Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 10649–10654, September 1998
Developmental Biology

Deletion of long-range regulatory elements upstream of *SOX9* causes campomelic dysplasia

VÉRONIQUE M. WUNDERLE*†‡, RICKY CRITCHER*, NICHOLAS HASTIE§, PETER N. GOODFELLOW*¶, AND ANDREAS SCHEDL§‖

*Department of Genetics, University of Cambridge, Downing Street, CB2 3EH Cambridge, United Kingdom; §Medical Research Council, Human Genetics Unit, Western General Hospital, Crewe Road, EH4 2XU Edinburgh, United Kingdom; and ¶SmithKline–Beecham Pharmaceuticals, New Frontiers Science Park, Third Avenue, CM19 5AW Harlow, United Kingdom

Edited by Shirley M. Tilghman, Princeton University, Princeton, NJ, and approved June 18, 1998 (received for review March 25, 1998)

ABSTRACT **Campomelic dysplasia (CD) is a rare, neonatal human chondrodysplasia characterized by bowing of the long bones and often associated with male-to-female sex-reversal. Patients present with either heterozygous mutations in the *SOX9* gene or chromosome rearrangements mapping at least 50 kb upstream of *SOX9*. Whereas mutations in *SOX9* ORF cause haploinsufficiency, the effects of translocations 5′ to *SOX9* are unclear. To test whether these rearrangements also cause haploinsufficiency by altering spatial and temporal expression of *SOX9*, we generated mice transgenic for human *SOX9-lacZ* yeast artificial chromosomes containing variable amounts of DNA sequences upstream of *SOX9*. We show that elements necessary for *SOX9* expression during skeletal development are highly conserved between mouse and human and reveal that a rearrangement upstream of *SOX9*, similar to those observed in CD patients, leads to a substantial reduction of *SOX9* expression, particularly in chondrogenic tissues. These data demonstrate that important regulatory elements are scattered over a large region upstream of *SOX9* and explain how particular aspects of the CD phenotype are caused by chromosomal rearrangements 5′ to *SOX9*.**

Major diagnostic criteria for the skeletal malformation syndrome, campomelic dysplasia (CD), are angulation of the tibiae and femura, hypoplastic scapulae, nonmineralization of the thoracic pedicles, 11 instead of 12 pairs of ribs, poor ossification of the pelvis, and bilateral talipes equinovaris (1–3). Other skeletal and nonskeletal defects are also associated with the disease, such as micrognathia, cleft palate, and low-set ears. Patients usually die soon after birth of respiratory distress, but the severity of the disease is variable and a few patients survive into adult life. Interestingly, male-to-female sex-reversal occurs in three-quarters of the XY CD patients, whose genitalia can be normal male, female, or ambiguous with various levels of male or female sexual differentiation (1, 3–5). The identification of *de novo* mutations in the *SOX9* gene of sex-reversed CD patients implicated *SOX9* as responsible for both skeletal and gonadal phenotypes. Only heterozygous mutations were detected in the patients, suggesting that CD has an autosomal dominant inheritance and haploinsufficiency is the probable cause of the CD and sex-reversal phenotypes (6, 7).

The human *SOX9* gene maps to chromosome 17q24 (6, 7) and belongs to the *SRY*-related HMG box (*SOX*) gene family (8, 9). *SOX* genes encode proteins with greater than 60% similarity at the amino acid level with the SRY DNA-binding domain or HMG box. The *SOX* genes have been isolated from a variety of organisms, and their role as transcription factors during embryogenesis has been suggested (9–12). *SOX9* encodes a putative 509 amino acid protein that contains an HMG box sharing 71% similarity with *SRY* HMG box and a transactivation domain at the C terminus, suggesting that SOX9 acts as a transcription activator (13, 14). Mouse *Sox9* gene shares 96% identity with its human homologue (15), indicating that the two proteins have similar functions. Analysis of *Sox9* expression during mouse development showed that *Sox9* is expressed at sites of chondrogenesis, suggesting its role in the initiation and maintenance of chondrocyte differentiation (15). Further *in vitro* and *in vivo* experiments confirmed this hypothesis showing activation by Sox9 of the collagen type II (Col2a1) gene, a major extracellular matrix component of cartilage (14, 16, 17). Expression during testis formation also was investigated, and differential expression during male and female sexual differentiation was demonstrated (18, 19).

A minority of CD patients do not show mutations in the ORF of *SOX9* but, instead, exhibit chromosome rearrangements mapping at least 50 kb upstream of *SOX9* (6, 7, 20). The relation between the translocation breakpoints and the CD phenotype has not been established, but several hypotheses have been suggested. The observed rearrangements may occur in another gene mapping upstream of *SOX9*. However, no expressed sequence has been identified in a 130-kb interval upstream of *SOX9* (7). Another possibility is that additional 5′ exons of the *SOX9* gene are interrupted by the breakpoints, resulting in the loss of *SOX9* function. 5′ RACE and RNase protection analysis have, however, confirmed the transcription start site for *SOX9* to a region mapping at least 50 kb 3′ to the observed translocation breakpoints (refs. 6 and 7; S. Guioli, personal communication). Alternatively, the rearrangements may affect spatial and temporal expression of *SOX9* by disrupting its control region. This could be caused either by silencing sequences brought into juxtaposition with *SOX9* or by the removal of regulatory sequences from the proximity of *SOX9*.

