UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie "L. Spallanzani"

Laurea Magistrale in Biologia Sperimentale e Applicata

Dipartimento di Medicina Molecolare
Unità di Biologia Generale e Genetica Medica

Teleangectasia Emorragica Ereditaria: analisi di varianti e studi funzionali nell'ambito di un progetto Whole Genome Sequencing (WGS)

Relatore:
Carla Olivieri

Tesi Sperimentale di
Lucia Spadaro
Matr. 565774

Anno Accademico 2025/2026

INDICE

HHT: CARATTERISTICHE GENERALI 3
SINTOMI 3
DIAGNOSI CLINICA E CRITERI DIAGNOSTICI 4
EPISTASSI 6
TELEANGECTASIE 7
LESIONI VASCOLARI VISCERALI 7
FAMILIARITÀ PER HHT 8
BASI GENETICHE E DIAGNOSI MOLECOLARE 9
PATHWAY DEL TGF-β E IL SUO COINVOLGIMENTO NELL'HHT 11
GENI COINVOLTI 15
ENG 15
CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT 15
ENDOGLINA E HHT1 16
VARIANTI DEL GENE *ENG* 17
ACVRL1 18
CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT 18
ACVRL1 E HHT2 19
VARIANTI DEL GENE *ACVRL1* 20
SMAD4 21
CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT 21
SMAD4 E JP-HHT 22
VARIANTI DEL GENE *SMAD4* 23
GDF2 24
CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT 24
GDF2 E HHT 25
RASA1 26
RASA1 E HHT 26
EPHB4 27
EPHB4 E HHT 28
BMPR2 28
BMPR2 E HHT 29

MODELLO “SECOND HIT” NELL’HHT 30
SEQUENZIAMENTO SANGER 32
NEXT GENERATION SEQUENCING (NGS) 33
WHOLE EXOME SEQUENCING (WES) 36
WHOLE GENOME SEQUENCING (WGS) 37
CRITERI ACMG/AMP PER LA CLASSIFICAZIONE DELLE VARIANTI NELL’HHT 38
MATERIALI E METODI 42
CAMPIONI BIOLOGICI 42
ESTRAZIONE DEL DNA 42
ESTRAZIONE DELL’RNA 44
DOSAGGIO SPETTROFOTOMETRICO DEGLI ACIDI NUCLEICI 45
RETROTRASCRIZIONE DELL’RNA IN cDNA 45
PCR PER IL DNA ESTRATTO O PER IL cDNA RETROTRASCRITTO 48
ELETTROFORESI SU GEL DI AGAROSIO 51
PURIFICA DELLE PCR E PREPARAZIONE DEI CAMPIONI PER IL SEQUENZIAMENTO SANGER 51
INVIO CAMPIONI PER SEQUENZIAMENTO 52
ANALISI E INTERPRETAZIONE DELLE VARIANTI 52
REAL-TIME PCR 53
RISULTATI 56
CAMPIONI BIOLOGICI, ESTRAZIONE E DOSAGGIO DEGLI ACIDI NUCLEICI 56
ANALISI MEDIANTE PCR E CORSA ELETTROFORETICA 56
SEQUENZIAMENTO SANGER E ANALISI DELLE VARIANTI 58
ANALISI MEDIANTE REAL-TIME PCR 61
DISCUSSIONE 64
CONCLUSIONI 71
BIBLIOGRAFIA 74

INTRODUZIONE

HHT: CARATTERISTICHE GENERALI

La Teleangectasia Emorragica Ereditaria, o HHT dall'acronimo inglese (Hereditary Hemorrhagic Teleangiectasia), è una rara patologia vascolare a trasmissione autosomica dominante, caratterizzata dalla presenza di anomalie nella formazione e nell'organizzazione dei vasi sanguigni.
Il nome storico della malattia, sindrome di Rendu-Osler-Weber (ROW) deriva dai medici H. Rendu, W. Osler e F. P. Weber che tra la fine del XIX e l'inizio del XX secolo contribuirono alla sua caratterizzazione clinica.

Le "Second International Guidelines for HHT" riportano una prevalenza stimata di circa 1 individuo ogni 5.000 persone; tuttavia, la prevalenza reale potrebbe essere superiore, poiché la patologia è considerata sottodiagnosticata. Tale sottostima è attribuibile all'elevata variabilità clinica e alla penetranza età-dipendente (*Faughnan et al., 2020*).

L'HHT presenta in alcune regioni una distribuzione geografica peculiare, che suggerisce un probabile effetto del fondatore, cioè la presenza di specifiche varianti patologiche particolarmente frequenti in determinate popolazioni. Esempi significativi di questo fenomeno sono stati descritti in alcune popolazioni delle Antille Olandesi, Francesi e Italiane (nella provincia di Bergamo) (*Gallione et al., 2000; Lesca et al., 2008; Sbalchiero et al., 2022*).

SINTOMI

L'HHT è caratterizzata principalmente dalla formazione di teleangectasie e malformazioni artero-venose (MAV). In condizioni normali, arterie e vene sono collegate attraverso una rete di capillari, mentre nell'HHT possono formarsi connessioni dirette tra il circolo arterioso e quello venoso. Queste strutture risultano estremamente fragili e, di conseguenza, sono maggiormente predisposte al sanguinamento (*Begbie et al., 2003*).

Le anomalie che interessano i vasi di piccole dimensioni nei tessuti mucocutanei vengono definite teleangectasie e si presentano come piccole dilatazioni vascolari superficiali, generalmente di colore rosso-violaceo. Queste teleangectasie possono comparire soprattutto a livello delle labbra, della lingua, della mucosa orale e nasale, del viso e dei polpastrelli (*Letteboer et al., 2008*). Quando le alterazioni coinvolgono vasi di dimensioni maggiori negli organi profondi, si parla invece di malformazioni artero-venose, che possono interessare diversi organi interni, in particolare fegato, polmoni e cervello, ma anche il tratto gastrointestinale, e più raramente i vasi intorno al midollo spinale (*Faughnan et al., 2020*).

Da un punto di vista clinico, l'HHT presenta una notevole variabilità di manifestazione: i sintomi, la loro gravità e gli organi interessati possono essere molto diversi da un individuo all'altro, anche all'interno della stessa famiglia o tra individui con la stessa mutazione. Per questo motivo, non tutti i pazienti presentano contemporaneamente l'intero quadro delle alterazioni associate alla malattia (*Faughnan et al., 2020*).

La manifestazione più comune dell'HHT è l'epistassi spontanea e ricorrente, cioè il sanguinamento spontaneo dal naso, dovuto alla presenza e alla rottura delle teleangectasie della mucosa nasale. Gli episodi possono variare notevolmente per frequenza e intensità e, in casi estremi, possono essere la causa dello sviluppo di anemia da carenza di ferro. Anche le teleangectasie gastrointestinali possono provocare sanguinamenti, che tendono a manifestarsi soprattutto con l'avanzare dell'età e possono determinare una perdita cronica di sangue, aggravando l'anemia. Le conseguenze cliniche delle MAV viscerali dipendono invece dall'organo coinvolto. Le complicanze possono essere invalidanti e, in alcuni casi, pericolose per la vita (*Faughnan et al., 2020*).

DIAGNOSI CLINICA E CRITERI DIAGNOSTICI

La diagnosi della Teleangectasia Emorragica Ereditaria si basa principalmente sulla valutazione clinica del paziente attraverso i criteri di Curaçao, standardizzati a livello internazionale nel 2000 e, successivamente, confermati nel 2011: con l'obiettivo di uniformare la diagnosi della patologia (*Shovlin et al., 2000; Faughnan et al., 2011*).
Questi criteri rappresentano ancora oggi il principale riferimento per la diagnosi clinica dell'HHT e prendono in considerazione quattro manifestazioni caratteristiche della malattia:

- Epistassi spontanee e ricorrenti
- Teleangectasie mucocutanee multiple e in siti caratteristici
- Presenza di lesioni vascolari viscerali come teleangectasie gastrointestinali, MAV polmonari, cerebrali, epatiche o spinali
- Familiarità per HHT con almeno un parente di primo grado affetto secondo questi criteri.

In base al numero di criteri di Curaçao soddisfatti, la diagnosi clinica può essere classificata come:

- Definitiva, se sono presenti 3 o 4 criteri
- Possibile o sospetta, se sono presenti 2 criteri
- Improbabile, se sono presenti 1 o 0 criteri

Questa classificazione permette di attribuire un diverso grado di probabilità diagnostica sulla base delle manifestazioni cliniche osservate nel paziente. Tuttavia, non è soltanto il numero complessivo dei criteri presenti ad assumere rilevanza, ma anche la specifica combinazione delle manifestazioni cliniche può influenzare la probabilità che il quadro osservato sia effettivamente riconducibile all'HHT (*Shovlin et al., 2000*).

La relazione tra la classificazione clinica e la diagnosi genetico-molecolare può essere valutata considerando la Mutation Detection Rate, ovvero la percentuale di soggetti sottoposti ad analisi genetica nei quali viene identificata una variante responsabile della malattia. Questo parametro consente quindi di valutare con quale frequenza, in soggetti caratterizzati da determinate manifestazioni cliniche, l'analisi molecolare permetta di identificare una variante associata all'HHT.

La Mutation Detection Rate varia sia in relazione al numero di criteri soddisfatti sia alla loro specifica combinazione. La percentuale più elevata di identificazione di una variante si osserva nei soggetti che presentano tutti e quattro i criteri, superando l'87%. La presenza contemporanea di tutti i criteri risulta quindi associata a un'elevata probabilità di identificare una variante responsabile della malattia (*Richards-Yutz et al., 2010*).

Nei soggetti che soddisfano tre criteri, la percentuale di identificazione delle varianti rimane relativamente elevata, sebbene presenti delle differenze in relazione alla specifica combinazione delle manifestazioni cliniche.

Una maggiore variabilità si trova nei soggetti che soddisfano due criteri di Curaçao, corrispondenti a una diagnosi probabile o sospetta. In questo gruppo, alcune combinazioni di manifestazioni cliniche sono associate a percentuali di identificazione delle varianti più elevate di altre. Questo evidenzia come, oltre al numero complessivo di criteri soddisfatti, anche la specifica combinazione dei criteri presenti possa avere rilevanza.

Anche nei soggetti che presentano un solo criterio, e che secondo la classificazione di Curaçao rientrano nella categoria di diagnosi improbabile, possono essere identificate varianti associate all'HHT. In questi casi, tuttavia, la percentuale di identificazione risulta molto variabile a seconda della manifestazione clinica considerata (*Richards-Yutz et al., 2010*).

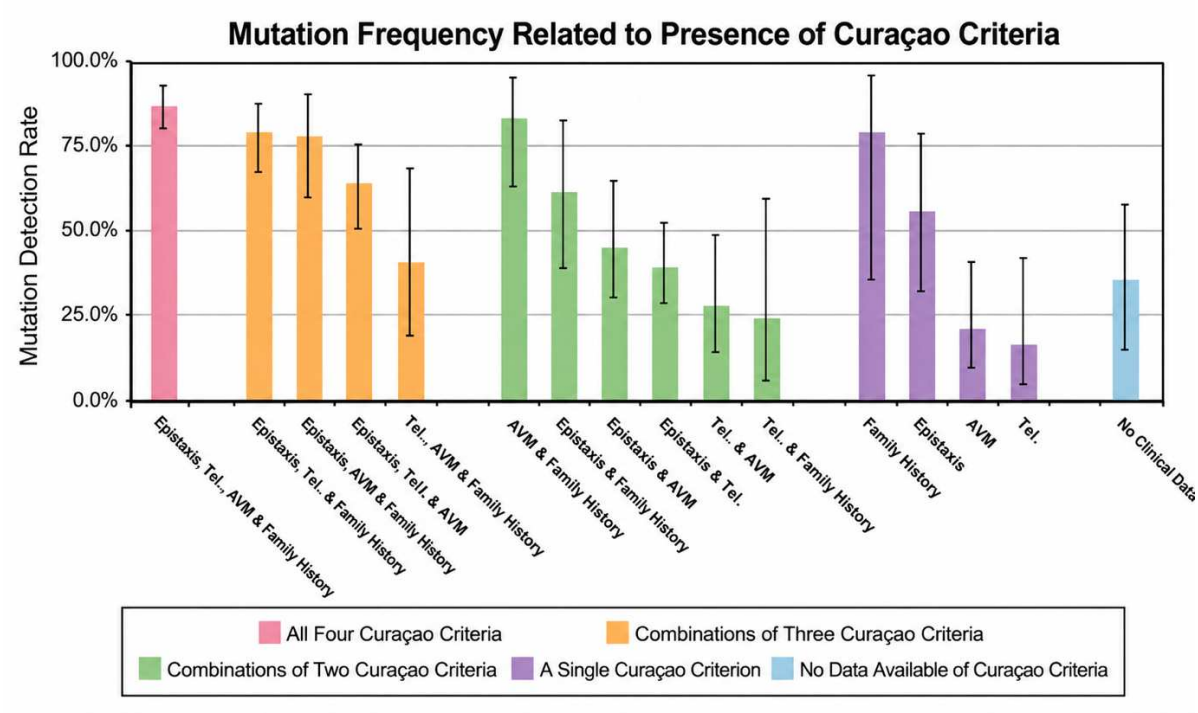


Figura 1. Relazione tra criteri diagnostici di Curaçao e Detection Rate delle varianti associate all'HHT (*Richards-Yutz et al., 2010*).

EPISTASSI

Il primo criterio, come visto, è rappresentato dall'epistassi spontanea e ricorrente, che costituisce una delle manifestazioni cliniche più frequenti dell'HHT. Il sanguinamento è principalmente riconducibile alla presenza di teleangectasie a livello della mucosa nasale, caratterizzate da un'elevata fragilità vascolare. Nei casi più severi, gli episodi emorragici possono determinare perdite significative che rendono necessarie trasfusioni di sangue. Nella formulazione originale dei criteri di Curaçao non viene stabilito un numero minimo di episodi né una particolare gravità del sanguinamento, ma viene richiesto che l'epistassi sia spontanea e ricorrente (*Shovlin et al., 2020*).

Poiché la valutazione dell'epistassi può essere influenzata dalla percezione soggettiva del paziente, nella pratica clinica sono stati sviluppati strumenti standardizzati che consentono di valutarne quantitativamente la gravità. Tra questi rientra l'Epistaxis Severity Score (ESS), che prende in considerazione diversi parametri relativi al sanguinamento, tra cui la frequenza e la durata degli episodi, l'intensità del flusso ematico, la presenza di anemia e l'eventuale necessità di interventi medici o trasfusioni. Anche il gruppo italiano ha proposto una classificazione delle epistassi secondo un sistema definito FID (Frequenza-Intensità-Durata) (*Pagella et al., 2020*). Questi strumenti permettono, quindi, di ottenere una valutazione più oggettiva della severità dell'epistassi e del suo impatto clinico (*Hayama et al., 2022*).

TELEANGECTASIE

Il secondo criterio riguarda la presenza di teleangectasie multiple localizzate in sedi caratteristiche, oltre a quelle della mucosa nasale viste in precedenza, in particolare a livello delle labbra, della cavità orale, dei polpastrelli delle dita delle mani.

La loro localizzazione assume particolare rilevanza diagnostica, poiché la presenza di teleangectasie in questi distretti costituisce una delle manifestazioni maggiormente caratteristiche dell'HHT. Studi più recenti hanno inoltre proposto di considerare indicativa la presenza di tre o più teleangectasie nelle sedi caratteristiche, sebbene numero e distribuzione possano variare in relazione al fenotipo del paziente (*McDonald et al., 2023*).

LESIONI VASCOLARI VISCERALI

Il terzo criterio è rappresentato dal coinvolgimento viscerale e comprende sia le teleangectasie gastrointestinali sia le malformazioni artero-venose (MAV) localizzate a livello epatico, polmonare, cerebrale o spinale. L'identificazione di queste alterazioni riveste particolare importanza non soltanto a fini diagnostici, ma anche per la gestione clinica del paziente, poiché le malformazioni viscerali possono rimanere a lungo asintomatiche, ma determinare successivamente complicanze severe (*Ahlhelm et al., 2013*).

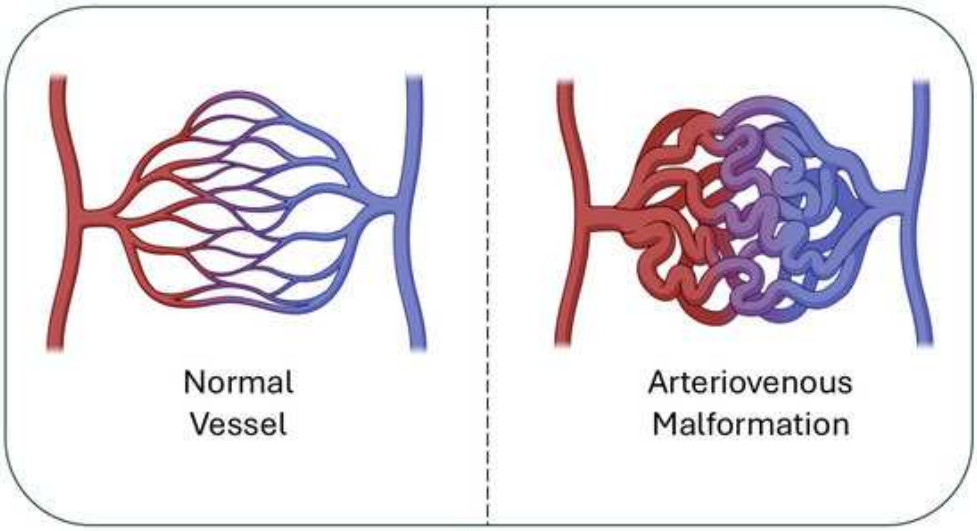


Figura 2. Vascolarizzazione normale e vascolarizzazione con MAV (*Le et al., 2025*).

Il coinvolgimento del tratto gastrointestinale interessa prevalentemente i pazienti adulti e tende ad acquisire maggiore rilevanza clinica con l'avanzare dell'età. Le teleangectasie sono frequentemente localizzate a livello dello stomaco e dell'intestino tenue e possono essere responsabili di

sanguinamenti cronici. La perdita ematica protratta nel tempo può determinare una progressiva carenza di ferro e, talvolta anche anemia, dovendo ricorrere nei casi più severi alla somministrazione di ferro o trasfusioni. Un sanguinamento gastrointestinale deve essere sospettato soprattutto quando la carenza di ferro o l'anemia risultano sproporzionate rispetto all'entità dell'epistassi (*Locke et al., 2026*).

L'organo maggiormente interessato da malformazioni vascolari è il fegato, in cui la frequenza e la rilevanza clinica delle malformazioni tende ad aumentare con l'età. Le alterazioni del flusso ematico epatico possono determinare diverse complicanze, tra cui insufficienza cardiaca ad alta gittata, ipertensione portale, encefalopatia e ischemia biliare o mesenterica, al punto da richiedere un trapianto. Le modificazioni vascolari possono inoltre associarsi a fenomeni proliferativi del parenchima epatico, come l'iperplasia nodulare rigenerativa e l'iperplasia nodulare focale, che in alcuni casi possono complicare l'interpretazione del quadro clinico (*Buscarini et al., 2004*).

Un'altra manifestazione caratteristica è rappresentata dalle malformazioni artero-venose polmonari (PAVM), nelle quali si stabilisce una comunicazione anomala tra arterie e vene polmonari, che determina uno shunt destro-sinistro. Il quadro clinico è estremamente variabile: alcuni pazienti rimangono asintomatici, mentre altri possono presentare affaticamento, dispnea, ipossiemia ed emicrania. Una delle complicanze più rilevanti è rappresentata dall'embolia paradossa, associata a un aumento del rischio di ictus, attacco ischemico transitorio e ascesso cerebrale. Di conseguenza, questo tipo di MAV richiede un trattamento di chiusura endoscopico (*Fish et al., 2021*).

Anche le malformazioni vascolari cerebrali associate all'HHT comprendono diverse tipologie di alterazioni, tra cui malformazioni artero-venose e fistole artero-venose. Le MAV cerebrali associate alla malattia possono essere multiple e spesso di piccole dimensioni. Nonostante molte di queste lesioni possano rimanere asintomatiche, la loro presenza espone il paziente al rischio di complicanze neurologiche ed emorragiche, rendendo molto importante la loro identificazione attraverso un adeguato programma di screening (*Easey et al., 2003; Palermo e Sturiale, 2026*).

FAMILIARITÀ PER HHT

Il quarto criterio prende in considerazione la storia familiare, ed è soddisfatto quando è presente almeno un familiare di primo grado affetto da HHT secondo gli stessi criteri diagnostici. La familiarità rappresenta un elemento particolarmente rilevante in considerazione della trasmissione autosomica dominante della patologia.

Come visto, in base al numero di criteri soddisfatti, la diagnosi clinica viene classificata in definita, probabile o sospetta e improbabile. L'elevato valore predittivo di questa classificazione clinica è stato supportato anche dal confronto con i risultati delle analisi genetico-molecolari.

In una coorte di 152 soggetti con diagnosi clinica definita di HHT, una variante causativa o sospettata come tale è stata identificata in *ENG* nel 44,1% dei soggetti e in *ACVRL1* nel 51,0%. Considerando complessivamente i due principali geni responsabili di HHT1 e HHT2, la Variant Detection Rate raggiungeva il 96,1%. L'analisi di *SMAD4* ha consentito di identificare ulteriori varianti nell'1,3% dei soggetti, mentre soltanto il 2,6% non presentava una variante identificabile (*McDonald et al., 2020*).

L'identificazione della variante patogenetica responsabile della malattia in un individuo affetto consente di effettuare un'analisi mirata nei familiari a rischio risultando particolarmente utile nei bambini, nei giovani adulti e nei pazienti asintomatici. L'integrazione tra criteri clinici di Curaçao e l'analisi genetico-molecolare permette, pertanto, di ottenere una diagnosi più accurata e, soprattutto, di identificare precocemente i soggetti che necessitano di specifici programmi di screening per le manifestazioni vascolari associate all'HHT (*McDonald et al., 2020; Faughnan et al., 2020*).

BASI GENETICHE E DIAGNOSI MOLECOLARE

La diagnosi molecolare dell'HHT è finalizzata principalmente all'identificazione di varianti patogenetiche nei geni coinvolti nella regolazione della segnalazione endoteliale, mediata dalla famiglia del TGF-β. I geni maggiormente coinvolti sono *ENG* e *ACVRL1*, responsabili rispettivamente delle forme HHT1 e HHT2 e codificanti per le proteine Endoglina e ALK1. Entrambe partecipano alla regolazione dell'omeostasi vascolare e dell'angiogenesi e la presenza di varianti Loss-of-function può determinare una riduzione della quantità di proteina funzionale disponibile, secondo un meccanismo prevalentemente riconducibile all'aploinsufficienza (*McDonald et al., 2020*).

Sebbene meno frequentemente coinvolto, *SMAD4* riveste una particolare importanza dal punto di vista clinico, poiché le sue varianti patogenetiche sono associate alla sindrome combinata Poliposi giovanile-HHT o JP-HHT. In questi pazienti, alle manifestazioni vascolari caratteristiche dell'HHT si associa la predisposizione allo sviluppo di polipi gastrointestinali e di neoplasie del tratto gastrointestinale (*Gallione et al., 2004*).

Oltre ai geni responsabili delle principali forme di HHT, sono stati descritti altri geni associati a fenotipi vascolari parzialmente sovrapponibili. Tra questi è compreso *GDF2*, che codifica per BMP9, un ligando della via di segnalazione che coinvolge ALK1 ed Endoglina. Varianti patogenetiche di *GDF2* sono state descritte in soggetti con manifestazioni cliniche HHT-like sebbene rappresentino una causa molto più rara rispetto alle alterazioni di *ENG* e *ACVRL1* (*Wooderchak-Donahue et al., 2013*).

La variabilità fenotipica dell'HHT rende inoltre necessaria la diagnosi differenziale con altre malattie vascolari ereditarie. Particolare rilevanza assume la Sindrome da malformazioni capillari-malformazioni arterovenose (CM-AVM), associata principalmente a varianti nei geni *RASA1* ed *EPHB4*. Questa condizione può presentare malformazioni capillari cutanee e malformazioni arterovenose che, in alcuni pazienti, possono determinare un quadro clinico sovrapponibile a quello dell'HHT. In particolare, varianti patogenetiche di *EPHB4* sono state identificate in pazienti inizialmente valutati per sospetta HHT, evidenziando l'importanza di considerare tali geni nell'iter diagnostico differenziale. Nella diagnosi differenziale può essere considerato anche *BMPR2*, le cui varianti patogenetiche sono principalmente associate all'ipertensione arteriosa polmonare (PAH), condizione che può presentare una sovrapposizione con alcune manifestazioni vascolari associate all'HHT (*Wooderchak-Donahue et al., 2019*).

L'eterogeneità genetica e la sovrapposizione tra differenti sindromi vascolari hanno favorito l'introduzione di pannelli multigenici mediante Next Generation Sequencing (NGS). Questi consentono di analizzare simultaneamente i principali geni associati all'HHT e alle condizioni correlate, tra cui *ENG*, *ACVRL1*, *SMAD4*, *GDF2*, *RASA1*, *EPHB4* e *BMPR2.* Tale approccio risulta particolarmente utile nei pazienti con manifestazioni cliniche atipiche o nei quali i soli criteri non consentono di raggiungere una diagnosi conclusiva.

L'analisi genetica assume particolare importanza anche nei familiari a rischio e nei soggetti pediatrici. Le manifestazioni cliniche dell'HHT presentano infatti una penetranza dipendente dall'età e possono non essere ancora evidenti durante l'infanzia. Di conseguenza, i criteri di Curaçao mostrano una sensibilità inferiore nei pazienti più giovani.

Quando nel paziente è stata identificata la variante responsabile della malattia, la ricerca mirata della stessa variante nei familiari consente di riconoscere precocemente i soggetti portatori, anche in

assenza di manifestazioni cliniche, e di indirizzarli verso gli appropiati programmi di screening delle malformazioni vascolari (*Aguilera et al., 2023*).

PATHWAY DEL TGF-β E IL SUO COINVOLGIMENTO NELL'HHT

La superfamiglia del Transforming Growth Factor-β o TGF-β comprende un ampio gruppo di fattori di crescita che regolano numerosi processi biologici fondamentali, tra cui proliferazione, differenziamento, migrazione e apoptosi cellulare, oltre a svolgere un ruolo importante nello sviluppo embrionale e nel mantenimento dell'omeostasi tissutale. La risposta biologica indotta da questi fattori è strettamente dipendente dal contesto cellulare e deriva dall'interazione tra ligandi, recettori di membrana e mediatori intracellulari, principalmente appartenenti alla famiglia delle proteine SMAD (*Hata e Chen, 2016*).

