Human Molecular Genetics, 2003, Vol. 12, No. 19 **2491–2501**
DOI: 10.1093/hmg/ddg248

Identification of *SATB2* as the cleft palate gene on 2q32–q33

David R. FitzPatrick[1,*], Ian M. Carr[2], Lorna McLaren[3], Jack P. Leek[2], Patrick Wightman[1], Kathy Williamson[1], Philippe Gautier[1], Niolette McGill[1], Caroline Hayward[1], Helen Firth[4], Alex F. Markham[2], Judy A. Fantes[1] and David T. Bonthron[2]

[1]Cell and Molecular Genetics, MRC Human Genetics Unit, Edinburgh, UK, [2]Molecular Medicine Unit, University of Leeds, Clinical Sciences Building, St James's University Hospital, Leeds LS9 7TF, UK, [3]Molecular Medicine Centre, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK and [4]Department of Clinical Genetics, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK

Received May 9, 2003; Revised July 1, 2003; Accepted July 16, 2003

Cytogenetic evidence, in the form of deletions and balanced translocations, points to the existence of a locus on 2q32–q33, for which haploinsufficiency results in isolated cleft palate (CPO). Here we show by high-resolution FISH mapping of two *de novo* CPO-associated translocations involving 2q32–q33 that one breakpoint interrupts the transcription unit of the gene encoding the DNA-binding protein *SATB2* (formerly KIAA1034). The breakpoint in the other translocation is located 130 kb 3′ to the *SATB2* polyadenylation signal, within a conserved region of non-coding DNA. The *SATB2* gene is transcribed in a telomeric to centromeric direction and lies in a gene-poor region of 2q32–q33; the nearest confirmed gene is 1.26 Mb centromeric to the *SATB2* polyadenylation signal. *SATB2*-encoding transcripts are assembled from 11 exons that span 191 kb of genomic DNA. They encode a protein of 733 amino acids that has two CUT domains and a homeodomain and shows a remarkable degree of evolutionary conservation, with only three amino acid substitutions between mouse and human. This protein belongs to the same family as SATB1, a nuclear matrix-attachment region binding protein implicated in transcriptional control and control of chromatin remodelling. There are also sequence similarities to the *Drosophila* protein DVE. Whole mount *in situ* hybridization to mouse embryos shows site- and stage-specific expression of *SATB2* in the developing palate. Despite the strong evidence supporting an important role for *SATB2* in palate development, mutation analysis of 70 unrelated patients with CPO did not reveal any coding region variants.

INTRODUCTION

Fusion of the secondary palate is one of the last morphogenetic processes to occur in the human embryo (1). Cleft palate is the result of failure of this fusion process, and occurs in 0.05–0.1% of live births (2). Cleft palate associated with cleft lip is generally considered to be a secondary phenomenon and is genetically distinct from cleft palate only (CPO) (3). CPO shows considerable a etiological heterogeneity, being seen in 15% of all cytogenetically visible autosomal deletions (4) and in at least 270 non-chromosomal malformation syndromes (5). This suggests that CPO may result from many different morphological anomalies in the embryo. Nonetheless, CPO most commonly presents as an isolated malformation, and studies of twin concordance (6,7) and familial aggregation (8,9) show a significant genetic contribution to the etiology of isolated CPO. Despite this, Mendelian inheritance of CPO is rare and specific causative loci have been difficult to identify. *TBX22* (10) and *IRF6* (11) have recently been identified as genes mutated in syndromes which include CPO as a major component of the phenotype (OMIM 303400 and 119300 respectively).

We previously reported two children with *de novo* apparently balanced chromosomal rearrangements involving distal 2q32, clinically presenting with CPO and strikingly similar minor craniofacial and digital dysmorphisms (12). FISH mapping using YAC clones revealed that both 2q breakpoints were located between markers D2S311 and D2S116 (~5 Mb apart).

*To whom correspondence should be addressed at: MRC Human Genetics Unit, Western General Hospital, Edinburgh, EH4 2XU, UK. Tel: +44 1313322471; Fax: +44 1313432620; Email: david.fitzpatrick@hgu.mrc.ac.uk

Further evidence that this region of 2q contained a gene critical for normal palatogenesis was also provided by a statistical analysis showing that 2q32–q33 was one of only three regions of the genome for which haploinsufficiency is significantly associated with CPO (4,12). Here we report high resolution mapping of the 2q breakpoints, which has resulted in the demonstration of disruption of the gene encoding the DNA-binding protein SATB2 (formerly KIAA1034). Together with the expression pattern of *SATB2* in embryos during the period of palatogenesis, the results indicate a critical role for the *SATB2* gene in palate formation.

RESULTS

Identification of breakpoint-spanning YAC, PAC and BAC clones

Our previous work (12) positioned the 2q32–q33 breakpoint of both translocations in an interval defined by D2S311 proximally and D2S116 distally. This earlier FISH work had utilized mostly YACs from the ICI library, chosen for their low rates of chimerism. We next used the marker D2S348, which maps between D2S311 and D2S116, to isolate ICI YACs 32EB9 and 19ID10. Both these clones hybridized centromeric to both 2q breakpoints. However, the larger CEPH YAC 854H8, which also contains D2S348, and the overlapping CEPH YAC 920H2 were found to span both breakpoints (not shown). These results localized the two breakpoints to a region of overlap between YACs 854H8 and 920H2, which includes D2S2392, D2S2684 (WI-2923), D2S1740 (WI-5648), D2S2396, D2S2708 (WI-6481), and D2S1837 (WI-5610). These markers were used to isolate further ICI YACs for FISH. In addition, PAC clones were isolated from a flow-sorted chromosome 2 library (LL02NP04) by hybridization to YAC 854H8, and from the RPCI-1 library by screening for individual STS markers.

End sequences from clones from these screens were used for second and third round screening of the RPCI1 PAC library. Combining this information established a PAC contig 50G8–316M7–61E16–104J10–202E4–30G1 from centromere to telomere. FISH analysis showed that whereas 50G8 mapped centromeric to both breakpoints, the last four of these overlapping PACs all mapped between the 2;11 and 2;7 breakpoints (data not shown). However, YAC 31AG11, which has a size of ~500 kb on PFGE, was found by FISH to span both the 2;11 and 2;7 breakpoints, indicating that the distance between them on chromosome 2 is likely to be substantially less than 500 kb, barring a large interstitial deletion within this YAC. Taken together, these results indicated that the small PAC contig spanned most or all of the region between the breakpoints.

For transcript identification, PACs 104J10 and 202E4 were subcloned and a total of ~105 kb of first-pass sequence was generated. Database searching revealed the presence of a number of exons corresponding to part of the *SATB2* transcript (see below), but no other identifiable genes. As draft human genome sequence became available for this region during the course of our study, we were further able to locate these breakpoint interval sequences onto an overlapping contig of four RPCI11 BAC clones (Table 1). Further FISH analysis using these BACs then showed that 530J6 spanned the t(2;7) and 505H14 the t(2;11) breakpoint (Figs 1A and 2A). Analysis of the complete sequence of this contig (NT_005229) reveals only one confirmed gene (*SATB2*), which has telomeric–centromeric orientation (Fig. 2C). The combined FISH and BAC contig data indicate that the t(2;11) and t(2;7) breakpoints lie ~300 kb apart.

Table 1. RPCI11 BAC clones spanning the K1AA1034 region

Library	Grid	Acc No.	t(2;7)	t(2;11)	*SATB2* Fragment
RPCI-11	105 K 1	AC079609	Tel	Tel	Exon 1A-3
RPCI-11	530 J 6	AC017096	BPS	Tel	Exon 1A-6
RPCI-11	404 F 23	AC027171	Cen	Tel	Exon 1B-7
RPCI-11	486 F 17	AC016746	Cen	Tel	Exon 4-11
RPCI-11	505 H 14	AC018717	Cen	BPS	3′ to polyadenylation signal

Tel = telomeric to 2q33 breakpoint; Cen = centromeric to 2q33 breakpoint; BPS = breakpoint spanning.

