Distinct transcriptional programs underlie Sox9 regulation of the mammalian chondrocyte

Shinsuke Ohba[1,3,*], **Xinjun He**[2,3], **Hironori Hojo**[2], and **Andrew P. McMahon**[2,*]

[1]Department of Bioengineering, The University of Tokyo Graduate School of Engineering, Tokyo 113-0033, Japan

[2]Department of Stem Cell Biology and Regenerative Medicine, Eli and Edythe Broad-CIRM Center for Regenerative Medicine and Stem Cell Research, W.M. Keck School of Medicine of the University of Southern California, CA 90089

SUMMARY

Sox9 encodes an essential transcriptional regulator of chondrocyte specification and differentiation. When Sox9 nuclear activity was compared with markers of chromatin organization and transcriptional activity in primary chondrocytes, we identified two distinct categories of target association. Class I sites cluster around the transcriptional start sites of highly expressed genes with no chondrocyte-specific signature. Here, Sox9 association reflects protein-protein association with basal transcriptional components. Class II sites highlight evolutionarily conserved active enhancers directing chondrocyte-related gene activity through direct binding of Sox9 dimer complexes to DNA. Sox9 binds through sites with sub-optimal binding affinity; the number and grouping of enhancers into super-enhancer clusters likely determines the levels of target gene expression. Interestingly, comparison of Sox9 action in distinct chondrocyte lineages points to similar regulatory strategies. In addition to providing insights into Sox family action, our comprehensive identification of the chondrocyte regulatory genome will facilitate study of skeletal development and human disease.

Graphical Abstract

*Correspondence: ohba@bioeng.t.u-tokyo.ac.jp (S.O.), amcmahon@med.usc.edu (A.P.M.).
[3]These authors contributed equally to this work.
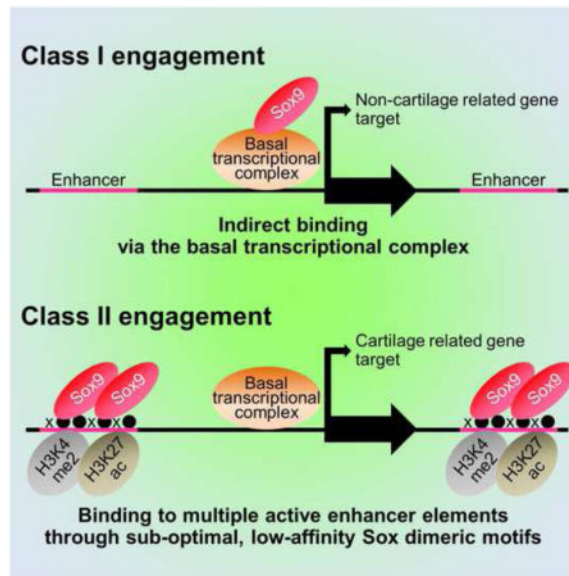
INTRODUCTION

The mammalian skeleton is synthesized by cartilage-secreting chondrocytes and bone-forming osteoblasts. Different skeletal structures arise from distinct cell lineages: neural crest cells form much of the cranial vault and face; paraxial mesoderm derivatives generate additional head structures, vertebrae and ribs, while lateral plate mesoderm derivatives generate the sternum and limb skeleton (Olsen et al., 2000). Direct intramembranous ossification and cartilage-templated endochondral ossification represent alternate modes of bone formation for specific skeletal structures (Helms and Schneider, 2003; Kronenberg, 2003). In endochondral ossification, mesenchymal cells initially differentiate into mitotic chondrocytes that deposit extracellular matrix to form cartilage molds. Mitotic chondrocytes transit to postmitotic hypertrophic chondrocytes that eventually undergo cell death, leaving a matrix that is converted to bone by invading osteoblasts.

SRY-box containing gene 9 (Sox9) is a key regulator of vertebrate endochondral skeletal development (Akiyama and Lefebvre, 2011). *Sox9* is initially expressed in mesenchymal condensations that differentiate into both chondrocytes and osteoblasts; early expression is essential for further development of both cartilage and bone (Akiyama et al., 2002; Akiyama et al., 2005; Bi et al., 1999). Subsequently, *Sox9* expression resolves exclusively to chondrocytes; here Sox9 activity is essential for the chondrogenic program (Akiyama et al., 2002; Bi et al., 1999). In this, *Sox9* is expressed at highest levels in mitotic and early pre-hypertrophic chondrocytes; *Sox9* is down-regulated as chondrocytes undergo hypertrophic expansion and in joint forming regions and joint-associated articular cartilage (Akiyama et al., 2002; Dy et al., 2012; Wright et al., 1995). Recent studies have extended Sox9 action from chondrocyte specification and early chondrocyte differentiation to the initiation of chondrocyte hypertrophy (Dy et al., 2012). In summary, Sox9 functions at multiple stages of the cartilage program from mesenchymal condensation to chondrocyte hypertrophy. The importance of *Sox9* levels to normal development is exemplified by the fact that

haploinsufficiency for *Sox9* is lethal in both mouse and man (Bi et al., 2001; Schafer et al., 1996), and causes campomelic dysplasia (Foster et al., 1994; Wagner et al., 1994).

Several *cis*-regulatory enhancer elements associated with direct, Sox9-dependent regulation have been identified by random and bioinformatically-driven analysis of non-coding regions flanking *Sox9* (Mead et al., 2013), *Ctgf* (connective tissue growth factor) (Huang et al., 2010), and seven genes encoding cartilage matrix proteins: *Col2a1* (Bell et al., 1997; Lefebvre et al., 1996; Leung et al., 1998; Zhou et al., 1998), *Col9a1* (Genzer and Bridgewater, 2007; Zhang et al., 2003), *Col10a1* (Dy et al., 2012), *Col11a2* (Bridgewater et al., 1998; Bridgewater et al., 2003; Liu et al., 2000), *Col27a1* (Jenkins et al., 2005), *Hapln1* (Kou and Ikegawa, 2004), *Matn1* (Nagy et al., 2011; Rentsendorj et al., 2005), and *Acan* (Han and Lefebvre, 2008; Hu et al., 2012). However, as there has been no systematic study of Sox9 actions, Sox9's broad regulatory functions in cartilage development are not well understood.

To this end, we performed a detailed analysis of Sox9 binding, chromatin organization and transcriptional programs within mammalian chondrocytes isolated directly from the neonatal mouse rib. In addition, Sox9 programs were compared between rib and nasal chondrocytes to understand Sox9 actions in chondrocytes arising from distinct cell lineages. Our studies demonstrate different modes of Sox9 engagement at cartilage specific gene targets compared with expressed non-cartilage specific genes.

RESULTS

Identification of Sox9 Class I targets

Rib chondrocytes were manually dissected from post-natal day 1 (P1) mouse to include proliferative and prehypertrophic zones and to exclude mature hypertrophic regions (Supplemental Experimental Procedures and Figures S1A–S1C). Chondrocytes were subjected to chromatin immunoprecipitation with a variety of antibodies (Table S1) followed by high-throughput sequencing (ChIP-seq); the key features of each ChIP-seq dataset are summarized in Table S2.

For the Sox9 rib ChIP-seq, 27,656 raw peaks met the peak-calling criterion. The first clear feature of Sox9 ChIP-seq peaks is a striking enrichment around the TSS; approximately 24.6% (6,794/27,656) of all peaks lie +/− 500 base pairs (bp) around the TSS even though this region represents only 0.001% of the genome (McLean et al., 2010) (Figure 1A). In addition, a large fraction of Sox9 bound regions lie at considerable distance from the TSS; about 50% of all peaks map to an interval between +/− 50 kb and 500 kb from the TSS (Figure 1B).

We first examined the characteristics of the TSS associated Sox9 dataset within the +/− 500 bp window around the TSS, hereafter referred to as Class I peaks (Sheet 1 of Table S3). The peak quality scores for Class I peaks ranked lower than non TSS associated Sox9 regions (Figure 2A): few Class I peaks were mapped within the top 2,000 peaks while many were found within the lowest ranked 2,000 peak regions (Figures 2B and 2C). These data suggest either a generally weaker direct binding interaction, or an indirect mode of interaction close

to the TSS. Consistent with the latter argument, Sox9 motifs were not enriched in TSS proximal peak regions whereas enrichment was observed in Sox9-associated regions more distant from the promoter (Table S4 and Figure 2D).

We analyzed gene expression within the same region of the rib by microarray analysis and compared this data with predicted Class I peak regions. The Sox9 ChIP-seq signal intensity around the TSSs was highly correlated with the expression level of the associated genes, and with the binding of RNA polymerase II and p300, an activating component of the transcriptional complex (Figure 2E); RNA polymerase II and p300 show a similar level of co-enrichment (Figure S2A–S2C). To understand the biological importance of this peak set, we performed a GREAT Gene Ontology analysis (GREAT GO) (McLean et al., 2010). Although some chondrogenesis-related genes are present in the nearest genes to Class I peaks (Sheet 1 of Table S3), Class I targets fall within terms related to general cellular processes, not chondrocyte specific activities (Figure 2F) though comparison of MGI mouse expression data does show a significant correlation with rib gene expression (Figure 2F).

In summary, for Class I targets, Sox9 associates around the TSS of what comprises a large set of genes controlling general cellular functions not specific to chondrocytes. The level of Sox9 binding reflects the levels of core transcriptional complex engagement, and with this, the levels of expression of the associated gene.

