Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22

Nandita A. Quaderi[1], Susann Schweiger[2], Karin Gaudenz[3], Brunella Franco[1], Elena I. Rugarli[1], Wolfgang Berger[2], George J. Feldman[3], Manuela Volta[1], Grazia Andolfi[1], S. Gilgenkrantz[4], Robert W. Marion[5], Raoul C.M. Hennekam[6], John M. Opitz[7], Maximilian Muenke[3], H. Hilger Ropers[2,8] & Andrea Ballabio[1]

Opitz syndrome (OS) is an inherited disorder characterized by midline defects including hypertelorism, hypospadias, lip-palate-laryngotracheal clefts and imperforate anus. We have identified a new gene on Xp22, *MID1* (Midline 1), which is disrupted in an OS patient carrying an X-chromosome inversion and is also mutated in several OS families. *MID1* encodes a member of the B-box family of proteins, which contain protein–protein interaction domains, including a RING finger, and are implicated in fundamental processes such as body axis patterning and control of cell proliferation. The association of *MID1* with OS suggests an important role for this gene in midline development.

Opitz G/BBB syndrome (OS; McKusick 145410 and 300000) is an inherited multiple-organ disorder primarily affecting midline structures. The disorder was first reported as two separate entities, BBB syndrome[1] and G syndrome[2]. Since then it has become apparent that the two syndromes are in fact a single entity, now named the Opitz G/BBB syndrome. OS is characterized by a constellation of symptoms, including hypertelorism; clefts of lip, palate and uvula; laryngotracheo-oesophageal abnormalities, leading to swallowing difficulty and hoarse cry; genito-urinary defects, such as hypospadias in males and splayed labia majora in females; imperforate anus; developmental delay; and congenital heart defects. The phenotype (Fig. 1) is more complex and more severe in male than in female patients[3–16]. OS is not only one of several X-linked syndromic mental retardation disorders but also one of the few mendelian inherited forms of cleft lip and palate.

A recent linkage study conducted on ten families segregating the OS phenotype revealed genetic heterogeneity for this disorder. In that study, the disorder was linked to *DXS987* in Xp22 in three families, with a lod score of 3.53 at zero recombination, while it was linked to *D22S345* from chromosome 22q11.2 in five families (ref. 17). Confirmation of the Xp22 localization came from a second study, in which linkage of the OS locus to *DXS7104* in Xp22 was reported[18]. The X-linked and autosomal forms of the disease cannot be differentiated on the basis of the clinical phenotype[4]. Further evidence for an X-linked form of OS was provided by the description of a large French family in which the disease appeared to co-segregate with a pericentric inversion of the X chromosome inv(X)(p22.3q26). In this three-generation family, four boys, their mothers and their maternal grandmother showed the typical manifestations of OS, and all of the symptomatic patients had the X-chromosome inversion[19]. The combination of linkage and inversion breakpoint data indicated Xp22.3 as the critical region for the OS gene.

As the anomalies observed in OS primarily affect midline structures, it has been proposed that genes important to human development, such as zinc-finger genes, would be good candidates for OS[17]. To identify the X-linked OS gene, we undertook a positional cloning approach to precisely localize the inversion breakpoint and identify transcribed sequences from the region[20,21]. Our efforts led to the identification of a new gene, *MID1* (Midline 1), which spans the OS inversion breakpoint. The MID1 protein belongs to a family of transcriptional regulators that contain protein–protein interaction domains and have been implicated in fundamental processes such as body axis patterning[22–25] and cell transformation[26,27]. The identification of mutations in affected individuals demonstrates that *MID1* is the gene for X-linked OS.

Fine localization of the Xp22 inversion breakpoint

An X-chromosome pericentric inversion, with breakpoints in Xp22 and Xq26, was found in affected individuals from a twenty-member OS family[19]. To localize the Xp22 inversion breakpoint, we performed fluorescence *in situ* hybridization (FISH) on chromosome spreads from individuals carrying the inverted X chromosome and a normal control, using as probes YAC clones known to map to Xp22 (ref. 20). The overlapping YACs 184F6 and 61B3 (ref. 20) gave two signals on the inverted X chromosome (on Xp22 and Xq26), indicating that they spanned the breakpoint (data not shown). Cosmid clones from the region covered by YAC 184F6 (ref. 28 and Cox *et al.*, submitted) were then hybridized by FISH to chromosome spreads from affected individuals carrying the inverted X chromosome, and subsequently to digested genomic DNA derived from affected and normal individuals. The results indicate that the Xp22 inversion

[1]Telethon Institute of Genetics and Medicine (TIGEM), San Raffaele Biomedical Science Park, 20132 Milan, Italy. [2]Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [3]The Children's Hospital of Philadelphia, Division of Human Genetics and Molecular Biology, Departments of Pediatrics and Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-4399, USA. [4]Nancy, France. [5]Department of Pediatrics, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York 10467-2401, USA. [6]Institute of Human Genetics, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands. [7]Department of Pediatrics, Human Genetics and Obstetrics & Gynecology, University of Utah, Salt Lake City, Utah 84132, USA. [8]Department of Human Genetics, University of Nijmegen, 6500 HB Nijmegen, The Netherlands. Correspondence should be addressed to A.B. e-mail: ballabio@tigem.it

breakpoint lies within the region of overlap between cosmid U143A10 and U202D11 (Fig. 2). We observed an abnormal banding pattern by hybridizing a pool of three *Eco*RI subclones derived from cosmid U143A10 to digested genomic DNA from affected individuals carrying the inverted X chromosome, as compared to a normal female control (data not shown).

Isolation of *MID1* cDNA clones

During our efforts to construct a transcription map of the Xp22 region, we performed exon-trapping experiments on the cosmids contained in a previously described contig[28]. Several exon-trapping products were identified. Sequence analysis of two exon-trapping products (6H and 10a) derived from cosmid U232B2 (Fig. 2) revealed nucleotide sequence identity with a cluster of seventeen overlapping ESTs. This cluster of ESTs had previously been mapped to the large interval between *DXS1223* and *DXS999* in Xp22 (ref. 29). A 1.8-kb infant brain cDNA clone (c-1zh10) containing ESTs Z44417 and Z40343 was used as a hybridization probe against a differentiated hNT2 neuron cDNA library, and twenty-three primary positive clones were isolated, seven of which were randomly selected for further study. Using both cDNA-specific and vector primers, we characterized these seven cDNA clones by end-sequencing, restriction mapping and PCR. A consensus cDNA sequence of 3,452 bp was assembled. We validated the authenticity of the 5′ end of the cDNA by sequencing the corresponding genomic region. The putative initiation codon was identified at position 187 bp and is located within a nucleotide sequence that fulfils Kozak's criteria[30]. An in-frame stop codon (TGA) was detected 48 nucleotides upstream of this putative initiation codon. The first downstream in-frame stop codon (TGA) was identified at nucleotide position 2188, resulting in a coding region of 2,001 bp with a predicted protein product of 667 amino acids (Fig. 3*a*). Preliminary characterization of additional cDNA clones suggests that *MID1* is subject to extensive alternative splicing in the 5′ UTR (data not shown). No splice variants in the *MID1* coding region have been detected.

