


## ORIGINAL ARTICLE

# Hereditary hemorrhagic telangiectasia: First demonstration of a founder effect in Italy; the *ACVRL1* c.289\_294del variant originated in the country of Bergamo 200 years ago

Anna Sbalchiero<sup>1</sup>  | Yasmin Abu Hweij<sup>1</sup> | Tommaso Mazza<sup>2</sup> |  
Elisabetta Buscarini<sup>3</sup> | Claudia Scotti<sup>4</sup> | Fabio Pagella<sup>5,6</sup> | Guido Manfredi<sup>3</sup> |  
Elina Matti<sup>6</sup> | Giuseppe Spinozzi<sup>6</sup> | Carla Olivieri<sup>1</sup> 

<sup>1</sup>General Biology and Medical Genetics Unit, Department of Molecular Medicine, University of Pavia, Pavia, Italy

<sup>2</sup>Laboratory of Bioinformatics, Fondazione IRCCS Casa Sollievo della Sofferenza-Mendel, San Giovanni Rotondo (FG), Italy

<sup>3</sup>UOC of Gastroenterology-Reference Centre for HHT, ASST Ospedale Maggiore di Crema, Crema (CR), Italy

<sup>4</sup>Immunology and General Pathology Unit, Department of Molecular Medicine, University of Pavia, Pavia, Italy

<sup>5</sup>UOC of Otorhinolaryngology, Department of Surgical Sciences, University of Pavia, Pavia, Italy

<sup>6</sup>UOC of Otorhinolaryngology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

## Correspondence

Carla Olivieri, General Biology and Medical Genetics Unit, Department of Molecular Medicine, University of Pavia, Via Forlanini 14 – 27100, Pavia, Italy.

Email: [carla.olivieri@unipv.it](mailto:carla.olivieri@unipv.it)

## Funding information

Italian Ministry of Education, University and Research to the DMM-University of Pavia “Dipartimenti di Eccellenza (2018-2022)”; Italian Ministry of Health (Ricerca Corrente 2018-2020) and 5x1000 voluntary contribution grant

## Abstract

**Background:** Hereditary hemorrhagic telangiectasia (HHT) is an autosomal dominant vascular disorder, affecting 1:5000 individuals worldwide. All the genes associated to the disease (*ENG*, *ACVRL1*, *SMAD4*, *GDF2*) belong to the TGF- $\beta$ /BMPs signaling pathway.

We found 19 HHT unrelated families, coming from a Northern Italy region and sharing the *ACVRL1 in-frame* deletion c.289\_294del (p.H97\_N98).

**Methods:** To test the hypothesis of a founder effect, we analyzed 88 subjects from 19 families (66 variant carriers, showing clinical signs of HHT, and 22 non-carriers, unaffected) using eight microsatellite markers within 3.7 Mb around the *ACVRL1* locus. After the haplotype reconstruction, age estimation of the variant was carried out.

**Results:** We observed a common disease haplotype in 16/19 families, while three families showed evidence of recombination around the *ACVRL1* locus. The subsequent age estimation analyses suggested that the mutation occurred about 8 generations ago, corresponding to about 200 years ago. We also present novel in silico and modeling data supporting the variant pathogenicity: the deletion alters the protein stability and removes the unique extracellular glycosylation site.

**Conclusion:** We have demonstrated, for the first time, a “founder effect” for a HHT pathogenic variant in Italy.

## KEYWORDS

*ACVRL1*, age estimation, common ancestor, founder effect, HHT, Rendu-Osler-weber syndrome

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals LLC.

## 1 | INTRODUCTION

Hereditary hemorrhagic telangiectasia (HHT), also known as Rendu–Osler–Weber syndrome, is a rare autosomal dominant vascular disorder affecting about 1 in 5000 individuals worldwide (Govani & Shovlin, 2009).

The vascular dysplasia in HHT is characterized by the presence of arteriovenous malformations (AVMs) in internal organs and telangiectases on mucocutaneous surfaces. Telangiectases are frequently and typically found in patient's nasal and oral cavities, on face, fingertips and gastrointestinal mucosa, whereas AVMs are observed in the lungs, liver, and central nervous system, with a variable incidence, related to the disease-causing gene (Lesca et al., 2007).

The clinical diagnosis of HHT is established if at least three of the following features (the so called “Curaçao criteria”) are present: epistaxis, mucocutaneous telangiectases, visceral AVMs, and the presence of a first degree relative diagnosed according to the same criteria (Shovlin et al., 2000). Epistaxis is the most common clinical sign: it is present in more than 90% of patients and frequently begins in childhood (Lesca et al., 2007).

HHT-causing variants have been found in four genes belonging to the TGF- $\beta$ /BMPs (Transforming Growth Factor- $\beta$ / bone morphogenetic proteins) signaling pathway: *ENG*, *ACVRL1*, *SMAD4*, *GDF2*.

*ENG* is located on the long arm of chromosome 9 (9q34.11) and codes for ENDOGLIN (or *ENG*), a type III auxiliary TGF- $\beta$  receptor. Pathogenic variants in this gene are associated to HHT1 (OMIM #187300) and to date more than 340 pathogenic variants have been reported in *ENG* ([http://arup.utah.edu/database/ENG/ENG\\_welcome.php](http://arup.utah.edu/database/ENG/ENG_welcome.php)).

*ACVRL1* defects are instead responsible for HHT2 (OMIM #600376); to date more than 260 pathogenic variants have been described for this gene ([http://arup.utah.edu/database/ACVRL1/ACVRL1\\_welcome.php](http://arup.utah.edu/database/ACVRL1/ACVRL1_welcome.php)). *ACVRL1* chromosomal locus is 12q13.13. This gene codes for Activin A Receptor Type II - Like I (or *ALK1*), a type I serine/threonine kinase receptor.

