

Dottorato di ricerca in: Scienze Pediatriche
Curriculum: Malattie Muscolari,
neurodegenerative e metaboliche dell'età
evolutiva (ciclo XXXI).
Coordinatore Prof. C. Minetti

**IDENTIFICAZIONE DI NUOVI GENI RESPONSABILI DI MALATTIE
RARE DEL NEUROSVILUPPO TRAMITE HOMOZYGOSITY MAPPING
E/O SEQUENZIAMENTO DI NUOVA GENERAZIONE**

Relatore

Prof. Federico Zara



Dottorando S3996214

Dott. Michele Iacomino



SOMMARIO

INTRODUZIONE	3
1 Sequenziamento di nuova generazione	3
1.1 Avvento delle tecnologie di nuova generazione	3
1.2 Applicazione delle tecnologie NGS	4
1.3 Whole Exome Sequencing (WES)	5
2. RAZIONALE	5
3 MATERIALI E METODI	6
3.1 Estrazione del DNA	6
3.2 Genotipizzazione e Homozygosity Mapping	7
3.3 Sequenziamento NGS	7
3.4 Analisi bioinformatica	9
3.4.1 Controllo qualità e Allineamento	10
3.4.2 Variants Calling	10
3.4.3 Annotazione e Filtraggio	10
3.4.4 Pipeline	10
3.5 Validazione Sanger	12
3.6 Coltura di fibroblasti	12
3.7 Estrazione delle proteine	13
3.8 Western blot	13
3.9 Estrazione RNA	14
3.10 Retrotrascrizione	14
3.11 qPCR e primers ER stress	14
3.12 Immunistoichimica (IHC)	14
3.13 Immunocitoichimica (ICC)	15
4. RISULTATI	15
4.1 Famiglia 1	16
4.2 Famiglia 2	18
4.3 Famiglia 3	22
4.4 Famiglia 4	29
5. CONCLUSIONI	31
6. BIBLIOGRAFIA	32

INTRODUZIONE

1. SEQUENZIAMENTO DI NUOVA GENERAZIONE

1.1 Avvento delle tecnologie di nuova generazione

Negli ultimi anni si è osservato un notevole incremento delle conoscenze riguardanti il genoma umano e il codice genetico, dalla scoperta della struttura del DNA sino al primo progetto sul genoma umano.

Nel 1990 coordinato dal National Institutes of Health degli Stati Uniti e dal U.S. Department of Energy (DOE) era stato avviato lo Human Genome Project, con l'intento di sequenziare il genoma umano ed identificare e mappare i geni in esso contenuti. Il progetto, basato dal punto di vista tecnico sul sequenziamento di tipo Sanger, ha visto il suo compimento in circa tredici anni di lavoro^{1,2}, la maggior parte dei quali sono stati necessari per la parte sperimentale del progetto stesso.

Il progetto Genoma umano ha consentito una evoluzione nell'innovazione scientifica nel campo biologico ed informatico, determinando il progresso delle piattaforme per il sequenziamento massivo di nuova generazione (NGS), tecnologia che negli anni ha consentito di abbattere i costi e i tempi di sequenziamento ed individuare geni sconosciuti che sono stati associati a malattie.

Questo tipo di tecnologia ha consentito il sequenziamento dell'esoma che corrisponde solo alle regioni codificanti dei geni comprese le regioni fiancheggianti che fino a questo momento era impensabile identificare geni nuovi malattia senza conoscere i geni target che erano sequenziati per Sanger^{3,4}.

Per anni la tecnica proposta da Sanger⁵ è stata il gold standard per il sequenziamento del DNA. Dall'avvento del NGS c'è stata una vera e propria rivoluzione nel campo del sequenziamento degli acidi nucleici. In tempi brevissimi, diverse piattaforme tecnologiche basate su questo nuovo sistema sono state sviluppate e perfezionate, dando origine ad una notevole quantità di dati prodotti e la velocità con cui essi vengono generati. Inoltre, negli ultimi anni si è assistito ad una considerevole diminuzione dei costi di processamento dei campioni, fattore essenziale che ha favorito l'ampia diffusione di questi sistemi.

Nonostante ciò, ad oggi il sequenziamento di Sanger, per la sua affidabilità, mantiene ancora la sua preminenza nella pratica clinica.

Le tecnologie NGS presentano diversi vantaggi rispetto al sequenziamento di tipo Sanger:

- 1) Produce molte gigabasi di sequenze a costi e tempi inferiori
- 2) Determina in maniera accurata la presenza di alleli a bassa frequenza

- 3) Capacità di processare diversi campioni in un unico esperimento attraverso marcatura con barcoding delle sequenze⁶

Nonostante tutti gli aspetti positivi, le tecnologie NGS presentano delle caratteristiche limitanti che ne condizionano l'applicazione:

- 1) Lunghezza dei frammenti (reads) che possono essere sequenziati, generalmente intorno alle 200- 300 bp, contro le 1000 bp di un sequenziatore capillare.
- 2) Aumentata percentuale di errore (falsi positivi e negativi), ma tale limite è superato attraverso l'aumento del coverage (quantità sequenze generate)

E' questo uno dei principali motivi per cui, nella pratica clinico-diagnostica, il sequenziamento Sanger mantiene la sua preminenza. Strettamente legata a questo aspetto è la necessità di continuo sviluppo di algoritmi bioinformatici, che consentano di analizzare con efficacia i dati prodotti. L'aspetto informatico ha profonda rilevanza nelle tecnologie di Next Generation.

Questa complessità di livelli di analisi informatica, complica notevolmente il compito, in particolar modo nei casi in cui sia necessario confrontare risultati ottenuti attraverso differenti piattaforme, poiché ciascuna tecnologia sfrutta programmi e parametri di analisi differenti, ma che alle volte permettono di identificare varianti che non risulterebbero con singoli algoritmi. Inoltre, l'interpretazione dei dati e la gestione di un sistema informatico così complesso prevede competenze specialistiche e richiede la presenza di figure professionali specifiche.

1.2 Applicazione delle tecnologie NGS

Durante gli ultimi 20 anni i determinanti genetici delle malattie mendeliane monogeniche sono stati identificati principalmente attraverso studi genetici tradizionali (<http://www.ncbi.nlm.nih.gov/omim>). Il cosiddetto "clonaggio posizionale" è stato essenziale per l'identificazione della maggioranza dei circa 3000 geni mutati in patologie mendeliane.

Tuttavia ancora per molte di queste malattie le limitazioni tecnologiche, come per il clonaggio posizionale, hanno impedito di giungere all'identificazione del difetto genetico causativo.

Recentemente lo sviluppo delle tecnologie NGS ha dimostrato di poter superare queste limitazioni e di poter catturare in un singolo esperimento tutte le varianti genetiche presenti in un genoma.

Diversi sono gli ambiti in cui viene a essere utilizzato il Next Generation Sequencing, uno di questi è il Whole Exome Sequencing (WES).

1.3 Whole Exome Sequencing (WES)

Il Whole Exome Sequencing è il sequenziamento di tutte le regioni codificanti dei geni di un individuo (esoma). Esistono per questa tecnologia differenti kit di sequenziamento, grazie ai quali vengono analizzati tutti i geni del genoma ed altri ai soli finora notoriamente associati a malattia (circa 6000) nella pratica dell'esoma clinico.

Quest'approccio è il più diretto per comprendere le basi molecolari coinvolte nella patogenesi di malattie ereditarie e presenta un riscontro anche a livello clinico. L'identificazione di un gene malattia, infatti, permette di poter svolgere una diagnosi molecolare diretta con una ricaduta immediata, sia per quanto riguarda gli aspetti di prevenzione della malattia, sia per un'accurata consulenza genetica e sia per un corretto calcolo del rischio di ricorrenza.

L'analisi molecolare di un gene può consentire di determinare il fenotipo atteso sulla base della correlazione genotipo/fenotipo e/o la prognosi della malattia. La comprensione dei meccanismi patogenetici responsabili della patologia, inoltre, può talvolta indirizzare verso nuovi approcci terapeutici per la cura di difetti biochimici alla base delle malattie genetiche.

2. RAZIONALE

La genetica di numerose malattie mendeliane del neurosviluppo è stata studiata ampiamente negli ultimi venti anni. Tuttavia, l'eziologia di molte malattie rare o ultrarare a trasmissione mendeliana è ancora sconosciuta, con significative implicazioni per la diagnosi, il counselling genetico e la terapia.

Finora, l'identificazione di geni associati a malattie rare del neurosviluppo è stata significativamente limitata da difficoltà metodologiche come la limitata informatività di famiglie per studi di linkage, di marcatori polimorfici ed il costo elevato per la genotipizzazione e il sequenziamento.

Tuttavia, gli enormi progressi nella ricerca genomica hanno oggi reso disponibile per la comunità scientifica la sequenza consenso del genoma umano e una mappa ad alta densità di varianti polimorfiche osservate in differenti popolazioni, e hanno reso disponibili potenti tecnologie per esplorare il DNA a livello di singole basi, rapidamente ed a basso costo. Negli ultimi anni, il completo sequenziamento del DNA codificante (cosiddetto "esoma") è stato applicato con successo per identificare mutazioni patogenetiche mediante il confronto delle varianti esoniche tra pazienti non imparentati ed affetti dalla stessa malattia.

Il sequenziamento dell'esoma può essere applicato a pazienti affetti da patologie rare non classificate e geneticamente eterogenee dovute sia a meccanismi genetici recessivi che dominanti. Al fine di chiarire l'eterogeneità genetica di alcune rare patologie neurodegenerative, abbiamo deciso di combinare l'approccio di "exome sequencing" con "homozygosity mapping", metodica che permette di localizzare loci associati a malattie mendeliane a trasmissione autosomica recessiva. Il progetto si basa sul sequenziamento di 7 famiglie selezionate riferite dalle unità operative di Neurologia Pediatrica e Malattie Muscolari e di Neurochirurgia dell'Istituto Giannina Gaslini.

Per questo tipo di studio abbiamo focalizzato l'attenzione su famiglie nucleari, con un parente affetto da malattia rara non diagnosticata del neurosviluppo, in relazione al bacino di utenza nel nostro istituto.

E' stato necessario il reclutamento di popolazioni di controllo di adeguate dimensioni in modo da poter, infine, differenziare varianti rare codificanti e fenotipicamente neutre da mutazioni patogenetiche.

Sebbene, ci si aspetti che tale approccio permetta di individuare in maniera inequivoca le mutazioni che sottendono alle malattie in esame, non è possibile escludere la possibilità che più di una variante possa essere identificata in una singola famiglia con possibile significato patogenetico.

Per discriminare varianti fenotipicamente neutre da mutazioni patogenetiche, abbiamo analizzato il grado di conservazione delle sequenze, l'espressione tissutale, la possibile funzione molecolare dei geni, ed il possibile effetto delle mutazioni mediante strumenti bioinformatici.

