



UNIVERSITÀ DEGLI  
STUDI DEL MOLISE

Dipartimento di Medicina e di Scienze Della Salute  
“Vicenzo Tiberio”

---

Corso di Dottorato in Medicina Traslazionale e Clinica  
(XXXI ciclo)

*SVILUPPO DI UN PERCORSO DIAGNOSTICO  
MEDIANTE TECNOLOGIA NGS PER LA  
CARATTERIZZAZIONE MOLECOLARE DI PAZIENTI  
ATASSO SPASTICI E PAZIENTI CON FENOTIPI  
NEUROLOGICI COMPLESSI.*

**RELATORE**

Prof Giuseppe LEMBO  
S.S.D. Med/50

**COORDINATORE**

Prof Ciro COSTAGLIOLA  
S.S.D Med/30

**CORRELATORE**

Dott. Stefano GAMBARDELLA

**CANDIDATO**

Dott.ssa Rosa CAMPOPIANO  
Matr.: 155938

---

ANNO ACCADEMICO 2017-2018

# INDICE

<b>SINOSSI</b>	Pag.1
<b>1. SEQUENZIAMENTO DI NUOVA GENERAZIONE</b>	Pag.4
1.1 <i>Avvento della tecnologia di nuova generazione: piattaforme e principi</i>	Pag.4
1.2 <i>Analisi dei dati</i>	Pag.6
1.3 <i>Applicazione delle tecnologie NGS</i>	Pag.8
<b>2. CARATTERISTICHE CLINICHE DI PAZIENTI CON ATASSIA E SPASTICITÀ</b>	Pag.11
2.1 <i>Pazienti affetti da Atassia</i>	Pag.11
2.2 <i>Pazienti affetti da Paraparesi spastica</i>	Pag.15
<b>3. MATERIALI E METODI</b>	Pag.19
3.1 <i>Pazienti</i>	Pag.19
3.2 <i>Pannello NGS con geni selezionati</i>	Pag.20
3.2.1 <i>Selezione dei geni</i>	Pag.20
3.2.2 <i>Preparazione delle librerie e sequenziamento</i>	Pag.21
3.2.3 <i>Analisi dei dati</i>	Pag.22
3.2.4 <i>Validazione pannello</i>	Pag.23
3.3 <i>Pannello esoma clinico</i>	Pag.26
3.3.1 <i>Preparazione delle librerie e sequenziamento</i>	Pag.26
3.3.2 <i>Analisi dei dati</i>	Pag.27
3.3.3 <i>Validazione del pannello</i>	Pag.28
<b>4. RISULTATI</b>	Pag.30
4.1 <i>Casistica e saggio NGS target</i>	Pag.30
4.1.1 <i>Varianti e geni</i>	Pag.34
4.1.2 <i>Pazienti in cui sono state identificate una o più varianti</i>	Pag.37
4.1.3 <i>Descrizione genetica e clinica dei pazienti</i>	Pag.38
→ <i>Varianti identificate in geni con modalità di trasmissione autosomico dominante (SPG4)</i>	Pag.38
→ <i>Varianti identificate in geni con modalità di trasmissione autosomico recessivo (SPG11, SPG7)</i>	Pag.46
→ <i>Varianti identificate in geni rari</i>	Pag.53
4.2 <i>Esoma clinico</i>	Pag.67
4.2.1 <i>Descrizione genetica e clinica dei pazienti</i>	Pag.68
<b>5. DISCUSSIONE</b>	Pag.77
<b>6. CONCLUSIONI</b>	Pag.85
<b>BIBLIOGRAFIA</b>	Pag.87

## SINOSI

Le malattie neurologiche e neurodegenerative sono un gruppo eterogeneo di patologie caratterizzate, in molti casi, da una forte eterogeneità clinica (età di esordio, progressione, sintomatologia) e genetica. Ne sono un esempio le atassie e le paraparesi spastiche, in cui la componente genetica risulta in molti casi di difficile inquadramento. Se da un lato in molti di questi pazienti non si riesce ad identificare tale componente, altri invece ricevono una diagnosi clinica nella quale viene identificato il gene responsabile del fenotipo. L'identificazione di questi geni ha contribuito nel tempo ad una migliore comprensione dei pathway molecolari maggiormente coinvolti.

Ad oggi sono note più di 60 forme di atassie geneticamente determinate, e circa 50 forme di paraparesi spastiche ereditarie. In tali pazienti, l'analisi genetica deve anche tener conto della penetranza incompleta, e della presenza di mutazioni "private", ovvero descritte solamente in alcune famiglie. Oltretutto, la difficoltà nell'eseguire studi multicentrici con pazienti selezionati in modo omogeneo, rende difficile sviluppare dei database specifici in cui siano riportate sia le varianti geniche che il fenotipo clinico di cui sono responsabili.

Le tecnologie di tipo NGS (Next Generation Sequencing) hanno risolto la problematica della "poligenicità", rendendo possibile l'analisi contemporanea di più geni, consentendo anche l'analisi di tutto il genoma. Ad oggi, pannelli accuratamente disegnati e validati sono entrati a far parte della pratica clinica.

L'applicazione di queste tecnologie nelle patologie neurologiche è più complessa perché richiede una accurata selezione dei geni da analizzare. Inoltre, considerata l'elevata eterogeneità allelica, risulta difficile stabilire una corretta detection rate (% dei positivi identificati), e quindi sviluppare dei pannelli di facile utilizzo in diagnostica molecolare.

Da queste osservazioni è nata la necessità di sviluppare delle soluzioni diagnostiche basate su tecnologia NGS per l'analisi di fenotipi atassici e/o spastici e di fenotipi neurologici più complessi in cui la spasticità e l'atassia sono solo una delle componenti del quadro sindromico.

Infatti, anche se il sequenziamento massivo (NGS) è ormai entrato nella pratica clinica, la sua applicazione nelle patologie neurologiche risulta ancora complesso e necessita lo sviluppo di saggi ad-hoc che consentano una elevata detection rate, con costi e tempi contenuti.

Alla luce di queste osservazioni, la tesi ha avuto i seguenti obiettivi:

- Disegnare, sviluppare e validare un pannello di geni basato su tecnologia NGS (Pannello NGS con 105 geni selezionati) per la diagnosi di pazienti caratterizzati da atassia e/o spasticità
- Valutare l'utilità dell'analisi NGS di tutti i geni codificanti associati a patologie mendeliane (Esoma clinico, circa 4500 geni) per un corretto inquadramento di fenotipi neurologici complessi.

Lo studio è stato condotto su 78 pazienti atassici e/o spastici reclutati presso l'IRCCS Neuromed. È stata identificata la componente genetica potenzialmente responsabile del fenotipo clinico nel 21% dei pazienti sottoposti ad analisi mediante Pannello NGS con geni selezionati, e nel 13% dei pazienti sottoposti ad esoma clinico.

I dati di questo studio dimostrano che pannelli NGS, accuratamente disegnati e validati, possano essere applicati alla diagnosi molecolare di patologie neurologiche caratterizzate da poligenicità e variabilità clinica. La validazione di questi percorsi richiede la selezione di una casistica accurata ed omogenea. Questo consente di delinearne il corretto utilizzo diagnostico, aumentando così la detection rate.

Risulta inoltre fondamentale l'utilizzo di un iter di interpretazione delle varianti molto stringente e riproducibile. Questo viene confermato nel presente studio, dove un corretto utilizzo congiunto di analisi di segregazione familiare (ove possibile reperire i familiari del paziente affetto) ed analisi bioinformatica, è fondamentale per l'identificazione delle varianti patologiche, soprattutto in geni per i quali sono disponibili dati limitati nei database pubblici.

Risulta inoltre utile applicare pannelli più estesi, come l'esoma clinico, in casi complessi di difficile inquadramento. In questi casi, la diagnosi molecolare si è rivelata, insieme a quella clinica, fondamentale per un corretto inquadramento diagnostico.

# 1. SEQUENZIAMENTO DI NUOVA GENERAZIONE (NGS)

## ***1.1 Avvento della tecnologia di nuova generazione: piattaforme e principi***

Dalla scoperta della struttura del DNA ad oggi si è osservato un notevole incremento delle conoscenze riguardo la sua variabilità ed il relativo impatto sul fenotipo. Nell'ultimo decennio, il progresso scientifico e tecnologico ha aperto le porte ad un nuovo tipo di sequenziamento, chiamato sequenziamento di seconda generazione o "Next Generation Sequencing" (NGS), che ha offerto tanti vantaggi rispetto al classico sequenziamento Sanger. Infatti, queste nuove tecnologie hanno la capacità di fornire milioni di sequenze a prezzi relativamente contenuti, stravolgendo completamente le conoscenze della genomica ed il suo utilizzo nella pratica clinica.

Ad oggi, sono stati sviluppati diversi sistemi di sequenziamento, che differiscono principalmente per la chimica utilizzata e per la tecnica di amplificazione clonale. In funzione di queste caratteristiche, i sistemi si differenziano per la lunghezza delle reads (letture o sequenze) prodotte, i tempi di corsa, la quantità di basi prodotta, e l'accuratezza dei dati grezzi.

La prima piattaforma ad essere immessa nel mercato fu la piattaforma 454 dalla Life Sciences-Roche nel 2005. Tale piattaforma prevede l'amplificazione clonale dei frammenti mediante emulsion PCR. Tale tecnologia consiste nella preparazione di un'emulsione di acqua e olio in cui le microbolle d'acqua funzionano da microreattori per le reazioni di PCR. Nella soluzione acquosa vengono dispensati tutti i reagenti per la reazione di amplificazione unitamente a delle biglie di supporto (beads) sulle quali si legano i frammenti di DNA. In base alla legge di distribuzione di Poisson quando i reagenti sono presenti in quantità bilanciate tra loro e rispetto ai volumi di acqua e olio si

distribuiranno in maniera tale che ogni microreattore contenga un frammento di DNA, la polimerasi, i reagenti necessari per l'amplificazione e una biglia di supporto. I microreattori vengono quindi sottoposti a variazioni cicliche di temperatura e ciò permetterà la formazione di migliaia di frammenti clonali che andranno ad aderire sulla superficie delle biglie grazie a specifiche corte sequenze complementari. La fase di sequenziamento viene effettuata col metodo del pirosequenziamento. Tale metodo si basa sulla determinazione della presenza del pirofosfato. Il pirofosfato, che funge da substrato per una luciferasi, viene rilasciato nel momento in cui si forma un nuovo legame fosfodiesterico e il rilascio determina una emissione di energia luminosa che viene rilevata da un sensore. Non vi è una differenza di segnale immesso tra i quattro nucleotidi ma questi vengono immessi nel sequenziatore in flussi separati seguiti da cicli di lavaggio per eliminare quello non incorporato. La piattaforma 454 si differenzia dalle altre piattaforme NGS per l'elevata velocità di processamento e per la maggiore lunghezza delle reads prodotte, 700 bp contro circa 300 bp delle altre piattaforme.

In seguito sono nate altre soluzioni. Ad esempio, il sistema Illumina è basato sull'utilizzo di terminatori fluorescenti reversibili. In questo caso l'amplificazione clonale dei frammenti di DNA avviene tramite la "bridge PCR". Ai frammenti di DNA in primo luogo vengono unite delle corte sequenze specifiche dette "adattatori". I frammenti uniti agli adattatori vengono poi distribuiti su un supporto rigido a cui si legano grazie alla presenza di sequenze complementari agli adattatori, ancorate alla superficie. Ciascun frammento viene poi amplificato in modo clonale con la conseguente formazione dei "cluster", ovvero microaree circoscritte del supporto nelle quali sono presenti i frammenti clonati da un unico template di partenza. Il sequenziamento si basa sull'utilizzo di nucleotidi con gruppi terminatori fluorescenti reversibili. Ad ogni ciclo (flows) si ha l'incorporazione di uno dei dNTP fluorescenti con conseguente emissione di fluorescenza che viene rilevata e convertita dal

sistema in immagini globali dell'intera superficie del vetrino di supporto. In seguito, avviene il rilascio per via chimica dei gruppi terminatori e ciò consente la prosecuzione del processo, con l'incorporamento di ulteriori nucleotidi e l'allungamento della sequenza.

Un'altra tipologia di sequenziamento è lo Ion Torrent, caratterizzato da fasi di preparazione e di sequenziamento concettualmente simili al pirosequenziamento Roche / 454. L'emulsion-PCR viene utilizzata per amplificare clonalmente i frammenti di DNA ligati agli adattatori presenti sulla superficie delle biglie. Le biglie vengono successivamente distribuite nei micropozzetti dove si verifica una reazione di sequenziamento per sintesi. A differenza del pirosequenziamento, che costituisce l'integrazione base delle coppie con la produzione di luce emessa da una luciferasi, il sequenziamento dei semiconduttori di Ion Torrent misura i cambiamenti di pH indotti dal rilascio di ioni idrogeno durante l'estensione del DNA. Questi cambiamenti di pH sono rilevati da un sensore posizionato nella parte inferiore del micropozzetto e convertiti in un segnale di tensione. Tale segnale è proporzionale al numero di basi incorporate e la discriminazione di base è permessa grazie all'aggiunta sequenziale di singoli nucleotidi durante ciascun ciclo di sequenziamento. A differenza delle piattaforme precedentemente descritte non si ha la scansione ottica per distinguere i nucleotidi durante i cicli di sequenziamento e ciò accelera notevolmente il processo e ne riduce i costi. Tuttavia, tale tecnologia ha lo svantaggio di generare sequenze più corte 200bp, contro le 400bp che si riescono ad ottenere con la piattaforma Illumina.

## ***1.2 Analisi dei dati***

Le piattaforme NGS producono un enorme mole di dati, nell'ordine delle Terabyte (TB). Questo causa una grande difficoltà nella conservazione, e nella successiva analisi dei dati, le cui tecniche richiedono continue implementazioni di software specifici. Anche se i software di analisi sono differenti a seconda della tecnologia NGS

utilizzata, le operazioni che svolgono sono le stesse. Infatti in questi programmi viene dapprima eseguito il processo di “base calling”, ovvero la conversione di immagini di luminescenza o fluorescenza acquisite in sequenze nucleotidiche (“reads”). A tale processo segue l’assegnazione del “quality score”, ovvero l’attribuzione di parametri per la valutazione della qualità e quindi la probabilità di errore nella lettura del nucleotide. Le reads che hanno superato questo step vengono allineate col genoma di riferimento. I software d’analisi annotano le differenze che riscontrano tra le reads ottenute dall’esperimento ed il genoma di riferimento sul quale le hanno allineate. Per poter essere allineate, le reads devono essere superiori ai 30 bp poiché soltanto il 90% del genoma umano può essere allineato in modo univoco con reads di 30 bp. Le sequenze ripetute rappresentano un grosso limite della fase di allineamento in quanto non sono presenti in maniera univoca nel genoma di riferimento e quindi possono essere collocate in più posizioni. Errori nella fase di allineamento generano i cosiddetti “gap”, ovvero regioni in cui il software non è riuscito ad allineare alcuna reads. I tassi di errore, pur sembrando maggiori nella tecnologia NGS rispetto alla metodica Sanger, vengono notevolmente abbattuti da una lettura ripetuta e massiva di ogni frammento (“coverage”). Il coverage rappresenta quindi il numero di volte che un nucleotide è stato “letto” durante il sequenziamento ( $\text{Coverage} = \frac{\text{Throughput}}{\text{Regione di interesse}}$ ). Quante più volte la regione di interesse sarà stata letta tanto più bassa è la probabilità che ci siano stati errori durante il sequenziamento. È stato stimato che un coverage di 20X è sufficiente per ritenere valida la lettura.

Una volta considerati tutti i parametri qualitativi si può passare a quella che è la fase dell’analisi dei dati ovvero l’interpretazione delle varianti annotate. L’interpretazione può variare a seconda di quello che è lo scopo dell’analisi. Nella genetica molecolare i filtri più comunemente utilizzati per l’identificazione di varianti patogenetiche sono la rimozione di tutte le varianti frequenti nella popolazione generale le cui frequenze vengono ricercate in banche

dati pubbliche quali dbSNP, 1000 genomes o ExAC; rimozione o meno di varianti già note in letteratura per essere o meno correlate a quadri patologici mediante la consultazione di banche dati biomediche quali ad esempio OMIM o PubMed; utilizzo di strumenti bioinformatici per analisi di predizione in silico quali ad esempio SIFT, POLYPHEN-2, Mutation taster. La gestione e l'interpretazione dei dati richiede quindi figure professionali altamente specializzate.

### ***1.3 Applicazione delle tecnologie NGS***

I sequenziatori di nuova generazione sono diventati strumenti d'elezione per analisi approfondite nel campo della biologia e della medicina. Questi strumenti consentono di approcciarsi ad una serie di applicazioni come sequenziamento de novo, resequencing, studi di metagenomica, di epigenomica e sequenziamento del trascrittoma.

Il sequenziamento de novo si propone di determinare genomi di organismi non ancora sequenziati, per i quali quindi non si dispone di genomi di riferimento. Di contro, il risequenziamento di un intero genoma già sequenziato può essere analizzato allo scopo di identificare mutazioni di un singolo nucleotide (SNP), inserzioni e delezioni di sequenze più o meno lunghe in determinate posizioni del genoma e variazioni nel numero di copie di determinati geni (CNV).

Ad oggi, molti approcci NGS soprattutto in diagnostica molecolare, analizzano selettivamente determinate regioni genomiche o un gruppo selezionato di geni (targeted resequencing) associati ad una particolare condizione clinica. Altri approcci prendono in considerazione l'esoma, ovvero l'insieme di tutte le porzioni del genoma che "codificano" per proteine. Pur rappresentando una quota vicina all'1% dell'intero genoma, è all'interno dell'esoma che avviene oltre l'85% delle mutazioni oggi note come clinicamente rilevanti. Le tecniche di sequenziamento di nuova generazione (NGS) consentono di analizzare le sequenze codificanti degli oltre

20,000 geni di un individuo in un unico test, in un approccio meglio conosciuto come whole exome-sequencing (WES). L'analisi dell'esoma è dunque uno strumento diagnostico potentissimo in grado di fornire informazioni di interesse clinico di livello e rilevanza impensabili con i tradizionali strumenti della genetica clinica. L'analisi dell'esoma consta nel sequenziamento dei soli geni nell'interezza di tutti i loro esoni, delle parti introniche che confinano con gli esoni e talvolta delle regioni regolatorie a monte e a valle del gene (dette rispettivamente 5'-UTR e 3'-UTR, lunghe da alcune centinaia ad alcune migliaia di nucleotidi), allo scopo di identificare mutazioni esoniche (missenso, nonsense, frame shifting, in frame o regolatorie) o mutazioni introniche che impattino il processo di formazione del messaggero maturo (mutazioni di splicing).

Una versione più contenuta WES è meglio conosciuta come esoma clinico (CES). Infatti, sebbene siano oltre 20 mila i geni che in ciascun individuo vengono espressi in proteine o RNA funzionali, sono circa 4800 quelli le cui mutazioni sono ad oggi state associate con condizioni patologiche ("geni-malattia", si veda la banca dati OMIM). L'analisi CES trae beneficio da queste recenti evidenze, ed è concepito in modo da assicurare la massima copertura e la maggior sensibilità analitica possibile nella valutazione dei circa 4800 geni-malattia attualmente identificati.

L'analisi del genoma intero (Whole-Genome Sequencing – WGS) è invece molto più esteso rispetto all'analisi dell'esoma. Proprio per questo, è utilizzato per l'identificazione di nuovi geni-malattia, ma è ancora difficile un suo utilizzo diretto per la diagnostica molecolare. Una analisi WGS consta nel sequenziamento dell'intero genoma, cioè di tutto quanto il DNA contenuto nel nucleo cellulare (3 miliardi di nucleotidi). Con questa metodica il patrimonio genetico di un individuo viene completamente sequenziato, sia nelle sue regioni codificanti che in quelle non codificanti. Come tale, dunque, il sequenziamento del genoma include il sequenziamento dell'intero esoma (cioè di tutte le regioni codificanti di ogni gene, dette esoni) e di tutte le regioni intergeniche, cioè le sequenze interposte fra un

gene e l'altro e che, da un punto di vista quantitativo, rappresentano la strangrande maggioranza del DNA (ben il 98%).

Oltre questi approcci di sequenziamento, meglio conosciuti ed utilizzati in diagnostica molecolare, le tecnologie NGS hanno permesso l'avvento su larga scala di:

- meta-genomica, cioè lo studio dell'insieme del materiale genetico di tutti gli organismi viventi di un determinato campione ambientale.
- epigenetica, ovvero lo studio dei meccanismi di regolazione dell'espressione genica attraverso, ad esempio, una tecnica di immunoprecipitazione della cromatica (ChIP-Seq)
- RNA-Seq. il Sequenziamento del trascrittoma o dei microRNA (piccoli RNA non codificanti ad attività regolativa), utilizzato per lo studio delle funzioni biologiche di una cellula, in quanto fornisce informazioni quantitative sulle differenze nei livelli di espressione dei geni in tessuti diversi o in uno stesso tessuto analizzato in diverse condizioni sperimentali o in differenti stadi di sviluppo.

## 2. CARATTERISTICHE CLINICHE DI PAZIENTI CON ATASSIA E SPASTICITÀ

### ***2.1 Pazienti affetti da Atassia***

Il termine “atassia” deriva dal greco e significa “assenza di ordine”, clinicamente sotto tale termine si raggruppano una serie di patologie neurologiche caratterizzate da disturbi dell’equilibrio della coordinazione e del linguaggio dovute ad alterazioni del cervelletto e/o delle sue connessioni con le strutture spinali, pontine e corticali. La manifestazione fenotipica è estremamente eterogenea. Oltre ai disturbi atassici propriamente detti, disturbi dell’equilibrio, difficoltà nel mantenimento della posizione eretta, mancanza di coordinazione della marcia e degli arti, tremore cinetico, si devono aggiungere altre manifestazioni cliniche quali nistagmo, dismetria saccadica, alterazioni dei riflessi oculo-vestibolari, oscillazioni saccadiche, disartria, disfagia, disgrafia, disfunzioni sfinteriali, disturbi del sonno, facile affaticamento, depressione e deficit cognitivi. La classificazione su base fenotipica è complicata in quanto tali sintomi possono presentarsi parzialmente e con gravità variabile non solo nelle diverse forme di atassia ma anche all’interno dello stesso sottogruppo. Su base eziologica le atassie possono essere suddivise in forme ereditarie e forme sporadiche. Le forme sporadiche tendono a manifestarsi dopo i 50 anni con esordio acuto o subacuto. Le possibili cause delle forme sporadiche sono molteplici, tra cui l’uso cronico di alcol o droghe, esposizione ad agenti tossici, anomalie congenite, infiammazione immuno-mediata, carenza di vitamine, deposizione leptomeningea cronica di ferro, infezione cronica del sistema nervoso centrale (SNC), lesioni neoplastiche e danni vascolari. La determinazione dell’eziologia e quindi la diagnosi delle forme sporadiche può essere effettuata mediante un accurato esame neurologico studi di neuroimaging ed esami di biochimica clinica. Le forme genetiche hanno, in genere, un esordio graduale tra i 30 e i 50 anni anche se sono note forme con esordio infantile e/o

adolescenziale. In base al meccanismo di ereditarietà le forme ereditarie possono essere a loro volte suddivise in autosomico dominante, autosomico recessivo, legate all'X e mitocondriali.

Le atassie autosomico dominante sono anche dette atassie spinocerebellari (SCA) e si manifestano principalmente con atassia cerebellare progressiva.

Sono caratterizzate da una ampia eterogeneità clinica e genetica. Tipicamente l'esordio avviene tra la 3°-4° decade; raramente sono descritti casi con esordio più precoce.

Secondo la classificazione di Harding possono essere suddivise in tre sottogruppi:

- atassie autosomico dominante di tipo I in cui l'atassia si associa a disturbi neurologici extracerebellari (segni piramidali ed extrapiramidali e neuropatia).
- atassie autosomico dominante di tipo II in cui l'atassia si associa a maculopatia, possono inoltre presentare sintomi delle atassie di tipo I.
- atassie autosomico dominante di tipo III in cui si manifesta solo atassia.

