**ORIGINAL ARTICLE**

**Genome sequencing reveals a deep intronic splicing ACVRL1 mutation hotspot in Hereditary Haemorrhagic Telangiectasia**

Whitney L. Wooderchak-Donahue,1,2 Jamie McDonald,2,3 Andrew Farrell,4 Gulsen Akay,1,5 Matt Velinder,4 Peter Johnson,1 Chad VanSant-Webb,1 Rebecca Margraf,1 Eric Briggs,1 Kevin J Whitehead,3,6 Jennifer Thomson,2,7 Angela E Lin,8 Reed E Pyeritz,9 Gabor Marth,4 Pinar Bayrak-Toydemir1,2

**ABSTRACT**

**Introduction** Hereditary haemorrhagic telangiectasia (HHT) is a genetically heterogeneous disorder caused by mutations in the genes ENG, ACVRL1, and SMAD4. Yet the genetic cause remains unknown for some families even after exhaustive exome analysis. We hypothesised that non-coding regions of the known HHT genes may harbour variants that disrupt splicing in these cases.

**Methods** DNA from 35 individuals with clinical findings of HHT and 2 healthy controls from 13 families underwent whole genome sequencing. Additionally, 87 unrelated cases suspected to have HHT were evaluated using a custom designed next-generation sequencing panel to capture the coding and non-coding regions of ENG, ACVRL1 and SMAD4. Individuals from both groups had tested negative previously for a mutation in the coding region of known HHT genes. Samples were sequenced on a HiSeq2500 instrument and data were analysed to identify novel and rare variants.

**Results** Eight cases had a novel non-coding ACVRL1 variant that disrupted splicing. One family had an ACVRL1 intron 9:chromosome 3 translocation, the first reported case of a translocation causing HHT. The other seven cases had a variant located within a ~300 bp CTRich ‘hotspot’ region of ACVRL1 intron 9 that disrupted splicing.

**Conclusions** Despite the difficulty of interpreting deep intronic variants, our study highlights the importance of non-coding regions in the disease mechanism of HHT, particularly the CTRich hotspot region of ACVRL1 intron 9. The addition of this region to HHT molecular diagnostic testing algorithms will improve clinical sensitivity.

**INTRODUCTION**

Hereditary haemorrhagic telangiectasia (HHT) is an autosomal dominant vascular dysplasia that occurs in 1 in 5000 individuals.1 HHT is characterised by (1) recurrent epistaxis; (2) telangiectases of the nasal mucosa, fingertips, lips and oral cavity, (3) arteriovenous malformations (AVMs) in the lungs, liver, gastrointestinal tract and brain and (4) a family history of HHT. Patients who present with three or more of these criteria are considered clinically diagnosed with HHT according to ‘Curaçao criteria’.2 HHT is a genetically heterogeneous disorder caused by mutations in one of multiple genes in the transforming growth factor-beta signalling pathway. Endoglin (ENG), activin A receptor type II-like 1 (ACVRL1/ALK1) and SMAD4 mutations cause HHT1 (OMIM 187300), HHT2 (OMIM 600376) and the combined juvenile polyposis/HHT syndrome (OMIM 175050), respectively.3–5 Mutations in these genes lead to an underproduction of their respective proteins and results in excessive abnormal angiogenesis.6 ENG and ACVRL1 mutations account for roughly equal percentages of the disorder,7 and are found in ~85% of cases submitted for molecular genetic testing for HHT.8–9 Mutations in SMAD4 cause only 1%–2% of cases.10 Numerous private mutations have been described across all coding regions of ENG and ACVRL1,11 including large exonic deletions or duplications that occur in ~6%–10% of cases.9,10,12 More recently, non-coding mutations in the 5’ untranslated region (UTR) of ENG were shown to cause HHT in ~1%–2% of cases,11,13,14 indicating the need to include this non-coding region in routine HHT molecular testing algorithms.15 Molecular diagnostic testing is important in the medical management of individuals with HHT. If a pathogenic ENG, ACVRL1 or SMAD4 mutation is identified in an affected proband, diagnostic testing is then available for at-risk family members. HHT is difficult to diagnose on physical examination and medical history alone in the first two decades of life, so identification of a familial mutation allows for early screening and detection of internal AVMs that can be treated prior to the development of serious complications or death.

