



UNIVERSITÀ DEGLI STUDI DI GENOVA

DOTTORATO IN SCIENZE PEDIATRICHE

Curriculum di Malattie Muscolari, Neurodegenerative e Metaboliche dell'Età Evolutiva

**NEXT-GENERATION SEQUENCING NELLA DIAGNOSTICA
DELLE EPILESSIE IN ETA' PEDIATRICA**

Relatore: Prof. Carlo Minetti

Candidato: Dott.ssa Maria Stella Vari

Anno accademico 2016-2017

INDICE

Capitolo 1 Introduzione

Capitolo 2 La Genetica delle Encefalopatie Epilettiche

2.1 Eterogeneità genetica della Sindrome di Dravet

Capitolo 3 Diagnostica Molecolare

3.1 Tecniche convenzionali

3.2 Tecniche di sequenziamento

Capitolo 4 Descrizione del progetto di ricerca

4.1 Materiali e Metodi

Capitolo 5 Risultati

Capitolo 6 Discussione

Capitolo 7 Conclusioni

Bibliografia

Ringraziamenti

CAPITOLO 1 INTRODUZIONE

La parola epilessia deriva dal verbo greco *ἐπιληψία* (*epilambanein*) che significa “essere sopraffatti, essere colti di sorpresa”. È un condizione che si manifesta con crisi epilettiche ricorrenti e non provocate, è la più comune malattia neurologica dopo la cefalea, con una prevalenza di 5-10/1000 persone ed un’incidenza di 50-120/100.000 persone per anno (Hauser et al. 1993).

È una condizione clinica alquanto eterogenea che riconosce, secondo la più recente classificazione ILAE 2010 (Berg et al, 2010), forme genetiche dove l'epilessia è il risultato di un difetto genetico conosciuto o presunto, forme strutturali/metaboliche, riconducibili ad un danno cerebrale, e , “epilessie da causa sconosciuta”, riferito a tutti quei casi in cui l'eziologia rimane indeterminata. Anche dal punto di vista clinico si esprime in forme molto diverse tra loro, alcune compatibili con una qualità di vita pressoché normale; altre di maggiore gravità.

Anche per l’epilessia, così come per molte altre patologie, è stata dimostrata una considerevole componente genetica (Steinlein OK. 2002) e, a partire dalla metà del 20th secolo, in seguito al lavoro di Lennox, sono stati identificati i primi geni specifici con eredità mendeliana. In particolare, dal 1995 grazie all’impressionante progresso nella ricerca genetica, oltre che allo sviluppo di nuove strategie interdisciplinari che combinano biologia molecolare con le tecniche neurofisiologiche, sono state identificate almeno una dozzina di epilessie geneticamente determinate. Inoltre è stato chiarito il potenziale contributo dei fattori genetici tra i meccanismi multifattoriali responsabili delle epilessie, anche in assenza della mutazione di un singolo gene. Molti dei geni coinvolti codificano per i canali ionici e recettori.

A partire dai primi studi nel campo della genetica dell’epilessia, sono state quindi individuate varie possibili modalità di eredità (Claes et al. 2001, Weber & Lerche 2008, Baulac & Baulac 2010, Klassen et al. 2011, Mulley & Mefford 2011):

- singolo gene, monogenica o Mendeliana;
- poligenica;

- multifattoriale o eredità complessa (interazione tra geni e fattori ambientali);
- mitocondriale (materna, citoplasmatica);
- anomalie cromosomiche (identificate da studi citogenetici, FISH o CGH array)
- fenocopie (ambientale). Quest'ultimo termine viene impiegato per riferirsi agli individui in una famiglia con lo stesso fenotipo ma che non hanno la stessa mutazione.

I principali tipi di mutazioni genetiche che contribuiscono alle malattie sono le seguenti:

- monogeniche, mutazioni con grandi effetti sul fenotipo;
- varianti genetiche rare che sono presenti in meno dell'1% della popolazione, possono essere patogene ed hanno effetto sul fenotipo;
- varianti genetiche comuni o polimorfismi con piccolo effetto sul fenotipo (SNP) che interagiscono con altri geni e/o con fattori ambientali per produrre il fenotipo;
- variazione nel numero di copie come delezioni, inserzioni, amplificazioni che coinvolgono uno o più esoni, delezioni che coinvolgono un numero di geni, rilevate con MLPA e array CGH;
- siti di splicing alternativi che determinano la trascrizione di isoforme amminoacidiche diverse, talvolta patogenetiche;
- fattori epigenetici come metilazione del DNA, modificazione delle proteine istoniche.

In aggiunta ai fattori genetici, il fenotipo può essere influenzato da fattori ambientali, di sviluppo o maturazione, endogeni o ormonali.

Nonostante il progresso scientifico nel campo della genetica dell'epilessia, ad oggi la genetica complessa delle sindromi epilettiche comuni ed i fattori genetici che determinano la risposta individuale ai farmaci antiepilettici (farmacogenetica) restano in gran parte sconosciuti ed in circa il 40% dei casi di epilessia in età pediatrica, l'eziologia rimane indeterminata (Steinlein OK., 2002).

Nel presente elaborato, presenterò i dati relativi all'efficacia e affidabilità di un pannello diagnostico mediante targeted resequencing per encefalopatie epilettiche (EE) in età pediatrica.

CAPITOLO 2 LA GENETICA DELLE ENCEFALOPATIA EPILETTICHE

Le encefalopatie epilettiche, secondo la definizione di Dulac del 2001, sono un ampio gruppo di condizioni cliniche in cui le funzioni cognitive, sensoriali e/o motorie si deteriorano come conseguenza dell'attività epilettica, costituita da frequenti crisi e/o da importante attività parossistica cosiddetta intercritica (Dulac et al. 2001). Si tratta di un gruppo eziologicamente composito di situazioni in cui spesso anche la patologia sottostante è coinvolta nel determinare la compromissione neurologica. In tale prospettiva è stata formulata da Engel nel 2001, e successivamente rivista dallo stesso autore nel 2006, una diversa definizione di encefalopatia epilettica, intesa come una condizione in cui il processo epilettico in sé è ritenuto contribuire al disturbo della funzione cerebrale (Engel et al. 2001, 2006), rispetto al quale anche la patologia sottostante può svolgere un ruolo preminente (Dravet et al. 2009). Secondo la più recente definizione quindi, tale condizione può essere il prodotto della causa sottostante, il risultato del processo epilettico o una combinazione di entrambi (Berg et al, ILAE 2010).

Questo gruppo di patologie può interessare qualsiasi età, ma sono più comuni e gravi durante l'infanzia e la prima infanzia, in cui si può determinare un deterioramento cognitivo globale e profondo.

Numerose sono le sindromi epilettiche, definite e riportate nella classificazione ILAE (Engel 2001, 2006), che configurano un quadro di encefalopatia epilettica: la Sindrome di Ohtahara (OS), l'Encefalopatia Mioclonica precoce (EME), l' Epilessia Parziale Migrante dell'Infanzia (MPSI), la Sindrome di West, (WS) la Sindrome di Dravet (DS o Epilessia Mioclonica Severa dell' Infanzia SMEI), lo Stato Mioclonico in Encefalopatie non progressive, l' Epilessia Mioclono-Astatica (EMA), la Sindrome di Lennox-Gastaut (LGS) , la sindrome di Landau-Kleffner, l' Encefalopatia con Punte-Onde Continue in Sonno Lento (ESES). Vi è però una quota di soggetti che, pur presentando un'epilessia ad esordio in età infantile configurante un quadro clinico di encefalopatia epilettica, non può essere inquadrata in nessuna di queste forme. Tali condizioni, pur molto eterogenee dal punto di vista elettroclinico, sono tutte età correlate e possono presentare caratteristiche comuni con la possibilità di un'evoluzione da una forma all'altra.

Per quanto riguarda l'aspetto eziologico, numerose ed eterogenee sono le cause identificabili alla base di tali condizioni, sia acquisite che genetiche (anomalie strutturali, malattie metaboliche, mutazioni in

specifici geni) (Nabbout e Dulac, 2003; Wolf, Bast e Surtess 2005). Non sempre, però, risulta possibile identificare e riconoscere una causa o una malattia neurologica sottostanti ed in più di un terzo dei casi l'eziologia rimane indeterminata nonostante approfondite indagini (come ripetute indagini neuroradiologiche, test genetici, esami metabolici, biopsia muscolare o della pelle, puntura lombare, ecc) con elevati costi socio-economici ed emotivi.

A tale proposito si tenga presente che, considerando semplicemente la Sindrome di West, di gran lunga la più frequente encefalopatia epilettica dell'infanzia, in circa il 30% dei soggetti affetti non è possibile formulare una diagnosi eziologica (Primec et al. 2002; Roger et al. 2005).

Le forme monogeniche di sindromi epilettiche sono legate a diversi geni con diverse funzioni e sono sempre crescenti le segnalazioni del coinvolgimento di nuovi geni implicati nella patogenesi di disordini neurologici, ad insorgenza precoce, caratterizzati dall'associazione di ritardo mentale ed epilessia, spesso severa e tale da configurare un quadro di encefalopatia epilettica.

Esiste una classificazione genetica delle forme monogeniche di Encefalopatie Epilettiche ad insorgenze precoce (Early Infantile Epileptic Encephalopathy, EIEE) nel database OMIM (Online Mendelian Inheritance in Man) che ad oggi individua 19 forme ([EIEE1-19](#)) alcune con un rilevato impatto clinico ed altre estremamente rare per cui sono riportati solo casi sporadici (Figura 1)

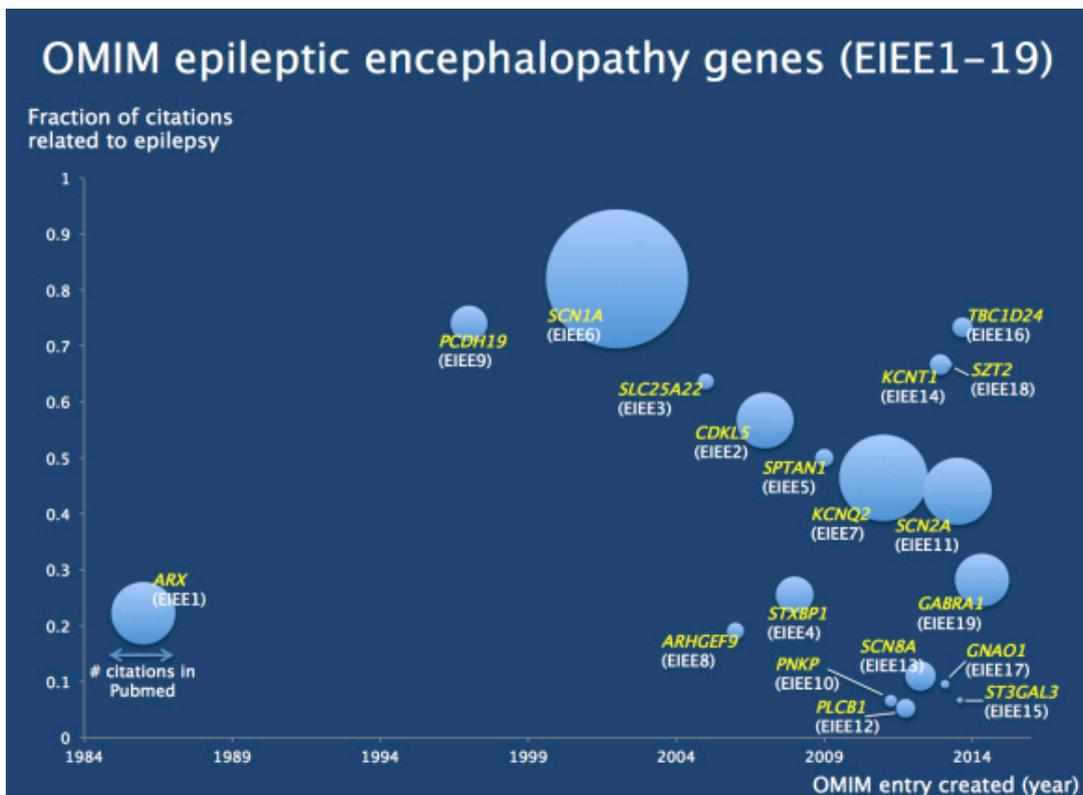


Figura 1 Sono rappresentati i geni delle EE per anno di entrata in OMIIM e per frazione di pubblicazione connesse all'epilessia in PubMed per un dato gene. Il diametro della bolla si riferisce al numero di pubblicazioni epilessia-correlate per un dato gene. Si noti la predominanza di SCN1A a conferma dell'importanza di questo gene nell'epilessia.

Il gene ARX (Aristaless-related homeobox, OMIM # 3008350 EIEE1) è localizzato a livello della regione Xp22.13. Mutazioni a carico del gene ARX sono state identificate come responsabili di numerosi e distinti fenotipi clinici neurologici (Suri, 2005), che comunemente vengono distinti in due grandi gruppi: sindromi con malformazioni, se sono presenti appunto malformazioni a carico del sistema nervoso centrale e/o dei genitali, e sindromi o disordini senza malformazioni (Kato et al. 2006). Nell'ambito del secondo gruppo, rientrano i seguenti fenotipi (riportati con il riferimento all'autore che per primo o meglio li ha definiti o correlati ad ARX), che interessano in genere solo soggetti di sesso maschile: la Sindrome degli Spasmi Infantili X-linked, ISSX (Stromme et al. 2002a, 2002b); epilessia mioclonica X-linked con spasticità e disabilità intellettiva (Scheffer et al. 2002); la Sindrome di Partington ovvero ritardo mentale X-linked, epilessia e distonia distale lieve (Stromme et al. 2002a, 2002b); ritardo mentale X-linked con o senza epilessia (possibile anche nelle femmine) (Bienvenu et al. 2002). Sempre di questo secondo gruppo fa parte il fenotipo descritto da Guerrini nel 2007 (Guerrini et

al. 2007) denominato epileptic-dyskinetic encephalopathy e caratterizzato da spasmi infantili, grave compromissione cognitiva e tetraparesi discinetica, nonostante la presenza di minute alterazioni di segnale alla RMN cerebrale suggestive di una perdita multifocale di tessuto, in assenza però di chiari aspetti malformativi. Recentemente, inoltre, lo spettro fenotipico dei disordini legati ad ARX si è ulteriormente allargato per la segnalazione di mutazioni di tale gene in soggetti affetti da Sindrome di Ohtahara (Kato et al. 2007, Holland e Hallinan 2010). Solo in una minoranza dei casi le mutazioni del gene ARX sono de novo (Bienvenu et al. 2002; Gronskov et al. 2004), anche se la presenza di mosaicismo germinale potrebbe rendere conto di alcuni di questi casi (Poirer et al. 2005). Nella maggior parte dei casi, infatti, si tratta invece sicuramente di mutazioni trasmesse, con una modalità di trasmissione X-linked recessiva: i maschi sono affetti, mentre in genere le femmine sono portatrici sane. Tutti i tipi di mutazioni sono stati descritti, missense, nonsense, mutazioni di splicing, e soprattutto duplicazioni e inserzioni che determinano espansioni dei tratti polialaninici. Queste ultime quando di dimensioni maggiori (che portano da 16 a 23 il numero di alanine in sequenza) sono state invece descritte associate alla Sindrome di West X-linked e alla più severa epileptic-dyskinetic encephalopathy descritta da Guerrini (Guerrini et al. 2007) e prima ancora da Wohlrab (Wohlrab et al. 2005). Ulteriori espansioni (che portano da 16 a 27 il numero di alanine in sequenza) sono state invece correlate alla Sindrome di Ohtahara (Kato et al. 2007), suggerendo l'esistenza di correlazione tra estensione del tratto polialaninico e gravità dell'encefalopatia.

Il gene CDKL5 (Cyclin-Dependent Kinase-like 5, OMIM #300203 EIEE2), detto anche STK9 (serine-threonine kinase 9), mappa a livello della regione Xp22.3 ed è stato inizialmente descritto in pazienti femmine con Sindrome di Rett atipica. La storia epilettologica di queste pazienti riconosce tre fasi: una fase precoce (1-10 settimane di vita) con frequenti crisi convulsive, riferite toniche generalizzate ed un EEG sostanzialmente normale, una seconda fase di encefalopatia epilettica con spasmi e crisi toniche ed un EEG ipsaritmico (dai 6 mesi ai 3 anni), ed una fase tardiva dominata dalla persistenza di crisi toniche e dalla comparsa di crisi miocloniche con un EEG caratterizzato da burst di anomalie epilettiformi amplissime ad andamento pseudo periodico (Bahi-Buisson et al 2008b). Come

succede nel contesto di ogni spettro fenotipico, accanto alla maggior parte dei soggetti riportati in letteratura che presentano un'epilessia severa ed una grave compromissione neurologica, vi sono soggetti mutati con un quadro neurologico, in particolare motorio, relativamente risparmiato (Weaving et al. 2004), e soggetti con un'epilessia meno grave (Psoni et al. 2009). La quasi totalità dei soggetti riportati è rappresentata da casi sporadici, ovvero da mutazioni de novo, che interessano prevalentemente soggetti di sesso femminile (trasmissione X-linked).

STXBP1 (Syntaxin-binding protein 1, OMIM: #602926 EIEE4) è localizzato sulla regione 9q34.11. Nel 2008 né è stata individuato il coinvolgimento (Saito et al.) in una paziente con S. di Ohtahara, successivamente confermato anche in altri casi con fenotipo simile. La Sindrome di Ohtahara è un'encefalopatia epilettica ad esordio precoce, in genere nei primi 10 giorni di vita, che si manifesta con diversi tipi di crisi convulsive (prevalentemente spasmi tonici e crisi parziali motorie). Gli aspetti elettroencefalografici sono peculiari e caratterizzano la sindrome per la presenza di un pattern tipico di scariche periodiche altovoltate diffuse che si alternano a tratti di tracciato piatto (suppression-burst). Il quadro clinico neurologico è rappresentato da una gravissima e precoce compromissione dello sviluppo psichico e motorio, associato spesso ad anomalie del tono muscolare e spasticità. Recentemente mutazioni in STXBP1 sono state riscontrate in altri fenotipi di epilessia (Hamdan et al. 2009; Carvill et al. 2014) e, sorprendentemente, anche in pazienti con deficit cognitivo senza epilessia (Hamdan et al. 2011)

SCN1A (Sodium channel, brain type I, alpha subunit OMIM: #182389 EIEE6) codifica per un canale del sodio espresso soprattutto a livello corticale (Nav1.1) ed è il più importante gene per le epilessie (Mulley et al. 2005; Harkin et al. 2007; Miller and Sotero de Menezes 2007; Escayg et al. 2010). Il fenotipo associato ad alterazioni di tale gene è molto ampio e varia da forme lievi di epilessia come l' Epilessia generalizzata con Convulsioni Febrili plus (Generalized epilepsy with febrile seizures plus' GEFS+) a forme con importante compromissione clinica come la Sindrome di Dravet. (DS o Epilessia Mioclonica Severa dell' infanzia SMEI). Recentemente, è stato descritto il coinvolgimento di tale gene, in altre entità cliniche come l' Eemicrania Emiplegica Familiare (FHM3), l' Epilessia parziale migrante

dell'infanzia (MPSI) (Carranza et al. 2011; Freilich et al. 2011), la Sindrome di West (Wallace et al. 2003), l' Epilessia mioclono-astatica (Wallace et al. 2001), la Sindrome di Lennox-Gastaut e forme di epilessie generalizzate o focali ad eziologia sconosciuta (Harkin et al. 2007).

In particolare, Harkin et al (2007) ha valutato l'incidenza di mutazioni in SCN1A in 188 pazienti con encefalopatia epilettica, ad esordio nei primi due anni di vita. Un maggiore percentuale di soggetti mutati è stata identificata in pazienti con DS (52/66; 79%) rispetto a pazienti con fenotipo Dravet like; inoltre mutazioni in SCN1A sono state identificate in soggetti con epilessia generalizzata criptogenetica (24%) e epilessia focale criptogenetica (22%). In particolare nelle MAE e le LGS è stata riscontrata una bassa percentuale di soggetti mutati (rispettivamente 2/10 e 1/12 casi positivi) ad sottolineare il possibile seppur raro coinvolgimento di tale gene in queste condizioni (Wallace et al. 2001; Nabbout et al. 2003b; Ebach et al. 2005). Fra i pazienti con epilessia focale è stato individuato un sottogruppo di soggetti con una forma devastante di encefalopatia epilettica multifocale dove in 3 soggetti su 5 è stata riscontrata una mutazione in SCN1A. L'autore ha inquadrato l' epilessia di questo sottogruppo di pazienti con esordio nel primo anno di vita con crisi polimorfe prevalentemente focali e importante declino cognitivo come Epilessia Multifocale Severa dell' Infanzia (Severe infantile multifocal epilepsy-SIMFE) (Harkin et al, 2007). Sulla base di questo elaborato, molti autori hanno successivamente studiato il coinvolgimento di SCN1A in altre encefalopatie epilettiche oltre che la DS (Miller and Sotero de Menezes, 2007; Zucca et al, 2008). Zucca et al. ha scrennato 60 pazienti con epilessia criptogenetica per mutazioni in SCN1A e ha identificato 12 soggetti mutati: 10 con DS, 1 con GEFS+ e 1 paziente con epilessia focale (Zucca et al, 2008). Dal confronto del fenotipo dei pazienti Dravet mutati rispetto a quello con GEFS+ ha individuato nei pazienti Dravet un esordio precoce delle crisi (3 mesi) con crisi generalizzate tonico-cloniche ed emilateralizzate sia in febbre che in apiressia, uno stato di male convulsivo in apiressia e la comparsa di segni neurologici prima dell'anno di vita con grave deficit cognitivo (Zucca et al, 2008).