Le potenziali varianti di incerto significato clinico richiedono una validazione sperimentale per dimostrare il probabile significato biologico.

L'identificazione di nuovi geni responsabili delle patologie che coinvolgono lo sviluppo neuronale potrà far luce sui meccanismi che regolano le funzioni cerebrali e sugli eventi patogenetici che sottendono ad alterazioni dello sviluppo cerebrale.

3. MATERIALI E METODI

3.1 Estrazione del DNA

Per l'estrazione e la purificazione del DNA da sangue periferico è stato utilizzato il QIAmp DNA Blood Midi Kit della Qiagen, in modo da ottenere DNA ad alto peso molecolare, non degradato e con rapporti 260/280 di circa 1.8 e 260/230 di circa 2.0.

La qualità del DNA è stata valutata mediante Nanodrop ND-1000 (NanoDrop Technologies, Wilmington, DE).

Ulteriore controllo di qualità del DNA prima del sequenziamento dell' esoma è stata la corsa elettroforetica su gel di agarosio allo 0,8% che ci esclude la presenza di eventuali degradazioni del campione.

3.2 Genotipizzazione e Homozygosity Mapping

I pazienti ed i loro parenti di primo grado sono stati genotipizzati utilizzando un vetrino ad alta densità Human610-Quad Infinium HD BeadChip utilizzando la piattaforma Illumina iSCAN e i dati elaborati mediante il Genotyping Module GenomeStudio.

È stato calcolato un rapporto Z-scores nonparametrico (ZLR scores) con il software Allegro per identificare SNPs omozigoti. I segmenti omozigoti con almeno 100 SNPs o di dimensione superiore ad 1 Mb saranno successivamente analizzati mediante NGS.

Inoltre, è stato calcolato il coefficiente di incrocio (coefficiente F) di ogni individuo in base alle informazioni genomiche e i loro LOD scores sono stati usati mediante la funzione statistica FLOD sviluppato appositamente per il mappaggio per omozigosi ed implementata nel programma FEstim che consente l'inclusione di individui consanguinei senza conoscenza della loro genealogia.

Le mutazioni responsabili di malattie recessive si trovano su un frammento cromosomico che viene trasmesso attraverso le generazioni ("identical-by-descent", IBD) e può variare di dimensioni in base al numero di ricombinazioni avvenute durante la trasmissione. Tale segmento IBD è caratterizzato da un insieme di polimorfismi in omozigosi ed è rilevabile pertanto mediante genotipizzazione.

Le attuali metodologie microarray-based permettono la tipizzazione di centinaia di migliaia di SNPs per poter identificare segmenti IBD molto piccoli originari da progenitori molto lontani⁷.

3.3 Sequenziamento NGS

Il DNA di interesse così ottenuto è stato trattato in service per alcune famiglie con un kit Illumina (Nextera Exome Enrichment Kit 62MB) ed altre con un kit Agilent Technologies (SureSelect Clinical research exome 54Mb).

I frammenti generati sono stati sequenziati paired-end tramite piattaforma Illumina HiSeq2000 e HiSeq1000, sequenziatori che utilizzano la tecnologia ad epifluorescenza che, mediante il software estrae l'intensità delle immagini ed esegue qualitativamente l'identificazione delle basi.

Il sequenziatore è in grado di leggere in una sola corsa 6 miliardi di reads della lunghezza di circa 200 bp in 8-10 giorni.

Il flusso di lavoro con HiSeq è suddiviso in tre fasi principali: generazione della libreria, preparazione dei cluster e sequenziamento (Fig.1)

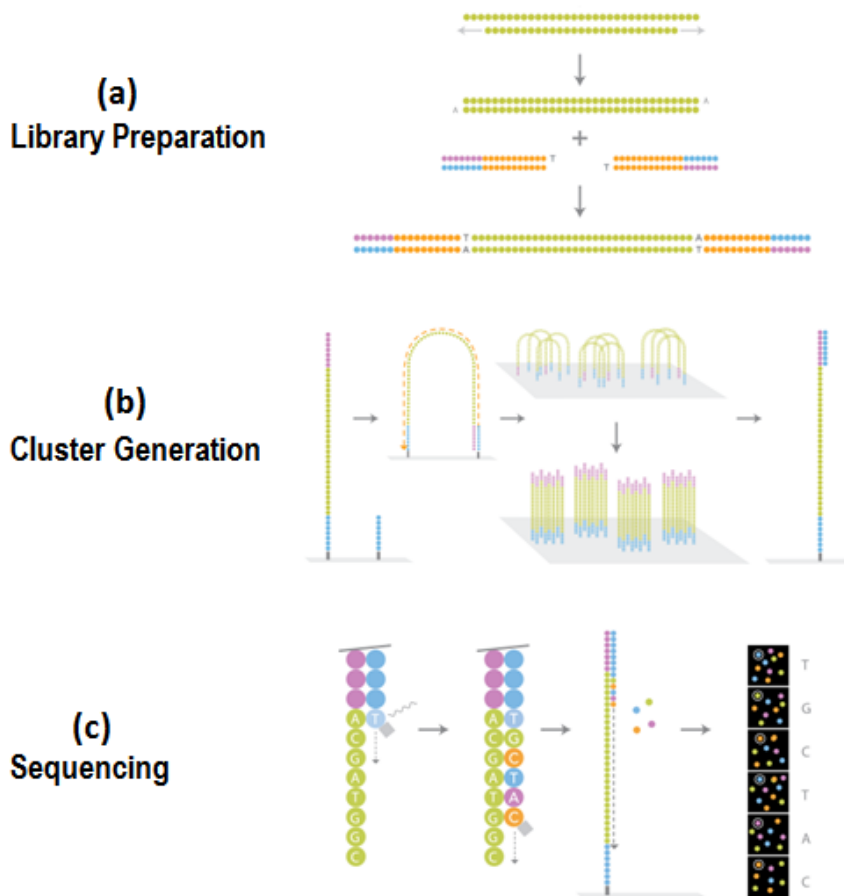


Fig.1: Rappresentazione schematica della procedura di sequenziamento della piattaforma Illumina HiSeq2000 (Illumina Inc., 2013).

In una prima fase (a), i campioni di DNA vengono frammentati in maniera randomizzata per la produzione di librerie di frammenti della lunghezza di 100-300 bp, alle cui estremità vengono aggiunti degli oligonucleotidi adattatori necessari ad ancorare e immobilizzare i frammenti sul supporto dove avverrà la successiva amplificazione.

La seconda fase (b) è rappresentata dall'amplificazione necessaria per generare il numero di copie di ogni singolo frammento di DNA (generazione di cluster). La generazione di copie dei singoli frammenti avviene su un supporto solido (*flow cell*) sul quale sono presenti oligonucleotidi complementari agli adattatori ligati. I frammenti di DNA si legano con entrambe le estremità a questi oligonucleotidi, formando delle strutture a ponte (*bridge PCR amplification*). A questo punto la DNA polimerasi sintetizza i filamenti complementari a quelli presenti, formando strutture a doppio filamento, che vengono poi denaturate. Con la rottura dei legami idrogeno che tengono unito il doppio filamento, si ottengono nuovamente due filamenti separati, numericamente raddoppiati rispetto allo stato iniziale. La catena di processi appena descritta si ripete un gran numero di volte,

fino ad ottenere un cluster (o fascio di frammenti), ciascuno composto da circa 100-200 milioni di ampliconi clonali separati. Poiché i clusters contengono sia la copia identica all'originale, sia il filamento inverso, è necessario ripulire il cluster dai frammenti antisenso prima di procedere con l'operazione di sequenziamento vera e propria.

L'ultima fase (c) prevede che ai frammenti di ogni cluster venga eseguita la sintesi del primer (innesco), il frammento di DNA che dà il via alla reazione di sequenziamento; questa, grazie alla tecnologia, può essere effettuata su centinaia di milioni di cluster in parallelo.

Ogni ciclo coinvolge una DNA polimerasi e i quattro dNTP (deossinucleotidi, usati dalla polimerasi per formare la doppia catena) a cui sono state apportate due modificazioni: l'incorporazione di un marcatore fluorescente e l'aggiunta di un terminatore reversibile. Il marcatore reagisce in maniera diversa per ognuno dei quattro nucleotidi quando sottoposto a onde laser e permette dunque l'identificazione della base che è stata sequenziata. Il terminatore è una molecola che blocca il gruppo ossidrilico impedendo l'ulteriore sintesi, in modo da garantire l'incorporazione di una sola base; è detto reversibile in quanto può essere dissociato chimicamente, riattivando la sintesi. Dopo ogni incorporazione, un laser eccita il fluorescente del dNTP generando un'emissione luminosa che ne permette l'identificazione. Dopodiché il terminatore e l'etichetta fluorescente vengono rimossi, in modo da abilitare il sequenziamento della base successiva.

Terminato il sequenziamento, le sequenze nucleotidiche vengono "estratte" dai dati di immagine generati dalla piattaforma di sequenziamento tramite algoritmi di base-calling che consentono di convertire l'informazione di immagine in informazione di sequenza.

3.4 Analisi bioinformatica

Per quanto le varie tecnologie di sequenziamento seguano strade diverse, forniscono tutti come output delle sequenze.

Per ciascun paziente vengono generate due sequenze grezze, forward e reverse, in formato Fastq che subiscono un processo bioinformatico di:

1. Controllo qualità e Allineamento
2. Variants calling
3. Annotazione e filtraggio

3.4.1 Controllo qualità e Allineamento

Il primo step dell'analisi è un controllo di qualità delle reads (Fastq format) effettuata mediante FastQC [<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>] tramite cui le reads di bassa qualità vengono scartate. Le sequenze grezze che passano il controllo di qualità vengono allineate e mappate ad un genoma umano di riferimento (GrCh37/Hg19) usando il tool BWA (Burrows-Wheeler Aligner)⁸.

Il software alla fine del processo genererà un file BAM (binary alignment mapping), il quale verrà sottoposto ad annotazioni e filtraggi.

3.4.2 Variants Calling

Una volta che le sequenze dei geni sono state ricostruite tramite l'allineamento, occorre individuare tutti i punti nei quali i geni differiscono dalle sequenze di riferimento del genoma umano archiviate nei database.

Tecnicamente il variant calling viene eseguito in automatico grazie all'ausilio di un software detto Genome Analysis Toolkit (GATK). Una volta terminata l'operazione, il risultato viene generalmente salvato in un file VCF (Variant Call Format). Uno dei problemi maggiori di un'operazione di variant calling è data dalla difficoltà nel riuscire a distinguere le varianti vere dalle varianti irreali, dovute ad artefatti del sequenziamento o ad errori nella fase di allineamento^{9,10}.

3.4.3 Annotazione e Filtraggio

Con i passaggi precedenti viene generata una lista di migliaia di potenziali differenze tra il genoma di riferimento e il genoma preso in esame, con valori di qualità della chiamata delle varianti. Il passaggio successivo è quello di determinare quale di queste varianti, non dovute a errori di sequenziamento, contribuiscono alla patologia.

