Serveur Académique Lausannois SERVAL serval.unil.ch

Author Manuscript
Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but dos not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

Title: Identification of novel craniofacial regulatory domains located far upstream of SOX9 and disrupted in Pierre Robin sequence.
Authors: Gordon CT, Attanasio C, Bhatia S, Benko S, Ansari M, Tan TY, Munnich A, Pennacchio LA, Abadie V, Temple IK, Goldenberg A, van Heyningen V, Amiel J, FitzPatrick D, Kleinjan DA, Visel A, Lyonnet S
Journal: Human mutation
Year: 2014 Aug
Volume: 35
Issue: 8
Pages: 1011-20
DOI: 10.1002/humu.22606

In the absence of a copyright statement, users should assume that standard copyright protection applies, unless the article contains an explicit statement to the contrary. In case of doubt, contact the journal publisher to verify the copyright status of an article.

serval
serveur académique lausannois

UNIL | Université de Lausanne
Faculté de biologie et de médecine

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Hum Mutat. Author manuscript; available in PMC 2015 August 01.

Published in final edited form as:

Hum Mutat. 2014 August ; 35(8): 1011–1020. doi:10.1002/humu.22606.

Identification of novel craniofacial regulatory domains located far upstream of *SOX9* and disrupted in Pierre Robin sequence

Christopher T. Gordon[1,#], **Catia Attanasio**[2,3], **Shipra Bhatia**[4], **Sabina Benko**[1,5], **Morad Ansari**[4], **Tiong Y. Tan**[6], **Arnold Munnich**[1,7], **Len A. Pennacchio**[2,8], **Véronique Abadie**[9], **I. Karen Temple**[10], **Alice Goldenberg**[11], **Veronica van Heyningen**[4], **Jeanne Amiel**[1,7], **David FitzPatrick**[4], **Dirk A. Kleinjan**[4], **Axel Visel**[2,8,12], and **Stanislas Lyonnet**[1,7,#]

[1]Université Paris Descartes–Sorbonne Paris Cité, Institut Imagine, INSERM U1163, Paris, France. [2]Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [3]Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, Switzerland. [4]MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom. [5]Department of Structural and Chemical Biology, Ichan School of Medicine at Mount Sinai, New York, USA. [6]Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Australia. [7]Hôpital Necker-Enfants Malades AP-HP, Paris, France. [8]U.S. Department of Energy Joint Genome Institute, Walnut Creek, California, USA. [9]Service de Pédiatrie Générale, Université Paris Descartes, Hôpital Necker-Enfants Malades, Paris, France. [10]Human Genetics and Genomic Medicine, Faculty of Medicine, University of Southampton, Southampton, United Kingdom. [11]Service de Génétique Médicale, CHU Charles Nicolle, Rouen, France. [12]School of Natural Sciences, University of California, Merced, California, USA.

Abstract

Mutations in the coding sequence of *SOX9* cause campomelic dysplasia (CD), a disorder of skeletal development associated with 46,XY disorders of sex development (DSDs). Translocations, deletions and duplications within a ~2 Mb region upstream of *SOX9* can recapitulate the CD-DSD phenotype fully or partially, suggesting the existence of an unusually large cis-regulatory control region. Pierre Robin sequence (PRS) is a craniofacial disorder that is frequently an endophenotype of CD and a locus for isolated PRS at ~1.2-1.5 Mb upstream of *SOX9* has been previously reported. The craniofacial regulatory potential within this locus, and within the greater genomic domain surrounding *SOX9*, remains poorly defined. We report two novel deletions upstream of *SOX9* in families with PRS, allowing refinement of the regions harbouring candidate craniofacial regulatory elements. In parallel, ChIP-Seq for p300 binding sites in mouse craniofacial tissue led to the identification of several novel craniofacial enhancers at the *SOX9* locus, which were validated in transgenic reporter mice and zebrafish. Notably, some of the functionally validated elements fall within the PRS deletions. These studies suggest that multiple

[#] - **Corresponding authors:** CTG and SL INSERM U1163 Institut Imagine, 24 Boulevard du Montparnasse, 75015, Paris France. chris.gordon@inserm.fr, stanislas.lyonnet@inserm.fr.

The authors have no conflict of interest to declare.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

non-coding elements contribute to the craniofacial regulation of *SOX9* expression, and that their disruption results in PRS.

Keywords

SOX9; craniofacial; enhancer; Pierre Robin; long-range regulation; campomelic dysplasia

Introduction

SOX9 (MIM# 608160) is an HMG-box transcription factor with multiple roles during embryonic development (reviewed in Pritchett et al., 2011). Conditional deletion in mice has demonstrated that *Sox9* is essential for the development of progenitors of chondrocytic, neuronal, pancreatic and Sertoli lineages, amongst others. *Sox9* also plays a role in the production of neural crest cells (NCCs) from the dorsal neural tube in several species (reviewed in Lee and Saint-Jeannet, 2011), and NCC-specific knockout mice fail to develop craniofacial cartilage (Mori-Akiyama et al., 2003). In humans, haploinsufficiency of *SOX9*, due to coding mutations, results in campomelic dysplasia (CD; MIM# 114290), a skeletal dysplasia often associated with disorders of sex development (DSDs). A frequent component of CD is Pierre Robin sequence (PRS; MIM# 261800), which involves micro- and/or retrognathia, glossoptosis and cleft palate. PRS in its isolated form is a relatively frequent craniofacial anomaly, with the initiating event thought to be defective mandibular outgrowth, although the underlying genetic causes are unknown in the vast majority of cases (Tan et al., 2013). Non-coding lesions (translocations, deletions, duplications) upstream of *SOX9* have suggested a complex regulatory domain controlling tissue-specific expression of *SOX9* during embryonic development. These lesions typically result in endophenotypes of complete CD, and approximate genotype-phenotype correlations can be made: within a 1 Mb domain upstream of *SOX9*, translocation breakpoints falling close to *SOX9* typically result in CD, while those further upstream tend to result in acampomelic campomelic dysplasia (ACD; MIM# 114290), a milder version of CD without limb involvement (reviewed in Gordon et al., 2009; Leipoldt et al., 2007). A series of large duplications 0.78-1.99 Mb upstream were associated with brachydactyly anonychia (BA; MIM# 106995) (Kurth et al., 2009), while duplications, a triplication or deletions of a region 517-595 kb upstream lead to DSDs (Benko et al., 2011; Cox et al., 2011; Vetro et al., 2011). Finally, a locus for isolated PRS has been identified at ~1.2-1.5 Mb upstream of *SOX9*, based on the position of a translocation breakpoint cluster (TBC) (Benko et al., 2009; Jakobsen et al., 2007) and deletions falling immediately proximal to this TBC (Benko et al., 2009; Fukami et al., 2012; Sanchez-Castro et al., 2013).

Regulatory elements that drive reporter expression in specific tissues in transgenic mice have been identified in the region upstream of *SOX9* (Bagheri-Fam et al., 2006; Benko et al., 2009; Mead et al., 2013; Sekido and Lovell-Badge, 2008; Wunderle et al., 1998). The overlap of reporter expression with endogenous *Sox9* expression supports the idea that disruption of these elements in patients with non-coding lesions result in CD endophenotypes due to perturbation of *SOX9* expression in a restricted range of tissues. Mechanistically, individual elements, or regions containing many elements, may make

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

contact with the *SOX9* promoter to regulate expression, despite being separated from the gene by large distances in *cis* (Maass et al., 2012; Smyk et al., 2013; Velagaleti et al., 2005; Zhang et al., 2012) or in *trans* (Maass et al., 2012). Candidate regulatory elements at the *SOX9* locus have typically been selected for functional studies on the basis of conservation between distantly related species. However, only a small fraction of the conserved non-coding elements (CNEs) have been tested for regulatory activity. In the PRS locus ~1.2-1.5 Mb upstream of *SOX9*, two enhancers capable of driving reporter gene expression in the embryonic mandible have been reported, consistent with their loss, and consequent disruption of *SOX9* expression in the mandible, being causal in PRS patients with deletions or translocations of this region (Benko et al., 2009). Within this PRS locus there exist at least 14 elements conserved between human and chicken, the regulatory potential of which has not been tested. Also, reporter studies have demonstrated that some craniofacial enhancers exist much closer to *SOX9* (Bagheri-Fam et al., 2006; Mead et al., 2013; Sekido and Lovell-Badge, 2008; Wunderle et al., 1998); restricted disruption of these elements could theoretically lead to craniofacial phenotypes similar to those associated with the lesions far upstream.

Here we screened patients with non-syndromic PRS for microdeletions at the *SOX9* locus in order to highlight non-coding regions that may harbour novel craniofacial regulatory elements. We combined this with a global analysis of chromatin marks at the *Sox9* locus in mouse embryonic craniofacial tissue. Our findings suggest that appropriate expression of *SOX9* during craniofacial development is achieved via the action of multiple discrete regulatory elements.