La superfamiglia comprende diversi gruppi di ligandi, tra cui i TGF-β, le activine e le Bone Morphogenetic Proteins (BMP). Nonostante appartengano alla stessa superfamiglia e condividano alcuni elementi della cascata di segnalazione, questi ligandi possono attivare differenti combinazioni recettoriali e differenti proteine SMAD, determinando risposte cellulari specifiche.

La trasduzione del segnale ha inizio a livello della membrana plasmatica attraverso l'interazione del ligando con complessi recettoriali costituiti da recettori di tipo I e di tipo II, entrambi dotati di attività serina/treonina-chinasica. In seguito al legame del ligando, il recettore di tipo II determina la fosforilazione e l'attivazione del recettore di tipo I. Quest'ultimo trasferisce quindi il segnale all'interno della cellula mediante la fosforilazione di specifiche proteine SMAD regolatorie (R-SMAD) (*Vander Ark et al., 2018*).

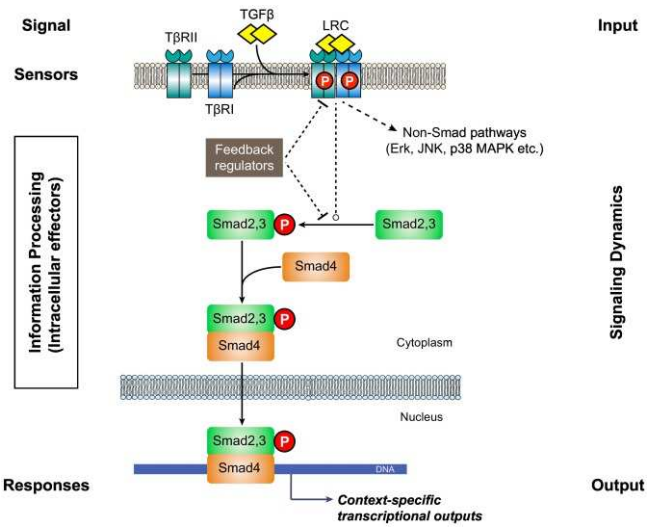


Figura 3. Rappresentazione schematica della rete di segnalazione del TGF-β (*Zi, 2019*).

Nel pathway canonico del TGF-β, il ligando si lega al recettore di tipo II, TGFBR2, che recluta e fosforila il recettore di tipo I, TGFBR1, noto anche come ALK5. L'attivazione di ALK5 determina prevalentemente la fosforilazione di SMAD2 e SMAD3. Una volta fosforilate, queste proteine formano un complesso con SMAD4, definito Co-SMAD o SMAD mediatore comune. Il complesso SMAD così formato trasloca nel nucleo, dove interagisce con diversi fattori di trascrizione, co-attivatori e co-repressori, modulando così l'espressione dei geni bersaglio e determinando la risposta cellulare al segnale.

La famiglia SMAD comprende inoltre proteine con funzione inibitoria, principalmente SMAD6 e SMAD7, che contribuiscono alla regolazione negativa della cascata di segnalazione. La presenza di meccanismi di attivazione e di inibizione permette quindi di controllare finemente intensità e durata della risposta cellulare (*Hata e Chen, 2016; Zi, 2019*).

La complessità di questo sistema di segnalazione spiega perché una qualunque alterazione possa causare numerose condizioni patologiche. Di conseguenza, diversi componenti del pathway sono studiati come possibili bersagli farmacologici, attraverso strategie dirette contro i ligandi, i recettori o i mediatori intracellulari della cascata. Tuttavia, la molteplicità delle funzioni fisiologiche svolte dalla superfamiglia TGF-β rende necessario un controllo estremamente selettivo della sua modulazione terapeutica (*Guo et al., 2024*).

All'interno della superfamiglia TGF-β, le BMP utilizzano un sistema di trasduzione del segnale analogo, ma coinvolgono prevalentemente R-SMAD differenti. Mentre il ramo TGF-β/attivina determina principalmente l'attivazione di SMAD2 e SMAD3, il ramo BMP determina la fosforilazione di SMAD1, SMAD5 e SMAD8/9. Le R-SMAD attivate di associano anche in questo caso a SMAD4 e il complesso risultante trasloca nel nucleo, dove regola la trascrizione dei geni bersaglio (*Guo et al., 2024*).

Nelle cellule endoteliali assume particolare importanza uno specifico ramo di questa via, mediato dai ligandi circolanti BMP9 e BMP10. Questi presentano un'elevata affinità per un complesso recettoriale che comprende il recettore di tipo I ALK1, codificato dal gene *ACVRL1*, recettori di tipo II, principalmente BMPRII e ActRIIA, e il corecettore Endoglina codificato dal gene *ENG*. Endoglina è una glicoproteina transmembrana espressa abbondantemente nelle cellule endoteliali, e pur non

possedendo una propria attività chinasica, partecipa al complesso recettoriale, potenziando la risposta a BMP9 e BMP10.

Come nello schema generale visto sopra, il legame di BMP9 o BMP10 determina l'assemblaggio e l'attivazione del complesso recettoriale. Il recettore di tipo II fosforila e attiva ALK1, che a sua volta determina la fosforilazione principalmente di SMAD1 e SMAD5. Le SMAD fosforilate si associano quindi a SMAD4, formando un complesso che trasloca nel nucleo e, insieme ad altri fattori di trascrizione, modula l'espressione di numerosi geni coinvolti nella regolazione dell'endotelio (*Tillet e Bailly, 2015*).

In condizioni fisiologiche, l'asse BMP9/10-ENG-ALK1-SMAD1/5-SMAD4 svolge un ruolo fondamentale nel mantenimento dell'omeostasi e della quiescenza vascolare. La sua attività contribuisce a controllare la proliferazione e la migrazione delle cellule endoteliali e a mantenere un corretto equilibrio tra formazione, rimodellamento e stabilizzazione dei vasi sanguigni (*Al Tabosh et al., 2024*).

Questo pathway assume particolare importanza nella Teleangectasia Emorragica Ereditaria (HHT), poiché i principali geni responsabili della malattia codificano per proteine appartenenti alla cascata di segnalazione. Infatti, *ENG* codifica per Endoglina, che agisce come co-recettore, *ACVRL1* per il recettore ALK1, *SMAD4* per un effettore intracellulare, e *GDF2* per BMP9, cioè il ligando. Varianti di questi geni sono associate alle diverse forme di HHT o a fenotipi vascolari HHT-simili (*Gallione et al., 2004; Wooderchak-Donahue et al., 2013*) .

Questa convergenza genetica evidenzia come l'HHT sia correlata a un'alterazione dell'asse BMP9/10-ENG-ALK1-SMAD. Per questo, sebbene l'HHT sia stata storicamente descritta come una patologia legata al pathway del TGF-β, le conoscenze molecolari più recenti permettono di definirla più precisamente come una malattia che coinvolge uno specifico ramo BMP della superfamiglia TGF-β, particolarmente importante nelle cellule endoteliali (*Tillet e Bailly, 2015*).

Nella maggior parte dei casi, le varianti patogenetiche associate all'HHT determinano una perdita di funzione (LOF) della proteina. Per *ENG* e *ACVRL1*, il principale meccanismo patogenetico è riconducibile all'aploinsufficienza, per cui la presenza di una sola copia funzionale del gene non è sufficiente a garantire livelli adeguati della proteina e una normale attività della via di segnalazione (*Abdalla e Letarte, 2006*).

La riduzione della funzionalità di Endoglina o ALK1 compromette quindi la corretta trasmissione del segnale indotto da BMP9/10, con conseguente riduzione dell'attivazione di SMAD1/5 e alterazione della regolazione trascrizionale mediata da SMAD4. Ne deriva una compromissione dei meccanismi che normalmente contribuiscono al mantenimento della quiescenza e dell'omeostasi endoteliale (*Gousmans et al., 2009*).

La perdita di questa regolazione rende l'endotelio maggiormente suscettibile agli stimoli pro-angiogenici. In particolare, la riduzione della segnalazione BMP9/BMP10-ENG-ALK1 è associata a una minore capacità di contrastare pathway pro-angiogenici, tra cui quello mediato dal VEGF. L'alterazione dell'equilibrio tra segnali pro- e anti- angiogenici favorisce quindi una proliferazione e una migrazione anomala delle cellule endoteliali e un rimodellamento vascolare alterato (*Al Tabosh et al., 2024*).

Il ruolo centrale di questo pathway nella patogenesi delle malformazioni vascolari è supportato anche da evidenze sperimentali. La perdita endoteliale di ENG, ALK1 o SMAD4, così come l'alterazione della segnalazione mediata da SMAD1/5, è associata nei modelli sperimentali alla formazione di MAV. Tali osservazioni indicano che il corretto funzionamento dell'asse è essenziale per mantenere una normale organizzazione della rete vascolare (*Arthur e Roman, 2022*).

L'alterazione della segnalazione non determina quindi semplicemente la formazione di vasi più fragili, ma compromette i meccanismi che regolano il comportamento delle cellule endoteliali durante l'angiogenesi e il rimodellamento vascolare. La conseguente perdita dell'equilibrio angiogenico contribuisce alla formazione delle teleangectasie e delle malformazioni artero-venose caratteristiche dell'HHT.

L'HHT può essere quindi interpretata come una patologia legata all'integrità dell'omeostasi vascolare, nella quale la ridotta attività dell'asse BMP9/BMP10-ENG-ALK1-SMAD1/5-SMAD4 determina un'alterata risposta delle cellule endoteliali agli stimoli angiogenici (*Arthur e Roman, 2022*).

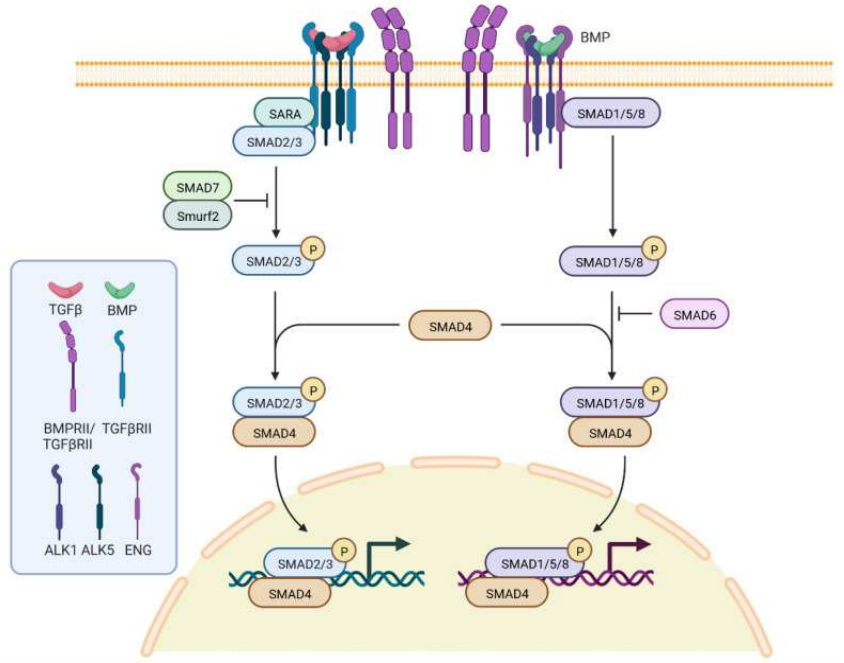


Figura 4. Pathway di segnalazione del TGF-β (*Danesino et al., 2023*).

GENI COINVOLTI

ENG

CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT

ENG o *Endoglina* (OMIM *131196) è un gene localizzato sul braccio lungo cromosoma 9, in corrispondenza della banda 9q34.11, e codifica per ENDOGLINA (ENG), una glicoproteina transmembrana di 658 amminoacidi, espressa prevalentemente a livello delle cellule endoteliali, sui monociti e sui macrofagi. Come visto, ENG è una componente del sistema di segnalazione del TGF-β, e partecipa alla regolazione di diversi processi coinvolti nell'omeostasi vascolare e nell'angiogenesi (*Abdalla e Letarte, 2006*).

Nella descrizione classica della sua organizzazione strutturale, il gene *ENG* si estende per circa 40 kb di DNA genomico, ed è costituito da 15 esoni codificanti. Il gene *ENG* può inoltre, mediante splicing alternativo, dare origine a due isoforme che differiscono per la porzione citoplasmatica della proteina.

La caratterizzazione della regione 5' di *ENG* ha evidenziato una complessa organizzazione della regione del promotore. Questo, infatti, è privo delle classiche sequenze TATA e CAAT box e presenta diversi siti di legame per fattori di trascrizione e molteplici siti di inizio della trascrizione. Queste caratteristiche contribuiscono alla regolazione dell'espressione di *ENG*, particolarmente importante a livello delle cellule endoteliali (*Rius et al., 1998*).

L'associazione tra *ENG* e l'HHT è stata dimostrata nel 1994, quando il gene è stato identificato come responsabile dell'HHT di tipo 1. L'identificazione di varianti patogenetiche in famiglie affette ha permesso di stabilire il ruolo di *ENG* nella malattia (*McAllister et al., 1994*).

ENG è una glicoproteina transmembrana di tipo I, organizzata in un ampio dominio extracellulare, un singolo dominio transmembrana idrofobico e una breve regione citoplasmatica. La proteina è espressa sulla superficie cellulare prevalentemente sotto forma di omodimero, nel quale i monomeri sono associati mediante ponti disolfuro (*Gougos e Letarte, 1990*).

Il dominio extracellulare costituisce la porzione maggiore della proteina ed è responsabile delle principali interazioni extracellulari di endoglina. La caratterizzazione tridimensionale della proteina ha fornito una base strutturale per interpretare gli effetti di alcune varianti identificate nei pazienti con HHT1 (*López-Novoa e Bernabeu, 2010; Saito et al., 2017*).

A livello della regione citoplasmatica sono state descritte due isoforme di ENG, la Long-endoglina (L-ENG), forma predominante nelle cellule endoteliali, e la Short-endoglina (S-ENG), generate mediante splicing alternativo. Le due isoforme condividono il dominio extracellulare e quello transmembrana, ma differiscono per la lunghezza e la composizione della porzione citoplasmatica. Questa differenza determina funzioni diverse nella modulazione del segnale e del comportamento delle cellule endoteliali (*Bellon et al., 1993*).

Accanto alle forme associate alla membrana, è stata descritta anche una forma solubile di ENG (sENG), costituita solo dalla porzione extracellulare della proteina rilasciata dalla superficie cellulare tramite proteolisi ad opera di metallo proteasi, in particolare MMP-14 . La forma solubile può interferire con la normale segnalazione vascolare ed è stata studiata in differenti condizioni patologiche di disfunzione endoteliale (*Venkatesha et al., 2006*).

ENDOGLINA E HHT1

Le varianti patogenetiche del gene *ENG* sono responsabili dell'HHT di tipo 1, una delle principali forme della malattia.

Il meccanismo patogenetico classicamente associato all'HHT1 è rappresentato dall'aploinsufficienza. In un individuo eterozigote per una variante patogenetica di *ENG*, la presenza di una sola copia funzionale del gene non permette di produrre una quantità sufficiente di proteina funzionale sulla superficie delle cellule endoteliali (*McAllister et al., 1994*).

Molte varianti determinano la produzione di trascritti o proteine che non raggiungono la membrana cellulare. Ad esempio, varianti che introducono codoni di stop prematuri possono causare la degradazione prematura dell'mRNA mediante il meccanismo di nonsense-mediated decay (NMD), mentre alcune varianti possono produrre proteine mal ripiegate, che non sono in grado di svolgere la loro normale funzione all'interno della cellula (*Gallione et al., 1998*).

La riduzione dell'espressione di ENG può quindi compromettere la risposta delle cellule endoteliali agli stimoli emodinamici, alterando il corretto rimodellamento vascolare (*Baeynes et al., 2016*)

VARIANTI DEL GENE *ENG*

Il gene *ENG* presenta una notevole eterogeneità mutazionale e non è stato possibile identificare un hotspot in cui si trovino concentrate la maggior parte delle mutazioni che causano l'HHT1. Le varianti patogenetiche sono distribuite lungo le diverse regioni codificanti del gene, e tra queste troviamo varianti missenso, non-senso, frameshift, varianti dei siti di splicing, ma anche inserzioni e delezioni, oltre a riarrangiamenti che possono coinvolgere più esoni (*Shovlin et al., 1997*).

Numerose varianti patogenetiche di *ENG* determinano l'introduzione di un codone di stop prematuro, causando l'interruzione anticipata della sintesi proteica. È stato riportato che circa l'80% delle mutazioni identificate presenta in alcuni casi la comparsa di codoni di stop prematuri e la produzione di proteine tronche, mentre nel resto dei casi si ha la degradazione del messaggero per NMD (*Gallione et al., 1998*).

Le varianti missenso sono alterazioni della sequenza nucleotidica che determinano la sostituzione di un amminoacido con un altro all'interno della proteina e possono causare differenze nella struttura e nella funzionalità della proteina. In alcuni casi viene compromesso il corretto folding della proteina, determinando la ritenzione di essa a livello intracellulare, e quindi una ridotta espressione della proteina stessa. La caratterizzazione strutturale di ENG si è quindi rivelata fondamentale per comprendere le conseguenze funzionali di queste varianti (*Lux et al., 2000; Saito et al., 2017*).

Le varianti patogenetiche di splicing sono alterazioni della sequenza di DNA che interferiscono con il normale processo di maturazione dell'mRNA. Durante lo splicing, gli introni vengono rimossi e gli esoni vengono uniti tra loro per formare l'mRNA maturo. Una variante che alteri le sequenze dei siti di splicing va a compromettere l'intero processo, determinando in alcuni casi l'eliminazione di un

esone o in altre la ritenzione di sequenze introniche. Queste alterazioni possono quindi causare un frameshift, che può poi determinare la comparsa di un codone di stop prematuro o la produzione di una proteina alterata, più corta o più lunga, che non funziona correttamente (*Abdalla e Letarte, 2006*).

Oltre alle alterazioni nelle regioni codificanti, sono state anche individuate varianti localizzate nelle regioni regolatorie o non codificanti del gene, che sono comunque in grado di influenzarne l'espressione. Queste regioni assumono particolare rilevanza nei pazienti con fenotipo HHT nei quali la ricerca di mutazioni all'interno delle regioni codificanti non ha permesso di identificare alcuna variante causativa. Le regioni non codificanti di maggiore interesse sono la regione del promotore e la regione 5' UTR, entrambe coinvolte nella regolazione dell'espressione di *ENG*.

L'importanza dello studio di queste regioni è stata ulteriormente promossa in seguito all'identificazione di varianti nella regione 5' UTR di *ENG* in pazienti con fenotipo HHT1. Alcune di queste varianti determinano la formazione di nuovi siti di inizio della traduzione a monte del sito canonico, con conseguente generazione di upstream open reading frame (uORF), che interferiscono con la normale traduzione dell'mRNA, causando una riduzione della produzione di ENG, e rappresentando quindi un ulteriore meccanismo causativo dell'HHT1 (*Soukarieh et al., 2023*).

L'interpretazione clinica delle varianti identificate richiede infine l'integrazione di diverse evidenze, tra cui la frequenza della variante nella popolazione, dati funzionali, segregazione familiare e informazioni disponibili nei database clinici, utili per la raccolta delle varianti genomiche e delle relative interpretazioni (*Landrum et al., 2018*).

ACVRL1

CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT

Il gene *ACVRL1* (Activin A Receptor Type II-Like 1) (OMIM *601284)**,** è uno dei principali geni coinvolti nella Teleangectasia Emorragica Ereditaria, in particolare nell'HHT di tipo 2.
Il gene è localizzato sul braccio lungo del cromosoma 12, nella regione 12q13.13, si estende per circa 16,4 kb ed è suddiviso in 10 esoni. Annotazioni genomiche successive hanno permesso di identificare ulteriori regioni esoniche non codificanti e differenti trascritti del gene. Inoltre, la caratterizzazione della regione 5' ha evidenziato la presenza di differenti siti di inizio della trascrizione e di un ulteriore esone non codificante, indicando una maggior complessità della regolazione trascrizionale di *ACVRL1* (*Berg et al., 1997*). L'espressione di *ACVRL1* è particolarmente rilevante a livello delle

cellule endoteliali, in quanto la proteina espressa ha un ruolo importante nella regolazione dell'angiogenesi e nel mantenimento dell'omeostasi vascolare (*Johnson et al., 1996*).

ALK1 appartiene alla famiglia dei recettori di tipo I della superfamiglia del TGF-β, ed è una serina/treonina chinasi transmembrana. Tra i principali ligandi di ALK1 sono stati identificati BMP9 e BMP10, che svolgono un importante ruolo nella regolazione della risposta delle cellule endoteliali. Come visto, il legame del ligando determina la formazione di un complesso tra recettori di tipo I e di tipo II, dove questi ultimi fosforilano e attivano ALK1, che a sua volta promuove principalmente la fosforilazione delle proteine SMAD1/5/8. Le SMAD attivate formano quindi complessi con SMAD4, fosforilandolo, raggiungendo il nucleo e regolando la trascrizione di geni coinvolti nel comportamento delle cellule endoteliali e nel rimodellamento vascolare (*David et al., 2007*).

L'attività di ALK1 contribuisce in particolare alla regolazione della proliferazione e della migrazione delle cellule endoteliali e al controllo dell'angiogenesi. Studi sperimentali hanno infatti dimostrato che BMP9, attraverso ALK1 può inibire la proliferazione, la migrazione endoteliale e contrastare l'angiogenesi indotta da VEGF, contribuendo al mantenimento di uno stato vascolare quiescente (*Scharpfnecker et al., 2007*).

La proteina prodotta dal gene *ACVRL1*, ALK1, ha la struttura caratteristica dei recettori di tipo I della superfamiglia del TGF-β. È costituito da un dominio extracellulare N-terminale, una singola regione transmembrana e un ampio dominio intracellulare con attività serina/treonina chinasica.
Tra la regione transmembrana e il dominio chinasico è presente una regione ricca in glicina e serina, denominata dominio GS, fondamentale per l'attivazione del recettore. Il dominio extracellulare contiene residui di cisteina conservati ed è coinvolto nell'interazione con i ligandi, mentre il dominio intracellulare media la trasduzione del segnale attraverso la fosforilazione delle proteine SMAD (*Abdalla e Letarte, 2006*).

ACVRL1 E HHT2

Il coinvolgimento di *ACVRL1* nell'HHT è stato dimostrato nel 1996, grazie all'identificazione di diverse varianti del gene in famiglie affette dalla malattia. Da allora, le alterazioni patogenetiche di *ACVRL1* sono state riconosciute come responsabili dell'HHT di tipo 2 (*Johnson et al., 1996*).

Il principale meccanismo molecolare proposto è rappresentato dall'aploinsufficienza, per cui la presenza di una sola copia funzionale di *ACVRL1* non garantisce un'attività sufficiente del pathway

nelle cellule endoteliali. Studi funzionali condotti su varianti identificate in pazienti HHT2 hanno infatti dimostrato che numerosi recettori ALK1 mutati presentano una ridotta o assente capacità di trasdurre il segnale indotto da BMP9, senza esercitare necessariamente un effetto dominante negativo sul recettore wild-type (*Ricard et al., 2010*).

La compromissione della segnalazione BMP9/BMP10-ALK1-SMAD1/5/8 altera la risposta delle cellule endoteliali agli stimoli angiogenici e il corretto rimodellamento vascolare. La perdita di attività di ALK1 può quindi favorire un comportamento endoteliale anomalo, caratterizzato da alterazioni della proliferazione, della migrazione e dell'organizzazione dei vasi, contribuendo alla formazione delle teleangectasie e delle malformazioni artero-venose caratteristiche dell'HHT (*Kim et al., 2012*).

Da un punto di vista clinico, pur presentando una notevole variabilità individuale, l'HHT2 mostra alcune differenze rispetto all'HHT1. In particolare, diversi studi di correlazione genotipo-fenotipo hanno evidenziato una maggiore frequenza di MAV epatiche nei pazienti con HHT2, mentre le MAV polmonari e cerebrali risultano generalmente più frequenti nell'HHT1 (*Bayrak-Toydemir et al., 2006*).

VARIANTI DEL GENE *ACVRL1*

Anche il gene *ACVRL1* presenta un'elevata eterogeneità mutazionale e non è stato identificato un singolo hotspot responsabile della maggiorparte dei casi di HHT2. Sono state descritte diverse tipologie di varianti patogenetiche, come varianti missenso, non-senso, frameshift, varianti di splicing, piccole inserzioni e delezioni, oltre a delezioni o duplicazioni che possono interessare uno o più esoni.

Una grande percentuale di mutazioni nel gene *ACVRL1* è rappresentata dalle varianti missenso. Queste alterazioni sono localizzate sia nel dominio extracellulare, sia nella porzione intracellulare della proteina, in particolare nel dominio serina/treonina chinasi. La sostituzione di residui importanti per la struttura o per l'attività catalitica del recettore può compromettere la capacità di ALK1 di trasdurre correttamente il segnale intracellulare (*Abdalla e Letarte, 2006; Scotti et al., 2011*).

La presenza della proteina mutata sulla membrana cellulare non implica necessariamente il mantenimento della sua funzionalità. L'analisi sperimentale di diverse varianti missenso associate all'HHT2 ha infatti dimostrato che molti recettori ALK1 mutati possono essere correttamente espressi

e raggiungere la superficie cellulare, ma presentano una ridotta o assente risposta alla stimolazione con BMP9. La valutazione funzionale della risposta a BMP9 rappresenta quindi uno strumento importante per stabilire la patogenicità di nuove varianti missenso o di varianti il cui significato clinico non sia immediatamente evidente (*Ricard et al., 2010*).

Le varianti nonsense e frameshift possono introdurre un codone di stop prematuro, determinando la produzione di una proteina tronca oppure la degradazione del trascritto mediante NMD. Anche le varianti di splicing possono interferire con la corretta maturazione dell'mRNA di *ACVRL1*, determinando trascritti anomali e compromettendo la produzione di ALK1 funzionale. Studi basati su saggi funzionali e minigene hanno dimostrato che alcune varianti introniche, e persino alcune all'interno di esoni, possono alterare lo splicing e determinare frameshift e codoni di stop prematuri (*Alaa El Din et al., 2015*).