Identification of Sox9 Class II targets

Next, we examined the features of the remaining Class II peaks (20,862 regions, Sheet 2 of Table S3). Class II peaks showed a highly enriched recovery of Sox9 motifs, even amongst the most distant peaks sets relative to the TSS (Figure 3A), indicative of direct Sox9 binding, and a high conservation score across vertebrate species around the peak center consistent with an expected conservation of the *cis*-regulatory genome (Figure 3B). A clear association of Class II regions with enhancer signatures was evident; Sox9 bound regions showed an H3K4me2^{high}/H3K4me3^{low} enhancer signature (He et al., 2010), a strong association of p300 and RNA polymerase II consistent with active enhancers (Heintzman et al., 2007; Kim et al., 2010; Visel et al., 2009) and peaks of H3K27 acetylation (H3K27Ac) flanking Sox9, indicative of open chromatin (Heintzman et al., 2007; Rada-Iglesias et al., 2011; Zentner et al., 2011) (Figures 3C and 3D). In striking contrast to the Class I data set, GREAT GO showed a highly significant recovery of expected terms for a Sox9-regulated skeletal program (Figure 3E).

We observed substantial Sox9 ChIP-seq signal at 7 of 9 cartilage enhancers that have been identified through rigorous transgenic analysis, though the one site in *Acan* was below the statistical cutoff of the stringent peak calling process (Table S5). In contrast, we observed Sox9 binding at only one out of 4 regulatory elements associated with Sox9 regulation in testis cells (Table S5). Figures 3F and 3G show screen shots of ChIP-seq data around two essential cartilage matrix protein encoding genes, *Acan* and *Col2a1*.

At a higher order level, a marked clustering of Sox9 binding was evident around key chondrocyte genes expressed at high levels in chondrocytes (Figure S2D). This is supported by further analysis using the algorithm to detect so-called super-enhancers (Whyte et al.,

2013). Among 13,648 adjusted rib chondrocyte enhancers in this analysis, we identified 502 super-enhancer groupings around 422 genes. Expression correlation analysis showed that expression levels of super-enhancer-associated genes were significantly higher than those of typical enhancer-associated genes ($p < 2.2e-16$) (Figures 3H and 3I, and Sheet 3 and 4 in Table S3), suggesting that multiple cooperative Sox9 bound enhancers may have evolved to ensure appropriate expression of chondrocyte genes.

Identification of putative target genes for Sox9

We connected Class II Sox9 binding events to nearest genes and intersected this data set with microarray generated gene expression profiles of comparable chondrocytes. Class II peak associated-genes were positively expressed in chondrocytes relative to mouse embryo fibroblasts (MEF). Further, summing multiple sites of engagement around a putative target, it is clear that expression levels of that target increase with increasing levels of Sox9 engagement (Figure 4A).

We utilized a distance weighted regulatory potential analysis tool, S-score (Tang et al., 2011) (Supplemental Experimental Procedures), to predict direct target genes for Sox9. The analysis extracted 735 genes with high confidence (Figure 4B and Sheet 5 in Table S3), including all known Sox9 chondrocyte target genes mentioned in the Introduction. David GO analysis on this gene list showed that skeletal system development and cartilage related terms were highly enriched (Figure 4C). Together, the data are consistent with a central role of Sox9 in the activation of chondrocyte targets. Furthermore, the sum of Sox9 engagement around a gene rather than Sox9 binding at any individual site provides the strongest prediction of gene expression levels (Figures S3A and S3B).

De novo motif analysis of Sox9 peaks

Sox9 has been shown to activate transcription of chondrocyte specific genes through its binding to the inverted repeat of a quite variable Sox recognition sequence and that dimerization is required for activation (Bernard et al., 2003; Bridgewater et al., 2003; Mead et al., 2013; Sock et al., 2003). We performed *de novo* motif analysis for Class II Sox9 peak regions using two algorithms, CisGenome (Ji et al., 2008) and MEME-ChIP (Bailey et al., 2009). The primary motifs recovered were similar to a previously predicted Sox-dimeric motif, a pair of the Sox consensus sequences (A/T)(A/T)CAA(A/T)G oriented head to head and separated by 4 nucleotides (Bridgewater et al., 2003; Sock et al., 2003) (Figures 4D and S3C). As expected for direct binding of Sox9 at Class II targets, this dimeric motif is highly enriched at the predicted center of Sox9 ChIP-seq peaks (Figure 4E). A Sox-monomer motif was also recovered though these exhibited a lower enrichment score and were only weakly centered about Sox9 peaks (Figures 4F, 4G, and S3C). Thus, the Sox monomer motif does not likely represent the preferred primary site for Sox9 engagement in chondrocytes. We also recovered significant enrichment of motifs predicting engagement of AP-1, Nfat, Fox, Runx and Hox transcription factor families, consistent with the integration of multiple regulatory inputs through Sox9 directed enhancer modules (Figure S3C).

In vivo enhancer activity of Sox9 peak regions and *in vitro* verification of recovered motifs

To verify Sox9-regulated enhancers from our data, we selected a set of predicted enhancers around likely Sox9 target genes, analyzing enhancer activity *in vivo* in a Zebrafish Enhancer Detection (ZED) system (Bessa et al., 2009). Given strong vertebrate conservation of enhancer modules, we expected that many would likely operate outside of the mouse, in a system better suited for large-scale transgenic approaches. Fourteen of 17 tested regions showed chondrocyte specific reporter gene expression (Figure 5A and Table S6). While some enhancers were active as a single copy, others required multimerization to reveal enhancer activity. Most enhancers were not active in all chondrocytes but displayed mosaic activity in individual skeletal elements. Increasing the number of copies of an individual element increased both the likelihood of enhancer activity and the level of reporter gene expression, and decreased mosaicism within the observed expression pattern (Figure S4A).

To explore Sox9 interactions in more depth, we focused on an enhancer in the intron of the cartilage matrix-encoding gene, *Col9a1*. Sox9 binding was associated with an evolutionary conserved motif predicted to bind Sox9 dimers within the identified chondrocyte enhancer element (Figures 5A and 5B). Enhancer-driven chondrocyte specific gene expression (Figure 5C, WT) was lost on mutation of the Sox9 site (Figure 5C, MT). Interestingly, when the endogenous motif was substituted with the optimum consensus predicted from the analysis of the entire Sox9 Class II data set, we observed stronger, less mosaic chondrocyte specific reporter activity (Figure 5C, OP).

Direct Sox9 interaction with target DNA was examined by an Electrophoresis Mobility Shift Assay (EMSA) on the predicted Sox dimeric motif for the wild-type *Col9a1* enhancer (WT, Figure 5D) and one substituting the optimized Sox9 motif from whole data analysis motif recovery (OP, Figure 5D). Sox9 bound to both sequences; however, binding was markedly stronger with the optimized Sox9 target sequence (Figure 5D). As expected, enhancer binding was dependent on the HMG-box of Sox9 (Figure 5D). Mutations in either half site abolished Sox9 binding (MT1 and MT2 in Figure 5D). Further, a mutated Sox9 oligomer failed to compete with Sox9 binding to the wild-type Sox9 motif (MT, Figure 5E). As expected, the wild-type motif competed effectively but only at a high (five-fold) molar excess (WT, Figure 5E). In contrast, the optimal Sox9 site containing oligonucleotide competed effectively at equimolar levels leading to the loss of most Sox9-binding (OP, Figure 5E).

To determine whether other endogenous Sox9 sites showed sub-optimal binding relative to the recovered optimum Sox9 motif, we examined Sox9 binding to oligonucleotides incorporating 10 additional predicted endogenous Sox9 binding motifs and assayed each by EMSA. All 10 probes were bound by Sox9 protein with different levels of binding ability in the EMSA assay (Figure S4B). All were competed as effectively, or more effectively by 3-fold molar excess of optimal oligonucleotide compared with oligonucleotides for each wild-type binding site (Figure S4C).

Together these data make a strong case for a direct mode of Sox9 engagement through a dimeric binding motif. Further, the data suggest that regulation may have favored weaker binding regions than those predicted from optimal nucleotide positioning, and the reiteration

of multiple weaker, *cis*-regulatory modules to maximize output. Consistent with this view, a statistical analysis of the occurrence of two optimal half-sites (ACAAAG, 298 in 2,474 Sox motifs) predicts 34 optimal motifs in this dataset, while the observed number is 13 (binomial p-value 4e-5), and none of these are associated with any likely chondrocyte target gene.

Functional validation of Class I and Class II engagement

We sought functional roles for Class I and Class II engagements through two approaches: (1) examining gene expression changes of targets upon Sox9 manipulation and (2) intersection of human SNP data with our data set.

Regarding the first approach, we ectopically expressed *SOX9* in human dermal fibroblasts (hDFs) with an adenoviral vector; hDFs are reported to acquire chondrocytic phenotypes upon the introduction of *SOX9* (Ikeda et al., 2004). We confirmed no endogenous SOX9-activity in hDFs by comparing activation of a 48-bp *COL2A1* intron 1 enhancer reporter construct (Kan et al., 2009) in hDFs with two Sox9+ chondrogenic cell lines (Figure 6A). Next, we infected hDFs with Sox9 containing virus, and performed RNA sequencing (RNA-seq) 2 days post infection. A box plot analysis of expression changes revealed that *SOX9* overexpression induced both Class I and Class II targets (Class I targets, p<2.2e-16; Class II targets, p<2.2e-16), though as expected Class II targets showed much larger expression changes compared to Class I targets (p=3.4e-16) upon *SOX9* overexpression (Figure 6B). Furthermore, the top 200 highly expressed genes within hDFs displayed a significant elevation in mRNA levels when *SOX9* overexpression was compared with an control (p=1.7e-05, Figure 6C). GO analysis revealed that skeletal system development related terms were the most significantly enriched in gene sets displaying a 2-fold or greater increase on *SOX9* overexpression (p= 7.2e-06, Figure 6D and Sheet 3 in Table S7). In contrast to the above gene sets comprising 235 genes (fold change rank 1 to 235), the next set of 235 genes (fold change rank 236 to 470), a low responder gene sets to SOX9 input, showed little association with skeleton related terms though ossification related terms were enriched with a low p value (p=0.048, Figure 6D and Sheet 4 in Table S7). In summary, Sox9 is likely to up-regulate a subset of Class II chondrocyte targets in hDFs. Further, Sox9 can broadly elevate transcript levels for non-chondrocyte genes consistent with a general interaction with basal transcriptional components.

