Functional analysis of a novel *ENG* variant in a patient with Hereditary Hemorrhagic Telangiectasia (HHT) identifies a new Sp1 binding-site

Sara Plumitallo^a, Lidia Ruiz-Llorente^b, Carmen Langa^b, Jacopo Morini^c, Gabriele Babini^c, Donata Cappelletti^d, Laura Scelsi^e, Alessandra Greco^e, Cesare Danesino^{a,f}, Carmelo Bernabeu^b, Carla Olivieri^a

^aMolecular Medicine Department, General Biology and Medical Genetics Unit, University of Pavia, Via Forlanini 14, 27100, Pavia, Italy. sara.plumitallo01@universitadipavia.it; cidi@unipv.it; carla.olivieri@unipv.it.

^bCentro de Investigaciones Biológicas – Consejo Superior de Investigaciones Científicas and Centro de Investigación Biomédica en Red de Enfermedades Raras, Calle Ramiro de Maeztu 9, 28040, Madrid, Spain. lruiz.llorente@cib.csic.es; langa.poza@cib.csic.es; bernabeu.c@cib.csic.es.

^cPhysics Department, Radiation Biophysics and Radiobiology Lab, University of Pavia, Via Bassi 6, 27100, Pavia, Italy. jacopo.morini@unipv.it; gabriele.babini@unipv.it.

^dCoronary Care Unit and Laboratory of Clinical and Experimental Cardiology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy; donata.cappelletti@unipv.it.

^eCardiothoracic-Vascular Department, Cardiology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy. l.scelsi@smatteo.pv.it; grecale1985@gmail.com.

^fGenetic Counselling Service, IRCCS Fondazione Policlinico San Matteo, Piazzale Golgi 2, 27100, Pavia, Italy. cidi@unipv.it.

Corresponding Author:

Sara Plumitallo,

e-mail sara.plumitallo01@universitadipavia.it

phone +39 0382 987730

Abstract

Hereditary Hemorrhagic Telangiectasia (HHT) is a rare disease, with an autosomal dominant inheritance and a worldwide incidence of about 1: 5,000 individuals. In more than 80% of patients, HHT is caused by mutations in either *ENG* or *ACVRL1*, which code for ENDOGLIN and Activin A Receptor Type II-Like Kinase 1 (ALK1), belonging to the TGF- β /BMP signalling pathway. Typical HHT clinical features are mucocutaneous telangiectases, arteriovenous malformations, spontaneous and recurrent epistaxis, as well as gastrointestinal bleedings. An additional, but less frequent, clinical manifestation in some HHT patients is the presence of Pulmonary Arterial Hypertension (PAH).

The aim of this work is to describe the functional role of a novel *ENG* intronic variant found in a patient affected by both HHT and PAH, in order to assess whether it has a pathogenic role. We proved that the variant lies in a novel binding-site for the transcription factor Sp1, known to be involved in the regulation of *ENG* and *ACVRL1* transcription. We confirmed a pathogenic role for this intronic variant, as it significantly reduces *ENG* transcription by affecting this novel Sp1 binding-site.

Key words: Rendu-Osler-Weber Syndrome; Pulmonary Arterial Hypertension; *ENG* transcription; intronic variant

1. INTRODUCTION

Hereditary Hemorrhagic Telangiectasia (HHT) is a rare genetic disease, also known as Rendu-Osler-Weber Syndrome (ROW), with an autosomal dominant inheritance. HHT affects approximately 1:5,000 individuals worldwide. The main feature of HHT is the presence of abnormal vascular structures such as telangiectases and ArterioVenous Malformations (AVMs), defined as direct connections between arteries and veins, with the loss of the interposed capillary beds (McDonald et al., 2015; Shovlin, 2010). HHT-related telangiectases are small and arise at mucocutaneous sites, such as nasal and oral mucosae, the lips, upper fingertips and the gastrointestinal tract. AVMs are larger lesions that typically affect solid organs including the lungs (Pulmonary AVMs – PAVM), liver (Hepatic AVMs – HAVMs), brain (Cerebral AVMs – CAVMs) and spinal cord (Spinal Cord AVMs – SCAVMs) (Bayrak-Toydemir et al., 2004; Shovlin, 2010). Both telangiectases and AVMs are characterized by thin walls of the vessels, which are prone to bleed when exposed to high perfusion pressure (Govani and Shovlin, 2009). In fact, as a consequence of the nasal telangiectases rupture, chronic epistaxes are observed in almost all HHT patients after the age of 40 (Lesca et al., 2007).

HHT clinical diagnosis is definite when three out of the following four “Curaçao criteria” are observed in patients: i) spontaneous and recurrent epistaxes; ii) telangiectases at the characteristic sites; iii) AVMs in the above-mentioned internal organs and iv) a positive family history, i.e. a first degree relative with an HHT diagnosis based on the same criteria (Shovlin et al., 2000).

HHT was initially described as a familial disorder characterized by severe and recurrent nosebleeds, gastrointestinal bleedings and anaemia (Osler, 1901; Rendu, 1896; Weber, 1907). Today, the phenotypic spectrum of the disease has been expanded to also include Juvenile Polyposis (JP) (Gallione et al., 2004) and Pulmonary Arterial Hypertension (PAH) (Trembath et al., 2001). In particular, PAH (MIM#178600) is a subtype of Pulmonary Hypertension (PH) characterized by the presence of pre-capillary PH, demonstrated by a mean Pulmonary Artery Pressure (mPAP) ≥ 25 mmHg, a pulmonary artery wedge pressure ≤ 15 mmHg and Pulmonary Vascular Resistance (PVR) > 3 Wood units, evaluated by right heart catheterization, in the absence of other causes (Galiè et al., 2016). The term PAH includes different disease subtypes: idiopathic PAH (IPAH), heritable PAH (HPAH) and PAH associated with other systemic diseases or drug/toxin exposures (Austin and Loyd, 2013). Right heart catheterization remains essential for PH or PAH diagnosis (Galiè et al., 2016).