Sono state inoltre identificate delezioni e duplicazioni di interi esoni o di più esoni, non sempre rilevabili mediante il semplice sequenziamento della regione codificante. L'impiego di metodiche, come l'MLPA e la q-PCR, in grado di identificare variazioni del numero di copie ha permesso di ampliare lo spettro delle alterazioni molecolari associate all'HHT2 (*Gedge at al., 2007*).

SMAD4

CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT

Il gene *SMAD4* (SMAD Family Member 4) (OMIM * 600993), precedentemente noto come *MADH4* o *DPC4*, è localizzato sul braccio lungo del cromosoma 18, nella regione 18q21.2. Il gene si estende per circa 55 kb di DNA genomico, ed è costituito da 12 esoni, la cui porzione codificante determina la produzione di una proteina di 552 amminoacidi. Appartiene alla famiglia dei geni *SMAD*, i cui prodotti proteici svolgono un ruolo fondamentale nella trasduzione dei segnali appartenenti alla superfamiglia del TGF-β.

A differenza di *ENG* e *ACVRL1*, che codificano per un co-recettore e un recettore di membrana espressi a livello dell'endotelio, *SMAD4* codifica per una proteina intracellulare che agisce come mediatore comune di diversi pathway della superfamiglia del TGF-β. SMAD4 è infatti classificata come co-SMAD (common mediator SMAD) e svolge la propria funzione formando complessi con le SMAD attivate dai recettori, permettendo la trasmissione del segnale dalla membrana cellulare al nucleo.

Attraverso questo meccanismo, SMAD4 partecipa alla regolazione di numerosi processi cellulari, tra cui la proliferazione, differenziamento, apoptosi e omeostasi tissutale. Nell'endotelio vascolare svolge inoltre un ruolo importante nella regolazione dell'angiogenesi e nel mantenimento della rete vascolare (*Lagna et al., 1996*).

La proteina SMAD4 presenta un'organizzazione strutturale caratteristica delle proteine della famiglia SMAD ed è costituita da due domini altamente conservati, denominati MH1 (Mad Homology 1) e MH2 (Mad Homology 2), separati da una regione intermedia meno conservata, la linker region.

Il dominio MH1, localizzato nella porzione N-terminale della proteina, è coinvolto principalmente nell'interazione con il DNA e contribuisce quindi al riconoscimento delle sequenze regolatorie dei geni bersaglio. Il dominio MH2, localizzato nella porzione C-terminale, è invece fondamentale per le interazioni proteina-proteina e permette a SMAD4 di associarsi alle SMAD attivate dal recettore e ad altri cofattori coinvolti nella regolazione della trascrizione. La regione linker che separa i due domini contribuisce alla regolazione dell'attività della proteina (*Heldin et al., 1997*).

Nel pathway rilevante per l'HHT, SMAD4 rappresenta un elemento comune e centrale della trasduzione del segnale, posto a valle del complesso recettoriale BMP9/BMP10-ENG-ALK1 (*David et al., 2007*).

SMAD4 **E JP-HHT**

Le varianti germinali patogeniche di *SMAD4* sono state inizialmente associate alla Sindrome da poliposi giovanile o JPS (Juvenile Polyposis Syndrome), una malattia ereditaria autosomica dominante caratterizzata dalla predisposizione allo sviluppo di polipi amartomatosi del tratto gastrointestinale e da un aumentato rischio di neoplasie gastrointestinali (*Zbuk e Eng, 2007*).

È stato successivamente dimostrato che varianti germinali di *SMAD4* determinano un fenotipo combinato caratterizzato dalla presenza contemporanea di manifestazioni della poliposi giovanile e della Teleangectasia Emorragica Ereditaria. Questa condizione è stata definita sindrome JP-HHT o Juvenile Polyposis - Hereditary Hemorrhagic Telangiectasia.

Per arrivare a questa conclusione, sono state inizialmente studiate sette famiglie non correlate tra loro, nelle quali segregavano caratteristiche cliniche riconducibili sia alla JPS sia all'HHT e varianti di SMAD4 negli individui affetti. I pazienti presentavano, oltre alla poliposi gastrointestinale, anche

manifestazioni cliniche caratteristiche dell'HHT, tra cui epistassi ricorrenti, teleangectasie mucocutanee e malformazioni artero-venose, soprattutto a livello polmonare e epatico. Questi risultati hanno dimostrato che *SMAD4* rappresenta il collegamento genetico tra le due condizioni (*Gallione et al., 2004*).

Successivamente, varianti di *SMAD4* sono state identificate anche in pazienti sottoposti ad analisi genetica per HHT e risultati negativi per varianti di *ENG* e *ACVRL1*. Questo ha evidenziato l'importanza dell'analisi di *SMAD4* nei pazienti con sospetto clinico di HHT in cui non è stata identificata alcuna variante nei due principali geni associati alla malattia (*Gallione et al., 2006*).

La sindrome JP-HHT presenta quindi un fenotipo più complesso rispetto alle forme classiche di HHT1 e HHT2, poiché alle manifestazioni vascolari tipiche dell'HHT si associa la predisposizione alla poliposi e allo sviluppo di tumori gastrointestinali, in particolare del colon-retto. Per questo, l'identificazione di una variante patogenetica in *SMAD4* in un paziente con HHT assume particolare rilevanza clinica, poiché comporta anche la necessità di considerare anche il rischio associato alla JPS.

VARIANTI DEL GENE *SMAD4*

Il gene *SMAD4* presenta una notevole eterogeneità mutazionale. Sono state descritte differenti tipologie di varianti germinali patogenetiche, comprendenti varianti missenso, non-senso, frameshift, alterazioni dei siti di splicing e delezioni, comprese alterazioni che possono coinvolgere porzioni più estese del gene.

Una particolare concentrazione di varianti patogenetiche è stata osservata nel dominio MH2, fondamentale per l'interazione di SMAD4 con le altre proteine SMAD e con diversi cofattori trascrizionali (*Gallione et al., 2010*).

Le prime osservazioni nei pazienti con JP-HHT avevano suggerito una probabile correlazione genotipo-fenotipo, poiché le varianti identificate risultavano prevalentemente concentrate nella porzione C-terminale della proteina, comprendente il dominio MH2. Questa osservazione aveva portato a ipotizzare che specifiche varianti localizzate in questo dominio fossero responsabili del fenotipo combinato JP-HHT (*Gallione et al., 2004*).

L'analisi di un numero maggiore di pazienti ha poi dimostrato che, sebbene molte delle varianti associate alla JP-HHT siano effettivamente localizzate nel dominio MH2, anche alterazioni presenti in altre regioni di *SMAD4* possono determinare il fenotipo combinato. Lo spettro delle varianti osservate nella JPS e nella JP-HHT risulta infatti ampiamente sovrapponibile, rendendo difficile stabilire una correlazione rigorosa tra la porzione della variante e lo sviluppo delle manifestazioni vascolari (*Gallione et al., 2010*).

Di conseguenza, la presenza di una variante patogenetica germinale di *SMAD4* deve essere considerata potenzialmente associata alle manifestazioni di entrambe le condizioni. Un paziente identificato inizialmente per JPS può essere a rischio di sviluppare manifestazioni dell'HHT e, viceversa, un paziente con HHT e una variante di *SMAD4* deve essere considerato a rischio di poliposi e neoplasie gastrointestinali.

GDF2

CARATTERISTICHE GENERALI E RELAZIONE CON L'HHT

Il gene *GDF2* (Growth Differentiation Factor 2) (OMIM *605120) è localizzato sul braccio lungo del cromosoma 10, nella regione 10q11.22, si estende per 5,1 kb e presenta un'organizzazione genomica relativamente semplice, costituita da 2 esoni.

Il gene codifica per BMP9 (Bone Morphogenetic Protein 9), una proteina di 429 amminoacidi appartenente alla famiglia delle Bone Morphogenetic Proteins e alla superfamiglia del TGF-β. Le BMPs sono proteine di segnalazione extracellulari coinvolte nella regolazione di numerosi processi biologici attraverso l'attivazione di complessi recettoriali di tipo I e tipo II, e la conseguente trasduzione intracellulare del segnale mediata principalmente dalle proteine SMAD (*Miyazono et al., 2010*).

BMP9 viene sintetizzata come una proteina precursore, costituita da un peptide segnale, un pro-dominio e una regione C-terminale che dà poi origine alla proteina matura. In seguito al taglio proteolitico, la forma matura di BMP9 forma un dimero biologicamente attivo, in cui il pro-dominio può rimanere associato alla proteina matura anche dopo la secrezione. BMP9 viene prodotta principalmente a livello epatico, in particolare dagli epatociti, e viene rilasciata nella circolazione sanguigna, dove è presente prevalentemente in una forma matura biologicamente attiva associata al pro-dominio (*Brown et al., 2005; Bidart et al., 2012*).

A livello vascolare, BMP9 svolge un ruolo importante nella regolazione dell'omeostasi endoteliale. BMP9 lega con elevata affinità il recettore di tipo I ALK1, codificato da *ACVRL1*, e la formazione del complesso determina l'attivazione delle proteine SMAD1/5/8, coinvolte nella regolazione della risposta delle cellule endoteliali. Questo colloca BMP9 nello stesso pathway di segnalazione in cui sono coinvolti altri geni associati all'HHT, tra cui *ENG*, *ACVRL1* e *SMAD4* (*Miyazono et al., 2010*).

GDF2 E HHT

Il coinvolgimento di *GDF2* nei fenotipi correlati all'HHT è emerso dall'identificazione di varianti eterozigoti missenso del gene in soggetti con manifestazioni cliniche compatibili con la malattia, ma privi di varianti causative nei principali geni associati all'HHT. In particolare, l'analisi di una coorte di 191 individui non correlati ha permesso di identificare varianti di *GDF2* in tre soggetti.

I soggetti portatori presentavano i segni classici della malattia come epistassi e lesioni cutanee poi identificate come teleangectasie. Tuttavia, la localizzazione e l'aspetto di queste lesioni non erano completamente compatibili con l'HHT classica, ma mostravano alcune somiglianze con le lesioni osservate nella Sindrome da malformazione capillare, e questo ha permesso di descrivere il quadro clinico come una sindrome da anomalie vascolari con sovrapposizione fenotipica con l'HHT (o HHT-like) (*Wooderchak-Donahue et al., 2013*).

Un'ulteriore evidenza del coinvolgimento di *GDF2* nell'HHT è stata fornita grazie all'identificazione di una variante eterozigote di *GDF2* segregante in tre membri affetti della stessa famiglia, negativi all'analisi dei geni *ENG*, *ACVRL1* e *SMAD4*. Tutti i soggetti presentavano epistassi e teleangectasie tipiche dell'HHT, ma anche gravi MAV polmonari.

La caratterizzazione funzionale della variante identificata ha mostrato che la proteina mutata era sintetizzata sotto forma di pro-proteina, ma non era in grado di raggiungere la forma di dimero maturo, ossia la forma biologicamente attiva di BMP9. Inoltre, nei soggetti portatori della variante erano presenti livelli plasmatici di BMP9 significativamente inferiori rispetto ai controlli. Questi risultati hanno fornito evidenze genetiche e funzionali a sostegno del fatto che varianti patogenetiche di GDF2 possono rappresentare una causa di HHT (*Balachandar et al., 2022*).

Nel complesso, *GDF2* rappresenta quindi un gene raramente coinvolto nei fenotipi HHT o HHT-like. Le alterazioni nella proteina BMP9 possono compromettere la corretta attivazione della via di

segnalazione BMP9-ALK1-SMAD, fornendo un ulteriore supporto al ruolo centrale di questo pathway nella patogenesi delle anomalie vascolari caratteristiche dell'HHT.

RASA1

Il gene *RASA1* (RAS p21 protein activator 1) (OMIM *139150) è localizzato sul braccio lungo del cromosoma 5, nella regione 5q14.3, si estende per circa 124 kb divisi in 25 esoni. Il gene codifica per p120-RasGAP, una proteina di 1.047 amminoacidi appartenente alla famiglia delle proteine attivanti le GTPasi di Ras (RasGAP). p120-RasGAP svolge un ruolo importante nella regolazione della via di segnalazione di Ras, coinvolta nel controllo di diversi processi cellulari, tra cui proliferazione, differenziamento, migrazione e sopravvivenza cellulare (*Pamonsinlapatham et al., 2009*).

La proteina p120-RasGAP agisce come regolatore negativo dell'attività di Ras, favorendo l'idrolisi del GTP legato a Ras e determinandone quindi il passaggio dalla forma attiva, legata al GTP, alla forma inattiva, legata al GDP. Attraverso questo meccanismo, *RASA1* contribuisce alla modulazione dei segnali intracellulari attivati da diversi recettori per fattori di crescita. La sua funzione risulta particolarmente importante anche a livello delle cellule endoteliali e nello sviluppo e mantenimento della rete vascolare.

Varianti patogeniche eterozigoti di *RASA1* sono principalmente associate alla Sindrome da malformazione capillare-malformazione arterovenosa di tipo 1 (CM-AVM1), una condizione autosomica dominante caratterizzata dalla presenza di malformazioni capillari cutanee multiple, che possono associarsi a malformazioni arterovenose (AVM) e fistole arterovenose localizzate in differenti distretti dell'organismo. In alcuni pazienti può essere presente anche la sindrome di Parkes Weber, caratterizzata da malformazioni vascolari associate a crescita eccessiva dell'arto interessato (*Eerola et al., 2003*).

RASA1 E HHT

La relazione tra il gene *RASA1* e HHT deriva principalmente dalla parziale sovrapposizione fenotipica tra CM-AVM e l'HHT. Entrambe le condizioni possono essere caratterizzate dalla presenza di anomalie vascolari, comprese malformazioni arterovenose e lesioni vascolari cutanee, che in alcuni pazienti possono rendere più complessa la distinzione clinica tra le due patologie (*Revencu et al., 2013*).

Il possibile coinvolgimento di RASA1 in pazienti con caratteristiche cliniche compatibili con l'HHT è stato ulteriormente evidenziato dall'identificazione di varianti patogenetiche del gene in soggetti precedentemente risultati negativi per varianti causative nei principali geni associati all'HHT. In particolare, l'analisi di 93 individui non correlati, tutti con sospetto clinico di HHT, ma negativi all'analisi dei geni *ENG*, *ACVRL1* e *SMAD4*, ha permesso di identificare due varianti patogenetiche di *RASA1* in due soggetti. Entrambi presentavano epistassi e lesioni cutanee descritte clinicamente come teleangectasie (*Hernandez et al., 2015*).

Questi risultati evidenziano come le manifestazioni associate a varianti di *RASA1* possano in alcuni casi sovrapporsi al fenotipo dell'HHT. Tuttavia, *RASA1* è associato alla CM-AVM1 e il suo coinvolgimento in pazienti con sospetta HHT deve essere interpretato soprattutto nell'ambito della diagnosi differenziale delle anomalie vascolari ereditarie, piuttosto che come una causa dall'HHT classica.

EPHB4

Il gene *EPHB4* (EPH receptor B4) (OMIM *600011) è localizzato sul braccio lungo del cromosoma 7, nella regione 7q22.1, si estende per 25 kb ed è composto da 17 esoni. Il gene codifica per EPHB4, un recettore tirosin-chinasico di 987 amminoacidi, appartenente alla famiglia dei recettori Eph, coinvolti nella regolazione di diversi processi cellulari e nello sviluppo del sistema vascolare.

EPHB4 è una proteina transmembrana il cui principale ligando è ephrin-B2 (EFNB2). Durante lo sviluppo vascolare, l'espressione di questi due componenti contribuisce alla corretta organizzazione della componente vascolare arteriosa e venosa della rete vascolare. Ephrin-B2 è prevalentemente associata alle cellule endoteliali arteriose, mentre EPHB4 rappresenta un marcatore delle cellule endoteliali venose. L'interazione tra Ephrin-B2 e EPHB4 svolge un ruolo importante nella regolazione della morfogenesi e del rimodellamento vascolare (*Adams et al., 1999*).

La segnalazione mediata da EPHB4 è funzionalmente collegata anche alla proteina p120-RasGAP, codificata dal gene *RASA1*. P120-RasGAP agisce come effettore di EPHB4 e contribuisce alla regolazione della via RAS-MAPK. La compromissione di questa interazione può determinare un'alterata attivazione della segnalazione RAS-MAPK, contribuendo allo sviluppo di anomalie vascolari (*Amyere at al., 2017*).

Varianti germinali eterozigoti loss-of-function di *EPHB4* sono associate alla Sindrome da malformazione capillare-malformazione arterovenosa di tipo 2 (CM-AVM2), una malattia vascolare autosomica dominante caratterizzata principalmente dalla presenza di malformazioni capillari multifocali, che possono associarsi a teleangectasie , malformazioni arterovenose e fistole arterovenose. Il fenomeno presenta una notevole sovrapposizione con quello della CM-AVM1 associata a varianti di *RASA1* (*Amyere et al., 2017*).

EPHB4 E HHT

Il coinvolgimento di *EPHB4* nel contesto dell'HHT è legato al fatto che alcune manifestazioni cliniche della CM-AVM2 possono sovrapporsi a quelle dell'HHT. In particolare, nei soggetti in cui sono state trovate varianti patogenetiche di *EPHB4*, spesso sono presenti fenomeni di epistassi, teleangectasie e malformazioni arterovenose, sintomi caratteristici che possono orientare verso una diagnosi clinica di HHT (*Wooderchak-Donahue et al., 2019*).

L'analisi di pazienti con CM-AVM2 ha tuttavia evidenziato alcune differenze rispetto all'HHT classica. Sebbene siano presenti epistassi e teleangectasie anche molto frequenti, la distribuzione delle lesioni vascolari, così come alcune caratteristiche cliniche consentono di distinguere i due fenotipi.

Inoltre, *EPHB4* non appartiene direttamente al pathway BMP9/BMP10-ENG-ALK1-SMAD, responsabile delle principali forme di HHT, ma si inserisce in una via di segnalazione coinvolta nella regolazione dello sviluppo vascolare (*Amyere at al., 2017*).

EPHB4 non viene quindi considerato uno dei principali geni alla base dell'HHT classica, ma rappresenta un gene di interesse nella diagnosi differenziale dei pazienti con manifestazioni HHT-like, soprattutto nei soggetti nei quali non vengono identificate varianti patogenetiche nei principali geni associati alla malattia. Inoltre, la sovrapposizione fenotipica tra CM-AVM2 e HHT sottolinea l'importanza dell'analisi genetica per riuscire a distinguere condizioni vascolari ereditarie clinicamente simili.

BMPR2

Il gene *BMPR2* (Bone Morphogenetic Protein Receptor Type 2) (OMIM * 600799) è un gene localizzato sul braccio lungo del cromosoma 2, nella regione 2q33.1-q33.2, si estende per circa 191 kb ed è composto da 13 esoni. Il gene codifica per BMPR2, un recettore transmembrana di 1.038

amminoacidi, appartenente alla famiglia dei recettori di tipo II per le Bone Morphogenetic Proteins (BMPs), a loro volta appartenenti alla superfamiglia del TGF-β.

BMPR2 è un recettore transmembrana con attività serina/treonina chinasica e partecipa alla trasduzione del segnale mediato dalle BMP. In seguito al legame con il ligando, BMPR2 forma complessi con i recettori BMP di tipo I, promuovendone l'attivazione attraverso fosforilazione, e la trasduzione del segnale mediante il coinvolgimento delle proteine SMAD1/5/8 prima, e SMAD4 come visto precedentemente. Attraverso questo pathway, *BMPR2* partecipa alla regolazione di processi cellulari come omeostasi e rimodellamento vascolare (*Miyazono et al., 2010*).

Varianti germinali eterozigoti patogenetiche di *BMPR2* rappresentano la principale causa genetica di ipertensione arteriosa polmonare ereditaria (HPAH). Le prime evidenze dell'associazione tra *BMPR2* e la malattia sono state ottenute attraverso l'identificazione di varianti germinali del gene in famiglie con ipertensione polmonare primaria, oggi classificata nell'ambito della PAH. Le varianti patogenetiche determinano prevalentemente una perdita di funzione del recettore e compromettono la corretta segnalazione mediata dalle BMP, contribuendo all'alterazione dell'omeostasi vascolare polmonare (*Deng et al., 2000*).

BMPR2 **E HHT**

Il coinvolgimento di *BMPR2* assume interesse anche nel contesto dell'HHT per la correlazione molecolare che c'è tra i pathway alla base delle due patologie. L'HHT e la PAH sono entrambe associate ad alterazioni di componenti della superfamiglia del TGF-β/BMP, e alcuni geni principalmente coinvolti nell'HHT, in particolare *ACVRL1*, possono essere associati anche allo sviluppo di PAH. In famiglie nelle quali l'HHT e la PAH coesistono, le varianti di *ACVRL1* rappresentano una delle principali cause genetiche della combinazione dei due fenotipi (*Trembath et al., 2001*).

Più raramente, varianti patogenetiche di *BMPR2* sono state identificate in pazienti con PAH, ma che presentavano manifestazioni cliniche caratteristiche dell'HHT. È stato descritto, ad esempio, un soggetto con PAH, epistassi, teleangectasie nasali e malformazioni arterovenose polmonari che all'analisi genetica è risultato negativo per varianti patogenetiche di *ENG*, *ACVRL1* e *SMAD4*, e nel quale è stata identificata una variante germinale non-senso di *BMPR2* (*Rigelsky et al., 2008*).

MODELLO "SECOND HIT" NELL'HHT

Il principale meccanismo genetico alla base dell'insorgenza dell'HHT è, come visto, l'aploinsufficienza, processo in cui la presenza di una variante germinale loss-of-function in eterozigosi determina la perdita di una delle due copie funzionali del gene, con conseguente riduzione della quantità di proteina disponibile o funzionante.

Il modello dell'aploinsufficienza non è tuttavia sufficiente a spiegare alcune caratteristiche della malattia. Nonostante la variante germinale sia presente in tutte le cellule dell'organismo, le teleangectasie e le MAV si sviluppano in maniera localizzata, in specifiche regioni del sistema vascolare. Inoltre, le manifestazioni cliniche mostrano un'elevata variabilità in termini di localizzazione, numero e severità, e questa variabilità è notevole anche tra individui affetti appartenenti alla stessa famiglia. Questo fenomeno ha portato a ipotizzare che, oltre all'aploinsufficienza, siano necessari ulteriori eventi locali in grado di favorire lo sviluppo delle malformazioni vascolari (*Bernabeu et al., 2020*).

Una possibile spiegazione è rappresentata dal modello del "Second hit", un ipotesi proposta inizialmente da Alfred Knudson nel contesto dello studio di geni oncosoppressori. Secondo questo modello, un individuo eredita una prima variante patogenetica dalla linea germinale, e questo costituisce il primo hit, ed è quindi presente in tutte le cellule dell'organismo. Successivamente, in una sottopopolazione cellulare, si può verificare un secondo evento mutazionale che compromette la copia corretta del gene, portando così a una condizione di perdita bi-allelica della funzione del gene (*Snelling et al., 2019*).

Nel contesto dell'HHT, il primo hit corrisponde alla variante germinale nei geni *ENG*, *ACVRL1* o *SMAD4*, mentre il secondo hit è rappresentato da un evento somatico che inattiva l'allele wild-type. Questo modello permette di mettere in relazione la predisposizione genetica e le manifestazioni localizzate delle malformazioni vascolari (*DeBose-Scarlett e Marchuk; 2025*).

Una delle prime evidenze molecolari a sostegno di questo meccanismo nell'HHT è stata fornita grazie al sequenziamento di 19 teleangectasie provenienti da pazienti affetti dalla malattia. Il sequenziamento ha mostrato mutazioni somatiche a bassa frequenza nei geni *ENG* e *ACVRL1* in 9 su 19 campioni analizzati. Un elemento significativo è stato ottenuto in sette campioni in cui è stato possibile analizzare in maniera più approfondita le varianti. In tutti i casi, la variante germinale e quella somatica erano localizzate in trans, su alleli diversi dello stesso gene. Questa osservazione

indica che l'evento somatico interessa l'allele wild-type, determinando la perdita dell'unica copia funzionale, e quindi una completa perdita di funzione dei geni interessati a livello di quella che poi sarà la lesione vascolare (*Snellings et al., 2019*).

Analisi successive di altri tessuti, in cui è frequente la manifestazione di teleangectasie, come teleangectasie nasali o MAV degli organi interni, hanno avvalorato questa teoria. In uno studio condotto su tessuti provenienti da pazienti con HHT, le varianti somatiche trovate con l'analisi delle lesioni erano presenti come mosaici, con frequenze alleliche comprese tra 0,7 e 2%. La presenza di second hit a frequenza così ridotta suggerisce che solo una piccola popolazione di cellule della lesione debba perdere completamente la funzione dei geni HHT (*Whitehead et al., 2024*).

Ulteriori evidenze sono state ottenute direttamente dallo studio delle MAV associate all'HHT. L'identificazione di varianti somatiche loss-of-function in *ENG* e *ACVRL1* nelle malformazioni di organi interni ha mostrato che il meccanismo non è limitato alle piccole teleangectasie mucocutanee, ma può contribuire anche alla formazione delle malformazioni artero-venose maggiori. In particolare, sono state individuate mutazioni germinali e somatiche responsabili della perdita biallelica di *ENG* o *ACVRL1* anche in MAV cerebrali, fornendo un ulteriore supporto al modello genetico dei due eventi (*DeBose-Scarlett et al., 2024*).

L'analisi genetica è stata anche ampliata per il gene *SMAD4*, grazie all'analisi di una MAV craniofacciale di grandi dimensioni proveniente da un paziente affetto da sindrome combinata poliposi giovanile-HHT (JP-HHT), portatore di una variante germinale confermata in *SMAD4*. Nella lesione è stata identificata una seconda variante somatica nello stesso gene, in trans rispetto alla variante germinale. Nella stessa MAV è stata anche rilevata una perdita di eterozigosi somatica nella regione cromosomica del gene *SMAD4*. Questo dimostra che anche la perdita biallelica di *SMAD4* può essere associata alla formazione di lesioni vascolari, suggerendo che il modello del second hit possa essere un meccanismo condiviso in tutti e tre i principali geni responsabili dell'HHT (*DeBose-Scarlett et al., 2025*).

