*, nuclear receptor 2f1; *Nr2f2*, nuclear receptor 2f2; SD, standard deviation; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.g005

The *Meis1* knockout is lethal at E11.5 with hematopoietic defects, but no body axis or limb defects are observed [66]. The *Meis2* knockout is lethal at E14.5 with defects in cranial and cardiac neural crest, but no defects in body axis or limb formation were observed [67]. As redundancy may have masked a body axis or limb defect, we generated *Meis1/Meis2* double mutants via CRISPR/Cas9 gene editing of fertilized mouse oocytes employing sgRNAs designed to generate frameshift knockout deletions in the second exons of both *Meis1* and *Meis2*. Embryos were dissected at E10.5 and stained for somite *Uncx* expression. Genotyping showed that E10.5 embryos carrying 1 or 2 knockout alleles for *Meis1/Meis2* were normal in size with normal-size somites compared with E10.5 wild type (Fig 6A). However, E10.5 embryos carrying 3 or 4 knockout alleles for *Meis1/Meis2* exhibited a body axis extension defect and were either similar in size to *Uncx*-stained E9.5 wild-type embryos ($n = 3$) or smaller ($n = 4$); comparison of somite size along the anteroposterior axis for 5 of these E10.5 mutants shows that somite sizes range from that seen in E9.5 wild type to about 60% of normal (Fig 6B–6D). We observed that E10.5 *Meis1/Meis2* mutants carrying 3–4 knockout alleles that grew similar in size and somite number to E9.5 embryos exhibited a lack of forelimb bud outgrowth; $n = 3$ (Fig 6E).

Overall, our findings show that loss of 3 or 4 alleles of *Meis1* and *Meis2* hinders body axis and forelimb formation, thus providing further evidence that our method of identifying candidate RA target genes can identify genes essential for development. *Meis1/Meis2* double mutants all exhibit a growth defect but display a variable range of somite sizes (from normal to 60% of normal), which is not as severe as for *Aldh1a2*-/- embryos (consistent 50% reduction) [65]; however, *Meis1/Meis2* double mutants that develop to E9.5 (20–25 somites) fail to develop forelimbs (Fig 6E), similar to E9.5 *Aldh1a2*-/- embryos [68]. In the future, more detailed studies of *Meis1/Meis2* double mutants can be performed to determine how these genes control body axis and limb formation; plus, additional studies can be performed to determine how RARE enhancers function along with other factors to control *Meis1* and *Meis2* expression during early development.

Discussion

Our epigenetic ChIP-seq studies combined with RNA-seq on wild-type versus *Aldh1a2*-/- RA-deficient trunk tissue provides a means for identifying new candidate RA target genes that may be required for development. By focusing on RA-regulated genes that also have changes in nearby RA-regulated H3K27ac and/or H3K27me3 epigenetic marks associated with highly conserved RARE enhancers or silencers, our approach can be used to identify excellent candidates for gene knockout studies to learn more about gene function.

Here, in our studies on *Aldh1a2*-/- trunk tissue, we were able to narrow down 4,298 genes identified with RNA-seq that have significant changes in gene expression following loss of RA to 38 excellent candidate RA target genes in E8.5 trunk that also have significant changes in H3K27ac and/or H3K27me3 marks (located nearby or further away in the same TAD) associated with highly conserved RAREs. Our method allows one to identify genes that are most likely to be transcriptional targets of the RA signaling pathway as opposed to those whose expression is changed by effects downstream of RARs and RA signaling such as changes in expression or activity of other transcription factors or posttranscriptional changes in mRNA abundance. Our findings allow us to predict that some genes are likely to be indirect transcriptional targets of RA, as they have nearby RA-regulated peaks for H3K27ac or H3K27me3 but no RAREs, i.e., *Pax6* that is transcriptionally regulated by factors whose expression is altered by loss of RA including *Sox2* [45], *Cdx* [46], and *Fgf8* [43].

Our findings provide evidence for additional RARE silencers. Previous methods designed to identify RAREs favored discovery of RARE enhancers, as studies were designed to find