The genetic bases of HHT lie in the TGF- β /BMP signalling pathway, which is involved in the regulation of several physiological processes, including angiogenesis. To date, four genes encoding for proteins belonging to this pathway have been related to HHT. Mutations in *ENG* (MIM*131195), coding for ENDOGLIN (ENG), and

ACVRL1 (MIM *601284), coding for Activin A Receptor Type II-Like Kinase 1 (ALK1), are associated with HHT type 1 (MIM#187300) and type 2 (MIM#600376), respectively, and account for more than 80% of patients (Olivieri et al., 2007). Mutations in *MADH4* (MIM*600993), coding for Mothers Against Decapentaplegic Homolog 4 (SMAD4), are observed in 2-3% of patients with a syndrome called JPHT (MIM#175050), characterised by both HHT and JP symptoms (Gallione et al., 2010). Lastly, mutations in *GDF2* (MIM *605120), coding for Bone Morphogenetic Protein 9 (BMP9), are described in only a few families and are associated with an HHT-like syndrome named HHT5 (MIM#615506) (Wooderchak-Donahue et al., 2013). Thus, in about 7-13% of individuals with an HHT clinical definite diagnosis, mutations in the previously listed genes are not found.

The development of PAH is related to the dysregulation of the TGF- β /BMP signalling pathway. In fact, mutations in *BMPR2* (Bone Morphogenetic Protein Receptor Type 2 – MIM*600799), encoding a receptor of the BMPs family, account for 75% of HPAH cases and for 15% of apparent IPAH cases (Austin and Loyd, 2013; West et al., 2014).

Here we present data about the pathogenic role of a novel *ENG* intronic Variant of Uncertain Significance (VUS), c.1852+42 C>T, observed in a patient with a PAH diagnosis followed by the identification of typical HHT clinical features. Thanks to different functional studies, we proved a decrease of *ENG* transcription due to the presence of this variant. In particular, this VUS impairs a novel intronic binding-site for the transcription factor Sp1, which is already known to be involved in *ENG* and *ACVRL1* transcription (Albiñana et al., 2017; Botella et al., 2001; Garrido-Martin et al., 2010; Rius et al., 1998).

2. MATERIALS AND METHODS

2.1. Patients

Patients' blood samples were collected at the Cardiology Unit at IRCCS Fondazione Policlinico “San Matteo” in Pavia. Each patient had a complete cardiologic examination and signed an informed consent for molecular analyses. The study was performed in accordance with the ethical principles of the Declaration of Helsinki (<http://www.wma.net/en/30publications/10policies/b3/>).

2.2. DNA extraction, Polimerase Chain Reaction (PCR) and Multiplex Ligation Probe Amplification (MLPA)

Genomic DNA was extracted from peripheral blood using the “GenElute™ Blood Genomic DNA Kit” by SIGMA, according to manufacturer's instructions. We used the online software “Primer 3” (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) to design the *BMPR2*, *ENG* and *ACVRL1* primers

used to amplify coding exons and flanking introns. The dimensions of the flanking intron sequences were evaluated individually for each exon to cover all regions that may potentially include known variants, either pathogenic or benign (<http://arup.utah.edu/database/HHT/>). The PCR, purified using MultiScreen®PCR_μ96 Filter Plate (Merck Millipore), underwent Sanger sequencing using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem™), following the manufacturer's instructions.

MLPA of *ENG*, *ACVRL1* and *BMPR2* was performed to assess the presence of large deletions or duplications, using the SALSA MLPA probemix P093-C2 HHT/PPH1 (MRC-Holland).

Both Sanger sequencing and fragment separation analyses were performed on a 3500 series Genetic Analyzer 8-Capillary (Applied Biosystem™). Regarding MLPA, a GS500 size standard was adopted and the results were analysed using Coffalyser software (<https://coffalyser.wordpress.com/>).

2.3. Plasmids site-directed mutagenesis

Site-directed mutagenesis was performed on two constructs: pcEXV-ENG (Bellón et al., 1993) and pDisplay-MinigEND (Blanco and Bernabeu, 2011). The first plasmid contains the entire *ENG* coding sequence and intron 14, while the second one contains the *ENG* sequence from the beginning of exon 14 to the end of exon 15. The mutagenesis was carried out using primers specifically designed with the variant of interest (Forward: 5'–GGGCCCCCTTCATCCACCC–3' and Reverse: 5'–GGGTGGATGAAGGGGCCC–3'; in bold the nucleotide change).

2.4. Bacterial transformation and cell transfection

Mutated constructs were used to transform One Shot TOP10 Chemically Competent *E. coli*. Bacteria were grown in LB agar plates (100μg/mL ampicillin). Single colonies were inoculated in LB medium (100μg/mL ampicillin) and incubated overnight at 37°C. DNA was extracted from bacteria by using PureYield™ Plasmid Miniprep System (Promega). The *ENG* insert of both constructs was fully sequenced and the mutation, as well as the absence of any error introduced by the amplification, was verified.