An X-chromosome inversion associated with OS disrupts *MID1*

Hybridization of the complete *MID1* cDNA to YAC clones assigned to the Xp22 region revealed three positive clones (YACs 184F6, 61B3 and NB2E12) known to map in the interval between *DXS1136* and *DXF22S5* (ref. 20), containing the OS inversion breakpoint (data not shown). FISH analysis with the full-length *MID1* cDNA on a chromosome spread derived from a male cell line showed a unique signal on Xp22, indicating that *MID1* represents a single-copy X-linked locus (data not shown). *Eco*RI-digested cosmid clones from the region covered by YAC 184F6 were hybridized with the full-length *MID1* cDNA. The cosmid clones U232B2, U11D6, U239F12, U8D11, U202D11 and U76G12 were observed to be positive (Fig. 2). The *MID1* genomic locus spans a total of eleven overlapping cosmid clones, representing approximately 300 kb (Fig. 2). Hybridization using subclones/oligonucleotide probes derived from the 5′ and 3′ regions of the *MID1* cDNA demonstrated that the direction of transcription is orientated from the centromere to Xpter, with the initial methionine codon located in cosmid U8D11 (Fig. 2). Together, these data demonstrate that the 5′ end of *MID1* spans the OS inversion breakpoint, suggesting that the disease phenotype displayed by individuals carrying the inverted X chromosome is caused by impaired *MID1* transcription. Furthermore, Southern-blot analysis of DNA from somatic-cell hybrids retaining the abnormal X chromosome revealed that *MID1* is deleted in two female patients (BA95 and BA333; ref. 31, pers. comm.) with *de novo* Xp22;Yq11 translocations (data not shown).

Fig. 1 A patient with X-linked OS. This patient belongs to family OS16 (patient 1644). The hypertelorism, telecanthi, mild entropion of lower eyelid, high nasal bridge and surgically corrected cleft lip are evident.

MID1 expression

The *MID1* cDNA was hybridized against commercially available northern blots containing polyA$^+$ RNA extracted from a variety of fetal and adult human tissues (Fig. 4*a*,*b*). A ~7-kb transcript and a fainter ~3.5-kb transcript were observed in all tissues present. In fetal organs, the kidney showed the highest level of *MID1* expression, followed by brain and lung. *MID1* appears to be expressed at relatively low levels in fetal liver. In the adult tissues, *MID1* expression is most abundant in the heart, placenta and brain. We investigated *MID1* expression in an affected male from the family carrying the pericentric X-chromosome inversion. RT-PCR analysis showed absence of *MID1* expression in the affected male, whereas an amplification product of expected size was seen in the control (data not shown). A control RT-PCR, using primers derived from an independent locus, was performed in parallel and amplified both the affected and unaffected individuals (data not shown).

To begin investigating the spatio-temporal expression pattern of *MID1* during development, we isolated a partial murine cDNA and performed preliminary RNA *in situ* hybridization studies on whole-mount mouse embryos. These experiments confirmed that *Mid1* is ubiquitously transcribed at an early stage (embryo day 10.5), with the exception of the developing heart (Fig. 5*c*). Notably, *Mid1* expression was found to be highest in the first and second branchial arches, along both the proximo-distal (Fig. 5*a*,*b*) and medio-lateral axes (Fig. 5*c*). We did not detect any expression signals above background level by using the sense probe as a control (data not shown).

MID1 encodes a RING-finger protein

Analysis of the consensus *MID1* cDNA sequence showed that the predicted MID1 amino-acid sequence shares homology with a family of proteins containing a tripartite protein–protein interaction motif consisting of a RING-finger, B-box (one or two copies) and

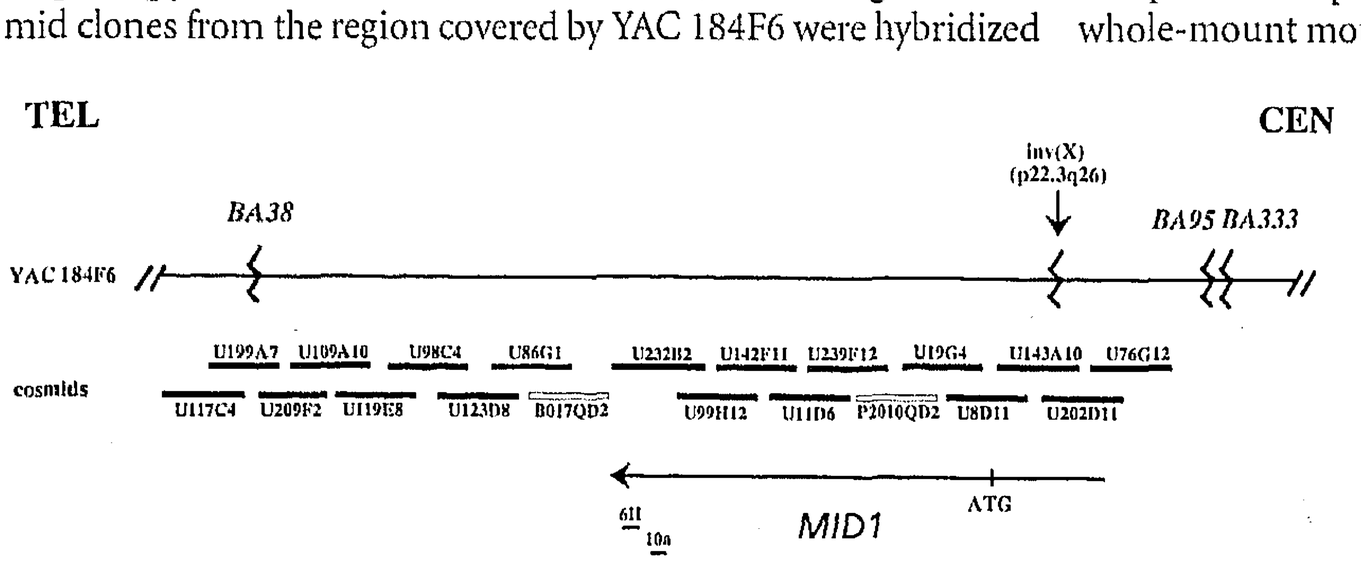

Fig. 2 Physical map of the Xp22.3 region containing the *MID1* gene. The position, start codon and direction of transcription of the *MID1* gene are indicated. The exon trapping products 6H and 10a are shown as short horizontal lines. The Xp22.3 breakpoint of the pericentric inversion (X)(p22.3q26) associated with OS is indicated by a jagged vertical line, as are the deletion/translocation breakpoints of patients BA38, BA95 and BA333 (ref. 60). Cosmid clones are indicated as filled (LLNL library) or hollow (RLDB library) horizontal bars (ref. 28 and Cox *et al.*, submitted) . The genomic region illustrated is contained in YAC clone 184F6.