Lo spettro di alterazioni riportate a carico di SCN1A è alquanto eterogeneo, vengono descritte mutazioni nonsense, mutazioni di splicing, e inserzioni in grado di alterare significativamente la funzionalità del canale e che sono state correlate a fenotipi più gravi come la DS o l' MPSI. Le

mutazioni missense di un singolo nucleotide determinano in genere la sostituzione di un singolo amminoacido e, nonostante determinino una minore compromissione funzionale del canale del sodio, sono responsabili di circa la metà dei casi di DS (Zuberi et al., 2011). Le mutazioni in SCN1A identificate in GEFS+ sono prevalentemente missense. (Zuberi et al., 2011). Anche riarrangiamenti genomici in SCN1A, tipicamente microdelezioni e microduplicazioni che spesso coinvolgono geni contigui, sono stati descritti in pazienti con DS e in altre forme di epilessia generalizzata.

Le mutazioni in SCN1A individuate in genere sono allo stato eterozigote, colpiscono quindi un solo allele e sono de novo (in circa il 95% dei casi dei pazienti con fenotipo Dravet), non riscontrate quindi in nessuno dei due genitori, a differenza di quanto avviene invece nei casi di GEFS+ dove molti casi mutati in SCN1A sono familiari. Recentemente è stato riscontrato in genitori asintomatici o paucisistomatici (con pregresse convulsioni febbrili) di pazienti con S. di Dravet un mosaicismo germinale o somatico per mutazioni in SCN1A (Depienne et al. 2006; Gennaro et al. 2006; Marini et al., 2006; Morimoto et al., 2006).

Le varianti di SCN1A individuate sono state raccolte in un database informatizzato consultabile dai clinici: SCN1A Infobase (<http://www.scn1a.info/>) e SCN1A Variant Database (<http://www.molgen.ua.ac.be/SCN1AMutations/Statistics/Mutations.cfm>) (Hirose et al. 2013) (Figura 2).

Table 1. SCN1A mutations and interpretations			
Type of variations	Common phenotypes	Uncommon/rare phenotypes	Explanatory examples for the description (Lossin, 2009)
Small scale variations			
Missense mutation	DS, GEFS+	MPSI, SIMFE, ICEGTC	c.3820T>A: p.Y1274N c.5075T>C: p.F1692S
Nonsense mutation	DS	SIMFE	c.1834 C>T: p.R612X c.3858 G>A: p.W1286X
Splice-site mutation	DS		c.265-1G>A c.1662+1G>T
Deletion mutation	DS	MPSI	c.429_430delGT: p.V143VfsX149 c.5296_5298delTTT: p.F1766del
Insertion mutation	DS		c.992_993insT: p.L331fsX340 c.1640_1641insA: p.K547fsX549
Large-scale variations			
Microdeletion	DS		
Microduplication	DS	SGE	
cf) Nonpathogenic small-scale variations			
Polymorphism	Nonpathogenic coding variant found in general population	c.1662G>A: p.Q554Q c.5771G>A: p.R1924H	

Figura 2 DS, Dravet syndrome; MPSI, migrating partial seizures of infancy; GEFS+, genetic epilepsy with febrile seizures plus; SIMFE, severe infantile multifocal epilepsy; ICEGTC, intractable childhood epilepsy with generalized tonic-clonic seizures, SGE, symptomatic generalized epilepsy.

PCDH19 (protocadherin-19 OMIM :#[300460](#) EIEE9) mappa a livello della regione Xp22.1 Mutazioni a carico della protocaderina 19 sono state inizialmente individuate in famiglie con epilessia e ritardo mentale limitata alle femmine (EFMR) (Dibbens et al 2008). Tale sindrome ha un'ampia variabilità fenotipica sia inter- che intrafamigliare, l'esordio dell'epilessia varia dai 6 ai 36 mesi con crisi sia febbrile che in apiressia, si associa un deficit cognitivo di grado variabile con spesso tratti autistici (Scheffer et al. 2008). Recentemente mutazioni a carico di tale gene sono state descritte in pazienti, in genere femmine, con fenotipo Dravet-like (Depienne et al. 2009). Le pazienti mutate in PCDH19 mostrano caratteristiche cliniche simili a quelli con mutazione in SCN1A come le convulsioni febbrili precoci, le crisi in cluster, la regressione dello sviluppo psicomotorio, il deficit cognitivo e le problematiche comportamentali; manca in genere il polimorfismo del fenotipo epilettico e le crisi, tipico invece delle Dravet classiche (Depienne et al- 2009).

Le encefalopatie correlate a mutazioni in PCDH19 interessano prevalentemente le femmine anche se è segnalato in letteratura il caso di un paziente maschio con fenotipo Dravet in cui è stata individuata una delezione emizigote sul cromosoma Xq22.1 che include il gene PCDH19 (Depienne et al. 2009). Lo spettro delle alterazioni a carico di tale gene è ampio, include mutazioni missense, mutazioni troncanti, piccole delezioni e inserzioni frameshift che determinano una perdita di funzione dell'allele mutato.

In genere le mutazioni sono de novo o ereditate da padri asintomatici. Tale modalità di trasmissione è differente rispetto ad altre malattie X-linked come ARX e CDKL5, infatti in questo caso solo le femmine eterozigoti esprimono il fenotipo mentre i maschi emizigoti per la mutazione sono asintomatici. Ciò può essere spiegato dal mosaicismo cellulare che si viene a creare nelle femmine mutate per la coesistenza di cellule PCDH19 positive e negative che determinerebbe un'alterazione dei meccanismi di interazione cellulare; al contrario i soggetti normali e i maschi mutati hanno un pattern cellulare omogeneo rispettivamente per cellule non mutate e mutate e non esprimono, quindi, il fenotipo. L'identificazione di un maschio affetto con identificazione nei fibroblasti di un mosaicismo in PCDH19 sostiene l'ipotesi dell'importanza di meccanismi di interazione cellulare come principale meccanismo patogenetico associato a mutazioni PCDH19.

Il gene [SCL25A22](#) (GC1 Glutamate carrier 1 OMIM #6093304 EIEE3) codifica per un trasportatore del glutammato, mutazioni recessive a carico di tale gene sono causa di una grave encefalopatia ad esordio precoce come la *Sindrome di Ohtahara*. È stato per la prima volta mappato sul cromosoma 11p15.5 in una famiglia araba con genitori consanguinei dove il probando affetto presentava crisi neonatali farmacoresistenti con suppression-burst ed importante ipotonia. È stata riscontrata una mutazione missense (p.Pro206Le) allo stato omozigote a carico del gene SCL25A22 (Molinari et al. 2005). Successivamente è stato identificato un altro paziente con mutazione recessiva di tale gene (Molinari et al 2009) e studi in vitro hanno confermato la patogenicità delle mutazioni rilevate (Molinari et al 2009).

[SPTAN1](#) (Alpha-II spectrin OMIM [#613477](#) EIEE5) mappa sul cromosoma 9q33. Mutazioni de novo a carico di tale gene sono state riscontrate in pazienti con encefalopatia epilettica ad esordio precoce, grave ipomielinizzazione e riduzione della sostanza bianca cerebrale. Nel 2008 sono stati identificati i primi 4 casi con quadro elettroclinico (in 3 pazienti su 4) suggestivo di S. di West (Tohyama et al. 2008, Saiatsu et al 2010).

[KCNQ2](#) (Potassium channel subunit Kv7.2 OMIM [#61372](#) EIEE7) mappa sul cromosoma 20q13.33 e sue mutazioni de novo sono state inizialmente descritte in pazienti con *Convulsioni Neonatali Familiari Benigne* (Benign Familial Neonatal Seizures, BFNS) (Biervert et al., 1998; Singh et al., 1998). Successivamente sono stati riportati casi sporadici e familiari con crisi neonatali ma outcome negativo associati a mutazioni in KCNQ2 (Borgatti et al. 2004; Steinlein et al. 2007). Recentemente mutazioni de novo in questo gene sono state riscontrate in pazienti con encefalopatia epilettica ad esordio precoce (S. di Othara e Encefalopatia non classificata) (Weckhuysen et al. 2012, Kato et al 2013) a confermarne l'associazione con forme di epilessia ad elevato impatto clinico.

[ARHGEF9](#) (guanine nucleotide exchange factor OMIM [#300607](#) EIEE8) mappa sul cromosoma Xq11.1-q11.2. Mutazioni o delezioni in tale gene possono essere responsabili di un fenotipo

caratterizzato da deficit cognitivo, epilessia ad esordio precoce e hyperekplexia (Harvey et al 2004, Marco et al 2008).

[PNKP \(OMIM #613402 EIEE10\)](#) mappa sul cromosoma 19q13.4. Mutazioni recessive in tale gene, che codifica per un enzima di riparazione del DNA, causano un'encefalopatia epilettica con microcefalia e spasticità progressiva agli arti inferiori (Shen et al 2010).

[SCN2A](#) (Sodium channel, brain type ii, alpha subunit OMIM#613721 EIEE11) mappa sul cromosoma [2q24.3](#). Mutazioni in SCN2A sono state inizialmente individuate in pazienti con BFNS e successivamente anche in casi di encefalopatia epilettica sia ad esordio precoce (S. di Ohtahara e S. di West) (Kamiya et al 2004, Ogiwara et al. 2009) sia in casi di S. di Lennox-Gastaut (Epi4K, 2013) Inoltre mutazioni de novo sono state identificate in soggetti con autismo e ritardo mentale ([Krumm et al 2014](#)) a sottolineare l'importanza di questo gene in meccanismo di sviluppo neurocognitivo.

[PLCB1](#) (Phospholipase C-beta 1 OMIM #613722 .EIEE12) mappa sul cromosoma [20p12.3](#) Delezioni omizigoti in questo gene sono state riscontrate in 2 casi con grave encefalopatia epilettica ad esordio precoce fra cui l' Epilessia parziale migrante dell'infanzia (Malignant Migrating Partial Seizures of Infancy [MMPSI](#)) (Kurian et al 2010, Poduri et al 2012).

[SCN8A](#) (Voltage-dependent sodium channel OMIM#614558 EIEE13) mappa sul cromosoma 12q13 e ed è espresso, diversamente da SCN1A, sulla membrana di neuroni eccitatori. Recente sono stati riportati 2 casi con epilessia ad esordio infantile e deficit cognitivo mutati in SCN8A (Veeramah et al 2012, Carvill et al 2013).

[KCNT1](#) (Sodium-activated potassium channel OMIM #614959 EIEE14) mappa sul cromosoma 9q34. Mutazioni de novo eterozigoti sono state individuate in casi con MMPSI (Barcia et al. 2012) e Epilessia Notturna del Lobo Frontale autosomica dominante (ENFL5) (Heron et al. 2012).

Il gene [ST3GAL3](#) (Beta-galactoside-alpha-2,3-sialyltransferase-III OMIM#615006 EIEE15) mappa sul cromosoma 1p34.1. Mutazioni in questo gene con trasmissione autosomica recessiva sono state individuate in una famiglia palestinese di consanguinei con 4 membri affetti da S. di West (Edvardson. et al 2013).

[TBC1D24](#) (Tre2-Bub2-Cdc16d24 OMIM#615338 EIEE16) mappa sul cromosoma 16p13; mutazioni recessive sono state identificate nel 2010 con next generation sequencing in 2 famiglie con grave encefalopatia epilettica con esordio dell' epilessia nella prima settimana di vita con crisi polimorfe farmacoresistenti. I pazienti affetti presentano una regressione dello sviluppo psicomotorio associato a segni extrapiramidali e ipotonia (Duru et al. 2010 and Milh et al. 2013).

[GNAO1](#) (Go, alpha subunit guanine nucleotide-binding protein, OMIM#615473 EIEE17) Mutazioni de novo eterozigoti sono state individuate in 4 casi con grave encefalopatia epilettica ad esordio precoce associata ad anomalie cerebrali quali atrofia o assottigliamento del corpo calloso (Nakamura et al. 2013).

[SZT2](#) (OMIM#615476 EIEE18) mappa sul cromosoma 1p34.2, Mutazioni recessive sono state descritte in due pazienti con grave encefalopatia epilettica e displasia del corpo calloso (Basel-Vanagaite et al. 2013).

[GABRA1](#) (Gamma-aminobutyric acid receptor, alpha-1 OMIM#615744 EIEE19) mappa sul cromosoma 5q34 e mutazioni in questo gene sono state precedentemente identificate in pazienti con Epilessia Mioclonica Giovanile ed Epilessia Assenza (Cossette et al. 2002, Maljevic et al 2006). Recentemente mutazioni in GABRA1 sono state identificate in pazienti con fenotipo Dravet like.

2.1 ETEROGENEITÀ GENETICA DELLA SINDROME DI DRAVET

La Sindrome di Dravet (DS o Epilessia Mioclonica Severa dell' Infanzia SMEI) ha un'incidenza di 1 ogni 20-40.000 nati vivi, esordisce nel primo anno di vita, con un picco a 5 mesi. L'esordio è

classicamente costituito da crisi in corso di febbre in bambini in precedenza sani, che vengono poste in diagnosi differenziale con le convulsioni febbrili. Caratteristiche distintive sono però l'esordio precoce, la lunga durata con tendenza a evolvere in stati di male febbrili, la presenza di elementi focali con lato variabile (crisi cloniche, emisomiche, e più raramente crisi apparentemente generalizzate) e la ricorrenza, spesso a temperatura corporea non elevata. In una seconda fase sopraggiungono altri tipi di crisi: mioclonie, assenze atipiche, stati di male non convulsivi e crisi parziali complesse, ad andamento pluriquotidiano e tendenza ad essere prolungate. Vi è poi una terza fase in cui si assiste ad un parziale miglioramento delle crisi (mioclonie e assenze possono scomparire, mentre persistono le crisi focali) ed il deterioramento cognitivo e motorio proprio della seconda fase tende a stabilizzarsi. Le mioclonie, pur essendo ritenute inizialmente indispensabili per la diagnosi di malattia, in realtà sono assenti in circa 1/5 dei casi e talora possono precedere la fase delle crisi febbrili.

L'EEG intercritico è usualmente normale all'esordio. In seguito tra il primo e quinto anno di età, durante la seconda fase, vi può essere un rallentamento dell'attività con comparsa di attività parossistica intercritica in genere scarsa (complessi punta-onda e polipunta onda generalizzati, anomalie focali e multifocali) e riscontro di fotosensibilità. Col tempo si assiste ad un progressivo rallentamento dell'attività di fondo (non costante), diminuzione delle scariche di complessi punta e polipunta onda, persistenza delle anomalie focali e multifocali, e fluttuazione della fotosensibilità.

L'EEG critico durante le mioclonie è caratterizzato da scariche più o meno generalizzate di punte e polipunte onda, mentre variabile è l'attività critica durante le crisi focali.

La prognosi è generalmente sfavorevole, con persistenza di crisi farmaco resistenti, scatenate principalmente da rialzi della temperatura corporea. Vi è deterioramento cognitivo e motorio, con sviluppo di ritardo mentale, stereotipie, atassia e segni piramidali. La mortalità complessiva è del 16%.

La terapia più efficace è costituita dall'associazione iniziale di acido valproico e clobazam, cui possono essere associati topiramato e stiripentolo. Tale terapia può consentire di ridurre la frequenza ed intensità delle crisi, anche se è in discussione se ciò influisca in modo determinante nel contenere il deterioramento cognitivo e la compromissione motoria.

Sono ad oggi molti gli studi che hanno confermato l'alta percentuale di mutazioni in SCN1A in pazienti con Sindrome di Dravet (Sugawara et al. 2002, Ohmori et al. 2002, Nabbout et al. 2003, Claes et al. 2003, Marini et al. 2007, Depienne et al. 2009). Sono state riscontrate tutti i tipi di mutazioni, sia troncanti che missense, in genere *de novo*, con trasmissione autosomica dominante. Le mutazioni troncanti (nonsense, di splicing, piccole delezioni e inserzioni) sono quelle più frequentemente identificate (50%) e in genere coinvolgono più domini del gene. (Mulley et al., 2005). Recentemente sono state individuate delezioni intrageniche o di tutto il gene in alcuni casi in assenza di altre mutazioni (troncanti, missense o di splicing) (Madia et al, 2006; Mulley et al, 2006; Suls et al, 2006.) (Figura 3).

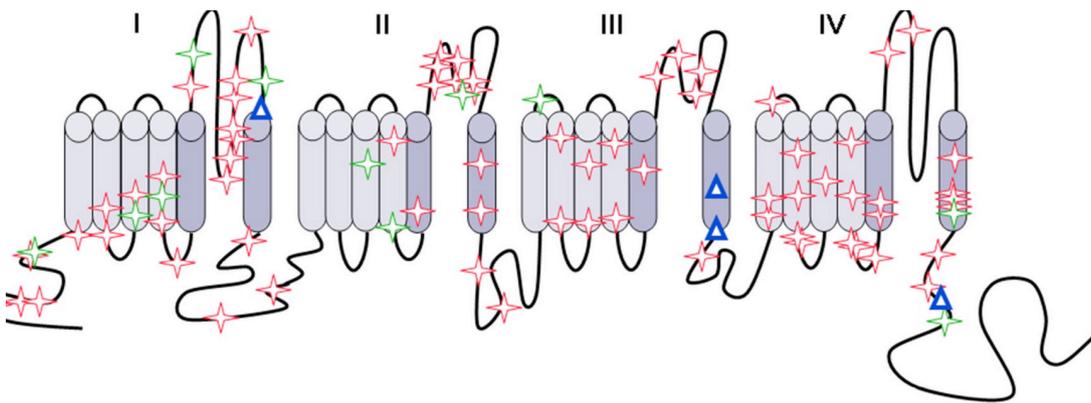
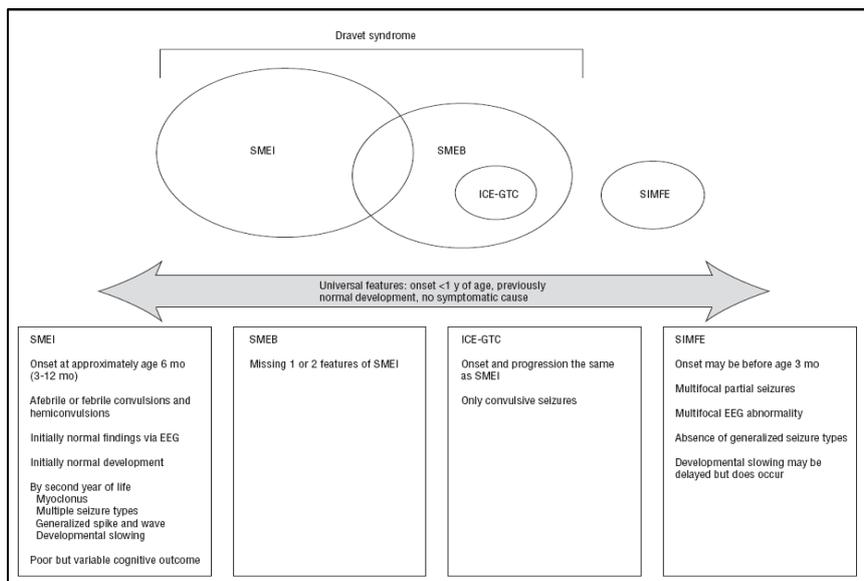


Figura 3 Rappresentazione schematica delle mutazioni missense e delle delezioni intrageniche nella proteina Nav1.1. Ogni stellina rappresenta una mutazione missense: quelle rosse sono quelle identificate in un singolo paziente e quelle verdi indicano le mutazioni ricorrenti. I triangoli blu rappresentano le delezioni in frame (Depienne et al. 2009).

