



Published in final edited form as:

Cell Rep. 2024 March 26; 43(3): 113907. doi:10.1016/j.celrep.2024.113907.

Deletion of *Pax1* scoliosis-associated regulatory elements leads to a female-biased tail abnormality

Aki Ushiki^{1,2,8}, Rory R. Sheng^{1,2,8}, Yichi Zhang^{1,2,3}, Jingjing Zhao^{1,2}, Mai Nobuhara^{1,2}, Elizabeth Murray^{1,2}, Xin Ruan^{1,2}, Jonathan J. Rios^{4,5,6,7}, Carol A. Wise^{4,5,6,7}, Nadav Ahituv^{1,2,9,*}

¹Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA 94158, USA

²Institute for Human Genetics, University of California San Francisco, San Francisco, CA 94158, USA

³School of Pharmaceutical Sciences, Tsinghua University, Beijing 100084, China

⁴Center for Translational Research, Scottish Rite for Children, Dallas, TX 75390, USA

⁵Department of Orthopedic Surgery, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁶Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁷Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁸These authors contributed equally

⁹Lead contact

Abstract

Adolescent idiopathic scoliosis (AIS), a sideways curvature of the spine, is sexually dimorphic, with increased incidence in females. A genome-wide association study identified a female-specific AIS susceptibility locus near the *PAX1* gene. Here, we use mouse enhancer assays, three mouse enhancer knockouts, and subsequent phenotypic analyses to characterize this region. Using mouse enhancer assays, we characterize a sequence, PEC7, which overlaps the AIS-associated variant, and find it to be active in the tail tip and intervertebral disc. Removal of PEC7 or Xe1, a known sclerotome enhancer nearby, or deletion of both sequences lead to a kinky tail phenotype only in the Xe1 and combined (Xe1+PEC7) knockouts, with only the latter showing a female

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

*Correspondence: nadav.ahituv@ucsf.edu.

AUTHOR CONTRIBUTIONS

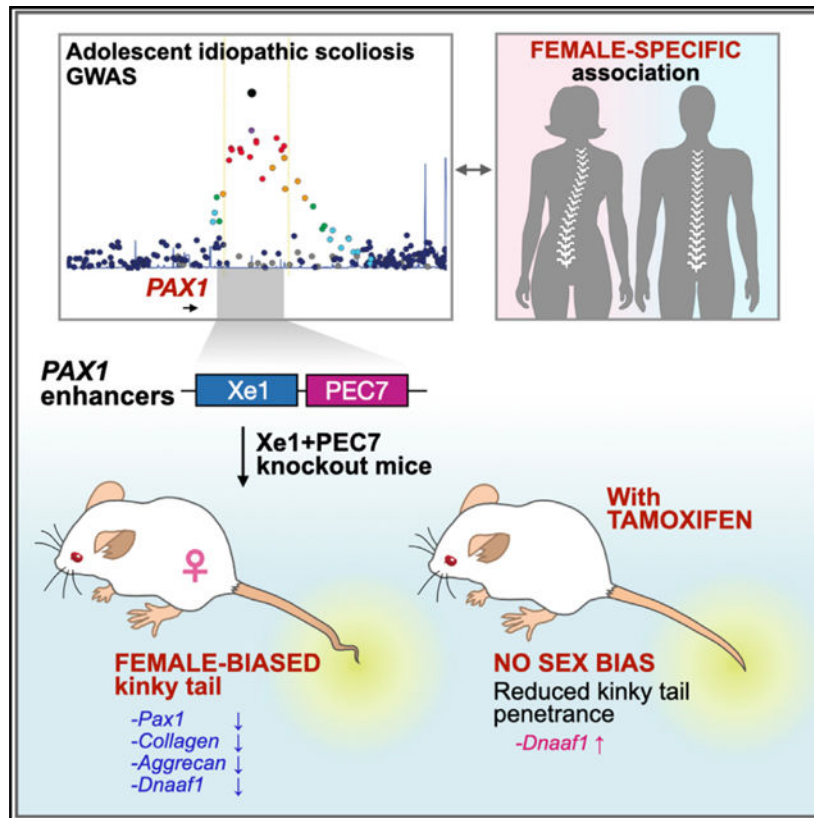
Conceptualization, A.U., R.R.S., and N.A.; methodology, A.U., R.R.S., Y.Z., J.Z., M.N., E.M., and X.R.; samples, A.U., R.R.S., Y.Z., J.Z., M.N., E.M., and X.R.; writing, review, and editing, A.U., R.R.S., J.J.R., C.A.W., and N.A., with contributions from all authors; visualization, A.U., R.R.S., and N.A.; supervision, N.A.; funding acquisition, N.A. and C.A.W.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2024.113907>.

sex dimorphic phenotype. Extensive phenotypic characterization of these mouse lines implicates several differentially expressed genes and estrogen signaling in the sex dimorphic bias. In summary, our work functionally characterizes an AIS-associated locus and dissects the mechanism for its sexual dimorphism.

Graphical Abstract



In brief

Ushiki et al. use mouse enhancer assays and knockouts to characterize a genomic region associated with adolescent idiopathic scoliosis (AIS) in females, finding that its removal leads to a female-biased kinky tail phenotype, which is reversed with tamoxifen administration, suggesting a role for estrogen signaling in AIS.

INTRODUCTION

Adolescent idiopathic scoliosis (AIS), characterized by a lateral curvature of the spine that occurs during early adolescence along with spine growth, affects ~3% of the population worldwide.¹ AIS is a sexually dimorphic disease.² Girls have approximately 10-fold higher risk of developing progressive curves that require operative treatment.³ AIS is caused by both environmental and genetic factors. A study of monozygotic and dizygotic twins from the Swedish Twin Registry estimated that the overall genetic effects for AIS accounted for 38% of the observed phenotypic variance.⁴ However, the genetic basis and pathogenic

mechanism of AIS have remained largely unknown. Genome-wide association studies (GWASs) have identified more than a dozen genetic loci associated with AIS, with the majority located in noncoding regions of the genome.^{5–14} These AIS-associated single-nucleotide polymorphisms (SNPs) potentially overlap gene regulatory sequences, such as enhancers, and could alter their target gene expression.¹⁵

In one of these GWASs, we identified an AIS susceptibility locus that was associated only with females, downstream of the paired box 1 transcription factor (*PAX1*), a transcription factor involved in sclerotome development.¹⁰ *Pax1* is expressed in the developing spine in mouse embryos. Both heterozygous and homozygous *Pax1* knockout (KO) mice exhibit a kinky tail phenotype^{16,17} and various spinal malformations, including scoliosis.^{18,19} *PAX1* regulates extracellular matrix (ECM) genes, such as collagen and aggrecan, and is crucial for mesenchyme condensation and intervertebral disc (IVD) development.^{20,21} Despite the growing understanding of the function of PAX1 in spine development, how it leads to AIS remains largely unknown.

Several *Pax1*-associated gene regulatory elements have been characterized. One such element, Xe1, which resides near the AIS-associated GWAS SNP (Figure 1A), was found to drive enhancer activity in the mouse sclerotome.²² Additionally, using zebrafish enhancer assays, our lab functionally characterized ten enhancer candidate sequences in the *PAX1* AIS-associated locus, termed *PAX1* enhancer candidates (PECs).¹⁰ These PECs were chosen based on their evolutionary conservation and/or having enhancer-associated marks in ENCODE datasets.²³ Only two of these sequences drive enhancer activity in the developing zebrafish spine and somitic muscle, the previously characterized Xe1 and PEC7, which is in close vicinity (1,036 base pairs) to Xe1 (Figure 1A). Of the two, only PEC7 harbors an AIS GWAS-associated variant, rs169311, which is in strong linkage disequilibrium with the AIS lead SNP rs6137473. Introducing the AIS-associated SNP rs169311 into PEC7 abolishes its enhancer activity in zebrafish,¹⁰ suggesting that rs169311 downregulates *PAX1* gene expression in the developing somites. However, the mechanisms of how this SNP alters PEC7 function and how this altered function leads to AIS have not been characterized.

In this study, we further characterized the role of this female AIS-associated region using multiple mouse models. Using transgenic mouse enhancer assays, we found PEC7 to be active in neonate tail tip region and adult IVD. We next generated three putative *Pax1* enhancer KO mouse lines—Xe1, PEC7, and both Xe1 and PEC7 (Xe1+PEC7)—using CRISPR-Cas9 genome editing. We found that approximately 20% of Xe1 homozygous KO mice have a kinky tail phenotype. This phenotype becomes more severe in Xe1+PEC7 mice, and, intriguingly, it has a higher penetrance in Xe1+PEC7 KO females (60%) than males (42%). To characterize the mechanism leading to the kinky tail phenotype, we performed RNA sequencing (RNA-seq) on embryonic day 12.5 (E12.5) tails, finding ECM genes, including collagen- and aggrecan-encoding genes, to be downregulated in all three homozygous KO mice. To further characterize the observed sexual dimorphic phenotype, we injected tamoxifen, an estrogen receptor (ESR) antagonist, into E17.5 Xe1+PEC7 homozygous pregnant females and found that the kinky tail sex ratio is reduced to similar levels to the Xe1 KO mice (25% in females and 27% in males), suggesting that estrogen could be involved in this sex bias. RNA-seq analysis of postnatal day 2 (P2) tail samples

of tamoxifen-injected versus un-injected Xe1+PEC7 mice identifies genes involved in the reversal of this sexual dimorphic kinky tail phenotype. Combined, our results show that PEC7 functions as an enhancer in the neonate tail tip region and adult IVD and contributes to the onset and progression of a sex-differential kinky tail phenotype, likely due to the estrogen signaling pathway.