SATB2 genomic structure

SATB2 was originally identified as KIAA1034, a near full-length cDNA (AB028957), isolated from a human fetal brain library as part of the HUGE project (13). Comparison with the genomic sequence shows that this transcript is assembled from 11 exons spanning 191 kb (Fig. 1B). The *SATB2* open reading frame starts in exon 2, with the first stop codon in exon 11. (Comparison with the sequence of the related transcription factor SATB1 suggests the use of this ATG, even though the *SATB2* ORF remains open to the 5′ end of exon 1A.) A variant cDNA, identified by database EST177825, also appeared to represent a product of the *SATB2* gene. We completely sequenced this clone (AJ438951, 2.2 kb). It contains a novel non-coding 5′ exon (which we designate 1B), *SATB2* exons 2 and 3, and two additional novel exons that lie upstream of exon 4 (Fig. 1B). The last of these (1175 bp) terminates at a *bona fide* polyadenylation site, and an AATAAA sequence lies in the usual position upstream of this point. The RNA splice variant represented by this clone is probably not functional, since its open reading frame is truncated; only the first quarter of the SATB1 conserved domain (see below) would be produced, followed by 10 novel amino acids encoded in the new exon. Despite this, exon 1B may well be a functional part of the gene, since it appears from available EST sequences to be more commonly used than exon 1A. Also, a similar 5′ exon was found in the mouse *SATB2* cDNA (see below). There is no alternative upstream in-frame ATG in exon 1B (and also no other long open reading frame in the EST177825 cDNA).

A region of ~7 kb with the characteristics of a CpG island surrounds and extends upstream of both exon 1A and exon 1B. Evidence that might suggest a third transcription start site in intron 5 comes from ESTs represented by AW351662 and AW375361. There is no CpG island upstream of this region, however.

Fine mapping of the translocation breakpoints

The complete genomic sequence of the region allowed a series of 10 kb FISH probes to be generated by PCR amplification of

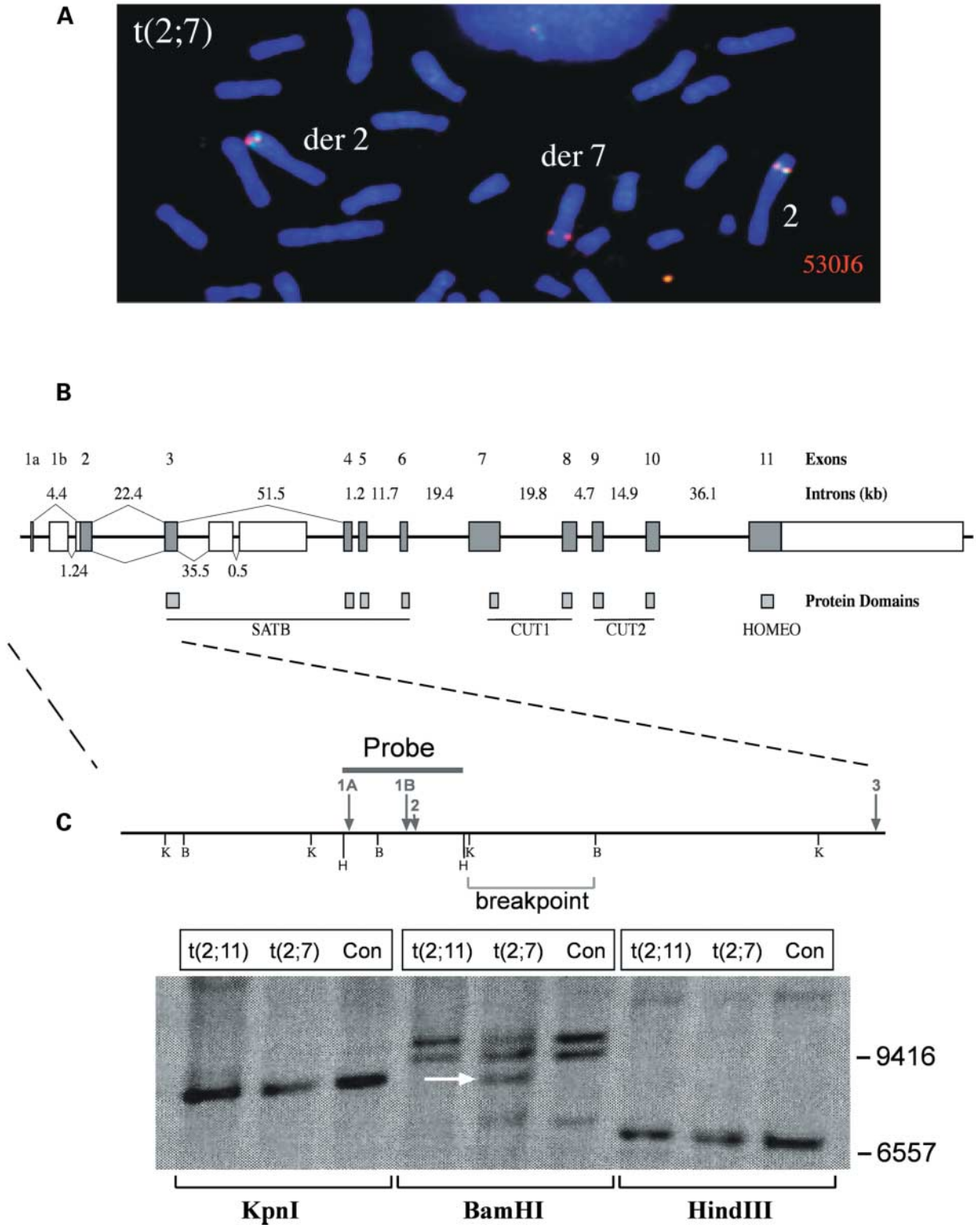

Figure 1. Identifying the t(2;7) breakpoint region. (**A**) FISH analysis showing BAC RPCI-11 530J6 (red signal), spanning t(2;7) breakpoint. (**B**) *SATB2* gene structure. Intron sizes are given in kb. The exon sizes (but not introns) are to scale with each other. Coding regions are shaded, non-coding open. Below the exon–intron diagram, the four conserved sequence motifs are positioned relative to the exon structure. For the first four exons, the splice pattern seen in the *SATB2* mRNA (AB028957) is shown above the line. The splice pattern in the variant cDNA (AJ438951) is shown below the line. (**C**) Southern blot analysis of the breakpoint within intron 2: t(2;11) = BO0145, t(2;7) = BB0213, Con = Control. Ten micrograms of each DNA was digested with *Kpn*I, *Bam*HI and *Hind*III and probed with the 6.8 kb *Hind*III fragment of RP11-530J6. The relative positions of the probe all the restriction sites is indicated on the cartoon above the blot. The probe encompasses exons 1A, 1B and 2 of *SATB2*. *Bam*HI generates two bands of 10.2 and 11.3 kb in all three lanes. In addtion t(2;7) shows a presumed junction fragment of ~9.2 kb on *Bam*HI digestion (arrowed) with an apparent reduction in intensity of the 11.3 kb band. No variant pattern is observed with *Kpn*I or *Hind*III.