To investigate the effects of the distant chromosome rearrangements on *SOX9* expression, chromosome breakpoints observed in CD patients were simulated by constructing mice transgenic for human *SOX9* yeast artificial chromosomes (YACs) containing variable amounts of DNA sequence upstream of *SOX9*. Our results show that *SOX9* expression derived from a 350-kb YAC with a breakpoint similar to the ones observed in two CD patients is dramatically downregulated particularly in the tissues where chondrogenesis

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviations: YAC, yeast artificial chromosome; CD, campomelic dysplasia; dpc, day(s) post coitus.

†Present address: Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, M5G 1X5 Toronto, ON, Canada.
‡To whom reprint requests and correspondence should be addressed at the present address. e-mail: wunderle@mshri.on.ca.
‖Present address: Max Delbrück Centrum for Molecular Medicine, Robert Rössle Strasse 10, 13122 Berlin Buch, Germany.

occurs. We also demonstrate that the CD phenotype observed in patients with such rearrangement is caused by the deletion of tissue-specific regulatory elements scattered over, at least, 200 kb upstream of *SOX9*.

MATERIALS AND METHODS

Case Report. Patient cu004 presents clinical and radiological features typical for CD with normal male gonads and external genitalia. Clinical features of patient cu002 were reported previously (3).

Deletion of YAC Sequences and Isolation of New STS Markers. Chromosome fragmentation vectors, pBP108 and pBP109 (Stratagene), containing *Alu* repetitive elements in either orientation were used as described previously (21) to delete sequences from the acentric URA arm of YAC ICRF-946e12. Terminal sequences from YACs 108.42 and 108.22 as well as ends from cosmids F11143, G1128, and C0456 located in the *SOX9* region (gift from J. Foster, University of Cambridge, Cambridge, U.K.) were isolated by using the vectorette protocol (26, 27).

The sequences of the STS markers isolated in this study are the following: 42HcL for 5′-GCC AAC ATG AGA CTT CTA TTC-3′; 42HcL rev, 5′-GCC AAT TCT GAT GGC CAA ATC-3′; 22RsL for 5′-CTC TGG TCA TGC AGT AAG TCT C-3′; 22RsL rev, 5′-CTC AGA TCA AAC CCA AGC TAC-3′; c456HcR for 5′-CTT TAT ACT TCA GTA CTC GG-3′; c456HcR rev, 5′-CAG CCT ATA GGG TGT ATT AGC-3′; F11143R for 5′-GTA TCA TAA CAG GGG CAG AGC-3′; F11143R rev, 5′-CAG CCT ATA GGG TGT ATT AGC-3′; F11143L for 5′-GCT CTA GGA AGT CAG TAT TG-3′; F11143L rev, 5′-GAG TTG AAG CCC AGC CTC TG-3′; G1128HcL for 5′-GTT AAA AAC AAT GGA GGC AGG-3′; G1128HcL rev, 5′-CAA AAC AAG ATG GCT ATG AGG-3′.

Construction of the Yeast Targeting Vector. A 0.5-kb *Xba*I–*Sal*I fragment containing the polyadenylation site of simian virus 40 was isolated from pGT1.8IRESβgeo vector (28) and ligated into pBluescript II SK (−) (Boehringer). After removal of the *Bam*HI site, the resulting *Xba*I–*Sal*I 0.5-kb SVpA fragment was cloned into an *Xba*I–*Sal*I-cut pSAβ-gal vector (gift from C. Boulter, University of Cambridge), and the *Xba*I site subsequently was removed. *Bam*HI–*Sal*I digestion of the resulting vector released a 1.9-kb fragment containing the PGK promoter upstream of the neomycin gene and SVpA. This fragment was cloned into the yeast shuttle vector pRS405 (Stratagene), and this construct was named pRS/neoSVpA. In parallel, a 3-kb *Eco*RI–*Hin*cII fragment containing 2.3 kb of genomic sequence mapping upstream of *SOX9* and 646 bp of *SOX9* first exon was cloned into an *Eco*RI–*Hin*cII pSK− vector. After removal of a *Bam*HI site, annealed oligonucleotides, NBB1 and NBB2, containing *Nru*I, *Bgl*II, and *Bam*HI restriction sites were ligated into this *Hin*cII–*Kpn*I-digested vector (NBB1, 5′-TCG AAT CGC GAG GAG ATC TAT GAC GGA TCC GTA C-3′; NBB2, 5′-GGA TCC GTC ATA GAT CTC CTC GCG AT-3′). This ligation reconstituted a *Sal*I site at the *Hin*cII cloning site. A 3.1-kb *Bam*HI fragment obtained from pSAβ-gal vector and containing a *lacZ* gene lacking its transcription start site was then cloned in the correct orientation into the *Bgl*II–*Bam*HI vector, placing *lacZ* in frame with *SOX9* coding sequence. This vector once digested with *Xba*I and *Bam*HI, released a 4.1-kb fragment containing 350 bp of genomic sequence upstream of *SOX9*, 646 bp of *SOX9* first exon, and the β-galactosidase gene in frame with *SOX9*. This fragment was ligated into *Xba*I–*Bam*HI-cut pRS/neoSVpA, and the resulting plasmid was digested with *Not*I and *Xba*I. A 0.5-kb *Not*I–*Xba*I fragment from *SOX9* third exon subsequently was ligated, leading to the formation of the final construct 5′3′S9β-neoSVpA. The vector was then linearized with *Xba*I and used for yeast transformation.