Si eseguono annotazioni attraverso database di popolazioni come dbSNP147, ExAC, GnomAD e di predizione di varianti come il dbNSFP che contiene i dati relativi a tools come Polyphen, Sift, Mutation Taster, CADD, GERPP+, fondamentali per il successivo filtraggio delle varianti patogenetiche di un certo impatto funzionale.

3.4.4 Pipeline

Il workflow utilizzato per lo scopo del progetto si articola in 10 fasi:

- MapReads to Reference: in questa fase le reads vengono mappate, ossia allineate alla versione più recente del genoma umano di riferimento hg19 (GRCh37), al fine di determinare l'esatta localizzazione e rilevare eventuali mismatch di sequenza.
- Local realignment: consente la rifinitura del mappaggio delle reads a livello locale in funzione della presenza di polimorfismi noti di tipo indel.
- Coverage Analysis: consente di analizzare la percentuale di quanto è coperto l'esoma, stimando il grado di efficienza dell'esperimento eseguito.
- Remove duplicate: è una plug-in che consente al sistema di eliminare tutti i duplicati di sequenza generati da amplificazione di PCR.
- Probabilistic variants detection: attraverso calcolo logaritmico il programma fornisce la probabilità che vi sia o meno una variante. Generalmente, si richiede una coverage minima di 10X (ovvero di 10 sequenze) al fine di ottenere la presenza o meno della variante d'interesse ottenuta da almeno una sequenza forward ed una reverse.
- Filter Base on Overlap: si basa su un sistema di confronto, che sfrutta un file di riferimento contenente le informazioni relative alle posizioni genomiche del kit di arricchimento utilizzato per l'allestimento della library. A tale scopo, il programma scarta tutte le varianti "off target" che non rientrano in questi parametri.
- Variant or Annotation Track: consiste nell'annotare ovvero dare un'identità ai polimorfismi rari e comuni, assegnando loro un codice numerico identificativo (detto RS). Si possono così distinguere polimorfismi frequenti o rari nella popolazione, mentre per le varianti non presenti nei database di popolazione non verrà assegnato alcun codice identificativo corrispondenti a varianti ultrarare.
- Annotate with Overlap Information: tale plug-in permette di annotare, attraverso un file di riferimento di testo (bed file), le posizioni genomiche di tutti i geni sequenziati consentendo di assegnare ad ogni variante il gene di riferimento.

- *Filter against known in Variants* (filtraggio dei polimorfismi comuni): dai passaggi precedenti si ottengono varianti rare, polimorfismi comuni e rari. Con questo ulteriore filtraggio si eliminano i polimorfismi con incidenza nella popolazione >1% individuando solo i polimorfismi e le varianti rare.
- *Amino Acid Changes/Splice sites* (annotazione predizione degli amminoacidi e siti di splicing): è un plug-in che usa due bed files, di cui uno relativo alle coordinate delle coding-sequences (CDS) dei geni che saranno tradotti in proteine e l'altro con le sequenze di riferimento di predizione di splicing.

Con tale sistema è possibile capire, in base al tipo di variante, quale nuovo amminoacido si viene a generare e quindi se si tratta di una mutazione sinonima, non-sinonima o una variante di splicing.

Il risultato sarà l'annotazione di eventuali varianti di splicing, sinonime, missenso, non-senso e frameshift, che determinano la possibile patogenicità dell'alterazione genica. Ulteriori filtri che vengono effettuati si basano sul probabile meccanismo genetico, isolando soltanto le varianti d'interesse e sul possibile impatto funzionale.

Le varianti candidate sono state poi validate, attraverso l'analisi di segregazione in Sanger sequencing.

3.5 Validazione Sanger

Tutte le varianti individuate sono state sequenziate nei probandi e nei genitori attraverso sequenziamento Sanger⁵ che utilizza il metodo a terminazione di catena (DNA sequencing with chain-terminating inhibitors).

Durante il processo elettroforetico i segmenti vengono eccitati da una luce laser, passano davanti a un rivelatore che capta la lunghezza d'onda e l'intensità delle emissioni fluorescenti, identificando quale nucleotide è presente.

Le informazioni sono trasformate in picchi di colore diverso (uno per ogni nucleotide) con aree proporzionali all'intensità di emissione, vengono poi lette in modo automatico.

3.6 Coltura di fibroblasti

È stata eseguita una biopsia cutanea dei pazienti (punch biopsy), utilizzando un bisturi a lama circolare in anestesia locale. I frammenti di tessuto ottenuti, si adagiano su piastra petri con terreno di coltura dando così origine ai fibroblasti.

I fibroblasti ottenuti sono stati messi in coltura utilizzando come terreno RPMI contenente FBS (16%), Glutamax (100X), Pen/Strep (100X), Fungizone (100X), il quale va a mimare le condizioni presenti nel derma *in vivo* e che consente ai fibroblasti di crescere e moltiplicarsi in due settimane.

3.7 Estrazione delle proteine

La procedura di estrazione delle proteine consta di diverse fasi. Inizialmente i fibroblasti dei pazienti cresciuti in coltura, vengono lisati con una soluzione di RIPA Lysis Buffer, rompendo così le cellule e mantenendo l'integrità delle proteine.

Successivamente, il campione viene sonicato e centrifugato andando così a ottenere due fasi: una fase liquida, surnatante, e una fase solida, pellet.

Infine si preleva il materiale di interesse e si effettua un dosaggio proteico, con metodo Bradford.

3.8 Western blot

Per dimostrare l'impatto della variante a livello proteico abbiamo effettuato un Western Blot.

Tecnica che permette di valutare la presenza o assenza delle proteine¹¹.

Queste vengono separate in un gel di poliacrilammide in base al proprio peso molecolare e immobilizzate su una membrana per poi essere acquisita attraverso anticorpi specifici. Le fasi principali sono:

- 1) *Elettroforesi su gel di poliacrilammide (SDS-PAGE)*
- 2) *Trans-blotting su nitrocellulosa o PVDF*
- 3) *Blocking o saturazione*
- 4) *Incubazione con anticorpo primario*
- 5) *Lavaggi in PBSt*
- 6) *Incubazione con anticorpo secondario*
- 7) *Lavaggi in PBSt*
- 8) *Acquisizione di proteine con ECL*

3.9 Estrazione RNA

L'estrazione di RNA totale da fibroblasti è stato ottenuto attraverso risospensione in Trizol che separa una miscela in una fase organica apolare, composta da fenolo cloroformio, un'interfase, contenente DNA, e una fase acquosa superiore, in cui si trova esclusivamente l'RNA. È stato estratto l'RNA con il QIAGEN RNeasy mini kit utilizzando una DNasi.

3.10 Retrotrascrizione

Una volta estratto l'RNA dai fibroblasti è stato effettuato una retro-trascrizione per ottenere il cDNA.

L'enzima che catalizza questa particolare reazione è la trascrittasi inversa, e consiste in una particolare DNA polimerasi che utilizza come filamento stampo una molecola di RNA.

La trascrittasi inversa, per catalizzare la sintesi di una catena di cDNA, necessita di un RNA stampo, deossiribonucleotidi trifosfati (dNTPs), e random examer primer, i quali fungono da inneschi per l'enzima. È stato utilizzato il kit della ditta Bio-Rad iScript cDNA Synthesis Kit.

3.11 qPCR e primers ER stress

La qPCR è stata effettuata in triplicato con 2x Sso Fast EvaGreen Supermix in CFX96 Real Time detection system (biorad) usando oligonucleotidi disegnati secondo criteri e programmi bioinformatici (Primers3 , Primer-BLAST, UCSC-Blat tool) specifici, generando frammenti di ~100bp ciascuno:

- BIP:
 - 5'-TGTTCAACCAATTATCAGCAAACCTC-3';
 - 5'-TTCTGCTGTATCCTCTTCACCAG-3';
- CHOP:
 - 5'-AGAACCAGGAAACGGAAACAGA-3';
 - 5'-TCTCCTTCATGCGCTGCTTT-3'.

3.12 Immunoistochimica (IHC)

Questa procedura consente di evidenziare la formazione di complessi antigene-anticorpo, all'interno di un tessuto grazie ad appositi sistemi di rivelazione.

Sono state tagliate con criostato sezioni di 7 um del tessuto d'interesse, processati con protocolli di colorazione di ematossilina-eosina.

3.13 Immunocitochimica (ICC)

Questa tecnica permette di confermare la presenza di una proteina in cellule in coltura e di determinare la posizione a livello cellulare. Si utilizzano anticorpi monoclonali o policlonali.

- 1:200 Novocastra Rabbit Polyclonal Antibody S-100 Protein (NCL-L-S100);
- 1:30 Mouse Monoclonal 1H2G8 anti-human EMILIN1;
- 1:200 Rabbit Polyclonal AS556 anti-human EMILIN1.

4. RISULTATI

I risultati ottenuti si sono focalizzati su casi familiari di malattie rare del neurosviluppo non ancora diagnosticate geneticamente, composte da almeno due soggetti affetti. Per ogni famiglia sono state create diverse pipeline bioinformatiche sulla base della modalità di trasmissione genetica attesa e dei possibili meccanismi patogenetici, in modo da selezionare tra le molteplici varianti identificate con l'esoma, quelle con significato clinico.

Dopo aver ricevuto il consenso delle 7 famiglie per lo studio ed in accordo con il centro di sequenziamento del TIGEM di Napoli e del CRS4 di Pula, sono stati sequenziati il DNA di pazienti e di genitori attraverso piattaforma Illumina Hiseq1000/2000.

L'analisi dei dati grezzi ottenuti al next generation sequencing (NGS) mediante pipeline bioinformatiche selezionate per le diverse famiglie; sono state individuate in 4 famiglie mutazioni patogenetiche, supportate anche da dati sperimentali. Inoltre, non è stato possibile raggiungere un risultato positivo nelle restanti 3 famiglie, probabilmente dovuto all'implicazione di meccanismi genetici non valutabili attraverso exome sequencing.

4.1 Famiglia 1

La prima famiglia è composta da 3 pazienti affetti con microcefalia, ipoplasia pontocerebellare con difetti dello sviluppo, epilessia ed anomalie dei denti (Fig.2).

Abbiamo eseguito un homozygosity mapping attraverso un vetrino ad alta densità Human610-Quad Infinium HD BeadChip su piattaforma Illumina iSCAN, individuando diverse regioni di omozigosi sui cromosomi 1-2-3-11 con regioni di circa 1Mb come descritto in Figura 3.