*SMAD4* (or *MADH4*: Mothers Against Decapentaplegic, *Drosophila*, Homolog of, 4) is positioned on chromosome 18 (18q21.2) and encodes for the homonymous protein, *SMAD4*, a nuclear effector of the signal transduction pathway. Pathogenic variants in *SMAD4* lead to the Juvenile Polyposis/Hereditary Hemorrhagic Telangiectasia Syndrome (JPHT; OMIM #175050) and 60 pathogenic variants are now known ([http://arup.utah.edu/database/SMAD4/SMAD4\\_welcome.php](http://arup.utah.edu/database/SMAD4/SMAD4_welcome.php)).

*GDF2* (Growth Differentiation Factor 2) maps on chromosome 10 (10q11.22) and codes for BMP9, the main physiological ligand of *ALK1*. Only few HHT-related variants have been referred to BMP9 (Wooderchak-Donahue et al., 2013).

After ligand/receptor type I and II binding, supported by type III receptor, the signal is transduced, by phosphorylation cascades involving *SMAD1/5*, to *SMAD4*; the *SMAD4/SMAD1,5* complex translocates to the nucleus and, together with other transcription factors, regulates target genes expression (Fernández et al., 2006). Target genes are not fully known yet.

It is currently accepted that the majority of the pathogenic variants found in HHT Patients are “family-exclusive” or “private” (Faughnan et al., 2011), although in several cases, the same mutation is reported, with low incidence in different populations (Lesca et al., 2008; Olivieri et al., 2007; ARUP Mutation Databases, 2022). In our database, which collects more than 500 families, mutations recurring in more than one family are about 20%.

Although HHT is widespread worldwide, in some regions a much higher prevalence has been observed suggesting a founder effect (Faughnan et al., 2020). Examples are the County of Fyn in Denmark (Kjeldsen et al., 1999), the Netherlands Antilles (Gallione et al., 2000), the Northern part of Japan (Dakeishi et al., 2002), the Haut-Jura mountains in France (Lesca et al., 2008) and Norway (Heimdal et al., 2016); in some cases, a founder effect was proven by molecular analyses (Dakeishi et al., 2002; Gallione et al., 2000; Heimdal et al., 2016; Lesca et al., 2008).

We report here the findings on 19 Italian families, all coming from the same Northern Italy restricted area, and carrying the same disease-causing variant: *ACVRL1* (NM\_000020.2) c.289\_294delCACAAAC (p.H97\_N98del). To the best of our knowledge, this variant has never been reported in non-Italian patients and, in Italy, it has been found only in individuals from nearby the city of Bergamo. We provide additional molecular evidence for the pathogenicity of the variant and data supporting the hypothesis of a founder effect. We also underline the relevance of our findings for the National Health Service as the incidence of this rare disease can be higher in the Bergamo area than in other regions. Evidence for the presence of a founder variant in this territory is useful to accelerate the genetic diagnosis of younger or mildly affected patients coming/having ancestors from the abovementioned province.

## 2 | MATERIALS AND METHODS

### 2.1 | Ethical compliance

This study was approved by the Ethics committee in the area of Pavia (reference number 1–29/1/14) and conforms the declaration of Helsinki.

According to the Italian Bioethics laws, each Subject signed a written informed consent to allow the use of his/her biological sample for diagnostic and research purposes.

## 2.2 | Patients

Subjects involved in the study were diagnosed and classified as “affected”/“non affected”, following the “Curaçao Criteria” (Shovlin et al., 2000) by clinicians (EB, FP, GM and EM) from two HHT Italian Reference Centers in Pavia and in Crema.

In a previous work by our laboratory (Olivieri et al., 2007), we observed an unusual “cluster” for the *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) (NM\_000020.2) variant: the variant was identified only in Italy, only in our cohort, and only in families coming from/having ancestors from the area of Bergamo (Northern Italy). For this reason, we called it the “Bergamasca variant” (“Bergamasca”, in Italian, is the feminine adjective for “coming from Bergamo”).

For this study, we searched the HHT database of the General Biology and Medical Genetics Unit (more than 2000 samples including Patients and relatives) for families in which the Index Case carries the “Bergamasca variant” and we found 98 individuals belonging to 19 apparently unrelated families. Of them, 88 were included in this study, whereas 10 subjects were excluded because DNA was no more available.

To the purpose of both co-segregation analysis and haplotype reconstruction, we included both carriers and non-carrier family members. Of note, for the co-segregation studies, we excluded subjects without any additional relatives and subjects younger than 40 years (24/88 subjects), while for haplotype reconstruction we considered all the subjects (Pedigrees in Figure 4 and Supplementary Figure S1).

## 2.3 | Molecular genetics analyses

The *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) (NM\_000020.2) variant was identified by Sanger sequencing starting from 3 ml of peripheral blood (EDTA) as previously reported (Olivieri et al., 2007). MutationTaster, MutPred Value and Varsome tools were used to define the variant pathogenicity (Kopanos et al., 2019; Pejaver et al., 2020; Schwarz et al., 2014).

## 2.4 | Protein modeling and glycosylation prediction

The ALK1 extracellular domain, where the c.289\_294delCACAAC variant falls, is involved in the

receptor/ligand binding and receptor/receptor interactions. Modeling of the extracellular domain of ALK1 carrying the H97\_N98del (ALK1 H97\_N98del) was performed by the Swiss-Model server (Waterhouse et al., 2018), providing both the protein sequence carrying the deletion and the structure of the extracellular domain of the wild-type ALK1 (ALK1-EC) derived from its complex with bone morphogenetic protein 10 (BMP10; PDB entry: 6SF1) as a template. Validation was based on the Global Quality Estimate provided by the Swiss Prot server, including QMEAN, (Qualitative Model Energy Analysis) (Benkert et al., 2011), which is a composite scoring function describing the major geometrical aspects of protein structures. Models were visualized with Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). PROVEAN (Protein Variation Effect Analyzer) is a software tool which predicts whether an amino acid substitution or indel has an impact on the biological function of a protein (Choi et al., 2012). Prediction of glycosylation sites was performed by the NetOGlyc 1.0 server (Steenftoft et al., 2013).