Tuttavia nella maggior parte dei casi non è possibile applicare una classificazione così netta solo sulla base della clinica, la maggior parte dei soggetti con SCA presenta un quadro clinico sovrapponibile, consistente all'esordio in disturbi della deambulazione con andatura atassica lievemente progressiva, disartria e incoordinazione motoria e ciò ne rende difficile la classificazione.

Ad oggi sono state individuate oltre 30 forme d'atassia dominante del tipo SCA geneticamente e clinicamente distinte. Alcune di queste forme sono dette atassie da espansione in quanto la mutazione genetica è data dall'espansione di alcuni nucleotidi all'interno del gene (ad esempio SCA1 data dall'espansione della tripletta CAG nella regione codificante del gene ATXN1). In questo tipo di SCA l'età di esordio è inversamente correlata al numero di ripetizioni, tuttavia non è possibile stabilire l'età di insorgenza in

funzione della lunghezza dell'espansione. Inoltre essendo tratti genici ad elevata instabilità l'espansione tende ad aumentare da una generazione all'altra assistendo, quindi, al fenomeno dell'anticipazione. Solo negli ultimi anni sono state identificate forme di SCA dovute a mutazioni puntiformi ma grazie all'avvento delle nuove tecniche di sequenziamento questo gruppo SCA è in continua crescita. Tale forma risulta essere molto meno comune rispetto a quelle da espansione, spesso si tratta di mutazioni private riscontrate in poche famiglie, e per questo motivo non è possibile, attualmente effettuare una correlazione tra fenotipo e genotipo. Inoltre la maggior parte dei geni responsabili sono stati identificati recentemente e pertanto il loro meccanismo patogenetico rimane ancora poco chiaro. La prevalenza globale di queste atassie è stimata in circa 3/100000 tuttavia tale dato varia a seconda dell'etnia e dell'aria geografica considerata. In Italia, le forme più frequenti sono SCA1 e SCA2 che rappresentano circa il 50% dei casi seguite da SCA3, SCA7, SCA17, SCA12, SCA8, SCA15, SCA28 che sono invece identificate circa nel 5% dei casi.

Le atassie spinocerebellari autosomiche recessive (ARCA) sono anch'esse di difficile di classificazione a causa dell'elevata eterogeneità clinica. Il fenotipo varia da sindromi cerebellari a neuropatie sensorio-motorie, disturbi oftalmologici, movimenti involontari, convulsioni, anomalie cutanee e scheletriche. Da un punto di vista patogenetico possono essere suddivise in forme da disturbi mitocondriali, forme date da difetti riparazione del DNA, da difetti del trasporto, ripiegamento e degradazione delle proteine, forme metaboliche e da canalopatie.

La maggior parte hanno un esordio precoce prima dei 20 anni, tuttavia ne sono note forme ad esordio in età adulta. Sono principalmente caratterizzate da atassia spinocerebellare, che coinvolge cervelletto, tronco encefalico e fasci spinocerebellari. Fenotipicamente si manifestano con scarso equilibrio, perdita della coordinazione negli arti, tremore e disartria generalmente associate a neuropatia periferica.

L'Atassia di Friedreich (FRDA) è la forma più comune ed in generale è anche la forma di atassia ereditaria più comune con una prevalenza stimata di 1/30000-50000 e la frequenza di portatori sani di 1/85. Nella maggior parte dei casi è causata dall'espansione in omozigosi di tripletta GAA nel primo introne del gene della frataxina (FXN) ma in piccola percentuale (tra il 2 -4%) la FRDA può essere causata da mutazioni puntiformi nel gene FXN. Clinicamente è caratterizzata da atassia degli arti e della marcia associata a disartria alterazioni della sensibilità vibratoria e propriocettiva, areflessia anomalie dei movimenti oculari, ipostenia piramidale in un quadro multisistemico dove spesso compaiono cardiomiopatia ipertrofica, diabete mellito e ipoacusia. Numerose sono però le ARCA che presentano un quadro clinico sovrapponibile e pertanto che vanno in diagnosi differenziale quali ad esempio l'atassia cerebellare ad esordio precoce con riflessi tendinei conservati (EOCARR). Le forme metaboliche includono forme con atassia progressiva e forme con atassia intermittente in generale si tratta di forme sindromiche in cui l'atassia è uno dei tanti sintomi. Esempi di queste forme sono la leucodistrofia metacromatica e l'adrenoleucodistrofia. Tra le forme date da difetti di riparazione del DNA, l'atassia-teleangiectasia è la forma a maggior prevalenza (1/100000) ed è dovuta a mutazioni nel gene *ATM*. Tipicamente si manifesta con teleangiectasie cutanee e oculari ed aprassia oculomotoria, l'atrofia cerebellare spesso è poco evidente agli esordi e più marcata col progredire della malattia.

Tra le atassie legate al cromosoma X, la più comune è la sindrome dell'X-fragile associata a tremore/atassia (FXTAS), causata da espansione di triplette nucleotidiche CCG nel gene *FMR1*. L'esordio della malattia avviene dopo i 50 anni e in genere si manifesta con tremore, parkinsonismo, disfunzioni autonome, polineuropatia e deficit cognitivi. Nelle donne la manifestazione è più lieve e maggiormente variabile.

Le forme sopra elencate sono solo alcune delle forme più frequenti di atassia geneticamente determinate. Ad oggi, la banca dati Gene Nomenclature Committee (HGNC) riporta 62 forme di atassia, e tale numero è in continua evoluzione.

Le tecnologie NGS stanno “avvicinando” il complesso spettro delle patologie atassiche alla diagnostica molecolare, risolvendo il problema della “poligenicità”, ovvero permettendo l’analisi contemporanea di più geni tutti potenzialmente coinvolti o responsabili del quadro clinico osservato. Infatti, pannelli multigenici accuratamente disegnati e validati che tengono conto sia delle forme frequenti che quelle più rare, spesso identificate anche in una sola famiglia, stanno entrando a far parte della pratica clinica.

## ***2.2 Pazienti affetti da Paraparesi spastica***

La malattia di Strümpell-Lorrain è una patologia neurodegenerativa dovuta ad una degenerazione delle fibre dei neuroni motori. Prende il nome dal neurologo tedesco Adolf Strümpell che la identificò agli inizi del 1900 e il medico francese, Maurice Lorrain, che ne fornì una descrizione più approfondita dei sintomi.

Ad oggi tale patologia viene più comunemente identificata con due sigle FSP (Familial Spastic Paraparesis) e HSP (Hereditary Spastic Paraplegia). Le HSP sono un gruppo di malattie neurodegenerative ereditarie clinicamente eterogenee il cui sintomo cardine è la spasticità bilaterale agli arti inferiori è legato ad una assonopatia distale coinvolgente il pattern motorio. Tale sintomatologia è dovuta al graduale deterioramento dei motoneuroni superiori dei tratti corticospinali discendenti e delle fibre ascendenti della corda spinale. Il processo di deterioramento inizia all’estremità dell’assone e procede in senso retrogrado verso il soma cellulare, in un processo che è stato denominato “dying back axonal degeneration”. In funzione della clinica, le HSP vengono suddivise in due grandi gruppi, le forme pure in cui la spasticità è l’unico sintomo clinico e forme complicate in cui oltre alla spasticità si

riscontrano sintomi secondari quali incontinenza urinaria, alterazione della sensibilità e piede cavo. Il 70% dei casi di HSP sono date da forme a trasmissione autosomica dominante, mentre le forme autosomiche recessive rappresentano circa il 29% dei casi e in genere presentano un'insorgenza più precoce e un decorso più grave, il restante 1% è rappresentato da varianti legate al cromosoma X.

La forma più comune di HSP a trasmissione autosomica dominante è SPG4 ed è data da mutazioni all'interno del gene *SPAST*. Mutazioni all'interno del gene *SPAST* vengono riscontrate circa nel 50% dei casi con familiarità per la patologia e nel 10-15% dei casi sporadici. Sebbene il gene presenti un meccanismo di trasmissione autosomico dominante è stato stimato che circa un 20% di portatori di mutazioni all'interno di tale gene siano asintomatici. In genere tale forma esordisce intorno ai 40 anni ma l'esordio può variare dall'età infantile ai 60 anni. In 1/3 dei casi la spasticità può essere accompagnata da deficit della sensibilità profonda e disturbi urinari. La forma SPG3A è riscontrata in meno del 10% dei casi ad esordio adulto e nel 30% dei casi dominanti con esordio precoce. In questa forma la spasticità può essere accompagnata da atrofia muscolare distale, atrofia ottica, epilessia, atassia e neuropatia periferica assonale. Frequentemente i potenziali evocati motori evidenziano un prolungamento del tempo di conduzione centrale. Infine tra le forme più comuni di HSP autosomico dominante troviamo SPG31. Nella popolazione europea mutazioni nel gene *REEP1* vengono identificate nel 7% dei casi negativi all'analisi del gene *SPAST*. SPG31 esordisce tipicamente intorno ai 20 anni. In genere si tratta di una forma pura raramente vengono identificate forme "silver like" in cui la spasticità è accompagnata dall'amiotrofia dei muscoli delle mani. Circa il 40 % dei casi di HSP autosomico recessiva è data dalla forma SPG11, si tratta di una forma complicata caratterizzata dall'assottigliamento del corpo calloso spesso accompagnata da alterazioni della sostanza bianca ed atrofia corticale. La spasticità è spesso accompagnata da progressivo declino cognitivo, disartria,

ed atrofia dei muscoli tenar ed ipotenar. La patologia esordisce tipicamente entro i 30 anni, può esordire con un quadro di parkinsonismo seguito a pochi mesi dall'esordio dalla comparsa della spasticità agli arti inferiori. Clinicamente la forma SPG11 non può essere distinta dalla forma SPG15 data da mutazioni nel gene *ZFYVE26* la cui frequenza nella popolazione è minore del 3%. Si tratta di una forma progressiva che in genere nell'arco di 16 anni da dipendenza della sedia a rotelle. Altra forma frequente di HSP autosomico recessiva è la forma SPG7 che rappresenta circa il 5-7% dei casi. Nella maggior parte dei pazienti si osserva una lieve o moderata atrofia cerebellare, raramente si osservano ulteriori sintomi quali atrofia ottica, paralisi sopranucleare e declino cognitivo. Nonostante sia tipicamente considerata una forma autosomica recessiva sono stati riscontrati casi in cui la presenza di una sola mutazione desse il fenotipo malato. Attualmente, due sono le forme di HSP legate all'X di cui è noto il gene responsabile. La SPG1 è causata da mutazioni nel gene *L1CAM*. *L1CAM* è stato il primo gene malattia identificato per le HSP. In realtà, mutazioni in *L1CAM* oltre ad essere responsabili di HSP tipo SPG1 possono determinare una serie di patologie quali idrocefalo congenito da stenosi dell'acquedotto di Silvio (HSAS), sindrome MASA (ritardo mentale, afasia, andatura trascinata, pollici addotti) e agenesia del corpo calloso (ACC). Tutte queste patologie sono correlate tra loro e possono essere indicate col termine di sindrome L1.

Come per le atassie le forme sopra elencate sono solo alcune delle forme più frequenti di paraplegia spastica ma ad oggi nella banca dati Gene Nomenclature Committee (HGNC) ne sono elencate 50 forme e tale numero è in continua crescita soprattutto per quanto riguarda le forme complicate dove il quadro sintomatologico si sovrappone a quello di molte altre malattie neurodegenerative. Sovrapposizione che rende quindi difficile la diagnosi differenziale sulla sola base del quadro clinico ed è in questa problematica che trovano ampia applicazione le tecnologie NGS. Tecnologie che permettendo l'analisi contemporanea di più geni superano il vecchio

flusso di lavoro della diagnostica molecolare in cui veniva analizzato un gene per volta partendo da quelli più frequenti e man mano a quelli più rari protraendo a tempi estremamente lunghi la diagnosi.

## 3. MATERIALI E METODI

### 3.1 Pazienti

Lo studio è stato condotto su 78 pazienti reclutati presso il Centro di Genetica Molecolare dell'IRCCS Neuromed. Il consenso informato scritto è stato ottenuto da tutti i pazienti arruolati nello studio.

Ogni paziente è stato sottoposto a visita neurologica, consulenza genetica, analisi strumentali. Nello studio sono stati inclusi tutti i pazienti che presentavano:

- Esame neurologico che ponesse un sospetto clinico di Atassia o di paraparesi spastica o forma mista atasso/spastica o quadri neurologici complessi con una forte componente atassica e/o spastica;
- Studio neurofisiopatologico compatibile con una diagnosi di sindrome atassica e/o spastica;
- Studio di neuroimmagini compatibile con una diagnosi di sindrome atassica e/o spastica
- Negativi ad analisi molecolare di primo livello per espansioni di triplette all'interno dei geni *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7* e *FXN* (variazioni frequenti ma non identificabili mediante sequenziamento).

I pazienti sono stati classificati in funzione della sintomatologia in:

- Atassici, Spastici, Atasso/spastici, Fenotipi complessi (fenotipi in cui l'atassia e la spasticità erano solo una delle manifestazioni cliniche di un quadro sindromico)
- Modalità di ereditarietà (casi sporadici e casi familiari)

In funzione di tali criteri la casistica è composta da:

- 20 pazienti con fenotipo atassico di cui 8 sporadici e 12 con familiarità.
- 42 pazienti con fenotipo spatico di cui 27 sporadici e 15 con familiarità.

- 2 pazienti con fenotipo misto atasso/spastico entrambi sporadici.
- 14 pazienti con fenotipo complesso di cui 8 sporadici e 6 con familiarità.

Inoltre, ove possibile, un prelievo di sangue è stato eseguito, previo consenso informato, anche sui familiari dei pazienti.

## **3.2 Pannello NGS con geni selezionati**

### **3.2.1 Selezione dei geni**

Il DNA dei 78 pazienti è stato sottoposto ad analisi di sequenza delle regioni codificanti e delle giunzioni esone-introne di un pannello custom di 105 geni responsabili di atassia e/o paraplegia spastica mediante la piattaforma MiSeq dell'Illumina.

La selezione dei geni è stata effettuata mediante la ricerca e lo studio di articoli scientifici sui principali database di letteratura biomedica e di malattie geneticamente determinate utilizzando un elenco di parole chiave che meglio definivano il fenotipo di nostro interesse. Le parole chiave utilizzate sono state: "ataxia", "cerebellar ataxia", "spinocerebellar ataxia", "episodic ataxia", "spastic paraplegia", "spastic quadriplegic", "Strümpell-Lorrain disease", "familial spastic paraparesis", "hereditary spastic paraplegia" e "spastic ataxia".

I database consultati sono stati PubMed, "Online Mendelian Inheritance in Man" (OMIM) e "the Human Gene Mutation Database Professional Version" (HGMD).

PubMed è una banca dati di letteratura scientifica biomedica sviluppata dal National Center of Biotechnology Information (NCBI). Comprende oltre 29 milioni di citazioni di letteratura biomedica tratte dalla banca dati "Medical Literature Analysis and Retrieval System Online" (MEDLINE), riviste scientifiche e libri online. Le citazioni possono collegarsi a contenuti "full-text" da PubMed Central a siti web di editori.

OMIM è una banca dati che fornisce la descrizione delle malattie e dei caratteri fenotipici ereditari e dei geni ad esse correlate. Questo database è stato sviluppato all'inizio degli anni '60 dal Dr. Victor A. McKusick ed oggi è costantemente aggiornato da parte dell'NCBI. HGMD è una banca dati di mutazioni responsabili di alterazioni e disfunzioni di geni nucleari in malattie ereditarie umane. Ogni mutazione, all'interno della banca dati, è presente una sola volta ed è riportata in associazione alla prima fonte bibliografica in cui è stata descritta.

Il pannello è stato disegnato in modo da ottenere ampliconi su tutte le regioni codificanti e le giunzioni esone introne dei geni di interesse. Infine ogni amplicone ha un coverage stimato di almeno 50X. Il coverage è una misura di ridondanza ovvero rappresenta il numero di volte che un nucleotide è stato "letto" durante il sequenziamento ( $\text{Coverage} = \text{Throughput} / \text{Regione di interesse}$ ). Il coverage viene, quindi, considerato come un indice di accuratezza dell'analisi. Quante più volte la regione di interesse sarà stata letta tanto più bassa è la probabilità che ci siano stati errori durante il sequenziamento.

### **3.2.2 Preparazione delle librerie e sequenziamento**

Le librerie per il sequenziamento sono state preparate mediante il protocollo standard del kit SeqCap EZ Inherited Disease Panel (Roche) e sequenziate con protocollo paired-end sulla piattaforma MiSeq dell'Illumina.

Il protocollo di lavoro ha previsto i seguenti step:

1. preparazione delle librerie a partire da DNA genomico (gDNA) col kit KAPA HyperPlus Library Preparation (Roche).
2. amplificazione delle librerie mediante il kit KAPA HiFi HotStart ReadyMix (2X) (Roche);
3. cattura delle regioni target mediante ibridazione delle librerie con il pool di sonde SeqCap EZ (Roche).
4. purifica del campione catturato usando il kit di ibridazione e lavaggio SeqCap Hybridization and Wash (Roche)

5. amplificazione del campione di DNA catturato usando KAPA HiFi HotStart ReadyMix (2X).

6. sequenziamento del campione di DNA catturato e amplificato mediante la piattaforma MiSeq dell'Illumina.

### **3.2.3 Analisi dei dati**

I dati ottenuti sono stati elaborati mediante MiSeq Reporter Software 1.3.17. A fine della fase di sequenziamento il software converte le immagini di fluorescenza acquisite in sequenze nucleotidiche dette "reads". Le reads generate vengono poi sottoposte ad analisi qualitativa mediante un'analisi statistica che calcola la probabilità di errore nella lettura del nucleotide. Una corsa può essere definita di buona qualità quando almeno il 50% delle reads hanno superato il controllo qualitativo e il 75% delle basi ha uno Q-Score (previsione delle probabilità di errore nell'identificazione delle basi)  $\geq 30$ .

Le reads generate e i Quality Score ad esse associate sono organizzate in un file di output in formato FASTQ. I file FASTQ vengono poi utilizzati per la fase di allineamento delle reads generate con il genoma di riferimento (GRC37/hg19). Terminata la fase di allineamento si ha il processo di "variant calling". Durante questa fase ogni qual volta che il software rileva una discrepanza tra la sequenza di riferimento e la sequenza del nostro campione la variante viene annotata. Tale processo identifica tutte le varianti presenti nel campione in esame annotando dai polimorfismi alle mutazioni, ma anche le eventuali varianti derivanti da errori di sequenziamento. Per ogni variante viene annotata la coordinata cromosomica, il gene, il nucleotide wildtype e quello mutato, l'eventuale cambio aminoacidico, la profondità di lettura per quella base, la percentuale di letture che contengono la variante e i parametri qualitativi.

Al fine di valutare l'eventuale effetto patogeno di tutte le varianti che avessero passato i principali filtri qualitativi sono stati consultati database quali Exome Aggregation Consortium (ExAC), Short Genetic Variation (dbSNP) ClinVar, Exome Variant Server e 1000

Genomes ed effettuate analisi di predizione in silico mediante tools bioinformatici quali Mutation Taster, SIFT, Provean, Polyphen-2. Infine il tool di allineamento multiplo di sequenze GERP è stato utilizzato per valutare la conservazione del nucleotide “wild type” nei mammiferi mentre per le sole varianti missenso è stato utilizzato il tool PhyloP20 per valutare la conservazione dell’amminoacido “wild type” nei mammiferi. Una variante è stata classificata come patogena qualora presentasse almeno due dei seguenti criteri:

- Descritta in letteratura e/o database di riferimento in associazione a fenotipo atassico e/o spastico.
- Bassa frequenza nella popolazione generale dell’allele  
le varianti non devono essere presenti con una frequenza maggiore del 15% tra i pazienti inclusi nello studio
- Classificata come dannosa in almeno due su tre siti di predizione in-silico.
- Localizzata in una regione cromosomica la cui sequenza nucleotidica è conservata nelle specie.
- Variante de novo o ereditata da un parente affetto.
- Il fenotipo clinico correla con quello riportato in letteratura relativamente al gene in questione.

### **3.2.4 Validazione pannello**

Dall’analisi dei dati sono state identificate 22 varianti con potenziale significato patogeno. Di queste, 4 sono state confermate mediante metodica sanger nel probando, 7 in 4 pazienti per i quali si disponeva di un familiare per lo studio di segregazione. Il sequenziamento Sanger è preceduto dall’amplificazione delle regioni di interesse mediante PCR. Mediante il software Integrative Genomics Viewer (IGV) è stato identificato l’esone in cui queste varianti fossero localizzate. Con il tool gratuito disponibile on line Primer Blast della NCBI sono stati disegnati i primer per l’amplificazione della regione genomica di interesse. (Tab. 1)

GENE	TRASCritto	ESONE	SEQUENZA PRIMER (5' →3')
<i>SPAST</i>	NM_014946.3	10-11-12	agtagtactctcccccttct
			tgtagttggtaacaagaggc
		15	gcatgcctataatcccagc
			ttcgaggctgcagtgaact
		13	ttctaacaataccgaattaagg
			taggagtgaaattgctggct
9	gattacaggcatgagccac		
	gtacataactttctactctgaa		
<i>SPG7</i>	NM_003119.2	2	acctattgctcagactatta
			ccgtgcaacagggcgaga
		15	tgcccccggttctagattt
			ACACCGAGCAATAGACCCAG
<i>SPTBN2</i>	NM_006946	8	at gtcagccttg gaggactg
			gcagaagacagggacgtgaca
		27-28	ctctccccctgcttct tttg
			gggaggcagaaacagcatca
<i>SPG11</i>	NM_025137.3	18	tccctctgtgatgtttcca
			tcagctgagatctagacaatcca
		16	agcttgagtctgtcgttgcta
			tgtaccaactgttgagatggaga
<i>PNKP</i>	NM_007254	1	cttctgcagcaaattgg gg
			caagcgtccctctggattgt

**Tabella 1:** In tabella sono riportati il gene in cui è stata identificata la variante in esame, il trascritto canonico, l'esone e le sequenze dei primer utilizzate.

L'amplificazione è stata eseguita usando 25 µl di reagenti con 50 ng di DNA. La miscela di reazione conteneva 1X PCR Buffer, 1 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 25 pmol/µl dello specifico primer senso, 25 pmol/µl primer antisenso, 1.25 u/µl GoTaq Flexi DNA polymerase (Promega), H<sub>2</sub>O MilliQ a volume. Il profilo termico utilizzato per la reazione di amplificazione è stato denaturazione iniziale a 95°C per 4 minuti, seguita da 35 cicli di 30 secondi a 95°C, 30 secondi alla temperatura di melting (compresa tra i 56 e i 60°C), 30 secondi a 72°C, al termine dei 35 cicli un'estensione finale di 6 minuti 72°C. L'avvenuta reazione di amplificazione è stata verificata mediante elettroforesi su gel di agarosio (1,8% agarosio in 1X TBE buffer). I prodotti PCR sono stati purificati con il Kit MinElute PCR Purification (Qiagen). La reazione di sequenza è stata eseguita con 10 µl di reagenti e 20ng del prodotto purificato. La miscela di reazione

conteneva 1X Buffer, 1X BigDye Terminator v 3.1 (Applied Biosystems) 10 pmol/ $\mu$ l dello specifico primer senso o antisenso, H<sub>2</sub>O MilliQ a volume. Il prodotto di sequenza è stato purificato con il kit DyeEx 2.0 Spin Kit (Qiagen) ed analizzato mediante elettroforesi capillare su ABI 3100 Genetic Analyzer (Applied Biosystems).

### **3.3 Pannello esoma clinico**

L'analisi di sequenziamento mediante esoma clinico è stata condotta in 16 dei 62 pazienti risultati negativi all'analisi di sequenziamento mediante pannello NGS con geni selezionati, e in cui è stato possibile reperire il campione biologico di un familiare.