Table 2. Primers and regions corresponding to PCR-amplified 10 kb FISH probes

Probe	Primers	BAC used for PCR	Position in SATB2 gene	Position in NT_022148 contig	FISH result
530 A	TCTACACCCACTAGAGCCTTGGCATTCCTG CCCTCACACAAGAAGGAAAGTCGTTGGAAA	RP11-530J6	Exon 3	681028–691027	Proximal to BP in t(2;7)
530 B	TGGGTTGTAAAATATCTGAGTGGGCCTTGG AGGGAGCCAACTAGGATCTCATGACAGGAA	RP11-530J6	Exons 4, 5, 6	741102–754605	Proximal to BP in t(2;7)
486 A	CATTCTTCATTGCCATCTTCATGCCCATTT ATCCTTTCAGCATTCCTCCATGCACACATT	RP11-486F17	Exon 11	843901–854010	Distal to BP in t(2;11)
505 B	TCCTTCCTTCCCTTCCTCCCTCTTTCTCAG GAAGCATGGTGGAAGAATGGAAACAGCATC	RP11-505H14	3′ flanking	858714–869314	Distal to BP in t(2;11)
505 H	TGTTCCCTAATGAATGCCTTGCCATTTTCA ACAGGAGGGGTTGTAGCTCATTGCACAGAC	RP11-505H14	3′ flanking	961475–972119	Distal to BP in t(2;11)
505 C	GCCAGTCATATGATCAAAAGGGGGTTTTTCATTT GTAAGAGGAGCTTGGCCTAAGCAGAGGATAGTGG	RP11-505H14	3′ flanking	977108–986953	Crosses BP in t(2;11)
505 E	AATGAACACAGACCCTGGGGTCCCAAATAC AGAAATCATGGTGGACAGGCAGAAATCCAA	RP11-505H14	3′ flanking	988141–997151	Proximal to BP in t(2;11)
505 G	TCCCACCTGCAAACTGGAATTGAAGAAAGA TTATCAGTGGTGGGATGGGCATAGAAGTGG	RP11-505H14	3′ flanking	1025869–1026331	Proximal to BP in t(2;11)

BAC DNA (Table 2). These enabled high-resolution mapping of both 2q32 breakpoints relative to the *SATB2* gene. The breakpoint in the t(2;7) case was shown in this way to disrupt *SATB2* between exons 2 and 3 (data not shown). Southern blot analysis (Fig. 1C) localized the breakpoint within intron 2 to between 1474 and 7112 bp from the end of exon 2. In the t(2;11) case, the breakpoint is located 123.5–133.5 kb 3′ of the polyadenylation signal (ttgttaaAATAAAcgcattc) (Fig. 2C). Since this breakpoint lies outside the *SATB2* transcription unit, we considered whether other genes might be nearby, or alternatively whether regions of conserved non-coding sequence exist that could represent long-range control elements disrupted by the translocation. 3′ to *SATB2*, the closest annotated gene is *PLCL1*, which lies ∼1.3 Mb from *SATB2*. Searches of the EST database using the 200 kb of genomic sequence 3′ to *SATB2* revealed no hits. To assess of the level of sequence conservation in the t(2;11) breakpoint region, the orthologous mouse gene was identified in the draft genome sequence, on chromosome 1 ∼57.5 Mb from ptel. In addition to the high level of conservation seen within the *SATB2* gene itself, there is a region 159 kb downstream of the gene that shows a very high level of sequence conservation (Fig. 2B and C). This region would be separated from *SATB2* by the t(2;11) translocation.

SATB2 protein structure and conservation

The long ORF of the *SATB2* transcript predicts a 733 amino-acid protein with a molecular weight of 82.5 kDa (SWISSPROT Y34_Human). This protein has no previously identified function, but shows considerable sequence similarity to the putative matrix attachment-mediating DNA-binding protein SATB1 (14). The similarity between these two proteins lies in three types of motif (Fig. 3A). Homology searches indicate the presence of two CUT domains (352–437 and 482–560) and a homeodomain (614–677). Each of the CUT domains is encoded by the adjacent portions of two exons, 7–8 or 9–10, while the homeo domain is encoded within the last exon (11). In addition to these motifs, a large region, encoded by exons 3–6 in their entirety, from *SATB2* residues 57–231, is highly conserved between SATB2 and SATB1 [142/163 (81%) identical residues]. It corresponds to the Pfam-B_10016 domain. It is present in all the SATB2 and SATB1 orthologous proteins, as well as in the *Drosophila* gene product DVE (defective proventriculus) and a predicted *C. elegans* protein, zk1193.5 (Q23413). These latter two proteins do not have CUT domains, but do have one or two homeodomains. This conserved domain is known to be required for dimerization of the SATB1 protein (15), and may therefore have a similar role in SATB2. Pfam-B_10016 is removed by caspase 6-mediated cleavage of SATB1 during T-cell apoptosis (15), but the caspase cleavage site is not conserved in SATB2 (Fig. 3A).

We assembled a full-length cDNA sequence for the murine *SATB2* homologue using RT–PCR of mouse brain RNA. This sequence (AF319623) revealed an extraordinarily high degree of sequence conservation, with only three predicted amino-acid substitutions in the 733 residue protein (I481, A590 and I730 being changed to V481, T590 and T730 respectively in the mouse protein). We have also identified orthologous proteins in *Fugu*, zebrafish and chicken, which show 65, 68 and 95% identity, respectively (the chicken protein is partial and covers only 58% of the full-length sequence). The homology is highest in the homeodomain, the 2 CUT domains and the Pfam-B_10016 domain. Because SATB1 and SATB2 proteins are so similar, we ensured that we had the correct orthologue by identifiying both proteins in *Fugu* and zebrafish and using these sequences to generate a multiple alignment and a phylogenetic tree (Fig. 3B).

Developmental expression pattern

Whole mount *in situ* hybridization of antisense riboprobes to staged mouse embryos revealed a site- and stage-specific expression pattern of *SATB2*. The 5′ and 3′ *SATB2* riboprobes gave identical results. No expression could be detected before 10.5 GD, when staining could be seen in the maxillary component of the first pharyngeal arch and the lateral aspect of the frontonasal process in the regions that will subsequently to fuse to form the primary palate (Fig. 4A and B). At 11–11.5 GD

