Deletion of the sclerotome-enriched lncRNA *PEAT* augments ribosomal protein expression

David A. Stafford[a], Darwin S. Dichmann[a], Jessica K. Chang[b], and Richard M. Harland[a,1]

[a]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720; and [b]Department of Genetics, Stanford University, Stanford, CA 94305-5120

Contributed by Richard M. Harland, September 23, 2016 (sent for review July 22, 2016; reviewed by Margaret Buckingham and Chen-Ming Fan)

To define a complete catalog of the genes that are activated during mouse sclerotome formation, we sequenced RNA from embryonic mouse tissue directed to form sclerotome in culture. In addition to well-known early markers of sclerotome, such as *Pax1*, *Pax9*, and the *Bapx2*/*Nkx3-2* homolog *Nkx3-1*, the long-noncoding RNA *PEAT* (*Pax1* enhancer antisense transcript) was induced in sclerotome-directed samples. Strikingly, *PEAT* is located just upstream of the *Pax1* gene. Using CRISPR/Cas9, we generated a mouse line bearing a complete deletion of the *PEAT*-transcribed unit. RNA-seq on *PEAT* mutant embryos showed that loss of *PEAT* modestly increases bone morphogenetic protein target gene expression and also elevates the expression of a large subset of ribosomal protein mRNAs.

sclerotome | CRISPR/Cas9 | long-noncoding RNA | ribosomal proteins | BMP pathway

The ribs, vertebrae, and centra of the axial skeleton derive from the sclerotome, a transient mesodermal population specified from the medial-ventral aspect of the embryonic somite. Sclerotomal cells undergo an epithelial-to-mesenchymal transition and migration to surround the spinal cord and extend processes and ribs. Multiple genes and pathways required for the specification of the sclerotome, as well as the subsequent differentiation and elaboration of axial skeletal progenitors, have been identified. Hedgehog (Hh) signals, principally Sonic (Shh), instruct sclerotome formation from the naïve somite (1). The bone morphogenetic protein (BMP) antagonists Noggin and Gremlin also permit Hh-mediated activation of transcription factors required for axial skeleton development (2-4).

Pax1 encodes a paired-domain transcription factor expressed in early sclerotome, and responds directly to Hh within hours of somite formation. Along with *Pax9*, *Pax1* is required for axial skeleton development (5). Other transcription factors have more restricted roles; for example, *Uncx* is required for formation of the transverse processes, proximal ribs, and pedicles (6). As the sclerotome progenitors mature, BMP signaling loses its inhibitory activity and instead assumes a positive role in axial skeleton development, where *Nkx3-2* and *Sox9* drive cartilage gene expression in an autoregulatory loop maintained by BMP (7).

The axial skeleton provides support and locomotive potential, as well as protection for the internal organs. Variation of axial skeletal developmental programs has enabled the successful radiation of vertebrates into diverse environments. In humans, the axial skeleton supports the entire body anterior to the pelvis, a task for which it functions imperfectly. Globally, back pain is the single most prevalent disability (8), with an economic burden of billions of dollars annually (9). In addition, an aging population will increase the incidence of skeletal diseases along with costs to the healthcare system. To date, there is no effective therapy to replace lost bone. Despite the evolutionary and biomedical significance of this organ system, our knowledge of sclerotome formation and axial skeleton development lacks a comprehensive catalog of transcripts, so our knowledge of the transcriptional network underlying sclerotome differentiation is incomplete.

Next-generation sequencing is the ideal tool to quantitatively define transcripts in the sclerotome lineage. However, obtaining developing sclerotome by dissection from early embryos at 8- to 9.5-d postconception is technically difficult. Although a model for presomitic mesoderm culture from stem cells has been developed (10), this has not been exploited as an in vitro model of the sclerotome. To obtain sufficient sclerotomal transcripts for RNA-seq, we chemically treated explants of embryonic tissue to bias differentiation of fluorescently marked presomitic mesoderm toward the sclerotome. Following FACS, we compared control and induced cultures to define the sclerotome transcriptome. In addition to the previously described transcription factors, we identified an uncharacterized noncoding RNA transcript that we named *PEAT* (*Pax1* enhancer antisense transcript; also known as *AI646519*) that is expressed in the reverse orientation to the *Pax1* promoter. To test the hypothesis that *PEAT* influences sclerotome specification through *Pax1*, we deleted the *PEAT* locus using the CRISPR/Cas9 system and performed RNA-seq on *PEAT* mutant embryos.

Results

The Sclerotome Transcriptome. Our objective was to comprehensively catalog the genes activated during sclerotome development. To isolate embryonic sclerotome precursors, we genetically marked the somite and manipulated this tissue toward the sclerotome fate with Hh agonist and BMP antagonist. The Notch ligand *Dll1* is expressed in precursors of the paraxial, intermediate, and lateral plate, as well as restricted populations in the nervous system (11). The regulatory elements directing *Dll1* (12) have been used to generate a mouse line expressing Cre (*Dll1-msd*) (13). We crossed these *Dll1-msd Cre* animals to *R26-stop-EYFP* cre reporter mice to fluorescently mark all derivatives of cells bearing *Dll1-msd* enhancer activity (Fig. 1*A*). We dissected competent posterior

Significance

The majority of transcription generates noncoding RNAs, most of which are uncharacterized. Using RNA-seq on cultured mouse sclerotome, we identified *PEAT*, a long-noncoding RNA (lncRNA) adjacent to a key regulator of sclerotome, *Pax1*. We deleted the entire *PEAT*-transcribed unit using CRISPR/Cas9 and analyzed RNA-seq from mutant embryos. While some lncRNAs regulate the expression of their proximal genes, our analysis showed *Pax1* expression to be unchanged. However, we identified 60 ribosomal proteins with elevated expression, and found evidence that bone morphogenetic protein signaling is slightly elevated in *PEAT* mutants. This study reveals a role for the lncRNA *PEAT* in sclerotome development and shows next-generation sequencing to be a powerful tool to reveal surprising functions for lncRNAs.