a

```
 1 M E T L E S E L T C P I C L E L F E D P  20
21 L L L P C A H S L C F N C A H R I L V S  40
41 H C A T N E S V E S I T A F Q C P T C R  60
61 H V I T L S Q R G L D G L K R N V T L Q  80
81 N I I D R F Q K A S V S G P N S P S E T 100
101 R R E R A F D A N T M T S A E K V L C Q 120
121 F C D Q D P A Q D A V K T C V T C E V S 140
141 Y C D E C L K A T H P N K K P F T G H R 160
161 L I E P I P D S H I R G L M C L E H E D 180
181 E K V N M Y C V T D D Q L I C A L C K L 200
201 V G R H R D H Q V A A L S E R Y D K L K 220
221 Q N L E S N L T N L I K R N T E L E T L 240
241 L A K L I Q T C Q H V E V N A S R Q E A 260
261 K L T E E C D L L I E I I Q Q R R Q I I 280
281 G T K I K E G K V M R L R K L A Q Q I A 300
301 N C K Q C I E R S A S L I S Q A E H S L 320
321 K E N D H A R F L Q T A K N I T E R V S 340
341 M A T A S S Q V L I P E I N L N D T F D 360
361 T F A L D F S R E K K L L E C L D Y L T 380
381 A P N P P T I R E E L C T A S Y D T I T 400
401 V H W T S D D E F S V V S Y E L Q Y T I 420
421 F T G Q A N V V S L C N S A D S W H I V 440
441 P N I K Q N H Y T V H G L Q S G T K Y I 460
461 F M V K A I N Q A G S R S S E P G K L K 480
481 T N S Q P F K L D P K S A H R K L K V S 500
501 H D N L T V E R D E S S S K K S H T P V E 520
521 R F T S Q G S Y G V A G N V F I D S G R 540
541 H # Y W E V V I S G S T W Y A I G L A Y K 560
561 S A P K H E W I G K N S A S W A L C R C 580
581 N N N W V V R H N S K E I P I E P A P H 600
601 L R R V G I L L D Y D N G S I A F Y D A 620
621 L N S I H L Y T F D V A F A Q P V C P T 640
641 F T V W N K C L T I I T G L P I P D H L 660
661 D C T E Q L P * 667
```

RING FINGER

B BOXES

COILED-COIL DOMAIN

C-TERMINAL DOMAIN

b

PROTEIN	POSITION	MOTIF
MID1	478-510	KLKTNSQPFK **LDPKSAHRKLKVS** HDNLTVERDE
Ro/SSA	281-313	MLRTCAVHIT **LDPDTANPWLILS** EDRRQVRLGD
Efp	451-483	ELLEYYIKVI **LDYNTAHNKVALS** ECYTVASVAE
PWA33	440-472	VVQPGLAPLT **LDPNTAHPNLVLS** EGLTSVKYTD
RFP	310-342	EAQLYSVDVT **LDPDTAYPSLILS** DNLRQVRYSY
Xnf7	426-458	VVIPSLTPML **LDPTSAHPNLHLS** DGLTSVRYGE
butyrophilin	297-329	RATLHAVDVT **LDPDTAHPHLFLY** EDSKSVRLED
MID1	529-556	GVAGNVFIDS **GRHYWEVV** ISGSTWYAIG
Ro/SSA	330-357	MVLGAQHFHS **GKHYWEVD** VTGKEAWDLG
Efp	500-527	QVLGLHCYKK **GIHYWEVE** LQKNNFCGVG
PWA33	489-516	LVLGAEGFDS **GKHYWEVE** VGNKTAWDVG
RFP	359-386	CVLGSPCFIA **GRHYWEVE** VGDKAKWTIG
Xnf7	475-502	LVLGSQGFDS **GRHYWEVE** VGDKTAWDVG
butyrophilin	346-373	CVMGREAFTS **GRHYWEVE** VGDRTDWAIG
MID1	593-628	IPIEPAPHLR **RVGILLDYDNGSIAFY** DALNSIHLYT
Ro/SSA	396-431	TPLHLQVPPC **QVGIFLDYEAGMVSFY** NITDHGSLIY
Efp	566-601	EKTLPSTKAT **RVGVLLNCDHGFVIFF** AVADKVHLMY
PWA33	555-590	KTLNLTSKPS **KIGVYLDYEGGQVSFY** NADDMSPIYT
RFP	425-460	TALPLRTPLQ **RVGIFLDYDAGEVSFY** NVTERCHTFT
Xnf7	541-576	KSLSLSSHPR **KIGVYVDYEGGQISFY** NADDMTIIYT
butyrophilin	411-446	TPLPLAGPPR **RVGVFLDYESGDIFFY** NMTDGSHIYT

Fig. 3 MID1 protein sequence. ***a***, Predicted MID1 amino-acid sequence. The conserved cysteine and histidine residues in the RING-finger and B-box domains are boxed. The predicted coiled-coil domain is underlined, as are the three conserved regions of the C-terminal domain. The in-frame 3-bp deletion (ΔM438) is shown in bold; the site of the 24-bp tandem duplication (1787ins24bp) is indicated by #; the site of a single base pair insertion (1744insG) is indicated by V. ***b***, Multiple-sequence alignment of conserved motifs in the C-terminal domain. Residues conserved in more than 50% of proteins are shown in bold. The proteins MID1, Ro/SSA, Efp, PWA33, RFP and Xnf7 all belong to the B-box family. Butyrophilin also contains the conserved C-terminal motifs, although it does not contain the other characteristic features of the B-box family.

coiled-coil domains (Fig. 3*a*). The consensus RING-finger sequence contains seven conserved cysteine residues and a single conserved histidine residue in the following format: CXXCX(9–27)CX(1–3)HXXCXXCX(4–48)CXXC, where X may be any amino acid[32]. MID1 also contains two copies of a B-box zinc-binding domain, immediately followed by a leucine-rich coiled-coil domain. The B-box consensus sequence consists of seven conserved cysteine/histidine residues. The first MID1 B-box contains an additional conserved cysteine residue (position 137), present in a subset of tripartite domain–containing proteins. An additional stretch of amino-acid conservation in the C-terminal region was also noted between MID1 and five-tripartite motif–containing proteins. The functional significance of this region is unknown. This conserved C-terminal region is also present in butyrophilin, an unrelated protein lacking all the characteristic elements of the tripartite motif[33] (Fig. 3*b* and Discussion).

Mutation analysis in OS patients

To partly characterize the genomic structure of the *MID1* gene, we performed direct sequencing of cosmid clones U232B2, U11D6, U239F12 and U8D11, or of *Eco*RI plasmid subclones derived from them, using oligonucleotides designed from the *MID1* consensus cDNA sequence. Oligonucleotide primers were designed in the flanking intronic sequences of two characterized exons to allow PCR amplification and subsequent SSCP analysis of the *MID1* coding regions and splice sites from OS patients and controls. We performed preliminary mutation analysis of the two characterized exons on 22 independent OS families in which X-linked inheritance was possible but not confirmed. Three SSCP band-shifts segregating with affected individuals from three OS pedigrees were identified. The first band-shift was observed with a 246-bp PCR product (using primers TO-3835 and TO-3870) spanning nucleotides 1472–1633 of the coding region. The two other band-shifts were both observed with a 345-bp PCR product (using primers TO-3869 and TO-3825) spanning nucleotides 1634–1844 of the coding region. Sequencing the *MID1* PCR products derived from affected individuals revealed three mutations, all of which result in abnormalities in the C-terminal region of the MID1 protein (Fig. 3*a*). The first mutation to be detected (using primers TO-3835 and TO-3870) in an affected male (1290, family OS5) was an in-frame 3-bp deletion (ΔM438), resulting in the absence of a methionine residue from the MID1 protein (Fig. 3*a*). On SSCP analysis, the band-shift caused by the deletion was seen to segregate with all five of the affected individuals tested and was absent in all six unaffected individuals in family OS5. The addition of a single G nucleotide (1745insG) was observed in an affected male (1644, family OS16; Figs 3*a*,6*a*). The resulting frameshift produces a premature stop codon at position 1795, truncating the *MID1* gene product by 131 amino acids in the conserved C-terminal domain (Figs 3*a*,*b*). The affected mother (1642) was found to be heterozygous for the mutation (Fig. 6*a*). An affected male (1676, family OS20) displayed a tandem duplication of 24 bp (1787ins24bp), resulting in the insertion of eight amino acids (FIDSGRHL) at position 534 of