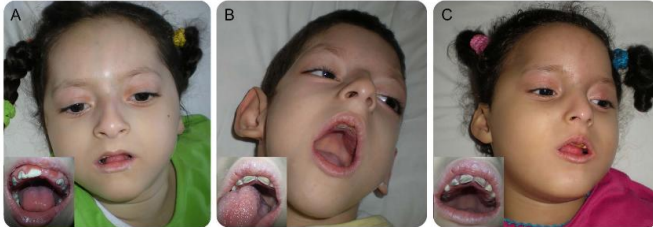


Fig.2: pazienti appartenenti alla famiglia 1 con microcefalia, epilessia ed anomalie dei denti

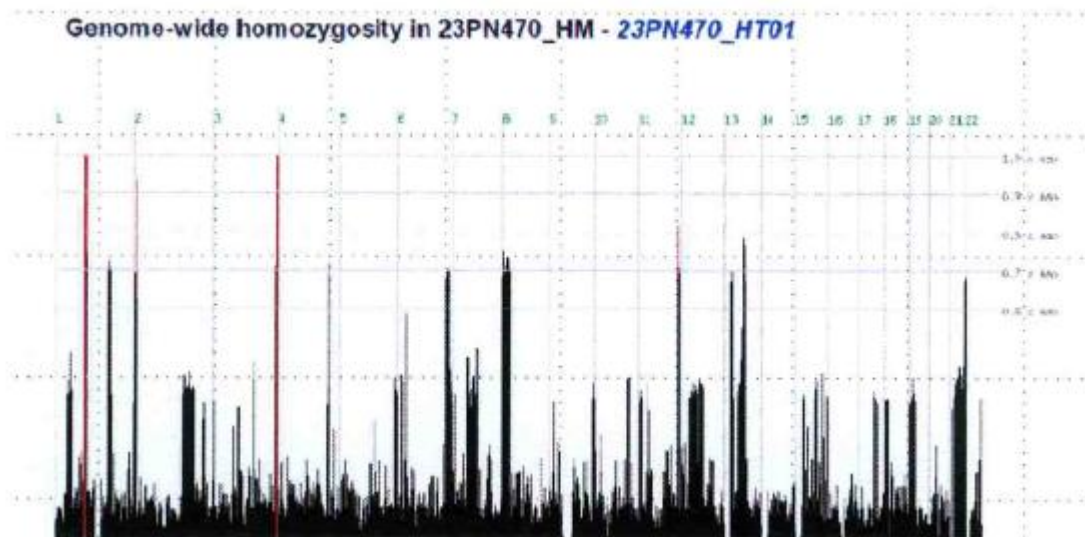


Fig.3: regioni di omozigosi presenti nel paziente sui cromosomi 1-2-3-11

Il DNA genomico della famiglia è stato sottoposto a NGS e trattato con un kit di arricchimento della ditta Agilent Technologies (SureSelect Clinical research exome 54Mb) e sequenziato attraverso sequenziatore HiSeq1000 dell' Illumina, tale sequenziamento ha garantito una copertura media dell'intero esoma del 95% a 10X (ogni base è coperta da almeno 10 sequenze) riuscendo ad arricchire diverse varianti all'interno di queste regioni.

Le sequenze grezze ottenute sono state analizzate attraverso una pipeline informatica costruita nel nostro laboratorio con il software CLC Bio Genomics Workbench 7.5.1.

Il numero di varianti ottenute dal processamento dei pazienti III.1, III.2, III.3 è rispettivamente di circa 40.738, 51.161 e 48.328.

Le varianti sono state selezionate sulla base del grado di intolleranza dei geni alla variabilità e sulla base di database di frequenza di popolazione come dbSNP147, ExAC e GnomAD con frequenza dell'allele minoritario $\leq 0,001\%$.

Inoltre, sono state selezionate le varianti condivise tra i pazienti e filtrate in base al possibile meccanismo ereditario, al tipo di mutazione (non-sinonima, insdel, splicing) e sull'impatto di patogenicità degli aminoacidi predetto attraverso programmi bioinformatici come Polyphen, Sift, Mutation Taster, CADD, GERP++. Infine, sono state effettuate le correlazioni genotipo/fenotipo sulla base delle varianti nei geni individuati.

Dopo aver effettuato i filtraggi, l'analisi ha messo in evidenza una variante patogenetica condivisa nei tre pazienti affetti presente nella regione di omozigosi del cromosoma 1.

Questa variante identificata è una delezione frameshift in omozigosi di una guanina a livello dell'esone 5 nel gene *AMPD2* (adenosina monofosfato deaminasi) #NM_001257360.1, c.495delG p.R165fs.

Tale gene codifica per una proteina che ha funzione di enzima deaminasico, importante nel metabolismo delle purine per la conversione di AMP in IMP. Un deficit di *AMPD2* comporta un accumulo di nucleotidi di adenosina e una deplezione dei nucleotidi di guanosina provocando una alterata traduzione delle proteine e difetti nel metabolismo^{14,15,16,17}.

La variante è stata validata attraverso Sanger sequencing in tutta la famiglia, nella quale i soggetti affetti presentavano una delezione in omozigosi ed i genitori erano portatori sani di tale variante (Fig.4).

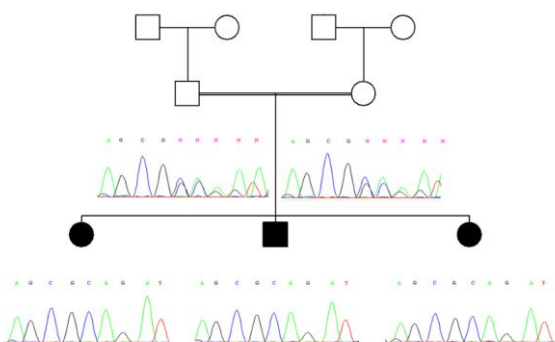


Fig.4: Albero genealogico e sanger sequencing della famiglia 1

Tale gene era stato precedentemente riportato da Aziu et al. nel 2013 sulla rivista Cell, in diverse famiglie con sindrome pontocerebellare (PCH)¹⁵. Il nostro studio espande il quadro fenotipico di questa sindrome la quale non era stata ancora associata a malformazione dentaria. Questa nuova variante patogenetica individuata espande lo spettro delle mutazioni e sottolinea quanto sia importante l'impatto di *AMPD2* per questa sindrome. Tale studio è stato oggetto di pubblicazione nella rivista Neurology Genetics.

Novel *AMPD2* mutation in pontocerebellar hypoplasia, dysmorphisms, and teeth abnormalities.

Accogli A, **Iacomino M**, Pinto F, Orsini A, Vari MS, Selmi R, Torella A, Nigro V, Minetti C, Severino M, Striano P, Capra V, Zara F. **Neurol Genet.** 2017 Aug 9;3:e179. doi: 10.1212/NXG.0000000000000179.

4.2 Famiglia 2

La seconda famiglia studio riguarda un paziente di nove mesi, originario di un piccolo paese siciliano (Fig.5) che all'ammissione presso il nostro ospedale nel 2015, presentava un quadro neurodegenerativo complesso e una grave encefalopatia ed atrofia muscolare, gravi ipotono, artrogriposi e ritardo dello sviluppo.

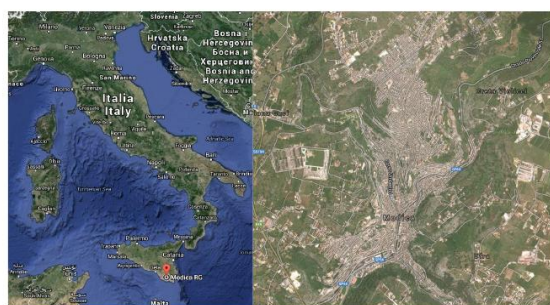


Fig.5: localizzazione del piccolo paese siciliano da 50000 abitanti

All'età di 6 mesi, soffriva di spasmi epilettici che coinvolgevano sia gli arti superiori che quelli inferiori e non ha risposto al trattamento con corticosteroidi.

La valutazione neurologica ha dimostrato una grave ipotonia globale con movimenti ridotti, nessun controllo della testa, plagiocefalia frontale, problemi respiratori e pugno debole (Fig.7 A).

Il paziente in esame soffriva di sporadici spasmi tonici degli arti superiori e inferiori.

Abbiamo eseguito lo screening per anomalie metaboliche risultato essere normale tranne per l'aumento di CK (976 U / L).

L'elettroencefalogramma (EEG) ha mostrato uno sfondo rallentato e anomalie epilettiche multifocali, più evidenti sulle aree occipitali sinistro e destra del cervello temporale (Fig.7 B).

La risonanza magnetica (MRI) del cervello ha mostrato atrofia corticale diffusa, grave assottigliamento della sostanza bianca e cambiamenti del segnale nella materia bianca periventricolare e ponte (Fig.7 C,D).

L'elettromiografia (EMG) ha confermato un danno neurogeno con presenza di maggiori ampiezze e potenziali di unità motoria polifasici (Fig.7 E). Attraverso una colorazione di Ematossilina/Eosina (H&E) su tessuto muscolare è stato messo in luce un danno della rete miofibrillare, comprese le fibre anulari dove sono in evidenza anche i nuclei (Fig.7 F).

La biopsia muscolare ha indicato un grave coinvolgimento del secondo motoneurone con sostituzione fibrosa dei tessuti muscolari, la presenza di fibre atrofiche raggruppate e l'ipertrofia di fibre di tipo I (Fig.7 G).

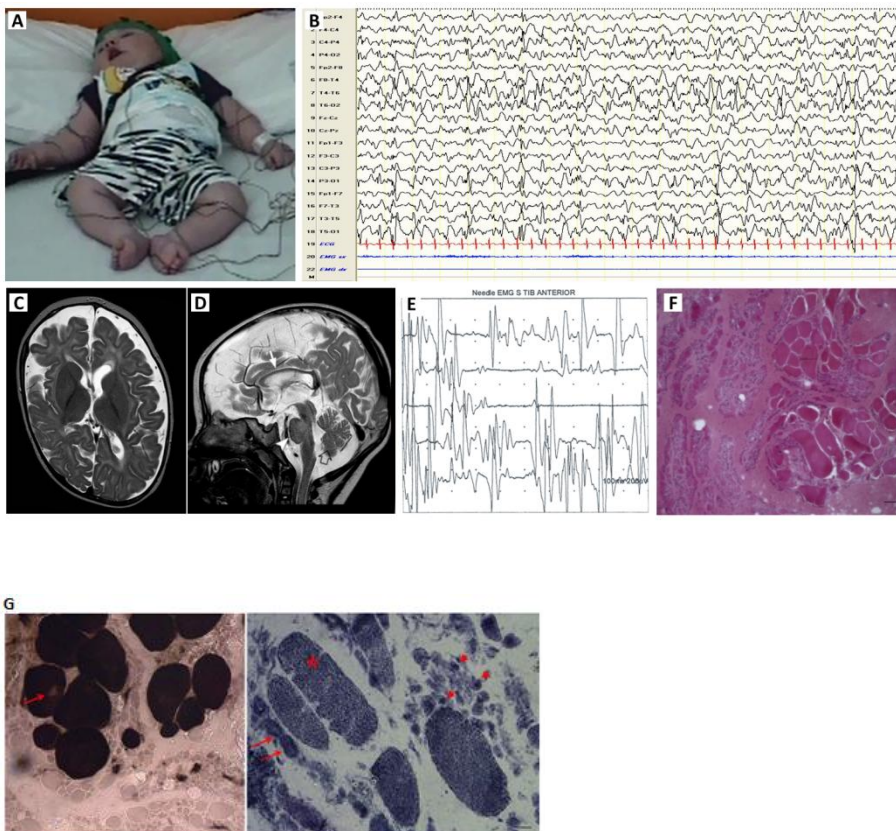


Fig7. (A-G): A) Paziente con variante PRUNE1; B) EEG presenta anomalie epilettiche multifocali; C-D) MRI, atrofia corticale e assottigliamento sostanza bianca con alterazioni del ponte; E) EMG presenta maggiori ampiezze e potenziali polibasici; F) H&E presenza di nuclei con danno

miofibrillare; G)ATPasi e NADH staining con presenza di fibre tipo I e II (freccia rossa) nell'immagine sinistra, mentre in quella di destra lesioni di fibre grandi (stella rossa) e piccole fibre che indicano lesioni (testa-freccia)con depositi citoplasmatici.