To examine the relationship of Sox9 binding with 73 million human SNPs in the NCBI SNP database we mapped 523,295 SNPs to the mouse genome (mm9), and 100 SNPs were then matched to homologous mouse Sox9 binding sites in rib chondrocytes (Sheet 6 in Table S3). We analyzed the closest genes to the SNPs overlapping with Sox9 binding sites, and 8 genes (*Nfatc1*, *Ucma*, *Col9a3*, *Itga5*, *Grasp*, *Col2a1*, *Glt25d2*, and *Mia1*) were identified from the list of Sox9 putative targets identified by the S-score analysis of Class II targets. This set of SNPs may be useful foci for future exploration in human disease.

Comparison of Sox9 targets and gene expression between mesodermal- and neural crest-derived chondrocytes

To investigate whether Sox9 regulatory networks are similar between chondrocytes derived from neural crest and mesoderm, we performed expression profiling and Sox9 ChIP-seq on

cartilage of E18.5 neural crest-derived nasal septum. Nasal chondrocytes at this time are proliferative and $Sox9^+$, $Runx2^-$ (Figures S5A and S5B).

We observed a strong correlation in gene expression between rib and nasal chondrocytes (Figure 7A). Sox9 bound Class II regions recovered from rib chondrocytes were largely bound by Sox9 in nasal chondrocytes and vise versa (Figure 7B). The major differences reflect differences in the weakest component of the dataset (Figure 7B, intensity plots); the higher ranking Sox9 peaks in rib chondrocyte data show a higher relative association with the top peaks in nasal chondrocytes than the entire Sox9 Class II peak set (Figure S5C). GREAT GO showed skeletal related terms highly enriched in nasal chondrocytes (Figure S5D). S-score analysis predicted a similar set of Sox9 target genes to those identified in rib chondrocytes (Figure S5E and Sheet 7 in Table S3) and skeletal development-related terms are significantly enriched in the target gene set (Figure S5F). These findings suggest that when Sox9 is activated in distinct skeletal lineages, Sox9 engages a similar set of target genes through a highly similar set of *cis*-regulatory modules.

Interestingly, despite the high correlation amongst predicted Sox9 regulated enhancers some differences are observed. We found a small number of peaks with differential signal intensities for Sox9 ChIP-seq between these two populations. To distinguish the biological significance of regions that had significantly higher Sox9 peak intensity in one population compared to the other, we performed GREAT GO analysis on such peak regions identified by peak intensity MA plots (Figure 7C). Remarkably, when nasal chondrocytes were compared to rib chondrocytes, Sox9 peak regions that had significantly higher peak intensity ($\log_2$ (nasal/rib) > 2, $-\log_{10}P > 10$) associated with disease terms reflecting abnormal palate and craniofacial development (Figure 7D, Blue panel; Sheet 8 in Table S3). Conversely, regions with higher Sox9 binding intensity in rib than in nasal chondrocytes ($\log_2$(rib/nasal) > 2, $-\log_{10}P > 10$) were associated with genes whose mutation caused abnormal skeletal or cartilage development (Figure 7D, Red panel; Sheet 9 in Table S3). For example, among genes associated with each term in Figure 7D (Sheet 10 in Table S3), the *Eya1* gene showed a higher expression level and stronger Sox9 association in nasal chondrocytes whereas *Hoxc8* expression and Sox9 association in flanking regions were higher in rib chondrocytes (Figure 7E and Sheet 11 in Table S3). In summary, general Sox9 binding profiles and identified Sox9 targets are similar between rib and nasal chondrocytes though variation in the levels of Sox9 binding, and a small number of qualitative differences, may play a significant role in regulatory outcomes within distinct chondrocyte populations.

DISCUSSION

Whereas a significant body of information has accrued on epigenetic control of mammalian gene activity from systematic study of some well-studied cell types in culture, much less is known about relevant cell types *in vivo* (Consortium et al., 2007; Maher, 2012; Rosenbloom et al., 2010). The mammalian skeleton shows a relatively simple, spatially defined differentiation program where cell-lineage and genetic studies have established cellular and regulatory hierarchies. Chondrocytes and osteoblasts arise from multiple lineages where a small number of transcriptional regulators play a predominant role in specifying skeletal cell types (Akiyama et al., 2002; Bi et al., 1999; Dy et al., 2012; Komori et al., 1997). In

particular, Sox9 has emerged as a key regulator of chondrogenic programs (Akiyama et al., 2002; Bi et al., 1999; Dy et al., 2012) and Runx2 and Sp7/Osx critical determinants of osteoblast development (Komori et al., 1997; Nakashima et al., 2002). The skeletal system and these cell-type specific regulatory factors are attractive targets for developing a deep understanding of regulatory interactions shaping the mammalian body plan.

Our studies provide many important insights into the Sox9-directed process of chondrogenesis. The analysis of Class I peaks shows broad, non-specific engagement of Sox9 at promoter regions within chondrocytes that is likely mediated through protein-protein interactions with the basal transcription apparatus. Ectopic expression studies suggest that Sox9 engagement may elevate overall expression levels. The analysis of Sox9 at non-promoter associated Class II peaks indicates direct DNA binding through dimeric Sox9 recognition within evolutionarily conserved enhancer elements. These enhancers cluster around highly expressed cartilage-related genes, multiple enhancers displaying active chromatin signatures extending for many 10's to 100's of kbs from the TSS of the target gene. *In vitro* analysis of recovered Sox9 motifs and *in vivo* enhancer organization argues for a mode of regulation favoring the use of multiple enhancers each with a sub-optimal Sox9 binding ability that form super-enhancer-like groupings, which likely underpins the high levels of gene expression observed for the many chondrocyte matrix encoding genes. Interestingly, Sox9 utilizes a qualitatively similar enhancer set regardless of chondrocyte origin though where there are key genes that distinguish cranial and axial chondrocytes, notably *Hox* genes in the latter, novel patterns of Sox9 engagement are observed in the chondrocyte lineages.

These data provide a comprehensive, genome-scale assessment of Sox9's action within the nucleus of the normal chondrocyte. A limited analysis of Sox9 engagement around promoter regions of a selected sub-set of targets has been reported from the study of a rat chondrosarcoma line (Oh et al., 2010). These studies were recently extended by RNA-seq and Sox9 ChIP-seq and included some analysis of primary chondrocytes (Oh et al., 2014). Unfortunately, we were unable to obtain the primary data underpinning these studies so we have been unable to corroborate, compare or incorporate the findings of these authors into our work.

Distinct engagement of Sox9 in chondrocytes

The two classes of Sox9 binding on the genome point to distinct function of Sox9 in the regulation of gene activity in chondrocytes. Class I binding correlates strikingly, both quantitatively and spatially, with engagement of p300, RNA polymerase II and presumably a spectrum of other components linked to the basal transcriptional apparatus. Further, the absence of enrichment of Sox9 binding motifs in this data set suggests an indirect mode of engagement with one or more of the above. Sox9 is known to associate with p300; consequently, their interaction is likely to account for some, and possibly all, of the Class I data (Furumatsu et al., 2005; Tsuda et al., 2003). This data, and analysis of gene expression following ectopic Sox9 expression in fibroblasts, suggests a model wherein Sox9 engagement with the basal transcription apparatus generally elevates gene expression. As the levels of Sox9 binding correlate with transcriptional activity rather than cell type

specificity of the target gene, it is likely that Sox9's action on Class I targets would be most significant for highly expressed genes, including house-keeping genes, independent of any particular function within the cell.

Class II binding associates with broad range of skeletal enhancers; the major cartilage matrix producing genes are prominent targets of the enhancer network. Amongst this group, we confirmed Sox9 engagement in most established enhancers previously linked to Sox9's regulation of early chondrocyte differentiation (Akiyama et al., 2002; Akiyama and Lefebvre, 2011; Bell et al., 1997), and Sox9's auto-regulation (Mead et al., 2013); Sox9 appears as a major target of its own regulation in our data with discrete binding sites extending for 100's of kilobases. Two other Sox factors, *Sox5* and *Sox6* act downstream of Sox9 (Akiyama et al., 2002) and are also predicted to be direct targets of Sox9 regulation in the current data.

