UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN AREA DEL FARMACO E TRATTAMENTI INNOVATIVI

CICLO XXX

COORDINATORE Prof.ssa Teodori Elisabetta

"Contribution of whole exome sequencing and bioinformatic analysis to the molecular diagnosis of pediatric epilepsies"

Settore Scientifico Disciplinare MED/39

Dottorando
Dott.ssa Bianchini Claudia
(firma)

Tutore scientifico
Prof. Guerrini Renzo
(firma)

Tutore teorico
Dott.ssa Parrini Elena
(firma)

Coordinatore
Prof.ssa Teodori Elisabetta
(firma)

Anni 2014/2017

Summary

1. Introduction

1.1 Epilepsy

Epilepsy is a common neurological disorder that affects 40-50 million people in the world and it is most common in childhood[1], 70 per 100000 children aged younger than 2 years[2]. Epilepsy is defined by the occurrence of at least two unprovoked seizures occurring more than 24 hours apart.

Pediatric epilepsies are a highly heterogeneous group of neurologic disorders that encompass a number of syndromes with variable severity, ranging from benign with early spontaneous remission to progressive encephalopathy refractory to treatment. Epilepsy syndromes are highly heterogeneous in terms of seizure types, age of onset, clinical features, electroencephalographic expression, and response to treatment.

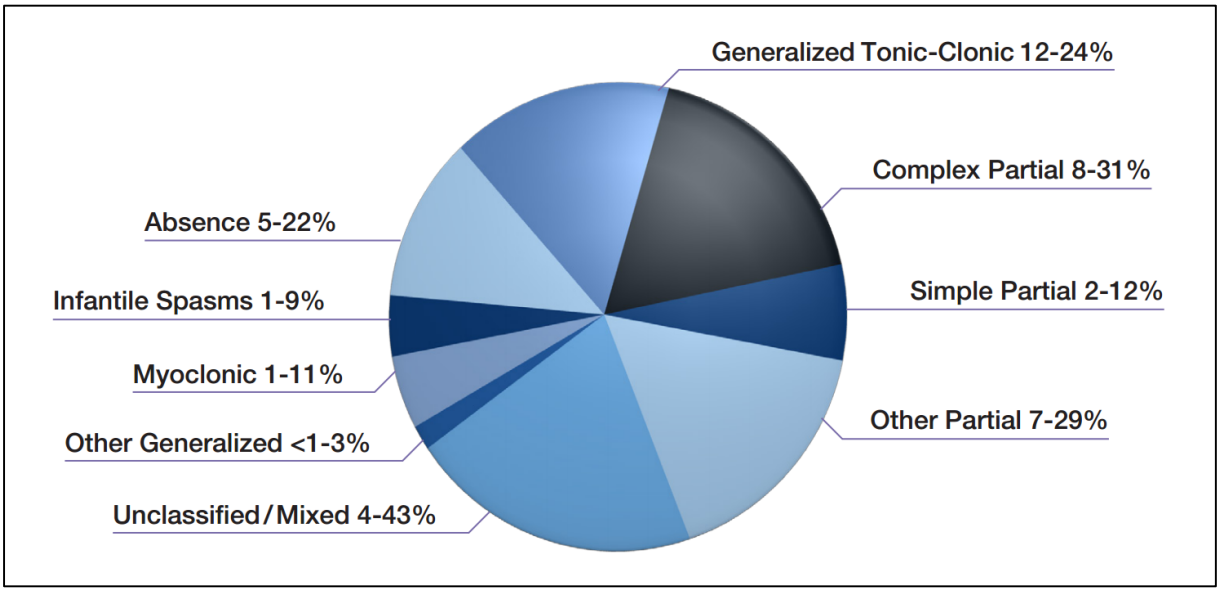

Figure 1. *Types of seizures.*

Seizures are the manifestation of abnormal hypersynchronous discharges of cortical neurons lasting for several seconds, minutes, or longer. The clinical presentation of seizures depends upon both the location of the epileptic discharges in the cortex and the extent and pattern of propagation of the electrical discharges in the brain. In many cases, seizures result in convulsions and loss of consciousness. Seizures may also manifest in other ways that affect personality, mood, memory, sensation, and/or movement. Blank stares, lip smacking, intermittent eye movements, and jerking movements of the extremities are all examples of possible manifestations of seizures. Seizures can be classified into various categories (Figure 1).

Epilepsy may be an isolated neurological symptom, or it may occur as part of a more complex syndrome. There are many different causes of epilepsy, including genetic disorders, metabolic diseases, and structural brain abnormalities. However, in some cases, the cause of epilepsy is not

known. The International League against Epilepsy (ILAE) has described the underlying causes of epilepsy as the following[3]:

Genetic: "The concept of genetic epilepsy is that epilepsy is, as best as understood, the direct result of a known or presumed genetic defect(s) in which seizures are the core symptom of the disorder[3]." The genetic contribution must be evident by extensive and replicable molecular studies or familial studies (e.g., *SCN1A* gene has been demonstrated to be associated with GEFS+[4]).

Structural/metabolic: If there is a specific structural or metabolic disorder that has been proven in various studies to be associated with increased risk of epilepsy, the cause of epilepsy can be defined as structural or metabolic. The cause for such disorder can be acquired (e.g., trauma, stroke, infection), genetic (e.g., malformations of cortical development) or both (e.g., West syndrome).

Unknown cause: This means that the underlying cause is yet unknown. Epilepsy might be presenting because of either some genetic defect or as a result of disorder that is not yet recognized. Epilepsies of unknown cause are classically divided into two main aetiological categories, symptomatic and idiopathic:

- *"idiopathic"* epilepsies are thought to be caused by a genetic predisposition;
- *"presumed symptomatic"* epilepsies are those whose cause is suspected to be induced by a pathology that is below the limit of detection of the available diagnostic tests;
- *"symptomatic"* epilepsies are caused by an obvious brain abnormality, which can in turn be genetically determined or by external factors acting prenatally or after birth[5]. The task is much more complex when the genetic cause is hidden and will only be recognizable or hypothesizable later during the course, when a given electro-clinical pattern or syndrome will become obvious. For this reason, the category of "genetic epilepsy" which has recently been proposed is controversial and, perhaps, somewhat uncertain as what is presumed to be symptomatic today, may become genetic after a molecular screening, or remain as such if screening is unavailable.

Genetic research in epilepsy represents an area of great interest for both clinical purposes and for understanding of the mechanisms underlying epilepsy.

Epilepsy represents a clinical manifestation that is often observed as part of complex neurodevelopmental disorders including intellectual disability and autism[6]. Although ion channel genes represent the gene family most frequently causally related to epilepsy, other genes have gradually been associated with complex developmental epilepsy conditions, revealing the pathogenic role of mutations affecting diverse molecular pathways that regulate membrane excitability, synaptic plasticity, presynaptic neurotransmitter release, postsynaptic receptors, transporters, cell metabolism,

and many formative steps in early brain development, such as the proliferation and migration of neuronal precursors, dendritogenesis, synaptogenesis, cell and glial biology[7].

1.2 Genetic epilepsy syndromes without structural brain abnormalities

The idiopathic epilepsies, including 30% of all epilepsies[8], are characterized by age-related onset, normal neurological and cognitive development, and absence of brain damage. The study of electroencephalogram (EEG) anomalies while awake and during sleep is useful to characterize the different forms. For some forms with Mendelian inheritance the genetic defect has been identified. In a limited number of patients rare forms of idiopathic epilepsies with a dominant pattern of inheritance are observed, which occur with repeated seizures in the neonatal or early infantile period, and with febrile, generalized, or partial seizures that persist into adulthood. They are of great interest as they are caused by alterations in genes that encode voltage-gated ion channel subunits or receptor subunits. The identified mutations are located in the neuronal nicotinic acetylcholine receptor in familial forms of frontal epilepsy, the K^+ channels in benign familial neonatal seizures, the Na^+ channels in a particular form of generalized epilepsy with febrile seizures, and the gamma aminobutyric acid (GABA) receptor in a variant of juvenile myoclonic epilepsy[9]. However, for the most common forms of idiopathic epilepsy, that do not follow a simple Mendelian pattern of inheritance, the molecular bases have not yet been defined. It is possible that different mutations can cause similar phenotypes in different families or in patients from different geographical areas. The phenotypic variability in some families has been ascribed to genetic modifiers (polymorphisms) or environmental factors that influence phenotypic expression[10].

1.2.1 Severe myoclonic epilepsy of infancy or Dravet syndrome

Dravet syndrome (DS), otherwise known as severe myoclonic epilepsy of infancy (SMEI) [MIM 607208], is an epileptic encephalopathy presenting in the first year of life[11]. DS starts at about 6 months of age in previously healthy infants, typically with prolonged generalized or hemiclonic febrile seizures. Between 1 and 4 years, other types of seizures appear, including myoclonic, partial, and absence seizures. Hyperthermia, such as fever or a warm bath, often precipitates seizures[12]. Development in the first year of life is normal but subsequently slows and may regress. The EEG may be normal until age 2 years when generalized spike wave activity is seen; approximately 10% of patients are photosensitive. The magnetic resonance imaging (MRI) scan is either normal or shows non-specific features[13]. Therefore, Dravet syndrome is considered an EE, that is a disorder in which epileptic seizures and epileptiform electrical activity impair brain function, although this causal link has not yet been clearly demonstrated. The discovery of mutations in the neuronal sodium channel

α1 subunit gene (*SCN1A*) on chromosome 2q in DS[14] prompted deeper molecular analysis of this severe epilepsy syndrome. Neuronal Na+ channels are essential for excitability, and consist of a principal α subunit and accessory β subunits. The frequency of *SCN1A* mutations in DS is approximately 85%[15]. The majority of patients exhibiting mutations of this gene carry *de novo* mutations (90%); about 40% of these mutations are truncation and 40% are missense mutations. 10% of patients who are negative on sequence-based mutational analysis, have copy number variations including exonic deletions or duplications that can involve several exons or the whole gene[16]. Some rare patients have a mutation in the *GABRG2* gene[17] or in *SCN1B* sodium channel β1 subunit[18]. Germline mosaicism may result in siblings with Dravet syndrome born from an unaffected or mildly affected parent carrying a low level of mosaicism for the mutation[19,20]. *SCN1A* mutations are also commonly found in the border line variant of severe myoclonic epilepsy of infancy (SMEI), whose separation from Dravet syndrome may be arbitrary[21]. Mutations are less commonly found in patients that have been categorized within different subgroups exhibiting various elements of Dravet syndrome[11,22,23].

In vitro functional studies of many missense *SCN1A* mutations in human embryonic kidney cells or Xenopus oocytes has revealed both gain- and loss-of-function mechanisms, but loss-of function without negative dominance seems to be the predominant mechanism of action of both truncations and missense mutations causing FS and GEFS+ , which is in agreement with genetic and functional studies in Dravet syndrome[24]. Studies on animal model suggest that Dravet syndrome, and maybe the other *SCN1A* linked seizure disorders, are caused by a decreased excitability of GABAergic interneurons owing to *SCN1A* haploinsufficiency[25,26]. GABA receptors were long suspected to be involved in epileptogenesis. Functional expression of some *GABRG2* (Gamma-aminobutyric Acid Receptor Type A gamma 2 Subunit) mutations, identified in patients with GEFS+, revealed a pronounced loss-of-function by altered gating or defective trafficking and reduced surface expression as a common pathogenic mechanism[24]. Hence, these mutations reduce the main mechanism for neuronal inhibition in the brain, which can explain the occurrence of seizures.

Recent functional studies carried out using human neurons differentiated form induced pluripotent stem cells have generated controversial results[27–29], similarly to the first functional studies in transfected cells. Phenotypes with some clinical features resembling Dravet syndrome have also been associated with protocadherin 19 (*PCDH19*) mutations in a few female patients. See the description of the spectrum of *PCDH19*-related epilepsies in the paragraph below.

1.2.2 Generalized epilepsy with febrile seizures plus

Generalized epilepsy with febrile seizures plus (GEFS+) is a familial epilepsy syndrome. It describes families who have at least two individuals from different generations with phenotypic heterogeneity including: febrile seizures (FS), usually tonic-clonic associated, with a high temperature nearly always happen during a period of illness, such as a viral infection, and febrile seizures plus (FS+), this is similar to febrile seizures, but the child has seizures beyond the normal age range. Overlap with classical idiopathic generalized epilepsy (IGE) is also seen. The course and response to antiepileptic drugs may be considerably variable within the same family. GEFS+ was originally recognized because of remarkable large autosomal dominant pedigrees with 60-70% penetrance. It is likely that most cases, however, occur in small families or are sporadic[30]. Mutations in genes that encode alpha and beta subunits of the voltage-gated Na^+ channel (*SCN1A* and *SCN1B*) have been associated with GEFS+[31]. In about 10% of families are found missense mutations in *SCN1A, resulting* the commonest identified molecular abnormalities[32]. In some families, mutations in the gene encoding the gamma 2 subunit of the $GABA_A$ receptor (*GABRG2*) have been identified[33]. The phenotypic variability observed in GEFS+ could be linked to the combined action of mutations in different genes. Some *SCN1A* mutations can cause phenotypes extending from different types of epilepsy to familial hemiplegic migraine[34]. It has also been shown that a *SCN1A* missense mutations, identified in families with extreme phenotypes, including Dravet syndrome, cause loss of function because of folding defects, which can be rescued by molecular interactions with associated proteins or pharmacological chaperones[35]. These results have been confirmed and extended to typical GEFS+ families and Dravet syndrome patients with *de novo* mutations[36]. This mechanism may generate phenotypic variability and also possibly be used in the development of therapeutic approaches. Animal models have confirmed that GEFS+ mutations cause loss of function of *SCN1A* and reduced excitability of GABAergic neurons[37]. Functional expression of some *GABRG2* mutations, identified in patients with GEFS+ revealed a pronounced loss of function by altered gating or defective trafficking, and reduced surface expression as a common pathogenic mechanism[38]. Hence, these mutations reduce the main mechanism for neuronal inhibition in the brain, similarly to *SCN1A* mutations, which can explain the occurrence of seizures.

1.2.3 Benign epilepsies of the first year of life

The clinical manifestations, which occur in asymptomatic babies, regress and eventually disappear spontaneously in benign epilepsies of the first year of life. The regression or the disappearance of symptoms is the reason because this group of syndromes are defined as “benign”. These forms are

transmitted with an autosomal dominant pattern of inheritance and are quite rare. Molecular diagnosis, where possible, is important in order to avoid unnecessary invasive testing and support genetic counselling. The clinical manifestations of these forms of epilepsy are rather similar, the age of onset of these forms of epilepsy is variable but the clinical manifestations are quite similar.

Benign familial neonatal seizures (BFNS)

Recurrent seizures appear from the first days of life up to the 3rd month and disappear spontaneously after weeks to months. The seizures can be focal, when involve only one side of the brain, or generalized when involve both sides.

Many infants have generalized tonic-clonic seizures that involve both sides of the brain and affects the entire body, causing muscle rigidity, convulsions, and loss of consciousness. EEG is usually normal. The risk of seizures recurring later in life is about 15%[39]. An increasing number of cases with learning disability have recently been described[40], but usually psychomotor development is normal.

The inheritance model that characterized BFNS is autosomal dominant with a penetrance of 85%. The most mutated gene in patients with BFNS is *KCNQ2*, encoding the K^+ channel voltage-dependent, KQT-like subtype member, and the mutations are mostly deletions/duplications[41].

In literature a small proportion of BFNS families carrying mutations in *KCNQ3* that encodes for voltage-dependent K^+ channel, KQT-like subtype, member 3, have been also described.

Both genes form a heteromeric K^+ channel, which determines the M-current, influencing the membrane potential at rest[42].

The reduction of about 20-30% in the resulting potassium current due to co-expression of the heteromeric wild-type and mutant KCNQ2/3 channels is apparently sufficient to cause BFNS[43].

It is not still fully understood why seizures preferentially occur in neonates[44], although the reduction of the potassium current can cause epileptic seizures by a subthreshold membrane depolarization, which increases neuronal firing. Probably this occurs because or the neonatal brain could be more vulnerable to changes, even small, of neuronal excitability; or KCNQ2 and KCNQ3 channels, when mutated, are replaced by other K^+ channels that become functional after the first months of life.

Benign familial neonatal-infantile seizures (BFNIS)

Seizures observed are similar to seizures observed in children with BFNS. The mean onset age of seizures is 3 months and it ranges from the neonatal period to infancy. Remission occurs by 12 months with a very low risk of later seizures[45,46]. *SCN2A* is the most mutated gene in families with BFNS[47].

Benign familial infantile seizures (BFIS)

The seizures are characterized by onset age of seizures around 6 months[39] but are very similar to seizure observed in BNFS , and in some children in later childhood appear paroxysmal dyskinesia (infantile convulsions and choreoathetosis)[48].

Mutations of the *PRRT2* gene, in the pericentromeric region of chromosome 16, have been associated with familial infantile convulsions, paroxysmal kinesigenic or exercise-induced dyskinesia, migraine, or hemiplegic migraine, or in various combinations[49]. *PRRT2* mutations appear to have a high prevalence in patients with benign familial infantile seizures (or epilepsy), either alone or in association with paroxysmal kinesigenic orexercise-induced dyskinesia (from 40% to 90%), but are only rarely detected in sporadic cases and only rarely cause other types of epilepsy.

A recurrent and frequent frameshift mutation (c.649_650insC, p.R217Pfs*8), found in a large number of patients from different ethnic backgrounds, as well as a considerable number of loss-of-function and missense amino acid-changing mutations, were identified in *PRRT2*. The location and type of mutation within *PRRT2* do not appear to predict the clinical phenotype[50].

Confirmation of *PRRT2* mutations in infants with infantile seizures, paroxysmal dyskinesia, or both, provides families and clinicians with reassurance that seizures are likely to be self-limited, with an excellent prognosis[50]. Linkage studies have identified another locus on chromosome 19[48,51]. Several families with mutations in the *ATP1A2* gene have benign infantile seizures in conjunction with hemiplegic migraine[52]. Rare families with mutations in *SCN2A* or *KCNQ2* have also been described in which only infantile seizures occur[53,54].

1.2.4 Autosomal dominant nocturnal frontal lobe epilepsy

Autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) includes frequent, usually brief, seizures occurring from childhood (average seizures onset around 10 years of age), with hyperkineticor tonic manifestations, typically in clusters at night during slow-wave sleep. The phenotype include also paroxysmal arousals, dystonia-like attacks, and epileptic nocturnal wanderings[55]. Penetrance is estimated at approximately 70-80%. Ictal video-EEG studies suggested a defect of a broader network reveling partial seizures originating from the frontal lobe but also from parts of the insula and temporal lobe[56,57].

This form of epilepsy is characterized by mutations in the *CHRNA4* gene encoding the a4-subunit of a neuronal nicotinic acetylcholine receptor[58], the *CHRNB2* gene encoding the beta2-subunit of neuronal nicotinic acetylcholine receptor, and the *CHRNA2* gene, encoding the neuronal nicotinic acetylcholine receptor alpha2-subunit[4,59–61].

These receptors are hetero pentamers consisting of various combinations of subunits. The alpha4-beta2 combination is the most represented in the thalamus and cerebral cortex. All the identified mutations are located in the pore-forming M2 transmembrane segments. Functional studies of nAchR have produced controversial results, which makes the underlying pathogenic mechanisms unclear.
Heron et al. reported 4 unrelated families with a severe form of ADNFLE, in which patients had an earlier mean age of onset (6 years) than other ADNFLE forms, and frequently showed psychiatric features and intellectual disability[62]. They carried missense mutations in the sodium-gated potassium channel gene *KCNT1*, but no functional analysis was performed.

1.2.5 Autosomal dominant temporal lobe epilepsy

Autosomal dominant temporal lobe epilepsy (ADTLE) is a form of autosomal dominant partial epilepsy characterized by focal seizures with prominent auditory symptoms and audiogenic seizures. The first clinical manifestations, usually occurring during childhood or adolescence, are auditory hallucinations, sometimes accompanied by vision or olfactory manifestations, or dizziness[63,64]. A pathogenic variant in the *LGI1* gene encoding a secreted leucine-rich protein that is expressed in brain and plays a role in regulating postnatal glutamatergic synapse development[65], has been identified in approximately one third of families with ADTLE[66,67]. The *LGI1* pathogenetic mechanism remains to be clarified. Functional inactivation of one allele leads to ADTLE, whereas silencing of both alleles has been observed in several high-gradegliomas[68].
A pathogenic variant in *RELN*, encodes a secreted protein, Reelin, which has important functions in both the developing and adult brain and is also found in the blood serum, has been identified also in several affected families[69]. ADLTE-related mutations significantly decrease serum levels of Reelin, suggesting an inhibitory effect of mutations on protein secretion[69].
Data obtained from animal models are more consistent with a direct presynaptic epileptogenic mechanism, whereas postsynaptic effects may be more involved in glutamatergic synapse maturation and dendritic pruning[70,71].

1.2.6 Familial Focal Epilepsy with Variable Foci

Seizures may be generated in different cortical regions (frontal, temporal, frontotemporal, parietal, or occipital) and the age of onset varies from infancy to adulthood in affected member within the same family. Affected individuals do not show cognitive deficits, although some family members exhibit intellectual disability.

Recent studies have identified in some of these families mutations in Dishevelled, Egl-10 and Pleckstrin domain-containing protein (*DEPDC5*), the function of which is still unclear, but it might be involved in membrane trafficking, G protein signaling and/or modulation of the mammalian target of rapamycin complex[72–74]. Most mutations resulted in a truncated protein and are consistent with loss of function.

1.2.7 Febrile seizures

Febrile seizures (FS) affect 2-4% of children between the ages of 3 months and 6 years[75]. The differential diagnosis includes seizures with fever in a child with epilepsy, seizures with central nervous system infection, or an acute metabolic disturbance[75]. FS are classified as a nosological entity distinct from epilepsy. The recurrence risk ratios is of 3-5 in first-degree relatives[76] where positive family history is one of the few known risk factors. About 3-6% of children with FS will have epilepsy later in life, including idiopathic generalized epilepsy and temporal lobe epilepsy with hippocampal sclerosis that might be related to prolonged FS[77]. Segregation analysis of a population-based sample showed that most FS have polygenic inheritance, although a small proportion have autosomal-dominant inheritance[78]. Using large, multigeneration families with presumed autosomal dominant inheritance, four loci for FS have been identified (FEB1-FEB4), but no gene defect has yet been identified[79].

1.2.8 Single gene severe epilepsies and epileptic encephalopathies

With the molecular revolution a genetic basis is also being recognized for an increasing number of epileptic encephalopathies of infancy and childhood. Epileptic encephalopathies comprise a heterogeneous group of severe epilepsies characterized by several seizure types, frequent epileptiform activity on EEG, and developmental slowing or regression that are often having early seizure onset and developmental delay, but diverse, sometime distinctive, phenotypes. The number of known monogenic determinants underlying epileptic encephalopathies has grown rapidly. *De novo* dominant mutations are frequently identified; somatic mosaicism and recessive disorders are also seen. Several genes can cause one electroclinical syndrome, and, conversely, one gene might be associated with phenotypic pleiotropy.

In this section some examples of mutated genes identified in patients with particular forms of severe epilepsy are described.

***PCDH19* (protocadherin 19)**

Protocadherin 19 is a 1148 amino acid transmembrane protein belonging to the protocadherin delta2 subclass of the cadherin superfamily, which is highly expressed in neural tissues and at different developmental stages[80]. The precise functions of the protein remain so far unknown.

PCDH19 mutations were first reported to cause epilepsy and mental retardation limited to females (EFMR)[80], an X-linked disorder surprisingly affecting only females and sparing transmitting males, but also often appearing *de novo*[81]. *PCDH19* associated phenotype is characterized by a variable clinical presentation, onset of seizures is between 6 and 36 months, variable degree of psychomotor delay and cognitive impairment, ranging from mild to severe mental retardation[80,82]. The epilepsy spectrum associated with *PCDH19* mutations is in turn variable and includes mild focal epilepsy starting in infancy or epilepsy with recurrent episodes of focal or generalized seizures in series or status epilepticus triggered by fever[81].

Males which are hemizygous for *PCDH19* mutations have a normal cognitive level and no epilepsy. Cellular interference has been proposed to explain the discrepancy between the clinical manifestations of heterozygous females and hemizygous males[83]. This model suggests that if an individual has an admixture of two populations of protocadherin cells (mutated and non-mutated) then a pathological phenotype occurs. According this model, a normal female or a transmitting male have only one protocadherin population of cells, protocadherin wildtype or protocadherin-mutant cells, respectively, and therefore they do not present a pathological phenotype. The development of genetically modified animal models will allow this puzzling disease mechanism to be better explored.

***ARX* (Aristaless-Related Homeobox)**

The *ARX* gene, on chromosome Xp22, is a transcription factor that encodes the Aristaless-related homeobox protein, which belongs to the Aristaless-related subset of the paired (Prd) class of homeodomain proteins. *ARX* acts as both a transcriptional repressor and activator in an incompletely characterized biochemical cascade and plays an important role in embryogenesis, especially in the development of the central nervous system[84].

ARX mutations have been identified in about 10 different clinical conditions, with or without brain malformations[85–87]. A phenotype-genotype correlation has been proposed, malformation phenotypes, including X-linked lissencephaly with abnormal genitalia(XLAG), X-linked lissencephaly with abnormal genitalia with severe hydrocephalus, and Proud syndrome (agenesis of the corpus callosum with abnormal genitalia) are associated with protein truncation mutations and missense mutations in the homeobox[85,88]. On the other hand, non malformation phenotypes, including X-linked infantile spasms (ISSX)/West syndrome, Partington syndrome (mental retardation with mild distal dystonia), and non-specific X-linked mental retardation (XLMR)[85,88] are known to be associated with missense mutations outside of the homeobox or expansion of the second polyA tract.

Expansions in the first polyA tract are known to cause ISSX/West syndrome (tonic spasms with clustering, severe psychomotor delay, and hypsarrhythmia on EEG)[89,90], a severe epileptic-dyskinetic encephalopathy[86], tonic seizures and dystonia without infantile spasms[91] and Ohtahara syndrome[92].