Per la concomitanza dei segni sia del sistema nervoso centrale che periferico ha permesso inizialmente di escludere una serie di geni coinvolti in un quadro SMA (*SMN1*, *EXOSC3* e *EXOSC8*) attraverso sequenziamento di singoli geni, risultati negativi. Abbiamo quindi sottoposto il campione del soggetto ad homozygosity mapping, attraverso un vetrino ad alta densità Human610-Quad Infinium HD BeadChip su piattaforma Illumina iSCAN, selezionando una regione di omozigosi specifica sul cromosoma 1 (Fig. 8).

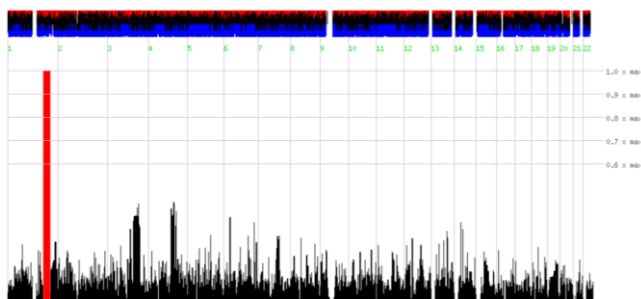


Fig. 8: Analisi di Homozygosity mapping della regione di omozigosi sul Cromosoma 1.

Dopo aver ricevuto il consenso, abbiamo eseguito uno studio di NGS garantendo una copertura media del 90% ad una coverage del 10X, per una migliore comprensione dei dati abbiamo sottoposto a sequenziamento anche i genitori.

Abbiamo trattato il DNA genomico per l'arricchimento delle sequenze codificanti con un kit della ditta Agilent Technologies (SureSelect Clinical research exome 54Mb) e sequenziati attraverso un sequenziatore HiSeq1000 della ditta Illumina.

Il numero di varianti ottenute dal processamento del paziente II.2 è di circa 135.738, tali varianti hanno subito una selezione sulla base della frequenza dell'allele minoritario ($\leq 0.001\%$), dell'impatto funzionale (non-sinonima, insdel, splicing) e del meccanismo genetico. Abbiamo quindi individuato 3 varianti rare nei geni *PRUNE1*, *SPAG17* e *ECM1* nella regione di omozigosi (Tabella 1), validate attraverso metodica Sanger (Fig.9).

Chr.	Region	Type	Ref.	Alt.	Zygoty	PolyPhen	Gene	Coding Region change	Amino acid change	MAF (ExAC)
1	118550780	SNV	G	A	Homozygous	Damaging	SPAG17	ENST0000036338: c.4474C>T	ENSP00000337804:p.R1492W	6,25x10 ⁵
1	150483454	SNV	G	T	Homozygous	Damaging	ECM1	ENST00000369049: c.569G>T	ENSP00000358045:p.R190L	8,94x10 ⁶
1	150991126	SNV	G	A	Homozygous	Damaging	<u>PRUNE1</u>	ENST00000271620: c.316G>A	ENSP00000271620:p.D106N	4,46x10 ⁵

Tabella 1. Lista delle varianti ereditate nel probando

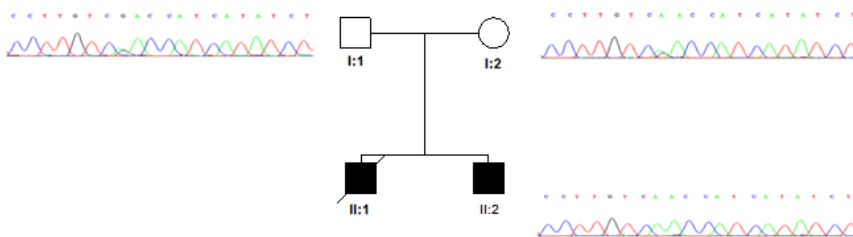


Fig.9: Albero genealogico con relativo sanger della variante omozigote in PRUNE1.

Recentemente sono state evidenziate mutazioni nel gene *PRUNE1* in una piccola coorte di pazienti con un complesso fenotipo neurologico del neurosviluppo, caratterizzato da ipotonia centrale, gravi ritardi nello sviluppo, microcefalia, convulsioni, atrofia e atrofia corticale e/o cerebellare^{19,20,21}.

Tali mutazioni sono state identificate e descritte nel 2015 sulla rivista *Neuron* da Karaca et al. in tre famiglie consanguinee.

Sulla base di questi dati ci siamo concentrati sul gene *PRUNE1*, unica variante che potesse spiegare il fenotipo del paziente. Ad oggi, i restanti due geni *SPAG17* e *ECM1* non sono stati descritti essere associati a malattia o dimostrati essere causativi di patologie umana.

Le famiglie descritte da Karaca et al condividono con il nostro paziente la stessa variante c.316G>A p.D106N, che ricade all'interno di uno dei 4 domini fondamentali della proteina, il dominio DHH (Asp-His-His) per la ripetizione di questi tre amminoacidi^{18,19,20}.

Nel 2017 è stato dimostrato da Zollo et al. il coinvolgimento della proteina nella regolazione della motilità e della proliferazione cellulare attraverso le vie dei microtubuli²⁰.

Tale lavoro pubblicato sulla rivista *Brain* ha evidenziato anche il ruolo nello sviluppo fisiologico del neurite, nella formazione di sinapsi e nella sopravvivenza del precursore del neurone.

A differenza dei lavori descritti in precedenza il nostro paziente associa, oltre che un interessamento del sistema nervoso centrale anche un quadro con interessamento muscolare e quindi del sistema periferico.

Questo paziente dimostra come il fenotipo sotteso da mutazioni in *PRUNE1* presenti difetti del secondo neurone motorio, come documentato dai risultati della biopsia clinica e muscolare.

Sebbene la combinazione del difetto del secondo motoneurone spinale e il disturbo dello sviluppo cerebrale non siano rari, questo coinvolgimento specifico non è mai stato riportato in altri pazienti *PRUNE1*.

Pertanto questo gene deve essere testato in pazienti con fenotipo complesso che coinvolge sia i neuroni centrali che periferici. Tale studio è stato pubblicato nella rivista *European Journal of Pediatric Neurology*.

Spinal Motor Neuron Involvement In a Patient With Homozygous *PRUNE1* Mutation.

Michele Iacomino MS, Chiara Fiorillo MD, Annalaura Torella PhD, Mariasavina Severino MD, Paolo Broda BS, Catia Romano MD, Raffaele Falsaperla MD, Giulia Pozzolini BS, Carlo Minetti MD, Pasquale Striano MD, PhD, Vincenzo Nigro MD, Federico Zara PhD

European Journal of Pediatric Neurology

4.3 Famiglia 3

La terza famiglia è costituita da padre (I-1) e tre figli (II-1, II-2, II-3) (Fig. 10), i quali presentavano, già nella prima infanzia, un quadro clinico caratterizzato da una amiotrofia e una lassità distale dei lembi inferiori²².

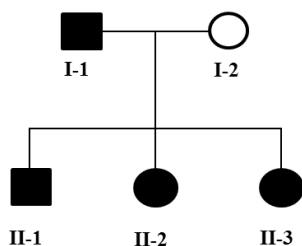


Fig.10: Albero genealogico famiglia 3

In comparazione con il padre, i due figli maggiori dispongono di un quadro clinico più severo, caratterizzato da una anormale andatura, una lassità articolare marcata ai piedi e un aumento del riflesso tendineo.

Studi di elettromiografia mostrano un cambiamento neurogeno e una ridotta ampiezza dei potenziali di azione dei muscoli degli arti inferiori, le conduzioni nervose sensoriali sono del tutto normali. L'aumento del riflesso tendineo e modello di andatura, con ginocchia addotte e piegate, suggeriscono anche il coinvolgimento del motoneurone superiore.

Il fenotipo difficilmente può essere inquadrato in una diagnosi specifica; tuttavia la gravità clinica è più pronunciata nei bambini in comparazione con il padre, suggerisce una espressività variabile dell'allele mutante.

La lassità articolare è marcata ai piedi e alle caviglie e minima agli arti superiori, quindi non può soddisfare i criteri per il tipo di ipermobilità di Ehlers Danlos (EDS)²³.

Sono inoltre documentati una bassa ampiezza di CMAP del nervo peroneo e, in associazione con alterazioni neurogeniche all'EMG e biopsia muscolare, possibili segni di una neuropatia motoria distale²².

Abbiamo eseguito un NGS della famiglia attraverso piattaforma Illumina Hiseq2000 con un kit di arricchimento Nextera Rapid Capture Expanded Exome 62Mb.

Abbiamo selezionato le varianti attraverso una pipeline disegnata per questo caso specifico.

Il numero totale di varianti per i pazienti II.1,II.2,II.3 sono rispettivamente 18.508, 19.043, 18.628 di cui 945 condivise allo stato eterozigote secondo un modello ereditario autosomico dominante, infine selezionate secondo criteri specifici come varianti ultra rare con una MAF $\leq 0,0001\%$ sulla base di database di popolazione e sul tipo di mutazione (non-sinonima, insdel, splicing) con caratteristiche patogenetiche.

Tra le varianti non sinonime riscontrate nei 3 pazienti, sono state selezionate due varianti patogenetiche:

- Una variante non-sinonima c.748C>T p.R250C in *EMILIN1* #NM_007046;
- Una variante di splicing c.5478+1G>T p.Q1826fs in *SPTAN1* #NM_001130438;

Il sequenziamento Sanger ha confermato la presenza di entrambe le mutazioni nei 3 soggetti affetti (II-1, II-2, II-3) e nel padre (I-2), mentre ne ha confermato l'assenza nella madre sana (I-2).