RESULTS

PEC7 shows enhancer activity in the distal part of the neonatal tail and adult IVD

We first set out to test the mouse enhancer activity of PEC7, as it encompasses rs169311, an AIS GWAS-associated variant that abolishes its somite enhancer activity when tested in zebrafish.^{10,24} We cloned mouse PEC7 DNA fragment into the Hsp68-LacZ vector, upstream of an *Hsp68* minimal promoter followed by the mCherry reporter gene.²⁵ This plasmid was injected into mouse zygotes, and two independent transgenic mouse lines were established (line 5339 and line 5358). We checked whole-mount mCherry fluorescence at E12.5, E15.5, E18.5, and P0–P4 and observed comparable enhancer activity in both lines (Figures 1B and S1A). From E18.5, mCherry was observed in the nails, followed by the distal part of the tail in neonatal stages (P0–P3), with P1–P2 showing the highest mCherry fluorescence, which gradually decreases in P3 and P4 (Figures 1B and S1B). We also quantified mCherry expression in P2 tail by qPCR and found that female mice show higher mCherry expression (Figure S1C). Analysis of mCherry fluorescence in 10-week-old mice observed fluorescence in the IVD but not in skeletal muscle (Figures 1C and S1D). These results suggest that PEC7 has enhancer activity in the distal part of the neonate tail and adult IVD.

Xe1+PEC7 homozygous mice have a kinky tail with altered sex penetrance

To characterize the function of this female-associated AIS locus, we generated various mouse KOs for this region. PEC7 is located nearby Xe1 (the distance between them is 1 kb in human and 4.3 kb in the Friend Virus B NIH Jackson (FVB/NJ) mouse strain; Figure 1A), a previously characterized sclerotome enhancer,²² and might be associated with its regulation. Therefore, we generated three different deletion lines: (1) PEC7 only, (2) Xe1 only, and (3) PEC7 and Xe1. We first identified the conserved region between humans and mice using the Ensemble genome browser,²⁴ finding both Xe1 and PEC7 to be sufficiently conserved between humans and mice (Figures 1A and S2A–S2C), including the region around the AIS-associated SNP (rs169311) but not the SNP itself (Figure S2A). We designed three gRNA targeting the 5' of Xe1, the 3' of Xe1, and the 3' of PEC7 (Figure 1A) and independently injected a pair of gRNA (5' and 3' of Xe1, 3' of Xe1 and 5' of PEC7, and 5' of Xe1 and 3' of PEC7) along with Cas9 protein into zygotes. We obtained founder lines for all three manipulations, which were further validated by PCR, sequencing, and Southern blot analyses (Figure S3).

In homozygous Xe1 and Xe1+PEC7 KO mouse lines, we observed kinks in the distal part of the tail due to a bent caudal vertebra, as observed in micro-computed tomography (micro-CT) (Figures 2A and 2B). This kinky tail phenotype has partial penetrance in Xe1 and Xe1+PEC7 lines (Figures 2C; Table S1). Homozygous Xe1 mice have around 20%

penetrance of kinky tails in both females and males. Interestingly, homozygous Xe1+PEC7 mice display higher penetrance with a significant sex difference, with 60% Xe1+PEC7 females and 42% males having kinky tails (Figures 2C; Table S1). To test whether these mice have any additional skeletal abnormalities, we carried out micro-CT at both 4 and 6 months of age for all genotypes, including mice with and without a kinky tail. We did not observe any other apparent skeletal abnormalities other than the kinky tail (Figures 2D and S4). We also checked the kinky tail progression daily after birth (P0–P20; Figure 2E) and found that the tail curvature starts from P0, with bending progressing to the distal part of the tail at later stages (>P7). This is consistent with the PEC7 enhancer activity observed in the neonatal stage (Figures 1B and S1A). In summary, our results suggest that deletion of Xe1 leads to a partially penetrant kinky tail phenotype that is intensified both in terms of penetrance and sex difference when the PEC7 region is deleted along with Xe1.

Xe1 and Xe1+PEC7 homozygous mice have reduced *Pax1* embryonic tail expression

We next set out to test the expression of *Pax1* in all three lines. As *Pax1* has been reported to be highly expressed at E10.5–E13.5 in the tail,^{16,17} we first analyzed its expression in this tissue. We carried out RT-qPCR and whole-mount *in situ* hybridization (WISH) on E12.5 embryos. We observed, via RT-qPCR, a ~50% down-regulation of *Pax1* expression in Xe1 and Xe1+PEC7 homozygous mice but no significant changes in expression for the PEC7 homozygous mice compared to wild-type (WT) mice (Figure 2F). Consistent with the RT-qPCR data, WISH also shows reduction of *Pax1* gene expression in the tail for both these homozygous lines (Figure 2G). To determine whether *Pax1* downregulation occurs at the initiation of its expression in skeletal development,¹⁶ E10.5, we carried out RT-qPCR on E10.5 tails, observing significant downregulation of *Pax1* in Xe1+PEC7 homozygous embryos (Figure S5A). We also checked *Pax1* gene expression in 10-week-old thymus and IVD, as it has been reported that *Pax1* is highly expressed in these tissues.²⁶ PEC7 homozygous female and Xe1 male show slightly higher expression compared to WT mice in thymus; however, other lines did not show significant *Pax1* gene expression differences in these tissues (Figures S5B–S5C).

RNA-seq identifies ECM genes to be downregulated in enhancer KO mice

To identify downstream gene expression changes in the various enhancer KOs, we performed RNA-seq on E12.5 embryonic tails for all three homozygous KOs (Xe1, PEC7, Xe1+PEC7) and WT mice, each in three females and three males, for a total of 24 samples. We chose E12.5 because *Pax1* is strongly expressed at the tail at this time point.^{16,17} We first identified differentially expressed genes between the transcriptomes of male and female mice between WT mice and the various mouse lines using the Gene Ontology (GO) term enrichment analysis that is part of the Partek genomics suite (see STAR Methods). We found a strong enrichment for “embryonic skeletal system morphogenesis,” “muscle tissue development,” “cell adhesion,” and “collagen fibril organization” in females (Figure 3A). Xe1+PEC7 females showed the highest association with embryonic skeletal system morphologies.

We next set out to identify differentially expressed genes between males and females, specifically in the Xe1+PEC7 homozygous KOs, to identify candidate genes that might

be responsible for the sex differences observed for the kinky tail penetrance. Analysis of differentially expressed genes between males and females identified only sex-chromosome differential genes (Figure S6A). Next, we compared the gene expression differences between WT and Xe1+PEC7 in both sexes separately. We observed significant downregulation of *Pax1* in the Xe1+PEC7 KO mice as expected (Figure 3B). In addition, we saw downregulation of ECM genes (*Acan*, *Coll1a1*, *Coll1a2*, and *Coll14a1*) and protocadherin genes involved in cell adhesion (*Pcdhb16* and *Pcdhb20*) (Figure 3B). Conversely, we observed upregulation of *Pax9* and *Foxd2os* in both sexes (Figure 3B). *Pax9* and *Pax1* have redundant functions in axial skeletogenesis,^{16,18,27} and *Pax9* expression was shown to be upregulated to compensate for *Pax1* expression in *Pax1* KO mice,²⁷ fitting with our observation of lower *Pax1* expression in our RT-qPCR and RNA-seq. *FOXD2-AS1* (human ortholog of *Foxd2os*) is known to be highly expressed and regulates chondrocyte proliferation in patients with osteoarthritis.^{28,29} We next compared the gene expression changes between Xe1 and Xe1+PEC7 homozygous female mice, as Xe1 homozygous mice do not have a kinky tail sex bias phenotype. We found *Sox2* and *Sox3* to be downregulated in Xe1+PEC7 KO females compared to Xe1 KO females (Figure 3C). *Sox2* and *Sox3* are known to have a redundant function in the development of otic and epibranchial tissue,³⁰ and *Sox2* is involved in mesoderm differentiation in zebrafish tail buds.³¹

To validate our RNA-seq results and analyze specific genotypes in more detail, we performed RT-qPCR using at least five mice for all genotypes and sex. We first confirmed that the ECM genes (*Acan*, *Coll1a1*, *Coll1a2*, and *Coll14a1*) are downregulated in all homozygous KO mouse lines compared to WT (Figure 3D). Interestingly, in the PEC7 homozygous KO mice, we also observed downregulation of *Acan*, *Coll1a2*, and *Coll14a1* (Figure 3D), despite not observing any expression changes in *Pax1* in these mice (Figure 2F). *Pax9* and *Foxd2os* are both upregulated in Xe1 and Xe1+PEC7 homozygous KO mice. Interestingly, upregulation of both these genes is higher in Xe1+PEC7 mice compared to Xe1 (Figure 3D), although *Pax1* expression is similar in these mice (Figure 2F). In PEC7 homozygous mice, we only observed *Foxd2os* to be significantly upregulated in females. Combined, these results suggest that Xe1 and PEC7 might work synergistically and that the deletion of PEC7 might affect the expression of other genes that are not directly regulated by *Pax1*.

Tamoxifen injection in Xe1+PEC7 pregnant mothers abolishes kinky tail sex difference

As this region is associated with AIS only in females,¹⁰ and as female Xe1+PEC7 KO mice show higher kinky tail penetrance, we hypothesized that estrogen could be involved in this process, based on previous reports associating it with AIS.³² Previous work showed that tamoxifen, a selective ESR modulator, decreases the rate of curve progression in a melatonin-deficient bipedal scoliosis mouse model.³³ We thus injected tamoxifen into Xe1+PEC7 KO E17.5 pregnant females and compared their kinky tail phenotypic ratio of pups at day 21 to un-injected mice. Interestingly, we observed significantly reduced penetrance for the kinky tail phenotype and that the previous sex difference for Xe1+PEC7 KO is ablated, showing an almost equal number of males and females with a kinky tail, 25% in females and 27% in males (Figure 4A; Table S1) compared to 60% and 42% in noninjected mice (Figures 2C; Table S1).