Efforts to identify new HHT causing genes and regions have proven difficult. Linkage analysis identified two additional HHT loci at chromosome 5q3116 and chromosome 7p1417 over a decade ago, but the genes remain unknown. In 2013, our group reported mutations in the bone morphogenetic 9 (BMP9)/GDF2 gene in three unrelated individuals suspected to have HHT who had previously tested negative for ENG, ACVRL1 and SMAD4.18 Although this has been designated as HHT, type 5 (OMIM 615306), these cases represented <1% of samples submitted to our laboratory for suspected...
HHT, and the clinical findings in these individuals differed from those found in ‘classical HHT’ as defined by the Curacao criteria. RASA1 mutations which cause capillary malformation-AVM (CM-AVM) syndrome (OMIM 608354) have been reported in several individuals clinically suspected to have HHT who had epistaxis and dermal lesions described as telangiectases.\(^\text{19, 20}\) A second form of CM-AVM (CM-AVM2) caused by germline loss-of-function mutations in \(EPHB4\) was recently reported.\(^\text{21}\) CM-AVM2 overlaps with RASA1-related CM-AVM as well as HHT in that clinical features include multifocal capillary malformations, AVMs, telangiectases and epistaxis (reported in several cases).\(^\text{21, 22}\) Because of the phenotypic overlap of these disorders with HHT, next-generation sequencing (NGS) molecular diagnostic panels for HHT and other vascular malformation syndromes (see online supplementary table 1) have become more widely used. Despite these advancements, the molecular diagnosis is still unknown for some HHT cases.

We hypothesized that deep intronic non-coding variants in \(ENG\) and \(ACVRL1\) may disrupt splicing and cause HHT in individuals without an identifiable \(ACVRL1\), \(ENG\) or \(SMAD4\) mutation by current molecular diagnostic testing protocols. Here, we used genome sequencing to evaluate 13 multigeneration families with HHT and a custom NGS panel designed to interrogate the coding and non-coding regions of \(ENG\), \(ACVRL1\) and \(SMAD4\) in 87 additional unrelated individuals suspected to have HHT with no detectable mutation. Although deep intronic variants can be difficult to interpret and characterise, such variants identified in these cases may alter splicing and reveal a hidden disease mechanism of HHT.

**MATERIALS AND METHODS**

**Subjects**

Individuals and families included in the study were clinically confirmed or suspected to have HHT, and had tested negative at ARUP Laboratories for a mutation in a known HHT gene by sequencing of coding regions and intron/exon borders, and duplication/deletion analysis. The majority of individuals from 13 families who underwent genome sequencing were examined at the University of Utah HHT Center of Excellence (Salt Lake City, Utah, USA). The other 10 families were considered suspicious for a ‘telangiectasia syndrome’, but not necessarily HHT. None of the probands or family members of these ten families met three or more diagnostic criteria for HHT. These families had also previously undergone exome sequencing with negative results. Eighty-seven additional individuals with clinical suspicion for HHT were identified from samples submitted to our laboratory for HHT testing. Patient history forms are required with submission of samples and ask ordering clinicians to specify the type, location and number of vascular malformations, as well as other related clinical findings such as epistaxis. Based on information provided on this form, 25/87 had three diagnostic criteria for HHT, 47/87 had two criteria and 15/87 had one criteria. In some cases, ARUP genetic counsellors contacted the ordering clinician to clarify reported findings or obtain more information.

**Genome sequencing and data analysis using RUFUS software**

Genomic DNA was extracted from peripheral blood for 37 individuals using a Gentra Puregene Blood Kit (Qiagen, Valencia, California, USA). Whole genomes from 35 individuals confirmed with or suspicious for HHT, and 2 unaffected individuals from 13 families were sequenced on a HiSeq2500 instrument (Illumina, San Diego, California, USA) using 2×151 paired-end sequencing to ~60× average read depth. Variants were identified using RUFUS, a k-mer-based method capable of detecting variants independent of alignment to a reference sequence (https://github.com/andrewrfarrell/RUFUS). Variants were prioritised through population allele frequency, predicted functional consequence and segregation within and between families.