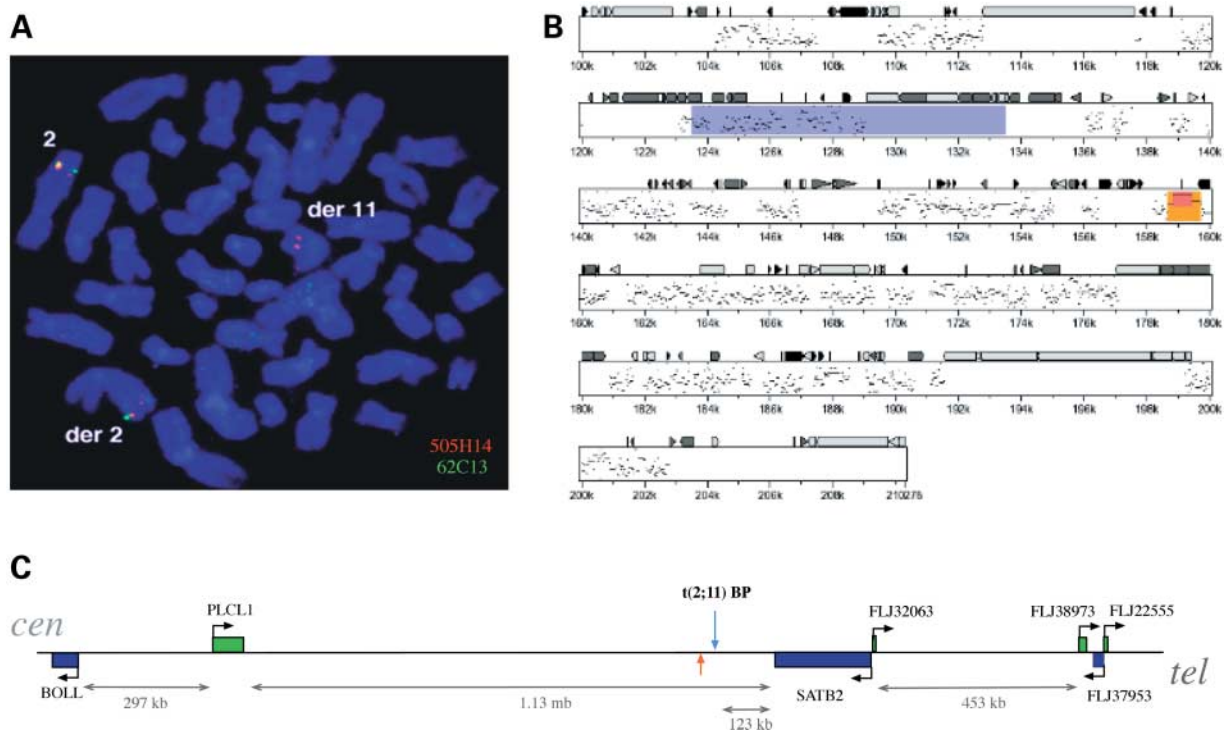

Figure 2. Identifying the t(2;11) breakpoint region. (**A**) FISH analysis showing BAC RPCI-11 505H14, spanning t(2;11) breakpoint (red signal). (**B**) PIP plot of the conserved region 100–210 kb downstream of SATB2. The baseline is the human sequence. Position '0' is at the end of SATB2. The dots and bars are hits with the mouse sequence. The baseline is 50% identity and the top line is 100% identity. The Blue box represents the t(2;11) case breakpoint region. The orange box represents a segment of 88% identity over more than 1043 bp. The red box is a segment of 94% identity over 601 bp. (**C**) Summary map of the SATB2 genomic region. The map is drawn to scale with the endpoints of transcription units of confirmed genes, represented by coloured boxes, positioned accurately according to genome sequence data (June 2002 freeze). The gene names are indicated above the boxes and the direction of transcription is indicated by a single arrow. The intergenic distances are represented by the grey double-headed arrows. The size in kilobases (kb) or megabases (Mb) is given below these arrows. The position of the t(2;11) breakpoint is indicated by the blue arrow. The orange arrow represents the position of the highly conserved sequence shown in orange in the PIP plot.

the expression pattern demarcates the region of the medial aspect of the maxillary process within the primitive oral cavity, which will form the palatal shelf (Fig. 4A and B). By 12.5 GD, symmetrical expression is seen in the medial edges of the developing palatal shelves and this continues until 13.5 GD when the strongest expression is in the mesenchyme underlying the medial edge epithelia (Fig. 4A and B). By the time of palatal shelf fusion at 14.5 GD the expression is dramatically down-regulated (data not shown). There was no expression detected elsewhere in the embryo at any stage examined.

Mutation analysis

Exons 3–11 of *SATB2* were PCR amplified from a panel of 70 genomic DNA samples from children with isolated CPO and subjected to heteroduplex based mutation analysis. No point mutations in the coding sequence were identified. Heterozygosity for a g/c SNP at position IVS4-35 was identified in 28/72 (39%) of cleft palate samples and 36/93 (39%) control samples.

DISCUSSION

Development of the secondary palate in mammals has been the subject of many excellent descriptive, biochemical and molecular studies over the last 25 years. This wealth of information makes cleft palate an attractive model for investigation of complex human genetic disease. Association studies using candidate genes have been the main approach used in human CPO studies. Although some interesting associations exist, no definitive causative mutations have been identified. The use of cytogenetic pointers to aid in the identification of genes required for particular developmental processes is another, powerful method, which we used in a systematic way for the present study. This approach may involve three phases of investigation. The first involves the identification of chromosomal regions that show statistically significant association with a specific malformation, using human deletion and duplication maps (4,16). The second relies on the availability of apparently balanced translocation breakpoints mapping to the region identified in the previous phase, and associated with the same phenotype. Breakpoint mapping is then used to identify the disruption of one or more genes hypothesized to be involved in that specific developmental pathway. Finally, the pathogenic role of this gene may be reinforced by the identification of mutations in affected cases with normal karyotypes.

Deletion of 2q31–q33 was one of only three segmental chromosomal aneuploidies identified as significantly associated with CPO using a catalogue of published clinical and