Materials and Methods

Comparative genomic hybridisation

The PRS183 deletion was identified using a custom-designed oligonucleotide microarray (Agilent Technologies) consisting of 44,000 60-mer oligonucleotide probes (eArray, Agilent Technologies). The design included the following genomic regions (hg19): *FBN2* (MIM# 612570) (chr5:125,172,101-128,172,101), *SOX9* (chr17:67,988,405-71,988,405), *SATB2* (MIM# 608148) (chr2:198,591,755-201,191,755), *TBX1* (MIM# 602054) (chr22:19,720,000-19,820,000) and *TBX22* (MIM# 300307) (chrX:78,613,344-79,963,344) with an average probe spacing of 260 bp. “Dye-swap” experiments were performed for each patient sample to reduce the variation related to labelling and hybridisation efficiencies. Briefly, 1 μg of genomic DNA from the patient and the control (pool of 5 samples) was digested with *Alu*I and *Rsa*I enzymes, and the digested DNA was labelled with Cyanine 3-dUTP or Cyanine 5-dUTP in dye-swap reactions, followed by hybridization, as per the manufacturer's instructions (Agilent Technologies). Slides were scanned using an Axon GenePix 4000B scanner (Molecular Devices). Images were extracted and normalised using the linear-and-Lowess normalisation module implemented in the Feature Extraction software (Agilent Technologies). An R-script was used to correct for systematic differences in probe efficiencies seen on a particular array using the “self-self” microarray data (T. Fitzgerald, personal communication). The GC bias (or wave profile) was also corrected using a 500 bp window around each probe (Marioni et al., 2007). Copy number analysis was

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

carried out with DNA Analytics software (Agilent Technologies) using the aberration detection module (ADM)-2 algorithm (Lipson et al., 2006) with a threshold of 6 and at least 5 consecutive probes showing a change in the copy number.

19 non-syndromic PRS patients were screened for copy number variants at the *SOX9* locus using a NimbleGen fine-tiling custom array covering chr17:37,612,234-73,596,758 (these and all co-ordinates herein refer to assembly hg19), with a mean probe spacing of 37 bp (used for the identification of the PRS116 deletion). (Note that different custom-designed CGH arrays were used for screening case PRS183 versus the subsequent 19 non-syndromic PRS patients because each microarray was designed independently by the FitzPatrick and Lyonnet labs, respectively). Approval for human genetic research was obtained from the Comité de Protection des Personnes Ile-de-France II.

Genomic qPCR

Quantitative real-time PCR was used to assess the presence or absence of the upstream *SOX9* deletion in family PRS183. Reactions were carried out in triplicate and consisted of 10 ng of genomic DNA, 1 μM of each of the forward (TCTCCCAGGAGATCTTTCAATG) and reverse (GCAGCTGCGAGCCATTAT) primers, 5 μl of 2X TaqMan™ Universal Master Mix (Applied Biosystems), 0.5 μl of 20X VIC-labelled RNaseP internal control (Applied Biosystems), 0.1 μM of FAM-labelled Universal Probe Library (UPL) probe 59 (Roche Diagnostics) and 0.5 μl of PCR-grade water. Amplification and real-time data collection was performed on an HT7900 instrument (Applied Biosystems) using the following conditions: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. After completion of PCR, fluorescence was read using the software SDS (Applied Biosystems) and the resulting Ct values were exported and DNA copy number determined.

Sequencing

Purification of genomic DNA from lymphocytes and Sanger sequencing were performed following standard protocols. Sequences of primers used for amplification and sequencing of the *FOXC2* (MIM# 602402) coding region and the family F1 and PRS116 deletion breakpoints are available upon request. The *FOXC2* variant identified in family F2 has been submitted to a relevant locus-specific database (http://www.lovd.nl/FOXC2).

ChIP-Seq, mouse transgenesis, optical projection tomography (OPT)

ChIP-Seq analysis of genomic regions immunoprecipitated by an anti-p300 antibody from E11.5 mouse craniofacial tissue (including nasal, maxillary, and mandibular prominences), generation of transient transgenic mice and OPT imaging were performed as previously described (Attanasio et al., 2013). Primers used for amplification of elements tested in transgenic reporter mice are available on request (element co-ordinates are listed in Table 1). Transgenic reporter data described herein will be available for viewing via the Vista Enhancer Browser at enhancer.lbl.gov (Visel et al., 2007) and FaceBase at www.facebase.org (Hochheiser et al., 2011). The E11.5 craniofacial H3K27ac ChIP-Seq dataset was generated using similar techniques as for that of the p300 data, and will be presented in greater detail elsewhere (C.A., L.A.P and A.V., unpublished data). Details

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

regarding the E13.5 palatal tissue p300 ChIP-Seq data can be obtained at the FaceBase website; see https://www.facebase.org/content/chip-seq-analysis-p300-bound-regions-e135-mouse-palates.

Zebrafish transgenesis

Stable lines of transgenic zebrafish were generated and analysed following methods described elsewhere (Bhatia et al., in preparation and Ravi et al., 2013). Briefly, the primers attB4-CACTTGATGAATTTCGGCA, attB1r-CAATTACATTTCATACTGGAG and attB1r-CAGTTACATTTCATACTGGAG were used for amplification of wildtype and variant versions of HCNE-F2 from control and patient genomic DNA respectively. Fragments were cloned into a P4P1r entry vector before being recombined with a pDONR221 construct containing either a gata2 promoter-eGFP-polyA or a gata2 promoter mCherry-polyA cassette into a destination vector with a Gateway R4-R2 cassette flanked by Tol2 recombination sites. The peak 17 fragment was cloned into the gata2 promoter-eGFP-pA reporter construct according to the same strategy using primers attB4-GCCCTCCTGAGGAAGAGTG and attB1r-CATAATGACAGTGACCAGTG.

Immunofluorescence

Cryosections of paraformaldehyde-fixed wildtype mouse embryos were processed by standard techniques for Sox9 immunofluorescence. Primary antibody: rabbit polyclonal anti-Sox9 (Chemicon, AB5535). Secondary antibody: Alexa Fluor 555 donkey anti-rabbit (Invitrogen, A31572).

Results

As a means to identify novel *SOX9* craniofacial regulatory regions, we screened 20 patients with non-syndromic PRS for microdeletions at the *SOX9* locus using custom-designed fine-tiling CGH arrays. In one familial case (PRS183), the proband presented with cleft palate, micrognathia, crowded teeth, flat facial features and marked pectus deformity (Fig 1A). X-rays showed hypoplastic clavicles. His brother also had cleft palate, micrognathia, flat facial appearance and mild pectus, while his father had a cleft soft palate and a flat midface. The oral abnormalities in both boys impacted on early feeding, causing failure to thrive. In the PRS183 proband we identified a deletion of 280 kb at 1,156 kb upstream of the *SOX9* transcription start site (maximum deletion co-ordinates, Hg19: chr17:68,680,882-68,965,323; minimum deletion co-ordinates: chr17:68,681,261-68,960,704), and genomic qPCR demonstrated segregation of the deletion with affected family members (Fig 1B, C). (We note that although a high-density custom array was used to identify this deletion, the deletion is large enough that it would likely have been detected using a number of commercially available microarray platforms). The telomeric boundary of the deletion falls within the PRS TBC, while the centromeric boundary falls within the F1 deletion of Benko et al. (2009) (Fig 2 and Supp. Fig S1A).

Screening of 19 additional non-syndromic PRS patients by high-density custom CGH led to the identification of two CNVs not listed in the Database of Genomic Variants, in the region surrounding *SOX9*. Firstly, in the female proband of a familial case of isolated PRS (her

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sister also presented with PRS), a 6.78 kb duplication was identified within an intron of *SLC39A11*, the next coding gene telomeric to *SOX9*. The duplicated region is 563 kb downstream of the *SOX9* transcription start site. However, a duplication junction-spanning PCR product used to genotype other family members indicated that the duplication was inherited from the proband's unaffected mother, and was not present in her affected sister (data not shown), suggesting that the duplication was not a major pathogenic factor, and the case was not considered further. The other novel CNV consisted of a 6 kb deletion 407 kb upstream of *SOX9* in the proband of family PRS116 (Fig 3A). This patient presented with isolated PRS. PRS was also present in her mother and maternal aunt (the latter died of asphyxiation three days after birth), while her maternal grandfather had feeding difficulties in the neonatal period, without PRS (Fig 3B). The proband's mother and maternal grandfather also had adult-onset hearing loss and Dupuytren's contracture (MIM# 126900). No radiographic anomalies were detected in the proband. The mother had supernumerary digital folds without radiographic anomalies of the hands. The deletion was verified in the proband by amplification and sequencing of a breakpoint-spanning PCR product (Fig 3A; chr17:69,704,137-69,710,167 deleted, inclusive). A three basepair region of microhomology was identified at the 5' and 3' breakpoints (sequence TTT); microhomology of this size is a frequent signature of CNV breakpoints (Conrad et al., 2010; Hastings et al., 2009). Using the breakpoint-spanning PCR product to genotype the family, the deletion was present in the proband and her affected mother and grandfather, but not in her healthy father and brothers (Fig 3B). The deletion falls much closer to *SOX9* than the previously described PRS cases and does not contain any previously reported regulatory elements, but harbours a single CNE conserved from humans to non-mammalian vertebrates (Fig 2 and Supp. Fig S1B).