**Custom NGS panel and data analysis**

Genomic DNA was extracted from peripheral blood using a Gentra Puregene Blood Kit (Qiagen) for 87 individuals on whom testing in our clinical lab had been ordered which included the HHT genes, with negative results. Custom 120 nucleotide RNA baits were designed to specifically target the coding and non-coding regions (including introns, 5'UTRs and 3'UTRs) for \(ACVRL1\), \(ENG\) and \(SMAD4\). Coding regions for \(GDF2\) and \(RASA1\) were also included. RNA baits were tiled at 5x spacing and were in replicates of 10 to increase hybridisation efficiency of the targeted region (~0.1 Mb). Genomic DNA (3μg) was sheared using a Covaris S220 ultrasonicator instrument (Covaris, Woburn, Massachusetts, USA) to 180bp fragments. Illumina adapters were added using the Bravo automated instrument and SureSelect XT kit reagents (Agilent Technologies, Santa Clara, California, USA). Adapter ligated DNA underwent hybridisation with the biotinylated RNA baits for 24 hours at 65°C. Hybridised DNA targets of interest were captured using streptavidin-coated magnetic beads. DNA targets of interest were eluted and barcode/indexed after a series of washes to remove the non-targeted, unbound genome. DNA quality and quantity were assessed using the TapeStation (Agilent Technologies). Samples were pooled (1:1) and sequenced on a HiSeq2500 instrument (Illumina) using 2×100 paired-end sequencing to ~500× average read depth.

Sequences were aligned to the human genome reference (hg19) sequence using the Burrows-Wheeler Alignment tool (0.5.9) with default parameters.\(^\text{23}\) PCR duplicates were removed using the Samtools package,\(^\text{24}\) and base quality score recalibration, local realignment and variant calling were performed using the Genome Analysis Toolkit (GaTK V.1.3).\(^\text{25}\) Alamut software was used to predict the effect of rare and novel variants on splicing. Potential pathogenic variants were confirmed using Sanger sequencing and further characterised using cDNA sequencing.

**Sanger sequencing**

Custom primers were designed to confirm non-coding variants of interest identified. Primer sequences are available on request. Amplicon fragments were bidirectionally sequenced with universal M13 primers using the Big Dye Terminator V3.1 cycle sequencing kit and an ABI 3730 DNA Analyzer (Life Technologies, Carlsbad, California, USA). Sequences were compared with the \(ACVRL1\) and \(ENG\) reference sequences (NM_000020.2 and NM_000118.3, respectively) using Mutation Surveyor (SoftGenetics, State College, Pennsylvania, USA).

**cDNA sequencing**

RNA was extracted from peripheral blood and converted into cDNA for select cases using reverse transcriptase with random primers. Exonic primers were used to specifically amplify and Sanger sequence the region of \(ACVRL1\) that had the potential splicing defect. Primer sequences are available on request.
RESULTS

Genome sequencing reveals non-coding variants in ACVRL1 that disrupt splicing

Thirty-five symptomatic individuals and 2 asymptomatic family members from 13 families with HHT or suspected to have HHT, who had previously tested negative by exome sequencing, underwent genome sequencing. All families had also previously tested negative for a mutation in ENG, ACVRL1 and SMAD4 by Sanger sequencing and deletion/duplication analysis by multiplex ligation-dependent probe amplification (MLPA). Genomes were sequenced to ~60× average coverage and novel and rare variants were analysed.

A novel deep intronic heterozygous variant (ACVRL1 c.1378-216C>G) was identified and confirmed by Sanger sequencing in one four-generation family (family 1) with HHT (figure 1A and table 1, case 1). Eighteen members of this family had been evaluated at the University of Utah HHT Clinic. Nine were affected with HHT on clinical grounds, while six were suspected affected and three considered unknown. In addition to epistaxis and characteristic telangiectasias, six of the affected had pulmonary AVM, five requiring treatment by transcatheter embolisation. Previous linkage analysis had shown that the disease in the family linked to chromosome 12q, consistent with linkage to ACVRL1 on chromosome 12q13 (data not shown). This deep intronic ACVRL1 variant (ACVRL1 c.1378-216C>G) was predicted by multiple splice site prediction programmes to create a new ‘AG’ splice acceptor site and alter splicing (table 1).