Mutated constructs were used to transfect Human Embryonic Kidney 293T (HEK293T) cells, grown in Dulbecco's Modified Eagle Medium (DMEM - Gibco) supplemented by 10% heat-inactivated Fetal Calf Serum (FCS), 2mM L-glutamine, 100U/mL penicillin and 100U/mL streptomycin. Cell transfection was carried out using Lipofectamine® LTX Reagent (Invitrogen), following the manufacturer's instructions.

2.5. Western Blot analyses

Forty-eight hours after transfection, the cells were incubated at 4°C for 10 minutes with a lysis solution containing 50 mM Hepes pH 7.5, 150 mM NaCl, 1% NP-40 and a cocktail of protease and phosphatase

inhibitors. Lysates were centrifuged at 13,200g, 4°C, for 15 minutes; subsequently, supernatants were stored at -20°C. The protein concentration was quantified by Bradford assay and 60µg of total protein was separated by SDS-PAGE and electrotransferred to a PVDF membrane using an iBlot Gel Transfer System (Invitrogen).

Immunodetection was performed with the following monoclonal primary antibodies: anti-CD105 (ab169545; Abcam), 1: 1,000, for pcEXV-ENG; anti-HA (12CA5; Roche), 1:50, for pDisplay-minigEND and anti-β-actin (AC-15; Sigma), 1: 4,000. Protein bands were revealed using SuperSignal® West Pico chemiluminescent substrate (Thermo Scientific).

2.6. Reverse Transcriptase PCR (RT-PCR) and quantitative PCR (qPCR)

Peripheral Blood Mononuclear Cells (PBMCs) were isolated using Histopaque®-1077 (Sigma®) and total RNA was extracted from PBMCs using RNeasy® Mini Handbook (Qiagen) according to the manufacturer's instructions. RNA samples were stored at -20°C until use.

Reverse transcriptase was performed using iScript cDNA kit (Bio-Rad) to retrotranscribe 350ng of total RNA for either patient or control, according to the manufacturer's instructions. The resulting cDNA was amplified using iProof High-Fidelity DNA Polymerase (Bio-Rad) and primers (For 5' – CAACTGTGTGAGCCTGCTGT – 3'; Rev 5' – TATGCCATGCTGCTGGTG – 3'). These primers were specifically designed from the middle sequence of exon 12 to the end of the coding region of exon 15 in order to include intron 14. In this way, we excluded the presence of genomic DNA amplification and distinguished between the two mRNA variants coding for S-ENG or L-ENG.

qPCR experiments were performed using EXPRESS One-Step SYBR® GreenER™ with Premixed ROX (Invitrogen). β-actin (*ACTB*) was used as reference gene.

Primers used for *ENG*, *ACVRL1* and *ACTB* amplification have been described elsewhere (Zucco et al. 2014).

2.7. Electrophoretic Mobility Shift Assay (EMSA)

Nuclear Extracts (NE) were prepared from AG01522 human fibroblasts by using the Nuclear Extraction Kit (Cayman chemical). DNA probes were labelled at the 3'-end with biotin using the Biotin 3'End DNA Labelling Kit (Thermo Scientific) according to the manufacturer's instructions. Three different double-stranded probes were used:

- control (CTRL). These probes have already been reported (Botella et al., 2001):

For: 5' – GCAGGCGGCCTGGCCCAGCCCCTTCTC – 3'

Rev: 3' – CGTCCGCCGGACCGGGTCGGGGAAGAG – 5'

- wild type (WT) (designed for this study):

For: 5' – GCATGCCGGGGCCCCTCCATCCA – 3'

Rev: 3' – CGTACGGCCCCGGGGAGGTAGGT – 5';

– mutated (MUT) (designed for this study):

For: 5' – GCATGCCGGGGCCCCTTCATCCA – 3'

Rev: 3' – CGTACGGCCCCGGGGAAGTAGGT – 5'.

Both WT and MUT probes are the *ENG* segment IVS14 +27/+48 and differed only with regard to the IVS14 +42 nucleotide (highlighted in bold), which was mutated in the patient. Each forward and reverse oligonucleotide was labelled as single stranded and annealed with the complementary probe using the following program: 95°C for 5'; 95°C – 1°C/cycle for 1' (28/36 cycles, depending on the melting temperature of the probes); 67/59°C* for 30'; 67/59°C* – 1°C/cycle for 1' (63/55 cycles, depending on the probes); 4°C pause.

The DNA binding reaction was performed using the Lightshift Chemiluminescent EMSA Kit (Thermo Scientific). 30µg of NE was incubated for 20 minutes at room temperature with 20fmol of labelled probes and 200-fold molar excess of the corresponding unlabelled probe was added as a specific competitor. The samples were run on 6% non-denaturing polyacrilamide gel TBE 0.5X (running buffer: pre-cooled TBE 0.5X) at 80V for 100 minutes. The electrophoretic transfer to Byodine® B pre-cut modified nylon membrane (0.45µm; Thermo Scientific) was performed at 320mA for 45 minutes at 4°C (transfer buffer: pre-cooled TBE 0.5X). Membranes were crosslinked on a transilluminator ($\lambda=312\text{nm}$) for 15 minutes. Biotin-labelled DNA was detected using the Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific). Antibody (Ab)-induced inhibition was performed by incubating, for an additional 20 minutes at room temperature, 4µg of anti-Sp1 (sc-59X – Santa Cruz) with the reaction mix after the addition of the labelled probe.