More contentious is Sox9's function in maturation of hypertrophic chondrocytes. Initial studies based on mouse genetics and *in vitro* studies suggested Sox9 inhibits the hypertrophic program (Akiyama et al., 2002; Ikeda et al., 2004). Further, Sox9 is reported to repress transcription of *Col10a1*, a key gene in the production of a collagen-type restricted to post-mitotic and hypertrophic chondrocyte fates. Here, repression was reported to map to a Sox motif upstream of the *Col10a1* gene (Leung et al., 2011). However, in contrast to these findings, recent studies disrupting *Sox9* within the growth plate suggest Sox9 promotes hypertrophy, acting cooperatively with Mef2c, a master regulator of hypertrophy, to promote transcription of *Col10a1* (Dy et al., 2012). We observe a weak Sox9 peak adjacent to the region identified upstream of *Col10a1* but no Sox9 binding to the regions around *Mef2c*. Our ChIP population is predominantly comprised of mitotic chondrocytes, the prehypertrophic zone is a small fraction and hypertrophic chondrocytes are absent. Clearly, Sox9 protein levels remain elevated within *Col10a1*expressing hypertrophic chondrocytes so Sox9 may play a continuing role in maturing chondrocytes. Thus, given potential heterogeneity of our ChIP population, future ChIP-seq studies that can overcome the current barrier to extensive dissection of primary cell types and the requirement for substantial amounts of nuclear material to generate comparable robust data to that presented here, will help in clarifying Sox9 action in hypertrophic chondrocyte development.

Maintenance of expression of key chondrocyte genes through Sox9 engagement

Sox9 acts predominantly as a homodimer in chondrocytes (Bridgewater et al., 2003; Coustry et al., 2010; Genzer and Bridgewater, 2007; Jenkins et al., 2005). Our findings lend strong support to Sox9 dimers as the key regulatory input into chondrocyte *cis*-regulatory modules, which is consistent with the identification of mutations that disrupt the dimerization capability of Sox9 in patients with campomelic dysplasia (Bernard et al., 2003; Sock et al., 2003). Biochemical and statistical analysis of Sox9 motifs indicates that enhancers favor lower affinity Sox9 sites compared with the highly symmetric homodimer motif that represents the motif preference from the entire Sox9 bound enhancer dataset. The reason for this is not clear but lower affinity interactions may favor cooperative engagement at regulatory regions. As an example, Sox5 and Sox6, members of the distinct SoxD sub-family are known to act downstream of Sox9 in chondrocyte regulation. While these

members can dimerize with each other, they do not dimerize with Sox9, but engage with Sox9 in the cooperative regulation of cartilage enhancers for *Col2a1* and *Acan* (Han and Lefebvre, 2008; Lefebvre, 2010; Lefebvre et al., 1998).

In examining the enhancer landscape around target genes, and the results of enhancer analysis in transgenic studies, collectively our data suggest that the transcriptional outcome in terms of the probability of expression and the level of expression of a given target gene is likely determined through the use of multiple enhancer elements. Though our transgenic studies are limited to facial structures in the zebrafish all enhancers have broad, general cartilage specificity. Thus, at this level of analysis, enhancers are likely "level-determining modules" that sum to give appropriate levels of expression of the skeletal target gene rather than "spatial-determining modules" each acting in a unique range of skeletal elements that sum to cover the entire skeleton.

Tissue specificity of Sox9 binding modes on the genome

Sox9 is a master regulator of male sex determination and skeletal development. Further, Sox9 action is critical in different aspects of lung, hair, kidney, ear, hair and gut development (reviewed in Lefebvre et al., 2007). How are these specific outcomes realized? A study of Sox9 interactions in male sex determination focused on ChIP-chip analysis of promoter regions (Bhandari et al., 2012). However, on the basis of data here, this will likely have selected for Class I interactions and consistent with this conclusion motif recovery did not provide compelling evidence for a Sox9 binding motif. A comparative analysis of the limited set of *cis*-regulatory modules engaged by Sox9 in the testis that include Sox9-itself (Sekido and Lovell-Badge, 2008), *miR202* (Wainwright et al., 2013), and *Amh* (Arango et al., 1999; De Santa Barbara et al., 1998), none of these enhancers were bound by Sox9 in chondrocytes.

Interestingly, reporter constructs comprising 5 copies of optimized Sox9 dimer motifs show no chondrocyte activity in the zebrafish assay. This may reflect accessibility of the reporter construct if for example specific histone modification of an enhancer is required to enable accessibility of Sox9 (Zhou et al., 2011). Indeed, the testis enhancers linked to Sox9's testis regulatory program did not display poised or active histone marks (H3K4me2 or H3K27acetyation, respectively) in chondrocytes (data not shown). Sox9 binding may also be insufficient for chondrocyte expression (data not shown). Cooperative partnerships between Sox9 and Sox5/6 discussed earlier may be one important factor in determining a specific regulatory outcome (Han and Lefebvre, 2008; Lefebvre, 2010; Lefebvre et al., 1998). Cooperative interaction with non Sox-family members is also suggested by our study. Motifs for AP-1, NFAT, Fox, Runx, and Hox transcription factor families were all significantly enriched in the Sox9 ChIP-seq datasets. Given that several of these families have been implicated in formation and maintenance of cartilage (Behrens et al., 2003; Gross et al., 2012; Karreth et al., 2004; Kimura et al., 2010; Nifuji et al., 2001; Rodova et al., 2011; Wang et al., 2009; Yoshida et al., 2004; Yueh et al., 1998), their actions may modulate Sox9 engagement, and cooperate with Sox9 to regulate enhancer activity. The identification and validation of the core Sox9 network provides a regulatory framework for

extending our understanding of the transcriptional mechanisms at play in mammalian skeletogenesis.

EXPERIMENTAL PROCEDURES

ChIP-seq

Rib chondrocytes isolated from newborn mice were immediately subjected to chromatin preparation and ChIP as described (Odom et al., 2004; Vokes et al., 2007). ChIP-seq libraries were constructed using ChIP-seq DNA Sample Prep Kit (IP-102-1001; Illumina Inc., San Diego, CA). Sequencing was performed on Genome Analyzer II (Illumina).

Expression profiling

RNA was isolated form rib or nasal chondrocytes using TRIZOL reagent (15596-026; Life Technologies) and RNeasy Mini Kit (74104; QIAGEN) according to manufacturers' instruction. Microarray analysis were performed on GeneChip mouse Gene 1.0 ST array (901169; Affymetrix, Santa Clara, CA). RNA-seq analysis was performed on the Illumina platform with Nextseq 500.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Akiyama H, Chaboissier MC, Martin JF, Schedl A, de Crombrugghe B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. Genes Dev. 2002; 16:2813–2828. [PubMed: 12414734]

Akiyama H, Kim JE, Nakashima K, Balmes G, Iwai N, Deng JM, Zhang Z, Martin JF, Behringer RR, Nakamura T, et al. Osteo-chondroprogenitor cells are derived from Sox9 expressing precursors. Proc Natl Acad Sci U S A. 2005; 102:14665–14670. [PubMed: 16203988]

Akiyama H, Lefebvre V. Unraveling the transcriptional regulatory machinery in chondrogenesis. J Bone Miner Metab. 2011; 29:390–395. [PubMed: 21594584]

Arango NA, Lovell-Badge R, Behringer RR. Targeted mutagenesis of the endogenous mouse Mis gene promoter: in vivo definition of genetic pathways of vertebrate sexual development. Cell. 1999; 99:409–419. [PubMed: 10571183]

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009; 37:W202–208. [PubMed: 19458158]

Behrens A, Haigh J, Mechta-Grigoriou F, Nagy A, Yaniv M, Wagner EF. Impaired intervertebral disc formation in the absence of Jun. Development. 2003; 130:103–109. [PubMed: 12441295]

Bell DM, Leung KK, Wheatley SC, Ng LJ, Zhou S, Ling KW, Sham MH, Koopman P, Tam PP, Cheah KS. SOX9 directly regulates the type-II collagen gene. Nat Genet. 1997; 16:174–178. [PubMed: 9171829]

Bernard P, Tang P, Liu S, Dewing P, Harley VR, Vilain E. Dimerization of SOX9 is required for chondrogenesis, but not for sex determination. Hum Mol Genet. 2003; 12:1755–1765. [PubMed: 12837698]

Bessa J, Tena JJ, de la Calle-Mustienes E, Fernandez-Minan A, Naranjo S, Fernandez A, Montoliu L, Akalin A, Lenhard B, Casares F, et al. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. Dev Dyn. 2009; 238:2409–2417. [PubMed: 19653328]

Bhandari RK, Haque MM, Skinner MK. Global genome analysis of the downstream binding targets of testis determining factor SRY and SOX9. PLoS One. 2012; 7:e43380. [PubMed: 22984422]

Bi W, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B. Sox9 is required for cartilage formation. Nat Genet. 1999; 22:85–89. [PubMed: 10319868]

Bi W, Huang W, Whitworth DJ, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B. Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. Proc Natl Acad Sci U S A. 2001; 98:6698–6703. [PubMed: 11371614]

Bridgewater LC, Lefebvre V, de Crombrugghe B. Chondrocyte-specific enhancer elements in the Col11a2 gene resemble the Col2a1 tissue-specific enhancer. J Biol Chem. 1998; 273:14998–15006. [PubMed: 9614107]

Bridgewater LC, Walker MD, Miller GC, Ellison TA, Holsinger LD, Potter JL, Jackson TL, Chen RK, Winkel VL, Zhang Z, et al. Adjacent DNA sequences modulate Sox9 transcriptional activation at paired Sox sites in three chondrocyte-specific enhancer elements. Nucleic Acids Res. 2003; 31:1541–1553. [PubMed: 12595563]

Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007; 447:799–816. [PubMed: 17571346]

Coustry F, Oh CD, Hattori T, Maity SN, de Crombrugghe B, Yasuda H. The dimerization domain of SOX9 is required for transcription activation of a chondrocyte-specific chromatin DNA template. Nucleic Acids Res. 2010; 38:6018–6028. [PubMed: 20484372]