I primi studi hanno correlato le mutazioni troncanti in SCN1A esclusivamente alla S. di Dravet (Claes et al, 2001; Sugawara et al, 2002). Successivamente, tali mutazioni sono state descritte anche in pazienti con fenotipo più lieve (Harkin et al. 2007; Yu et al. 2010) o in rari casi familiari con genitore asintomatico o lievemente sintomatici (Gennaro et al. 2003). Questi risultati sottolineano la difficoltà di prevedere il fenotipo sulla base del tipo di alterazioni in SCN1A individuata suggerendo l'importanza di altri fattori (background genetico di base e fattori ambientali) nella determinazione del fenotipo. (Yu et al, 2010). Data l'identificazione nell'ultimo decennio di casi mutati con solo alcune delle caratteristiche della DS, lo spettro fenotipico dei pazienti mutati in SCN1A è stato esteso identificando una nuova entità sindromica denominata SMEI di confine o SMEB (Ohmori et al. 2003, Fukuma et al. 2004, Dravet et al. 2005) (Figura 4)

Figura 4 Spettro fenotipico dei pazienti mutati in SCN1A



Questi pazienti mostrano percentuali di mutazioni in SCN1A simili a quelli dei pazienti con fenotipo completo e possono presentare come quest'ultimi tutti i tipi di mutazioni (Ohmori et al. 2003, Fukuma et al. 2004, Miller e Sotero de Menezes 2007). Recentemente mutazioni in SCN1A sono state riscontrate in pazienti con epilessia dell'infanzia con crisi generalizzate tonico cloniche intrattabile (ICEGTC) dove sono presenti crisi generalizzate tonico-cloniche a partire dal primo anno di vita senza l'evoluzione verso altri tipi di crisi e progressivo deterioramento neuro cognitivo simile a quello riscontrato nella DS (Fujiwara et al. 2003). Dal momento che tali entità sindromiche, SMEI classica, SMEI di confine e ICEGTCS, sono simili per molti aspetti (insorgenza di convulsioni durante il primo anno in bambini precedentemente normali, convulsioni prolungate, sensibilità febbre, povero di sequestro e di risultato mentale) e la frequenza di rilevazione di mutazione SCN1A è lo stesso (Miller e Sotero de Menezes 2007) molti autori, infine, hanno proposto l'uso del solo eponimo "Sindrome di Dravet" (DS), per definire l'intero gruppo di pazienti con diversa compromissione clinica ma uguale esordio clinico e decorso (Dravet et al. 2005, Miller e Sotero 2007). Nella più recente classificazione ILAE, rimane solo il termine "Sindrome di Dravet" in quanto la parola "epilessia mioclonica" può essere confondente dato che, come abbiamo visto, si riferisce ad un aspetto che può mancare (Arzimanoglou, 2009). Definiamo DS quella condizione clinica dove sono presenti:

- normale sviluppo psicomotorio sino ad un anno di vita
- esordio dell' epilessia nel primo anno di vita con crisi emicloniche, focali o generalizzate, in genere associate alla febbre
- rapido sviluppo di un epilessia farmacoresistente
- negatività del quadro neuroradiologico e degli accertamenti metabolici
- evidenza di ritardo dello sviluppo psicomotorio prima dei 2 anni di vita
- possibile sviluppo di crisi polimorfiche (miocloniche, assenze, atoniche, focali) ma incostante.

La proporzione dei pazienti con DS mutati in SCN1A è altamente variabile nei diversi studi con un range che va dal 33% (Nabbout et al. 2003, Wallace et al. 2003) all' 80-100%(Claes et al 2003, Fujiwara et al 2006, Fukuma et al. 2004) e ciò è spiegabile in parte dalla diversa dimensione del campione, dalla metodologia usata e dall'uso di criteri clinici più o meno selettivi. Considerando infatti solo i pazienti con fenotipo Dravet classico, in circa il 20% dei casi non sono presenti mutazioni né riarrangiamenti in SCN1A .

Al fine di identificare nuovi geni responsabili della DS, in un lavoro del 2009, Depienne et al. 2009 hanno screenato con microarray ad alta densità 41 casi con DS negativi per SCN1A individuando casi con mutazioni in PCDH19, dato che è stato successivamente confermato anche da altri studi (Depienne et al. 2012).

I pazienti mutati in PCDH19 presentano caratteristiche cliniche simili ai casi mutati in SCN1A come l' insorgenza precoce delle crisi sia in febbre che in apiressia, le crisi in clusters, il ritardo dello sviluppo psicomotorio e del linguaggio, le problematiche comportamentali e la regressione neurocognitiva. Esistono però alcune ma rilevanti differenze fra questi due gruppi, maggiormente per ciò che concerne l'evoluzione dell'epilessia in particolare, i pazienti mutati in PCDH19 sviluppano crisi febbrile alquanto polimorfiche (crisi miocloniche e assenze atipiche) spesso prolungate sino a configurare uno stato di male.

La letteratura internazionale ha ormai dimostrato come PCDH19 giochi un ruolo fondamentale nella patogenesi di encefalopatie epilettiche ad esordio infantile con uno spettro fenotipo simile a quello della DS (Depienne et al. 2009).

Inoltre, a conferma dell'eterogeneità genetica di tale entità sindromica, sono stati riportati casi con fenotipo Dravet-like e familiarità per GEFS+ mutati in GABRG2 (Harkin et al 2002, Hirose et al. 2006). Inoltre grazie allo sviluppo delle nuove tecniche di next generation sequencing, sono state recentemente individuate in pazienti con DS alterazioni in altri geni come SCN1B (2 casi con mutazione omozigote) (Patino et al. 2009, Ogiwara et al. 2012) e CHD2 (3 pazienti con mutazione de novo) (Suls et al. 2013). In particolare i pazienti mutati in CHD2 presentano alcune caratteristiche comuni ai pazienti con mutazione in SCN1A, come le convulsioni generalizzate scatenate dalla febbre, ma è descritto un esordio più tardivo dell' epilessia (dopo l'anno) e la presenza di ritardo dello sviluppo psicomotorio precedente l'esordio delle crisi.

A rendere ancora più complessa la genetica della DS, gli ultimi lavori scientifici (Carvill et al. 2014, Nava et al. 2014) mostrano un possibile coinvolgimento di nuovi geni (GABRA1, STXBP1 e HCN1) nella patogenesi di questa malattia che verranno discussi nei prossimi capitoli.

3 DIAGNOSTICA MOLECOLARE

3.1 TECNICHE CONVENZIONALI

Citogenetica

L'analisi cromosomica rappresenta la tecnica più consolidata per determinare i difetti genetici dei pazienti con disabilità intellettiva e l'epilessia. Tuttavia, le sindromi causate da aberrazioni cromosomiche, dove l'epilessia è una caratteristica costante, significativa e ricorrente della stessa patologia, che possono essere facilmente determinate con l'analisi del cariotipo convenzionale sono un numero esiguo.

A questa categoria appartengono la trisomia 21, il cromosoma 20 ad anello, l'inversione/duplicazione del cromosoma 15, il cromosoma 14 ad anello.

Rauch et al. 2006 ha riportato su una coorte di 600 pazienti affetti da ritardo mentale con o senza epilessia i dati relativi all'analisi del cariotipo convenzionale: anomali

e del numero di cromosomi in 68 casi (11,3%), riarrangiamenti/traslocazioni sbilanciate in 19 casi

(3,2%) e 3 casi (0,5%) con traslocazioni de novo apparentemente bilanciate.

Con le tecniche di citogenetica convenzionali, quindi, si possono identificare sino al 15% dei casi affetti da ritardo mentale ed epilessia ma tale dato non è invece valido per la diagnostica delle sindromi epilettiche. Inoltre, molte delle aberrazioni cromosomiche possono essere diagnosticata fenotipicamente e quasi tutti le aberrazioni cromosomiche sono anche rilevabile con tecniche di array.CGH.

In un recente trattato, Miller et al. 2012 hanno suggerito l' array- CGH, piuttosto che l'analisi cromosomica tradizionale, come test diagnostico di primo livello per le persone affette da ritardo mentale.

Diagnostica Array

L'ibridazione genomica comparativa su microarray (Array -Comparative GenomicHybridization o Array-CGH) è una tecnica sviluppata per identificare anomalie cromosomiche di tipo numerico (aneuploidie) a carico dei 22 autosomi (cromosomi dal nr. 1 al nr. 22) e dei cromosomi sessuali (X e Y), o anche variazioni (Variazioni del numero di copie "CNV") del contenuto di piccole porzioni cromosomiche, come duplicazioni/amplificazioni (presenza di copie in eccesso di segmenti di DNA), o delezioni (perdite di porzioni di genoma). Il principio su cui si basa la tecnica dell'Array CGH è la comparazione quantitativa del DNA in esame o DNA test (estratto dalle cellule fetali, in caso di diagnosi prenatale, o dal prelievo ematico del paziente, in caso di diagnosi post-natale) e del DNA genomico di riferimento proveniente da un soggetto sano (reference DNA). Durante il processo analitico questi DNA sono marcati in maniera differenziale con molecole fluorescenti (generalmente si utilizza un fluorocromo rosso il DNA test ed un fluorocromo verde per il *reference DNA*) e, successivamente, vengono mescolati in parti uguali e fatti incubare (Ibridazione) su un microarray, costituito da un supporto di vetro la cui superficie è coperta di frammenti di DNA, noti come sonde o cloni. Ognuno di questi cloni rappresenta una specifica regione del genoma umano, fino a ricomprendere l'intero assetto cromosomico umano. Tanto più è elevato il numero di cloni maggiore è l'efficacia dell'array nell'identificazione delle variazioni del numero di copie corrispondenti a piccole porzione di ciascun cromosoma. Il potere risolutivo della piattaforma utilizzata può variare in funzione

della densità e della tipologia delle sonde utilizzate; attualmente per scopi diagnostici vengono impiegati array tra 1 Mb e 100 kb. Al termine della suddetta incubazione, sia il DNA in esame che quello di controllo si legheranno ai cloni presenti sull'array. Il risultato sarà l'emissione di due distinti segnali fluorescenti le cui intensità saranno misurate a seguito di lettura degli arrays mediante un apposito strumento (scanner). Sull'immagine ottenuta verrà poi effettuata l'analisi comparativa tra le intensità di fluorescenza emesse dai due DNA e la relativa elaborazione dei dati mediante un apposito software, al fine di evidenziare eventuali variazioni del numero di copie del DNA test.

La citogenetica tradizionale, pur utilissima nell'individuare un gran numero di anomalie cromosomiche, numeriche e strutturali, è necessariamente limitata nelle sue possibilità diagnostiche dal potere di risoluzione del microscopio, l'analisi del genoma basata su array ha infatti una risoluzione molto più elevata (100 volte). Ciò consente di evidenziare anomalie del DNA che normalmente non potrebbero essere rilevate, incrementando notevolmente le possibilità di raggiungere una diagnosi certa. Inoltre, la tecnica Array-CGH non solo è in grado di valutare la presenza di eventuali anomalie cromosomiche a livello dell'intero genoma in un unico esperimento, senza sapere in anticipo cosa cercare, ma è in grado anche di definire esattamente la regione genomica alterata e quindi anche i geni in essa contenuti, migliorando la comprensione delle relazioni esistenti tra variazioni del numero di copie e patologia.

Vari gruppi di studio hanno valutato l'uso delle tecniche array nella diagnostica di diverse sindromi epilettiche (tabella 1)

FENOTIPO	NUM SOGGETTI	CNVs	%	FONTE
Malformazioni cerebrali	169	38	22.5%	Kariminejad et al. 2011
Ritardo Mentale + Epilessia	265	45	17%	Striano et al. 2012
Epilessia idiomatica	517	46	8.9%	Mefford et al., 2010
Encefalopatie epilettiche	315	25	7.9%	Mefford et al., 2011
Epilessia Generalizzata Idiopatica	1234	22	1.8%	de Kovel et al. 2010

Tabella 1

Kariminejad et al. 2011 in una coorte di pazienti con malformazioni cerebrali, la maggior parte dei quali con associato ritardo mentale ed epilessia sintomatica, ha riscontrato una positività dell'array-CGH nel 22,5 % dei casi. Striano et al. 2012 ha rilevato tassi simili di Copy NumberVariation (CNVs) in pazienti con Epilessia associata a Ritardo Mentale e pazienti con Ritardo Mentale senza epilessia. Tra i soggetti con epilessia idiopatica, Mefford et al. 2010 ha rilevato CNV nel 8,9 % dei casi, tale dato è sovrapponibile a quello riscontrato in pazienti con encefalopatie epilettiche (Mefford et al., 2011). Anche nel gruppo affetto da epilessia generalizzata idiopatica è stata riscontrata in una minor percentuale di soggetti (1,8%) CNV, a suggerirne la minima implicazione eziopatogenetica (de Kovel et al. 2010).

Dall'analisi di tali dati sembra quindi che in pazienti affetti da epilessia associata ad altri disturbi come ritardo mentale o malformazioni cerebrali, l'array-CGH possa essere diagnostico in una percentuale sufficiente di casi. Al contrario quando l'epilessia è l'unico sintomo, l'impatto diagnostico di tale tecnica sembra essere inferiore.

3.2 TECNICHE DI SEQUENZIAMENTO

La scoperta dell'uso dei di-deossi nucleotidi applicati al metodo a terminazione di catena sviluppato da Sanger (Sanger, Nicklen et al. 1977) ha segnato una pietra miliare nella storia del sequenziamento del DNA. Questo concetto ha fornito una base per lo sviluppo di sistemi automatizzati di sequenziamento Sanger (Smith, Sanders et al 1986; Ansorge, Sproat et al 1987) che è stato il metodo di scelta per il sequenziamento del DNA per quasi 20 anni. In questi anni, la tecnologia è stata migliorata per l'utilizzo di frammenti di DNA più lunghi e per un più elevato livello di parallelismo. Attualmente, tale tecnologia, supportando un sequenziamento simultaneo di 1000 paia di basi (bp) per frammento di DNA in 96 capillari, ha raggiunto un livello limitato di parallelizzazione e gli approcci basati sul Sanger non sono in grado di analizzare il DNA ad alta capacità.

Il sequenziamento Sanger era la tecnologia di base del Progetto Genoma Umano, che è stato finanziato nel 1990 con l'obiettivo di determinare i tre miliardi di paia di basi che costituiscono il genoma umano. Ci sono voluti dieci anni per produrre una prima bozza di risultati (Lander, Linton et al 2001; Venter, Adams et al 2001) e ulteriori tre anni per completare il progetto (Roberts 2003).

Durante la fase finale del progetto e i primi anni successivi, sono stati avviati numerosi studi tra cui il Progetto Internazionale HapMap ed il prominente 1000 Genomes Project. Il primo progetto mira a sviluppare una mappa degli aplotipi del genoma umano, che descrive i modelli comuni di variazione di sequenza del DNA umano (International HapMap Project, 2006). Il secondo si è concentrato sul sequenziamento di genomi di almeno mille partecipanti anonimi, di gruppi etnici diversi per fornire una adeguata variabilità genetica umana (Genomes). Entrambi i progetti sono stati accompagnati dalla necessità di un sequenziamento rapido, completo, affidabile ed economico di interi genomi aventi dimensioni paragonabili ai mammiferi (Fredlake, Hert et al 2006; Service 2006), hanno portato allo sviluppo di nuove tecnologie di sequenziamento. Questi cosiddetti sequenziamenti di nuova generazione (Next Generation Sequencing, NGS) hanno permesso il sequenziamento ad una velocità senza precedenti, in combinazione con bassi costi. Di conseguenza, il numero di dati relativi al sequenziamento, archiviati in banche dati pubbliche, è aumentato significativamente nel corso degli ultimi anni ed è destinato a crescere ancora più velocemente. Approfittando delle macchine di nuova

generazione, il progetto 1000 Genomi ha completato il sequenziamento dell'intero genoma di 185 individui da quattro popolazioni, ed ha analizzato le sequenze esoniche di 697 individui provenienti da sette popolazioni in soli due anni (Consortium, 2010). Visto che le macchine NGS per ogni corsa generano milioni di sequenze da leggere, la difficoltà nel sequenziamento è spostata dalla generazione della sequenza alla gestione e analisi dei dati. Il volume di dati rappresenta oggi una sfida importante per l'archiviazione, il backup e l'analisi. Nuovi approcci algoritmici sono richiesti per superare gli inconvenienti di queste macchine di poter generare e quindi leggere brevi sequenze (di circa un centinaio di pb). Lo sviluppo di pipelines snelle, altamente automatizzate per l'analisi dei dati è fondamentale per il passaggio dalle nuove tecnologie alla ricerca accelerata e conseguenti pubblicazioni (Koboldt, Ding et al. 2010). (Figura 5).

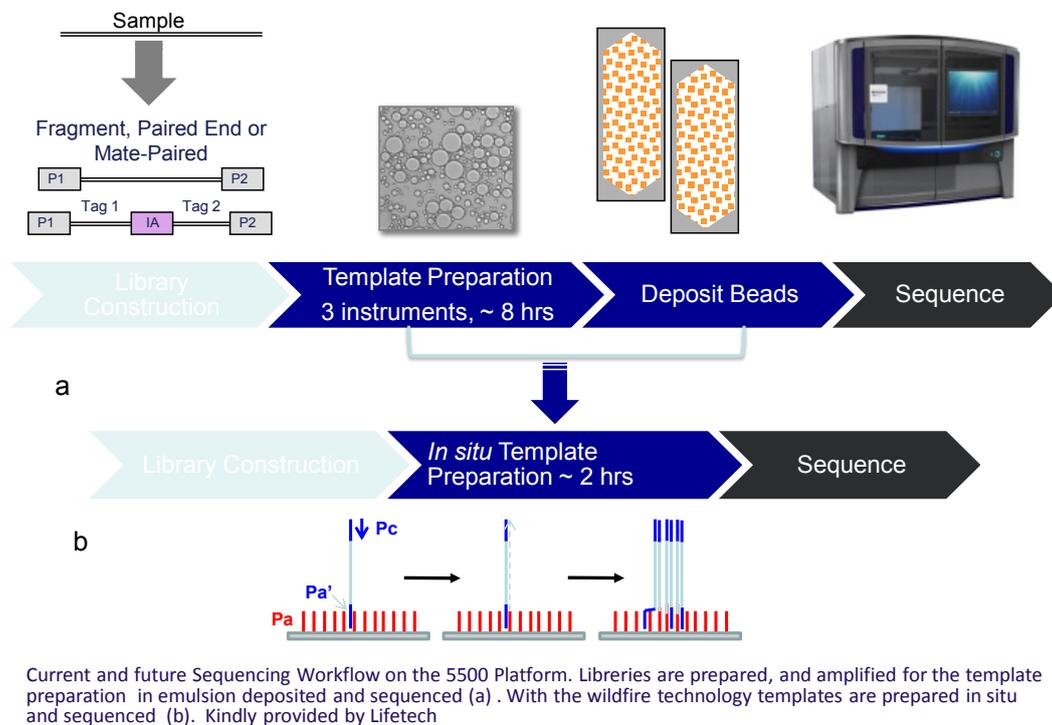


Figura 5. The WildFiretechnology semplifica radicalmente i tempi di preparazione del campione (da 8 a 2 ore circa) riducendo inoltre del 50% i costi.

Metodo Sanger

Il sequenziamento diretto Sanger rappresenta il gold standard delle diagnostica molecolare. Storicamente, nel 1975, il Premio Nobel F. Sanger ha introdotto il concetto di metodo di sequenziamento del DNA nella sua pionieristica lettura Croonian (Sanger 1975) ed in seguito, ha pubblicato un metodo rapido per la determinazione delle sequenze di DNA mediante sintesi innescata con DNA polimerasi (Sanger e Coulson 1975). Nel corso del 1977 sono stati pubblicati due articoli di riferimento per il sequenziamento del DNA, la tecnica di sequenziamento del DNA dideossi enzimatica di F. Sanger basata sugli analoghi terminali di catena del dideossinucleotide (Sanger, Nicklen et al. 1977) e la tecnica di degradazione chimica per il sequenziamento del DNA di A. Maxam e W. Gilbert in cui i frammenti di DNA terminali marcati vengono chimicamente scissi in specifiche basi e separati mediante elettroforesi su gel (Maxam e Gilbert 1977) (Figura 6).

