Am. J. Hum. Genet. 61:68–79, 1997

Characterization of Endoglin and Identification of Novel Mutations in Hereditary Hemorrhagic Telangiectasia

Claire L. Shovlin,[1,3,*] J. M. B. Hughes,[3] J. Scott,[3] Christine E. Seidman,[2] and J. G. Seidman[1]

[1]Department of Genetics and Howard Hughes Medical Institute, Harvard Medical School, and [2]Howard Hughes Medical Institute and Department of Medicine, Brigham and Women's Hospital, Boston; and [3]Department of Medicine, Royal Postgraduate Medical School, Hammersmith Hospital, London

Summary

To identify mutations that cause hereditary hemorrhagic telangiectasia (HHT, or Rendu-Osler-Weber syndrome), clinical evaluations and genetic studies were performed on 32 families. Linkage studies in four of eight families indicated an endoglin (ENG) gene mutation. *ENG* sequences of affected members of the four linked families and probands from the 24 small families were screened for mutations, by Southern blot analyses and by cycle sequencing of PCR-amplified DNA. Seven novel mutations were identified in eight families. Two mutations (a termination codon in exon 4 and a large genomic deletion extending 3′ of intron 8) did not produce a stable ENG transcript in lymphocytes. Five other mutations (two donor splice-site mutations and three deletions) produce altered mRNAs that are predicted to encode markedly truncated ENG proteins. Mutations in other families are predicted to lie in *ENG*-regulatory regions or in one of the additional genes that may cause HHT. These data suggest that the molecular mechanism by which *ENG* mutations cause HHT is haploinsufficiency. Furthermore, because the clinical manifestation of disease in these eight families was similar, we hypothesize that phenotypic variation of HHT is not related to a particular *ENG* mutation.

Introduction

Hereditary hemorrhagic telangiectasia (HHT, or Rendu-Osler-Weber syndrome) is an autosomal dominant disorder characterized by aberrant vascular development (Rendu 1896) and has an incidence of ~1/10,000 (Plauchu et al. 1989). Clinical manifestations include mucocutaneous telangiectasia and hemorrhage from nasal and gastrointestinal vascular beds. Nosebleeds (epistaxes) are usually the earliest clinical symptom, mucocutaneous telangiectases develop in approximately one-third of cases at age <30 years, whereas anemia due to gastrointestinal bleeding occurs later in life. Twenty percent of affected individuals (Plauchu et al. 1989) develop pulmonary arteriovenous malformations (PAVMs), which may result in hypoxemia, and paradoxical embolism.

The vascular pathology of HHT may be distinguished histologically from other forms of hereditary telangiectases and vasodilatation (Braverman et al. 1983, 1990; Shovlin and Scott 1996). However, no specific ultrastructural or histochemical features have been observed to account for the vascular pathology. The small HHT mucocutaneous telangiectasia are focal dilatations of postcapillary venules in the dermal upper-horizontal plexus, with abnormal stress fibers in the venule pericytes (Braverman et al. 1990). More complex vascular abnormalities, such as arteriovenous malformations, exhibit direct communication between dilated arterial and venous elements and are devoid of intervening capillary beds. The lesions appear to progress by remodeling (Clark and Clark 1940) via recruitment of mural cells and involvement of further vascular segments.

Molecular-genetic analyses of HHT have identified disease loci on chromosomes 9 and 12, and at least one other *HHT* locus is predicted (Piantanida et al. 1996). The disease gene on chromosome 9q encodes endoglin (ENG; McAllister et al. 1994), which is expressed predominantly on endothelial cells as a heavily glycosylated disulfide-linked dimer that binds TGF-β1 and -β3 (Cheifetz et al. 1992). ENG associates with the TGF-β signaling receptors (Yamashita et al. 1994; Zhang et al. 1996) and modifies TGF-β1 signal transduction (Lastres et al. 1996). The HHT disease gene on chromosome 12 encodes an activin receptor–like kinase (ALK1), a type I receptor of the TGF-β receptor family (Johnson et al. 1996). The relative frequencies of mutations in *ENG*, *ALK1*, and other loci in HHT are not known. Others (Guttmacher et al. 1995; Berg et al. 1996) have suggested that PAVMs are seen most frequently in affected individuals whose disease is caused by *ENG* mutations.

Received December 4, 1996; accepted for publication April 23, 1997.

Address for correspondence and reprints: Dr. Jonathon Seidman, Department of Genetics, Harvard Medical School, 533 Alpert Building, 200 Longwood Avenue, Boston, MA 02115. E-mail: Seidman@Rascal.med.harvard.edu

*Present affiliation: Respiratory Medicine Unit, Department of Medicine (RIE), The University of Edinburgh (RIE), Edinburgh.

0002-9297/97/6101-0012$02.00

To better understand the relationship between *ENG* mutations and the incidence of PAVMs and to elucidate the mechanism by which *ENG* mutations cause HHT, we studied 32 HHT families. *ENG* sequences were screened for defects in probands from 28 families; and in 8 families a causal mutation was identified. Affected individuals from these eight families exhibited comparable clinical features of HHT, indicating little effect of genotype on the variable phenotype of HHT. Furthermore, two of the mutations demonstrated a marked reduction in *ENG* transcription from the mutant allele, thereby providing information on the molecular mechanism of HHT pathogenesis.

Material and Methods

Clinical Evaluations

All studies were performed with the approval of local committees for research on human subjects. Thirty-two study families of northern-European descent were identified when one affected member was referred to a tertiary-care hospital (Hammersmith Hospital, London) for therapy. Referrals were usually for PAVMs; five probands were referred for other HHT manifestations. Clinical evaluations, including history, physical examination, and portable oximetry to screen for PAVMs (Chilvers et al. 1990), were performed on all available family members of each proband, by C.L.S. Additional medical history was obtained from patient records. Diagnostic criteria for disease status have been described elsewhere (Shovlin et al. 1994*a*).

Genetic Analyses

Genomic DNA was derived from peripheral blood, as described elsewhere (Shovlin et al. 1994*a*). mRNA was prepared from Epstein-Barr virus (EBV)–transformed lymphoblastoid lines (Ausubel et al. 1989) or from buffy-coat leukocyte preparations from 10 ml of fresh blood by use of RNAzol^{TM} (Cinna Biotecx). *ENG* cDNA was transcribed from these templates by Moloney murine leukemia virus reverse transcriptase (Gibco BRL) (Ausubel et al. 1989) and by SuperscriptII^{TM} (Gibco BRL), by use of an oligo $(dT)_{14}$ primer, *ENG*-specific primers, and, for family I, RKdt (ccgcgaattcggtacc$[dT]_{17}$).