To confirm the predicted splicing defect in this family, RNA was obtained from an affected family member and converted to cDNA. Exon-specific primers showed the presence of two mRNA species, one correlating in size to the wild-type band and another higher mutant band (figure 1B). Sanger sequencing of the higher band confirmed the predicted partial retention of ACVRL1 intron 9 in the aberrantly spliced product (figure 1C). This variant tracked with the disease in the family, as all nine affected family members had the variant (figure 1D). Of note, a rare benign variant (rs111710113, 0.004 minor allele frequency) at the same location but a different nucleotide change (ACVRL1 c.1378-216C>T) was not predicted to affect splicing as it does not create a new ‘AG’ acceptor site.

Novel ACVRL1 translocation identified using genome sequencing

Another family (family 2) that underwent genome sequencing was found to carry a novel heterozygous ACVRL1 intron 9:chromosome 3 translocation (figure 1 and table 1, case 2). The proband aged 27 years has epistaxis, telangiectasias in characteristic locations and a pulmonary AVM. Her mother has epistaxis and telangiectasias. PCR primers flanking the t(12,3)(q13;p21) translocation were used to confirm the translocation in the affected family members by Sanger sequencing their genomic DNA across the breakpoints (chr12:52313774(chr3:48252144) (figure 1A). The chromosome 3 breakpoint was located in an intergenic region; however, the chromosome 12 breakpoint was located in ACVRL1 intron 9 (c.1348-769). The translocation is expected to interrupt the ACVRL1 gene, because it would lead to a loss of the last coding exon (exon 10). The proband and her affected mother carried the variant, whereas her unaffected grandmother did not carry the variant (figure 1B). These translocation-specific primers did not give amplification in the unaffected family member or wild-type control (see online supplementary figure 1). This is the first report of a translocation causing HHT. Although Sanger sequencing confirmed the breakpoints of the translocation for chromosome 12, no additional blood sample was available for karyotyping to confirm at the chromosomal level.

Based on genome sequencing results in these original 13 families, two families with clinically confirmed HHT had a deep intronic ACVRL1 variant involving intron 9 that caused HHT. The remaining 11 families (10 of whom had been considered to be atypical in some way for HHT) did not have an identifiable causal variant in the non-coding regions of ENG, ACVRL1 or SMAD4, and genes examined as part of the HHT clinical
differential (GDF2 and RASA1) were also negative. Genome data are still being analysed for these families for gene discovery purposes.

**ACVRL1 intron 9 mutation hotspot discovered using NGS panel sequencing**

Given the genome sequencing results in the initial 13 families, 87 additional unrelated individuals whose samples had been submitted to our clinical laboratory for testing which included the HHT genes, with a negative result, were evaluated using an ‘HHT genome’ panel. This small custom NGS panel was designed to specifically target the coding and non-coding regions of the ACVRL1 intron 9 variant predicted by multiple SpliceFinder classifications compared with native splice site (89.5, 78.8).