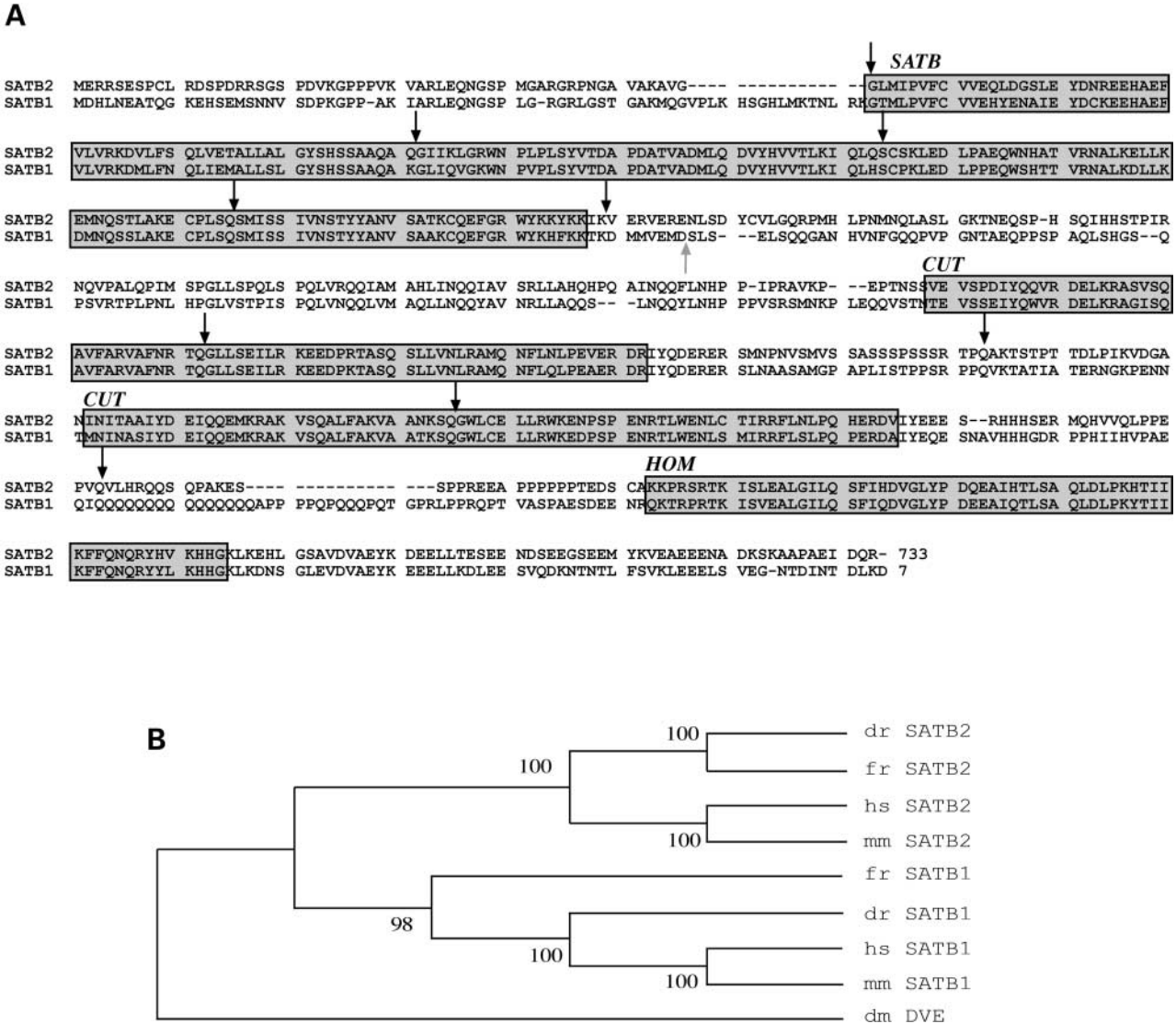

Figure 3. SATB2; protein domains and conservation. (**A**) Alignment of predicted SATB2 (above) and SATB1 peptide sequences, generated using ClustalW. Four domains have been highlighted: a large conserved N-terminal 'SATB' region that may mediate dimerization, the two CUT repeats, and the atypical homeodomain. The positions of *SATB2* splice junctions are indicated by down-pointing arrows. The caspase 6 cleavage site of SATB1 is indicated by an upward arrowhead. (**B**) Phylogenetic tree made from the SATB2 and SATB1 protein alignment. Sequences used are: hs_satb2 (Y34_HUMAN, Q9UPW6), mm_satb2 (KIAA1034-like DNA binding protein, NP_631885), hs satb1 (SAT1_HUMAN, Q01826), mm satb1 (AT-rich sequence binding protein, NP_033148), dr_SATB2, fr_SATB2, fr satb1 and dr satb1 are predictions from genome assemblies. The tree is rooted with *Drosophila* DVE protein (NP_726183). Species abbreviations are as follow: hs = *Homo sapiens*; mm = *Mus musculus*; dr = *Danio rerio*; fr = *Fugu rubripes*; and dm = *Drosophila melanogaster*.

cytogenetic data (4,12). We were then able to ascertain two apparently balanced translocations involving 2q33 for further study. These cases were of particular interest because they were *de novo* mutations and both children had very similar clinical features associated with CPO (12). Characterization of the breakpoints has now shown disruption of the coding region of a gene, *SATB2*, between exons 2 and 3 in one case. The function of this gene is previously undefined. However, it shows considerable similarity to the nuclear matrix attachment region (MAR)-binding protein SATB1 (14) and also significant homology in its N-terminal region to the *Drosophila* protein DVE.

On structural grounds, SATB2 belongs to the CUT superclass of homeodomain proteins, characterized by a conserved 70–75 residue DNA-binding CUT domain in addition to a homeodomain. Within this superclass, three classes, CUX, SATB and ONECUT, have been defined, containing respectively three, two or one CUT domains. The SATB class, of which SATB1 and SATB2 are the only recognized members, is the most divergent from the other families in the sequence of both its CUT and homeo domains, the latter being particularly divergent in its first third (17). SATB1 was identified by expression screening for proteins binding to sequences present in nuclear MARs (14). It shows specificity for AT-rich sequence motifs devoid of G residues on one strand (ATC regions). Such regions have a high propensity to unwind under superhelical stress, a property dependent on a core unwinding

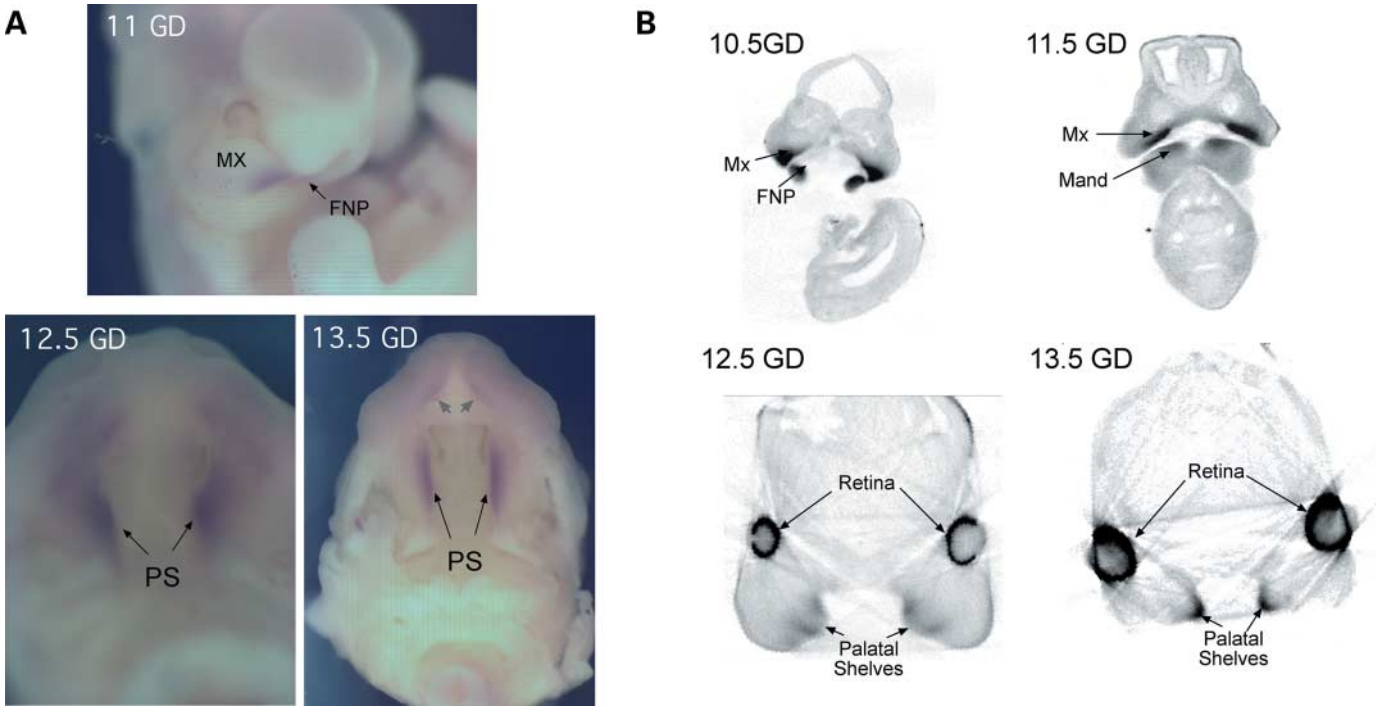

Figure 4. Developmental expression of *SATB2* in mouse embryos. (**A**) Embryo whole-mount *in situ* hybridization with antisense *SATB2* riboprobe. The top picture shows a ventrolateral view of 11 GD embryo showing strong expression in medial and caudal aspects of the maxillary process. Some non-consistent staining is seen in the dorsal head region of this embryo. The lower images are of embryos at 12.5 and 13.5 GD, respectively, that have been dissected to remove the lower jaw and tongue and viewed looking up to the roof of the primitive oral cavity. Symmetrical expression is seen in the medial edges of the developing palatal shelves (PS) at both stages. Some expression is seen in the developing tooth buds at 13.5 GD (grey arrowheads). No consistent expression is detected elsewhere in these embryos. (**B**) Virtual sectioning of optical projection tomography reconstructions of embryos following whole mount *in situ* hybridization with antisense *SATB2* riboprobes. The upper panel shows coronal sections from whole embryos. The lower panel shows dissected embryo upper heads at the stages indicated. Expression is shown by dark staining on these grayscale images, however pigment from the retina results in an interference pattern on OPT most prominently seen at 13.5 GD. The 10.5 GD image shows staining in the maxillary component of the first pharyngeal arch (MX) and the lateral aspect of the frontonasal process (FNP) in the regions that will subsequently to fuse to form the primary palate. The 11.5 GD image shows expression which demarcates the region of the maxillary process (MX) destined to form the palatal shelf. The 12.5 GD image reveals symmetrical expression is seen in the medial edges of the developing palatal shelves and this continues until 13.5 GD.

element that SATB1 binds specifically (18). SATB1 is involved in regulating several T cell-specific genes. During T cell maturation, for example, induction of DNase hypersensitivity upstream of the TCRβ enhancer is accompanied by binding of SATB1 and CDP (19). Another MAR element mediates transcriptional regulation of the CD8a gene in T cells, and here displacement of CDP by SATB1 favours CD8a transcription (20). DNA regions to which SATB1 binds are located at the bases of chromatin loops and tightly matrix-associated (21). Synthetic SATB1-binding MAR elements insulate transgenes from position effects and exert tissue-specific transcriptional repression on them (22), consistent with the idea that MARs can demarcate autonomous regions of gene regulation. Heterologously expressed SATB1 can suppress transcription from reporter genes with *cis*-linked MAR elements (23). Knockout models indicate that these properties reflect a role for SATB1 as a transcriptional regulator of multiple T-cell specific loci during thymocyte development (24). This locus control activity may be mediated through the ability of the MAR-binding SATB1 to act as a focus for recruitment of chromatin remodelling complexes (25).