De Santa Barbara P, Bonneaud N, Boizet B, Desclozeaux M, Moniot B, Sudbeck P, Scherer G, Poulat F, Berta P. Direct interaction of SRY-related protein SOX9 and steroidogenic factor 1 regulates transcription of the human anti-Mullerian hormone gene. Mol Cell Biol. 1998; 18:6653–6665. [PubMed: 9774680]

Dy P, Wang W, Bhattaram P, Wang Q, Wang L, Ballock RT, Lefebvre V. Sox9 directs hypertrophic maturation and blocks osteoblast differentiation of growth plate chondrocytes. Dev Cell. 2012; 22:597–609. [PubMed: 22421045]

Foster JW, Dominguez-Steglich MA, Guioli S, Kwok C, Weller PA, Stevanovic M, Weissenbach J, Mansour S, Young ID, Goodfellow PN, et al. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature. 1994; 372:525–530. [PubMed: 7990924]

Furumatsu T, Tsuda M, Yoshida K, Taniguchi N, Ito T, Hashimoto M, Ito T, Asahara H. Sox9 and p300 cooperatively regulate chromatin-mediated transcription. J Biol Chem. 2005; 280:35203–35208. [PubMed: 16109717]

Genzer MA, Bridgewater LC. A Col9a1 enhancer element activated by two interdependent SOX9 dimers. Nucleic Acids Res. 2007; 35:1178–1186. [PubMed: 17264118]

Gross S, Krause Y, Wuelling M, Vortkamp A. Hoxa11 and Hoxd11 regulate chondrocyte differentiation upstream of Runx2 and Shox2 in mice. PLoS One. 2012; 7:e43553. [PubMed: 22916278]

Han Y, Lefebvre V. L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer. Mol Cell Biol. 2008; 28:4999–5013. [PubMed: 18559420]

He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q, Zhang Y, Xu K, Ni M, Lupien M, et al. Nucleosome dynamics define transcriptional enhancers. Nat Genet. 2010; 42:343–347. [PubMed: 20208536]

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007; 39:311–318. [PubMed: 17277777]

Helms JA, Schneider RA. Cranial skeletal biology. Nature. 2003; 423:326–331. [PubMed: 12748650]

Hu G, Codina M, Fisher S. Multiple enhancers associated with ACAN suggest highly redundant transcriptional regulation in cartilage. Matrix Biol. 2012; 31:328–337. [PubMed: 22820679]

Huang BL, Brugger SM, Lyons KM. Stage-specific control of connective tissue growth factor (CTGF/CCN2) expression in chondrocytes by Sox9 and beta-catenin. J Biol Chem. 2010; 285:27702–27712. [PubMed: 20571031]

Ikeda T, Kamekura S, Mabuchi A, Kou I, Seki S, Takato T, Nakamura K, Kawaguchi H, Ikegawa S, Chung UI. The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. Arthritis Rheum. 2004; 50:3561–3573. [PubMed: 15529345]

Jenkins E, Moss JB, Pace JM, Bridgewater LC. The new collagen gene COL27A1 contains SOX9-responsive enhancer elements. Matrix Biol. 2005; 24:177–184. [PubMed: 15922909]

Ji H, Jiang H, Ma W, Johnson DS, Myers RM, Wong WH. An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nat Biotechnol. 2008; 26:1293–1300. [PubMed: 18978777]

Kan A, Ikeda T, Saito T, Yano F, Fukai A, Hojo H, Ogasawara T, Ogata N, Nakamura K, Chung UI, et al. Screening of chondrogenic factors with a real-time fluorescence-monitoring cell line ATDC5-C2ER: identification of sorting nexin 19 as a novel factor. Arthritis Rheum. 2009; 60:3314–3323. [PubMed: 19877062]

Karreth F, Hoebertz A, Scheuch H, Eferl R, Wagner EF. The AP1 transcription factor Fra2 is required for efficient cartilage development. Development. 2004; 131:5717–5725. [PubMed: 15509771]

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010; 465:182–187. [PubMed: 20393465]

Kimura A, Inose H, Yano F, Fujita K, Ikeda T, Sato S, Iwasaki M, Jinno T, Ae K, Fukumoto S, et al. Runx1 and Runx2 cooperate during sternal morphogenesis. Development. 2010; 137:1159–1167. [PubMed: 20181744]

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M, et al. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. Cell. 1997; 89:755–764. [PubMed: 9182763]

Kou I, Ikegawa S. SOX9-dependent and -independent transcriptional regulation of human cartilage link protein. J Biol Chem. 2004; 279:50942–50948. [PubMed: 15456769]

Kronenberg HM. Developmental regulation of the growth plate. Nature. 2003; 423:332–336. [PubMed: 12748651]

Lefebvre V. The SoxD transcription factors--Sox5, Sox6, and Sox13--are key cell fate modulators. Int J Biochem Cell Biol. 2010; 42:429–432. [PubMed: 19647094]

Lefebvre V, Dumitriu B, Penzo-Mendez A, Han Y, Pallavi B. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int J Biochem Cell Biol. 2007; 39:2195–2214. [PubMed: 17625949]

Lefebvre V, Li P, de Crombrugghe B. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. Embo J. 1998; 17:5718–5733. [PubMed: 9755172]

Lefebvre V, Zhou G, Mukhopadhyay K, Smith CN, Zhang Z, Eberspaecher H, Zhou X, Sinha S, Maity SN, de Crombrugghe B. An 18-base-pair sequence in the mouse proalpha1(II) collagen gene is sufficient for expression in cartilage and binds nuclear proteins that are selectively expressed in chondrocytes. Mol Cell Biol. 1996; 16:4512–4523. [PubMed: 8754852]

Leung KK, Ng LJ, Ho KK, Tam PP, Cheah KS. Different cis-regulatory DNA elements mediate developmental stage- and tissue-specific expression of the human COL2A1 gene in transgenic mice. J Cell Biol. 1998; 141:1291–1300. [PubMed: 9628886]

Leung VY, Gao B, Leung KK, Melhado IG, Wynn SL, Au TY, Dung NW, Lau JY, Mak AC, Chan D, et al. SOX9 governs differentiation stage-specific gene expression in growth plate chondrocytes via direct concomitant transactivation and repression. PLoS Genet. 2011; 7:e1002356. [PubMed: 22072985]

Liu Y, Li H, Tanaka K, Tsumaki N, Yamada Y. Identification of an enhancer sequence within the first intron required for cartilage-specific transcription of the alpha2(XI) collagen gene. J Biol Chem. 2000; 275:12712–12718. [PubMed: 10777565]

Maher B. ENCODE: The human encyclopaedia. Nature. 2012; 489:46–48. [PubMed: 22962707]

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28:495–501. [PubMed: 20436461]

Mead TJ, Wang Q, Bhattaram P, Dy P, Afelik S, Jensen J, Lefebvre V. A far-upstream (−70 kb) enhancer mediates Sox9 auto-regulation in somatic tissues during development and adult regeneration. Nucleic Acids Res. 2013; 41:4459–4469. [PubMed: 23449223]

Nagy A, Kenesi E, Rentsendorj O, Molnar A, Szenasi T, Sinko I, Zvara A, Oommen ST, Barta E, Puskas LG, et al. Evolutionarily conserved, growth plate zone-specific regulation of the matrilin-1 promoter: L-Sox5/Sox6 and Nfi factors bound near TATA finely tune activation by Sox9. Mol Cell Biol. 2011; 31:686–699. [PubMed: 21173167]

Nakashima K, Zhou X, Kunkel G, Zhang Z, Deng JM, Behringer RR, de Crombrugghe B. The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. Cell. 2002; 108:17–29. [PubMed: 11792318]

Nifuji A, Miura N, Kato N, Kellermann O, Noda M. Bone morphogenetic protein regulation of forkhead/winged helix transcription factor Foxc2 (Mfh1) in a murine mesodermal cell line C1 and in skeletal precursor cells. J Bone Miner Res. 2001; 16:1765–1771. [PubMed: 11585339]

Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, et al. Control of pancreas and liver gene expression by HNF transcription factors. Science. 2004; 303:1378–1381. [PubMed: 14988562]

Oh CD, Lu Y, Liang S, Mori-Akiyama Y, Chen D, de Crombrugghe B, Yasuda H. SOX9 regulates multiple genes in chondrocytes, including genes encoding ECM proteins, ECM modification enzymes, receptors, and transporters. PLoS One. 2014; 9:e107577. [PubMed: 25229425]

Oh CD, Maity SN, Lu JF, Zhang J, Liang S, Coustry F, de Crombrugghe B, Yasuda H. Identification of SOX9 interaction sites in the genome of chondrocytes. PLoS One. 2010; 5:e10113. [PubMed: 20404928]

Olsen BR, Reginato AM, Wang W. Bone development. Annu Rev Cell Dev Biol. 2000; 16:191–220. [PubMed: 11031235]

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011; 470:279–283. [PubMed: 21160473]

Rentsendorj O, Nagy A, Sinko I, Daraba A, Barta E, Kiss I. Highly conserved proximal promoter element harbouring paired Sox9-binding sites contributes to the tissue- and developmental stage-specific activity of the matrilin-1 gene. Biochem J. 2005; 389:705–716. [PubMed: 15804237]