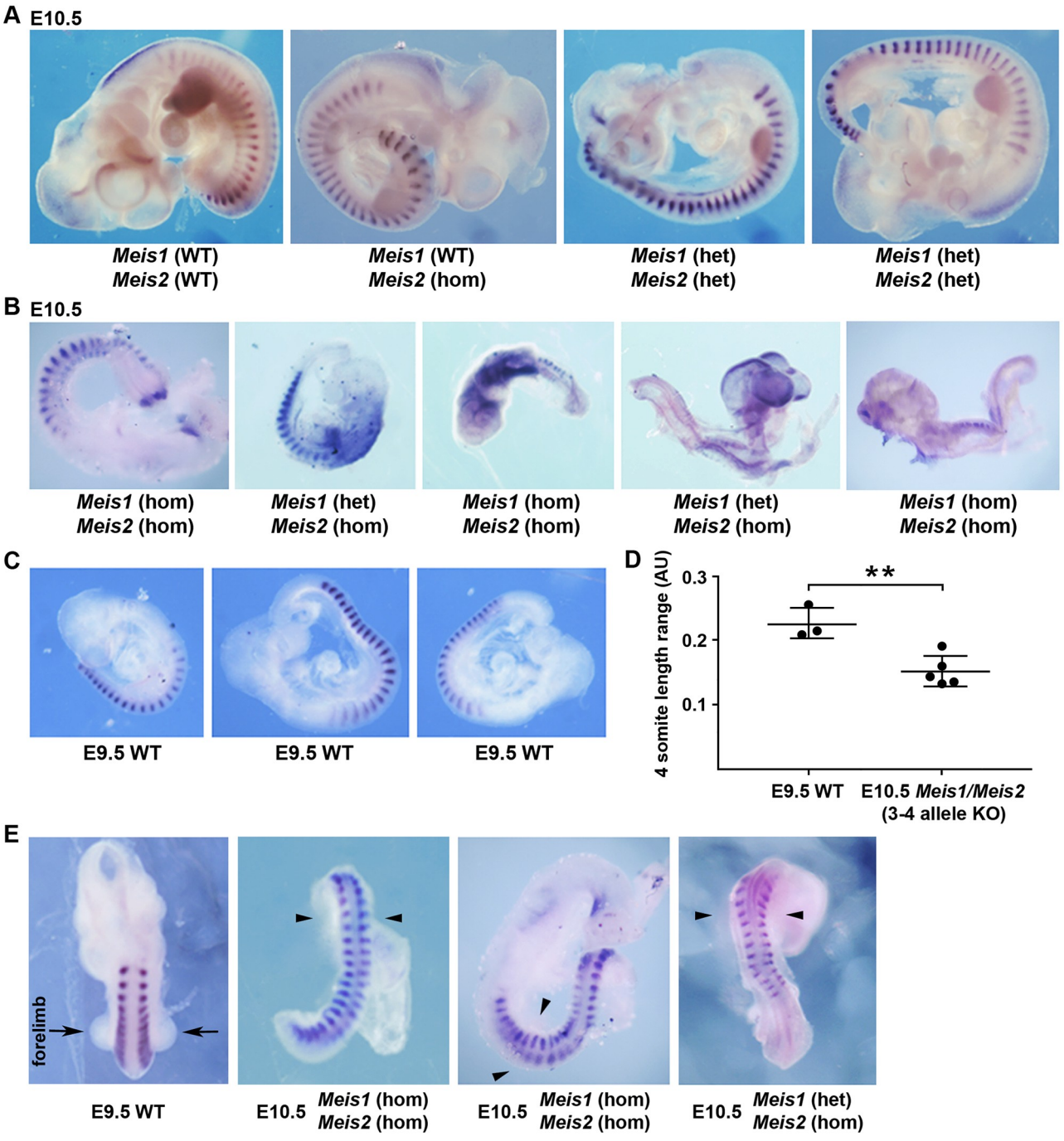

Fig 6. *Meis1/Meis2* double mutants exhibit defects in body axis and forelimb formation. (A) Embryos dissected at E10.5 carrying 0–2 KO alleles for *Meis1* or *Meis2* have normal somites and body axis formation based on expression of the somite marker *Uncx*; also, limb formation is normal. (B) Embryos dissected at E10.5 and stained for *Uncx* that carry 3 or 4 KO alleles for *Meis1* or *Meis2* exhibit small somites and reduced body axis growth resembling the size of embryos at E9.5. (C) WT E9.5 embryos stained for *Uncx* expression. (D) Comparison of somite size along the anteroposterior axis between E9.5 WT and E10.5 *Meis1/Meis2* KO embryos (3–4 KO alleles); $^{*}p < 0.05$, data expressed as mean ± SD, 1-way ANOVA (nonparametric test); WT, $n = 3$ biological replicates; *Meis1/Meis2* 3–4 allele double KO, $n = 5$ biological replicates. (E) Forelimb buds (arrows) normally observed in an E9.5 WT embryo are absent in E10.5 *Meis1/Meis2* KO embryos with 3–4 knockout alleles (arrowheads). Data associated with this figure can be found in S1 Data. AU, arbitrary units; E, embryonic day; het, heterozygous; hom, homozygous; *Meis1*, *Meis* homeobox 1; *Meis2*, *Meis* homeobox 2; KO, knockout; SD, standard deviation; WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000719.g006

DNA elements that, when fused to a heterologous promoter and marker gene, would stimulate expression of the marker gene in the presence of RA. Also, when NCOAs and NCORs that control RA signaling were originally discovered, the model proposed for their function suggested that binding of RA to RAR favored binding of NCOA to activate transcription, with unliganded RAR favoring release of NCOA and binding of NCOR to repress transcription [69]. However, analysis of the *Fgf8* RARE silencer at −4.1 kb demonstrated that RARs bound to RAREs can recruit NCOR in an RA-dependent manner; plus, this RARE is required for normal body axis extension [7]. The *Fgf8* RARE silencer was also found to recruit PRC2 and histone deacetylase 1 (HDAC1) in an RA-dependent manner, providing further evidence that RA can directly control gene silencing [17]. Here, we identified additional RARE silencers near *Fgf8* and *Cdx2* plus several additional genes. Our studies indicate that RARE silencers are less common than RARE enhancers, and we found that *Fgf8* is the only gene associated with a RARE silencer conserved beyond mammals. These additional RARE silencers can be further examined in comparison to the *Fgf8* RARE silencer to determine the mechanism through which RA directly represses transcription. It will be important to determine how RAREs can function as RA-dependent enhancers for some genes but RA-dependent silencers for other genes.

RA has been shown to be required for balanced NMP differentiation during body axis formation by favoring a neural fate over a mesodermal fate [19,29,32]. Our studies provide evidence that RA directly regulates several genes at the trunk/caudal border needed for NMP differentiation—i.e., activation of *Sox2* in the neural plate that favors neural differentiation, repression of *Fgf8* that favors mesodermal differentiation, and repression of *Cdx2* that helps define the location of NMPs. We now provide evidence for a candidate RARE enhancer that activates *Sox2*, 3 candidate RARE silencers that repress *Cdx2*, and 2 additional candidate RARE silencers for *Fgf8*. As the knockout of the original *Fgf8* RARE silencer at −4.1 kb exhibited a body axis phenotype less severe than loss of RA in *Aldh1a2*-/- embryos [7], it is possible that the additional 2 candidate RARE silencers found here provide redundant functions for *Fgf8* repression.

Our observation of highly conserved candidate RARE enhancers near 2 members of 2 different gene families (*Nr2f* and *Meis*) was intriguing, as it suggested that these gene family members may play redundant roles in body axis formation downstream of RA. As we were not sure whether the lack of previous studies on *Nf2f1;Nf2f2* double-knockout and *Meis1;Meis2* double-knockout mouse lines may be due to defects in double heterozygote adults that prevent generation of double homozygote embryos by conventional genetic approaches, we employed CRISPR gene editing to directly generate F0 double knockouts. Our *Nr2f1/Nr2f2* double-knockout studies indeed revealed a defect in body axis formation and small somites that is not observed in each single knockout. Interestingly, zebrafish *nr2f1a/nr2f2* double-knockout embryos reported recently exhibit a heart defect more severe than each single knockout but not a body axis defect or body growth defect [52]. This observation is consistent with studies showing that RA is not required for NMP differentiation or body axis formation in zebrafish [70,71]. Thus, it appears that the ancestral function of *Nr2f* genes in fish was to control heart formation but that during evolution, another function to control body axis formation was added. Future studies can be directed at understanding the mechanism through which *Nr2f1* and *Nr2f2* control body axis formation.

The *Meis1/Meis2* double knockouts we describe here revealed an unexpected function for *Meis* genes in body axis extension and forelimb initiation. *Meis1* and *Meis2* are markers of the proximal limb during forelimb and hindlimb development and were proposed to be activated by RA in the proximal limb as part of the proximodistal limb patterning mechanism in chick embryos [54,72,73]. However, knockout of *Rdh10* required to generate RA demonstrated that

complete loss of RA in the limb fields prior to and during limb development did not affect hindlimb initiation or patterning, whereas forelimbs were stunted but with *Meis1* and *Meis2* expression still maintained in a proximal position in both stunted forelimbs and hindlimbs [74,75] (reviewed in [5]). Our epigenetic results here support the previous proposal that RA can up-regulate *Meis1* and *Meis2* (but in the body axis prior to limb formation as opposed to the limb itself), and we provide evidence that *Meis1* and *Meis2* are transcriptional targets of RA in the body axis. Future studies can be directed at understanding the mechanism through which *Meis1* and *Meis2* control body axis and limb formation.