To further characterize the genes involved in this process, we conducted RNA-seq using the distal part of the P2 tail tip, due to it being the time point and location where the kinky tail initiates (Figure 2E) and also the tamoxifen-injected condition. We extracted RNA from WT, Xe1, PEC7, Xe1+PEC7, and Xe1+PEC7 pups from tamoxifen-injected dams (Xe1+PEC7(Tam)). We used four samples per condition and sex, totaling 40 samples for RNA-seq. In contrast to our E12.5 analyses, we observed no significant gene expression changes for *Pax1* between the different genotypes and sexes when compared to WT at this P2 time point (Figure S6B). These results suggest that the functional effect and activity of Xe1 and/or PEC7 are likely earlier. Next, we compared males and females to identify differential GO term enrichment. We found Xe1+PEC7-specific enrichment in females for “dynein heavy-chain binding” and “ciliary-based body organization” (Figure 4B). Intriguingly, KO of axonemal dynein and its assembly factor genes, such as *dnah10*, *dnaaf1*, and *zmynd10*, is known to lead to a scoliosis phenotype in zebrafish.^{34,35} Defects of motile cilia have also been shown to cause scoliosis in zebrafish.³⁶ Next, to identify differentially expressed genes in these processes, we compared gene expression differences between males and females in Xe1+PEC7 KO mice, finding only sex-chromosome-associated genes (Figure S6C). We next focused only on females and analyzed gene expression differences between the various genotypes, including WT versus Xe1+PEC7 and Xe1 versus Xe1+PEC7. Interestingly, *Foxj1* and *Efcab1*, which are involved with cilium function,^{37–39} were found to be upregulated in Xe1+PEC7 female mice compared to WT females (Figure 4C). We also found *Dnaaf1* to be downregulated in Xe1+PEC7 (Figures 4C and 4D). The dynein axonemal heavy chain family genes, such as *Dnah11* and *Dnah3*, are differentially expressed (Figures 4C and 4D). In addition, *Mmp9*, an enzyme involved in cartilage degradation during endochondral ossification,⁴⁰ is found to be downregulated in Xe1+PEC7 female mice compared to Xe1 female mice.

We next set out to assess the effects of tamoxifen on gene expression. We first compared Xe1+PEC7 female mice with or without tamoxifen. Intriguingly, *Dnaaf1*, *Dnah2*, and *Dnah3* expression levels are upregulated in the Xe1+PEC7(Tam) females (Figure 4E). These results suggest that dynein and cilia function could be associated with kinky tail progression in Xe1+PEC7 KO females. To further identify genes that could be involved in the sex difference phenotype in Xe+PEC7 mice, we characterized genes that are downregulated in Xe1+PEC7 KO mice compared to WT and upregulated in Xe1+PEC7(Tam) mice compared to un-injected mice. We identified 39 genes that were downregulated in Xe1+PEC7 KO and restored with tamoxifen injection specifically in females (Figures 4F and 4G). While we did not observe any specific GO enriched term for these genes, we did identify some interesting AIS-associated genes. They include the aforementioned *Dnaaf1*, the deletion of which causes a scoliosis phenotype in zebrafish³⁴; *Hoxa1*, a homeobox gene essential for the development of head and neck structures, including hindbrain, ear, and occipital and hyoid bones^{41,42}; *Mesp1*, a transcription factor involved in mesoderm specification, somite boundary formation, and somite polarity regulation^{43,44}; and *Tnfrsf14*, a receptor for *Tnfsf14*. *Tnfsf14* is primarily expressed in lymphocytes, as a critical regulator of key enzymes that control lipid metabolism,⁴⁵ and it resides in a region on chromosome 19p13.3 that has been linked to AIS.⁴⁶ In summary, our RNA-seq identified genes that show differential expression between the various genotypes and with or without tamoxifen

treatment, several of which are involved in ciliary or dynein function and could be involved in AIS sexual dimorphic phenotypes.

DISCUSSION

Using mouse transgenics and KOs, we characterized an AIS susceptibility locus downstream of *PAX1* that is associated with female AIS.¹⁰ We found that while both Xe1 and PEC7 function as enhancers, only removal of Xe1 leads to a kinky tail phenotype. Interestingly, though, removal of PEC7 along with Xe1 leads both to an increased penetrance of this phenotype and sexual dimorphism, with females showing a higher prevalence. Using RNA-seq, we extensively characterized the genes and pathways associated with this phenotype, finding genes involved in ECM, ciliary, or dynein function. Using tamoxifen injections in Xe1+PEC7 pregnant females, we show that the sex bias is likely associated with estrogen signaling. Combined, our results further dissect this female AIS-associated locus and identify candidate genes and pathways for AIS.

A kinky tail phenotype has been observed for several genes associated with scoliosis,⁴⁷ including *Wnt3a*,⁴⁸ *Axin1*,^{49,50} *Sox9*,⁵¹ and *Pax1*.¹⁹ In our RNA-seq data, we did not observe significant changes in the expression of these three genes. While *Pax1* gene KO mice also show lumbar scoliosis, we only observed a kinky tail in our enhancer KO mice. In addition, the kinky tail in the *Pax1* gene KO phenotype shows complete penetrance and is much more severe, with several bones affected throughout the tail.¹⁹ Our enhancer KO leads to a less severe phenotype with reduced penetrance, in line with previous enhancer KO phenotypes, which usually tend to show a more subtle phenotype than that of their target gene KO.⁵² This could be due to genes having multiple functions in various tissues versus an enhancer being tissue-/cell-type specific,⁵³ promoters having a stronger regulatory role than enhancers, and/or enhancer redundancy,^{54,55} i.e., having multiple enhancers for a specific target gene with similar function. This more subtle and not fully penetrant phenotype is also in line with AIS-associated GWAS variants being associated with AIS susceptibility, where having a specific variant along with environmental and/or other genetic factors combined lead to AIS.

Our mouse enhancer assays show that PEC7 functions as an enhancer in the distal part of the neonate tail and adult IVD. However, its removal does not lead to an observable phenotype, as assessed by micro-CT up to 6 months of age. In contrast, independent removal of Xe1 leads to a kinky tail phenotype with 20% penetrance, with no differences between males and females. In line with this locus being associated with AIS in females,¹⁰ we observed an increased penetrance of a kinky tail phenotype in females only in our Xe1+PEC7 homozygous mice. While Xe1 does not contain an AIS-associated variant, the homologous human PEC7 region does, suggesting that PEC7 could be the cause of the sexual dimorphism in this region. Interestingly, this locus has also been associated with male pattern baldness,^{56,57} suggesting that it can also affect male-specific phenotypes. Our tamoxifen injections into E17.5 pregnant Xe1+PEC7 mice significantly reduce the kinky tail penetrance in both males and even more in females, bringing it to similar levels as Xe1 KOs, suggesting that PEC7 might be regulated by estrogen signaling. Utilizing JASPAR,⁵⁸ we found that the mouse PEC7 region contains an *ESR2* motif (Figure 1A). However, this

motif is not conserved in humans, nor did we find other *ESR2* motifs in the human *PEC7* sequence. *PEC7* could be regulated by factors downstream to *ESR* and/or the androgen receptor. Further characterization of how this region leads to various sexual dimorphic phenotypes, including scoliosis or male pattern baldness, would be of extreme interest.

Our RNA-seq analyses in both E12.5 and P2 tails identified numerous genes and pathways that could be linked to AIS in general and to a sexual dimorphic phenotype. We observed several ECM genes (*Acan*, *Col11a1*, *Col11a2*, *Col14a1*, *Pcdhb16*, and *Pcdhb20*) to be significantly downregulated in both *Xe1+PEC7* male and female KOs. The ECM has been strongly associated with AIS, with many other AIS GWAS loci residing near ECM-associated genes and zebrafish ECM gene KOs having a scoliosis phenotype.⁵⁹ In terms of sexually dimorphic genes, we analyzed both *Xe1* versus *Xe1+PEC7* homozygous female mice, as *Xe1* homozygous did not have a kinky tail sex biased. We also analyzed tamoxifen-treated versus untreated *Xe1+PEC7* KO P2 tails, as tamoxifen treatment demolishes the kinky tail sex bias in these *Xe1+PEC7* KOs. In particular, we identified the dynein-associated gene *Dnaaf1*, which is known to cause scoliosis in zebrafish,³⁴ to be downregulated in *Xe1+PEC7* compared to *Xe1* female KO P2 tails. In addition, defects of motile cilia are known to cause scoliosis in zebrafish,³⁶ fitting with our observation of several cilia-associated genes. We also found *Sox2* and *Sox3*, which are known to be involved in the development of otic and epibranchial tissue³⁰ and mesoderm differentiation in zebrafish tail buds,³¹ to be downregulated in *Xe1+PEC7* homozygous female mice compared to *Xe1* homozygous female mice at E12.5 tails. In addition, for *SOX3*, altered gene regulation, due to an interchromosomal insertion downstream of this gene, is thought to lead to congenital generalized hypertrichosis along with spina bifida and scoliosis.⁶⁰ As their name implies, sex-determining region Y-box transcription factors have been associated with various sex dimorphic phenotypes. *SOX2*, for example, is known to be involved in sexual dimorphic differences in the peripheral nervous system,⁶¹ which has been shown to have a major role in scoliosis etiology.⁶² Interestingly, *Tnfrsf14*, which was found to be significantly downregulated in *Xe1+PEC7* homozygous KOs but not in the tamoxifen treatment, was found to be associated with sexual dimorphic expression in osteoprogenitors in progesterone receptor KO mice.⁶³ Gene regulatory elements were found to be a major cause of sex-associated gene expression⁶⁴ and sexual dimorphic phenotypes, such as inguinal hernia⁶⁵ or cancer.⁶⁶ Further work dissecting the sexual dimorphic function and mechanism of these genes and regulatory elements could shed light on AIS prevalence differences between males and females specifically and in general for other sexual dimorphic phenotypes.

