Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1

K.A. McAllister[1], K.M. Grogg[1], D.W. Johnson[1], C.J. Gallione[1], M.A. Baldwin[1], C.E. Jackson[2,3], E.A. Helmbold[3], D.S. Markel[3], W.C. McKinnon[4], J. Murrell[5], M.K. McCormick[5], M.A. Pericak-Vance[6], P. Heutink[7], B.A. Oostra[7], T. Haitjema[8], C.J.J. Westerman[8], M.E. Porteous[9], A.E. Guttmacher[4], M. Letarte[10] & D.A. Marchuk[1]

[1]Department of Genetics, Duke University Medical Center, Box 3175, Durham, North Carolina 27710, USA
[2]Division of Clinical and Molecular Genetics, Henry Ford Hospital, Detroit, Michigan 48202, USA
[3]Family Studies Core, Human Genome Center, University of Michigan Medical Center, Ann Arbor, Michigan 48109, USA
[4]Department of Pediatrics, University of Vermont College of Medicine, Burlington, Vermont 05401, USA
[5]Molecular Neurogenetics Laboratory, Massachusetts General Hospital, Boston, Massachusetts 02129, USA
[6]Division of Neurology, Duke University Medical Center, Durham, North Carolina 27710, USA
[7]Department of Clinical Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
[8]St Antoniusziekenhuis, P.O. Box 2500, 3430 EM Nieuwegein, The Netherlands
[9]Department of Human Genetics, University of Edinburgh, Edinburgh EH4 2XU, UK
[10]Division of Immunology and Cancer, Research Institute, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada

Correspondence should be addressed to D.A.M.

Hereditary haemorrhagic telangiectasia (HHT) is an autosomal dominant disorder characterized by multisystemic vascular dysplasia and recurrent haemorrhage. Linkage for some families has been established to chromosome 9q33–q34. In the present study, endoglin, a transforming growth factor β (TGF-β) binding protein, was analysed as a candidate gene for the disorder based on chromosomal location, expression pattern and function. We have identified mutations in three affected individuals: a C to G substitution converting a tyrosine to a termination codon, a 39 base pair deletion and a 2 basepair deletion which creates a premature termination codon. We have identified endoglin as the HHT gene mapping to 9q3 and have established HHT as the first human disease defined by a mutation in a member of the TGF-β receptor complex.

Hereditary haemorrhagic telangiectasia (HHT) or Osler-Weber-Rendu disease (OMIM #18730) is an autosomal dominant disorder characterized by multisystemic vascular dysplasia and recurrent haemorrhage. The disorder is named after the recurrent haemorrhage from vascular lesions, especially in the nasal mucosa and gastrointestinal tract, and for the presence of mucosal, dermal and visceral telangiectases. Pulmonary arteriovenous malformations (PAVMs) occur in approximately 20% of patients and are associated with serious complications including stroke and brain abscess. Other neurological manifestations include cerebral arteriovenous malformation, aneurysm and migraine headache.

Ultrastructural analyses of the vascular dysplasia seen in affected individuals have failed to demonstrate a unique pathological abnormality that might suggest the nature of the primary biochemical defect. Studies indicate that the dilated channels of telangiectases are lined by a single layer of endothelium attached to a continuous basement membrane[1,2]. The earliest event in the formation of telangiectases appears to be dilation of post-capillary venules[3]. Eventually the dilated venules connect to enlarging arterioles through capillary segments which later disappear, creating direct arteriolar–venular connections. This sequence of events is associated with a perivascular mononuclear infiltrate[3]. Various explanations have been put forward to explain the angiodysplasia seen in HHT including endothelial cell degeneration[4], defects in endothelial junctions[2], lack of elastic fibres and incomplete smooth muscle cell coating of the vessels[1], and weak connective tissue surrounding the vessel[4].

Genetic linkage for some HHT families was recently established to markers on chromosome 9q33–q34 (refs 5,6), and the locus was named *OWR1*. Genetic heterogeneity was established with the identification of some families clearly not linked to this region[6–9]. However, a predisposition to PAVMs in multiple affected members of a family correlates with 9q3 linkage[7–9], establishing a valuable diagnostic criterion for OWR1. The identification of key obligate recombinants in affected individuals allowed refinement of the *OWR1* locus and placed the most likely candidate interval between *D9S60* and *D9S61* in a 2 centiMorgan (cM) interval[6,7,9]. The *COL5A1* gene mapping to chromosome 9q34 has been considered a candidate gene for this disorder, but our mapping studies indicate it lies distal to the candidate interval with obligate recombinants identified in *OWR1* kindreds[10].

Here, we investigate a strong candidate for the *OWR1* disease gene. Endoglin is a homodimeric integral membrane glycoprotein expressed at high levels on human vascular endothelial cells of capillaries, arterioles and venules in all tissues examined[11]. On endothelial cells, endoglin is the most abundant transforming growth factor β (TGF-β) binding protein[12]. In the presence of TGF-β ligand, endoglin can associate with the signaling receptors RI and RII and is thought to initiate response to the growth factor[13] (Letarte *et al.*, unpublished observations). TGF-β is the prototype of a family of at least 25 growth factors which regulate growth, differentiation, motility, tissue remodeling, wound repair and programmed cell death in many cell types[14].