The structural similarity between SATB1 and SATB2 suggests that the latter may perform analogous roles in transcriptional regulation, although in different cell types. As expected, both these proteins are highly conserved; murine and human SATB1 differ at only 14 positions (98.2% identity), while SATB2 is even more highly conserved (99.6% identity).

The *Drosophila* protein to which SATB2 shows similarity (DVE) was identified as the product of the gene mutated in two allelic enhancer-trap mutants *defective proventriculus* (*DVE*) (26). The proventriculus is a valve structure, which regulates the passage of food from foregut to midgut in *Drosophila* larvae and could, therefore, be fancifully described as a palate-like structure. Using a series of double mutants, it was shown that *DVE* is regulated by both *wingless* (*wg*) and *decapentaplegic* (*dpp*, a member of the TGFβ superfamily). This is particularly interesting, since TGFβ3 is known to play a critical role in development of the secondary palate in mammals (27). Taken together, these data suggest that the role of SATB2 in palatogenesis may involve mediating some of the TGFβ-induced changes in medial edge epithelium that enable midline fusion to occur.

The most convincing support for SATB2's role in palatogenesis comes from the site- and stage-specific expression pattern we have demonstrated in mouse embryos. Mammalian secondary palate development can be divided into four stages: (1) initiation of the primordial palatal shelves from the medial

aspect of the maxillary processes within the primitive oral cavity; (2) rapid vertical down growth of the palatal shelves lateral to the tongue; (3) reorientation of the shelves to come into apposition above the tongue, and (4) fusion of the medial edge epithelia to form the midline epithelial seam, which subsequently disrupts, thereby establishing continuity in the palate. *SATB2* expression in the 11.5 GD embryo would be consistent with a role in shelf initiation. Later expression is restricted to the mesenchyme underlying the medial edge epithelium of the palatal shelves and decays rapidly following shelf contact. This would suggest a role in preparation of the medial edge epithelium or sub-epithelial mesenchyme for either shelf elevation or adhesion. Interestingly there is also evidence from the 10.5–11.5 GD embryos examined for a role of *SATB2* in primary palate development. The primary and secondary palate are the sites of the two, true fusion events in facial development and this may suggest that there is a common role for SATB2 in this process. It is interesting that no convincing expression of *SATB2* was detected out-with the developing facial processes. It is obviously important to confirm whether the same expression patterns observed here are seen in human embryos, however, SAGE data from the COGENE project (hg.wustl.edu/COGENE/) suggests that *SATB2* is present in the embryonic face at a low level. One-third of *SATB2* cDNAs in the human EST databases are derived from adult brain libraries with the remainder from a wide variety of libraries from normal tissue and cancerous lesions. We have not examined *SATB2* expression in the adult brain or during fetal brain development but this may reveal interesting site- and stage-specific expression. This is particularly important due to the mild cognitive problems that were apparent in the two translocation cases (12).

No *SATB2* pathogenic mutations were identified in a cohort of 70 unrelated isolated CPO cases and there was no evidence of association detected using an intragenic intronic SNP typed in cases and controls. The mutation analysis we performed could not exclude intragenic deletions incorporating whole exons or primer sites. It is striking that we also failed to identify a single exonic polymorphism during this mutation screening. This, and the extraordinary degree of mouse–human conservation, suggests that the *SATB2* gene is under extreme evolutionary pressure. Whether this reflects biological roles distinct from palatogenesis awaits future knockout experiments. The lack of exonic polymorphisms prevented us from confirming monoallelic expression using RT–PCR in either of the cell lines. This would have been particularly valuable in the case of the t(2;11) cell line, in which the translocation breakpoint was found to lie 130 kb 3′ of the *SATB2* polyadenylation signal. In this region, there is no evidence for transcribed genes. (Indeed, the closest neighbour is >1 Mb 3′ to *SATB2*.) However, a large (~50 kb) region showing particularly high mouse–human sequence conservation lies immediately downstream of the t(2;11) breakpoint. We therefore hypothesize that the t(2;11) translocation may separate *SATB2* from a necessary 3′ locus control element, thereby resulting in functional haploinsufficiency equivalent to that resulting from the t(2;7), which bisects the gene itself. Similarly conserved 3′ sequences have been shown to control the spatio-temporal expression pattern of *Pax6* in transgenic mice (28). Correspondingly, in humans, translocations involving this region downstream of *PAX6* are known to cause the developmental eye defect aniridia that is phenotypically identical to that resulting from deletion or truncating mutation of *PAX6* itself (29). An alternative explanation for a haploinsufficiency effect in the t(2;11) case would be the result of an epigenetic modification of the regions surrounding the chromosomal rearrangement. Such an effect has been recently demonstrated in a inversion of chromosome 18 associated with Tourette syndrome (30).

In spite of the lack of germline point mutations in sporadic CPO patients, *SATB2* is a very good candidate gene for involvement in the process of palatal fusion that is disturbed in CPO. Disruption of *SATB2* function appears to be the cause of the CPO displayed by both of the children with 2q32–q33 translocations featured in the present study. Future studies involving mutation analysis of larger numbers of CPO individuals, including the highly conserved non-coding region identified 3′ to the t(2;11) breakpoint and targetted inactivation of the highly conserved mouse gene promise to define the nature of this developmental role in more detail.

MATERIALS AND METHODS

YAC, PAC and BAC clones: screening and analysis

YAC clones were isolated from the ICI (31) and CEPH (32) YAC libraries. PAC clones were isolated from the LLNL human chromosome 2-specific library LL02NP04, obtained from the UK HGMP-RC as follows: 100 ng ^{32}P-labelled YAC DNA was preannealed to 1 μg unlabelled Cot1 DNA (Gibco BRL) and hybridized (in the presence of 200 μg/ml denatured total human genomic DNA) to a filter carrying 5760 PAC clones (1.9 genome equivalents at an average insert size of ~85 kb). The probe used was the CEPH megaYAC 854H8 (insert size ~1760 kb), that had been shown to generate FISH signals spanning both translocation breakpoints. Additional PACs were isolated from the RPCI-1 library (obtained from the UK HGMP-RC) by PCR screening for individual STSs. Clones from the RPCI11 human BAC library were obtained from the BACPAC Resources (Oakland, CA, USA). YAC, PAC and BAC DNA was prepared by a standard mini-prep method.

Sequence sampling of the PACs was performed using dye terminator chemistry on an ABI377 sequencer. PACs were subcloned into plasmid vectors after digesting with various combinations of enzymes: dJ104J10; *Bgl*II+*Bam*HI, *Hin*dIII, *Eco*RI, *Xba*I+*Sty*I, *Sau*3AI (partial); dJ202E4; *Xba*I+*Sty*I, *Hin*dIII, *Sau*3AI (partial).

RT–PCR analysis

Full-length murine *SATB2* cDNA was isolated and sequenced after RT–PCR and 5′-RACE on mouse brain RNA, using previously described methods (33).