the MID1 protein (Fig. 6*b*). This duplication disrupts a highly conserved C-terminal motif (Fig. 3*a*,*b*). As expected, the affected mother (1674) was found to be heterozygous for the mutation (Fig. 6*b*). None of the three identified mutations were detected in 100 X chromosomes from normal females of European descent (data not shown).

Discussion

MID1 is a member of the B-box family of zinc-finger proteins. MID1 encodes a putative 667-amino-acid protein containing a RING-finger motif, followed by two B-boxes and a coiled-coil domain. This tripartite motif identifies a subfamily of zinc-finger proteins called the B-box family[31,34]. The RING and B-box motifs are defined by the presence of conserved cysteines and histidines, which are active in zinc binding. At present, the function of the RING finger and B-box domains is unknown. Many of the proteins containing this tripartite motif are found in multiprotein macromolecular assemblages within cells[35]. There is evidence that these motifs mediate protein–protein interactions, probably acting as potential homo-heterodimerization domains[34].

Genes belonging to the B-box family encode putative transcriptional regulators. Three of them have oncogenic potential in man and mouse when involved in a chromosomal translocation, resulting in the fusion of the tripartite motif with other proteins. These three genes are *PML*, which is associated with acute promyelocytic leukaemia when fused to the retinoic-acid receptor–α gene[26,27]; *RFP*, the ret finger gene, which acquires transforming activity when fused with the ret proto-oncogene[36]; and the mouse *Tif1* gene, which, after being fused to the B-raf proto-oncogene, encodes a transforming fusion protein[37]. Other members of the B-box family are the *Xenopus laevis Xnf7* gene, encoding a maternally derived transcriptional regulator engaged in early development[38]; *Rpt1*, which encodes a potential transcription factor that regulates expression of the interleukin-2 receptor[39]; *Efp*, an oestrogen-responsive gene[40]; PWA33, a *Pleurodeles waltl* gene encoding the nuclear protein that binds to lampbrush chromosome loops[41]; and the gene encoding the 52-kD auto-antigen Ro/SSA involved in Sjögren syndrome[42].

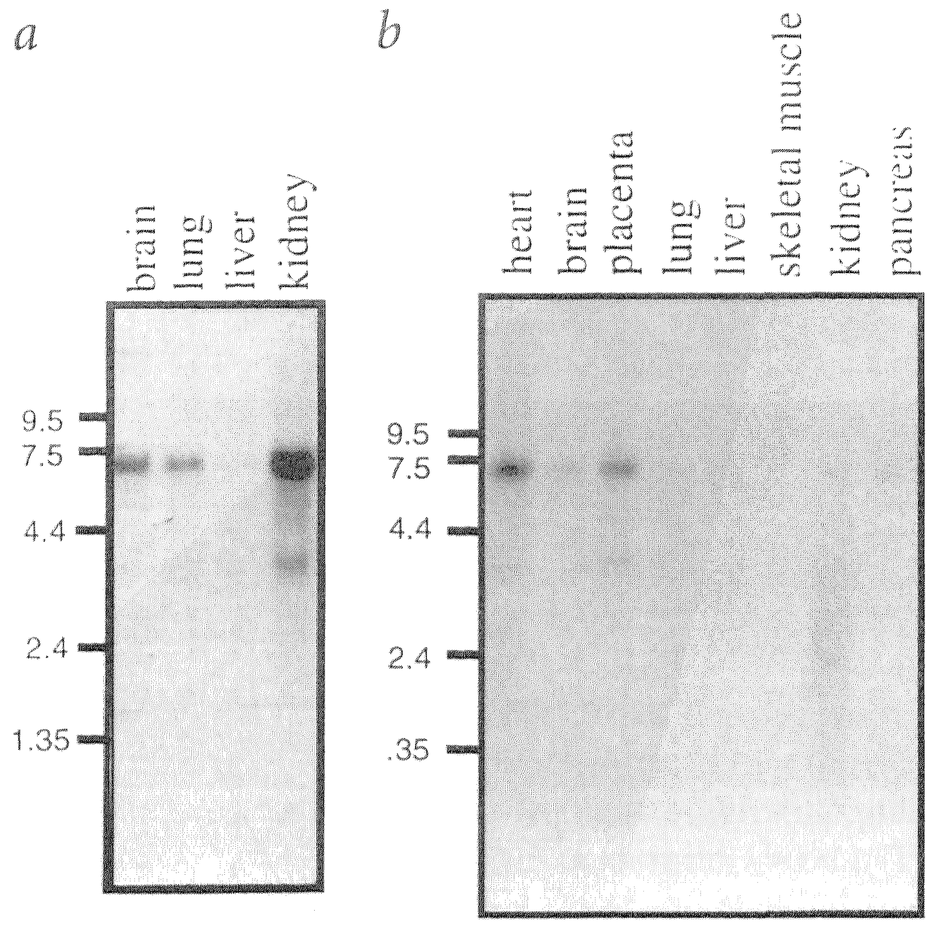

Fig. 4 Northern-blot analysis of the *MID1* gene. The full-length *MID1* cDNA was hybridized to multiple-tissue northern blots of fetal (***a***) and adult (***b***) human tissues. A transcript of approximately 7-kb and a fainter transcript of approximately 3.5-kb were observed in all tissues present.