2.8. *In silico* analyses

The MutationTaster2; FATHMM/mklv2.3 and Ensembl/VEP tools were used to assess the potential pathogenic role of the variant. The TomTom4.11.2; Patch1.0 and AliBaba2.1 tools were used to predict the transcription factors binding-sites affected by the studied variant.

2.9. Statistics

A two-tail Student's t-test was used for statistical computing, with a *p-value* < 0.05 considered as statistically significant.

3. RESULTS

3.1. Patients

The proband, P19, a 64 years old Italian man, was hospitalized in 1996 in the Cardiology Unit (IRCCS, Fondazione Policlinico “San Matteo”, Pavia, Italy) for the evaluation of an atrial flutter. Right heart catheterization revealed severe pre-capillary PH, with an mPAP of 58mmHg (83/33mmHg). No right to left shunt was identified. The first diagnosis was IPAH. In 2014, the patient was hospitalized in the Otorhinolaryngology Unit (IRCCS, Fondazione Policlinico “San Matteo”, Pavia, Italy) for the presence of recurrent epistaxes; telangiectases of nasal mucosa, lips and fingertips were also observed. These symptoms led to the diagnosis of “suspected” HHT, and a genetic counselling was offered to the patient. The presence of HHT-related symptoms was identified in other family members (Fig. 1).

3.2. Mutation analyses

The analyses of *BMPR2*, *ENG* and *ACVRL1* coding regions and flanking introns by direct sequencing and MLPA failed to identify a known disease-causing mutation. We found the VUS c.1852+42 C>T in intron 14 of *ENG*, 16bps after the STOP codon of the short isoform of the protein (Bellón et al., 1993). To evaluate the co-segregation of the variant and the disease within the family, we collected additional blood samples from relatives, both symptomatic and asymptomatic. The variant was detected only in the individuals with clinical symptoms related to HHT (Fig. 1). The relatives carrying the mutation undergo scheduled clinical screening for both diseases even though, to date, none of them have shown any sign of PAH.

3.3. *In silico* analyses of the VUS

The variant is not listed in the HHT International database (<http://www.arup.utah.edu/database/HHT/>), but it is present in The Single Nucleotide Polymorphism database (dbSNP; <https://www.ncbi.nlm.nih.gov/projects/SNP/>) (rs779974705) and in The Exome Aggregation Consortium (ExAC; <http://exac.broadinstitute.org/>). The ExAC database reports the presence only of the heterozygous genotype with a very low worldwide frequency (0.0001331), including the European (Non-Finnish) population (0.0002139). To evaluate the pathogenic role of this variant, we used different *in silico* tools, such as MutationTaster (<http://www.mutationtaster.org/>), FATHMM/mkl (<http://fathmm.biocompute.org.uk/fathmmMKL.htm>) and Ensembl/VEP (<http://www.ensembl.org/info/docs/tools/vep/index.html>). MutationTaster classified the variant as a polymorphism capable of changing a splice site. According to FATHMM/mkl, the variant had a damaging effect. Ensembl/VEP described the impact of the variant as a modifier while indicating that the prediction was complicated by the localization of the VUS within a non-coding region.

3.4. Protein expression

Protein expression experiments were performed using the pcEXV-ENG (Bellón et al., 1993) and pDisplay-MinigEND (Blanco and Bernabeu, 2011) plasmids. Mutated and wild type constructs were transfected into HEK293T cells and western-blotting was used to evaluate ENG expression. We observed that ENG was expressed at lower levels in the mutant construct, with a statistically significant reduction when compared to the control. In particular, ENG expression showed a mean reduction of more than 40% (standard deviation = $\pm 3.5\%$; p -value = 0.0001) and 25% (standard deviation = $\pm 7.0\%$; p -value = 0.02) in the pc-EXV-ENG and pDisplay-MinigEND constructs, respectively (Fig. 2a-d).

Since the VUS lies in the intron IVS14, retained in the biosynthesis of the short isoform of the protein (S-ENG) (Bellón et al., 1993), we assessed the possibility of a splice site alteration. HEK293T cells were co-transfected with the pDisplay construct, either wild type or mutated, together with the expression vector of the splicing factor ASF/SF2, known to be involved in the retention of *ENG* IVS14 (Blanco and Bernabeu, 2011). A significant increase in S-ENG levels was obtained when ASF/SF2 was over-expressed, in line with data from the literature (Blanco and Bernabeu, 2011). We excluded the impairment of *ENG* alternative splicing, as both the wild type and the mutated constructs did not lead to differences in protein expression levels (Fig. 3a). In addition, short-ENG expression was evaluated *in vivo* using PBMCs. The RT-PCR analysis of the Patient's total RNA showed the presence of a single transcript of 435bp, corresponding to the long isoform of the protein, but not the predicted 570bp band corresponding to S-ENG (Fig. 3b). In the same experiment, the amplified band from the patient was clearly less intense than the band of the control subject, suggesting that the rate of transcription in the mutant *ENG* was impaired. Taken together, both results allowed us to reject the hypothesis of a splice site alteration.

3.5. Sp1 and *ENG* expression

The transcript levels of *ENG* and *ACVRL1* were subsequently evaluated by qRT-PCR in PBMCs derived from the HHT patient (Fig. 4). We observed that *ENG* was transcribed 65% less in the patient compared to the control (standard deviation = $\pm 20\%$; p -value = 0.046). Similarly, the patient's mRNA levels of *ACVRL1* showed a trend of decreased expression (Fig. 4), which is in line with data from the literature (Albiñana et al., 2017; Zucco et al., 2014), although our results did not achieve statistical significance. These results were concordant with those obtained by western blotting, suggesting a possible role of the VUS in the regulation of *ENG* transcription.