The size of the deletion in the PRS family F1 was previously reported as ~75 kb, at ~1.38 Mb upstream of *SOX9* (Benko et al., 2009). As depicted in Supp. Fig S1A, the F1 and PRS183 deletions are partly overlapping, suggesting that important craniofacial regulatory elements may fall within this region of overlap. In order to more precisely define the extent of overlap, and the CNEs that this region contains, we amplified and sequenced a deletion-spanning PCR product from genomic DNA of one affected individual from the F1 family. This analysis indicated a deletion of 84 kb (chr17:68,663,923-68,748,011 inclusive; Hg19), accompanied by an insertion of 9 bp (Supp. Fig S2). The region of overlap between the F1 and PRS183 deletions (67 kb) contains four CNEs displaying conservation of at least 70% over 300 bp between human and each of opossum and chicken (selected at the ECR Browser; http://ecrbrowser.dcode.org/) (Supp. Fig S1A and Table 1). Although two previously reported craniofacial enhancers, HCNE-F2 and 9CE4Z, lie centromeric to the PRS TBC (Benko et al., 2009), neither falls within the F1-PRS183 overlap (Supp. Fig S1A). (Despite several attempts, we were unable to amplify a PRS183 deletion-spanning PCR product; therefore the precise boundaries of the PRS183 deletion were not mapped. However, more than 10 non-deleted CGH probes fell between HCNE-F2 and the first deleted probe of the PRS183 deletion interval (a region of approximately 5 kb), indicating that HCNE-F2 falls outside the PRS183 deletion and hence outside the F1-PRS183 overlap.) This suggests that other craniofacial regulatory elements may exist within the F1-PRS183 deletion overlap.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The HCNE-F2 element was previously investigated because it harbours a single base variant in several affected individuals of a PRS family (family F2 in Benko et al., 2009). HCNE-F2 is not conserved outside placental mammals, and was the only CNE within the F1 deletion that was tested for regulatory activity (Benko et al., 2009). Subsequent to that report, we identified features suggestive of lymphoedema-distichiasis syndrome (LDS; MIM# 153400) in several members of the F2 family (Fig 4). Sequencing of the coding region of the major candidate gene for LDS, *FOXC2*, revealed a 10 bp deletion, leading to a frameshift and a premature stop (RefSeq transcript NM_005251.2:c.975_984delCGCCGGGGGC, p.Ala326ThrfsX8), in the proband, his brother and his father, each of whom have PRS, and in the paternal grandfather who does not have PRS (Fig 4). The proband and his brother have not yet developed lymphoedema. PRS or cleft palate has been reported in a proportion of LDS patients harbouring *FOXC2* mutations (Bahuau et al., 2002; Brice et al., 2002; Erickson et al., 2001; Fang et al., 2000; Finegold et al., 2001; Tanpaiboon et al., 2010), and *Foxc2* homozygous null mice display mandibular dysplasia and cleft palate (Iida et al., 1997; Winnier et al., 1997). The single base change in HCNE-F2 was inherited from the paternal grandmother, who did not present with PRS (Benko et al., 2009); therefore the three members of this family with PRS are the only three with both the *FOXC2* mutation and the HCNE-F2 variant, suggesting that both these events may have contributed to the PRS phenotype. Although the variant was initially not detected in a group of control patients (Benko et al., 2009), the same base change has recently been reported as a SNP (rs78542003, T>C), with a C allele frequency of 1.433% (33 counts/2302 alleles) (from UCSC data last updated 2012-11-09). Previous *in vitro* assays suggested that the variant sequence modified the binding of MSX1 (MIM# 142983), a transcriptional regulator with essential roles during craniofacial development. To support the argument that the HCNE-F2 variant can alter enhancer activity *in vivo*, we performed transgenic analysis in zebrafish embryos, testing the ability of the wildtype and variant versions of the enhancer to drive mCherry or GFP reporter expression, respectively, in four independent stable lines (each line harbouring both transgenes).

We found that wildtype HCNE-F2 could drive expression in the developing pharyngeal arches and craniofacial cartilage (but not Meckel's cartilage), while this activity was decreased or lost for the version harbouring the variant (representative images are shown in Supp. Fig S3). Collectively these results suggest that the HCNE-F2 variant is likely to play a less important role in the pathogenesis of PRS than previously suspected, however it remains possible that it may contribute to the phenotype by acting as a modifying allele in the presence of the *FOXC2* coding mutation.

We also conducted a genome-wide analysis of craniofacial cis-regulatory activity, by performing ChIP-Seq for p300- or acetylated H3K27 (H3K27ac)-associated genomic elements in E11.5 mouse craniofacial tissue and p300-associated elements in E13.5 mouse palatal tissue. H3K27ac and p300 chromatin marks have been used previously as predictors of developmental enhancers (Rada-Iglesias et al., 2011; Visel et al., 2009). We focused on p300 and H3K27ac binding events (peaks) within a ~3 Mb interval encompassing *Sox9*. Following conversion from mouse to human co-ordinates (where possible), we compared the distribution of peaks with genomic lesions associated with PRS in previous studies and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the present report, with the hypothesis that the peaks represent elements involved in the normal expression of *SOX9* during early craniofacial development, and that their loss could contribute to the PRS phenotype. Nine E11.5 p300 peaks were identified in the 1.9 Mb interval between *KCNJ2* (MIM# 600681) and *SOX9*, and one fell within a 1 Mb region downstream of *SOX9* (Fig 2). Apart from peak23, which falls on a portion of the *SOX9* proximal promoter, all eight other peaks in the *KCNJ2-SOX9* interval overlapped a CNE with conservation between human and chicken (according to the Multiz alignment track in the UCSC browser), supporting the functional relevance of these peaks.

Importantly, five of these eight peaks (peaks 15-19) were clustered in a domain of 115 kb, proximal to the PRS TBC (Fig 2 and Supp. Fig S1A). Peak22 falls within the 6 kb deletion identified in PRS116 (Supp. Fig S1B); this is particularly striking considering that it is the only p300 peak within a 0.9 Mb interval upstream of the *SOX9* proximal promoter (Fig 2). Two peaks from the E11.5 craniofacial H3K27ac and three from the E13.5 palatal p300 ChIP-Seq studies fell in the *KCNJ2-SOX9* interval (Fig 2, Supp. Fig S1A and S1B); several of these were also either within or adjacent to regions deleted in PRS patients.

For each of the 10 E11.5 p300 peaks, *in vivo* reporter activity was tested in transient transgenic mouse assays at E11.5, using constructs harbouring a given peak region upstream of *lacZ* (genomic co-ordinates of the regions used and a summary of *in vivo* craniofacial reporter activity are listed in Table 1). Two peaks demonstrated *lacZ* expression in mandibular mesenchyme (peaks 16 and 22; Fig 5A and Supp. Fig S4). Peak 16 drove expression in the proximal first arch in 4/8 embryos (Fig 5A, upper row), and peak 22 drove relatively weak but reproducible expression in a specific region of the mandible (5/10 embryos; Fig 5A, lower row). Examination of OPT slices from peak 16- and peak 22-*lacZ* embryos indicated an overlap of *lacZ* expression with a portion of the endogenous Sox9 protein expression domain in the mandible (Supp. Fig S4 and see Supp. Movies S1 and S2). Peak 23 drove reporter expression in the lower lip in 5/13 embryos and peak 24 drove expression in the lateral nasal prominence in 3/9 embryos. Peak 22-*lacZ* also displayed strong relative staining in distal limb bud mesenchyme (9/10 embryos; Fig 5A, lower row and Supp. Fig S4E). Peak 24 drove expression in an anterior domain of the limb bud in 6/9 embryos and peak 15-*lacZ* showed staining in a specific region of the limb bud in 6/11 embryos (data not shown).