Our studies demonstrate the power of combining gene knockouts, ChIP-seq on epigenetic marks, and RNA-seq to identify transcription factor target genes required for a particular developmental process. In addition to H3K27ac and H3K27me3 epigenetic marks that are quite commonly observed near genes during activation or repression, respectively, it is likely that further ChIP-seq studies that identify RA-regulated binding sites for coactivators and corepressors will provide additional insight into RA target genes and transcriptional pathways. Such knowledge is essential for determining the mechanisms through which RA controls developmental pathways and should be useful to address RA function in adult organs. A similar epigenetic approach can be used to determine the target genes for any transcriptional regulator for which a knockout is available, thus accelerating the ability to understand gene regulatory networks in general.

Methods

Ethics statement

All mouse studies conformed to the regulatory standards adopted by the Institutional Animal Care and Use Committee at the SBP Medical Discovery Institute, which approved this study under Animal Welfare Assurance Number A3053-01 (approval #18–092). Animal care and use protocols adhered to the guidelines established by the National Institutes of Health (USA).

Generation of *Aldh1a2*-/- mouse embryos and isolation of trunk tissue

Aldh1a2-/- mice have been previously described [3]. E8.5 *Aldh1a2*-/- embryos were generated via timed matings of heterozygous parents; genotyping was performed by PCR analysis of yolk sac DNA. E8.5 trunk tissue was released from the rest of the embryo by dissecting across the posterior hindbrain (to remove the head, anterior hindbrain, pharyngeal region, and heart) and just posterior to the most recently formed somite (to remove the caudal progenitor zone) as previously described [17].

RNA-seq analysis

Total RNA was extracted from E8.5 trunk tissue (2 wild-type trunks and 2 *Aldh1a2*-/- trunks), and DNA sequencing libraries were prepared using the SMARTer Stranded Total RNA-Seq Kit v2 Pico Input Mammalian (Takara). Sequencing was performed on Illumina NextSeq 500, generating 40 million reads per sample with single read lengths of 75 bp. Sequences were aligned to the mouse mm10 reference genome using TopHat splice-aware aligner; transcript abundance was calculated using expectation-maximization approach; FPKM was used for sample normalization; generalized linear model likelihood ratio test in edgeR software was used as a differential test. High-throughput DNA sequencing was performed in the Sanford Burnham Prebys Genomics Core.

qRT-PCR analysis

Total RNA was extracted from 20 trunks of either E8.5 wild-type or *Aldh1a2*-/- embryos with the RNeasy Micro Kit (Qiagen #74004). Reverse transcription was performed with the High-Capacity cDNA RT Kit (Thermo Fisher Scientific #4368814). Quantitative PCR (qPCR) was performed using Power SYBR Green PCR Master Mix (Life Tech Supply #4367659). Relative quantitation was performed using the ddCt method with the control being expression of *Rpl10a*. Primers used for PCR (5′-3′):

Rpl10a-F ACCAGCAGCACTGTGATGAA

Rpl10a-R cAGGATACGTGGgATCTGCT

Rarb-F CTCTCAAAGCCTGCCTCAGT

Rarb-R GTGGTAGCCCGATGACTTGT

Nr2f1-F TCAGGAACAGGTGGAGAAGC

Nr2f1-R ACATACTCCTCCAGGGCACA

Nr2f2-F GACTCCGCCGAGTATAGCTG

Nr2f2-R GAAGCAAGAGCTTTCCGAAC

Meis1-F CAGAAAAAGCAGTTGGCACA

Meis1-R TGCTGACCGTCCATTACAAA

Meis2-F AACAGTTAGCGCAAGACACG

Meis2-R GGGCTGACCCTCTGGACTAT

Spry4-F CCTGTCTGCTGTGCTACCTG

Spry4-R AAGGCTTGTCAGACCTGCTG

ChIP sample preparation for ChIP-seq

For ChIP-seq, we used trunk tissue from E8.5 wild-type or *Aldh1a2*-/- embryos dissected in modified PBS, i.e., phosphate-buffered saline containing 1X complete protease inhibitors (concentration recommended by use of soluble EDTA-free tablets sold by Roche #11873580001) and 10 mM sodium butyrate as a histone deacetylase inhibitor (Sigma # B5887). Samples were processed similar to previous methods [76]. Dissected trunks were briefly centrifuged in 1.5-ml tubes, and excess PBS dissection buffer was removed. For cross-linking of chromatin DNA and proteins, 500 μl 1% formaldehyde was added, the trunk samples were minced by pipetting up and down with a 200-μl pipette tip and then incubated at room temperature for 15 minutes. To stop the cross-linking reaction, 55 μl of 1.25 M glycine was added, and samples were rocked at room temperature for 5 minutes. Samples were centrifuged at 5,000 rpm for 5 minutes, and the supernatant was carefully removed and discarded. A wash was performed in which 1,000 μl of ice-cold modified PBS was added and mixed by vortex, followed by centrifugation at 5,000 rpm for 5 minutes and careful removal of supernatant that was discarded. This wash was repeated. Cross-linked trunk samples were stored at −80 ˚C until enough were collected to proceed, i.e., 100 wild-type trunks and 100 *Aldh1a2*-/- trunks to perform ChIP-seq with 2 antibodies in duplicate.

Chromatin was fragmented by sonication. Cross-linked trunk samples were pooled, briefly centrifuged, and excess PBS removed. A 490-μl lysis buffer (modified PBS containing 1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8.0]) was added, mixed by vortexing, then samples were incubated on ice for 10 minutes. Samples were divided into 4 sonication microtubes (Covaris AFA Fiber Pre-Slit Snap-Cap 6x16 mm, #520045) with 120 μl per tube. Sonication was performed with a Covaris Sonicator with the following settings: Duty, 5%; Cycle, 200; Intensity, 4; #Cycles, 10; 60 seconds each for a total sonication time of 14 minutes. The contents of the 4 tubes were recombined by transfer to a single 1.5-ml microtube, which was then centrifuged for 10 minutes at 13,000 rpm, and the supernatants transferred to a fresh 1.5-ml microtube. These conditions were found to shear trunk DNA to an average size of 300 bp using a 5-μl sample for Bioanalyzer analysis. At this point, 20 μl was removed for each sample (wild-type trunks and *Aldh1a2*-/- trunks) and stored at −20 ˚C to serve as input DNA for ChIP-seq.