Genomic Characterization of ENG

Four clones containing *ENG* were identified from a pAd10sacBII ("P1") library of human genomic DNA (DuPont-Merck), as described elsewhere (Shovlin 1996). Clone 22 contained the entire coding sequence of *ENG* and the adjacent adenylate kinase 1 gene (*AK1*). The clone was characterized further by Southern blots using either PCR-amplified *ENG* exons (Saiki et al. 1988) labeled with ^{32}P αdCTP (Du Pont/NEN) by PCR or random hexanucleotide labeling (Boehringer-Mannheim) (Ausubel et al. 1989). Clone 104 contained the 3′ exons of *ENG* and a novel CA repeat described below. P1 rearrangements were excluded by both analysis of genomic DNA and agreement with the *AK1* fragment sizes predicted by the GenBank sequence.

Linkage Analyses

Six polymorphic loci, ordered *D9S60/D9S315/AK1/*ENG*/SPTAN/D9S61*, span 2 cM on chromosome 9q (Shovlin et al. 1995*a*). Linkage studies were performed as described elsewhere (Shovlin et al. 1994*a*), by use of published sequences to amplify *D9S60, D9S315, D9S61* (GenBank), and *AK1* (Puffenberger and Francomano 1991). Two novel CA repeats associated with P1 clones containing *ENG* and *SPTAN* were identified in P1 clones by use of Quick Light (Life-Codes) probes and protocols. The primer sequences, derived by vectorette-mediated PCR (Edwards et al. 1991) of P1 clones, were *ENG*-CA 5′atctcagctactgtgagctgaggca and 5′aagccccacggacagcactgga and *SPTAN*-CA 5′accgtgttagccaggatggtcttga and 5′tatccccagcactttgcatggtattt. *ENG*-CA has four alleles, size 215–221 bp and PIC .35; *SPTAN*-CA amplified three alleles, size 185–193 bp and PIC .26.

ENG *Primer Sequences*

All *ENG* cDNA primers (C1–C19) were derived from the published GenBank (NIH sequence databases) *ENG* cDNA sequence (GenBank accession number X72012) and are described in the legend to figure 1. Intronic primers (E1 forward–E14 reverse; fig. 2) were derived either from the work of McAllister et al. (1994) or by sequencing templates derived from the P1 clones, including subclones, long-PCR products, inverse-PCR templates (Ausubel et al. 1989), vectorette ("bubble")-mediated PCR (Edwards et al. 1991), and *alu*-exon PCR products, by use of *alu* primers complementary to the reverse 5′ and forward 3′ of the Alu consensus sequence: Tagle3 (gatcgcgccactgcactcc) and alu450 (aaagtgctgggattacagg).

ENG-*Mutation Screens*

Southern blot analyses were performed by use of genomic DNA digested with *Eco*RI, *Hin*dIII, and *Taq*I and were probed either with ^{32}P-labeled *ENG* 18A cDNA (Gougos and Letarte 1990) (kindly supplied by Michelle Letarte) or with a ^{32}P-labeled C1/C12 PCR product. PCR products spanning exons or groups of exons were screened for mutations, by manual cycle sequencing. *ENG* sequences were amplified from genomic DNA or cDNA templates by use of a final 25-μl PCR reaction composed of 0.6 units of r*Tth* DNA polymerase XL (Perkin Elmer), 85 mM potassium acetate, 25 mM tricine pH 8.7, 8% glycerol, 1% dimethylsulfoxide (XL buffer II; Perkin Elmer), 1 mM magnesium acetate, and

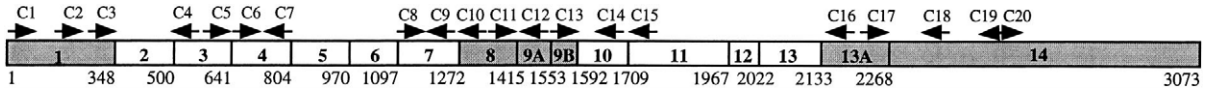

Figure 1 *ENG* cDNA structure and the exons that encode it. The location of oligonucleotide primers used to amplify *ENG* sequences are indicated (*arrows*). The number and order of exons are as described by McAllister et al. (1994) and our unpublished data; for each exon, the last nucleotide is numbered according to GenBank accession number X72012. The variably retained intron (13A) between exons 13 and 14 (*shaded*) encodes the ENG-S form (Bellon et al. 1993). Primers are as follows (start nucleotide"c'" [if sequence is complementary] length [in nucleotides]): C1 (63;26), C2 (257;24), C3 (282;23), C4 (534;c′;28), C5 (604;28), C6 (724;25), C7 (775;c′;24), C8 (1149;24), C9 (1225;c′;25), C10 (1342;c′;23), C11 (1328;24), C12 (1441;c′;25), C13 (1552;24), C14 (1650;c′;25), C15 (1759;c′;25), C16 (2178;c′;26), C17 (2213,28), C18 (2463;c′;22), C19 (2591;c′;25), and C20 (2567;23).

0.2 mM each dNTP. The enzyme was added in a manual "hot start" after the first denaturation step (20 s at 94°C). The subsequent 30 cycles were 94°C for 10 s and 68°C for 1–5 min (according to product length). For amplification of cDNA reverse transcribed from *ENG* mRNA (reverse-transcription PCR [RT-PCR]), the reaction included a first cycle of 94°C for 1 min and 68°C for 10 min. Exons 1 and 3 required an annealing step of 58°C, including, for exon 1, a step-down procedure from an initial temperature of 68°C (Hecker and Roux 1996). The quality of PCR-amplified products was assessed by agarose-gel electrophoresis. The Cyclist™ *Taq* DNA sequencing kit (Stratagene) was used to sequence 2 μl of PCR products directly: gel purification was required only for early *Taq* PCR products in unnested sequencing reactions or to separate wild-type and mutant alleles, as an alternative to T-vector cloning (pGEM-SZ; Promega). Sequence analyses were performed by use of the GCG software package (program manual for the Wisconsin Package, version 8).

Results

Since HHT can be caused by mutations in more than one gene, linkage analyses were initially performed in the study families whose size was sufficient to provide information. Linkage studies excluded an *ENG* mutation in four families (families S, T, V, and Y; LOD score <-1.32); HHT in four families (A, B, F, and W; LOD score >1.32) was linked to the *ENG* locus on chromosome 9 (table 1). DNA samples from probands of the 4 families whose HHT was linked to chromosome 9 and from probands from the remaining 24 small HHT families were studied further.