Dei 16 pazienti selezionati:

- 3 con paraplegia spastica (2 familiari, 1 sporadico)
- 6 con fenotipo atassico (1 familiare, 5 sporadici)
- 1 con fenotipo atasso-spastico senza familiarità
- 6 con fenotipo complesso (3 familiari, 3 sporadici)

L'esoma clinico è l'analisi contemporanea di tutti i geni ad oggi associati a malattia. Per l'analisi dell'esoma clinico è stato utilizzato il kit CE-IVD Clinical Exome Solution (SOPHiA Genetics) e sequenziate con protocollo paired-end sulla piattaforma MiSeq dell'Illumina. Il kit permette l'analisi contemporanea di 4493 geni associati a patologia è inoltre implementato da una piattaforma dedicata, SOPHiA DDM (SOPHiA Genetics) per l'analisi e l'interpretazione dei dati.

#### **3.3.1 Preparazione delle librerie e sequenziamento**

Le librerie per il sequenziamento sono state preparate mediante il protocollo standard del kit Clinical Exome Solution (SOPHiA GENETICS) e sequenziate con protocollo paired-end sulla piattaforma MiSeq dell'Illumina.

Il protocollo di lavoro ha previsto i seguenti step:

1. preparazione delle librerie a partire da DNA genomico (gDNA)
2. amplificazione delle librerie
3. cattura delle regioni target mediante ibridazione delle librerie
4. purifica del campione catturato
5. amplificazione del campione di DNA catturato
6. sequenziamento del campione di DNA catturato e amplificato mediante la piattaforma MiSeq dell'Illumina.

### **3.3.2 Analisi dei dati**

Come per l'analisi del pannello NGS target ci si è avvalsi del MiSeq Reporter Software 1.3.17 per la generazione dei file FASTQ. I file FASTQ generati sono poi stati caricati sulla piattaforma SOPHiA DDM (SOPHiA Genetics). Il software esegue quindi le fasi di "alignment" e di "variant calling", un'analisi qualitativa delle reads ed un'analisi bioinformatica di tutte le varianti con un coverage  $\geq 20$ . La piattaforma fornisce informazioni sulla frequenza della variante nella popolazione generale annotata nelle principali banche dati di riferimento, consulta tre banche dati (ESP, 1000 genomes, ExAC) inoltre verifica la frequenza della variante identificata in tutta la casistica delle analisi precedentemente caricate fino a fornire anche l'elenco dei pazienti in cui è stata identificata, e nella casistica di tutti gli utilizzatori della piattaforma. Il software inoltre consulta tre siti di predizione in silico (Mutation taster, Polyphen-2 e SIFT) per verificare l'effetto patogeno della variante. In funzione dall'analisi bioinformatica effettuata fornisce già una classificazione delle varianti in varianti di classe A ovvero varianti ad alto significato patogeno, varianti a potenziale effetto patogeno, classe B, classe C varianti a significato incerto infine varianti di classe D ovvero varianti verosimilmente benigne. La piattaforma permette di collegarsi direttamente sulla pagina di riferimento ExAC, IGV, ClinVar NCBI, Google, OMIM e ALAMUT. Per l'analisi dei dati è possibile sfruttare tutte queste informazioni per impostare filtri personalizzati in funzione delle proprie esigenze d'analisi. Nel nostro caso è stato impostato un filtro che escludesse tutte le varianti introniche ad eccezione di quelle situate a meno di 25bp dall'esone, tutte le varianti sinonime e tutte le varianti presenti in ExAC con una frequenza  $\geq 0.001$ . Un ulteriore filtro si è basato sull'esclusione di tutte le varianti identificate nei precedenti esperimenti di CES in soggetti con fenotipo differente da quello in studio. Sono, poi, state escluse tutte le varianti presenti nel familiare sano, qualora analizzato e incluse tutte quelle presenti nel familiare affetto, qualora analizzato. Sono state prese in considerazione le sole

varianti in cui almeno due delle analisi di predizione in silico da parte dei tool indagati dal software (Mutation taster, Polyphen2, SIFT) ne conferissero un effetto patogeno. Infine sono state incluse le varianti che avessero una frequenza  $\leq 0.001$  anche nelle ulteriori due banche dati consultate dal software (1000 genomes ed ESP). L'analisi bioinformatica col Sophia DDM è stata integrata mediante i tool di allineamento multiplo di sequenza GERP e PhyloP20 per verificare il grado di conservazione del punto in cui ricadeva la variante identificata nei mammiferi.

In questo tipo di approccio fondamentale è stata la comparazione del fenotipo del paziente col fenotipo causato dal gene candidato. Per la classificazione della variante sono stati utilizzati gli stessi criteri utilizzati per le varianti identificate mediante pannello NGS target.

### 3.3.3 Validazione del pannello

D'all'analisi dei dati sono state identificate 3 varianti con potenziale significato patogeno in due differenti pazienti. Mediante metodica Sanger è stata studiata la segregazione delle varianti identificate all'interno della famiglia. Il sequenziamento Sanger è preceduto dall'amplificazione delle regioni di interesse mediante PCR con il disegno dei primer mediante il tool Primer Blast della NCBI (Tab. 2).

GENE	TRASCritto	ESONE	SEQUENZA PRIMER (5' →3')
<i>POLR3A</i>	NM_007055	17	aatcgggcatctgcctca
			tgtctgtttggagctgtgac
		19	tcagaatcgtggccttgaa
			cagctcatgtgcaaacgtg
<i>ABCDI</i>	NM_000033	7	cga tgtgagcgtgtggatg
			gtctaaagtgccaggtgcc

**Tabella 2:** In tabella sono riportati il gene in cui è stata identificata la variante in esame, il trascritto canonico, l'esone amplificato, e le sequenze dei primer utilizzate.

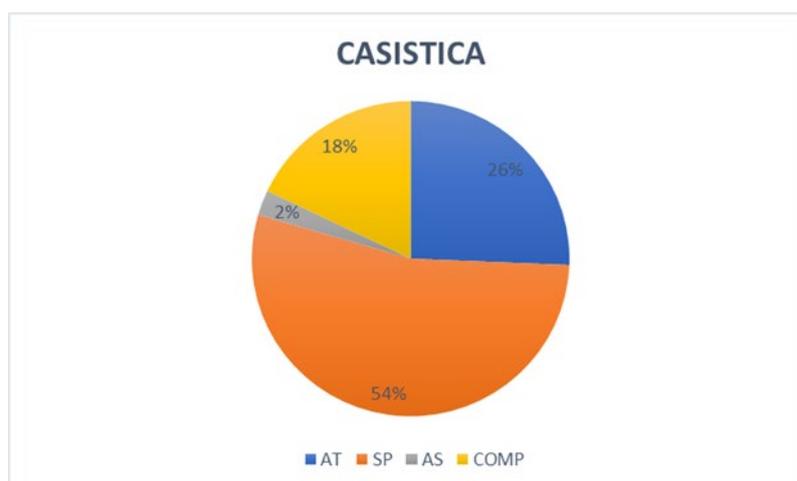
L'amplificazione è stata eseguita usando 25 µl di reagenti con 50 ng di DNA. La miscela di reazione conteneva 1X PCR Buffer, 1 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 25 pmol/µl dello specifico primer senso, 25

pmol/μl primer antisenso, 1.25 u/μl GoTaq Flexi DNA polymerase (Promega), H<sub>2</sub>O MilliQ a volume. Il profilo termico utilizzato per la reazione di amplificazione è stato denaturazione iniziale a 95°C per 4 minuti, seguita da 35 cicli di 30 secondi a 95°C, 30 secondi alla temperatura di melting (compresa tra i 56 e i 60°C), 30 secondi a 72°C, al termine dei 35 cicli un'estensione finale di 6 minuti 72°C. L'avvenuta reazione di amplificazione è stata verificata mediante elettroforesi su gel di agarosio (1,8% agarosio in 1X TBE buffer). I prodotti PCR sono stati purificati con il Kit MinElute PCR Purification (Qiagen). La reazione di sequenza è stata eseguita con 10 μl di reagenti e 20ng del prodotto purificato. La miscela di reazione conteneva 1X Buffer, 1X BigDye Terminator v 3.1 (Applied Biosystems) 10 pmol/μl dello specifico primer senso o antisenso, H<sub>2</sub>O MilliQ a volume. Il prodotto di sequenza è stato purificato con il kit DyeEx 2.0 Spin Kit (Qiagen) ed analizzato mediante elettroforesi capillare su ABI 3100 Genetic Analyzer (Applied Biosystems).

## 4.RISULTATI

### 4.1 Casistica e saggio NGS target

Sono stati analizzati 78 pazienti di cui 20 appartenenti al gruppo Atassico (AT), 42 al gruppo spastico (SP), 2 al gruppo atasso-spastico (AS) e 14 al gruppo con fenotipo complesso con componente atasso-spastica (COMP) (Fig.1).



**Figura 1:** Distribuzione dei pazienti nei differenti fenotipi. (AT: atassico, SP: spastico, AS: atasso-spastico, COMP: complesso).

I 78 campioni sono stati sottoposti ad analisi di sequenza mediante tecnologia NGS di 105 geni responsabili di atassia paraplegia spastica o sindromi atasso-spastiche (Tab.3).

GENE	FENOTIPO	MECCANISMO TRASMISSIONE
<i>ABCB7</i>	Anemina sideroblastica con atassia	XL
<i>ABHD12</i>	Polineuropatia, ipoacusia, atassia, retinite pigmentosa e cataratta	AR
<i>ADCK3</i>	Atassia autosomica recessiva da deficit di ubiquinone	AR
<i>AFG3L2</i>	Atassia spastica di tipo 5 o Atassia spinocerebellare di tipo 28	AD/AR
<i>ALDH18A1</i>	Paraplegia spastica di tipo 9A o 9B	AD/AR
<i>ALS2</i>	Paralisi spastica ascendente ereditaria a esordio infantile	AR
<i>ANO10</i>	Atassia spinocerebellare di tipo 10	AR
<i>AP4B1</i>	Paraplegia spastica tipo 47	AR
<i>AP4E1</i>	Paraplegia spastica tipo 51	AR
<i>AP4M1</i>	Paraplegia spastica tipo 50	AR

<i>AP4S1</i>	Paraplegia spastica tipo 52	AR
<i>AP5Z1</i>	Paraplegia spastica tipo 48	AR
<i>APTX</i>	Atassia con aprassia oculomotoria, ipoalbuminemia ad esordio precoce	AR
<i>ATCAY</i>	Atassia cerebellare di tipo Cayman	AR
<i>ATL1</i>	Paraplegia spastica tipo 3°	AD
<i>ATM</i>	Atassia – telangectasia	AR
<i>ATN1</i>	Atrofia dentato-rubro-pallido-luisiana	AD
<i>ATP1A3</i>	Sindrome di Capos	AD
<i>ATP8A2</i>	Sindrome da disequilibrio	AR
<i>ATXN1</i>	Atassia spinocerebellare di tipo 1	AD
<i>ATXN10</i>	Atassia spinocerebellare di tipo 10	AD
<i>ATXN2</i>	Atassia spinocerebellare di tipo 2	AD
<i>ATXN3</i>	Malattia di Machado-Joseph	AD
<i>ATXN7</i>	Atassia spinocerebellare di tipo 7	AD
<i>BEAN1</i>	Atassia spinocerebellare di tipo 31	AD
<i>BSC12</i>	Paraplegia spastica tipo 17	AD
<i>C9orf72</i>	Demenza frontotemporale e o sclerosi laterale amiotrofica di tipo 1	AD
<i>CA8</i>	Atassia con disabilità cognitiva, ritardo nella deambulazione e ipoplasia cerebellare.	AR
<i>CACNA1A</i>	Atassia spinocerebellare di tipo 6	AD
<i>CACNA1G</i>	Atassia spinocerebellare tipo 42	AD
<i>CACNB4</i>	Atassia episodica tipo 5	AD
<i>CAMTA1</i>	Atassia cerebellare non progressiva con disabilità cognitiva	AD
<i>CASK</i>	Deficit cognitivo legato all'X, tipo Najm	XLD
<i>CCT5</i>	Neuropatia autonoma e sensoriale ereditaria - paraplegia spastica	AR
<i>CLCN2</i>	Leucoencefalopatia con atassia	AR
<i>CLN5</i>	Ceroidolipofuscinosi neuronale dell'infanzia tardiva	AR
<i>CYP27A1</i>	Xantomatosi cerebrotendinea	AR
<i>CYP7B1</i>	Paraplegia spastica tipo 5°	AR
<i>DNMT1</i>	Atassia cerebellare, ipoacusia e narcolessia	AD
<i>ELOVL4</i>	Atassia spinocerebellare di tipo 34	AR/AD
<i>ENTPD1</i>	Paraplegia spastica tipo 64	AR
<i>ERLIN2</i>	Paraplegia spastica tipo 18	AR
<i>EXOSC3</i>	Ipoplasia pontocerebellare tipo 1B	AR
<i>FA2H</i>	Paraplegia spastica tipo 35	AR
<i>FGF14</i>	Atassia spinocerebellare tipo 27	AD
<i>FLVCR1</i>	Atassia del cordone posteriore - retinite pigmentosa	AR
<i>FMR1</i>	Sindrome da X fragile con tremore e atassia	XLD
<i>FXN</i>	Atassia di Friedreich	AR
<i>GAD1</i>	Tetraplegia spastica congenita ereditaria	AR
<i>GJC2</i>	Paraplegia spastica tipo 44	AR

<i>GOSR2</i>	Epilessia mioclonica progressiva tipo 6	AR
<i>GRM1</i>	Paraplegia spastica tipo 44	AD/AR
<i>ITPR1</i>	Paraplegia spastica tipo 15 o tipo 29	AD
<i>KCNA1</i>	Atassia episodica	AD
<i>KCNC3</i>	Atassia spinocerebellare tipo 13	AD
<i>KCND3</i>	Atassia spinocerebellare tipo 19	AD
<i>KCNJ10</i>	Sindrome di SESAME	AR
<i>KIAA0196</i>	Paraplegia spastica tipo 8	AD
<i>KIAA0226</i>	Spinocerebellare atassia di tipo 15	AR
<i>KIF1A</i>	Paraplegia spastica tipo 30	AR
<i>KIF5A</i>	Paraplegia spastica tipo 10	AD
<i>L1CAM</i>	sindrome di MASA	XLR
<i>LAMA1</i>	Sindrome di Poretti-Boltshauser	AR
<i>LAMB1</i>	Lissencefalia di tipo 5	AR
<i>MARS2</i>	Atassia spastica tipo 3	AR
<i>MTPAP</i>	Atassia spastica tipo 4	AR
<i>NOP56</i>	Atassia spinocerebellare tipo 36	AD
<i>OPHN1</i>	Deficit cognitivo legato all'X - ipoplasia cerebellare	XLR
<i>PDYN</i>	Atassia spinocerebellare tipo 23	AD
<i>PEX7</i>	Disordine della biogenesi dei perossisomi tipo 9b	AR
<i>PHYH</i>	Malattia di Refsum	AR
<i>PNKP</i>	Atassia con aprassia oculomotoria tipo 4	AR
<i>PNPLA6</i>	Paraplegia spastica tipo 39	AR
<i>POLG</i>	Atassia spinocerebellare con epilessia	AR
<i>PPP2R2B</i>	Atassia spinocerebellare tipo 12	AD
<i>PRKCG</i>	Atassia spinocerebellare tipo 14	AD
<i>REEP1</i>	Paraplegia spastica tipo 31	AD
<i>SACS</i>	Atassia spastica tipo Charlevoix-Saguenay	AR
<i>SCN2A</i>	Encefalopatia epilettica tipo 11	AD
<i>SETX</i>	Atassia spinocerebellare associata a neuropatia assonale, tipo 2	AD
<i>SIL1</i>	Atassia spastica tipo Charlevoix-Saguenay	AR
<i>SLC16A2</i>	Sindrome di Allan-Herndon-Dudley	XL
<i>SLC1A3</i>	Atassia episodica tipo 6	AD
<i>SLC33A1</i>	Paraplegia spastica tipo 42	AD
<i>SLC9A6</i>	Sindrome di Christianson	XLD
<i>SPAST</i>	Paraplegia spastica tipo 4	AD
<i>SPG11</i>	Paraplegia spastica tipo 11	AR
<i>SPG20</i>	Sindrome di Troyer	AR
<i>SPG21</i>	Sindrome di Mast	AR
<i>SPG7</i>	Paraplegia spastica tipo 7	AR
<i>SPTBN2</i>	Atassia spinocerebellare tipo 5 o tipo 14	AD/AR
<i>SYNE1</i>	Atassia spinocerebellare tipo 8	AR
<i>SYT14</i>	Atassia spinocerebellare di tipo 11	AR

<i>TBP</i>	Atassia spinocerebellare di tipo 17	AD
<i>TDP1</i>	Atassia spinocerebellare associata a neuropatia assonale	-
<i>TGM6</i>	Atassia spinocerebellare di tipo 35	AD
<i>TRPC3</i>	Atassia spinocerebellare di tipo 41	AD
<i>TSFM</i>	Malattia mitocondriale fatale da difetto combinato della fosforilazione ossidativa, tipo 3	AR
<i>TTBK2</i>	Atassia spinocerebellare di tipo 11	AD
<i>TTPA</i>	Atassia con deficit della vitamina E	AR
<i>VLDLR</i>	Sindrome da disequilibrio	AR
<i>WFS1</i>	Sindrome di Wolfram	AR/AD
<i>WWOX</i>	Atassia spinocerebellare di tipo 1	AR
<i>ZFYVE26</i>	Paraplegia spastica di tipo 15	AR
<i>ZFYVE27</i>	Paraplegia spastica di tipo 33	AD

**Tabella 3:** In tabella sono riportati i geni selezionati, il fenotipo associato e il meccanismo di trasmissione (AD: autosomico dominante, AR: autosomico recessivo, XL: legato al cromosoma X, XLD: legato al cromosoma X ma con trasmissione dominante, -: ancora non nota).

Il saggio è stato disegnato in modo da sequenziare le regioni codificanti e le giunzioni esone-introne dei 105 geni con una copertura media percentuale del target  $\geq 50X$ . Le reads generate hanno una lunghezza media di 150 bp. In media il 90% delle reads presentavano un Q-Score  $\geq 30$ . Per ogni paziente sono state identificate una media di 1000 varianti con un minimo di 800 ad un massimo di 1200 varianti a campione.

### 4.1.1 Varianti e geni

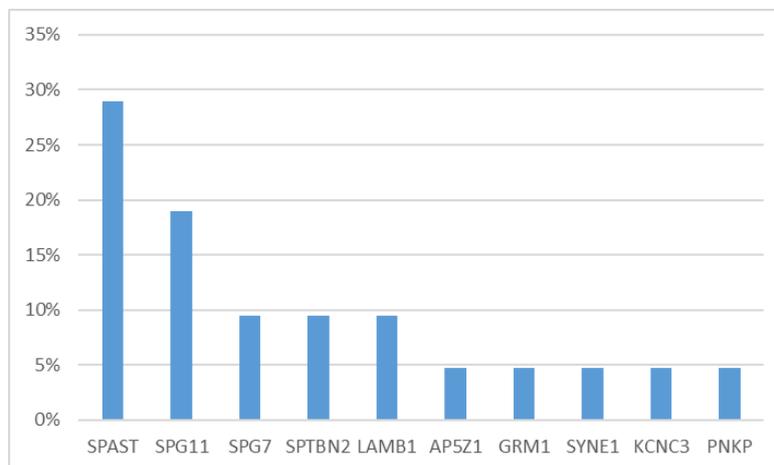
Il 20,5% (16 di 78) dei pazienti è risultato positivo all'analisi, per un totale di 23 varianti identificate (Tab.4).

PAZIENTE (ID)	GENE	VARIANTE NUCLEOTIDICA	VARIANTE PROTEICA	dbSNP ID	GENOTIPO
Paziente 1 (Z35)	<i>SPAST</i>	c.1378 C>T	p.Arg460Cys	rs878854990	hz
Paziente 2 (Z90)	<i>SPAST</i>	c.1679C>T	p.Pro560Leu	nr	hz
Paziente 3 (Z143)	<i>SPAST</i>	c.1537-8T>G	-	nr	hz
Paziente 4 (z265)	<i>SPAST</i>	c.1537-8T>G	-	nr	hz
Paziente 5 (Z192)	<i>SPAST</i>	c.1245+1G>A	-	rs875989878	hz
Paziente 6 (Z424)	<i>SPAST</i>	c.1414-2A>T	-	nr	hz
Paziente 7 (Z286)	<i>SPG11</i>	c.6331dupG	p.Glu2111GlyfsTer38	nr	hz
	<i>SPG11</i>	c.6175C>T	p.Arg2059Trp	nr	hz
	<i>SPG11</i>	c.5986dupT	p.Cys1996LeufsTer4	rs312262775	hz
Paziente 8 (Z766)	<i>SPG11</i>	c.3291+3A>G	-	nr	hz
		c.2842dupG	p.Val948GlyfsTer6	rs312262750	hz
Paziente 9 (Z261)	<i>SPG7</i>	c.233T>A	p.Leu78Ter	rs121918358	hz
		c.1940C>A	p.Ala647Glu	nr	hz
Paziente 10 (Z578)	<i>KCNC3</i>	c.2197C>A	p.Gln733Lys	nr	hz
Paziente 11 (Z436)	<i>SPTBN2</i>	c.5662G>T	p.Ala1888Ser	nr	hz
Paziente 12 (Z450)	<i>AP5Z1</i>	c.985G>A	p.Val329Met	rs764235114	hz
	<i>GRM1</i>	c.3388G>C	p.Asp1130His	nr	hz
Paziente 13 (Z503)	<i>SYNE1</i>	c.19893+2T>C	-	nr	oz
Paziente 14 (Z770)	<i>SPTBN2</i>	c.830A>G	p.Tyr277Cys	nr	hz
Paziente 15 (Z781)	<i>LAMB1</i>	c.3317T>A	p.Met1106Lys	nr	hz
		c.313C>G	p.Pro105Ala	nr	hz
Paziente 16 (Z1107)	<i>PNKP</i>	c.148C>G	p.Gln50Glu	rs756746191	oz

**Tabella 4:** In tabella sono riportati il paziente in cui è stata identificata una variante potenzialmente patologica, il gene in cui è stata identificata, la variante nucleotidica in funzione della regione codificante, la variante proteica, l'eventuale ID della banca dati dbSNP, e il genotipo (nr: non riportato; hz: eterozigote; oz: omozigote).

Le varianti identificate sono state riscontrate in 9 geni responsabili di atassia e/o paraplegia spastica e una variante è stata identificata nel gene *LAMB1* responsabile di Lissencefalia di tipo 5 in cui l'atassia e la spasticità possono essere sintomi di un quadro sindromico più complesso

Il 29% delle varianti è stata identificata all'interno del gene *SPAST* confermando i dati di letteratura in cui tale gene risulta il più frequentemente mutato in pazienti con fenotipo spastico puro. Il 19% delle varianti sono state identificate nel gene *SPG11* il 10% nei geni *SPG7*, *SPTBN2* e *LAMB1* infine il 5% nei geni *AP5Z1*, *GRM1*, *SYNE1*, *KCNC3* e *PNKP* (Fig.2).



**Figura 2:** Distribuzione delle varianti potenzialmente patogene: 29% nel gene *SPAST*, 19% nel gene *SPG11* il 10% nei geni *SPG7*, *SPTBN2* e *LAMB1* infine il 5% nei geni *AP5Z1*, *GRM1*, *SYNE1*, *KCNC3* e *PNKP*.

Delle 23 varianti identificate 7 sono localizzate nel sito di splicing e pertanto putativamente responsabili di splicing alternativo. Le restanti 16 varianti sono tutte localizzate all'interno della regione codificante del gene. Di queste 12 sono varianti missenso, ovvero che causano un cambio amminoacidico nella proteina codificata dal gene, 3 sono varianti frameshift, ovvero che causano uno slittamento del codice di lettura con conseguente cambio amminoacidico fino alla formazione di un codone di stop prematuro,

infine 1 variante è una stop gained, ovvero che causa la formazione di un codone di stop prematuro (Tab.5).