The only E11.5 p300 peak falling within the minimal region of overlap between the F1 and PRS183 deletions is peak 17. This peak covers one of the four human-opossum-chicken CNEs in this minimal region of overlap. However, the peak 17 region displayed no reproducible lacZ reporter activity in transient transgenic mice at E11.5. Given the potential importance of this element, we tested its activity in a different model. We generated four independent lines of zebrafish stably expressing a peak 17-GFP reporter gene. We observed consistent expression of GFP in the region surrounding the oral cavity at 48 hpf and in first branchial arch-derived cartilages, including Meckel's, at 96 hpf (Fig 5B), supporting the idea that peak 17 functions as a craniofacial regulatory element.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Discussion

In this report we have identified two novel deletions upstream of *SOX9* in PRS patients, one of which (in family PRS183) partly overlaps a previously described deletion associated with isolated PRS, and the other (in family PRS116) falling in a region much closer to *SOX9*, not previously associated with PRS. These deletions contain or fall in close proximity to novel candidate craniofacial enhancers of *SOX9*, identified here by ChIP-Seq and validated in transgenic reporter assays. For each deletion, PRS was associated with other phenotypes. In the PRS183 family, other features included pectus deformity and flat midface (we note however that several features in this family displayed variable expression, with the two deletion-positive siblings more severely affected than their father). Although thorax abnormalities are common in CD, pectus deformities are not frequently reported. However, pectus was described in several members of a family affected with ACD (Stalker and Zori, 1997) and harbouring a translocation breakpoint 932 kb upstream of *SOX9* (Hill-Harfe et al., 2005). The 280 kb deletion in family PRS183 falls 224 kb further proximal to this breakpoint, suggesting that both cases involve the disruption of regulatory elements that direct appropriate expression of *SOX9* during thoracic skeletal development. In addition to family PRS183, a flattened face was mentioned in two PRS cases harbouring balanced translocations within the PRS TBC (Benko et al., 2009; Jakobsen et al., 2007; Vintiner et al., 1991). A flat face being a typical feature of CD, non-mandibular craniofacial enhancers of *SOX9* may also exist proximal to the PRS TBC. A 3 Mb deletion upstream of *SOX9* was recently reported in a patient with PRS, scapula hypoplasia and clubfeet (Fukami et al., 2012); in this case the telomeric breakpoint boundary also fell within the PRS TBC (see Fig 2). It has been reported that a ~2.4 Mb region centred on *Sox9* undergoes chromatin decompaction in the embryonic mouse mandible (Benko et al., 2009) and that *SOX9* co-localises with a region 1.1 Mb upstream in human interphase nuclei (Velagaleti et al., 2005). In the Odd Sex mouse mutant, a transgene insertion influences expression from almost 1 Mb upstream of *Sox9* (Qin et al., 2004). Recently it has been demonstrated by chromatin conformation capture that an element 1 Mb upstream loops to the *SOX9* promoter in a prostate cancer cell line (Zhang et al., 2012), while chromosome conformation capture-on-chip analysis identified interactions between the *SOX9* promoter and regions ~1.13 Mb and ~967 kb upstream in Sertoli cells and lymphoblasts (Smyk et al., 2013). The above data support the idea that *SOX9* can be regulated by a region further than 1 Mb upstream, but that this region is not solely involved in mandibular *SOX9* expression. In family PRS116, two individuals harbouring the deletion had Dupuytren's contracture, a fibrotic connective tissue disorder that to our knowledge has not previously been described in the context of *SOX9*-related disease. The relevance of this phenotype to the deletion remains unclear, however it is worth noting that *SOX9* has been implicated in various fibrotic pathologies, where it may regulate extracellular matrix deposition (Pritchett et al., 2011). The same two PRS116 individuals also had late-onset deafness, which is intriguing given the role of Sox9 in mice in both ossicular and otic vesicle development (Barrionuevo et al., 2008; Mori-Akiyama et al., 2003). Further analysis of peak 22-*lacZ* transgenic mice at multiple stages of development will be required to determine whether this element can drive expression in the ear.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Our ChIP-Seq analysis of regions binding p300 in the *KCNJ2-SOX9* interval led to the identification of several novel candidate enhancers of *SOX9* expression. Transgenic analysis in mice and zebrafish demonstrated that peaks 16, 17 and 22 can drive reporter expression within the developing mandible, consistent with their loss playing a role in the pathogenesis of PRS in patients harbouring deletions of these elements. The PRS183 deletion reported here does not include the previously identified craniofacial enhancer HCNE-F2 (Benko et al., 2009), while the family F1 deletion (Benko et al., 2009) did not include the 9CE4Z enhancer. Therefore, the region of overlap between the PRS183 and F1 deletions (67 kb) may harbour other craniofacial *SOX9* enhancers – indeed, the region contains four CNEs conserved from human to chicken. Notably, peak 17, which displayed craniofacial regulatory activity in zebrafish, falls on one of these four CNEs. Although peak 22 drove reproducible but relatively weak expression in the E11.5 mandible, it was strongly active in the distal limb bud at this stage – no previously reported *SOX9* enhancer shows this degree of specificity within the developing limb, and it will be of interest to explore the expression pattern of peak 22-*lacZ* (along with the other limb bud-active peaks identified here; peaks 15 and 24) at later stages of chondrogenic development. No single enhancer thus far identified at the *SOX9* locus drives expression specifically in all chondrogenic tissue. It is possible that *SOX9* expression in developing cartilage is achieved by the action of many regulatory elements, each providing maximal activity in small subdomains. Interestingly, of the two H3K27ac craniofacial peaks within the 1.9 Mb *KCNJ2-SOX9* interval, one falls 8 kb away from p300 peak 22, which displayed *in vivo* reporter activity and falls within the PRS116 deletion. (Note the only other H3K27ac peak falls on the *SOX9* proximal promoter). Further upstream, p300 craniofacial peaks 15-19, p300 palate peak 1, HCNE-F2 and 9CE4Z all fall within a 210 kb interval proximal to the PRS TBC; note that none of these eight elements are overlapping, suggesting that many individual inputs may contribute to the craniofacial regulatory activity of the region. In addition, previous studies showed craniofacial regulatory activity closer to *SOX9* (Bagheri-Fam et al., 2006; Mead et al., 2013; Sekido and Lovell-Badge, 2008; Wunderle et al., 1998), and the peak 22 craniofacial enhancer identified here in the PRS116 deletion falls at 408 kb upstream. A major question is how these dispersed activities are coordinated with respect to each other and to enhancers for other cell types.

In conclusion, we have identified novel PRS-associated deletions upstream of *SOX9*, and we have utilised p300 ChIP-Seq and *in vivo* reporter assays to define novel craniofacial regulatory elements falling within these or previously reported deletions. Our data are consistent with the hypothesis that these elements are required for appropriate expression of *SOX9* during craniofacial development, and that their loss contributes to the PRS phenotype.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

C.T.G., J.A., and S.L. were supported by funding from the ANR grants EvoDevoMut, CRANIRARE and IHU-2010-001. T.Y.T. was supported by a NHMRC Training Fellowship (#607431). A.V. and L.A.P. were supported by NIH grants R01HG003988 and U01DE020060. C.A. was supported by a SNSF advanced researchers

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

fellowship. Research by A.V., L.A.P. and C.A. was conducted at the E.O. Lawrence Berkeley National Laboratory and performed under Department of Energy Contract DE-AC02-05CH11231, University of California.

Grant numbers: ANR: EvoDevoMut, CRANIRARE and IHU-2010-001. NHMRC Training Fellowship #607431. NIH grants R01HG003988 and U01DE020060. SNSF advanced researchers fellowship. Department of Energy Contract DE-AC02-05CH11231.

References

Attanasio C, Nord AS, Zhu Y, Blow MJ, Li Z, Liberton DK, Morrison H, Plajzer-Frick I, Holt A, Hosseini R, Phouanenavong S, Akiyama JA, et al. Fine tuning of craniofacial morphology by distant-acting enhancers. Science. 2013; 342(6157):1241006. [PubMed: 24159046]

Bagheri-Fam S, Barrionuevo F, Dohrmann U, Gunther T, Schule R, Kemler R, Mallo M, Kanzler B, Scherer G. Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. Dev Biol. 2006; 291(2):382–97. [PubMed: 16458883]

Bahuau M, Houdayer C, Tredano M, Soupre V, Couderc R, Vazquez MP. FOXC2 truncating mutation in distichiasis, lymphedema, and cleft palate. Clin Genet. 2002; 62(6):470–3. [PubMed: 12485195]

Barrionuevo F, Naumann A, Bagheri-Fam S, Speth V, Taketo MM, Scherer G, Neubuser A. Sox9 is required for invagination of the otic placode in mice. Dev Biol. 2008; 317(1):213–24. [PubMed: 18377888]

Benko S, Fantes JA, Amiel J, Kleinjan DJ, Thomas S, Ramsay J, Jamshidi N, Essafi A, Heaney S, Gordon CT, McBride D, Golzio C, et al. Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat Genet. 2009; 41(3):359–64. [PubMed: 19234473]

Benko S, Gordon CT, Mallet D, Sreenivasan R, Thauvin-Robinet C, Brendehaug A, Thomas S, Bruland O, David M, Nicolino M, Labalme A, Sanlaville D, et al. Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. J Med Genet. 2011; 48(12):825–30. [PubMed: 22051515]