Author contributions: R.M.H. and D.A.S. designed research; D.A.S. performed research; R.M.H., D.A.S., D.S.D., and J.K.C. analyzed data; and R.M.H., D.A.S., D.S.D., and J.K.C. wrote the paper.

Reviewers: M.B., Pasteur Institute; and C.-M.F., Carnegie Institution of Washington.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE86112).

[1]To whom correspondence should be addressed. Email: harland@berkeley.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1612069113/-/DCSupplemental.

paraxial mesoderm from *Dll1-msd Cre;R26-stop-EYFP* embryos at 9-d postconception and cultured explants for 12 h with a "midline cocktail" containing sonic agonist (SAG), to mimic Shh from the notochord and floorplate (1), and the BMPR1 inhibitor LDN 193189 (LDN), to mimic notochord-derived Noggin and dermatome derived Gremlin (2–4). Because we isolated explants from the posterior mesoderm before somite patterning, we hypothesized that any effects would include changes in gene expression coincident with or upstream of sclerotome induction. By quantitative PCR (qPCR) we observed a 14-fold relative induction of *Pax1* in LDN and SAG-treated samples and a concomitant reduction in markers of other cell types derived from the *Dll1* lineage (Fig. 1*B*), confirming that sclerotome was induced. We collected EYFP$^+$ cells from dissociated SAG/LDN-treated and vehicle control cultures and analyzed gene expression by RNA-seq from four biological replicates (Fig. 1*C*).

Using DESeq2, we identified 264 genes with statistically significant (adjusted $P < 0.05$), greater than twofold differential expression between control and sclerotome-directed cultures, with 98 transcripts elevated in response to SAG and LDN and 166 transcripts reduced by the treatment (Table S1). Known early sclerotome-specific genes featured prominently in the sclerotome-directed dataset. *Pax1* and *Pax9* expression was elevated in directed samples 6.2- and 2.4-fold, respectively. Expression of the paralog *of Bapx/Nkx3-2*, *Nkx3-1*, was 7.2 times that of control samples. Similar to *Pax1*, *Nkx3-1* expression was detected in somites 2 h after formation of the somite (Fig. 2*B*). Although *Nkx3-1* is dispensable for skeletal formation, the severity of the *Bapx/Nkx3-2* skeletal phenotype is enhanced in the absence of *Nkx3-1* (14). *Heyl*, a member of the hairy/enhancer-of-split related family of transcription factors, was expressed at 6.0 times the level of controls. *Heyl* is expressed in the early sclerotome but is prominently associated with cardiovascular development (15); no somite-formation or somite-patterning role has been ascribed. Notably, only eight transcription factors were induced greater than twofold. In addition to *Pax1*, *Pax9*, and *Nkx3-1*, these transcription factors included *Heyl*, *Eya2*, *Barx2*, *Olig2*, and *Bhlhe41*. Our analysis also further highlighted the role of BMP inhibition (2–4); among the 10 genes with greatest differential expression were *BMP3*, an inhibitory BMP (16), and *Neuroblastoma1* (*Nbl1*, also known as *Dan*), a DAN family BMP antagonist (17).

Genes that respond to SAG and LDN with diminished expression underscore the importance of BMP inhibition in sclerotome formation. Members of the inhibitor of DNA binding (Id) family control differentiation by interfering with the DNA binding activity of basic helix–loop–helix transcription factors. Our analysis showed *Id1*, *Id2,* and *Id3* expression to be reduced (0.36-, 0.12-, and 0.16-fold, respectively). *Id* genes are targets of BMP (18), so their reduced expression in response to LDN was

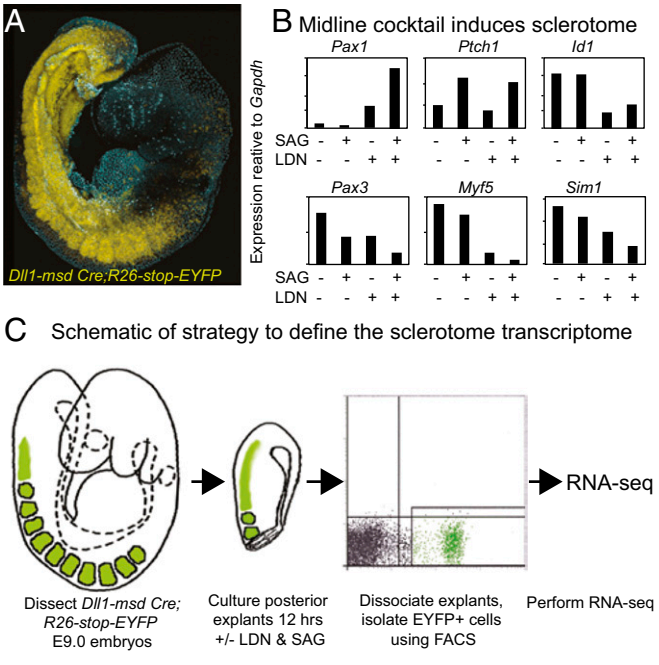

Fig. 1. Activation of Hh and inhibition of BMP directs paraxial mesoderm toward sclerotome formation. (*A*) Confocal *z*-stack (magnification: 100×) of E9 *Dll1-msd Cre;R26-stop-EYFP* mouse embryo. (*B*) qPCR on mouse embryo explants cultured with or without LDN and SAG: *Pax1* marks the sclerotome; *Ptch1* indicates Hh pathway activation; *Id1* is a direct BMP target; *Pax3* marks dermomyotome; *Myf5* marks skeletal muscle progenitors; and *Sim1* is expressed in lateral dermomyotome. (*C*) Schematic of methodology.