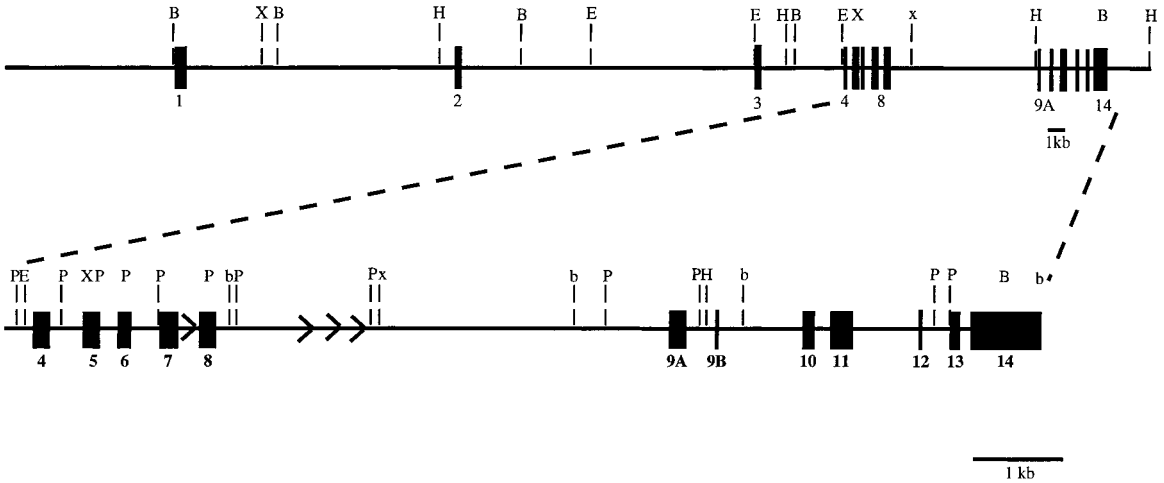

Figure 2 Schematic of *ENG* spanning 35 kb, and expanded view of the 11 kb spanning exons 4–14. Restriction-enzyme sites are indicated as follows: B = *Bam*HI; X = *Xho*I; H = *Hin*dIII; E = *Eco*RI; x = *Xba*I; P = *Pst*I; and b = *Bgl*I. Arrows represent the position and orientation of the *alu* elements flanking exon 8. Exons were amplified with intronic primers E2 forward and reverse, E4 forward and reverse, E5 forward, E6 forward and reverse, E7 forward and reverse, E10 reverse, E11 forward and reverse, E12 forward and reverse, and E13 forward, described by McAllister et al. (1994). In addition intronic primers E1 forward (ccagccccttctctaaggaa), E1 reverse (tccccaccctgggtccctggaca), E3 forward (aacctatacaaatctgact), E3 reverse (tgacagtaggacttcccat), E5 reverse (tctcgggntggggactagtgtca), E8 forward (gccgcctggcctgcctctgcta), E8 reverse (tgagcnngaggggcaggagtt), E9 forward (aatggctgtgacttgggacccctg), E9A reverse (gacaacagctggtcctgatac), E9B reverse (aggagtttcccgaggcctgctccca), and E14 reverse (aatgtcactgcccttctcccagc) were used. Exons 9A, 9B, and 10 were amplified by use of E9 forward/E10 reverse (or, for exon 10, by use of C13/E10 reverse), and exons 13, 13A, and 14 were amplified by use of E13 forward/E14 reverse.

ENG *Structure and Transcripts*

As a first step toward identification of HHT mutations, the *ENG* structure (fig. 2) and RNA transcripts were characterized. Distances between exons were assessed by use of long PCR (only the first and second introns were not spanned by this method) and were refined by restriction mapping of P1 clone 22. *ENG* spans ~35 kb and is organized into 15 exons. *ENG* sequences previously assigned (McCallister et al. 1994) to exons 8 and 9 were found to be organized into three exons. Exon 8 contains nucleotides 1391–1415. Exon 9 sequences are contained in two exons, here designated "9A" and "9B." Exon 9B contains only 39 bp. Comparison of *ENG* mRNA transcripts and genomic structure demonstrated that the long (L-ENG) and short (S-ENG) forms of ENG protein (Bellon et al. 1993) result from alternative splicing of the final intron between exons 13 and 14 (designated "13A").

Two polymorphisms were mapped during these studies. The RFLPs in *AK1* (Puffenberger and Francomano 1991) were identified on the same *Eco*RI fragment as was *ENG* exon 1. Furthermore, a novel CA-dinucleotide repeat was located <60 kb from *ENG* exon 14.

ENG *Mutations*

ENG was analyzed in probands from 28 HHT families, by three independent methods: Southern blot analyses, genomic amplification and sequencing of each exon, and reverse transcription of *ENG* mRNA and PCR amplification of *ENG* cDNA (rt-PCR) for sequence analysis (Methods). Novel mutations were identified in eight HHT families (fig. 3). These defects included four deletions and three point mutations, one shared by two families.

Southern blot analyses of the proband in family R (fig. 4*Ai*) demonstrated a mutant allele that deleted ~1.5 kb of DNA, spanning exon 8 and the flanking intronic sequences. The introns flanking exon 8 were found to contain *alu* sequences (fig. 2, *arrows*), the site and orientation of which could account for the deletion by homologous recombination, as has been described for other sequences containing *alu* repeats (Shovlin et al. 1994*b*). The deletion in family R results in a mutant *ENG* transcript (fig. 5*Aii*) that lacks exon 8 and, because of a frameshift, predicts the incorporation of 25 novel amino acids before a TGA stop codon is reached in exon 9A.

An *ENG* deletion was also found in family I. Southern blot analyses identified a large deletion extending from intron 8, deleting exons 9A–14 (fig. 4*B*). Although the 3′ boundary of the deletion has not been defined, it does not extend to an *ENG* CA repeat located <60 kb distant, since affected family members demonstrate heterozygosity at this polymorphism. In addition to protein-encoding sequences, this mutation deletes the *ENG* polyadenylation addition site, thereby predicting that no stable mRNA species would result. RT-PCR (data not shown) of samples from affected family members failed to amplify product from the mutant allele; only wild-type *ENG* cDNA was detected; that is, *ENG* mRNA was reversed transcribed (see Material and Methods), PCR amplified, and cloned into a plasmid vector, and the clones were characterized by nucleotide-sequence analysis; no mutant *ENG* cDNA clones were detected, and only wild-type *ENG* cDNA clones were found.

The deletion mutation in family B was first detected as a short cDNA species: sequencing of this product demonstrated a deletion of 152 nucleotides encoded in exon 2 (fig. 4*Ci*). This deletion produces a frameshift that predicts the incorporation of 27 novel amino acids before a TGA stop codon is reached in exon 3. Southern blot analyses of DNAs (fig. 4*Cii*) from affected but not from unaffected members of family B demonstrated reduced hybridization to an *ENG* exon 2 probe, compared with an *ADA* probe (Shovlin et al. 1994*b*) derived from chromosome 20, indicating that the exon skip occurred because of a genomic deletion spanning exon 2.