Brice G, Mansour S, Bell R, Collin JR, Child AH, Brady AF, Sarfarazi M, Burnand KG, Jeffery S, Mortimer P, Murday VA. Analysis of the phenotypic abnormalities in lymphoedema-distichiasis syndrome in 74 patients with FOXC2 mutations or linkage to 16q24. J Med Genet. 2002; 39(7): 478–83. [PubMed: 12114478]

Conrad DF, Bird C, Blackburne B, Lindsay S, Mamanova L, Lee C, Turner DJ, Hurles ME. Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. Nat Genet. 2010; 42(5): 385–91. [PubMed: 20364136]

Cox JJ, Willatt L, Homfray T, Woods CG. A SOX9 duplication and familial 46,XX developmental testicular disorder. N Engl J Med. 2011; 364(1):91–3. [PubMed: 21208124]

Erickson RP, Dagenais SL, Caulder MS, Downs CA, Herman G, Jones MC, Kerstjens-Frederikse WS, Lidral AC, McDonald M, Nelson CC, Witte M, Glover TW. Clinical heterogeneity in lymphoedema-distichiasis with FOXC2 truncating mutations. J Med Genet. 2001; 38(11):761–6. [PubMed: 11694548]

Fang J, Dagenais SL, Erickson RP, Arlt MF, Glynn MW, Gorski JL, Seaver LH, Glover TW. Mutations in FOXC2 (MFH-1), a forkhead family transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome. Am J Hum Genet. 2000; 67(6):1382–8. [PubMed: 11078474]

Finegold DN, Kimak MA, Lawrence EC, Levinson KL, Cherniske EM, Pober BR, Dunlap JW, Ferrell RE. Truncating mutations in FOXC2 cause multiple lymphedema syndromes. Hum Mol Genet. 2001; 10(11):1185–9. [PubMed: 11371511]

Fukami M, Tsuchiya T, Takada S, Kanbara A, Asahara H, Igarashi A, Kamiyama Y, Nishimura G, Ogata T. Complex genomic rearrangement in the SOX9 5' region in a patient with Pierre Robin sequence and hypoplastic left scapula. Am J Med Genet A. 2012; 158A(7):1529–34. [PubMed: 22529047]

Gordon CT, Tan TY, Benko S, Fitzpatrick D, Lyonnet S, Farlie PG. Long-range regulation at the SOX9 locus in development and disease. J Med Genet. 2009; 46(10):649–56. [PubMed: 19473998]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Hastings PJ, Ira G, Lupski JR. A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS Genet. 2009; 5(1):e1000327. [PubMed: 19180184]

Hill-Harfe KL, Kaplan L, Stalker HJ, Zori RT, Pop R, Scherer G, Wallace MR. Fine mapping of chromosome 17 translocation breakpoints > or = 900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. Am J Hum Genet. 2005; 76(4):663–71. [PubMed: 15717285]

Hochheiser H, Aronow BJ, Artinger K, Beaty TH, Brinkley JF, Chai Y, Clouthier D, Cunningham ML, Dixon M, Donahue LR, Fraser SE, Hallgrimsson B, et al. The FaceBase Consortium: a comprehensive program to facilitate craniofacial research. Dev Biol. 2011; 355(2):175–82. [PubMed: 21458441]

Iida K, Koseki H, Kakinuma H, Kato N, Mizutani-Koseki Y, Ohuchi H, Yoshioka H, Noji S, Kawamura K, Kataoka Y, Ueno F, Taniguchi M, et al. Essential roles of the winged helix transcription factor MFH-1 in aortic arch patterning and skeletogenesis. Development. 1997; 124(22):4627–38. [PubMed: 9409679]

Jakobsen LP, Ullmann R, Christensen SB, Jensen KE, Molsted K, Henriksen KF, Hansen C, Knudsen MA, Larsen LA, Tommerup N, Tumer Z. Pierre Robin sequence may be caused by dysregulation of SOX9 and KCNJ2. J Med Genet. 2007; 44(6):381–6. [PubMed: 17551083]

Kurth I, Klopocki E, Stricker S, van Oosterwijk J, Vanek S, Altmann J, Santos HG, van Harssel JJ, de Ravel T, Wilkie AO, Gal A, Mundlos S. Duplications of noncoding elements 5′ of SOX9 are associated with brachydactyly-anonychia. Nat Genet. 2009; 41(8):862–3. [PubMed: 19639023]

Lee YH, Saint-Jeannet JP. Sox9 function in craniofacial development and disease. Genesis. 2011; 49(4):200–8. [PubMed: 21309066]

Leipoldt M, Erdel M, Bien-Willner G, Smyk M, Theurl M, Yatsenko S, Lupski J, Lane A, Shanske A, Stankiewicz P, Scherer G. Two novel translocation breakpoints upstream of SOX9 define borders of the proximal and distal breakpoint cluster region in campomelic dysplasia. Clin Genet. 2007; 71(1):67–75. [PubMed: 17204049]

Lipson D, Aumann Y, Ben-Dor A, Linial N, Yakhini Z. Efficient calculation of interval scores for DNA copy number data analysis. J Comput Biol. 2006; 13(2):215–28. [PubMed: 16597236]

Maass PG, Rump A, Schulz H, Stricker S, Schulze L, Platzer K, Aydin A, Tinschert S, Goldring MB, Luft FC, Bahring S. A misplaced lncRNA causes brachydactyly in humans. J Clin Invest. 2012; 122(11):3990–4002. [PubMed: 23093776]

Marioni JC, Thorne NP, Valsesia A, Fitzgerald T, Redon R, Fiegler H, Andrews TD, Stranger BE, Lynch AG, Dermitzakis ET, Carter NP, Tavare S, et al. Breaking the waves: improved detection of copy number variation from microarray-based comparative genomic hybridization. Genome Biol. 2007; 8(10):R228. [PubMed: 17961237]

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(8):4459–69. [PubMed: 23449223]

Mori-Akiyama Y, Akiyama H, Rowitch DH, de Crombrugghe B. Sox9 is required for determination of the chondrogenic cell lineage in the cranial neural crest. Proc Natl Acad Sci U S A. 2003; 100(16): 9360–5. [PubMed: 12878728]

Pritchett J, Athwal V, Roberts N, Hanley NA, Hanley KP. Understanding the role of SOX9 in acquired diseases: lessons from development. Trends Mol Med. 2011; 17(3):166–74. [PubMed: 21237710]

Qin Y, Kong LK, Poirier C, Truong C, Overbeek PA, Bishop CE. Long-range activation of Sox9 in Odd Sex (Ods) mice. Hum Mol Genet. 2004; 13(12):1213–8. [PubMed: 15115764]

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(7333):279–83. [PubMed: 21160473]

Ravi V, Bhatia S, Gautier P, Loosli F, Tay BH, Tay A, Murdoch E, Coutinho P, van Heyningen V, Brenner S, Venkatesh B, Kleinjan DA. Sequencing of Pax6 loci from the elephant shark reveals a family of Pax6 genes in vertebrate genomes, forged by ancient duplications and divergences. PLoS Genet. 2013; 9(1):e1003177. [PubMed: 23359656]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sanchez-Castro M, Gordon CT, Petit F, Nord AS, Callier P, Andrieux J, Guerin P, Pichon O, David A, Abadie V, Bonnet D, Visel A, et al. Congenital heart defects in patients with deletions upstream of SOX9. Hum Mutat. 2013; 34(12):1628–31. [PubMed: 24115316]

Sekido R, Lovell-Badge R. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature. 2008; 453(7197):930–4. [PubMed: 18454134]

Smyk M, Szafranski P, Startek M, Gambin A, Stankiewicz P. Chromosome conformation capture-on-chip analysis of long-range cis-interactions of the SOX9 promoter. Chromosome Res. 2013; 21(8): 781–8. [PubMed: 24254229]

Stalker HJ, Zori RT. Variable expression of rib, pectus, and scapular anomalies with Robin-type cleft palate in a 5-generation family: a new syndrome? Am J Med Genet. 1997; 73(3):247–50. [PubMed: 9415678]

Tan TY, Kilpatrick N, Farlie PG. Developmental and genetic perspectives on Pierre Robin sequence. Am J Med Genet C Semin Med Genet. 2013; 163(4):295–305. [PubMed: 24127256]

Tanpaiboon P, Kantaputra P, Wejathikul K, Piyamongkol W. 595-596 insC of FOXC2 underlies lymphedema, distichiasis, ptosis, ankyloglossia, and Robin sequence in a Thai patient. Am J Med Genet A. 2010. c; 152A(3):737–40. [PubMed: 20186799]