Rodova M, Lu Q, Li Y, Woodbury BG, Crist JD, Gardner BM, Yost JG, Zhong XB, Anderson HC, Wang J. Nfat1 regulates adult articular chondrocyte function through its age-dependent expression mediated by epigenetic histone methylation. J Bone Miner Res. 2011; 26:1974–1986. [PubMed: 21452283]

Rosenbloom KR, Dreszer TR, Pheasant M, Barber GP, Meyer LR, Pohl A, Raney BJ, Wang T, Hinrichs AS, Zweig AS, et al. ENCODE whole-genome data in the UCSC Genome Browser. Nucleic Acids Res. 2010; 38:D620–625. [PubMed: 19920125]

Schafer AJ, Foster JW, Kwok C, Weller PA, Guioli S, Goodfellow PN. Campomelic dysplasia with XY sex reversal: diverse phenotypes resulting from mutations in a single gene. Ann N Y Acad Sci. 1996; 785:137–149. [PubMed: 8702120]

Sekido R, Lovell-Badge R. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature. 2008; 453:930–934. [PubMed: 18454134]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sock E, Pagon RA, Keymolen K, Lissens W, Wegner M, Scherer G. Loss of DNA-dependent dimerization of the transcription factor SOX9 as a cause for campomelic dysplasia. Hum Mol Genet. 2003; 12:1439–1447. [PubMed: 12783851]

Tang Q, Chen Y, Meyer C, Geistlinger T, Lupien M, Wang Q, Liu T, Zhang Y, Brown M, Liu XS. A comprehensive view of nuclear receptor cancer cistromes. Cancer Res. 2011; 71:6940–6947. [PubMed: 21940749]

Tsuda M, Takahashi S, Takahashi Y, Asahara H. Transcriptional co-activators CREB-binding protein and p300 regulate chondrocyte-specific gene expression via association with Sox9. J Biol Chem. 2003; 278:27224–27229. [PubMed: 12732631]

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009; 457:854–858. [PubMed: 19212405]

Vokes SA, Ji HK, McCuine S, Tenzen T, Giles S, Zhong S, Longabaugh WJR, Davidson EH, Wong WH, McMahon AP. Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. Development. 2007; 134:1977–1989. [PubMed: 17442700]

Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Bricarelli FD, Keutel J, Hustert E, et al. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. Cell. 1994; 79:1111–1120. [PubMed: 8001137]

Wainwright EN, Jorgensen JS, Kim Y, Truong V, Bagheri-Fam S, Davidson T, Svingen T, Fernandez-Valverde SL, McClelland KS, Taft RJ, et al. SOX9 regulates microRNA miR-202-5p/3p expression during mouse testis differentiation. Biol Reprod. 2013; 89:34. [PubMed: 23843232]

Wang J, Gardner BM, Lu Q, Rodova M, Woodbury BG, Yost JG, Roby KF, Pinson DM, Tawfik O, Anderson HC. Transcription factor Nfat1 deficiency causes osteoarthritis through dysfunction of adult articular chondrocytes. J Pathol. 2009; 219:163–172. [PubMed: 19526482]

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013; 153:307–319. [PubMed: 23582322]

Wright E, Hargrave MR, Christiansen J, Cooper L, Kun J, Evans T, Gangadharan U, Greenfield A, Koopman P. The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. Nat Genet. 1995; 9:15–20. [PubMed: 7704017]

Yoshida CA, Yamamoto H, Fujita T, Furuichi T, Ito K, Inoue K, Yamana K, Zanma A, Takada K, Ito Y, et al. Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. Genes Dev. 2004; 18:952–963. [PubMed: 15107406]

Yueh YG, Gardner DP, Kappen C. Evidence for regulation of cartilage differentiation by the homeobox gene Hoxc-8. Proc Natl Acad Sci U S A. 1998; 95:9956–9961. [PubMed: 9707582]

Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. 2011; 21:1273–1283. [PubMed: 21632746]

Zhang P, Jimenez SA, Stokes DG. Regulation of human COL9A1 gene expression. Activation of the proximal promoter region by SOX9. J Biol Chem. 2003; 278:117–123. [PubMed: 12399468]

Zhou G, Lefebvre V, Zhang Z, Eberspaecher H, de Crombrugghe B. Three high mobility group-like sequences within a 48-base pair enhancer of the Col2a1 gene are required for cartilage-specific expression in vivo. J Biol Chem. 1998; 273:14989–14997. [PubMed: 9614106]

Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nature reviews Genetics. 2011; 12:7–18.
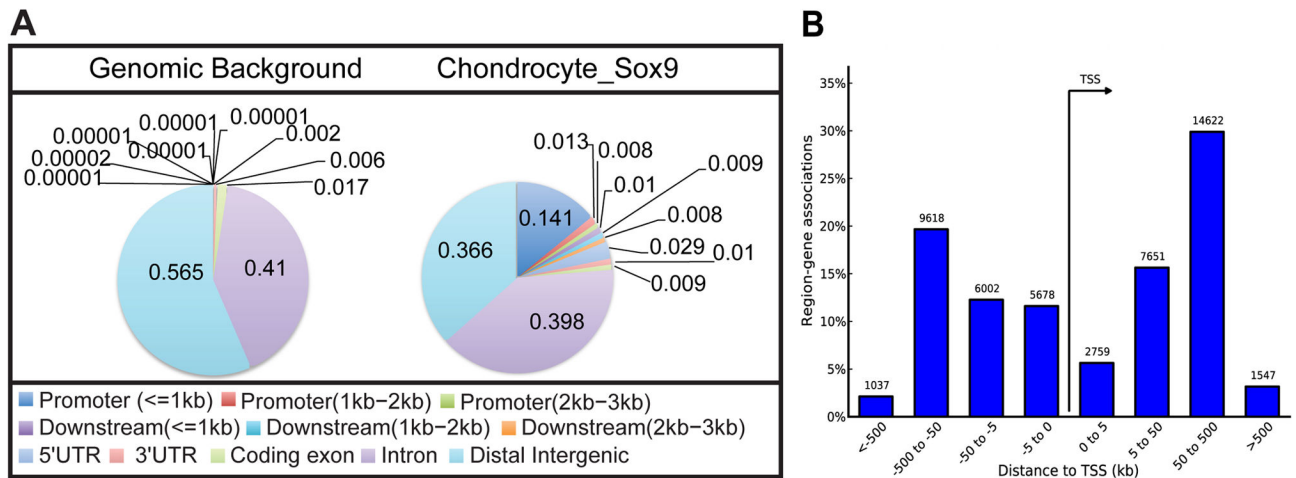
Figure 1. Peak statistics and annotation of Sox9 ChIP-seq data sets obtained from rib chondrocytes

(A) Distribution of Sox9 peaks relative to predicted ref-seq genes in comparison with the representation of different regions within the mouse genome. Color-coded distance windows are shown in the same clockwise order as in pie chart.

(B) Distribution of Sox9 peaks relative to TSSs analyzed by GREAT GO tools. The Y-axis represents the percentage of peaks in each distance category; peak counting is indicated on each bar.
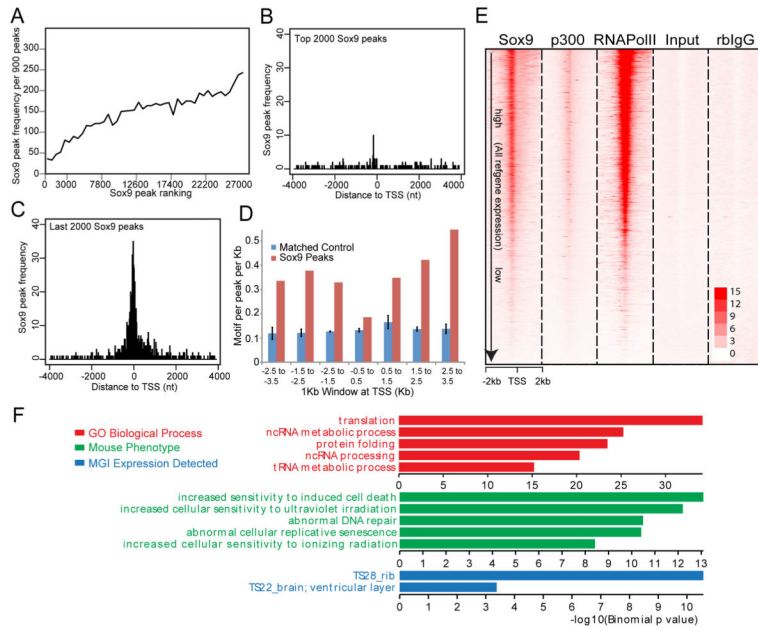
Figure 2. TSS-associated Class I Sox9 ChIP-seq peaks

(A) Correlation of frequencies and peak calling scores at Class I Sox9 peaks. The X-axis represents consecutive bins of high (left) to low (right) peaks with 900 peaks per bin. The Y-axis plots the frequency of Class I Sox9 peaks within each of the binned peak sets.

(B and C) Mapping of top 2,000 ranked and bottom 2,000 ranked Sox9 peaks in relation to the nearest TSS. The X-axis represents distance from the peak center to the nearest TSS and, the Y-axis represents the frequency of Sox9 peak enrichment.

(D) Sox dimeric motif frequency in Sox9 peaks around the TSS. Sox9 peaks around TSSs (+/− 3.5 kb from TSSs) were grouped into bins of 1 kb, and the frequencies of Sox dimeric motifs per peak per kb were plotted.