Each sample was divided into four 100-μl aliquots to perform immunoprecipitation with 2 antibodies in duplicate. Immunoprecipitation was performed using the Pierce Magnetic ChIP Kit (Thermo Scientific, #26157) following the manufacturer's instructions and ChIP-grade antibodies for H3K27ac (Active Motif, Cat#39133) or H3K27me3 (Active Motif, Cat#39155). The immunoprecipitated samples and input samples were subjected to reversal of cross-linking by adding water to 500 μl and 20 μl 5 M NaCl, vortexing and incubation at 65 ˚C for 4 hours; then addition of 2.6 μl RNase (10 mg/ml), vortexing and incubation at 37 ˚C for 30 min; then addition of 10 μl 0.5 M EDTA, 20 μl 1 M Tris-HCl (pH 8.0), 2 μl proteinase K (10 mg/ml), vortexing and incubation at 45 ˚C for 1 hour. DNA was extracted using ChIP DNA Clean & Concentrator (Zymo, # D5201 & D5205). After elution from the column in 50 μl of elution buffer, the DNA concentration was determine using 2-μl samples for Bioanalyzer analysis. The 2 input samples ranged from 16–20 ng/l and the 8 immunoprecipitated samples ranged from 0.1 to 0.2 ng/μl (5–10 ng per 100 trunks). For ChIP-seq, 2 ng was used per sample to prepare libraries for DNA sequencing.

ChIP-seq genomic sequencing and bioinformatic analysis

Libraries for DNA sequencing were prepared according to the instructions accompanying the NEBNext DNA Ultra II kit (catalog # E7645S; New England Biolabs). Libraries were sequenced on the NextSeq 500 following the manufacturer's protocols, generating 40 million reads per sample with single read lengths of 75 bp. Adapter remnants of sequencing reads were removed using cutadapt v1.18 [77]. ChIP-Seq sequencing reads were aligned using STAR aligner version 2.7 to Mouse genome version 38 [78]. Homer v4.10 [79] was used to call peaks from ChIP-seq samples by comparing the ChIP samples with matching input samples. Homer v4.10 was used to annotate peaks to mouse genes and quantify reads count to peaks. The raw reads count for different peaks were compared using DESeq2 [80]. *p* Values from DESeq2 were corrected using the Benjamini and Hochberg (BH) method for multiple testing errors [81]. Peaks with BH-corrected p-value < 0.05 (BHP < 0.05) were selected as significantly differentially marked peaks. Transcription factor binding sites motif enrichment analyses were performed using Homer v4.10 [79] to analyze the significant RA-regulated ChIP-seq peaks; DR1 RAREs were found by searching for TR4(NR),DR1; DR2 RAREs by Reverb(NR),DR2; and DR5 RAREs by RAR:RXR(NR),DR5. Evolutionary conservation of RAREs was performed via DNA sequence homology searches using the UCSC genome browser software. TAD analysis was performed using the 3D Genome Browser (http://promoter.bx.psu.edu/hi-c/view.php); we used the TAD database from Hi-C data for mouse ES cells reported for the mouse mm10 genome. IPA was used to identify pathways for our list of target genes; from IPA results, heatmaps were designed with Prism software, and associated networks were created using STRING

software. High-throughput DNA sequencing was performed in the Sanford Burnham Prebys Genomics Core, and bioinformatics analysis was performed in the Sanford Burnham Prebys Bioinformatics Core.

Generation of mutant embryos by CRISPR/Cas9 mutagenesis

CRISPR/Cas9 gene editing was performed using methods similar to those previously described by others [82,83] and by our laboratory [7]. sgRNAs were generated that target exons to generate frameshift null mutations, with 2 sgRNAs used together for each gene. sgRNAs were designed with maximum specificity using the tool at crispr.mit.edu to ensure that each sgRNA had no more than 17 out of 20 matches with any other site in the mouse genome and that those sites are not located within exons of other genes. DNA templates for sgRNAs were generated by PCR amplification (Phusion DNA Polymerase; New England Biolabs) of ssDNA oligonucleotides (purchased from Integrated DNA Technologies) containing on the 5′ end a minimal T7 promoter, then a 20-nucleotide sgRNA target sequence (underlined below), and finally the tracrRNA sequence utilized by Cas9 on the 3′ end, shown as follows:

5′-GCGTAATACGACTCACTATAGGNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT-3′

The 20-nucleotide target sequences used were as follows:

Nf2f1 exon 2 (#1) TTTTTATCAGCGGTTCAGCG

Nf2f1 exon 2 (#2) GGTCCATGAAGGCCACGACG

Nf2f2 exon 2 (#1) GGTACGAGTGGCAGTTGAGG

Nf2f2 exon 2 (#2) CGCCGAGTATAGCTGCCTCA

Meis1 exon 2 (#1) CGACGACCTACCCCATTATG

Meis1 exon 2 (#2) TGACCGAGGAACCCATGCTG

Meis2 exon 2 (#1) GATGAGCTGCCCCATTACGG

Meis2 exon 2 (#2) CGACGCCTTGAAAAGAGACA

sgRNAs were then transcribed from templates using HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs) and purified using Megaclear Kit (Life Technologies). sgRNAs were tested in vitro for their cleavage ability in combination with Cas9 nuclease (New England Biolabs); briefly, genomic regions flanking the target sites were PCR amplified, then 100 ng was incubated with 30 nM Cas9 nuclease and 30 ng sgRNA in 30 μl for 1 hour at 37 ˚C, followed by analysis for cleavage by gel electrophoresis.

For injection into mouse embryos, a solution containing 50 ng/μl Cas9 mRNA (Life Technologies) and 20 ng/μl for each sgRNA used was prepared in nuclease free water. Fertilized oocytes were collected from 3–4-week-old superovulated C57Bl6 females prepared by injecting 5 IU each of pregnant mare serum gonadotrophin (PMSG) (Sigma Aldrich) and human chorionic gonadotropin (hCG) (Sigma Aldrich). Fertilized oocytes were then transferred into M2 medium (Millipore) and injected with the Cas9 mRNA/sgRNA solution into the cytoplasm. Injected embryos were cultured in KSOMaa medium (Zenith) in a humidified atmosphere with 5% CO_2 at 37 ˚C overnight to maximize the time for CRISPR/Cas9 gene editing to occur at the 1-cell stage, then reimplanted at the 2-cell stage into recipient pseudopregnant ICR

female mice. Implanted females were sacrificed to obtain F0 E9.0 embryos (*Nr2f1/Nr2f2*) or F0 E10.5 embryos (*Meis1/Meis2*). As fertilized mouse oocytes spend a long time at the 1-cell and 2-cell stages, this facilitates CRISPR/Cas9 gene editing at early stages and allows many F0 embryos to be examined for mutant phenotypes [7]. For genotyping, yolk sac DNA was collected, and PCR products were generated using primers flanking the sgRNA target sites; PCR products were subjected to DNA sequence analysis from both directions using either upstream or downstream primers. For each gene analyzed, embryos were classified as heterozygous (het) if the DNA sequence contained both a wild-type allele and a frameshift allele; embryos were classified as homozygous (hom) if only frameshift alleles were detected but no wild-type sequence.