Velagaleti GV, Bien-Willner GA, Northup JK, Lockhart LH, Hawkins JC, Jalal SM, Withers M, Lupski JR, Stankiewicz P. Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. Am J Hum Genet. 2005; 76(4):652–62. [PubMed: 15726498]

Vetro A, Ciccone R, Giorda R, Patricelli MG, Della Mina E, Forlino A, Zuffardi O. XX males SRY negative: a confirmed cause of infertility. J Med Genet. 2011; 48(10):710–2. [PubMed: 21653197]

Vintiner GM, Temple IK, Middleton-Price HR, Baraitser M, Malcolm S. Genetic and clinical heterogeneity of Stickler syndrome. Am J Med Genet. 1991; 41(1):44–8. [PubMed: 1683158]

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009; 457(7231):854–8. [PubMed: 19212405]

Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res. 2007; 35(Database issue):D88–92. [PubMed: 17130149]

Winnier GE, Hargett L, Hogan BL. The winged helix transcription factor MFH1 is required for proliferation and patterning of paraxial mesoderm in the mouse embryo. Genes Dev. 1997; 11(7): 926–40. [PubMed: 9106663]

Wunderle VM, Critcher R, Hastie N, Goodfellow PN, Schedl A. Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. Proc Natl Acad Sci U S A. 1998; 95(18):10649–54. [PubMed: 9724758]

Zhang X, Cowper-Sal lari R, Bailey SD, Moore JH, Lupien M. Integrative functional genomics identifies an enhancer looping to the SOX9 gene disrupted by the 17q24.3 prostate cancer risk locus. Genome Res. 2012; 22(8):1437–46. [PubMed: 22665440]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

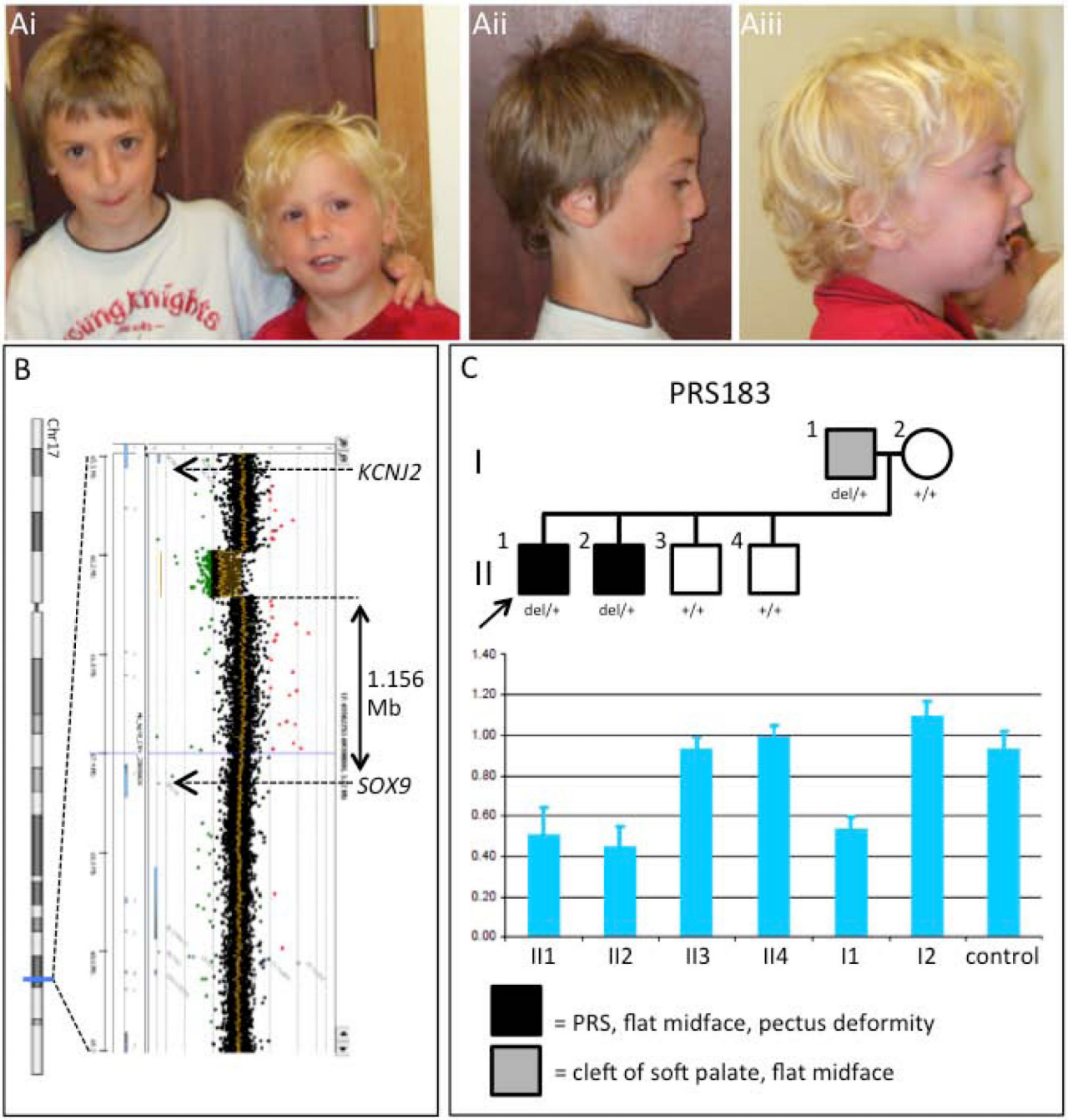

Fig 1.

A, the family PRS183 proband (panels i, left and ii) and his affected brother (panels i, right and iii). B, a 280 kb deletion 1,156 kb upstream of *SOX9* was identified by custom high-density 44K CGH array in the proband of family PRS183. C, the deletion was shared by his brother who had a similar phenotype (PRS and pectus deformity), and by his father with mild features of PRS, as determined by qPCR using genomic DNA from family members. The vertical axis of the graph represents genomic copy number relative to control. Bars of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the graph correspond to respective family members in the tree directly above. Error bars are standard error of the mean copy number.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

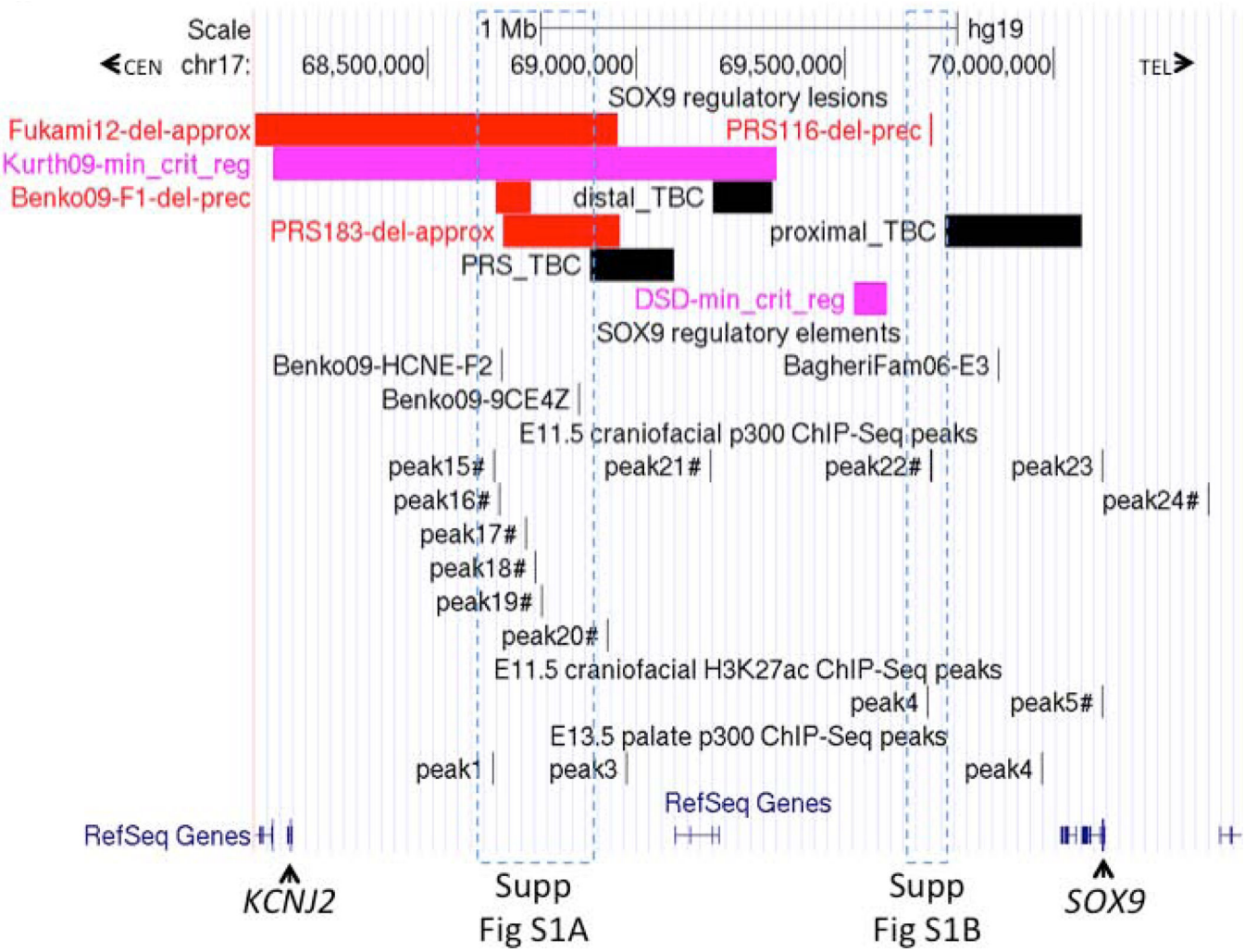

Fig 2.