Entrambe le varianti nei suddetti geni sono risultate essere damaging da tools bioinformatici e assenti in database come ExAC e GnomAD.

Non riuscendo a discriminare le varianti patogenetiche sono stati effettuati studi funzionali, in quanto il gene *SPTAN1* era già stato associato ad un quadro di encefalopatia epilettica a differenza del gene *EMILIN1*, descritto in una unica famiglia di tre generazioni che causa un disturbo del tessuto connettivo associato ad una neuropatia sensoriale-motoria²⁴.

Attraverso studi funzionali con tecnica Western Blot (WB) per Sptan1, su proteine estratte da fibroblasti, i livelli di proteina Sptan1 marcata con anticorpo anti-Sptan1 erano sovrapponibili al confronto tra pazienti (II.2, II.1) e controlli (CN1,CN2), e per tale motivo questa variante di splicing non comportava nessun difetto a livello della proteina (Fig.11).

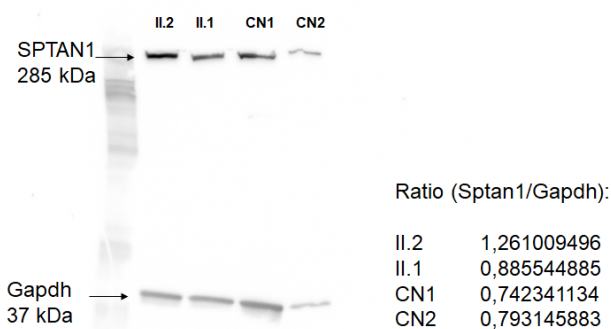


Fig.11: Western blot SPTAN1: linea 1: Marker; linea 2-3 rispettivamente pazienti (II.2,II.1); linee 4-5 rispettivamente controlli (CN1,CN2); nella parte inferiore la proteina di controllo GAPDH; sulla destra il rapporto tra i valori di SPTAN1/GAPDH.

Risultando negativa l'indagine su *SPTAN1*, sono stati concentrati gli sforzi sulla variante p.R250C di Emilin1 che presenta un residuo di extracisteina a livello della regione coiled-coil. Attraverso il tool Lupas^{12,13,25} che esamina le strutture coiled-coil è stata evidenziata una riduzione della formazione del potenziale coiled-coil e quindi un probabile difetto di folding della proteina (Fig. 12).

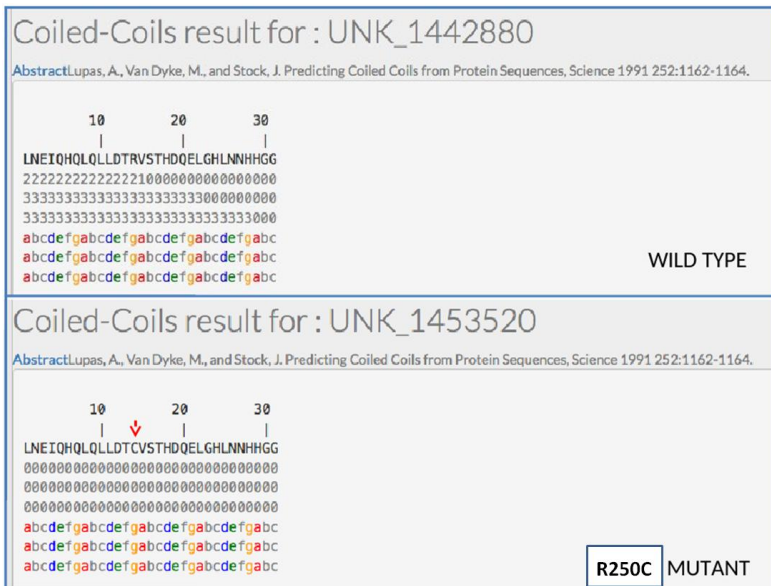
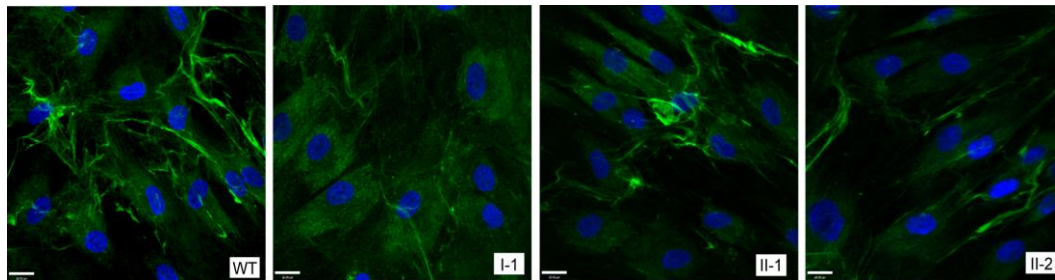


Fig 12: Tools Lupas coiled-coil valori di probabilità tra amminoacido wt con variante in *EMILIN1*

Conseguentemente, è stata effettuata una analisi funzionale su fibroblasti cutanei dei tre pazienti, e attraverso immunofluorescenza con anticorpo anti-Emilin1. Lo studio ha evidenziato un difetto di deposizione di Emilin1 nello spazio extracellulare, e disorganizzazione delle fibre di emilina scarsamente ramificate. Tali dati suggeriscono che la mutazione potrebbe compromettere la normale secrezione e/o la scorretta organizzazione della proteina nella matrice extracellulare (Fig.13a). Infatti, attraverso analisi di WB con anti-emilin1 e anti-tubulin (proteina di controllo) si può notare una riduzione massiccia dei livelli di Emilina nella matrice extracellulare dei pazienti (I-1,II-1,II-2) rispetto ai controlli (Wt) (Fig.13b).

a)



b)

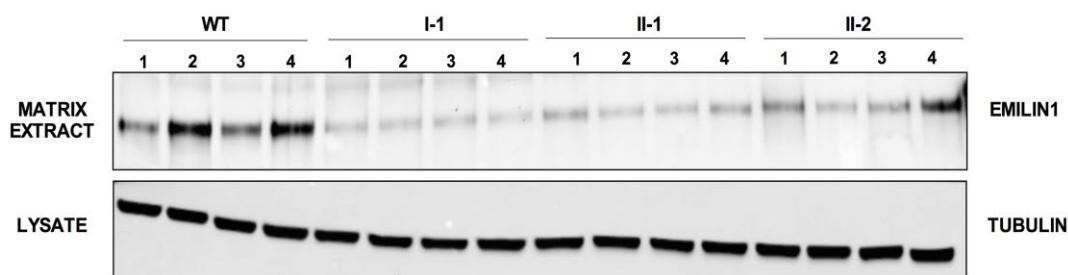


Fig. 13: (a) Immunofluorescenza con anticorpo anti-emilin1 su fibroblasti, una riduzione del rilascio di Emilin1 nei pazienti (I-1,II-1,II,2) rispetto al controllo Wt. (b) Western blot su matrice extracellulare con anticorpo anti-emilin1 su pazienti e controllo, con proteina di controllo Tubulina su lisati cellulari.

Dal precedente lavoro di Capuano et al, la variante individuata p.A22T presente nella regione N-terminale del peptide segnale della proteina, mediante analisi qPCR dei geni CHOP e BIP che sono dei marker molecolari dello stress del reticolo endoplasmatico, hanno dimostrato che la variante provocava difetti di accumulo e rilascio da parte del reticolo²⁴.

Per tale motivo, è stata eseguita una qPCR da RNA totale di II-1 e II-2 con CHOP e BIP, dove non sono state ottenute variazioni significative, probabilmente dovuto al fatto che la variante p.R250C fosse collocata a livello del coiled-coil, regione proteica non implicata nel meccanismo di rilascio del reticolo endoplasmatico (Fig. 14 a-b).

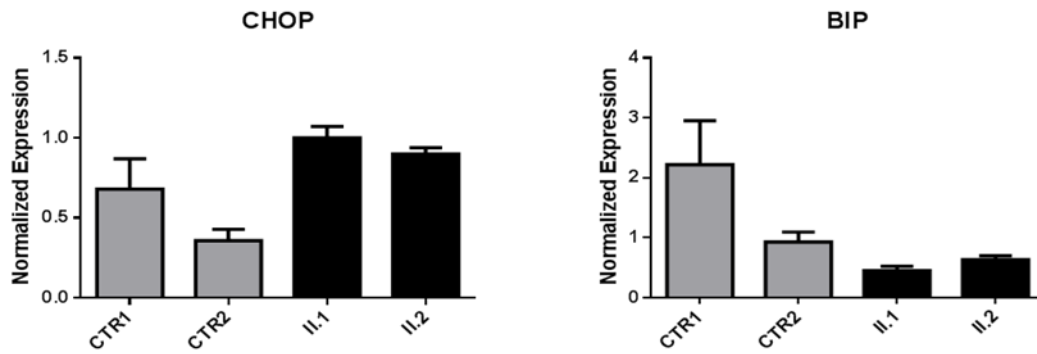


Fig.14: qPCR su RNA da fibroblasti per i geni responsabili dello stress del reticolo endoplasmatico: (a) CHOP e (b)BIP, mentre PPIA e RPL13A sono stati scelti come geni housekeeping. Infine i livelli di espressione sono stati normalizzati usando il metodo comparativo Ct con i due housekeeping PPIA e RPL13A, da Biorad CFX manager software.

L'analisi statistica è stata eseguita mediante il test di Kruskal-Wallis con le correzioni di Dunn's per le comparazioni multiple (n=3).

Infine, per sostenere il ruolo dell'emilina nel tessuto connettivo umano e/o nei nervi periferici, abbiamo studiato le sezioni muscolari del paziente II-2 e notato una forte presenza della proteina nel perimisio e in alcuni fasci nervosi intramuscolari.

La proteina S100 è stata utilizzata come marker di componenti neuroectodermici.

Abbiamo effettuato una co-immunofluorescenza delle proteine Emilin1 e S100, all'interno della struttura nervosa del muscolo.

L'anticorpo monoclonale contro Emilin1 utilizzato in verde, mentre l'anticorpo policlonale contro S100 in rosso.

Abbiamo notato che S100 è fortemente presente negli assoni nervosi, mentre emilin1 sembra marcare prevalentemente gli strati del tessuto connettivo intorno ai fascicoli nervosi (Fig.15).