## 2.5 | Short tandem repeats location and haplotype analysis

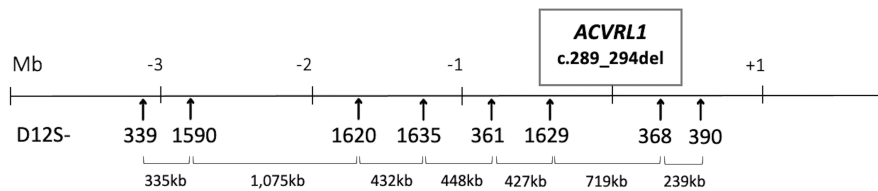
We studied eight polymorphic STR markers spanning 3.7 Mb around the *ACVRL1* locus for the haplotype analysis: DS12S339, D12S1590, D12S1620, D12S1635, D12S361, D12S1629, D12S368, and D12S390. Markers were selected using the UCSC genome browser (<http://genome.ucsc.edu>). The range of heterozygosity for these eight STR markers was 52%–85%. Schematic distribution of markers and relative distances in kb are reported in Figure 1.

The genomic DNA from each individual was amplified by PCR using fluorescently 6-Fam or Hex labeled forward primers of each microsatellite (Supplementary Table S1).

The amplicons were run on a 2% w/v agarose gel containing GelRed® intercalating agent (Biotinum) and visualized by ChemiDoc® (Biorad) to check the quality and quantity of PCR products.

Fragments from each subject were mixed and analyzed by capillary electrophoresis (3500 Dx Genetic Analyzer-Applied Biosystem®) using GeneScan™ 600 LIZ™ (ThermoFisher) as size standard. Two different mixes were run to better define single microsatellite. Mix 1 included: D12S1590, D12S1635, D12S390, D12S368, D12S361, D12S1620; mix 2 included D12S339 and D12S1629.

The fragments lengths were then established using the GeneMapper™ software (ThermoFisher), and haplotypes were manually reconstructed and confirmed by analyses reported in the “Estimate of mutation age” section.



**FIGURE 1** STR markers location compared to the *ACVRL1* variant locus. The microsatellites analyzed ordered by their locations are: D12S339, D12S1590, D12S1620, D12S1635, D12S361, D12S1629, D12S368, D12S390. The total length of the included region corresponds to 3.7 mb.

## 2.6 | Estimate of mutation age

Age estimation of the *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) variant was carried out with the DMLE+2.3 software package (Reeve & Rannala, 2002). Input data to DMLE included an encoded description of the full haplotypes of the mutation carriers and non-carriers, the latter being used as controls from the general population for the eight examined markers. Genetic distances between markers were retrieved from the Marshfield Comprehensive Human Genetic Maps (<https://www.biost.wisc.edu/~kbroman/publications/mfdmaps/>) or estimated based on physical distances (in Mb) given in UCSC Genome Browser (GRCh37/hg19 assembly) and considering 1 Mb~1 cM. The population growth rate for the province of Bergamo from the year 1300 to the present was estimated to be 0.0526 (Caleca et al., 2014). Three different estimates for the proportion of sampled mutation-carrying chromosomes were used: 0.015, 0.01, and 0.005. These indices were subsequently used for mutation age estimates, which were averaged over 10 independent simulations.

## 3 | RESULTS

### 3.1 | Genetic classification of the variant

In all families, we observed, with no exceptions, cosegregation of the variant with the presence of the clinical picture of HHT (See Supplementary Figure S1). We obtained as a total result, 66 subjects, clinically affected and carriers of the variant, and 22 relatives, clinically unaffected and non-carriers.

The *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) variant is classified as of “uncertain significance” according to the ARUP HHT database and has never been reported in non-Italian families.

The variant is not present in GnomAD or in ClinVar (Karczewski et al., 2020; Landrum et al., 2014).

Varsome software predicts this variant as “Likely Pathogenic” according to ACMG (American College of Medical Genetics) classification (Kopanos et al., 2019).

MutationTaster predicts this variant as “disease causing” (Schwarz et al., 2014) and MutPred-Indels returns a

pathogenicity score of 0.882 in a scale from 0 (benign) to 1 (pathogenic) (Pejaver et al., 2020).

### 3.2 | Effects of the variant on the ALK1 protein structure

The structural model of ALK1 H97\_N98del generated by Swiss Model (proMod 3.2.0) showed a global model quality estimate (GMQE score) of 0.56 and a Quality Mean (QMEAN) score of  $-1.06$ , indicating a good reliability. Not unexpectedly, however, the local QMEAN scores were worse for the stretch of C-terminal residues where the H97\_N98 deletion was located (pink salmon regions, Supplementary Figure S2) and where the deleted polypeptide chain tried to reach an alternative, energetically acceptable, new conformation compared to the wild type. As QMEAN is mainly a geometrical indicator, this suggests a suboptimal geometry of this protein region in the deleted version of the polypeptide.

Structural analysis shows that the deletion removes two residues forming a short loop located on the surface of ALK1-EC (Figure 2a, pink salmon). In Figure 2b the prominent structural change induced by the mutation on the surface profile is shown (disappearance of the pink salmon-colored surface). PROVEAN analysis of the mutated protein returned a value of  $-17.498$ , consistent with the totally new configuration of hydrogen bonds compared to the wild type: the side chain of residue N96 in ALK1 H97\_N98del has an orientation similar to the one of H97 in the wild type and can replace it in the formation of the conserved hydrogen bond to T35, while the V99 carbonyl atom establishes contacts with R67 (Figure 3). Interestingly, the NetNGlyc 1.0 server predicted a glycosylation site at N98 (score 0.7182), also reported by the UNIPROT database, which is completely removed by the “Bergamasca variant.”