YAC DNA Purification and Microinjection. Gel-purified YAC DNA prepared as described previously (29) was microinjected into CBA × C57BL/6J F_1 fertilized oocytes. Transgenic lines were obtained by crossing the transgenic animals with CD1 albino mice.

Determination of Copy Number. A semiquantitative PCR assay was performed on DNA from the progeny of each transgenic line by using primers amplifying, with equal efficiency, a 318-bp homologous region of mouse and human *SOX9* loci. Four hundred nanograms of DNA extracted from mouse tails was used in a 100-μl PCR. After 20 and 22 cycles, 18-μl aliquots were collected. One microliter of 0.2 M NaCl and 2.5 units *Afl*III were added, and the reactions were incubated for 2 hr at 37°C. Whereas mouse PCR products remained undigested, human-amplified sequence resulted in two fragments of 200 and 118 bp. After Southern blotting and hybridization with an oligonucleotide detecting mouse 318-bp and human 200-bp amplified fragments, the intensity of each signal was quantified and compared by using PHOSPHORIMAGER software (Molecular Dynamics). SOX9M/H for CTT CTC GCC TTT CCC GGC CAC; SOX9M/H rev, CAT GGG CAC CAG CGT CCA GTC; SOX9M/H probe, TCG GGC TCC GGC TCG GAC AC.

β-Galactosidase Staining and Sectioning. Embryos were fixed and stained as described (30). Color development was performed from 2 to 5 hr. Embryos were then fixed overnight in 4% paraformaldehyde in PBS, dehydrated, and embedded in paraffin wax.

***In Situ* Hybridization.** Mouse embryos were collected and processed as described (18, 31). Labeled sense and antisense riboprobes were generated from cDNA subclone pb*Sox9* as described previously (19). No signal was obtained with the sense probe.

RESULTS

Mapping of Patient cu002 and cu004 Breakpoints. We concentrated our study on two CD patients with chromosomal rearrangements upstream of *SOX9*: an XY female (cu002) (3) with a chromosome 17 inversion and an XY male (cu004) with a t(9,17) translocation. Using fluorescence *in situ* hybridization analysis, a 2-Mb YAC (ICRF-946e12) encompassing the *SOX9* locus (6) was demonstrated to cross the breakpoints in both patients (Fig. 1). Deletions upstream of *SOX9* were performed on the 2-Mb YAC by using a chromosome fragmentation strategy (21). A panel of YACs containing various amounts of DNA upstream of *SOX9* hence was produced, and six new STS markers (Fig. 1) were identified by cloning the distal terminal sequences from two YAC derivatives as well as from three cosmids mapping to this region (6). To map the breakpoint in cu004 patient, these markers were tested on a somatic cell hybrid derived from this patient, which retained the derivative chromosome 9 in the absence of the normal 17 and reciprocal translocated chromosome. This analysis mapped the cu004 breakpoint to 110–140 kb upstream of *SOX9* (Fig. 1). Two probes previously used for cytogenetic mapping (M. Dominguez-Steglich, personal communication) also were tested on the YAC derivative panel locating the cu002 breakpoint to 75–350 kb upstream of *SOX9* (Fig. 1).

Construction of the YAC Transgenes and Generation of the Transgenic Mice. From the mapping data previously described and fluorescence *in situ* hybridization analysis, we determined that the 600-kb YAC derivative, 108.43, crossed both breakpoints (data not shown). With sequences extending 350 kb 5′ of *SOX9*, this YAC was likely to contain most of *SOX9* regulatory elements. In contrast, the 350-kb YAC, 108.15, contained 75 kb of *SOX9* upstream sequences with a similar breakpoint as patients cu002 and cu004, and patient E, whose breakpoint was mapped 88 kb 5′ of *SOX9* (6). Both YACs were used to generate transgenic mice. To facilitate