(E) Signal intensity plots of ChIP-seq data for Sox9, p300, and RNA Pol II in promoter regions (+/− 2 kb from TSSs) compared to input and rabbit (rb) IgG controls in relation to expressions of their associated genes. Class I Sox9 peaks were ordered according to the expression level of the nearest associated gene (highest expression level at top). ChIP-seq signals for RNA Pol II and p300 were mapped to corresponding Sox9 peak regions.

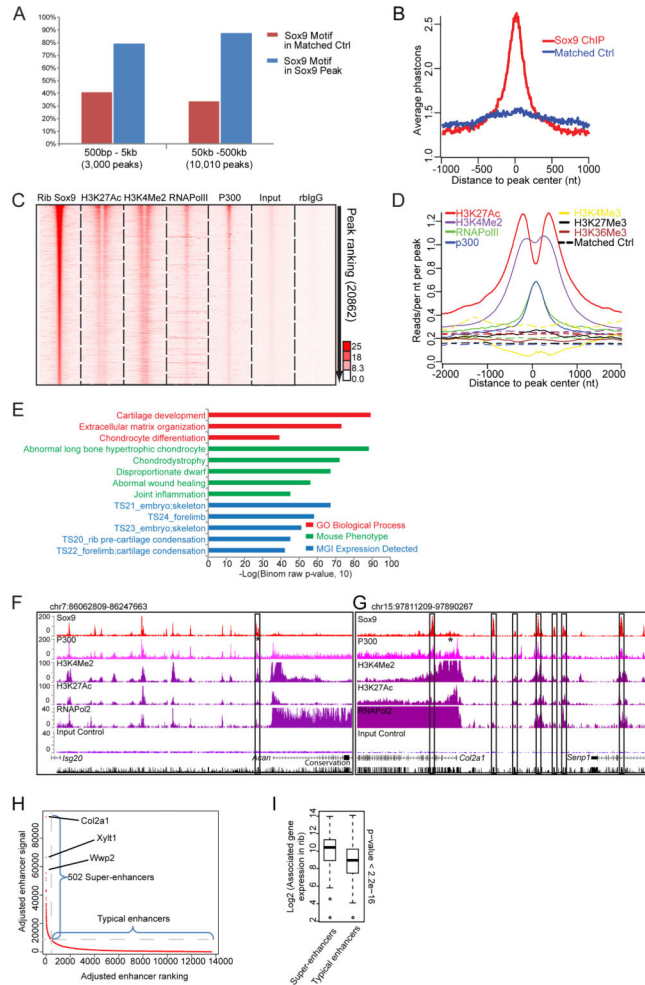(F) GREAT GO analysis of Class I Sox9 peaks.

Figure 3. Engagement of the Class II Sox9 ChIP-seq peaks in chondrocyte development

(A) Sox9 motif enrichment in the indicated windows from the nearest associated TSS relative to position matched random control regions that are calculated from the genome. Blue, percentage of peaks carrying at least one Sox dimeric motif in Sox9 peaks; red, percentage of regions carrying Sox dimeric motifs in the control regions.

(B) Conservation of Class II Sox9 peak regions at +/− 1 kb from peak centers analyzed by phastcons in comparison with position matched control regions in the genome.

(C) Peak intensity plots of ChIP-seq reads for Sox9, histone modification marks, components of transcriptional machinery, and ChIP controls (input and rb-IgG samples). Normalized read signals from each ChIP-seq were plotted on a 4 kb window centered on the Sox9 peak center. Class II Sox9 peaks were ordered according to their peak ranking (highest ranking at top). ChIP-seq signals for H3K27Ac, H3K4Me2, RNA Pol II and p300 were mapped to corresponding Sox9 peak regions.

(D) Normalized mean signal intensity for the following ChIP-seq reads: H3K27Ac (red), H3K4Me2 (purple), RNA Pol II (green), p300 (blue), H3K4Me3 (yellow), H3K27Me3 (black), and H3K36Me3 (brown) centered around Class II Sox9 peaks. The Y-axis

represents normalized read counts per bp per peak per 10 million reads. Background signal intensities were determined from complementary position matched control data.

(E) GO biological processes, mouse phenotypes, and MGI expression annotations of Class II Sox9 peaks analyzed by GREAT GO tools.

(F and G) CisGenome browser views of Sox9 peaks and associated chromatin, p300 and polymerase binding around *Acan* (*Aggrecan*) and *Col2a1*, as well as input control values and a conservation index. Asterisks indicate enhancer regions verified in previous studies (see also Table S6). Boxes highlight functionally validated enhancers in our study that have not been previously reported in the literature.

(H) Supper-enhancer predictions from Class II Sox9 ChIP-seq data sets. Adjusted enhancer ranking and signal intensity were plotted to show the distribution of super-enhancers relative to typical enhancers. The top three super-enhancer target genes are highlighted. See also Table S3.

(I) Comparison of relative expression of genes associated with Sox9-bound super-enhancers and typical enhancers. Box plots show gene expressions in each enhancer group. See also Table S3.
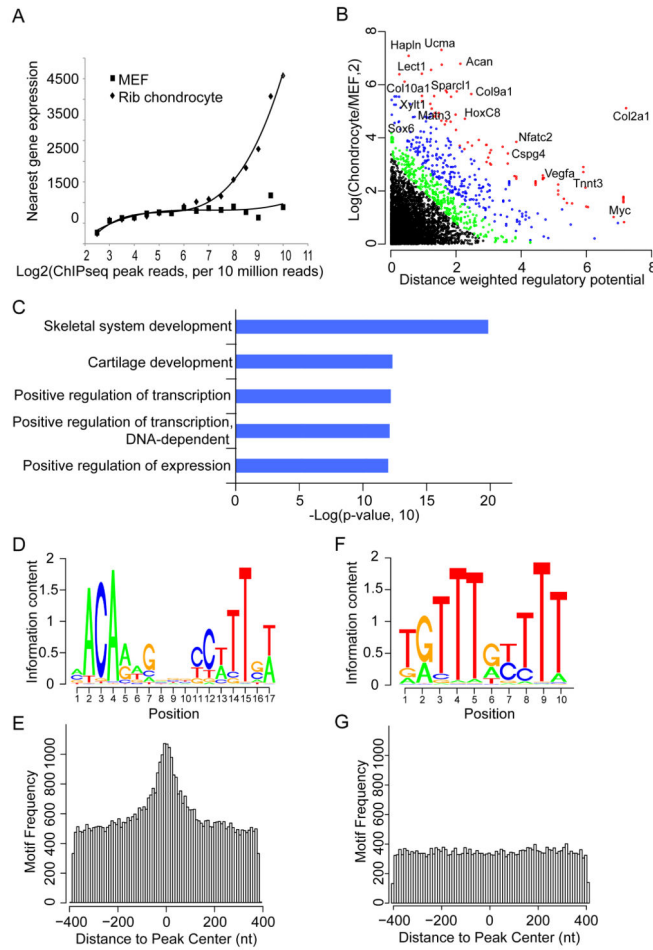
Figure 4. Prediction of Sox9 direct targets and *de novo* motif recovery in Sox9 ChIP-seq peaks

(A) Correlation of Sox9 peak signal with expression of the nearest gene. Sox9 ChIP-seq signal for all nearby peaks around each ref-seq gene was compared to the expression levels of the relevant gene in mouse chondrocytes and embryonic mouse fibroblasts.

(B) Intersection of Sox9 peak reads and gene expression levels using S-score analysis. Red, FDR 1%; Blue, 1%<FDR 5%; Green, 5%<FDR 10%; Black, FDR>10%.

(C) DAVID GO analysis for putative Sox9 target genes predicted by the S-score analysis from FDR1% to FDR10%. Top five enriched GO terms are shown.

(D–G) Enrichment of Sox dimeric motifs (D and E) and monomer motifs (F and G) recovered *de novo* from Sox9 peak regions. Motif logos of the recovered Sox motifs, which display nucleotide frequencies (scaled relative to the information content) at each position, are shown in (D) and (F); enrichment level of the Sox motifs mapped to Class II Sox9 peaks is shown in (E) and (G), where the X- and Y-axis represent distance from mapped motifs to the peak center and frequency of mapped motifs, respectively. See also Figure S3.
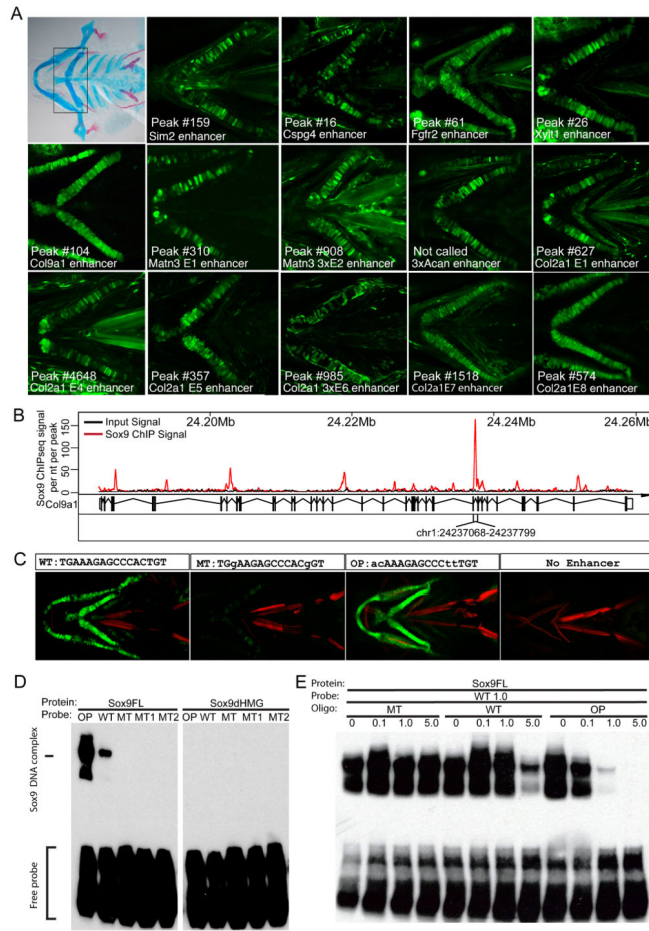
Figure 5. Validation of enhancer activities of Class II Sox9 ChIP-seq peak regions *in vivo* and *in vitro*

(A) Enhancer activity testing for regions predicted by Sox9 ChIP-seq analysis utilizing the ZED assay. Alcian blue and Alizarin red staining of a 5-day-old zebrafish embryo pharyngeal skeleton is shown in the top left panel. Boxed area in the panel indicates ceratohyal cartilage used to determine *in vivo* enhancer activities. In some instances, three tandem copies (3x) of the enhancer sequence were assayed. Note that the one test region for *Acan* was not identified as a Sox9 peak region by our stringent peak calling process, although substantial Sox9 ChIP-seq signal was observed. The numbers (#) indicate the ranking of the Sox9 ChIP-seq peak in that region.