### 3.3 | Microsatellite analyses, haplotype reconstruction and age estimate

The results of microsatellite analyses, including haplotype reconstruction, are available as Supplementary Figure S1.



Examples of family trees and haplotype reconstruction is reported in Figure 4.

Comparing the results of microsatellite markers analysis, we found that the great majority of affected individuals (49/66) carrying the *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) variant exhibit a “conserved” haplotype; 11/66 subjects exhibit only one recombinant allele in marker D12S339 or D12S368; a “mixed” haplotype is observed in 4/66 subjects, while a haplotype partially overlapping the “mixed” one, but with a further recombination, was found in the remaining two patients (Table 1).

Ten replicas of three distinct analyses were performed setting the population growth rate per generation to 0.052 but using different estimates of the proportion of sampled mutation-carrying chromosomes (0.015, 0.01, and 0.005). Average age estimations for the three analyses were similar: eight generations for the first two batches of analyses and seven generations for the latter. Hence, the age was estimated to be ~eight generations (95% credible set: 6–11). Assuming an interval of 25 years per generation, this corresponds to the mutation being ~200 (165–265) years old (Figure 5b).

## 4 | DISCUSSION

### 4.1 | The variant, its structure and functional significance

We identified the *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) variant several years ago (Olivieri et al., 2002). Although reported as “Pending Classification” in the ARUP HHT database ([http://arup.utah.edu/database/ACVRL1/ACVRL1\\_welcome.php](http://arup.utah.edu/database/ACVRL1/ACVRL1_welcome.php)), its potential pathogenicity was supported when several bioinformatic tools as MutationTaster, MutPred Value and Varsome were interrogated, as they mostly interpreted it as damaging.

This *in silico* evaluation perfectly fits with our previous observation that the variant is pathogenic as it always co-segregates with the disease in all the pedigrees we observed (Supplementary Figure S1).

Three-dimensional modeling studies added further evidence supporting the pathogenicity of the variant. In fact, the deletion greatly alters protein stability, causing a different configuration of hydrogen bonds as compared to the normal counterpart. The two deleted amino acids, H97 and N98 are not involved in ligand binding (see PDB ID: 6SF1, structure of the BMP10 complexed with the extracellular domain of ALK1). However, structural analysis predicts that this *in frame* deletion is deleterious (See Supplementary Figure S2). Moreover, the deletion removes the unique putative N-glycosylation site of ALK1 at N98. This could be destructive in terms both of folding and stability, and could impair the correct membrane localization of the protein (Alt et al., 2012; Hammond et al., 1994; Mitra et al., 2006).

### 4.2 | Founder effect of the variant and its age of occurrence

When we first observed the *ACVRL1* c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) variant, it was immediately clear that it was quite common among the variants we were collecting in our database. Indeed, the database available at the General Biology and Medical Genetics Unit includes 244 different pathogenic variants, 100 for *ENG* and 144 for *ACVRL1*; among them, c.289\_294delCACAAC (p.H97\_N98 *del in-frame*) is the most frequent, as it is present in 76 individuals (6.7% of the total; 10% considering *ACVRL1* variants only). This variant has never been reported outside of Italy, and in Italy it was identified only in a restricted area in the surroundings of the city of Bergamo. All these evidences strongly suggested the hypothesis of a founder effect, probably due to the appearance of a *de novo* variant.

The haplotype analysis we here report confirms this hypothesis, with most patients (60/66) sharing the same haplotype including individuals showing recombination for one microsatellite allele.

In fact, the most common haplotype is present in 16/19 families: in 11 of them the haplotype is conserved in all analyzed generations (families A, B, C, E, G, H, J, K, Q, R, S; Supplementary Figure S1, red alleles) whereas in five