Body axis length analysis of embryos

ImageJ software (https://imagej.net) [84] was used to measure body axis length along a 3-somite region (*Nr2f1/Nr2f2* double mutants) or 4-somite region (*Meis1/Meis2* double mutants) compared with wild type, with each specimen photographed at the same magnification. Statistical analysis was performed using 1-way ANOVA (nonparametric test) with data presented as mean ± standard deviation (SD) and with $p > 0.05$ indicating significance.

In situ gene expression analysis

Embryos were fixed in paraformaldehyde at 4 ˚C overnight, dehydrated into methanol, and stored at −20 ˚C. Detection of mRNA was performed by whole-mount in situ hybridization as previously described [85].

Supporting information

S1 Table. Comparison of *Aldh1a2*-/- and wild-type E8.5 trunk tissue for H3K27ac ChIP-seq and RNA-seq results to identify RA-regulated H3K27ac ChIP-seq peaks near genes with RA-regulated expression. *Aldh1a2*, aldehyde dehydrogenase 1A2; ChIP-seq, chromatin immunoprecipitation sequencing; E, embryonic day; H3K27ac, histone H3 K27 acetylation; RA, retinoic acid.
(DOCX)

S2 Table. Comparison of *Aldh1a2*-/- and wild-type E8.5 trunk tissue for H3K27me3 ChIP-seq and RNA-seq results to identify RA-regulated H3K27me3 ChIP-seq peaks near genes with RA-regulated expression. *Aldh1a2*, aldehyde dehydrogenase 1A2; ChIP-seq, chromatin immunoprecipitation sequencing; E, embryonic day; H3K27me3, histone H3 K27 trimethylation; RA, retinoic acid.
(DOCX)

S3 Table. DNA sequences of all RAREs located in RA-regulated ChIP-seq peaks for H3K27ac or H3K27me3 near all RA-regulated genes in same TAD. RAREs contain no more than 2 mismatches to Homer consensus DR5, DR2, or DR1 RARE motifs shown here. ChIP-seq, chromatin immunoprecipitation sequencing; DR, direct repeat; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; RA, retinoic acid; RARE, RA response element; TAD, topologically associated domain.
(DOCX)

S4 Table. DNA sequences of highly conserved RAREs and relationship to panRAR ChIP-seq peaks for mouse EBs [8] and F9 embryonal carcinoma cells [9] reported using mm9 genomic coordinates. ChIP-seq, chromatin immunoprecipitation sequencing; EB, embryoid

body; panRAR, including all RA receptor subtypes; RARE, retinoic acid response element.
(DOC)

S1 Fig. Bioinformatic analysis of genes identified as RA target genes. (A) Venn diagram showing the number of genes that have both RA-regulated expression and RA-regulated deposition of nearby H3K27a or H3K27me3 marks following loss of RA. (B) The heatmap was designed with Prism software (left panel) from the list of genes involved in "Development of Body Trunk" obtained by IPA analysis of RA target genes identified by loss of RA, and the associated network was created using STRING software (right panel). Individual values for H3K27ac and H3K27me3 (wild type versus *Aldh1a2*-/-) can be found in ChIP-seq data deposited at GEO under accession number GSE131624. Individual values for RNA-seq (wild type versus *Aldh1a2*-/-) can be found in the data deposited at GEO under accession number GSE131584. *Aldh1a2*, aldehyde dehydrogenase 1A2; ChIP-seq, chromatin immunoprecipitation sequencing; GEO, Gene Expression Omnibus; H3K27ac, histone H3 K27 acetylation; H3K27me3, histone H3 K27 trimethylation; IPA, Ingenuity Pathway Analysis; RA, retinoic acid.
(TIF)

S2 Fig. Analysis of candidate RA target gene expression by qRT-PCR. Analysis of differential expression of several candidate RA target genes by qRT-PCR analysis of E8.5 WT versus *Aldh1a2*-/- trunk tissue. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.005$; data are expressed as mean ± SD (Student *t* test); WT and *Aldh1a2*-/-, $n = 3$. Data associated with this figure can be found in S1 Data. *Aldh1a2*, aldehyde dehydrogenase 1A2; RA, retinoic acid; qRT-PCR, quantitative reverse transcription-polymerase chain reaction; SD, standard deviation; WT, wild type.
(TIF)

S3 Fig. Analysis of *Nr2f* and *Meis* expression by in situ hybridization. Analysis of *Nr2f1*, *Nr2f2*, *Meis1*, and *Meis2* expression by whole-mount in situ hybridization of E8.5 WT versus *Aldh1a2*-/- trunk tissue; $n = 3$ for both WT and *Aldh1a2*-/-; brackets point to trunk tissue; for each gene analyzed, both lateral (left) and dorsal (right) views are shown, except ventral for *Meis1*. *Aldh1a2*, aldehyde dehydrogenase 1A2; E, embryonic day; *Meis1*, *Meis* homeobox 1; *Meis2*, *Meis* homeobox 2; *Nr2f1*, nuclear receptor 2f1; *Nr2f2*, nuclear receptor 2f2; WT, wild type.
(TIF)

S1 Data. Original numerical values for Figs 5E and 6D, and S2 Fig.
(XLSX)

Acknowledgments

We thank the Genomics Core Facility and Bioinformatics Core Facility at SBP Medical Discovery Institute for help with ChIP-seq and RNA-seq analysis. We thank the Animal Resources Core Facility at SBP Medical Discovery Institute for conducting timed matings to generate mouse embryos.

Author Contributions

Conceptualization: Marie Berenguer, Gregg Duester.

Data curation: Marie Berenguer, Jun Yin.

Formal analysis: Marie Berenguer, Gregg Duester.

Funding acquisition: Gregg Duester.

Investigation: Marie Berenguer, Karolin F. Meyer.

Methodology: Marie Berenguer, Karolin F. Meyer, Jun Yin.

Project administration: Gregg Duester.

Resources: Jun Yin.

Software: Jun Yin.

Supervision: Gregg Duester.

Validation: Marie Berenguer, Jun Yin.

Writing – original draft: Marie Berenguer, Gregg Duester.

Writing – review & editing: Marie Berenguer, Jun Yin, Gregg Duester.