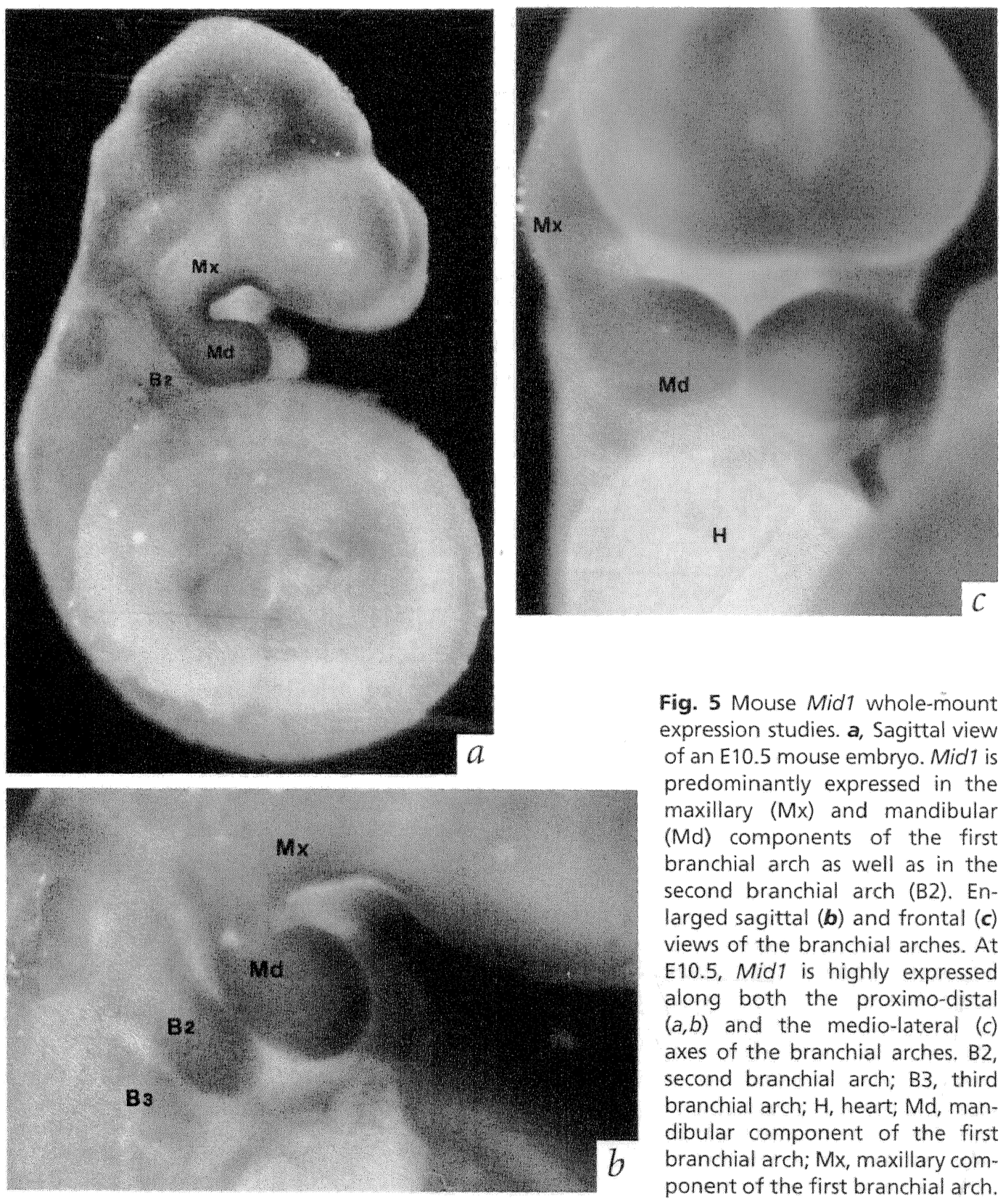

Fig. 5 Mouse *Mid1* whole-mount expression studies. ***a***, Sagittal view of an E10.5 mouse embryo. *Mid1* is predominantly expressed in the maxillary (Mx) and mandibular (Md) components of the first branchial arch as well as in the second branchial arch (B2). Enlarged sagittal (***b***) and frontal (***c***) views of the branchial arches. At E10.5, *Mid1* is highly expressed along both the proximo-distal (*a*,*b*) and the medio-lateral (*c*) axes of the branchial arches. B2, second branchial arch; B3, third branchial arch; H, heart; Md, mandibular component of the first branchial arch; Mx, maxillary component of the first branchial arch.

Additionally, sequence analysis of MID1 showed high levels of identity in the C-terminus region with a subgroup of the previously mentioned proteins—namely, RFP, Efp, Xnf7, PWA33 and Ro/SSA. In particular, three regions containing a highly conserved core sequence, flanked by conserved residues, are found within this C-terminal region (Fig. 3*b*). The spacing between these three motifs is highly conserved in all of the proteins, suggesting that it is important for their function. These motifs are also found in another protein, the bovine butyrophilin, a secreted molecule completely unrelated to the B-box family, which associates with the fat globule membrane of bovine milk *via* a single transmembrane domain[33]. The C-terminal region of butyrophilin is cytoplasmic and shares high homology with the C-terminal domain of MID1. Moreover, it has been demonstrated that the cytoplasmic region of murine butyrophilin is able to specifically associate with a 150-kD protein[43]. This suggests that the C-terminal domain detected in all of these proteins is a novel protein–protein interacting domain.

***MID1* is mutated in Opitz syndrome.** We have demonstrated that *MID1* is the X-linked OS gene. *MID1* maps to Xp22, where the X-linked form of OS has been mapped[17]; furthermore, *MID1* is disrupted in a pericentric inversion of the X chromosome inv(X)(p22.3q26), co-segre-