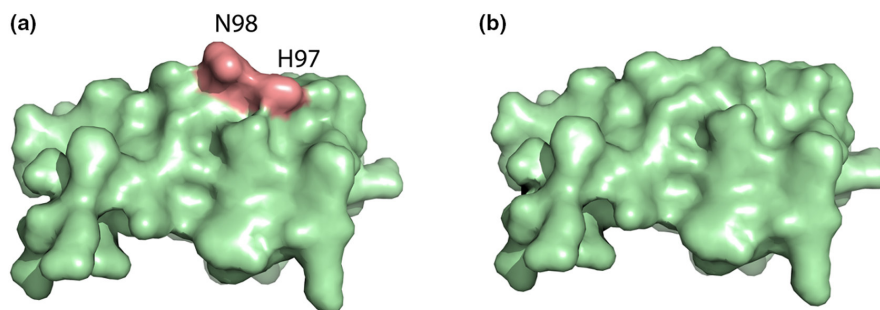
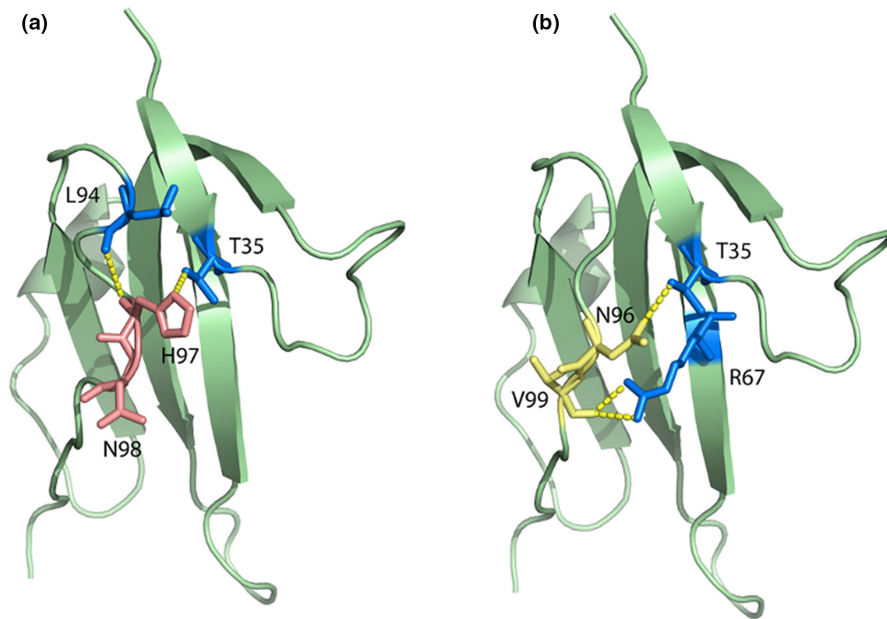
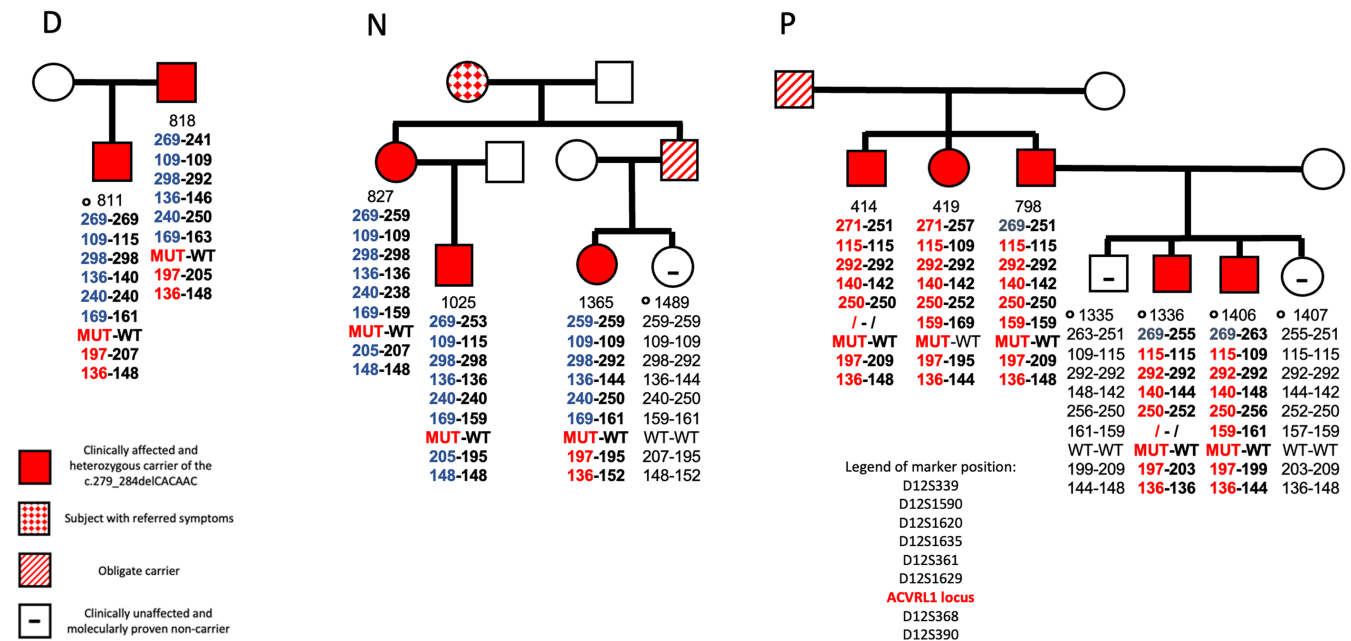


FIGURE 2 Surface representation of the structure of the extracellular domain of wild type ALK1 (a, PDB ID: 6SF1) and of the model of ALK1 H97\_N98del (b) Pink salmon: Residues deleted in ALK1 H97\_N98del. Figure generated by Pymol.



**FIGURE 3** Ribbon representation of the extracellular domain of the structure of ALK1 (a, PDB ID: 6SF1) and of the model of ALK1 H97\_N98del (b) Pink salmon sticks: Residues deleted by the “Bergamasca mutation.” Blue sticks: Residues interacting with H97 and N98 in the wild type and with residues N96 and V99 in ALK1 H97\_N98del, respectively. Figure generated by Pymol.



**FIGURE 4** Family trees representing the three different haplotypes observed. The common haplotype associated to the “Bergamasca variant” and shared by 60/66 patients is written in red (P), the “mixed” haplotype is written in blue and red (D). In family N there are two individuals with a haplotype partially overlapping the “mixed” one written all in blue (family N; 827 and 1025) and a subject carrying the “mixed haplotype” (family N; 1365). Subjects marked with  $\circ$  were excluded from the co-segregation analysis.

(families F, L, M, O, P; Figure 4; Supplementary Figure S1, red and gray alleles) expansion or recombination occurs for a single microsatellite marker.

In a single family we observed two individuals with a haplotype partially overlapping the “mixed” one, but with a further recombination (family N [827; 1025]; Figure 4; Supplementary Figure S1, blue alleles). This haplotype is clearly recombinant as in the same family is present a subject carrying a “mixed haplotype” (family N [1365]; Supplementary Figure S1) including the same alleles of

the most common haplotype in the distal part starting from *ACVRL1* locus and a recombinant haplotype in the region centromeric to *ACVRL1*.

This “mixed haplotype” is also found in other two families (families D and I; Figure 4; Supplementary Figure S1) suggesting that these three families represent a different branch in the whole genealogical tree.