Endoglin has been mapped to human chromosome 9q34 using fluorescence *in situ* hybridization (FISH)[15]. It lies in a broad region on 9q33–q34 that shows conserved

```
                                   Exon 1                                    ▼
    1     ATGGACCGCGGCACGCTCCCTCTGGCTGTTGCCCTGCTGCTGGCCAGCTGCAGCCTCAGCCCCACAAGTCTTGCAGAAACAGTCCATTGT
    1     MetAspArgGlyThrLeuProLeuAlaValAlaLeuLeuLeuAlaSerCysSerLeuSerProThrSerLeuAlaGluThrValHisCys
                                                            Exon 2
   91     GACCTTCAGCCTGTGGGCCCCGAGAGGGGCGAGGTGACATATACCACTAGCCAGGTCTCGAAGGGCTGCGTGGCTCAGGCCCCCAATGCC
   31     AspLeuGlnProValGlyProGluArgGlyGluValThrTyrThrThrSerGlnValSerLysGlyCysValAlaGlnAlaProAsnAla

                                                 ▼
  181     ATCCTTGAAGTCCATGTCCTCTTCCTGGAGTTCCCAACGGGCCCGTCACAGCTGGAGCTGACTCTCCAGGCATCCAAGCAAAATGGCACC
   61     IleLeuGluValHisValLeuPheLeuGluPheProThrGlyProSerGlnLeuGluLeuThrLeuGlnAlaSerLysGlnAsnGlyThr
                          Exon 3                                                                    ▼
  271     TGGCCCCGAGAGGTGCTTCTGGTCCTCAGTGTAAACAGCAGTGTCTTCCTGCATCTCCAGGCCCTGGGAATCCCACTGCACTTGGCCTAC
   91     TrpProArgGluValLeuLeuValLeuSerValAsnSerSerValPheLeuHisLeuGlnAlaLeuGlyIleProLeuHisLeuAlaTyr
                                                    Exon 4
  361     AATTCCAGCCTGGTCACCTTCCAAGAGCCCCCGGGGGTCAACACCACAGAGCTGCCATCCTTCCCCAAGACCCAGATCCTTGAGTGGGCA
  121     AsnSerSerLeuValThrPheGlnGluProProGlyValAsnThrThrGluLeuProSerPheProLysThrGlnIleLeuGluTrpAla

                                                                                   ▼
  451     GCTGAGAGGGGCCCCATCACCTCTGCTGCTGAGCTGAATGACCCCCAGAGCATCCTCCTCCGACTGGGCCAAGCCCAGGGGTCACTGTCC
  151     AlaGluArgGlyProIleThrSerAlaAlaGluLeuAsnAspProGlnSerIleLeuLeuArgLeuGlyGlnAlaGlnGlySerLeuSer
                                                            Exon 5
  541     TTCTGCATGCTGGAAGCCAGCCAGGACATGGGCCGCACGCTCGAGTGGCGGCCGCGTACTCCAGCCTTGGTCCGGGGCTGCCACTTGGAA
  181     PheCysMetLeuGluAlaSerGlnAspMetGlyArgThrLeuGluTrpArgProArgThrProAlaLeuValArgGlyCysHisLeuGlu

                                                                     ▼
  631     GGCGTGGCCGGCCACAAGGAGGCGCACATCCTGAGGGTCCTGCCGGGCCACTCGGCCGGGCCCCGGACGGTGACGGTGAAGGTGGAACTG
  211     GlyValAlaGlyHisLysGluAlaHisIleLeuArgValLeuProGlyHisSerAlaGlyProArgThrValThrValLysValGluLeu
                          Exon 6
  721     AGCTGCGCACCCGGGGATCTCGATGCCGTCCTCATCCTGCAGGGTCCCCCCTACGTGTCCTGGCTCATCGACGCCAACCACAACATGCAG
  241     SerCysAlaProGlyAspLeuAspAlaValLeuIleLeuGlnGlyProProTyrValSerTrpLeuIleAspAlaAsnHisAsnMetGln

                ▼             *
  811     ATCTGGACCACTGGAGAATACTCCTTCAAGATCTTTCCAGAGAAAAACATTCGTGGCTTCAAGCTCCCAGACACACCTCAAGGCCTCCTG
  271     IleTrpThrThrGlyGluTyrSerPheLysIlePheProGluLysAsnIleArgGlyPheLysLeuProAspThrProGlnGlyLeuLeu
                                                    Exon 7
  901     GGGGAGGCCCGGATGCTCAATGCCAGCATTGTGGCATCCTTCGTGGAGCTACCGCTGGCCAGCATTGTCTCACTTCATGCCTCCAGCTGC
  301     GlyGluAlaArgMetLeuAsnAlaSerIleValAlaSerPheValGluLeuProLeuAlaSerIleValSerLeuHisAlaSerSerCys

            ▼                                                                Exon 8
  991     GGTGGTAGGCTGCAGACCTCACCCGCACCGATCCAGACCACTCCTCCCAAGGACACTTGTAGCCCGGAGCTGCTCATGTCCTTGATCCAG
  331     GlyGlyArgLeuGlnThrSerProAlaProIleGlnThrThrProProLysAspThrCysSerProGluLeuLeuMetSerLeuIleGln

                                                               ▼
 1081     ACAAAGTGTGCCGACGACGCCATGACCCTGGTACTAAAGAAAGAGCTTGTTGCGCATTTGAAGTGCACCATCACGGGCCTGACCTTCTGG
  361     ThrLysCysAlaAspAspAlaMetThrLeuValLeuLysLysGluLeuValAlaHisLeuLysCysThrIleThrGlyLeuThrPheTrp
                                                          Exon 9
 1171     GACCCCAGCTGTGAGGCAGAGGACAGGGGTGACAAGTTTGTCTTGCGCAGTGCTTACTCCAGCTGTGGCATGCAGGTGTCAGCAAGTATG
  391     AspProSerCysGluAlaGluAspArgGlyAspLysPheValLeuArgSerAlaTyrSerSerCysGlyMetGlnValSerAlaSerMet

                                                             ▼
 1261     ATCAGCAATGAGGCGGTGGTCAATATCCTGTCGAGCTCATCACCACAGCGGAAAAAGGTGCACTGCCTCAACATGGACAGCCTCTCTTTC
  421     IleSerAsnGluAlaValValAsnIleLeuSerSerSerSerProGlnArgLysLysValHisCysLeuAsnMetAspSerLeuSerPhe
                            Exon 10                                                         ▼
 1351     CAGCTGGGCCTCTACCTCAGCCCACACTTCCTCCAGGCCTCCAACACCATCGAGCCGGGGCAGCAGAGCTTTGTGCAGGTCAGAGTGTCC
  451     GlnLeuGlyLeuTyrLeuSerProHisPheLeuGlnAlaSerAsnThrIleGluProGlyGlnGlnSerPheValGlnValArgValSer

 1441     CCATCCGTCTCCGAGTTCCTGCTCCAGTTAGACAGCTGCCACCTGGACTTGGGGCCTGAGGGAGGCACCGTGGAACTCATCCAGGGCCGG
  481     ProSerValSerGluPheLeuLeuGlnLeuAspSerCysHisLeuAspLeuGlyProGluGlyGlyThrValGluLeuIleGlnGlyArg
                   Exon 11
 1531     GCGGCCAAGGGCAACTGTGTGAGCCTGCTGTCCCCAAGCCCCGAGGGTGACCCGCGCTTCAGCTTCCTCCTCCACTTCTACACAGTACCC
  511     AlaAlaLysGlyAsnCysValSerLeuLeuSerProSerProGluGlyAspProArgPheSerPheLeuLeuHisPheTyrThrValPro

                                                                            ▼
 1621     ATACCCAAAACCGGCACCCTCAGCTGCACGGTAGCCCTGCGTCCCAAGACCGGGTCTCAAGACCAGGAAGTCCATAGGACTGTCTTCATG
  541     IleProLysThrGlyThrLeuSerCysThrValAlaLeuArgProLysThrGlySerGlnAspGlnGluValHisArgThrValPheMet
            Exon 12                      ▼                                            Exon 13
 1711     CGCTTGAACATCATCAGCCCTGACCTGTCTGGTTGCACAAGCAAAGGCCTCGTCCTGCCCGCCGTGCTGGGCATCACCTTTGGTGCCTTC
  571     ArgLeuAsnIleIleSerProAspLeuSerGlyCysThrSerLysGlyLeuValLeuProAlaValLeuGlyIleThrPheGlyAlaPhe

                                                              ▼
 1801     CTCATCGGGGCCCTGCTCACTGCTGCACTCTGGTACATCTACTCGCACACGCGTTCCCCCAGCAAGCGGGAGCCCGTGGTGGCGGTGGCT
  601     LeuIleGlyAlaLeuLeuThrAlaAlaLeuTrpTyrIleTyrSerHisThrArgSerProSerLysArgGluProValValAlaValAla
                                Exon 14
 1891     GCCCCGGCCTCCTCGGAGAGCAGCAGCACCAACCACAGCATCGGGAGCACCCAGAGCACCCCCTGCTCCACCAGCAGCATGGCATAGCCC
  631     AlaProAlaSerSerGluSerSerSerThrAsnHisSerIleGlySerThrGlnSerThrProCysSerThrSerSerMetAla***
          CGGCCCCCCGCGCTCGCCCAGCAGGAGAGACTGAGCAGCCGCCAGCTGGGAGCACTGGTGTGAACTCACCCTGGGAGCCAGTCCTCCACT
          CGACCCAGAATGGAGCCTGCTCTCCGCGCCTACCCTTCCCGCCTCCCTCTCAGAGGCCTGCTGCCAGTGCAGCCACTGGCTTGGAACACC
          TTGGGGTCCCTCCACCCCACAGAACCTTCAACCCAGTGGGTCTGGGATATGGCTGCCCAGGAGACAGACCACTTGCCACGCTGTTGTAAA
          AACCCAAGTCCCTGTCATTTGAACCTGGATC
```