La presenza di una perdita biallelica, tuttavia, potrebbe rappresentare soltanto una parte del processo patogenetico. È stato ipotizzato che anche diversi fattori non genetici possono contribuire alla comparsa di lesioni vascolari in un endotelio geneticamente già predisposto. Tra questi fattori sono stati considerati danni vascolari, traumi locali, infiammazione, stimoli angiogenici, aumento dell'attività del VEGF e modificazioni delle forze emodinamiche, in particolare a causa dello shear

stress. Anche fattori ambientali e modificatori genetici potrebbero contribuire alla variabilità nella distribuzione e nella severità delle manifestazioni cliniche (*Bernabeu et al., 2020*).

L'importanza degli stimoli locali è sostenuta anche dai modelli sperimentali. La perdita della funzione dei geni HHT nelle cellule endoteliali può predisporre alla formazione delle anomalie vascolari, ma in diversi modelli animali adulti, la comparsa delle lesioni viene fortemente favorita dalla presenza di stimoli angiogenici o di un danno tissutale locale. In questo contesto, fattori come VEGF, infiammazione e alterazioni del flusso ematico potrebbero creare un microambiente favorevole all'espansione e al rimodellamento delle cellule endoteliali che hanno perso completamente la funzione dei geni HHT (*DeBose-Scarlett e Marchuk; 2025*).

Le evidenze attualmente disponibili suggeriscono quindi di considerare la patogenesi delle lesioni HHT secondo un modello multifattoriale, più complesso rispetto alla semplice aploinsufficienza. Mentre la variante germinale determina una predisposizione, eventi somatici locali possono determinare la perdita biallelica del gene in una porzione di cellule. Qui possono successivamente agire fattori angiogenici, infiammatori ed emodinamici, che contribuiscono all'inizio, espansione e rimodellamento delle malformazioni vascolari. Partendo da queste nuove conoscenze, si può affermare che il modello Second hit può andare ad ampliare quello dell'aploinsufficienza, in cui alla predisposizione determinata dalla variante germinale, si aggiungono eventi somatici e fattori legati al microambiente vascolare che contribuiscono all'insorgenza localizzata delle lesioni (*Bernabeu et al., 2020; DeBose-Scarlett e Marchuk, 2025*).

SEQUENZIAMENTO SANGER

Il sequenziamento Sanger, sviluppato da Frederick Sanger nel 1977, è una tecnica di sequenziamento del DNA basata sul principio di terminazione della catena. Prima dell'introduzione delle tecnologie di Next Generation Sequencing, il metodo Sanger ha rappresentato il principale approccio utilizzato per il sequenziamento del DNA.

Il metodo si basa sulla sintesi di un nuovo filamento di DNA complementare a una sequenza utilizzata come stampo. La reazione richiede quindi un DNA usato come templato, un primer, una DNA polimerasi e i quattro deossiribonucleotidi trifosfato (dNTP). A questi vengono aggiunte piccole quantità di dideossiribonucleotidi trifosfato (ddNTP) che sono l'elemento fondamentale del metodo.

La differenza tra dNTP e ddNTP riguarda la struttura dello zucchero. I ddNTP sono infatti privi del gruppo ossidrilico in posizione 3' (3'-OH) necessario alla formazione del legame fosfodiesterico. Durante la sintesi, la DNA polimerasi incorpora i normali dNTP e permette l'allungamento del filamento, fino a che non viene incorporato casualmente un ddNTP, che blocca la sintesi della catena. La presenza contemporanea di dNTP e ddNTP determina quindi la produzione di una popolazione di frammenti di DNA di lunghezza differente, ciascuno terminante in corrispondenza di uno specifico nucleotide (*Sanger et al., 1977*).

Nel sequenziamento Sanger automatizzato, i quattro ddNTP sono marcati con fluorocromi differenti, permettendo di distinguere la base terminale di ciascun frammento. I prodotti ottenuti dalla reazione vengono successivamente separati in funzione delle loro dimensioni mediante elettroforesi capillare. I frammenti più corti attraversano il capillare più rapidamente rispetto a quelli più lunghi e durante la migrazione, un sistema di rivelazione identifica il segnale fluorescente associato al ddNTP terminale.

I segnali rilevati vengono elaborati e rappresentati sotto forma di elettroferogramma, nel quale ogni picco corrisponde a uno dei quattro nucleotidi. L'ordine dei segnali viene ricostruito partendo dal frammento più corto al frammento più lungo, e infine viene ricostruita l'intera sequenza del filamento.

Il sequenziamento Sanger presenta un'elevata accuratezza e risulta particolarmente adatto per l'analisi di singole regioni genomiche o di un numero limitato di sequenze. Tuttavia, le regioni analizzabili contemporaneamente sono ridotte rispetto all'NGS, limitandone l'impiego quando è necessario analizzare un elevato numero di geni o estese porzioni del genoma. Nonostante queste limitazioni, il sequenziamento Sanger mantiene ancora oggi un ruolo importante nella diagnostica e nella ricerca genetica (*Slatko et al., 2018*).

NEXT GENERATION SEQUENCING (NGS)

Il Next Generation Sequencing (NGS) comprende un insieme di tecnologie di sequenziamento ad alta capacità, sviluppate per permettere l'analisi in contemporanea di un elevato numero di frammenti di DNA. Rispetto al tradizionale sequenziamento Sanger, le tecnologie NGS permettono di generare contemporaneamente milioni di sequenze, le reads, aumentando notevolmente la quantità di dati ottenuti con una singola analisi.

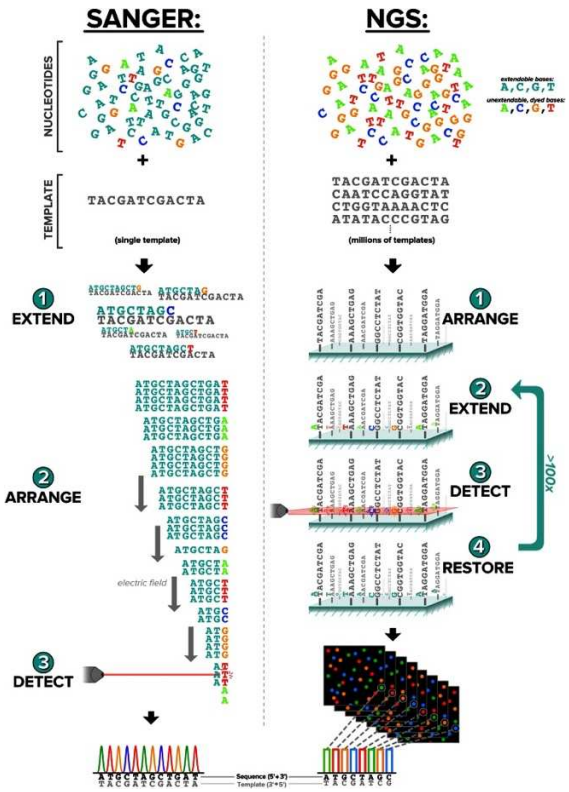


Figura 5. Confronto tra i meccanismi del sequenziamento Sanger e NGS (*Muzzey et al., 2015*).

L'introduzione dell'NGS ha cambiato significativamente l'approccio all'analisi genomica, rendendo possibile lo studio simultaneo di numerosi geni o di porzioni molto estese del genoma.
In ambito medico e diagnostico, queste tecnologie vengono utilizzate per identificare varianti genetiche causative di patologie ereditarie o acquisite, e possono essere applicate con differenti strategie a seconda dell'estensione della regione genomica che si intende analizzare. È possibile limitare l'analisi a gruppi selezionati di geni mediante pannelli multigenici oppure estenderla a tutte le regioni codificanti attraverso il Whole Exome Sequencing (WES). Con il progredire delle tecniche, ora è possibile analizzare l'intero genoma mediante il Whole Genome Sequencing (WGS) (Mandlik et al., 2024).

Il processo di sequenziamento mediante NGS può essere suddiviso schematicamente in tre fasi principali:

- Preparazione del campione e della libreria
- Sequenziamento
- Analisi bioinformatica dei dati.

La prima fase prevede l'estrazione del DNA dal campione biologico e la successiva preparazione della libreria di sequenziamento. Nelle piattaforme Short-read, il DNA viene frammentato, e legato

ad appositi adattatori oligonucleotidici. Queste brevi sequenze consentono ai frammenti di ancorarsi alla piattaforma utilizzata e forniscono i siti di innesco per la reazione di sequenziamento. In base all'obiettivo dell' analisi, la libreria può comprendere l'intero DNA genomico, oppure essere sottoposta a una fase di arricchimento per selezionare specifiche regioni d'interesse (*Muzzey et al., 2015*).

Una volta costruita la libreria, questa viene sottoposta al sequenziamento, durante il quale i numerosi frammenti vengono analizzati contemporaneamente. Il principio specifico utilizzato per determinare la sequenza nucleotidica dipende dalla piattaforma utilizzata, ma il risultato è in ogni caso la produzione di un elevato numero di reads contenenti la sequenza dei singoli frammenti di DNA (*Metzker, 2010*).

I dati prodotti dal sequenziatore vengono infine sottoposti a un'analisi bioinformatica. Dopo una prima valutazione della qualità delle sequenze ottenute, le reads vengono allineate a un genoma di riferimento, in modo da stabilire la loro posizione genomica. Successivamente, viene effettuato il confronto tra la sequenza analizzata e quella di riferimento per individuare eventuali differenze attraverso un processo definito Variant calling. In questo modo, vengono individuate le varianti, che dovranno essere successivamente annotate, filtrate e interpretate per valutarne la possibile rilevanza biologica e clinica (*Muzzey et al., 2015*).

Le applicazioni dell'NGS variano in base allo scopo dell'analisi. Il sequenziamento *De novo* consente di determinare la sequenza di un genoma in assenza di un riferimento. Quando questo è disponibile, l'analisi delle varianti può essere condotta mediante WGS, WES o approcci mirati a specifiche regioni di interesse (*Satam et al., 2023*).

L'NGS torna utile anche nello studio dell'espressione genica con l'analisi del trascrittoma tramite RNA-seq, che prevede la conversione dell'RNA in cDNA prima del sequenziamento (*Kandpal et al., 2012*). Per quanto riguarda l'epigenomica e la regolazione genica, l'NGS torna particolarmente utile, grazie all'impiego della ChIP-seq (immunoprecipitazione della cromatina seguita da sequenziamento), una tecnica che consente di mappare con precisione le interazioni DNA-proteina (*Park, 2009*).

L'NGS rappresenta quindi una piattaforma tecnologica alla base di differenti strategie di analisi genomica. La scelta delle regioni da sequenziare determina l'ampiezza dell'indagine e consente di

distinguere approcci mirati. Nonostante i continui progressi e la crescente accuratezza delle tecnologie NGS, le varianti identificate mediante questo approccio richiedono, secondo le linee guida, una successiva conferma mediante una seconda metodica, generalmente tramite sequenziamento Sanger.

WHOLE EXOME SEQUENCING (WES)

Il Whole Exome Sequencing (WES) è una strategia di sequenziamento, basata sulle tecnologie NGS che consente di analizzare principalmente le regioni codificanti del genoma, rappresentate dagli esoni. Sebbene l'esoma costituisca solo l'1-2% del genoma umano, le regioni codificanti contengono la quota principale delle varianti note responsabili di malattie mendeliane.

Il WES rappresenta quindi un importante strumento per l'identificazione di varianti associate a patologie genetiche, soprattutto quando il fenotipo non permette di individuare a priori un singolo gene candidato.

Dal punto di vista tecnico, il WES segue il workflow generale delle tecnologie NGS, ma presenta una fase caratteristica di arricchimento delle regioni esoniche, necessaria per selezionare le sequenze di interesse rispetto al resto del genoma (*Bamshad et al., 2011*)

Il WES è particolarmente efficace nell'identificazione di varianti a singolo nucleotide (SNV) e di piccole inserzioni e delezioni (indel) localizzate nelle regioni esoniche coperte. I dati ottenuti possono inoltre essere analizzati mediante specifici approcci bioinformatici per la ricerca di altre tipologie di varianti genomiche (*Bao et al., 2024*).

Uno dei principali vantaggi del WES deriva quindi dalla possibilità di concentrare il sequenziamento su una porzione relativamente ridotta del genoma, ottenendo un'elevata quantità di informazioni sulle regioni codificanti, con un volume di dati inferiore rispetto al sequenziamento dell'intero genoma. Questo comporta generalmente costi e spazi di archiviazione inferiori rispetto al WGS e permette di raggiungere un'elevata profondità di sequenziamento delle regioni target.

Il WES presenta tuttavia alcune importanti limitazioni. L'efficienza della cattura non è uniforme per tutte le regioni, e alcuni esoni possono essere scarsamente rappresentati a causa di caratteristiche della sequenza, come un elevato contenuto in GC, la presenza di sequenze ripetute o un'elevata omologia con altre regioni genomiche. La copertura dell'esoma può quindi risultare insufficiente, e alcune

regioni possono essere rappresentate da un numero basso di reads, riducendo la capacità di identificare eventuali varianti presenti in tali regioni (*Belkadi et al., 2015*).

Un ulteriore limite deriva dalla selezioni delle regioni target. Il WES è rivolto alle sequenze esoniche e alle regioni immediatamente adiacenti incluse nel disegno di cattura, quindi gran parte delle regioni introniche, intergeniche e regolatorie rimangono escluse dall'analisi. Inoltre, il WES presenta una capacità più limitata nell'identificazione di alcune classi di varianti, in particolare determinate varianti strutturali e riarrangiamenti genomici complessi. Pertanto, un risultato negativo al WES non comporta necessariamente l'assenza di una variante causativa della patologia (*Burdick et al., 2020*).

La necessità di superare alcuni di questi limiti e di ottenere una rappresentazione più completa della variabilità genetica ha favorito il crescente utilizzo del WGS (*Belkadi et al., 2015*).

WHOLE GENOME SEQUENCING (WGS)

Il Whole Genome Sequencing (WGS) è una tecnica di NGS che consente di analizzare in maniera estesa l'intero genoma, comprese regioni codificanti e non codificanti come regioni introniche, intergeniche e regolatorie. A differenza del WES è quindi un'analisi più completa e permette di ottenere una rappresentazione più ampia della variabilità genetica di un paziente.

Dopo l'estrazione del DNA genomico, questo viene frammentato e utilizzato per la preparazione della libreria, che viene successivamente sottoposta a sequenziamento. Le reads ottenute vengono ancora una volta sottoposte a controllo qualità e allineate a un genoma di riferimento, consentendo l'identificazione delle varianti tramite il variant calling. Ma l'assenza della fase di cattura, presente invece nel WES, riduce alcune delle distorsioni causate dall'arricchimento e permette di ottenere una copertura migliore (*Belkadi et al., 2015*).

L'estensione dell'analisi consente al WGS di identificare non solo varianti a singolo nucleotide (SNV), piccole inserzioni e delezioni (indel), ma anche copy number variants (CNV) e varianti strutturali (SV) come delezioni, duplicazioni, inversioni e traslocazioni. La presenza di sequenze delle regioni non codificanti permette anche di individuare varianti *deep intronic* e varianti localizzate in regioni regolatorie, che non vengono analizzate nel WES (*Austin-Tse et al., 2022*).

Il WGS risulta particolarmente utile nei casi in cui precedenti analisi mediante pannelli multigenici o WES non hanno consentito di raggiungere una diagnosi molecolare.

La maggiore quantità di informazioni ottenuta mediante WGS ha tuttavia delle limitazioni. Il sequenziamento dell'intero genoma genera un volume di dati considerevolmente superiore rispetto al WES, aumentando la difficoltà nel contesto di archiviazione, elaborazione dei dati e analisi bioinformatica. Inoltre, l'identificazione di un numero elevato di varianti rende più complesso il processo di filtraggio e interpretazione, soprattutto per le alterazioni localizzate nelle regioni non codificanti, per le quali il significato funzionale può essere meno evidente.

Inoltre, il WGS non garantisce l'identificazione di tutte le possibili alterazioni genetiche. Alcune regioni genomiche caratterizzate da elevata ripetitività, elevata omologia o particolare complessità strutturale possono essere difficili da analizzare, soprattutto mediante tecnologie basate sulle short reads. Analogamente, la capacità di rilevare determinate espansioni di sequenze ripetute, varianti strutturali complesse o alterazioni presenti a bassi livelli, come nei casi di mosaicismo, può dipendere dalla piattaforma utilizzata, dalla profondità di sequenziamento e dagli algoritmi bioinformatici (*Austin-Tse et al., 2022*). Una parte di queste limitazioni viene oggi superata dall'avvento del sequenziamento genomico con le long reads.

Nel complesso, il WGS rappresenta un approccio più completo rispetto al WES per lo studio della variabilità genomica, poiché amplia l'analisi delle regioni non codificanti e consente di investigare un insieme più ampio di tipologie di varianti.

CRITERI ACMG/AMP PER LA CLASSIFICAZIONE DELLE VARIANTI NELL'HHT

L'identificazione di una variante mediante le tecniche di sequenziamento è solo la prima fase del processo diagnostico, in quanto successivamente è necessario determinare il possibile significato biologico della variante stessa.

A questo scopo, l'American College of Medical Genetics and Genomics (ACMG) e l'Association of Molecular Pathology (AMP) hanno elaborato delle linee guida che forniscono un sistema standardizzato per l'interpretazione delle varianti al fine di poterle classificare in cinque categorie (*Richards et al., 2015*):

- Benigne (B)
- Probabilmente benigne (LB)
- Di significato incerto (VUS)
- Probabilmente patogenetiche (LP)

- Patogenetiche (P).

Questi criteri devono essere applicati considerando le caratteristiche specifiche del gene e della patologia analizzata. La disponibilità di informazioni relative al meccanismo patogenetico, alla frequenza delle varianti nella popolazione, alle caratteristiche cliniche e alla funzione della proteina possono modificare il peso attribuito alla singola variante, e permettono di aumentare l'affidabilità e migliorare l'interpretazione delle varianti. Per questo sono stati creati gruppi di esperti per singole patologie ai quali viene richiesto di rivedere i criteri a seconda delle caratteristiche proprie della patologia.

In questo contesto, il ClinGen Hereditary Hemorrhagic Teleangiectasia Variant Curation Expert Panel, che annovera fra i suoi membri collaboratori ed ex-collaboratori del laboratorio presso il quale questa tesi è stata condotta, ha rivalutato i criteri ACMG/AMP nel contesto delle varianti dei due principali geni associati all'HHT, *ENG* e *ACVRL1*, sfruttando le conoscenze disponibili per questi geni e per la patologia.

La classificazione integra diverse categorie, come dati di popolazione, informazioni cliniche e fenotipiche, studi di segregazione familiare, caratteristiche molecolari della variante, predizioni virtuali e studi funzionali. Queste regole elaborate sono state successivamente applicate a un gruppo di varianti di *ENG* e *ACVRL1* per verificarne la corretta applicabilità.

La stesura di queste regole ha rappresentato quindi un passo verso una classificazione più standardizzata e specifica delle varianti associate all'HHT, riducendo la possibilità di interpretazioni discordanti e facilitando la distinzione tra varianti clinicamente rilevanti e varianti per le quali non si hanno ancora dati sufficienti. Una corretta classificazione risulta particolarmente importante nel contesto dell'HHT, poiché l'identificazione di una nuova variante patogenetica o probabilmente patogenetica può contribuire alla conferma della diagnosi molecolare e consentire ulteriori studi funzionali (*DeMille et al., 2024*)

SCOPO DEL LAVORO

La Teleangectasia Emorragica Ereditario (HHT) è una patologia vascolare ereditaria caratterizzata da un'elevata variabilità clinica e genetica. I principali geni coinvolti sono *ENG* e *ACVRL1*, responsabili rispettivamente delle forme HHT1 e HHT2, e codificanti per proteine coinvolte nella regolazione dell'omeostasi vascolare e dell'angiogenesi. L'identificazione e la corretta interpretazione delle varianti presenti in questi geni rappresentano quindi un passo fondamentale nell'ambito della diagnosi molecolare dell'HHT.

Questo lavoro di tesi, svolto presso il Laboratorio di Ricerca Molecolare HHT del Dipartimento di Medicina Molecolare dell'Università di Pavia, ha avuto come obiettivo principale la conferma e caratterizzazione di varianti associate all'HHT, ottenute da sequenziamento WGS, attraverso l'analisi del DNA genomico e, per una specifica variante, lo studio dell'RNA, poi retrotrascritto in cDNA. Parallelamente, ci si è posti l'obiettivo di individuare potenziali geni endogeni di riferimento con livelli di espressione adeguati e riproducibili nel sangue periferico, da validare e utilizzare poi come riferimento nelle analisi di espressione dei geni *ENG* e *ACVRL1* mediante Real-Time PCR.

A tale scopo, la prima parte del lavoro è stata dedicata all'analisi del DNA di sei pazienti, ciascuno portatore di una variante in *ENG* o in *ACVRL1*. Le regioni genomiche di interesse sono state amplificate mediante PCR e i prodotti ottenuti sono stati successivamente sottoposti a sequenziamento Sanger, con lo scopo di verificare la presenza delle varianti. I risultati ottenuti sono stati poi analizzati attraverso database e strumenti bioinformatici, per ottenere una classificazione e un'informazione sul possibile impatto sullo splicing. Particolare attenzione è stata rivolta alle varianti in cui le predizioni bioinformatiche suggerivano una possibile alterazione della maturazione del trascritto.

Da queste osservazioni si è sviluppata la seconda parte del lavoro, incentrata sull'analisi di una delle varianti analizzate, identificata precedentemente in altre famiglie, per la quale per la prima volta si rendeva disponibile un campione di RNA. Questo ha permesso di estendere l'analisi dal DNA genomico al trascritto, mediante la retrotrascrizione dell'RNA in cDNA e la successiva amplificazione di specifiche regioni di *ACVRL1*. Lo scopo di questa parte è stato quindi quello di approfondire il possibile effetto della variante sullo splicing, verificando direttamente a livello del trascritto le alterazioni suggerite dalle predizioni bioinformatiche.

La terza linea sperimentale è stata dedicata all'individuazione di geni endogeni di riferimento da utilizzare nelle future analisi di espressione di *ENG* e *ACVRL1* mediante Real-Time PCR. A tale scopo è stata utilizzata la TaqMan™ Array Human Endogenous Control Plate, che ha permesso di analizzare il cDNA ottenuto mediante retrotrascrizione dell'RNA, estratto da sangue periferico di un soggetto sano usato come controllo. Questa analisi ha avuto lo scopo di effettuare una prima selezione di geni con livelli di espressione più vicini a quelli dei geni target, e caratterizzati da una buona riproducibilità tra i triplicati, così da individuare potenziali geni endogeni da sottoporre a successive analisi di validazione.

MATERIALI E METODI

CAMPIONI BIOLOGICI

Gli studi sono stati condotti utilizzando campioni di sangue periferico, raccolti in specifiche provette, in base al protocollo di estrazione utilizzato successivamente.

Per l'estrazione del DNA genomico, 3-5 mL di sangue sono stati raccolti in provette contenenti EDTA (acido etilendiamminotetraacetico), che agisce come anticoagulante attraverso la chelazione degli ioni calcio necessari per il processo di coagulazione, consentendo quindi di mantenere il campione in condizioni idonee per la successiva estrazione del DNA.

Per l'estrazione dell'RNA totale, il sangue periferico è stato invece raccolto in provette con Na Citrato (5-10 mL) per il protocollo che prevedeva l'isolamento dei leucociti mediante Histopaque, oppure in provette Tempus^{TM} Blood RNA (2-4 mL), contenenti uno specifico reagente utile per la stabilizzazione dell'RNA nei campioni. Quest'ultimo tipo di prelievo consente una rapida lisi delle cellule del sangue e inattiva le RNAsi endogene limitando i processi di degradazione e consentendo di preservare l'integrità dell'acido nucleico e il profilo di espressione genica presente al momento del prelievo. La stabilizzazione immediata dei campioni permette inoltre di ridurre eventuali variazioni associate al trasporto e alla conservazione prima delle successive fasi di analisi.

Tutti i prelievi sono stati effettuati in seguito alla firma, da parte dei soggetti coinvolti, di un consenso informato, nell'ambito di uno studio approvato da un Comitato Etico. Le successive analisi sono state svolte presso il Laboratorio di Ricerca Molecolare HHT del Dipartimento di Medicina Molecolare dell'Università di Pavia.

ESTRAZIONE DEL DNA

Una prima tecnica per l'estrazione del DNA genomico si è basata su un protocollo del tipo "salting out", che prevede la lisi cellulare, digestione delle proteine e successiva precipitazione del DNA.

In una prima fase, il sangue è stato sottoposto a ripetuti lavaggi con tampone TE 20:5 freddo, costituito da Tris ed EDTA in rapporto 20mM:5mM rispettivamente. Il Tris è un tampone che contribuisce a mantenere stabile il pH della soluzione e che risulta ipotonico per i globuli rossi causandone la lisi, mentre l'EDTA, con la sua capacità chelante per gli ioni positivi bivalenti, contribuisce a proteggere il DNA dall'azione delle nucleasi (che lavorano grazie a ioni Mg^{2+}). I

lavaggi e le successive centrifugazioni consentono di eliminare progressivamente le componenti non necessarie e di ottenere un pellet cellulare composto da cellule nucleate (leucociti).

Il pellet è stato poi trattato con Sarcosyl, un detergente che favorisce la lisi delle membrane e la solubilizzazione delle componenti cellulari, fino a una concentrazione finale di 1%, e proteinasi K, un enzima che determina la digestione delle proteine, comprese quelle associate al DNA, favorendone la purificazione, fino a una concentrazione finale di 1mg/mL. La digestione è stata condotta a 37°C over-night in stufa.
Successivamente è stato aggiunto NaCl 6,5 M, fino a una concentrazione finale di 57 g/L , che serve per separare le proteine legate al DNA e a neutralizzare le cariche negative della molecola, seguito da isopropanolo assoluto per causare la precipitazione del DNA. In queste condizioni il DNA diventa insolubile e può essere recuperato sotto forma di aggregato filamentoso ("medusa").

Il DNA precipitato viene poi recuperato e sottoposto a lavaggio con etanolo al 70%, allo scopo di rimuovere sali e altre impurità. Dopo l'asciugatura il DNA viene risospeso in tampone TE 20:1, costituito da Tris e EDTA in rapporto 20:1 (20 mM di Tris e 1 mM EDTA), che mantiene stabile il pH e protegge il DNA dalla degradazione, consentendone la conservazione.