UCSC screenshot depicting the distribution of genomic lesions between *SOX9* and *KCNJ2*, relative to positions of ChIP-Seq peaks converted from mouse coordinates. A 2.3 Mb interval on chr17q24.3 is shown (co-ordinates are from Hg19). Centromeric (CEN) and telomeric (TEL) ends are indicated. The *SOX9* regulatory lesions track depicts deletions (del) reported in Fukami et al. (2012), Benko et al. (2009) and the present study; minimal critical regions (min crit reg) for brachydactyly anonychia duplications from Kurth et al. (2009) and for DSD duplications and deletions (see text); and PRS, distal and proximal translocation breakpoint clusters (TBCs) (see Gordon et al. (2009)). Whether deletion breakpoints are precisely (prec) or approximately (approx) mapped is indicated. HCNE-F2 and 9CE4Z from Benko et al. (2009) and E3 from Bagheri-Fam et al. (2006) are three previously described craniofacial enhancers. ChIP-Seq peak tracks: regions from the mouse genome corresponding to p300 or H3K27ac binding events were converted, where possible, to human coordinates. A peak name followed by # indicates that the peak region overlaps a CNE conserved between human and chicken in the UCSC Multiz alignment. The dashed boxes are magnified in Supp. Figures S1A and S1B.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

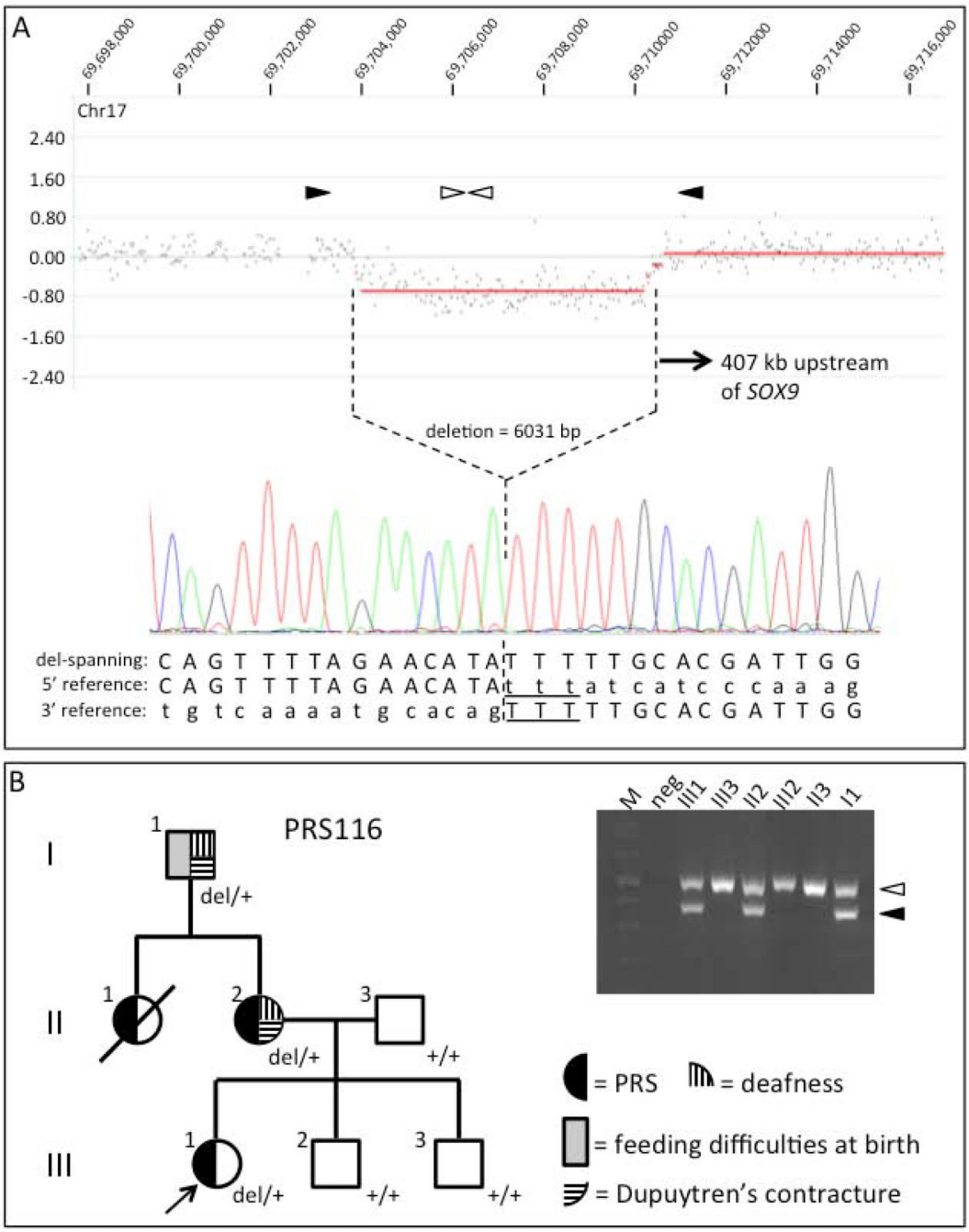

Fig 3.

A, a 6 kb deletion 407 kb upstream of *SOX9* was identified by custom high-density CGH array in the proband of family PRS116. Black arrowheads indicate the approximate position of primers used to amplify a breakpoint-spanning PCR product. White arrowheads indicate the position of primers used as a control for multiplex PCR in B. The electropherogram displays the breakpoint-spanning sequence. Uppercase sequence, outside the deletion; lowercase sequence, within the deletion. A three basepair region of microhomology at the breakpoint is underlined. B, segregation of the 6 kb deletion in family PRS116. Genotyping

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

for the breakpoint-spanning PCR product (black arrowhead) indicates that the deletion segregates with affected family members. A white arrowhead indicates an internal PCR control for DNA quality.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

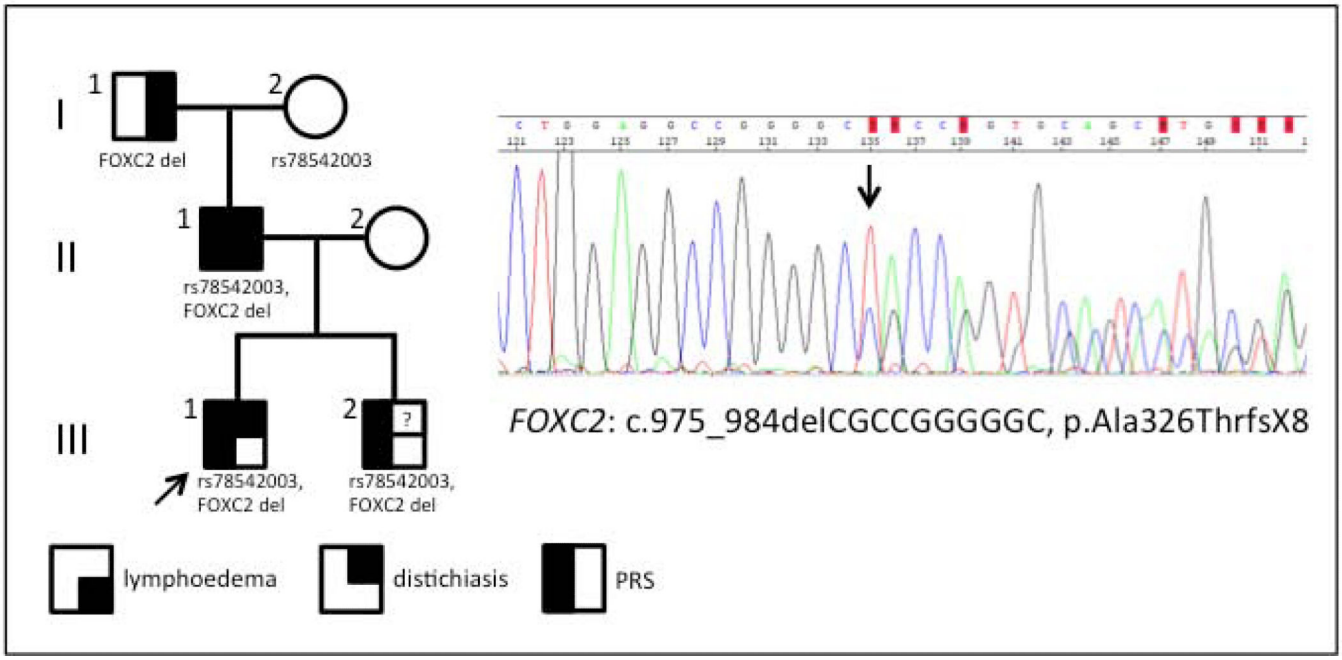

Fig 4.