Limitations of the study

We used mice to analyze how these regulatory sequences are associated with AIS. It is important to note that modeling AIS in mice had been limited.⁴⁷ This is likely due to mice being quadrupedal compared to bipedal humans, leading to inherent differences of their spinal anatomy. In addition, while the human-mouse sequences are conserved for both *Xe1* and *PEC7* (Figure S2), there are sequence differences that could potentially lead to differential function. While we observed sex-associated phenotypic differences that are

rescued by tamoxifen injections, suggesting they are associated with estrogen signaling, we did not identify the mechanism/pathway involved in these differences.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Nadav Ahituv (nadav.ahituv@ucsf.edu).

Materials availability—This study did not generate new unique reagents

Data and code availability

- RNA-seq data have been deposited at SRA and are publicly available as of the date of publication. Accession number NCBI: PRJNA951902
- This paper does not report original code. Refer to STAR Methods details for RNA-seq analyses pipeline.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice—All mouse work was approved by the UCSF IACUC, protocol number AN197608, and was conducted in accordance with AALAC and NIH guidelines. Animals were maintained in temperature- and humidity-controlled facilities with 12 h light-dark cycle and *ad libitum* access to water and standard chow. FVB/NJ mouse strain (Jackson Laboratory; 001800) was used in all mouse experiments. Studies were performed with age- and sex-matched male and female mice at embryonic day 10.5 to 6 months. Fluorescence of PEC7-HSP68-mCherry transgenic mice was analyzed from E12.5 to postnatal day 3. Kinky tail ratio of knockout mice was analyzed at post-natal day 21. Gene expression analysis was performed at E10.5, E12.5, postnatal day 2 and 10 weeks of age. Scoliosis phenotype was analyzed at 4 and 6 months of age. Details of age and sex used in the experiment are described in the figures or figure legends. All sex differences were described in the Results section.

METHOD DETAILS

Generation of transgenic mice—For the mouse transgenic enhancer assay, we used an *Hsp68*-mCherry (hCR) plasmid,⁶⁷ a kind gift provided by Drs. Len Pennacchio, Axel Visel and Dianne Dickel (LBNL). This plasmid was digested with *KpnI* and the PCR fragment was cloned, including multi cloning sites amplified by HSP68-MCS primer set (Table S2), using *Hsp68*-LacZ plasmid as a template.²⁵ This modified *Hsp68*-mCherry plasmid was digested with *HindIII*, and mouse PEC7 (chr2:147,566,137–147,567,924; *mm10*) amplified by PEC7-HSP68 primer set with FVB/NJ mouse genome as a template was cloned into the plasmid. The plasmid was digested with *SaI*, and the assayed DNA fragment was released from the backbone vector and used for pronuclear injection, which was performed by the Transgenic Gene Targeting Core at the Gladstone Institute.

Generation and validation of knockout mice—Mouse Xe1 (chr2:147,560,271–147,561,812; *mm10*) and PEC7 (chr2:147,566,137–147,567,924; *mm10*) homology sequence was identified with Ensemble genome browser.²⁴ To generate Xe1, PEC7 or Xe1+PEC7 knockout mice, three gRNAs were designed (Figure 1A, and Table S2) using the gRNA design tool on the Integrated DNA Technologies (IDT) Website and selected based on low off-target and high on-target scores. Two crRNA, tracrRNA and Cas9 protein (IDT; 1081058) were injected to zygotes via the Transgenic Gene Targeting Core at the Gladstone Institute. All mouse work was approved by the UCSF IACUC, protocol number AN197608, and was conducted in accordance with AALAC and NIH guidelines. FVB/NJ mouse strain (Jackson Laboratory, catalog no. 001800) was used.

PCR-Sanger sequencing (primers provided in Table S2) was performed using standard techniques. For Southern blot analyses, genomic DNA was treated with *Bgl*II or *Pvu*II (New England Biolabs; R0144 or R0151) and fractionated by agarose gel electrophoreses. Following capillary transfer onto nylon membranes, blots were hybridized with Digoxigenin (DIG)-labeled DNA probes (Table S2) amplified by the PCR DIG Probe Synthesis Kit (Sigma-Aldrich; 11636090910). The hybridized probe was immunodetected with anti-digoxigenin Fab fragments conjugated to alkaline phosphatase (Sigma-Aldrich; 11093274910) and visualized with a CDP star (Sigma-Aldrich; 11685627001), according to the manufacturer's protocol. Chemiluminescence was detected using FluorChem E (ProteinSimple; 92-14860-00).

Fluorescence quantification—The fluorescent images were analyzed using Fiji software⁶⁸ using a calculation that corrected for total cell fluorescence (CTCF) = integrated density–(area of selected cell × mean fluorescence of background readings), as described by McCloy et al.⁷⁰ The fluorescence intensity of each tail was calculated.

Micro-computed tomography (microCT)—MicroCT scans were performed on fixed mouse skeletons using a Scanco Medical μ CT50 at the UCSF Core Center of Musculoskeletal Biology and Medicine. Specimens were scanned at 30.0 μ m resolution with scanner settings of 55kVP, 109 μ A, 6W, 0.5mm AI filter as well as an integration time of 500ms. Reconstructions were converted into DICOM files using Scanco Medical's integrated μ CT Evaluation Program V6.5–3, then converted into 3D volumes using μ CT Ray V4.0–4.

MicroCT scans were performed on living mice with U-CT (MILabs B.V.), part of VECT or 4CT at the pre-Clinical Imaging core in the Department of Radiology and Biomedical Imaging at UCSF. The scanning parameters were: X-ray tube voltage at 55 kVp and current at 0.19 mA and 75 ms exposure time at each angular step over 360° and 480 steps. Once the data were acquired, the manufacturer-provided cone beam FDK algorithm was used for image reconstruction with the voxel size of 0.160 mm × 0.160 mm × 0.160 mm.

Quantitative RT-PCR—Total RNA was collected from mouse tissues using TRIzol (Thermo Fisher Scientific; 15596026) and converted to cDNA using ReverTra Ace qPCR-RT master mix with genomic DNA (gDNA) remover (Toyobo; FSQ-301). qPCR was

performed using SsoFast EvaGreen supermix (Bio Rad; 1725205). Primer sequences used for qPCR are listed in Table S2.

Whole-mount *in situ* hybridization—Mouse E12.5 embryos were fixed in 4% paraformaldehyde. A plasmid containing mouse *Pax1* (GenScript; OMu21524) was used as template for DIG-labeled probes. Mouse whole-mount *in situ* hybridization was performed according to standard procedures.⁷¹ Briefly, embryos were rehydrated and treated with Proteinase K (Promega; V3021). Following re-fixation and prehybridization, embryos were hybridized with a DIG-labeled probe. The hybridized probe was immunodetected with anti-digoxigenin Fab fragments conjugated to alkaline phosphatase (Sigma-Aldrich; 11093274910) and visualized with a Bmpurple (Sigma-Aldrich; 11442074001) according to the manufacturer's protocol.

Tamoxifen injection—Tamoxifen (Sigma; T5648) was dissolved in sunflower oil (Sigma; S5007d). Pregnant females were injected with tamoxifen subcutaneously (30 mg/kg body weight) at gestational day 17.5. Pups' tail morphology was visually analyzed at P21.

RNA-seq—Total RNA was extracted as described above. RNA-seq was conducted by Novogene. The library was generated with NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB; 7770), and sequencing was done using the NovaSeq 6000 S4 platform with PE150. The data were analyzed using Partek Flow (Version 10.0). After primary quality assessment was performed, bases and reads with low quality were filtered out and the reads were aligned to mouse reference genome (mm10) using the STAR (Version 2.7.8a).⁷² The final BAM files were quantified using the Partek E/M algorithm.⁶⁹ Normalization of read count was performed by the total number of counts (count per million) plus 0.0001, and all genes with less than ten normalized read counts were excluded from subsequent analyses. Differentially expressed genes were identified using Partek gene-specific analysis (GSA) algorithm. Geneontology⁷³ and Morpheus (RRID: SCR_017386) were used for data analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were analyzed with GraphPad Prism 10. Results are presented as mean \pm S.E.M. unless otherwise noted. The number of observations is shown as dots in the figures. Two-tailed Student's t-tests or Fisher's test (sample size <5) or Chi-squared test (sample size >5) were used to determine statistical differences.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported in part by the National Institute of Child Health and Human Development (NICHD) grant number 1P01HD084387 (C.A.W. and N.A.). A.U. was supported by the Japan Society for the Promotion of Science (JSPS) postdoctoral fellowships for research abroad, the Uehara Memorial Foundation postdoctoral fellowship, and the National Human Genome Research Institute (NHGRI).

DECLARATION OF INTERESTS

N.A. is a co-founder and on the scientific advisory board of Regal Therapeutics. N.A. receives funding from BioMarin Pharmaceutical Incorporated.