Come seconda tecnica, il DNA genomico è stato estratto da sangue periferico mediante il kit NucleoSpin® Blood, basato sulla cromatografia su colonna con membrana di silice in grado di legare selettivamente il DNA.

Il campione viene sottoposto a lisi mediante l'aggiunta di proteinasi K e di un Buffer (B3) che permette la disgregazione delle cellule e il rilascio del DNA. Successivamente il campione viene incubato per 15 minuti a 70°C per favorire il completamento della lisi.

Al lisato è stato successivamente aggiunto etanolo per favorire il legame del DNA alla membrana di silice. Il campione viene poi trasferito nella NucleoSpin Blood Column e centrifugato per trattenere il DNA nella membrana, mentre altre componenti del campione attraversano la colonna e vengono eliminate.

La membrana viene poi sottoposta a successivi lavaggi utili per la rimozione delle proteine, sali e altri contaminanti. Una successiva centrifugazione permette inoltre di asciugare la membrana e rimuovere l'etanolo residuo.

Infine, il DNA viene recuperato mediante eluizione con Buffer BE. Il tampone applicato direttamente sulla membrana, modifica le condizioni che ne favoriscono il legame alla silice, consentendo al DNA di essere rilasciato e raccolto mediante centrifugazione.

ESTRAZIONE DELL'RNA

L'RNA totale, del paziente e del controllo utilizzato per la q-PCR, è stato estratto a partire da sangue periferico mediante un protocollo che prevede una prima fase di isolamento delle cellule mononucleate mediante Histopaque®-1077, seguita dall'estrazione dell'RNA mediante NucleoZOL.

Inizialmente, il sangue periferico è stato stratificato su Histopaque®-1077, un mezzo a densità definita che consente di separare le diverse popolazioni cellulari del sangue mediante centrifugazione su gradiente di densità. Durante la centrifugazione, gli eritrociti e i granulociti, che possiedono maggiore densità, sedimentano verso il fondo della provetta, mentre le cellule mononucleate, principalmente linfociti e monociti si dispongono tra il plasma e l'Histopaque. Lo strato contenente queste cellule può essere quindi recuperato e sottoposto a lavaggi per eliminare residui di plasma, Histopaque e altre componenti cellulari, ottenendo un pellet arricchito in cellule mononucleate.

Le cellule isolate sono state poi lisate mediante NucleoZOL, un reagente che determina la rottura delle membrane cellulari e consente il rilascio degli acidi nucleici, creando al tempo stesso condizioni idonee alla preservazione dell'RNA. Successivamente viene aggiunta acqua RNAse-free e alla centrifugazione, il DNA, proteine e altri contaminanti precipitano formando un pellet, mentre l'RNA rimane solubilizzato nel surnatante. Quest'ultimo può essere recuperato per poi proseguire l'estrazione.

L'RNA presente nel surnatante viene precipitato mediante isopropanolo, che ne riduce la solubilità e ne permette il recupero sotto forma di pellet dopo la centrifugazione. Il pellet di RNA è stato poi sottoposto a lavaggi con etanolo al 75%, necessario per rimuovere sali e contaminanti.
Infine, l'RNA viene risospeso in acqua RNAse-free, in modo da evitare l'introduzione di RNAsi che potrebbero determinare la degradazione, ottenendo così una soluzione di RNA totale utilizzabile per le successive analisi molecolari.

La seconda tecnica di estrazione dell'RNA viene svolta a partire da sangue periferico raccolto direttamente in provette Tempus, in modo da stabilizzare l'RNA già al momento del prelievo.

Al campione di sangue viene poi aggiunto PBS privo di Ca2+ e Mg2+, viene miscelato e poi centrifugato a 4° per permettere la formazione di un pellet contenente gli acidi nucleici. Il pellet viene poi risospeso con l'RNA Purification Resospension Solution, e il tutto viene poi trasferito su una provetta contenente una colonnina cromatografica con filtro, precedentemente reidratato con la Wash Solution 1. Il campione viene poi centrifugato per permettere all'RNA di essere trattenuto dal filtro.

La membrana del filtro è stata poi sottoposta a due lavaggi con la Wash Solution 1 e due lavaggi con la Wash Solution 2, finalizzati alla rimozione delle componenti cellulari, dei sali e degli altri contaminanti presenti nel campione, mantenendo l'RNA associato alla membrana.

Al termine dei lavaggi viene fatta un'ulteriore centrifugazione per asciugare la membrana ed eliminare le soluzioni di lavaggio. Infine, l'RNA viene recuperato grazie all'aggiunta della Nucleic Acid Purification Elution Solution, che consente il rilascio dell'RNA dalla membrana e il suo recupero mediante centrifugazione, e l'eluato è stato fatto passare nuovamente attraverso il filtro per favorire il recupero dell'RNA, è stato poi trasferito in una nuova provetta e conservato a -80°C.

DOSAGGIO SPETTROFOTOMETRICO DEGLI ACIDI NUCLEICI

La concentrazione e la purezza del DNA e dell'RNA estratti sono state valutate mediante spettrofotometro NanoDrop 1000 (ND-1000). Prima della misurazione, lo strumento viene calibrato usando come bianco la stessa soluzione in cui è sospeso o eluito il campione. Successivamente 1 µL di campione viene posta sul piedistallo e viene misurata l'assorbanza a 260nm, lunghezza d'onda alla quale gli acidi nucleici presentano il massimo assorbimento.

La concentrazione degli acidi nucleici è stata determinata sulla base dell'assorbanza misurata a 260nm (A260), applicando uno specifico fattore di conversione. Per il DNA a doppio filamento, un valore di 260 pari a 1 corrisponde a una concentrazione di 50 ng/µL, mentre per l'RNA corrisponde a 40 ng/µL. La purezza dei campioni è stata invece valutata mediante i rapporti di assorbanza A268/280 e A260/230. Il primo fornisce un'indicazione dell'eventuale presenza di contaminazione proteica, mentre il secondo permette di evidenziare la possibile presenza di contaminanti organici, Sali o reagenti residui derivanti dalle procedure di estrazione.

RETROTRASCRIZIONE DELL'RNA IN cDNA

In seguito all'estrazione e alla quantificazione dell'RNA, questo è stato sottoposto a retrotrascrizione mediante High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems™) e Reliance Select cDNA

Synthesis Kit (Bio Rad), al fine di ottenere DNA complementare (cDNA) utilizzabile nelle successive analisi molecolari.

La retrotrascrizione è il processo mediante il quale una trascrittasi inversa utilizza l'RNA come stampo per la sintesi di un filamento di DNA complementare. In questo modo, le molecole di RNA presenti nel campione vengono convertite in una forma di DNA più stabile e adatta alle successive analisi mediante PCR o Real-Time PCR.

Per le retrotrascrizioni svolte con il kit High-Capacity RNA-to-cDNA™, per ciascun campione è stata preparata una miscela di reazione che permette di garantire le condizioni chimiche necessarie per il corretto svolgimento della reazione, e in cui sono presenti le componenti enzimatiche e non, necessarie alla sintesi del cDNA (Tabella 1). Per ogni esperimento di retrotrascrizone è stata anche preparato un No-Template-Control (NTC), con tutti i reagenti, ma in cui la quota di RNA è stata sostituita con H_2O per verificare che non ci sia contaminazione dei reagenti. La miscela di reazione viene preparata mantenendo i reagenti in ghiaccio, e successivamente distribuita nelle provette apposite.

Componenti	**Volume per reazione**	
	RT+	NTC
2x RT Buffer Mix	10 µL	10 µL
20x RT Enzyme Mix	1 µL	1 µL
RNA	Fino a 9 µL	-
H_2O RNAse Free	Fino a volume	9 µL
Volume finale	**20 µL**	**20 µL**

Tabella 1. Composizione e volumi dei reagenti utilizzati per la reazione di retrotrascrizione dell'RNA (RT+) e per il controllo negativo (NTC) con il kit High-Capacity RNA-to-cDNA™.

Una volta preparati i campioni vengono brevemente centrifugati per raccogliere l'intero volume sul fondo delle provette. La reazione viene eseguita in un Termociclatore Gene Amp PCR 9700 (Applied Biosystems™), utilizzando il protocollo termico fornito insieme al kit:

- 37°C per 60 minuti
- 95°C per 5 minuti
- 4°C ∞ (fino al recupero dei campioni).

Per le retrotrascrizioni eseguite con il Reliance Select cDNA Synthesis Kit, ciascuna reazione è stata preparata una miscela di reazione contenente i componenti necessari alla sintesi del cDNA, l'RNA estratto e i primer oligo(dT) inclusi nel kit, fino a un volume finale di 20 µL. Gli oligo(dT) sono brevi sequenze di deossitimidine in grado di appaiarsi alla coda poli-A presente all'estremità 3' degli mRNA. Gli oligo(dT) forniscono quindi l'estremità di innesco necessaria alla trascrittasi inversa per avviare la sintesi del filamento di cDNA (Tabella 2).

Componenti	**Volume per reazione**	
	RT+	NTC
5x Reliance Select cDNA Synthesis Reaction Buffer	4 µL	4 µL
Reliance Reverse Transcriptase	1 µL	1 µL
Oligo(dT) primer Mix	2 µL	2 µL
RNA	Fino a 13 µL	-
H_2O RNAse Free	Fino a volume	13 µL
Volume finale	**20 µL**	**20 µL**

Tabella 2. Composizione e volumi dei reagenti utilizzati per la reazione di retrotrascrizione dell'RNA (RT+) e per il controllo negativo (NTC) con il kit Reliance Select cDNA Synthesis Kit.

Alla fine, le provette sono state centrifugate e sottoposte al protocollo termico previsto dal protocollo:

- 50°C per 90 minuti
- 95°C per 1 minuto
- 4°C ∞ (fino al recupero dei campioni).

La quantità di RNA utilizzata per ciascun campione è stata definita sulla base della concentrazione di RNA precedentemente determinata, in modo da utilizzare la quantità richiesta per la reazione di retrotrascrizione. Al fine di garantire un adeguato controllo delle analisi e rendere confrontabili i risultati ottenuti, i campioni sono stati sottoposti a retrotrascrizione utilizzando la stessa concentrazione di RNA di partenza. Nelle successive reazioni di PCR è stato poi utilizzato un volume di cDNA uguale per ciascun campione.

Al termine della reazione, il cDNA ottenuto può essere utilizzato per le successive analisi oppure conservato a -20°C.

PCR PER IL DNA ESTRATTO O PER IL cDNA RETROTRASCRITTO

Una volta ottenuto il DNA o il cDNA è stata svolta la PCR (Polymerase Chain Reaction), una tecnica che permette di produrre numerose copie di una regione genica attraverso una serie di cicli di amplificazione condotti in condizioni controllate.

Per condurre una PCR, il DNA stampo viene unito agli appositi reagenti, la cui composizione e dosaggio dipende dal volume finale desiderato e dai parametri dell'esperimento (Tabella 3).

Sia per il DNA genomico, sia per il cDNA ottenuto mediante retrotrascrizione, sono state inizialmente svolte delle reazioni di PCR in un volume finale di 10 µL, utilizzate come controllo preliminare dell'avvenuta amplificazione. Una volta verificata la corretta amplificazione mediante PCR di controllo, per il DNA genomico la PCR è stata ripetuta in un volume finale di 30 µL, al fine di ottenere una quantita di prodotto sufficiente per le successive fasi di purificazione e sequenziamento.

Anche per il cDNA è stata successivamente allestita una PCR in un volume finale di 30 µL, utilizzando la PrimeSTAR Max DNA Polymerase (Takara Bio), che consente anche l'amplificazione di frammenti di DNA di lunghezza elevata.

Fa eccezione il campione utilizzato per l'analisi mediante Real-time PCR, per cui è stata allestita solo una PCR di controllo con volume finale di 10 µL, per verificare l'avvenuta retrotrascrizione.

Reagenti PCR	**Volume per reazione PCR di controllo**	**Volume per reazione PCR per sequenziamento**	**Volume per reazione PCR con kit Takara**
H_2O	3,9 µL	11,7 µL	9,4 µL
Buffer 10x	1 µL	3 µL	6 µL
$MgCl_2$	1 µL	3 µL	-
dNTPs mix 200µM	1 µL	3 µL	2,4 µL
Primer Forward 2µM	1 µL	3 µL	3 µL
Primer Reverse 2µM	1 µL	3 µL	3 µL
Taq DNA Polimerasi FIREpol® (5 U/ µL)	0,1 µL	0,3 µL	1,2 µL
DNA o cDNA	1 µL	3 µL	5 µL
Volume finale	**10 µL**	**30 µL**	**30 µL**

Tabella 3. Composizione delle miscele di reazione utilizzate per l'amplificazione mediante PCR, nei volumi di 10 µL per le PCR di controllo e di 30 µL per le PCR destinate al sequenziamento.

Per queste analisi, le coppie di primers utilizzate sono quelle disegnate da tempo nel laboratorio (*Olivieri et al., 2007*), oppure specificatamente disegnate per l'esperimento, usando i software Primer3 e Primer-BLAST. I primer sono stati quindi utilizzati per amplificare la regione comprendente la variante di interesse, ottenendo il prodotto di PCR successivamente sottoposto a sequenziamento Sanger (Tabella 4).

Gene	Primer	Sequenza
ENG	E4_N_F → Forward	5' - GCAAGGGAGAAGGGGCTGAT - 3'
ENG	E4_N_R → Reverse	5' - GGGCACCAAGAGTGCAGGAG - 3'
ENG	E5_N_F → Forward	5' - TGGCAGGTGAGTGGTGGAA - 3'
ENG	E5_N_R → Reverse	5' - GCAGCCCTGTGAGAGCACA - 3'
ENG	E8_N_F → Forward	5' - GGTGCGGGCACACAGTGAT - 3'
ENG	E8_N_R → Reverse	5' - GCCAATCACGTGGCACAGTAA - 3'
ENG	E9_N_F → Forward	5' - GACTTGGGACCCCTGGGTTGT - 3'
ENG	E9_N_R → Reverse	5' - GGAAACTCCTGGCAAGCCATG - 3'
ACVRL1	A7_N_F → Forward	5' - CCCACCCCCAGACCTAGCTT - 3'
ACVRL1	A7_N_R → Reverse	5' - TTCTCCTCCGCAAGACCCC - 3'
ACVRL1	A9_N_F → Forward	5' - ATTGCATTATACTGTCCCTCTCAG - 3'
ACVRL1	A9_N_R → Reverse	5' - TGTCCTGGTTAGGGCACCTC - 3'

Tabella 4. Sequenze dei primer utilizzati per la verifica delle varianti nel DNA genomico.

Per la PCR di analisi del cDNA è stata utilizzata una coppia di primer specifici per il gene *ACVRL1*, già disponibile in laboratorio, in grado di amplificare una regione compresa tra gli esoni 7, 8, 9 e 10 (Tabella 5).

Primer	Sequenza
mAeF → Forward	5' - CCATCTGGCTCTGAGGCT - 3'
mAeR → Reverse	5' - AATGGCATCGTGGAGGACT - 3'
mAfF → Forward	5' - GCACGGACTGCTTTGAGTC - 3'
mAfR → Reverse	5' - CAGGCAGAAAGGAATCAGGT - 3'
mAgR → Reverse	5' - ATACAGCTGGGCTGAAAC- 3'

Tabella 5. Sequenze dei primer utilizzati per la verifica della variante nel cDNA.

Per l'amplificazione dei campioni è stato utilizzato il termociclatore S1000™ Thermal Cycler (Bio-Rad Laboratories), e le reazioni di PCR sono state eseguite utilizzando il medesimo programma termico per tutte le coppie di primer, fatta eccezione per le PCR fatte con il kit Takara, variando esclusivamente la temperatura di annealing (3) in funzione della specifica coppia utilizzata (Tabella 6).

1. 95ºC per 5 minuti
2. 95ºC per 40 secondi
3. Variabile in base alla coppia di primer per 40 secondi } 40 cicli
4. 72ºC per 40 secondi
5. 72ºC per 5 minuti
6. 4ºC ∞ (fino al recupero dei campioni)

Per le reazioni di PCR eseguite sul cDNA utilizzando la Prime STAR Max DNA Polymerase (Takara Bio) è stato invece impiegato un diverso programma termico.

1. 98ºC per 10 secondi
2. 60ºC per 15 secondi
3. Variabile in base alla coppia di primer per 40 secondi } 30 cicli
4. 4ºC ∞ (fino al recupero dei campioni)

Coppia di primer	Temperatura di annealing	Amplificato atteso
E4_N_F / E4_N_R	63°C	458 pb
E5_N_F / E5_N_R	65°C	238 pb
E8_N_F / E8_N_R	65°C	418 pb
E9_N_F / E9_N_R	63°C	627 pb
A7_N_F / A7_N_R	65°C	393 pb
A9_N_F / A9_N_R	61°C	407 pb
mAeF / mAeR	64°C	374 pb
mAfF / mAfR	63°C	385 pb
mAeF / mAgR	60°C	797 pb
mAeF / mAfR	60°C	713 pb

Tabella 6. Temperature di annealing utilizzate per le diverse coppie di primer impiegate nelle reazioni di PCR.

ELETTROFORESI SU GEL DI AGAROSIO

Al termine della PCR, la presenza e la dimensione dei prodotti ottenuti sono state verificate mediante elettroforesi su gel di agarosio, che permette di distinguere frammenti di DNA di diversa lunghezza sfruttandone la migrazione attraverso una matrice porosa sottoposta a un campo elettrico.

Durante la corsa, l'applicazione di una differenza di potenziale determina la migrazione del DNA, che, per la presenza dei gruppi fosfato carichi negativamente si sposta verso il polo positivo. La matrice di agarosio agisce come un setaccio molecolare, in cui i frammenti di dimensioni inferiori attraversano più facilmente i pori del gel e percorrono quindi una distanza maggiore rispetto ai frammenti più lunghi.

Per l'analisi è stato utilizzato un gel di agarosio all'1,5%, ottenuto aggiungendo l'agarosio al tampone Tris-Borato-EDTA (TBE 1x) e riscaldando la soluzione fino al completo scioglimento. Dopo aver lasciato raffreddare parzialmente la soluzione, è stato aggiunto GelRed^{TM}, intercalante fluorescente che si lega agli acidi nucleici e ne consente la successiva rilevazione. Il gel è stato poi colato nell'apposito supporto e lasciato solidificare utilizzando un pettine per formare i pozzetti per il caricamento dei campioni.

Dopo la solidificazione, il gel è stato trasferito nella camera elettroforetica contenente TBE 1x come tampone di corsa. Prima del caricamento, ai prodotti di PCR è stato aggiunto Orange G, utilizzato come buffer di caricamento e come indicatore dell'avanzamento della corsa elettroforetica. La dimensione delle bande è stata valutata mediante confronto con un marcatore di peso molecolare (DNA Ladder), costituito da DNA di lunghezza nota.

Terminata la separazione elettroforetica, i prodotti sono stati visualizzati mediante il sistema di acquisizione ChemiDoc® (Bio-Rad).

PURIFICA DELLE PCR E PREPARAZIONE DEI CAMPIONI PER IL SEQUENZIAMENTO SANGER

I prodotti ottenuti dall'amplificazione del DNA genomico sono stati purificati mediante il kit NucleoSpin® Gel and PCR Clean-up (MACHEREY-NAGEL) prima di essere sottoposti a sequenziamento Sanger. La purificazione permette di recuperare il DNA amplificato eliminando i componenti residui della reazione di PCR che potrebbero interferire con le successive analisi.

Al prodotto di PCR è stato inizialmente aggiunto il Buffer NTI, utilizzato per creare le condizioni necessarie affinché il DNA possa legarsi alla membrana di silice presente nella colonnina NucleoSpin®. La miscela è stata quindi trasferita sulla colonnina e centrifugata, per fare in modo che il DNA venga trattenuto dalla membrana, mentre parte delle sostanze presenti nella miscela attraversa la colonna e viene eliminata.

La membrana è stata successivamente sottoposta a due lavaggi con Buffer NT3, allo scopo di rimuovere i contaminanti ancora presenti tramite una successiva centrifugazione.
Terminati i lavaggi, è stata effettuata un'ulteriore centrifugazione senza l'aggiunta di alcun prodotto per asciugare completamente la membrana ed eliminare eventuali residui del buffer di lavaggio.

Infine, la colonnina è stata trasferita in una nuova provetta e il DNA purificato è stato recuperato mediante eluizione con 15 µL di Buffer NE, pre-riscaldato a 70°C, aggiunto direttamente sulla membrana, seguito da un'incubazione di 5 minuti a 70°C, seguita da una centrifugazione di 1 minuto. In queste condizioni il DNA legato alla membrana viene rilasciato e raccolto nella provetta sottostante, ottenendo così il prodotto di PCR purificato e utilizzabile per il successivo Sequenziamento Sanger.

INVIO CAMPIONI PER SEQUENZIAMENTO

Al termine della purificazione, i prodotti di PCR sono stati preparati per il successivo sequenziamento Sanger.
Per ciascun campione sono stati prelevati 5 µL dell'eluato ottenuto dalla purificazione, poi uniti a 5 µL delle aliquote dei primer Forward (F) e Reverse (R) utilizzati per l'amplificazione della regione di interesse, ciascuno a una concentrazione di 5 µM, seguendo quanto richiesto dal service di sequenziamento.

Per ogni prodotto di PCR purificato sono state svolte due sequenze, una con il primer Forward e l'altra con il primer Reverse, per permettere l'analisi della sequenza nelle due direzioni. Ogni provetta è stata poi etichettata per consentirne l'identificazione, e successivamente spedita al sequenziamento presso la EUROFINS GENOMICS.

ANALISI E INTERPRETAZIONE DELLE VARIANTI

Le varianti identificate grazie al sequenziamento sono state analizzate mediante diversi database e strumenti bioinformatici, utili per valutarne il significato clinico e il possibile impatto funzionale.

Un primo sito è VarSome, una piattaforma che integra informazioni provenienti da diversi database e strumenti di predizione, consentendo di ottenere una visione complessiva della variante e della sua classificazione clinica. Un'altra piattaforma utile è ClinVar, un database pubblico che raccoglie le interpretazioni cliniche delle varianti genetiche e permette di verificare se una variante è già stata descritta in letteratura, e il significato clinico che le è stato attribuito. È stato anche consultato gnomAD (Genome Aggregation Database), che raccoglie dati di sequenziamento provenienti da ampie popolazioni e consente di verificare la presenza e la frequenza delle varianti nella popolazione generale.

Per valutare il possibile effetto delle varianti sullo splicing sono stati usati SpliceAI, SpliceAid e Human Splicing Finder (HSF). SpliceAI permette di predire se una variante possa determinare la perdita di un sito di splicing esistente o la formazione di un nuovo sito donatore o accettore. SpliceAid fornisce informazioni sulle sequenze coinvolte nella regolazione dello splicing e sulla loro interazione con proteine regolatrici. Human Splicing Finder consente infine di valutare il possibile effetto delle varianti sui siti di splicing e sugli elementi regolatori coinvolti nel riconoscimento degli esoni.

REAL-TIME PCR

La Real-Time PCR (qPCR) è una tecnica che consente di monitorare l'amplificazione di una specifica sequenza di DNA in tempo reale, attraverso la rilevazione di un segnale fluorescente che aumenta parallelamente all'accumulo del prodotto di amplificazione.

In questo studio, la Real-Time PCR è stata eseguita a partire dal cDNA precedentemente ottenuto mediante retrotrascrizione dell'RNA estratto da sangue periferico, con lo scopo di valutare l'espressione di una serie di geni endogeni e individuare quelli caratterizzati da un'espressione sufficientemente stabile nei campioni analizzati. La scelta di un adeguato controllo endogeno è infatti necessaria per la successiva normalizzazione dei dati di espressione genica.

L'analisi è stata effettuata utilizzando la TaqMan™ Array Human Endogenous Control Plate, 96-well (Applied Biosystems™), una piastra da 96 pozzetti contenenti 32 geni umani candidati come controlli endogeni, ciascuno presente in triplicato (Figura 6). In ogni pozzetto sono già presenti in forma essiccata il primer e la sonda TaqMan™, necessari per l'amplificazione e la rilevazione del rispettivo target.

Gene	1	2	3	4	5	6	7	8	9	10	11	12
A	18S	GAPDH	HPRT1	GUSB	18S	GAPDH	HPRT1	GUSB	18S	GAPDH	HPRT1	GUSB
B	ACTB	B2M	HMBS	IPO8	ACTB	B2M	HMBS	IPO8	ACTB	B2M	HMBS	IPO8
C	PGK1	RPLP0	TBP	TFRC	PGK1	RPLP0	TBP	TFRC	PGK1	RPLP0	TBP	TFRC
D	UBC	YWHAZ	PPIA	POLR2A	UBC	YWHAZ	PPIA	POLR2A	UBC	YWHAZ	PPIA	POLR2A
E	CASC3	CDKN1A	CDKN1B	GADD45 A	CASC3	CDKN1A	CDKN1B	GADD45 A	CASC3	CDKN1A	CDKN1B	GADD45 A
F	PUM1	PSMC4	EIF2B1	PES1	PUM1	PSMC4	EIF2B1	PES1	PUM1	PSMC4	EIF2B1	PES1
G	ABL1	ELF1	MT-ATP6	MRPL19	ABL1	ELF1	MT-ATP6	MRPL19	ABL1	ELF1	MT-ATP6	MRPL19
H	POP4	RPL37A	RPL30	RPS17	POP4	RPL37A	RPL30	RPS17	POP4	RPL37A	RPL30	RPS17

Office on the web Frame

Figura 6. Organizzazione dei geni TaqMan™ Array Human Endogenous Control Plate.

Il sistema TaqMan™ utilizzato in questa piastra si basa sull'impiego di una sonda fluorescente specifica per la sequenza bersaglio, oltre alla coppia di primer necessaria per l'amplificazione. Durante la reazione, la sonda si lega alla sequenza target compresa nella regione amplificata. Con il procedere dell'amplificazione viene generato un segnale fluorescente, la cui intensità aumenta in relazione alla quantità di prodotto che si accumula. La fluorescenza viene rilevata dallo strumento a ogni ciclo, consentendo di seguire l'amplificazione in tempo reale.

Per ciascun campione è stata quindi preparata una miscela contenente l'H_2O Nuclease Free, il cDNA e la TaqMan™ Gene Expression Master Mix necessarie alla reazione. La soluzione è stata distribuita nei pozzetti della piastra, permettendo la risospensione degli assay TaqMan™ (Tabella 7).