GENE	VARIANTE NUCLEOTIDICA	VARIANTE PROTEICA	dbSNP ID	EFFETTO	LOCALIZZAZIONE
<i>SPAST</i>	c.1378 C>T	p.Arg460Cys	rs878854990	Missenso	Esonica
<i>SPAST</i>	c.1679C>T	p.Pro560Leu	Nr	Missenso	Esonica
<i>SPAST</i>	c.1537-8T>G	-	Nr	Splicing	Sito di splicing
<i>SPAST</i>	c.1537-8T>G	-	Nr	Splicing	Sito di splicing
<i>SPAST</i>	c.1245+1G>A	-	rs875989878	Splicing	Sito di splicing
<i>SPAST</i>	c.1414-2A>T	-	Nr	Splicing	Sito di splicing
<i>SPG11</i>	c.6331dupG	p.Glu2111GlyfsTer38	Nr	Frameshift	Esonica
<i>SPG11</i>	c.5986dupT	p.Cys1996LeufsTer4	rs312262775	Frameshift	Esonica
<i>SPG11</i>	c.3291+3A>G	-	Nr	Splicing	Sito di splicing
	c.2842dupG	p.Val948GlyfsTer6	rs312262750	Frameshift	Esonica
<i>SPG7</i>	c.233T>A	p.Leu78Ter	rs121918358	Stop gained	Esonica
	c.1940C>A	p.Ala647Glu	Nr	Missenso	Esonica
<i>KCNC3</i>	c.2197C>A	p.Gln733Lys	Nr	Missenso	Esonica
<i>SPTBN2</i>	c.5662G>T	p.Ala1888Ser	Nr	Missenso	Esonica
<i>AP5Z1</i>	c.985G>A	p.Val329Met	rs764235114	Missenso	Esonica
<i>GRM1</i>	c.3388G>C	p.Asp1130His	Nr	Missenso	Esonica
<i>SYNE1</i>	c.19893+2T>C	-	Nr	Splicing	Sito di splicing
<i>SPTBN2</i>	c.830A>G	p.Tyr277Cys	Nr	Missenso	Esonica
<i>LAMB1</i>	c.3317T>A	p.Met1106Lys	Nr	Missenso	Esonica
	c.313C>G	p.Pro105Ala	Nr	Missenso	Esonica
<i>PNKP</i>	c.148C>G	p.Gln50Glu	rs756746191	Missenso	Esonica

**Tabella 5:** In tabella sono riportati il gene, la variante nucleotidica in funzione della regione codificante, la variante proteica, l'eventuale ID della banca dati dbSNP, l'effetto, e la localizzazione.

#### **4.1.2 Pazienti in cui sono state identificate una o più varianti**

È stata identificata una variante nel 20,5 % dei pazienti (16 di 78). Considerando le varie sottoclassi cliniche (non considerando il gruppo AS con un numero statisticamente non significativo), la percentuale dei pazienti positivi oscilla da un minimo di 14% dei pazienti COMP (2 di 14), ad un massimo di 21,5% dei pazienti SP (9 di 42), passando per il 20% dei pazienti AT (4 di 20).

Il maggior contributo di pazienti con mutazione identificata è fornito dal gruppo SP, (56 %, 9/16), seguito in ordine dagli AT (25%, 4 /16), COMP (12%, 2 di 16); e AS (6 1 di 16).

Dei 78 pazienti, il 42% ha una storia familiare (33/78) mentre il restante 58% non ha una anamnesi familiare, e viene considerata quindi sporadica (45/78). Non è stata identificata una differenza significativa tra i pazienti sporadici e familiari né considerando i 16 pazienti in cui è stata identificata almeno una variante, (sporadici 9/16, 56%, familiari 7/16. 44%,  $p=0.48$ ), né considerando gli interi sottogruppi di sporadici vs familiari (sporadici positivi 9/45, 20%, familiari positivi 7/33, 21,2%,  $p=0.89$ )

Tuttavia, sono state identificate differenze statisticamente significative nei vari gruppi clinici. Infatti, gli SP familiari sono molto più rappresentati degli sporadici (sporadici 3/27, 11,1%, familiari 6/15, 40%,  $p=0.03$ ), mentre gli AT sporadici sono molto più rappresentati dei familiari (sporadici 3/8, 37,5%, familiari 1/12, 8,3%,  $p=0.13$ )

Dei 16 pazienti, 10 hanno mostrato una segregazione autosomica dominante (sia in pazienti familiari che sporadici), mentre 5 una segregazione autosomica recessiva (solo in casi sporadici).

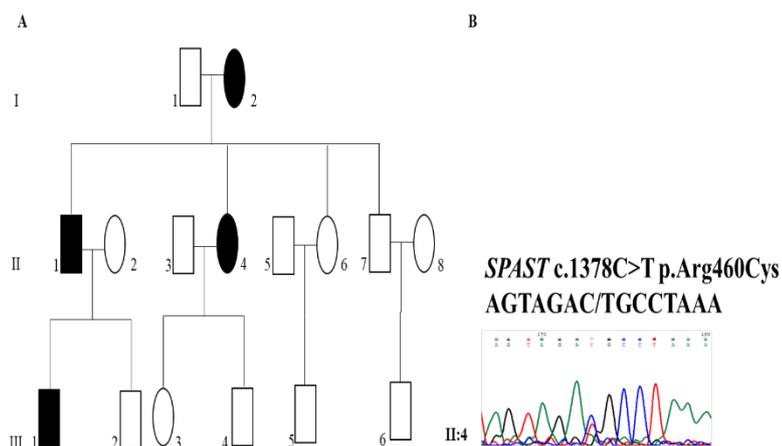
Dei 10 con AD 1 SP sporadici, 6 SP familiare, 1 AT familiari e 2 AT sporadici. Dei 6 con AR 2 SP sporadici, 1 AT sporadico, 1 AS sporadico e 2 COMP sporadici.

### 4.1.3 Descrizione genetica e clinica dei pazienti

→Varianti identificate in geni con modalità di trasmissione autosomica dominante (SPG4)

#### Paziente 1, IdZ35

Paziente di sesso femminile con andatura di tipo spastico e debolezza muscolare. Le indagini strumentali elettromiografia (EMG), elettroencefalografia (ENG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati visivi (PEV), potenziali evocati somatosensoriali (SEP) ai quattro arti e i potenziali evocati motori (PEM) non mostrano alterazioni significative. Sulla base della sintomatologia e degli esiti negativi delle indagini strumentali viene eseguita una diagnosi clinica di una forma pura di Paraparesi Spastica con esordio a 33 anni (gruppo SP). In sede di consulenza genetica, viene evidenziata familiarità per la patologia nella mamma (I:2), nel fratello (II:1), e nel figlio del fratello (III:1) (Fig.3A).



**Figura 3: (A)** Albero genealogico: (II:4) probanda, (I:2; II:1 e III:1) rispettivamente madre fratello e nipote della probanda con diagnosi clinica di paraparesi spastica. **(B)** Elettroferogramma della variante c.1378C>T p.Arg460Cys nel gene SPAST (NM\_014946) identificata nella paziente (II:4).

L'analisi di sequenziamento ha identificato una variante missenso in eterozigosi nel gene SPAST (NM\_014946.3) rs878854990

c.1378C>T p.Arg460Cys, che causa la sostituzione di un'arginina con una cisteina nella posizione 460 della proteina. Tale variante è stata considerata come patologica in quanto già descritta in letteratura (Falco et al (2004) in associazione a paraplegia spastica ereditaria. Inoltre la frequenza della variante non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tutti e quattro i tool utilizzati (Mutation taster, SIFT, PROVEAN, e Polyphen-2.). L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (0,9) e un NR pari a 5.6 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.6)

VARIANTE	SPAST	c.1378 C>T	p.Arg460Cys	rs878854990
ExAC	FREQUENZA			
	nr			
1000 Genomes	nr			
	Prediction		Accuracy	
MUTATION TASTER	Disease causing		1	
SIFT	Damaging		0	
PROVEAN	Damaging		-8	
POLYPHEN-2	Probably damaging		1	
	mammalian		mamm. rankscore	
phyloP20	0,9		0,4	
	NR			
GERP	5,6			

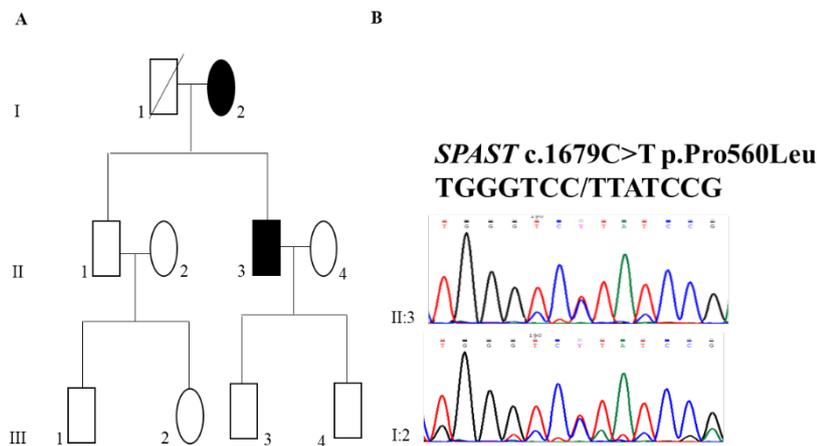
**Tabella 6:** Risultati analisi bioinformatica.

La variante è stata confermata mediante sequenziamento di tipo Sanger nella paziente (II:4) (Fig.3B). Non è stato possibile effettuare studi di segregazione.

### **Paziente 2, IdZ90**

Paziente di sesso maschile con andatura di tipo spastico e debolezza muscolare. Le indagini strumentali di elettromiografia (EMG), elettroencefalografia (ENG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati visivi (PEV), potenziali evocati somatosensoriali (SEP) ai quattro arti e i potenziali evocati motori (PEM) non mostrano alterazioni significative. Sulla base della sintomatologia e degli esiti negativi

delle indagini strumentali viene eseguita una diagnosi clinica di una forma pura di Paraparesi Spastica con esordio a 44 anni (gruppo SP). In sede di consulenza genetica, viene evidenziata familiarità per la patologia nella mamma (I:2) (Fig.4A).



**Figura 4:** (A) Albero genealogico: (II:3) probando (I:2) madre del probando con diagnosi clinica di paraparesi spastica. (B) Analisi di segregazione: Elettroferogrammi della variante c.1679C>T p.Pro560Leu nel gene SPAST (NM014946) identificata nel paziente (II:3) e confermata nella mamma (I:2).

L'analisi di sequenza ha identificato una variante missenso in eterozigosi nel gene *SPAST* (NM\_014946.3) c.1679C>T p.Pro560Leu che causa la sostituzione di una prolina con una leucina nella posizione 560 della proteina. La variante è stata considerata patologica in funzione di tre criteri. L'analisi di segregazione ha confermato la presenza della variante anche nella mamma (I:2) (Fig.4B). La frequenza della variante non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tutti e quattro i tool utilizzati (Mutation taster, SIFT, PROVEAN, e Polyphen-2.). L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (0,8) e un NR pari a 5.6 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.7)

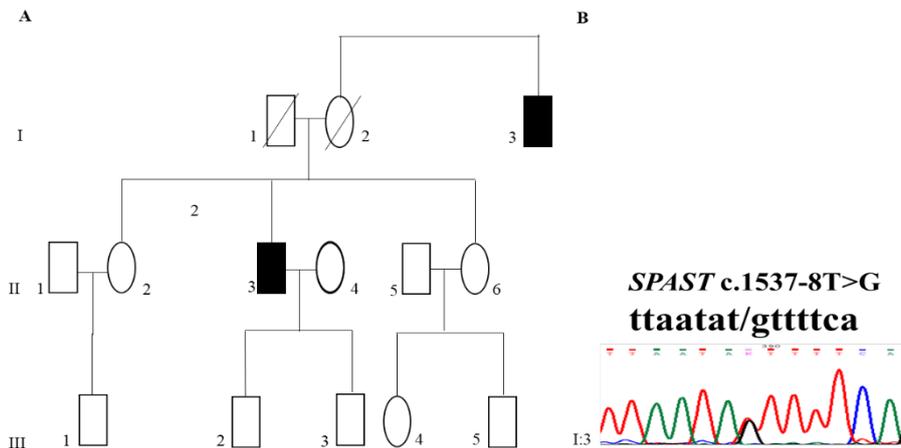
<b>VARIANTE</b>	<b>SPAST</b>	c.1679C>T	p.Pro560Leu
<b>ExAC</b>	<b>FREQUENZA</b>		
	nr		
<b>1000 Genomes</b>	nr		
	<b>Prediction</b>	<b>Accuracy</b>	
<b>MUTATION TASTER</b>	Disease causing	0,9	
<b>SIFT</b>	Damaging	0	
<b>PROVEAN</b>	Damaging	-8	
<b>POLYPHEN-2</b>	Probably damaging	0,9	
	<b>mammalian</b>	<b>mamm. rankscore</b>	
<b>phyloP20</b>	<b>0,8</b>	0,3	
	<b>NR</b>		
<b>GERP</b>	5,6		

**Tabella 7:** Risultati analisi bioinformatica.

### **Paziente 3, IdZ143**

Paziente di sesso maschile con difficoltà nella deambulazione di tipo spastico.

Durante un ricovero esegue una serie di indagini che non mostrano alterazioni significative: elettromiografia (EMG), elettroencefalografia (ENG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati visivi (PEV), potenziali evocati somatosensoriali (SEP) ai quattro arti. I potenziali evocati motori (PEM) mostrano agli arti inferiori una latenza aumentata bilateralmente delle risposte corticali ed un tempo di conduzione motoria centrale aumentato. Su tali informazioni viene eseguita una diagnosi clinica di Paraparesi Spastica con esordio a 35 anni (gruppo SP). In sede di consulenza genetica, viene evidenziata familiarità per la patologia nello zio materno (I:3) (Fig.5A).



**Figura 5: (A)** Albero genealogico: (II:3) probando (I:3) zio materno del probando con sintomatologia assimilabile a quella del probando. **(B)** Analisi di segregazione. Elettroferogramma della variante c.1537-8T>G nel gene SPAST (NM014946) nel paziente I:3.

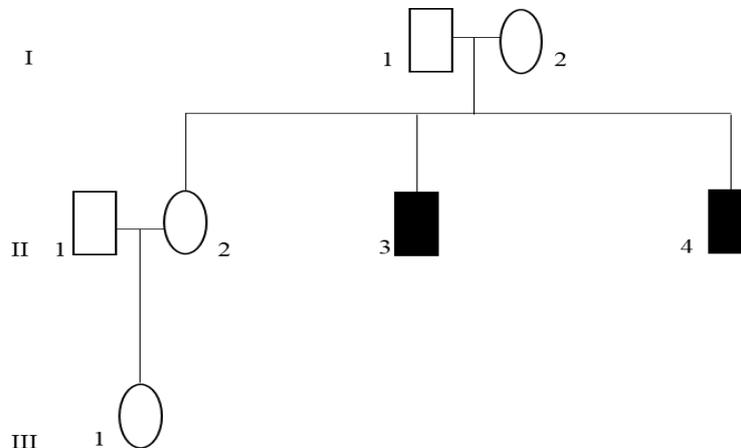
L'analisi di sequenziamento ha identificato una variante di splicing in eterozigosi nel gene *SPAST* (NM\_014946.3) c.1537-8T>G. La variante è stata considerata patologica in funzione di tre criteri. L'analisi di segregazione ha confermato la presenza della variante anche nello zio affetto (Fig.5B). La frequenza della variante non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. I punteggi ADA e RF del tool dbSCSNV sono stati entrambi di 0.9 confermando l'ipotesi che la variante possa essere causa di splicing alternativo. L'NR del tool GERP è pari a 5.6 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.8).

VARIANTE	<i>SPAST</i>	c.1537-8T>G
<b>FREQUENZA</b>		
<b>ExAC</b>	nr	
<b>1000 Genomes</b>	nr	
	<b>ADA Score</b>	<b>RF Score</b>
<b>dbSCSNV</b>	0,9	0,9
<b>NR</b>		
<b>GERP</b>	5,6	

**Tabella 8:** Risultati analisi bioinformatica.

### **Paziente 4, IdZ265**

Paziente di sesso maschile con disturbo della deambulazione di tipo spastico. Il paziente è stato sottoposto ad analisi strumentali quali elettromiografia (EMG), elettroencefalografia (ENG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati visivi (PEV), potenziali evocati somatosensoriali (SEP) ai quattro arti e potenziali evocati motori (PEM) ai quattro arti, tali indagini sono risultate tutte nella norma tranne i PEM dove veniva evidenziato un lieve aumento bilaterale del tempo di conduzione motoria centrale. Su tali informazioni viene eseguita una diagnosi clinica di Paraparesi Spastica con esordio a 30 anni (gruppo SP). In sede di consulenza genetica, viene evidenziata familiarità per la patologia nel fratello (II:4) e nel padre (I:1), per i quali viene riportata una condizione clinica sovrapponibile (Fig.6).

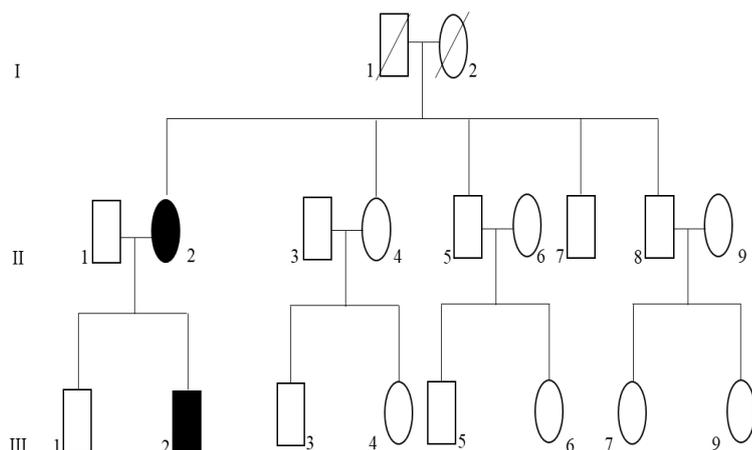


**Figura 6:** Albero genealogico: (II:3) probando, (I:1; II:4) rispettivamente padre e fratello con diagnosi clinica di paraparesi spastica.

L'analisi di sequenziamento ha identificato una variante in eterozigosi nel gene *SPAST* (NM\_014946.3) c.1537-8T>G, una mutazione di splicing già identificata nel Paziente 3 idZ143. Non è stato possibile eseguire una analisi di segregazione nella famiglia.

### **Paziente 5, IdZ192**

Paziente di sesso maschile con disturbo della deambulazione di tipo spastico. Sulla base di informazioni cliniche viene eseguita una diagnosi di Paraparesi Spastica con esordio neonatale (gruppo SP). In sede di consulenza genetica, viene evidenziata familiarità per la patologia nella mamma (II:1) (Fig.7).



**Figura 7:** Albero genealogico: (III:2) probando (Z192), (II:2) madre (Z191) con diagnosi clinica di paraparesi spastica.

L'analisi di sequenziamento ha identificato una variante di splicing in eterozigosi nel gene *SPAST* (NM\_014946.3) rs87598987 c.1245+1G>A.

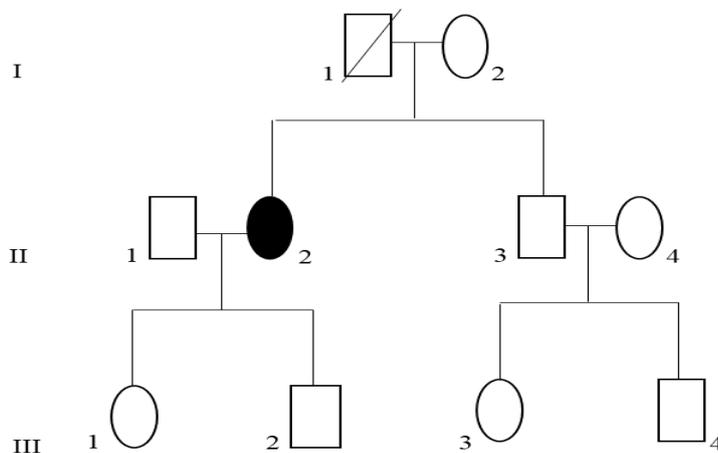
La variante è stata considerata patologica in funzione di tre criteri. L'analisi di segregazione ha confermato la presenza della variante anche nella mamma caratterizzata però da un esordio in età adulta (42 aa). La frequenza della variante non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. I punteggi ADA e RF del tool dbSNV sono stati entrambi di 0.9 confermando l'ipotesi che la variante possa essere causa di splicing alternativo. L'NR del tool GERP è pari a 5.7 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.9).

VARIANTE	SPAST	c.1245+1G>A	rs875989878
FREQUENZA			
ExAC	nr		
1000 Genomes	nr		
	ADA Score	RF Score	
dbscSNV	0,9	0,9	
NR			
GERP	5,7		

**Tabella 9:** Risultati analisi bioinformatica.

### **Paziente 6, IdZ424**

Paziente di sesso femminile con difficoltà della deambulazione di tipo spastico. Durante un ricovero esegue una serie di indagini: elettroencefalografia (ENG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati somatosensoriali (SEP) ai quattro arti e potenziali evocati motori (PEM) ai quattro arti. L'RM mostra una tenue e diffusa iperintensità in T2 della sostanza bianca. Su tali informazioni viene eseguita una diagnosi clinica di Paraparesi Spastica con esordio a 47 anni (gruppo SP). In sede di consulenza genetica, non viene evidenziata familiarità per la patologia (Fig.8).



**Figura 8:** Albero genealogico: (II:2) probanda.

L'analisi di sequenziamento ha identificato una variante di splicing in eterozigosi nel gene SPAST (NM\_014946.3) c.1414-2A>T. La variante è stata considerata patologica sulla base della frequenza e dell'analisi bioinformatica. La frequenza della variante non è nota

né sulla banca dati ExAC né sulla banca dati 1000 Genomes. I punteggi ADA e RF del tool dbscSNV sono stati entrambi di 0.9 confermando l'ipotesi che la variante possa essere causa di splicing alternativo. L'NR del tool GERP è pari a 5.4 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.10).

VARIANTE	<i>SPAST</i> c.1414-2A>T	
	FREQUENZA	
ExAC	nr	
1000 Genomes	nr	
	ADA Score	RF Score
dbscSNV	0,9	0,9
	NR	
GERP	5,4	

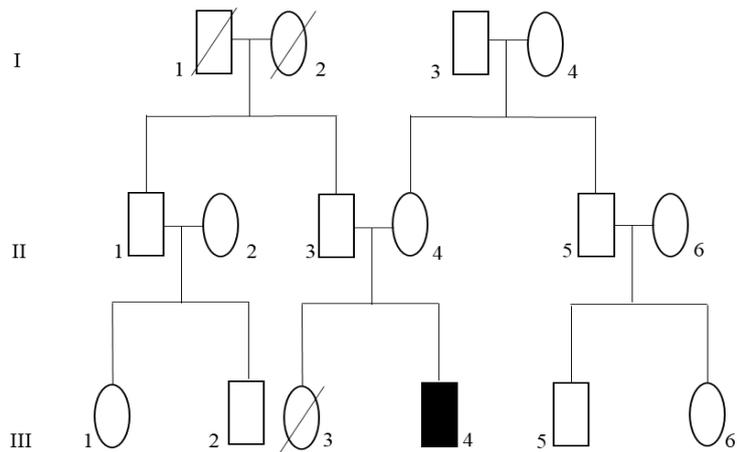
**Tabella 10:** Risultati analisi bioinformatica.

**→Varianti identificate in geni con modalità di trasmissione autosomico recessivo (SPG11, SPG7)**

**Paziente 7, IdZ286**

Paziente di sesso maschile con disturbo dell'andatura di tipo spastico, disturbo del linguaggio e lieve ritardo intellettivo.

Durante un ricovero esegue una serie di indagini. L'elettromiografia (EMG), elettroencefalografia (ENG) ed i potenziali evocati somatosensoriali (SEP) ai quattro arti sono risultati nella norma. I potenziali evocati motori (PEM) hanno evidenziato agli arti inferiori un aumento bilaterale del tempo di conduzione motoria centrale associato a notevole riduzione di ampiezza delle risposte corticali. La risonanza magnetica (RM) dell'encefalo ha evidenziato sfumata iperintensità marginale periferica dei corni ventricolari frontali, presenza di asimmetria dei trigoni per prevalenza a destra e minima cisterna magna retrocerebellare. Su tali informazioni cliniche viene eseguita una diagnosi clinica di Paraparesi Spastica con esordio a 15 anni (gruppo SP). In sede di consulenza genetica, non viene evidenziata familiarità per la patologia (Fig.9).