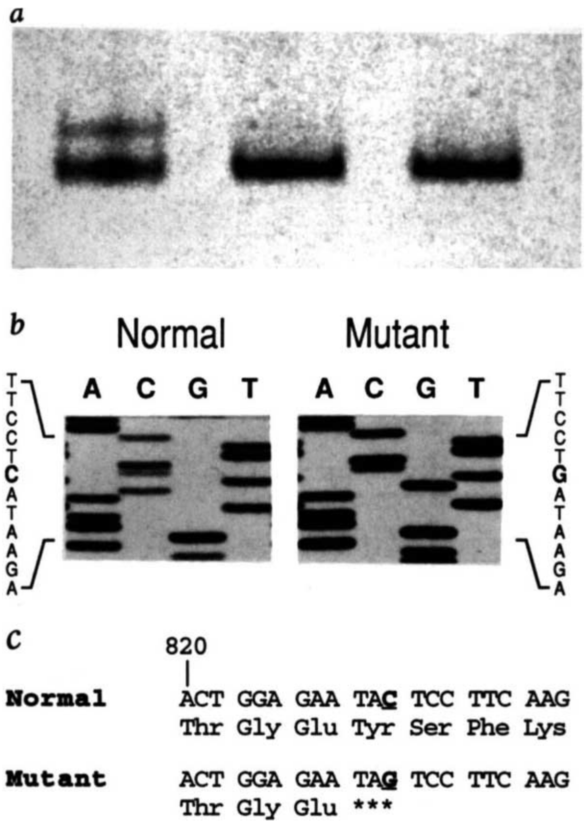


Fig. 2 Point mutation in sample 1159. *a*, Heteroduplex analysis showing a shift in an affected proband (sample 1159; lane 1) next to two samples (lanes 2 and 3) not displaying this anomaly. *b*, Sequence of representative clones revealing the normal (C) and mutant sequence (G) of amplified exon 7 in sample 1159. *c*, Consequence of the substitution in sample 1159. This C to G substitution converts a tyrosine to a stop codon at amino acid 277.

synteny with mouse chromosome 2. The mouse endoglin (*End* or *Eng*) locus is genetically inseparable with the genetic marker adenylate kinase 1, *AK1*[16]. Human *AK1* is itself genetically inseparable from *D9S60*[17], which forms the proximal border of the candidate region for *OWR1* (refs 6,7,9). FISH using interphase nuclei places *AK1* between *D9S60* and *D9S61* on human chromosome 9q[18], within the smallest candidate interval. We inferred that the tightly linked human endoglin gene would also lie within this interval and be a candidate gene for *OWR1*.