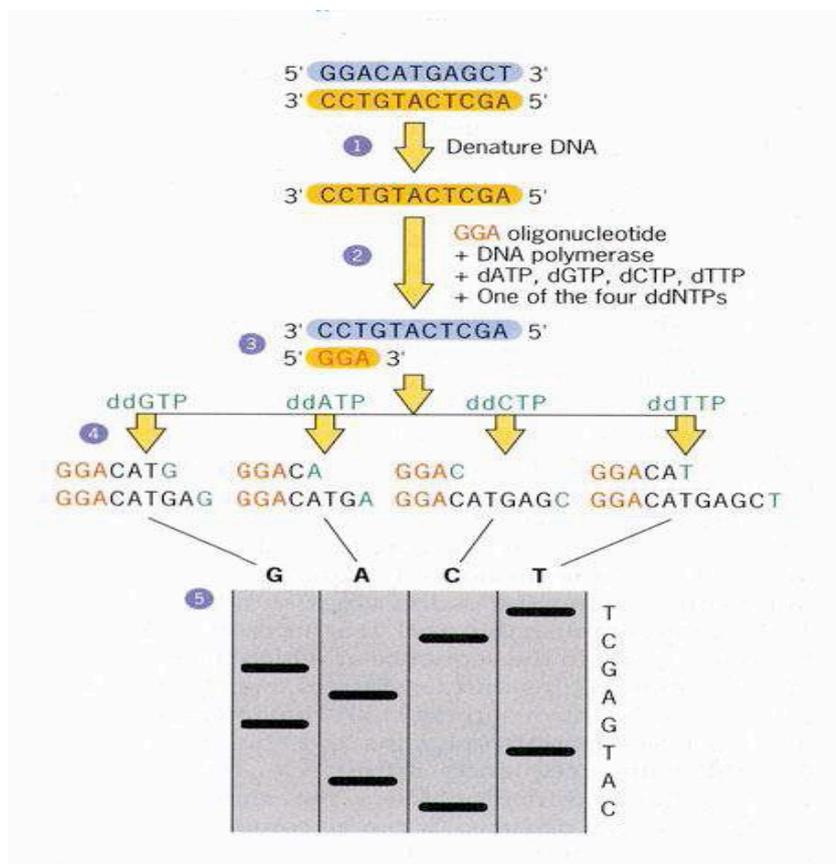


Figura 6 Schema del metodo Sanger per il sequenziamento del DNA.

Due importanti laboratori sono stati responsabili dell'introduzione dei primi sequenziatori automatici di DNA guidati dal California Institute of Technology o Caltech (Smith, Sanders et al. 1986), che è stato successivamente commercializzato da AppliedBiosystems (ABI), il Laboratorio europeo di Biologia Molecolare (EMBL) (Ansorge, Sproat et al. 1986; Ansorge, Sproat et al. 1987) e da Pharmacia Amersham, dopo General Electric (GE). La commercializzazione del metodo di sequenziamento ha condotto alla sua ampia diffusione in tutta la comunità di ricerca mondiale. Attraverso la prima apparecchiatura automatizzata per il sequenziamento di DNA a fluorescenza, è stato sequenziato un locus genico completo per il gene ipoxantina-guanina fosforibosiltransferasi (HPRT), utilizzando per la prima volta l'approccio di sequenziamento bidirezionale (Edwards, Voss et al. 1990). Nel 1996, ABI ha introdotto il primo sequenziatore di DNA commerciale che utilizzava una lastra di gel per elettroforesi dalla ABI Prism 310. Due anni più tardi, ABI Prism 3700 ha sostituito il notevole lavoro di versare il gel nelle lastre con ricarica automatica dei capillari con la matrice polimerica tramite 96 capillari. Questo sequenziatore di DNA automatizzato è stato impiegato con successo per il sequenziamento del primo genoma umano nel 2003, tenendo conto di 13 anni di sforzi del consorzio del progetto genoma umano, e con un costo stimato di 2,7 miliardi di dollari. Negli anni seguenti, un altro punto di riferimento è stato ottenuto dal sequenziamento del DNA del primo piccolo genoma del fago (5386 basi di lunghezza), seguito dal sequenziamento del genoma umano fino a 3 miliardi di basi (Lander, Linton et al. 2001; Venter, Adams et al. 2001). È notevole che tale progresso sia stato realizzato utilizzando metodi che sono perfezionamenti del metodo di base "dideoossi" introdotto da Sanger nel 1977.

Prima Generazione

Nel 2000, Jonathan Rothberg ha fondato 454 Life Sciences, che ha ulteriormente sviluppato la prima piattaforma NGS disponibile in commercio, la GS 20, chiamata J Appl Genetics. J Appl Genetica introdotto nel 2005, sviluppato da 454 Life Sciences (www.454.com), come il primo sistema NGS sul mercato. La tecnica sviluppata è stata implementata con successo combinando la PCR in emulsione di singole molecole con il pyro-sequenziamento (procedura di sequenziamento shotgun) di tutte le 580069

bp del genoma del *Mycoplasma genitalium* ottenendo il 96% di copertura e 99,96% di precisione in un'unica corsa GS 20 (Margulies et al. 2005). Negli anni successivi, Roche Applied

Science ha acquisito 454 Life Sciences ed esteso ulteriormente la nuova versione dello strumento 454, il GS FLX titanio. Condividendo lo stesso principio tecnologico sia in GS 20 che in GS FLX titanio, la flowcell è indicata come piastra "Pico-Titer", che è costituito da un fascio di fibre ottiche fuse. Su un fronte separato, è stata sviluppata la piattaforma Roche HT che si basa sulla PCR di singole molecole in microcompartimenti costituiti da emulsione olio-acqua (Tawfik e Griffiths 1998). In generale, il principio della tecnica pirosequenziamento si basa sul "sequenziamento di sintesi", ovvero sul rilevamento del rilascio di pirofosfato durante l'incorporazione del nuovo nucleotide, a differenza del metodo Sanger che funziona tramite la reazione di terminazione della catena con dideossinucleotidi. La tecnica è stata sviluppata dagli sforzi congiunti di un gruppo svedese (il team di M. Ronaghi, M. Uhlen, e P. Nyren) a Stoccolma (Ronaghi, Karamohamed et al. 1996). In primo luogo hanno messo a punto un approccio di sequenziamento basato sul rilevamento chemiluminescente del pirofosfato rilasciato durante l'incorporazione polimerasi-mediata del deossinucleosidtrifosfato (dNTP) (Nyren, Pettersson et al 1993; Nyren 2007) e che hanno poi sfruttato per l'identificazione delle singole basi costituenti la catena di DNA in tempo reale (Ronaghi, Uhlen et al. 1998). Nel pirosequenziamento la sintesi del DNA viene eseguita all'interno di una complessa reazione che comprende l'ATP sulfurilasi e l'enzima luciferasi ed i substrati adenosina 5' fosfosolfato e luciferina in modo tale che, il gruppo pirofosfato sia rilasciato dopo aggiunta di un nucleotide, con conseguente produzione di luce rilevabile.

Le tecniche HT-NGS, che sono nuove opportunità e hanno un grande impatto nell'ambito della genomica dei mammiferi, sono stati selezionati come il metodo dell'anno 2007 (Schuster 2008).

Il primo passo della tecnica HT-NGS consiste appunto nel rilevare la base successiva, marcata con fluorescenza (Reversible Terminator), aggiunta alla catena di DNA crescente mediante un sensitive charge-coupled device camera (CCDs camera). Questa tecnica è stata applicata all'analisi di un gran numero di campioni di DNA in parallelo, sia essa eseguita mediante l'uso di un supporto solido o su biglie, minimizzando così i volumi di reazione e lavorando in un microsistema miniaturizzato. Nella

fase successiva il Terminator viene convertito in un nucleotide standard e il fluorocromo viene rimosso. Questo ultimo ciclo e il processo precedente vengono ripetuti per determinare la base successiva nella sequenza.

Il principio di HT-NGS coinvolge le molecole di DNA, che sono sequenziate in modo parallelo in una flowcell (Mardis 2008; Metzker 2010). Il sequenziamento è condotto sia in un processo suddiviso in diversi passaggi o in maniera continua in tempo reale, in virtù del fatto che le reazioni avvengono in ogni templatato clonato o singola molecola, viene sequenziato “individualmente” e può essere quantificato all’interno di tutte le sequenze che vengono generate. L’elevata resa, combinata con la qualità e quantità delle sequenze generate ha portato l’analisi del genoma ad un livello che in precedenza non era possibile sia per la mancanza delle adeguate tecnologie sia per l’elevato costo con gli strumenti che erano disponibili.

Seconda Generazione

La seconda generazione di piattaforme HT-NGS è in grado di generare circa cinquecento milioni di basi di sequenze (Roche) fino a miliardi di basi in un’unica corsa (Illumina, SOLiD). Questi nuovi metodi si basano sull’analisi in parallelo e a cicli di sequenze, separate tra loro da ampliconi; il metodo Roche adotta una chimica basata sul pirosequenziamento, quello SOLiD si avvale invece di un sequenziamento basato sul legame sequenziale di sonde oligonucleotidiche e la tecnologia che utilizza terminatori fluorescenti reversibili. Attualmente, queste tre piattaforme leader HT-NGS di seconda generazione sono disponibili in commercio e la gara per più piattaforme aggiuntive è sempre all’orizzonte.

Tra le piattaforme di seconda generazione troviamo:

- Sistema di pirosequenziamento 454: è stata la prima piattaforma di sequenziamento nextgeneration disponibile come prodotto commerciale. Attualmente, il costo-base del sequenziamento con la piattaforma 454 è molto maggiore di quello delle altre piattaforme (SOLiD e Solexa), ma può essere il metodo di scelta per alcune applicazioni dove è necessario avere sequenze di lettura molto lunghe (assemblaggio de novo e metagenomica) (Mardis 2008).

- Illumina Genome Analyzer: comunemente indicato come “ Solexa”, questa piattaforma grazie al lavoro di Turcatti e colleghi, e la fusione di quattro aziende: Solexa (Essex, UK), Lynx Therapeutics (Hayward, CA, USA), ManteiaPredictive Medicine (Coinsins, Svizzera) e Illumina. Questa macchina è in grado di sequenziare librerie generate attraverso uno dei tanti metodi che prevede la generazione di frammenti singolo filamento in maniera random a cui vengono legati degli adattatori ottenendo frammenti di lunghezza fino a di diverse centinaia di coppie di basi (bp). Gli adattatori inseriti sono complementari alle sequenze presenti sulla flowcell di sequenziamento e, durante la cosiddetta fase di “Cluster Generation” avviene un’amplificazione a ponte delle sequenze target arricchite. (Mardis 2008)
- AB SOLiD: questa piattaforma è nata dal sistema descritto da J.S. e colleghi nel 2005 e nel lavoro di McKernan e colleghi Agencourt Personal Genomics (Beverly, MA, USA) (acquisita da AppliedBiosystems (Foster City, CA, USA) nel 2006). Il costo dello strumento è sostanzialmente inferiore a quello di altri strumenti di sequenziamento di seconda generazione. Inoltre, lo strumento è open source e programmabile, consentendo potenzialmente all’utente un’implementazione (ad esempio, l'uso di biochimica alternativa). Le attuali lunghezze lette, tuttavia, possono limitarne fortemente l’applicazione in ambito di ricerca. (Shendure e Ji 2008).

Illumina Genome Analyzer

Comunemente indicato come “ Solexa”, questa piattaforma grazie al lavoro di Turcatti e colleghi, e la fusione di quattro aziende: Solexa (Essex, UK), Lynx Therapeutics (Hayward, CA, USA), ManteiaPredictive Medicine (Coinsins, Svizzera) e Illumina. Questa macchina è in grado di sequenziare librerie generate attraverso uno dei tanti metodi che prevede la generazione di frammenti singolo filamento in maniera random a cui vengono legati degli adattatori ottenendo frammenti di lunghezza fino a di diverse centinaia di coppie di basi (bp). Gli adattatori inseriti sono complementari alle sequenze presenti sulla flowcell di sequenziamento e, durante la cosiddetta fase di “Cluster Generation” avviene un’amplificazione a ponte delle sequenze target arricchite. (Mardis 2008). (Figura 7)

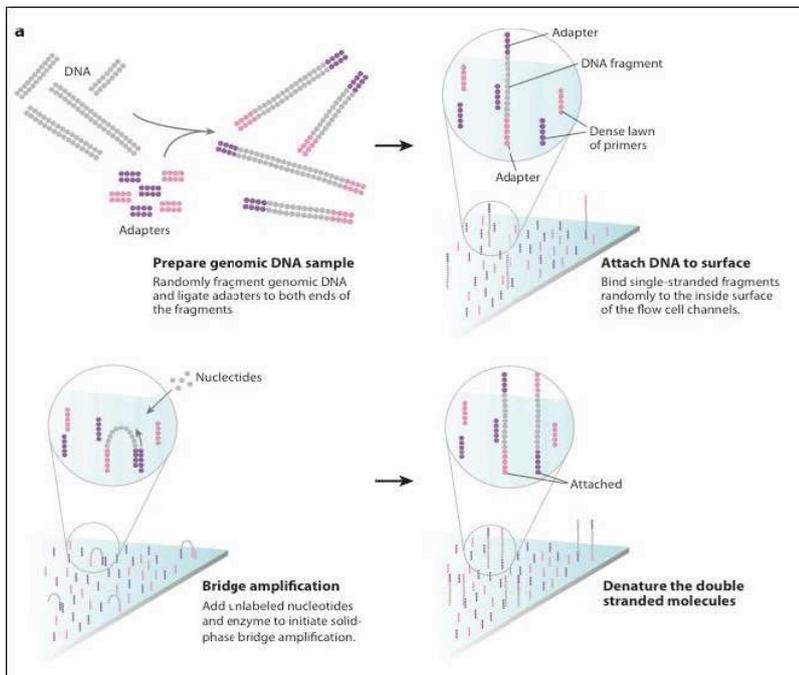


Figura 7: Illumina Genome Analyzer “Solexa” Workflow a.

In questo approccio, primers di PCR sono legati ad un substrato solido da un linker flessibile in modo tale che tutti gli ampliconi derivanti da ogni singola molecola durante l'amplificazione rimangono immobilizzati e raggruppati in un unico luogo fisico sul supporto stesso.

Sulla piattaforma Illumina, “l'amplificazione a ponte” avviene grazie all'alternanza tra cicli di estensione mediati da polimerasi e cicli di denaturazione con formamide. Ne risulta quindi la formazione di clusters, ognuno dei quali consiste in circa mille cloni amplificati. Diverse milioni di distinti clusters possono essere amplificati all'interno di una delle otto lanes della flowcell (otto librerie indipendenti possono essere sequenziate in parallelo durante la stessa corsa dello strumento).

Dopo la generazione dei clusters, gli ampliconi sono a singolo filamento (linearizzazione) e un primer di sequenziamento viene ibridato ad una sequenza universale (adattatori) fiancheggiante la regione di interesse. Ogni ciclo di sequenza consiste nell'estensione di una singola base con una DNA polimerasi modificata e una miscela di quattro nucleotidi. Questi nucleotidi sono modificati in due modi. Sono “terminators reversibili”, tali per cui una porzione rimuovibile chimicamente in posizione del 3'idrossile consente poi l'incorporazione di una singola base per ogni ciclo, e per ogni nucleotide abbiamo un fluororomo diverso corrispondente, anch'esso chimicamente removibile dalla sequenza in sintesi-sequenziamento. Dopo l'estensione alla singola base e l'acquisizione di immagini in quattro canali, la

scissione chimica di entrambi i gruppi prepara la sequenza al ciclo successivo. Leggere lunghezze fino a 36 bp è attualmente routine; letture più lunghe sono possibili, ma si potrebbe incorrere in un tasso di errore superiore (Mardis 2008).

La lunghezza delle sequenze lette sono limitate da molteplici fattori che causano il decadimento del segnale e il defasamento, come scissione incompleta delle etichette fluorescenti o frazioni di terminazione. Il tipo di errore dominante è la sostituzione, anziché inserzioni o delezioni (gli omopolimeri sono certamente un problema minore rispetto ad altre piattaforme). La media del tasso di errore è nell'ordine di 1-1,5%, ma le basi di maggiore precisione con un tasso di errore di 0.1% o meno possono essere identificate attraverso metriche di qualità associate a ciascuna base-call. Come con altri sistemi, modifiche hanno recentemente attivato letture appaiate (Shendure e Ji 2008).

Terza Generazione

Nelle precedenti piattaforme HT-NGS di seconda generazione, il principio è basato sulla PCR in emulsione per l'amplificazione di frammenti di DNA, per rendere il segnale luminoso abbastanza forte per il rilevamento affidabile della base con telecamere CCD. Sebbene l'amplificazione PCR ha rivoluzionato l'analisi del DNA, in alcuni casi può introdurre errori di sequenza di base o favorire certe sequenze rispetto ad altre, modificando così la frequenza relativa e l'abbondanza di vari frammenti di DNA esistenti prima dell'amplificazione. Per superare questo, l'ideale sarebbe poter generare la sequenza partendo direttamente da una singola molecola di DNA, senza la necessità di amplificazione tramite PCR e i bias conseguenti. Le tecnologie che offrono l'opportunità di poter generare sequenze da una singola molecola di DNA vengono definite "terza generazione di HTNGS" (Schadt, Turner et al. 2010). Il concetto di sequenziamento per sintesi senza uno step di amplificazione precedente e il sequenziamento di una singola molecola sono attualmente perseguiti da molte aziende ma non esiste ancora nessuna tecnologia in grado di offrire risultati affidabili, riproducibili e a costi sostenibili.

I DIVERSI APPROCCI AL SEQUENZIAMENTO

L'introduzione nella pratica dei laboratori delle varie tecniche di sequenziamento ha sicuramente alimentato la curiosità dei ricercatori verso una comprensione sempre più dettagliata del genoma umano; da qui la nascita per esempio del Progetto Genoma Umano (HGP). Le discussioni negli anni 1980 che hanno portato al HGP sono state motivate dal fatto che conoscere la sequenza del genoma umano sarebbe straordinariamente utile per comprendere la biologia umana e le malattie.

Nel 1988, un rapporto dell'US National Research Council (NRC) ha articolato un piano audace per uno sforzo che sarebbe culminato nel sequenziamento dell'intero genoma umano, il rapporto afferma che un tale "progetto aumenterebbe notevolmente la nostra comprensione della biologia umana e consentirebbe un rapido progresso nelle diagnosi e nel controllo di molte malattie umane". Nel quarto di secolo scorso, la lungimiranza di questa visione è stata confermata. Il progresso nella genomica è stato grandioso. Anche se le sfide restano impressionanti, gli obiettivi fondamentali non sono cambiati: la genomica e i relativi studi biologici su larga scala porteranno, in modi non disponibili in precedenza, ad una profonda conoscenza sulla biologia dei genomi e quindi delle malattie, a progressi inimmaginabili nel campo della scienza medica, oltre che della farmagogenomica ed a nuovi modi efficaci per migliorare la salute umana.

Con queste premesse è nato il Progetto Genoma Umano, un imponente progetto che aveva obiettivi complessi:

- identificare tutti i circa 20.000-25.000 geni nel DNA umano;
- determinare le sequenze dei 3 miliardi di paia di basi che compongono il DNA umano;
- memorizzazione di queste informazioni nei database;
- migliorare gli strumenti di analisi dei dati;
- trasferimento delle tecnologie relative al settore privato;
- affrontare le questioni etiche, legali e sociali (ELSI) che possono derivare dal progetto.

Naturalmente il "genoma" di ogni individuo è unico a se stesso, la mappatura del genoma umano, include quindi anche l'identificazione delle molteplici varianti che ogni individuo "sano" può

presentare. La mappatura in più individui della sequenza nucleotidica che costituisce il loro genoma ha dato la possibilità di creare un genoma aploide riferimento.

Il Progetto Genoma Umano è stato dichiarato compiuto nel 2003, tuttavia ci sono ancora una serie di regioni del genoma umano che possono essere considerate incomplete:

- le regioni centrali di ciascun cromosoma, note come centromeri, sono sequenze di DNA altamente ripetitive e sono difficili da sequenziare utilizzando la tecnologia attuale. I centromeri sono milioni di coppie di basi lunghe, e per la maggior parte sono assolutamente non sequenziati;
- le estremità dei cromosomi, chiamate telomeri, sono anch'esse altamente ripetitive. Per la maggior parte dei 46 cromosomi le regioni telomeriche sono incomplete. È noto quanto rimane da sequenziare ma il sequenziamento specifico di queste regioni, con la tecnologia attuale, non è possibile;
- ci sono diversi loci nel genoma di ogni individuo che contengono i membri delle famiglie multigeniche che sono difficili da districare con metodi di sequenziamento shotgun, queste famiglie multigeniche spesso codificano proteine per le funzioni immunitarie e sono incomplete;
- oltre a quanto sopra ci sono ancora alcune regioni chiamate "gaps" che contengono sequenze altamente ripetitive, e quindi l'identificazione specifica della corretta alternanza delle quattro basi azotate diviene complessa.