REFERENCES

- Weinstein SL, Dolan LA, Wright JG, and Dobbs MB (2013). Effects of bracing in adolescents with idiopathic scoliosis. *N. Engl. J. Med* 369, 1512–1521. [PubMed: 24047455]
- Raggio CL (2006). Sexual dimorphism in adolescent idiopathic scoliosis. *Orthop. Clin. North Am* 37, 555–558. [PubMed: 17141012]
- Karol LA, Johnston CE 2nd, Browne RH, and Madison M (1993). Progression of the curve in boys who have idiopathic scoliosis. *J. Bone Joint Surg. Am* 75, 1804–1810. [PubMed: 8258551]
- Grauers A, Rahman I, and Gerdhem P (2012). Heritability of scoliosis. *Eur. Spine J* 21, 1069–1074. [PubMed: 22094388]
- Khanshour AM, Kou I, Fan Y, Einarsdottir E, Makki N, Kidane YH, Kere J, Grauers A, Johnson TA, Paria N, et al. (2018). Genome-wide meta-analysis and replication studies in multiple ethnicities identify novel adolescent idiopathic scoliosis susceptibility loci. *Hum. Mol. Genet* 27, 3986–3998. [PubMed: 30395268]
- Kou I, Takahashi Y, Johnson TA, Takahashi A, Guo L, Dai J, Qiu X, Sharma S, Takimoto A, Ogura Y, et al. (2013). Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. *Nat. Genet* 45, 676–679. [PubMed: 23666238]
- Londono D, Kou I, Johnson TA, Sharma S, Ogura Y, Tsunoda T, Takahashi A, Matsumoto M, Herring JA, Lam T-P, et al. (2014). A meta-analysis identifies adolescent idiopathic scoliosis association with LBX1 locus in multiple ethnic groups. *J. Med. Genet* 51, 401–406. [PubMed: 24721834]
- Miyake A, Kou I, Takahashi Y, Johnson TA, Ogura Y, Dai J, Qiu X, Takahashi A, Jiang H, Yan H, et al. (2013). Identification of a susceptibility locus for severe adolescent idiopathic scoliosis on chromosome 17q24.3. *PLoS One* 8, e72802. [PubMed: 24023777]
- Ogura Y, Kou I, Miura S, Takahashi A, Xu L, Takeda K, Takahashi Y, Kono K, Kawakami N, Uno K, et al. (2015). A functional SNP in BNC2 is associated with adolescent idiopathic scoliosis. *Am. J. Hum. Genet* 97, 337–342. [PubMed: 26211971]
- Sharma S, Londono D, Eckalbar WL, Gao X, Zhang D, Mauldin K, Kou I, Takahashi A, Matsumoto M, Kamiya N, et al. (2015). A PAX1 enhancer locus is associated with susceptibility to idiopathic scoliosis in females. *Nat. Commun* 6, 6452. 10.1038/ncomms7452. [PubMed: 25784220]
- Sharma S, Gao X, Londono D, Devroy SE, Mauldin KN, Frankel JT, Brandon JM, Zhang D, Li Q-Z, Dobbs MB, et al. (2011). Genome-wide association studies of adolescent idiopathic scoliosis suggest candidate susceptibility genes. *Hum. Mol. Genet* 20, 1456–1466. [PubMed: 21216876]
- Takahashi Y, Kou I, Takahashi A, Johnson TA, Kono K, Kawakami N, Uno K, Ito M, Minami S, Yanagida H, et al. (2011). A genome-wide association study identifies common variants near LBX1 associated with adolescent idiopathic scoliosis. *Nat. Genet* 43, 1237–1240. [PubMed: 22019779]
- Zhu Z, Tang NL-S, Xu L, Qin X, Mao S, Song Y, Liu L, Li F, Liu P, Yi L, et al. (2015). Genome-wide association study identifies new susceptibility loci for adolescent idiopathic scoliosis in Chinese girls. *Nat. Commun* 6, 8355. [PubMed: 26394188]
- Kou I, Otomo N, Takeda K, Momozawa Y, Lu H-F, Kubo M, Kamatani Y, Ogura Y, Takahashi Y, Nakajima M, et al. (2019). Genome-wide association study identifies 14 previously unreported susceptibility loci for adolescent idiopathic scoliosis in Japanese. *Nat. Commun* 10, 3685. [PubMed: 31417091]
- Makki N, Zhao J, Liu Z, Eckalbar WL, Ushiki A, Khanshour AM, Wu J, Rios J, Gray RS, Wise CA, and Ahituv N (2021). Genomic characterization of the adolescent idiopathic scoliosis-associated transcriptome and regulome. *Hum. Mol. Genet* 29, 3606–3615. [PubMed: 33179741]
- Wallin J, Wilting J, Koseki H, Fritsch R, Christ B, and Balling R (1994). The role of Pax-1 in axial skeleton development. *Development* 120, 1109–1121. [PubMed: 8026324]

17. Sivakamasundari V, Kraus P, Jie S, and Lufkin T (2013). Pax1(EGFP): new wildtype and mutant EGFP mouse lines for molecular and fate mapping studies. *Genesis* 51, 420–429. [PubMed: 23377878]
18. Wilm B, Dahl E, Peters H, Balling R, and Imai K (1998). Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency. *Proc. Natl. Acad. Sci. USA* 95, 8692–8697. [PubMed: 9671740]
19. Adham IM, Gille M, Gamel AJ, Reis A, Dressel R, Steding G, Brand-Saberi B, and Engel W (2005). The scoliosis (sco) mouse: a new allele of Pax1. *Cytogenet. Genome Res* 111, 16–26. [PubMed: 16093716]
20. Sivakamasundari V, Kraus P, Sun W, Hu X, Lim SL, Prabhakar S, and Lufkin T (2017). A developmental transcriptomic analysis of Pax1 and Pax9 in embryonic intervertebral disc development. *Biol. Open* 6, 187–199. [PubMed: 28011632]
21. Takimoto A, Kokubu C, Watanabe H, Sakuma T, Yamamoto T, Kondoh G, Hiraki Y, and Shukunami C (2019). Differential transactivation of the upstream aggrecan enhancer regulated by PAX1/9 depends on SOX9-driven transactivation. *Sci. Rep* 9, 4605. [PubMed: 30872687]
22. Kokubu C, Horie K, Abe K, Ikeda R, Mizuno S, Uno Y, Ogiwara S, Ohtsuka M, Isotani A, Okabe M, et al. (2009). A transposon-based chromosomal engineering method to survey a large cis-regulatory landscape in mice. *Nat. Genet* 41, 946–952. [PubMed: 19633672]
23. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. [PubMed: 22955616]
24. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov AG, Barnes I, Bennett R, et al. (2022). Ensembl 2022. *Nucleic Acids Res.* 50, D988–D995. [PubMed: 34791404]
25. Kothary R, Clapoff S, Brown A, Campbell R, Peterson A, and Rossant J (1988). A transgene containing lacZ inserted into the dystonia locus is expressed in neural tube. *Nature* 335, 435–437. [PubMed: 3138544]
26. Wallin J, Eibel H, Neubüser A, Wilting J, Koseki H, and Balling R (1996). Pax1 is expressed during development of the thymus epithelium and is required for normal T-cell maturation. *Development* 122, 23–30. [PubMed: 8565834]
27. Peters H, Wilm B, Sakai N, Imai K, Maas R, and Balling R (1999). Pax1 and Pax9 synergistically regulate vertebral column development. *Development* 126, 5399–5408. [PubMed: 10556064]
28. Wang Y, Cao L, Wang Q, Huang J, and Xu S (2019). LncRNA FOXD2-AS1 induces chondrocyte proliferation through sponging miR-27a-3p in osteoarthritis. *Artif. Cells, Nanomed. Biotechnol* 47, 1241–1247. [PubMed: 30945573]
29. Cao L, Wang Y, Wang Q, and Huang J (2018). LncRNA FOXD2-AS1 regulates chondrocyte proliferation in osteoarthritis by acting as a sponge of miR-206 to modulate CCND1 expression. *Biomed. Pharmacother* 106, 1220–1226. [PubMed: 30119190]
30. Gou Y, Guo J, Maulding K, and Riley BB (2018). Sox2 and sox3 cooperate to regulate otic/epibranchial placode induction in zebrafish. *Dev. Biol* 435, 84–95. [PubMed: 29355522]
31. Kinney BA, Al Anber A, Row RH, Tseng Y-J, Weidmann MD, Knaut H, and Martin BL (2020). Sox2 and canonical Wnt signaling interact to activate a developmental checkpoint coordinating morphogenesis with mesoderm fate acquisition. *Cell Rep.* 33, 108311. [PubMed: 33113369]
32. Zheng S, Zhou H, Gao B, Li Y, Liao Z, Zhou T, Lian C, Wu Z, Su D, Wang T, et al. (2018). Estrogen promotes the onset and development of idiopathic scoliosis via disproportionate endochondral ossification of the anterior and posterior column in a bipedal rat model. *Exp. Mol. Med* 50, 1–11.
33. Demirkiran G, Dede O, Yalcin N, Akel I, Marcucio R, and Acaroglu E (2014). Selective estrogen receptor modulation prevents scoliotic curve progression: radiologic and histomorphometric study on a bipedal C57Bl6 mice model. *Eur. Spine J* 23, 455–462. [PubMed: 24136418]
34. Wang Y, Troutwine BR, Zhang H, and Gray RS (2022). The axonemal dynein heavy chain 10 gene is essential for monocilia motility and spine alignment in zebrafish. *Dev. Biol* 482, 82–90. [PubMed: 34915022]
35. Wang Y, Liu Z, Yang G, Gao Q, Xiao L, Li J, Guo C, Troutwine BR, Gray RS, Xie L, and Zhang H (2020). Coding variants coupled with rapid modeling in zebrafish implicate dynein genes, dnaaf1