***CDKL5* (cyclin-dependent, kinase-like 5)**

CDKL5 gene, on chromosome Xp22.13, is a member of Ser/Thr protein kinase family and encodes a phosphorylated protein with protein kinase activity.

Mutations in this gene cause early onset intractable seizures, severe developmental delay and, often, subsequent appearance of Rett syndrome (RTT) like features[93]. The phenotypic spectrum associated with *CDKL5* mutations also includes X-linked infantile spasms syndrome (ISSX)[94], a form of myoclonic encephalopathy[95], and severe encephalopathy with refractory seizures[96]. Mutations in *CDKL5* are mainly found in females, suggesting gestational lethality in males, although in some papers males harboring mutations in *CDKL5* have been described[97]. Parental germline mosaicism has been reported[98]. In affected females, a seemingly normal early development, followed by onset of intractable seizures between the first days and 4th month of life are the early key diagnostic criteria. Severe developmental delay with regression become apparent after seizures onset. Seizures are usually manifested as infantile spasms, or prolonged tonic seizures followed by spasms and myoclonus, with a peculiar electro-clinical pattern[99] variably associated with migrating focal seizures during the course. Overall, 16.3% of girls with early-onset intractable epilepsy, with or without infantile spasms, exhibit either mutations or genomic deletions involving *CDKL5*[100].

***STXBP1* (syntaxin binding protein 1)**

STXBP1 gene encodes a syntaxin-binding protein that appears to play a role in release of neurotransmitters via regulation of syntaxin, a transmembrane attachment protein receptor[101].

Mutations in this gene have been associated with Ohtahara syndrome or early infantile epileptic encephalopathy (EIEE), characterized by early onset of tonic spasms, seizure intractability, a characteristic suppression-burst pattern on the EEG and poor outcome with severe psychomotor retardation[102,103]. *STXBP1* mutations have also been found in some children with infantile spasms[104,105]. The transition from EIEE to West syndrome occurs in 75% of individuals with EIEE[102,103].

1.2.9 Epilepsy Aphasia Syndromes

Epilepsy aphasia syndromes including Landau-Kleffner syndrome and continuous spike-waves in slow sleep are a group of rare, severe EEs with a characteristic EEG pattern and developmental regression affecting language, in particular.

The lack of evident brain magnetic resonance imaging abnormalities, in addition to a frequent positive family history of epilepsy or EEG abnormalities, has always raised the suspicion of a genetic etiology. Mutations in the *GRIN2A* gene encoding the α2 subunit of the N-methyl-D-aspartate glutamate receptor (also known as GluNR2A) have been identified in 5% of patients with centro-temporal spikes, and 20% of nonlesional continuous spike-waves in slow sleep and Landau-Kleffner syndrome; functional analysis using in vitro systems has shown a gain of function mechanism[106–108]. Mutations in the *GRIN2B* gene, encoding the β2 subunit of the N-methyl-D-aspartate glutamate receptor, have also been recently identified. The *GRIN2B* associated phenotypes include West syndrome and focal epilepsy with intellectual disability, and appear to be very different from that due to *GRIN2A* mutations[109] .

1.2.10 Progressive myoclonus epilepsies

Progressive myoclonus epilepsies (PMEs) represent less than 1% of all epilepsy cases and have a variable geographic and ethnic distribution.

The age onset, the rapid progression of the symptoms, and the prognosis are different and depend on the specific aetiology and on the type of causative mutations.

These genetically heterogeneous disorders share clinical features that include action myoclonus, myoclonic jerks that are segmental and arrhythmic, appearing both at rest and as stimulus sensitive manifestations, epileptic seizures, predominantly generalized tonic–clonic. Progressive mental deterioration, cerebellar and extrapyramidal signs are also present in a variable proportion of patients[110]. Most PMEs result from the intracellular accumulation of abnormal deposit material[111].

Many gene defects underlying PMEs have been identified such as CSTB (Unverricht-Lundborg disease), EPM2A, NHLRC1 (Lafora disease), and CLN (3,4,5,6,8,9) gene family (Neuronal ceroid-lipofuscinoses), but in a considerable proportion of patients the cause remains unknown.

All forms progress relentlessly, leading to severe disability or death in matter of years, with the exception of the Unverricht-Lundborg disease, although disabling, can reach a degree of stabilization.

1.2.11 Genetic Generalized Epilepsy (GGE) with Complex Inheritance

Mode of inheritance

20-30% of all epilepsies are represented by genetic generalized epilepsies (GGEs) and are characterized by absence seizures, myoclonus, and generalized tonic-clonic seizures.

There is a partial overlap in age of onset, type, and frequency of seizures, prognosis, and response to treatment. In the GGEs are included several sub-syndromes such as: childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE), juvenile myoclonic epilepsy (JME), and epilepsy with generalized tonic–clonic seizures alone. Generalized spike-wave activity and normal background are observed in the EEG.

Among epilepsies with complex inheritance, GGEs seem very suitable for genetic studies because they are common, have a relatively well-defined phenotype and occur in familial clusters.

Close relatives of GGE probands have 4-10% risk of developing epilepsy[112]. Higher risk is seen in sibling and offspring, and is lower in second-degree relatives. In contrast to monogenic inheritance, polygenic inheritance leads to a more rapid decrease of the risk in relatives as the distance from the affected individuals increases[113]. GGEs have a genetic aetiology with complex inheritance. It has been suggested that they might result from the interaction of two or more genes[114]. However a high degree of complexity is operating as large-scale exome sequencing of ion channels reveals that rare missense variation in known Mendelian disease genes are equally prevalent in healthy individuals and in those with idiopathic generalized epilepsy, revealing that even deleterious ion channel mutations confer an uncertain risk to an individual depending on the other variants with which they are combined[115].

Molecular basis

Linkage studies on a large number of families with GGE have identified several susceptibility loci (18q, 2q, 3q, and 14q)[116]. In rare families, pathogenic mutations in single genes have been reported such as: mutations in the *GABRG2* gene in families with febrile seizures and childhood absence epilepsy[117], mutations in the *GABRA1* gene in family with dominantly inherited juvenile myoclonic epilepsy[118], and mutations in *CLCN2* in families with heterogeneous GGE phenotypes, including childhood absences epilepsy[119]. Rare variants in *CACNA1H* have been identified in childhood absence epilepsy and other generalized epilepsy phenotypes[120]. Finally, other susceptibility genes for GGE are *GABRD*[121], *ME2*[122], *BRD2*[123], and *NEDD4L*[124].

1.2.12 Phenotypic and Genetic heterogeneity

In epilepsy phenotypic heterogeneity or pleiotropy is increasingly recognized, mutations in a single gene cause different phenotypes (Table 1).

Epilepsy syndromes	**Onset age**	**Gene**				
		>= 50% of cases	**10–50% of case**	**5–10% of cases**	**< 5% of cases**	**unknown % of cases**
Early myoclonic encephalopathy	0-3 months					*PIGA, SETBP1, SIK1, SLC25A22*
Early-onset epileptic encephalopathy	0-3 months		*KCNQ2, CDKL5, SCN2A, STXBP1*		*GNAO1*	*AARS, CACNA2D2, NECAP1, PIGA, QARS, SCN8A, ARX, DOCK7, SLC25A22, SLC35A2, WWOX*
Early infantile epileptic encephalopathy (Ohtahara syndrome)	0-3 months		*CDKL5, SCN2A, STXBP1*		*GNAO1, KCNQ2*	*KCNT1, PIGQ, ARX, DOCK7, SLC25A22, SLC35A2, WWOX*
Dravet syndrome	6-12 months	*SCN1A*				*GABRA1, GABRG2, HCN1, KCNA2, SCN1B, STXBP1*
Infantile spasms	3-12 months		*CDKL5, SCN2A, STXBP*		*GNAO1*	*ALG13, DNM1, FOXG1 duplications, GABRA1, GABRB3, GRIN1, GRIN2A, GRIN2B, IQSEC2, KCNT1, MAGI2, MEF2C, NEDDL4, NDP, NRXN1, PIGA, PLCB1, PTEN, SCA2, SCN1A, SCN2A, SCN8A, SETBP1, SIK1, SLC25A22, SLC35A2, SPTAN1, ST3GAL3, STXBP1, TBC1D24, TCF4, ARX, DOCK7, SLC25A22,*

						SLC35A2, WWOX,
Epilepsy of infancy with migrating focal seizures	0-6 months	*KCNT1*	*SCN2A, SCN1A*			*PLCB1, QARS, SCN8A, SLC25A22, TBC1D24, SLC12A5*
Other predominantly myoclonic epilepsies	0–1 years					*EEF1A2, MEF2C, SCN1A, SLC2A1, SPTAN1, SYNGAP1, TBC1D24*
	>1 year					*CHD2, MEF2C, SYNGAP1, UBE3A*
Other predominantly focal or multifocal epilepsies	0–6 months					*ARHGEF9, DEPDC5, SCN1A, TBC1D24, PNKP, SLC2A1*
	6–12 months					*ARHGEF9, DEPDC5, FOXG, MBD5, PIGO, SLC13A5*
	>1 year					*ARHGEF9, DEPDC5, MBD5, PCDH19, POLG, TNK2, ZEB*
Epilepsy with myoclonic–atonic seizures	>1 year			*SLC2A1*	*SLC6A1*	*GABRA1, GABRG2, SCN1A, SCN1B*
Lennox-Gastaut syndrome	>1 year					*ALG13, DNM1, FLNA, GABRB3, GLI3, HNRNPU, SCN1A, SCN2A, SCN8A, STXBP1*
Epilepsy–aphasia spectrum	>1 year		*GRIN2A*			

Table 1. *Genetic causes of epilepsy syndromes. Only non-chromosomal, non-malformative, and non-metabolic disorders are included. Only genes with more than one case reported are included*[125].

Many factors contribute to phenotypic heterogeneity, including the following:

- type and timing of mutations during development;
- timing and location of physiological gene expression;
- epigenetic factors;
- modifier genes[126–128].

The epilepsy syndromes associated with a gene might range from a benign seizure disorder to an epileptic encephalopathy.

For example, *KCNQ2* mutations cause the benign familial neonatal epilepsy syndrome[129]. This disorder results from autosomal dominant inheritance with high penetrance. *KCNQ2* mutations are also associated with a severe neonatal onset epileptic encephalopathy characterized by tonic seizures and profound developmental impairment, sometimes presenting as Ohtahara syndrome[130]. These patients often have *de novo KCNQ2* mutations, which might be dominant negative, with far more profound negative effect on channel function than those that cause benign familial neonatal epilepsy[131].

The severity of the epilepsy syndrome not only might differ with mutations in a specific gene, but the nature of the syndrome might be surprisingly different.

Mutations in *KCNT1* gene, that encode for sodium activated potassium channel, cause one of the most severe epileptic encephalopathies, epilepsy of infancy with migrating focal seizures. *De novo* mutations in *KCNT1* are observed in 50% of cases and leading to a three times gain in channel function[132]. *KCNT1* mutations were identified also in severe autosomal dominant nocturnal frontal lobe epilepsy[133] and rarely observed also in case of Ohtahara syndrome and West syndrome[134]. Functional studies show a possible phenotype-genotype association, with a 13 times increase in channel amplitude in Ohtahara syndrome[135]. Although autosomal dominant nocturnal frontal lobe epilepsy and epilepsy of infancy with migrating focal seizures share focal seizures and a genetic aetiology, their electroclinical pattern is quite distinct.

Genetic heterogeneity occurs in every epilepsy syndrome. Even in the prototypical genetic epileptic encephalopathy, Dravet syndrome, in which more than 80% of patients have a *SCN1A* mutation, other genes (eg, *STXBP1* and *GABRA1*) account for a small proportion of cases[136].

Analysis of larger numbers of genetically homogeneous cases could demonstrate clinical features that distinguish the phenotype. For example, epilepsy with myoclonic-atonic seizures, described by Doose[137], is associated with mutations in *CHD2* or *SLC2A1* in a small proportion (4%) of cases; *CHD2* is associated with clinical photosensitivity and *SLC2A1* is associated with paroxysmal exercise-induced dyskinesia[138]. Each genetic entity might show subtly different phenotypic features that help diagnosis and, in turn, might have treatment implications, such as the ketogenic diet in glucose transporter 1 deficiency due to mutations in *SLC2A1*.

The knowledge of genetic causes of epilepsy is likely to further expand, since WES and multigene panels are applied to epilepsy syndromes and identification of causative mutations becomes more straight forward from a bioinformatic perspective.

1.3 Next Generation Sequencing

The sequencing revolution began right about 10 years ago with the invention of massive parallel sequencing techniques[139]. When only a few genes, mostly for ion channels, had been causally associated with epilepsy in large, multigenerational pedigrees, and the overall feeling was that the Mendelian epilepsy syndrome was the consequence of mutations affecting one gene alone and very few genes were involved.

In the Next Generation Sequencing (NGS) technologies, fragmented DNA is amplified and sequenced in parallel, then aligned to the reference genome and evaluated for nucleotide changes and small insertions/deletions by bioinformatics software. The first human genome sequencing published in 2001 took 15 years to complete and cost approximately $3 billion[140]. The new Illumina HiSeq X Ten System can sequence 40 to 50 whole genomes per day at a cost of around $1000 per genome with an average coverage depth of approximately 30X[141].

DNA sequencing capabilities have been increasing at an outstanding rate, with the data output of NGS more than doubling each year. NGS sequencing experiments produce millions to billions of short sequence “reads” at a high speed generating a huge amounts of raw data. For example, the raw data produced by Illumina HiSeq2500 platform add up to 1TB per run, covering 150-180 human whole exome sequencing (WES) samples at a depth of 50x or higher (Illumina Inc.). The increasing of data production has shifted the bottleneck of NGS from producing sequence data to data management, analysis, and summarization.

Three basic approaches have been used to study the role of rare variation in mendalian disorders: trio-based studies, case-control studies, and familial segregation studies.

Trio-based studies: To identify *de novo* mutations in patients, the patient and both unaffected parents (referred to as a trio) are sequenced to identify genomic sites where the patient, or "proband," has an allele that is not observed in either parent.

Trio-based assessments of inherited mutations: The same exome sequence data used to identify *de novo* mutations can, of course, also be used to study inherited variants. Because trios are ascertained on the basis of the affected proband, mutations that increase disease risk will be preferentially transmitted.

Case-Control Studies: Rare variation from exome and genome sequencing can be studied using a case-control study design that looks for enrichment at the variant, gene, or gene-set level. There are generally insufficient copies of the rare alleles to have sufficient power to test a single variant. Consequently, rare variants must be grouped together across a gene, or sets of biologically linked genes, for case-control association testing. The most straightforward way to study rare variants collectively is a collapsing analysis, whereby the total number of rare alleles in cases for a particular gene or pathway is tested against the number of equivalent alleles in controls[142].

Many patients, who receive conventional clinical evaluation and targeted genetic testing performed with Sanger sequencing, remain without specific diagnosis even after extensive workup. This has serious consequences for patients and their families, preventing the access to the right treatment or accurate counseling for pregnancies and prognosis.

Not surprisingly, NGS has led to a large number of genetic discoveries and increase the knowledge of epilepsy genetics[143]. From 2007 to 2013, OMIM entries for new Mendelian diseases increased by 87%, with many of those discoveries associated with neurologic disorders[144].

Next-generation sequencing technology first changed the approach for genetic research and discovery and is now transforming clinical testing[145,146].

The emergence of faster NGS technologies has led to new diagnostic strategies with increasing levels of complexity, each with its pros and cons. These include panel testing, whereby between two and several hundred genes are sequenced in parallel, as well as whole exome sequencing (WES) that is not limited to selected genes, as for gene panels, but includes the coding of almost all the known human genes (about 20.000) and full-scale whole genome sequencing (WGS) which sequencing the entire genome.

1.3.1 Whole Genome Sequencing (WGS)

Whole-genome sequencing delivers a comprehensive view of the entire genome, unlike focused approaches such as exome sequencing or targeted resequencing, which analyze a limited portion of the genome. It is ideal for discovery applications, such as identifying causative variants and novel genome assembly.

WGS has been mostly used to identify the genetic bases of disease in single families[147,148].

A single WGS study has been published in a small cohort of patients with epilepsy[134] which makes an assessment of this technique in genetic epilepsies still premature (Table 2).

Whole-genome sequencing can detect single nucleotide variants (SNV), insertions/deletions, copy number variants (CNVs), and large structural variants (SV).

Rapidly dropping sequencing costs and the ability to produce large volumes of data with today's sequencers make whole-genome sequencing a powerful tool for genomics research. Advantages of WGS are so many, i.e.: provides a high-resolution, base-by-base view of the genome, captures both large and small variants that might otherwise be missed, identifies potential causative variants for further follow-on studies of gene expression and regulation mechanisms and delivers large volumes of data in a short amount of time to support assembly of novel genomes.

This approach produces library without any capture related-issues because overcame the targeting capture, and leads to obtain very homogenous coverage. Coverage is the average number of reads that align known reference bases, and it is calculated at a single nucleotide level and more the reads overlapping a region higher will be the coverage. This peculiarity permits to detect large structural variants and copy number variants.

The pro of WGS, the huge amount of data produced, become also its con, because we are not still so strength in the evaluation of single nucleotide polymorphisms (SNPs) or CNV in the no coding regions of human genome. Usually the data analysis of WGS is focused only on the 1-2% of genome, coding for proteins. WGS is only for research use so far, although probably it will become soon a new diagnostic tool[147,149].

1.3.2 Whole Exome Sequencing (WES)

Whole exome sequencing (WES), focusing on the most informative regions of the genome, and scanning thousands of genes simultaneously, is an alternative to gene-panel testing and locus specific analysis to investigate the molecular basis of genetic disorders in research and clinical diagnostics set-ups.

WES is less complex and expensive than WGS since sequencing the coding regions only reduces wet-lab, computational and data storage costs. An obvious advantage of WES is that all coding regions of all genes (about 20.000) are sequenced simultaneously, but technological limitations mean that some regions of the genome are not covered as well as others, and larger deletions/duplications usually cannot be detected. CES indicates a WES analysis performed in a laboratory certified by Clinical Laboratory Improvement Amendment (CLIA), College of American Pathology (CAP), or by the International Organization for laboratories. These quality and competence management systems ensure that the quality of the results is sufficient to be reported to the patient by a physician and entered in the patient's medical record[150]. Whole exome sequencing (WES) is typically seen the final stop on the diagnostic odyssey; however, it may make sense to bypass panels and go straight to WES for patients with additional features suggestive of a genetic syndrome (e.g. dysmorphic features, multiple congenital anomalies, etc.).

The WES diagnostic yield observed in clinical diagnostic testing of suspected monogenic diseases ranges from 22 to 31% but depends on several factors including phenotype (with early age of onset, consanguinity or recurrence in sibling considerably raising the rate of positive findings) and the sequencing strategy used (trios analysis for *de novo* mutations; quartets analysis for recessive or compound heterozygous conditions)[145]. In the epilepsies of suspected genetic origin, the molecular yield of WES ranges from 17% to 72%. These considerably different values are probably due to both the small number of enrolled patients in some studies and characteristics of the population studied. In a large cohort of unselected and consecutively enrolled patients with epilepsy, which might have provided a reasonably unbiased estimate of the diagnostic yield of WES/CES in epilepsy, Helbig et al. reported 38.2% solved cases[151] (Table 2).

The rate of variants with uncertain significance (VUS) compared to NGS panels is much lower because most WES is done on a trio basis (patient plus both parents). Systematic reanalysis of WES data often leads to a diagnosis, so reanalysis of WES data may be beneficial, typically at least 12 months after the initial uninformative result, when WES results come back as "negative".

If we considered diagnostic yield alone, WES would seem like the appropriate first step for all patients. However, for example for a patient with Dravet syndrome, if we tested only with WES, you would run the risk of missing a clinically relevant deletion of the *SCN1A* gene that could be detected using different testing strategies. So when deciding which test to order, it's best to take some time to decide which option is best for each case.

Although WES and CES have been already proposed as diagnostic tools in epilepsy[146], gene panel sequencing represents, at present, the most widely adopted method for the clinical diagnosis of genetic epilepsies, offering a rapid and cost-effective approach to identify mutations.

Reference	N° of patients	Phenotype(s)	NGS method	N° of cases with likely pathogenic mutation	Diagnostic yield (%)
[134]	6	sporadic severe early-onset epilepsy	WGS	4	66%
[147]	10	epilepsy (7 out of 10 with EE)	WES	7	70%
[152]	11	epilepsy	WES	8	72%
[153]	18	IS	WES	5	28%
[154]	50	EOEE	WES (data analysis restricted to an in silico panel of 137 genes)	11	22%
[155]	84	PME	WES	26	31%
[156]	264	IS & LGS	WES	44	17%
[151]	293	epilepsy	DES/CES	112	38.2%
[157]	356	EE	WES	61	17%

Table 2. *Summary of the WGS/WES/CES studies (sorted according to the number of patients) Abbreviations: IS infantile spasms; EOEE early onset epileptic encephalopathy; PME progressive myoclonic epilepsy; LGS Lennox-Gastaut syndrome; WGS whole genome sequencing* [158].

1.3.3 Gene panels

Gene panels contain a select set of genes or gene regions that have known or suspected associations with the disease or phenotype under study. Gene panels can be purchased with preselected content or custom designed to include genomic regions of interest. Multiple genes can be assessed across many samples in parallel, saving time and reducing costs associated with running multiple separate assays. Targeted gene sequencing also produces a smaller, more manageable data set compared to broader approaches such as whole-genome sequencing, making analysis easier.

Regions included in the panel design must be enriched by target capture approaches that are usually based upon a multiplexed PCR amplification reaction, DNA hybridization to a capture oligonucleotide (either on an array or in solution) or DNA capture via molecular inversion probe circularization[159].

Gene panels are increasingly used in a diagnostic setting, replacing a single-gene approach with a panel approach of various candidate genes. Gene panels have been used widely for testing for a genetic etiology in various epilepsies.

The first gene panel study was published by Lemke et al. in 2012 using a gene panel of 265 genes on a cohort of 33 index patients with various epilepsy phenotypes[106]. The authors identified presumed disease-causing mutations in 16 of 33 patients, demonstrating that gene panel approaches are useful diagnostic tools.

Several studies utilizing gene panels of 35-265 genes have been reported in the literature, with diagnostic yields ranging between 10% and 48.5%[160–163].

The diagnostic yield likely depends on the number of genes included on the panel and the phenotypes of the patients selected for the analysis. A gene panel with a high number of genes is invariably likely to give a high diagnostic yield. However, a higher number of variants of unknown significance may also be expected, blurring the interpretation of causative findings due to excessive validation analysis.

Most studies applying gene panels have not used systematic exon-level deletion and duplication testing, possibly missing small structural genomic variants in causative genes.

The variety of gene panel approaches used in the literature raises the issue of optimal gene panel design, which remains an unanswered question to date.

Considering the rapid advances in the area of epilepsy genetics, a gene panel might become obsolete very soon. A possible strategy to keep pace with rapid advances in gene discovery is to develop and validate custom gene panels. A better solution seems to be the use of virtual gene panel approach, where patients are genotyped by exome sequencing, but only a selected number of known epilepsy genes are analyzed and reported. The existing exome data can be reanalyzed when new genes are published in the literature. Due to technical reasons, the coverage of individual genes on a designed gene panel is superior to the coverage on exomes, but it is not known at this point whether this impacts diagnostic yield[146]. However a well-developed custom gene panel offers a better coverage of the targeted genes near to 100%, and this outweighs, at present, these WES/CES potential benefits in a diagnostic setting[164].

1.3.4 Impact of NGS in epilepsy

NGS technology has certainly improved the turnaround time for molecular diagnosis and therefore favored more timely decision making for management. A major impact of NGS, has arisen from the growing evidence that many sporadic developmental disorders, especially severe early onset epilepsies, once considered as potentially recessive in nature, have been demonstrated to arise from *de novo* mutations of dominant genes, with remarkable consequences on genetic counselling and prenatal diagnostic procedures[165,166]. Although inherited mutations in known epilepsy genes only account for a small minority of patients[7], following the advent of NGS in the field of genetic epilepsies, the number of genes that have been associated with monogenic epilepsies has rapidly increased to exceeded the number of[150,167], highlighting the role of an increasingly broad spectrum of biological pathways. NGS has also contributed to delineating the phenotypic spectrum associated with known monogenic epilepsy genes. The mass effect of these discoveries is rapidly changing our comprehension of the relations between early onset severe epilepsy and the associated neurological impairment, progressively narrowing the concept of epileptic encephalopathies and thereby influencing treatment choices.

Several genes have been identified sequencing large cohorts of sporadic patients, such as the Epi4K and Euroepinomics Consortia, but also through the study of small cohorts of patients[168]. The most successful NGS studies have been performed in sporadic patients with severe epilepsies (especially the early-onset epileptic encephalopathies) using a trios analysis consisting in the WES or WGS sequencing of the proband and healthy parents. Through this approach several novel genes have been identified carrying *de novo* mutations causing sporadic epilepsies, thus further highlighting the genetic heterogeneity of these conditions[125,168].

Focusing on the analysis to those genes identified in the last four years, Myers and Mefford[168] highlighted that, with a few exceptions, the majority of novel identified epilepsy genes account for less than a few dozen of patients. All these findings make it clear how the attempts to correlate epileptic syndromes to single genes are, at most, not easy. In addition, there is growing evidence that the same gene can be associated to a broader phenotypic spectrum than originally believed[169]. The traditional molecular genetics and the related conceptual approach that have guided clinicians in addressing genetic testing has long been one of single-gene sequencing test. The initial gene to be tested was selected based of the clinical picture and in patients with epilepsy, on electroencephalographic (EEG) and imaging findings.

If a disease-causing mutation was not identified, further genetic testing could be subsequently requested with specific genes being tested in succession based of their likely association with the phenotype. However, this process is expensive and time consuming, at times leading to late identification of potentially treatable conditions.

NGS has significantly changed the diagnostic workflow making it possible to simultaneously sequence a large number of genes and to obtain, in a short time, a molecular diagnosis in an increasing number of patients.