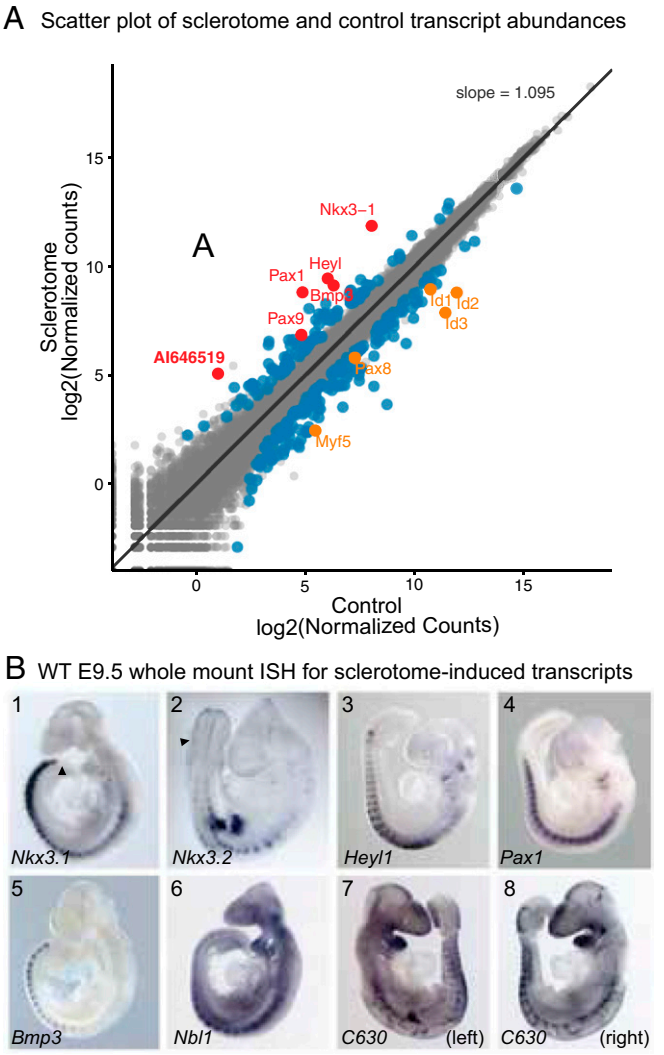

Fig. 2. The sclerotome transcriptome. (*A*) Scatter plot of sclerotome-directed RNAseq data. Blue dots indicate transcripts with greater than twofold differential expression ($P < 0.05$), red dots highlight select genes elevated in directed samples, and orange dots show repressed examples. (*B*) E9 whole-mount in situ hybridizations (ISHs) for select transcripts elevated following sclerotome induction. (1 and 2) *Nkx3-1* appears in the second most recently formed somite (S1; panel 1), whereas *Nkx3-2* is detected in somites 6+ (2). Arrowheads indicate S1. (3) *Heyl1* is activated in S1. (4) *Pax1*. (5 and 6) *Bmp3* (5) and *Nbl1* (6) are extracellular inhibitors of BMP signaling. (7 and 8) The left and right sides, respectively, of embryos probed for *C630043F03Rik* (*C630*). (Magnification: 30×.)

expected. The potential functional activity for Id proteins in restricting axial skeleton development has not been examined. We also observed diminished expression of genes with roles in the specification of other cell types in the *Dll1* lineage, including *Myf5* for demomyotome (19) and *Pax8* lateral plate mesoderm (20).

Long-noncoding RNAs (lncRNAs) have a wide range of regulatory functions that have only begun to be explored (21–23). We identified several uncharacterized, differentially expressed (DE) lncRNAs with expression in somitic tissues. For example, *C630043F03Rik* (C630), expressed 2.2-fold over control cultures, exhibits dynamic embryonic expression in the head, first pharyngeal arch, base of the right limb bud, and somites. Among the transcripts with greatest relative induction in the SAG- and LDN-treated samples was a noncoding RNA gene *AI646519* (MGI), with 6.7-fold greater expression over controls. *AI646519* is spliced and polyadenylated and is transcribed unidirectionally from the minus strand 500 bp 5′ to the *Pax1* transcription start site (TSS). The proximity of this lncRNA to *Pax1*, the transcription factor most associated with early sclerotome, was particularly intriguing. We named this gene *PEAT*. Expression of *PEAT* is identical to *Pax1* (Figs. 2 *B*, *4*, and 3 *C* and *D*). Because lncRNAs have been shown to affect transcription of an adjacent gene (24), we hypothesized that *PEAT* may regulate the early development of the sclerotome through *Pax1*.

Deletion of the lncRNA PEAT. The *Pax1* gene is located in a 870-kb gene desert between the developmental transcription factors *Nkx2-2* and *FoxA2* (Fig. 3). This locus has a rich history of genetic analysis. The naturally occurring, semidominant *Undulated-Short Tail* (*Pax1un-s*) mutation is a 122-kb deletion that includes all *Pax1* coding sequences and presents with severe defects of the axial skeleton (25). However, animals lacking just the *Pax1* gene (*Pax1null*) are viable and fertile and exhibit only tail defects (26). The explanation for this difference came from deletion analysis that demonstrated the *Pax1un-s* phenotype to be a consequence of ectopic *Nkx2-2* expression in the sclerotome as a result of a repositioned *Pax1* 3′ enhancer element (27). Because no targeted mutation of the *Pax1* locus ablated 5′ sequences but retained the *Pax1* coding region, the contribution of *PEAT* is unknown.

To address the role of *PEAT* in the early development of the sclerotome, we designed single-guide RNAs (sgRNAs) to delete the *PEAT*-transcribed unit using the CRISPR/Cas9 system. Because *PEAT* lacks an ORF and consensus TSS, we used the existing annotation (mm10) and our mapped RNA-seq reads to predict the transcribed region. We injected one-cell mouse embryos with *Cas9* mRNA and the sgRNAs and transferred these to pseudopregnant females. Seven of the 9 resulting pups carried the designed deletion. The *ΔPEAT* deletion followed Mendelian inheritance, and homozygous mutant animals are viable and fertile. *ΔPEAT* homozygous mutant embryos processed for in situ hybridization for *PEAT* lacked detectable signal (Fig. 3).