Componenti	Volume per reazione
TaqMan™ Gene Expression Master Mix	10 µL
cDNA	2,2 µL
H2O Nuclease Free	7,8 µL
Volume finale	**20 µL**

Tabella 7. Composizione e volumi dei reagenti utilizzati per l'allestimento della reazione di Real-Time PCR.

La piastra è stata successivamente sigillata e inserita nel Bio-Rad CFX Connect Real-Time PCR Detection System per l'esecuzione del programma di amplificazione e la rilevazione della fluorescenza. Il seguente programma di amplificazione ci è stato fornito dal produttore del kit:

- Hold, 95°C per 10 minuti
- Denaturazione, 95°C per 15 secondi
- Estensione, 60°C per 1 minuto

} 40 cicli

Al termine dell'analisi, per ciascun gene viene determinato il valore di Cycle Threshold (Ct), corrispondente al ciclo nel quale il segnale fluorescente supera una determinata soglia. Il confronto dei valori ottenuti per i diversi geni endogeni consente quindi di valutarne il comportamento nei campioni analizzati e di identificare quelli più adatti a essere utilizzati come riferimento nelle successive analisi di espressione.

RISULTATI

CAMPIONI BIOLOGICI, ESTRAZIONE E DOSAGGIO DEGLI ACIDI NUCLEICI

Per lo studio sono stati analizzati campioni di sangue periferico provenienti da sei pazienti (A-F), dai quali è stato estratto il DNA genomico, utilizzato per l'analisi delle varianti. Per un ulteriore paziente (G) è stato invece estratto l'RNA, al fine di approfondire il possibile effetto della variante identificata a livello del trascritto. È stato inoltre estratto l'RNA da sangue periferico, da tre diversi prelievi, proveniente da un soggetto sano (Ctrl_1- Ctrl_3), utilizzato come controllo per l'analisi mediante Real-Time PCR.

In seguito all'estrazione, i campioni di DNA e RNA sono stati sottoposti a dosaggio spettrofotometrico mediante NanoDrop, al fine di determinarne la concentrazione e valutarne la purezza (Tabella 8).

Campione	Acido Nucleico	ng/µL	260/280	260/230
A	DNA	93,5	1,78	1,22
B	DNA	16,8	1,73	1,82
C	DNA	15,8	1,7	1,63
D	DNA	12,6	1,8	1,95
E	DNA	8,2	1,9	1,1
F	DNA	19,3	1,64	1,47
G	RNA	60,4	2,2	1,09
Ctrl_1	RNA	99,9	2,12	1,41
Ctrl_2	RNA	118,75	2,15	1,46
Ctrl_3	RNA	101,05	2,09	1,40

Tabella 8. Valori medi di concentrazione e purezza dei campioni estratti.

ANALISI MEDIANTE PCR E CORSA ELETTROFORETICA

Per l'analisi delle varianti a livello del DNA genomico, le regioni di interesse cono state amplificate mediante PCR utilizzando coppie di primer specifiche, selezionate in funzione della variante da analizzare. Nella tabella di seguito sono riportati, per ciascun paziente, il gene, la variante oggetto di studio e la relativa coppia di primer utilizzata per l'amplificazione (Tabella 9).

Campione	Gene	Variante	Coppia di primer
A	*ENG*	c. 626 delT	E5_N_F / E5_N_R
B	*ACVRL1*	c. 696 delA	A7_N_F / A7_N_R
C	*ENG*	c. 392 delC	E4_N_F / E4_N_R
D	*ENG*	c. 1103 T>C	E8_N_F / E8_N_R
E	*ACVRL1*	c. 1377+4 A>G	A9_N_F / A9_N_R
F	*ENG*	c.1311+5 G>A	E9_N_F / E9_N_R

Tabella 9. Varianti analizzate e relative coppie di primer utilizzate per ciascun campione.

Per l'analisi del campione G e dei controlli (Ctrl), l'RNA è stato estratto, retrotrascritto in cDNA, e quest'ultimo è stato utilizzato per effettuare una PCR di controllo per verificare, tramite l'amplificazione, l'effettiva riuscita della retrotrascrizione.

Per il campione G, portatore della variante *ACVRL1* c. 1377+4 A>G, sono stati utilizzati primer specifici in grado di amplificare diverse regioni del trascritto allo scopo di studiare gli effetti dovuti alla presenza della variante, così da consentirne la successiva analisi (Tabella 10 e Figura 7).

Campione	Gene	Variante	Coppia di primer
G	*ACVRL1*	c. 1377+4 A>G	mAfF / mAfR
			mAeF / mAgR
			mAeF / mAfR

Tabella 10. Variante analizzata e relative coppie di primer utilizzate.

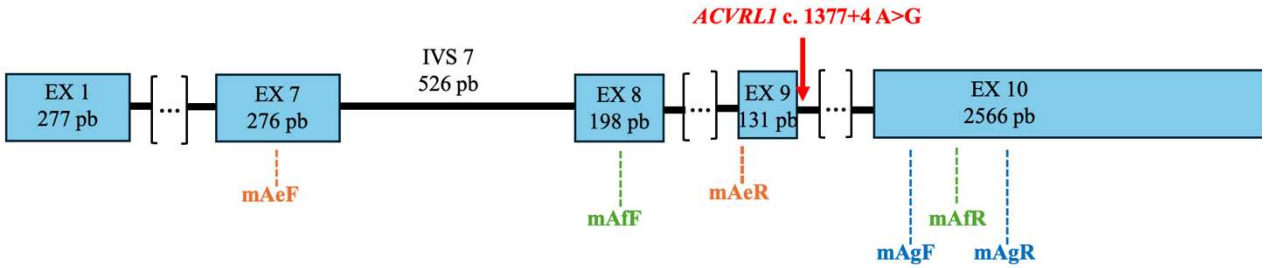
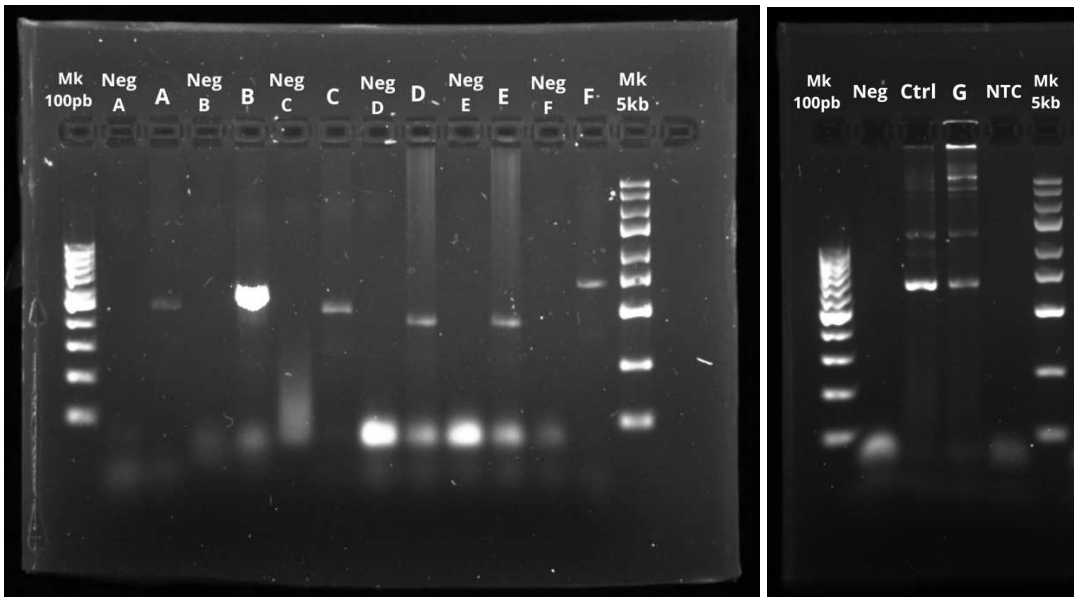


Figura 7. Rappresentazione schematica, in scala, della regione del gene *ACVRL1* analizzata e delle coppie di primer impiegate nello studio.

I prodotti di PCR ottenuti sono stati successivamente analizzati mediante elettroforesi su gel di agarosio, al fine di studiare la presenza dei prodotti di amplificazione attesi.



A B

Figura 8. Corsa elettroforetica su gel di agarosio delle PCR di prova dei campioni A, B, C, D, E, F (A) e della PCR del campione G e del controllo con il kit Takara (B). Le colonne indicate come "Neg" contengono il controllo negativo, la colonna "NTC" indica il No Template Control.

L'analisi mediante PCR sul cDNA ha evidenziato, sia nel campione G sia nel controllo, la presenza delle bande di amplificazione attese. Oltre a queste sono state osservate anche bande aggiuntive di circa 1000 bp, successivamente identificate come prodotti di amplificazione aspecifici.

SEQUENZIAMENTO SANGER E ANALISI DELLE VARIANTI

I prodotti di amplificazione ottenuti mediante PCR sono stati purificati e sottoposti a sequenziamento Sanger. L'analisi degli elettroferogrammi ottenuti ha permesso di valutare, nei diversi pazienti, la presenza delle varianti oggetto di studio.

Le varianti identificate sono state successivamente analizzate mediante database e strumenti bioinformatici, al fine di valutarne la classificazione clinica e il possibile effetto sui meccanismi di splicing (Tabella 11). Come riportato, tutte le varianti identificate sono considerate LP/P. l'unica variante classificata come VUS (Hot) è quella che abbiamo quindi analizzato anche a livello dell'RNA.

Campione	Gene	Variante	Risultato Sanger	VarSome/ClinVar	Predizione sullo Splicing
A	*ENG*	c. 626 delT (p.Leu209Trp fs*13)	Variante presente	Likely Pathogenic	Basso impatto predetto sullo splicing
B	*ACVRL1*	c. 696 delA (p.Lys323Thr fs*68)	Variante presente	Likely Pathogenic	New acceptor splice site, activation of cryptic acceptor site
C	*ENG*	c. 392 delC (p.Pro131Arg fs*17)	Variante presente	Pathogenic	New donor splice site, activation of cryptic donor site
D	*ENG*	c. 1103 T>C (p.Met268Thr)	Variante presente	Likely Pathogenic	Nessun effetto sullo splicing predetto
E	*ACVRL1*	c. 1377+4 A>G	Variante presente	Uncertain Significance (VUS hot)	Broken WT donor site
F	*ENG*	c.1311+5 G>A	Variante presente	Likely Pathogenic	Donor shift, partial intron 10 retention

Tabella 11. Risultati del sequenziamento Sanger, classificazione delle varianti identificate e predizione del loro possibile effetto sullo splicing.

Nelle immagini di seguito, sono riportati a scopo esemplificativo gli elettroferogrammi relativi a tre dei sei pazienti analizzati, nello specifico A (Figura 9), D (Figura 10), E (Figura 11).

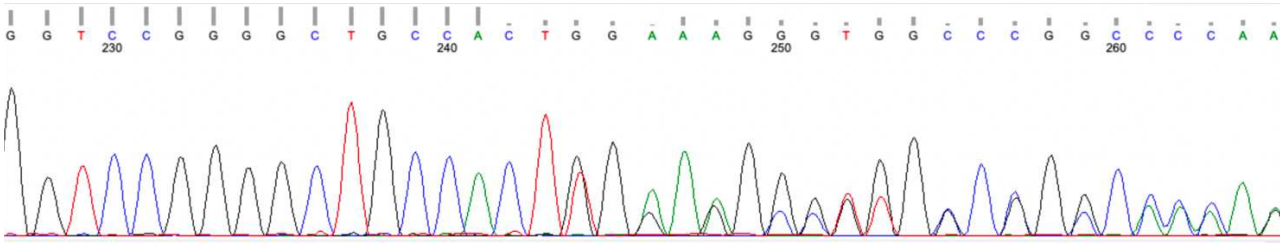


Figura 9. Elettroferogramma del campione A (*ENG* c. 626 delT (p.Leu209Trp fs*13)), in cui si può vedere l'inizio del frameshift a livello del nucleotide 244 della sequenza.

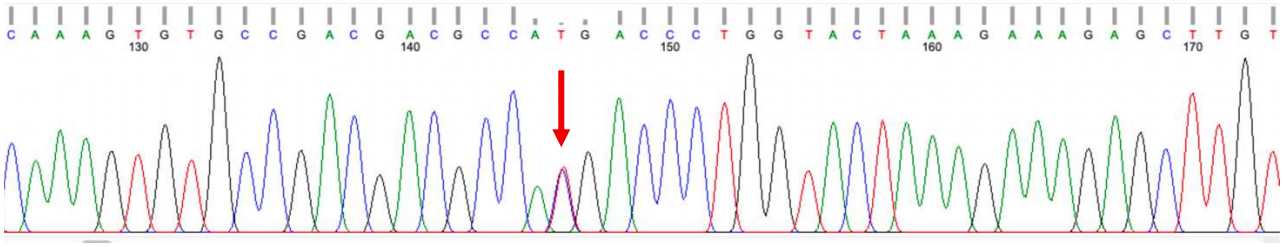


Figura 10. Elettroferogramma del campione D (*ENG* c. 1103 T>C (p.Met268Thr)). La freccia rossa indica la posizione della variante.

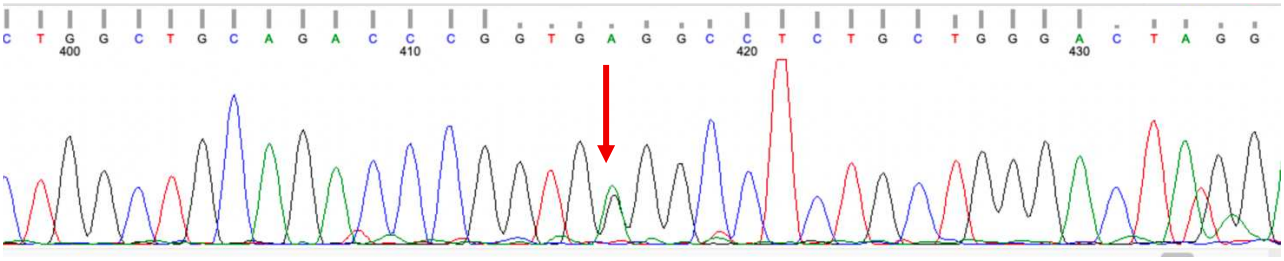


Figura 11. Elettroferogramma del campione E (*AVRL1* c. 1377+4 A>G). La freccia rossa indica la posizione della variante.

Il sequenziamento Sanger dei prodotti di amplificazione eseguito per il campione G e per il controllo ha confermato che le bande delle dimensioni attese corrispondevano esclusivamente ai prodotti derivanti dallo splicing fisiologico, senza evidenziare differenze tra il campione e il controllo. Alcuni esempi degli elettroferogrammi ottenuti sono riportati nelle Figure 12, 13, 14 e 15. Anche le bande aggiuntive sono state sequenziate e sono risultate, in entrambi i casi, prodotti di amplificazione aspecifici.

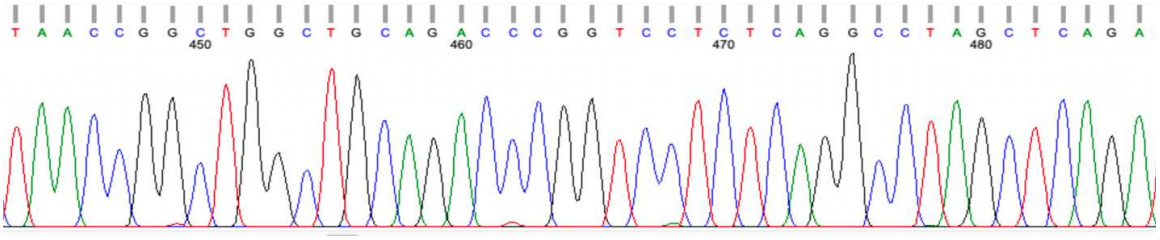


Figura 12. Elettroferogramma ottenuto mediante sequenziamento Sanger, con il primer Forward, del prodotto di amplificazione atteso nel campione di controllo.

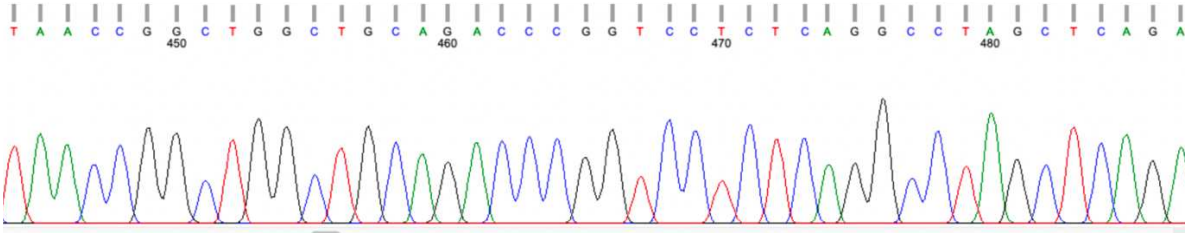


Figura 13. Elettroferogramma ottenuto mediante sequenziamento Sanger, con il primer Forward, del prodotto di amplificazione atteso nel campione G.

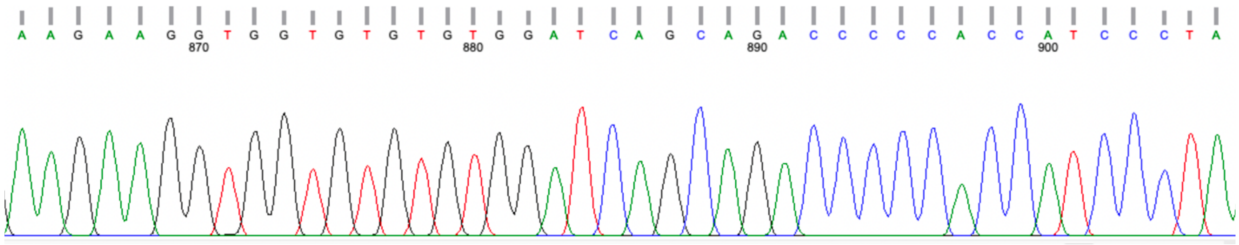


Figura 14. Elettroferogramma ottenuto mediante sequenziamento Sanger, con il primer Reverse, del prodotto di amplificazione atteso nel campione di controllo.

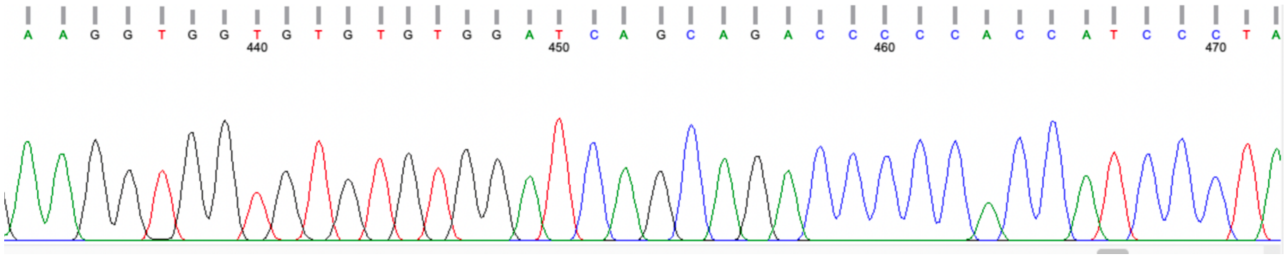


Figura 15. Elettroferogramma ottenuto mediante sequenziamento Sanger, con il primer Reverse, del prodotto di amplificazione atteso nel campione G.

ANALISI MEDIANTE REAL-TIME PCR

Al fine di identificare un trascritto la cui espressione possa essere paragonabile a quella dei geni da noi studiati, in particolare *ENG* e *ACVRL1*, ed avere quindi uno specifico controllo endogeno, è stata effettuata un'analisi mediante Real-Time PCR, utilizzando cDNA ottenuto dall'RNA estratto dal sangue periferico di un soggetto sano. Le precedenti analisi di riferimento erano state condotte utilizzando cDNA retrotrascritto dall'RNA di fibroblasti, mentre il cDNA dei pazienti analizzati derivava da sangue periferico.

L'utilizzo di un controllo ottenuto dallo stesso tessuto biologico dei pazienti ha quindi consentito di effettuare il confronto in condizioni fisiologicamente simili. I risultati sono stati confrontati con una precedente analisi di Real-Time PCR, fatta utilizzando specifiche sonde TaqMan™ per i geni *ENG* e *ACVRL1*, effettuata sempre su campioni di controllo.

Per ciascun target è stato considerato il valore di Cq (quantification cycle), che rappresenta il ciclo di amplificazione in cui il segnale di fluorescenza supera una determinata soglia, e diventa rilevabile. Il valore di Cq è inversamente proporzionale alla quantità iniziale del target; quindi, valori più bassi indicano una maggiore quantità iniziale del trascritto, mentre valori più elevati sono associati a una minore quantità.

Nel complesso, i geni endogeni analizzati hanno mostrato valori di Cq generalmente inferiori rispetto a quelli precedentemente ottenuti per *ENG* e *ACVRL1* mediante sonde TaqMan specifiche. Questa differenza è compatibile con il fatto che questi geni presentano livelli di espressione nel sangue più bassi rispetto a molti dei geni endogeni presenti nella piastra.

		1	2	3	4	5	6	7	8	9	10	11	12
A	CtrlntentA	18S	GAPDH	HPRT1	GUSB	18S	GAPDH	HPRT1	GUSB	18S	GAPDH	HPRT1	GUSB
	SampleA	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqA	35,47	33,38	39,94			20,27	38,59		17,01	28,14		33,77
B	CtrlntentB	ACTB	B2M	HMBS	IPO8	ACTB	B2M	HMBS	IPO8	ACTB	B2M	HMBS	IPO8
	SampleB	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqB	21,54	14,78	25,72	25,91	15,45	18,38	25,46	24,25	16,82	15,26	30,83	
C	CtrlntentC	PGK1	RPLP0	TBP	TFRC	PGK1	RPLP0	TBP	TFRC	PGK1	RPLP0	TBP	TFRC
	SampleC	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqC	20,57	18,20	24,98	22,78	21,02	18,02	24,54	25,12	21,00	19,35	38,05	
D	CtrlntentD	UBC	YWHAZ	PPIA	POLR2A	UBC	YWHAZ	PPIA	POLR2A	UBC	YWHAZ	PPIA	POLR2A
	SampleD	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqD	29,55	22,53	18,63	22,40	19,13	21,94	18,42	22,67	19,09	3,25	20,81	38,92
E	CtrlntentE	CASC3	CDKN1A	CDKN1B	GADD45A	CASC3	CDKN1A	CDKN1B	GADD45A	CASC3	CDKN1A	CDKN1B	GADD45A
	SampleE	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqE	21,53	24,82	20,86	26,30	21,30	24,59	21,62	26,09	21,08	24,93	20,67	
F	CtrlntentF	PUM1	PSMC4	EIF2B1	PES1	PUM1	PSMC4	EIF2B1	PES1	PUM1	PSMC4	EIF2B1	PES1
	SampleF	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqF	25,18	24,31	23,20	24,53	21,57	23,18	33,94	25,22	21,66	23,19	25,89	29,52
G	CtrlntentG	ABL1	ELF1	MT-ATP6	MRPL19	ABL1	ELF1	MT-ATP6	MRPL19	ABL1	ELF1	MT-ATP6	MRPL19
	SampleG	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqG	28,25	20,63	16,11	23,86	27,33	20,29	15,72	23,40	24,35	31,48	16,41	
H	CtrlntentH	POP4	RPL37A	RPL30	RPS17	POP4	RPL37A	RPL30	RPS17	POP4	RPL37A	RPL30	RPS17
	SampleH	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
	CqH	35,53	16,76		26,94	23,14	18,31	17,03	32,07	23,32	35,13	17,63	

Figura 16. Valori di ottenuti mediante TaqMan™ Array Human Endogenous Control Plate, 96-well, utilizzando cDNA ottenuto da RNA estratto da sangue periferico di un soggetto sano.

Nonostante ciò, dall'analisi dei valori di Cq è stato possibile individuare alcuni geni endogeni caratterizzati da livelli di espressione relativamente bassi, e quindi potenzialmente più adatti ad essere utilizzati come riferimento per il confronto nelle successive analisi. Un altro parametro utilizzato per la scelta è stata la deviazione standard, quale misura della ripetibilità delle analisi.

I risultati ottenuti ci hanno permesso di focalizzare la nostra attenzione sui geni *CDKN1A*, *CASC3* e *PGK1*. In particolare, *CDKN1A* presenta il numero di Cq più vicino a quello osservato nei geni *ENG* e *ACVRL1*, e una buona deviazione standard.

	A	B	C	D	E	F	G	H	I
1	Well	Fluor	Target	Sample	Cq		Cq mean		Dev st
2	E02	FAM	**CDKN1A**	Ctrl	24,82		24,783651		0,1729402
3	E06	FAM	**CDKN1A**	Ctrl	24,59				
4	E10	FAM	**CDKN1A**	Ctrl	24,93				
5									
6	E01	FAM	**CASC3**	Ctrl	21,53		21,302081		0,2276004
7	E05	FAM	**CASC3**	Ctrl	21,30				
8	E09	FAM	**CASC3**	Ctrl	21,08				
9									
10	C01	FAM	**PGK1**	Ctrl	20,57		20,862134		0,2552234
11	C05	FAM	**PGK1**	Ctrl	21,02				
12	C09	FAM	**PGK1**	Ctrl	21,00				
13									
14	G03	FAM	**MT-ATP6**	Ctrl	16,11		16,079106		0,346366
15	G07	FAM	**MT-ATP6**	Ctrl	15,72				
16	G11	FAM	**MT-ATP6**	Ctrl	16,41				
17									
18	E03	FAM	**CDKN1B**	Ctrl	20,86		21,052117		0,5001297
19	E07	FAM	**CDKN1B**	Ctrl	21,62				
20	E11	FAM	**CDKN1B**	Ctrl	20,67				
21									
22	F02	FAM	**PSMC4**	Ctrl	24,31		23,558777		0,647194
23	F06	FAM	**PSMC4**	Ctrl	23,18				
24	F10	FAM	**PSMC4**	Ctrl	23,19				
25									
26	C02	FAM	**RPLP0**	Ctrl	18,20		18,52323		0,7230673
27	C06	FAM	**RPLP0**	Ctrl	18,02				
28	C10	FAM	**RPLP0**	Ctrl	19,35				
29									
30	D03	FAM	**PPIA**	Ctrl	18,63		19,28567		1,326086
31	D07	FAM	**PPIA**	Ctrl	18,42				
32	D11	FAM	**PPIA**	Ctrl	20,81				

Figura 17. Valori di Cq ottenuti per i geni endogeni analizzati in triplicato mediante Real-Time PCR, ordinati secondo valori crescenti di deviazione standard.