References

1. Sandell LL, Sanderson BW, Moiseyev G, Johnson T, Mushegian A, Young K et al. RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. Genes Dev. 2007; 21:1113–1124. https://doi.org/10.1101/gad.1533407 PMID: 17473173
2. Niederreither K, Subbarayan V, Dollé P, Chambon P. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. Nature Genet. 1999; 21:444–448. https://doi.org/10.1038/7788 PMID: 10192400
3. Mic FA, Haselbeck RJ, Cuenca AE, Duester G. Novel retinoic acid generating activities in the neural tube and heart identified by conditional rescue of *Raldh2* null mutant mice. Development. 2002; 129:2271–2282. PMID: 11959834
4. Rhinn M, Dolle P. Retinoic acid signalling during development. Development. 2012; 139:843–858. https://doi.org/10.1242/dev.065938 PMID: 22318625
5. Cunningham TJ, Duester G. Mechanisms of retinoic acid signalling and its roles in organ and limb development. Nature Rev Mol Cell Biol. 2015; 16:110–123. https://doi.org/10.1038/nrm3932
6. Mark M, Ghyselinck NB, Chambon P. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. Ann Rev Pharmacol Toxicol. 2006; 46:451–480.
7. Kumar S, Cunningham TJ, Duester G. Nuclear receptor corepressors *Ncor1* and *Ncor2* (Smrt) are required for retinoic acid-dependent repression of *Fgf8* during somitogenesis. Dev Biol. 2016; 418:204–215. https://doi.org/10.1016/j.ydbio.2016.08.005 PMID: 27506116
8. Moutier E, Ye T, Choukrallah MA, Urban S, Osz J, Chatagnon A et al. Retinoic Acid Receptors Recognize the Mouse Genome through Binding Elements with Diverse Spacing and Topology. J Biol Chem. 2012; 287:26328–26341. https://doi.org/10.1074/jbc.M112.361790 PMID: 22661711
9. Chatagnon A, Veber P, Morin V, Bedo J, Triqueneaux G, Semon M et al. RAR/RXR binding dynamics distinguish pluripotency from differentiation associated cis-regulatory elements. Nucleic Acids Res. 2015; 43:4833–4854. https://doi.org/10.1093/nar/gkv370 PMID: 25897113
10. Dupé V, Davenne M, Brocard J, Dollé P, Mark M, Dierich A et al. In vivo functional analysis of the *Hoxa-1* 3' retinoic acid response element (3'RARE). Development. 1997; 124:399–410. PMID: 9053316
11. Houle M, Sylvestre JR, Lohnes D. Retinoic acid regulates a subset of Cdx1 function in vivo. Development. 2003; 130:6555–6567. https://doi.org/10.1242/dev.00889 PMID: 14660544
12. Nishimoto S, Wilde SM, Wood S, Logan MP. RA Acts in a Coherent Feed-Forward Mechanism with Tbx5 to Control Limb Bud Induction and Initiation. Cell Reports. 2015; 12:879–891. https://doi.org/10.1016/j.celrep.2015.06.068 PMID: 26212321
13. Cunningham TJ, Lancman JJ, Berenguer M, Dong PDS, Duester G. Genomic knockout of two presumed forelimb Tbx5 enhancers reveals they are nonessential for limb development. Cell Reports. 2018; 23:3146–3151. https://doi.org/10.1016/j.celrep.2018.05.052 PMID: 29898387
14. Duester G. Knocking out enhancers to enhance epigenetic research. Trends Genet. 2019; 35:89. https://doi.org/10.1016/j.tig.2018.10.001
15. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011; 470:279–283. https://doi.org/10.1038/nature09692 PMID: 21160473
16. Laugesen A, Helin K. Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell. 2014; 14:735–751. https://doi.org/10.1016/j.stem.2014.05.006 PMID: 24905164

17. Kumar S, Duester G. Retinoic acid controls body axis extension by directly repressing *Fgf8* transcription. Development. 2014; 141:2972–2977. https://doi.org/10.1242/dev.112367 PMID: 25053430

18. Ribes V, Le Roux I, Rhinn M, Schuhbaur B, Dolle P. Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. Development. 2009; 136:665–676. https://doi.org/10.1242/dev.016204 PMID: 19168680

19. Cunningham TJ, Brade T, Sandell LL, Lewandoski M, Trainor PA, Colas A et al. Retinoic acid activity in undifferentiated neural progenitors Is sufficient to fulfill Its role in restricting Fgf8 expression for somitogenesis. PLoS ONE. 2015; 10:e0137894. https://doi.org/10.1371/journal.pone.0137894 PMID: 26368825

20. Cunningham TJ, Colas A, Duester G. Early molecular events during retinoic acid induced differentiation of neuromesodermal progenitors. Biology Open. 2016; 5:1821–1833. https://doi.org/10.1242/bio.020891 PMID: 27793834

21. Niederreither K, Dolle P. Retinoic acid in development: towards an integrated view. Nature Rev Genet. 2008; 9:541–553. https://doi.org/10.1038/nrg2340 PMID: 18542081

22. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485:376–380. https://doi.org/10.1038/nature11082 PMID: 22495300

23. Mendelsohn C, Ruberte E, LeMeur M, Morriss-Kay G, Chambon P. Developmental analysis of the retinoic acid-inducible RAR-b2 promoter in transgenic animals. Development. 1991; 113:723–734. PMID: 1668276

24. Durand B, Saunders M, Leroy P, Leid M, Chambon P. All-trans and 9-cis retinoic acid induction of CRABPII transcription is mediated by RAR-RXR heterodimers bound to DR1 and DR2 repeated motifs. Cell. 1992; 71:73–85. https://doi.org/10.1016/0092-8674(92)90267-g PMID: 1327537

25. Gaunt SJ, Paul YL. Origins of Cdx1 regulatory elements suggest roles in vertebrate evolution. Int J Dev Biol. 2011; 55: 93–98. https://doi.org/10.1387/ijdb.103252sg PMID: 21425084

26. Edri S, Hayward P, Jawaid W, Martinez Arias A. Neuro-mesodermal progenitors (NMPs): a comparative study between pluripotent stem cells and embryo-derived populations. Development. 2019; 146:24.

27. Wilson V, Olivera-Martinez I, Storey KG. Stem cells, signals and vertebrate body axis extension. Development. 2009; 136:1591–1604. https://doi.org/10.1242/dev.021246 PMID: 19395637

28. Kondoh H, Takemoto T. Axial stem cells deriving both posterior neural and mesodermal tissues during gastrulation. Curr Opin Genet Dev. 2012; 22:374–380. https://doi.org/10.1016/j.gde.2012.03.006 PMID: 22575683

29. Henrique D, Abranches E, Verrier L, Storey KG. Neuromesodermal progenitors and the making of the spinal cord. Development. 2015; 142:2864–2875. https://doi.org/10.1242/dev.119768 PMID: 26329597

30. Kimelman D. Tales of tails (and trunks): forming the posterior body in vertebrate embryos. Curr Top Dev Biol. 2016; 116:517–536. https://doi.org/10.1016/bs.ctdb.2015.12.008 PMID: 26970638

31. Koch F, Scholze M, Wittler L, Schifferl D, Sudheer S, Grote P et al. Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. Dev Cell. 2017; 42:514–526. https://doi.org/10.1016/j.devcel.2017.07.021 PMID: 28826820