We predicted that the *ΔPEAT* phenotype would be subtle, like that of the *Pax1*-null (26), so we assessed changes in gene expression in differentiating sclerotome by RNA-seq. We dissected trunk tissue corresponding to somites 1–24 from embryonic day (E) 9 embryos generated by *ΔPEAT* heterozygote intercrosses. Because committed sclerotome is mesenchymal, we included flanking tissues so as not to lose any of this population. We sequenced libraries from three wild-type and three homozygous mutants. Reads mapping to regions 5′ to *Pax1* were absent, confirming that *ΔPEAT* eliminates transcription in or near the sgRNA targets (Table S2 and Fig. S1). Despite this finding, we observed no significant difference in *Pax1* expression between wild-type and *ΔPEAT* samples. Although alternative analyses might reveal a consistent effect, at this point we have no indication of a role for *PEAT* in *Pax1* expression, nor do we see the subtle phenotype of the *Pax1* mutant (26).

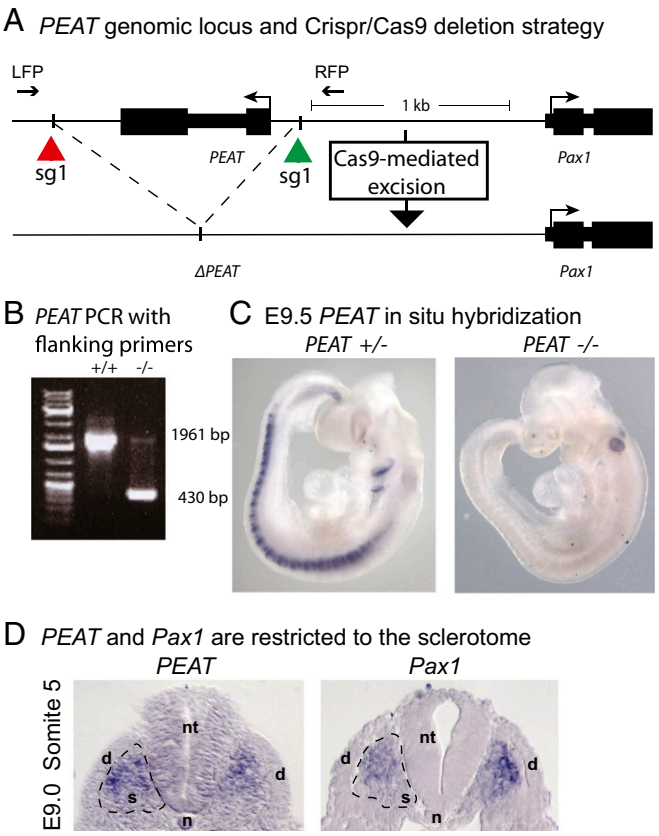

Fig. 3. The lncRNA *PEAT*. (*A*) *PEAT* is transcribed off the minus-strand from ∼1,000 bp 5′ to the *Pax1* TSS (chr2, 127.05 Mb). Cas9 and sgRNAs targeting sites flanking the *PEAT* transcribed unit were used to generate the *ΔPEAT* line. (*B*) PCR on genomic DNA from *ΔPEAT* homozygotes using primers flanking the sgRNA target sites produces a 430-bp amplicon. (*C*) E9 whole-mount ISHs (Magnification: 30×.) for *PEAT* on *PEAT* embryos (heterozygous and homozygous *ΔPEAT* from left to right). *PEAT* is expressed identically to *Pax1* (compare with Fig. 2 *B*, *4*). (*D*) E9 transverse sections of wild-type *PEAT* and *Pax1* ISH at level of fifth most recently formed somite. d, dermomyotome; n, notochord; Nt, neural tube; left sclerotomes (s) delineated by dashed lines.

In contrast to *Pax1*, we did find 700 differentially expressed genes (DEGs) (Fig. 4) (adjusted $P < 0.05$), although none showed greater than twofold difference in expression between conditions. *PEAT* ablation resulted in more genes with elevated expression (482) than reduced expression (218) (Fig. 4). Strikingly, Gene Ontology (GO) analysis showed the preponderance of these genes encoded ribosomal subunit proteins (Fig. 4), members of a large family of proteins that modulate ribosome biogenesis and translation. Of the 105 expressed ribosomal protein genes with greater than 10 mapped reads in directly sequenced *Dll1*-lineage samples (*Materials and Methods*), 55 displayed significantly elevated expression in *ΔPEAT*; only one exhibited reduced expression (Table S2). We confirmed by qPCR that the five ribosomal proteins with greatest differential activation all trended higher in independently generated cDNA samples from embryos (*SI Appendix*, Fig. S2).

Several targets of BMP signaling also appeared in the list of genes with significant elevated expression. These genes included *Id3* and *Id1* (18), which were expressed at 1.4× that of wild-type controls, and *JunD* (28), expressed 1.3× over controls. *Dpysl2*, which was reduced in ΔPEAT samples by 14%, is repressed by BMP4 (29). Because BMP signaling interferes with early sclerotome development, these changes could indicate that *PEAT* helps maintain a permissive environment in the developing trunk favorable to early axial skeleton development. The lack of an overt