(B) Sox9 ChIP-seq signal at *Col9a1* gene region. The strong Sox9 peak between exon 29 and 30 of *Col9a1* (chr1: 24237068–24237799) was subjected to enhancer analysis in (C).

(C) The effect of two mutations in Sox9 motif (MT) or the replacement of endogenous Sox9 motif with predicted optimized Sox9 motif (OP) on the enhancer activity of the *Col9a1* enhancer. Control lines carried the reporter construct but no test enhancer sequence.

(D) Validation of Sox9 binding to Sox9 motif in *Col9a1* enhancer tested in (A) and (C). EMSA was performed using labeled probes carrying endogenous (WT), mutated (MT, MT1, and MT2; Supplemental Experimental Procedures), and optimal (OP) Sox dimeric motif.

Full-length Sox9 protein (Sox9FL) and HMG domain-deleted Sox9 protein (Sox9dHMG) were examined.

(E) EMSA competition experiment using unlabeled MT, WT, and OP oligos with biotin-labeled WT probe used in (D). Numbers indicate relative molarity amount of unlabeled competitor to the labeled WT probe.
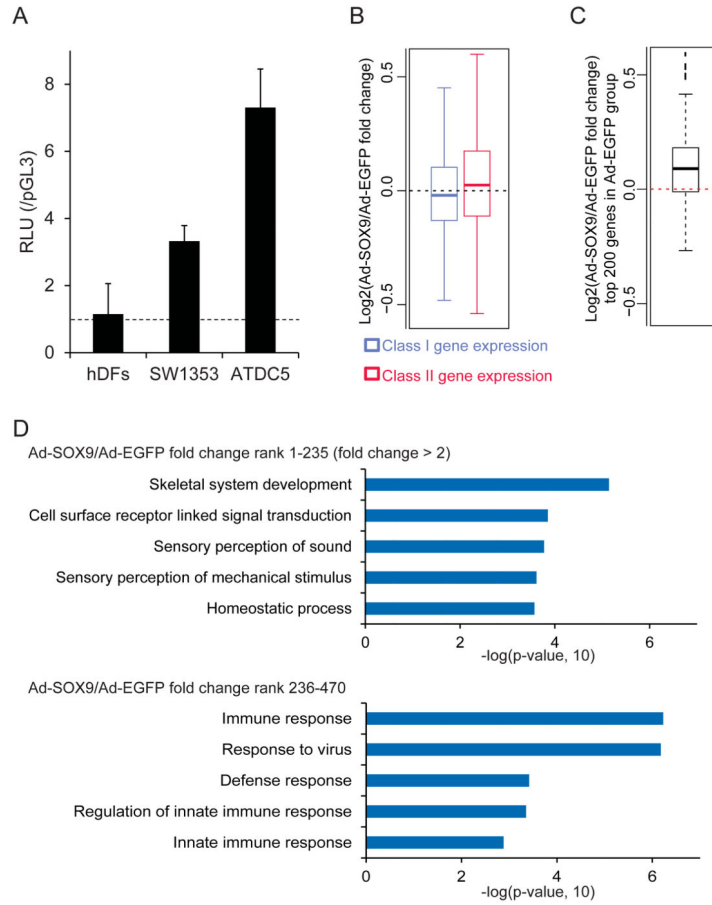
Figure 6. Induction of Class I and Class II targets upon ectopic *SOX9* expression in human dermal fibroblasts

(A) Luciferase reporter assay using the luciferase reporter construct driven by 4 copies of the 48-bp *COL2A1* intron 1 enhancer in human dermal fibroblasts (hDFs), SW1353, and ATDC5 cells. RLU, relative light units of experimental and control reporter vector.

(B) Comparison of expression changes of Class I and Class II targets upon *SOX9* overexpression in hDFs. Box plots show log$_2$ fold changes of gene expression in the Ad-*SOX9*-infected group compared to the Ad-*EGFP*-infected one.

(C) *SOX9* overexpression-induced expression changes for highly expressed genes in hDFs deduced from RNA-seq analysis. The top 200 highly expressed genes in the Ad-*EGFP*-infected hDFs were compared between Ad-*EGFP* and Ad-*SOX9*-infected groups. Box plot shows log$_2$ fold changes of gene expression in the Ad-*SOX9*-infected group compared to the Ad-*EGFP*-infected group.

(D) David GO analysis of the gene set displaying a 2-fold or greater increase on *SOX9* overexpression (upper, top 235 genes in the ranking of Ad-SOX9/Ad-EGFP fold change) and the next set of 235 genes (lower, rank 236 – 470 in the ranking). Top 5 enriched GO terms are shown with associated p-values. See also Table S7.
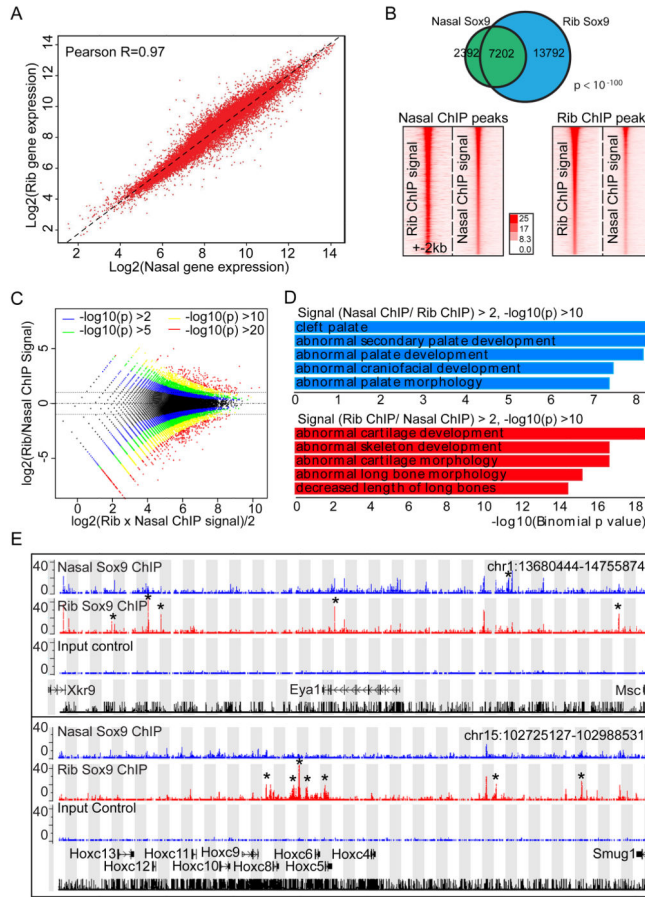
Figure 7. Comparison of gene expression profile and Sox9 binding regions between rib and nasal chondrocytes

(A) Correlation of gene expression profile between rib and nasal chondrocytes. Different dots represent individual genes. The X- and Y-axis show log2 expression level of each gene in rib chondrocytes and nasal chondrocytes, respectively.

(B) Comparison of Sox9 peaks between rib and nasal chondrocytes. Venn diagram for Sox9 peak intersection between rib and nasal chondrocytes (upper) and intensity plots of Sox9 ChIP-seq signal in nasal and rib chondrocytes (lower) are shown. In lower panels, rib and nasal Sox9 peaks are plotted on a 4 kb window from peak centers according to their ranking (highest ranked at top). Lower left panel indicates Sox9 ChIP-seq signal obtained from rib chondrocytes and lower right panel Sox9 ChIP-seq signal from nasal chondrocytes.

(C) Log-intensity ratios (M-values) versus log-intensity averages (A-values) (MA) plot of rib and nasal chondrocyte Sox9 ChIP-seq peaks. Colors represent different p-value levels for confidence levels.

(D) GREAT GO analysis on peak regions identified by peak intensity MA plots shown in (C). Blue bars, enrichments of terms in regions where Sox9 peak intensities are higher in nasal chondrocytes than in rib chondrocytes; Red bars, enrichments of terms in regions where Sox9 peak intensities were higher in rib chondrocytes than in nasal chondrocytes.

(E) CisGenome browser views of Sox9 ChIP-seq peaks around *Eya1* (upper) and *Hoxc8* (lower) in nasal and rib chondrocytes. Asterisks highlight differences in Sox9 peak predictions between nasal and rib chondrocyte data sets.