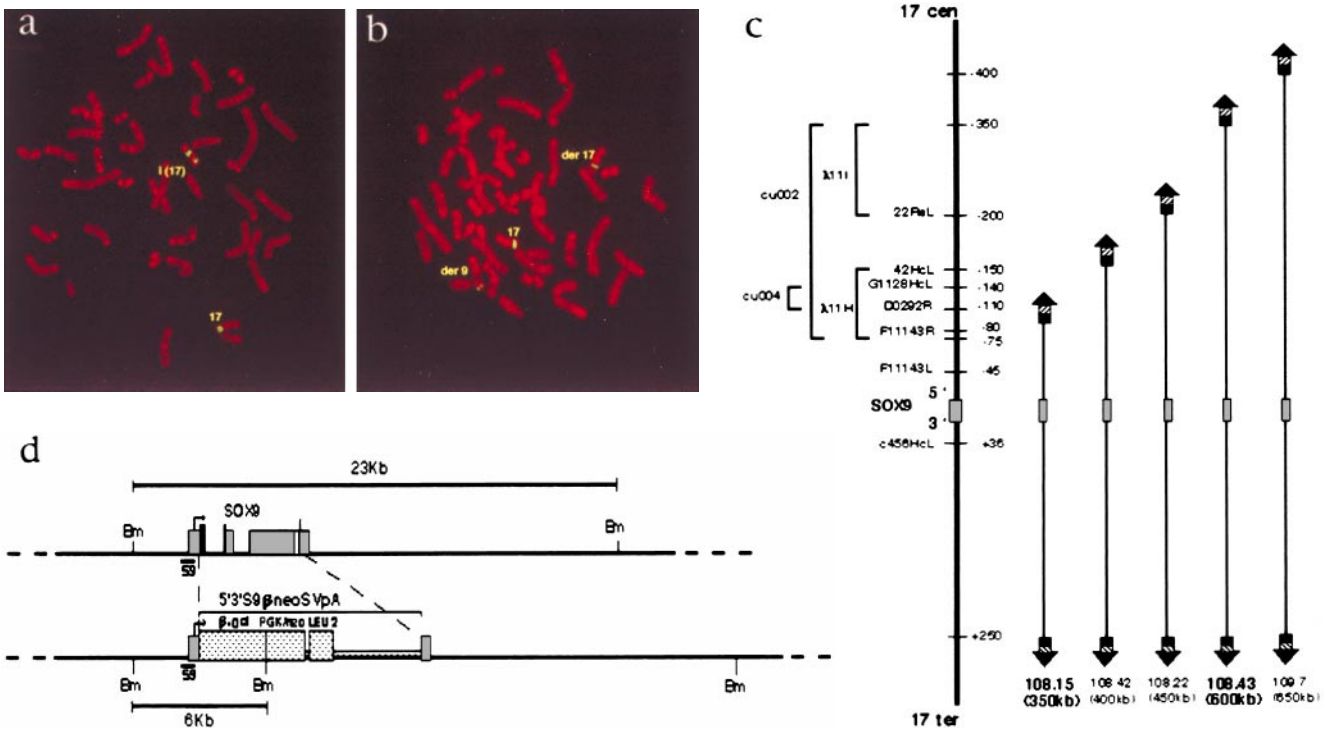

FIG. 1. Molecular analysis of CD patients and construction of the *SOX9/lacZ* YACs. The YAC ICRF-946e12 was used as probe on metaphase spreads prepared from lymphoblastoid cell lines from sex-reversed patient cu002: 46,XY, inv17(q11.2;q24.3–25.1) (*a*) and patient cu004: 46,XY, t(9;17)(p13;q23.3–24.1) (*b*). (*c*) Deletion of YAC sequences was performed by using chromosome fragmentation vectors, and YAC derivatives still positive for *SOX9* were selected. Fragmented YACs are depicted as vertical lines, and the names of identified markers, including their distances in kb, relative to *SOX9*, are given on a schematic representation of human chromosome 17q. Fluorescence *in situ* hybridization analysis performed on both patients and hybridization using the YAC derivative panel mapped the two cytogenetic probes λ11H, λ11I, to 75–150 kb and 200–350 kb 5′ of *SOX9*, respectively. (*d*) Schematic representation of the yeast-targeting vector 5′3′S9β-neoSVpA after homologous recombination with the YACs. The vector 5′3′S9β-neoSVpA was designed to integrate a *lacZ* reporter gene in frame with *SOX9* coding sequence and to delete *SOX9* HMG box (large, solid box) and transactivation domain (open box) via one event of homologous recombination with the 108.15 and 108.43 YACs.

detection of transcription from the transgenes, we introduced a *lacZ* reporter gene in frame with exon 1 of *SOX9*. A yeast-targeting construct was designed to replace the *SOX9* HMG box and transactivation domain by *lacZ*, to avoid possible dominant effects caused by overexpression of the human *SOX9* gene (Fig. 1). After selection of correctly targeted YACs, gel-purified YAC DNA was microinjected into fertilized oocytes. After reimplantation, DNA from tail tips of the newborn mice was analyzed by PCR and Southern blotting, revealing that 8 (4.3%) and 2 (9.1%) of the founder mice were transgenic for the 600-kb and 350-kb YACs, respectively. Copy numbers of the transgenes were determined in the offspring by using a semiquantitative PCR assay (data not shown), and integrity of the YAC transgenes was investigated by PCR screening and Southern blotting (Fig. 2). This analysis indicated that lines A46/A63 and A74/A75 were transgenic for intact copies of the 600-kb and 350-kb YACs, respectively.

Expression Derived from the 600-kb YAC. Comparison of *SOX9/lacZ* expression derived from the 600-kb YAC with mouse endogenous *Sox9* shows similar patterns in most tissues. As with mouse *Sox9*, expression derived from the 600-kb YAC is detected in the mesenchyme condensations of the bones of the skull, the branchial arches and their subsequent derivatives, and the developing ear (Fig. 2). Moreover, *SOX9/lacZ* expression is found in the neuroepithelial layer of the neural tube and developing brain as well as in the notochord. Expression in the paraxial mesoderm is also detected within the developing vertebrae and ribs. Therefore, most of *SOX9* regulatory elements are present on the 600-kb YAC and must be conserved between mouse and human. However, differences in the expression patterns are observed in particular tissues. Indeed, in the limbs, *SOX9/lacZ* expression appears first in the developing long bones, conversely to *Sox9*, whose expression was stronger in the hand and foot plates. Moreover, although mouse *Sox9* is expressed in genital ridges from 10.5 days post coitus (dpc), no β-galactosidase activity is detected in the developing gonads at any stage of development.

Comparison Between Expression from the 600-kb and 350-kb YACs. We next compared the expression from the 600-kb YAC with the 350-kb construct to explore the consequence of 5′ deletions on *SOX9* expression. Embryos carrying a single copy of the short construct (line A45.1) show, throughout development, a dramatic decrease or absence of *SOX9/lacZ* expression in all tissues except the neuroectoderm (Fig. 3). These modifications were further analyzed in embryos from lines A74 and 75 where, because of the higher copy numbers of the transgene, β-galactosidase activity was easily detectable.