A 21-bp deletion in exon 5 (fig. 4*D*) was detected in affected members of family Q. This in-frame deletion removes cDNA nucleotides 855–875, which encode amino acids Arg-Thr-Leu-Glu-Trp-Arg-Pro. Four of these residues are highly conserved *ENG* transcripts sequenced from man, mouse, and pig (fig. 6).

The HHT disease haplotype (loci *D9S315/ENG/SPTAN*, <2 cM from *ENG*) segregating in families F

Table 1

LOD Scores Reflecting Linkage between *ENG* and HHT in Eight Families

		LOD Score at θ =		
Family	Marker[a]	.00	.01	.05
A	*D9S60*	3.64	3.64	3.54
F	*D9S61*	1.79	1.75	1.61
W	*D9S61*	3.29	3.24	3.02
B	*SPTAN*	1.46	1.43	1.33
S	*D9S60*	−3.79	−1.75	−.89
T	*D9S61*	−4.13	−2.06	−1.19
V	*D9S315*	−2.58	−1.4	−.23
Y	*D9S61*	−6.12	−3.95	−1.98

Note.—HHT is linked to *ENG* in families A, F, W, and B and is unlinked to it in families S, T, V, and Y. The linkage relationships between HHT and *ENG* in families A, F, and T have been described elsewhere (Shovlin et al. 1994*a*).

[a] Markers were selected because they were informative in the families. All markers are <2 cM of *ENG* (Shovlin et al. 1995*a*; authors' unpublished data).

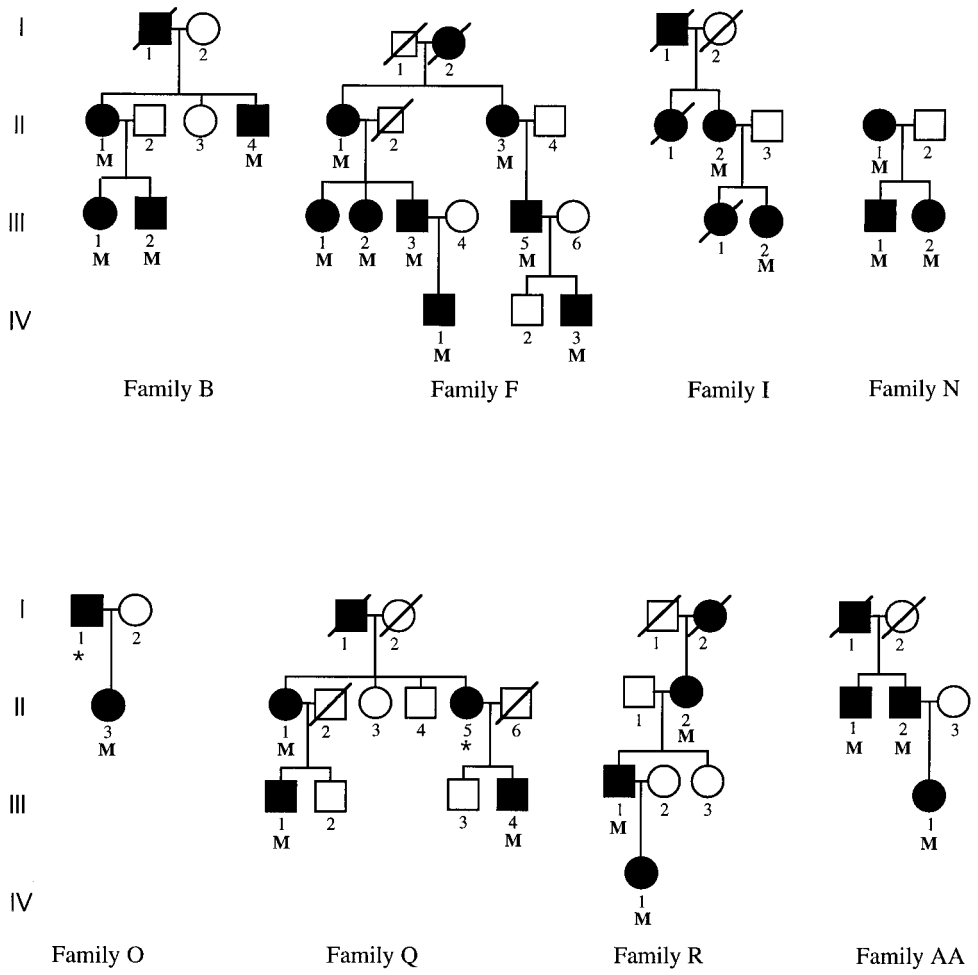

Figure 3 Pedigrees of HHT families. Individuals are designated by sex, disease status (blackened symbols denote affected individuals, and unblackened symbols denote unaffected individuals), and pedigree numbers. Deceased individuals are denoted by diagonal slashes. M = presence of *ENG* mutation causing HHT. Asterisks (*) denote living individuals who did not participate in genetic studies.

and N was the same, suggesting that these two families, which are of similar southern-English ancestry and currently reside ~100 miles apart, share a founding mutation that arose in a shared ancestor. This hypothesis was confirmed by sequence analyses of *ENG* in the probands of families F and N, which demonstrated that these individuals shared the same HHT-causing mutation. A G→A substitution was identified in the donor splice signal of intron 8 (fig. 4*A*) in affected individuals from both family F and family N. Although this gene defect might generate several ENG transcripts because of exon skipping and/or use of cryptic splice sites, only one abnormal cDNA species was amplified by PCR. The mutant, truncated *ENG* transcript lacked 24 nucleotides (1391–1415) that encode eight poorly conserved amino acids (fig. 6). These nucleotides, encoded in exon 8 (fig. 1), are deleted from the mutant cDNA sequence by activation of a cryptic splice site within exon 8, located 24 nucleotides upstream from the natural splice site. This cryptic splice site shares only 5 of the 11 residues of the optimal donor splice-site consensus sequence, C/AAG-GTT/ATGTAT (Zhang et al. 1995). However, an adjacent polypurine tract comprising nucleotides 1397–1406 may enhance the use of this splicing site (Laviguer et al. 1993; Manley and Tacke 1996) and thereby account for the absence of transcripts that lack exon 8 sequences.

A donor splice-site mutation was also detected in family O. The A→G substitution in intron 3 was identified by genomic sequencing (data not shown). Although there is a potential cryptic donor splice site upstream (nucleotides 560–571) that contains 7 of 11 residues of the optimal donor splice-site consensus sequence, this did not appear to be active. RT-PCR using primers C2 and C7 amplified only one mutant transcript, which lacked exon 3 (fig. 4*B*). This mutant transcript predicts the loss of 47 amino acids from the translated protein.