- and *zmynd10*, as adolescent idiopathic scoliosis candidate genes. *Front. Cell Dev. Biol* 8, 582255. [PubMed: 33251213]
36. Grimes DT, Boswell CW, Morante NFC, Henkelman RM, Burdine RD, and Ciruna B (2016). Zebrafish models of idiopathic scoliosis link cerebrospinal fluid flow defects to spine curvature. *Science* 352, 1341–1344. 10.1126/science.aaf6419. [PubMed: 27284198]
 37. Yu X, Ng CP, Habacher H, and Roy S (2008). *Foxj1* transcription factors are master regulators of the motile ciliogenic program. *Nat. Genet* 40, 1445–1453. [PubMed: 19011630]
 38. Stubbs JL, Oishi I, Izpisua Belmonte JC, and Kintner C (2008). The forkhead protein *Foxj1* specifies node-like cilia in *Xenopus* and zebrafish embryos. *Nat. Genet* 40, 1454–1460. [PubMed: 19011629]
 39. Sasaki K, Shiba K, Nakamura A, Kawano N, Satouh Y, Yamaguchi H, Morikawa M, Shibata D, Yanase R, Jokura K, et al. (2019). Calaxin is required for cilia-driven determination of vertebrate laterality. *Commun. Biol* 2, 226. [PubMed: 31240264]
 40. Kojima T, Hasegawa T, de Freitas PHL, Yamamoto T, Sasaki M, Horiuchi K, Hongo H, Yamada T, Sakagami N, Saito N, et al. (2013). Histochemical aspects of the vascular invasion at the erosion zone of the epiphyseal cartilage in MMP-9-deficient mice. *Biomed. Res* 34, 119–128. [PubMed: 23782745]
 41. Tischfield MA, Bosley TM, Salih MAM, Alorainy IA, Sener EC, Nester MJ, Oystreck DT, Chan W-M, Andrews C, Erickson RP, and Engle EC (2005). Homozygous *HOXA1* mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. *Nat. Genet* 37, 1035–1037. [PubMed: 16155570]
 42. Muscarella LA, Guarnieri V, Sacco R, Militerni R, Bravaccio C, Trillo S, Schneider C, Melmed R, Elia M, Mascia ML, et al. (2007). *HOXA1* gene variants influence head growth rates in humans. *Am. J. Med. Genet. B Neuropsychiatr. Genet* 144B, 388–390. [PubMed: 17171652]
 43. Ajima R, Sakakibara Y, Sakurai-Yamatani N, Muraoka M, and Saga Y (2021). Formal Proof of the Requirement of *MESP1* and *MESP2* in Mesoderm Specification and Their Transcriptional Control via Specific Enhancers in Mice. *Development* 148. 10.1242/dev.194613.
 44. Takahashi Y, Kitajima S, Inoue T, Kanno J, Saga Y, and Saga Y (2005). Differential contributions of *Mesp1* and *Mesp2* to the epithelialization and rostro-caudal patterning of somites. *Development* 132, 787–796. [PubMed: 15677726]
 45. Lo JC, Wang Y, Tumanov AV, Bamji M, Yao Z, Reardon CA, Getz GS, and Fu Y-X (2007). Lymphotoxin β Receptor–Dependent control of lipid homeostasis. *Science* 316, 285–288. [PubMed: 17431181]
 46. Chan V, Fong GCY, Luk KDK, Yip B, Lee M-K, Wong M-S, Lu DDS, and Chan T-K (2002). A genetic locus for adolescent idiopathic scoliosis linked to chromosome 19p13.3. *Am. J. Hum. Genet* 71, 401–406. [PubMed: 12094330]
 47. Ouellet J, and Odent T (2013). Animal models for scoliosis research: state of the art, current concepts and future perspective applications. *Eur. Spine J* 22 (Suppl 2), S81–S95. [PubMed: 23099524]
 48. The PLOS ONE Staff (2014). Correction: Transcriptional profiling of *Wnt3a* mutants identifies *Sp* transcription factors as essential effectors of the *wnt*/ β -catenin pathway in neuromesodermal stem cells. *PLoS One* 9, e95947.
 49. Reed SC (1937). The inheritance and expression of fused, a new mutation in the house mouse. *Genetics* 22, 1–13. [PubMed: 17246824]
 50. Transgenerational Inheritance of Epigenetic States at the Murine *AxinFu* Allele Occurs after Maternal and Paternal Transmission.
 51. Au TYK, Yip RKH, Wynn SL, Tan TY, Fu A, Geng YH, Szeto IYY, Niu B, Yip KY, Cheung MCH, et al. (2023). Hypomorphic and dominant-negative impact of truncated *SOX9* dysregulates Hedgehog-Wnt signaling, causing campomelia. *Proc. Natl. Acad. Sci. USA* 120, e2208623119. [PubMed: 36584300]
 52. Chatterjee S, and Ahituv N (2017). Gene Regulatory Elements, Major Drivers of Human Disease. *Annu. Rev. Genomics Hum. Genet* 18, 45–63, Epub 2017 Apr 7. 10.1146/annurev-genom-091416-035537. [PubMed: 28399667]

53. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112. [PubMed: 19295514]
54. Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, Lee EA, Zhu Y, Plajzer-Frick I, Pickle CS, et al. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* 554, 239–243. [PubMed: 29420474]
55. Hong JW, Hendrix DA, and Levine MS (2008). Shadow enhancers as a source of evolutionary novelty. *Science* 321, 1314. [PubMed: 18772429]
56. Richards JB, Yuan X, Geller F, Waterworth D, Bataille V, Glass D, Song K, Waeber G, Vollenweider P, Aben KKH, et al. (2008). Male-pattern baldness susceptibility locus at 20p11. *Nat. Genet* 40, 1282–1284. [PubMed: 18849991]
57. Hillmer AM, Brockschmidt FF, Hanneken S, Eigelshoven S, Steffens M, Flaquer A, Herms S, Becker T, Kortüm AK, Nyholt DR, et al. (2008). Susceptibility variants for male-pattern baldness on chromosome 20p11. *Nat. Genet* 40, 1279–1281. [PubMed: 18849994]
58. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranaši D, et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 48, D87–D92. [PubMed: 31701148]
59. Wise CA, Sepich D, Ushiki A, Khanshour AM, Kidane YH, Makki N, Gurnett CA, Gray RS, Rios JJ, Ahituv N, and Solnica-Krezel L (2020). The cartilage matrisome in adolescent idiopathic scoliosis. *Bone Res.* 8, 13. [PubMed: 32195011]
60. Zhu H, Shang D, Sun M, Choi S, Liu Q, Hao J, Figuera LE, Zhang F, Choy KW, Ao Y, et al. (2011). X-linked congenital hypertrichosis syndrome is associated with interchromosomal insertions mediated by a human-specific palindrome near SOX3. *Am. J. Hum. Genet* 88, 819–826. [PubMed: 21636067]
61. Chernov AV, and Shubayev VI (2022). Sexually dimorphic transcriptional programs of early-phase response in regenerating peripheral nerves. *Front. Mol. Neurosci* 15, 958568. [PubMed: 35983069]
62. Blecher R, Krief S, Galili T, Biton IE, Stern T, Assaraf E, Levanon D, Appel E, Anekstein Y, Agar G, et al. (2017). The proprioceptive system masterminds spinal alignment: Insight into the mechanism of scoliosis. *Dev. Cell* 42, 388–399.e3. [PubMed: 28829946]
63. Kot A, Zhong ZA, Zhang H, Lay Y-AE, Lane NE, and Yao W (2017). Sex dimorphic regulation of osteoprogenitor progesterone in bone stromal cells. *J. Mol. Endocrinol* 59, 351–363. [PubMed: 28871061]
64. Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, Wucher V, Gewirtz ADH, Cotter DJ, Parsana P, Kasela S, Balliu B, Viñuela A, et al. (2020). The impact of sex on gene expression across human tissues. *Science* 369, eaba3066. 10.1126/science.aba3066. [PubMed: 32913072]
65. Choquet H, Li W, Yin J, Bradley R, Hoffmann TJ, Nandakumar P, 23 and Me Research Team; Mostaedi R, Tian C, Ahituv N, et al. (2022). Ancestry- and sex-specific effects underlying inguinal hernia susceptibility identified in a multiethnic genome-wide association study meta-analysis. *Hum. Mol. Genet* 31, 2279–2293. [PubMed: 35022708]
66. Kfoury N, Qi Z, Prager BC, Wilkinson MN, Broestl L, Berrett KC, Moudgil A, Sankaraman S, Chen X, Gertz J, et al. (2021). Brd4-bound enhancers drive cell-intrinsic sex differences in glioblastoma. *Proc. Natl. Acad. Sci. USA* 118, e2017148118. [PubMed: 33850013]
67. Dickel DE, Ypsilanti AR, Pla R, Zhu Y, Barozzi I, Mannion BJ, Khin YS, Fukuda-Yuzawa Y, Plajzer-Frick I, Pickle CS, et al. (2018). Ultraconserved enhancers are required for normal development. *Cell* 172, 491–499.e15. [PubMed: 29358049]
68. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. [PubMed: 22743772]
69. Xing Y, Yu T, Wu YN, Roy M, Kim J, and Lee C (2006). An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. *Nucleic Acids Res.* 34, 3150–3160. [PubMed: 16757580]

70. McCloy RA, Rogers S, Caldon CE, Lorca T, Castro A, and Burgess A (2014). Partial inhibition of Cdk1 in G 2 phase overrides the SAC and decouples mitotic events. *Cell Cycle* 13, 1400–1412. [PubMed: 24626186]
71. Hargrave M, Bowles J, and Koopman P (2006). In situ hybridization of whole-mount embryos. *Methods Mol. Biol* 326, 103–113.
72. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. [PubMed: 23104886]
73. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet* 25, 25–29. [PubMed: 10802651]

Highlights

- A scoliosis-associated region, PEC7, is an enhancer in the mouse tail and intervertebral disc
- Deletion of Xe1 or both Xe1 and PEC7 enhancers leads to a kinky tail phenotype
- Deletion of Xe1 and PEC7 together leads to a female-biased kinky tail
- Tamoxifen suppresses the sex dimorphic phenotype, suggesting a role for estrogen signaling