### Table 1 Clinical and molecular findings of eight unrelated probands with a novel ACVRL1 intron 9 variant

<table>
<thead>
<tr>
<th>Case #</th>
<th>Gene</th>
<th>IVS #</th>
<th>Nucleotide change</th>
<th>Protein effect</th>
<th>Predicted splicing scores*</th>
<th>Variant classification</th>
<th>Age, sex</th>
<th>Fam Hx</th>
<th>E</th>
<th>T</th>
<th>AVM</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACVRL1</td>
<td>9</td>
<td>c.1378-216C&gt;G</td>
<td>Splicing† Partial retention of IVS9</td>
<td>89.2, 84.8</td>
<td>Pathogenic</td>
<td>63 F</td>
<td>Yes</td>
<td>x</td>
<td>PAVM, HAVM</td>
<td>Tracked in family</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ACVRL1</td>
<td>9</td>
<td>t(12,3)(q13,p21) translocation</td>
<td>Splicing†</td>
<td>NA</td>
<td>Pathogenic</td>
<td>27 F</td>
<td>Yes</td>
<td>x</td>
<td>PAVM</td>
<td>Tracked in family</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ACVRL1</td>
<td>9</td>
<td>c.1378-274C&gt;G</td>
<td>Splicing† Partial retention of IVS9</td>
<td>87.9, 79.5</td>
<td>Pathogenic</td>
<td>41 M</td>
<td>Unknown</td>
<td>x</td>
<td>PAVM</td>
<td>Multiple T in characteristic locations</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ACVRL1</td>
<td>9</td>
<td>(c.1378-156C&gt;A; 1378-155T&gt;G)</td>
<td>Splicing† Partial retention of IVS9</td>
<td>88.5, 85.1</td>
<td>Pathogenic</td>
<td>64 M</td>
<td>Yes</td>
<td>x</td>
<td>Colon AVM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ACVRL1</td>
<td>9</td>
<td>c.1378-131C&gt;G</td>
<td>Splicing† Partial retention of IVS9</td>
<td>80.3, 73.8</td>
<td>Pathogenic</td>
<td>47 F</td>
<td>Yes</td>
<td>x</td>
<td>PAVM, HAVM</td>
<td>E monthly in childhood; T on lips</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ACVRL1</td>
<td>9</td>
<td>c.1378-78T&gt;G</td>
<td>Partial retention of IVS9 predicted</td>
<td>91.7, 80.0</td>
<td>Pathogenic</td>
<td>70 F</td>
<td>Yes</td>
<td>x</td>
<td>PAVM, HAVM</td>
<td>GI and skin T</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>ACVRL1</td>
<td>9</td>
<td>c.1378-69C&gt;A</td>
<td>Splicing† Partial retention of IVS9</td>
<td>91.5, 84.5</td>
<td>Pathogenic</td>
<td>49 M</td>
<td>Yes</td>
<td>x</td>
<td>Diffuse PAVM</td>
<td>Mother has E and a stroke</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>ACVRL1</td>
<td>9</td>
<td>c.1378-69C&gt;A</td>
<td>Splicing† Partial retention of IVS9</td>
<td>91.5, 84.5</td>
<td>Pathogenic</td>
<td>80 F</td>
<td>Yes</td>
<td>x</td>
<td>PAVM</td>
<td>T on lips and hands; Fam Hx of E, T, CAVM and PAVM</td>
<td></td>
</tr>
</tbody>
</table>

*Human Splicing Finder and SpliceSite Finder score compared with native splice site (89.5, 78.8).†Molecular characterisation using cDNA sequencing was performed to confirm the splicing defect.

IVS, intervening sequencing; ESE, exon-splicing enhancer; Fam Hx, family history; E, epistaxis; T, telangiectasia; AVM, arteriovenous malformation; PAVM, pulmonary AVM; HAVM, hepatic AVM; CAVM, cerebral AVM; GI, gastrointestinal.

For case 6 with the ACVRL1 c.1378-78T>G variant, the Human Splicing Finder and SpliceSite Finder prediction scores (91.7 and 80, respectively) were stronger compared with the native intron 9 splice acceptor (89.5, 78.8). The remaining 81 cases did not have an identifiable causal variant in the non-coding regions of ENG, ACVRL1 or SMAD4, and genes examined as part of the HHT clinical differential (GDF2 and RASA1) were also negative.

**DISCUSSION**

These results reveal that non-coding region variants play a larger role in HHT than previously thought. Overall, 8% of these cases (8/100) with some suspicion of HHT, with previous negative molecular testing, had a novel non-coding ACVRL1 intron 9 variant that was proven to disrupt splicing. However, it should be emphasised that this cohort consists of cases that were considered suspicious for HHT by a clinician; but only 28/100 were reported to have three or more criteria for HHT, 47/100 with two and 15/100 with only one clinical criteria. Thus, 8% clearly underestimates the detection rate of pathogenic intron 9 variants in individuals who meet clinical diagnostic criteria for HHT, but have previously tested ‘negative’ in a clinical genetics laboratory. The detection rate for a non-coding region variant of ACVRL1 intron 9 was 29% (8/28) for individuals/families reported to have three or more Curaçao criteria.