1.3.5 Bioinformatic behind NGS

Bioinformatics is a recently defined discipline that develops and applies advanced computational tools to analyze and interpret high dimensional biological data.
NGS-based bioinformatics tools are designed to convert signals to data, data to interpretable information, and information into actionable knowledge. This process can be conceptualized as[170] (Figure 2):

- Primary analyses: consists of processing raw sequencing instrument signals into nucleotide base and short-read data;
- Secondary analyses: analysis involves the alignment to a reference sequence or *de novo* assembly of the NGS nucleotide reads and subsequent variant detection;
- Tertiary analyses: provide context to the information generated during an NGS experiment by associating the sample-specific genomic profile with disparate descriptive annotations.

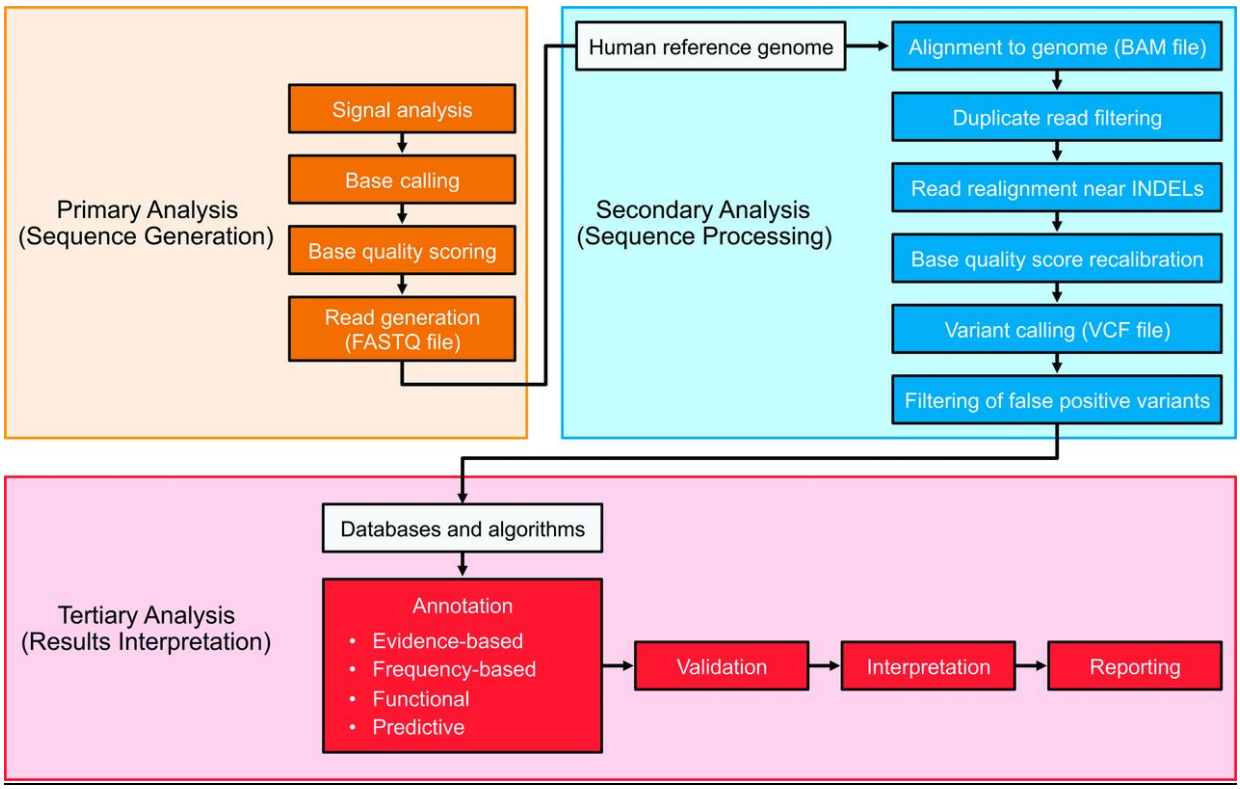

Figure 2. *Flow diagram illustrating the major components of a clinical NGS analytical pipeline*[171].

Primary analysis

Primary analysis convert the raw signals generated by the sequencing instruments into nucleotide bases with associated quality scores, and ultimately, short nucleotide sequences or "reads." In some instances, the primary analysis also includes demultiplexing of multiple samples indexed and pooled into a single sequencing run. It is a process that has become highly integrated with the sequencing instruments and associated onboard software.

Secondary analysis

Secondary analysis consists of a variable collection of methods that operate together to detect genomic aberrations from quality-scored sequence data (Table 3).

The class of genomic variation profiled can includes:

- single nucleotide variants,
- small insertions and deletions,
- larger alterations like structural rearrangements and copy number changes

Furthermore, genomic variations can be either constitutional (*de novo* or inherited) or somatic (acquired), affecting only a subset of the body's cells, such as in cancer.

Category	Tools	Available via
Aligners		
	BWA	http://bio-bwa.sourceforge.net/
	Novoalign	http://www.novocraft.com
	Bowtie	http://bowtie-bio.sourceforge.net/bowtie2
Germline callers		
	GATK (UnifiedGenotyper/ Haplotype Caller)	https://www.broadinstitute.org/gatk/download
	SAMtools	http://samtools.sourceforge.net/
	Varscan2	http://varscan.sourceforge.net
	SNVMix	http://compbio.bccrc.ca/software/snvmix/
Somatic callers		
	SomaticSniper	http://gmt.genome.wustl.edu/somatic-sniper
	JoinSNVMix2	https://code.google.com/p/joint-snv-mix/
	Mutect	
SV[b] tools		
	Delly	https://github.com/tobiasrausch/delly
	Breakdancer	http://breakdancer.sourceforge.net/
	PINDEL	http://gmt.genome.wustl.edu/pindel/current
	CREST	http://www.stjuderesearch.org/site/lab/zhang

Table 3. *Commonly used tools for NGS-based DNA analysis by functional category.* [a] *A selection of commonly utilized NGS analysis tools used in secondary analysis work flows. All tools are freely available, although licensing may be necessary to enable for-profit use.* [b] *SV, structural variant.*

The first analysis step usually involves the collective alignment of reads to a reference human genome. Once reads have been aligned to the genome, are often performed[172] the flag or filter of duplicate reads likely to be PCR artifacts, and realignment, which leverages a collective view of reads around putative insertion/deletion (indel) sites to minimize erroneous alignment of read ends. Variant calling involves the comparison of the sequenced reads to their point of alignment on the human genome to determine areas that differ on the basis of statistical modeling techniques that aim to distinguish genuine genomic variations from errors[173].

Variant calling errors are common, as NGS technologies are inherently less accurate than traditional sequencing methods and, therefore, artifacts occur with greater regularity[174].

This problem is partially corrected for by increasing sequencing depth (i.e., sequencing each base position multiple times). Prioritization is often preferred to removal of candidate variants to avoid the incorrect and irreversible filtering of a genuine variant call. The filtering or prioritization process can involve computational or human efforts, including visual inspection of variant alignments, and can be based on empirical cutoffs or more advanced statistical approaches. Criteria used to assess the quality of variant calls varies but in general include the frequency with which a variant allele is observed in a sample, the base quality of the variant alleles as predicted by the sequencing instrument, and the ability of a read containing a variant allele to map uniquely to a single location on the human reference genome. Variant callers are also widely recognized as generating very different results and often each caller will detect its own distinct set of unique, correct calls[175]. It is possible under some circumstances that no one approach or configuration will be sufficient to achieve acceptable performance for a given application, and in such instances the use of parallel and complementary methods is advisable in many cases to achieve the required level of sensitivity.

Tertiary analysis

In tertiary analysis the variants are annotated to determine their biological significance and enable functional prioritization and downstream interpretation. This characterization is generally achieved using a combination of biological annotation sources including frequency, structural, prediction, or evidence-based data (Table 4).

	Annotation source	Description	Available via
Population frequency based	1000 Genomes Project	Low-coverage whole genome sequencing of 2500 healthy humans.	http://www.1000genomes.org
	NHLBI Cohort	6500 Sequenced exomes from heart, lung, and blood disorder patients.	https://esp.gs.washington.edu/drupal/
	HapMap Project	SNPa-based data set to define haplotypes across 270 ethnically diverse humans.	http://hapmap.ncbi.nlm.nih.gov
	gnomAD	123136 exome sequences and 15496 whole-genome sequences from unrelated individuals sequenced as part of various disease-specific and population genetic studies.	http://gnomad.broadinstitute.org/
Structural based	SnpEff	Variant impact on codon and gene structure.	http://snpeff.sourceforge.net/SnpEff.html
	VEP	Variant impact on gene, transcript, protein sequence.	http://www.ensembl.org/info/docs/tools/vep/index.html
Prediction based	SIFT	Sequence conservation.	http://sift.jcvi.org/
	POLYPHEN	Phylogenetic and structural characteristics.	http://genetics.bwh.harvard.edu/pph/
	CONDEL	Meta-prediction aggregator.	http://omictools.com/sequencing/genome-resequencing/ driver-mutations/condel-s654.html
	MutPred	Random forest prediction method.	http://mutpred.mutdb.org/
	CADD	Meta-prediction and annotation score.	http://cadd.gs.washington.edu
	VAAST	Phylogenetic and disease-based conservation.	http://www.yandell-lab.org/software/vaast.htm
	MutationTaster	Meta-data type integration.	http://www.mutationtaster.org
	ANNOVAR	Meta-data, meta-prediction aggregator	http://www.openbioinformatics.org/annovar/
Evidence based	OMIM	Disease phenotype–gene relationships.	http://www.omim.org

	Human Gene Mutation Database	Human inherited disease gene lesions.	http://www.hgmd.org
	ClinVar	Clinical human variation to phenotype relationships.	http://www.ncbi.nlm.nih.gov/clinvar/

Table 4. *Commonly used tertiary analysis annotation resources.* [a] *SNP, single nucleotide polymorphism.*

Population frequency-based annotations are often a core component of tertiary analysis because variants that are common in the general population are unlikely to have biological relevance in the context of a clinical assay. Frequency thresholds are generally applied to remove benign polymorphisms from variant lists.

Structural-based annotations assign the effect of a variant on the transcripts and encoded protein(s) based on the resulting amino acid change[176]. The effect on the encoded protein sequence is subsequently categorized using clearly defined rules; for example, nonsense mutations are categorized as highly impactful.

Prediction-based annotations use nucleotide and/or amino acid changes integrated with additional contextual data, including evolutionary conservation scores, amino acid substitution matrices, and impact on 3D protein structures to infer the variant's impact on the resulting sequence product. Finally, evidence-based annotations are derived from the literature and other historical data[171].

1.3.6 Bioinformatic workflow in NGS

When the sequencing on NGS platform was finished, we had a lot of small sequences of DNA, called reads.

The reads were stored in FASTq format, that was a raw format with unaligned reads, consisting in 4 elements (Figure 3)

- Identifier: begins with a '@' character and is followed by a sequence identifier and an optional description;
- Sequence letter: is the raw sequence letters;
- Plus: begins with a '+' character and is optionally followed by the same sequence identifier (and any description) again;
- Base qualities score: encodes the quality values for the sequence in line 2, and must contain the same number of symbols as letters in the sequence.

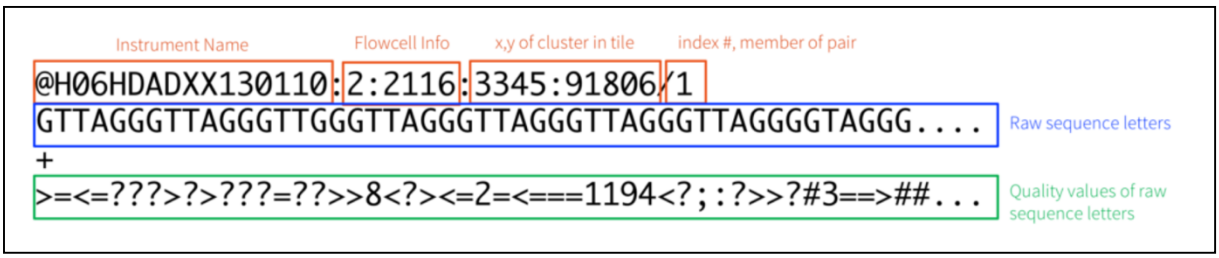

Figure 3. *Example of Fastq format.*

The Base qualities score of the fourth row is generated from NGS platform and is represented as the character with an ASCII code equal to its value plus 33.

ASCII stands for American Standard Code for Information Interchange. Computers can only understand numbers, so an ASCII code is the numerical representation of a character such as 'a' or '@' or an action of some sort.

This format has been used for storing both a biological sequence (usually nucleotide sequence) and its corresponding quality scores.

The second step is represented by the alignment of the reads to the reference and results in the generation of the BAM file, that also stores aligned reads. At the end of the analysis we obtained a variant call format (VCF), which contains all information about variants. VCF is a tab-separated file format with columns for the chromosome, position, Single Nucleotide Polymorphism Database (dbSNP) ID, reference allele, alternative allele, variant quality, a filter column containing information on whether the variant has passed or failed various filters, a column containing extra annotations, and finally columns that define the genotype of the variant. One we obtained the VCF format we can proceed we the annotation of the data. We gave a list of variants with chromosome, start position, end position, reference nucleotide and observed nucleotides, to an annotation tools (i.e. ANNOVAR, an efficient software tool that performs the annotation of variants) (Figure 4). In this way we are capable to have for all variants, a value of frequency in population, conservation and the prediction of their benign or damaging impact.

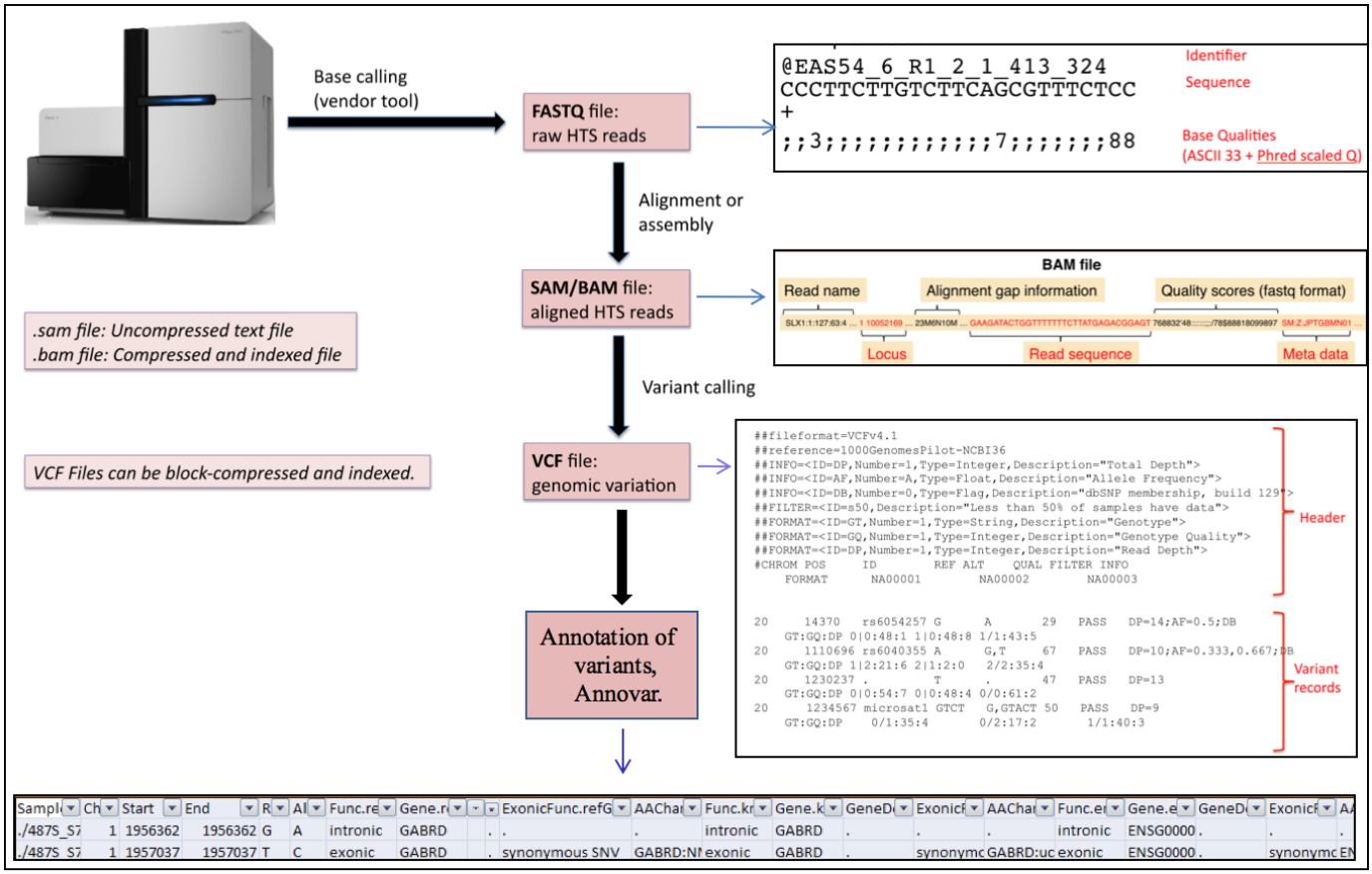

Figure 4. *Summary of common bioinformatic workflow.*

2. Aim of the study

Pediatric epilepsies are a highly heterogeneous group of neurologic disorders that encompass a number of syndromes with variable severity, ranging from benign with early spontaneous remission to progressive encephalopathy refractory to treatment.

In recent years, major improvements in phenotyping and recognition of distinctive epilepsy syndromes, paralleled by new methods in molecular genetics and gene sequencing, especially NGS methods, have led to the identification of a large number of new monogenic epilepsies, both inherited and *de novo* in origin, especially in the category of the epileptic encephalopathies, a group of the most severe early onset epilepsies, for which *de novo* mutations have been identified in a plethora of genes. Genetic research in epilepsy represents an area of great interest for both clinical purposes and for the understanding of mechanisms underlying epilepsy.

Many patients who received conventional clinical evaluation and targeted genetic testing performed with Sanger sequencing remain without specific diagnosis even after extensive workup. This has serious consequences for patients and their families, preventing the access to the right treatment or accurate counseling for pregnancies and prognosis. Hence, there is growing interest in implementing Next-Generation Sequencing (NGS) approaches that deliver fast, and detailed genetic information, providing an effective approach for identifying causal variants in Mendelian disease genes. The advent of NGS allowed the identification of many new genes and a wide number of mutations associated with epilepsy, but despite for many cases the genetic basis remain still unknown.

In this project we explored the utility of whole exome sequencing (WES) for identifying candidate causal mutations in a cohort of children with heterogeneous epilepsies, without etiologic diagnoses and complex phenotypes. We selected a cohort of children with complex phenotypes whom were already been tested with target sequencing and resulted to be negative. We investigated the genetic architecture of epilepsy to understand how mutations are involved in patients with complex phenotypes and performed an accurate genotype-phenotype correlation to better define the phenotypic spectrum associated with the identified variants.

3. Materials and Methods

We selected a cohort of children with complex phenotype characterized by different types of epilepsy sometimes associated with additional feature resulted negative to previously molecular tests. We assembled two different groups of patients, the first one consisting of trios samples (patient, mother and father) and the second one composed only by the patient.
We sent our samples to two of the most important genomic centers to perform WES: the Beijing Genomics Institute (BGI) and the Broad Institute. The trios cohort was sent to the Beijing Genomics Institute (BGI) and the singleton cohort to the Broad Institute.
Patients were referred from the Pediatric Neurology Unit of the Anna Meyer Children's Hospital, as well as from other national and European neurology centers, and exhibiting a wide spectrum of severe epilepsy. WES was performed in all patients after obtaining informed consent. We obtained clinical information and blood samples of patients. We obtained approval for this study from the Institutional Review Board of the Meyer Children's Hospital.

3.1 Sample

The peripheral blood samples of patients were collected by physicians of the Pediatric Neurology of the AOU Meyer or came from other Italian centers that recruited the samples. We used blood samples or genomic DNA of patients and when possible of their parents for molecular tests. The first step in the processing of the samples was their storage in our laboratory database giving each patient an alphanumeric code. Blood from families was taken in EDTA (Ethylenediaminetetraacetic acid) tube. Genomic DNA (gDNA) was extracted from whole blood using a QIAsymphony SP automated robot (Qiagen, Hilden, Germany).

3.2 DNA Quantification

Genomic DNA was quantified using a Quantus Fluorometer and the QuantiFluor ONE dsDNA System (Promega, Madison, WI, USA).

DNA quality was assessed with a Qiaxcel spectrophotometer (Qiagen, Hilden, Germany). Absorbance measurements will measure any molecules absorbing at a specific wavelength, DNA will absorb at 260nm and will contribute to the total absorbance. For quality control, the ratios of absorbance at 260/280 nm and 260/230nm were used to assess the purity of DNA.

Ratios of 260/280 + ~1.8 and 260/230 = 2.0-2.2 denote pure DNA. Ratios significantly lower than expected may indicate the presence of contaminants. The Qiaxpert Spectrophotometer is loaded with 2 µl of solution and reads the DNA concentration that will be expressed in ng/µl. gDNA were purified and resuspended in correct buffer using the DNA Clean & Concentrator-5 columns (Zymo Research Corporation, Irvine, CA, USA).

3.3 Pipeline to bioinformatic analysis: GATK best practice

In this section we reported the recommendations according to Genome Analysis Toolkit (GATK) for bioinformatic analysis (https://software.broadinstitute.org/gatk/bestpractices). according to

Sequence data obtained from sequencing provider is typically in a raw state (one or several FASTQ files) that is not immediately usable for variant discovery analysis. Even if you receive a BAM file (file in which the reads have been aligned to a reference genome) you still need to apply some data processing in order to maximize the technical correctness of the data.

This first phase is the pre-processing steps that are necessary in order to prepare data for analysis, starting with FASTQ or uBAM files and ending in an analysis-ready BAM file.

The begin is the mapping of the sequence reads to the reference genome to produce a file in SAM/BAM format sorted by coordinate. Next, the duplicates are marked to mitigate biases introduced by data generation steps such as PCR amplification. Finally, the recalibration of the base quality scores, because the variant calling algorithms rely heavily on the quality scores assigned to the individual base calls in each sequence read.

Pre processing

This first phase is the pre-processing steps that are necessary in order to prepare data for analysis, starting with FASTQ or uBAM files and ending in an analysis-ready BAM file.

Map to Reference

The first step is of course to map of the reads to the reference to produce a file in SAM/BAM format (in our case we used human genome hg19/b37). Generally, it is recommend to use BWA MEM for DNA, but depending on data and how it was sequenced. After the mapping of the reads, the reads are sorted in the proper order (by coordinate). Read group information is typically added during this step, but can also be added or modified after mapping using Picard AddOrReplaceReadGroups.

Mark Duplicates

Once the data has been mapped to the reference genome, we can proceed to mark duplicates. The idea here is that during the sequencing process, the same DNA fragments may be sequenced several times. The resulting duplicate reads are not informative and should not be counted as additional evidence for or against a putative variant. The duplicate marking process does not remove the reads, but identifies them as duplicates by adding a flag in the read's SAM record. Duplicate marking should NOT be applied to amplicon sequencing data or other data types where reads start and stop at the same positions by design.

Recalibrate Bases

Variant calling algorithms rely heavily on the quality scores assigned to the individual base calls in each sequence read. These scores are per-base estimates of error emitted by the sequencing machines. Unfortunately, the scores produced by the machines are subject to various sources of systematic technical error, leading to over-or-under-estimated base quality scores in the data. Base quality score recalibration (BQSR) is a process in which GATK developers applied machine learning to model these errors empirically and adjust the quality scores accordingly. This allows BQSR to get more accurate base qualities, which in turn improves the accuracy of variant calls. The base recalibration process involves two key steps: first the program builds a model of covariation based on the data and a set of known variants (which you can bootstrap if there is none available for your organism), then it adjusts the base quality scores in the data based on the model. There is an optional but highly recommended step that involves building a second model and generating before/after plots to visualize the effects of the recalibration process. This is useful for quality control purposes.

Variant Discovery

Once the data has been pre-processed, can start the variant discovery process. Unfortunately, some observed variation is caused by mapping and sequencing artifacts, so the greatest challenge here is to balance the need for sensitivity (to minimize false negatives, i.e. failing to identify real variants) vs. specificity (to minimize false positives, i.e. failing to reject artifacts). For this complicated balance the variant discovery process was divide into two separate steps: variant calling and variant filtering. The first step is designed to maximize sensitivity, while the filtering step aims to deliver a level of specificity that can be customized for each project. For DNA, the variant calling step is further subdivided into two separate steps (per-sample calling followed by joint genotyping across samples) in order to enable scalable and incremental processing of cohorts comprising many individual samples.

This workflow involves running HaplotypeCaller on each sample separately in GVCF mode, to produce an intermediate file format called GVCF (for Genomic VCF). The GVCFs of multiple samples are then run through a joint genotyping step to produce a multi-sample VCF call-set, which can then be filtered to balance sensitivity and specificity as desired. The best way to filter the resulting variant call-set is to use variant quality score recalibration (VQSR), which uses machine learning to identify annotation profiles of variants that are likely to be real. The drawback of this sophisticated method is that it requires a large call-set (minimum 30 exomes, more than 1 whole genome if possible) and highly curated sets of known variants. This makes it difficult to apply to small experiments, RNAseq experiments, and non-model organisms; for those it is typically necessary to develop hard filtering parameters manually.

Call Variants

In the past, variant callers specialized in either SNPs or Indels (like the GATK's own UnifiedGenotyper), could call both but had to do so them using separate models of variation. The HaplotypeCaller is capable of calling SNPs and Indels simultaneously via local *de novo* assembly of haplotypes in an active region. In other words, whenever the program encounters a region showing signs of variation, it discards the existing mapping information and completely reassembles the reads in that region. This allows the HaplotypeCaller to be more accurate when calling regions that are traditionally difficult to call, for example when they contain different types of variants close to each other. It also makes the HaplotypeCaller much better at calling Indels than position-based callers like UnifiedGenotyper. In the so-called GVCF mode used for scalable variant calling in DNA sequence data, HaplotypeCaller runs per-sample to generate an intermediate genomic (GVCF), which can then be used for joint genotyping of multiple samples in a very efficient way, which enables rapid incremental processing of samples as they roll off the sequencer, as well as scaling to very large cohort sizes (e.g. the 92K exomes of ExAC).