Genomic structure of endoglin

As an initial screen for gross abnormalities in the endoglin gene in affected HHT individuals, Southern blots of DNA from the probands of 33 unrelated families were probed with a nearly complete cDNA of endoglin, clone 18A[11]. This analysis using three restriction endonucleases revealed no gross abnormalities of the endoglin gene in these samples. RT-PCR was attempted using RNA prepared from several EBV-transformed lymphoblast lines established from our patient cohort, but expression levels of endoglin appeared to be too low to allow routine amplification in a single round (35 cycles) of PCR. As the expression of endoglin is restricted to endothelial cells, activated monocytes[19], syncytiotrophoblasts[20], and certain stromal cells[21], screening for mutations within endoglin cDNA was not feasible. We therefore began to determine the genomic structure of endoglin.

A gridded cosmid chromosome 9 library was screened with the 18A cDNA probe and 17 cosmids were obtained. Southern analysis of these clones in comparison with total genomic DNA revealed that one cosmid, 21c10, appeared to contain most of the gene. This cosmid was subcloned into a phagemid library which was screened for positive plaques with the 18A cDNA probe. Hybridizing clones were converted to plasmids and sequenced using vector primers flanking the cloning site to identify intron–exon borders.

Preliminary sequence analysis suggests that the coding region of endoglin contains 14 exons (Fig. 1). One or more splice junctions may remain unidentified within the 5' end of the gene, as the sequence denoted exon 1, which contains the putative signal peptide, was found to be missing in the 21c10 cosmid. There is also evidence for alternative splicing variants of the endoglin transcript (Pichantes *et al.*, unpublished observations). Since only one variant was used to identify subclones for genomic sequencing, it is possible that additional exons exist within the depicted coding region (Fig. 1). (The exon number assignments must be regarded as preliminary until the entire gene structure is resolved.)

The 14 exons are sufficiently small to allow for PCR amplification of each as a single unit (Table 1). The smallest is exon 12 which contains the complete membrane spanning domain and is 55 basepairs (bp) in length. The longest exon completely contained within the coding region is exon 11 (258 bp). Exon 14 contains at least 429 bp but contains only 125 bases of coding information, the remainder being the 3' untranslated region.

Identification of HHT mutations

In an initial screen for mutations, primers located within the introns flanking exons 7 and 11 (the first exons to be identified) were designed to establish PCR assays for each exon (see Methodology). A panel of 68 DNA samples was used for the mutation screen. These were collected from probands of unrelated families, most of which were members of kindreds with PAVM involvement, increasing

◀ Fig. 1 Genomic structure of endoglin. The cDNA sequence of endoglin is shown with the amino acid sequence below. The nucleotide and amino acid positions are based on numbering the A in the ATG start codon of the full-length L-form of endoglin[36] as nucleotide number 1. Exons 1 through 14 are labelled above the cDNA sequence in bold and the intron/exon borders are marked with arrows. Exon 7 codes for amino acids 273–331; exon 11 spans residues 477–562. The four potential N-linked glycosylation sites are in boldface, italicized type and are underlined. The membrane spanning domain is double underlined. The positions of the mutations described in this report are shown in relation to the gene structure; the C to G change at nucleotide 831 is indicated by a star and the positions of the two deletions are underlined (nucleotides 882–920 and 1553–1554). The 2 bp deletion creates a premature termination codon which is indicated by bold type.

Table 1 PCR assays for endoglin exons

Exon	Size	Forward primer Sequence (5'–3')	Reverse prime Sequence (5'–3')	PCR prod. size	Buffer	[MgCl$_2$] (mM)	Annealing temp. (°C)	No. cycles
1	> 67							
2	152	cctcataaggtggctgtgatgatg	catctgccttggagcttcctct	413	BMB	1.5	60	35
3	141							
4	163	ttcctgacctcctacatgg	ctcttggtgcccaagttt	330	TNK 50	1.0	51	30
5	166	gggctctgttaggtgcag	gggtggggctttataaggga	294	TNK 25	1.0	57	35
6	127	ctgtccgcttcagtgttccatc	ggaaacttccctgatccagaggtt	230	TNK 100	1.5	59	40
7	176	gaggcctggcataaccct	gtggccactgatccaagg	315	BMB	1.5	60	35
8	141							
9	178							
10	117	agattgaccaagtctccctccc	aggctgtctccctcctgactct	227	TNK 50	1.0	61	35
11	258	actcaggggtgggaactctt	ccttccatgcaaaccacag	430	TNK 50	1.0	57	32
12	55	gagtaaacctggaagccgc	gccactagaacaaacccgag	164	TNK 100	1.0	55	35
13	111	ccagcacaacagggtaggggat	ctcagaggcttcactgggctcc	255	TNK 50	1.0	61	35
14	> 429	tgaagcctctgagggattgagg	gagttcacaccagtgctcccag	267	TNK 50	1.0	57	40

the likelihood that the individuals would harbour mutations at the *OWR1* locus[7–9]. Included in our analysis was one member from each of eight 9q3-linked families previously described[5,7,9]. Heteroduplex analysis was performed on amplified products from this cohort as a screen for potential mutations. Abnormal PCR products seen on these gels were directly sequenced for further analysis.