Whereas expression from the 350-kb YAC occurs at the correct location within the otic vesicle, an important delay in the establishment of the expression pattern takes place (Fig. 3). This delay may correspond to the time necessary for SOX9/LacZ proteins produced by the 350-kb YAC to accumulate in the cells until detection of sufficient β-galactosidase activity. In the branchial arches of the embryos transgenic for the short YAC, expression is detected mainly in the maxillary process of the first arch, indicating that the deletion of sequences upstream of *SOX9* results in the loss of *SOX9/lacZ* expression in the mandibular process and second branchial arch (Fig. 3). Analysis of expression in the paraxial mesoderm revealed that *SOX9/lacZ* expression is absent from the developing ribs and cervical vertebrae. Furthermore, expression in the developing limbs showed that *SOX9/lacZ* is present mainly in the mes-

a

	YAC TRANSGENIC LINES								
	A46 (1-2)	A63 (1-3)	A53.1 (1)	A53 (1)	A45.1 (1)	A52 (1)	A45 (1)	A74 (3)	A75 (5)
MARKERS									
HIS3 (-350Kb) (YAC left arm)	+	+	-	-	-	-	-	+	+
22ReL (-200Kb)	+	+	+	-	-	-	+		
42HcL (-150Kb)	+	+	+	+	-	-	-		
G1128L (-140Kb)	+	+	+	+	-	-	-		
D0292R (-110Kb)	+	+	+	+	-	-	-		
F11143R (-80Kb)	+	+	+	+	-	-	-		
F11143L (-45Kb)	+	+	+	+	+	+	-	+	+
SOX9 (1)	+	+	+	+	+	+	+	+	+
C456R (+36Kb)	+	+	+	+	+	-	+	+	+
TRP1 (+250Kb) (YAC right arm)	+	+	-	-	+	+	-	+	+
TISSUES									
limb	+	+	-	-	-			-	-
vertebral column	+	+	-	-	-			-	-
branchial arches	+	+	+	+	-			-	-
otic vesicle	+	+	+	+	-			-	-

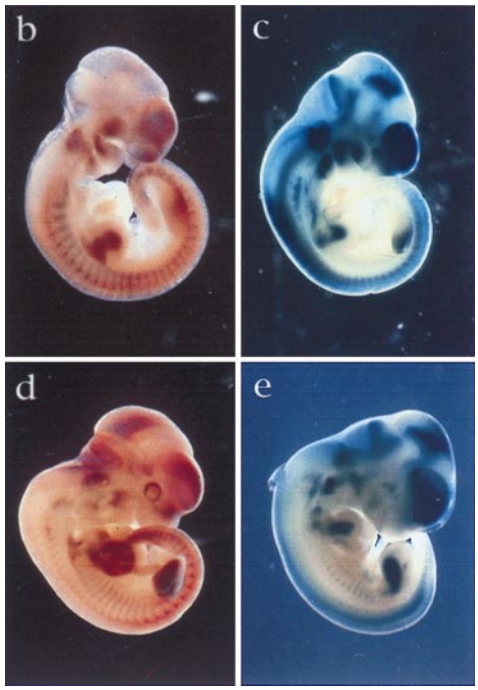

FIG. 2. (*a*) Analysis of transgenic offspring. After microinjection of purified YAC DNA, integrity and copy number of the transgenes were tested in the transgenic mice. Copy numbers are indicated between brackets below the names of each transgenic line. Markers used for testing YAC transgene integrity are listed, and their positions, relative to *SOX9*, are indicated. +, Marker present in the tested mice; −, absent marker. Tissues showing modifications in *SOX9*/*lacZ* expression are also reported. +, Normal tissue; −, altered tissue. (*b–e*) Comparison between expression derived from mouse endogenous *Sox9* gene and the 600-kb human YAC transgene. (*b* and *d*) *In situ* hybridization using a *Sox9* RNA probe on wild-type embryos. (*c* and *e*) β-Galactosidase staining of embryos from line A46 transgenic for one copy of the 600-kb YAC. (*b* and *c*) Embryos (10.5 dpc). (*d* and *e*) Embryos (11.5 dpc). *SOX9*/*lacZ* expression is detected at the correct time in the mesenchyme of the head, the otic vesicles, and branchial arches. Expression is also found in the neural tube, notochord, migrating sclerotomal cells, and developing ribs. Whereas expression in the limb buds is highly similar at 10.5 dpc, differences in the patterns occur from 11.5 dpc. At that stage, *SOX9*/*lacZ* is strongly expressed in the mesenchymal condensation of the long bones and expression in the hand and foot plates is detected only from 12.5 dpc. In contrast, from 11.5 dpc, mouse *Sox9* expression is strongly expressed in the hand and foot plates, the developing long bones only expressing weakly *Sox9*. This difference may be caused by species-specific elements or by the lack, on the YAC, of control elements involved in *SOX9*/*Sox9* expression during limb development.

enchyme condensations of the hand and foot plates rather than the developing long bones (Fig. 3).

DISCUSSION

The detection of cytogenetic rearrangements, such as balanced translocations, in patients suffering from genetic disorders has greatly facilitated the identification of genes involved in diseases. However, the modification of molecular and cellular mechanisms caused by these rearrangements is difficult to analyze, as modification of gene expression may be subtle and tissues from patients are rarely available. The transgenic approach described in this study allowed us to directly investigate the relationship between chromosomal rearrangements and their effects on gene expression.