It is noteworthy that all eight cases in which a non-coding region variant was identified in this study met clinical diagnostic criteria for HHT. Seven of the eight cases/families were diagnosed at an HHT Center of Excellence in North America, with the eighth case diagnosed by a medical geneticist with a particular interest in HHT. This despite the fact that overall the cases sent from an HHT Center or Genetics Clinic were in the minority. It is our experience that the accuracy and rigor of which the Curaçao diagnostic criteria are applied is very dependent on...
a clinician’s experience with vascular malformation syndromes; in particular, knowing what constitutes cutaneous lesions and nosebleeds that are typical for HHT.

Based on our results, we estimate that pathogenic variants in ACVRL1 intron 9 alone account for about one-third (~29%) of cases with clinically confirmed HHT who do not have a genetic diagnosis using testing currently available in most molecular diagnostic laboratories. Furthermore, we estimate that ACVRL1 intron 9 variants account for approximately 1% of all HHT cases. This is based on our experience that ~97% of cases with clinically confirmed HHT have a causative variant in the coding region or exon/intron boundaries of ACVRL1, ENG or SMAD4 (McDonald, unpublished results). To increase clinical sensitivity, molecular diagnostic testing algorithms should be expanded to include this ~300bp region. Any novel variants identified that create a new ‘AG’ splice acceptor site should be investigated further using confirmatory RNA studies to confirm the splicing defect.

One family had an ACVRL1 intron 9:chromosome 3 translocation, the first translocation identified to cause HHT. The discovery of this translocation represents a new disease mechanism in HHT. In this family, ACVRL1 transcription is predicted to be disrupted since the translocation is located before the last exon of the gene. While this novel translocation was identified using genome sequencing, this event could have been detected using karyotyping. It is likely that the prevalence of translocations in HHT is very rare.

Our findings suggest that deep intronic splice variants in ACVRL1 intron 9 indicates a previously unrecognised mutational hotspot. Intriguingly, all variants identified in our cohort except for the translocation were located in a CT-rich region of ACVRL1 intron 9 (figure 3B), and all result in the generation of a new ‘AG’ splice acceptor. It is well known that the canonical acceptor splice site region contains high CT-rich sequences; however, this CT-rich region usually does not extend to deep intronic regions. All identified splice site variants in our clinical cases are located within 275 bases from exon 10 within intron 9. A sequence structure analysis up to 317 bases from exon 10 revealed that the CT ratio of this region is 85%, which is similar to the canonical acceptor splice site. The high CT ratio in ACVRL1 intron 9 decreases after 317bp.

It is likely that any variant residing in the ~300bp CT-rich region that creates a new ‘AG’ sequence in ACVRL1 intron 9 will activate a new cryptic splice site. To test this hypothesis, we simulated this effect by altering every possible variant in this region to create a new ‘AG’ site using Alamut software in which multiple splice prediction scores were obtained and compared with the native ‘AG’ intron 9 acceptor splice site. In nearly all (42) ‘AG’ variant simulations, the simulated variants were predicted to create a strong cryptic acceptor splice site (figure 3C). When
Figure 3  ACVRL1 intron 9 mutation hotspot. (A) ACVRL1 intron 9 variant (c.1378-274C>G) identified by panel sequencing depicts the CT-rich region below. (B) Diagram of ACVRL1 intron 9 variants (*) in which most create a new 'AG' splice acceptor in the CT-rich region causing aberrant splicing and hereditary haemorrhagic telangiectasia. (C) Sequence of the ACVRL1 intron 9 hotspot region depicting the bases in bold in which variants at that location were simulated by mutation to create a new 'AG' acceptor site and evaluated using Alamut. Starred (*) bases had a splice site prediction score from either the Human Splicing Finder or the SpliceSite Finder that was higher than the score from the native splice acceptor, AG. Boxed bases represent the location of a variant that affected splicing in a clinical case. Two base locations are indicated with two stars because that one location could create two 'AG' sites depending on the in silico variant. IVS, intervening sequencing.

we compared the prediction scores of these simulated ‘AG’ variants with the scores from the native acceptor site, we found that 30 of 42 simulated AG sites had a higher prediction score than the native site (figure 3C, see starred bases). This effect was not observed when the same exercise was applied to other introns in ACVRL1 such as in intron 6 where only 13 ‘AG’ simulated variants were predicted to create a new cryptic splice site.