With this initial screen, we have identified three mutations in affected individuals. The first mutation was identified by a heteroduplex shift in the exon 7 PCR product from sample 1159 (Fig. 2*a*). The products of two independent PCR reactions were directly sequenced and a C (normal) and a G (mutation) at nucleotide position 831 were clearly visible. PCR products amplified from this individual were then cloned and individual clones sequenced to validate the results of the direct sequencing (Fig. 2*b*). This change converts a tyrosine at codon 277 to a termination codon (Fig. 2*c*). This mutation is present in the proband of a pedigree with multiple affected members with documented PAVMs. However, additional members of this family were not available for analysis. The truncated protein resulting from this mutation would comprise only half of the extracellular domain and lack the membrane spanning and cytoplasmic domains.

Amplification of exon 7 in sample 8019 revealed a second mutation in a family (Family 3186) previously linked to 9q3 (ref. 9). A second PCR fragment smaller than the wild-type fragment was visible in both agarose gels and heteroduplex analysis, suggesting the existence of a deletion. The smaller fragment was not seen in 278 normal chromosomes and is unlikely to be a polymorphism. Sequence analysis of the PCR products revealed a 39 bp deletion in the exon beginning at nucleotide position 882 of endoglin (Fig. 3*a*). This in-frame deletion removes 13 amino acids (amino acids 295 to 307) and alters the first amino acid of a potential N-linked glycosylation site (see Fig. 1). Amplification of this exon revealed the presence of the deletion in all affected family members, but no unaffected members (Fig. 3*b*).

Heteroduplex analysis of amplified exon 11 revealed a very pronounced band in sample 2061 that was not visible with agarose gel electrophoresis (Fig. 4*a*). Independent clones of the PCR product were sequenced and revealed the wild-type sequence and a 2 bp deletion beginning with nucleotide 1553 of endoglin (Fig. 4*b*). This deletion creates a *Mae*III restriction site. This sample was from the proband of a family with multiple affected members displaying PAVMs. Exon 11 was amplified from all available family members and digested with *Mae*III. All affected family members share the additional *Mae*III site, whereas the unaffected members do not, establishing linkage of this mutation to the disease phenotype in this family (Fig. 4*c*). The mutation creates a frame shift that results in a premature termination codon 8 amino acids beyond the deletion (Fig. 4*d*). The predicted truncated protein would lack the membrane spanning and cytoplasmic domains of endoglin.

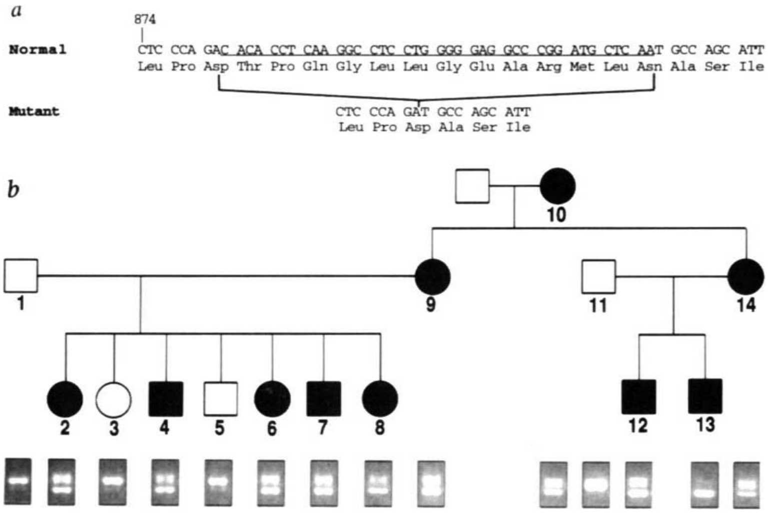


Fig. 3 *a*, Deletion mutation in sample 8019. *a*, This 39 bp deletion, found in the 9q3-linked Family 3186 (ref. 9), is located at nucleotide positions 882 though 920 in exon 7, removing 13 amino acids from the protein and altering the first amino acid (position 307) in a potential N-linked glycosylation site. *b*, Segegation of 39 bp deletion in Family 3186. Amplification of exon 7 in Family 3186 on an agarose gel reveals the presence of a lower band (the 39 bp deletion product) in affected family members only. Preferential amplification of the smaller fragment is sometimes observed (see individual 13).

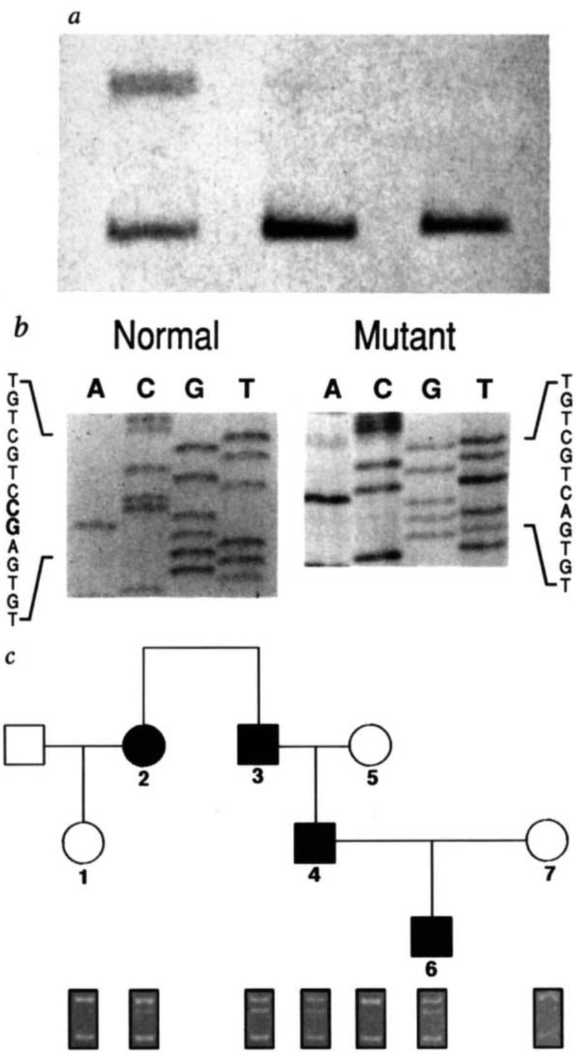
Discussion

We have established endoglin as the *OWR1* disease locus mapping to 9q3 with the identification of three independent mutations in affected HHT individuals. The gene maps to the tightest *OWR1* candidate interval on 9q33–q34 based on evidence from the mouse and human genetic and physical maps. The restricted tissue distribution of endoglin and its expression at high levels on the surface of endothelial cells is consistent with the pathology of the disorder. Two of the three mutations described in this report create premature termination codons and would be expected to lead to reduced message levels that if translated would encode severely truncated proteins, suggestive of loss-of-function alleles. The third would remove 13 amino acids from the extracellular domain of the receptor and would likely have a deleterious effect on receptor function. Finally, a defect in a cell surface binding protein would account for the limited and localized nature of the vascular lesions present in this disease.