Considerando la sequenza finita non ridondante, più o meno $(2,86 \text{ Gb}) / (3,1 \text{ Gb}) = 92,3\%$ del genoma umano è stato sequenziato, coprendo le regioni codificanti, ovvero gli esoni, il Progetto Genoma Umano può essere considerato completo; a parte piccoli dettagli.

Recenti miglioramenti nella tecnologia di sequenziamento hanno portato alla cosiddetta nuova generazione di macchinari che hanno drasticamente ridotto il costo del sequenziamento e che ha portato al 1000 Genomes Project.

Ci sono altri importanti progetti di riferimento del genoma umano che alla fine hanno l'obiettivo di creare una mappa completa delle variazioni del genoma umano a disposizione della comunità scientifica mondiale attraverso banche dati pubbliche liberamente accessibili. Questo è molto importante, in quanto si tratta di una perfetta realizzazione della collaborazione umana per la conoscenza, senza chiedere un risarcimento. L'obiettivo del 1000 Genomes Project è trovare più varianti genetiche che

hanno frequenza di almeno l'1% nelle popolazioni studiate. Questo obiettivo può essere raggiunto mediante il sequenziamento di molti individui mediante la medesima tecnologia. Per sequenziare il genoma di una persona, molte copie del DNA vengono suddivise (in maniera random) in piccoli frammenti, ognuno dei quali viene sequenziato. Questo implica che le sequenze generate saranno più o meno casualmente distribuite in tutto il genoma-exoma, garantendo una lettura omogenea dello stesso. I frammenti vengono quindi allineati alla sequenza di riferimento e uniti tra loro in modo da ricostruire l'intero target. Per trovare la sequenza genomica completa di una persona con le piattaforme di sequenziamento vigenti si richiede il sequenziamento del DNA della stessa per l'equivalente di circa 28 volte (chiamato 28X). Se le letture della sequenza venissero fatte solo una media di una volta in tutto il target (1X), mancherebbe gran parte della sequenza, perché non ci sarebbe un equilibrio nel numero di frammenti (e quindi di letture) che coprono una determinata regione genomica rispetto a quelli che sono complementari ad altre regioni; un'adeguata copertura dell'intero target da sequenziare garantisce la corretta identificazione non solo di variazioni di singolo nucleotide (SNVs) ma in alcuni casi può anche consentire di rilevare varianti strutturali.

Il progetto prevede attualmente di sequenziare ciascun campione per 4X coperture; con questa profondità di sequenziamento non si può fornire il genotipo completo di ciascun campione, ma dovrebbe consentire il rilevamento della maggior parte delle varianti con frequenze attorno all'1%. Combinare i dati di 2500 campioni dovrebbe consentire in definitiva, di avere una stima altamente accurata relativa alle varianti geniche di ciascun soggetto. La creazione di una banca dati delle SNVs comuni degli esoni umani, come il dbSNP, il Progetto HapMap, il 1000 Genomes Project e altri dati di sequenziamento dell'esoma permette ai ricercatori di controllare e distinguere se le SNV identificate possono essere o meno patogenetiche. Questo si è dimostrato efficace sia per filtrare un numero consistente di varianti con bassa probabilità di essere causative, sia nel riconoscere quelle varianti potenzialmente causative delle malattie mendeliane.

Il goal finale resta quello del Whole Genome Sequencing (WGS), ovvero il sequenziamento completo del DNA (sia nucleare sia mitocondriale) del genoma di un organismo in un unico esperimento. (Bonfond, Durand et al. 2010).

Fin dai suoi primi giorni, la ricerca medica ha cercato di individuare le cause dei disturbi con l'obiettivo finale di istituire trattamenti terapeutici e trovare una cura. Oggi gli approcci (WGS) sono stati progettati per scoprire variazioni genetiche che contribuiscono alle malattie rare o comuni. Nonostante la diminuzione dei costi di sequenziamento, le spese di routine per ottenere e analizzare i genomi completi di un gran numero di individui rimangono proibitivi (Hedges, Burges et al. 2009).

I metodi alternativi, anche definiti “targhettati”, si concentrano solo su regioni specifiche del genoma umano, rappresentando approcci a prezzi accessibili per individuare eventuali varianti genetiche associate alla malattia. Il sequenziamento di tutte le regioni codificanti le proteine del genoma, indicato anche come exomesequencing, è il candidato promettente: esoma costituisce circa l'1% del genoma umano richiedendo la sequenza di sole 30 Mb circa (Ng, Turner et al. 2009), il whole exome sequencing (WES) è solo 1/20 rispetto al WGS (Ng, Bigham et al. 2010); i polimorfismi dei singoli nucleotidi (SNPs) che si verificano nelle regioni codificanti, sono le cause più comuni per le malattie mendeliane (Horner, Pavesi et al. 2010). Rare malattie monogeniche sono di interesse sostanziale perché l'identificazione delle loro basi genetiche fornisce importanti conoscenze sui meccanismi della malattia, pathway proteici e potenziali bersagli terapeutici. Tuttavia, ad oggi, sono state scoperte meno della metà delle varianti alleliche sottostanti tutte le malattie monogeniche. Ciò è perché l'identificazione delle variazioni del singolo nucleotide, (SNVs) per molte malattie rare è fondamentalmente limitata da fattori quali la disponibilità solo di un piccolo numero di individui affetti (casi) o famiglie, l'eterogeneità del locus, o sostanzialmente la ridotta capacità riproduttiva; ciascuno di questi fattori diminuisce il potere delle tradizionali strategie di clonazione posizionale e spesso limita l'analisi a geni candidati identificati a priori. Al contrario, l'approfondito risequenziamento di tutti i geni umani per la scoperta di varianti alleliche potrebbe identificare il gene alla base di una rara malattia monogenica. Le tecnologie di sequenziamento in parallelo del DNA hanno reso sempre più pratico il risequenziamento dell'intero genoma di individui umani, anche se il costo resta un fattore fondamentale.

Il sequenziamento dell' esoma, piuttosto che l'intero genoma umano, risulta quindi ben giustificato come una strategia efficace per la ricerca di mutazioni causative dei disordini mendeliani. In primo luogo gli studi di clonazione posizionale, che sono incentrati sulle sequenze proteina-codificanti, se

adeguatamente eseguiti, dimostrano grande successo nell'identificazione di varianti alla base di malattie monogeniche. In secondo luogo, la netta maggioranza delle varianti alleliche conosciute alla base dei disturbi mendeliani interviene sulle sequenze proteina-codificanti. I siti di splicing accettori e donatori rappresentano una classe aggiuntiva di sequenze che sono arricchite da varianti altamente funzionali e si mira pertanto anche a questi. In terzo luogo, una grande frazione di rare varianti non-sinonime del genoma umano si prevede che siano deleterie. Ciò contrasta con le sequenze non codificanti, dove le varianti con più probabilità non hanno effetto sul fenotipo, o solo debole effetto, anche se ben conservate. L'esoma rappresenta quindi un sottoinsieme altamente ricco del genoma in cui cercare le varianti con grande effetto sul fenotipo.

Lo studio di linkage, che è stato utilizzato attraverso l'ultimo mezzo secolo per comprendere e descrivere i disordini mendeliani, non è adatto per le estremamente rare malattie mendeliane a causa della difficoltà nella raccolta di un numero adeguato di individui affetti (di famiglie multigenerazionali) e famiglie per un studio statisticamente significativo. Questo approccio non è applicabile anche per i casi sporadici, per esempio la sindrome di Kabuki, una rarissima autosomica dominante mendeliana disturbo con un'incidenza stimata di 1 in 32.000, dove la maggior parte dei casi segnalati sono sporadici (Ng, Buckingham et al. 2010). Come risultato, la variante causativa ed il gene candidato per la sindrome di Kabuki sono rimasti sconosciuti fino a poco tempo fa. Un totale di 33 diverse varianti causative in MLL2 sono stati identificati dalla Ng et al. (2010) in 35 di 53 individui affetti da sindrome di Kabuki. Inoltre in 12 di questi individui, di cui erano disponibili i campioni dei genitori, le varianti in MLL2 sono state identificate come mutazioni de novo. Nello studio solo dieci di questi individui sono stati studiati utilizzando exome sequencing per identificare le varianti causali in MLL2, e gli esoni di questo gene sono stati poi screenati in ulteriori 43 casi mediante sequenziamento Sanger (Ng, Bigham et al. 2010). Analogamente, la maggior parte dei casi di sindrome di Schinzel-Giedion si sono verificati sporadicamente suggerendo che varianti de novo in eterozigosi possono aver causato il disturbo. Ciò è stato ora ulteriormente supportato dall'identificazione di varianti de novo causative in SETBP1 in quattro individui affetti da questo disturbo attraverso exome sequencing (Hoischen, van Bon et al. 2010). Queste varianti de novo causative non sarebbero state altrimenti identificate senza exome

esequencing. Al contrario, anche se nessuna delle varianti causali in DHODH sembra essersi verificato de novo per la sindrome di Miller, è ancora una malattia estremamente rara (Ng, Buckingham et al. 2010). Nell'insieme, questi studi hanno dimostrato i vantaggi di exome sequencing oltre lo studio di linkage in situazioni in cui un piccolo numero di campioni indipendenti o casi sporadici sono disponibili. Inoltre lo studio di linkage non è utile per i disturbi mendeliani con eterogeneità genetica (le varianti causative sono presenti in geni diversi) ed eterogeneità fenotipica (diverse manifestazioni cliniche o fenotipiche conducono all'incertezza nella diagnosi del disturbo). Allo stesso modo, questi problemi sono ben rappresentati nella sindrome di Kabuki, che è probabilmente una malattia geneticamente eterogenea in quanto non tutti gli individui colpiti hanno varianti causali nel gene singolo candidato (MLL2) (Ng, Bigham et al 2010; Paulussen, Stegmann et al 2010). Tuttavia, a sostenere ulteriormente la sua eterogeneità genetica, le varianti in diversi geni causali non sono ancora state trovate.

Exome sequencing è più utile per i disturbi con una probabile eterogeneità genetica. Questa strategia può essere estesa anche alle malattie con una genetica più complessa, attraverso l'analisi di campioni di dimensioni più grandi e lo studio dell'impatto funzionale delle varianti non sinonime identificate. Inoltre, il whole exome sequencing (WES) rappresenta una base promettente per la medicina genomica personalizzata, infatti grazie all'identificazione del gene causativo sarà possibile individuare il pathway proteico alterato che può fornire il target per una terapia mirata.

Si può concludere che, consentendo il sequenziamento contemporaneo dell'intera porzione codificante di tutti i geni candidati, o, se necessario, dell'intero esoma, le tecniche recentemente sviluppate del Next Generation Sequencing (NGS), ed in particolare il Whole Exome Sequencing (WES), possano essere l'approccio migliore volto all'identificazione delle anomalie genetiche sottostanti alle patologie con elevata eterogeneità genetica, come le epilessie. Se, da una parte, con il WES è possibile evidenziare alterazioni geniche in qualsiasi parte dell'esoma indipendentemente dal quesito clinico, dall'altra parte è maggiore il rischio, rispetto al sequenziamento di uno o più geni selezionati, di rivelare incidentalmente varianti geniche non correlate al quesito diagnostico primario. Dorschner et al. 2013 in una recente review ha analizzato 1.000 esomi scelti a casa

dal Server ExomeVariant (EVS, NHLBI Exome Sequencing Project <http://evs.gs.washington.edu/EVS>) e ha confrontato le 114 varianti geniche riscontrate nel campione a carico di geni-malattia (per esempio geni associati alla malattia conduzione cardiaca, predisposizione al cancro, ecc) con le mutazioni descritte nel Human Gene Mutation Database (HGMD, <http://www.hgmd.cf.ac.uk/ac/index.php>) identificando, nel 2,3 % dei casi, varianti considerate patogenetiche. Inoltre, oltre all'identificazioni di reperti incidentali, si evidenziano una notevole quantità di nuove varianti in geni noti e sconosciuti, che complicano ulteriormente l'interpretazione dei risultati. Per ridurre al minimo i problemi interpretativi, molti studi di ricerca hanno optato per l'analisi esomica del trio (caso indice con entrambi i genitori) facilitando notevolmente la rilevazione di mutazioni causative de novo e eterozigoti composti sia in geni dominati che recessivi. Tuttavia, in numerosi studi, sono state riscontrate 1-2 mutazioni de novo per ogni individuo, talvolta anche di più, ad indicare che la conferma di un origine de novo non è sufficiente a dimostrare la patogenicità di una variante genetica.

WES è uno strumento molto utile per l'analisi dell'esoma in studi di ricerca ed è anche uno degli strumenti più importanti nell'identificazione di nuovi geni causativi di malattia. Tuttavia, l'interpretazione dei dati è tuttora un compito molto impegnativo soprattutto in assenza del materiale genetico dei genitori. La validazione delle varianti sconosciute o resequencing di tutte le regioni sottorappresentate non è possibile. Tuttavia i database sulle varianti di codifica e del sito di splicing e le informazioni sulle numerose varianti attualmente conosciute sono in rapida crescita in modo da facilitare l'interpretazione dei risultati WES nel prossimo futuro .

CAPITOLO 4. DESCRIZIONE DEL PROGETTO DI RICERCA

Le epilessie si presentano con un'ampia gamma di caratteristiche cliniche, sono riconosciute oltre 50 sindromi epilettiche distinte. Sono condizioni eterogenee anche dal punto di vista eziologico, possono infatti essere sintomatiche, caratterizzate da causa identificabile come i disturbi metabolici, traumi o

tumori cerebrali intracranici; secondarie a disordini monogenici; o idiopatiche, che si verificano cioè in assenza di fattori causali identificabili e si pensa possano avere un forte contributo genetico, anche se attualmente l'eziologia genetica è determinata solo per una piccola parte di esse.

In letteratura sono presenti numerosi studi volti all'identificazione delle cause molecolari, sia attraverso studi di linkage sia con l'adozione della più recente ed efficace tecnica degli array-CGH; quest'ultima ha permesso a Mefford et al (2010) di evidenziare la presenza di variazioni del numero di copie (Copy Number Variants, CNVs) come delezioni o amplificazioni, in non meno del 3% dei casi studiati.

La diagnosi molecolare di epilessia è importante soprattutto in ambiente pediatrico sia per rassicurare i genitori che per utilizzare le terapie mirate, ma anche per la precisa determinazione del rischio di ricorrenza nelle gravidanze successive, tenendo presente che alcune anomalie sono caratterizzate da bassa penetranza/variabilità fenotipica (Escayg et al., 2010).

L'identificazione dell'anomalia genetica patologica nelle epilessie idiopatiche può essere complicata, perché lo stesso fenotipo caratterizzato da anomalie elettroencefalografiche può essere riconducibile a diversi geni. La tendenza generale nella ricerca genetica delle epilessie è lo sviluppo di piattaforme sempre più potenti volte all'identificazione di CNVs e lo studio approfondito della sequenza genomica.

Inevitabilmente, questo approccio porterà rapidamente alla scoperta di anomalie genetiche potenzialmente correlate alle epilessie genetiche, e questo avrà importanti implicazioni per la nostra capacità di diagnosticare e prevedere con precisione il follow up. Inoltre, tale sforzo in ambito di diagnosi molecolare porterà alla stratificazione dei pazienti in funzione del pathway proteico alterato in modo da consentire la conduzione di trials clinici volti al recupero di funzione delle macromolecole compromesse dall'alterazione genica.

In tale contesto scientifico nasce il progetto oggetto di questa tesi: verificare l'efficacia e affidabilità di un pannello diagnostico mediante targeted resequencing per encefalopatie epilettiche (EE) in età pediatrica.

Attraverso tale pannello genetico è stato possibile dapprima lo screening mutazionale di 19 geni (Versione 1.0, Giugno 2013); successivamente, in rapporto alle sempre più attuali conoscenze in

letteratura, il pannello è stato ampliato e ha permesso lo screening mutazionale di 37 geni (Versione 2.0, Settembre 2016).

I geni considerati nel pannello diagnostico sono tutti attualmente noti come associati a malattie genetiche o candidati scelti perché appartengono a “pathways” simili.

4.1 MATERIALI E METODI

Il test “pannello genetico” è una complessa procedura diagnostica basata su nuove tecniche di sequenziamento (NGS), che consentono l'analisi di numerosi geni in un unico saggio. Il test si basa sulla capacità di sequenziare in parallelo numerosi geni di interesse di un individuo per determinare in maniera accurata il profilo genomico. Il profilo genomico individuale viene poi paragonato ad un profilo genomico di riferimento che rispecchia la popolazione generale al fine di identificare variazioni geniche del possibile significato clinico. Il sequenziamento è stato effettuato con gli strumenti Ion Torrent PGM System (Life Technologies) e MiSeq Sequencing System (Illumina) come da procedura. L'utilizzo del chip 316 su piattaforma PGM ha permesso di sequenziare contemporaneamente 16 pazienti; il kit V2 (300 cicli) su piattaforma MiSeq ha consentito di sequenziare invece 24 pazienti EE a esperimento. Le reads di sequenza (fastq) sono state analizzate utilizzando in parallelo due pipeline: Ion Reporter (Life Technologies) e CLC Genomics Workbench 6.5.1 (Qiagen) per i dati provenienti dal PGM, Sure Call (Agilent) e Biomedical Genomics Workbench (Qiagen) per i dati provenienti dal MiSeq. La chiamata delle varianti è stata effettuata con un coverage minimo di 20 reads ($\geq 20X$) e filtrando le varianti con frequenza $>1\%$ (dbSNP138 and EVS database). Le varianti chiamate sono state validate con metodo Sanger.

L'effetto patogenico delle varianti è stato attribuito sulla base di:

- a) effetto funzionale putativo (Polyphen 2.0, Mutation Taster)
- b) correlazione genotipo-fenotipo
- c) analisi di segregazione nei genitori, ove presenti.

La prima versione del test (Ver 1.0, Giugno 2013) per EE prevedeva lo screening di mutazioni puntiformi dei seguenti 19 geni:

SCN1A, PCDH19, STXBP1, SLC22A25, CDKL5, KCNQ2, SLC2A1, PNPO, ALDH7A1, KCNT1, SCN2A, SCN8, ARHGEF9, GRIN2A, TBC1D24, PNKP, PLCB1, ST3GAL3, SPTAN1

La seconda e attuale versione del test (Ver 2.0, Settembre 2016) prevede lo screening mutazionale di 37 geni:

ALDH7A1, ALG13, BRAT1, CDKL5, CHD2, DNM1, GABRA1, GABRA3, GNAO1, GRIN2A, HCN1, IQSEC2, KCNA2, KCNB1, KCNQ2, KCNQ3, KCNT1, MEF2C, PCDH19, PIGA, PNKP, PNPO, PRRT2, QARS, SCN1A, SCN2A, SCN8A, SLC2A1, SLC25A22, SCL35A2, SLC6A1, SPTAN1, ST3GAL3, STXBP1, SYNGAP1, TBC1D24, WWOX

I pazienti sono stati selezionati mediante i seguenti requisiti fondamentali:

- Quadro clinico compatibile con una delle sindromi epilettiche in studio, secondo la classificazione ILAE del 2010
- Dati clinico-anamnestici completi
- Dati antropometrici e valutazione neurologica approfondita
- Anamnesi epilettologica e quadro elettroencefalografico all'esordio dell'epilessia, all'esordio di altri tipi di crisi ed all'ultimo controllo (rivisti e interpretati secondo gli standard ILAE)
- Informazioni relative a eventuali altri test genetici eseguiti
- Esclusione di quadri sintomatici con Risonanza Magnetica dell'encefalo (MRI)
- Esclusione di disordini congeniti del metabolismo con esami ematici di primo livello
- Acquisizione del consenso informato.

Questi dati hanno permesso di individuare cinque grandi raggruppamenti nell'ambito delle encefalopatie epilettiche:

1. EE ad esordio nel primo anno di vita (tra cui S. Ohtahara e le crisi piridossina-dipendenti) (EOEE)
2. Spettro spasmi infantili-Sindrome di West
3. Sindrome di Dravet e forme correlate
4. Spettro CSWS-Sindrome di Landau Kleffner
5. EE con esordio successivo al primo anno di vita

5. RISULTATI

Sono stati analizzati 258 pazienti (Tab. 2) di cui 94 appartenenti al gruppo EOEE; 48 pazienti con epilessia ad insorgenza dopo il 1° anno di vita; 66 pazienti con Sindrome di Dravet e forme correlate, 37 pazienti appartenenti allo spettro Spasmi Infantili /S. West; 13 pazienti allo spettro CSWS. Mediante l'utilizzo del pannello EE, sono state identificate, nel periodo compreso tra Giugno 2014 e Settembre 2017, 73 varianti geniche in 69/258 pazienti analizzati (Tab. 3):

- n° 47 (64,3%) varianti patogenetiche su 73 varianti geniche riscontrate in 258 pazienti (10 SCN1A, 10 KCNQ2, 6 CDKL5, 4 STXBP1, 3 KCNT1, 2 SCN8A, 2 PCDH19, 2 SCN2A, 2 TBC1D24, 2 CHD2, 1 GRIN2A, 1 PNPO, 1 SPTAN1, 1 SLC2A1)
- 7 varianti probabilmente patogenetiche ma delle quali la segregazione è mancante (1 ALDH7A1, 1 SCN1A, 2 SCN2A, 1 SCN8A, 1 SLC2A1, 1 SLC35A3)
- 19 (26 %) varianti di significato incerto (VUSs) su 73 varianti geniche riscontrate in 15 pazienti

Il rate diagnostico (Tab2) per EOEE è risultata del 24%; per lo spettro Dravet del 22,7%, per il gruppo Spasmi Infantili /S.West del 13,5%; per EE dopo il I anno di vita del 5%, per lo spettro CSWS del 7,6%.