The relationship between *SOX9* expression and chromosome rearrangements upstream of *SOX9* was investigated by constructing mice transgenic for a human 600-kb YAC crossing the breakpoint in two CD patients and for a 350-kb YAC truncated 5′ to *SOX9*, with a breakpoint similar to the ones observed in CD patients. The expression pattern derived from the 600-kb YAC appeared comparable to mouse *Sox9* in most of the skeletal tissues. Thus, most of the elements necessary for *SOX9* proper expression in these tissues are located up to 350 kb upstream of *SOX9* and are highly conserved between mouse and human. However, no *SOX9*/*lacZ* expression derived from the YACs was detected in the genital ridges of the transgenic mice. This result can be explained either by the lack, on the YAC, of gonad-specific regulatory elements located further 5′ or 3′ of *SOX9* or by the removal of such elements during *lacZ* integration. Alternatively, species-specific elements may be involved in the activation of the sex-determination pathway. Interestingly, previous experiments with human *WT1* transgenes (A. Moore, personal communication) also failed to show proper expression in the genital ridges, suggesting the rapid evolution of elements involved in mouse and human sex determination. This hypothesis is further supported by the observation that a human *SRY* transgene is unable to induce sex reversal in mice (22).

When compared with embryos transgenic for the 600-kb YAC, embryos carrying the 350-kb YAC showed a dramatic decrease in the level of *SOX9* expression, particularly in tissues where chondrogenesis occurs. Additional analysis performed on the transgenic lines carrying multiple copies of the short construct showed a delay in the onset of *SOX9*/*lacZ* expression in the otic vesicles and revealed that only a reduced number of cells expressed the transgene in the branchial arches. Furthermore, timing and level of expression were modified particularly in the limb buds and no expression was detected in the cervical vertebrae and ribs. These results clearly demonstrate that the deletions not only affect the level of expression but also the onset of *SOX9* expression in particular tissues. Careful comparison of expression patterns derived from the different YAC transgenes suggests the presence of several tissue-specific cis-acting elements scattered over a large region upstream of *SOX9*. Transgenic lines A53.1 and A53, for example, do not show proper expression in the developing ear and branchial arches or limb and vertebral column, respectively, suggesting the presence of elements regulating expression in those tissues to 45–150 kb and 200–350 kb 5′ to *SOX9* (Fig. 2). Tissues with modified *SOX9* expression correspond to sites of frequent