TGF-β *in vivo* is a potent angiogenic factor and a mediator of vascular remodelling as it controls extracellula matrix production by endothelial cells, smooth muscl cells and pericytes[22]. Following soft tissue injury or i response to angiogenic factors, microvascular endothelia cells detach from their basement membrane, migrate an proliferate in the interstitial stroma, and form ne microvessels. When grown *in vitro* in three-dimensiona gels and in the presence of TGF-β, these endothelial cell form tube-like cellular aggregates with a lumen and tigh junctions, and deposit an organized basement membrane mimicking vessel formation[22]. However, TGF-β, almos exclusively in the β1 isoform, will inhibit the proliferatio of endothelial cells grown on plastic[22,23]. The response o endothelial cells to TGF-β depends on the interactio with the surrounding extracellular matrix via integrin expressed on their surface[24]. The production of matri proteins by stromal interstitial cells, smooth muscle cells pericytes and endothelial cells and the expression o integrins on endothelial cells are also regulated by TGF- β[22,24].

We would expect that endothelial cells lacking endogli would respond poorly to TGF-β1 and form abnorma vessels, particularly in response to injury. TGF-β signalin is mediated by TGF-β receptors RI and RII, which form heteromeric complex upon binding TGF-β[25,26]. Endogli binds TGF-β1 and -β3 with high affinity but does not bin -β2 (ref. 12), and is structurally related to betaglyca which binds all three isoforms of TGF-β[27,28]. Betaglycan i the presence of ligand interacts with the signaling kinas complex of RI and RII and potentiates the response to al three isoforms of the growth factor[29]. Endoglin als interacts with the kinase complex suggesting a potentiatin role similar to that of betaglycan. Endothelial cells expres very low levels of betaglycan, which may explain thei poor response to TGF-β2 (ref.30). Thus endoglin-deficien endothelial cells, as observed in *OWR1*-linked patients would only express the signaling RI and RII complex an would lack the regulatory co-receptor capable o controlling the response. This might alter cell adhesio properties, leading to the vascular anomalies seen in thi disorder. Stromal cells in several tissues[21] and activate monocytes[19] also express endoglin and could be impaire in their response to TGF-β1 in the vascular lesions o OWR1 patients.

The biological consequences of a defect within an member of the TGF-β ligand-receptor complex are onl beginning to be elucidated. TGF-β1 null mice were initiall found to die within three weeks of birth from sever



d

```
          1543
          |
Normal    AAC TGT GTG AGC CTG CTG TCC CCA AGC CCC GAG GGT GAC
          Asn Cys Val Ser Leu Leu Ser Pro Ser Pro Glu Gly Asp

Mutant    AAC TGT GTG ACT GCT GTC CCC AAG CCC CGA GGG TGA
          Asn Cys Val Thr Ala Val Pro Lys Pro Arg Gly ***
```

Fig. 4 Deletion mutation in sample 2061. *a*, Heteroduplex analysis showing a shift in an affected proband (sample 2061; lane 1) next to two samples (lanes 2 and 3) not displaying this anomaly. *b*, Sequence of 2 bp deletion in sample 2061. Sequence of the two independently cloned PCR products of affected individual sample 2061 revealing the normal sequence and the 2 bp deletion in exon 11 beginning at nucleotide position 1553. *c*, Segregation of 2 bp deletion in family 63. The 2 bp deletion creates an additional *Mae*III restriction site. Affected family members exhibit an additional novel fragment visible as the middle band in each lane, with half the intensity of the other two bands. A second novel band produced by digestion at this site is not visible on this gel. *d*, Consequence of the 2 bp deletion in sample 2061. This mutation causes a frameshift and a premature termination after an additional seven amino acids.

inflammatory disease[31,32]. However, maternal TGF-β1 was later shown to contribute to the survival of the embryos. When a null female was treated with dexamethasone to prevent the inflammatory response and mated to a heterozygous male, the null offspring showed abnormal heart formation with unusual atrioventricular junctions and disordered myocyte proliferation[33].

Our results are particularly significant because they show a direct link between a human genetic disorder and defined mutations within the TGF-β binding protein endoglin. As HHT is a genetically heterogeneous disease, the observation that endoglin is defective in *OWR1*-linked families suggests that loci encoding other components of the TGF-β ligand-receptor complex might explain the locus heterogeneity. Before determining the genomic structure of endoglin for mutation analysis, we performed genetic linkage analysis on three non-9q3-linked families, using genetic markers located near the map positions of the TGF-β ligands (β1, β2 and β3) and the only other mapped TGF-β receptor, the TGF-β type II receptor. One of these families was linked to 3p22 (D.W.S. *et al.*, manuscript in preparation), where the TGF-β II receptor is located[34]. This supports our hypothesis that the locus heterogeneity in this disorder may be due to mutations within other members of the TGF-β receptor complex or other endothelial cell components of the TGF-β signal transduction pathway.