Merge (optional)

This is an optional step that applies only to the DNA workflow. It consists of merging GVCF files hierarchically in order to reduce the number of files that must be handled simultaneously in the next step. This is only necessary if you are working with more than a few hundred samples.

Joint Genotype

At this step, all the per-sample GVCFs (or combined GVCFs if we are working with large numbers of samples) are gather and pass them all together to the joint genotyping tool, GenotypeGVCFs. This produces a set of joint-called SNP and indel calls ready for filtering. This cohort-wide analysis empowers sensitive detection of variants even at difficult sites, and produces a squared-off matrix of genotypes that provides information about all sites of interest in all samples considered, which is important for many downstream analyses.

Filter Variants

The best way to filter the raw variant call-set is to use variant quality score recalibration (VQSR), which uses machine learning to identify annotation profiles of variants that are likely to be real, and assigns a VQSLOD score to each variant that is much more reliable than the QUAL score calculated by the caller. In the first step of this two-step process, the program builds a model based on training variants, then applies that model to the data to assign a well-calibrated probability to each variant call. We can then use this variant quality score in the second step to filter the raw call set, thus producing a subset of calls with our desired level of quality, fine-tuned to balance specificity and sensitivity.

Callset Refinement

Once the call-set is generated and filtered according to GATK recommendations, there are several options for evaluating and refining the variant and genotype calls further, before moving on with the study. None of next steps are absolutely required, and the workflow should be adapted to each project's requirements.

Annotate Variants

We can distinguish two types of variant annotations that differ in both origin and purpose: context annotations and functional annotations.

- Context annotations are typically determined during the variant discovery steps. They provide information about the properties of the sequencing data and the genomic context of the variant from a technical point of view. They are largely used for filtering purposes. When is required to add variant context annotations to a VCF file, whether before or after the filtering step, we can use the VariantAnnotator, which can call most annotation modules listed in the tool documentation. Note that some annotation modules may have some special input requirements, as they may need to see e.g. the original BAM file, a pedigree file, or a minimum number of samples.

- Functional annotations are predictions of what effect a variant may have on biological function based on its position within coding regions, regulatory regions and so on. These depends on external databases of known variants and functional predictions. At this time GATK does not include any tools that perform functional annotation. The annotation tool that we use internally is called Oncotator and is supported in a subset of the GATK forum. Other tools that are widely used in the community are SnpEff and Annovar, but they cannot provide any support for their use.

3.4 Human Phenotype Ontology (HPO)

The Human Phenotype Ontology (HPO) provides comprehensive bioinformatics resources for the analysis of human diseases and phenotypes, offering a computational bridge between genome biology and clinical medicine[177].

The description of phenotypic variation has become a central topic for translational research and genomic medicine[178], and "computable" descriptions of human disease using HPO phenotypic profiles (also known as "annotations") have become a key element in a number of algorithms being used to support genomic discovery and diagnostics.

There are now 123724 annotations of HPO terms to rare diseases and 132620 to common diseases.

Using HPO to generate individualized phenotype-driven gene panels for diagnostics led to an increase in the diagnostic yield[179].

The algorithm, Phenotypic Interpretation of Exomes (PhenIX) contributed to a diagnostic rate of 28% in children in whom previous extensive workups had failed to reveal a diagnosis[180].

The HPO has been used in many ways in research on disease pathophysiology, diagnostics and gene-discovery projects. It has been used to provide lists of genes associated with one or more HPO terms in order to filter lists of candidate genes[181], to prioritize candidate genes in exome sequencing studies via PhenIX, Phevor or Exomiser[182], and to identify known or novel disease genes or to analyze structural variation in large cohorts[183].

3.5 BGI project

Beijing Genomics Institute (BGI), is one of the world's genome sequencing centers, headquartered in Shenzhen, China.

We sent to BGI a cohort of 63 samples:

- 17 trios (patient, father and mother)
- 2 quartets (patient, sibling affected, father and mother)
- 1 singleton (patient)

The patients have a very severe form of epilepsy syndrome and all patients are already prescreening with other molecular technology and resulted all negative.

3.6 BGI sequencing

The 63 exons were captured from fragmented genomic DNA samples using the SureSelect Human All Exon v4 (Agilent Technologies, Santa Clara, CA, USA), and paired-end 90-base massively parallel sequencing was carried out on an Illumina HiSeq 2000, according to the manufacturer's protocols (Illumina, San Diego, CA, USA).

Bioinformatics analysis

The bioinformatics analysis was carried out using an in-house developed pipeline. Sequencing reads passing quality filtering were aligned to the human reference genome (hg19) with Burrows-Wheeler Aligner (BWA)[184]. The Genome Analysis Toolkit (GATK)[185] was used for base quality score recalibration, indel realignment, duplicate removal, and to perform SNP and INDEL discovery and genotyping across all samples simultaneously using variant quality score recalibration according to GATK Best Practices recommendations[172,186]. For annotating and filtering data, the SNPeff program was used[176]. The data was further filtered prioritizing the variants with three different inheritance models: dominant *de novo*, heterozygous compound and homozygous recessive. The *de novo* variants were called using the DeNovoGear tool with a 0.9 threshold at the posterior probability of the most likely *de novo* genotype configuration[187]. All samples had a mean depth of target region covered at 112X and >97% of bases in the consensus coding sequences (CCDSs) covered by at least 20 reads.

Filtration rules

We excluded variants with minor allele frequency >1% in either the 1000 Genomes Project (1000g) or the Exome Aggregation Consortium (ExAC v0.3) databases.

We evaluated whether the human protein-coding gene is likely to harbor disease-causing mutations using different gene-level prediction tools such as the ExAC (Exome Aggregation Consortium) constraint metrics[188], the residual variation intolerance score (RVIS, based on ExAC v2 release 2.0)[189] and the gene damage index (GDI)[190]. Furthermore, to improve the use of existing variant-level methods such as the CADD, we also used the mutation significance cutoff (MSC) server[191], a quantitative approach that provides gene-level and gene-specific phenotypic impact cutoff values.

Mutations' pathogenicity for the identified variants has been evaluated through in silico prediction using the dbNSFP database (v3.0a), which provides functional prediction scores on more than 20 different algorithms (https://sites.google.com/site/jpopgen/dbNSFP). To assess the effects of variants, we used both the dbNSFP ensemble rank scores MetaSVM and MetaLR[192]. In addition, we

used the scores obtained from Revel[193], M-CAP[194] and Eigen[195], three different bioinformatics tools to evaluate the pathogenicity of rare variants.
We submitted for validation and segregation testing by Sanger sequencing candidate *de novo* variants that were predicted to alter protein function (non-synonymous, stop-gain, stop-loss, frameshift and splice-junction mutations).

3.7 Epi25k Consortium

The Epi25k Consortium represents the unification of various consortia dedicated to the study of epilepsy, including Epi4K, EPIGEN, EuroEPINOMICS, the Epilepsy Phenome/Genome Project, EpiPGX, SANAD, and EpiCURE. Each of these large, multinational projects maintains a DNA repository. The objective of this collaboration is to combine genotype, phenotype, and genomic sequencing data from Epi25K GWAS and genome sequencing (WES and WGS) studies, and to perform joint analyses of these datasets. It is the largest sequencing project in the epilepsies, responding to a call by the National Human Genome Research Institute (NHGRI). Epi25K samples are currently being selected and shipped. Many of the samples have been ideally consented, and for some of the collections the DNA is verified to be of high quality (and in all cases is collected from a blood sample can be sequenced PCR-free). The prioritization of samples from a phenotypic standpoint is decide by the Epi25K Consortium, in consultation with the CCDG at the Broad Institute. The Epi25 project should be able to collect and sequence around 25,000 individuals with epilepsy, helping us to achieve the next, necessary level for gene discovery in human epilepsies. At this time only the first 6000 patients have been sequenced and the project is still ongoing.
Together, the consortium's sample collection is the best resource in the world for genetic studies of epilepsy, both in terms of numbers of patients included and the quality of phenotypic information available.

We sent to Broad institute a cohort of 117 patients with severe form of epilepsy.

3.8 Epi25k sequencing

The 117 samples exon were captured from fragmented genomic DNA samples using the Rapid Illumina's TruSeq Exome Capture (Illumina, Inc., San Diego, CA, USA) and for all samples sequenced before November of Nextera Rapid Capture Exomes. There is only a slight difference in hybridization between the two kits, but both utilize Illumina Nextera sample prep and Illumina Rapid Capture Enrichment with ICE bait.
The kits cover 214405 exons with a total size of about 37 Mb. Sequencing was done using either the

NextSeq500 or HiSeq4000 sequencers (Illumina, Inc.) to produce 2 × 150 bp reads, and always pooling up to 9 WES per lane.

Raw sequencing reads were converted to standard fastq format using bcl2fastq software 2.17.1.14 (Illumina, Inc.), and fed to developed pipeline for the analysis of WES data that is based on the 1000 Genomes Project (1000G) data analysis pipeline and GATK best practice recommendations, which includes widely used open source software projects.

In the data of singleton cohort, the patients losing the genetic familiar background and can't perform a classical trios analysis.
We decided to process the data analyzing and prioritizing the genes that have high HPO match score using Exomiser tools. HPO terms translate the peculiar phenotype features of the patients in particular alfa numeric code, recognized from the major exomes tools analyzers. With the help of physicians of Pediatric Neurology Unit of the Anna Meyer Children's hospital, for each patient have been assigned the appropriate HPO terms, that explain in the better way the peculiar phenotypic features of patients.

3.9 Exomiser

Exomiser is an application that prioritizes genes and variants in next-generation sequencing projects for novel disease-gene discovery or differential diagnostics of Mendelian disease. Exomiser comprises a suite of algorithms for prioritizing exome sequences using random-walk analysis of protein interaction networks, clinical relevance and cross-species phenotype comparisons, as well as a wide range of other computational filters for variant frequency, predicted pathogenicity and pedigree analysis[196]. Exomiser requires ~3GB of ram and roughly 15-90s of computing time on a standard desktop computer to analyze a variant call format (VCF) file. Exomiser is freely available for academic use from http://www.sanger.ac.uk/science/tools/exomiser.
The inputs to Exomiser are the called variants resulting from exome sequencing of a rare disease patient and, optionally, other affected and unaffected family members. These variants are stored using VCF[197]. The Exomiser scans these VCF files using a single command to first filter the variants and then to prioritize the remaining candidates to help researchers identify the causative variant or variants. The filtering step is critical in order to reduce the 30,000 variants seen in a typical exome to a more manageable size. However, this still typically leaves more candidates (up to 1,000) than can be reasonably manually assessed, so some sort of ranking by a prioritization algorithm is necessary (Figure 5). The Exomiser comprises a suite of four algorithms for the analysis of NGS data for diagnostics or novel disease-gene discovery in the field of rare disease. If the VCF file comprises samples from multiple family members, then a PED file is additionally required.

The Exomiser initially annotates variants using Jannovar27, and then it removes variants that are off the exomic target or that are more common than a user-supplied threshold; it then ranks the remaining variants according to their predicted pathogenicity. Finally, the clinical relevance of the genes harboring these variants is assessed using one of three phenotype-driven algorithms (PHIVE, PhenIX or hiPHIVE) or by a random-walk algorithm that assesses the vicinity of the genes to members of disease-gene family on the protein-protein interactome. Users are required to supply HPO terms for the phenotype-driven algorithms or a list of seed genes representing the disease-gene family for ExomeWalker.

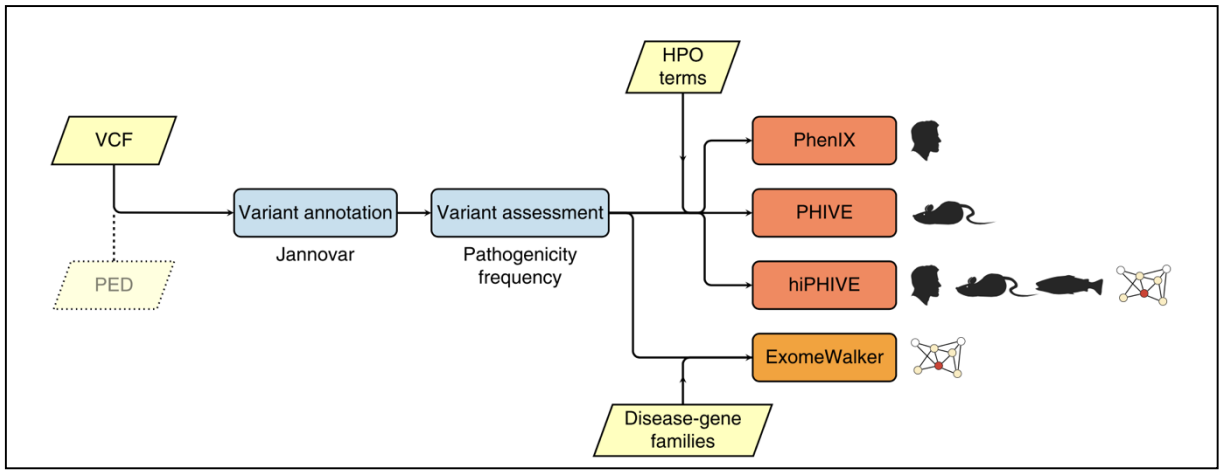

Figure 5. *Overview of the processing steps of Exomiser*[196].

Exomiser can be run using a set of seed genes if a disease-gene family can be identified. For instance, numerous genes mutated in retinitis pigmentosa (RP) have been identified. If it is suspected that a patient may have a mutation in a novel gene for RP, then Exomiser can be run with the ExomeWalker prioritization method together with a list of NCBI Entrez gene IDs for genes already known to be involved in RP. Alternatively, Exomiser can be run using phenotypic similarity–based algorithms (PHIVE, PhenIX or hiPHIVE). Here the user needs to enter a list of HPO terms representing the clinical manifestations observed in the individual being investigated. The choice of the prioritization method and parameters will depend on the clinical or research goal. PhenIX will interrogate only known Mendelian disease genes, and it will use only human phenotypic data to calculate phenotypic similarities. PHIVE will use mouse phenotypic data to identify candidates. Finally, hiPHIVE can use mouse, zebrafish and human clinical data, and it can additionally integrate further candidates on the basis of an analysis of the protein-protein interaction network. All of the analyses can be combined with additional filter[196].

3.10 Variants confirmation

We confirmed and tested the identified variants in the probands and parents by Sanger sequencing. For conventional Sanger sequencing, we amplified 50 ng of genomic DNA from patients and their parents using the FastStart Taq DNA Polymerase (Roche) and specific primers designed using Primer3 plus software (http://www.bioinformatics.nl/cgi- bin/primer3plus/primer3plus.cgi). After checking PCR products by 1.5% agarose gel electrophoresis and a purification step with the ExoSAP-IT (Affymetrix, Santa Clara CA), we analyzed the amplicons by direct sequencing, on both strands, using the BigDye Terminator V1.1 chemistry (Life Technologies, Carlsbad, CA), on an ABI Prism 3130XL automated capillary sequencer (Life Technologies).

Criteria for Pathogenicity of Rare or Novel Variants

We classified rare or novel variants as being "pathogenic", "likely pathogenic", "variant of uncertain significance (VOUS)" or "benign" according to the international guidelines of the ACMG Laboratory Practice Committee Working Group[198].

4. Results

Overall, we analyzed 140 patients for clinical WES, with a comparable number of males and females. We solved 16% (22/140) of the total cases, identifying 22 causative mutations, 18 of which (81%) are *de novo* mutations.

We classified all variants where the parents are not available for segregation testing as potentially causative mutation. If we consider also the potentially causative mutations, we obtained a higher yield of 24% (34/140).

The trios cohort, including 63 exomes, consisted in 18 trios (proband, father and mother), 2 quartets (proband, sibling affected, father and mother) and one singleton. The major part was analyzed with a trio design (parents and index), allowing analyses consistent with all possible patterns of inheritance of the disease.

The singleton cohorts counted 117 exomes, composed only from probands, and in only 20% of cases were available.

We reported all causative or potentially causative mutations in the Table 5.

Cohort	Gene	Mutation	Status	Inheritance	Phenotype	In HGMD
Trios	*GABRA1*	NM_000806.5c.865A>C p.(Thr289Pro)	solved	*de novo*	Neonatal onset epileptic encephalopathy and severe a developmental delay. Died	yes[199]
Trios	*ATP6V1A*	NM_001690.3:c.298G>T p.(Asp100Tyr)	solved	*de novo*	Late infantile EE with heterogeneous seizures	no
Trios	*CACNA1E*	NM_001205293.1:c.2104G>A p.(Ala702Thr)	solved	*de novo*	Early onset epileptic encephalopathy, severe intellectual disability, spastic tetraparesis and movement disorder	no
Trios	*DMXL2*	NM_001174116.1:c.5135C>T p.(Ala1712Val) NM_001174116.1:c.4478C>G p.(Ser1493*)	solved	het comp	Early onset epileptic encephalopathy, severe intellectual disability, spastic tetraparesis. Died	no
Singleton	*UBE3A*	NM_130838.1:c.510_515del p.(Glu171_Lys172del)	solved	x-linked	Epileptic spasms, Hypoplasia of the corpus callosum	no
Singleton	*KCNB1*	NM_004975:c.1747C>T p.(Arg583*)	solved	*de novo*	Developmental regression, Epileptic spasms, Hypsarrhythmia	no
Singleton	*GNAO1*	NM_138736.2:c.607G>A p.(Gly203Arg)	solved	*de novo*	Intellectual disability, Dystonia, Spastic	yes[200]

					tetraparesis, Congenital nystagmus, Focal seizures	
Singleton	*SCN2A*	NM_021007.2:c.2576A>G p.(Lys859Arg)	solved	*de novo*	Generalized tonic-clonic seizures	no
Singleton	*PDHA1*	NM_000284.3:c.904C>T p.(Arg302Cys)	solved	*de novo*	Epileptic spasms, Hypsarrhythmia	yes[163]
Singleton	*PURA*	NM_005859.4:c.768dup. p.(Ile257Hisfs*37)	solved	*de novo*	Intellectual disability, Tetraparesis, EEG with spike-wave complexes, Seizures, Generalized hypotonia	yes[201]
Singleton	*ST3GAL5*	NM_003896.3:c.1000C>T p.(Arg334*) NM_003896.3:c.1166A>G p.(His389Arg)	solved	het comp	Intellectual disability, severe, Focal motor seizures, Dyskinesia, Stereotypic behavior	no
Singleton	*PURA*	NM_172107.2:c.977A>C p.(Gln326Pro)	solved	*de novo*	Intellectual disability, severe, Tetraparesis, Generalized tonic seizures, Epileptic encephalopathy, Infantile axial hypotonia, Uni and bilateral multifocal epileptiform discharges , Focal seizures	no
Singleton	*TBC1D24*	NM_001199107.1:c.671A>G p.(Tyr224Cys) NM_001199107.1:c.1008del p.(His336Glnfs*12)	solved	het comp	Cortical myoclonus, Versive seizures, EEG with polyspike wave complexes, Intellectual disability, borderline, Obtundation status.	no
Singleton	*MEF2C*	NM_002397.4:c.52_53del p.(Gln18Glyfs*16)	solved	*de novo*	Infantile muscular hypotonia, Abnormality of the cerebral white matter, Febrile seizures, Macrocephaly at birth, Megalencephaly, Abnormal facial shape	no
Singleton	*HCN1*	NM_021072.3:c.790A>T p.(Ser264Cys)	solved	*de novo*	Seizures, Delayed speech and language development	no
Singleton	*PURA*	NM_005859.4:c.649del p.(Glu217Serfs*8)	solved	*de novo*	Severe global developmental delay, Epileptic encephalopathy, Neonatal hypotonia, Focal tonic seizures	no
Singleton	*TBC1D24*	NM_001199107.21:c.430T>C p.(Tyr144His)	solved	Homo recessive	Decreased activity of mitochondrial complex III, Spastic tetraparesis, Status epilepticus, Decreased activity of mitochondrial complex II, Intellectual disability, profound, Focal myoclonic seizures, Focal tonic seizures	no
Singleton	*NALCN*	NM_052867.2:c.3542G>A p.(Arg1181Gln)	solved	*de novo*	Generalized tonic seizures, Atonic	yes[202]

					seizures, Epileptic encephalopathy, Atypical absence seizures, Developmental regression, Expressive language delay	
Singleton	*IQSEC2*	NM_001111125.2:c.2890-2A>G	solved	*de novo*	Focal seizures with impairment of consciousness or awareness, Intellectual disability moderate, Postnatal microcephaly	no
Singleton	*SCN1A*	NM_001165963.1:c.2983T>C p.(Phe995Leu)	solved	*de novo*	Developmental stagnation at onset of seizures, Premature birth, Intellectual disability profound, Absence seizures with eyelid myoclonia	yes[203]
Singleton	*SCN1A*	NM_001165963.1:c.1624C>T p.(Arg542*)	solved	*de novo*	Febrile seizures, Generalized tonic-clonic seizures, Focal seizures, Epileptic encephalopathy, Posterior fossa cyst	yes[19]
Singleton	*SYNGAP1*	NM_006772.2:c.1388_1393del p.(Asp463_Leu465delinsVal)	solved	*de novo*	Intellectual disability severe, Generalized hypotonia, Phimosis, Stereotypic behavior, Focal seizures	no
Singleton	*CACNA1A*	NM_001127222.1:c.7423G>A p.(Gly2475Ser)* *canonical transcript, no HGMD transcript	potentially solved	untested	Broad-based gait, Aplasia-Hypoplasia of the corpus callosum, focal seizures, focal tonic seizures, anormality of the corpus callosum, Intellectual disability, Epileptic encephalopathy	no
Singleton	*TSC2*	NM_000548.3:c.2921C>T p.(Ala974Val)	potentially solved	untested	Intellectual disability, Seizures, Dystonia, EEG with continuous slow activity, Tetraplegia/tetraparesis, Microcephaly	no
Singleton	*SCN2A*	NM_021007:c.1976G>A p.(Gly659Asp)	potentially solved	father inheritance	Early onset absence seizures, EEG with polyspike wave complexes, Absence seizures with eyelid myoclonia	no
Singleton	*GRIN1*	NM_007327.2:c.1909G>T p.(Ala637Ser)	potentially solved	untested	Autistic behavior, Stereotypic behavior, Global developmental delay	no
Singleton	*ARID1B*	NM_020732.2:c.1009G>A p.(Ala337Thr)	potentially solved	untested	Intellectual disability, Versive seizures , Epileptic spasms	no
Singleton	*NOTCH3*	NM_000435.2:c.3944G>A p.(Cys1315Tyr)	potentially solved	untested	Intellectual disability, Generalized tonic seizures, Autistic behavior, Epileptic spasms,	yes[204]

					Hypsarrhythmia, Focal seizures	
Singleton	*COL4A1*	NM_001845.4:c.4982G>A p.(Arg1661His)	potentially solved	untested	Seizures, Aphasia, EEG with centrotemporal focal spike waves, Language impairment	no
Singleton	*GPHN*	NM_020806.4:c.757G>T p.(Val253Phe)	potentially solved	untested	Absence seizures, Generalized tonic-clonic seizures	no
Singleton	*KCNQ2*	NM_172107.2:c.977A>C p.(Gln326Pro)	potentially solved	untested	Intellectual disability severe, Dystonia, Status epilepticus, Bilateral convulsive seizures, Multifocal epileptiform discharges, Postnatal microcephaly, Spastic tetraplegia	no
Singleton	*STXBP1*	NM_003165.3:c.1012C>T p.(Gln338*)	potentially solved	untested	Neonatal respiratory onset, Neonatal hypotonia, Infantile spasms, focal tonic seizures, abnormality of the pereventrivular white matter, abnormality of the cerebral ventricles, Intellectual disability severe, Epileptic encephalopathy	no
Singleton	*COL4A2/ MBD5*	NM_001846.2:c.1019C>T p.(Ala340Val); NM_018328.3:c.2979G>C p.(Gln993His)	potentially solved	untested	Intellectual disability severe, Generalized seizures, Central hypotonia, Aggressive behavior	no
Singleton	*CDKL5*	NM_003159.2:c.2490_2491insA p.(Thr831Asnfs*79)	potentially solved	untested	Multifocal epileptiform discharges, cortical dysplasia, epileptic encephalopathy	no

Table 5*. Positive cases found in the global cohort of 140 patients.*

4.1 Singleton cohort, results in details

We analyzed the singleton cohort prioritizing the variants using the HPO system.
In this cohort, we lack the inheritance family information because unfortunately the parents are not available for the major of patients. A prioritization system is necessary to reduce the 30,000 variants seen in a typical exome to a more manageable size: typically, after filtering steps it leaves more candidates (up to 1,000) than can be reasonably manually assessed. We analyzed the data using the Exomiser algorithm which ranks the remaining variants according to their predicted pathogenicity and HPO ranking. The first step prioritized following the known epilepsy genes and the second considered the best phenotype matching.

We applied the HPO system to classify the clinical indications for WES. We had a mean of 4,1 HPO terms for each patient. The majority of the patients had focal seizures (n = 25; 21%), epileptic encephalopathy (n= 24; 20%), epileptic spasms (n= 22; 19%) and intellectual disability (n=13; 11%) as the most common indications. A summary of the 19 major categories of HPO is shown in the graph below (Figure 6).