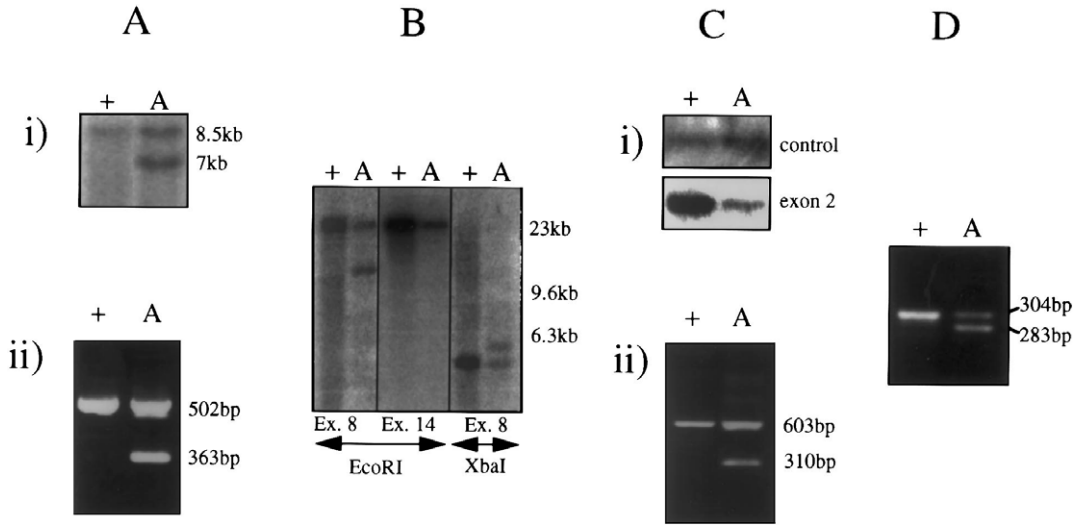

Figure 4 *ENG* deletions causing HHT in four families. DNA samples from affected individuals (lanes A) and unaffected individuals (lanes +) were analyzed. *Ai, Hin*dIII digest of genomic DNA from individuals in family R. A Southern blot filter probed with PCR product which spans exon 7. The shorter *Hin*dIII fragment is due to a deletion, which also created a novel *Pst*I fusion fragment of 2 kb, by the apposition of the truncated fragment containing exons 7 and 8 (5′) to a fragment normally 3′ of exon 8. *Aii,* C8/C12 RT-PCR (using primers C8 and C12), demonstrating the wild-type product (502 bp) and a shorter product (363 bp) that lacked all exon 8 sequence. *B,* Genomic Southern blot of samples from family I. *Eco*RI-digested DNA was probed with PCR products containing exon 8 (E8 forward/E8 reverse) and exon 14 (E13 forward/E14 reverse). Although these exons lie on the same *Eco*RI fragment, in DNA from affected individuals a small fragment detected by the exon 8 probe was not detected by the exon 14 probe, indicating that the deletion includes exon 14. The larger *Xba*I fragment defines the 5′ extent of the deletion, since intron 8 normally contains an *Xba*I site (data not shown). *Ci,* Genomic Southern blot of samples from family B. DNA samples were digested with *Hin*dIII, and the filter was probed sequentially with PCR products containing a part of the 5′ region of *ADA* on chromosome 20 (Shovlin et al. 1994*b*) and with exon 2 (E2 forward and reverse). Quantitative assessment of these and additional data not illustrated, by use of a Molecular Dynamics Phosphorimager™ (model 4255) and the ImageQuant™ program, confirmed that the *ENG* exon 2 bands from affected individuals were half the predicted intensity, compared with those in equally loaded wild-type controls. *Cii,* RT-PCR (using primers C2 and C7), which amplified the wild-type product (603 bp) and a shorter fragment (310 bp) that lacked exon 2 nucleotides. *D,* Exon 5 amplified from genomic DNA in samples from family Q (E5 forward and reverse). The 283-bp product results from a 21-bp genomic deletion.

Affected members of family AA were heterozygous for a C→T substitution at residue 792 in exon 4 (fig. 4*A*). This mutation changes the normal CGA codon (encoding arginine) to a TGA stop codon. mRNAs derived from either peripheral blood lymphocytes or EBV-transformed lymphoblast cell lines derived from two affected family members were amplified by RT-PCR using poly(dT)$_{14}$ and primer C14 (fig. 1). The amplified PCR product was subjected to sequence analysis to determine the ratio of mutant and wild-type cDNA. In three independent analyses, only wild-type sequence was detected. Absence of the mutated transcript may reflect early degradation of mRNAs containing nonsense codons (Belgrader and Maquat 1994), especially because residues 758–771 (tgctgagctgaatg) have homology to a *Saccharomyces cerevisiae* sequence that is responsible for premature mRNA degradation (Zhang et al. 1995).

Only one *ENG* defect was identified in affected members from these eight families. There was complete concordance between the clinical status of family members and the presence of each *ENG* defect. More than 100 normal chromosomes were screened for each of these seven *ENG* defects; no defects were detected. We conclude that these *ENG* defects caused HHT in each family.

Two additional sequence variants were identified. These appear to be rare polymorphisms in *ENG,* rather than disease-causing mutations. Members of one family were heterozygous for the transition C295T, which alters a nonconserved codon, in exon 1, from ACG (threonine) to ATG (methionine). However, this allele did not segregate with disease. In another family, C401T was present only in an affected father and daughter but did not predict an amino acid substitution. Two further silent nucleotide variations were found to be polymorphic in the normal population: G488A in exon 2 (in 11% of the control population) and C1310T in exon 8 (in 5% of the control population).

Southern blot analyses and sequence analyses of *ENG* exons and cDNAs failed to identify *ENG* mutations in two families (A and W), although, in each of them, genetic studies indicated linkage between HHT and the *ENG* locus (LOD scores >3, $\theta = 0$). We suspect that