6. DISCUSSIONE

I pannelli NGS sono, allo stato attuale, tecniche ampiamente utilizzate per identificare le cause genetiche di epilessia, sostituendo il tradizionale approccio gene-by-gene. Questo strumento si mostra particolarmente utile e di rapida esecuzione e consente il sequenziamento contemporaneo dell'intera porzione codificante di tutti i geni candidati (pannello diagnostico), o, se necessario, dell'intero esoma (Whole Exome Sequencing) e numerosi sono gli studi che identificano tale metodiche come l'approccio migliore volto all'identificazione delle anomalie sottostanti nelle malattie con elevata eterogeneità genetica, come le epilessie (Veeramah et L. 2013, Epi4K Consortium , Nature 2013) . Inoltre tale approccio consente l'identificazione non solo di mutazioni classiche già note per essere associate con

un determinato tipo di epilessia, ma anche differenti tipi di mutazioni, nello stesso gene, eventualmente associato ad un diverso tipo di epilessia.

In un recente lavoro pubblicato su *The New England Journal of Medicine* del gruppo americano del Baylor College di Houston (Y. Yang et al. Oct. 2013) è stato valutato l'impatto diagnostico del Whole Exome Sequencing in pazienti con fenotipo suggestivo di patologie genetiche. Sono stati studiati retrospettivamente 250 probandi di cui 213 affetti da un disordine neurologico associato o meno a altre patologie organiche (140 pazienti) e 27 affetti da patologie non neurologiche, al momento dello studio, senza inquadramento eziologico. Sono state individuate circa da 200,000 a 400,000 SNVs e microriarrangiamenti/microdelezioni nel genoma umano del campione. Successivamente sono stati studiati i geni codificati dalle regioni mutate per verificare se le mutazioni dei geni candidati «possano fittare» con il fenotipo clinico dei pazienti. Più del 86% delle SNVs sono state confermate mediante sequenziamento Sanger, il restante 14% è risultato un falso positivo. In 62 pazienti sono state individuate mutazioni patogenetiche (36 autosomico dominanti, 16 autosomico recessive, 9 x-linked) e in 4 pazienti è stata formulata diagnosi di 2 patologie genetiche diverse. L'83% mutazioni è risultato de novo, il 67% mutazioni identificate non erano state precedentemente riportate in letteratura. In tale lavoro il WES ha permesso di formulare una diagnosi nel 25% dei casi, percentuale che risulta essere maggiore rispetto ai tassi positivi (dal 3,5 al 10%) riscontrati con metodiche standard (come l'analisi del cariotipo e l'array-CGH) in bambini con disordini neurologici. Gli autori validano, quindi, l'utilizzo di tale procedura come test diagnostico per pazienti con patologie senza caratteristiche specifiche o con fenotipo atipico di patologie che indirizzano verso una possibile causa genetica, e per i pazienti con diagnosi clinica di malattie genetiche eterogenee. Tuttavia viene evidenziato il grosso limite di questa metodica e cioè la difficoltà e la complessità di analisi ed interpretazione dei dati ottenuti e proposta un'alternativa quale appunto il sequenziamento mirato di geni «candidati» attraverso pannelli diagnostici con relativo contenimento dei costi, e maggiore sensibilità nell'identificare alterazioni della sequenza genica nei geni testati con minor problemi interpretativi.

La problematica dell'interpretazione dei dati ottenuti con il WES è stata valutata da diversi gruppi di studio. Dorschner et al. 2013 in un recente lavoro ha studiato le SNVs riscontrate incidentalmente in

1000 partecipanti al progetto “National Heart, Lung, and Blood Institute ExomeSequencing Project”. Il DNA è stato screenato per varianti geniche note in 114 geni selezionati per la loro implicazione in condizioni genetiche spesso diagnosticate solo in età adulta (per esempio geni associati alla malattia conduzione cardiaca, predisposizione al cancro, ecc), i risultati ottenuti sono stati confrontati con Human Gene Mutation Database (HGMD). Sono state individuate in 23 partecipanti (2,3% del campione) varianti geniche patogeniche o probabilmente patogenetiche, complicando notevolmente il processo di counselling genetico. Gli autori sottolineano la necessità di sviluppare criteri selettivi per la segnalazione delle SNVs soprattutto quando riscontrate incidentalmente. A tal proposito l'American College of Medical Genetics and Genomics (ACMG) ha redatto linee guida sulla comunicazione dei risultati genetici incidentali ottenuti con tecniche di NGS aprendo interessanti punti di discussione. Inoltre, oltre ai reperti incidentali, le tecniche di NGS rivelano un elevato numero di nuove varianti geniche sia in geni noti che sconosciuti complicando ulteriormente l'interpretazione dei risultati.

I pannelli diagnostici rappresenterebbero, in accordo con i dati di letteratura, un'alternativa più efficiente ed economica.

Con questo studio, appunto, abbiamo voluto verificare l'efficacia diagnostica di due pannelli multigenici custom finalizzati alla diagnosi delle encefalopatie epilettiche (EE) in età pediatrica, mediante tecnica NGS.

Le finalità diagnostiche del pannello hanno indirizzato la scelta su geni per i quali erano disponibili 1) evidenze multiple ed indipendenti del ruolo patogenetico; 2) robuste correlazioni genotipo-fenotipo. Sono stati pertanto esclusi geni il cui ruolo nei fenotipi in esame non era comprovato o risultava controverso.

I nostri risultati sono conformi a quelli precedentemente riportati da altri autori, con rate diagnostiche che vanno dal 10% al 48,5% (Lemke et al, 2012; Della Mina et al, 2015; Parrini et al, 2016) In accordo con i dati della letteratura le mutazioni de novo svolgono un ruolo importante nelle EE (Allen AS, 2013; Epi4K- Consortium 2014; Helbig, 2016) e i risultati positivi sono relativi a geni noti perché causali di grave epilessia precoce e con stretta relazione genotipo-fenotipo.

L'utilizzo del pannello EE nel nostro studio ha permesso di identificare la causa genetica di epilessia nel

18,2% dei casi (47 mutazioni patogenetiche su 258 pz analizzati); con una detection rate del 24,4% nel gruppo EOEE; del 22,7% per lo spettro Dravet, del 13,5 % per il gruppo Spasmi Infantili /S.West , del 5% per EE ad insorgenza dopo il I anno di vita, del 7,6%, per lo spettro CSWS (Tab2).

Nell' analisi del nostro pannello EE il rate diagnostico più basso si è verificato nel gruppo di epilessie ad insorgenza dopo il I anno di vita, verosimilmente per la maggiore eterogeneità fenotipica ed elettroencefalografica; per alcuni pazienti un secondo approfondimento neuroradiologico (anche con tecniche più avanzate quali RM 3T) ha dimostrato anomalie cerebrali strutturabili non note in età precoce e che avrebbero escluso il paziente da tale indagine. Inoltre in alcuni pazienti, la successiva maggiore caratterizzazione fenotipica (per esempio l'individuazione di morfismi specifici con l'accrescimento) ha permesso la diagnosi attraverso altre tecniche diagnostiche (es. array-CGH) e quindi il riscontro di riarrangiamenti cromosomici riguardanti altri geni coinvolti nell'epilessia e non presenti nel nostro pannello.

Dall'analisi del pannello EE un altro dato interessante è il riscontro del 73% di mutazioni in SCN1A tra le mutazioni individuate in pazienti con fenotipo Dravet-like (11 mutazioni in SCN1A su 15 riscontrate in 66 pazienti con fenotipo Dravet-like) (Tab.3). Questo fenomeno conferma come il fenotipo Dravet sia strettamente correlato a mutazioni del gene SCN1A. Tuttavia a conferma dell'eterogeneità genetica della Sindrome di Dravet, abbiamo identificato mutazioni in altri geni la cui correlazione con tale sindrome è già nota (PCDH19, SCN2A, SCN8A) (Depienne C. et al 2009, Suls A. et al 2013).

È ormai noto come in una minor percentuale di femmine o di maschi con mosaicismo germinale affetti da epilessia con fenotipo Dravet-like, si riscontri una mutazione in PCDH19 (Depienne C. et al. 2009 e 2012) o in SCN8A (Hirose S. et al 2007; Harkin LA. et al. 2002; Patino GA. Et al 2009). Recentemente in letteratura sono stati riportati 3 casi con mutazione in CHD2 affetti da encefalopatia epilettica con caratteristiche cliniche e elettroencefalografiche sovrapponibili ai pazienti con S. di Dravet (Suls A. et al 2013).

Nel lavoro di Carvill et al. su un campione di 80 soggetti con epilessia Dravet-like sono stati individuati con metodiche di NGS, 4 casi con mutazione in GABRA1 e 3 casi mutati in STXBP1, tutte de novo, non precedentemente descritte in associazione con tale quadro clinico. Rispetto ai casi con fenotipo

Dravet classico, in tali pazienti vengono riportate con maggiore incidenza (2/4 pazienti mutati in GABRA1 e 2/3 pazienti mutati in STXBP1) crisi atoniche con caduta e crisi toniche già dopo pochi mesi dall'esordio. Inoltre 2/4 pazienti mutati in GABRA1 presentano la stessa sostituzione amminoacidica (c355 G>A, p.Arg112Gin) a suggerire che possa essere ricorrente in pazienti con S. di Dravet. Tali dati sembrano ridefinire il ruolo di GABRA1 (inserito nella nuova versione del pannello 2.0) da gene associato a epilessie generalizzate con scarsa compromissione clinica e a convulsioni febbrili a gene implicato in encefalopatie epilettiche devastanti come la S. di Dravet.

Fra i geni associati alle encefalopatie epilettiche ad esordio precoce, si annovera STXBP1, (nella nostra casistica individuate 4 mutazioni in STXBP1 de novo), che è descritto in casi con esordio dell'epilessia prima dei 4 mesi con crisi toniche, spasmi o crisi miocloniche ed un quadro elettroencefalografico tipo suppression-bursts (Hiroto S. et al 2008; Holland KD. et al. 2010). I pazienti mutati in STXBP1 individuati da Carvill GL. et al. hanno tutti un esordio più tardivo delle crisi e l'ampia variabilità delle mutazioni descritte in tale gene non ha permesso una correlazione genotipo-fenotipo fra il tipo di mutazione e le caratteristiche cliniche. Gli autori concludono che GABRA1 e STXBP1 possono essere considerati responsabili di molti casi con fenotipo Dravet e che devono essere studiati nei casi non diagnosticati dove le mutazioni nei geni noti (SCN1A e PCDH19) sono già state escluse.

Nella nuova versione del test (Ver 2.0, Settembre 2016) è stato incluso il gene HCN1 (hyperpolarization activated cyclic nucleotide-gated potassium channel 1); ad oggi sono stati individuati dal gruppo di Nava C. et al. 6 casi mutati con caratteristiche cliniche simili con esordio fra i 4 e i 13 mesi con crisi polimorfe sia focali che generalizzate sia in febbre che in apiressia e successivo sviluppo di assenze atipiche, crisi miocloniche e crisi tonico-cloniche generalizzate; è costata la compromissione cognitiva da moderata a severa così come l'associazione con atipie comportamentali e tratti autistici. HCN1 appartiene alla famiglia dei pore-loop cation channel costituita da 4 sottotipi espressi principalmente nel cuore e nel SNC (Bennarroch E. 2013), mappa sul cromosoma 5p12. HCN1 gioca un ruolo fondamentale nel controllo dell'eccitabilità neuronale, nell'integrazione dendritica e nella trasmissione dei potenziali sinaptici e nella oscillazione ritmica del potenziale di membrana dei neuroni e nel network neuronale. Questo canale sarebbe inoltre coinvolto in meccanismi che regolano la plasticità

sinaptica e la memoria e studi sperimentali ne hanno evidenziato la possibile implicazione nell'epilettogenesi e nei meccanismi che regolano il dolore (Noan Y. Et al 2011). Il ruolo di HCN1 nell'epilessia sembra alquanto complesso; studi sperimentali su modelli animali hanno mostrato che sia un "upregulation" che una "downregulation" del canale possono essere associate a crisi epilettiche e che l'effetto finale dipende dal contesto fisiologico e dal tipo di cellula in cui il canale è espresso, così come accade per il canale del sodio voltaggio-dipendente (Nav 1.1 e Nav 1.7) (Kase D. et al. 2012). Dall'analisi dei dati preliminari su modelli sperimentali wilde-type (Figura 8), sembrerebbe che le mutazioni de novo individuate nei 6 pazienti mutati HCN1 determinerebbero principalmente un guadagno di funzione del canale piuttosto che una perdita.

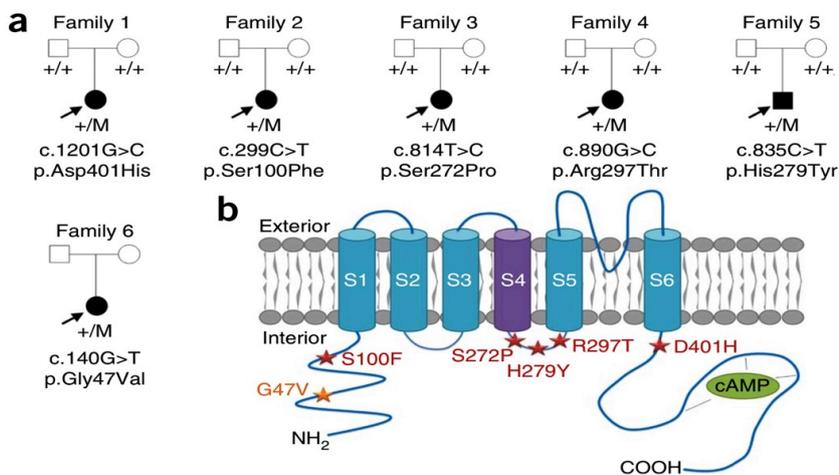


Figura 8: Pedigree e analisi di segregazione delle 6 mutazioni in HCN1 individuate nello studio

Tale studio, pertanto, sembrerebbe validare l'implicazione di tale gene nelle epilessie e il fenotipo presentato dai casi mutati in HCN1, con crisi febbrili all'esordio e successivo sviluppo di crisi focali e assenze, sono simili a quelle presentate dai modelli sperimentali animali, sottolineando la possibile implicazione di questo gene nei casi di epilessia Dravet-like.

7. CONCLUSIONI:

I pannelli custom utilizzati in questo studio si sono dimostrati strumenti efficaci ed affidabili per individuare la causa genetica di pazienti con EE ed MCD; la diagnosi certa è stata ottenuta nel 18,2% dei casi per i pazienti con EE. Nuovi geni non ancora definiti potrebbero giocare un ruolo nei casi non diagnosticati; a tal fine, il pannello “custom” si rivela uno strumento assai efficace e dinamico, poiché permette un continuo aggiornamento del disegno in accordo con gli ultimi dati della letteratura. NGS si è rivelato, inoltre, uno strumento efficace per individuare i difetti genetici presenti nei pazienti con fenotipi epilettici altamente eterogenei con un’ampia e poco chiara eziologia genetica, come la Sindrome di Dravet SCN1A. L’alta percentuale di rilevazione di mutazioni note dimostra che un fenotipo clinico specifico indichi già la possibile presenza di geni specifici e di conseguenza un utilizzo di metodi tradizionali potrebbe essere la scelta diagnostica più facile da effettuare. Tuttavia, quando un particolare fenotipo può essere causato da mutazioni in diversi geni, lo screening diretto con i metodi convenzionali è lungo e costoso e di conseguenza si può pensare di utilizzare il NGS.

La conferma di mutazioni già rilevate con il sequenziamento tradizionale ha comunque fortemente sottolineato l’attendibilità di queste nuove metodiche diagnostiche.

Anche dati della letteratura (Veeramah et L. 2013, Epi4K Consortium 2013) mostrano che in pazienti con fenotipo per cui il background genetico è largamente sconosciuto, come nel caso degli Spasmi Infantili non lesionali e delle Encefalopatie Epilettiche non classificate, in cui il sequenziamento tradizionale può essere molto complesso a causa della moltitudine di geni potenzialmente causativi di un certo fenotipo, l’utilizzo di NGS può essere la metodica preferenziale. I dati in nostro possesso, pur ancora preliminari, ci permettono di ipotizzare che l’utilizzo di NGS, rappresenti un veloce e conveniente metodo di screening diagnostico per analizzare le basi genetiche della epilessia.

Questo approccio innovativo ha così dimostrato di essere un potente strumento di diagnostica che può contribuire alla raccolta di informazioni sulle forme di epilessia sia comuni che ignote e nel delineare l’associazione dei fenotipi con geni mutati meno frequentemente.

Le grandi quantità di dati di sequenza in uscita rappresentano una sfida per la bioinformatica clinica di laboratorio. Oltre al trattamento dei dati, l’interpretazione del sequenziamento dei risultati richiede

un'ulteriore caratterizzazione delle varianti genomiche presenti nelle regioni analizzate. Anche se molto lavoro ci attende per l'utilizzo di NGS in diagnostica clinica, le potenziali applicazioni sono entusiasmanti e numerose.

Lo studio mirato all'identificazione dei geni che possono determinare o anche solo influenzare il rischio di sviluppare epilessia oltre ad aiutare a migliorare le conoscenze della patogenesi dell'epilettogenesi, della sua variabilità e suscettibilità genetica, ha implicazioni di rilievo nella diagnosi per il paziente e gli altri membri della famiglia, nella gestione clinica, nella cura e nella prevenzione.