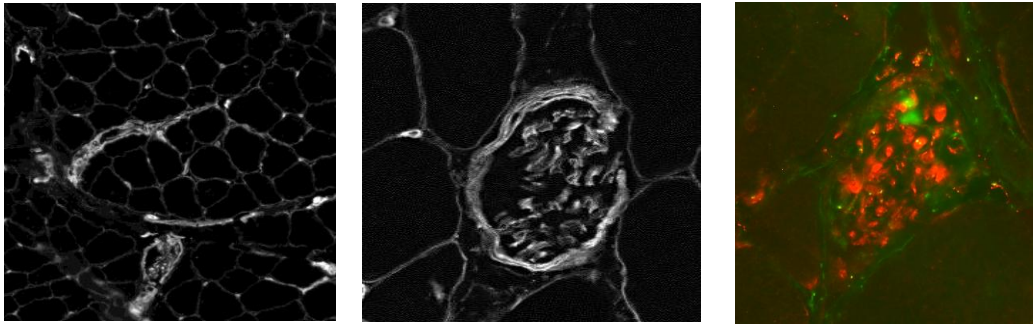


Fig 15: Immunofluorescenza con anticorpo policlonare per Emilin1 su un nervo di una biopsia muscolare del paziente II.2. Immunofluorescenza su nervo II.2 con anticorpo monoclonare in verde per emilin1 e rosso per S100.

Per sostenere ulteriormente un ruolo di EMILIN-1 nella funzionalità e nello sviluppo dei nervi periferici abbiamo inattivato emilin1a in zebrafish mediante un morfolino già testato.²⁶

Abbiamo osservato che emilin1a morphants a 48hpf mostrava difetti morfologici, come la curvatura del corpo (Fig.16A) e una compromissione dello sviluppo del cervello, con una differenza statisticamente significativa della fluorescenza nell'area cerebrale rispetto ai fratelli di controllo (Fig.16B-C). L'immunofluorescenza con l'anticorpo ZNP1 ha confermato che la down-regolazione di emilin1a era associata a difetti del motoneurone mostrando uno sviluppo anormale degli assoni, con diminuzione delle arborizzazioni (Fig.16D). Difetti locomotori sono stati osservati a 30hpf, attraverso l'aumento della frequenza di avvolgimento almeno tre volte superiore rispetto agli embrioni di controllo, suggerendo una compromessa eccitabilità neuronale (Fig.16E); inoltre, a 48hpf (Fig.16F) e 120hpf (Fig.16G-H), gli embrioni e le larve di morphants hanno mostrato difetti motori, con una diminuzione dell'attività locomotoria rispetto al wild-type. I dati ricavati dagli esperimenti attraverso la co-iniezione di spliceMO e cRNA di mutanti umani portatori di p[R250C] hanno avvalorato i possibili effetti patogeni della mutazione (Fig.16I), poiché non era in grado di recuperare il difetto locomotore a 48hpf. Al contrario, questo è stato in parte osservato con la co-iniezione di spliceMO e cRNA di umani e zebrafish WT (Fig.16L-M).

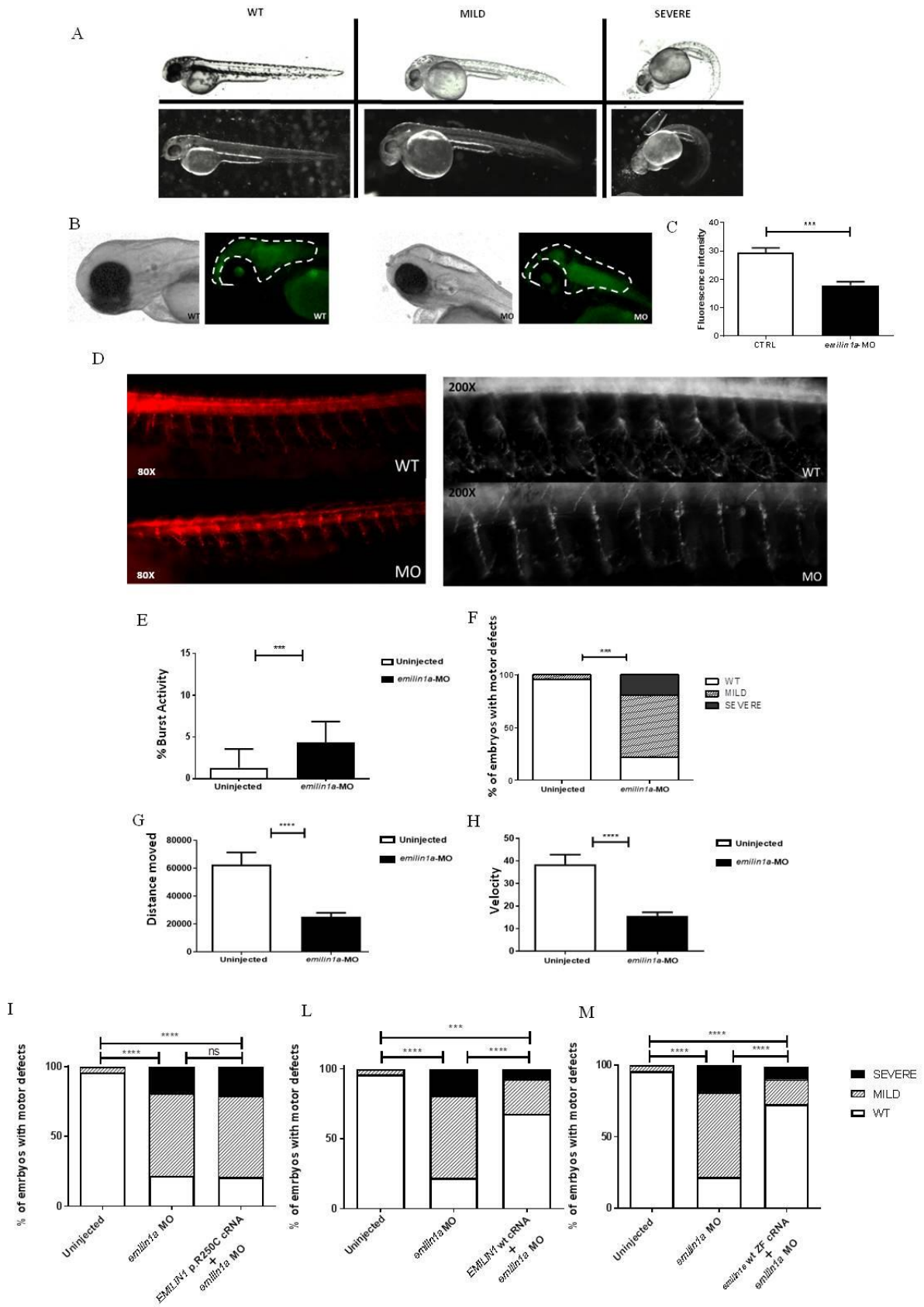


Fig.16: A) *emilin1a* morphants a 48hpf mostrano difetti morfologici. B) Sviluppo del cervello alterato di larve morphant a 3dpf rispetto ai fratelli di controllo. C) Quantificazione della

*fluorescenza cerebrale che mostra un'intensità di fluorescenza ridotta nelle larve emlin1a. (*** $P \leq 0,001$). D) emilin1a morphants a 48hpf mostrano uno sviluppo anormale generale di assoni di motoneuroni spinali (80x), con una ramificazione assonale ridotta rispetto ai fratelli di controllo (200x). E) Analisi di avvolgimento di embrioni 30hpf che mostrano un aumento ipereccitabilità neuronale degli embrioni emilin1a iniettati (*** $P \leq 0,001$). F) Diminuzione del movimento in risposta allo stimolo tattile dei morphants a 48hpf rispetto ai controlli non iniettati (*** $P \leq 0,001$). G-H) Compromissione locomotoria di larve 5 dpf iniettate con emilin1a morfolino rispetto ai fratelli wild-type, utilizzando un sistema di localizzazione (**** $P \leq 0,0001$). I) Analisi di risposta evocata tattile di embrioni co-iniettati a 48hpf con EMILIN1 p.R250C cRNA ed emilin1a MO umani che rivelano che il cRNA mutante non è in grado di recuperare il fenotipo locomotore. L-M) Recupero del fenotipo locomotore utilizzando cRNA di wt umani e zebrafish. (ns $P > 0,05$; *** $P \leq 0,001$; **** $P \leq 0,0001$).*

Gli studi funzionali eseguiti hanno dimostrato l'effetto patogenetico della mutazione in *EMILIN1*. Emilin1, è stato identificato come componente delle fibre elastiche ed è fortemente espresso nei tessuti dove flessibilità ed elasticità sono fondamentali. Attraverso lo studio del modello di zebrafish è stato dimostrato il coinvolgimento di EMILIN1 nel SNC (Sistema Nevoso Centrale) attraverso studi di locomozione e arborizzazione assonale da parte dei motoneuroni. È noto che la famiglia in esame abbia rilevato anomalie minori di rotazione e lieve deterioramento cognitivo che possa far pensare ad un coinvolgimento del SNC, seppure non grave, ed in linea con alterazioni lievi del SNC osservata in zebrafish. Lo studio confermato su embrioni di zebrafish può far pensare ad un coinvolgimento di EMILIN-1 nello sviluppo del sistema nervoso. Sicuramente con ulteriori pazienti mutati in questo gene sarà possibile espandere ulteriormente il fenotipo e giungere ad una correlazione genotipo-fenotipo definitiva. Tale studio è oggetto di un manoscritto in corso di pubblicazione.

Proposal paper:

A NOVEL EMILIN1 MUTATION IMPAIRS THE COILED COIL FORMING POTENTIAL AND CAUSES AXONAL MOTOR NEUROPATHY

Michele Iacomino, Roberto Doliana, Maria Marchese, Alessandra Capuano, Pasquale Striano, Paola Spessotto, Rosa Iodice, Fiore Manganelli, Giulia Bosisio, Paola Lanteri, Simona Baldassari,

Serena Baratto, Floriana Fruscione, Valeria Prada, Angelo Schenone, Paolo Broda, Alessandra Tessa, Alfonso Colombatti, Carlo Minetti, Filippo Maria Santorelli, Federico Zara*, Chiara Fiorillo*.

4.4 Famiglia 4

La quarta famiglia (Fig.17) è composta da un soggetto affetto con una grave ipoplasia del ponte e del verme, un quadro epilettico grave, con spasticità e profondo ritardo nello sviluppo.

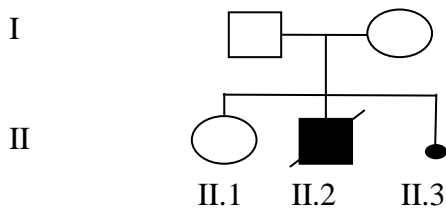


Fig.17: Albero genealogico famiglia 4 con due soggetti affetti.

Abbiamo eseguito exome sequencing con un kit Agilent SureSelect Clinical exome 54Mb su DNA genomico del probando II.2, il numero di varianti generate erano 45.848 . Tali varianti hanno subito diversi filtraggi specifici, consentendo di individuare una nuova variante in omozigosi c.790C>T p.R264X nel gene *WWOX* #NM_016373.3 già associato a certe forme di encefalopatia^{27,28,29,30,31}.