**Figura 9:** Albero genealogico: (III:4) probando.

L'analisi di sequenziamento ha identificato due varianti in eterozigosi nel gene SPG11 (NM\_025137.3). La variante frameshift c.6331dupG p.Glu2111GlyfsTer38 che causa sostituzioni amminoacidiche a partire dalla posizione 2111 della proteina fino la formazione di un codone di stop prematuro nel 38° codone successivo e la variante frameshift rs312262775 c.5986dupT p.Cys1996LeufsTer4 che causa sostituzioni amminoacidiche a partire dalla posizione 1996 della proteina fino la formazione di un codone di stop prematuro nel 4° codone successivo.

Le due varianti sono state considerate patologiche in funzione della frequenza e dell'analisi di predizione in silico. La frequenza della variante c.6331dupG p.Glu2111GlyfsTer38 non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico è stata effettuata mediante il tool Polyphen-2 che ne conferisce un effetto dannoso. L'NR del tool GERP è pari a 5.4 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.11).

VARIANTE	SPG11	c.6331dupG	p.Glu2111GlyfsTer38
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	Prediction		Accuracy
POLYPHEN-2	Possibly Damaging		0,9
	NR		
GERP	5,4		

**Tabella 11:** Risultati analisi bioinformatica.

La frequenza della variante rs312262775 c.5986dupT p.Cys1996LeufsTer4 è pari a 0,000016 sulla banca dati ExAC e non nota nella banca dati 1000 Genomes. L'analisi di predizione in silico è stata effettuata mediante il tool Polyphen-2 che ne conferisce un effetto dannoso. L'NR del tool GERP è pari a 5.4 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.12).

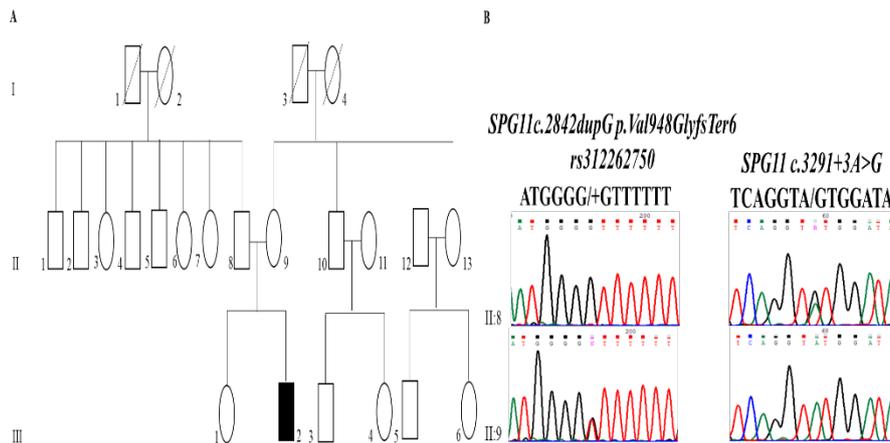
VARIANTE	SPG11	c.5986dupT	p.Cys1996LeufsTer4 rs312262775
ExAC	FREQUENZA		
	0,000016		
1000 Genomes	nr		
	Prediction		Accuracy
POLYPHEN-2	Probably damaging		0,9
	NR		
GERP	5,3		

**Tabella 12:** Risultati analisi bioinformatica.

### **Paziente 8, IdZ766**

Paziente di sesso maschile con difficoltà della deambulazione e andatura di tipo spastica. L'elettromiografia (EMG) e l'elettroencefalografia (ENG) hanno evidenziato segni neurogeni ai 4 arti.

Su tali informazioni cliniche viene eseguita una diagnosi clinica di Paraparesi Spastica con esordio a 33 anni (gruppo SP). In sede di consulenza genetica, non viene evidenziata familiarità per la patologia (Fig.10A).



**Figura 10: (A)** Albero genealogico: (II:8, I:9) padre e madre sani. (III:2) probando. **(B)** Analisi di segregazione: (II:8) padre wild type per la variante c.2842dupG e eterozigote per la variante c.3291+3A>G; (II:9) madre eterozigote per la variante c.2842dupG e wild type per la variante c.3291+3A>G.

L'analisi di sequenziamento ha identificato due varianti in eterozigosi nel gene *SPG11* (NM\_025137.3). La variante di splicing c.3291+3A>G e la variante frameshift rs312262750 c.2842dupG p.Val948GlyfsTer6 che causa sostituzioni amminoacidiche a partire dalla posizione 948 della proteina fino la formazione di un codone di stop prematuro nel 6° codone successivo. Stevanin et al (2007) hanno descritto la variante frameshift rs312262750 c.2842dupG p.Val948GlyfsTer6 nel gene *SPG11* (NM\_025137.3) in associazione a paraplegia spastica ereditaria con modalità di ereditarietà autosomica recessiva. La patogenicità della variante è inoltre confermata dall'analisi di segregazione, dalla frequenza e dall'analisi bioinformatica. Questi ultimi tre criteri conferiscono patogenicità anche alla seconda variante identificata. Gli studi di segregazione mediante metodica sanger hanno identificato la variante c.3291+3A>G nel padre (II:8) e la variante rs312262750 c.2842dupG p.Val948GlyfsTer6 nella madre (II:9) entrambi quindi portatori sani di una delle due varianti (Fig10B).

La frequenza della variante c.3291+3A>G è pari a 0,000008 nella banca dati ExAC e non è nota nella banca dati 1000 Genomes. I punteggi ADA e RF del tool dbSNV sono stati entrambi di 0.9 confermando l'ipotesi che la variante possa essere causa di splicing

alternativo. L'NR del tool GERP è pari a 5.4 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.13).

VARIANTE	SPG11	c.3291+3A>G
	<b>FREQUENZA</b>	
ExAC	0,000008	
1000 Genomes	nr	
	<b>ADA Score</b>	<b>RF Score</b>
dbscSNV	0,9	0,9
	<b>NR</b>	
GERP	5,5	

**Tabella 13:** Risultati analisi bioinformatica.

La frequenza della variante rs312262750 c.2842dupG p.Val948GlyfsTer6 non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico è stata effettuata mediante il tool Polyphen-2 che ne conferisce un effetto dannoso. L'NR del tool GERP è pari a 5 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.14).

VARIANTE	SPG11	c.2842dupG	p.Val948GlyfsTer6	rs312262750
	<b>FREQUENZA</b>			
ExAC	0,000016			
1000 Genomes	nr			
	<b>Prediction</b>		<b>Accuracy</b>	
POLYPHEN-2	Possibly Damaging		0,9	
	<b>NR</b>			
GERP	5			

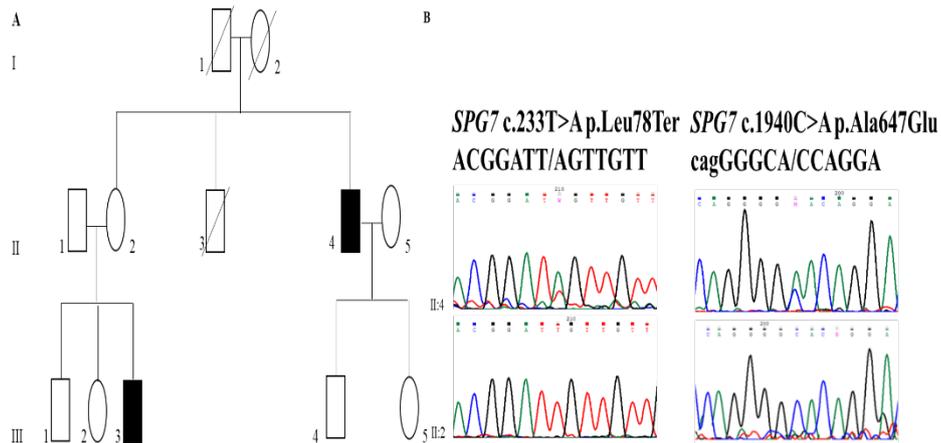
**Tabella 14:** Risultati analisi bioinformatica.

### **Paziente 9, IdZ261**

Paziente di sesso maschile con andatura atasso-spastica anserina, marcia su punte e talloni e postura instabile. Il paziente è stato sottoposto ad analisi strumentali quali, elettroencefalogramma (EEG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati visivi (PEV), potenziali evocati somatosensoriali (SEP) ai quattro arti e potenziali evocati motori (PEM) ai quattro arti. L'RM ha dato esito di atrofia cerebellare vermiana-paravermiana, ENG,

PEV e SEP sono risultati nella norma mentre i PEM sono risultati alterati sia centralmente che perifericamente.

Sulla base di queste osservazioni cliniche viene eseguita una diagnosi di sindrome atasso-spastica con esordio a 26 anni (gruppo AS). In sede di consulenza genetica non si evidenzia familiarità per la condizione patologica (Fig.11A).



**Figura 11: (A)** Albero genealogico: (II:4) probando, (II:2) sorella sana. **(B)** Analisi di segregazione il probando (II:4) eterozigote sia per la variante c.233T>A p.Leu78Ter che per la variante c.1940C>A p.Ala647Glu nel gene *SPG7* (NM\_003119.2) la sorella (II:2) “wild type” per entrambe le varianti.

L’analisi di sequenziamento ha identificato due varianti in eterozigosi nel gene *SPG7* (NM\_003119.2). La variante stop gained rs121918358 c.333T>A p.Leu78Ter che causa la formazione di un codone di stop prematuro nella posizione 78 della proteina, e la variante missenso c.1940C>A p.Ala647Glu che causa la sostituzione di una Alanina con un acido glutammico nella posizione 647 della proteina.

Arnoldi et al (2008) hanno descritto per la prima volta la variante rs121918358 c.333T>A p.Leu78Ter nel gene *SPG7* (NM\_003119.2) in associazione a soggetti affetti da paraplegia spastica in uno studio svolto nella popolazione italiana. Successivamente tale variante è stata descritta in soggetti con fenotipo misto atasso-spastico da Choquet et al (2016).

La patogenicità della variante è inoltre confermata dall'analisi di segregazione dalla frequenza e dall'analisi bioinformatica. Questi ultimi tre criteri conferiscono patogenicità anche alla seconda variante identificata. Gli studi di segregazione mediante metodica sanger hanno escluso la presenza di entrambe le varianti nella sorella sana (II:2) (Fig.11B).

La frequenza della variante rs121918358 c.333T>A p.Leu78Ter è pari a 0.0004 nella banca dati ExAC e non è nota sulla banca dati 1000 Genomes. Per l'analisi di predizione in silico è stato possibile consultare solo il tool Mutation taster che ne conferisce effetto patogeno.

L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5.3 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.15).

VARIANTE	SPG7	c.233T>A	p.Leu78Ter	rs121918358
ExAC	<b>FREQUENZA</b>			
	0,0004			
1000 Genomes	nr			
	<b>Prediction</b>			<b>Accuracy</b>
MUTATION TASTER	Disease causing automatic			1
	<b>mammalian</b>		<b>mamm. rankscore</b>	
phyloP20	1		0,7	
	<b>NR</b>			
GERP	5,3			

**Tabella 15:** Risultati analisi bioinformatica.

La frequenza della variante c.1940C>A p.Ala647Glu non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tutti e quattro i tool utilizzati (Mutation taster, SIFT, PROVEAN, e Polyphen-2.). L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (0,8) e un NR pari a 5 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.16).

VARIANTE	SPG7	c.1940C>A	p.Ala647Glu
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	<b>Prediction</b>	<b>Accuracy</b>	
MUTATION TASTER	Disease causing	1	
SIFT	Damaging	0	
PROVEAN	Damaging	-4,55	
POLYPHEN-2	Probably damaging	0,9	
	<b>mammalian</b>	<b>mamm. rankscore</b>	
phyloP20	0,8	0,3	
	NR		
GERP	5		

**Tabella 16:** Risultati analisi bioinformatica.

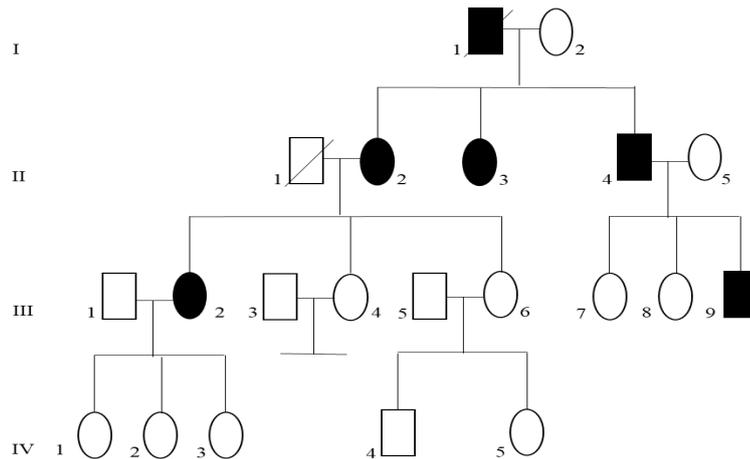
### →Varianti identificate in geni rari

Nel pannello sono stati presi in considerazione anche tutti quei geni di recente identificazione le cui varianti sono state spesso identificate in una o solo poche famiglie e pertanto i cui dati sulla penetranza e la frequenza non sono ancora noti.

#### **Paziente 10, IdZ578**

Paziente di sesso femminile con difficoltà deambulazione con andatura di tipo spastico. I potenziali evocati somatosensoriali (SEP) ed i potenziali evocati motori (PEM) ai quattro arti e la risonanza magnetica (RM) sono risultati nella norma. In sede di consulenza genetica si evidenzia familiarità per la condizione patologica nel fratello (III:9), mamma (II:2), zio (II:4), e il nonno materno(I:1). (Fig.12).

Sulla base della andatura e dell'anamnesi familiare viene eseguita una diagnosi di paraparesi spastica con esordio a 50 anni (gruppo SP).



**Figura 12:** Albero genealogico: (III:2) probanda.

L'analisi di sequenziamento ha identificato una variante missenso in eterozigosi nel gene *KCNC3* (NM\_004977) c.2197C>A p.Gln733Lys che causa la sostituzione di una glutammina con una lisina nella posizione 733. La patogenicità di tale variante rimane in dubbio. La frequenza della variante non è nota né sulla banca dati EXaC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico mediante i tool Mutation Taster e Polyphen2 ne conferisce un effetto neutrale mentre non ha dato esiti con i tool SIFT e PROVEAN.

L'analisi mediante il tool PhyloP20 ha dato un punteggio negativo (-0.07) evidenziando un tasso di evoluzione più veloce rispetto all'atteso e quindi tale amminoacido non risulta essere conservato nei mammiferi. L'NR del tool GERP ha un punteggio intermedio (2,3) pertanto il nucleotide ricade in un punto moderatamente conservato (Tab.17). Non è stato possibile effettuare analisi di segregazione.

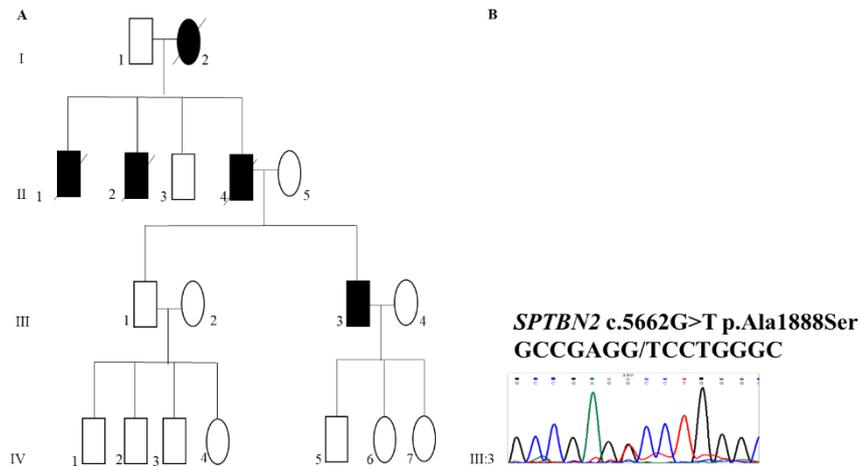
VARIANTE	KCNC3	c.2197C>A	p.Gln733Lys
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	Prediction	Accuracy	
MUTATION TASTER	Polymorphism	0,9	
SIFT	nr		
PROVEAN	nr		
POLYPHEN-2	Benign	0,9	
	mammalian	mamm. rankscore	
phyloP20	-0,07	0,9	
	NR		
GERP	2,3		

**Tabella 17:** Risultati analisi bioinformatica.

### **Paziente 11, IdZ436**

Paziente di sesso maschile con andatura regolare con lieve ipopendolarismo a destra.

L'elettroencefalografia (ENG) i potenziali evocati somatosensoriali (SEP) ai quattro arti ed i potenziali evocati motori (PEM) sono risultati nella norma. La risonanza magnetica (RM) ha evidenziato marcata atrofia cerebellare con dilatazione degli spazi subaracnoidei adiacenti. Sulla base di queste osservazioni cliniche viene eseguita una diagnosi di atassia spinocerebellare con esordio a 56 anni (gruppo AT). In sede di consulenza genetica si evidenzia familiarità per la condizione patologica nel padre (II:4), due zii (II:1; II:2), e nella nonna paterna (I:2), (Fig.13).



**Figura 13: (A)** Albero genealogico: (III:3) probando (Z436), (I:2; II:1, II:2, II:4) nonna due zii e padre con diagnosi clinica di atassia spinocerebellare. **(B)** Elettroferogramma della variante c.5662G>T p.Ala1888Ser nel gene *SPTBN2* (NM\_006946) del campione Z436.

L'analisi di sequenziamento ha identificato una variante in eterozigosi nel gene *SPTBN2* (NM\_006946) la variante missenso c.5662G>T p.Ala1888Ser, che causa una sostituzione di una alanina con una serina nella posizione 1888 della proteina. La presenza della variante è stata confermata mediante metodica sanger nel paziente (III:3) (Fig.13B).

La patogenicità della variante è stata attribuita in funzione della frequenza e dell'analisi bioinformatica. La frequenza della variante c.5662G>T p.Ala1888Ser non è nota né sulla banca dati ExAC né sulla banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tutti e quattro i tool utilizzati (Mutation taster, SIFT, PROVEAN, e Polyphen-2.). L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (0,9) e un NR pari a 4.7 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.18).

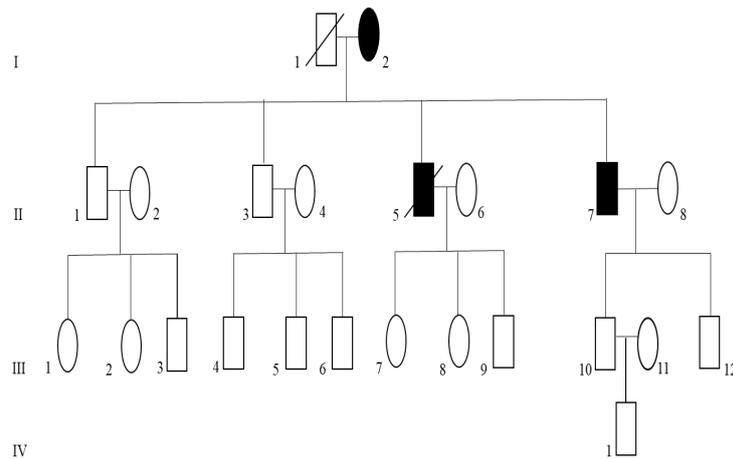
VARIANTE	<i>SPTBN2</i>	c,5662G>T	p.Ala1888Ser
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	Prediction	Accuracy	
MUTATION TASTER	Disease causing	1	
SIFT	Damaging	0	
PROVEAN	Damaging	-2,6	
POLYPHEN-2	Probably damaging	0,8	
	mammalian	mamm. rankscore	
phyloP20	0,9	0,4	
	NR		
GERP	4,7		

**Tabella 18:** Risultati analisi bioinformatica.

### **Paziente 12, IdZ450**

Paziente di sesso maschile con disturbo atassico, marcia allargata e disturbi dell'equilibrio. I potenziali evocati somatosensoriali (SEP) ed i potenziali evocati motori (PEM) ai quattro arti sono risultati nella norma. L'elettromiografia (EMG) ed l'elettroencefalografia (ENG) hanno evidenziato lieve sofferenza del nervo mediano al canale carpale bilateralmente, sofferenza neurogena cronica senza denervazione in atto a distribuzione (C5)-C6 bilaterale ed L5-S1 bilaterale. La risonanza magnetica (RM) ha evidenziato verosimile falda di igroma bilaterale.

Sulla base di queste osservazioni cliniche viene eseguita una diagnosi di atassia spinocerebellare con esordio a 52 anni (gruppo AT). In sede di consulenza genetica si evidenzia familiarità per la condizione nella madre (I:2), e nel fratello (II:5) (Fig.14)



**Figura 14:** Albero genealogico: (II:7) probando (Z450), (I:2; II:5) mamma e fratello con diagnosi clinica di atassia spinocerebellare.

L'analisi di sequenziamento ha identificato due varianti missenso in eterozigosi in due geni differenti. Nel gene AP5Z1 (NM\_014855) la variante missenso rs764235114 c.985G>A p.Val329Met che causa la sostituzione di una valina con una metionina nella posizione 329 della proteina. Nel gene GRM1 (NM\_001278064) la variante missenso c.3888G>C p.Asp1130His che causa la sostituzione di un acido aspartico con una istidina nella posizione 1130 della proteina. La patogenicità delle varianti è stata attribuita in funzione della frequenza e dell'analisi bioinformatica. La frequenza della variante rs764235114 c.985G>A p.Val329Met nel gene AP5Z1 è pari a 0.00007 nella banca dati ExAC e non nota nella banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tre dei quattro i tool utilizzati (Mutation taster, SIFT, e Polyphen-2.) la sola predizione del tool PROVEAN ne conferisce un effetto neutrale. L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5,4 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.19).

VARIANTE	AP5ZI	c.985G>A	p.Val329Met	rs764235114
ExAC	FREQUENZA			
	0,00007			
1000 Genomes	nr			
	Prediction		Accuracy	
MUTATION TASTER	Disease causing		0,9	
SIFT	Damaging		0,002	
PROVEAN	Neutral		-2,09	
POLYPHEN-2	Probably damaging		1	
	mammalian		mamm. rankscore	
phyloP20	1		0,7	
	NR			
GERP	5,4			

**Tabella 19:** Risultati analisi bioinformatica.

La frequenza della variante c.3888G>C p.Asp1130His nel gene *GRM1* non è nota né nella banca dati ExAC né nella banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tre dei quattro i tool utilizzati (Mutation taster, SIFT, e Polyphen-2.) la sola predizione del tool PROVEAN ne conferisce un effetto neutrale. L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5,8 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.20). Non è stato possibile eseguire uno studio di segregazione

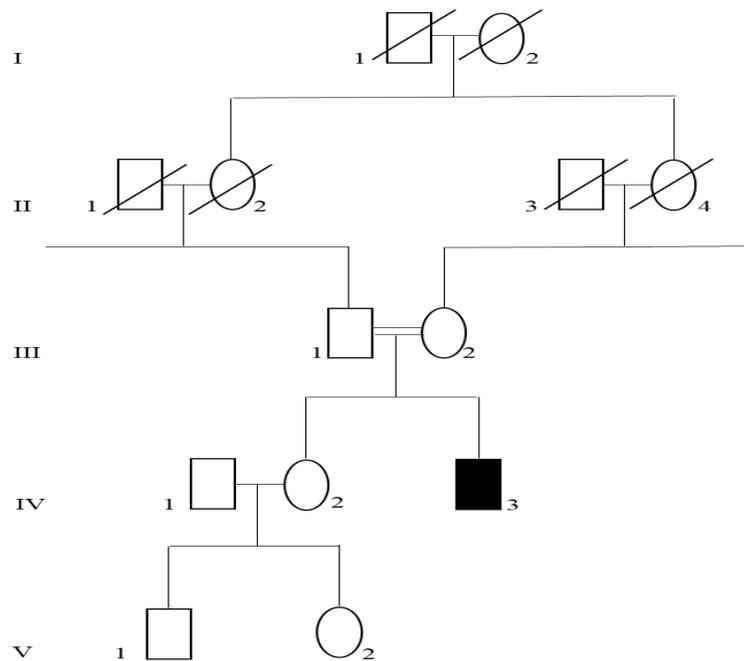
VARIANTE	GRM1	c.3388G>C	p.Asp1130His
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	Prediction		Accuracy
MUTATION TASTER	Disease causing		1
SIFT	Damaging		0,003
PROVEAN	Neutral		-1
POLYPHEN-2	Probably damaging		0,8
	mammalian		mamm. rankscore
phyloP20	1		0,7
	NR		
GERP	5,8		

**Tabella 20:** Risultati analisi bioinformatica.

### **Paziente 13, IdZ503**

Paziente di sesso maschile con disturbi della deambulazione tipo atassico ad andamento lentamente progressivo, disartria e disturbi della mobilità oculare. La risonanza magnetica evidenziava atrofia cerebellare. In funzione del quadro clinico viene eseguita diagnosi di atassia con esordio a 36 anni (gruppo AT).