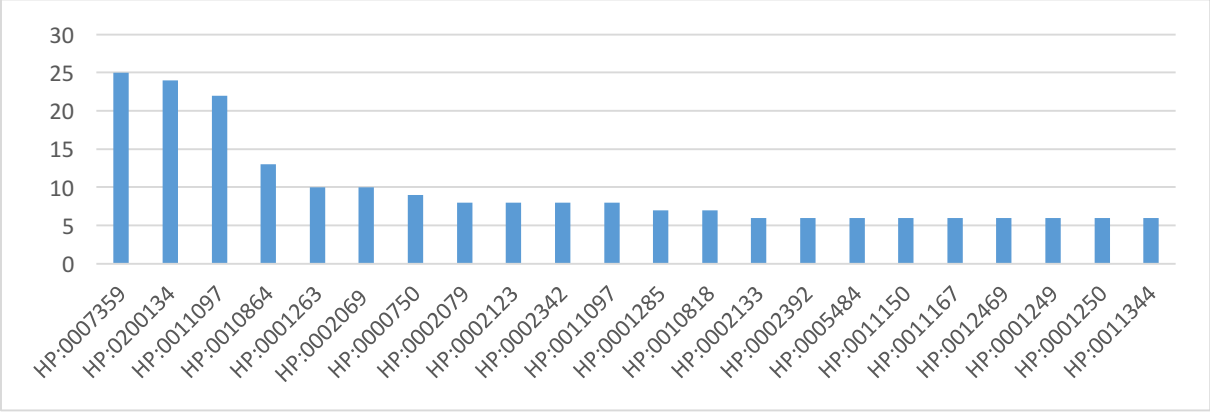

Figure 6. *Major 20 HPO categories. Legend: Focal seizures (HP:0007359), Epileptic encephalopathy (HP:0200134), Epileptic spasms (HP:0011097), Intellectual disability (HP:0010864) ,severe global developmental delay (HP:0001263), Generalized tonic-clonic seizures (HP:0002069), Delayed speech and language development (HP:0000750), Hypoplasia of the corpus callosum (HP:0002079), Generalized myoclonic seizures (HP:0002123), Intellectual disability moderate (HP:0002342), Epileptic spasms (HP:0011097), Spastic tetraparesis (HP:0001285), Generalized tonic seizures (HP:0010818), Status epilepticus (HP:0002133), EEG with polyspike wave complexes (HP:0002392), Postnatal microcephaly (HP:0005484), Myoclonic absences (HP:0011150), Focal tonic seizures (HP:0011167), Infantile spasms (HP:0012469), Intellectual disability (HP:0001249), Cortical Dysplasia (HP:0001250), Severe global developmental delay (HP:0011344).*

Analyzing the 117 patients using the HPO approach. We identified the underlying pathogenic variants in 18 patients (15%), 15 *de novo* mutations (83%) and 3 compound heterozygous mutations state (17%). The most mutated gene resulted to be *PURA*, with 3 *de novo* mutations in 3 unrelated patients.

Other recurrently mutated genes included: *SCN1A*, *SCN2A* and *TBC1D24* (with two patients each) and *GNAO1, HCN1, IQSEC2, KCNB1, MEF2C, NALCN, PDHA1, ST3GAL5, SYNGAP1, UBE3A* (with one patient each). If we also consider the variants of likely pathogenic significance, possibly explaining the clinical symptoms of the index patient without parents available to test, we had a 26% (30/117) of molecular yield (Figure 7). The proportion of mutations identified in our cohort might be slightly underestimated because all patients were already screened by panel analysis. We can also found 2 mutations in *SCN1A, SCN2A* genes already screened with sanger sequencing for all these patients, it is due to technical limitations of this technique.

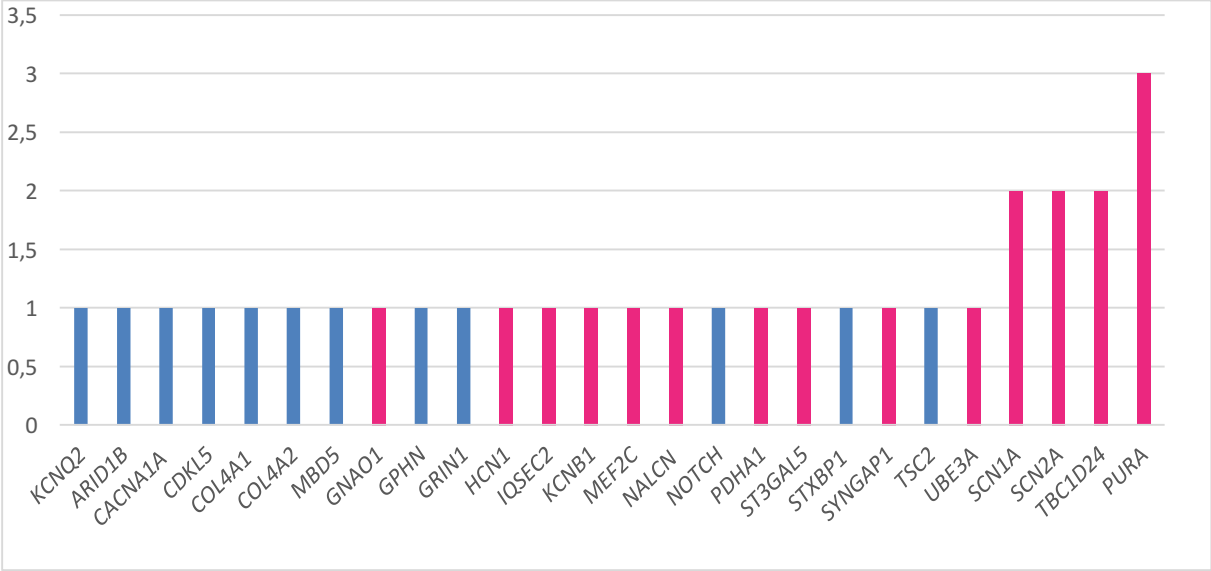

Figure 7. *All likely pathogenic variants found in singleton cohort. In pink the pathogenic variants confirmed by family segregation.*

We found 3 *de novo* mutations in the *PURA* gene in 3 unrelated patients, 2 of which resulted to be novel mutations and were deletions. The third one was a duplication (c.768dup), and already reported in the HGMD database, and associated to hypotonia and developmental delay (Table 6).
We found three *de novo* mutations in heterozygous state. The 3 patients showed common clinical features of developmental delay and hypotonia. The two patients with novel mutations showed more severe phenotype with also epileptic encephalopathy.

Patient	Gene	Mutation	Inheritance	Phenotype	In HGMD
1	*PURA*	NM_005859.4:c.768dup. p.(Ile257Hisfs*37)	*de novo*	Intellectual disability, Tetraparesis, EEG with spike-wave complexes, Seizures, Generalized hypotonia	yes[201]
2	*PURA*	NM_005859.4:c.675_676del p.(Val226Glyfs*67)	*de novo*	Intellectual disability severe, Tetraparesis, Generalized tonic seizures, Epileptic encephalopathy, Infantile axial hypotonia, Uni and bilateral multifocal epileptiform discharges , Focal seizures	no, novel
3	*PURA*	NM_005859.4:c.649del p.(Glu217Serfs*8)	*de novo*	Severe global developmental delay, Epileptic encephalopathy, Neonatal hypotonia, Focal tonic seizures	no, novel

Table 6. *Causative mutations in PURA gene.*

PURA encodes for the Pur-α, a highly conserved multifunctional protein that has an important role in normal postnatal brain development in animal models. The associated human phenotype of *de novo* heterozygous mutations in this gene is variable, but moderate to severe neurodevelopmental delay and learning disability are common to all. Neonatal hypotonia, early feeding difficulties and seizures, or 'seizure-like' movements, were also common. Additionally, it is suspected that anterior pituitary dysregulation may be within the spectrum of this disorder. Psychomotor developmental outcomes appear variable between patients, a more severe phenotype seems to be connected to a disruption of Pur repeat III[205].

In the Figure 8, we report a schematic representation of the structure of the Purα protein with mutations reported in literature (HGMD and Clinvar), our mutations and the variants present in gnomaAD database as negative control. To ensure that each variant in this study was based on the same *PURA* transcript, we focused on the Uniprot canonical transcript (http://www.uniprot.org/) (The UniProt Consortium 2017). We focused on a single transcript for *PURA*: (Uniprot Q00577; ENST00000331327.4; NM_005859.4) to cluster the variants in *PURA*.

We represent the domains of the protein following the Uniprot indications: the N-terminus containing a glycine-rich domain that contains a stretch of 18 glycine residues interrupted only by a single serine (Gly_D), the central DNA-binding domain containing 3 class of first repeats (I) and 2 class of second repeats (II), the "psycho" domain that has homology to polyomavirus large-T antigen from SV40, JCV or BKV and other proteins (Psyco); the C-terminal glutamine-rich domains (Gln_D), and the glutamate-rich domains (Glu_D)[206].

We can appreciate that almost all variants cluster within or very close to the fuctional domains. We can also note that our two novel deletions (NM_005859.4:c675_676del p.(Val226Glyfs*67); NM_005859.4:c649del p.(Glu217Serfs*8) cluster within the two last repeats of the DNA-binding domain.

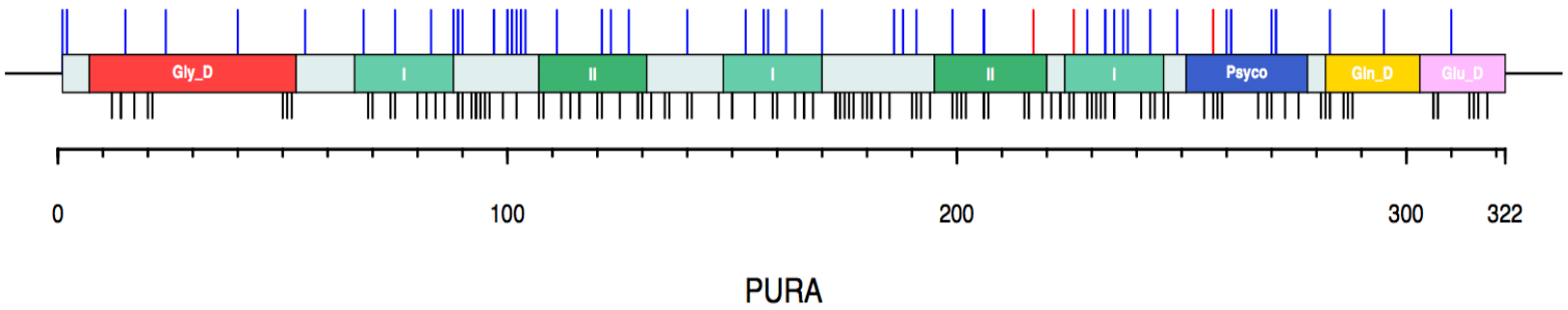

Figure 8. *Schematic representation of the structure of the Pura protein. The N-terminus contains a glycine-rich domain (Gly_D, red) that contains a stretch of 18 glycine residues interrupted only by a single serine. The central DNA-binding domain containing 3 class I repeats (I, aquamarine) and 2 class II repeats (II, light green). The "psycho" domain has homology to polyomavirus large-T antigen from SV40, JCV or BKV and other proteins (Psyco, blue).The C-terminal glutamine-rich (Gln_D, yellow), and glutamate-rich domains (Glu_D, pink)*[206] *are also shown . On the bottom in black are represented the GNOMAD control, on the top are reported in red our mutations, and in blue mutations already reported in HGMD and in Clinvar as pathogenic.*

In the singleton cohort we identified pathogenic variants in 15% (18/117) of patients and 17% (3/18) of these are in recessive genes: *ST3GAL5* (ST3 Beta-Galactoside Alpha-2,3-Sialyltransferase 5) and *TBC1D24* (TBC1 Domain Family Member 24).

We identified two heterozygous variants in *ST3GAL5* (NM_003896.3:c.1000C>T p.(Arg334*); NM_003896.3:c.1166A>G p.(His389Arg)) in a patient with severe intellectual disability, focal seizures, dyskinesia, hyperpigmented skin macules and stereotypic behavior. We tested the inheritance of variants in the family with Sanger sequencing and resulted one inherited form the father and one from the mother, in a classical compound heterozygous state.

Ganglioside GM3 is known to participate to the induction of cell differentiation, modulation of cell proliferation, maintenance of fibroblast morphology, signal transduction, and integrin-mediated cell adhesion. *ST3GAL5* codes for a protein that is a type II membrane protein which catalyzes the formation of GM3 using lactosylceramide as the substrate. The encoded protein is a member of glycosyltransferase family 29 and may be localized to the Golgi apparatus. Mutations in *ST3GAL5* have been associated with Amish infantile epilepsy syndrome. It is an autosomal recessive neurocutaneous disorder characterized by infantile onset of refractory and recurrent seizures associated with profoundly delayed psychomotor development and/or developmental regression as well as abnormal movements and visual loss[207]. Affected individuals develop hypo or hyperpigmented skin macules on the trunk, face, and extremities in early childhood. According with literature, we can affirm that the variants we identified can explain the complex phenotype of the patient.

In addition, in two unrelated patients we identified mutations in the *TBC1D24* gene. *TBC1D24* codes for a protein with a conserved domain, referred to as the TBC domain, characteristic of proteins which interact with GTPases. TBC domain proteins may serve as GTPase-activating proteins for a particular group of GTPases, the Rab (Ras-related proteins in brain) small GTPases which are involved in the regulation of membrane trafficking. Mutations in this gene are associated with familial infantile myoclonic epilepsy, early infantile epileptic encephalopathy 16[208].

We identified a variant (NM_001199107.21:c.430T>C p.(Tyr144His)) in an homozygous state in a patient with decreased activity of mitochondrial complex II and III, spastic tetraparesi, status epilepticus, profound intellectual disability and focal myoclonic seizures.

Sanger sequencing parent's segregation revealed that both parents were heterozygous for the variant, according with the recessive pattern of inheritance. A second patient was compound heterozygous for two *TBC1D24* variants (NM_001199107.1:c.671A>G p.(Tyr224Cys) and NM_001199107.1:c.1008delT p.(His336Glnfs*12) and presented cortical myoclonus seizures,

abnormal EEG pattern with polyspike wave complexes, intellectual disability and obtundation status. According with literature, the *TBC1D24* mutations can explain the severe phenotype of these patients.

4.2 Trios result in details

In the trios cohort, we identified causative mutations in the *GABRA1, ATP6V1A, CACNA1E* and *DMXL2* genes, with a mutation yield of 20% (4/20). The *DMXL2* has been identified in a recessive quartet.

GABRA1

We found a *GABRA1* variant (Gamma-Aminobutyric Acid Type A Receptor Alpha1 Subunit) in a patient, prematurely dead at the age of 6 years old. She presented neonatal onset epileptic encephalopathy and severe developmental delay. The variant: NM_000806.5c.865A>C(p.Thr289Pro) is predicted to be damaging by bioinformatics tools as Mutation taster, SIFT and Poliphen and it is not present in all human database (1000genome, EXAC and gnomAD). The variant was confirmed by Sanger sequencing and confirmed to be *de novo* variant. This gene encodes a gamma-aminobutyric acid (GABA) receptor. GABA is the major inhibitory neurotransmitter in the mammalian brain where it acts at $GABA_A$ receptors, which are ligand-gated chloride channels. Chloride conductance of these channels can be modulated by agents such as benzodiazepines that bind to the $GABA_A$ receptor. $GABA_A$ receptors are pentameric, consisting of proteins from several subunit classes: alpha, beta, gamma, delta and rho. Mutations in this gene cause juvenile myoclonic epilepsy and childhood absence epilepsy type 4. Recently, *GABRA1* gene mutations have been associated to a severe phenotype: epileptic encephalopathy early infantile.
The variant we found is very near to another mutation that create a change in electric function of the channel already described in literature[136]. This might suggest that our mutation create a similar electric change in the channel. In a recent paper of Johannesen K et al, our mutation was been subjected to functional analysis using the Xenopus laevis oocyte expression system. Revealing a loss of function, without a clear genotype-phenotype correlation[199] . According with literature, we can affirm that the variant we identified can explain the complex phenotype of the patient.

ATP6V1A

In a 12 years old girl with late infantile EE characterized by heterogeneous seizures including spasms, dystonic quadriplegia, dyskinesia, severe developmental delay, and congenital microcephaly we identified a variant in *ATP6V1A* (ATPase H+ Transporting V1 Subunit A) gene. This heterozygous variant NM_001690.3:c.298G>T (p.Asp100Tyr) resulted to be *de novo*. The variant is predicted to be damaging by bioinformatics tools as Mutation taster, SIFT and Poliphen and it is not present in human polymorphisms databases (1000genome, EXAC and gnomAD).

Mutations' pathogenicity of *ATP6V1A* variants has been evaluated through in silico prediction using the dbNSFP database (v3.0a), which provides functional prediction scores on more than 20 different algorithms (https://sites.google.com/site/jpopgen/dbNSFP). To assess the effects of missense *ATP6V1A* substitution, we used both the dbNSFP ensemble rank scores MetaSVM and MetaLR[192]. In addition, we used the scores obtained from Revel[193], M-CAP[194] and Eigen[195], three different bioinformatics tools that evaluate the pathogenicity of rare variants.

ATP6V1A encodes a component of vacuolar ATPase (V-ATPase), a multisubunit enzyme that mediates acidification of eukaryotic intracellular organelles. V-ATPase dependent organelle acidification is necessary for such intracellular processes as protein sorting, zymogen activation, receptor-mediated endocytosis, and synaptic vesicle proton gradient generation. V-ATPase is composed of a cytosolic V1 domain and a transmembrane V0 domain. The V1 domain consists of three A and three B subunits, two G subunits plus the C, D, E, F, and H subunits. The V1 domain contains the ATP catalytic site. The V0 domain consists of five different subunits: a, c, c', c", and d. This encoded protein is one of two V1 domain A subunit isoforms and is found in all tissues. Diseases associated with *ATP6V1A* include Histoplasmosis and Baylisascariasis. Pathways related to *ATP6V1A* include insulin receptor recycling and transport of glucose and other sugars, bile salts and organic acids, metal ions and amine compounds. This gene is also associated in HGMD database with cutis laxa[209].

ATP6V1A is expressed in the brain, in particular in hippocampus, and it is involved in neurotransmitter release and has been linked to schizophrenia. The mutation we identified has been included in a paper recently submitted where three *ATP6V1A* mutations have been studied in patients with severe forms of epilepsy. In this paper we suggest *ATP6V1A* as a new epilepsy candidate gene with relevant neurobiological functions.

CACNA1E

The patient with *CACNA1E* variant is a 9 years old female, with early onset epileptic encephalopathy, severe intellectual disability, spastic tetraparesis and movement disorder.

We followed the patient for several years and we performed different genetic tests. We performed as a first test a panel of 35 genes and we identified a variant in *SLC2A1 (Solute Carrier Family 2 Member 1)* NM_006516.2:c.1001G>A p.(Arg334Gln). However, it resulted inherited from her mother and it couldn't explain the complex phenotype of patient.

We decided to re-analyze the patient with larger panel of 95 genes, and we identified a variant in the *DEPDC5 (DEP Domain Containing 5)* gene (NM_001242896.1:c.194-3T>C), but the variant resulted inherited from her unaffected father.

Both *SLC2A1* and *DEPDC5* variants couldn't explain the complex phenotype of the patient and have been classified as variant of uncertain significance (VOUS).

To further investigate this patient we decided to perfom by WES, analysis that lead to identification of a mutation in the *CACNA1E* gene (Calcium Voltage-Gated Channel Subunit Alpha1 E) NM_001205293.1, c.2104G>A p.(Ala702Thr) that resulted *de novo* and was confirmed by Sanger sequencing. The variant is predicted to be damaging by bioinformatics tools as Mutation taster, SIFT and Poliphen. It is not present in human variants databases (1000genome, EXAC and gnomAD).

This gene encode for the Calcium Voltage-Gated Channel Subunit Alpha1 E. Voltage-dependent calcium channels are multi-subunit complexes consisting of alpha-1, alpha-2, beta, and delta subunits in a 1:1:1:1 ratio. These channels mediate the entry of calcium ions into excitable cells, and are also involved in a variety of calcium-dependent processes, including muscle contraction, hormone or neurotransmitter release, gene expression, cell motility, cell division and cell death. The alpha-1E subunit of the R-type calcium channels, which belong to the 'high-voltage activated' group that maybe involved in the modulation of firing patterns of neurons important for information processing.

This gene has not been demonstrated to be associated to specific phenotypes, despite in literature there are functional studies that correlate this gene to epilepsy[210] and in a recent paper mutation in this gene are described in a patient with epileptic encephalopathy[151]. For this reason, we can suppose a causative role of this variant in the complex phenotype of our patient. This case underlines a long clinical procedure to identify the real cause of heterogeneous phenotype, and the strength of WES to find a new causative gene.

DMXL2

We identified two variants in the *DMXL2* (Dmx Like 2) gene in a patient belonging to a recessive quartet, and his affected sister.

The siblings, shared a very similiar phenotype characterized by early onset epileptic encephalopathy, severe intellectual disability, spastic tetraparesis, both died prematurely.

We performed the whole exome sequencing only in parents and their daughter, because the DNA sample of the boy didn't overcome preliminary quality controls.

We analyzed the data with all inheritance models. With recessive model, we identified two variants in *DMXL2* gene in the affected daughter resulted inherited one from the father and one from the mother. We confirmed and tested by Sanger sequencing the two variants in all the quartet and they resulted present in both siblings, one inherited from the mother and one from the father, as expected in the compound heterozygous recessive model.

This *DMXL2* gene encodes a protein with 12 WD domains. Proteins with WD domains are involved in many functions including participation in signal transduction pathways. Encoded protein in regulation of the Notch signaling pathway (promotes neurogenesis, neural differentiation, role in the regulation of embryonic development) has been demonstrated in vitro using several human cell lines[211].

The hortologue of this gene as Rav1in yeast, Rabconnectin3a (Rbcn-3a), Rbcn3a in Zebrafish and Dmxl2 in mouse.

The *DMXL2* gene encodes the synaptic protein rabconnectin-3a which has been identified as a putative scaffold protein for Rab3-GAP and Rab3-GEP, two regulators of the GTPase Rab3a.

Regulator of the H^+-ATPase of the vacuolar and endosomal membranes (RAVE) would play an important role in adjusting V-ATPase activity[212]. When the outer environment, specifically the pH and ionic salt concentration, of living cells changes, RAVE can regulate V-ATPase assembly by separating or combining the V1 and V0 domains[213]. Currently, it is generally accepted that RAVE is a novel protein complex composed of the following subunits: Skp1p (~ 22 kDa), Rav1p (~ 155 kDa) and Rav2p (~ 40 kDa). In the RAVE complex, Rav1p binds to Rav2p and Skp1p, while there is no direct interaction between Rav2p and Skp1p[214]. When RAVE regulates reversible V-ATPase assembly, it ordinarily binds to subunit E or G of V1[215]. Active RAVE can contribute to an efficient assembly of the C subunit in V-ATPase, whereas with RAVE deficiency, the C subunit is not able to bind stably to V-ATPase[214].

In literature correlation of *DMXL2* and epilepsy has not been reported yet, but there is several evidence that suggest an important role of this gene in severe epilepsy. Indeed *DMXL2* interacts with the family of *RAB3GAP* gene, that is already known to be associated with epilepsy and with the V-ATPases (H^+ ATPases) complex, (ATP-dependent proton pumps that regulate pH homeostasis in virtually all eukaryotes) of interest, *DMXL2* interacts with *ATP6V1A*, a new epilepsy candidate gene with relevant neurobiologic functions (Figure 9). These data suggest a causative role of *DMXL2* in the complex phenotype of our patients.

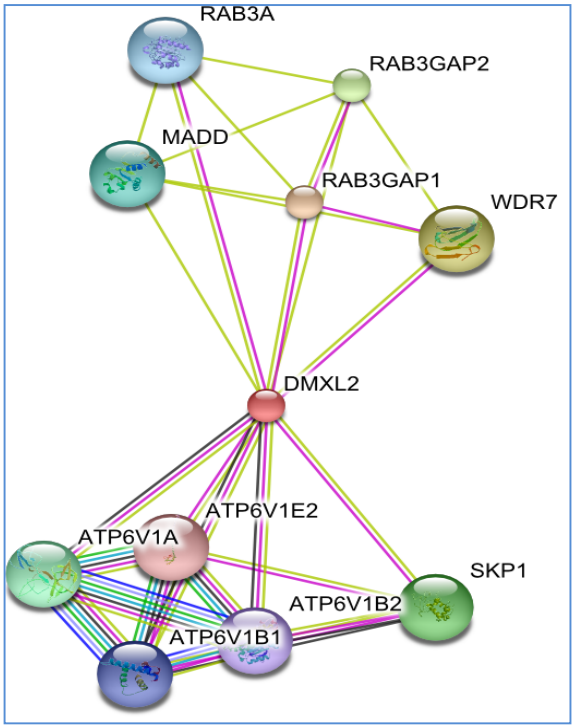

***Figure 9**. Interacting Proteins for DMXL2 Gene: STRING Interaction Network Preview.*

5. Discussions

Since 2010, whole-exome sequencing (WES) emerged as a powerful tool for the identification of genetic variants associated with Mendelian conditions[216].

To date, reported WES detection rates for deleterious variants in rare and heterogeneous genetic disorders showed a diagnostic yield between 25 and 31%, making it by far the most powerful individual diagnostic test for such conditions, especially in a context of no clear clinical diagnosis[217,218]. Other studies that used different classification criteria and highly selected phenotypes or populations reported higher numbers[219,220].

Here we reported the analysis of 140 cases with heterogeneous form of severe epilepsy and family structure. Altogether, we found pathogenic or likely pathogenic variants in 34 of the 140 families; an overall molecular diagnostic yield of 24%.

The diagnostic rate of WES we obtained is a little lower the common molecular yield reported in literature, but we have to take into account, in order to compare our detection yield among studies previously reported in literature, that other studies with a patient cohort similar to our cohort don't. For patients of our cohort, WES was not the first choice of analysis, since single gene testing and big panel for epilepsy genes have already been performed. For this reason, our molecular yield resulted underestimated in comparison with other studies. However, using WES we were able to resolve very complex cases and help patients that had been undergone into a long diagnostic iter, where classical genetic diagnosis based on single gene or panel sequencing often led to negative or inconclusive results, and would have improved life quality or even benefit from medical treatment, if diagnosis was provided earlier.

To better understand our results, we divided our patients into two different cohorts according to the approach of analysis used to study the patients and we compared the diagnostic yield of the two cohorts.

In singleton cohort, we analyzed 117 WES cases with heterogeneous phenotypes of severe epilepsy. We identified pathogenic variants in 18 patients. If we considered also the likely pathogenic variants, we detected 30 variants. The diagnostic yield in this cohort ranges between 15% (18/117) if we consider only the pathogenic variants, and 26% (30/117) if we consider also the likely pathogenic variants. In this cohort, we found pathogenic variants in known epilepsy genes, most of these muations consisted in single-hits in single genes, with the exception of mutations in four genes (*PURA, SCN1A, SCN2A, TBC1D24*), in which we had recurrent mutations.