The age estimation suggests that the variant arose eight generations ago, corresponding to about 200 years (165–265). A number of papers reporting founder effect have

TABLE 1 Hereditary hemorrhagic telangiectasia–related haplotypes observed in our cohort of patients

STR markers	Alleles	Alleles	Alleles	Alleles	Alleles	Alleles	Alleles
D12S339	271	269	271	271	269	259	269
D12S1590	115	115	115	115	109	109	109
D12S1620	292	292	292	292	298	298	298
D12S1635	140	140	140	140	136	136	136
D12S361	250	250	250	250	240	240	240
D12S1629	159	159	159	159	169	169	169
ACVRL1 variant	+	+	+	+	+	+	+
D12S368	197	197	195/199	207	197	197	205
D12S390	136	136	136	136	136	136	148
# patients with the haplotype	49	9	1	1	3	1	2

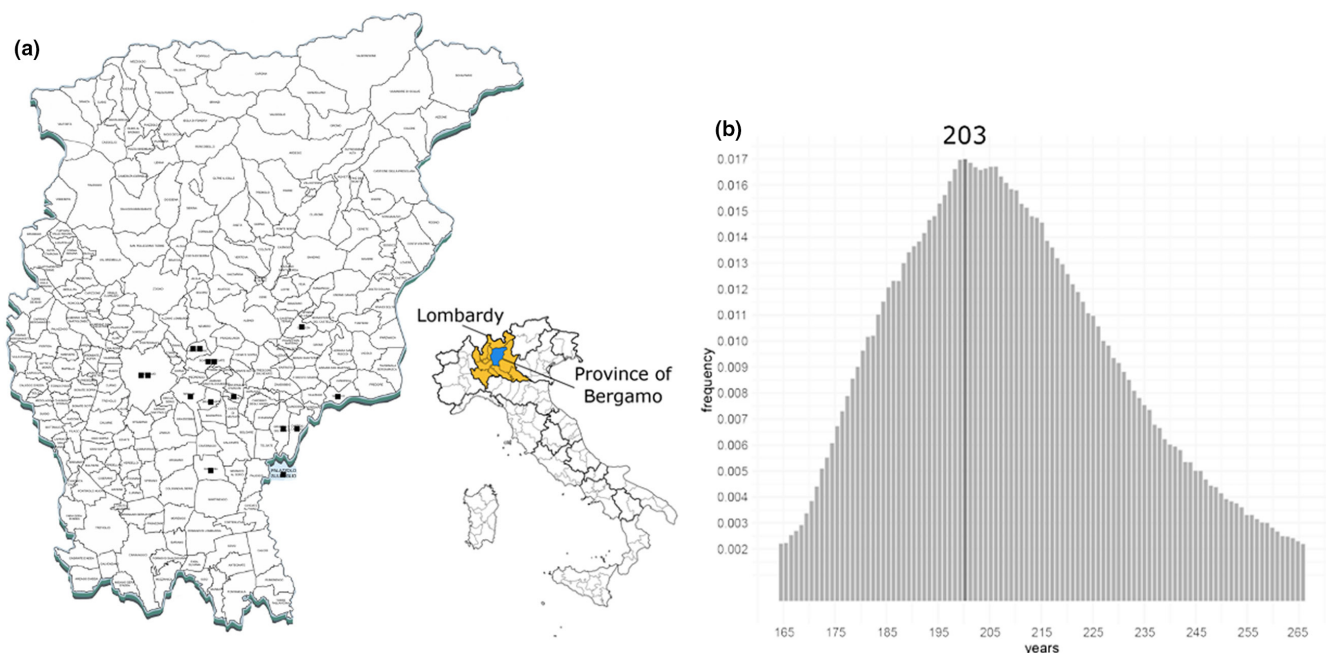


FIGURE 5 (a) Schematic representation of the province of Bergamo: Towns and villages where patients and ancestors came from are represented with black squares and localization of Bergamo Province in Italy. (b) Histograms for the averaged posterior distribution of the time since the original mutation over ten replicas; the highest bar identifies the mutation most probable age.

already been published; however, age estimation was performed only by Lesca et al. (2008) (Dakeishi et al., 2002; Gallione et al., 2000; Heimdal et al., 2016; Lesca et al., 2008).

Up to now, variants for which a founder effect has been suggested (including the “Bergamasca variant”) are 26 in *ACVRL1* and 12 in *ENG* (Reviewed by Major et al., 2021). However, in most cases haplotype analyses were not performed; this will lead to an underestimation of independent mutation events occurrence. No immediate explanation is available for the differences between the two genes. We feel that a hypothesis that could be tested might be related to the somehow more severe phenotype for HHT1, with a possible reduced fitness for patients carrying pathogenic variants in this gene.

### 4.3 | Final considerations: The variant and its relevance for families and public health

Knowledge of local founder pathogenic variants can be used to speed up and simplify the genetic diagnosis for diseases in specific populations, suggesting, in well-defined cases or families, a targeted mutation analysis instead of using panels or whole exome sequencing as the first choice.

Besides patients harboring this *in frame* deletion variant in their genome, our database collects additional 18 HHT-affected individuals (from 13 different families) coming from the province of Bergamo. We observed

seven different pathogenic variants in *ACVRL1* and six in *ENG*, confirming the hypothesis that most HHT-related variants are “private.” In this context, as already suggested by previous papers on founder effects in HHT, the presence of a founder effect leads to a higher prevalence of the disease.

“Raw data” from the regional rare disease registry (Registro Lombardo Malattie rare, 2019) report a total HHT prevalence of 1:20,000 in Lombardy (the region including the province of Bergamo). The same registry reports the prevalence of 1:8921 for the Province of Bergamo, more than doubled compared to the total, supporting our observations. No updated and reliable data are available for the total Italian population, but the clinical diagnostic work suggests that the disease is likely to be underdiagnosed.

The data we report impact the local health care picture. The prevalence quoted before (1:5000), in fact, originates from literature data from non-Caucasian population in a restricted geographical area and from a different type of data source (Dakeishi et al., 2002).