DEVELOPMENTAL BIOLOGY

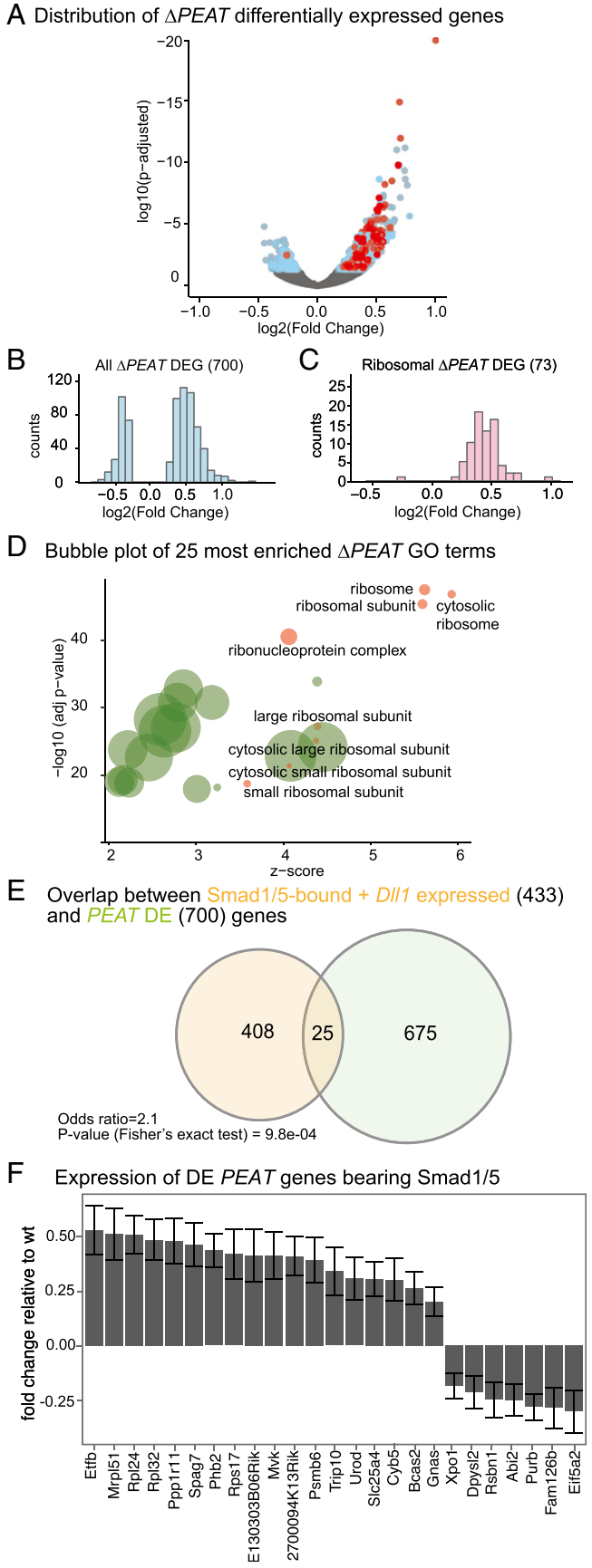

Fig. 4. Deletion of *PEAT* increases ribosomal protein and BMP pathway gene expression. (*A*) Volcano plot of *ΔPEAT* RNA-seq data. Blue dots signify significantly DEGs ($P < 0.05$). Red dots signify ribosomal protein DEGs. (*B*) Histogram showing distribution of all 700*ΔPEAT* DEGs. (*C*) Histogram showing distribution of the 73 ribosomal protein *ΔPEAT* DEGs. (*D*) Cellular component (CC) GO analysis on *ΔPEAT* DEGs. The bubble plot was generated using GO plot 1.0.2 from GO enrichment analysis (*Materials and Methods*). Circle size is proportional to the number of genes assigned to a GO term. The *z*-score is number of genes increased minus genes decreased divided by the square root of the number of genes assigned to a GO category. The GO terms specifically indicating ribosomal protein genes are clustered in the upper right (Table S4). (*E*) The *ΔPEAT* DEGs exhibit significant enrichment for genes with Smad1/5-bound proximal promoters. (*F*) Histogram showing expression of the 25 *ΔPEAT* DEGs bearing Smad1/5 relative to wild-type controls. Bars indicate log fold-change SE.

ΔPEAT phenotype is consistent with this model. Not only are the changes in gene expression modest, but there is precedent for tolerance of modestly increased BMP signal transduction in vivo; mice conditionally heterozygous for the BMP antagonist *Noggin* exhibit no negative consequences to a small yet measurable increase in expression of the direct BMP target *Id3* (*SI Appendix*, Fig. S3).

We reasoned that if *ΔPEAT* samples have elevated BMP signaling, expression of the activated genes might be reduced by interference with BMP signal transduction. We exploited our sclerotome-directed dataset to test this, because these conditions use the BMP antagonist LDN. Indeed, we found 16 such reduced targets, including *Id1*, *Id3* (Tables S1 and S2, and *SI Materials and Methods*). In addition, *Csrp2*, whose expression increases by 32% in *ΔPEAT* mutants and is induced by Bmp2 (30), is repressed to 66% of control levels in sclerotome-directed samples. Furthermore, although BMP has not directly been shown to activate the Wnt-inhibitory *Tcf3* (induced 17% in *ΔPEAT* samples), it is known that Wnt inhibition functions with BMP signaling to promote epithelial stem cell differentiation (31), so a positive *Tcf3*–BMP relationship is likely.

To substantiate a relationship between *PEAT* and the BMP pathway, we turned to published ChIP-seq data for the BMP signaling effector proteins, Smad1 and Smad5 (Smad1/5) (29). Fei et al. (29) performed their analysis on mouse embryonic stem cells; there is no existing R-Smad ChIP-seq analysis for the paraxial mesoderm. To generate a list more reflective of the midgestation mesoderm, we cross-referenced the list of genes with Smad1/5 bound in the proximal promoter with genes expressed in the *Dll1* mesodermal lineage (*Materials and Methods* and Table S3). Of 24,062 annotated transcripts, 433 had Smad1/5-bound promoters and were expressed in the *Dll1* lineage. Twenty-five genes on this list also appeared in our list of 700 DE*ΔPEAT* genes, which indicates a strong likelihood of association (Fisher's exact test $P = 9.8e^{-4}$; observed odds ratio = 2.1) (Table S3). This analysis strongly supports our hypothesis that *ΔPEAT* mutants exhibit altered BMP signaling.