This aspect is not surprisingly because we used a singleton approach for diseases exhibiting considerable locus heterogeneity, with mutations in many different genes. As reported by the Epi4K Consortium, the identification of long list of single-hit genes is very common for these type of phenotypes[221].

In this cohort, SCN1A and SCN2A were the two second most recurrently mutated genes, despite many of the patients had already been screened for these genes. However, it is not surprisingly, since the patients had been analyzed in the past with old molecular techniques that lead to lose some mutations for technical limitations[222].

Surprisingly, *PURA* resulted to be the most mutated gene. We identified three de novo mutations in three unrelated patients that share the same phenotypic features: intellectual disability, hypotonia and seizures. Two of these mutations resulted novel and both were deletions (c.675_676del and c.649del) leading to frameshift mutations (p.(Val226Glyfs*67) and p.(Glu217Serfs*8)). The patients with the deletions showed a severe phenotype with severe global developmental delay and epileptic encephalopathy. The two novel variants (p.(Val226Glyfs*67) and p.(Glu217Serfs*8)) cluster very close one to each other and are located in the central DNA-binding domain, in particular in the class I. PUR domains repeats, which are highly conserved throughout evolution[206]. Probably, the clusterization in this conserved domains can explain the severe phenotype of these patients. The third mutation (c.768dup. p.(Ile257Hisfs*37)) is located indeed in the "*psycho*" domain. This domain has homology to polyomavirus Large-T antigen from SV40, JCV or BKV and other proteins[206]. The identification of three additional unrelated patients with *de novo* mutations in *PURA* provide further evidence that *PURA* is largely, if not solely, responsible for the developmental delay, hypotonia, and seizures observed in the 5q31.3 microdeletion syndrome[223]. Our results expand the number and location of mutations in *PURA* and the associated phenotypes.

Another interesting result of the singleton cohort analysis, was the identification of pathogenic variants in recessive genes. We were able to identify these variants even without the family inheritance information, since with our prioritization method the variants were located in highest position of importance ranking. This aspect was unexpected but probably, thank to a detailed HPO terms correlation, we were able to identify also variants in recessive gene without segregation information.

Our results indicate that the singleton approach is a low-cost, fast but also efficient strategy for the identification of variants across different phenotypes. We can conclude that the singleton approach, using a HPO prioritization approach, is a promising diagnostic tool that can help in the diagnostic screening of patients with complex phenotypes resulted negative to other previously tests.

In the trios cohort, we obtained a molecular yield of 20%, 75% in dominant genes and 25% in recessive genes. In this cohort, we obtained a higher yield in comparison with the singleton cohort. We can explain this difference because in this cohort the phenotype of patients is more uniform (all patients present severe epileptic encephalopathy) and also because we have the family information that help in the identification of causative variants. The molecular yield of trios cohort obtained in our study is underestimated compared with the literature, but we have to consider that our patients have already been screened with Sanger sequencing and epilepsy genes panels.

In the trios cohort, we identified mutations in two known epilepsy genes: *GABRA1* and *CACNA1E* and mutations in two new causative epilepsy genes: *ATP6V1A* and *DMXL2*. The most interesting aspect of the trios approach is the possibility to discover new causative genes and new pathways involved in epileptic encephalopathy.

We found a heterozygous novel variant in the *ATP6V1A* gene (NM_001690.3:c.298G>T p.(Asp100Tyr)) in a patient presenting late infantile epileptic encephalopathy with heterogeneous seizures. This gene is not already described associated with epilepsy and also is a gene with a very low probability to have mutations. In the context of a scientific collaboration, three additional *de novo* *ATP6V1A* mutations in patients with similar phenotypes were identified leading to 3 the total number of mutations. Functional studies to prove the role of *ATP6V1A* mutations, including the mutation identified in our patient (p.Asp100Tyr), in epileptic encephalopathy were also performed. The experimental work on the p.Asp100Tyr missense substitutions indicates that *ATP6V1A* mutations cause developmental encephalopathies with epilepsy through a novel pathomechanism that includes both alteration of lysosomal homeostasis and the ensuing abnormal synaptogenesis. The main function of v-ATPase is proton transport and acidification of intracellular organelles, in particular lysosomes that have critical requirements for low pH. The functional experiments performed on the p.Asp100Tyr mutation, causative for a severe phenotype, demonstrated an increased degradation with impaired expression of ATP6V1A, and a defective lysosomal maturation with no significant impairment of endo-lysosomal pH. This suggests that the p.Asp100Tyr mutation induces activation of degradative pathways in cells and that loss of expression with consequent loss of function may be at the basis of the severe phenotype observed in this patient. These findings were supported by structural modeling results, suggesting that the p.Asp100Tyr mutation decreases protein stability. The modelling studies on the available prokaryotic and eukaryotic structures of v-ATPase predicted the p.Asp100Tyr (Asp85 in the model) mutation to cause steric hindrance and destabilize protein folding, necessary for proton transport. The mutation (p.Asp100Tyr) falls into the A subunit of the V1 domain (Figure 10A) and makes a hydrogen bond with Arg89, a guanidium group which makes van der Waals contacts with adjacent hydrophobic side chains (Figure 10B/C).

As a result, the side chain of Asp85 is closely surrounded by the hydrophobic residues, and its replacement with a bulkier tyrosine residue quite likely causes steric hindrance and destabilizes the protein folding. This is supported by the FoldX calculation, which predicted a significant increase in free energy (about 13 kcal/mol) upon the Asp85Tyr mutation.

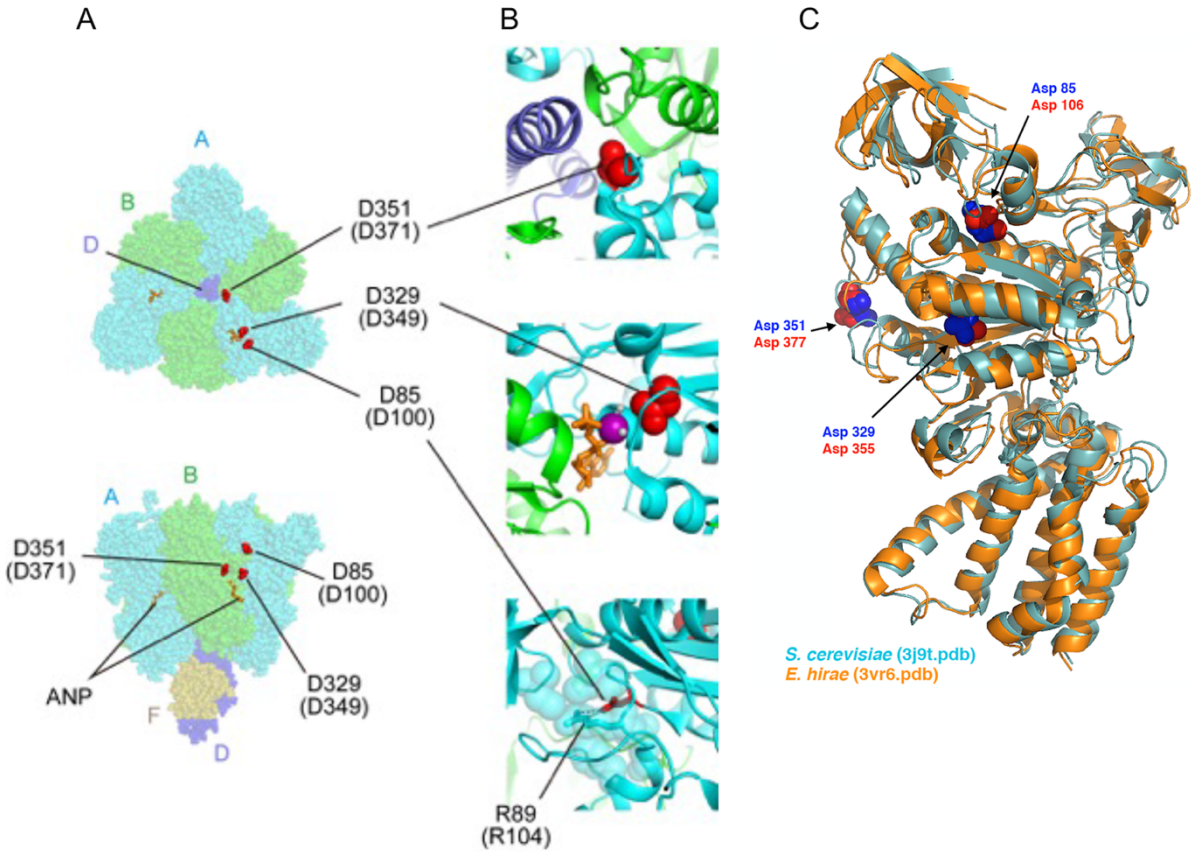

Figure 10. *Structural mapping of the mutation in the V-ATPase. Structural mapping of the mutation in the V-ATPase.* ***A****, Crystal structure of the V1 domain from Enterococcus hirae V-ATPase in a nucleotide-bound state (PDB code 3VR6), viewed from the extracellular side (upper) and the membrane plane (bottom), shown as the sphere representation. The A, B, D and F subunits are colored in cyan, green, blue and brown, respectively, and residues at the mutation sites are colored in red. The non-hydrolysable ATP analogue ANP (phosph-oaminophosphonic acid-adenylate ester) is depicted in orange sticks.* ***B,*** *Magnified views of the mutation sites presented in the ribbon model. Asp85 is depicted as red sticks, and Arg89, which makes a salt bridge with Asp85, is shown as sticks with translucent spheres. Some side chains of hydrophobic residues around Arg89 are shown in translucent spheres, and magnesium ion and its coordinated water molecules are depicted as a purple sphere and small gray dots, respectively. Black dotted lines indicate hydrogen bonds. Amino acid numbers in parentheses correspond to those of human ATP6V1A.* ***C****, Crystal structures of the A subunit from Enterococcus hirae (orange) and the cryo-EM structure of the A subunit from Saccaromyces cerevisiae (cyan, PDB code 3J9T) V-ATPase. Mutation sites are shown as spheres and colored red for S. cerevisiae and blue for E. hirae.*

When modeled in hippocampal neuronal cells, these pathogenic mutations produced the same significant and divergent effects on lysosomal biogenesis and endo-lysosomal pH that were observed in cell lines and probands lymphoblasts, demonstrating their impact also in central neurons. Moreover, in neuronal cells, p.Asp100Tyr mutant induced a dramatic and comparable defect in dendrite maturation and excitatory synapse formation. The phenotype of hippocampal neurons overexpressing the *de novo ATP6V1A* pathogenic variants demonstrated a previously unexplored role of the v-ATPase in the processes of neuronal development and connectivity, and suggests that any alterations in lysosomal homeostasis can bring about impairments in dendritogenesis and excitatory synapses formation. This hypothesis is supported by recent reports on novel roles of lysosomes at synapses for the turnover of synaptic proteins as well as for the plasticity of dendritic spines[224,225]. The loss of function mutation (p.Asp100Tyr) determined a comparable neuronal phenotype, suggesting that either loss of lysosomal structures or decreased endo-lysosomal pH could alter enzymatic function and result in common defects in dendrite maturation and synapse formation (Figure 11). In agreement with this hypothesis, it has been recently reported that inhibition of lysosomal function by the protease inhibitor leupeptin resulted in an analogous decrease of excitatory inputs in rat hippocampal cultures[226]. Although the precise molecular mechanisms through which the described *de novo* mutations result in a developmental encephalopathy with epilepsy remain to be clarified, these studies demonstrate that maintenance of lysosomal homeostasis is necessary to fulfill the complete neuronal developmental program and to establish a mature network connectivity. Alteration in the intracellular signaling pathways that regulate lysosomal homeostasis may therefore represent a novel pathogenetic mechanism of epileptogenic developmental encephalopathies. In conclusion, the functional studies can show that *de novo* heterozygous *ATP6V1A* mutations cause a developmental encephalopathy with epilepsy with a pathomechanism that involves the effects of v-ATPase in lysosomal homeostasis and neuronal connectivity. While the finding reveals a strong link between alterations of lysosomal homeostasis and neuronal development, it also proposes a novel role for v-ATPase in the formation and maturation of neural networks.

Interestingly we identified two variants with a compound heterozygous recessive model in the *DMXL2* (Dmx Like 2) gene in a patient belonging to a recessive quartet.

The siblings, shared a very similar phenotype characterized by early onset epileptic encephalopathy, severe intellectual disability, spastic tetraparesis. Both children died prematurely. In literature, correlation between *DMXL2* and epilepsy has not been reported yet, but there is several evidence that suggest an important role of this gene in severe epilepsy. The most interesting evidence is the interaction of *DMLX2* with the V-ATPases (H^+ ATPases) complex, (ATP-dependent proton pumps that regulate pH homeostasis in virtually all eukaryotes), in particular *DMXL2* interacts with *ATP6V1A,* a new epilepsy candidate gene with relevant neurobiologic functions, as we described

above. Thanks to an international call, we found an additional patient from Israel with a heterozygous compound *DMXL2* mutation and a similar phenotype. We are setting functional collaborative studies to clarify the role of this gene in the genesis of this severe phenotype and also to understand better the interaction pathway with the V-ATPases (H^+ ATPases) complex.

We can suggest WES trios approach as a first diagnostic choice if there is no clear differential diagnosis because permit to solve difficult cases in a short time. In addition, the WES trios approach represents a powerful method to discover new causative genes and also new pathogenic pathways that lead to very severe forms of epileptic encephalopathy.

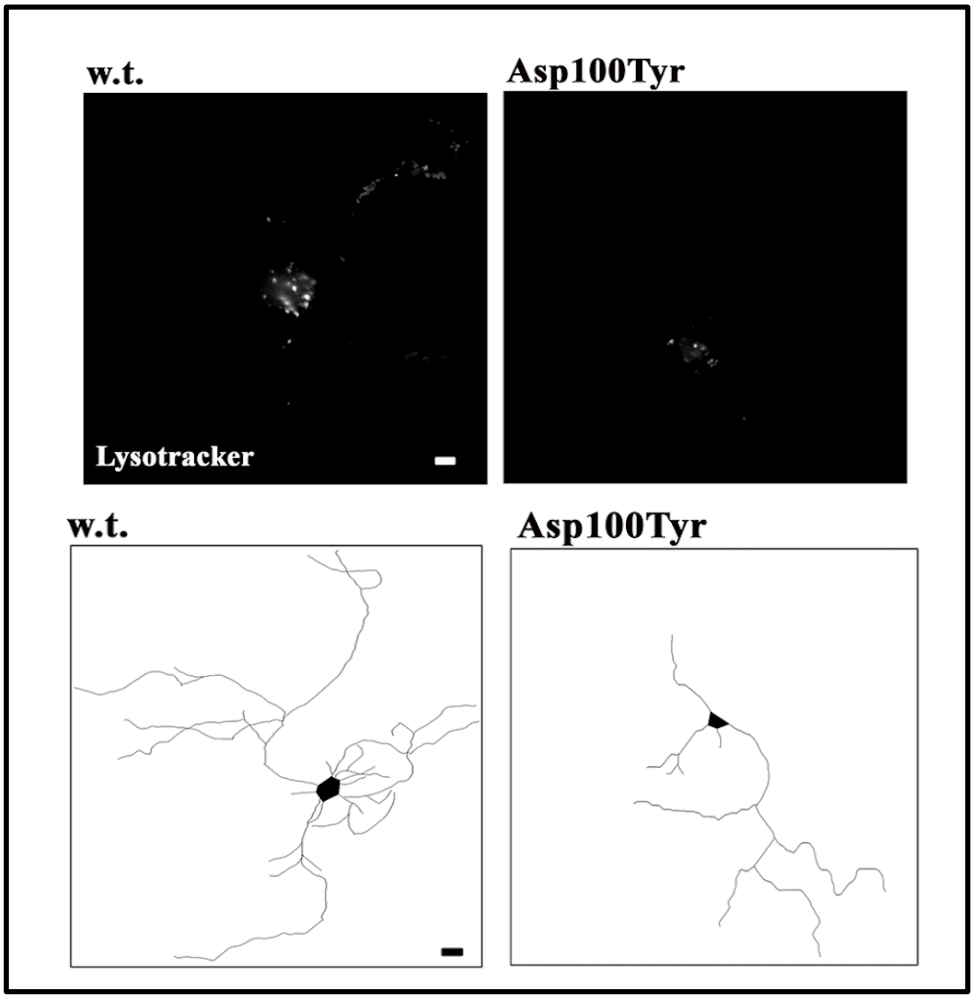

Figure 11. *Effect of ATP6V1A mutation on neuronal development. In the top of figure, a representative images of rat hippocampal neurons transfected with wild-type ATP6V1A (w.t.) and Asp100Tyr ATP6V1A (Asp100Tyr) variants at 7 DIV and incubated with LTR (50 nM, 30 min) at 10 DIV; in the bottom of figure, a representative traces of 10 DIV rat hippocampal neurons transfected.*

5.1 Therapeutic approach

Genetic testing with rapid turnaround would allow timely personalized treatment and, when possible, avoid needless investigations and unnecessary treatment in benign conditions. All these factors may contribute to better management, even when the impact of each one alone is limited. The advances in the genetics and neurobiology of the epilepsies are establishing the basis for a new era in the treatment, focused on each individual and their specific epilepsy, even later in life. Genetic counseling can be arranged once the genetic architecture and mode of inheritance of the epilepsy are fully understood. The discovery of a causative gene defect associated with a non-progressive course may reduce the need for further diagnostic investigations in the search for a progressive disorder at the biochemical and imaging level and allow more timely and straightforward treatment choices for specific conditions as well as avoiding needless investigations and inappropriate or unnecessary choices. There is, however, the need for a close collaboration between the geneticists and epileptologists to ensure the proper management of genetic investigations in patients with epilepsy. Disease-specific treatments are currently available for only a minority of genetic epilepsies whereas for the large majority of patients, treatment options comprise the usual drug armamentarium that do not address the underlying biological mechanism. In addition, epileptologists well known that there is high variability in the response to antiepileptic treatment and that several factors may contribute to such variability. Variation in response to antiepileptic drug treatment may arise from genetic variation in a range of gene categories, including genes affecting drug pharmacokinetics, and drug pharmacodynamics, but also genes held to actually cause the epilepsy itself[227]. There are a few examples in which the discovery of the genetic defect underlying a particular form of epilepsy may explain, in whole or in part, the response, whether positive or negative to certain anti-epileptic drugs (AEDs).

Dravet syndrome, which is caused by mutations of the gene encoding the sodium channel (*SCN1A*) can be worsened by treatment with sodium-blocking drugs, such as carbamazepine and phenytoin, which should therefore be avoided[228]. Controversy is the use of lamotrigine (LTG) that works as a sodium channel and N-type calcium channel blocker. Lamotrigine has been reported to cause seizure exacerbation, which has led to avoid ance of its use in patients with Dravet syndrome, but has been reported also a positive effect in some patients with Dravet Syndrome. This beneficial effect maybe explained with the action of LTG that involve hyperpolarization-activated cyclic nucleotide gated channels[229].Conversely, sodium blockers are considered first-choice therapy on epilepsy syndrome related to mutation of the gene *SCN8A* (another sodium channel gene) and *KCNQ2*[230].

Heterozygous *KCNQ2* gene mutations have been found in patients with a continuum of overlapping epileptic disorders of neonatal onset ranging from very benign familial neonatal seizures to severe EE with refractory seizures, an EEG pattern of burst suppression and severe developmental delay[231,232]. *KCNQ2* encodes the potassium voltage gated channel subfamily KQT member 2, also known as Kv7.2, and mutations in this gene cause a variable loss of function[233,234]. Recent reports have described several patients with *KCNQ2*-EE with seizures fully controlled by either phenytoin or carbamazepine, whose mechanism of action consists of blocking the flux of Na^+ ions through channels during propagation of the action potential, thus preventing the development of seizure activity[235]. Voltage-gated Na^+ channels and KCNQ potassium channels colocalize and are bound at critical locations of the neuronal membrane[236]. The response to Na-channel blockers in patients with K channel disorders could be explained by structure–function approaches showing that modulation of one channel may significantly impact the function of the channel complex[237]. Retigabine, which is not yet approved for use in the paediatric population and now apparently withdrawn from commerce, acts as a neuronal KCNQ/kv7 K channel opener and might provide a targeted treatment for *KCNQ2*-EE[238,239].

KCNT1 gene-related epilepsy is another example of a potential treatment based on known pathophysiological mechanisms, although initial attempts towards targeted treatment have been somewhat deceiving. Mutations in the *KCNT1* gene have been found in patients with both autosomal-dominant nocturnal frontal lobe epilepsy and malignant migrating partial seizures in infancy (MMPSI)[62,240]. The *KCNT1* gene encodes for subunits contributing to the sodium-activated K^+ current (KNa). Quinidine is an anti-arrhythmic agent that stabilizes the neuronal membrane by inhibiting the Na influx required for the initiation and conduction of impulses. Functional studies have thus shown that quinidine might counteract and reverse the dysfunction of *KCNT1*-activating mutations in patients with MMPSI. This finding prompted treatment with quinidine in the hope of achieving seizure control[135,241]. Following the initial report of effective treatment in some patients, subsequent reports[242] and personal unpublished trials in single patients with MPPS have been rather disappointing.

GLUT1-deficiency syndrome (*GLUT1*-DS) is perhaps the best example of a treatable, highly epileptogenic encephalopathy. *GLUT1*-deficiency syndrome is a genetic syndrome caused by insufficient transport of glucose from the blood to the brain which is characterized by the appearance of paroxysmal events in early childhood. Glucose is the essential fuel for brain energy, thus its impaired uptake into the brain causes hypoglycorrhachia and subsequent progressive symptoms, including microcephaly, epilepsy, movement disorders and intellectual disability, in various associations. The specific treatment is the ketogenic diet, an high fat, low-carbohydrate diet mimicking the fasting state. Nutritional fat is transformed into ketone bodies, which can be used as

metabolic substrate for the brain when lacking glucose. On ketogenic diet significant seizure reduction was seen in children with the most severe epilepsies; seizure free or reduction was seen in 38% to 100% of children on the diet[243]. Moreover, early diagnosis and prompt treatment are important for prognosis[244].

Other examples are pyridoxal 5-phosphate dependent epilepsy due to mutation of the *ALDH7A1* gene recently *PNPO* gene[245–247]. This enzyme converts aminoadipic semialdehyde (-AASA) into aminoadipate (AAA), a critical step in the lysine metabolism of the brain and is involved in the metabolism of vitamin B6 of nutritional origin (in the form of pyridoxine and pyridoxamine) in active form, pyridoxal 5-phosphate (PLP). PLP is necessary in many processes including the production of neurotransmitters[246]. This epilepsy syndrome is characterized by the onset of seizures shortly after birth, or in some cases, even intrafetal[248]. AEDs which are usually given to control seizures, are ineffective in patients with PLP-dependent epilepsy. These individuals respond instead to treatment with high dose on a daily use of PLP. If untreated can instead develop encephalopathy that can also have a fatal outcome[245,248].

Protocadherin 19 (*PCDH19*) gene-related epilepsy is an unusual X-chromosome disorder that primarily affects females. It encompasses a broad clinical spectrum from infantile EE resembling Dravet syndrome to epilepsy with or without autism spectrum disorders and intellectual disability. Recent functional studies have shown that *PCDH19* mutations lead to deficiency of the neurosteroid allopregnanolone, one of the most potent GABA-receptor modulators[249]. These findings provide evidence for a role of neurosteroids in *PCDH19*-related epilepsy, intellectual disability and autism and create realistic opportunities for targeted therapeutic interventions. Ganaxolone, a synthetic methyl analogue of allopregnanolone, has already been used to treat epilepsy and status epilepticus related to various aetiologies in open-label pilot studies, with encouraging results in some patients[250,251]; however, no randomized studies are available

Another interesting finding concerns the association between genetic mutations in the gene *DEPDC5*, a gene involved in mTOR pathway and tuberous sclerosis, and various forms of epilepsy, tumor and neurodevelopmental disease both familial and sporadic with wide phenotypic spectrum[72]. In these case a possibility is to use a specific drug known as rapamycin that act directly on mTOR pathway[252]. In particular, tuberous sclerosis complex, a disease that brings to more than 80% of drug resistant epilepsy patients other than disease in multiple organs, caused by mutations in *TSC1* and *TSC2* genes, encoding a complex with inhibitory activity against mTOR are susceptibility to rapamacin treatment. Early administration of the specific treatment can lead to better outcome not only in seizure but also in developmental delay[253]. In Table 7 are reported a summary of genes and related phenotypes for which specific indications or contraindications for treatment are available**.**

The increasing use of genetic sequencing could lead to a dramatic improvement of the effectiveness of epilepsy treatments, by targeting the biological mechanisms responsible for epilepsy in each specific individual. Precision medicine treatments represent a growing area of interest, focusing on reversing or circumventing the pathophysiological effects of specific gene mutations.