Tabella 2 PATIENTS ANALYZED WITH NGS EPILEPSY PANEL

ELECTROCLINICAL SYNDROMES	Number of patient analyzed	Number of pathogenic mutation	Diagnostic rate %
Early onset EEs (Ohtahara Syndrome and Epilepsies with pyridoxin-dependent seizures, other unclassified) (0-12 months)	94	23	24,4%
Infantile spasms and West syndrome	37	5	13,5%
Dravet Syndrome and related conditions	66	15	22,7%
CSWS and Landau Kleffner Syndromes	13	1	7,6%
Epilepsy with onset after the first year of life (Lennox-Gastaut, other unclassified)	48	3	5%
Total	258	47	18,2%

Figura 9. PATIENTS ANALYZED WITH NGS EPILEPSY PANEL

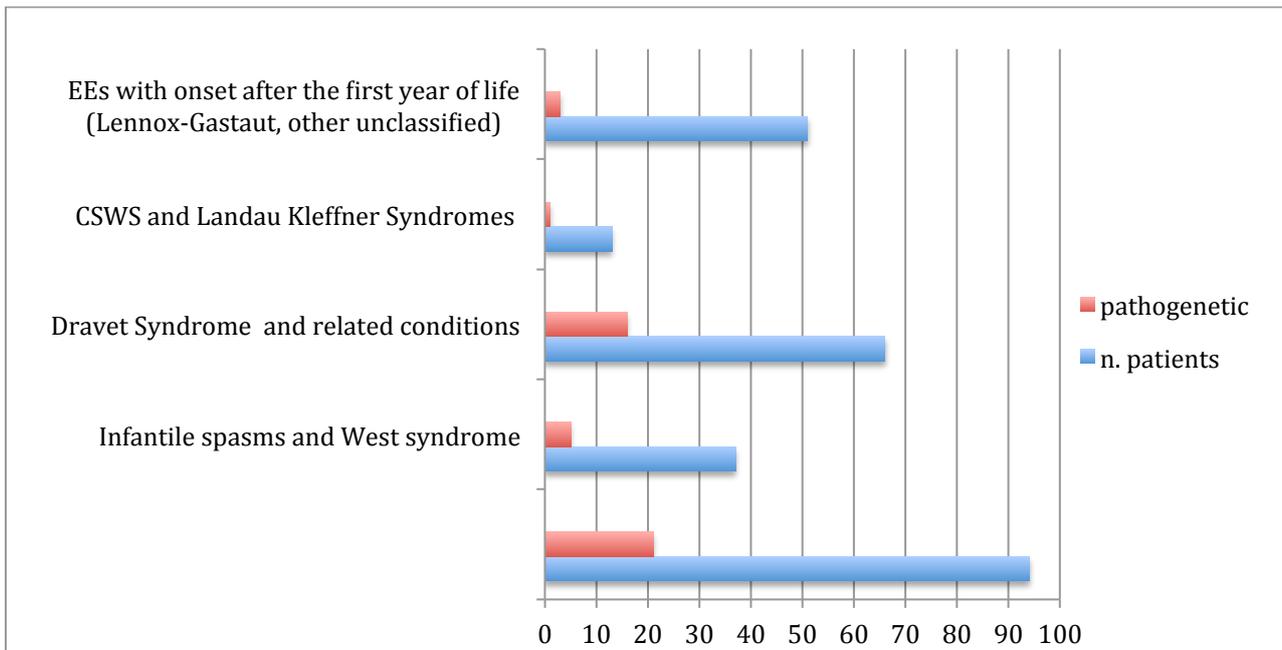


Figura 10. GENES WITH PATHOGENIC OR LIKELY PATHOGENIC AND POTENTIALLY CAUSATIVE IN 258 CLINICALLY REFERRED EPILEPSY PATIENTS

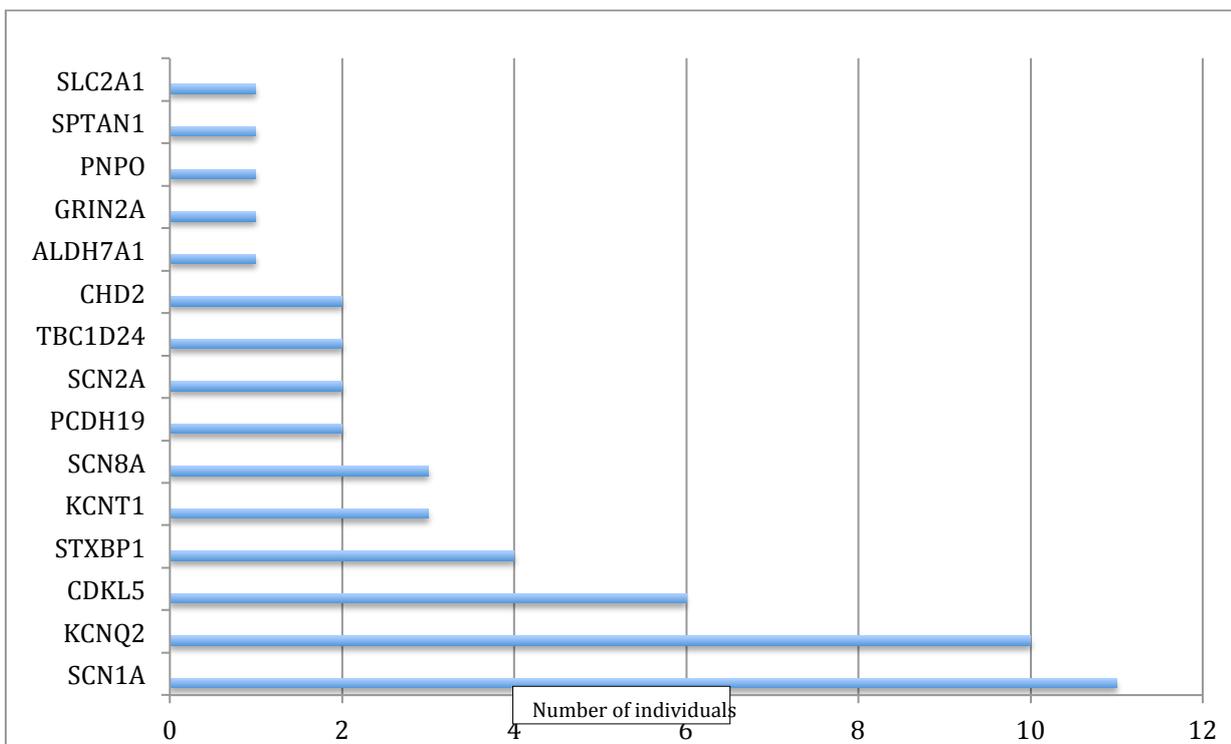


Tabella 3. PATHOGENIC AND VUS MUTATIONS IDENTIFIED BY A MULTIGENIC PANEL OF EPILEPSY.

CASE N°.	NGS PANEL	GENE	INHERITANCE	ZYGOSITY	VARIANT	CLASSIFICATION	PHENOTYPE	GENDER
1.	20 genes	ALDH7A	Unknown	Heterozygous	1 c.1235C>T	Likely pathogenic variant	EOEE	M
2.	20 genes	ALDH7A1	Maternal	Heterozygous	c.650+7C>A	Variant of unknown significance(VUS)	EOEE	M
3.	20 genes	ALDH7A1	Maternal	Heterozygous	c.359C>T	Variant of unknown significance(VUS)	EOEE	F
4.	20 genes	CDKL5	Unknown	Heterozygous	c.2249C>G [p.Ser750*]	Pathogenic	EOEE	F
5.	20 genes	CDKL5	De novo	Heterozygous	c.528G>A [p.Trp176*]	Pathogenic	EOEE	F
6.	20 genes	CDKL5	De novo	Heterozygous	c.171delT [p.Thr57fs]	Pathogenic	EOEE	F
7.	20 genes	CDKL5	De novo	Heterozygous	c.1670_1671insA [p.Ser557fs]	Pathogenic	EOEE	M
8.	37 genes	CDKL5	no maternal father unknown	Heterozygous	c.1930_1931delCT [p.Leu644fs];	Pathogenic	EOEE	F
9.	37 genes	CDKL5	De novo	Heterozygous	c.744+2T>C [splice site];	Pathogenic	EOEE	F
10.	20 genes	CDKL5	Mather wt father unknown	Heterozygous	c.1616G>C	Variant of unknown significance (VUS)	EOEE	F
11.	37 genes	CHD2	Unknown	Heterozygous	c.2425C>T [p.Arg809*]; no rs [0,000008]	Pathogenic	Epilepsy with onset after the first year of life	F
12.	20 genes	CHD2	De novo	Heterozygous	c.4164del (fshif) y	Pathogenic	Epilepsy with onset after the first year of life	F

13.	20 genes	CHD2 SLC35A3	Unknown	Double Heterozygous	c.3514C>T gene CHD2 c.212G>A gene SLC35A3	Variants of unknown significance(VUS)	Epilepsy with onset after the first year of life	M
14.	20 genes	GRIN2A	Unknown	Heterozygous	c.110C>T [p.Ala37Val]	Variant of unknown significance(VUS)	Epilepsy with onset after the first year of life	F
15.	20 genes	GRIN2A	Paternal	Heterozygous	c.4277A>G [p.Glu1426Gly]	Variant of unknown significance(VUS)	Epilepsy with onset after the first year of life	F
16.	20 genes	GRIN2A	de novo	Heterozygous	833 C>A	Pathogenic	Epilepsy with onset after the first year of life	M
17.	20 genes	KCNT1	Unknown	Heterozygous	c.2849G>A [p.Arg950Gln]	Pathogenic	EOEE	M
18.	20 genes	KCNT1	Unknown	Heterozygous	c.2220C>G [Asp740Glu]	Variant of unknown significance(VUS)	Epilepsy with onset after the first year of life	M
19.	37 genes	KCNT1	De novo	Heterozygous	c.2896G>A [p.Ala966Thr]; no rs [n/a]	Pathogenic	EOEE	M
20.	20 genes	KCNT1	De novo	Heterozygous	c.2677G>A agene KCNT1	Pathogenic	EOEE	M

21.	20 genes	KCNQ2	De novo	Heterozygous	c.830C>T [p.Thr277Ile]	Pathogenic	EOEE	M
22.	20 genes	KCNQ2	Unknown	Heterozygous	c.918G>A [synonymous]	Pathogenic	EOEE	M
23.	20 genes	KCNQ2	De novo	Heterozygous	c.812G>A [p.Gly271Asp]	Pathogenic	EOEE	F
24.	20 genes	KCNQ2	Unknown	Heterozygous	c.2105C>T [p.Ser702Leu]	Pathogenic	EOEE	M
25.	20 genes	KCNQ2	De novo	Heterozygous	c.587C>T [p.Ala196Val]	Pathogenic	EOEE	F
26.	20 genes	KCNQ2	Unknown	Heterozygous	637C>T [p.Arg213Trp]	Pathogenic	EOEE	F
27.	37 genes	KCNQ2	Unknown	Heterozygous	c.1639C>T [p.Arg547Trp]; no rs [n/a]	Pathogenic	EOEE	M
28.	37 genes	KCNQ2	Unknown	Heterozygous	c.1513C>T [p.Gln505*]; no rs [Pathogenic	EOEE	M
29.	37 genes	KCNQ2	Unknown	Heterozygous	c.1687G>A [p.Asp563Asn]; rs796052653	Pathogenic	EOEE	M
30.	37 genes	KCNQ2	Unknown	Heterozygous	c.1657C>T [p.Arg553Trp]; rs759584387 [0,00001]	Pathogenic	EOEE	F
31.	37 genes	PNKP	Maternal Paternal	Heterozygous Heterozygous	c.1360C>A [p.Leu454Met]; rs200611702 [0,00032] c.672C>T [silente]; rs151180981 [0,00016]	Variants of unknown significance(VUS)	EOEE	F
32.	20genes	PCDH19	Paternal	Heterozygous	c.446C>A	Pathogenic	Dravet Syndrome and related conditions	F
33.	37 genes	PCDH19	De novo	Heterozygous	c.814C>T [p.Gln272*] rs794726897	Pathogenic	EOEE	F
34.	20 genes	PNPO	Paternal+ Maternal	Omozygous	c.674G>A [p.Arg225His]	Pathogenic	EOEE	M
35.	20 genes	SCN1A	De novo	Heterozygous	c.5326G>T [p.Val1776Phe]	Pathogenic	Dravet Syndrome and related	F

							conditions	
36.	20 genes	SCN1A	Unknown	Heterozygous	c.302G>A [p.Arg101Gln]	Likely pathogenic variant	Dravet Syndrome and related conditions	F
37.	20 genes	SCN1A	Unknown	Heterozygous	c.4945C>T [synonymous]	Variant of unknown significance(VUS)	Dravet Syndrome and related conditions	M
38.	20 genes	SCN1A	De novo	Heterozygous	c.301C>T [p.Arg101Trp]	Pathogenic	Dravet Syndrome and related conditions	F
39.	20 genes	SCN1A	De novo	Heterozygous	c.3266delG [p.Gly1089fs]	Pathogenic	Dravet Syndrome and related conditions	F
40.	37 genes	SCN1A	De novo	Heterozygous	c.4547C>A [p.Ser1516*]; no rs	Pathogenic	Dravet Syndrome and related conditions	M
41.	37 genes	SCN1A	Unknown	Heterozygous	c.187T>C [p.Phe63Leu]; rs121917907	Pathogenic	Dravet Syndrome and related conditions	F
42.	20 genes	SCN1A	De novo	Heterozygous	c.25C>G	Pathogenic	Dravet Syndrome and related conditions	F
43.	20 genes	SCN1A	De novo	Heterozygous	c.2378C>T	Pathogenic	Dravet Syndrome and related conditions	M
44.	20 genes	SCN1A	Unknown	Heterozygous	c. 2650 G>A	Pathogenic	Dravet Syndrome and related conditions	M

							conditions	
45.	20 genes	SCN1A	De novo	Heterozygous	c. 4543 G>A>T	Pathogenic	Dravet Syndrome and related conditions	M
46.	20 genes	SCN1A	Unknown	Heterozygous	c.4319C>T	Pathogenic	Dravet Syndrome and related conditions	M
47.	20 genes	SCN2A	De novo	Heterozygous	c.4633A>G [p.Met1545Val]	Pathogenic	EOEE	M
48.	20 genes	SCN2A	Unknown	Heterozygous	c.3754A>G [p.Ile1252Val]	Likely pathogenic variant	Infantile spasms and West syndrome	M
49.	20 genes	SCN2A	Unknown	Heterozygous	c.1270G>T [p.Val424Leu]	Likely pathogenic variant	EOEE	F
50.	20 genes	SCN2A	Maternal	Heterozygous	c. 3043 G>A	Variant of unknown significance(VUS)	Dravet Syndrome and related conditions	F
51.	20 genes	SCN8A	Unknown	Heterozygous	c.3967G>A [p.Ala1323Thr]	Variant of unknown significance(VUS)	Dravet Syndrome and related conditions	F
52.	20 genes	SCN8A	De novo	Heterozygous	c.-8A>G [?]	Pathogenic	Dravet Syndrome and related conditions	M
53.	20 genes	SCN8A	Unknown	Heterozygous	c.5630A>G [p.Asn1877Ser]	Likely pathogenic variant	Dravet Syndrome and related conditions	M
54.	20 genes	SCN8A	Materna	Heterozygous	c.3101T>C [p.Leu1034Pro]	Variant of unknown significance(VUS)	Dravet Syndrome and related conditions	F

55.	20 genes	SCN8A	De novo	Heterozygous	c. 3943G>A	Variant of unknown significance(VUS)	EOEE	F
56.	20 genes	SCN8A SPTAN1	De novo De novo	Maternal Heterozygous	c.1201T>C c.2687T>G	Pathogenic	Dravet Syndrome and related conditions	M
57.	37 genes	SLC25A22:	Paternal Maternal	Heterozygous Heterozygous	c.736T>C [p.Cys246Arg]; no rs c.235G>C [p.Glu79Gln]; no rs	Pathogenic	EOEE	F
58.	20 genes	SLC2A1:	De novo	Heterozygous	c.511_512delGCinsTT [p.Ala171Phe]	Pathogenic	EOEE	F
59.	20 genes	SLC2A1	Unknown	Heterozygous	c.606C>G	Variant of unknown significance(VUS)	Epilepsy with onset after the first year of life	F
60.	20 genes	SLC2A1	Maternal	Heterozygous	c.180G>A	Likely pathogenic variant	EOEE	F
61.	20 genes	SLC2A1 SPTAN1	Unknown	Heterozygous	c.606C>G gene SLC2A1 c.5044-4C>T gene SPTAN1	Variants of unknown significance(VUS)	EOEE	M
62.	20 genes	SLC35A3	Unknown	Heterozygous	c. 310 G>A	Likely pathogenic variant	Epilepsy with onset after the first year of life	M
63.	20 genes	SPTAN1	Unknown	Heterozygous	c.1007G>A [p.Arg336Gln]	Variant of unknown significance(VUS)	Infantile spasms and West syndrome	F
64.	20 genes	STXBP1	De novo	Heterozygous	c.704G>A [p.Arg235Gln]	Pathogenic	EOEE	M
65.	20 genes	STXBP1	De novo	Heterozygous	c. 922A>T [p.Lys308*]	Pathogenic	EOEE	F
66.	20 genes	STXBP1:	De novo	Heterozygous	c.430-	Pathogenic	EOEE	M

					2_430+2delinsTGGGAGA			
67.	37 genes	STXBP1:	De novo	Heterozygous	c.57_59delAAA [p.Ile19_Lys20delinsMet], no rs [n/a]	Pathogenic	EOEE	M
68.	20 genes	TBC1D24:	Unknown	Heterozygous Heterozygous	619C>T [[Gln207*] 1512A>G [splicing]	Pathogenic	EOEE	M
69.	20 genes	TBC1D24:	Unknown	Omozygous	32A>G [p.Asp11Gly]	Pathogenic	EOEE	F

BIBLIOGRAFIA

- Akhras MS, Unemo M, Thiyagarajan S, Nyren P, Davis RW, Fire AZ, Pourmand N. Connector inversion probe technology: a powerful one-primer multiplex DNA amplification system for numerous scientific applications. *PLoS One*. **19**;2(9):e915 (2007).
- Allen AS, Berkovic SF, Cossette P, et al. De novo mutations in epileptic encephalopathies. *Nature* **501**:217–221 (2013).
- Andermann E, Metrakos JD. EEG studies of relatives of probands with focal epilepsy who have been treated surgically. *Epilepsia*. **10**(3):415 (1969).
- Andermann F, Kobayashi E, Andermann E. Genetic focal epilepsies: state of the art and paths to the future. *Epilepsia*. **46** Suppl 10:61-7 (2005).
- Avanzini G. and Noebels J. Genetics of Epilepsy and genetic epilepsies. *Mariani Foundation Paediatric Neurology*: **20**. John Libbey Eurotext. (2009).
- Baulac S, Baulac M. Advances on the genetics of Mendelian idiopathic epilepsies. *Clin Lab Med* **30**:911–929 (2010).
- Bahi-Buisson N. et al. The three stages of epilepsy in patients with CDKL5 mutations. *Epilepsia* **49**(6):1027–1037 (2008).
- Bahi-Buisson N. et al. Key clinical features to identify girls with CDKL5 mutations. *Brain* **131** (Pt 10): 2647–2661 (2008).
- Barcia, G. et al. De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. *Nat Genet* **44**, 1255-9 (2012).
- Basel-Vanagaite L. et al. SZT2 mutations cause infantile encephalopathy with epilepsy and dysmorphic corpus callosum. *Am. J. Hum. Genet.* **93**: 524-529 (2013).
- Baulac S, Baulac M- Advances on the genetics of Mendelian idiopathic epilepsies. *Clin Lab Med* **30**:911–929 (2010).
- Benlounis A. et al. Genetic predisposition to severe myoclonic epilepsy in infancy. *Epilepsia* **42**(2):204–209 (2001).
- Berg AT. et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia*. **51**(4):676- 85 (2010)
- Berkovic SF, Mulley JC, Scheffer IE, Petrou S. Human epilepsies: interaction of genetic and acquired factors. *Trends Neurosci.* Jul;**29**(7):391-7 (2006).
- Blake, J.A. et al. The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. *Nucleic Acids Res* **42**, D810-7 (2014).

Bienvenu T, Poirier K, Friocourt G, et al. ARX, a novel Prd-class-homeobox gene highly expressed in the telencephalon, is mutated in X-linked mental retardation. *Hum Mol Genet* **11**(8): 981–991 (2002).

Carranza Rojo D, Hamiwka L, McMahon JM, Dibbens LM. et al. De novo SCN1A mutations in migrating partial seizures of infancy. *Neurology*.**26**;77(4):380-3 (2011).

Carvill, G.L. et al. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat Genet* **45**, 825-30 (2013).

Carvill, G.L. et al. GABRA1 and STXBP1: novel genetic causes of Dravet syndrome. *Neurology*. 82(14):1245-53 (2014).

Claes L. et al. De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. *Am J Hum Genet* **68**:1327–1332(2001).

Claes L, Ceulemans B, Audenaert D, et al. De novo SCN1A mutations are a major cause of severe myoclonic epilepsy of infancy. *Hum Mutat* **J21**(6):615–621 (2003).

Cossette P, Liu L, Brisebois K, Dong H, Lortie A, Vanasse M, Saint-Hilaire JM, Carmant L, Verner A, Lu WY, Wang YT, Rouleau GA. Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. *Nat Genet* ;**31**(2):184-9 (2002).

Davidsson J, Collin A, Olsson ME, Lundgren J, Soller M. Deletion of the SCN gene cluster on 2q24.4 is associated with severe epilepsy: an array-based genotype-phenotype correlation and a comprehensive review of previously published cases. *Epilepsy Res* **81**(1):69–79 (2008).

De Ligt, J. et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* **367**, 1921-9 (2012).

De Kovel CG, Trucks H, Helbig I, Mefford HC, Baker C, Leu C et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain*. 133:23-32 (2010).

Depienne C., et al. Spectrum of SCN1A gene mutations associated with Dravet syndrome: analysis of 333 patients. *J Med Genet* **46**(3):183–191 (2009).

Depienne C, Bouteiller D, Keren B, et al. Sporadic infantile epileptic encephalopathy caused by mutations in PCDH19 resembles Dravet syndrome but mainly affects females. *PLoS Genet* **5**(2):e1000381 (2009).

Depienne C, Trouillard O, Gourfinkel-An I, et al. Mechanism or variable expressivity of inherited SCN1A mutations causing Dravet syndrome. *J Med Genet* **47**:404–410 (2010).

Depienne C, Trouillard O, Gourfinkel-An I, et al. Mechanisms for variable expressivity of inherited SCN1A mutations causing Dravet syndrome. *J Med Genet*. **47** (6):404–410 (2010).

Dibbens LM. et al. X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nat Genet*. **40**(6):782-8 (2008).