The Expression of *PEAT*-Regulated Genes During *Dll1*-Lineage Maturation. Recently, the impact of ribosomal proteins on specific developmental processes has become appreciated, particularly in the developing somite (32). We hypothesized that *PEAT* may potentiate sclerotome development by regulating genes that impact sclerotome development. To track changes in the expression of *PEAT*-regulated genes in the *Dll1*-lineage, we generated a gene-expression catalog for four successively more mature stages of differentiation (Fig. 5*A*). We isolated trunk segments from E9 embryos, each with a length corresponding to four somite diameters reflecting ∼8 h of development each. Using the first posterior somite as a landmark, segments A included the posterior mesoderm and presomatic mesoderm, as well as the noncommitted most recently formed somite (33). Segments B contained somites 2–5 and correspond to stages during which initial somite pattern formation occurs. We also collected two additional segments, C (somites 6–9) and D (somites 10–13), reflecting somite maturation (Fig. 5*A*). We restricted this

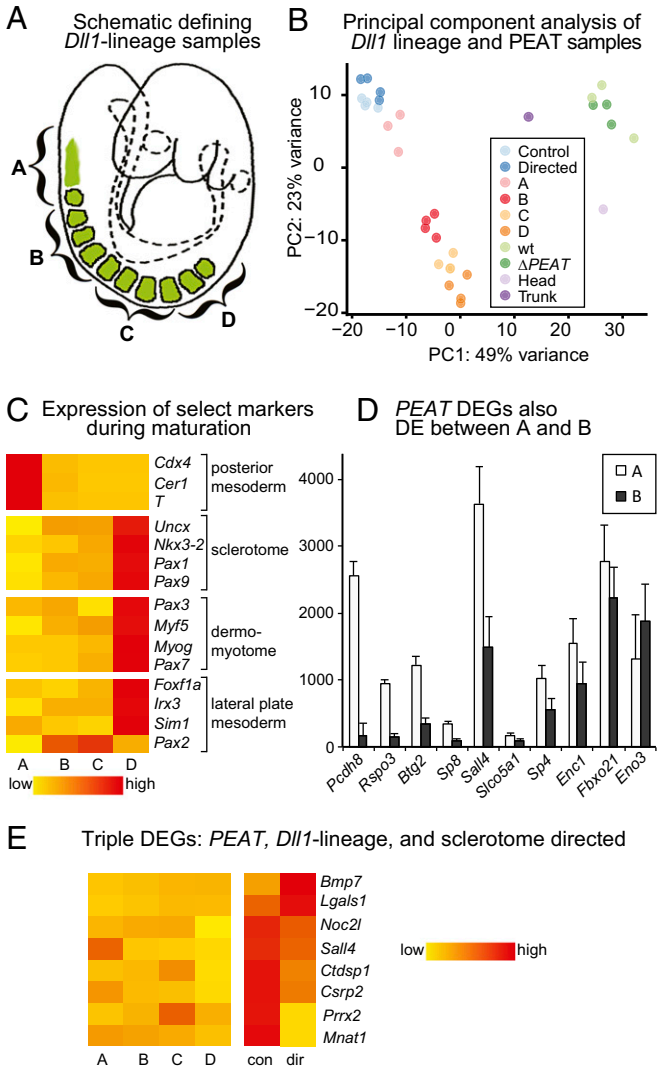

Fig. 5. Analysis of the *Dll1*-lineage. (*A*) Schematic describing the four segments, A to D, dissected from E9.0 *Dll1-msd Cre;R26-stop-EYFP* mouse embryo. (*B*) PCA for all RNA-seq samples, including single-replicate unsorted head and trunk samples. (*C*) Heat map showing average expression of select mesoderm-expressed genes for the 32 h following *Dll1-msd* initiation (Table S5). (*D*) Histogram showing the difference in average expression of the 10 *ΔPEAT* DEGs that are also variable between *Dll1*-lineage samples A and B. Bars indicate 2× SE. (*E*) Heat map displaying the variation in average expression of the eight *ΔPEAT* DEGs that are also DE between both *Dll1*-lineage samples and the control and sclerotome-directed samples.

study to embryos with between 16 and 20 somites to maintain consistency in anterior–posterior identity of dissected segments. Samples from multiple embryos were pooled according to anterior–posterior segment type, dissociated for 30 min, and FACS-sorted for GFP^+ cells. We generated RNA-seq libraries from batches of 2.5×10^5 cells. Principal component analysis showed that replicates from each group formed tight clusters, indicating highly consistent and reproducible gene-expression profiles (Table S6 and Fig. 5*B*).

The *Dll1*-lineage stage series tracks transcription during the specification and early development of all *Dll1*-derived somitic cell types, including dermomyotome and myotome and sclerotome. We found that the lineage-expression analysis faithfully tracks normal maturation of trunk lineages (Fig. 5 *B* and *C*); the posterior, early-expressed trunk genes are enriched as expected in the youngest, most posterior segment. In more anterior explants, markers of differentiated somite derivatives appear. This analysis also captured expression of anterior-posterior regionalized markers, notably the *Hox* genes (Table S7).

We then assessed changes in expression of *PEAT*-regulated genes across the four *Dll1*-lineage time points. This analysis identified 29 *PEAT*-regulated genes that exhibit changes in expression during the 32 h of *Dll1*-development. Because our primary interest is initial pattern formation within the paraxial mesoderm, we examined expression changes of the 10 *PEAT* DEGs from segments A to B (Fig. 5*D*). For 9 of the 10 DEGs, average normalized counts fell from A to B, coincident with the onset of *PEAT* expression. We also queried whether there were any genes that were DE between (*i*) *Dll1*-lineage segments, (*ii*) wild-type control and *ΔPEAT* embryos, and (*iii*) control and sclerotome-directed samples. Eight genes satisfied this triple condition (Fig. 5*E*) and are candidates to have unexplored activities promoting or repressing sclerotome development.

However, the most striking outcome of this *Dll1*-lineage expression analysis was the absence of DE ribosomal proteins, the most prominent DEG category in *PEAT* mutants. There was only one gene associated with translation, *Eif5a2.* This result does not support a model in which these ribosomal proteins function directly in *Dll1*-lineage tissue specification. We speculate that the changes in ribosomal protein gene expression in response to *PEAT* deletion are compensatory, and act to maintain cellular conditions required for normal development.