DISCUSSIONE

La Teleangectasia Emorragica Ereditaria (Hereditary Hemorrhagic Teleangiectasia, HHT) è una patologia vascolare a trasmissione autosomica dominante, caratterizzata dalla presenza di teleangectasie mucocutanee e malformazioni vascolari che possono interessare diversi organi. Le manifestazioni cliniche presentano un'espressività variabile e possono comparire progressivamente con l'età, rendendo in alcuni casi più complesso il riconoscimento della malattia, soprattutto nei soggetti più giovani o con un quadro clinico poco evidente (*Faughnan et al., 2020*). La diagnosi clinica dell'HHT si basa sui criteri di Curaçao, che comprendono epistassi spontanee e ricorrenti, teleangectasie in sedi caratteristiche, malformazioni vascolari viscerali e familiarità di primo grado per l'HHT. La presenza di almeno tre criteri consente di formulare una diagnosi clinica definita, mentre due criteri identificano una condizione sospetta (*Shovlin et al., 2000*). Tuttavia, la variabilità delle manifestazioni fenotipiche rende necessario, in alcuni casi, l'analisi genetica al fine di confermare la diagnosi. L'identificazione di una variante causativa permette quindi di confermare la diagnosi molecolare e successivamente di cercare, prima che la malattia si manifesti o che insorgano le complicanze ad essa legata, la stessa variante nei familiari (*Faughnan et al., 2020*).

L'analisi molecolare dell'HHT risulta quindi essenziale per confermare la possibile diagnosi nei soggetti molto giovani o con clinica dubbia. Le analisi sono maggiormente incentrate sui 2 principali geni causativi della malattia (circa il 90% dei casi), *ENG* e *ACVRL1*, responsabili rispettivamente dell'HHT1 e HHT2 (*McDonald et al., 2020*). Altri geni che vengono analizzati oltre a quelli già citati sono *SMAD4*, associata alla sindrome combinata JP-HHT, ma anche *GDF2*, *RASA1*, *EPHB4* e *BMPR2*, coinvolti in condizioni vascolari che possono presentare caratteristiche cliniche sovrapponibili a quelle dell'HHT (*Jelsig et al., 2016; Desroches-Castan et al., 2026; Wooderchak-Donahue et al., 2019*).

Approcci di sequenziamento come il Whole Exome Sequencing (WES), nonostante i progressi delle tecniche di sequenziamento e delle pipelines di analisi, non sempre riescono a identificare la variante causativa e il Whole Genome Sequencing (WGS) può contribuire ad ampliare la ricerca. Il WES consente l'analisi delle regioni codificanti del genoma, mentre il WGS permette di estendere l'analisi anche alle regioni non codificanti, con l'identificazione di varianti introniche o strutturali (*Belkadi et al., 2015*).

In questo contesto, si inserisce il presente lavoro di tesi sperimentale, che si è incentrato sullo studio molecolare dell'HHT in tre linee sperimentali parallele.

La prima parte ha riguardato la validazione, a livello del DNA genomico, di varianti precedentemente identificate da un sequenziamento WGS nei geni *ENG* e *ACVRL1*, mediante l'amplificazione delle regioni di interesse e sequenziamento Sanger. A questo scopo è stato seguito un workflow sperimentale, in cui ciascuna fase è stata necessaria per ottenere un prodotto di DNA di qualità e specificità adeguata.

Il sequenziamento Sanger ha confermato la presenza delle varianti precedentemente identificate mediante WGS. Le varianti sono state successivamente valutate mediante database e strumenti bioinformatici, che hanno permesso di raccogliere informazioni relative al loro significato clinico e di analizzarne il possibile impatto sullo splicing.

La classificazione delle varianti è stata quindi approfondita mediante l'applicazione dei criteri ACMG/AMP, considerando anche le indicazioni gene-specifiche disponibili per *ENG* e *ACVRL1*. In particolare, è stato effettuato un confronto tra la classificazione riportata nei database consultati e quella ottenuta applicando i criteri ACMG/AMP gene-specifici.

Le varianti identificate nei campioni A (*ENG* c. 626 delT (p.Leu209Trp fs*13)), B (*ACVRL1* c. 696 delA (p.Lys323Thr fs*68)) e C (*ENG* c. 392 delC (p.Pro131Arg fs*17) sono delezioni di un singolo nucleotide che determinano un frameshift, con conseguente introduzione di un codone di stop prematuro. Tali alterazioni possono determinare la perdita della funzione della proteina, attraverso la produzione di una proteina tronca o la degradazione del trascritto mediante NMD. La perdita di funzione rappresenta un meccanismo patogenetico noto per i geni *ENG* e *ACVRL1*, coerente con la classificazione Pathogenic/Likely Pathogenic ottenuta dai software. In particolare, le linee guida ACMG/AMP specifiche per *ENG* e *ACVRL1* indicano che le varianti che determinano un'interruzione prematura della cornice di lettura sono predette come suscettibili al nonsense-mediated decay (NMD) quando localizzate entro il codone 601 di *ENG* o entro il codone 442 di *ACVRL1* (*DeMille et al., 2024*). Le tre varianti da noi analizzate rientrano in questi intervalli e risultano pertanto compatibili con una possibile degradazione del trascritto mediante NMD.

La variante identificata nel campione A non risulta presente nel nostro database, così come quella del campione B. Per quest'ultimo, tuttavia, avendo a disposizione il materiale genetico del padre affetto del paziente, è stato possibile studiare la cosegregazione nella famiglia, confermando la presenza della variante anche nel genitore. La variante del campione C è stata invece ritrovata in altre tre

famiglie presenti all'interno del nostro database, due provenienti dalla Lombardia e una dalla Sicilia, ma non è stato possibile stabilire l'esistenza di un'eventuale relazione di parentela tra queste.

Un aspetto interessante emerge inoltre dall'analisi mediante i software di predizione dello splicing nei campioni B e C. In questi due casi, oltre all'effetto del frameshift, è stato predetto un possibile effetto anche sullo splicing del pre-mRNA, attraverso l'attivazione di un sito accettore criptico nel campione B e di un sito donatore criptico nel campione C. I siti criptici sono sequenze normalmente non utilizzate durante lo splicing, ma che in seguito all'introduzione della variante, possono acquisire caratteristiche tali da essere riconosciute dallo spliceosoma. Il loro utilizzo al posto dei normali siti di splicing potrebbe determinare un taglio del pre-mRNA in un sito anomalo.
Tuttavia, non essendo disponibili campioni di RNA dei pazienti interessati non è stato possibile verificare sperimentalmente tali predizioni mediante analisi del cDNA.

La variante identificata nel campione D (*ENG* c. 1103 T>C (p.Met268Thr)) è una variante missenso, in cui la sostituzione di una timina con una citosina determina, a livello proteico, la sostituzione della metionina in posizione 268 con una treonina. Si tratta quindi di una variante che, a differenza delle delezioni precedentemente descritte, non causa un frameshift, ma determina la sostituzione di un singolo amminoacido. In particolare, la metionina è un amminoacido prevalentemente apolare e idrofobico, mentre la treonina presenta una catena laterale polare, e quindi la sostituzione determina un cambiamento delle proprietà chimico-fisiche del residuo, che potrebbe influenzare localmente la struttura o la funzione della proteina. La variante, inoltre, è stata precedentemente riportata in un altro paziente nel nostro laboratorio.
Per quanto riguarda un possibile effetto sulla maturazione dell'RNA, i software di predizione dello splicing utilizzati non hanno evidenziato alterazioni significative, suggerendo che il principale effetto della variante sia riconducibile alla sostituzione amminoacidica piuttosto che a un'alterazione dello splicing.

Le varianti identificate nei campioni E (*ACVRL1* 1377+4A>G) e F (*ENG* 1311+5G>A) sono entrambe localizzate in regioni introniche prossime ai siti di splicing.

Per la variante F, i software predicono un'alterazione del sito donatore di splicing, con uno spostamento del sito di taglio e una conseguente ritenzione parziale dell'introne 10. La variante risulta inoltre classificata come Likely Pathogenic.

La variante non era presente nel nostro database, ma nella stessa regione di *ENG* sono presenti due varianti note, localizzate in posizione +2, che interessano il sito canonico di splicing. È interessate osservare come la variante in esame, pur essendo localizzata in posizione +5, e quindi al di fuori delle due basi canoniche del sito donatore, possa comunque avere un effetto sullo splicing. Tuttavia, non essendo disponibile l'RNA del paziente, non è stato possibile verificare sperimentalmente l'impatto della variante sul trascritto.

La variante E rappresenta invece l'unica classificata come Variante di Significato Incerto (Uncertain Significance, VUS). La sua localizzazione in prossimità del sito donatore suggerisce un possibile coinvolgimento nel processo di splicing e le analisi tramite software bioinformatici ne predicono una possibile alterazione. Tuttavia, non essendo disponibile RNA del campione E, non è stato possibile verificare sperimentalmente l'effetto della variante sul trascritto.
Questa variante è stata riscontrata in diverse famiglie presenti nel nostro database e, in particolare, in sette famiglie provenienti dal Veneto. Questa ricorrenza geografica potrebbe suggerire la presenza di un possibile effetto del fondatore, fenomeno già descritto per altre varianti associate all'HHT (*Gallione et al., 2000; Lesca et al., 2008; Sbalchiero et al., 2022*). Tre dei pazienti presenti nel database sono casi singoli, mentre in due famiglie è stato possibile effettuare uno studio di cosegregazione, rispettivamente tra padre e figlia, ma anche tra zia e pronipoti. Per il paziente E non è presente materiale biologico dei familiari, ma una futura acquisizione permetterebbe di valutare la cosegregazione della variante all'interno della famiglia.

La stessa variante è stata anche identificata anche nel campione G, per cui era disponibile un campione di RNA. La possibilità di analizzare direttamente il trascritto ha quindi permesso di approfondire il possibile effetto della variante sullo splicing. Lo studio della variante nel campione G costituisce quindi la seconda parte di questo lavoro di tesi sperimentale, finalizzata alla caratterizzazione del suo possibile effetto sulla maturazione dell'mRNA.

La seconda parte del mio lavoro di tesi si è quindi incentrata sullo studio di una variante di *ACVRL1* ma, a differenza dei campioni precedentemente analizzati, l'analisi è stata condotta a partire dall'RNA, retrotrascritto in cDNA. La retrotrascrizione si è resa necessaria poiché l'RNA non può essere utilizzato direttamente come stampo nella PCR, e deve quindi essere convertito in DNA complementare, che può essere amplificato e analizzato.

L'utilizzo del cDNA assume particolare rilevanza nello studio di varianti localizzate in prossimità dei siti di splicing, poiché consente di analizzare direttamente il trascritto derivante dal processo di maturazione dell'RNA. Lo splicing determina infatti la rimozione degli introni dal pre-mRNA, e la corretta congiunzione degli esoni, attraverso il riconoscimento di specifiche sequenze.
Una variante che interessa queste regioni può alterare il riconoscimento dai parte dei fattori di splicing, e determinare una maturazione anomala del trascritto, ad esempio, attraverso la perdita o acquisizione di un sito di splicing, che può comportare l'esclusione di un esone o la ritenzione, completa o parziale, di una sequenza intronica.
Tali alterazioni non possono essere dimostrate attraverso la sola analisi della sequenza del DNA genomico, che permette di identificare la variante, ma non di osservare direttamente le sue conseguenze sul trascritto. L'analisi del cDNA consente invece di esaminare ciò che avviene a livello dell'RNA maturo, e rappresenta quindi un approccio utile per verificare sperimentalmente se una variante coinvolta nello splicing determini effettivamente un'alterazione del trascritto.

Il cDNA ottenuto dal paziente è stato quindi utilizzato per amplificare mediante PCR la regione di *ACVRL1* comprendente il sito interessato, procedendo successivamente alla verifica della variante, secondo un approccio analogo a quello utilizzato per i campioni analizzati a partire dal DNA genomico. L'analisi dei prodotti di amplificazione non ha evidenziato la presenza di bande aggiuntive specifiche riconducibili ad alterazioni dello splicing, in quanto le bande aggiuntive sono risultate, in seguito al sequenziamento, prodotti di amplificazione aspecifici. Risultati sovrapponibili sono derivati sia da cDNA ottenuti tramite retrotrascrizione con Random Examers, sia con Oligo(dT). In quest'ultimo caso, la scelta del primer di innesco della retrotrascrizione ha portato a una selezione qualitativa dell'RNA retrotrascritto che dovrebbe essere rappresentato in grande percentuale da mRNA.

È stata tuttavia osservata una minore intensità della banda corrispondente al prodotto atteso nel campione del paziente rispetto al controllo. Tale osservazione è stata riscontrata in esperimenti indipendenti, effettuati a partire da differenti reazioni di retrotrascrizione. Poiché la PCR qualitativa non consente di stabilire se tale differenza di intensità rifletta una reale differenza quantitativa del trascritto o sia dovuta a variabili tecniche, questo dovrà essere ulteriormente verificato mediante Real-Time PCR, che consentirà una valutazione quantitativa del trascritto di *ACVRL1*, evidenziando eventuali differenze fra paziente e controllo.

La terza e ultima linea sperimentale ha invece riguardato l'analisi mediante Real-Time PCR di una piastra di geni endogeni, con lo scopo di individuare potenziali geni di riferimento da utilizzare nelle successive analisi di espressione di *ENG* e *ACVRL1* su campioni di sangue periferico.

Negli studi di espressione genica mediante RT-qPCR è necessario disporre di uno o più geni di riferimento, utilizzati per confrontare correttamente i livelli di espressione dei geni di interesse tra campioni diversi. Affinché il confronto sia attendibile, questi geni devono mantenere livelli di espressione costanti nei diversi campioni, e non essere influenzati dalla patologia analizzata.
Per questo motivo, prima di valutare eventuali variazioni nell'espressione di *ENG* e *ACVRL1* nei pazienti HHT, è necessario individuare geni di riferimento adeguati e stabili nei campioni di sangue periferico utilizzati nello studio.

Non esiste tuttavia un gene di riferimento universalmente stabile, e anche geni housekeeping comunemente utilizzati come controllo, ad esempio *ACTB*, *GAPDH* e *β-tubulina*, possono mostrare variazioni di espressione in funzione del tipo cellulare, del tessuto e della condizione sperimentale o patologica. Per questo motivo, la stabilità dei potenziali geni di riferimento dovrebbe essere verificata nello stesso sistema biologico (*Stamova et al., 2009)*.

È noto che endoglina, codificata dal gene *ENG*, svolge un ruolo nella regolazione del citoscheletro delle cellule endoteliali, e una sua ridotta espressione è stata associata ad alterazioni dell'organizzazione dell'F-actina. È stato anche osservato che il silenziamento del gene *ENG* può modificare la distribuzione e l'organizzazione tra filamenti di actina e tubulina nelle cellule endoteliali. Queste evidenze suggeriscono quindi di porre maggiore attenzione nell'utilizzo di geni come *ACTB* e geni codificanti per la tubulina negli studi di espressione dell'HHT (*Rossi et al., 2021*).
Anche GAPDH, nonostante sia frequentemente utilizzato come gene di riferimento, potrebbe non rappresentare una scelta ottimale, in quanto la sua espressione non è costante in tutte le linee cellulari, ma può essere influenzata da variazioni metaboliche, e in particolare dall'ipossia (*Graven et al., 1994*). Di conseguenza, eventuali variazioni della sua espressione potrebbero compromettere una corretta normalizzazione dei dati di Real-Time PCR, rendendo necessaria la valutazione sperimentale della sua stabilità prima di utilizzarlo come gene endogeno di riferimento.

In questo lavoro di tesi, questo aspetto ha particolare importanza, poiché le future analisi di espressione di *ENG* e *ACVRL1* saranno condotte su campioni di sangue periferico. A questo scopo è stato utilizzato il cDNA, ottenuto dall'RNA estratto dal sangue periferico di un soggetto sano, in

modo da effettuare la selezione dei geni endogeni, che verranno poi utilizzati per lo studio dell'espressione genica nei pazienti. Questo consente di evitare di basare il confronto su geni di riferimento, e su valori ottenuti da altri tipi cellulari, come i fibroblasti, nei quali il profilo di espressione può essere differente.

L'analisi della piastra ha permesso di confrontare i diversi geni endogeni sulla base dei rispettivi valori di Cq (quantification cycle), ovvero il ciclo in corrispondenza del quale il segnale di fluorescenza supera una determinata soglia, e diventa rilevabile. A parità di condizioni sperimentali, un valore di Cq più basso è indicativo di una maggiore quantità iniziale del trascritto, mentre un Cq più elevato corrisponde a una minore quantità iniziale. Poiché ciascun gene nella piastra è stato analizzato in triplicato, è stato inoltre possibile valutare la variabilità tra le repliche, attraverso la media e la deviazione standard dei valori di Cq.

Nel complesso, i geni endogeni analizzati hanno mostrato valori di Cq generalmente inferiori rispetto a quelli precedentemente ottenuti nel nostro laboratorio usando sonde TaqMan commerciali per *ENG* e *ACVRL1*, che presentavano dei Cq intorno ai 29 cicli. Dall'analisi della piastra è stato comunque possibile individuare alcuni geni caratterizzati da valori di Cq più vicini a quelli dei due geni target e da una buona riproducibilità tra i triplicati, e che quindi possono essere considerati buoni candidati per le successive fasi di validazione. Nel nostro caso, i principali geni che hanno mostrato buoni valori di Cq sono *CDKN1A*, *CASC3* e *PGK1*, che rappresentano quindi i principali candidati da sottoporre alle successive analisi di validazione.
CDKN1A codifica per p21, una proteina coinvolta nella regolazione del ciclo cellulare; *CASC3* codifica per un componente del Exon Junction Complex (EJC), coinvolto nei processi di maturazione dell'mRNA; *PGK1* codifica invece per la fosfoglicerato chinasi 1, un enzima coinvolto nella glicolisi. In questo contesto, *CDKN1A* potrebbe essere un candidato meno appropriato, poiché la sua espressione è direttamente regolata dalla via di segnalazione TGF-β/SMAD. *CASC3* e *PGK1* risultano invece candidati più interessanti da sottoporre alle successive analisi di validazione, nelle quali dovrà essere verificata la stabilità della loro espressione nelle specifiche condizioni sperimentali.

In esperimenti futuri, tali geni potranno essere analizzati sia nei soggetti sani, sia nei pazienti HHT, al fine di verificarne la stabilità di espressione e identificare quelli più idonei alla normalizzazione dei dati. I geni di riferimento selezionati potranno essere utilizzati per la valutazione dei livelli di espressione di *ENG* e *ACVRL1* nei pazienti.

CONCLUSIONI

Il presente lavoro di tesi ha permesso di approfondire la caratterizzazione molecolare di alcune varianti dei geni *ENG* e *ACVRL1*, i due principali geni coinvolti nella Teleangectasia Emorragica Ereditaria. L'integrazione di differenti processi sperimentali e bioinformatici ha consentito non solo di confermare la presenza delle varianti identificate, ma anche di raccogliere ulteriori informazioni utili alla loro interpretazione e al possibile coinvolgimento nei meccanismi molecolari alla base della patologia.

Tramite il sequenziamento Sanger è stato possibile confermare le varianti precedentemente identificate mediante WGS nei sei pazienti analizzati, e la successiva valutazione attraverso database e strumenti di predizione *in silico* ha permesso di approfondirne il significato clinico e, per alcune di esse, di evidenziare un possibile coinvolgimento nei meccanismi di splicing. In particolare, le analisi hanno mostrato come anche varianti localizzate al di fuori dei siti canonici possano interferire con la corretta maturazione del trascritto, sottolineando l'importanza di ampliare le regioni da studiare.

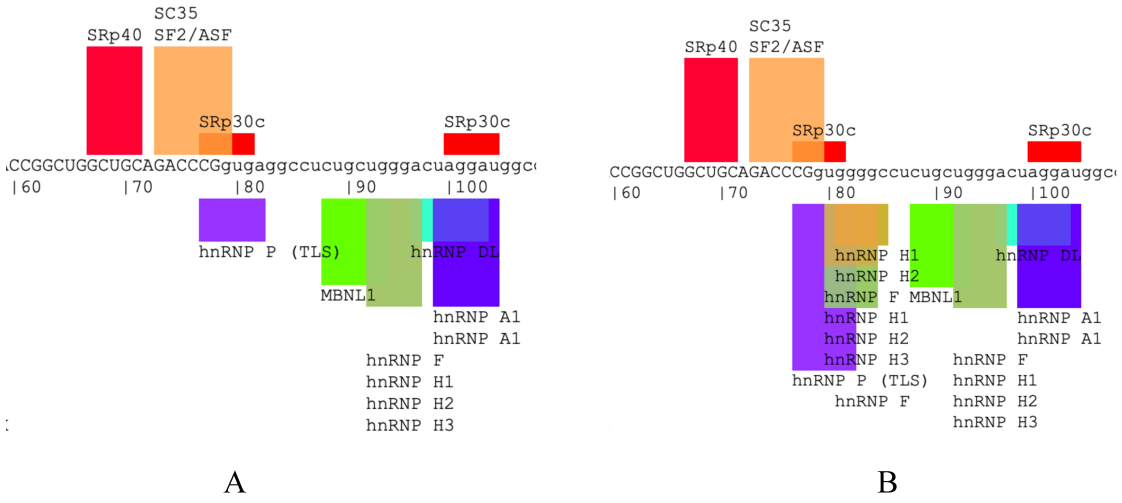
Proprio per questo, è stata sviluppata la seconda parte del lavoro, dedicata allo studio dell'RNA di un settimo paziente, che presentava una delle varianti precedentemente descritte. L'analisi del cDNA rappresenta infatti un passaggio importante per approfondire le conseguenze di varianti potenzialmente coinvolte nello splicing, poiché permette di studiare direttamente il trascritto maturo, e di andare oltre le sole predizioni ottenute mediante strumenti bioinformatici.

Sebbene le nostre analisi non abbiano evidenziato alterazioni qualitative chiaramente riconducibili a un evento di splicing aberrante, non è possibile escludere che la variante possa esercitare un effetto a livello quantitativo, determinando una variazione dei livelli del trascritto. Tale ipotesi dovrà essere verificata mediante analisi di espressione quantitativa, come la Real-Time PCR. Varianti localizzate nelle regioni introniche possono infatti esercitare effetti regolatori indipendenti dallo splicing, interferendo con elementi cis- regolatori o con siti di legame per fattori di trascrizione e determinando alterazioni dell'attività trascrizionale del gene.

Un meccanismo di questo tipo è stato già descritto nell'HHT. In particolare, una variante intronica di *ACVRL1* (c. 772+27G>C) è stata associata a una riduzione dei livelli di mRNA e proteina del gene in assenza di alterazioni dello splicing. Gli studi funzionali hanno dimostrato una riduzione dell'attività trascrizionale associata alla variante e una alterazione del legame del fattore di trascrizione Sp1, suggerendo quindi un meccanismo patogenetico basato sulla regolazione della trascrizione piuttosto

che sull'alterazione dello splicing (*Yu et al., 2013*). Anche il nostro gruppo ha identificato la variante c. 1852+42C>T nel gene *ENG*, dimostrando un meccanismo simile a questo (*Plumitallo et al., 2018*).

L'analisi in silico mediante SpliceAid è stata effettuata confrontando la sequenza wild-type con quella contenente la variante. Il confronto ha evidenziato una modificazione del profilo dei potenziali siti di legame per fattori regolatori dello splicing in prossimità della variante. Sulla base di tale predizione, si ipotizzava che la variante potesse interferire con il normale riconoscimento del sito di splicing e favorire eventi di splicing aberrante, come la ritenzione di una porzione o dell'intero introne, con conseguente formazione di un trascritto di dimensioni maggiori. Tuttavia, questa ipotesi non è stata finora confermata sperimentalmente dall'analisi dei prodotti ottenuti.



A B

Figura 18. Analisi mediante SpliceAid della regione di interesse di *ACVRL1* nel controllo (A) e nel paziente (G) con rappresentazione dei siti di legame predetti. Le barre colorate coprono le sequenze riconosciute dalle diverse proteine regolatrici dello splicing; la loro altezza è proporzionale alla forza di legame con queste sequenze, permettendo di confrontare differenze nel profilo di legame tra la sequenza WT e quella mutata.

Una terza linea sperimentale sviluppata nel corso del lavoro ha riguardato la ricerca di geni endogeni di riferimento da utilizzare nelle future analisi di espressione di *ENG* e *ACVRL1* su campioni di sangue periferico. L'analisi mediante TaqMan™ Array Human Endogenous Control Plate ha consentito di effettuare una prima selezione di potenziali geni candidati, sulla base dei livelli di espressione e della riproducibilità tra i triplicati. In particolare, i geni *CDKN1A*, *CASC3* e *PGK1* sono stati selezionati come principali geni candidati da approfondire nelle successive fasi di valutazione. Questa rappresenta una selezione preliminare e non consente ancora di considerarli geni di riferimento pienamente confermati.

Le prospettive future riguarderanno quindi, in primo luogo, l'estensione delle analisi funzionali delle varianti in cui si ipotizza un possibile effetto sullo splicing, nel momento in cui saranno disponibili campioni di RNA dei pazienti interessati. L'analisi del cDNA permetterebbe infatti di verificare sperimentalmente la presenza di trascritti anomali. Per alcune varianti potrebbe inoltre essere utile ampliare lo studio familiare, attraverso l'analisi di ulteriori soggetti appartenenti alle famiglie dei pazienti, così da ottenere informazioni sulla cosegregazione della variante con la malattia.

Parallelamente, sarà necessario proseguire la validazione dei geni endogeni individuati in questa prima fase. *CDKN1A*, *CASC3* e *PGK1* potranno essere analizzati su un numero maggiore di campioni, provenienti sia da soggetti sani, sia da pazienti con HHT, per verificarne la stabilità di espressione. Una volta individuato uno o più geni di riferimento adeguati, sarà possibile utilizzarli per la normalizzazione dei dati e procedere allo studio dell'espressione dei geni *ENG* e *ACVRL1* nel sangue periferico dei pazienti.