Our work demonstrates the successful integration of data from molecular genetics, population genetics and from structure analysis. We obtained not only a new piece of evidence on the natural history of a single variant, but also information relevant for the National Health Service, as they demonstrate a higher prevalence of a rare, often underdiagnosed disease as HHT, in a well-defined area as the city of Bergamo and its surroundings.

Our work provides data useful for an extensive and enhanced cooperation between Local Health Authorities and HHT patients and families thanks to a new and better awareness of the clinical and social impact of the disease in this area, with the aim to offer the population a precocious genetic and clinical diagnosis to prevent the complications of the disease and obtain a better patients' care.

## ACKNOWLEDGMENTS

Italian Ministry of Education, University and Research to the DMM University of Pavia “Dipartimenti di Eccellenza (2018-2022)” grant (CS; CO); Italian Ministry of Health (Ricerca Corrente 2018-2020) and 5 × 1000 voluntary contribution grant (TM). We gratefully acknowledge all the people who, behind the scenes, contributed to the realization of this study: Prof. Cesare Danesino, the two Italian HHT patient' association “HHT Onlus” and “Onilde Carini” and all the patients.

## AUTHOR CONTRIBUTIONS

AS: performed molecular genetic analyses; wrote and revised the manuscript. YAH: performed molecular genetic analyses. TM: performed age estimation analyses; revised the manuscript. EB: provided clinical data on patients. CS: performed structural protein analyses; revised the

manuscript. FP: provided clinical data on patients. GM: provided clinical data on patients. EM: provided clinical data on patients. GS: provided patients clinical database; wrote and revised the manuscript. CO: study design; molecular genetic analyses; wrote and revised the manuscript.

## CONFLICT OF INTEREST

All the authors have no conflict of interest to declare.

## ETHICAL STATEMENT

This study was approved by the Ethics committee in the area of Pavia (reference number 1-29/1/14) and conforms the declaration of Helsinki. All the involved subjects signed an informed consent prior to the analyses.

## DATA AVAILABILITY STATEMENT

The variant discussed in this study is openly available in the HHT mutation database (ARUP Database: [https://arup.utah.edu/database/ACVRL1/ACVRL1\\_welcome.php](https://arup.utah.edu/database/ACVRL1/ACVRL1_welcome.php)). All the other data that support the finding of this study and have not been included in the supplementary material section are available from the corresponding author upon reasonable request.

## ORCID

Anna Sbalchiero  <https://orcid.org/0000-0002-0010-0840>

Carla Olivieri  <https://orcid.org/0000-0001-5812-3175>

## REFERENCES

- Alt, A., Miguel-Romero, L., Donderis, J., Aristorena, M., Blanco, F. J., Round, A., Rubio, V., Bernabeu, C., & Marina, A. (2012). Structural and functional insights into endoglin ligand recognition and binding. *PLoS One*, 7(2), e29948.
- ARUP Mutation Databases. (2022). *ARUP scientific resource for research and education: Mutation databases*. University of Utah. <https://arup.utah.edu/database/>
- Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)*, 27(3), 343–350. <https://doi.org/10.1093/bioinformatics/btq662>
- Caleca, L., Putignano, A. L., Colombo, M., Congregati, C., Sarkar, M., Magliery, T. J., Ripamonti, C. B., Foglia, C., Peissel, B., Zaffaroni, D., Manoukian, S., Tondini, C., Barile, M., Pensotti, V., Bernard, L., Papi, L., & Radice, P. (2014). Characterization of an Italian founder mutation in the RING-finger domain of BRCA1. *PLoS One*, 9(2), e86924. <https://doi.org/10.1371/journal.pone.0086924>
- Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, 7(10), e46688. <https://doi.org/10.1371/journal.pone.0046688>
- Dakeishi, M., Shioya, T., Wada, Y., Shindo, T., Otaka, K., Manabe, M., Nozaki, J.-I., Inoue, S., & Koizumi, A. (2002). Genetic epidemiology of hereditary hemorrhagic telangiectasia in a local community in the northern part of Japan. *Human Mutation*, 19(2), 140–148. <https://doi.org/10.1002/humu.10026>