To assess this hypothesis, we compared wild type and mutated sequences with *in silico* prediction tools (TomTom4.11.2; Patch1.0 and AliBaba2.1), aiming at identifying putative transcription factor binding-sites centred in the VUS. Each prediction tool identified several differences between the two sequences and, of note,

each of them predicted that the C>T change leads to the alteration or to the complete loss of a putative Sp1 binding-site (Online Resource 1).

Therefore, we tested the wild-type and the mutated sequences for their ability to bind Sp1 by using the EMSA. As expected, in the experiments performed with the CTRL oligonucleotide, a shifted band appeared after the incubation of the labelled probe with the NE (Fig. 5a). Upon adding the Ab- α Sp1, we observed a marked decrease in the intensity of the shifted band, demonstrating that Sp1 was included in the DNA-protein complex (Fig. 5b). The experiments were repeated with the WT and MUT oligonucleotides. Both probes shifted their mobility in the presence of NE, likely involving the binding to Sp1, although the binding of each probe occurred with a different affinity, as shown by the Ab-induced inhibition. In fact, the reduction of the shifted band intensity was higher for the WT probes, suggesting that the Sp1 binding-site was impaired by the presence of the VUS (Fig. 5b).

4. DISCUSSION

Patient P19 was diagnosed with Hereditary Hemorrhagic Telangiectasia in 2014 at the age of 61, after an initial diagnosis of Idiopathic Pulmonary Arterial Hypertension in 1996 at the age of 43. The diagnostic delay was partly due to the late onset of the typical HHT symptoms and to the fact that the overlap between PAH and HHT was unknown at the time of the first diagnosis. This finding agrees with the data from the literature regarding a single patient showing PAH as the first symptom of HHT (Ishiwata et al., 2014).

Molecular screening of *BMPR2*, *ENG* and *ACVRL1* revealed the presence of a single variant of uncertain significance, c.1852+42 C>T, in intron 14 of *ENG*. The co-segregation, within the family, of HHT symptoms and this VUS (Fig. 1), in addition to its very low frequency, as reported in the ExAC database (0.0002139 in the European non-Finnish population), suggest that the variant has a putative pathogenic effect. In fact, the observed frequency is lower than the 1% threshold normally used to discern rare mutations from common polymorphisms (Brookes, 1999). Recently, Karki and collaborators have pointed out that in the era of the next generation sequencing may not be enough to classify the role of a DNA variant according to its frequency in the general population (Karki et al., 2015). Thus, using the 1% threshold alone without further investigation may be inappropriate. In fact, mutations with a higher frequency, rare variants which do not correlate with human diseases and variants with different effect in separated populations have been described (Auer et al., 2012; Karki et al., 2015; Myles et al., 2008).

Our functional studies clearly demonstrate, both *in vitro* and *in vivo*, the pathogenetic role of the *ENG* c.1852+42 C>T variant, which leads to a decreased expression of Endoglin. The VUS lies in a novel intronic

binding-site for the transcription factor Sp1, known as a key regulator of *ENG* and *ACVRL1* transcription (Albiñana et al., 2017; Botella et al., 2001; Garrido-Martin et al., 2010). In fact, the promoters of the two genes contain multiple Sp1 binding-sites. Previous reports have widely documented the existence of multiple Sp1 functional sites in *ENG*; these studies have already shown the involvement of Sp1 in basal transcription, as well as in stimulation-induced transcription (Albiñana et al., 2017; Botella et al., 2002, 2001; Garrido-Martin et al., 2010; Rfús et al., 1998; Sánchez-Elsner et al., 2002). The studied VUS changes the predicted Sp1 binding-site, thereby interfering with the protein-DNA interaction (Fig. 6). As shown by the EMSA experiments, the oligonucleotides containing the wild type allele bind to Sp1 with higher specificity and affinity when compared to the mutated ones. We observed that the unlabelled probes completely prevent the DNA-protein complex formation only for the WT samples (Fig. 5a). Moreover, a stronger inhibition by Ab- α Sp1 for the WT oligonucleotides than for the MUT ones is pointed out (Fig. 5b).

Our data are the first to highlight the presence of a novel intronic Sp1 binding-site which is involved in the regulation of *ENG* transcription. Even if Sp1 is generally described as a promoter-specific transcription factor, it can also concurrently bind to regulatory elements, distal to the promoter (Williams et al., 2007). Therefore, the alteration of the predicted binding-site, caused by the presence of the VUS, leads to Endoglin haploinsufficiency and, as a result, to the onset of the disease.

Since the variant lies in the intron retained during the alternative splicing of the precursor mRNA of *ENG*, it was important to rule out the hypothesis of an alteration of this process, especially considering that the two isoforms present contrasting functions. Both isoforms are mainly expressed on the surface of endothelial cells: L-ENG is the predominant isoform and is associated with cell proliferation, while S-ENG derives from the retention of intron 14 and its expression increases during cell senescence (Blanco et al., 2008). Hence, an increased expression of S-ENG and a lower expression of L-ENG could simulate the haploinsufficiency, the prevalent model for explaining HHT pathogenesis (McDonald et al., 2015). Results from western blotting and *in vivo* RT-PCR prove that the VUS is not involved in this alternative splicing (Fig. 3).