Continued analysis of endoglin mutations in *OWR1*-linked families will be necessary to determine the functional consequences of the mutations with regard to binding of endoglin with its ligands or disruption of the interactions of endoglin with other TGF-β receptors. There were no obvious differences in the clinical features seen in the three families with described mutations, suggesting that the molecular pathology for the potential loss-of-function mutations (premature termination) and the small deletion may be the same. We favour the hypothesis that OWR1, although inherited as an autosomal dominant disorder, exhibits a cellular-recessive pathology and requires inactivation of the normal allele as the initiating event in the formation of a vascular lesion. In support of this, the vascular lesions in this disorder are localized to discreet regions within the affected tissue, with no evidence of abnormal vessel structure or pathology outside the lesions themselves. Cutaneous lesions are most often located in exposed areas that might be subject to ultraviolet irradiation, and in some cases seem to increase in number with age. This hypothesis can be tested by uncovering a secondary (somatic) mutation or loss of heterozygosity in the wild-type allele in endothelial cells of an OWR1-derived vascular lesion. Alternatively, the intiating event in the formation of a vascular lesion might be damage to the vessel wall. In this scenario, mutations resulting in reduced endoglin expression might lead to defective repair of the vessel wall. Mutations disrupting endoglin dimerization might lead to a similar outcome due to a dominant-negative effect.

Received 31 October; accepted 8 November 1994.

Methodology

Clinical evaluation. The diagnostic criteria used for collection of family members was as described[5,7]. Descriptions and pedigrees of all 9q3-linked families have been published[5,7,9].

Genomic sequence determination. A nearly complete cDNA sequence of endoglin (18A) was used to screen a gridded chromosome 9 cosmid library (Los Alamos National Laboratory). One subclone that contained nearly all hybridizing bands that are seen with genomic DNA was subcloned using Lambda ZAP Express system (Stratagene). Plaque screens were performed by hybridization with 18A cDNA probe to identify positive clones. Intron–exon borders were identified by sequencing these clones using Sequenase Ver. 2.0 DNA sequencing kit (United States Biochemical) using both vector and exon primers.

PCR amplification of exons. Primers were designed from itron genomic sequences flanking exons 7 and 11 of endoglin. For exon 7 (nts 817–992), the forward primer is 5'–GAGGCCTGG-CATAACCCT, and the reverse primer is 5'–GTGGCCA-CTGATCCAAGG. The 315 bp product was amplified using a buffer consisting of 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl2, and 50 mM KCl. After initial denaturation, 35 cycles of the following program were run: 94 °C for 30 s; 60 °C for 60 s; 72 °C for 30 s. For exon 11 (nts 1429–1686), the forward primer is 5'–ACTCAGGGGTG-GGAACTCTT and the reverse is 5'–CCTTCCATGCAAACCACAG. The 430 bp product was amplified in 10 mM Tris-HCl, pH 8.3, 1 mM MgCl2, 50 mM KCl and 5 mM NH4Cl. After initial denaturation, 32 cycles of the following program were run: 94 °C for 30 s; 57 °C for 60 s; 72 °C for 30 s.

Ten of the other 14 exons can be amplified using conditions described in Table 1. Each amplification raction contains 100 ng of genomic DNA, 100 ng of each oligonucleotide primer, 0.20 mM of each dNTP, 1.25 U of *Taq* DNA polymerase in final volume of 25 µl. Reaction conditions were optimized individually for each primer pair by adjusting annealing temperatures and buffer conditions as described[35], or using the *Taq* polymerase beffer supplied by Boehringer Mannheim Biochemicals, Indianapolis, IN (BMB).

Mutation analysis. Heteroduplex analysis was carried out as described using MDE gel mix (AT Biochem) with the addition of 15% urea. Samples were denatured for 5 min and allowed to slow cool before fragments were separated by electrophoresis on non-denaturing gels. Products were visualized by ethidium bromide staining. Altered PCR products detected by heteroduplex analysis were directly sequenced using AmpliTaq Cycle sequencing kit (Perkin Elmer). Primers were end-labelled and samples run on 6% polyacrylamide gels. PCR products of the individuals containing the identified stop codon and the 2 bp deletion were cloned into pCR-Script Direct SK(+) cloning vector using pCR-Script Direct SK(+) Directional Cloning Kit (Stratagene) and sequenced using Sequenase Ver. 2.0 (United States Biochemical).

Acknowledgements

We gratefully acknowledge the family members who participated in this study as well as the support of the Hereditary Hemorrhagic Telangiectasia Foundation International. We thank S. Kiousis, S. Santos, A. Thukkani and M. Walsh for technical assistance and L. Deaven for use of the gridded cosmid library. This study was supported by NIH grant HL 49171 to D.A.M., a grant from the Medical Research Council to M.L. and the General Clinical Research Center of the University of Michigan Hospitals.