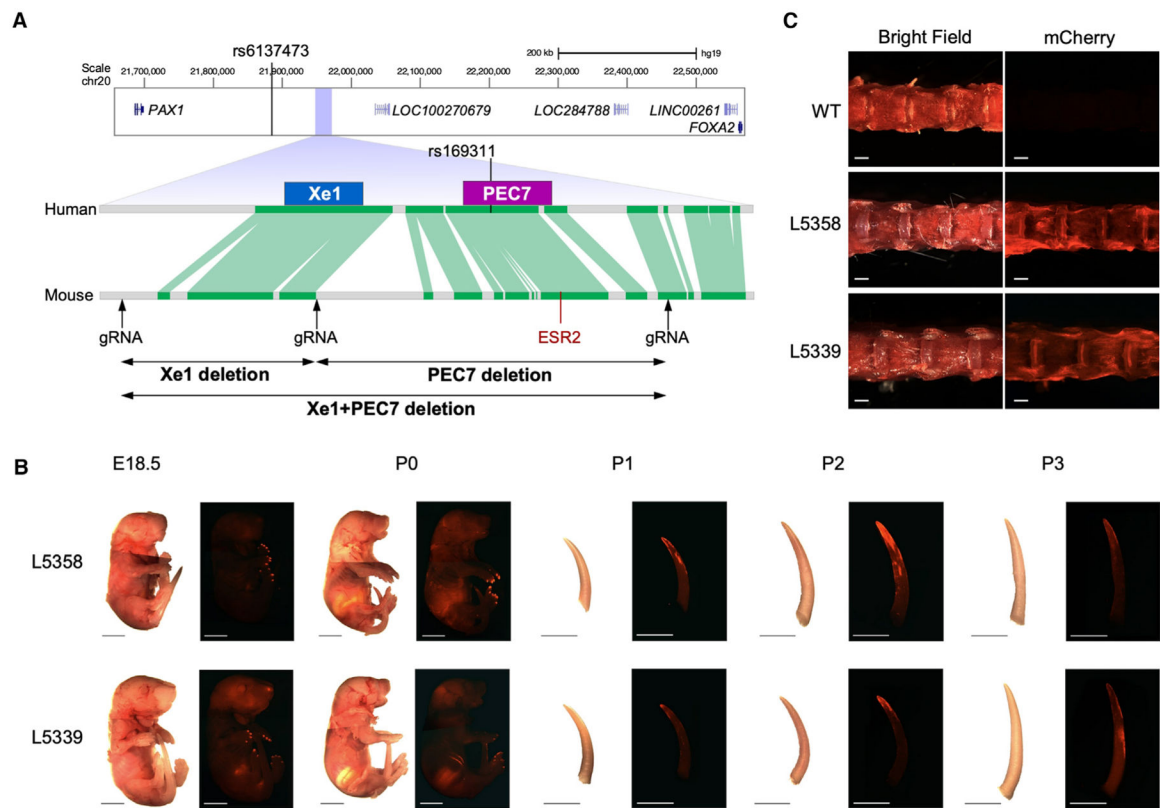


Figure 1. *PAX1* genomic locus in humans and mice and *PEC7* enhancer transgenic assay
 (A) Comparison of the *PAX1* locus between human and mouse. The conservation track is from the Ensembl Genome Browser, with green lines indicating conserved sequences between humans and mice. The location of the AIS-associated SNPs and gRNAs are also shown.

(B) mCherry fluorescence in *PEC7*-HSP68-mCherry transgenic mice from E18.5 to P3 (lines 5358 and 5339). Bar represents 5 mm.

(C) mCherry expression in lumbar IVD of 10-week-old mouse. White bars represent 1 mm.

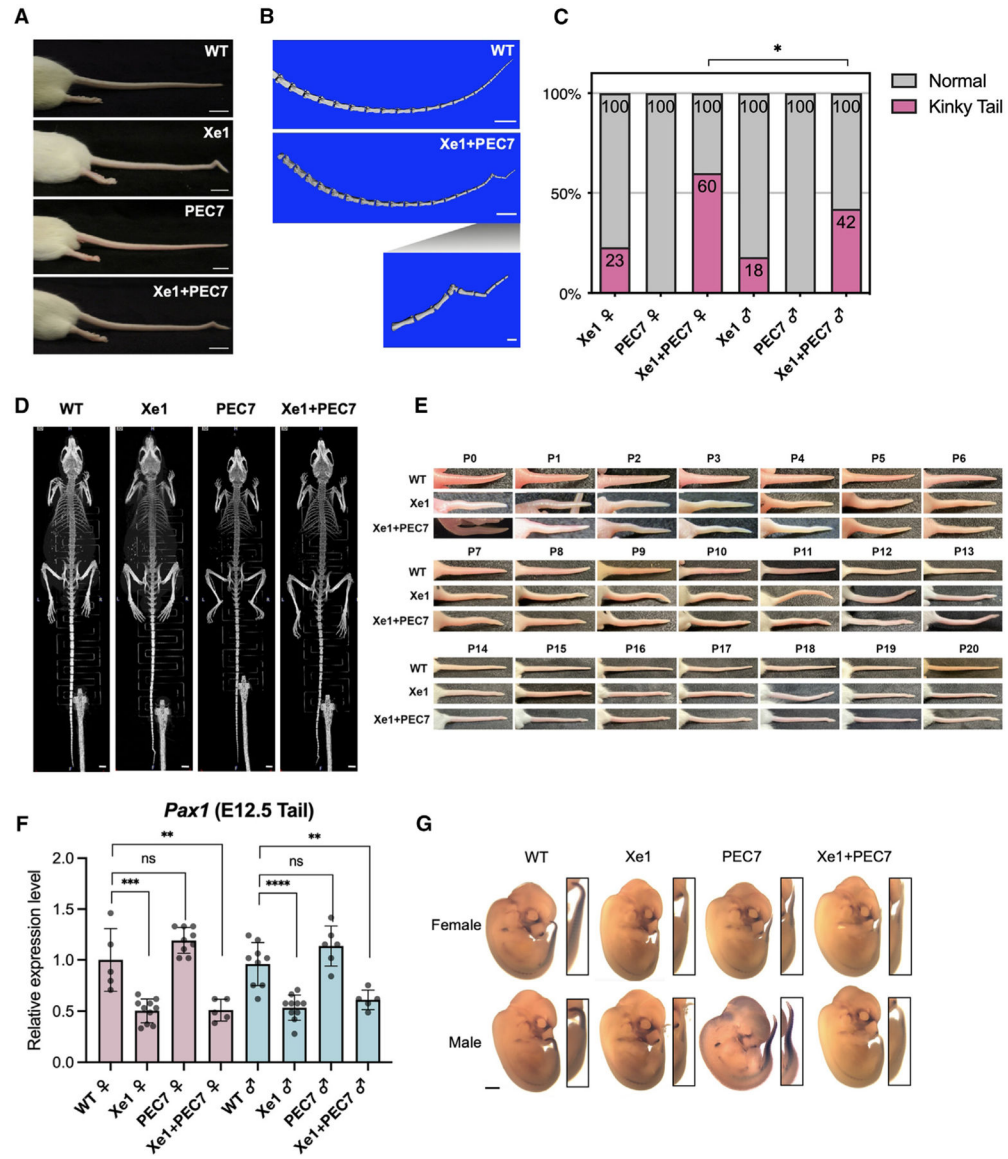


Figure 2. Phenotypic characterization of *Pax1* enhancer knockout mice

(A) Representative tails for each genotype. White bars represent 1 cm.

(B) Tail skeletal structure of 2.5-month-old mice scanned by micro-CT (the same mice as shown in A). White bars represent 5 mm (top two images) and 1 mm (bottom enlarged image).

(C) Kinky tail ratio observed at 3 weeks of age in each knockout mouse group. Statistical differences were determined using chi-squared test (* $p < 0.05$).

(D) Whole-body skeletal structure of 6-month-old mice scanned by micro-CT. White bars represent 3 mm.

(E) Photo of tail from P0 to P20 neonatal stages. Eight Xe1+PEC7 knockout and six Xe1 knockout pups were analyzed.

(F) *Pax1* gene expression levels from E12.5 mouse tail as determined by RT-qPCR. Each value represents the ratio of *Pax1* gene expression to that of β -Actin, and values are mean

± standard deviation. The expression value of wild type (WT) females was arbitrarily set at 1.0. Each dot represents one embryo, and statistical differences were determined using unpaired t test (****p < 0.001, ***p < 0.005, **p < 0.01, ns, not significant).

(G) Whole-mount *in situ* hybridization for *Pax1* of E12.5 mouse embryos. Black bar represents 1 mm.