5. CONCLUSIONS

This work describes the pathogenic role of a new *ENG* intronic variant observed in a patient affected with HHT and presenting PAH as the main clinical manifestation. The obtained results support the need to further investigate the non-coding regions of the known HHT-causing genes, as stated by Albiñana and collaborators (Albiñana et al., 2017), who recently described a novel *ENG* pathogenic variant in the proximal promoter of the gene. Similar conclusions were recently reported at the XII HHT International Scientific Conference (June 8th-

11th, 2017, in Dubrovnik, Croatia). To the best of our knowledge, this is the first time that the link between an intronic variant and the regulation of *ENG* transcription has been described.

Acknowledgements

We wish to thank the Italian HHT Charity “Onilde Carini”, along with all the patients and their families. We also thank Dr. V. Curti and Prof. M. Daglia for the qPCR facilities.

Funding: This work was supported by grants from the *Ministerio de Industria, Economia y Competitividad* (SAF2013-43421-R to CB) and the *Centro de Investigacion Biomedica en Red de Enfermedades Raras* (CIBERER; ISCIII-CB06/07/0038 and ER16PIAC707 to CB) in Spain, both supported by European Regional Development Funds (FEDER). CIBERER is an initiative of the Instituto *de Salud Carlos III* (ISCIII) in Spain. This work is part of the results of research activity by SP, who holds a research fellowship (2/R.C./2016, n.20160028774) from the IRCCS Fondazione Policlinico San Matteo, Pavia, Italy.

Conflict of interest: All the authors declare they do not have any conflict of interest.

REFERENCES

- Albiñana, V., Zafra, M.P., Colau, J., Zarrabeitia, R., Recio-Poveda, L., Olavarrieta, L., Pérez-Pérez, J., Botella, L.M., 2017. Mutation affecting the proximal promoter of Endoglin as the origin of hereditary hemorrhagic telangiectasia type 1. *BMC Med. Genet.* 18, 20. doi:10.1186/s12881-017-0380-0
- Auer, P.L., Johnsen, J.M., Johnson, A.D., Logsdon, B.A., Lange, L.A., Nalls, M.A., Zhang, G., Franceschini, N., Fox, K., Lange, E.M., Rich, S.S., O'Donnell, C.J., Jackson, R.D., Wallace, R.B., Chen, Z., Graubert, T.A., Wilson, J.G., Tang, H., Lettre, G., Reiner, A.P., Ganesh, S.K., Li, Y., 2012. Imputation of exome sequence variants into population-based samples and blood-cell-trait-associated loci in african americans: NHLBI GO exome sequencing project. *Am. J. Hum. Genet.* 91, 794–808.
- Austin, E., Loyd, J., 2013. Heritable forms of pulmonary arterial hypertension. *Semin. Respir. Crit. Care Med.* 34, 568–580. doi:10.1055/s-0033-1355443
- Bayrak-Toydemir, P., Mao, R., Lewin, S., McDonald, J., 2004. Hereditary hemorrhagic telangiectasia: an overview of diagnosis and management in the molecular era for clinicians. *Genet. Med.* 6, 175–191.
- Bellón, T., Corbí, a, Lastres, P., Calés, C., Cebrián, M., Vera, S., Cheifetz, S., Massague, J., Letarte, M., Bernabéu, C., 1993. Identification and expression of two forms of the human transforming growth factor-beta-binding protein endoglin with distinct cytoplasmic regions. *Eur. J. Immunol.* 23, 2340–2345.
- Blanco, F.J., Bernabeu, C., 2011. Alternative splicing factor or splicing factor-2 plays a key role in intron retention of the endoglin gene during endothelial senescence. *Aging Cell* 10, 896–907.
- Blanco, F.J., Grande, M.T., Langa, C., Oujó, B., Velasco, S., Rodríguez-Barbero, A., Perez-Gomez, E., Quintanilla, M., López-Novoa, J.M., Bernabeu, C., 2008. S-endoglin expression is induced in senescent endothelial cells and contributes to vascular pathology. *Circ. Res.* 103, 1383–1392.
- Botella, L.M., Sánchez-Elsner, T., Rius, C., Corbí, A., Bernabéu, C., 2001. Identification of a critical Sp1 site within the Endoglin promoter and its involvement in the transforming growth factor- β stimulation. *J. Biol. Chem.* 276, 34486–34494.
- Botella, L.M., Sanchez-Elsner, T., Sanz-Rodriguez, F., Kojima, S., Shimada, J., Guerrero-Esteo, M., Cooreman, M.P., Ratzliff, V., Langa, C., Vary, C.P., Ramirez, J.R., Friedman, S., Bernabeu, C., 2002. Transcriptional activation of endoglin and transforming growth factor-beta signaling components by cooperative interaction between Sp1 and KLF6: their potential role in the response to vascular injury. *Blood*. doi:10.1182/blood.V100.12.4001 100/12/4001 [pii]
- Brookes, A.J., 1999. The essence of SNPs. *Gene*.
- Galiè, N., Humbert, M., Vachiery, J.-L., Gibbs, S., Lang, I., Torbicki, A., Simonneau, G., Peacock, A., Vonk Noordegraaf, A., Beghetti, M., Ghofrani, A., Gomez Sanchez, M.A., Hansmann, G., Klepetko, W., Lancellotti, P., Matucci, M., McDonagh, T., Pierard, L.A., Trindade, P.T., Zompatori, M., Hoeper, M., Aboyans, V., Vaz Carneiro, A., Achenbach, S., Agewall, S., Allanore, Y., Asteggiano, R., Paolo Badano, L., Albert Barberà, J., Bouvaist, H., Bueno, H., Byrne, R.A., Carerj, S., Castro, G., Erol, Ç., Falk, V., Funck-Brentano, C., Gorenflo, M., Granton, J., Iung, B., Kiely, D.G., Kirchhof, P., Kjellstrom, B., Landmesser, U., Lekakis, J., Lionis, C., Lip, G.Y.H., Orfanos, S.E., Park, M.H., Piepoli, M.F., Ponikowski, P., Revel, M.-P., Rigau, D., Rosenkranz, S., Völler, H., Luis Zamorano, J., 2016. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur. Heart J.* 37, 67–119. doi:10.1093/eurheartj/ehv317
- Gallione, C., Aylsworth, A.S., Beis, J., Berk, T., Bernhardt, B., Clark, R.D., Clericuzio, C., Danesino, C., Drautz, J., Fahl, J., Fan, Z., Faughnan, M.E., Ganguly, A., Garvie, J., Henderson, K., Kini, U., Leedom, T., Ludman, M., Lux, A., Maisenbacher, M., Mazzucco, S., Olivieri, C., Van Ploos Amstel, J.K., Prigoda-Lee, N., Pyeritz, R.E., Reardon, W., Vandezande, K., Waldman, J.D., White, R.I., Williams, C.A., Marchuk, D.A., 2010. Overlapping spectra of SMAD4 mutations in juvenile polyposis (JP) and JP-HHT syndrome. *Am. J. Med. Genet. Part A* 152, 333–339. doi:10.1002/ajmg.a.33206
- Gallione, C.J., Repetto, G.M., Legius, E., Rustgi, A.K., Schelley, S.L., Tejpar, S., Mitchell, G., Drouin, É., Westermann, C.J.J., Marchuk, D.A., 2004. A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). *Lancet* 363, 852–859.
- Garrido-Martin, E.M., Blanco, F.J., Fernandez-L, A., Langa, C., Vary, C.P., Lee, U.E., Friedman, S.L., Botella, L.M., Bernabeu, C., 2010. Characterization of the human Activin-A receptor type II-like kinase 1 (ACVRL1) promoter and its regulation by Sp1. *BMC Mol. Biol.* 11, 51. doi:10.1186/1471-2199-11-51
- Govani, F.S., Shovlin, C.L., 2009. Hereditary haemorrhagic telangiectasia: a clinical and scientific review. *Eur. J. Hum. Genet.* 17, 860–71.
- Ishiwata, T., Terada, J., Tanabe, N., Abe, M., Sugiura, T., Tsushima, K., Tada, Y., Sakao, S., Kasahara, Y., Nakanishi, N., Morisaki, H., Tatsumi, K., 2014. Pulmonary Arterial Hypertension as the First Manifestation in a Patient with Hereditary Hemorrhagic Telangiectasia . *Intern. Med.* . doi:10.2169/internalmedicine.53.2850
- Karki, R., Pandya, D., Elston, R.C., Ferlini, C., 2015. Defining “ mutation ” and “ polymorphism ” in the era of