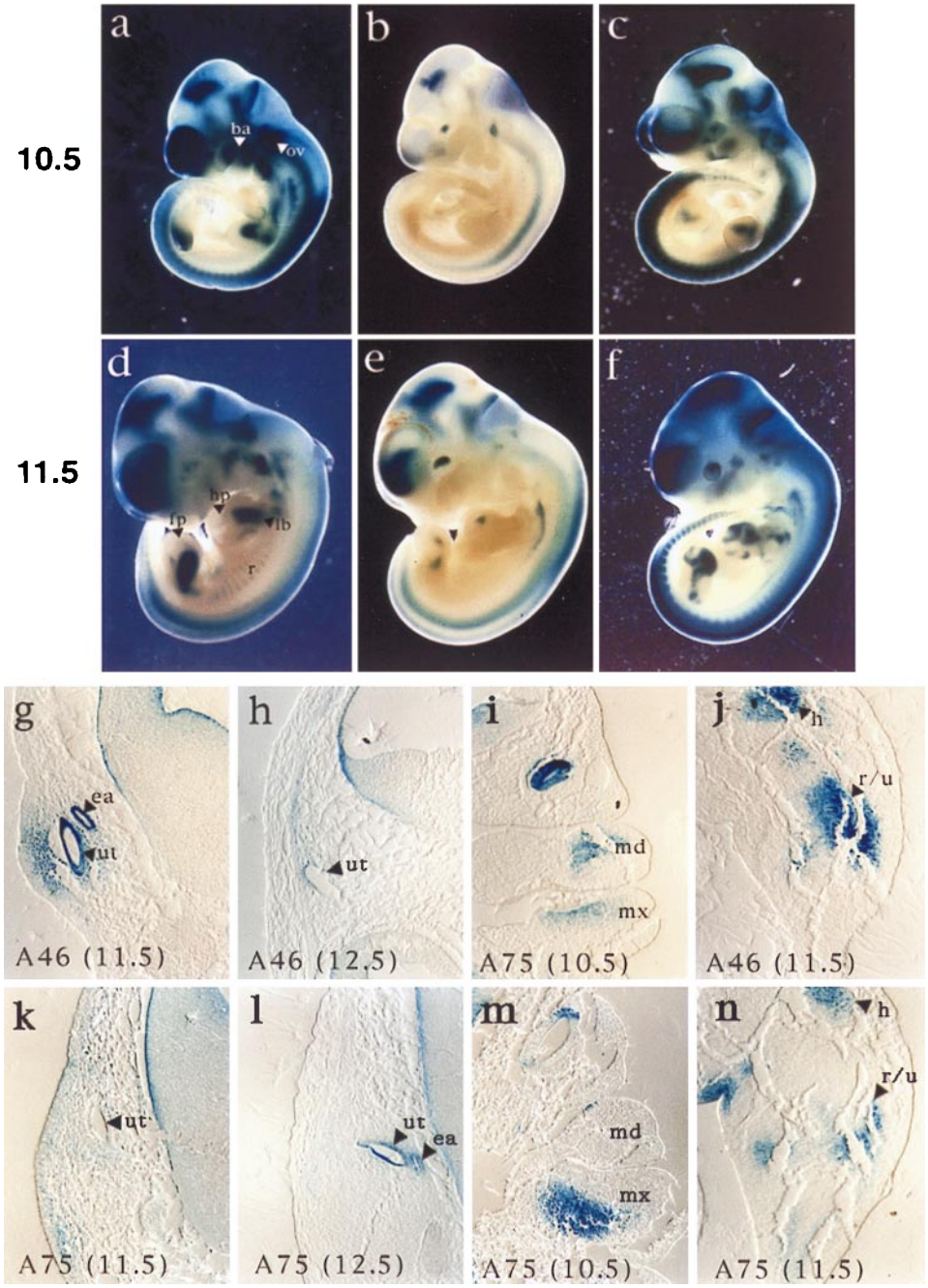

FIG. 3. Comparison between *SOX9/lacZ* expression derived from the 350-kb and 600-kb YACs. (*a–f*) β-Galactosidase staining of embryos from line A46 (one copy of the 600-kb YAC), line A45.1 (one copy of the short construct), and line A75 (multiple copies of the 350-kb YAC). (*g–n*) Microtome sections performed on 10.5- to 12.5-dpc embryos from lines A46 and A75. ba, branchial arches; ov, otic vesicle; lb, long bones; hp, hand plate; fp, foot plate; r, ribs; ea, endolymphatic duct; ut, utriculus; md, mandibular process of the first branchial arch; mx, maxillary process; h, humeral mesenchymal condensation; r/u, radio-ulna mesenchymal condensation. Throughout development, expression in line A45.1 is found dramatically decreased in all tissues normally expressing *SOX9/lacZ* (line A46). Expression in the neuroectoderm, however, is only mildly affected. (*g*, *h*, *k*, and *l*) Expression derived from the 350-kb YAC occurs at the correct location in the developing ear, but a significant delay in the establishment of the expression pattern takes place: *SOX9/lacZ* expression in the endolymphatic and utricular portions of the otocyst is only detected from 12.5 dpc instead of 11.5 dpc with the 600-kb YAC. This delay may correspond to the time necessary for SOX9/LacZ proteins to accumulate in the cells to obtain detectable levels of β-galactosidase activity. (*i* and *m*) In the developing branchial arches, the deletion of *SOX9* upstream sequences results in *SOX9/lacZ* expression only in the maxillary process of the first branchial arch. This suggests the involvement of multiple elements in the control of *SOX9* expression in distinct regions of the branchial arches. (*j* and *n*) In contrast to line A46, where expression in the developing long bones precedes expression in the hand and foot plates, expression derived from the 350-kb YAC is induced from 11.5 dpc in the hand and foot plates, with the expression in the long bones remaining weak. The 5′ deletion therefore alters onset and level of expression in particular regions of the limb bud.

skeletal abnormalities in CD patients. The decrease or lack of expression in the branchial arches and bone derivatives, as described in the transgenic mice, is likely to affect the formation of specific bones of the face and larynx/trachea cartilages, both common features in the CD phenotype. Alteration of *SOX9* expression in the developing vertebral column could also result in the formation of the CD-specific abnormalities of vertebrae and ribs. The type of modification encountered by the different tissues in the transgenic mice, considered in the view of the CD phenotype, therefore

provides information about the cause of particular skeletal defects characteristic of the CD condition. Moreover, evidence of *SOX9* expression remaining after 5′ deletion, even at a low level in particular tissues, may explain the milder phenotype observed in translocated patients (3), demonstrating the importance of a tight control of *SOX9* expression levels during embryogenesis.

The identification of tissue-specific elements scattered over at least 200 kb upstream of *SOX9* suggests that the phenotype of the CD patients depends on the position of the chromosome rearrangements. Therefore, we compared the position of published CD breakpoints (6, 7, 23) with the probable position of elements regulating *SOX9* expression. Although most of the breakpoints are located within large intervals, none of them seems to exceed 400 kb upstream of *SOX9*. This analysis indicates that although the breakpoints are not clustered, they all are potentially deleting the same regulatory elements leading to a similar phenotype in the CD patients.

Although the presence of regulatory elements at such long distances seems unusual for a mammalian gene, other examples have been described previously, such as the locus control region located 50–65 kb upstream of the human *β-globin* locus (24) and enhancer elements mapped at least 150 kb downstream of the human *PAX6* gene (25). Hence, the work described in this study not only provides evidence of the direct effect on *SOX9* expression of distant breakpoints observed in CD patients but also reveals the complex organization of *SOX9* regulatory region. Future experiments with additional constructs will allow us to further delineate and identify the long-distance regulatory elements involved in *SOX9* expression during embryogenesis.

We thank Prof. I. Young and Dr. F. Hinde for the cu004 case report, and A. Schafer, L. McInnes, and B. Doe for technical assistance. We are also grateful to J. Foster, S. Morais da Silva, C. Huxley, and C. Boulter for gifts of reagents. This work was supported by a Wellcome Trust Prize Studentship to V.M.W.