32. Gouti M, Delile J, Stamataki D, Wymeersch FJ, Huang Y, Kleinjung J et al. A gene regulatory network balances neural and mesoderm specification during vertebrate trunk development. Dev Cell. 2017; 41:243–261. https://doi.org/10.1016/j.devcel.2017.04.002 PMID: 28457792

33. Amin S, Neijts R, Simmini S, van Rooijen C, Tan SC, Kester L et al. Cdx and T brachyury co-activate growth signaling in the embryonic axial progenitor niche. Cell Reports. 2016; 17:3165–3177. https://doi.org/10.1016/j.celrep.2016.11.069 PMID: 28009287

34. Olivera-Martinez I, Harada H, Halley PA, Storey KG. Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. PLoS Biol. 2012; 10:e1001415. https://doi.org/10.1371/journal.pbio.1001415 PMID: 23118616

35. Tsakiridis A, Huang Y, Blin G, Skylaki S, Wymeersch F, Osorno R et al. Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. Development. 2014; 141:1209–1221. https://doi.org/10.1242/dev.101014 PMID: 24595287

36. Garriock RJ, Chalamalasetty RB, Kennedy MW, Canizales LC, Lewandoski M, Yamaguchi TP. Lineage tracing of neuromesodermal progenitors reveals novel Wnt-dependent roles in trunk progenitor cell maintenance and differentiation. Development. 2015; 142:1628–1638. https://doi.org/10.1242/dev.111922 PMID: 25922526

37. Martin BL, Kimelman D. Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. Dev Cell. 2012; 22:223–232. https://doi.org/10.1016/j.devcel.2011.11.001 PMID: 22264734

38. Naiche LA, Holder N, Lewandoski M. FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. Proc Natl Acad Sci USA. 2011; 108:4018–4023. https://doi.org/10.1073/pnas.1007417108 PMID: 21368122

39. Jurberg AD, Aires R, Novoa A, Rowland JE, Mallo M. Compartment-dependent activities of Wnt3a/beta-catenin signaling during vertebrate axial extension. Dev Biol. 2014; 394:253–263. https://doi.org/10.1016/j.ydbio.2014.08.012 PMID: 25152336

40. Takemoto T, Uchikawa M, Yoshida M, Bell DM, Lovell-Badge R, Papaioannou VE et al. Tbx6-dependent *Sox2* regulation determines neural or mesodermal fate in axial stem cells. Nature. 2011; 470:394–398. https://doi.org/10.1038/nature09729 PMID: 21331042

41. Wymeersch FJ, Huang Y, Blin G, Cambray N, Wilkie R, Wong FC et al. Position-dependent plasticity of distinct progenitor types in the primitive streak. eLife. 2016; 5:e10042. https://doi.org/10.7554/eLife.10042 PMID: 26780186

42. Diez del Corral R, Olivera-Martinez I, Goriely A, Gale E, Maden M, Storey K. Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. Neuron. 2003; 40:65–79. https://doi.org/10.1016/s0896-6273(03)00565-8 PMID: 14527434

43. Patel NS, Rhinn M, Semprich CI, Halley PA, Dolle P, Bickmore WA et al. FGF Signalling Regulates Chromatin Organisation during Neural Differentiation via Mechanisms that Can Be Uncoupled from Transcription. PLoS Genet. 2013; 9:e1003614. https://doi.org/10.1371/journal.pgen.1003614 PMID: 23874217

44. Zhao X, Duester G. Effect of retinoic acid signaling on Wnt/beta-catenin and FGF signaling during body axis extension. Gene Expr Patterns. 2009; 9:430–435. https://doi.org/10.1016/j.gep.2009.06.003 PMID: 19539783

45. Oosterveen T, Kurdija S, Enstero M, Uhde CW, Bergsland M, Sandberg M et al. SoxB1-driven transcriptional network underlies neural-specific interpretation of morphogen signals. Proc Natl Acad Sci USA. 2013; 110:7330–7335. https://doi.org/10.1073/pnas.1220010110 PMID: 23589857

46. Joshi P, Darr AJ, Skromne I. CDX4 regulates the progression of neural maturation in the spinal cord. Dev Biol. 2019; 449:132–142. https://doi.org/10.1016/j.ydbio.2019.02.014 PMID: 30825428

47. Lin FJ, Qin J, Tang K, Tsai SY, Tsai MJ. Coup d'Etat: an orphan takes control. Endocrine Reviews. 2011; 32:404–421. https://doi.org/10.1210/er.2010-0021 PMID: 21257780

48. Schulte D, Geerts D. MEIS transcription factors in development and disease. Development. 2019; 146:15.

49. Penkov D, San Martin DM, Fernandez-Diaz LC, Rossello CA, Torroja C, Sanchez-Cabo F et al. Analysis of the DNA-Binding Profile and Function of TALE Homeoproteins Reveals Their Specialization and Specific Interactions with Hox Genes/Proteins. Cell Reports. 2013; 3:1321–1333. https://doi.org/10.1016/j.celrep.2013.03.029 PMID: 23602564

50. Ishibashi T, Usami T, Fujie M, Azumi K, Satoh N, Fujiwara S. Oligonucleotide-based microarray analysis of retinoic acid target genes in the protochordate, *Ciona intestinalis*. DevDyn. 2005; 233:1571–1578.

51. Laursen KB, Mongan NP, Zhuang Y, Ng MM, Benoit YD, Gudas LJ. Polycomb recruitment attenuates retinoic acid-induced transcription of the bivalent NR2F1 gene. Nucleic Acids Res. 2013; 41:6430–6443. https://doi.org/10.1093/nar/gkt367 PMID: 23666625

52. Dohn TE, Ravisankar P, Tirera FT, Martin KE, Gafranek JT, Duong TB et al. Nr2f-dependent allocation of ventricular cardiomyocyte and pharyngeal muscle progenitors. PLoS Genet. 2019; 15:e1007962. https://doi.org/10.1371/journal.pgen.1007962 PMID: 30721228

53. Zhuang Y, Gudas LJ. Overexpression of COUP-TF1 in murine embryonic stem cells reduces retinoic acid-associated growth arrest and increases extraembryonic endoderm gene expression. Differentiation. 2008; 76:760–771. https://doi.org/10.1111/j.1432-0436.2007.00258.x PMID: 18177425

54. Mercader N, Leonardo E, Piedra ME, Martínez-A C, Ros MA, Torres M. Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of Meis genes. Development. 2000; 127:3961–3970. PMID: 10952894

55. Niederreither K, Vermot J, Schuhbaur B, Chambon P, Dollé P. Retinoic acid synthesis and hindbrain patterning in the mouse embryo. Development. 2000; 127:75–85. PMID: 10654602

56. Kashyap V, Laursen KB, Brenet F, Viale AJ, Scandura JM, Gudas LJ. RARg is essential for retinoic acid induced chromatin remodeling and transcriptional activation in embryonic stem cells. J Cell Sci. 2013; 126:999–1008. https://doi.org/10.1242/jcs.119701 PMID: 23264745