- personal genomics 1–7. doi:10.1186/s12920-015-0115-z
- Lesca, G., Olivieri, C., Burnichon, N., Pagella, F., Carette, M.-F., Gilbert-Dussardier, B., Goizet, C., Roume, J., Rabilloud, M., Saurin, J.-C., Cottin, V., Honnorat, J., Coulet, F., Giraud, S., Calender, A., Danesino, C., Buscarini, E., Plauchu, H., 2007. Genotype-phenotype correlations in hereditary hemorrhagic telangiectasia: data from the French-Italian HHT network. *Genet. Med.* 9, 14–22.
- McDonald, J., Wooderchak-Donahue, W., VanSant Webb, C., Whitehead, K., Stevenson, D.A., Bayrak-Toydemir, P., 2015. Hereditary hemorrhagic telangiectasia: Genetics and molecular diagnostics in a new era. *Front. Genet.* 5, 1–8. doi:10.3389/fgene.2015.00001
- Myles, S., Davison, D., Barrett, J., Stoneking, M., Timpson, N., 2008. Worldwide population differentiation at disease-associated SNPs. *BMC Med. Genomics* 1, 22. doi:10.1186/1755-8794-1-22
- Olivieri, C., Pagella, F., Semino, L., Lanzarini, L., Valacca, C., Pilotto, A., Corno, S., Scappaticci, S., Manfredi, G., Buscarini, E., Danesino, C., 2007. Analysis of ENG and ACVRL1 genes in 137 HHT Italian families identifies 76 different mutations (24 novel). Comparison with other European studies. *J. Hum. Genet.* 52, 820–829.
- Osler, W., 1901. On a family form of recurring epistaxis, associated with multiple telangiectasis of the skin and mucous membranes. *Bull Johns Hopkins Hosp* 28, 401–403.
- Rendu, H., 1896. Épistaxis répétées chez un sujet porteur de petits angiomes cutanés et muqueux. *Bull Soc méd Hosp* 13, 731–733.
- Ríus, C., Smith, J.D., Almendro, N., Langa, C., Botella, L.M., Marchuk, D. a, Vary, C.P., Bernabéu, C., 1998. Cloning of the promoter region of human endoglin, the target gene for hereditary hemorrhagic telangiectasia type 1. *Blood* 92, 4677–4690.
- Sánchez-Elsner, T., Botella, L.M., Velasco, B., Langa, C., Bernabéu, C., 2002. Endoglin expression is regulated by transcriptional cooperation between the hypoxia and transforming growth factor- β pathways. *J. Biol. Chem.* doi:10.1074/jbc.M207160200
- Shovlin, C.L., 2010. Hereditary haemorrhagic telangiectasia: Pathophysiology, diagnosis and treatment. *Blood Rev.* 24, 203–219.
- Shovlin, C.L., Guttmacher, A.E., Buscarini, E., Faughnan, M.E., Hyland, R.H., Westermann, C.J., Kjeldsen, A.D., Plauchu, H., 2000. Diagnostic criteria for hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber syndrome). *Am. J. Med. Genet.* 91, 66–67.
- Trembath, R.C., Thomson, J.R., Machado, R.D., Morgan, N. V, Atkinson, C., Winship, I., Simonneau, G., Galie, N., Loyd, J.E., Humbert, M., Nichols, W.C., Morrell, N.W., Berg, J., Manes, A., McGaughan, J., Pauculo, M., Wheeler, L., 2001. Clinical and molecular genetic features of pulmonary hypertension in patients with hereditary hemorrhagic telangiectasia. *N.Engl.J.Med.* 345, 325–334.
- Weber, F., 1907. Multiple hereditary developmental angiomas (telangiectases) of the skin and mucous membranes associated with recurring haemorrhages. *Lancet* 2, 160–162.
- West, J., Austin, E., Fessel, J.P., Loyd, J., Hamid, R., 2014. Rescuing the BMPR2 signaling axis in pulmonary arterial hypertension. *Drug Discov. Today*.
- Williams, A.O., Isaacs, R.J., Stowell, K.M., 2007. Down-regulation of human topoisomerase II α expression correlates with relative amounts of specificity factors Sp1 and Sp3 bound at proximal and distal promoter regions. *BMC Mol. Biol.* 8, 36. doi:10.1186/1471-2199-8-36
- Wooderchak-Donahue, W.L., McDonald, J., O’Fallon, B., Upton, P.D., Li, W., Roman, B.L., Young, S., Plant, P., Fülöp, G.T., Langa, C., Morrell, N.W., Botella, L.M., Bernabeu, C., Stevenson, D.A., Runo, J.R., Bayrak-Toydemir, P., 2013. BMP9 mutations cause a vascular-anomaly syndrome with phenotypic overlap with hereditary hemorrhagic telangiectasia. *Am. J. Hum. Genet.* 93, 530–537. doi:10.1016/j.ajhg.2013.07.004
- Zucco, L., Zhang, Q., Kuliszewski, M.A., Kandic, I., Faughnan, M.E., Stewart, D.J., Kutryk, M.J., 2014. Circulating angiogenic cell dysfunction in patients with hereditary hemorrhagic telangiectasia. *PLoS One* 9, 1–8. doi:10.1371/journal.pone.0089927