1. Houston, C. S., Opitz, J. M., Spranger, J. W., MacPherson, R. I., Reed, M. H., Gilbert, E. F., Herrmann, J. & Schinzel, A. (1983) *Am. J. Med. Genet.* **15,** 3–28.
2. McKusick, V. A. (1990) in *Camptomelic Dwarfism* (Johns Hopkins Univ. Press, Baltimore), pp. 1074–1075.
3. Mansour, S. C., Hall, C. M., Pembrey, M. E. & Young, I. D. (1995) *J. Med. Genet.* **32,** 415–420.
4. Hovmoller, M. L., Osuna, A., Eklof, O., Fredga, K., Hjerpe, A., Lindsten, J., Ritzen, M., Stanescu, V. & Svenningsen, N. (1977) *Hereditas* **86,** 51–62.
5. Dagna Bricarelli, F., Fraccaro, M., Lindsten, J., Muller, U., Baggio, P., Doria Lamba Carbone, L., Hjerpe, A., Lindgren, F., Mayerova, A., Ringertz, H., *et al.* (1981) *Hum. Genet.* **57,** 15–22.
6. Foster, J. W., Dominguez-Steglish, M. A., Guioli, S., Kwok, C., Weller, P. A., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I. D., Goodfellow, P. N., Brook, J. D. & Schafer, A. J. (1994) *Nature (London)* **372,** 525–530.
7. Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasentes, J., Dagna Bricarelli, F., Keutel, J., Hurstert, E., *et al.* (1994) *Cell* **79,** 1111–1120.
8. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P. & Lovell-Badge, R. (1990) *Nature (London)* **346,** 245–250.
9. Denny, P., Swift, S., Brand, N., Dabhade, N., Barton, P. & Ashworth, A. (1992) *Nucleic Acids Res.* **20,** 2887.
10. van de Wetering, M., Oosterwegel, M., van Norren, K. & Clevers, H. (1993) *EMBO J.* **12,** 3847–3854.
11. Connor, F., Wright, E., Denny, P., Koopman, P. & Ashworth, A. (1995) *Nucleic Acids Res.* **23,** 3365–3372.
12. Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. & Lovell-Badge, R. (1996) *Development (Cambridge, U.K.)* **122,** 509–520.
13. Sudbeck, P., Lienhard Schmitz, M., Baeuerle, P. A. & Scherer, G. (1996) *Nat. Genet.* **13,** 230–232.
14. Ng, L. J., Wheatley, S., Muscat, G. E. O., Conway-Campbell, J., Bowles J., Wright, E., Bell, D. M., Tam, P. P. L., Cheah, K. S. E. & Koopman, P. (1997) *Dev. Biol.* **183,** 108–121.
15. Wright, E., Hargrave, M. R., Christiansen, J., Cooper, L., Kun, J., Evans, T., Gangadharan, U., Greenfield, A. & Koopman, P. (1995) *Nat. Genet.* **9,** 15–20.
16. Lefebvre, V., Huang, W., Zhou, G., Bi, W., Harley, V., Goodfellow, P. N. & de Crombrugghe, B. (1997) *Mol. Cell. Biol.* **17,** 2336–2346.
17. Bell, D. M., Leung, K. K. H., Wheatley, S. C., Ng, L. J., Zhou, S., Ling, K. W., Sham, M. H., Koopman, P., Tam, P. P. L. & Cheah, K. S. E. (1997) *Nat. Genet.* **16,** 174–178.
18. Kent, J., Wheatly, S. C., Andrews, J. E., Sinclair, A. E. & Koopman, P. (1996) *Development (Cambridge, U.K.)* **122,** 2813–2822.
19. Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A. & Lovell-Badge, R. (1996) *Nat. Genet.* **14,** 62–68.
20. Kwok, C. P., Weller, P. A., Guioli, S., Foster, J. W., Mansour, S., Zuffardi, O., Punnett, H. H., Dominguez-Steglich, M. A., Brook, J. D., Young, I. D., *et al.* (1995) *Am. J. Hum. Genet.* **57,** 1028–1036.
21. Reeves, R. H., Pavan, W. J. & Hieter, P. (1992) *Methods Enzymol.* **216,** 584–603.
22. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Lovell-Badge, R. (1991) *Nature (London)* **351,** 117–121.
23. Wirth, J., Wagner, T., Meyer, J., Pfeiffer, R. A., Tietze, H.-U., Zimmer, J., Weissenbach, J., Schempp, W. & Scherer, G. (1996) *Hum. Genet.* **97,** 186–193.
24. Forrester, W. C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G. & Groudine, M. (1987) *Nucleic Acids Res.* **15,** 10159–10177.
25. Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., Mannens, M., *et al.* (1995) *Hum. Mol. Genet.* **4,** 415–422.
26. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C. & Markham, A. F. (1990) *Nucleic Acids Res.* **18,** 2887–2890.
27. Potier, M. C., Dutriaux, A. & Reeves, R. (1996) *Mamm. Genome* **7,** 85–88.
28. Mountford, P. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 4303–4307.
29. Schedl, A. (1996) *Methods Mol. Biol.* **54,** 293–306.
30. Price, J. (1994) in *Essential Development Biology: A Practical Approach*, eds. Stern, C. D. & Holland, P. (IRL, Oxford), pp. 187–190.
31. Wilkinson, D. (1993) in *Guide to Techniques in Mouse Development*, eds. de Wasserman, P. M. & de Pamphilis, M. L. (Academic, New York), pp. 361–372.