1. Jahnke, V. Ultrastructure of hereditary telangiectasia. *Arch. Otolaryngol.* **91**, 262–265 (1970).
2. Hashimoto, K. & Pritzker, M. Hereditary hemorrhagic telangiectasia: an electron microscope study. *Oral Surg., Oral Med., Oral Pathol.* **34**, 751–768 (1972).
3. Braverman, I.M., Keh, A. & Jacobson, B.S. Ultrastructure and three-dimensional organization of the telangiectases of hereditary hemorrhagic telangiectasia. *J. invest. Dermatol.* **95**, 422–427 (1990).
4. Menefee, M.G., Flessa, H.C., Glueck, H.I. & Hogg, S.P. Hereditary hemorrhagic telangiectasia: (Osler-Weber-Rendu disease): an electron microscopic study of the vascular lesions before and after therapy with hormones. *Arch. Otolaryngol.* **101**, 246–251 (1975).
5. McDonald, M.T. *et al.* A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33–34. *Nature Genet.* **6**, 197–204 (1994).
6. Shovlin, C. *et al.* A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nature Genet.* **6**, 205–209 (1994).
7. McAllister, K.A. *et al.* Genetic heterogeneity in hereditary hemorrhagic telangiectasia: possible correlation with clinical phenotype. *J. med. Genet.* **31**, 927–932 (1994).
8. Porteous, M.E.M. *et al.* Genetic heterogeneity in hereditary haemorrhagic telangiectasia. *J. med. Genet.* **31**, 925–926 (1994).
9. Heutink, P. *et al.* Linkage of hereditary hemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. *J. med. Genet.* **31**, 933–936 (1994).
10. Greenspan, D.S. *et al.* COL5A1: Fine genetic mapping and exclusion as candidate gene in families with nail-patella syndrome, tiberous sclerosis 1, hereditary hemorrhagic telangiectasia and Ehlers-Danlos syndrome type II. *Genomics* (in the press).
11. Gougos, A. & Letarte, M. Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J. biol. Chem.* **265**, 8361–8364 (1990).
12. Cheifetz, S. *et al.* Endoglin is a component of the transforming growth factor-b receptor system in human endothelial cells. *J. biol. Chem.* **267**, 19027–19030 (1992).
13. Yamashita, H. *et al.* Endoglin forms a heteromeric complex with the signaling receptors for transforming growth factor-β. *J. biol. Chem.* **269**, 1995–2001 (1994).
14. Massagué, J., Attisano, L. & Wrana, J.L. The TGF-β family and its composite receptors. *Trends cell Biol.* **4**, 172–178 (1994).
15. Fernández-Ruiz, E., St-Jacques, S., Bellón, T., Letarte, M. & Bernabeu, C. Assignment of the human endoglin gene (END) to 9q34–qter. *Cytogenet. Cell Genet.* **64**, 204–207 (1993).
16. Pilz, A., Woodward, K., Peters, J., Povey, S. & Abbot, C. Comparative mapping of 38 human chromosome 9 loci in the laboratory mouse. Abstract from the third international workshop on chromosome 9. *Ann. hum. Genet.* **58**, 231–232 (1994).
17. Povey, S. *et al.* Report on the third international workshop on human chromosome 9. *Ann. hum. Genet.* **58**, 177–199 (1994).
18. Leversha, M.A., Carter, N.P. & Ferguson-Smith, M.A. Physical mapping in 9q using fluorescence *in situ* hybridization. Abstract from the third international workshop on chromosome 9. *Ann. hum. Genet.* **58**, 224 (1994).
19. Lastres, P. *et al.* Regulated expression on human macrophages of endoglin, an RGD containing surface antigen. *Eur. J. Immunol.* **22**, 393–397 (1992).
20. Gougos, A. *et al.* Identification of distinct epitopes of endoglin, an RGD-containing glycoprotein of endothelial cells, leukemic cells, and syncytiotrophoblasts. *Int. Immunol.* **4**, 83–92 (1992).
21. St-Jacques, S., Cymerman, U., Pece, N. & Letarte, M. Molecular characterization and *in situ* localization of murine endoglin reveal that it is a transforming growth factor-β binding protein of endothelial and stromal cells. *Endocrinology* **134**, 2645–2657 (1994).
22. Madri, J.A. *et al.* Interactions of matrix components and soluble factors in vascular responses to injury. Modulation of cell phenotype. in *Endothelial cell dysfunctions* (eds Simionescu, N. & Simionescu, M.) (Plenum Press, New York, 1992).
23. Jennings, J.C., Mohan, S., Linkhart, T.A., Widstrom, R. & Baylink, D.J. Comparison of the biological actions of TGF beta-1 and TGF beta-2: differential activity in endothelial cells. *J. cell. Physiol.* **137**, 167–172 (1988).
24. Luscinkas, F.W. & Lawler, J. Integrins as dynamics regulators of vascular function. *FASEB J.* **8**, 929–938 (1994).
25. Wrana, J.L. *et al.* TGF-β signals through a heteromeric protein kinase receptor complex. *Cell* **71**, 1003–1014 (1992).
26. Wrana, J.L., Attisano, A., Wieser, R., Ventura, F., Massagué, J. Mechanisms of activation of the TGF-β receptor. *Nature* **370**, 341–347 (1994).
27. López-Casillas, F. *et al.* Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF-β receptor. *Cell* **67**, 785–795 (1991).
28. Wang, X-F. et al. Expression cloning and characterization of the TGF-β type III receptor. *Cell* **67**, 797–805 (1991).
29. López-Casillas, F., Wrana, J.L. & Massagué, J. Betaglycan presents ligand to the TGF-β signaling receptor. *Cell* **73**, 1435–1444 (1993).
30. Cheifetz, S. *et al.* Distinct transforming growth factor-β (TGF-β) receptor subsets as determinants of cellular responsiveness to three TGF-β isoforms. *J. biol. Chem.* **265**, 20533–20538 (1990).
31. Shull, M.M. *et al.* Targeted disruption of the mouse transforming growth factor-β 1 gene results in multifocal inflammatory disease. *Nature* **359**, 693–699 (1992).
32. Kulkarni, A.B. *et al.* Transforming growth factor β1 null mutation in mice causes excessive inflammatory response and early death. *Proc. natn. Acad. Sci. U.S.A.* **90**, 770–774 (1993).
33. Letterio, J.J. *et al.* Maternal rescue of transforming growth factor-β1 null mice. *Science* **264**, 1936–1938 (1994).
34. Mathew, S. et al. Transforming growth factor gene TGFBR2 maps to human chromosome band 3p22. *Genomics* **20**, 114–115 (1994).
35. Blanchard, M.M., Taillon-Miller, P., Nowotny, P. & Nowotny, V. PCR buffer optimization with uniform temperature regimen to facilitate automation. *PCR Meth. Applic.* **2**, 234–240 (1993).
36. Bellón, T. *et al.* Identification and expression of two forms of the human transforming growth-factor-β binding-protein endoglin with distinct cytoplasmic regions. *Eur. J. Immunol.* **23**, 2340–2345 (1993).