Nel complesso, i risultati ottenuti evidenziano l'importanza di integrare l'analisi della sequenza del DNA con lo studio delle possibili conseguenze delle varianti a livello dell'RNA. Questo approccio può contribuire a una caratterizzazione più completa delle varianti associate all'HHT, in particolare nei casi in cui il loro significato biologico o il loro effetto funzionale non siano totalmente chiari.

BIBLIOGRAFIA

Abdalla SA, Letarte M. Hereditary haemorrhagic telangiectasia: current views on genetics and mechanisms of disease. *J Med Genet*. 2006;43(2):97-110. doi:10.1136/jmg.2005.030833

Adams RH, Wilkinson GA, Weiss C, et al. Roles of ephrinB ligands and EphB receptors in cardiovascular development: demarcation of arterial/venous domains, vascular morphogenesis, and sprouting angiogenesis. *Genes Dev*. 1999;13(3):295-306. doi:10.1101/gad.13.3.295

Aguilera C, Padró-Miquel A, Esteve-Garcia A, et al. Improving Hereditary Hemorrhagic Telangiectasia Molecular Diagnosis: A Referral Center Experience. *Genes (Basel)*. 2023;14(3):772. Published 2023 Mar 22. doi:10.3390/genes14030772

Ahlhelm F, Lieb J, Schneider G, Müller U, Ulmer S. Morbus Osler [Osler's disease]. *Radiologe*. 2013;53(12):1084-1090. doi:10.1007/s00117-013-2552-z

Al Tabosh T, Al Tarrass M, Tourvieilhe L, Guilhem A, Dupuis-Girod S, Bailly S. Hereditary hemorrhagic telangiectasia: from signaling insights to therapeutic advances. *J Clin Invest*. 2024;134(4):e176379. Published 2024 Feb 15. doi:10.1172/JCI176379

Alaa El Din F, Patri S, Thoreau V, et al. Functional and splicing defect analysis of 23 ACVRL1 mutations in a cohort of patients affected by Hereditary Hemorrhagic Telangiectasia. *PLoS One*. 2015;10(7):e0132111. Published 2015 Jul 15. doi:10.1371/journal.pone.0132111

Amyere M, Revencu N, Helaers R, et al. Germline Loss-of-Function Mutations in EPHB4 Cause a Second Form of Capillary Malformation-Arteriovenous Malformation (CM-AVM2) Deregulating RAS-MAPK Signaling. *Circulation*. 2017;136(11):1037-1048. doi:10.1161/CIRCULATIONAHA.116.026886

Arthur HM, Roman BL. An update on preclinical models of hereditary haemorrhagic telangiectasia: Insights into disease mechanisms. *Front Med (Lausanne)*. 2022;9:973964. Published 2022 Sep 29. doi:10.3389/fmed.2022.973964

Austin-Tse CA, Jobanputra V, Perry DL, et al. Best practices for the interpretation and reporting of clinical whole genome sequencing. *NPJ Genom Med*. 2022;7(1):27. Published 2022 Apr 8. doi:10.1038/s41525-022-00295-z

Baeyens N, Larrivée B, Ola R, et al. Defective fluid shear stress mechanotransduction mediates hereditary hemorrhagic telangiectasia. *J Cell Biol*. 2016;214(7):807-816. doi:10.1083/jcb.201603106

Balachandar S, Graves TJ, Shimonty A, et al. Identification and validation of a novel pathogenic variant in GDF2 (BMP9) responsible for hereditary hemorrhagic telangiectasia and pulmonary arteriovenous malformations. *Am J Med Genet A*. 2022;188(3):959-964. doi:10.1002/ajmg.a.62584

Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011;12(11):745-755. Published 2011 Sep 27. doi:10.1038/nrg3031

Bao R, Huang L, Andrade J, et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform*. 2014;13(Suppl 2):67-82. Published 2014 Sep 21. doi:10.4137/CIN.S13779

Bayrak-Toydemir P, McDonald J, Markewitz B, et al. Genotype-phenotype correlation in hereditary hemorrhagic telangiectasia: mutations and manifestations. *Am J Med Genet A*. 2006;140(5):463-470. doi:10.1002/ajmg.a.31101

Begbie ME, Wallace GM, Shovlin CL. Hereditary haemorrhagic telangiectasia (Osler-Weber-Rendu syndrome): a view from the 21st century. *Postgrad Med J*. 2003;79(927):18-24. doi:10.1136/pmj.79.927.18

Belkadi A, Bolze A, Itan Y, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci U S A*. 2015;112(17):5473-5478. doi:10.1073/pnas.1418631112

Bellón T, Corbí A, Lastres P, et al. Identification and expression of two forms of the human transforming growth factor-beta-binding protein endoglin with distinct cytoplasmic regions. *Eur J Immunol*. 1993;23(9):2340-2345. doi:10.1002/eji.1830230943

Berg JN, Gallione CJ, Stenzel TT, et al. The activin receptor-like kinase 1 gene: genomic structure and mutations in hereditary hemorrhagic telangiectasia type 2. *Am J Hum Genet*. 1997;61(1):60-67. doi:10.1086/513903

Bernabeu C, Bayrak-Toydemir P, McDonald J, Letarte M. Potential Second-Hits in Hereditary Hemorrhagic Telangiectasia. *J Clin Med*. 2020;9(11):3571. Published 2020 Nov 5. doi:10.3390/jcm9113571

Bidart M, Ricard N, Levet S, et al. BMP9 is produced by hepatocytes and circulates mainly in an active mature form complexed to its prodomain. *Cell Mol Life Sci*. 2012;69(2):313-324. doi:10.1007/s00018-011-0751-1

Brown MA, Zhao Q, Baker KA, et al. Crystal structure of BMP-9 and functional interactions with pro-region and receptors. *J Biol Chem*. 2005;280(26):25111-25118. doi:10.1074/jbc.M503328200

Burdick KJ, Cogan JD, Rives LC, et al. Limitations of exome sequencing in detecting rare and undiagnosed diseases. *Am J Med Genet A*. 2020;182(6):1400-1406. doi:10.1002/ajmg.a.61558

Buscarini E, Danesino C, Plauchu H, et al. High prevalence of hepatic focal nodular hyperplasia in subjects with hereditary hemorrhagic telangiectasia. *Ultrasound Med Biol*. 2004;30(9):1089-1097. doi:10.1016/j.ultrasmedbio.2004.08.004

Danesino C, Cantarini C, Olivieri C. Hereditary Hemorrhagic Telangiectasia in Pediatric Age: Focus on Genetics and Diagnosis. *Pediatr Rep*. 2023;15(1):129-142. Published 2023 Feb 10. doi:10.3390/pediatric15010011

David L, Mallet C, Mazerbourg S, Feige JJ, Bailly S. Identification of BMP9 and BMP10 as functional activators of the orphan activin receptor-like kinase 1 (ALK1) in endothelial cells. *Blood*. 2007;109(5):1953-1961. doi:10.1182/blood-2006-07-034124

DeBose-Scarlett E, Marchuk DA. The Role of Somatic Mutation in Hereditary Hemorrhagic Telangiectasia Pathogenesis. *J Clin Med*. 2025;14(13):4479. Published 2025 Jun 24. doi:10.3390/jcm14134479

DeBose-Scarlett E, Ressler AK, Friday C, et al. Arteriovenous malformation from a patient with JP-HHT harbours two second-hit somatic DNA alterations in *SMAD4*. *J Med Genet*. 2025;62(4):281-288. Published 2025 Mar 20. doi:10.1136/jmg-2024-110569

DeBose-Scarlett E, Ressler AK, Gallione CJ, et al. Somatic mutations in arteriovenous malformations in hereditary hemorrhagic telangiectasia support a bi-allelic two-hit mutation mechanism of pathogenesis. *Am J Hum Genet*. 2024;111(10):2283-2298. doi:10.1016/j.ajhg.2024.08.020

DeMille D, McDonald J, Bernabeu C, et al. Specifications of the ACMG/AMP Variant Curation Guidelines for Hereditary Hemorrhagic Telangiectasia Genes-*ENG* and *ACVRL1*. *Hum Mutat*. 2024;2024:3043736. Published 2024 May 18. doi:10.1155/2024/3043736

Deng Z, Morse JH, Slager SL, et al. Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. *Am J Hum Genet*. 2000;67(3):737-744. doi:10.1086/303059

Desroches-Castan A, Beurier-Soulat L, Tusseau M, et al. Functional and Molecular Characterization of Novel *GDF2* (*BMP9*) and *BMP10* Variants From the French PAH and HHT Cohorts. *Circ Genom Precis Med*. 2026;19(4):e005626. doi:10.1161/CIRCGEN.125.005626

Easey AJ, Wallace GM, Hughes JM, Jackson JE, Taylor WJ, Shovlin CL. Should asymptomatic patients with hereditary haemorrhagic telangiectasia (HHT) be screened for cerebral vascular malformations? Data from 22,061 years of HHT patient life. *J Neurol Neurosurg Psychiatry*. 2003;74(6):743-748. doi:10.1136/jnnp.74.6.743

Eerola I, Boon LM, Mulliken JB, et al. Capillary malformation-arteriovenous malformation, a new clinical and genetic disorder caused by RASA1 mutations. *Am J Hum Genet*. 2003;73(6):1240-1249. doi:10.1086/379793

Faughnan ME, Mager JJ, Hetts SW, et al. Second International Guidelines for the Diagnosis and Management of Hereditary Hemorrhagic Telangiectasia. *Ann Intern Med*. 2020;173(12):989-1001. doi:10.7326/M20-1443

Faughnan ME, Palda VA, Garcia-Tsao G, et al. International guidelines for the diagnosis and management of hereditary haemorrhagic telangiectasia. *J Med Genet*. 2011;48(2):73-87. doi:10.1136/jmg.2009.069013

Fish A, Henderson K, Moushey A, Pollak J, Schlachter T. Incidence of Spontaneous Pulmonary AVM Rupture in HHT Patients. *J Clin Med*. 2021;10(20):4714. Published 2021 Oct 14. doi:10.3390/jcm10204714

Gallione C, Aylsworth AS, Beis J, et al. Overlapping spectra of SMAD4 mutations in juvenile polyposis (JP) and JP-HHT syndrome. *Am J Med Genet A*. 2010;152A(2):333-339. doi:10.1002/ajmg.a.33206

Gallione CJ, Klaus DJ, Yeh EY, et al. Mutation and expression analysis of the endoglin gene in hereditary hemorrhagic telangiectasia reveals null alleles. *Hum Mutat*. 1998;11(4):286-294. doi:10.1002/(SICI)1098-1004(1998)11:4<286::AID-HUMU6>3.0.CO;2-B

Gallione CJ, Repetto GM, Legius E, et al. A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). *Lancet*. 2004;363(9412):852-859. doi:10.1016/S0140-6736(04)15732-2

Gallione CJ, Richards JA, Letteboer TG, et al. SMAD4 mutations found in unselected HHT patients. *J Med Genet*. 2006;43(10):793-797. doi:10.1136/jmg.2006.041517

Gallione CJ, Scheessele EA, Reinhardt D, et al. Two common endoglin mutations in families with hereditary hemorrhagic telangiectasia in the Netherlands Antilles: evidence for a founder effect. *Hum Genet*. 2000;107(1):40-44. doi:10.1007/s004390000326

Garcia-Tsao G. Liver involvement in hereditary hemorrhagic telangiectasia (HHT). *J Hepatol*. 2007;46(3):499-507. doi:10.1016/j.jhep.2006.12.008

Gedge F, McDonald J, Phansalkar A, et al. Clinical and analytical sensitivities in hereditary hemorrhagic telangiectasia testing and a report of de novo mutations. *J Mol Diagn*. 2007;9(2):258-265. doi:10.2353/jmoldx.2007.060117

Gougos A, Letarte M. Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J Biol Chem*. 1990;265(15):8361-8364.

Goumans MJ, Liu Z, ten Dijke P. TGF-beta signaling in vascular biology and dysfunction. *Cell Res*. 2009;19(1):116-127. doi:10.1038/cr.2008.326

Graven KK, Troxler RF, Kornfeld H, Panchenko MV, Farber HW. Regulation of endothelial cell glyceraldehyde-3-phosphate dehydrogenase expression by hypoxia. *J Biol Chem*. 1994;269(39):24446-24453.

Guo W, Liu H, Yan Y, et al. Targeting the TGF-β signaling pathway: an updated patent review (2021-present). *Expert Opin Ther Pat*. 2024;34(3):99-126. doi:10.1080/13543776.2024.2346325

Hammill AM, Wusik K, Kasthuri RS. Hereditary hemorrhagic telangiectasia (HHT): a practical guide to management. *Hematology Am Soc Hematol Educ Program*. 2021;2021(1):469-477. doi:10.1182/hematology.2021000281

Hata A, Chen YG. TGF-β Signaling from Receptors to Smads. *Cold Spring Harb Perspect Biol*. 2016;8(9):a022061. Published 2016 Sep 1. doi:10.1101/cshperspect.a022061

Hayama M, Maeda Y, Shikina T, et al. Validation of epistaxis severity score for hereditary hemorrhagic telangiectasia in Japan. *Auris Nasus Larynx*. 2022;49(3):415-420. doi:10.1016/j.anl.2021.11.005

Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 1997;390(6659):465-471. doi:10.1038/37284

Hernandez F, Huether R, Carter L, et al. Mutations in RASA1 and GDF2 identified in patients with clinical features of hereditary hemorrhagic telangiectasia. *Hum Genome Var*. 2015;2:15040. Published 2015 Nov 5. doi:10.1038/hgv.2015.40

Jelsig AM, Tørring PM, Kjeldsen AD, et al. JP-HHT phenotype in Danish patients with SMAD4 mutations. *Clin Genet*. 2016;90(1):55-62. doi:10.1111/cge.12693

Johnson DW, Berg JN, Baldwin MA, et al. Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. *Nat Genet*. 1996;13(2):189-195. doi:10.1038/ng0696-189

Kandpal RP, Rajasimha HK, Brooks MJ, et al. Transcriptome analysis using next generation sequencing reveals molecular signatures of diabetic retinopathy and efficacy of candidate drugs. *Mol Vis*. 2012;18:1123-1146.

Kim JH, Peacock MR, George SC, Hughes CC. BMP9 induces EphrinB2 expression in endothelial cells through an Alk1-BMPRII/ActRII-ID1/ID3-dependent pathway: implications for hereditary hemorrhagic telangiectasia type II. *Angiogenesis*. 2012;15(3):497-509. doi:10.1007/s10456-012-9277-x

Lagna G, Hata A, Hemmati-Brivanlou A, Massagué J. Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. *Nature*. 1996;383(6603):832-836. doi:10.1038/383832a0

Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062-D1067. doi:10.1093/nar/gkx1153

Le N, Li Y, Walker G, et al. Arteriovenous Malformations (AVMs): Molecular Pathogenesis, Clinical Features, and Emerging Therapeutic Strategies. *Biomolecules*. 2025;15(12):1661. Published 2025 Nov 27. doi:10.3390/biom15121661

Lesca G, Genin E, Blachier C, et al. Hereditary hemorrhagic telangiectasia: evidence for regional founder effects of ACVRL1 mutations in French and Italian patients. *Eur J Hum Genet*. 2008;16(6):742-749. doi:10.1038/ejhg.2008.3

Letteboer TG, Mager HJ, Snijder RJ, et al. Genotype-phenotype relationship for localization and age distribution of telangiectases in hereditary hemorrhagic telangiectasia. *Am J Med Genet A*. 2008;146A(21):2733-2739. doi:10.1002/ajmg.a.32243

Locke T, Sequeira Campos MB, Chen P. Hereditary Hemorrhagic Telangiectasia (HHT). In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; July 5, 2026.

López-Novoa JM, Bernabeu C. The physiological role of endoglin in the cardiovascular system. *Am J Physiol Heart Circ Physiol*. 2010;299(4):H959-H974. doi:10.1152/ajpheart.01251.2009

Lux A, Gallione CJ, Marchuk DA. Expression analysis of endoglin missense and truncation mutations: insights into protein structure and disease mechanisms. *Hum Mol Genet*. 2000;9(5):745-755. doi:10.1093/hmg/9.5.745

Mandlik JS, Patil AS, Singh S. Next-Generation Sequencing (NGS): Platforms and Applications. *J Pharm Bioallied Sci*. 2024;16(Suppl 1):S41-S45. doi:10.4103/jpbs.jpbs_838_23

McAllister KA, Grogg KM, Johnson DW, et al. Endoglin, a TGF-beta binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nat Genet*. 1994;8(4):345-351. doi:10.1038/ng1294-345

McAllister KA, Lennon F, Bowles-Biesecker B, et al. Genetic heterogeneity in hereditary haemorrhagic telangiectasia: possible correlation with clinical phenotype. *J Med Genet*. 1994;31(12):927-932. doi:10.1136/jmg.31.12.927

McDonald J, Bayrak-Toydemir P, DeMille D, Wooderchak-Donahue W, Whitehead K. Curaçao diagnostic criteria for hereditary hemorrhagic telangiectasia is highly predictive of a pathogenic variant in ENG or ACVRL1 (HHT1 and HHT2). *Genet Med*. 2020;22(7):1201-1205. doi:10.1038/s41436-020-0775-8

McDonald J, Kornish J, Stevenson DA, et al. Frequency of epistaxis and telangiectasia in patients with hereditary hemorrhagic telangiectasia (HHT) in comparison with the general population: Curaçao diagnostic criteria revisited. *Genet Med*. 2023;25(8):100865. doi:10.1016/j.gim.2023.100865

Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. 2010;11(1):31-46. doi:10.1038/nrg2626

Miyazono K, Kamiya Y, Morikawa M. Bone morphogenetic protein receptors and signal transduction. *J Biochem*. 2010;147(1):35-51. doi:10.1093/jb/mvp148

Muzzey D, Evans EA, Lieber C. Understanding the Basics of NGS: From Mechanism to Variant Calling. *Curr Genet Med Rep*. 2015;3(4):158-165. doi:10.1007/s40142-015-0076-8

Olivieri C, Pagella F, Semino L, et al. 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*. 2007;52(10):820-829. doi:10.1007/s10038-007-0187-5

Pagella F, Maiorano E, Matti E, et al. FID Score: an effective tool in Hereditary Haemorrhagic Telangiectasia - related epistaxis. *Rhinology*. 2020;58(5):516-521. doi:10.4193/Rhin20.078

Palermo M, Sturiale CL. Cerebrovascular Malformations Associated With Hereditary Hemorrhagic Telangiectasia and HHT-Like Syndromes: A Comparative Overview. *Eur J Neurol*. 2026;33(2):e70523. doi:10.1111/ene.70523

Pamonsinlapatham P, Hadj-Slimane R, Lepelletier Y, et al. p120-Ras GTPase activating protein (RasGAP): a multi-interacting protein in downstream signaling. *Biochimie*. 2009;91(3):320-328. doi:10.1016/j.biochi.2008.10.010

Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet*. 2009;10(10):669-680. doi:10.1038/nrg2641

Plumitallo S, Ruiz-Llorente L, Langa C, et al. Functional analysis of a novel ENG variant in a patient with hereditary hemorrhagic telangiectasia (HHT) identifies a new Sp1 binding-site. *Gene*. 2018;647:85-92. doi:10.1016/j.gene.2018.01.007

Revencu N, Boon LM, Mendola A, et al. RASA1 mutations and associated phenotypes in 68 families with capillary malformation-arteriovenous malformation. *Hum Mutat*. 2013;34(12):1632-1641. doi:10.1002/humu.22431

Ricard N, Bidart M, Mallet C, et al. Functional analysis of the BMP9 response of ALK1 mutants from HHT2 patients: a diagnostic tool for novel ACVRL1 mutations. *Blood*. 2010;116(9):1604-1612. doi:10.1182/blood-2010-03-276881

Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424. doi:10.1038/gim.2015.30

Richards-Yutz J, Grant K, Chao EC, Walther SE, Ganguly A. Update on molecular diagnosis of hereditary hemorrhagic telangiectasia. *Hum Genet*. 2010;128(1):61-77. doi:10.1007/s00439-010-0825-4

Rigelsky CM, Jennings C, Lehtonen R, Minai OA, Eng C, Aldred MA. BMPR2 mutation in a patient with pulmonary arterial hypertension and suspected hereditary hemorrhagic telangiectasia. *Am J Med Genet A*. 2008;146A(19):2551-2556. doi:10.1002/ajmg.a.32468

Ríus C, Smith JD, Almendro N, et al. Cloning of the promoter region of human endoglin, the target gene for hereditary hemorrhagic telangiectasia type 1. *Blood*. 1998;92(12):4677-4690.

Rossi E, Kauskot A, Saller F, et al. Endoglin Is an Endothelial Housekeeper against Inflammation: Insight in ECFC-Related Permeability through LIMK/Cofilin Pathway. *Int J Mol Sci*. 2021;22(16):8837. Published 2021 Aug 17. doi:10.3390/ijms22168837

Saito T, Bokhove M, Croci R, et al. Structural Basis of the Human Endoglin-BMP9 Interaction: Insights into BMP Signaling and HHT1. *Cell Rep*. 2017;19(9):1917-1928. doi:10.1016/j.celrep.2017.05.011

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74(12):5463-5467. doi:10.1073/pnas.74.12.5463

Satam H, Joshi K, Mangrolia U, et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology (Basel)*. 2023;12(7):997. Published 2023 Jul 13. doi:10.3390/biology12070997

Sbalchiero A, Abu Hweij Y, Mazza T, et al. 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. *Mol Genet Genomic Med*. 2022;10(8):e1972. doi:10.1002/mgg3.1972

Scharpfenecker M, van Dinther M, Liu Z, et al. BMP-9 signals via ALK1 and inhibits bFGF-induced endothelial cell proliferation and VEGF-stimulated angiogenesis. *J Cell Sci*. 2007;120(Pt 6):964-972. doi:10.1242/jcs.002949

Scotti C, Olivieri C, Boeri L, et al. Bioinformatic analysis of pathogenic missense mutations of activin receptor like kinase 1 ectodomain. *PLoS One*. 2011;6(10):e26431. doi:10.1371/journal.pone.0026431

Shovlin CL, Guttmacher AE, Buscarini E, et al. Diagnostic criteria for hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber syndrome). *Am J Med Genet*. 2000;91(1):66-67. doi:10.1002/(sici)1096-8628(20000306)91:1<66::aid-ajmg12>3.0.co;2-p

Shovlin CL, Hughes JM, Scott J, Seidman CE, Seidman JG. Characterization of endoglin and identification of novel mutations in hereditary hemorrhagic telangiectasia. *Am J Hum Genet*. 1997;61(1):68-79. doi:10.1086/513906

Shovlin CL, Simeoni I, Downes K, et al. Mutational and phenotypic characterization of hereditary hemorrhagic telangiectasia. *Blood*. 2020;136(17):1907-1918. doi:10.1182/blood.2019004560

Slatko BE, Gardner AF, Ausubel FM. Overview of Next-Generation Sequencing Technologies. *Curr Protoc Mol Biol*. 2018;122(1):e59. doi:10.1002/cpmb.59

Snellings DA, Gallione CJ, Clark DS, Vozoris NT, Faughnan ME, Marchuk DA. Somatic Mutations in Vascular Malformations of Hereditary Hemorrhagic Telangiectasia Result in Bi-allelic Loss of ENG or ACVRL1. *Am J Hum Genet*. 2019;105(5):894-906. doi:10.1016/j.ajhg.2019.09.010

Soukarieh O, Tillet E, Proust C, et al. uAUG creating variants in the 5'UTR of ENG causing Hereditary Hemorrhagic Telangiectasia. *NPJ Genom Med*. 2023;8(1):32. Published 2023 Oct 17. doi:10.1038/s41525-023-00378-5

Stamova BS, Apperson M, Walker WL, et al. Identification and validation of suitable endogenous reference genes for gene expression studies in human peripheral blood. *BMC Med Genomics*. 2009;2:49. Published 2009 Aug 5. doi:10.1186/1755-8794-2-49

Tillet E, Bailly S. Emerging roles of BMP9 and BMP10 in hereditary hemorrhagic telangiectasia. *Front Genet*. 2015;5:456. Published 2015 Jan 8. doi:10.3389/fgene.2014.00456

Trembath RC, Thomson JR, Machado RD, et al. Clinical and molecular genetic features of pulmonary hypertension in patients with hereditary hemorrhagic telangiectasia. *N Engl J Med*. 2001;345(5):325-334. doi:10.1056/NEJM200108023450503

Vander Ark A, Cao J, Li X. TGF-β receptors: In and beyond TGF-β signaling. *Cell Signal*. 2018;52:112-120. doi:10.1016/j.cellsig.2018.09.002

Venkatesha S, Toporsian M, Lam C, et al. Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nat Med*. 2006;12(6):642-649. doi:10.1038/nm1429

Whitehead KJ, Toydemir D, Wooderchak-Donahue W, et al. Investigation of the Genetic Determinants of Telangiectasia and Solid Organ Arteriovenous Malformation Formation in Hereditary Hemorrhagic Telangiectasia (HHT). *Int J Mol Sci*. 2024;25(14):7682. Published 2024 Jul 12. doi:10.3390/ijms25147682

Wooderchak-Donahue WL, Akay G, Whitehead K, et al. Phenotype of CM-AVM2 caused by variants in EPHB4: how much overlap with hereditary hemorrhagic telangiectasia (HHT)?. *Genet Med*. 2019;21(9):2007-2014. doi:10.1038/s41436-019-0443-z

Wooderchak-Donahue WL, McDonald J, O'Fallon B, et al. BMP9 mutations cause a vascular-anomaly syndrome with phenotypic overlap with hereditary hemorrhagic telangiectasia. *Am J Hum Genet*. 2013;93(3):530-537. doi:10.1016/j.ajhg.2013.07.004

Yu Q, Shen XH, Li Y, Li RJ, Li J, Luo YY, Liu SF, Deng MY, Pei MF, Zhang GS. An intron mutation in the ACVRL1 may be associated with a transcriptional regulation defect in a Chinese family with hereditary hemorrhagic telangiectasia. PLoS One. 2013;8(2):e58031. doi: 10.1371/journal.pone.0058031. Epub 2013 Feb 27. PMID: 23460919; PMCID: PMC3584037.

Zbuk KM, Eng C. Hamartomatous polyposis syndromes. *Nat Clin Pract Gastroenterol Hepatol*. 2007;4(9):492-502. doi:10.1038/ncpgasthep0902

Zi Z. Molecular Engineering of the TGF-β Signaling Pathway. *J Mol Biol*. 2019;431(15):2644-2654. doi:10.1016/j.jmb.2019.05.022