FIGURE 1

FIGURE 2

FIGURE 3

FIGURE 4

FIGURE 5

FIGURE 6

FIGURE CAPTIONS

Fig. 1 Genealogical tree of the proband (arrow) family In black, the Patients with HHT symptoms.

Fig. 2 Effect of VUS on ENG expression in vitro (A-C) Experiments performed using the pc-EXV revealed a reduction of about 40% (standard deviation: $\pm 3.5\%$) in the expression of mutated ENG (p -value = 0.0001; ***). (B-D) Similar results were obtained using the pDisplay minigEND construct, showing a 25% (standard deviation: $\pm 7\%$) reduction in the expression of the mutated protein (p -value = 0.02; *). The plasmids without ENG constructs were used as a control and are indicated by $\emptyset$.

Fig. 3 Analysis of S-ENG expression using *in vitro* and *in vivo* experiments (A) HEK293T cells were transfected with pDisplay minigEND construct alone or together with the expression vector of ASF/SF2. The co-transfection led to the expression of both L-ENG and S-ENG, but there was no difference between the wild type and the mutated constructs. (B) PMBCs were obtained from the patient (P19) and a control subject (CTRL) and RT-PCR experiments were performed to assess, *in vivo*, the effect of the variant on ENG alternative splicing. Lad stands for DNA ladder and the minus symbol (-) indicates the negative control without cDNA. We observed, in both the patient and the control, a single 435bp band. If ENG intron 14 had been retained and, thus, S-ENG had also been expressed, the RT-PCR would have shown an additional band of 570bp.

Fig. 4 Reduced ENG and ACVRL1 expression levels in the patient Total RNA isolated from PBMCs of patient and control subjects underwent qRT-PCR. A decreased expression of both ENG and ACVRL1 transcript levels in the patient versus the control was observed. In particular, ENG was transcribed 65% (standard deviation: $\pm 20\%$) less in the patient compared to the control (p -value = 0.046; *) and ACVRL1 was transcribed 25% (standard deviation: $\pm 47\%$) less in the patient (n.s. = not significant).

Fig. 5 EMSA (A) Shift of the CTRL oligonucleotides, showing the binding to an undetermined transcription factor. A shift band appears in both the WT and MUT oligonucleotides. (B) The antibody-induced inhibition experiments revealed that Sp1 binds to all the analysed probes; however, a higher binding affinity for the WT oligonucleotides when compared to the MUT probes was observed.

Fig. 6 Schematic model of the effect of the VUS on ENG transcription On the left, Sp1 binds with high affinity to the wild type sequence, leading to normal gene transcription and ENG expression levels. On the right, the presence of the T allele reduces the strength of the DNA-protein binding and results in a decrease of mRNA levels and ENG biosynthesis.