In sede di consulenza genetica non si evidenzia familiarità per la condizione, ma si osserva consanguineità dei genitori (Fig.15).



**Figura 15:** Albero genealogico: (IV:3) probando, (III:1; III:2) mamma e padre sani consanguinei tra loro.

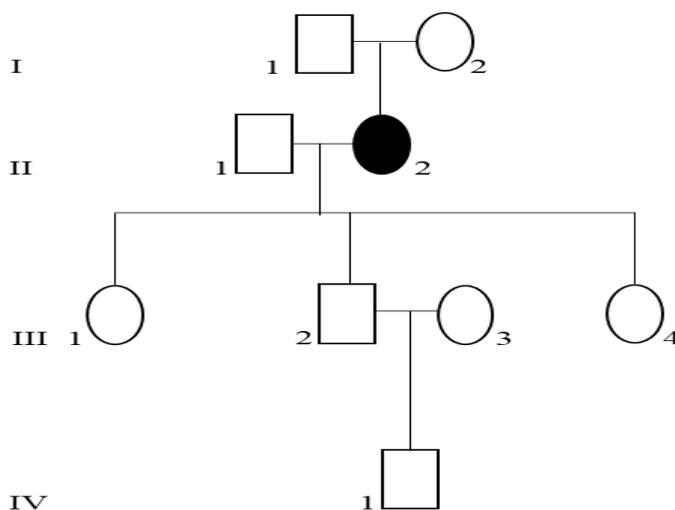
L'analisi di sequenziamento ha identificato, in omozigosi, la variante di splicing c.19893+2T>C nel gene *SYNE1* (NM\_182961). La frequenza della variante c.3291+3A>G non è nota né nella banca dati ExAC né nella banca dati 1000 Genomes. I punteggi ADA e RF del tool dbSCSNV sono stati entrambi di 0.9 confermando l'ipotesi che la variante possa essere causa di splicing alternativo. L'NR del tool GERP è pari a 5.6 indicando che il nucleotide wild type è conservato nei mammiferi (Tab.21). Non è stato possibile eseguire uno studio di segregazione.

<b>VARIANTE</b>	<b>SYNE1</b>	c.19893+2T>C
<b>FREQUENZA</b>		
<b>ExAC</b>	nr	
<b>1000 Genomes</b>	nr	
	<b>ADA Score</b>	<b>RF Score</b>
<b>dbSNV</b>	0,9	0,9
<b>NR</b>		
<b>GERP</b>	5,6	

**Tabella 21:** Risultati analisi bioinformatica.

### **Paziente 14, IdZ770**

Paziente di sesso femminile con andatura di tipo atassico e disturbi dell'equilibrio. La risonanza magnetica mostra un quadro significativo di atrofia cerebellare emisferica e soprattutto vermiana di tipo aspecifico. Viene eseguita diagnosi di atassia spinocerebellare con esordio a 52 anni (gruppo AT). In sede di consulenza genetica non si evidenzia familiarità per la condizione patologica. (Fig.16)



**Figura 16:** Albero genealogico: (II:2) probanda.

L'analisi di sequenziamento ha identificato, in eterozigosi nel gene *SPTBN2* (NM\_006946) la variante missenso c.830A>G p.Tyr277Cis che causa una sostituzione di una tirosina con una cisteina nella posizione 277 della proteina. La patogenicità della variante è stata attribuita in funzione della frequenza e dell'analisi

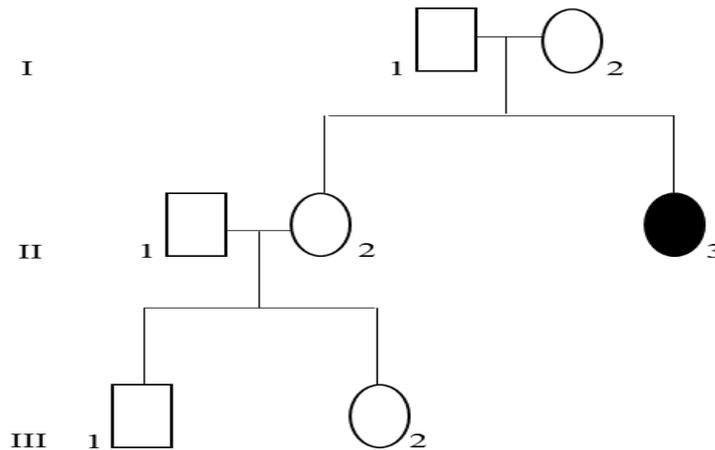
bioinformatica. La frequenza della variante c.830A>G p.Tyr277Cis non è nota né nella banca dati ExAC né nella banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tutti e quattro i tool utilizzati (Mutation taster, SIFT, e Polyphen-2 PROVEAN.) L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5,3 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.22).

VARIANTE	<i>SPTBN2</i>	<i>c.830A&gt;G</i>	<i>p.Tyr277Cys</i>
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	<b>Prediction</b>	<b>Accuracy</b>	
<b>MUTATION TASTER</b>	Disease causing	1	
<b>SIFT</b>	Damaging	0	
<b>PROVEAN</b>	Damaging	-4,55	
<b>POLYPHEN-2</b>	Probably damaging	0,9	
	<b>mammalian</b>	<b>mamm. rankscore</b>	
<b>phyloP20</b>	<b>1</b>	0,8	
	NR		
<b>GERP</b>	5,3		

**Tabella 22:** Risultati analisi bioinformatica.

### **Paziente 15, IdZ781**

Paziente di sesso femminile con marcata disartria e disfagia stazione eretta e deambulazione non valutabili in quanto permanente costretta sulla sedia a rotelle, moderata ipotrofia delle mani, ipotonia muscolare, ritardo nello sviluppo psico-fisico, epilessia sin da bambina ed intervento di cataratta bilaterale a 7 anni. La risonanza magnetica (RM) ha evidenziato marcata atrofia del verme e del cervelletto, dei peduncoli cerebellari, delle strutture ponto-mesencefaliche con una iperintensità delle vie spino-cerebellari bilateralmente. Viene eseguita diagnosi di atassia spino-cerebellare con esordio a 38 anni. In considerazione del quadro clinico considerata come fenotipo complesso (gruppo COMP). In sede di consulenza genetica non si evidenzia familiarità per la condizione patologica. (Fig.17).



**Figura 17:** Albero genealogico: (II:3) probanda (Z781).

L'analisi di sequenziamento ha identificato, in eterozigosi, due varianti missenso nel gene *LAMB1* (NM\_002291). La variante missenso c.3317T>A p.Met1106Lys che causa la sostituzione di una metionina con una lisina nella posizione 1106 della proteina, e la variante missenso c.313C>G p.Pro105Ala che causa la sostituzione di una prolina con una alanina nella posizione 105 della proteina. Le varianti sono state considerate patogene in funzione della frequenza e dell'analisi bioinformatica. La frequenza della variante c.3317T>A p.Met1106Lys non è nota né nella banca dati ExAC né nella banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tre dei quattro i tool utilizzati (Mutation taster, SIFT, e Polyphen-2.) il solo tool PROVEAN la predice come neutrale. L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.23).

VARIANTE	LAMBI	c.3317T>A	p.Met1106Lys
ExAC	FREQUENZA		
	nr		
1000 Genomes	nr		
	Prediction		Accuracy
MUTATION TASTER	Disease causing		1
SIFT	Damaging		0
PROVEAN	Neutral		1,3
POLYPHEN-2	Probably damaging		0,8
	mammalian		mamm. rankscore
phyloP20	1		0,9
	NR		
GERP	5		

**Tabella 23:** Risultati analisi bioinformatica.

La frequenza della variante c.313C>G p.Pro105Ala è pari a 0,000008 nella banca dati ExAC e non nota nella banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tutti i quattro i tool utilizzati (Mutation taster, SIFT, PROVEAN e Polyphen-2.) L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.24).

VARIANTE	LAMBI	c,313C>G	p.Pro105Ala
ExAC	FREQUENZA		
	0,000008		
1000 Genomes	nr		
	Prediction		Accuracy
MUTATION TASTER	Disease causing		1
SIFT	Damaging		0
PROVEAN	Damaging		-4,13
POLYPHEN-2	Probably damaging		0,9
	mammalian		mamm. rankscore
phyloP20	1		0,7
	NR		
GERP	5		

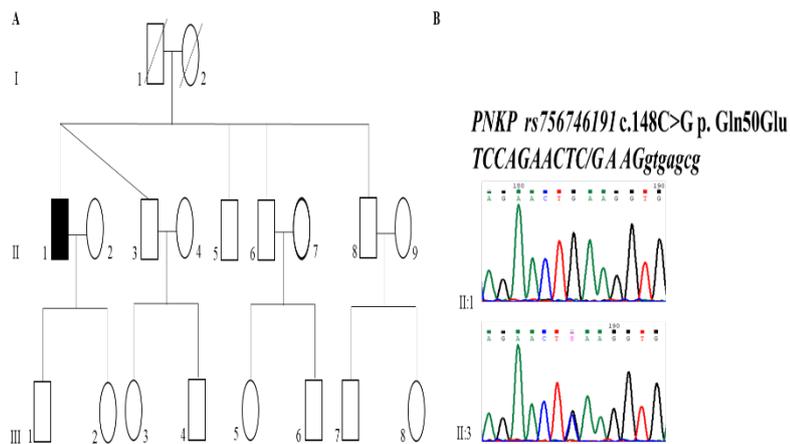
**Tabella 24:** Risultati analisi bioinformatica.

### **Paziente 16, IdZ1107**

Paziente di sesso maschile con disturbo dell'andatura di tipo atassico, aprassia oculomotoria, neuropatia demielinizzante sensitivo motoria.

Vengono eseguite indagini strumentali quali, elettroencefalografia (ENG) ed elettromiografia (EMG) ai quattro arti, risonanza magnetica (RM) dell'encefalo, potenziali evocati somatosensoriali (SEP) e potenziali evocati motori (PEM) ai quattro arti. L'EMG-ENG hanno evidenziato grave polineuropatia sensitivo-motoria di tipo assonale. I PEM hanno evidenziato tempo di conduzione centrale aumentato ai quattro arti. L'RM encefalo ha dato esito di non significative alterazioni dell'intensità di segnale nei tessuti dell'encefalo. Gli spazi periencefalici della base sono apparsi ampliati, mentre più lieve è apparsa la dilatazione di quelli della volta. Infine si è evidenziato un alterato segnale a livello dell'acquedotto cerebrale, come per accelerazione di flusso.

Viene eseguita diagnosi di sindrome spino-cerebellare a fenotipo complesso con esordio a 54 anni (gruppo COMP). In sede di consulenza genetica non si evidenzia familiarità per la condizione patologica. (Fig.18).



**Figura 18: (A)** Albero genealogico: (II:1) probando, (II:3) fratello gemello sano. **(B)** Analisi di sequenza: (II:1) il probando omozigote per la variante c.148C>G p.Gln50Glu nel gene PNKP (NM\_007254), (II:3) fratello eterozigote per la variante c.148C>G p.Gln50Glu nel gene PNKP (NM\_007254).

L'analisi di sequenziamento ha identificato nel gene *PNKP* (NM\_007254). in omozigosi la variante missenso rs756746191 c.148C>G p.Gln50Glu che causa la sostituzione di una glutammina

con un acido glutammico nella posizione 50 della proteina. Tale variante è stata considerata patogenetica in funzione dell'analisi di segregazioni della frequenza e dell'analisi bioinformatica. Gli studi di segregazione mediante metodica sanger hanno identificato la variante in eterozigosi nel suo fratello gemello (II:3) asintomatico (Fig.21).

La frequenza della variante rs756746191 c.148C>G p.Gln50Glu non è nota né nella banca dati ExAC né nella banca dati 1000 Genomes. L'analisi di predizione in silico ne conferisce un effetto dannoso in tre dei quattro i tool utilizzati (Mutation taster, SIFT, e Polyphen-2.) il solo tool PROVEAN la predice come neutrale. L'analisi mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 5,4 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.25).

VARIANTE	PNKP	c.148C>G	p.Gln50Glu	rs756746191
ExAC	FREQUENZA			
	nr			
1000 Genomes	nr			
	Prediction		Accuracy	
MUTATION TASTER	Disease causing		0,9	
SIFT	Damaging		0,007	
PROVEAN	Neutral		-2,11	
POLYPHEN-2	Probably damaging		0,9	
	mammalian		mamm. rankscore	
phyloP20	1		0,7	
	NR			
GERP	5,4			

**Tabella 25:** Risultati analisi bioinformatica.

## 4.2 Esoma clinico

L'analisi di sequenziamento mediante esoma clinico è stata condotta su 16 pazienti risultati negativi all'analisi di sequenziamento mediante pannello NGS target, per i quali erano disponibili campioni biologici di familiari. L'analisi permette il sequenziamento delle regioni codificanti e delle giunzioni esone-introne di 4493 ad oggi noti per essere associati a patologia. In media il 98% delle regioni target sono state sequenziate con una copertura del 20X. L'analisi ha portato all'identificazione di una media di 19000 varianti a campione di cui in media 6000 sono le varianti che hanno passato i principali filtri qualitativi (Qscore  $\geq 30$  e coverage  $\geq 20X$ ). L'analisi dei dati è stata effettuata mediante la piattaforma SOPHiA DDM (SOPHiA genetics). In primo luogo, sul software di analisi è stato impostato un filtro "custom" che escludesse tutte le varianti introniche ad eccezione di quelle situate a meno di 25bp dall'esone, tutte le varianti sinonime e tutte le varianti presenti in ExAC con una frequenza  $\geq 0.001$ . Un ulteriore filtro si è basato sull'esclusione di tutte le varianti identificate nei precedenti esperimenti di CES in soggetti con fenotipo differente da quello in studio. Sono, poi, state escluse tutte le varianti presenti nel familiare sano, qualora analizzato e incluse tutte quelle presenti nel familiare affetto, qualora analizzato. Sono state prese in considerazione le sole varianti in cui almeno due delle analisi di predizione in silico da parte dei tool indagati dal software (Mutation taster, Polyphen2, SIFT) ne conferissero un effetto patogeno. Infine sono state incluse le varianti che avessero una frequenza  $\leq 0.001$  anche nelle ulteriori due banche dati consultate dal software (1000 genomes ed exome variant server).

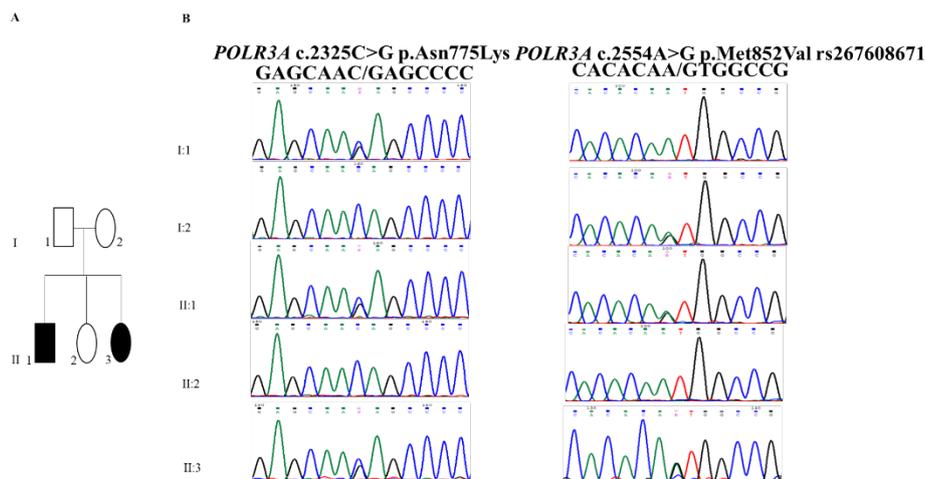
L'analisi mediante SOPHiA DDM (SOPHiA genetics) è stata integrata dall'analisi mediante i tool di allineamento multiplo PhyloP20 e GERP per valutare il grado di conservazione del nucleotide e dell'amminoacido wild type soggetto a variazione.

Questo procedimento ha permesso l'identificazione di varianti che potessero spiegare la patogenesi in due pazienti dei 16 analizzati.

#### 4.2.1 Descrizione genetica e clinica dei pazienti

##### **Paziente 1, IdZ232**

Paziente di sesso femminile con diagnosi di leucodistrofia ipomielinizzante (HL) con esordio della patologia a 19 anni, caratterizzato da un progressivo e lento ritardo cognitivo a disturbi dell'andatura e amenorrea. Attualmente, a 34 anni, soffre di una grave regressione cognitiva e costretta all'uso fisso della sedia a rotelle. In sede di consulenza è emersa familiarità per la condizione nel fratello (II:1) (Fig.19A).



**Figura 19:** (A) Albero genealogico: (I:1, I:2) padre e madre sani. (II:1, II:3), fratelli affetti; (II:2) sorella sana. (B) Analisi di segregazione: (I:1) padre eterozigote per la variante p.Asn775Val e wild type per la variante p.Met852Val; (I:2) wild type per la variante p.Asn775Val ed eterozigote per la variante p.Met852Val, (II:1, II:3) fratelli affetti eterozigoti per entrambe le varianti, (II:2) sorella sana wild type per entrambe le varianti.

La diagnosi nel fratello maschio è stata eseguita mediante risonanza magnetica effettuata in seguito ad un grave episodio di cefalea all'età di 41 anni. Al momento ha 46 anni e ha sviluppato un lieve deficit cognitivo, distonia con ipotonia a 4 arti e moderata

dismetria con compromissione dell'equilibrio e dell'andatura. Dall'analisi contemporanea dei due fratelli sono state identificate due varianti in eterozigosi nel gene *POLR3A* (NM\_007055). La variante missenso rs267608671 c.2554A>G p.Met852Val e la variante missenso c.2325C>G p.Asn775Lys. La frequenza della variante p.Met852Val è pari a 0.0082 nella banca dati ExAC mentre non è nota in ESP e 1000 genomes. I tre siti di predizione in silico (Mutation taster, Polyphen-2 e SIFT) ne conferiscono un effetto patogeno. Inoltre è stata precedentemente descritta in letteratura da Bernard et al (2011) in associazione a pazienti affetti da leucodistrofia. Essi la identificarono in eterozigosi composta con due differenti varianti. Con la variante p.Arg873AlafsX878 in due pazienti di sesso femminile con età di insorgenza a 12 anni con segni cerebellari e regressione cognitiva senza ritardo dello sviluppo. Con la variante p.Arg140X in una paziente affetta da leucodistrofia con oligodonzia, ed età di esordio intorno ai 3 anni. La frequenza della variante p.Asn775Lys non è nota in banca dati (ExAC, ESP, 1000genomes), ed i tre siti di predizione in silico (Mutation Taster, SIFT e Polyphen-2) ne conferiscono un effetto patogeno.

Gli studi di segregazione mediante metodica sanger hanno permesso di stabilire che ciascuna delle due varianti è stata eredita da uno dei due genitori sani. Infatti la variante p.Asn775Lys è stata identificata in eterozigosi nel padre (I:1) e la variante p.Met852Val è stata identificata in eterozigosi nella mamma (I:2) mentre nessuna delle due varianti è stata identificata nella sorella sana (II:2) (Fig.19B).

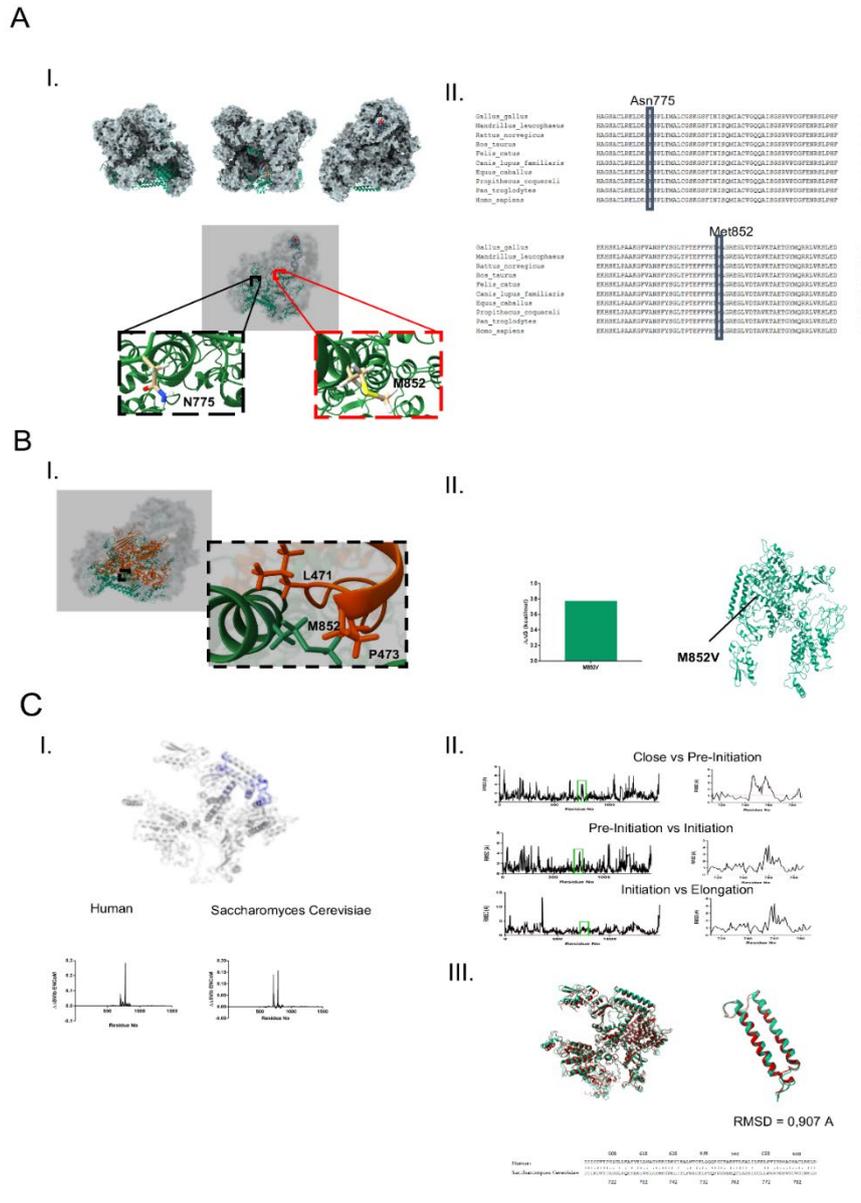
Al fine di confermare ulteriormente il ruolo patogenico delle due varianti identificate è stata eseguita un'analisi bioinformatica. Mediante I-TASSER server è stata ottenuta la struttura tridimensionale della proteina RPC1 (RNA polymerase subunit) codificata dal gene *POLR3A* (struttura non presente nella banca dati PDB) (Fig.23AI). Le analisi di allineamento multiplo su 10

specie di mammiferi evidenziano l'importanza dei residui Asn775 e Met852 nella funzionalità della proteina (Fig20AII). Il residuo Met852 è situato sulla superficie della subunità RPC1, nelle vicinanze dell'interfaccia RPC2 (Fig20BI), permettendo quindi di ipotizzare che la mutazione p.Met852Val possa compromettere la stabilità del legame RPC1-RPC2. Tale ipotesi è stata dimostrata eseguendo l'analisi strutturale del complesso. Il valore  $\Delta\Delta G$  previsto del mutante p.Met852Val era 0,77 kcal / mol (Fig20BII). Tale valore implica che la mutazione del sito potrebbe ridurre l'interazione della proteina RPC1-RPC2, compromettendo così la corretta funzione dell'RNA Pol III.