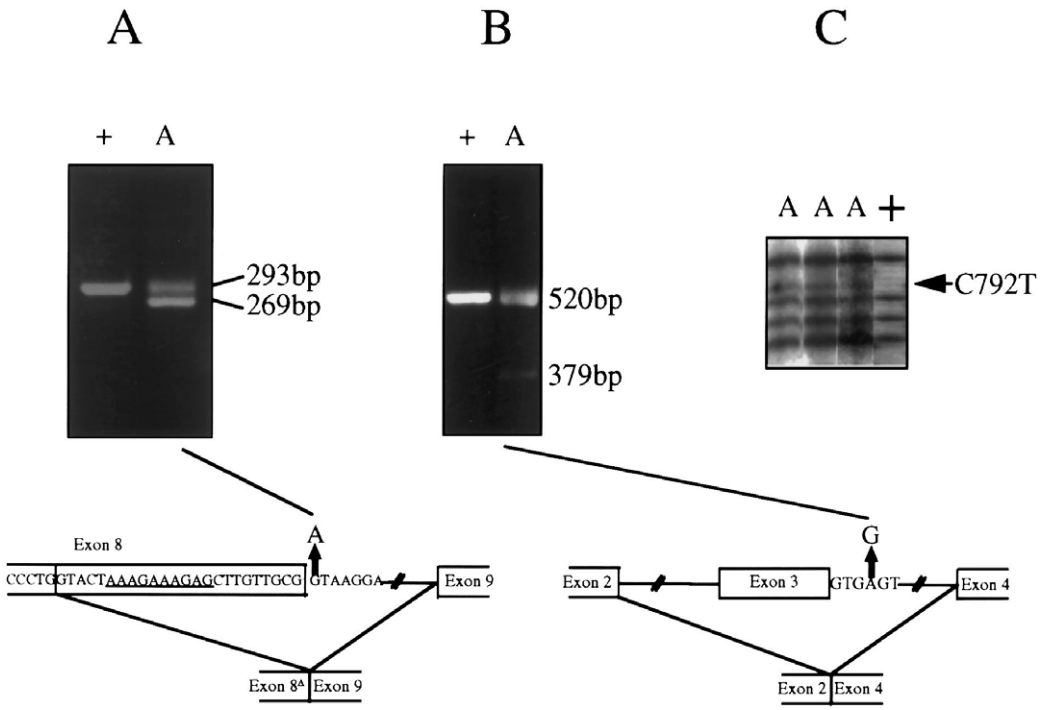

Figure 5 *ENG* point mutations causing HHT in four families. *A, ENG* transcripts, amplified by use of primers C8 and C12 (fig. 1), from affected members of families F and N, revealing the normal sized product (293 bp) and a truncated product (269 bp) resulting from a G→A splice-site mutation and use of a cryptic splice site. The genomic mutation is indicated, and the polypurine tract is underlined. *B,* A→G splice site mutation identified in family O, which produces a truncated transcript that lacks exon 3. RT-PCR (using primers C2 and C7) amplified the wild-type 520-bp product and the mutant 379-bp product. C, T lane of dideoxy sequence reactions, for DNA samples from three affected members of family AA and for one control sample. Note the absence of the band in the T lane (residue 792) in the control sample.

an undetected 5′ deletion, possibly involving ENG promoter sequences, may account for HHT in family W, since affected members have only one allele at each *AK1* polymorphism. Large deletions that might mutate important 3′ *ENG* sequences appear to be an unlikely cause of HHT in either family, since affected members from both families were heterozygous for the *ENG* CA repeat (data not shown).

Clinical Data

The clinical manifestations of HHT (table 2) that are associated with these novel *ENG* mutations were com-

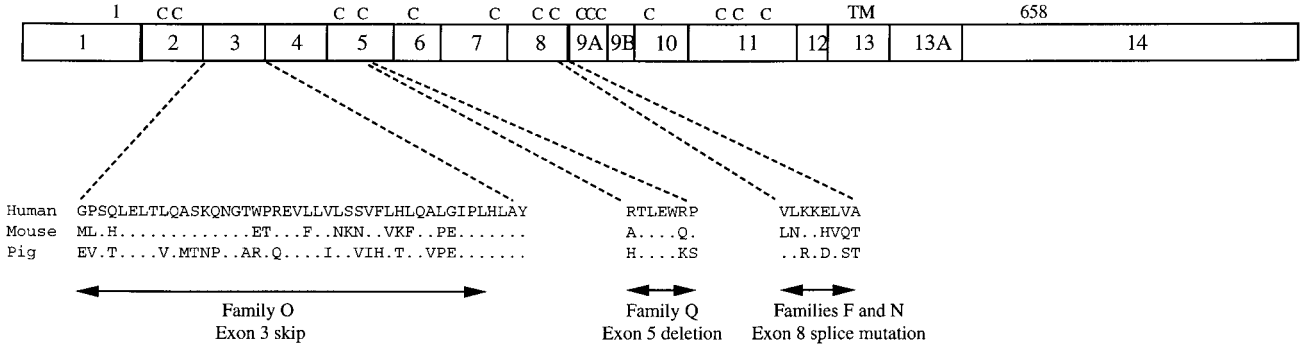

Figure 6 Human, mouse, and pig *ENG* amino acid sequences in the vicinity of HHT mutations (the sites of the stop codons in exons 3 and 9A, which predict truncated proteins, are not illustrated). The human *ENG* cDNA is illustrated as in figure 1 and defines the sites of encoded cysteine residues implicated in dimerization (C), as well as the position of the transmembrane domain (TM).

pared. Telangiectases and nosebleeds, often commencing in childhood, were common manifestations of disease in the 30 affected individuals from these eight HHT families. Seven individuals from five families had documented gastrointestinal involvement. Twenty-one individuals from eight families had PAVMs. Although a larger series (Shovlin et al. 1995*b*) suggested an increased incidence of PAVMs in female patients, the presence of PAVMs in affected men or women did not statistically differ in this study group. Three women, from families F, I, and O, had pregnancy-related deterioration in the pulmonary vascular bed. Parameters analyzed by analysis of variance (computed by use of Statview^{TM} MII) did not demonstrate differences—with respect to age of presentations, subjective severity of nose bleeds or telangiectasia, or PAVMs—in HHT caused by the seven different *ENG* mutations.

The incidence of PAVMs in individuals whose disease was caused by *ENG* mutations and in individuals whose disease was not caused by *ENG* mutations was assessed (table 2 and data not shown). Forty-one percent (23/56) of HHT patients with *ENG* mutations had PAVMs, whereas a significantly smaller fraction, 14% (5/35), of HHT patients in whom linkage analyses indicated non-*ENG* mutations had PAVMs ($P < .01$).

Discussion

We have characterized the structure of *ENG,* defined the alternative splice variants that result in L-ENG and S-ENG, and identified seven novel mutations that cause HHT in eight families. Two point mutations result in aberrant splicing. A third, which would predict premature termination in exon 4, was not detected in cDNA preparations. Four deletions were also identified, and at least one of these appears to function as a null allele.

These seven mutations have not been described elsewhere, and we did not detect in the 32 HHT study families any of the ENG defects reported elsewhere (McAllister et al. 1994, 1995). A greater diversity of different types of *ENG* mutations were identified in the present study than in previous studies, presumably because of the comprehensive screening methods that we applied, rather than because of differences in patient populations. cDNA analyses proved to be the most informative approach and identified five of the seven mutations; three of seven mutations would have been entirely missed without cDNA screening or without careful analysis for genomic deletions. Despite genomic and cDNA analyses, mutations were defined in only two of four families with HHT linked to *ENG*. Families with HHT linked to *ENG* in which mutations were not detected may have defects either in regulatory regions of *ENG* or in intronic sequences. We have demonstrated *ENG* mutations in at least one-quarter of the HHT probands studied here and suggest that we could have missed *ENG* mutations in some families. Taken together, these data suggest that *ENG* mutations may cause as much as 50% of HHT.