Cell culture and FISH analysis

Slides for FISH analysis of the two translocation breakpoints were made from fixed cell suspensions by standard methods,

Table 3. Primers used in mutation analysis

Exon	Forward	Reverse
3	CTGTTGCTTCCCTTCTCATC	TATTTCAGAAAACATATTCTTCC
4	GAGTGGGCCTTGGATGTTG	TGTCCAGAAATTAATAGACAGG
5	ATTCTTTAGCATTTCTTCTGAAG	TGATAGAGAGTCTACCAATGG
6	AGTTCAGTTTACAGGAGAGTG	CGTAAATTGTATCTAAAATGAGC
7	AAAGTATGATAATTCAGTTCTAC	TAACCTATCTCAATTCTGTCC
8	GGAAGACAGAAACCACACAC	GGGAAAACACAGTAACTAGTG
9	GATAATTTCTGCACAGATGGC	GTTTCTTGGGGGTTATTACTG
10	ATGTACTGTGATGGCACTGG	GTGGTGTGTGCCACTTGG
11	CTTACCTCCAGTGCTAGATAG	TAAAGAAATGAAAGCAGAAAATCC

from lymphoblastoid cell lines; B00213:t(2;7)(q32;p21) and BB00145:t(2;11)(q32;p14). FISH using ICI YAC clones was performed as previously described (12). For PACs and BACs, DNA was labelled with digoxigenin-11-dUTP or biotin-16-dUTP (Roche, Indianapolis, IN, USA) by nick translation. Some clones were labelled using a modified DOP–PCR protocol (34). DNA hybridization and antibody detection were carried out as described previously (35). At least five metaphases were analysed for each hybridization, using a Zeiss Axioplan 2 fluorescence microscope equipped with a triple band-pass filter (no. 83000 for DAPI, FITC and Texas Red; Chroma Technology, Brattleboro, VT, USA). Images were collected using a cooled CCD camera (Princeton Instruments Pentamax, Roper Scientific, Trenton, NJ, USA) and analysed using IPLab software (Scanalytics, Vienna, VA, USA).

Fine mapping of translocation breakpoint

Primers were designed using the Primer3 program (www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) from the genomic sequences AC016746 (RP11-486F17), AC018717 (RP11-505H14) and AC017096 (RP11-530J6) to generate 10 kb probes for *SATB2* and its 3′ flanking sequence (see Table 1 for details). PCR reactions were performed using the Expand Long Template PCR kit (Roche) using the manufacturer's conditions. Products were purified using a Qiaquick PCR purification kit (Qiagen) and labelled by nick translation with digoxigenin-11-dUTP. For the region including exons 1 and 2, long-range PCR was unsuccessful. A 6.8 kb *Hind*III fragment containing exons 1A, 1B and 2 was isolated from BAC RP11-530J6 and used for FISH and Southern blot analysis. Genomic DNA was extracted from each lymphoblastoid cell lines and digested with 40 U of *Bam*HI, Asp718 (*KpnI*) and *Hind*III (Roche) according to manufacturers recommendations. A 10 µg aliquot of digested DNA was separated in a 0.7% agarose gel and Southern blotting was performed by standard methods. A 100 ng aliquot of the probe (6.8 kb *Hind*III fragment containing *SATB2* exons 1A, 1B and 2) was labelled by direct alkaline phosphatase conjugation and hybridized and detected according to the manufacturer's recommendations. The filter was visulalized using a STORM 840 fluorescent scanner (Molecular Dynamics).

Bioinformatic analyses

Protein and nucleotide sequence analysis was performed using the PIX and UWGCG suites of programs (accessed at www.hgmp.mrc.ac.uk). Large-scale genomic comparisons between mouse (NT_039170.1) and human (NT_005403.13) sequence was performed using PipMaker (http://bio.cse.psu.edu/pipmaker/) (36). Sequences were masked using RepeatMasker (http://ftp.genome.washington.edu/RM/RepeatMasker.html) (37). *Fugu* and zebrafish protein sequences are predictions based on the assemblies available via the Ensembl website (www.ensembl.org/Fugu_rubripes/ and http://pre.ensembl.org/Danio_rerio/). Protein predictions have been made using the program genewise (www.ebi.ac.uk/Wise2/) (38). Both proteins were subsequently manually corrected with the help of individual shotgun sequence reads. Multiple alignments were done using CLUSTALW (version 1.82 with default settings) (39), and visualized with Genedoc (www.psc.edu/biomed/genedoc). Positions in alignments containing gaps were omitted from subsequent analyses. All phylogenetic trees were constructed by the neighbour-joining method (40). Phylogenetic trees were constructed using MEGA (version 2.1; www.megasoftware.net/) (41) with a 1000 replications bootstrap.

Whole-mount *in situ* hybridization and OPT imaging

Two sets of primers were designed to amplify murine *SATB2* cDNA. The two PCR products (5′ and 3′) contained non-overlapping 498 bp and 499 bp regions of the *SATB2* coding region. The primers incorporated T3 and T7 promotor sequences (underlined) on the sense and antisense strands, respectively: 5′S T3, AATTAACCCTCACTAAAGGCAGGCCCAAGGAATAATCAA; 5′AS T7, TAATACGACTCACTATAGGTGTGGTGGATTTGGCTATGA; 3′S T3, AATTAACCCTCACTAAAGGTGGTCTCCTCTGCCTCTAGC; 3′AS T7, TAATACGACTCACTATAGGTTTGGCACAGCTGTCTTCTG.

Sense and antisense riboprobes were generated using fluorescein RNA labelling mix (Roche), and whole mount *in situ* hybridization was performed according to previously described techniques (42). Transcripts were detected using an alkaline phosphatase anti-fluorescein antibody (Vector Laboratories Inc., Burlingame, CA, USA) and visualized with NBT/BCIP (Roche). For optical sectioning the embryos were clarified and scanned by optical projection tomography (OPT) as previously described (43). Analysis of the OPT reconstructions was performed using the MAPaint program (http://genex.hgu.mrc.ac.uk/Software/paint/paint_help/MAPaint_1.00/paint_help/paint_help.html).

Mutation analysis

DNA from 70 cleft palate patients was screened for mutations in *SATB2* coding region. These included 23 cases of Pierre Robin sequence but excluded any other known syndromes or chromosomal anomalies. PCR amplification was performed in 50 µl in Ampli*Taq* Gold buffer (Applied Biosystems) with 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM of each oligonucleotide, 1 U Ampli*Taq* Gold polymerase and 0.125 U *Pfu* polymerase (Stratagene). The oligonucleotides used map to the flanking intronic sequence or 3′UTR of each specified exon, were annealled at 53 or 55°C, and are listed in Table 3. Heteroduplex-based mutation analysis was performed by DHPLC using a Transgenomic WAVE System under conditions determined by Wavemaker 3.4.4 software (Transgenomic Inc.). Variant fragments were re-amplified, using template from an independent source, and sequenced using an ABI 377 to identify nucleotide changes.

ACKNOWLEDGEMENTS

The work was supported by an EU FP5 grant as part of the Eurocran project, and by grants from the M.R.C. and Wellcome Trust. L.M. was supported by a grant from the Birth Defects Foundation.