57. Lalevee S, Anno YN, Chatagnon A, Samarut E, Poch O, Laudet V et al. Genome-wide in Silico Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 bp). J Biol Chem. 2011; 286:33322–33334. https://doi.org/10.1074/jbc.M111.263681 PMID: 21803772

58. Will AJ, Cova G, Osterwalder M, Chan WL, Wittler L, Brieske N et al. Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). Nature Genet. 2017; 49:1539–1545. https://doi.org/10.1038/ng.3939 PMID: 28846100

59. Osterwalder M, Barozzi I, Tissieres V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY et al. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature. 2018; 554:239–243. https://doi.org/10.1038/nature25461 PMID: 29420474

60. Dickel DE, Ypsilanti AR, Pla R, Zhu Y, Barozzi I, Mannion BJ et al. Ultraconserved enhancers are required for normal development. Cell. 2018; 172:491–499. https://doi.org/10.1016/j.cell.2017.12.017 PMID: 29358049

61. Béland M, Lohnes D. Chicken ovalbumin upstream promoter-transcription factor members repress retinoic acid-induced *Cdx1* expression. J Biol Chem. 2005; 280:13858–13862. https://doi.org/10.1074/jbc.M412981200 PMID: 15677473

62. Vilhais-Neto GC, Maruhashi M, Smith KT, Vasseur-Cognet M, Peterson AS, Workman JL et al. Rere controls retinoic acid signalling and somite bilateral symmetry. Nature. 2010; 463:953–957. https://doi.org/10.1038/nature08763 PMID: 20164929

63. Qiu Y, Pereira FA, DeMayo FJ, Lydon JP, Tsai SY, Tsai MJ. Null mutation of mCOUP-TFI results in defects in morphogenesis of the glossopharyngeal ganglion, axonal projection, and arborization. Genes Dev. 1997; 11:1925–1937. https://doi.org/10.1101/gad.11.15.1925 PMID: 9271116

64. Pereira FA, Qiu YH, Zhou G, Tsai MJ, Tsai SY. The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. Genes Dev. 1999; 13:1037–1049. https://doi.org/10.1101/gad.13.8.1037 PMID: 10215630

65. Vermot J, Llamas JG, Fraulob V, Niederreither K, Chambon P, Dollé P. Retinoic acid controls the bilateral symmetry of somite formation in the mouse embryo. Science. 2005; 308:563–566. https://doi.org/10.1126/science.1108363 PMID: 15731404

66. Hisa T, Spence SE, Rachel RA, Fujita M, Nakamura T, Ward JM et al. Hematopoietic, angiogenic and eye defects in *Meis1* mutant animals. EMBO Journal. 2004; 23:450–459. https://doi.org/10.1038/sj.emboj.7600038 PMID: 14713950

67. Machon O, Masek J, Machonova O, Krauss S, Kozmik Z. Meis2 is essential for cranial and cardiac neural crest development. BMC Dev Biol. 2015; 15:40. https://doi.org/10.1186/s12861-015-0093-6 PMID: 26545946

68. Zhao X, Sirbu IO, Mic FA, Molotkova N, Molotkov A, Kumar S et al. Retinoic acid promotes limb induction through effects on body axis extension but is unnecessary for limb patterning. Curr Biol. 2009; 19:1050–1057. https://doi.org/10.1016/j.cub.2009.04.059 PMID: 19464179

69. Perissi V, Aggarwal A, Glass CK, Rose DW, Rosenfeld MG. A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. Cell. 2004; 116:511–526. https://doi.org/10.1016/s0092-8674(04)00133-3 PMID: 14980219

70. Begemann G, Schilling TF, Rauch GJ, Geisler R, Ingham PW. The zebrafish *neckless* mutation reveals a requirement for *raldh2* in mesodermal signals that pattern the hindbrain. Development. 2001; 128:3081–3094. PMID: 11688558

71. Berenguer M, Lancman JJ, Cunningham TJ, Dong PDS, Duester G. Mouse but not zebrafish requires retinoic acid for control of neuromesodermal progenitors and body axis extension. Dev Biol. 2018; 441:127–131. https://doi.org/10.1016/j.ydbio.2018.06.019 PMID: 29964026

72. Cooper KL, Hu JK, ten Berge D, Fernandez-Teran M, Ros MA, Tabin CJ. Initiation of proximal-distal patterning in the vertebrate limb by signals and growth. Science. 2011; 332:1083–1086. https://doi.org/10.1126/science.1199499 PMID: 21617075

73. Rosello-Diez A, Ros MA, Torres M. Diffusible signals, not autonomous mechanisms, determine the main proximodistal limb subdivision. Science. 2011; 332:1086–1088. https://doi.org/10.1126/science.1199489 PMID: 21617076

74. Cunningham TJ, Chatzi C, Sandell LL, Trainor PA, Duester G. *Rdh10* mutants deficient in limb field retinoic acid signaling exhibit normal limb patterning but display interdigital webbing. Dev Dyn. 2011; 240:1142–1150. https://doi.org/10.1002/dvdy.22583 PMID: 21360789

75. Cunningham TJ, Zhao X, Sandell LL, Evans SM, Trainor PA, Duester G. Antagonism between retinoic acid and fibroblast growth factor signaling during limb development. Cell Reports. 2013; 3:1503–1511. https://doi.org/10.1016/j.celrep.2013.03.036 PMID: 23623500

76. Voss AK, Dixon MP, McLennan T, Kueh AJ, Thomas T. Chromatin immunoprecipitation of mouse embryos. In: Vancura A, editor. *Transcriptional Regulation*: *Methods and Protocols*, *Methods in Molecular Biology*. New York: Humana Press; 2012. p. 335–352.

77. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. 2011; 17.

78. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21. https://doi.org/10.1093/bioinformatics/bts635 PMID: 23104886

79. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010; 38:576–589. https://doi.org/10.1016/j.molcel.2010.05.004 PMID: 20513432

80. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281

81. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995; 57:289–300.

82. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Cell. 2013; 153:910–918. https://doi.org/10.1016/j.cell.2013.04.025 PMID: 23643243

83. Tan EP, Li Y, Del Castillo Velasco-Herrera M, Yusa K, Bradley A. Off-target assessment of CRISPR-Cas9 guiding RNAs in human iPS and mouse ES cells. Genesis. 2015; 53:225–236. https://doi.org/10.1002/dvg.22835 PMID: 25378133

84. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature Methods. 2012; 9:671–675. https://doi.org/10.1038/nmeth.2089 PMID: 22930834

85. Sirbu IO, Duester G. Retinoic acid signaling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. Nature Cell Biol. 2006; 8:271–277. https://doi.org/10.1038/ncb1374 PMID: 16489341