Discussion

Here we report analysis of the paraxial mesoderm transcriptome, with emphasis on the sclerotome lineage. A highly activated lncRNA in the sclerotome lineage, *PEAT* (AI646519; MGI) is expressed in the reverse orientation from the Pax1 promoter. As Pax1 regulates sclerotome induction, we directly tested the contribution of *PEAT* to sclerotome development through a CRISPR/Cas9-mediated deletion. Using RNA-seq on *PEAT* mutant embryos, we identified a class of 60 ribosomal proteins with elevated expression, and present evidence that BMP signaling is significantly elevated in *PEAT* mutants.

Previous work has indicated that enhancer-associated lncRNAs can affect the transcription of the adjacent gene. *Pax1/9* enhancer-associated lncRNAs are conserved in location but not sequence, for example in the human genome, where (RP5-106502.4/LOC101929608) is expressed from the human *PAX1* enhancer region (chr20:21.203 Mb). *Pax9* also has an enhancer-associated lncRNA, *ENSMUST00000159874*. Although this conservation does strongly suggest functional importance, our *PEAT* knockout shows that the RNA is not needed for regulation of the adjacent *Pax1* gene. Nonetheless, the genomic position of *PEAT* and changes in gene expression following *PEAT* deletion suggest a role in sclerotome development. We speculate that the answer may lie in the joint utilization of an enhancer. We envision a situation in which two genes are involved in the development of the same cell type, but are not directly related to each other mechanistically. We think of this strategy to link the generation of transcripts involved in parallel in a developmental process as "car pooling." Such a mechanism would be an efficient strategy to generate transcripts that are involved in the same developmental trajectory, whether or not they are functionally related.

How *PEAT* deletion affects the expression of a large subset of ribosomal proteins is also unclear. One of the best-characterized factors in the regulation of ribosomal protein expression is the Target of Rapamycin (TOR) pathway, which activates ribosomal protein biogenesis under favorable conditions (34). We see no change in TOR pathway mRNAs in *PEAT* mutants, and although there could be a change in protein activity, it is unclear how this would only a subset of the ribosomal protein gene family.

PEAT mutants showed a significant but moderate increase in BMP target gene expression. Because BMP signaling inhibits sclerotome induction, this result is consistent with a role for

PEAT in restricting a pathway that interferes with sclerotome development. The concordant increase in a subset of ribosomal proteins may implicate these genes as BMP-regulated and similarly inhibitory to sclerotome. However, we do not see any decrease in the expression of genes activated in the sclerotome lineage in *PEAT* mutants. For this reason we propose that *PEAT* may be part of a buffering system where the steady-state expression of genes including ribosomal proteins is sufficient and does not change much in the lineage; stress on the system may cause a broad increase in the expression of these ribosomal proteins that promotes translation of proteins critical for development the lineage. We hypothesize that the differentially expressed *ΔPEAT* ribosomal proteins are thus prime candidates to have tissue-specific roles in development. Indeed, expression of *Rpl38*, the ribosomal protein known to function specifically in somitic *Hox* mRNA translation (32), is elevated in *PEAT* mutants.

Materials and Methods

Generation of Mice. *Dll1*-msd Cre animals on the FVB/N were kindly provided by Deborah Chapman, Department of Biological Sciences, University of Pittsburgh, Pittsburgh, and R26r-EYFP animals on a C57BL/6J background were obtained from Jackson Laboratories.

The confocal image of the *Dll1-msd Cre;R26r-EYFP* was recorded on a Zeiss LSM710. The *ΔPEAT* line was generated using albino C57BL/6J zygotes by Chulho Kang of the Berkeley Cancer Research Laboratory, Berkeley, CA. The sgRNA sequences were 5′: GCTCAAAGAGTAGCTAAAGT, 3′: TGTAAGGACAGACTACGCAA. A male pup bearing a deletion covering the entire transcribed unit was selected as the founder for these experiments. All animals were housed and manipulated under the University of California, Berkeley Animal Care and Use Committee guidelines.

Culturing, Dissociation, and FACS. E9 mouse embryos were isolated, screened to confirm *Dll1*-lineage GFP expression, and cut in the transverse plane at the anterior–posterior level of the most recently formed somite. Posterior explants were cultured in 20-μL hanging drops of 50% (vol/vol) DMEM:50% (vol/vol) Rat Serum (Harlan) at 37 °C in a standard tissue culture incubator [5% (vol/vol) CO_2] with 1 μm SAG (Cayman Chemical) and 1 μm LDN-193189 (Stemcell) or vehicle control (DMSO). After 12 h, explants were collected and rinsed 2× in PBS and digested 25 min in 1× Trypsin LE (ThermoFisher). Tissue was gently triturated in 5% (vol/vol) FBS, 10 U/mL DNase in Hepes-buffered HBSS with a 1-mm bore fire-polished Pasteur pipette and passed through a 40-μm cell strainer. Dissociated cells were sorted using a BD Influx cell sorter with GFP^+ cells sorted directly into TRIzol-LS (ThermoFisher). Approximately 2,000 GFP^+ cells were obtained per posterior explant, and 200,000 cells were collected for each biological replicate. Total RNA was extracted using the manufacturer's instructions and purified using RNeasy MinElute clean-up columns (Qiagen).

RNA-Seq Libraries. Triplicate sequencing libraries for control, sclerotome-directed, and *Dll1*-lineage stage series used the TruSeq RNA Library Preparation Kit V2 with 11 rounds of PCR. Indexed libraries were sequenced on a HiSeq. 2000 machine for 100-bp paired-end reads. Trunk RNA from three individual wild-type and *ΔPEAT*-embryos were used to generate sequencing libraries at the Functional Genomics Lab, a QB3-Berkeley Core Research Facility at the University of California, Berkeley. Illumina Ribo-Zero rRNA Removal Kits were used to deplete rRNA. Library preparation used an Apollo 324 with PrepX RNAseq Library Prep Kits (WaferGen Biosystems), and 18 cycles of PCR amplification was used for index addition and library-fragment enrichment. Libraries were sequenced for 50-base single reads on one lane of an Illumina HiSEq. 4000.