- Dorschner MO, Amendola LM, Turner EH, Robertson PD, Shirts BH, Gallego CJ et al. Actionable, incidental findings in 1,000 participants' exomes. *Am J Hum Genet.* **93**: 631-40 (2013).
- Dravet C, Catani C, Bureau M, Roger J. Partial epilepsies in infancy: a study of 40 cases. *Epilepsia.* **30**(6):807–812 (1989).
- Dulac O. Epileptic encephalopathy. *Epilepsia* **42**(s3):23-26 (2001).
- Duru N. et al. Early-onset progressive myoclonic epilepsy with dystonia mapping to 16pter-p13.3. *J. Neurogenet.* **24**: 207-215 (2010).
- Edvardson S. West syndrome caused by ST3Gal-III deficiency. *Epilepsia* **54**: e24-e27, (2013).
- Endele, S. et al. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* **42**, 1021-6 (2010).
- Engel J. ILAE classification of epilepsy syndromes. *Epilepsy Research* **70S**:S5-10 (2006).
- Engel J. Report of the ILAE classification Core Group. *Epilepsia* **47**(9):1558-1568 (2006).
- Epi4K Consortium; Epilepsy Phenome/Genome Project. *Nature.* **501**: 217-21. (2013).
- Escayg, Goldin AL Sodium channel SCN1A and epilepsy: mutations and mechanisms. *Epilepsia.* **51**(9):1650-8 (2010).
- Freilich ER, Jones JM, Gaillard WD, Conry JA, et al. Novel SCN1A mutation in a proband with malignant migrating partial seizures of infancy. *Arch Neurol.* **68**(5):665-71 (2011).
- Fujiwara T. Clinical spectrum of mutations in SCN1A gene: severe myoclonic epilepsy in infancy and related epilepsies. *Epilepsy Res* **70**(Suppl 1):S223–230 (2006).
- Fukuma G, Oguni H, Shirasaka Y, et al. Mutations of neuronal voltage-gated Na⁺ channel alpha 1 subunit gene SCN1A in core severe myoclonic epilepsy in infancy (SMEI) and in borderline SMEI (SMEB). *Epilepsia* **45**(2):140–148 (2004).
- Gambardella A, Marini C. Clinical spectrum of SCN1A mutations. *Epilepsia* **50**(Suppl. 5):20–3. (2009)
- Gecz J, Cloosterman D, Partington M. ARX: a gene for all seasons. *Curr Opin Genet Dev* **16**(3):308–316 (2006).
- Gennaro E, Veggiotti P, Malacarne M, Madia F, Cecconi M, Cardinali S, et al. Familial severe myoclonic epilepsy of infancy: truncation of Nav1.1 and genetic heterogeneity. *Epileptic Disord*; **5**:21–5. (2003).
- Gerber SH, Rah JC, Min SW, et al. Conformational switch of syntaxin-1 controls synaptic vesicle fusion. *Science* **321**(5895):1507–1510 (2008).
- Garofalo S, Cornacchione M, Di Costanzo A. From genetics to genomics of epilepsy. *Neurol Res Int.* **12**:876234 (2012).
- Gonzalez KD, Shahmirzadi L, Chao E, Gandomi S, Parra M, Tippin Davis B et al., *Annual Meeting of the American Society of Human Genetics*, Boston 2013, Poster 1111T.

- Guerrini R. Epilepsy in children. *Lancet* **11**;367(9509):499–524 (2006).
- Guerrini R, Moro F, Kato M, et al. Expansion of the first PolyA tract of ARX causes infantile spasms and status dystonicus. *Neurology* **69**(5):427–433 (2007) .
- Güven, A., Tolun, A. TBC1D24 truncating mutation resulting in severe neurodegeneration. *J. Med. Genet.* **50**: 199-202 (2013).
- Hamdan F. F. et al. De novo STXBP1 mutations in mental retardation and nonsyndromic epilepsy. *Ann. Neurol.* **65**: 748-753 (2009).
- Hamdan F. F. et al. Intellectual disability without epilepsy associated with STXBP1 disruption. *Europ. J. Hum. Genet.* **19**: 607-609 (2011).
- Harkin LA, Bowser DN, Dibbens LM, et al. Truncation of the GABA(A)-receptor gamma2 subunit in a family with generalized epilepsy with febrile seizures plus. *Am J Hum Genet* **70**(2): 530–536 (2002).
- Harkin LA, McMahan JM, Iona X, Dibbens L, Dibbens L, Pelekanos JT, et al. The spectrum of SCN1A-related infantile epileptic encephalopathies. *Brain*, **130**:843–52 (2007).
- Harvey K. et al. The GDP-GTP exchange factor collybistin: as essential determinant of neuronal gephyrin clustering. *J. Neurosci.* **24**: 5816-5826 (2004).
- Hauser WA, Annegers JF, Kurland LT. Incidence of epilepsy and unprovoked seizures in Rochester, Minnesota: 1935-1984. *Epilepsia.* **34**(3):453-68 (1993).
- Hedges DJ, Burges D, Powell E, Almonte C, Huang J, Young S, Boese B, Schmidt M, Pericak-Vance MA, Martin E, Zhang X, Harkins TT, Zuchner S. Exome sequencing of amultigenerational human pedigree. *PLoS One.* **14**;4(12):e8232 (2009).
- Heron S. et al. Missense mutations in the sodium-gated potassium channel gene KCNT1 cause severe autosomal dominant nocturnal frontal lobe epilepsy. *Nature Genet.* **44**: 1188-1190 (2012).
- Hynes K, Tarpey P, Dibbens LM, et al. Epilepsy and mental retardation limited to females with PCDH19 mutations can present de novo or in single generation families. *J Med Genet.* **47**(3):211–216 (2010).
- Holland KD, Hallinan BE. What causes epileptic encephalopathy in infancy?: the answer may lie in our genes. *Neurology.* **28**;75(13):1132-3 (2010).
- Horner DS, Pavesi G, Castrignano T, De Meo PD, Liuni S, Sammeth M, Picardi E, Pesole G. Bioinformatics approaches for genomics and post genomics applications of nextgeneration sequencing. *Brief Bioinform.* **11**(2):181-97 (2010).
- Ishii A. et al. A recurrent KCNT1 mutation in two sporadic cases with malignant migrating partial seizures in infancy. *Gene* **531**: 467-471 (2013).
- Jiang YH, Yuen RK, Jin X, Wang M, Chen N, Wu X et al., *Am J Hum Genet.* 2013; 93: 249-63 (2013).

- Jones MA, Bhide S, Chin E, Ng BG, Rhodenizer D, Zhang VW, Sun JJ, Tanner A, Freeze HH, Hegde MR. Targeted polymerase chain reaction-based enrichment and next generation sequencing for diagnostic testing of congenital disorders of glycosylation. *Genet Med* **13**:921–932. (2011)
- Kalscheuer VM. Et al. Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am J Hum Genet* **72**(6):1401–1411 (2003).
- Kamiya K. A nonsense mutation of the sodium channel gene SCN2A in a patient with intractable epilepsy and mental decline. *J. Neurosci.* **24**: 2690-2698 (2004).
- Kanai K, Hirose S, Oguni H, et al. Effect of localization of missense mutations in SCN1A on epilepsy phenotype severity. *Neurology* **63**(2):329–334 (2004).
- Kato M. et al. A longer polyalanine expansion mutation in the ARX gene causes early infantile epileptic encephalopathy with suppression-burst pattern (Ohtahara syndrome). *Am J Hum Genet* **81**(2):361–366 (2007).
- Kato M. et al. Frameshift mutations of the ARX gene in familial Ohtahara syndrome. *Epilepsia.* **51**(9):1679-84. (2010).
- Kitamura K. et al. Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nat Genet* **32**(3):359–369 (2002).
- Kariminejad R, Lind-Thomsen A, Tümer Z, Erdogan F, Ropers HH, Tommerup N et al. High frequency of rare copy number variants affecting functionally related genes in patients with structural brain malformations. *Hum Mutat.* **32**: 1427-35 (2011).
- Kim YO, Dibbens L, Marini C, Suls A, Chemaly N, Mei D et al. Do mutations in SCN1B cause Fravet syndrome. *Epilepsy Res.* **103**: 97-100 (2013).
- *
- Kjeldsen MJ, Corey LA, Christensen K, Friis ML. Epileptic seizures and syndromes in twins: the importance of genetic factors. *Epilepsy Res.* **55**(1-2):137-46 (2003).
- Klassen T, Davis C, Goldman A, Burgess D, Chen T, Wheeler D, McPherson J, Bourquin T, Lewis L, Villasana D, Morgan M, Muzny D, Gibbs R, Noebels J. Exome sequencing of ion channel genes reveals complex profiles confounding personal risk assessment in epilepsy. *Cell* **145**:1036–1048 (2011).
- Kodera, H. et al. De Novo Mutations in SLC35A2 Encoding a UDP-Galactose Transporter Cause Early-Onset Epileptic Encephalopathy. *Hum Mutat* **34**, 1708-14 (2013).
- Kodera, H. et al. Targeted capture and sequencing for detection of mutations causing early onset epileptic encephalopathy. *Epilepsia* **54**, 1262-9 (2013).
- Krumm N, O'Roak BJ, Shendure J, Eichler EE. A **de novo convergence of autism genetics and molecular neuroscience**. *Trends Neurosci.* **37**(2):95-105 (2014).
- Kurian M. A. et al. Phospholipase C beta 1 deficiency is associated with early-onset epileptic encephalopathy. *Brain* **133**: 2964-2970, 2010.
- Lander ES et al.; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature.* **15**;409(6822):860-921 (2001).

- Lemke JR, Riesch E, Scheurenbrand T, Schubach M, Wilhelm C, Steiner I et al., Targeted next generation sequencing as a diagnostic tool in epileptic disorders. *Epilepsia*. **53**: 1387-98 (2012).
- Lemke JR, Lal D, Reinthaler EM, Steiner I, Nothnagel M, Alber M et al. Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet*. **45**: 1067-72 (2013).
- Lemke, J.R. et al. GRIN2B mutations in west syndrome and intellectual disability with focal epilepsy. *Ann Neurol* (2013).
- Lesca, G. et al. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. *Nat Genet* **45**, 1061-6 (2013).
- Lossin C.A Catalog of SCN1A variants. *Brain Dev*. **31**(2):114-30 (2009).
- Madia F, Striano P, Gennaro E, Malacarne M, Paravidino R, Biancheri R, Budetta M, Cilio MR, Gaggero R, Pierluigi M, Minetti C, Zara F. Cryptic chromosome deletions involving SCN1A in severe myoclonic epilepsy of infancy. *Neurology*. **10**;67(7):1230-5 (2006).
- Maljevic S et al. A mutation in the GABA-A receptor alpha-1-subunit is associated with absence epilepsy. *Ann. Neurol*. **59**: 983-987 (2006).
- Marini C. et al. Idiopathic epilepsies with seizures precipitated by fever and SCN1A abnormalities. *Epilepsia* **48**(9):1678–1685 (2007).
- Marini C, Scheffer IE, Nabbout R, et al. SCN1A duplications and deletions detected in Dravet syndrome: Implications for molecular diagnosis. *Epilepsia*. Mar 11 (2009).
- Mefford HC, Muhle H, Ostertag P, von Spiczak S, Buysse K, Baker C et al. Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. *PLoS Genet*. 20;6(5):e1000962. (2010).
- Mefford HC, Yendle SC, Hsu C, Cook J, Geraghty E, McMahon JM et al., Rare copy number variants are an important cause of epileptic encephalopathies. *Ann Neurol*. 70:974-85 (2011).
- Milh M. et al. A. Novel compound heterozygous mutations in TBC1D24 cause familial malignant migrating partial seizures of infancy. *Hum. Mutat*. **34**: 869-872 (2013).
- Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*. **86**: 749-64 (2010).
- Miura H, Yanazawa M, Kato K, Kitamura K. Expression of a novel aristaless related homeobox gene 'Arx' in the vertebrate telencephalon, diencephalon and floor plate. *Mech Dev* 65(1–2):99– 109 (1997).
- Molinari F, Raas-Rothschild A, Rio M, et al. Impaired mitochondrial glutamate transport in autosomal recessive neonatal myoclonic epilepsy. *Am J Hum Genet* **76**(2):334–339 (2005) .
- Molinari F, Kaminska A, Fiermonte G, et al. Mutations in the mitochondrial glutamate carrier SLC25A22 in neonatal epileptic encephalopathy with suppression bursts. *Clin Genet* **76** (2):188–194 (2009).

- Mulley JC, Scheffer IE, Harkin LA, Berkovic SF, Dibbens LM. Susceptibility genes for complex epilepsy. *Hum Mol Genet.* 14 Spec No. 2:R243-9 (2005).
- Mulley JC, Mefford HC. Epilepsy and the new cytogenetics. *Epilepsia* **52**:423–432 (2011).
- Nabbout R, Dulac O. Epileptic encephalopathies: a brief overview. *J Clin Neurophysiol.* **20**(6):393-7 (2003).
- Nabbout R. et al. Spectrum of SCN1A mutations in severe myoclonic epilepsy of infancy. *Neurology.* **24**;60(12):1961-7 (2003).
- Nakamura, K. et al. De Novo Mutations in GNAO1, Encoding a G α Subunit of Heterotrimeric G Proteins, Cause Epileptic Encephalopathy. *Am J Hum Genet* **93**, 496-505 (2013).
- Nava et al. De novo mutations in HCN1 cause early infantile epileptic encephalopathy. *Nat Genet.* **46**(6):640-5 (2014).
- Neale, B.M. et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242-5 (2012).
- Ogiwara I. et al. De novo mutations of voltage-gated sodium channel alpha-II gene SCN2A in intractable epilepsies. *Neurology* **73**: 1046-1053 (2009).
- Ohmori I. et al. Significant correlation of the SCN1A mutations and severe myoclonic epilepsy in infancy. *Biochem Biophys Res Commun* **295**(1):17– 23 (2002).
- Ottman R, Hirose S, Jain S, Lerche H, Lopes-Cendes I, Noebels JL, Serratosa J, Zara F, Scheffer IE. Genetic testing in the epilepsies--report of the ILAE Genetics Commission. *Epilepsia.* **51**(4):655-70 (2010).
- Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S. & Goldstein, D.B. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* **9**, e1003709 (2013).
- Poduri A. et al. Homozygous PLCB1 deletion associated with malignant migrating partial seizures in infancy. *Epilepsia* **53**: e146-e150 (2012).
- Rauch, A. et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* **380**, 1674-82 (2012).
- Redies C, Vanhalst K, Roy F. delta-Protocadherins: unique structures and functions. *Cell Mol Life Sci* **62**(23):2840–2852 (2005).
- Saito H. et al. De novo mutations in the gene encoding STXBP1 (MUNC18-1) cause early infantile epileptic encephalopathy. *Nat Genet.* **40**(6):782-8 (2008).
- Saito H. et al. Dominant-negative mutations in alpha-II spectrin cause West syndrome with severe cerebral hypomyelination, spastic quadriplegia, and developmental delay. *Am J Hum Genet.* **11**;86(6):881–891 (2010).
- Saito H, Kato M, Okada I, et al. STXBP1 mutations in early infantile epileptic encephalopathy with suppressionburst pattern. *Epilepsia* **51**:2397–2405 (2010).

Saito H, Kato M, Shimono M, et al. Association of genomic deletions in the STXBP1 gene with Ohtahara syndrome. *Clin Genet* **81**:399–402 (2012).

Sanders, S.J. et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–41 (2012).

Scheffer IE, Berkovic SF. Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes. *Brain* **120**(Pt 3):479–490 (1997).

Scheffer IE, Harkin LA, Grinton BE, et al. Temporal lobe epilepsy and GEFS1 phenotypes associated with SCN1B mutations. *Brain* **130**:100–109 (2007).

Scheffer I. et al. Epilepsy and mental retardation limited to females: an under-recognized disorder. *Brain* **131**: 918–927 (2008).

Shen J. et al. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. (Letter) *Nature Genet.* **42**: 245–249 (2010).

Singh R, Andermann E, Whitehouse WP, et al. Severe myoclonic epilepsy of infancy: extended spectrum of GEFS+ *Epilepsia* **42**(7):837–844 (2001).

Shahmirzadi L, Chao EC, Palmaer E, Parra MC, Tang S, Gonzalez KD. Patient decisions for disclosure of secondary findings among the first 200 individuals undergoing clinical diagnostic exome sequencing. *Genet Med.* **16**(5):395–9 (2014).

Steinlein OK. Genes and mutations in human idiopathic epilepsy. *Brain Dev.* **26**(4):213 (2004).

Striano P. et al. Brain MRI findings in severe myoclonic epilepsy in infancy and genotype-phenotype correlations. *Epilepsia.* **48**(6):1092–6 (2007).

Striano P, Coppola A, Paravidino R, Malacarne M, Gimelli S et al. Clinical significance of rare copy number variations in epilepsy: a case-control survey using microarray-based comparative genomic hybridization. *ArchNeurol.* **69**: 322–30 (2012).

Sugawara T, Mazaki-Miyazaki E, Fukushima K, et al. Frequent mutations of SCN1A in severe myoclonic epilepsy in infancy. *Neurology* **58**(7):1122–1124 (2002).

Suls A. et al. Microdeletions involving the SCN1A gene may be common in SCN1A-mutation-negative SMEI patients. *Hum Mutat* **27**(9):914–920 (2006).

Suls A, Velizarova R, Yordanova I, et al. Four generations of epilepsy caused by an inherited microdeletion of the SCN1A gene. *Neurology.* **75**(1):72–76 (2010).

Suls A, Kecskés A, Weber Y, et al. De novo loss-of-function mutations in CHD2 cause a fever-sensitive myoclonic epileptic encephalopathy sharing features with Dravet syndrome. *Am J Hum Genet* **93**:967–975 (2013).

Tao J. et al. Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation. *Am J Hum Genet* **75**(6):1149–1154 (2004).

Tohyama J et al. Early onset West syndrome with cerebral hypomyelination and reduced cerebral white matter. *Brain Dev* **30**(5):349–355 (2008).

- Vanderver A. et al. Identification of a novel de novo p.Phe932Ile KCNT1 mutation in a patient with leukoencephalopathy and severe epilepsy. *Pediat. Neurol.* **50**: 112-114 (2014).
- Veeramah K. R. et al. De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. *Am. J. Hum. Genet.* **90**: 502-512 (2012).
- Veeramah KR, Johnstone L, Karafet TM, Wolf D, Sprissler R, Salogiannis J et al., Exome sequencing reveals new causal mutations in children with epileptic encephalopathies *Epilepsia.* **54**: 1270-81 (2013).
- Zucca C, Redaelli F, Epifanio R, et al. Cryptogenic epileptic syndromes related to SCN1A: twelve novel mutations identified. *Arch Neurol* **65**(4):489–494 (2008).
- Wang JW, Kurahashi H, Ishii A, et al. Microchromosomal deletions involving SCN1A and adjacent genes in severe myoclonic epilepsy in infancy. *Epilepsia.* **49**(9):1528-34 (2008).
- Wallace RH, Scheffer IE, Barnett S, Richards M, Dibbens L et al. Neuronal sodium-channel alpha1 subunit mutations in generalized epilepsy with febrile seizures plus. *Am J Hum Genet.* **68**(4):859-65. 2001
- Weaving LS. Et al. Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am J Hum Genet* **75**(6):1079–1093 (2004).
- Weber YG, Lerche H. Genetic mechanisms in idiopathic epilepsies. *Dev Med Child Neurol* **50**:648–654 (2008).
- Wolverton T, Lalande M. Identification and characterization of three members of a novel subclass of protocadherins. *Genomics* **76**(1–3):66–72 (2001).
- Yu MJ, Shi YW, Gao MM, Deng WY, Liu XR, Chen L, Long YS, Yi YH, Liao WP. Milder phenotype with SCN1A truncation mutation other than SMEI. *Seizure.* **19**(7):443-5 (2010).

Ringraziamenti

Ringrazio di cuore tutti coloro che hanno condiviso con me questi anni di importante crescita professionale e questo progetto di ricerca.

Un ringraziamento al Prof. Minetti, per la disponibilità accordatami, per aver creduto in me e sostenuto la mia crescita professionale.

Un grazie sentito a Francesca e Giovanna, colleghe e amiche ma soprattutto compagne di vita, a Marina, Cristina e Federico, pilastri e riferimenti per la mia professione e maestri di vita.

Un particolare ringraziamento a Pasquale per il prezioso aiuto scientifico, gli impagabili suggerimenti, il costante confronto, la condivisione quotidiana di un percorso professionale e umano.

Infine, grazie a Sandro e alla mia famiglia, ultimi ma non per importanza, per avermi permesso di spiegare le vele e seguire il vento del mio destino.

Stella