Recentemente, mutazioni in *WWOX* sono state identificate nell'ambito dei disordini del sistema nervoso centrale (SNC), evidenziando un ruolo precedentemente non rivelato di questo gene nel normale sviluppo e funzione del SNC^{27,28,29}.

Nel Gennaio 2018 il paziente è deceduto all'età di 13 anni, contestualmente la madre rimasta incinta di un altro figlio, presentava ai dati strumentali di laboratorio livelli di alfa-fetoproteina alterati ed all'analisi eco-tesla3D evidenti alterazioni strutturali, come assenza del corpo calloso e ipoplasia del verme.

È stata eseguita una diagnosi genetica su liquido amniotico con sequenziamento Sanger della variante già evidenziata nel fratello II.2, risultato essere positivo.

Per i risultati ottenuti, la madre ha optato per una interruzione di gravidanza alla 21^a settimana di gestazione.

All'analisi autoptica è stato osservata e confermata una mancanza del corpo calloso ed una ipoplasia del verme, tali risultati sono in linea con la letteratura³¹ e tale caso potrebbe essere il secondo caso di feto descritto in letteratura, ma il primo con una caratterizzazione dei difetti anatomici strutturali evidenziati attraverso una colorazione di Ematossilina/Eosina dell'encefalo che ha messo in evidenza un evidente difetto importante di organizzazione nella corteccia (Fig.18).

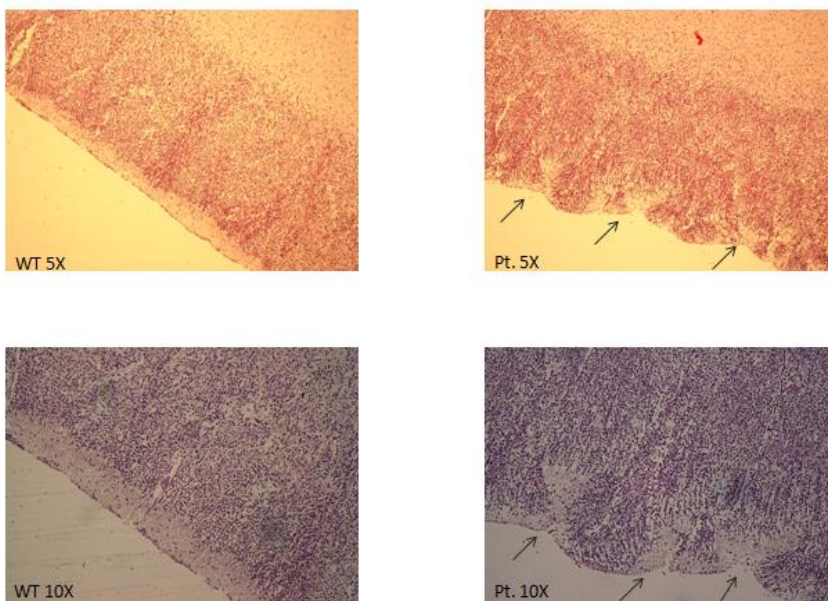


Fig18. Colorazione con Ematossilina/eosina di corteccia encefalica di un controllo e paziente II.3 di 21 settimane, si osservano difetti di corteccia e disorganizzazione; Acquisizione con ingrandimento a 5 e 10X.

5. CONCLUSIONI

Le tecnologie NGS hanno garantito alla comunità scientifica la possibilità di studiare una vasta gamma di malattie genetiche, garantendo un lavoro uniforme in ambito della ricerca clinica, con conseguente crollo dei costi e tempi di analisi, rispetto alle tecniche molecolari tradizionali.

Le tecnologie informatiche hanno giocato un ruolo chiave nel passaggio della semplice identificazione, all'interpretazione del ruolo patogenetico delle varianti genomiche determinando un miglioramento del lavoro del ricercatore nella elaborazione dei dati genetici.

La tecnologia WES è stata utile negli ultimi anni per i casi non diagnosticati clinicamente, ma che presentavano un sospetto diagnostico di malattia.

Nel breve futuro, il sequenziamento "whole exome sequencing" di malattie rare del neurosviluppo, negative ad analisi genetiche tradizionali, sostituirà i pannelli genetici custom per specifiche patologie, in quanto permette di ottenere sia informazioni su geni noti che geni non associati a malattia che potrebbero spiegare il fenotipo.

Per tale motivo, lo studio ha evidenziato le potenzialità delle nuove tecnologie genomiche nel chiarificare le basi genetiche di patologie neurologiche rare, difficilmente diagnosticabili su base clinica. Tale approccio genome-wide, inoltre, consente di mettere in luce nuove correlazioni genotipo-fenotipo e di espandere lo spettro clinico di specifiche condizioni genetiche.

Lo studio mette in luce anche l'identificazione dei portatori per un più efficace counselling genetico per le famiglie a rischio e per la diagnosi prenatale.

L'identificazione dei meccanismi patogenetici alla base delle patologie neurodegenerative è essenziale per lo sviluppo di terapie più efficaci per la futura medicina personalizzata.

6. BIBLIOGRAFIA

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nuskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferreira S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell

T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. *Science*. 2001 Feb 16;291(5507):1304-51.

2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Showkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglu S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI,

- Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001 Feb 15;409(6822):860-921
3. 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
 4. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next generation sequencing. *Nat Methods*. 2010 Feb;7(2):111-8
 5. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977 Dec;74(12):5463-7
 6. Tucker T, Marra M, Friedman JM. Massively parallel sequencing: the next big thing in genetic medicine. *Am J Hum Genet*. 2009 Aug;85(2):142-54
 7. Seelow D, Schuelke M, Hildebrandt F, Nürnberg P. HomozygosityMapper—an interactive approach to homozygosity mapping. *Nucleic Acids Res*. 2009 Jul;37:W593-9
 8. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754-60
 9. Whiteford N, Haslam N, Weber G, Prugel-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
 10. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 2008 Nov;18(11):1851-8
 11. Liu ZQ, Mahmood T, Yang PC. Western blot: technique, theory and trouble shooting. *N Am J Med Sci*. 2014 Mar;6(3):160

12. Lupas, A., Van Dyke, M., and Stock, J. Predicting Coiled Coils from Protein Sequences, *Science*, 1991, 252:1162-1164
13. Lupas, A., Prediction and Analysis of Coiled-Coil Structures *Meth. Enzymology*, 1996, 266:513-525
14. Rudnik-Schoneborn S, Barth PG, Zerres K. Pontocerebellar hypoplasia. *Am J Med Genet C Semin Med Genet* 2014; 166C:173–183
15. Akizu N, Cantagrel V, Schroth J, et al. AMPD2 regulates GTP synthesis and is mutated in a potentially treatable neurodegenerative brainstem disorder. *Cell* 2013;154:505–517
16. Marsh AP, Lukic V, Pope K, et al. Complete callosal agenesis, pontocerebellar hypoplasia, and axonal neuropathy due to AMPD2 loss. *Neurol Genet* 2015;1:e16
17. Marsh AP, Yap P, Tan T, et al. A novel AMPD2 mutation outside the AMP deaminase domain causes pontocerebellar hypoplasia type 9. *Am J Med Genet A* 2017;173:820–823
18. Aravind L, Koonin EV. A novel family of predicted phosphoesterases includes *Drosophila* prune protein and bacterial RecJ exonuclease. *Trends Biochem Sci.* 1998 Jan;23(1):17-9
19. Karaca E, Harel T, Pehlivan D, et al. Genes that affect brain structure and function identified by rare variant analyses of Mendelian neurologic disease. *Neuron.* 2015 Nov 4;88(3):499-513
20. Zollo M, Ahmed M, Ferrucci V, et al. PRUNE is crucial for normal brain development and mutated in microcephaly with neurodevelopmental impairment. *Brain.* 2017 Feb; 140(4):940-952
21. Costain G, Shugar A, Krishnan P, et al. Homozygous mutation in PRUNE1 in an Oji-Cree male with a complex neurological phenotype. *Am J Med Genet A.* 2017 Mar;173(3):740-743

22. Rossor AM, Carr AS, Devine H et al : Peripheral neuropathy in complex inherited diseases: an approach to diagnosis. *J Neurol Neurosurg Psychiatry* 2017;88:846-863
23. Voermans NC, van Alfen N, Pillen S et al : Neuromuscular involvement in various types of Ehlers-Danlos syndrome. *Ann Neurol* 2009;65:687–697
24. Capuano A, Bucciotti F, Farwell KD et al : Diagnostic exome sequencing identifies a novel gene, EMILIN1, associated with autosomal-dominant hereditary connective tissue disease. *Human Mutation* 2016;37:84-97
25. Fariselli P, Molinini D, Casadio R, Krogh A : Prediction of structurally-determined coiled-coil domains with hidden Markov models. *Bioinform Res Dev* 2007;4414:292-302
26. Tijssen, M.R., Cvejic, A., Joshi, A., Hannah, R.L., Ferreira, R., Forrai, A., Bellissimo, D.C., Oram, S.H., Smethurst, P.A., Wilson, N.K., Wang, X., Ottersbach, K., Stemple, D.L., Green, A.R., Ouwehand, W.H., Göttgens, B., 2011. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev Cell.* 20(5):597-609.
27. Ben-Salem S, Al-Shamsi AM, John A, Ali BR, Al-Gazali L. A novel whole exon deletion in WWOX gene causes early epilepsy, intellectual disability and optic atrophy. *J Mol Neurosci.* 2015 May;56(1):17-23
28. Tabarki B, Hashem A, Shahwan S, Alkuraya FS, Gedela S, Zuccoli G. Severe CNS involvement in WWOX mutations: Description of five new cases. *Am J Med Genet A.* 2015 Dec;167A(12):3209-13
29. Johannsen J, Kortüm F, Rosenberger G, Bokelmann K, Schirmer MA, Denecke J, Santer R. A novel missense variant in the SDR domain of the WWOX gene leads to complete loss of WWOX protein with early-onset epileptic encephalopathy and severe developmental delay. *Neurogenetics.* 2018 Aug;19(3):151-156

30. Bednarek AK, Keck-Waggoner CL, Daniel RL, Laflin KJ, Bergsagel PL, Kiguchi K, Brenner AJ, Aldaz CM. WWOX, the FRA16D gene, behaves as a suppressor of tumor growth. *Cancer Res.* 2001 Nov 15;61(22):8068-73
31. Bednarek AK, Laflin KJ, Daniel RL, Liao Q, Hawkins KA, Aldaz CM. WWOX, a novel WW domain-containing protein mapping to human chromosome 16q23.3-24.1, a region frequently affected in breast cancer. *Cancer Res.* 2000 Apr 15;60(8):2140-5
32. Valduga M, Philippe C, Lambert L, Bach-Segura P, Schmitt E, Masutti JP, François B, Pinaud P, Vibert M, Jonveaux P. WWOX and severe autosomal recessive epileptic encephalopathy: first case in the prenatal period. *J Hum Genet.* 2015 May;60(5):267-71