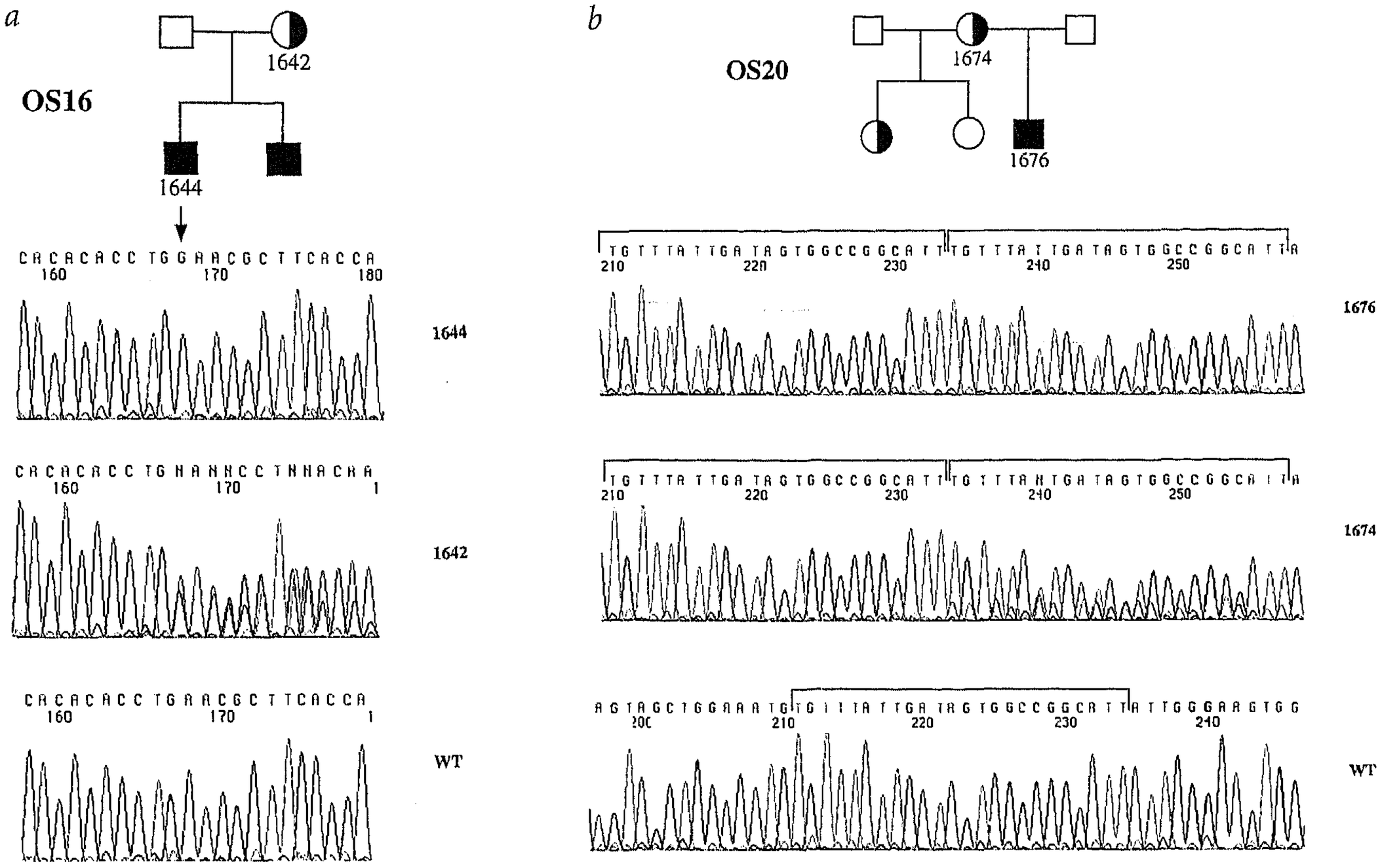

Fig. 6 Mutations of *MID1* in OS families. ***a***, Family OS16. Sequence analysis reveals the addition of a single G nucleotide (arrow) in affected son, 1644, and affected heterozygous mother, 1642. Note presence of double peaks in female patient 1642 downstream of the insertion site due to the presence of both normal and out-of-frame alleles. WT, control wild-type allele. ***b***, Family OS20. Sequence analysis reveals a 24-bp duplication (square brackets) in affected son, 1676, and affected heterozygous mother, 1674. Note presence of double peaks in patient 1674 downstream of the duplication due to the presence of both normal and mutant alleles. WT, control wild-type allele.

gating with the disease in a large pedigree[19], and is mutated in three independent OS families. Among the mutations found, a duplication and an insertion create a major change in the MID1 protein product by disrupting the highly conserved C-terminal domain. In the first case, the mutation causes a duplication of eight amino acids within the second conserved C-terminal motif. In the second case, the 1-bp insertion produces a frameshift that causes premature termination of the MID1 protein and deletion of the last two C-terminal motifs. These mutations strongly support an important functional role for the C-terminal domain. This is a new concept, as most previous studies on proteins bearing the tripartite motif have concentrated on the function of the RING finger and the B-box, not on the C-terminal region of these proteins.

MID1 is completely deleted in two female patients (BA95 and BA333) with *de novo* Xp22;Yq11 translocations, resulting in very large (more than 10 Mb) terminal Xp deletions (ref. 31, pers. comm). The X-chromosome breakpoints of these two translocations fall just centromeric to the 5′ end of the gene (Fig. 2). These female patients show symptoms highly suggestive of OS, such as dysmorphic face, abnormal external genitalia, imperforate anus and developmental delay. Several of these anomalies —such as imperforate anus, hypertelorism, laryngeal web and cardiac malformations—have been reported in other females carrying even larger terminal deletions of Xp22 and displaying more complex phenotypes, such as microphthalmia with linear skin lesions and Aicardi or Goltz syndrome[44–47]. These findings suggest that the complex phenotype of these patients is due to a contiguous-gene syndrome, allowing for the first time the dissection of the clinical spectrum of these patients by attribution of some symptoms to deletions to *MID1*.

Possible role of *MID1* in midline development. Most symptoms of OS are due to midline fusion defects. For instance, widely spaced eyes may be produced by a failure of migration of the orbits towards the midline, clefts in the lip, palate and trachea/oesophagus by a defect in closure of the facial and pharyngeal processes, and hypospadias by a defective fusion of urethral folds. In general, these symptoms suggest a defective patterning of midline elements along the medio-lateral axis, resulting in an expansion of the midline. A number of genes active in midline patterning have been identified, and some of them have been implicated in human genetic disorders. For example, mutations in the zinc-finger *GLI3* gene cause Greig syndrome, which is characterized by hypertelorism and pre- and post-axial polydactyly[48]. *GLI3* is negatively regulated by *sonic hedgehog* (shh)[49], which is expressed in the midline structures of both axial mesoderm and the ventral neural tube[50]. Notably, mutations of *SHH* in humans cause a loss rather than an expansion of midline tissue, resulting in holoprosencephaly, a combination of severe forebrain and midfacial defects[51,52].

Proteins containing the RING-finger and B-box domains are expressed from a very early stage during development. In particular, the *Xenopus laevis Xnf7* gene is maternally expressed and is retained in the cytoplasm until the midblastula transition, when it re-enters the nucleus[22–25]. Northern-blot analysis of *MID1* in fetal and adult human tissues revealed a broad pattern of expression. Preliminary data obtained by RNA *in situ* hybridization on

whole-mount mouse embryos determined that *Mid1* is almost ubiquitously expressed. We detected higher expression levels in the branchial arches, correlating with the defective patterning of the skeletal elements within the craniofacial structures.

Our data indicate that mutations in the gene encoding a novel member of the B-box family are responsible for X-linked OS. It could be anticipated that the study of the function and dysfunction of this gene in humans and model organisms will help clarify the mechanisms underlying the establishment and development of midline structures in vertebrates. Moreover, in identifying the *MID1* gene implicated in OS, we have discovered a new aetiologic factor in genetically determined mental retardation and cleft lip and palate. The search for genes homologous to *MID1* could be instrumental in the discovery of the molecular basis of other syndromes associated with mental retardation and cleft lip and palate, including the chromosome 22–linked form of OS.

Methods

Patients with Opitz syndrome. Three families segregating OS were identified as carrying a mutation in *MID1*. Detailed clinical findings, including photographs, were previously reported for kindred OS5 (refs 4,17). Kindred OS16 consists of two affected brothers and their affected mother (Fig. 6*a*). The brothers have similar craniofacial findings, including bilateral cleft lip and palate and other characteristic OS facial features, such as widely spaced eyes, widow's peak and high nasal bridge (Fig. 1). Both have hypospadias; prominent costal arches, giving the thorax the appearance of a pectus excavatum; and increased space between the third and fourth fingers. They have a mild developmental delay, especially in their speech. Both are hyperactive. Their mother has widely spaced eyes and is of normal intelligence. The proband in kindred OS20 (Fig. 6*b*) has hypertelorism, severe hypospadias (penile-scrotal hypospadias) and respiratory problems with inspiratory stridor due to trachaeomalacia (although no structural anomalies on examination). At 3 years of age, he was developmentally delayed, functioning at a level of mild mental retardation. His mother and half-sister both have hypertelorism and widow's peak. All individuals participating in this study were clinically evaluated and their disease status was ascertained before initiation of molecular studies. Informed consent was obtained in accordance with the standards set by local review boards.

Sequence analysis. Sequence analysis was performed essentially as previously described[53]. C-terminal conserved motifs were detected with MEME software[54]. The predicted coiled-coil domain was generated with the COILS program[55].

cDNA cloning. *MID1* cDNA clones were isolated according to standard procedures[56] from a mature hNT neuron cDNA library (Stratagene).