Bioinformatic Analysis. Details of the bioinformatic pipeline can be found in *SI Materials and Methods*. The data discussed in this publication have been deposited in National Center for Biotechnology Information's Gene Expression Omnibus and are accessible through GEO Series accession no. GSE86112.

qPCR. cDNA was synthesized from TRIzol or TRIzol-LS-extracted total RNA using iScript (Bio-Rad). qPCR with SYBR Green reagents on a BioRad CFX-100 Real Time System. Primers sequences can be found in the *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Rachel Kjolby for help with data submission. This work was supported by NIH Grants GM42341, GM49346, and HD065705 (to R.M.H.) and National Science Foundation Fellowship DGE-114747 (to J.K.C.).

1. Fan CM, Tessier-Lavigne M (1994) Patterning of mammalian somites by surface ectoderm and notochord: evidence for sclerotome induction by a hedgehog homolog. *Cell* 79(7):1175–1186.
2. McMahon JA, et al. (1998) Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev* 12(10):1438–1452.
3. Stafford DA, Brunet LJ, Khokha MK, Economides AN, Harland RM (2011) Cooperative activity of noggin and gremlin 1 in axial skeleton development. *Development* 138(5):1005–1014.
4. Stafford DA, Monica SD, Harland RM (2014) Follistatin interacts with Noggin in the development of the axial skeleton. *Mech Dev* 131:78–85.
5. Peters H, et al. (1999) Pax1 and Pax9 synergistically regulate vertebral column development. *Development* 126(23):5399–5408.
6. Leitges M, Neidhardt L, Haenig B, Herrmann BG, Kispert A (2000) The paired homeobox gene Uncx4.1 specifies pedicles, transverse processes and proximal ribs of the vertebral column. *Development* 127(11):2259–2267.
7. Zeng L, Kempf H, Murtaugh LC, Sato ME, Lassar AB (2002) Shh establishes an Nkx3.2/Sox9 autoregulatory loop that is maintained by BMP signals to induce somitic chondrogenesis. *Genes Dev* 16(15):1990–2005.
8. Hoy D, et al. (2014) The global burden of low back pain: Estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis* 73(6):968–974.
9. Dagenais S, Caro J, Haldeman S (2008) A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 8(1):8–20.
10. Zhao J, et al. (2014) Small molecule-directed specification of sclerotome-like chondroprogenitors and induction of a somitic chondrogenesis program from embryonic stem cells. *Development* 141(20):3848–3858.
11. Bettenhausen B, Hrabě de Angelis M, Simon D, Guénet JL, Gossler A (1995) Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to *Drosophila* Delta. *Development* 121(8):2407–2418.
12. Beckers J, et al. (2000) Distinct regulatory elements direct delta1 expression in the nervous system and paraxial mesoderm of transgenic mice. *Mech Dev* 95(1-2):23–34.
13. Wehn AK, Gallo PH, Chapman DL (2009) Generation of transgenic mice expressing Cre recombinase under the control of the Dll1 mesoderm enhancer element. *Genesis* 47(5):309–313.
14. Herbrand H, Pabst O, Hill R, Arnold H-H (2002) Transcription factors Nkx3.1 and Nkx3.2 (Bapx1) play an overlapping role in sclerotomal development of the mouse. *Mech Dev* 117(1-2):217–224.
15. Weber D, Wiese C, Gessler M (2014) Hey bHLH transcription factors. *Curr Top Dev Biol* 110:285–315.
16. Daluiski A, et al. (2001) Bone morphogenetic protein-3 is a negative regulator of bone density. *Nat Genet* 27(1):84–88.
17. Hsu DR, Economides AN, Wang X, Eimon PM, Harland RM (1998) The *Xenopus* dorsalizing factor Gremlin identifies a novel family of secreted proteins that antagonize BMP activities. *Mol Cell* 1(5):673–683.
18. Hollnagel A, Oehlmann V, Heymer J, Rüther U, Nordheim A (1999) Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *J Biol Chem* 274(28):19838–19845.
19. Ott MO, Bober E, Lyons G, Arnold H, Buckingham M (1991) Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo. *Development* 111(4):1097–1107.
20. Bouchard M, Souabni A, Mandler M, Neubüser A, Busslinger M (2002) Nephric lineage specification by Pax2 and Pax8. *Genes Dev* 16(22):2958–2970.
21. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 445(7128):666–670.
22. Grote P, et al. (2013) The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell* 24(2):206–214.
23. Yoon J-H, et al. (2013) Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat Commun* 4:2939.
24. Wang KC, et al. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472(7341):120–124.
25. Blandova ZK, Egorov I (1975) Sut allelic with un. *Mouse News Lett* 52:43.
26. Wilm B, Dahl E, Peters H, Balling R, Imai K (1998) Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency. *Proc Natl Acad Sci USA* 95(15):8692–8697.
27. Kokubu C, et al. (2003) Undulated short-tail deletion mutation in the mouse ablates Pax1 and leads to ectopic activation of neighboring Nkx2-2 in domains that normally express Pax1. *Genetics* 165(1):299–307.
28. Lee S-Y, et al. (2012) The role of heterodimeric AP-1 protein comprised of JunD and c-Fos proteins in hematopoiesis. *J Biol Chem* 287(37):31342–31348.
29. Fei T, et al. (2010) Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. *Genome Res* 20(1):36–44.
30. Clancy BM, et al. (2003) A gene expression profile for endochondral bone formation: Oligonucleotide microarrays establish novel connections between known genes and BMP-2-induced bone formation in mouse quadriceps. *Bone* 33(1):46–63.
31. He XC, et al. (2004) BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-beta-catenin signaling. *Nat Genet* 36(10):1117–1121.
32. Kondrashov N, et al. (2011) Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* 145(3):383–397.
33. Aoyama H, Asamoto K (1988) Determination of somite cells: Independence of cell differentiation and morphogenesis. *Development* 104(1):15–28.
34. Martin DE, Soulard A, Hall MN (2004) TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell* 119(7):969–979.