Gene name; inheritance; OMIM number	**Phenotype(s)**	**Beneficial**	**Anecdotal evidence**	**Proof of concept**	**Seizure worsening**
ALDH7A1; AR *107323	Pyridoxine-dependent epilepsy	Pyridoxine, folinic acid, lysine-restricted diet, L-arginine supplementation [245,254]			
BRAT1; AR; *136430	EE		ZNS[255]		
DEPDC5; AD; *614506	FFEVF, ADNFLE, FMTLE, ADEAF, IS			mTOR inhibitors (rapamycine)[256–258]	
FOLR1; AR; *136430	Folinic acid-responsive seizures	Leucovorin[259,260]			
FOXG1; (genomic duplications); AD (*de novo*); *164874	WS		Steroids[261–263]		
GABRG2; AD (de novo); *137164	DS, DS-like, GEFS+, GGE				LTG[264]
GRIN2A; AD (*de novo*); *138253	EAS, WS			NMDA inhibitors such as memantine[265,266]	
GRIN2B; AD (*de novo*); *138252	WS, LGS			NMDA inhibitors such as memantine[109,267]	
GRIN2D; AD; *602717	EE			NMDA inhibitors such as memantine[268]	
KCNQ2; AD (*de novo* in EE-inherited in	BFNS, EE	PHT, CBZ[130,233,269]	Retigabine (ezogabine-US		

"benign" epilepsies); *602235			adopted name)[131,270]		
KCNT1; AD (*de novo*); *608167	EIMFS, ADNFLE, EIEE, WS			Quinidine[270,271]	
MEF2C; AD (*de novo*); *600662	WS	VPA[272,273]			
MTOR; AD (*de novo*); *601231				mTOR inhibitors (rapamycin)[274,275]	
PCDH19; XL (limited to female and mosaic male); *300460	EFMR, DS-like		Clobazam, bromide[276]	Neurosteroids[249]	
PNPO; AR; *603287	Pyridoxal 5'-phosphate-responsive epilepsy	Pyridoxal 5'-phosphate[254,277,278]			
SCN1A; AD; *182389	DS, GEFS+, other epilepsies	Stiripentol, VPA, BDZ [279,280]	Fenfluramine, TPM[228,281,282]		CBZ, OXC, LTG [228,279,283]
SCN2A; AD (*de novo* in EE—inherited in 'benign' epilepsies); *182390	EIEE, EIMFS, EOEE, WS, BFNIS		PHT, other sodium channel blockers (in GoF mutations) [284]		
SCN8A; AD (*de novo*); *600702	EIMFS, WS		PHT, other sodium channel blockers (in GoF mutations) [285–287]		
SLC2A1; AD; *138140	GLUT1 deficiency syndrome, EOAE, EMA	KD[138,243,288–290]			
STRADA; AR; *608626	PMSE			mTOR inhibitors (rapamycin)[252,253]	

STXBP1; AD (*de novo*); *602926	EIEE, EOEE, WS, DS			PPI inhibitor[291]	

Table 7. *Genes and related phenotypes for which specific indications or contraindications for treatment are available. Abbreviations: AD autosomal dominant, ADEAF autosomal dominant epilepsy with auditory features, ADNFLE autosomal dominant nocturnal frontal lobe epilepsy, AR autosomal recessive, BDZ benzo- diazepine, BFNIS benign familial neonatal infantile seizures, BFNS benign familial neonatal seizures, CBZ carbamazepine, DS Dravet syndrome, DS-like Dravet-like syndrome, EAS epilepsy-aphasia syndromes, EE epileptic encephalopathy, EFMR epilepsy limited to females with mental retardation, EIEE early infantile epileptic encephalopathy, EIMFS epilepsy of infancy with migrating focal seizures, EMA epilepsy with myoclonic absence, EOAE early-onset absence epilepsy, EOEE early-onset epileptic encephalopathy, FFEVF familial focal epilepsy with variable foci, FMTLE familial mesial temporal lobe epilepsy, GEFS? genetic epilepsy with febrile seizures plus, GGE genetic generalized epilepsy, GoF gain of function, IS infantile spasms, KD ketogenic diet, LGS Lennox-Gastaut syndrome, LTG lamotrigine, mTOR mechanistic target of rapamycin, NMDA N-methyl-D-aspartate receptor, OMIM online Mendelian inheritance in man, OXC oxcarbazepine, PHT phenytoin, PMSE polyhydramnios, megalencephaly, symptomatic epilepsy syndrome-Pretzel syndrome, PPI protein-protein interaction, TPM topiramate, VPA valproate, WS West syndrome, ZNS zonisamide*[158].

6. Conclusions

The introduction of NGS has revolutionized genetic research as it has made cheap and massive sequencing available. Since its introduction, WES has been used to detect hundreds of new disease genes. In epilepsy, WES had the largest impact on gene discovery of sporadic severe epileptic encephalopathy, for which *de novo* mutations in many new genes are still being discovered[168].

Our results support diagnostic WES as a first diagnostic choice if there is no clear differential diagnosis. Overall the superiority of clinical WES over standard genetic tests is illustrated by the broad simultaneous coverage of thousands of genes, by a low-cost and fast turnaround approach, and also by the unique potential for dual molecular diagnosis and efficient identification of variants across diverse phenotypes and populations. Ion channel genes represent the gene family most frequently causally related to epilepsy, but other genes have gradually been associated with complex developmental epilepsy conditions, revealing the pathogenic role of mutations affecting diverse molecular pathways that regulate membrane excitability, synaptic plasticity, presynaptic neurotransmitter release, postsynaptic receptors, transporters, cell metabolism, and many formative steps in early brain development. Some of these discoveries are being followed by proof-of-concept laboratory studies that might open new pathways towards personalized treatment choices, not only for children but also for adult patients that may have a beneficial effect identifying the specific defects with a precision medicine treatment and a specific counseling for procreation. This increasing knowledge has led to the design of new drugs targeted to specific pathogenic mechanisms, or to a specific action of mutated proteins, up to a gene replacement therapy. However, further studies research will require an interdisciplinary and international collaboration approach, combining basic research with clinical studies.

7. Bibliography

1 Wirrell EC, Grossardt BR, Wong-Kisiel LCL, Nickels KC. Incidence and classification of new-onset epilepsy and epilepsy syndromes in children in Olmsted County, Minnesota from 1980 to 2004: a population-based study. *Epilepsy Res* 2011; **95**: 110–118.

2 Eltze CM, Chong WK, Cox T *et al.* A population-based study of newly diagnosed epilepsy in infants. *Epilepsia* 2013; **54**: 437–445.

3 Berg AT, Berkovic SF, Brodie MJ *et al.* Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia* 2010; **51**: 676–685.

4 Steinlein OK. Genetic mechanisms that underlie epilepsy. *Nat Rev Neurosci* 2004; **5**: 400–408.

5 Cowan LD. The epidemiology of the epilepsies in children. Ment. Retard. Dev. Disabil. Res. Rev. 2002; **8**: 171–181.

6 Vijayakumar NT, Judy MV. Autism spectrum disorders: Integration of the genome, transcriptome and the environment. *J Neurol Sci* 2016; **364**: 167–176.

7 Guerrini R, Noebels J. How Can Advances in Epilepsy Genetics Lead to Better Treatments and Cures? In: Scharfman HE, Buckmaster PS (eds). *Issues in Clinical Epileptology: A View from the Bench*. Springer Netherlands: Dordrecht, 2014, pp 309–317.

8 Heron SE, Scheffer IE, Berkovic SF, Dibbens LM, Mulley JC. Channelopathies in idiopathic epilepsy. *Neurotherapeutics* 2007; **4**: 295–304.

9 Gourfinkel-An I, Baulac S, Nabbout R *et al.* Monogenic idiopathic epilepsies. *Lancet Neurol* 2004; **3**: 209–218.

10 Martin MS, Tang B, Papale LA, Yu FH, Catterall WA, Escayg A. The voltage-gated sodium channel Scn8a is a genetic modifier of severe myoclonic epilepsy of infancy. *Hum Mol Genet* 2007; **16**: 2892–2899.

11 Fujiwara T. Clinical spectrum of mutations in SCN1A gene: Severe myoclonic epilepsy in infancy and related epilepsies. Epilepsy Res. 2006; **70**: 223–230.

12 Fukuda M, Morimoto T, Nagao H, Kida K. Clinical study of epilepsy with severe febrile seizures and seizures induced by hot water bath. *Brain Dev* 1997; **19**: 212–216.

13 Catarino CB, Liu JYW, Liagkouras I *et al.* Dravet syndrome as epileptic encephalopathy: evidence from long-term course and neuropathology. *Brain* 2011; **134**: 2982–3010.

14 Claes L, Del-Favero J, Ceulemans B, Lagae L, Van Broeckhoven C, De Jonghe P. De Novo Mutations in the Sodium-Channel Gene SCN1A Cause Severe Myoclonic Epilepsy of Infancy. *Am J Hum Genet* 2001; **68**: 1327–1332.

15 Mulley JC, Scheffer IE, Petrou S, Dibbens LM, Berkovic SF, Harkin LA. SCN1A mutations and epilepsy. *Hum Mutat* 2005; **25**: 535–542.

16 Marini C, Scheffer IE, Nabbout R *et al.* SCN1A duplications and deletions detected in Dravet syndrome: implications for molecular diagnosis. *Epilepsia* 2009; **50**: 1670–1678.

17 Macdonald RL, Kang J-Q. Molecular pathology of genetic epilepsies associated with GABAA receptor subunit mutations. *Epilepsy Curr* 2009; **9**: 18–23.

18 Patino GA, Claes LRF, Lopez-Santiago LF *et al.* A functional null mutation of SCN1B in a patient with Dravet syndrome. *J Neurosci* 2009; **29**: 10764–10778.

19 Depienne C, Arzimanoglou A, Trouillard O *et al.* Parental mosaicism can cause recurrent transmission of SCN1A mutations associated with severe myoclonic epilepsy of infancy. *Hum Mutat* 2006; **27**: 389.

20 Marini C, Mei D, Helen Cross J, Guerrini R. Mosaic SCN1A mutation in familial severe myoclonic epilepsy of infancy. *Epilepsia* 2006; **47**: 1737–1740.

21 Guerrini R, Oguni H. Borderline Dravet syndrome: a useful diagnostic category? *Epilepsia* 2011; **52 Suppl 2**: 10–12.

22 Wallace RH, Scheffer IE, Barnett S *et al.* Neuronal sodium-channel alpha1-subunit mutations in generalized epilepsy with febrile seizures plus. *Am J Hum Genet* 2001; **68**: 859–865.

23 Harkin LA, McMahon JM, Iona X *et al.* The spectrum of SCN1A-related infantile epileptic encephalopathies. *Brain* 2007; **130**: 843–852.

24 Lerche H, Weber YG, Jurkat-Rott K, Lehmann-Horn F. Ion channel defects in idiopathic epilepsies. *Curr Pharm Des* 2005; **11**: 2737–2752.

25 Yu FH, Mantegazza M, Westenbroek RE *et al.* Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. *Nat Neurosci* 2006; **9**: 1142–1149.

26 Ogiwara I, Miyamoto H, Morita N *et al.* Nav1.1 localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an Scn1a gene mutation. *J Neurosci* 2007; **27**: 5903–5914.

27 Higurashi N, Uchida T, Lossin C *et al.* A human Dravet syndrome model from patient induced pluripotent stem cells. *Mol Brain* 2013; **6**: 19.

28 Jiao J, Yang Y, Shi Y *et al.* Modeling Dravet syndrome using induced pluripotent stem cells (iPSCs) and directly converted neurons. *Hum Mol Genet* 2013; **22**: 4241–4252.

29 Liu Y, Lopez-Santiago LF, Yuan Y *et al.* Dravet syndrome patient-derived neurons suggest a novel epilepsy mechanism. *Ann Neurol* 2013; **74**: 128–139.

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

31 Escayg A, MacDonald BT, Meisler MH *et al.* Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. *Nat Genet* 2000; **24**: 343–345.

32 Marini C, Mei D, Temudo T *et al.* Idiopathic epilepsies with seizures precipitated by fever and SCN1A abnormalities. *Epilepsia* 2007; **48**: 1678–1685.

33 Baulac S, Huberfeld G, Gourfinkel-An I *et al.* First genetic evidence of GABA(A) receptor dysfunction in epilepsy: a mutation in the gamma2-subunit gene. *Nat Genet* 2001; **28**: 46–48.

34 Cestele S, Schiavon E, Rusconi R, Franceschetti S, Mantegazza M. Nonfunctional NaV1.1 familial hemiplegic migraine mutant transformed into gain of function by partial rescue of folding defects. *Proc Natl Acad Sci U S A* 2013; **110**: 17546–17551.

35 Rusconi R, Scalmani P, Cassulini RR *et al.* Modulatory proteins can rescue a trafficking defective epileptogenic Nav1.1 Na+ channel mutant. *J Neurosci* 2007; **27**: 11037–11046.

36 Sugiura Y, Ogiwara I, Hoshi A, Yamakawa K, Ugawa Y. Different degrees of loss of function between GEFS+ and SMEI Nav 1.1 missense mutants at the same residue induced by rescuable folding defects. *Epilepsia* 2012; **53**: e111-4.

37 Tang B, Dutt K, Papale L *et al.* A BAC transgenic mouse model reveals neuron subtype-specific effects of a Generalized Epilepsy with Febrile Seizures Plus (GEFS+) mutation. *Neurobiol Dis* 2009; **35**: 91–102.

38 Tian M, Mei D, Freri E *et al.* Impaired surface alphabetagamma GABA(A) receptor expression in familial epilepsy due to a GABRG2 frameshift mutation. *Neurobiol Dis* 2013; **50**: 135–141.

39 Vigevano F. Benign familial infantile seizures. *Brain Dev* 2005; **27**: 172–177.

40 Steinlein OK, Conrad C, Weidner B. Benign familial neonatal convulsions: always benign? *Epilepsy Res* 2007; **73**: 245–249.

41 Heron SE, Cox K, Grinton BE *et al.* Deletions or duplications in KCNQ2 can cause benign familial neonatal seizures. J. Med. Genet. 2007; **44**: 791–796.

42 Delmas P, Brown DA. Pathways modulating neural KCNQ/M (Kv7) potassium channels. *Nat Rev Neurosci* 2005; **6**: 850–862.

43 Schroeder BC, Kubisch C, Stein V, Jentsch TJ. Moderate loss of function of cyclic-AMP-modulated KCNQ2/KCNQ3 K+ channels causes epilepsy. *Nature* 1998; **396**: 687–690.

44 Weber YG, Lerche H. [Genetics of idiopathic epilepsies]. *Nervenarzt* 2013; **84**: 151–156.

45 Heron SE, Crossland KM, Andermann E *et al.* Sodium-channel defects in benign familial neonatal-infantile seizures. *Lancet (London, England)* 2002; **360**: 851–852.

46 Berkovic SF, Heron SE, Giordano L *et al.* Benign familial neonatal-infantile seizures: characterization of a new sodium channelopathy. *Ann Neurol* 2004; **55**: 550–557.

47 Herlenius E, Heron SE, Grinton BE *et al.* SCN2A mutations and benign familial neonatal-infantile seizures: the phenotypic spectrum. *Epilepsia* 2007; **48**: 1138–1142.

48 Szepetowski P, Rochette J, Berquin P, Piussan C, Lathrop GM, Monaco AP. Familial infantile convulsions and paroxysmal choreoathetosis: a new neurological syndrome linked to the pericentromeric region of human chromosome 16. *Am J Hum Genet* 1997; **61**: 889–898.

49 Marini C, Conti V, Mei D *et al.* PRRT2 mutations in familial infantile seizures, paroxysmal dyskinesia, and hemiplegic migraine. *Neurology* 2012; **79**: 2109–2114.

50 Guerrini R, Mink JW. Paroxysmal disorders associated with PRRT2 mutations shake up expectations on ion channel genes. Neurology. 2012; **79**: 2086–2088.

51 Guipponi M, Rivier F, Vigevano F *et al.* Linkage mapping of benign familial infantile convulsions (BFIC) to chromosome 19q. *Hum Mol Genet* 1997; **6**: 473–477.

52 Vanmolkot KRJ, Kors EE, Hottenga J-J *et al.* Novel mutations in the Na+, K+-ATPase pump gene ATP1A2 associated with familial hemiplegic migraine and benign familial infantile convulsions. *Ann Neurol* 2003; **54**: 360–366.

53 Zhou X, Ma A, Liu X *et al.* Infantile seizures and other epileptic phenotypes in a Chinese family with a missense mutation of KCNQ2. *Eur J Pediatr* 2006; **165**: 691–695.

54 Striano P, Bordo L, Lispi ML *et al.* A novel SCN2A mutation in family with benign familial infantile seizures. Epilepsia. 2006; **47**: 218–220.

55 Scheffer IE, Bhatia KP, Lopes-Cendes I *et al.* Autosomal dominant nocturnal frontal lobe epilepsy. A distinctive clinical disorder. *Brain* 1995; **118 (Pt 1**: 61–73.

56 Picard F, Baulac S, Kahane P *et al.* Dominant partial epilepsies. A clinical, electrophysiological and genetic study of 19 European families. *Brain* 2000; **123 (Pt 6**: 1247–1262.

57 Ryvlin P, Minotti L, Demarquay G *et al.* Nocturnal hypermotor seizures, suggesting frontal lobe epilepsy, can originate in the insula. *Epilepsia* 2006; **47**: 755–765.

58 Steinlein OK, Mulley JC, Propping P *et al.* A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* 1995; **11**: 201–203.

59 Aridon P, Marini C, Di Resta C *et al.* Increased sensitivity of the neuronal nicotinic receptor alpha 2 subunit causes familial epilepsy with nocturnal wandering and ictal fear. *Am J Hum Genet* 2006; **79**: 342–350.

60 Chen Y, Wu L, Fang Y *et al.* A novel mutation of the nicotinic acetylcholine receptor gene CHRNA4 in sporadic nocturnal frontal lobe epilepsy. *Epilepsy Res* 2009; **83**: 152–156.

61 Liu H, Lu C, Li Z *et al.* The identification of a novel mutation of nicotinic acetylcholine receptor gene CHRNB2 in a Chinese patient: Its possible implication in non-familial nocturnal frontal lobe epilepsy. *Epilepsy Res* 2011; **95**: 94–99.

62 Heron SE, Smith KR, Bahlo M *et al.* Missense mutations in the sodium-gated potassium channel gene KCNT1 cause severe autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* 2012; **44**: 1188–1190.

63 Ottman R, Risch N, Hauser WA *et al.* Localization of a gene for partial epilepsy to chromosome 10q. *Nat Genet* 1995; **10**: 56–60.

64 Michelucci R, Poza JJ, Sofia V *et al.* Autosomal dominant lateral temporal epilepsy: clinical spectrum, new epitempin mutations, and genetic heterogeneity in seven European families. *Epilepsia* 2003; **44**: 1289–1297.

65 Anderson MP. Arrested glutamatergic synapse development in human partial epilepsy. *Epilepsy Curr* 2010; **10**: 153–158.

66 Kalachikov S, Evgrafov O, Ross B *et al.* Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features. *Nat Genet* 2002; **30**: 335–341.

67 Ottman R, Winawer MR, Kalachikov S *et al.* LGI1 mutations in autosomal dominant partial epilepsy with auditory features. *Neurology* 2004; **62**: 1120–1126.

68 Chernova OB, Somerville RP, Cowell JK. A novel gene, LGI1, from 10q24 is rearranged and downregulated in malignant brain tumors. *Oncogene* 1998; **17**: 2873–2881.

69 Dazzo E, Fanciulli M, Serioli E *et al.* Heterozygous reelin mutations cause autosomal-dominant lateral temporal epilepsy. *Am J Hum Genet* 2015; **96**: 992–1000.

70 Zhou Y-D, Lee S, Jin Z, Wright M, Smith SEP, Anderson MP. Arrested maturation of excitatory synapses in autosomal dominant lateral temporal lobe epilepsy. *Nat Med* 2009; **15**: 1208–1214.

71 Yu YE, Wen L, Silva J *et al.* Lgi1 null mutant mice exhibit myoclonic seizures and CA1 neuronal hyperexcitability. *Hum Mol Genet* 2010; **19**: 1702–1711.

72 Dibbens LM, de Vries B, Donatello S *et al.* Mutations in DEPDC5 cause familial focal epilepsy with variable foci. *Nat Genet* 2013; **45**: 546–551.

73 Ishida S, Picard F, Rudolf G *et al.* Mutations of DEPDC5 cause autosomal dominant focal epilepsies. *Nat Genet* 2013; **45**: 552–555.

74 Bar-Peled L, Chantranupong L, Cherniack AD *et al.* A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science* 2013; **340**: 1100–1106.

75 Sadleir LG, Scheffer IE. Febrile seizures. *BMJ* 2007; **334**: 307–311.

76 Hauser WA, Annegers JF, Anderson VE, Kurland LT. The risk of seizure disorders among relatives of children with febrile convulsions. *Neurology* 1985; **35**: 1268–1273.

77 Maher J, McLachlan RS. Febrile convulsions. Is seizure duration the most important predictor of temporal lobe epilepsy? *Brain* 1995; **118 (Pt 6**: 1521–1528.

78 Johnson WG, Kugler SL, Stenroos ES *et al.* Pedigree analysis in families with febrile seizures. *Am J Med Genet* 1996; **61**: 345–352.

79 Baulac S, Gourfinkel-An I, Nabbout R *et al.* Fever, genes, and epilepsy. *Lancet Neurol* 2004; **3**: 421–430.

80 Dibbens LM, Tarpey PS, Hynes K *et al.* X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nat Genet* 2008; **40**: 776–781.

81 Marini C, Mei D, Parmeggiani L *et al.* Protocadherin 19 mutations in girls with infantile-onset epilepsy. *Neurology* 2010; **75**: 646–653.

82 Deprez L, Jansen A, De Jonghe P. Genetics of epilepsy syndromes starting in the first year of life. *Neurology* 2009; **72**: 273–281.

83 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* 2009; **5**: e1000381.

84 Ohira R, Zhang YH, Guo W *et al.* Human ARX gene: genomic characterization and expression. *Mol Genet Metab* 2002; **77**: 179–188.

85 Gecz J, Cloosterman D, Partington M. ARX: a gene for all seasons. *Curr Opin Genet Dev* 2006; **16**: 308–316.

86 Guerrini R, Moro F, Kato M *et al.* Expansion of the first PolyA tract of ARX causes infantile spasms and status dystonicus. *Neurology* 2007; **69**: 427–433.

87 Marsh E, Fulp C, Gomez E *et al.* Targeted loss of Arx results in a developmental epilepsy mouse model and recapitulates the human phenotype in heterozygous females. *Brain* 2009; **132**: 1563–1576.

88 Kato M, Das S, Petras K *et al.* Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. *Hum Mutat* 2004; **23**: 147–159.

89 Stromme P, Mangelsdorf ME, Scheffer IE, Gecz J. Infantile spasms, dystonia, and other X-linked phenotypes caused by mutations in Aristaless related homeobox gene, ARX. *Brain Dev* 2002; **24**: 266–268.

90 Wallerstein R, Sugalski R, Cohn L, Jawetz R, Friez M. Expansion of the ARX spectrum. *Clin Neurol Neurosurg* 2008; **110**: 631–634.

91 Shinozaki Y, Osawa M, Sakuma H *et al.* Expansion of the first polyalanine tract of the ARX gene in a boy presenting with generalized dystonia in the absence of infantile spasms. *Brain Dev* 2009; **31**: 469–472.

92 Kato M, Saitoh S, Kamei A *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* 2007; **81**: 361–366.

93 Mari F, Azimonti S, Bertani I *et al.* CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Hum Mol Genet* 2005; **14**: 1935–1946.

94 Kalscheuer VM, Tao J, Donnelly A *et al.* Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am J Hum Genet* 2003; **72**: 1401–1411.

95 Buoni S, Zannolli R, Colamaria V *et al.* Myoclonic encephalopathy in the CDKL5 gene mutation. *Clin Neurophysiol* 2006; **117**: 223–227.

96 Archer HL, Evans J, Edwards S *et al.* CDKL5 mutations cause infantile spasms, early onset seizures, and severe mental retardation in female patients. *J Med Genet* 2006; **43**: 729–734.

97 Elia M, Falco M, Ferri R *et al.* CDKL5 mutations in boys with severe encephalopathy and early-onset intractable epilepsy. *Neurology* 2008; **71**: 997–999.

98 Weaving LS, Christodoulou J, Williamson SL *et al.* Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am J Hum Genet* 2004; **75**: 1079–1093.

99 Melani F, Mei D, Pisano T *et al.* CDKL5 gene-related epileptic encephalopathy: electroclinical findings in the first year of life. *Dev Med Child Neurol* 2011; **53**: 354–360.

100 Mei D, Marini C, Novara F *et al.* Xp22.3 genomic deletions involving the CDKL5 gene in girls with early onset epileptic encephalopathy. *Epilepsia* 2010; **51**: 647–654.

101 Saitsu H, Kato M, Mizuguchi T *et al.* De novo mutations in the gene encoding STXBP1 (MUNC18-1) cause early infantile epileptic encephalopathy. *Nat Genet* 2008; **40**: 782–788.

102 Djukic A, Lado FA, Shinnar S, Moshe SL. Are early myoclonic encephalopathy (EME) and the Ohtahara syndrome (EIEE) independent of each other? *Epilepsy Res* 2006; **70 Suppl 1**: S68-76.

103 Ohtahara S, Yamatogi Y. Ohtahara syndrome: with special reference to its developmental aspects for differentiating from early myoclonic encephalopathy. *Epilepsy Res* 2006; **70 Suppl 1**: S58-67.

104 Mignot C, Moutard M-L, Trouillard O *et al.* STXBP1-related encephalopathy presenting as infantile spasms and generalized tremor in three patients. *Epilepsia* 2011; **52**: 1820–1827.

105 Milh M, Villeneuve N, Chouchane M *et al.* Epileptic and nonepileptic features in patients with early onset epileptic encephalopathy and STXBP1 mutations. *Epilepsia* 2011; **52**: 1828–1834.

106 Lemke JR, Riesch E, Scheurenbrand T *et al.* Targeted next generation sequencing as a diagnostic tool in epileptic disorders. *Epilepsia* 2012; **53**: 1387–1398.

107 Lesca G, Rudolf G, Bruneau N *et al.* GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. *Nat Genet* 2013; **45**: 1061–1066.

108 Carvill GL, Regan BM, Yendle SC *et al.* GRIN2A mutations cause epilepsy-aphasia spectrum disorders. *Nat Genet* 2013; **45**: 1073–1076.

109 Lemke JR, Hendrickx R, Geider K *et al.* GRIN2B mutations in West syndrome and intellectual disability with focal epilepsy. *Ann Neurol* 2014; **75**: 147–154.

110 Berkovic SF, Andermann F, Carpenter S, Wolfe LS. Progressive myoclonus epilepsies: specific causes and diagnosis. *N Engl J Med* 1986; **315**: 296–305.

111 Delgado-Escueta A V, Ganesh S, Yamakawa K. Advances in the genetics of progressive myoclonus epilepsy. *Am J Med Genet* 2001; **106**: 129–138.

112 Annegers JF, Hauser WA, Anderson VE, Kurland LT. The risks of seizure disorders among relatives of patients with childhood onset epilepsy. *Neurology* 1982; **32**: 174–179.