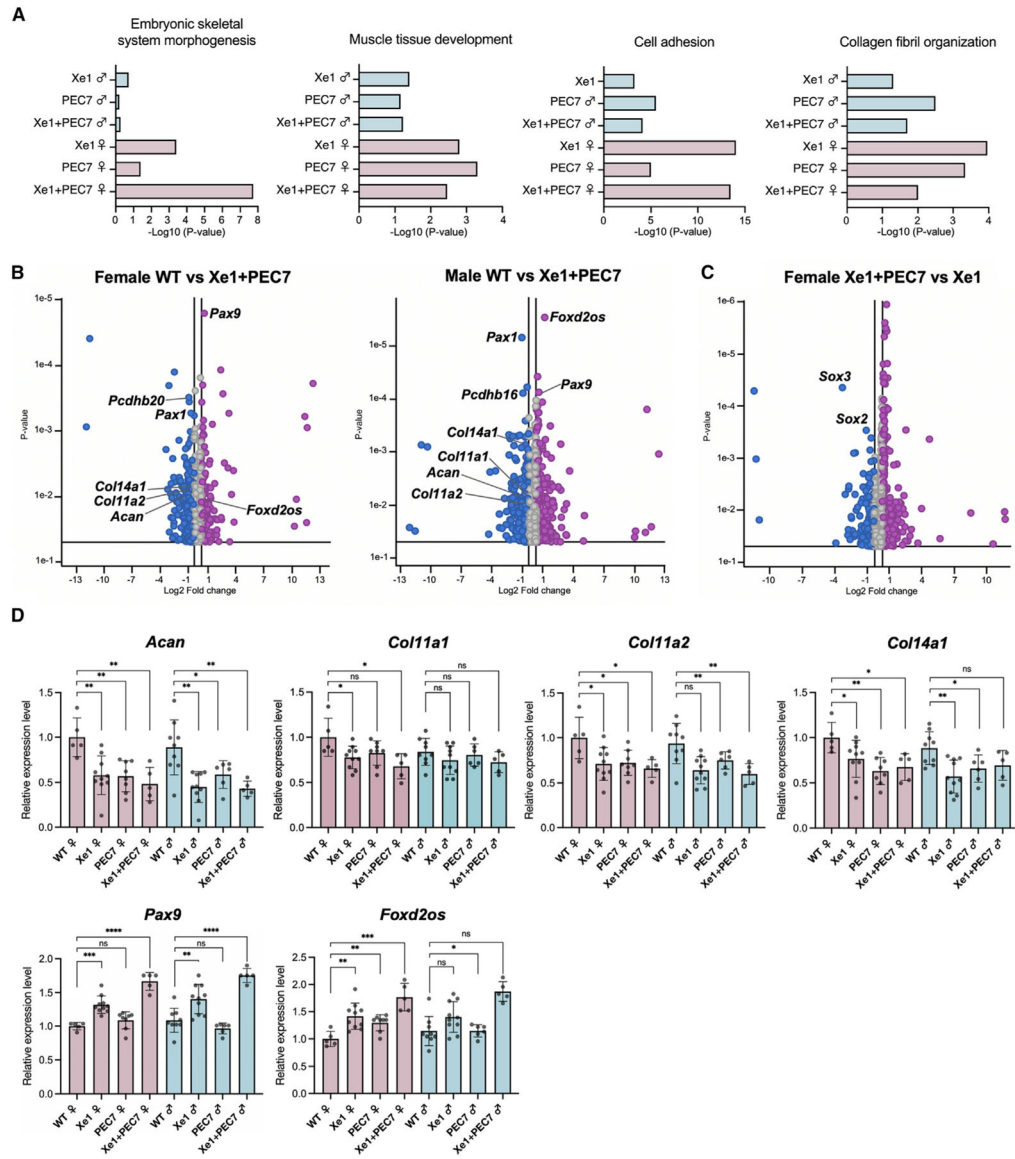


Figure 3. Gene expression profiling of E12.5 tails by RNA-seq and qPCR

(A) Enriched GO terms for gene sets expressed differentially between WT and enhancer knockout mice ($p \leq 0.05$).

(B and C) Volcano plots showing the global transcriptional changes for the indicated groups. Each circle represents one gene. The log₂ fold change in the indicated genotype is represented on the x axis. The y axis shows the p value. A p value of 0.05 and a fold change of 1.5 are indicated by lines.

(D) Gene expression levels dissected from E12.5 mouse tail as determined by RT-qPCR. Each value represents the ratio of gene expression to that of β -Actin, and values are mean \pm standard deviation. The expression value of WT females was arbitrarily set at 1.0. Each dot represents one embryo, and statistical differences were determined using unpaired t test (**** $p < 0.001$, *** $p < 0.005$, ** $p < 0.01$, * $p < 0.05$, ns, not significant).

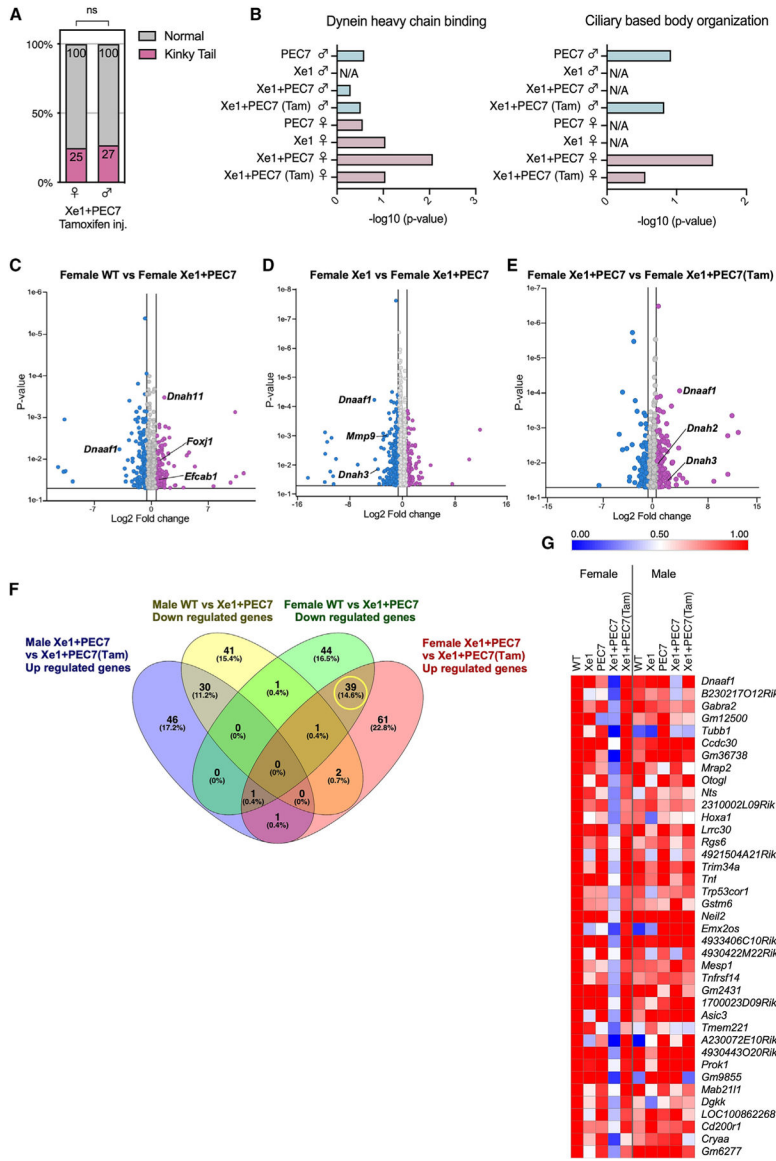


Figure 4. Gene expression profiling at P2 tail by RNA-seq

(A) Kinky tail ratio observed at 3 weeks of age in Xe1+PEC7 knockout mice after tamoxifen injection. Statistical differences were determined using chi-squared test (ns, not significant).

(B) Enriched GO terms for gene sets differentially expressed between WT and enhancer knockout mice ($p \leq 0.05$).

(C–E) Volcano plots showing the global transcriptional changes for the various groups. Each circle represents one gene. The log₂ fold change in the indicated genotype is represented on the x axis. The y axis shows the p value. A p value of 0.05 and a fold change of 1.5 are indicated by lines.

(F) Venn diagram showing genes using the following criteria: downregulated genes compared between WT and Xe1+PEC7 male or female and upregulated genes compared between Xe1+PEC7 and Xe1+PEC7(Tam) male or female (at least ± 2 -fold changes; $p \leq 0.05$). Each number shows gene number and percentage (%).

(G) Heatmap of 39 differentially expressed genes identified from the Venn diagram. The expression value of the WT female group was arbitrarily set at 1.0.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-digoxigenin Fab fragments conjugated to alkaline phosphatase	Sigma-Aldrich	Cat# 11093274910;RRID:AB_514497
Bacterial and virus strains		
Stellar Competent Cells	Takara	636766
Chemicals, peptides, and recombinant proteins		
HindIII	New England Biolabs	R3104
KpnI	New England Biolabs	R3142
SalI	New England Biolabs	R3138
BglII	New England Biolabs	R0144
PvuII	New England Biolabs	R0151
TRIzol	Thermo Fisher Scientific	15596026
Chloroform	Sigma-Aldrich	C2432
Proteinase K	Promega	V3021
Bmpurple	Sigma-Aldrich	11442074001
Tamoxifen	Sigma-Aldrich	T5648
Sunflower oil	Sigma-Aldrich	S5007d
16% Formaldehyde Solution	Thermo Scientific	28906
Cas9 protein	IDT	1081058
CDP star	Sigma-Aldrich	11685627001
Critical commercial assays		
PCR DIG Probe Synthesis Kit	Sigma-Aldrich	11636090910
ReverTra Ace qPCR-RT master mix with genomic DNA	Toyobo	FSQ-301
SsoFast EvaGreen supermix	Bio Rad	1725205
NEBNext Ultra II RNA Library Prep Kit for Illumina	New England Biolabs	7770
Deposited data		
RNA-seq data	Sequence Read Archive (SRA)	PRJNA951902
Experimental models: Organisms/strains		
FVB/NJ mouse	Jackson Laboratory	001800
PEC7-HSP68-mCherry transgenic mice	This paper	N/A
Xe1 knockout mice	This paper	N/A
PEC7 knockout mice	This paper	N/A
Xe1+PEC7 knockout mice	This paper	N/A
Oligonucleotides		
Primers for cloning and sanger sequencing, see Table S2	This paper	N/A
gRNA, see Table S2	This paper	N/A
DNA probe for Southern blot, see Table S2	This paper	N/A
Primers for qPCR, see Table S2	This paper	N/A
Recombinant DNA		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hsp68-mCherry (hCR) plasmid	Dickel et al. ⁶⁷	N/A
Hsp68-LacZ plasmid	Kothary et al. ²⁵	N/A
Mouse Pax1 cDNA plasmid	GenScript	OMu21524
Software and algorithms		
Fiji	Schindelin et al. ⁶⁸	https://imagej.net/software/fiji/
Partek Flow (Version 10.0)	Xing et al. ⁶⁹	https://www.partek.com/partek-flow/
Scanco Medical's integrated μ CT Evaluation Program V6.5–3	ScancoMedical	N/A
μ CT Ray V4.0–4	Scanco Medical	N/A
Cone beam FDK algorithm	MILabs B.V.	N/A
Prism 10	GraphPad	https://www.graphpad.com/
Morpheus	Broadinstitute	https://software.broadinstitute.org/morpheus/
Other		
FluorChem E	ProteinSimple	92-14860-00
Scanco Medical μ CT50	ScancoMedical	N/A
U-CT	MILabs B.V.	N/A
QuantStudio 6 Real-Time PCR Systems	ThermoFisher	N/A