- Faughnan, M. E., Mager, J. J., Hetts, S. W., Palda, V. A., Lang-Robertson, K., Buscarini, E., Deslandres, E., Kasthuri, R. S., Lausman, A., Poetker, D., Ratjen, F., Chesnutt, M. S., Clancy, M., Whitehead, K. J., Al-Samkari, H., Chakinala, M., Conrad, M., Cortes, D., Crocione, C., ... Zarrabeitia, R. (2020). Second international guidelines for the diagnosis and management of hereditary hemorrhagic telangiectasia. *Annals of Internal Medicine*, 173(12), 989–1001. <https://doi.org/10.7326/M20-1443>
- Faughnan, M. E., Palda, V. A., Garcia-Tsao, G., Geisthoff, U. W., McDonald, J., Proctor, D. D., Spears, J., Brown, D. H., Buscarini, E., Chesnutt, M. S., Cottin, V., Ganguly, A., Gossage, J. R., Guttmacher, A. E., Hyland, R. H., Kennedy, S. J., Korzenik, J., Mager, J. J., Zanne, A. P., ... Zarrabeitia, R. (2011). International guidelines for the diagnosis and management of hereditary haemorrhagic telangiectasia. *Journal of Medical Genetics*, 48(2), 73–87. <https://doi.org/10.1136/jmg.2009.069013>
- Fernández, L. A., Sanz-Rodríguez, F., Blanco, F. J., Bernabéu, C., & Botella, L. M. (2006). Hereditary hemorrhagic telangiectasia, a vascular dysplasia affecting the TGF- $\beta$  signaling pathway. *Clinical Medicine & Research*, 4(1), 66–78.
- Gallione, C. J., Scheessele, E. A., Reinhardt, D., Duits, A. J., Berg, J. N., Westermann, C. J. J., & Marchuk, D. A. (2000). Two common endoglin mutations in families with hereditary hemorrhagic telangiectasia in The Netherlands Antilles: Evidence for a founder effect. *Human Genetics*, 107(1), 40–44.
- Govani, F. S., & Shovlin, C. L. (2009). Hereditary haemorrhagic telangiectasia: A clinical and scientific review. *European Journal of Human Genetics*, 17(7), 860–871. <https://doi.org/10.1038/ejhg.2009.35>
- Hammond, C., Braakman, I., & Helenius, A. (1994). Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proceedings of the National Academy of Sciences of the United States of America*, 91(3), 913–917. <https://doi.org/10.1073/pnas.91.3.913>
- Heimdal, K., Dalhus, B., Rødningen, O. K., Kroken, M., Eiklid, K., Dheyauldeen, S., Røysland, T., Andersen, R., & Kulseth, M. A. (2016). Mutation analysis in Norwegian families with hereditary hemorrhagic telangiectasia: Founder mutations in ACVRL1. *Clinical Genetics*, 89(2), 182–186. <https://doi.org/10.1111/cge.12612>
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., ... MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809), 434–443. <https://doi.org/10.1038/s41586-020-2308-7>
- Kjeldsen, A. D., Vase, P., & Green, A. (1999). Hereditary haemorrhagic telangiectasia: A population-based study of prevalence and mortality in Danish patients. *Journal of Internal Medicine*, 245(1), 31–39. <https://doi.org/10.1046/j.1365-2796.1999.00398.x>
- Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., & Massouras, A. (2019). VarSome: The human genomic variant search engine. *Bioinformatics*, 35(11), 1978–1980. <https://doi.org/10.1093/bioinformatics/bty897>
- Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M., & Maglott, D. R. (2014). ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research*, 42(Database issue), D980–D985. <https://doi.org/10.1093/nar/gkt1113>
- Lesca, G., Genin, E., Blachier, C., Olivieri, C., Coulet, F., Brunet, G., Dupuis-Girod, S., Buscarini, E., Soubrier, F., Calender, A., Danesino, C., Giraud, S., & Plauchu, H. (2008). Hereditary hemorrhagic telangiectasia: Evidence for regional founder effects of ACVRL1 mutations in French and Italian patients. *European Journal of Human Genetics*, 9, 742–749.
- Lesca, G., Olivieri, C., Burnichon, N., Pagella, F., Carette, M., 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. *Genetics in Medicine*, 9(1), 14–22. <https://doi.org/10.1097/GIM.0b013e31802d8373>
- Major, T., Gindele, R., Balogh, G., Bárdossy, P., & Berezsky, Z. (2021). Founder effects in hereditary hemorrhagic telangiectasia. *Journal of Clinical Medicine*, 10(8), 1682. <https://doi.org/10.3390/jcm10081682>
- Mitra, N., Sinha, S., Ramya, T. N. C., & Suroliya, A. (2006). N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends in Biochemical Sciences*, 31(3), 156–163. <https://doi.org/10.1016/j.tibs.2006.01.003>
- Olivieri, C., Mira, E., Delù, G., Pagella, F., Zambelli, A., Malvezzi, L., Buscarini, E., & Danesino, C. (2002). Identification of 13 new mutations in the ACVRL1 gene in a group of 52 unselected Italian patients affected by hereditary haemorrhagic telangiectasia. *Journal of Medical Genetics*, 39(7), E39. <https://doi.org/10.1136/jmg.39.7.e39>
- 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. *Journal of Human Genetics*, 52(10), 820–829. <https://doi.org/10.1007/s10038-007-0187-5>
- Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K. A., Lin, G. N., Nam, H.-J., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., & Radivojac, P. (2020). Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nature Communications*, 11(1), 5918. <https://doi.org/10.1038/s41467-020-19669-x>
- Reeve, J. P., & Rannala, B. (2002). DMLE+: Bayesian linkage disequilibrium gene mapping. *Bioinformatics (Oxford, England)*, 18(6), 894–895. <https://doi.org/10.1093/bioinformatics/18.6.894>
- Regional Registry of Rare Diseases (Lombardy) - Registro Lombardo Malattie Rare Report of December 31st 2019, edited by the "Centro di Coordinamento della Rete Regionale per le Malattie Rare" <http://malattierare.marionegri.it/content/view/92/95/>
- Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. *Nature Methods*, 11(4), 361–362. <https://doi.org/10.1038/nmeth.2890>
- 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). *American Journal of Medical Genetics*, 91(1), 66–67. [https://doi.org/10.1002/\(sici\)1096-8628\(20000306\)91:1<66::aid-ajmg12>3.0.co;2-p](https://doi.org/10.1002/(sici)1096-8628(20000306)91:1<66::aid-ajmg12>3.0.co;2-p)

- Stentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vestergaard, M. B., Schjoldager, K. T.-B. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., & Clausen, H. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO Journal*, 32(10), 1478–1488. <https://doi.org/10.1038/emboj.2013.79>
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>
- 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. *The American*

*Journal of Human Genetics*, 93(3), 530–537. <https://doi.org/10.1016/j.ajhg.2013.07.004>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Sbalchiero, A., Abu Hweij, Y., Mazza, T., Buscarini, E., Scotti, C., Pagella, F., Manfredi, G., Matti, E., Spinozzi, G., & Olivieri, C. (2022). Hereditary hemorrhagic telangiectasia: First demonstration of a founder effect in Italy; the *ACVRL1* c.289\_294del variant originated in the country of Bergamo 200 years ago. *Molecular Genetics & Genomic Medicine*, 10, e1972. <https://doi.org/10.1002/mgg3.1972>