Infine abbiamo analizzato gli effetti della mutazione p.Asn775Lys sulla isomerizzazione della RNA polimerasi III. La trascrizione è un processo a più fasi che richiede una frequente isomerizzazione strutturale dell'intero complesso della RNA polimerasi pertanto RPC1 deve essere abbastanza flessibile da sostenere tutti i cambiamenti conformazionali che si susseguono durante questo processo. Per verificare l'ipotesi che la mutazione p.Asn775Lys potesse compromettere la plasticità RPC1, abbiamo eseguito un'analisi strutturale usando il sito DynaMut (Fig.20CI). I risultati mostrano una riduzione della flessibilità RCP1 localizzata tra i residui 698-774. Per capire se questa regione può essere richiesta per l'isomerizzazione, abbiamo analizzato l'RNA Pol III del lievito *Saccharomyces cerevisiae*, confrontando la struttura della subunità RPC1 durante il reclutamento (6F42), la pre-iniziazione (6EU0), l'iniziazione (6CNB) e l'allungamento (5FJ8) (Fig.20CII).

Per prima cosa abbiamo generato il mutante p.Asn775Lys nella corrispondente posizione della subunità RPC1 di *Saccharomyces Cerevisiae*. Come previsto, il mutante p.Asn775Lys mostra una riduzione della flessibilità localizzata nella stessa regione (712-788). Quindi, abbiamo analizzato il RMSD cercando il picco di fluttuazione all'interno della regione interessata dalla mutazione p.Asn775Lys. Si sono osservati cambiamenti conformazionali

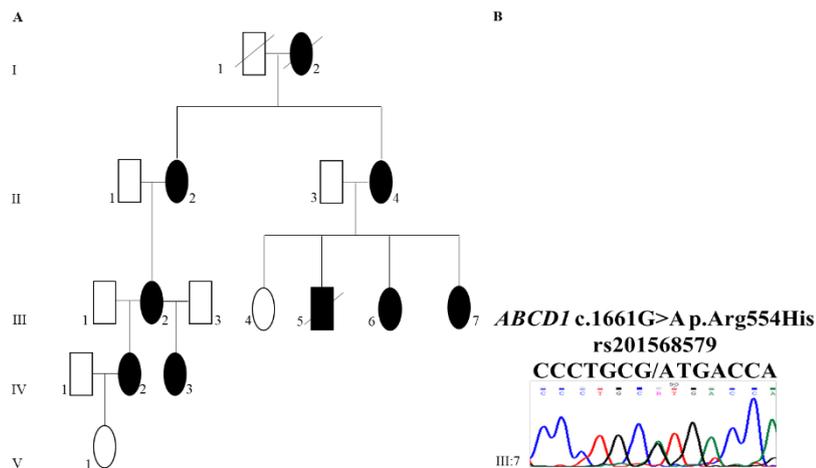
significativi tra la fase di pre-iniziazione / iniziazione e quella di inizio / allungamento, suggerendo che la mutazione p.Asn789Lys potrebbe ridurre la flessibilità di RPC1, compromettendo quindi la corretta funzione di RNA Pol III. Dal momento che esiste una similarità del 77,3% tra la sequenza primaria umana e *Saccharomyces Cerevisiae* di RPC1, considerando la topologia simile tra loro (RMSD = 0,907 Å), possiamo supporre che la mutazione p.Asn775Lys possa influenzare la RPC1 umana nello stesso modo.



**Figura 20.** Analisi bioinformatica delle varianti rs267608671 c.2554A>G p.Met852Val e c.2325C>G p.Asn775Lys nel gene *POLR3A*.

### **Paziente 2, IdZ433**

Paziente di sesso femminile con disturbo della deambulazione con andatura spastica. In sede di consulenza è emersa familiarità per la condizione [sorella (IV:3), mamma (II:2) due cugine materne (III:6, III:7), nonna (II:2), sorella della nonna (II:4) bisnonna materna (I:2)]. (Fig.21A).



**Figura 21: (A)** Albero genealogico: (I:2, II:2, II:4, III:2, IV:3) soggetti con diagnosi clinica di paraparesi spastica. (IV:2, III:6, III:7), probanda e zie di secondo grado in cui è stata identificata la variante rs201568579. **(B)** Analisi di sequenza: sequenza della variante rs201568579 effettuata mediante metodica sanger nella paziente (III:7).

Ai fini dell'inquadramento diagnostico è stato possibile comparare i quadri clinici della paziente (IV:2) e delle due zie materne (III:6, III:7) (Tab.26).

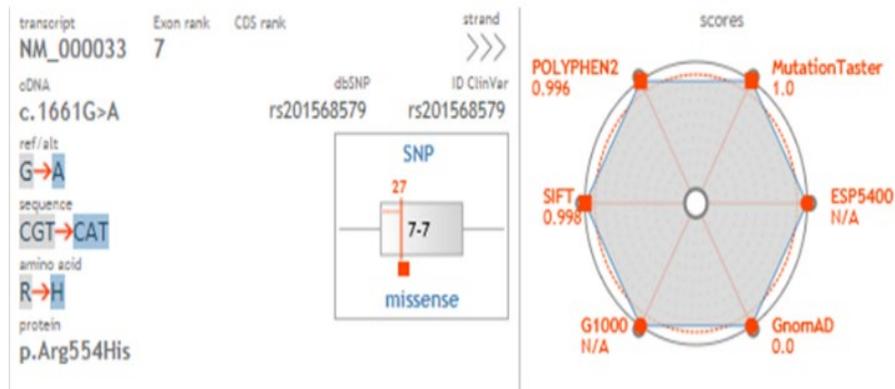
ID	ETÀ ESORDIO	RM	PEV	PEM		SEP		EMG-ENG	
				AASS	AAII	AASS	AAII	AASS	AAII
<b>Z433</b> <b>(IV:2)</b>	35	+	-	-	+	-	+	-	-
<b>Z203</b> <b>(III:6)</b>	40	+	-	-	+	-	+	+	-
<b>Z137</b> <b>(III:7)</b>	38	+	-	-	+	-	+	-	-

**Tabella 26:** In tabella sono stati paragonati gli esiti delle analisi strumentali e l'età di esordio delle tre pazienti (IV:2, III:6, III:7). (+: esito positivo, -: esito negativo, RM: risonanza magnetica, PEM: potenziali evocati motori, SEP: potenziali evocati somatosensoriali, EMG-ENG: elettromiografia-elettroencefalografia, AASS: arti superiori, AAll: arti inferiori).

Nelle tre pazienti l'età di esordio è compresa tra i 35 e i 40 anni. La risonanza magnetica è compatibile con una diagnosi di adrenoleucodistrofia. I PEM hanno un tempo di conduzione motoria centrale aumentato bilateralmente agli arti inferiori. I SEP evidenziano un aumento di latenza delle risposte corticali bilateralmente e tempo di conduzione sensitiva centrale aumentato bilateralmente agli arti inferiori. L'EMG-ENG sono risultate alterate, agli arti superiori in una delle due zie materne (III:6) in cui si evidenziava neuropatia canalicolare del mediano al polso bilateralmente di grado lieve.

L'analisi dell'esoma clinico è stata condotta in contemporanea nella paziente (IV:2) e in una delle due zie materne (III:6). È stata identificata la variante missenso in eterozigosi nel gene *ABCD1* (NM\_000033) rs201568579 c.1661G>A p.Arg554His che causa la sostituzione di un'arginina con una istidina nella posizione 554 della proteina. La presenza della variante è stata confermata mediante metodica sanger in III:7 (Fig.21B). La patogenicità della variante è stata confermata dall'analisi di segregazione in quanto presente in

tutti gli affetti (IV2, III:6, III:7) e in funzione della frequenza e dell'analisi bioinformatica. La frequenza di tale variante non è annotata nelle tre banche dati consultate (ESP Gnomad, 1000 genomes). Tutti i tre siti di predizione in silico (Mutation Taster, Polyphen 2 e SIFT) indagati dal software ne conferiscono un effetto patogeno (Fig.22).



**Figura 22** Interpretazione delle variante p.Arg554His nel gene ABCD1 mediante il software Sophia DDM. A sinistra è descritta la variante, codice del trascritto in cui è stata identificata, localizzazione esonica, posizione nella porzione codificante, cambio nucleotidico, cambio del codone, cambio proteico ed eventuale codice identificativo annotato nel dbSNP. A destra interpretazione grafica della patogenicità della variante data dal punteggio di patogenicità dei siti di predizione Mutation taster, Polyphen-2 e SIFT e l'annotazione della frequenza riportata nelle tre banche dati consultate Exome Variant Server (ESP5400), GnomAD, 1000 Genomes (G1000). (N/A= non rilevato).

L'analisi integrativa mediante i tool PhyloP20 e GERP hanno dato rispettivamente un punteggio positivo (1) e un NR pari a 4.7 indicando che sia il nucleotide wild type che l'amminoacido sono conservati nei mammiferi (Tab.27).

VARIANTE	<i>ABCDI</i>	<i>c.1661G&gt;A</i>	p.Arg554His	rs201568579
	mammalian		mamm. Rankscore	
phyloP20	1		0,8	
	NR			
GERP	5,3			

**Tabella 27:** Risultati analisi dei tool di allineamento multiplo.

## 5.DISCUSSIONE

Le malattie neurologiche e neurodegenerative sono un gruppo eterogeneo di patologie caratterizzate, in molti casi, da una forte eterogeneità clinica (età di esordio, progressione, sintomatologia) e genetica. Ne sono un esempio le atassie ereditarie e le paraplegie spastiche, un gruppo di malattie estremamente eterogeneo sia dal punto di vista fenotipico che da quello genotipico.

Le manifestazioni cliniche di queste patologie, pur condividendo alcune caratteristiche comuni, presentano uno spettro fenotipico molto ampio, ed in certi casi anche una differente sintomatologia che rende difficile l'individuazione delle mutazioni coinvolte. Dal punto di vista del genotipo, nell'eziopatogenesi sono coinvolti più di 100 geni che presentano tutte le modalità di trasmissione genetica (dominante, recessiva, legata all'X e mitocondriale). Pertanto, tale componente è spesso di difficile inquadramento. Se da un lato in molti di questi pazienti non si riesce ad identificare tale componente, altri invece ricevono una diagnosi clinica nella quale viene identificato il gene responsabile del fenotipo. L'identificazione di questi geni ha contribuito nel tempo ad una migliore comprensione dei pathway molecolari maggiormente coinvolti.

Ad oggi sono note numerose forme di atassie e spasticità geneticamente determinate. Inoltre, sono tantissime le malattie neurologiche rare in cui le atassie e la spasticità sono aspetti clinici di quadri molto più complessi. Oltre alla multigenicità, in tali pazienti una analisi genetica deve anche tener conto di altri aspetti che complicano l'indagine molecolare quali penetranza incompleta, e presenza di mutazioni "private", ovvero descritte solamente in alcune famiglie. Oltretutto, la difficoltà nell'eseguire studi multicentrici con pazienti selezionati in modo omogeneo, rende difficile sviluppare dei database specifici in cui siano riportate le varianti in questi geni, ed il fenotipo clinico di cui sono responsabili. A causa di questa grande eterogeneità, un'analisi mutazionale con il tradizionale sequenziamento Sanger risulta essere laboriosa e costosa, e spesso poco risolutiva. Oltre ad un discorso di scarsa

detection rate delle varianti, c'è anche il discorso delle lunghe tempistiche che vengono richieste, e che talvolta rendono inutile l'utilizzo di queste indagini nei percorsi clinici.

Le tecnologie di tipo NGS (Next Generation Sequencing) hanno risolto la problematica della "poligenicità", rendendo possibile l'analisi contemporanea di più geni, consentendo anche l'analisi di tutto il genoma. Tali tecnologie hanno anche risolto il problema delle tempistiche. La loro rapidità sta facendo sì che la diagnostica molecolare stia diventando sempre più un punto cardine per l'inquadramento clinico di pazienti con patologie neurologiche e neurodegenerativa, ancor più in casi complessi in cui la clinica non riesce a garantire un esaustivo inquadramento diagnostico.

Ad oggi, pannelli NGS accuratamente disegnati e validati sono entrati a far parte della pratica clinica. La loro applicazione nelle patologie neurologiche è comunque molto complessa perché richiede una accurata selezione dei geni da analizzare. Inoltre, considerata l'elevata eterogeneità allelica, risulta difficile stabilire una corretta detection rate (% dei positivi identificati), e quindi sviluppare dei saggi ad-hoc di facile utilizzo in diagnostica molecolare.

Da queste osservazioni è nata la necessità di sviluppare delle soluzioni diagnostiche basate su tecnologia NGS per l'analisi di fenotipi atassici e/o spastici e di fenotipi neurologici più complessi in cui la spasticità e l'atassia sono solo una delle componenti del quadro sindromico.

In questo studio sono stati analizzati 78 pazienti, reclutati presso l'Istituto IRCCS Neuromed. Ogni paziente è stato sottoposto a visita neurologica, consulenza genetica, ed analisi strumentali. Sono stati inclusi nello studio solamente pazienti negativi alla analisi genetica di primo livello per espansioni di triplette all'interno dei geni *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7* e *FXN* (variazioni frequenti e non identificabili mediante sequenziamento).

In base a tutto questo, i pazienti sono stati reclutati e suddivisi in quattro sottogruppi: Atassici (AT), Spastici (SP), Atasso/spastici

(AS), Fenotipi complessi (COMP). Per ogni paziente, ove possibile, è stata valutata l'eventuale familiarità della patologia (ulteriore suddivisione in sporadici vs familiari).

Dei 78 pazienti, 20 appartengono al gruppo Atassico (AT), 42 al gruppo spastico (SP), 2 al gruppo atasso-spastico (AS) e 14 al gruppo con fenotipo complesso con componente atasso-spastica (COMP). Dei 78 pazienti, il 42% ha una storia familiare (33/78) mentre il restante 58% non ha una anamnesi familiare, e viene considerata quindi sporadica (45/78).

Il DNA dei 78 pazienti è stato sottoposto ad analisi di sequenza massiva delle regioni codificanti e delle giunzioni esone-introne di 105 geni responsabili di atassia e/o paraplegia spastica accuratamente selezionati sulla base della letteratura scientifica e di database di riferimento.

Il saggio sviluppato ha consentito di raggiungere livelli soddisfacenti di qualità. Infatti, delle circa 1000 varianti identificate per ogni paziente, oltre il 98% sono qualitativamente soddisfacenti (Qscore  $\geq 30$  ed un coverage  $\geq 50X$ ).

Per considerare una variante patologica, sono stati utilizzati dei criteri molto stringenti. Infatti, doveva presentare almeno due dei seguenti criteri: (I) Descritta in letteratura e/o database di riferimento in associazione a fenotipo atassico e/o spastico; (II) Bassa frequenza nella popolazione generale dell'allele, (III) Classificata come dannosa in almeno due su tre siti di predizione in-silico, (IV) Localizzata in una regione cromosomica la cui sequenza nucleotidica è conservata nelle specie, (V) Variante de novo o ereditata da un parente affetto, (VI) Il fenotipo clinico correla con quello riportato in letteratura relativamente al gene in questione. In funzione di tali criteri sono state identificate 23 varianti potenzialmente patogene in 16 pazienti.

Le varianti identificate sono state riscontrate in 9 geni responsabili di atassia e/o paraplegia spastica e una variante è stata identificata nel gene *LAMB1* responsabile di Lissencefalia di tipo 5 in cui

l'atassia e la spasticità possono essere sintomi di un quadro sindromico più complesso.

Il 29% (6/23) delle varianti è stata identificata all'interno del gene *SPAST* confermando i dati di letteratura in cui tale gene risulta il più frequentemente mutato in pazienti con fenotipo spastico puro. Il 19% (4/23) delle varianti sono state identificate nel gene *SPG11* il 10% (2/23) nei geni *SPG7*, *SPTBN2* e *LAMB1* infine 1 variante è stata identificata nei geni *AP5Z1*, *GRM1*, *SYNE1*, *KCNC3* e *PNKP*.

È stata identificata una variante nel 20,5% dei pazienti (16 di 78). Il maggior contributo di pazienti con mutazione identificata è fornito dal gruppo SP, (56 %, 9/16), in cui i casi familiari sono molto più rappresentati degli sporadici (sporadici 3/27, 11,1%, familiari 6/15, 40%), seguito in ordine dai AT (25%, 4 /16), in cui gli sporadici sono più rappresentati dei familiari (sporadici 3/8, 37,5%, familiari 1/12, 8,3%). Questo dato suggerisce che questo percorso diagnostico deve tenere in considerazione, soprattutto per pazienti affetti da Atassie Spinocerebellare, anche i casi in cui non vi è una spiccata segregazione familiare della patologia.

Considerando le varie sottoclassi cliniche (non considerando il gruppo AS con un numero statisticamente non significativo), la percentuale dei pazienti positivi oscilla da un minimo di 14% dei pazienti COMP (2 di 14), ad un massimo di 21,5% dei pazienti SP (9 di 42), passando per il 20% dei pazienti AT (4 di 20). Questo suggerisce che un approccio con 105 geni non risponde a tutte le domande cliniche necessarie in casi più complessi di difficile inquadramento. Per questi, è più utile applicare sin da subito pannelli più estesi, come l'esoma clinico.

In questo studio, sono stati selezionati 16 pazienti negativi alla prima analisi con 105 geni, e per i quali si disponeva di altri soggetti affetti e sani all'interno della stessa famiglia. Questi sono stati analizzati tramite esoma clinico. È stato possibile eseguire una diagnosi molecolare in 2 pazienti, che alla fine hanno mostrato due fenotipi in cui la spasticità e la atassia sono degli aspetti clinici in quadri molto complessi, come la Leucodistrofia demielinizzante

autosomica recessiva di tipo 7 e la Adrenoleucodistrofia legata all'X (X-ALD).

L'utilizzo di questi pannelli, sia pannelli multigenici che esoma clinico, può avere diverse utilità diagnostiche. Infatti, mutazioni in alcuni geni servono unicamente ad avvalorare la diagnosi clinica. Al contrario, in altre situazioni le varianti identificate sono fondamentali per un corretto inquadramento diagnostico. In alcuni casi, le varianti identificate servono ad eseguire la diagnosi clinica, soprattutto nei casi in cui le informazioni cliniche e strumentali sono contrastanti o non significative. Di seguito vengono descritti degli esempi per ognuna delle classi.

a. Analisi genetiche di conferma della diagnosi clinica in Paraparesi spastiche Autosomico-dominanti (HSP-AD)

Pazienti con un quadro clinico di spasticità pura hanno molte probabilità che il loro quadro clinico sia dovuto a mutazioni nella Spastina (SPG4), che da sola rappresentano più del 40% delle HSP-AD.

Evidenze sperimentali suggeriscono che la Spastina è una proteina con una funzione complessa utile nel rimodellamento dinamico dell'assone adulto ed in crescita.

Nei pazienti appartenenti al gruppo SP (n= 42), sono state identificate varianti in 9 pazienti (21,5%), maggiormente identificate in pazienti con storia familiare positiva rispetto agli sporadici. Di questi 9, 6 mostravano una segregazione autosomica dominante, mentre 3 autosomica recessiva. Delle dominanti, tutte le 6 varianti sono state identificate in SPG4 (c.1378 C>T; c.1679C>T; c.1537-8T>G in due pazienti; c.1245+1G>A; c.1414-2A>T). Di queste, 5 sono nuove e 1 è già descritta, sono stati individuati nelle regioni codificanti (n=2) o nei siti di splicing(n=4).

La frequenza complessiva di mutazioni di SPG4 nel nostro studio era del 38% (6/16) raggiungendo il 67% (6/9) quando erano considerati i pazienti solo con HSP-AD. Questi dati ricadono nel range osservato in differenti popolazioni.

In questi pazienti, la diagnosi molecolare è utile alla conferma della diagnosi clinica.

b. Analisi genetiche che consentono un approfondimento di una diagnosi clinica

□ Diagnosi di Atassia Spinocerebellare Autosomica dominante di tipo 5 (SCA5)

L'analisi di sequenziamento ha permesso di eseguire diagnosi di SCA5 (Phenotype MIM number 600224), in due pazienti, le cui informazioni cliniche avevano identificato una atassia spinocerebellare con esordio tardivo (56aa, Paziente 11, 52aa, Paziente 14). La SCA5 è causata da mutazioni nel gene *SPTBN2* (Phenotype MIM number 604985) sul cromosoma 11q13. *SPTBN2* codifica per la beta-3 spettrina, coinvolta nella formazione di tetrameri di spettrina che stabilizzano giunzioni cellulari, canali e molecole di adesione lungo la parte citoplasmatica della membrana (Clarkson et al., 2010). Tale proteina è fondamentale per il corretto funzionamento e lo sviluppo delle cellule di Purkinje.

La atassia spinocerebellare tipo 5 (SCA5) è una forma rara di atassia cerebellare autosomica dominante tipo 3 (ADCA tipo 3), in cui il processo degenerativo può essere limitato al cervelletto. È caratterizzata da segni cerebellari ad esordio precoce con anomalie dei movimenti oculari (nistagmo evocato dallo sguardo, nistagmo verticale verso il basso e la compromissione dei movimenti lenti di inseguimento), ed una età mediana all'esordio è 33 anni (tra 10 e i 68 anni). I pazienti affetti da SCA5 presentano segni cerebellari (atassia, disartria, tremore intenzionale) e anomalie dei movimenti oculari. Sono rari i difetti del campo visivo e la paralisi dello sguardo orizzontale.

Sia il paziente 11 che 14 mostrano delle caratteristiche cliniche compatibili con la SCA5 (risonanza magnetica con marcata atrofia cerebellare) e progressione molto lenta. Non sono presenti anomalie dei movimenti oculari o altri disturbi oculomotori.

L'inquadramento diagnostico in una atassia spinocerebellare di tipo 5 permette di prevedere un andamento della malattia lento, con una durata superiore ai 30 anni, e con una aspettativa di vita molto simile a quello della popolazione generale

□ Diagnosi di Atassia Spinocerebellare Autosomica recessiva di tipo 8 (SCAR8)

L'analisi di sequenziamento ha permesso di eseguire diagnosi di SCAR8 (Phenotype MIM number 610743), in un paziente le cui informazioni cliniche avevano identificato una atassia spinocerebellare con esordio a 36 aa (Paziente 11). LA SCAR8, ereditata con modalità autosomica recessiva, è causata da mutazioni in *SYNE1*. In forma dominante può causare la distrofia di Emery-Dreifuss (Phenotype MIM number 612998). *SYNE1* codifica per una proteina della famiglia delle specttrine che collega la membrana plasmatica al citoscheletro di actina. Nel paziente 13, è stata identificata una mutazione in omozigosi ereditata da genitori consanguinei, asintomatici. Questo a conferma della modalità di segregazione autosomica recessiva. La SCA8 è caratterizzata da atassia cerebellare pura, a evoluzione lenta, associata a disartria. Questo quadro clinico è compatibile con quello osservato nel paziente 13, che evidenzia disturbi della deambulazione tipo atassico ad andamento lentamente progressivo, disartria e disturbi della mobilità oculare. La risonanza magnetica evidenziava atrofia cerebellare. Anche in questo caso, la diagnosi di SCAR8 in un paziente relativamente giovane (36aa) permette di prevedere un lento decorso della malattia.

c. Analisi genetiche che consentono di eseguire una diagnosi clinica

Come detto, in alcuni casi le le analisi genetiche servono ad eseguire la diagnosi clinica in casi in cui le informazioni cliniche e strumentali sono contrastanti o non significative.