Characterization of *MID1* genomic structure. Sequencing of *MID1* cDNA clones and genomic clones was performed according to the manufacturer's recommendations (Amersham). Positions of introns were determined by comparison of genomic and cDNA sequences.

Hybridization. Southern and northern hybridizations were performed as previously described[57].

Whole-mount expression studies. *In situ* hybridization experiments on whole-mount mouse embryos were performed as previously described[58]. Clone ME10.2 was used as a template for the digoxygenin-labelled RNA antisense and sense probes.

Mutation detection. Mutations were detected by SSCP analysis as previously described[59]. Samples showing variation were compared to those of other family members to assess allele segregation, and with at least 40 unrelated from people of European descent control samples to distinguish mutations from polymorphisms. PCR products amplified from total genomic DNA from patients demonstrating SSCP variants were run on 1% agarose gels and purified with the QIAquick gel extraction kit (Qiagen). Sequencing was performed on an ABI sequencer. The following primers (5′→3′) were used to detect *MID1* mutations: TO-3835, CTTGTGCCAAAGAACTGCAC; TO-3870, GAGCAGATAAGACATGACAG; TO-3869, TTGAAAATATGCTATTGGTTATG; and TO-3825, AGTACAGAATGAGATGTCC.

GenBank accession numbers. The sequences of the full-length human *MID1* cDNA (Y13667) and the mouse partial *Mid1* cDNA (Y14848) have been deposited in GenBank. Accession numbers for exon-trapping products 6H and 10a are Z44417, Z40343, F07967, F04215, R66237, R67582, N25620, N34218, W90080, W90206, AA009434, AA009915, AA195713, AA195750, AA235258, AA242882 and AA2521647.

Acknowledgements

We are grateful to all OS family members for their participation in the study. We thank A. Bulfone, T. Cox, L. Dal Zotto, E. Montini, V. Suckow, T. Todt, V. Valsecchi and S. van der Maarel for their valuable contributions. We thank L. Bernard for technical support. This work is supported by the Italian Telethon Foundation, by the EC under grant Nos. BMH4-CT96-1134 and BMH4-CT96-0889 (to A.B.), by the Deutsche Forschungsgemeinschaft grant No. R0389/17 (to H.H.R.) and by the Children's Hospital of Philadelphia Development Fund and NIH grant Nos. R29HD28732 and RO1HD29862 (to M.M.). N.A.Q. is supported by a Marie Curie postdoctoral fellowship (ERBFMBICT960649) from the EC.

Received 23 July; accepted 10 September 1997.

1. Opitz, J.M., Summitt, R.L. & Smith, D.W. The BBB syndrome familial telecanthus with associated congenital anomalies. *Birth Defects Orig. Art. Ser.* **(V)2**, 86–94 (1969).
2. Opitz, J.M., Frías, J.L., Guttenberger, J.E. & Pellet, J.R. The G syndrome of multiple congenital anomalies. *Birth Defects Orig. Art. Ser.* **(V)2**, 95–102 (1969).
3. Opitz, J.M. G syndrome (hypertelorism with esophageal abnormality and hypospadias, or hypospadias-dysphagia, or 'Opitz-Frias' or 'Opitz-G' syndrome): perspective in 1987 and bibliography. *Am. J. Med. Genet.* **28**, 275–285 (1987).
4. Robin, N.H., Opitz, J.M. & Muenke, M. Opitz G/BBB syndrome: clinical comparisons of families linked to Xp22 and 22q, and a review of the literature. *Am. J. Med. Genet.* **62**, 305–317 (1996).
5. Bershof, J.F., Guyuron, B. & Olsen, M.M. G syndrome: a review of the literature and a case report. *J. Craniomaxillofac. Surg.* **20**, 24–27 (1992).
6. Cappa, M., Borrelli, P., Marini, R. & Neri, G. The Opitz syndrome: a new designation for the clinically indistinguishable BBB and G syndromes. *Am. J. Med. Genet.* **28**, 303–309 (1987).
7. Cordero, J.F. & Holmes, L.B. Phenotypic overlap of the BBB and G syndromes. *Am. J. Med. Genet.* **2**, 145–152 (1978).
8. Farndon, P.A. & Donnai, D. Male-to-male transmission of the G syndrome. *Clin. Genet.* **23**, 446–448 (1983).
9. Fryns, J.P., Delooz, J. & van den Berghe, H. Posterior scalp defects in Opitz syndrome: another symptom related to a defect in midline development. *Clin. Genet.* **42**, 314–316 (1992).
10. Guion-Almeida, M.L. & Richieri-Costa, A. CNS midline anomalies in the Opitz G/BBB syndrome: report on 12 Brazilian patients. *Am. J. Med. Genet.* **43**, 918–928 (1992).
11. Howell, L. & Smith, J.D. G syndrome and its otolaryngological manifestations. *Ann. Otol. Rhinol. Laryngol.* **98**, 185–190 (1989).
12. MacDonald, M.R., Schaefer, G.B., Olney, A.H., Tamayo, M. & Frias, J.L. Brain magnetic resonance imaging findings in the Opitz G/BBB syndrome: extension of the spectrum of midline brain anomalies. *Am. J. Med. Genet.* **46**, 706–711 (1993).
13. Sedano, H.O. & Gorlin, R.J. Opitz oculo-genital-laryngeal syndrome (Opitz BBB/G compound syndrome). *Am. J. Med. Genet.* **30**, 847–849 (1988).
14. Stevens, C.A. & Wilroy, R.S., Jr. The telecanthus-hypospadias syndrome. *J. Med. Genet.* **25**, 536–542 (1988).
15. Tolmie, J.L., Coutts, N. & Drainer, I.K. Congenital anal anomalies in two families with the Opitz G syndrome. *J. Med. Genet.* **24**, 688–691 (1987).
16. Verloes, A., Le Merrer, M. & Briard, M.-L. BBBG syndrome or Opitz syndrome: new family. *Am. J. Med. Genet.* **34**, 313–316 (1989).
17. Robin, N.H. *et al.* Opitz syndrome is genetically heterogeneous, with one locus on Xp22, and a second locus on 22q11.2. *Nature Genet.* **11**, 459–461 (1995).
18. May, M., Huston, S., Wilroy, R.S. & Schwartz, C. Linkage analysis in a family with the Opitz GBBB syndrome refines the location of the gene in Xp22 to a 4 cM region. *Am. J. Med. Genet.* **68**, 244–248 (1997).
19. Verloes, A. *et al.* Opitz GBBB syndrome: chromosomal evidence of an X-linked form. *Am. J. Med. Genet.* **59**, 123–128 (1995).
20. Ferrero, G.B. *et al.* An integrated physical and genetic map of a 35 Mb region on chromosome Xp22.3–Xp21.3. *Hum. Mol. Genet.* **4**, 1821–1827 (1995).
21. Wapenaar, M.C. *et al.* The genes for X-linked ocular albinism (OA1) and microphthalmia with linear skin defects (MLS): cloning and characterization of the critical regions. *Hum. Mol. Genet.* **2**, 947–952 (1993).
22. Miller, M. *et al.* The nuclear-cytoplasmic distribution of the *Xenopus* nuclear