REFERENCES

1. Luke, D.A. (1976) Development of the secondary palate in man. *Acta Anat.*, **94**, 596–608.
2. FitzPatrick, D.R., Raine, P.A. and Boorman, J.G. (1994) Facial clefts in the west of Scotland in the period 1980–1984: epidemiology and genetic diagnoses. *J. Med. Genet.*, **31**, 126–129.
3. Dronamraju, K.R. (1971) Genetic studies of a cleft palate clinic population. *Birth Defects Orig. Artic. Ser.*, **7**, 54–57.
4. Brewer, C., Holloway, S., Zawalnyski, P., Schinzel, A. and FitzPatrick, D. (1998) A chromosomal deletion map of human malformations. *Am. J. Hum. Genet.*, **63**, 1153–1159.
5. Winter, R.M. and Barraitser, M. (2000) *London Dysmorphology Database*. Oxford University Press, Oxford.
6. Christensen, K. and Fogh-Andersen, P. (1996) Cleft-twin sets in Finland 1948–1987. *Cleft Palate Craniofac. J.*, **33**, 530.
7. Christensen, K. and Fogh-Andersen, P. (1993) Isolated cleft palate in Danish multiple births, 1970–1990. *Cleft Palate Craniofac. J.*, **30**, 469–474.
8. FitzPatrick, D. and Farrall, M. (1993) An estimation of the number of susceptibility loci for isolated cleft palate. *J. Craniofac. Genet. Dev. Biol.*, **13**, 230–235.
9. Christensen, K. and Mitchell, L.E. (1996) Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate—a Danish Registry study. *Am. J. Hum. Genet.*, **58**, 182–190.
10. Braybrook, C., Doudney, K., Marcano, A.C., Arnason, A., Bjornsson, A., Patton, M.A., Goodfellow, P.J., Moore, G.E. and Stanier, P. (2001) The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. *Nat. Genet.*, **29**, 179–183.
11. Kondo, S., Schutte, B.C., Richardson, R.J., Bjork, B.C., Knight, A.S., Watanabe, Y., Howard, E., de Lima, R.L., Daack-Hirsch, S., Sander, A. *et al.* (2002) Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat. Genet.*, **32**, 285–289.
12. Brewer, C.M., Leek, J.P., Green, A.J., Holloway, S., Bonthron, D.T., Markham, A.F. and FitzPatrick, D.R. (1999) A locus for isolated cleft palate, located on human chromosome 2q32. *Am. J. Hum. Genet.*, **65**, 387–396.
13. Kikuno, R., Nagase, T., Waki, M. and Ohara, O. (2002) HUGE: a database for human large proteins identified in the Kazusa cDNA sequencing project. *Nucl. Acids Res.*, **30**, 166–168.
14. Dickinson, L.A., Joh, T., Kohwi, Y. and Kohwi-Shigematsu, T. (1992) A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell*, **70**, 631–645.
15. Galande, S., Dickinson, L.A., Mian, I.S., Sikorska, M. and Kohwi-Shigematsu, T. (2001) SATB1 cleavage by caspase 6 disrupts PDZ domain-mediated dimerization, causing detachment from chromatin early in T-cell apoptosis. *Mol. Cell. Biol.*, **21**, 5591–5604.
16. Brewer, C., Holloway, S., Zawalnyski, P., Schinzel, A. and FitzPatrick, D. (1999) A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality—and tolerance of segmental aneuploidy—in humans. *Am. J. Hum. Genet.*, **64**, 1702–1708.
17. Lannoy, V.J., Burglin, T.R., Rousseau, G.G. and Lemaigre, F.P. (1998) Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding properties, contain a bifunctional homeodomain, and define the new ONECUT class of homeodomain proteins. *J. Biol. Chem.*, **273**, 13552–13562.
18. Dickinson, L.A., Dickinson, C.D. and Kohwi-Shigematsu, T. (1997) An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. *J. Biol. Chem.*, **272**, 11463–11470.
19. Chattopadhyay, S., Whitehurst, C.E. and Chen, J. (1998) A nuclear matrix attachment region upstream of the T cell receptor beta gene enhancer binds Cux/CDP and SATB1 and modulates enhancer-dependent reporter gene expression but not endogenous gene expression. *J. Biol. Chem.*, **273**, 29838–29846.
20. Banan, M., Rojas, I.C., Lee, W.H., King, H.L., Harriss, J.V., Kobayashi, R., Webb, C.F. and Gottlieb, P.D. (1997) Interaction of the nuclear matrix-associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene. *J. Biol. Chem.*, **272**, 18440–18452.
21. de Belle, I., Cai, S. and Kohwi-Shigematsu, T. (1998) The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB1) *in vivo* in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. *J. Cell. Biol.*, **141**, 335–348.
22. Neznanov, N., Kohwi-Shigematsu, T. and Oshima, R.G. (1996) Contrasting effects of the SATB1 core nuclear matrix attachment region and flanking sequences of the keratin 18 gene in transgenic mice. *Mol. Biol. Cell.*, **7**, 541–552.
23. Kohwi-Shigematsu, T., Maass, K. and Bode, J. (1997) A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes. *Biochemistry*, **36**, 12005–12010.
24. Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y. and Kohwi-Shigematsu, T. (2000) The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes. Dev.*, **14**, 521–535.
25. Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P. and Kohwi-Shigematsu, T. (2002) SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*, **419**, 641–645.
26. Nakagoshi, H., Hoshi, M., Nabeshima, Y. and Matsuzaki, F. (1998) A novel homeobox gene mediates the Dpp signal to establish functional specificity within target cells. *Genes. Dev.*, **12**, 2724–2734.
27. Proetzel, G., Pawlowski, S.A., Wiles, M.V., Yin, M., Boivin, G.P., Howles, P.N., Ding, J., Ferguson, M.W. and Doetschman, T. (1995) Transforming growth factor-beta 3 is required for secondary palate fusion. *Nat. Genet.*, **11**, 409–414.
28. Griffin, C., Kleinjan, D.A., Doe, B. and van Heyningen, V. (2002) New 3′ elements control Pax6 expression in the developing pretectum, neural retina and olfactory region. *Mech. Dev.*, **112**, 89–100.
29. Kleinjan, D.A., Seawright, A., Schedl, A., Quinlan, R.A., Danes, S. and van Heyningen, V. (2001) Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum. Mol. Genet.*, **10**, 2049–2059.
30. State, M.W., Greally, J.M., Cuker, A., Bowers, P.N., Henegariu, O., Morgan, T.M., Gunel, M., DiLuna, M., King, R.A., Nelson, C. *et al.* (2003) Epigenetic abnormalities associated with a chromosome 18(q21–q22) inversion and a Gilles de la Tourette syndrome phenotype. *Proc. Natl Acad. Sci. USA*, **100**, 4684–4689.
31. Anand, R., Riley, J.H., Butler, R., Smith, J.C. and Markham, A.F. (1990) A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. *Nucl. Acids Res.*, **18**, 1951–1956.
32. Dausset, J., Ougen, P., Abderrahim, H., Billault, A., Sambucy, J.L., Cohen, D. and Le Paslier, D. (1992) The CEPH YAC library. *Behring. Inst. Mitt.*, 13–20.

33. Hayward, B.E. and Bonthron, D.T. (2000) An imprinted antisense transcript at the human GNAS1 locus. *Hum. Mol. Genet.*, **9**, 835–841.
34. Huang, Q., Schantz, S.P., Rao, P.H., Mo, J., McCormick, S.A. and Chaganti, R.S. (2000) Improving degenerate oligonucleotide primed PCR-comparative genomic hybridization for analysis of DNA copy number changes in tumors. *Genes Chromosomes Cancer*, **28**, 395–403.
35. Chong, S.S., Pack, S.D., Roschke, A.V., Tanigami, A., Carrozzo, R., Smith, A.C., Dobyns, W.B. and Ledbetter, D.H. (1997) A revision of the lissencephaly and Miller-Dieker syndrome critical regions in chromosome 17p13.3. *Hum. Mol. Genet.*, **6**, 147–155.
36. Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. and Miller, W. (2000) PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res.*, **10**, 577–586.
37. Smit, A.F. (1996) The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.*, **6**, 743–748.
38. Birney, E. and Durbin, R. (2000) Using GeneWise in the Drosophila annotation experiment. *Genome Res.*, **10**, 547–548.
39. Higgins, D.G. and Sharp, P.M. (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, **73**, 237–244.
40. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
41. Kumar, S., Tamura, K., Jakobsen, I.B. and Nei, M. (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244–1245.
42. Henrique, D., Adam, J., Myat, A., Chitnis, A., Lewis, J. and Ish-Horowicz, D. (1995) Expression of a Delta homologue in prospective neurons in the chick. *Nature*, **375**, 787–790.
43. Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sorensen, J., Baldock, R. and Davidson, D. (2002) Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science*, **296**, 541–545.