È stata studiata una paziente con paraparesi spastica esordita a circa 35 anni, in cui la risonanza magnetica non ha offerto utili informazioni diagnostiche. L'analisi di sequenziamento mediante esoma clinico ha evidenziato una variante missenso in eterozigosi nel gene *ABCD1* (MIM number 300371). Questo gene è responsabile di Adrenoleucodistrofia legata all'X (X-ALD), un disturbo della beta ossidazione degli acidi grassi perossisomali che provoca l'accumulo dei VLCFA (acidi grassi a lunghissima catena) nei tessuti di tutto il corpo che si presenta più comunemente nei maschi.

Per valutare il possibile ruolo clinico del gene *ABCD1* in questa paziente, sono state studiate due sorelle, entrambe affette da paraparesi spastica con esordio tra i 35 ed i 40 anni. L'analisi di segregazione ha confermato la presenza della variante anche nelle sorelle.

A sostegno di questo, sebbene la malattia sia legata al cromosoma X, e quindi si presenta più comunemente nei maschi, anche il 50% delle femmine eterozigoti mostra alcuni sintomi nel corso della vita. In questo caso, le informazioni diagnostiche della analisi molecolare sono state confermate e sostenute dai valori elevati di VLCFA, e da risonanze magnetiche che nel corso del tempo hanno iniziato a mostrare segni di adrenoleucodistrofia.

In questo modo, la diagnosi molecolare può rivelarsi informativa in quadri clinici in cui alcuni sintomi compaiono in modo variabile e più avanti nel corso della vita, rendendo la diagnosi clinica difficoltosa se non utilizza il sostegno della analisi genetica. Oltretutto, questo conferma che è fondamentale disporre di altri soggetti affetti e sani all'interno della stessa famiglia. In questo studio, un corretto utilizzo congiunto di analisi di segregazione familiare ed analisi bioinformatica, è fondamentale per l'identificazione delle varianti patologiche.

## 6 CONCLUSIONI

I dati di questo studio dimostrano che pannelli NGS, accuratamente disegnati e validati, possano essere applicati alla diagnosi molecolare di patologie neurologiche caratterizzate da poligenicità e variabilità clinica. La validazione di questi percorsi richiede la selezione di una casistica accurata ed omogenea. Questo consente di delinearne il corretto utilizzo diagnostico, aumentando così la detection rate.

Ad esempio, questo studio dimostra che l'analisi genetica deve essere tenuta in considerazione anche nei pazienti affetti da Atassie Spinocerebellare in cui non vi è una spiccata segregazione familiare della patologia.

Risulta inoltre fondamentale l'utilizzo di un iter di interpretazione delle varianti molto stringente e riproducibile. Questo viene confermato nel presente studio, dove un corretto utilizzo congiunto di analisi di segregazione familiare (ove possibile reperire i familiari del paziente affetto) ed analisi bioinformatica, è fondamentale per l'identificazione delle varianti patologiche, soprattutto in geni per i quali sono disponibili dati limitati nei database pubblici.

Risulta inoltre utile applicare pannelli più estesi, come l'esoma clinico, in casi complessi di difficile inquadramento. In questi casi, la diagnosi molecolare si è rilevata, insieme a quella clinica, fondamentale per un corretto inquadramento diagnostico. Può inoltre rivelarsi informativa anche in quadri clinici con sintomi variabili e con esordi differenti, rendendo la diagnosi clinica impossibile senza beneficiare di quella genetica.

L'utilizzo di questi pannelli, sia pannelli multigenici che esoma clinico, può avere diverse utilità diagnostiche. Infatti mutazioni in alcuni geni servono unicamente ad avvalorare la diagnosi clinica, come nel caso delle Paraparesi spastiche Autosomico-dominanti. Al contrario, in altre situazioni le varianti identificate sono fondamentali per un corretto inquadramento diagnostico. È l'esempio della

Atassia spinocerebellare di tipo 5 permette di prevedere un andamento della malattia lento, con una durata superiore ai 30 anni, e con una aspettativa di vita molto simile a quello della popolazione generale

In altri casi, le varianti identificate servono ad eseguire la diagnosi clinica, soprattutto nei casi in cui le informazioni cliniche e strumentali sono contrastanti o non significative come nel caso della Adrenoleucodistrofia legata all'X (X-ALD).

## BIBLIOGRAFIA

- Adam E. Handel, Giulio Disanto, Sreeram V. Ramagopalan. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. *Expert Rev Neurother.* 2013 Feb; 13(2): 10.1586/ern.12.165. doi: 10.1586/ern.12.165
- Aicardi J Barbosa C Andermann E et al., (1988 )Ataxia oculo-motor 24; 497-502.
  - Alvarez V, Sánchez-Ferrero E, Beetz C, Díaz M, Alonso B, Corao AI, Gámez J, Esteban J, Gonzalo JF, Pascual-Pascual SI, López de Munain A, Moris G, Ribacoba R, Márquez C, Rosell J, Marín R, García-Barcina MJ, Del Castillo E, Benito C, Coto E; Group for the Study of the Genetics of Spastic Paraplegia. Mutational spectrum of the SPG4 (SPAST) and SPG3A (ATL1) genes in Spanish patients with hereditary spastic paraplegia. *BMC Neurol.* 2010 Oct 8;10:89. doi: 10.1186/1471-2377-10-89.
  - Arnoldi A, Tonelli A, Crippa F, Villani G, Pacelli C, Sironi M, Pozzoli U, D'Angelo MG, Meola G, Martinuzzi A, Crimella C, Redaelli F, Panzeri C, Renieri A, Comi GP, Turconi AC, Bresolin N, Bassi MT. A clinical, genetic, and biochemical characterization of SPG7 mutations in a large cohort of patients with hereditary spastic paraplegia. *Hum Mutat.* 2008 Apr;29(4):522-31. doi: 10.1002/humu.20682.
  - Ashley-Koch A., Bonner E.R., Gaskell P.C., West S.G., Tim R., Wolpert C.M., Jones R., Farrell C.D., Nance M., Svenson I.K., Marchuk D.A., Boustany R.-M. N., Vance J.M., Scott W.K., Pericak-Vance M.A. Fine mapping and genetic heterogeneity in the pure form of autosomal dominant familial spastic paraplegia. *Neurogenetics* (2001); 3:91-97
  - Bernard G, Chouery E, Putorti ML, Tétreault M, Takanohashi A, Carosso G, Clément I, Boespflug-Tanguy O, Rodriguez D, Delague V, Abou Ghoch J, Jalkh N, Dorboz I, Fribourg S, Teichmann M, Megarbane A, Schiffmann R, Vanderver A, Brais B. Mutations of POLR3A encoding a catalytic subunit of RNA

polymerase Pol III cause a recessive hypomyelinating leukodystrophy. *Am J Hum Genet.* 2011 Sep 9;89(3):415-23. doi: 10.1016/j.ajhg.2011.07.014.

- Bick D and Dimmock D. Whole exome and whole genome sequencing. *Curr Opin Pediatr.* 2011 Dec;23(6):594-600. doi: 10.1097/MOP.0b013e32834b20ec.
- Bird Thomas D. Hereditary Ataxia Overview. *GeneReviews.* 2018
- Bickerstaff ER. Hereditary spastic paraplegia. *J Neurol Neurosurg Psychiatry* (1950); 13:134-145.
- Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat Genet.* 2003 Mar;33 Suppl:228-37.
- Brent L. Fogel, MD, PhD, Hane Lee, PhD, Joshua L. Deignan, PhD, Samuel P. Strom, PhD, Sibel Kantarci, PhD, Xizhe Wang, BS, Fabiola Quintero-Rivera, MD, Eric Vilain, MD, PhD, Wayne W. Grody, MD, PhD, Susan Perlman, MD, Daniel H. Geschwind, MD, PhD, and Stanley F. Nelson, MD. Exome Sequencing in the Clinical Diagnosis of Sporadic or Familial Cerebellar Ataxia. *JAMA Neurol.* 2014 October ; 71(10): 1237–1246. doi:10.1001/jamaneurol.2014.1944.
- Bruyn RPM. The neuropathology of hereditary spastic paraparesis. *Clinical Neurology and Neurosurgery.* (1992); 94(Suppl.):S16-18.
- Campbell EM, Li H, Mori T, Osterweil P, Guise JM. The Impact of Health Information Technology on Work Process and Patient Care in Labor and Delivery. In: Henriksen K, Battles JB, Keyes MA, Grady ML, editors. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 4: Technology and Medication Safety).* Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 Aug.
- Chan S.S., Longley M.J. and Copeland W.C., (2005). The common A467T mutation in the human mitochondrial DNA

Polymerase (POLG) compromises catalytic efficiency and interaction with the accessory subunit. *J. Biol. Chem.*, 280: 31341-31346.

- Claes S, Devriendt K, Van Goethem G, Roelen L, Meireleire J, Raeymaekers P, Cassiman JJ, Fryns JP. Novel syndromic form of X-linked complicated spastic paraplegia. *Am J Med Genet.* (2000); 4;94(1):1-4.
- Coutinho P, Barros J, Zemmouri R, Guimaraes J, Alves C, Choro R, Lourenco E, Ribeiro P, Loureiro JL, Santos JV, Hamri A, Paternotte C, Hazan J, Silva MC, Prud'homme JF, Grid D. Clinical heterogeneity of autosomal recessive spastic paraplegias: analysis of 106 patients in 46 families. *Arch Neurol* (1999);56:943-9
- Coleman TM, Huang F. Optimal random libraries for the isolation of catalytic RNA. *RNA Biol.* 2005 Oct-Dec;2(4):129-36.
- Christensen KD, Dukhovny D, Siebert U, Green RC. Assessing the Costs and Cost-Effectiveness of Genomic Sequencing. *J Pers Med.* 2015 Dec 10;5(4):470-86. doi: 10.3390/jpm5040470.
- Damiani E., Borsani G., Giacomuzzi E. Amplicon-based semiconductor sequencing of human exomes: performance evaluation and optimization strategies. *Hum Genet.* 2016; 135: 499–511. Published online 2016 Mar 22. doi: 10.1007/s00439-016-1656-8.
- Denton K, Mou Y, Xu CC, Shah D, Chang J, Blackstone C, Li XJ. Impaired mitochondrial dynamics underlie axonal defects in hereditary spastic paraplegias. *Hum Mol Genet.* 2018 Jul 15;27(14):2517-2530. doi: 10.1093/hmg/ddy156.
- Di Resta Chiara, Galbiati Silvia, Carrera Paola, Ferrari Maurizio. Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities. *EJIFCC.* 2018 Apr; 29(1): 4–14.
- Durr A. Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. *Lancet Neurol.* 2010 Sep;9(9):885-94. doi: 10.1016/S1474-4422(10)70183-6.

- Dürr A., Camuzat A., Colin E., Tallaksen C., Hannequin D., Coutinho P., et al. Atlastin mutations are frequent in young-onset autosomal dominant spastic paraplegia. *Arch Neurol* (2004); 61:1867-1872.
- Ferrer I, Olivé M. Molecular pathology of myofibrillar myopathies. *Expert Rev Mol Med*. 2008 Sep 3;10: e25. doi: 10.1017/S1462399408000793
- Falco M, Scuderi C, Musumeci S, et al. Two novel mutations in the spastin gene (SPG4) found by DHPLC mutation analysis. *Neuromuscul Disord*. 2004;14(11):750–753.
- Federico A., Caltagirone C., Provinciali L., Tedeschi G. *Neurologia pratica*. EdiSES 2014.
- Gasser T, et al. EFNS guidelines of the molecular diagnosis of ataxia and spastic paraplegias. *Eur J Neurol*. 2010
- Greco G.S., Malandrini A., Comanducci G., et al.,(2004). Novel SACS mutations in autosomal recessive spastic ataxia of Charlevoix-Saguenay type. *Neurology* 13; 62; 103-106.
- Guergueltcheva V, Azmanov DN, Angelicheva D, Smith KR, Chamova T, Florez L, Bynevelt M, Nguyen T, Cherninkova S, Bojinova V, Kaprelyan A, Angelova L, Morar B, Chandler D, Kaneva R, Bahlo M, Tournev I, Kalaydjieva L. Autosomal-recessive congenital cerebellar ataxia is caused by mutations in metabotropic glutamate receptor 1. *Am J Hum Genet*. 2012 Sep 7;91(3):553-64. doi: 10.1016/j.ajhg.2012.07.019. Epub 2012 Aug 16.
- Hedera Peter, MD, PhD, FACMG. Hereditary Spastic Paraplegia Overview. *GeneReviews*. 2018
- Harding AE. Hereditary pure spastic paraplegia: a clinical and genetic study of 22 families. *J Neurol Neurosurg Psychiatry* (1981); 44:871-873
- Harding A.E. Classification of the hereditary ataxias and paraplegias. *Lancet* (1983) 1 (8334):1151-1155
- Hentati A, Deng HX, Zhai H, Chen W, Yang Y, Hung WY, Azim AC, Bohlega S, Tandan R, Warner C, Laing NG, Cambi F, Mitumoto H, Roos RP, Boustany RM, Ben Hamida M, Hentati F,

Siddique T. Novel mutations in spastin gene and absence of correlation with age at onset of symptoms. *Neurology*. 2000 Nov 14;55(9):1388-90.

- Hirst J, Madeo M, Smets K, Edgar JR, Schols L, Li J, Yarrow A, Deconinck T, Baets J, Van Aken E, De Bleecker J, Datiles MB 3rd, Roda RH, Liepert J, Züchner S, Mariotti C, De Jonghe P, Blackstone C, Kruer MC. Complicated spastic paraplegia in patients with AP5Z1 mutations (SPG48). *Neurol Genet*. 2016 Aug 25;2(5):e98. doi: 10.1212/NXG.000000000000098. eCollection 2016 Oct. Erratum in: *Neurol Genet*. 2016 Oct 06;2(5):e111.
- Indelicato E, Nachbauer W, Fauth C, Krabichler B, Schossig A, Eigentler A, Dichtl W, Wenning G, Wagner M, Fanciulli A, Janecke A, Boesch S. SYNE1-ataxia: Novel genotypic and phenotypic findings. *Parkinsonism Relat Disord*. 2018 Dec 11. pii: S1353-8020(18)30535-2. doi: 10.1016/j.parkreldis.2018.12.007.
- Jason A. Reuter, Damek Spacek, and Michael P. Snyder. High-Throughput Sequencing Technologies. *Mol Cell*. 2015 May 21; 58(4): 586–597. doi: 10.1016/j.molcel.2015.05.004
- Joanna M. Solowska and Peter W. Hereditary spastic paraplegia SPG4: what is known and not known about the disease. *Brain*. 2015 Sep; 138(9): 2471–2484.
- Kaplan JC. The 2012 version of the gene table of monogenic neuromuscular disorders. *Neuromuscul Disord*. 2011 Dec;21(12):833-61.
- Lin X, Tang W, Ahmad S, Lu J, Colby CC, Zhu J, Yu Q. Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. *Hear Res*. 2012 Jun;288(1-2):67-76. doi: 10.1016/j.heares.2012.01.004.
- Maarouf H, Taboada M, Rodriguez H, Arias M, Sesar Á, Sobrido MJ. An ontology-aware integration of clinical models, terminologies and guidelines: an exploratory study of the Scale for the Assessment and Rating of Ataxia (SARA). *BMC Med Inform*

Decis Mak. 2017 Dec 6;17(1):159. doi: 10.1186/s12911-017-0568-4.

- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. (Margulies, 2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005 Sep 15;437(7057):376-80. Epub 2005 Jul 31.
- McCorquodale DS 3, Ozomaro U, Huang J, Montenegro G, Kushman A, Citrigno L, Price J, Speziani F, Pericak-Vance MA, Züchner S. Mutation screening of spastin, atlastin, and REEP1 in hereditary spastic paraplegia. *Clin Genet*. 2011 Jun;79(6):523-30. doi: 10.1111/j.1399-0004.2010.01501.x.
- Matilla-Dueñas A, et al. The spinocerebellar ataxias: clinical aspects and molecular genetics. *Adv Exp Med Biol*. 2012.
- Moreira MC, Barbot C, Tachi N, Kozuka N, Uchida E, Gibson T, Mendonça P, Costa M, Barros J, Yanagisawa T, Watanabe M, Ikeda Y, Aoki M, Nagata T, Coutinho P, Sequeiros J, Koenig M. The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin. *Nat Genet*. 2001 Oct;29(2):189-93.
- Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*. 2010 Jan;42(1):30-5. doi: 10.1038/ng.499.
- Nick Dand, Reiner Schulz, Michael E. Weale, Laura Southgate, Rebecca J. Oakey, Michael A. Simpson, Thomas Schlitt. Network-Informed Gene Ranking Tackles Genetic

Heterogeneity in Exome-Sequencing Studies of Monogenic Disease. *Hum Mutat.* 2015 Dec; 36(12): 1135–1144. Published online 2015 Oct 7. doi: 10.1002/humu.22906

- Parolin Schneckenberg R, Perkins EM, Miller JW, Davies WI, D'Adamo MC, Pessia M, Fawcett KA, Sims D, Gillard E, Hudspith K, Skehel P, Williams J, O'Regan M, Jayawant S, Jefferson R, Hughes S, Lustenberger A, Ragoussis J, Jackson M, Tucker SJ, Németh AH. De novo point mutations in patients diagnosed with ataxic cerebral palsy. *Brain.* 2015 Jul;138(Pt 7):1817-32. doi: 10.1093/brain/awv117.
- Perkins E, Suminaite D, Jackson M. Cerebellar ataxias:  $\beta$ -III spectrin's interactions suggest common pathogenic pathways. *J Physiol.* 2016 Aug 15;594(16):4661-76. doi: 10.1113/JP271195. Epub 2016 Apr 24. Review.
- Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J. An integrated semiconductor device enabling non-optical genome sequencing. *Nature.* 2011 Jul 20;475(7356):348-52. doi: 10.1038/nature10242.
- Salinas S., Proukakis C., et al. Hereditary spastic paraplegia: clinical features and pathogenetic mechanism. *Salinas. Lancet Neurol.*, 2008.
- Schuurs-Hoeijmakers JH1, Vulto-van Silfhout AT, Vissers LE, van de Vondervoort II, van Bon BW, de Ligt J, Gilissen C, Hehir-Kwa JY, Neveling K, del Rosario M, Hira G, Reitano S, Vitello A, Failla P, Greco D, Fichera M, Galesi O, Kleefstra T, Grealley MT, Ockeloen CW, Willemsen MH, Bongers EM, Janssen IM, Pfundt R, Veltman JA, Romano C, Willemsen MA, van Bokhoven H, Brunner HG, de Vries BB, de Brouwer AP. Identification of pathogenic gene

variants in small families with intellectually disabled siblings by exome sequencing. *J Med Genet*. 2013 Dec;50(12):802-11.

- Scholz C, Golas MM, Weber RG, Hartmann C, Lehmann U, Sahm F, Schmidt G, Auber B, Sturm M, Schlegelberger B, Illig T, Steinemann D, Hofmann W. Rare compound heterozygous variants in PNKP identified by whole exome sequencing in a German patient with ataxia-oculomotor apraxia 4 and pilocytic astrocytoma. *Clin Genet*. 2018 Jul;94(1):185-186. doi: 10.1111/cge.13216
- Stevanin G, Santorelli FM, Azzedine H, Coutinho P, Chomilier J, Denora PS, Martin E, Ouvrard-Hernandez AM, Tessa A, Bouislam N, Lossos A, Charles P, Loureiro JL, Elleuch N, Confavreux C, Cruz VT, Ruberg M, Leguern E, Grid D, Tazir M, Fontaine B, Filla A, Bertini E, Durr A, Brice A. Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. *Nat Genet*. 2007 Mar;39(3):366-72.
- Swan L, Cardinal J, Coman D. SYNE1-related autosomal recessive cerebellar ataxia, congenital cerebellar hypoplasia, and cognitive impairment. *Clin Pract*. 2018 Aug 27;8(3):1071. doi: 10.4081/cp.2018.1071. eCollection 2018 Jul 10.
- Travaglini L, Aiello C, Stregapede F, D'Amico A, Alesi V, Ciolfi A, Bruselles A, Catteruccia M, Pizzi S, Zanni G, Loddo S, Barresi S, Vasco G, Tartaglia M, Bertini E, Nicita F. The impact of next-generation sequencing on the diagnosis of pediatric-onset hereditary spastic paraplegias: new genotype-phenotype correlations for rare HSP-related genes. *Neurogenetics*. 2018 May;19(2):111-121. doi: 10.1007/s10048-018-0545-9. Epub 2018 Apr 24.
- Tonduti D, Dorboz I, Renaldo F, Masliah-Planchon J, Elmaleh-Bergès M, Dalens H, Rodriguez D, Boespflug-Tanguy O. Cystic leukoencephalopathy with cortical dysplasia related to LAMB1 mutations. *Neurology*. 2015 May 26;84(21):2195-7. doi: 10.1212/WNL.0000000000001607.

- Umar Akbar, M.D. and Tetsuo Ashizawa, M.D. Ataxia. *Neurol Clin.* 2015 February ;33(1): 225–248. doi:10.1016/j.ncl.2014.09.004.
- Watson LM, Bamber E, Schnekenberg RP, Williams J, Bettencourt C, Lickiss J, Jayawant S, Fawcett K, Clokie S, Wallis Y, Clouston P, Sims D, Houlden H, Becker EBE, Németh AH. Dominant Mutations in GRM1 Cause Spinocerebellar Ataxia Type 44. *Am J Hum Genet.* 2017 Nov 2;101(5):866. doi: 10.1016/j.ajhg.2017.10.008. No abstract available.
- Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock Li H, Durbin R. The complete genome of an individual by massively parallel DNA sequencing. *Nature.* 2008 Apr 17;452(7189):872-6. doi: 10.1038/nature06884.
- Winner B, Uyanik G, Gross C, Lange M, Schulte-Mattler W, Schuierer G, Marienhagen J, Hehr U, Winkler J. Clinical progression and genetic analysis in hereditary spastic paraplegia with thin corpus callosum in spastic gait gene 11 (SPG11). *Arch Neurol* (2004); 61:117-21.
- Van Deerlin VM. The genetics and neuropathology of neurodegenerative disorders: perspectives and implications for research and clinical practice. *Acta Neuropathol*, 2012
- Zhi D, Chen R. Statistical guidance for experimental design and data analysis of mutation detection in rare monogenic mendelian diseases by exome sequencing. *PLoS One.* 2012;7(2):e31358. doi: 10.1371/journal.pone.0031358.
- Whiteford N, Haslam N, Weber G, Prügél-Bennett A, Essex JW, Roach PL, Bradley M, Neylon C. An analysis of the feasibility of short read sequencing. *Nucleic Acids Res.* 2005 Nov 7;33(19):e171.
- Karine Choquet, Martine Tétreault, Sharon Yang, Roberta La Piana,<sup>1</sup> Marie- Joséé Dicaire, Megan R Vanstone, Jean

Mathieu, Jean-Pierre Bouchard, Marie-France Rioux, Guy A Rouleau, Care4Rare Canada Consortium,4 Kym M Boycott,4 Jacek Majewski, and Bernard Brais.SP G7 mutations explain a significant proportion of French Canadian spastic ataxia cases. Eur J Hum Genet. 2016 Jul; 24(7): 1016–1021.

- <https://www.ncbi.nlm.nih.gov/pubmed/>
- <http://www.omim.org/>
- <http://www.hgmd.org/>
- <http://ExAC.broadinstitute.org/>
- <https://www.ncbi.nlm.nih.gov/projects/SNP/>
- <https://www.ncbi.nlm.nih.gov/clinvar/>
- <http://evs.gs.washington.edu/EVS/>
- <http://www.internationalgenome.org/>
- <http://www.mutationtaster.org/>
- <http://sift.bii.a-star.edu.sg/>
- <http://provean.jcvi.org/index.php>
- <http://genetics.bwh.harvard.edu/pph2/>

## **RINGRAZIAMENTI**

Ringrazio tutti coloro che mi sono stati vicini durante questo percorso e che con il loro sostegno mi hanno permesso di raggiungere tale obiettivo.

Desidero ringraziare il Prof. Giuseppe Lembo per l'opportunità concessami.

Desidero inoltre ringraziare il Dott. Stefano Gambardella che mi ha dato l'opportunità di far parte del suo laboratorio e che in questi anni mi ha supportata sopportata e soprattutto fatto crescere professionalmente.

Desidero infine ringraziare tutte le mie colleghe del Centro di Genetica Molecolare.