The polypyrimidine tract is a well-known motif for recruiting spliceosome proteins to the pre-mRNA. However, it has not been well studied if the polypyrimidine tract or simply the presence of a CT-rich sequence may increase the chance of a new AG acceptor site as a disease-causing mechanism. It requires more computational and some functional studies to answer this question.

In this study, we have shown the importance of non-coding region changes on the HHT disease mechanism as well as detailed clinical phenotyping to reveal these regions or variants. Genome and custom panel sequencing revealed a non-coding region pathogenic variant in 8 of 100 individuals/families suspected to have HHT, with no causative variant detected to date. We previously reported that non-coding mutations in the 5’UTR region of ENG cause HHT in ~1%–2% of cases. Although non-coding variants can be difficult to interpret and characterise, our studies highlight the importance of including these regions in molecular diagnostic testing for HHT.

CONCLUSION
Non-coding variants in ENG and ACVRL1 play a greater role in causing HHT than previously thought, suggesting molecular testing expansion to include non-coding regions, particularly the ACVRL1 intron 9 CT-rich mutation hotspot. Variants in these regions may explain many of the families with definite HHT according to Curaçao criteria who have not had a causative genetic variant detected to date in a clinical laboratory. Using RUFUS analysis of genome data, a new causal mechanism of
HHT was identified in one family who had a novel ACVRL1 intron 9: chromosome 3 translocation. Despite the difficulty of interpreting deep intrinsic variants, our study highlights the importance of non-coding regions in the disease mechanism of HHT. In particular, the addition of the −300bp CT-rich region of ACVRL1 intron 9 to HHT molecular diagnostic testing algorithms will improve clinical sensitivity.

Author affiliations
1 ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT, Salt Lake City, USA
2 Department of Pathology, University of Utah, Salt Lake City, Utah, USA
3 HHT Center, Department of Radiology, University of Utah, Salt Lake City, Utah, USA
4 USTAR Center for Genetic Discovery, University of Utah, Salt Lake City, Utah, USA
5 Department of Pediatric Genetics, Zeynep Kamil Maternity and Children’s Training and Research Hospital, Istanbul, Turkey
6 Division of Cardiovascular Medicine, Department of Medicine, University of Utah, Salt Lake City, Utah, USA
7 Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds, UK
8 Medical Genetics, Department of Pediatrics, Mass General Hospital for Children, Boston, Massachusetts, USA
9 Department of Internal Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correction notice
This article has been corrected since it was published online first. Details related to Dr Farrell’s funding have been added to the ‘Funding’ section.

Acknowledgements
The authors would like to thank the patients and their family members for participating in this research. The authors would also like to thank the members of the ARUP Molecular Genetics and Genomics Clinical Laboratories for assisting in the sequence analysis of these patients.

Contributors
WWD planned, conducted all experiments, analysed the data and drafted the manuscript. JM and PBT planned experiments, obtained all samples, interpreted the manuscript. AF, MV, RM and GM performed the bioinformatics analysis, assisted in the interpretation of results and reviewed the manuscript. GA, PJ, CWV and EB assisted with experiments, data analysis and reviewed the manuscript. KW, JT, AEL and REP saw the patients, obtained samples and reviewed the manuscript.

Funding
This study was funded by Cure HHT, the Chan Soon-Shiong Family Foundation (Heritage 1K Utah Genome Project) and the ARUP Institute for Clinical and Experimental Pathology. Dr Gulsen Akay was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK), with 22/19 Postdoctoral Research Fellowship. Dr Andrew Farrell was funded from a T1L fellowship (T1L TR001066).

Competing interests
None declared.

Patient consent
Not required.

Ethics approval
This study was approved by the University of Utah Institutional Review Board (IRB #7275, 35637 and 20480).

Provenance and peer review
Not commissioned; externally peer reviewed.

REFERENCES