Identification of a *FOXC2* deletion in family F2 from Benko et al. (2009). The 10 bp deletion induces a frameshift and a premature stop codon. rs78542003 is a SNP occurring in a CNE 1,441 kb upstream of *SOX9* (HCNE-F2); the variant was previously reported to segregate with PRS in family F2 (Benko et al. (2009)). The proband is indicated with an arrow. A question mark indicates that it is currently unknown whether the patient has distichiasis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

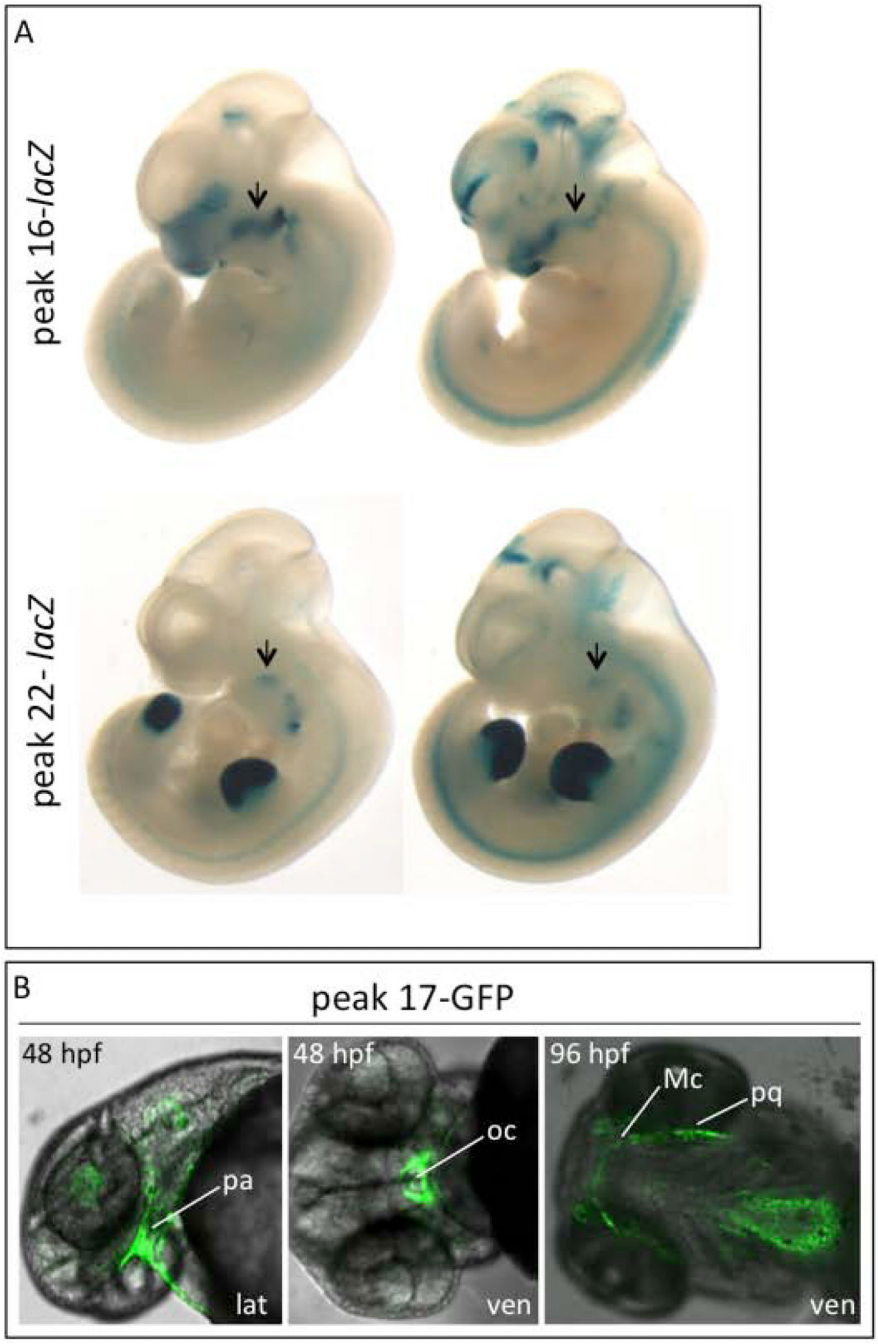

Fig 5.

p300 ChIP-Seq peak regions assayed for enhancer activity in transgenic animals. A, E11.5 transient transgenic mice harbouring reporter genes consisting of peak regions 16 (upper row) or 22 (lower row) upstream of *lacZ*. Arrows indicate mandibular expression. B, representative fluorescent confocal images of zebrafish embryos at 48 and 96 hpf expressing a stably integrated peak 17-*GFP* transgene. lat, lateral view; ven, ventral view; pa, pharyngeal arches; oc, oral cavity; Mc, Meckels's cartilage; pq, palatoquadrate.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Genomic co-ordinates of candidate *cis*-regulatory elements and reporter activity in transgenic mice

***cis*-regulatory element**[a]	**Co-ordinates (hg19)**[b]	**Co-ordinates of regions used for mouse transgenic assay constructs (mm9)**	**Vista Enhancer Browser ID**[c]	**Craniofacial enhancer activity *in vivo* at E11.5 (number embryos with activity/total embryos)**
E11.5 p300 peak 15	chr17:68,657,930-68,658,587	chr11:111,424,415-111,425,316	mm627	no activity
E11.5 p300 peak 16	chr17:68,670,660-68,671,244	chr11:111,429,837-111,430,753	mm628	positive (4/8)[d]
E11.5 p300 peak 17	chr17:68,734,773-68,735,865	chr11:111,488,426-111,489,423	mm629	no activity
E11.5 p300 peak 18	chr17:68,757,543-68,758,311	chr11:111,508,642-111,510,310	mm630	no activity
E11.5 p300 peak 19	chr17:68,772,358-68,773,236	chr11:111,527,366-111,528,791	mm631	no activity
E11.5 p300 peak 20	chr17:68,930,291-68,930,938	chr11:111,642,336-111,643,268	mm632	no activity
E11.5 p300 peak 21	chr17:69,177,260-69,178,008	chr11:111,822,417-111,824,278	mm633	no activity
E11.5 p300 peak 22	chr17:69,705,542-69,709,331	chr11:112,273,180-112,274,981	mm634	positive (5/10)[d]
E11.5 p300 peak 23	chr17:70,115,368-70,115,821	chr11:112,641,566-112,642,353	mm635	positive (5/13)[e]
E11.5 p300 peak 24	chr17:70,371,357-70,373,428	chr11:112,874,476-112,876,828	mm636	positive (3/9)[f]
E11.5 H3K27ac peak 4	chr17:69,696,698-69,697,764	-		
E11.5 H3K27ac peak 5	chr17:70,116,597-70,119,020	-		
E13.5 palate p300 peak 1	chr17:68,657,258-68,657,319	-		
E13.5 palate p300 peak 3	chr17:68,976,805-68,976,929	-		
E13.5 palate p300 peak 4	chr17:69,972,382-69,972,452	-		
hoc CNE-A	chr17:68,698,451-68,699,247	-		
hoc CNE-B	chr17:68,731,019-68,731,632	-		
hoc CNE-C	chr17:68,735,308-68,735,624	-		
hoc CNE-D	chr17:68,747,112-68,747,577	-		

[a] craniofacial (or palate, as indicated) ChIP-Seq peaks and/or conserved elements. In the latter case, elements are labelled hoc CNE-A to -D. hoc indicates human-opossum-chicken CNEs in the F1-PRS183 deletion overlap region. Note that hoc CNE-C falls within E11.5 p300 peak 17.

[b] theLiftOver tool at the UCSC genome browser was used for conversion of ChIP-Seq peak co-ordinates from mm9 to hg19.

[c] element identity available for viewing via the Vista Enhancer Browser at enhancer.lbl.gov.

[d] reporter expression in the mandible.

[e] reporter expression in the lower lip.

[f] reporter expression in the lateral nasal prominence.