Previous studies have suggested that PAVMs frequently occurred in HHT patients whose disease was caused by *ENG* mutations (Guttmacher et al. 1995). In the present study, HHT patients with PAVMs were identified in 19 unrelated families in which no *ENG* mutations were detected, including affected members of 4 families whose disease did not segregate with the *ENG* locus (table 1). Among the HHT patients studied here, the incidence of PAVMs was higher in those whose disease was caused by *ENG* mutations than in those whose disease was caused by mutations in other HHT disease genes.

Characterization of HHT-causing mutations provides insights into the molecular mechanism by which *ENG* mutations cause HHT, and it may provide novel approaches for therapeutic development. ENG is expressed as a disulfide-linked dimer on endothelial cells. Previously reported mutations predicted defects in the peptide's extracellular domain, and a dominant negative mechanism has been proposed (McAllister et al. 1995) to account for the pathogenesis of HHT. ENG deficiency has also been postulated to cause HHT, via a somatic, second event (a "two-hit" mechanism; McAllister et al. 1994). The data presented here demonstrate that mutations that block *ENG* mRNA production can also cause HHT. We conclude that haploinsufficiency of ENG can cause HHT.

The lack of correlation between genotype and clinical manifestations of HHT supports the haploinsufficiency model. If a dominant-negative mechanism accounted for HHT, different mutations might be expected to be associated with discrete clinical phenotypes. In contrast, if HHT mutations result in ENG haploinsufficiency, it would be surprising if clinical features were mutation specific. In our series, the clinical phenotypes (table 2) caused by seven mutations were not statistically different. Furthermore, the greatest trend toward differences in visceral involvement was observed between families F and N, which share the identical *ENG* mutation.

How does the loss of ENG protein result in HHT? The majority of vascular beds in which ENG is expressed (Cheifetz et al. 1992) appear to develop normally in HHT patients, implying that the quantitative requirement for ENG varies spatially and, perhaps, temporally. ENG binds TGF-β1 and -β3 (Cheifetz et al. 1992) and can modify a subset of cellular responses to TGF-β1, including the inhibition of cellular proliferation and the stimulation of extracellular adhesion molecules (Lastres et al. 1996). There is a wide range of known roles for the ENG ligands, which in the vasculature include development (Dickson et al. 1995; Kaartinen et al. 1995), angiogenesis (Koh et al. 1995), and vascular repair (Ma-

Table 2

Clinical Spectrum of HHT Caused by Seven *ENG* Mutations

Family and Individual[a]	Age[b] (years)	HHT Profile[c]				PAVM Profile[c]			Comments
		Age at Onset of Epistaxis[b] (years)	Severity of Nosebleeds	Severity of Facial Telangiectasia	Documented Visceral Involvement	Age at Diagnosis or at Documented Onset of Complications[b] (years)	No. of Lesions	Minimum Pretreatment Oxygen Saturation[d]	
B:									
I-1 (M)	Deceased	...	...	...	Pulmonary	30	...	...	One treatment by surgery
II-1 (F)	50	<10	+	+	Pulmonary, gastrointestinal	30	10	88	Three treatments by embolization
II-4 (M)	50	<10	+++	++					
III-1 (F)	20	<10	++		Pulmonary	20	2	97	Under investigation
III-2 (M)	10	<10	++	++					
F:									
II-1 (F)	70	<10	++	++	Pulmonary	...	...	77	Declined treatment
II-2 (F)	70	<10	++	++					
III-1 (F)	40	<10	+	+					
III-2 (F)	40	<10	++	++	Pulmonary	20	25	60	Pregnancy deterioration noted, six treatments by embolization
III-3 (M)	30	...	...	++					
III-5 (M)	40	<10		++					
IV-1 (M)	10	<10	+++	+	Pulmonary, cerebral	<10	25	74	Three treatments by embolization
IV-3 (M)	10	<10	++	+					
I:									
II-1 (F)		10	...	...	Pulmonary	10	...	...	Died before treatment
II-2 (F)	50	...	Nil	+					
III-1 (F)		10	...	...	Pulmonary	10	...	...	Died before treatment
III-2 (F)	30	20	+	+	Pulmonary	20	3	96	Pregnancy deterioration noted, two treatments by embolization

N:									
I-1 (F)	60	...	+++	++	Pulmonary, gastrointestinal	40	10	...	One treatment by surgery
II-1 (M)	30	<10	++	++	Gastrointestinal				
II-2 (F)	30	10	++	++					
O:									
II-1 (F)	20	<10	+++	++	Pulmonary, cerebral, hepatic	<10	25	52	Pregnancy deterioration noted, six treatments by embolization
Q:									
II-1 (F)	60	30		++	Gastrointestinal				
III-1 (M)	30	10		++					
III-4 (M)	20	10	++	+	Pulmonary	10	3	93	One treatment by embolization
R:									
II-1 (F)	60	10	++	++					
III-1 (M)	30	<10	++	++					
IV-1 (F)	<10	<10	++	++	Pulmonary, gastrointestinal	<10	1	83	One treatment by embolization
AA:									
II-1 (M)	60	40	+	+	Pulmonary, cerebral	40	2	82	Two treatments by embolization
II-2 (M)	60	<10	+++	+	Pulmonary, gastrointestinal	40	4	95	One treatment by embolization
III-1 (F)	30	<30	+	+					

[a] Designations of individuals correspond to those given in figure 3.
[b] By decade.
[c] + = Mild; ++ = moderate; and +++ = severe.
[d] Recorded either while individual was in erect position or after exercise.

dri et al. 1992). Future studies of the consequences of *ENG* mutations on these processes, as well the development of an animal model for HHT, may provide crucial insights into the mechanisms by which TGF-β signaling maintains the structural integrity of the vascular bed.

Acknowledgments

This work was supported by a Wellcome Trust Training Fellowship (to C.L.S.) and by the Howard Hughes Medical Institute (support to J.G.S. and C.E.S.). We thank Bernadette Farren and Evelyn Fuggle for assistance in obtaining clinical data and thank Carin Mordin and Richard Magee for pulmonary-function testing. We are grateful to Brian Harding and Mohammed Miri for transformation of lymphoblastoid cell lines and to Michelle Letarte for the provision of the *ENG* 18A cDNA clone. We would also like to express our thanks to the families, for their patience, insight, and cooperation in these studies.