factor, xnf7, coincides with its state of phosphorylation during early development. *Development* **113**, 569–575 (1991).
23. Li, X. & Etkin, L.D. *Xenopus* nuclear factor 7 (xnf7) possesses an NLS that functions efficiently in both oocytes and embryos. *J. Cell Sci.* **105**, 389–395 (1993).
24. Li, X., Shou, W., Kloc, M., Reddy, B.A. & Etkin, L.D. Cytoplasmic retention of *Xenopus* nuclear factor 7 before the mid blastula transition uses a unique anchoring mechanism involving a retention domain and several phosphorylation sites. *J. Cell Biol.* **124**, 7–17 (1994).
25. Shou, W. *et al.* Finely tuned regulation of cytoplasmic retention of *Xenopus* nuclear factor 7 by phosphorylation of individual threonine residues. *Mol. Cell. Biol.* **16**, 990–997 (1996).
26. de The, H. *et al.* The PML-RARα fusion mRNA generated by t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell* **66**, 675–684 (1991).
27. Grignani, F. *et al.* Acute promyelocytic leukemia: from genetics to treatment. *Blood* **83**, 10–25 (1994).
28. Wapenaar, M.C. *et al.* A YAC-based binning strategy facilitating the rapid assembly of cosmid contigs: 1.6 Mb of overlapping cosmids in Xp22. *Hum. Mol. Genet.* **3**, 1155–1161 (1994).
29.Schuler, G.D. et al. A gene map of the human genome. *Science* **274**, 540–546 (1996).
30. Kozak, M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857–873 (1984).
31. Johnston, K. *et al.* De novo X;Y translocation associated with imperforate anus and retinal pigmentary abnormalities. *Am. J. Med. Genet.* **27**, 603–611 (1987).
32.Borden, K.L.B. *et al.* Characterisation of a novel cysteine/histidine-rich metal binding domain from *Xenopus* nuclear factor XNF7. *FEBS Lett.* **335**, 255–260 (1993).
33. Mather, I.H. & Jack, L.J. A review of the molecular and cellular biology of butyrophilin, the major protein of bovine milk fat globule membrane. *J. Dairy Sci.* **76**, 3832–3850 (1993).
34. Borden, K.L.B. & Freemont, P.S. The RING finger domain: a recent example of a sequence-structure family. *Curr. Opin. Struct. Biol.* **6**, 395–401 (1996).
35. Borden, K.L.B. *et al. In vivo* and *in vitro* characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia protooncoprotein PML. *Proc. Natl. Acad. Sci. USA* **93**, 1601–1606 (1996).
36. Takahashi, M., Inaguma, Y., Hiai, H. & Hirose, F. Developmentally regulated expression of a human "finger"-containing gene encoded by the 5′ half of the *ret* transforming gene. *Mol. Cell. Biol.* **8**, 1853–1856 (1988).
37. LeDouarin, B. *et al.* The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *EMBO J.* **14**, 2020–2033 (1995).
38. Reddy, B.A., Kloc, M. & Etkin, L. The cloning and characterization of a maternally expressed novel zinc finger nuclear phosphoprotein (xnf7) in *Xenopus laevis*. *Dev. Biol.* **148**, 107–116 (1991).
39. Patarca, R. *et al.* rpt-1, an intracellular protein from helper/inducer T cells that regulates gene expression of interleukin 2 receptor and human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **85**, 2733–2737 (1988).
40. Inoue, S. *et al.* Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *Proc. Natl. Acad. Sci. USA* **90**, 11117–11121 (1993).
41. Bellini, M., Lacroix, J.-C. & Gall, J.G. A putative zinc-binding protein on lampbrush chromosome loops. *EMBO J.* **12**, 107–114 (1993).
42. Chan, E.K.L., Hamel, J.C., Buyon, J.P. & Tan, E.M. Molecular definition and sequence motifs of the 52-kD component of human SS-A/Ro autoantigen. *J. Clin. Invest.* **87**, 68–76 (1991).
43. Ishii, T. *et al.* Carboxy-terminal cytoplasmic domain of mouse butyrophilin specifically associates with a 150-kDa protein of mammary epithelial cells and milk fat globule membrane. *Biochim. Biophys. Acta.* **1245**, 285–292 (1995).
44. al-Gazali, L.I. *et al.* Two 46,XX,t(X;Y) females with linear skin defects and congenital microphthalmia: a new syndrome at Xp22.3. *J. Med. Genet.* **27**, 59–63 (1990).
45. Temple, I.K., Hurst, J.A., Hing, S., Butler, L. & Baraitser, M. De novo deletion of Xp22.2-pter in a female with linear skin lesions of the face and neck, microphthalmia, and anterior chamber eye anomalies. *J. Med. Genet.* **27**, 56–58 (1990).
46. Bird, L.M., Krous, H.F., Eichenfield, L.F., Swalwell, C.I. & Jones, M.C. Female infant with oncocytic cardiomyopathy and microphthalmia with linear skin defects (MLS): a clue to the pathogenesis of oncocytic cardiomyopathy? *Am. J. Med. Genet.* **53**, 141–148 (1994).
47. Lindsay, E.A. *et al.* Microphthalmia with linear skin defects (MLS) syndrome: clinical, cytogenetic and molecular characterization. *Am. J. Med. Genet.* **49**, 229–234 (1994).
48. Vortkamp, A., Gessler, M. & Grzeschik, K.-H. GLI3 zinc-finger gene interrupted by translocations in Greig syndrome families. *Nature* **352**, 539–540 (1991).
49. Marigo, V., Johnson, R.L., Vortkamp, A. & Tabin, C.J. Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development. *Dev. Biol.* **180**, 273–283 (1996).
50. Hammerschmidt, M., Brook, A. & McMahon, A.P. The world according to *hedgehog*. *Trends Genet.* **13**, 14–21 (1997).
51. Belloni, E. *et al.* Identification of *Sonic hedgehog* as a candidate gene responsible for holoprosencephaly. *Nature Genet.* **14**, 353–356 (1996).
52. Roessler, E. *et al.* Mutations in the human *Sonic Hedgehog* gene cause holoprosencephaly. *Nature Genet.* **14**, 357–360 (1996).
53. Banfi, S. *et al.* Identification and mapping of human cDNAs homologous to *Drosophila* mutant genes through EST database searching. *Nature Genet.* **13**, 167–174 (1996).
54. Bailey, T.L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Second International Conference on Intelligent Systems for Molecular Biology* 28–36 (AAAI Press, Menlo Park, California, 1994).
55. Lupas, A., Van Dyke, M. & Stock, J. Predicting coiled coils from protein sequences. *Science* **252**, 1162–1164 (1991).
56. Maniatis, T., Fritsch, E.S. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982).
57. Franco, B. *et al.* A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin embryopathy. *Cell* **81**, 15–25 (1995).
58. Riddle, R.D., Johnson, R.L., Laufer, E. & Tabin, C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401–1416 (1993).
59. Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA* **86**, 2766–2770 (1989).
60. Schaefer, L. *et al.* A high resolution deletion map of human chromosome Xp22. *Nature Genet.* **4**, 272–279 (1993).