References

Belgrader P, Maquat LE (1994) Nonsense but not missense mutations can decrease the abundance of nuclear mRNA for the mouse major urinary protein, while both types of mutation can facilitate exon skipping. Mol Cell Biol 14: 6326–6336

Bellon T, Corbi A, Lastres P, Cales C, Cebrian M, Vera S, Cheifetz S, et al (1993) Identification and expression of two forms of the human transforming growth factor-β-binding protein endoglin with distinct cytoplasmic domains. Eur J Immunol 23:2340–2345

Berg JN, Guttmacher AE, Marchuk DA, Porteous MEM (1996) Clinical heterogeneity in hereditary haemorrhagic telangiectasia—are pulmonary arteriovenous malformations more common in families linked to endoglin? J Med Genet 33:256–257

Braverman IM, Keh A, Jacobson BS (1990) Ultrastructure and three-dimensional organization of the telangiectases of hereditary hemorrhagic telangiectasia. J Invest Dermatol 95: 422–427

Braverman IM, Keh-Yen A (1983) Ultrastructure and three-dimensional reconstruction of several macular and papular telangiectases. J Invest Dermatol 81:489–497

Cheifetz S, Bellon T, Cales C, Vera S, Bernabeu C, Massague J, Letarte M (1992) Endoglin is a component of the transforming growth factor-β receptor system in human endothelial cells. J Biol Chem 267:19027–19030

Chilvers ER, Whyte MKB, Jackson JE, Allison DJ, Hughes JMB (1990) Effect of percutaneous catheter embolisation on pulmonary function, right-to-left shunt, and arterial oxygenation in patients with pulmonary arteriovenous malformations. Am Rev Respir Med 142:420–425

Clark ER, Clark EL (1940) Microscopic observations on the extra-endothelial cells of living mammalian blood vessels. Am J Anat 66:2–49

Dickson MC, Martin JS, Cousins FM, Kulkarni AB, Karlsson S, Akhurst RJ (1995) Defective haematopoeisis and vasculogenesis in transforming growth factor-β1 knock out mice. Development 121:1845–1854

Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756

Gougos A, Letarte M (1990) Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. J Biol Chem 265:8361–8364

Guttmacher AE, Marchuk DA, White RI (1995) Hereditary hemorrhagic telangiectasia. N Engl J Med 333:918–924

Hecker KH, Roux KH (1996) High and low annealing temperatures increase both specificity and yield in touchdown and step-down PCR. BioTechniques 20:478–485

Johnson DW, Berg JN, Baldwin MA, Gallione CJ, Marodel I, Yoon S-J, Stenzel TT, et al (1996) Mutations in the activin receptor-like kinase 1 gene in hereditary hemorrhagic telangiectasia type 2. Nat Genet 13:189–195

Kaartinen V, Voncken JW, Schuler C, Warburton D, Bu D, Heisterkamp N, Groffen J (1995) Abnormal lung development and cleft palate in mice lacking TGF-β3 indicates defects of epithelial-mesenchymal interaction. Nat Genet 11: 415–421

Koh GY, Kim S-J, Klug MG, Park K, Soonpaa MH, Field LJ (1995) Targeted expression of transforming growth factor-β1 in intracardiac grafts promotes vascular endothelial cell DNA synthesis. J Clin Invest 95:114–121

Lastres P, Letamendia A, Zhang H, Rius C, Almendro N, Raab U, Lopez LA, et al (1996) Endoglin modulates cellular responses to TGF-β1. J Cell Biol 133:1109–1121

Laviguer A, La Branche H, Kornblihtt AR, Chabot B (1993) A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. Genes Dev 7:2405–2417

Madri JA, Merwin JR, Bell L, Basson CT, Kocher O, Perlmutter R, Prinz C (1992) Interactions of matrix components and soluble factors in vascular responses to injury. In: Simionescu N, Simionescu M (eds) Endothelial cell dysfunctions. Plenum Press, New York

Manley JL, Tacke R (1996) SR proteins and splicing control. Genes Dev 10:1569–1579

McAllister KA, Baldwin MA, Thukkani AK, Gallione CJ, Berg JN, Porteous ME, Guttmacher AE, et al (1995) Six novel endoglin mutations in the endoglin gene in hereditary hemorrhagic telangiectasia type I suggest a dominant-negative effect of receptor function. Hum Mol Genet 4:1983–1985

McAllister KA, Grogg KM, Johnson DW, Gallione CJ, Baldwin MA, Jackson CE, Helmbold EA, et al (1994) Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. Nat Genet 8: 345–351

Piantanida M, Buscarini E, Dellavecchia C, Minelli A, Rossi A, Buscarini L, Danesino C (1996) Hereditary haemorrhagic telangiectasia with extensive liver involvement is not caused by either HHT1 or HHT2. J Med Genet 33:441–443

Plauchu H, de Chadarévian JP, Bideau A, Robert JH (1989) Age-related profile of hereditary hemorrhagic telangiectasia in an epidemiologically recruited population. Am J Med Genet 32:291–297

Puffenberger EG, Francomano CA (1991) PCR-based detec-

tion of polymorphic DdeI and KpnI sites in intron 5 of the adenylate kinase (AK1) gene. Nucleic Acids Res 19:1161

Rendu H (1896) Épistaxis répetées chez un sujet porteur de petits angiomes cutanés et muquez. Gaz Hop (Paris) 135: 132–133

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, et al (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Shovlin CL (1996) A streamlined method for screening a P1 library. BioTechniques 21:388–390

Shovlin CL, Hughes JMB, Tuddenham EGD, Temperley I, Perembelon YFN, Scott J, Seidman CE, et al (1994*a*) A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. Nat Genet 6:205–209

Shovlin CL, Kwiatkowski DJ, Hughes JMB, Seidman CE, Seidman JG (1995*a*) Mapping in the D9S59-67 interval on human chromosome 9. Ann Hum Genet 59:381–382

Shovlin CL, Scott J (1996) Inherited diseases of the vasculature. Annu Rev Physiol 58:483–507

Shovlin CL, Simmonds HA, Fairbanks LD, Deacock SJ, Lechler RI, Hughes JMB, Webster ADB, et al (1994*b*) Adult onset immunodeficiency caused by inherited adenosine deaminase deficiency. J Immunol 153:2331–2335

Shovlin CL, Winstock AR, Peters AM, Jackson JE, Hughes JMB (1995*b*) Medical complications of pregnancy in hereditary haemorrhagic telangiectasia. Q J Med 88:879–887

Yamashita H, Ichijoro H, Grimsby S, Moren A, ten Dijke P, Miyazono K (1994) Endoglin forms a heteromeric complex with the signaling receptors for transforming growth factor-β. J Biol Chem 269:1995–2001

Zhang H, Shaw ARE, Mak A, Letarte M (1996) Endoglin is a component of the transforming growth factor (TGF)-β receptor complex of human pre-B leukemic cells. J Immunol 56:565–573

Zhang S, Ruiz-Echevarria MJ, Quan Y, Peltz SW (1995) Identification and characterization of a sequence motif involved in nonsense-mediated mRNA decay. Mol Cell Biol 15: 2231–2244