113 Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 1990; **46**: 222–228.

114 Berkovic SF, Scheffer IE. Genetics of the epilepsies. *Epilepsia* 2001; **42 Suppl 5**: 16–23.

115 Klassen T, Davis C, Goldman A *et al.* Exome sequencing of ion channel genes reveals complex profiles confounding personal risk assessment in epilepsy. *Cell* 2011; **145**: 1036–1048.

116 Sander T, Schulz H, Saar K *et al.* Genome search for susceptibility loci of common idiopathic generalised epilepsies. *Hum Mol Genet* 2000; **9**: 1465–1472.

117 Wallace RH, Marini C, Petrou S *et al.* Mutant GABA(A) receptor gamma2-subunit in childhood absence epilepsy and febrile seizures. *Nat Genet* 2001; **28**: 49–52.

118 Cossette P, Liu L, Brisebois K *et al.* Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. *Nat Genet* 2002; **31**: 184–189.

119 D'Agostino D, Bertelli M, Gallo S *et al.* Mutations and polymorphisms of the CLCN2 gene in idiopathic epilepsy. *Neurology* 2004; **63**: 1500–1502.

120 Powell KL, Cain SM, Ng C *et al.* A Cav3.2 T-type calcium channel point mutation has splice-variant-specific effects on function and segregates with seizure expression in a polygenic rat model of absence epilepsy. *J Neurosci* 2009; **29**: 371–380.

121 Dibbens LM, Feng H-J, Richards MC *et al.* GABRD encoding a protein for extra- or peri-synaptic GABAA receptors is a susceptibility locus for generalized epilepsies. *Hum Mol Genet* 2004; **13**: 1315–1319.

122 Greenberg DA, Cayanis E, Strug L *et al.* Malic enzyme 2 may underlie susceptibility to adolescent-onset idiopathic generalized epilepsy. *Am J Hum Genet* 2005; **76**: 139–146.

123 Pal DK, Evgrafov O V, Tabares P, Zhang F, Durner M, Greenberg DA. BRD2 (RING3) is a probable major susceptibility gene for common juvenile myoclonic epilepsy. *Am J Hum Genet* 2003; **73**: 261–270.

124 Dibbens LM, Ekberg J, Taylor I *et al.* NEDD4-2 as a potential candidate susceptibility gene for epileptic photosensitivity. *Genes Brain Behav* 2007; **6**: 750–755.

125 McTague A, Howell KB, Cross JH, Kurian MA, Scheffer IE. The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol. 2016; **15**: 304–316.

126 Miceli F, Soldovieri MV, Ambrosino P *et al.* Genotype-phenotype correlations in neonatal epilepsies caused by mutations in the voltage sensor of K(v)7.2 potassium channel subunits. *Proc Natl Acad Sci U S A* 2013; **110**: 4386–4391.

127 Nakamura K, Kato M, Osaka H *et al.* Clinical spectrum of SCN2A mutations expanding to Ohtahara syndrome. *Neurology* 2013; **81**: 992–998.

128 Singh NA, Pappas C, Dahle EJ *et al.* A role of SCN9A in human epilepsies, as a cause of febrile seizures and as a potential modifier of Dravet syndrome. *PLoS Genet* 2009; **5**: e1000649.

129 Biervert C, Steinlein OK. Structural and mutational analysis of KCNQ2, the major gene locus for benign familial neonatal convulsions. *Hum Genet* 1999; **104**: 234–240.

130 Kato M, Yamagata T, Kubota M *et al.* Clinical spectrum of early onset epileptic encephalopathies caused by KCNQ2 mutation. *Epilepsia* 2013; **54**: 1282–1287.

131 Orhan G, Bock M, Schepers D *et al.* Dominant-negative effects of KCNQ2 mutations are associated with epileptic encephalopathy. *Ann Neurol* 2014; **75**: 382–394.

132 Ohba C, Kato M, Takahashi N *et al.* De novo KCNT1 mutations in early-onset epileptic encephalopathy. *Epilepsia* 2015; **56**: e121-8.

133 Moller RS, Heron SE, Larsen LHG *et al.* Mutations in KCNT1 cause a spectrum of focal epilepsies. *Epilepsia* 2015; **56**: e114-20.

134 Martin HC, Kim GE, Pagnamenta AT *et al.* Clinical whole-genome sequencing in severe early-onset epilepsy reveals new genes and improves molecular diagnosis. *Hum Mol Genet* 2014; **23**: 3200–3211.

135 Milligan CJ, Li M, Gazina E V *et al.* KCNT1 gain of function in 2 epilepsy phenotypes is reversed by quinidine. *Ann Neurol* 2014; **75**: 581–590.

136 Carvill GL, Weckhuysen S, McMahon JM *et al.* GABRA1 and STXBP1: novel genetic causes of Dravet syndrome. *Neurology* 2014; **82**: 1245–1253.

137 Doose H. Myoclonic-astatic epilepsy. *Epilepsy Res Suppl* 1992; **6**: 163–168.

138 Mullen SA, Marini C, Suls A *et al.* Glucose transporter 1 deficiency as a treatable cause of myoclonic astatic epilepsy. *Arch Neurol* 2011; **68**: 1152–1155.

139 Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet* 2010; **11**: 31–46.

140 Finishing the euchromatic sequence of the human genome. *Nature* 2004; **431**: 931–945.

141 Watson M. Illuminating the future of DNA sequencing. *Genome Biol* 2014; **15**: 108.

142 Heinzen EL, Neale BM, Traynelis SF, Allen AS, Goldstein DB. The genetics of neuropsychiatric diseases: looking in and beyond the exome. *Annu Rev Neurosci* 2014; **38**: 150403170110009.

143 Sisodiya SM. Genetic screening and diagnosis in epilepsy? *Curr Opin Neurol* 2015; **28**.http://journals.lww.com/co-neurology/Fulltext/2015/04000/Genetic_screening_and_diagnosis_in_epilepsy_.9.aspx.

144 Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic Acids Res* 2015; **43**: D789-98.

145 Warman Chardon J, Beaulieu C, Hartley T, Boycott KM, Dyment DA. Axons to Exons: the Molecular Diagnosis of Rare Neurological Diseases by Next-Generation Sequencing. *Curr Neurol Neurosci Rep* 2015; **15**: 64.

146 Moller RS, Dahl HA, Helbig I. The contribution of next generation sequencing to epilepsy genetics. *Expert Rev Mol Diagn* 2015; **15**: 1531–1538.

147 Veeramah KR, O'Brien JE, Meisler MH *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* 2012; **90**: 502–510.

148 Touma M, Joshi M, Connolly MC *et al.* Whole genome sequencing identifies SCN2A mutation in monozygotic twins with Ohtahara syndrome and unique neuropathologic findings. *Epilepsia* 2013; **54**: e81-5.

149 Nemirovsky SI, Cordoba M, Zaiat JJ *et al.* Whole genome sequencing reveals a de novo SHANK3 mutation in familial autism spectrum disorder. *PLoS One* 2015; **10**: e0116358.

150 Muthugovindan D, Hartman AL. Pediatric epilepsy syndromes. *Neurologist* 2010; **16**: 223–237.

151 Helbig KL, Farwell Hagman KD, Shinde DN *et al.* Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. *Genet Med* 2016; **18**: 898–905.

152 Dyment DA, Tetreault M, Beaulieu CL *et al.* Whole-exome sequencing broadens the phenotypic spectrum of rare pediatric epilepsy: a retrospective study. *Clin Genet* 2015; **88**: 34–40.

153 Michaud JL, Lachance M, Hamdan FF *et al.* The genetic landscape of infantile spasms. *Hum Mol Genet* 2014; **23**: 4846–4858.

154 Allen NM, Conroy J, Shahwan A *et al.* Unexplained early onset epileptic encephalopathy: Exome screening and phenotype expansion. *Epilepsia* 2016; **57**: e12-7.

155 Muona M, Berkovic SF, Dibbens LM *et al.* A recurrent de novo mutation in KCNC1 causes progressive myoclonus epilepsy. *Nat Genet* 2015; **47**: 39–46.

156 Consortium E, Project EP. De novo mutations in epileptic encephalopathies. *Nature* 2013; **501**: 217–221.

157 Appenzeller S, Balling R, Barisic N *et al.* Erratum: De Novo Mutations in Synaptic Transmission Genes Including DNM1 Cause Epileptic Encephalopathies (American Journal of Human Genetics (2014) 95(4) (360–370)(S0002929714003838)(10.1016/j.ajhg.2014.08.013)). *Am J Hum Genet* 2017; **100**: 179.

158 Mei D, Parrini E, Marini C, Guerrini R. The Impact of Next-Generation Sequencing on the Diagnosis and Treatment of Epilepsy in Paediatric Patients. *Mol Diagn Ther* 2017. doi:10.1007/s40291-017-0257-0.

159 Boyle EA, O'Roak BJ, Martin BK, Kumar A, Shendure J. MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing. *Bioinformatics* 2014; **30**: 2670–2672.

160 Kodera H, Kato M, Nord AS *et al.* Targeted capture and sequencing for detection of mutations causing early onset epileptic encephalopathy. *Epilepsia* 2013; **54**: 1262–1269.

161 Wang J, Gotway G, Pascual JM, Park JY. Diagnostic yield of clinical next-generation sequencing panels for epilepsy. *JAMA Neurol* 2014; **71**: 650–651.

162 Della Mina E, Ciccone R, Brustia F *et al.* Improving molecular diagnosis in epilepsy by a dedicated high-throughput sequencing platform. *Eur J Hum Genet* 2015; **23**: 354–362.

163 Parrini E, Marini C, Mei D *et al.* Diagnostic Targeted Resequencing in 349 Patients with Drug-Resistant Pediatric Epilepsies Identifies Causative Mutations in 30 Different Genes. *Hum Mutat* 2017; **38**: 216–225.

164 Trump N, McTague A, Brittain H *et al.* Improving diagnosis and broadening the phenotypes in early-onset seizure and severe

developmental delay disorders through gene panel analysis. *J Med Genet* 2016; **53**: 310–317.

165 Scheffer IE. Epilepsy genetics revolutionizes clinical practice. *Neuropediatrics* 2014; **45**: 70–74.

166 Nieh SE, Sherr EH. Epileptic encephalopathies: new genes and new pathways. *Neurotherapeutics* 2014; **11**: 796–806.

167 Holmes GL, Noebels JL. The Epilepsy Spectrum: Targeting Future Research Challenges. *Cold Spring Harb Perspect Med* 2016; **6**. doi:10.1101/cshperspect.a028043.

168 Myers CT, Mefford HC. Advancing epilepsy genetics in the genomic era. *Genome Med* 2015; **7**: 91.

169 Helbig I, Tayoun AAN. Understanding Genotypes and Phenotypes in Epileptic Encephalopathies. *Mol Syndromol* 2016; **7**: 172–181.

170 Moorthie S, Hall A, Wright CF. Informatics and clinical genome sequencing: opening the black box. *Genet Med* 2013; **15**: 165–171.

171 Oliver GR, Hart SN, Klee EW. Bioinformatics for Clinical Next Generation Sequencing. *Clin Chem* 2014; **61**: 124.

172 DePristo MA, Banks E, Poplin R *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011; **43**: 491–498.

173 Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet* 2011; **12**: 443–451.

174 Voelkerding K V, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem* 2009; **55**: 641–658.

175 Pabinger S, Dander A, Fischer M *et al.* A survey of tools for variant analysis of next-generation genome sequencing data. *Brief Bioinform* 2014; **15**: 256–278.

176 Cingolani P, Platts A, Wang LL *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 2012; **6**: 80–92.

177 Köhler S, Vasilevsky NA, Engelstad M *et al.* The human phenotype ontology in 2017. *Nucleic Acids Res* 2017; **45**: D865–D876.

178 Deans AR, Lewis SE, Huala E *et al.* Finding our way through phenotypes. *PLoS Biol* 2015; **13**: e1002033.

179 Ales M, Luca L, Marija V *et al.* Phenotype-driven gene target definition in clinical genome-wide sequencing data interpretation. *Genet Med* 2016; **18**: 1102–1110.

180 Zemojtel T, Kohler S, Mackenroth L *et al.* Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Sci Transl Med* 2014; **6**: 252ra123.

181 Pinto D, Delaby E, Merico D *et al.* Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet* 2014; **94**: 677–694.

182 Covone AE, Fiorillo C, Acquaviva M *et al.* WES in a family trio suggests involvement of TECPR2 in a complex form of progressive motor neuron disease. *Clin Genet* 2016; **90**: 182–185.

183 Yuen RKC, Thiruvahindrapuram B, Merico D *et al.* Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med* 2015; **21**: 185–191.

184 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; **26**: 589–595.

185 McKenna A, Hanna M, Banks E *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**: 1297–1303.

186 Van der Auwera GA, Carneiro MO, Hartl C *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma* 2013; **43**: 11.10.1-33.

187 Ramu A, Noordam MJ, Schwartz RS *et al.* DeNovoGear: de novo indel and point mutation discovery and phasing. *Nat Methods* 2013; **10**: 985–987.

188 Karczewski KJ, Weisburd B, Thomas B *et al.* The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res* 2017; **45**: D840–D845.

189 Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 2013; **9**: e1003709.

190 Itan Y, Shang L, Boisson B *et al.* The human gene damage index as a gene-level approach to prioritizing exome variants. *Proc Natl Acad Sci U S A* 2015; **112**: 13615–13620.

191 Itan Y, Shang L, Boisson B *et al.* The mutation significance cutoff: gene-level thresholds for variant predictions. Nat. Methods. 2016; **13**: 109–110.

192 Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat* 2016; **37**: 235–241.

193 Ioannidis NM, Rothstein JH, Pejaver V *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* 2016; **99**: 877–885.

194 Jagadeesh KA, Wenger AM, Berger MJ *et al.* M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet* 2016; **48**: 1581–1586.

195 Ionita-Laza I, McCallum K, Xu B, Buxbaum JD. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nat Genet* 2016; **48**: 214–220.

196 Smedley D, Jacobsen JOB, Jäger M *et al.* Next-generation diagnostics and disease-gene discovery with the Exomiser. *Nat Protoc* 2015; **10**: 2004–2015.

197 Danecek P, Auton A, Abecasis G *et al.* The variant call format and VCFtools. *Bioinformatics* 2011; **27**: 2156–2158.

198 Richards S, Aziz N, Bale S *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; **17**: 405–424.

199 Johannesen K, Marini C, Pfeffer S *et al.* Phenotypic spectrum of GABRA1: From generalized epilepsies to severe epileptic encephalopathies. *Neurology* 2016; **87**: 1140–1151.

200 Nakamura K, Kodera H, Akita T *et al.* De Novo mutations in GNAO1, encoding a Galphao subunit of heterotrimeric G proteins, cause epileptic encephalopathy. *Am J Hum Genet* 2013; **93**: 496–505.

201 Tanaka AJ, Bai R, Cho MT *et al.* De novo mutations in PURA are associated with hypotonia and developmental delay. *Cold Spring Harb Mol case Stud* 2015; **1**: a000356.

202 Chong JX, McMillin MJ, Shively KM *et al.* De novo mutations in NALCN cause a syndrome characterized by congenital contractures of the limbs and face, hypotonia, and developmental delay. *Am J Hum Genet* 2015; **96**: 462–473.

203 Ishii A, Watkins JC, Chen D, Hirose S, Hammer MF. Clinical implications of SCN1A missense and truncation variants in a large Japanese cohort with Dravet syndrome. *Epilepsia* 2017; **58**: 282–290.

204 Valenti R, Bianchi S, Pescini F *et al.* First report of a pathogenic mutation on exon 24 of the NOTCH3 gene in a CADASIL family. *J Neurol* 2011; **258**: 1632–1636.

205 Hunt D, Leventer RJ, Simons C *et al.* Whole exome sequencing in family trios reveals de novo mutations in PURA as a cause of severe neurodevelopmental delay and learning disability. *J Med Genet* 2014; **51**: 806–813.

206 Chai J, Li Y, Wang H. DNA Damage and Repair , Neurodegeneration and the Role of Purα in DNA Repair. 2015; **7**: 1–9.

207 Fragaki K, Ait-El-Mkadem S, Chaussenot A *et al.* Refractory epilepsy and mitochondrial dysfunction due to GM3 synthase deficiency. *Eur J Hum Genet* 2013; **21**: 528–534.

208 Milh M, Falace A, Villeneuve N *et al.* Novel compound heterozygous mutations in TBC1D24 cause familial malignant migrating partial seizures of infancy. *Hum Mutat* 2013; **34**: 869–872.

209 Van Damme T, Gardeitchik T, Mohamed M *et al.* Mutations in ATP6V1E1 or ATP6V1A Cause Autosomal-Recessive Cutis Laxa. *Am J Hum Genet* 2017; **100**: 216–227.

210 Weiergraber M, Kamp MA, Radhakrishnan K, Hescheler J, Schneider T. The Ca(v)2.3 voltage-gated calcium channel in epileptogenesis--shedding new light on an enigmatic channel. *Neurosci Biobehav Rev* 2006; **30**: 1122–1144.

211 Sethi N, Yan Y, Quek D, Schupbach T, Kang Y. Rabconnectin-3 is a functional regulator of mammalian Notch signaling. *J Biol Chem* 2010; **285**: 34757–34764.

212 Sennoune SR, Luo D, Martinez-Zaguilan R. Plasmalemmal vacuolar-type H+-ATPase in cancer biology. *Cell Biochem Biophys* 2004; **40**: 185–206.

213 Smardon AM, Tarsio M, Kane PM. The RAVE complex is essential for stable assembly of the yeast V-ATPase. *J Biol Chem* 2002; **277**: 13831–13839.

214 Smardon AM, Kane PM. RAVE is essential for the efficient assembly of the C subunit with the vacuolar H(+)-ATPase. *J Biol Chem* 2007; **282**: 26185–26194.

215 Seol JH, Shevchenko A, Shevchenko A, Deshaies RJ. Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. *Nat Cell Biol* 2001; **3**: 384–391.

216 de Ligt J, Willemsen MH, van Bon BWM *et al.* Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012; **367**: 1921–1929.

217 Lee H, Deignan JL, Dorrani N *et al.* Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 2014; **312**: 1880–1887.

218 Yang Y, Muzny DM, Reid JG *et al.* Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013; **369**: 1502–1511.

219 Yavarna T, Al-Dewik N, Al-Mureikhi M *et al.* High diagnostic yield of clinical exome sequencing in Middle Eastern patients with Mendelian disorders. *Hum Genet* 2015; **134**: 967–980.

220 Beaulieu CL, Majewski J, Schwartzentruber J *et al.* FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. *Am J Hum Genet* 2014; **94**: 809–817.

221 Kearney JA. Epi4K Phase I: Gene Discovery in Epileptic Encephalopathies by Exome Sequencing. *Epilepsy Curr* 2014; **14**: 208–210.

222 Djémié T, Weckhuysen S, von Spiczak S *et al.* Pitfalls in genetic testing: the story of missed SCN1A mutations. *Mol Genet Genomic Med* 2016; **4**: 457–464.

223 Lalani SR, Zhang J, Schaaf CP *et al.* Mutations in PURA cause profound neonatal hypotonia, seizures, and encephalopathy in 5q31.3 microdeletion syndrome. *Am J Hum Genet* 2014; **95**: 579–583.

224 Sambri I, D'Alessio R, Ezhova Y *et al.* Lysosomal dysfunction disrupts presynaptic maintenance and restoration of presynaptic function prevents neurodegeneration in lysosomal storage diseases. *EMBO Mol Med* 2017; **9**: 112–132.

225 Padamsey Z, McGuinness L, Bardo SJ *et al.* Activity-Dependent Exocytosis of Lysosomes Regulates the Structural Plasticity of Dendritic Spines. *Neuron* 2017; **93**: 132–146.

226 Goo MS, Sancho L, Slepak N *et al.* Activity-dependent trafficking of lysosomes in dendrites and dendritic spines. *J Cell Biol* 2017; **216**: 2499–2513.

227 Orsini A, Zara F, Striano P. Recent advances in epilepsy genetics. *Neurosci Lett* 2017. doi:10.1016/j.neulet.2017.05.014.

228 Wirrell EC. Treatment of Dravet Syndrome. *Can J Neurol Sci* 2016; **43 Suppl 3**: S13-8.

229 Andrade DM. Dravet syndrome, lamotrigine, and personalized medicine. *Dev Med Child Neurol* 2015; **57**: 118–119.

230 Boerma RS, Braun KP, van den Broek MPH *et al.* Remarkable Phenytoin Sensitivity in 4 Children with SCN8A-related Epilepsy: A Molecular Neuropharmacological Approach. *Neurotherapeutics* 2016; **13**: 192–197.

231 Gardella E, Becker F, Moller RS *et al.* Benign infantile seizures and paroxysmal dyskinesia caused by an SCN8A mutation. *Ann Neurol* 2016; **79**: 428–436.

232 Anand G, Collett-White F, Orsini A *et al.* Autosomal dominant SCN8A mutation with an unusually mild phenotype. *Eur J Paediatr Neurol* 2016; **20**: 761–765.

233 Pisano T, Numis AL, Heavin SB *et al.* Early and effective treatment of KCNQ2 encephalopathy. *Epilepsia* 2015; **56**: 685–691.

234 Numis AL, Angriman M, Sullivan JE *et al.* KCNQ2 encephalopathy: delineation of the electroclinical phenotype and treatment response. *Neurology* 2014; **82**: 368–370.

235 Rosati A, De Masi S, Guerrini R. Antiepileptic Drug Treatment in Children with Epilepsy. *CNS Drugs* 2015; **29**: 847–863.

236 Pan Z, Kao T, Horvath Z *et al.* A common ankyrin-G-based mechanism retains KCNQ and NaV channels at electrically active domains of the axon. *J Neurosci* 2006; **26**: 2599–2613.

237 Nguyen HM, Miyazaki H, Hoshi N *et al.* Modulation of voltage-gated K+ channels by the sodium channel beta1 subunit. *Proc Natl Acad Sci U S A* 2012; **109**: 18577–18582.

238 Gunthorpe MJ, Large CH, Sankar R. The mechanism of action of retigabine (ezogabine), a first-in-class K+ channel opener for the treatment of epilepsy. *Epilepsia* 2012; **53**: 412–424.

239 Barrese V, Miceli F, Soldovieri MV *et al.* Neuronal potassium channel openers in the management of epilepsy: role and potential of retigabine. *Clin Pharmacol* 2010; **2**: 225–236.

240 Barcia G, Fleming MR, Deligniere A *et al.* De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. *Nat Genet* 2012; **44**: 1255–1259.

241 Rizzo F, Ambrosino P, Guacci A *et al.* Characterization of two de novoKCNT1 mutations in children with malignant migrating partial seizures in infancy. *Mol Cell Neurosci* 2016; **72**: 54–63.

242 Chong PF, Nakamura R, Saitsu H, Matsumoto N, Kira R. Ineffective quinidine therapy in early onset epileptic encephalopathy with KCNT1 mutation. Ann. Neurol. 2016; **79**: 502–503.

243 Ramm-Pettersen A, Nakken KO, Haavardsholm KC, Selmer KK. Occurrence of GLUT1 deficiency syndrome in patients treated with ketogenic diet. *Epilepsy Behav* 2014; **32**: 76–78.

244 Akman CI, Yu J, Alter A, Engelstad K, De Vivo DC. Diagnosing Glucose Transporter 1 Deficiency at Initial Presentation Facilitates Early Treatment. *J Pediatr* 2016; **171**: 220–226.

245 van Karnebeek CDM, Tiebout SA, Niermeijer J *et al.* Pyridoxine-Dependent Epilepsy: An Expanding Clinical Spectrum. *Pediatr Neurol* 2016; **59**: 6–12.

246 Coci EG, Codutti L, Fink C *et al.* Novel homozygous missense mutation in ALDH7A1 causes neonatal pyridoxine dependent epilepsy. *Mol*

Cell Probes 2017; **32**: 18–23.

247 Baumgart A, Spiczak S von, Verhoeven-Duif NM *et al.* Atypical vitamin B6 deficiency: a rare cause of unexplained neonatal and infantile epilepsies. *J Child Neurol* 2014; **29**: 704–707.

248 Jaeger B, Abeling NG, Salomons GS *et al.* Pyridoxine responsive epilepsy caused by a novel homozygous PNPO mutation. *Mol Genet Metab reports* 2016; **6**: 60–63.

249 Tan C, Shard C, Ranieri E *et al.* Mutations of protocadherin 19 in female epilepsy (PCDH19-FE) lead to allopregnanolone deficiency. *Hum Mol Genet* 2015; **24**: 5250–5259.

250 Pieribone VA, Tsai J, Soufflet C *et al.* Clinical evaluation of ganaxolone in pediatric and adolescent patients with refractory epilepsy. *Epilepsia* 2007; **48**: 1870–1874.

251 Kerrigan JF, Shields WD, Nelson TY *et al.* Ganaxolone for treating intractable infantile spasms: a multicenter, open-label, add-on trial. *Epilepsy Res* 2000; **42**: 133–139.

252 Sadowski K, Kotulska-Jozwiak K, Jozwiak S. Role of mTOR inhibitors in epilepsy treatment. *Pharmacol Rep* 2015; **67**: 636–646.

253 Jeong A, Wong M. mTOR Inhibitors in Children: Current Indications and Future Directions in Neurology. *Curr Neurol Neurosci Rep* 2016; **16**: 102.

254 Stockler S, Plecko B, Gospe SMJ *et al.* Pyridoxine dependent epilepsy and antiquitin deficiency: clinical and molecular characteristics and recommendations for diagnosis, treatment and follow-up. *Mol Genet Metab* 2011; **104**: 48–60.

255 Saitsu H, Yamashita S, Tanaka Y *et al.* Compound heterozygous BRAT1 mutations cause familial Ohtahara syndrome with hypertonia and microcephaly. *J Hum Genet* 2014; **59**: 687–690.

256 Marsan E, Ishida S, Schramm A *et al.* Depdc5 knockout rat: A novel model of mTORopathy. *Neurobiol Dis* 2016; **89**: 180–189.

257 Weckhuysen S, Marsan E, Lambrecq V *et al.* Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. *Epilepsia* 2016; **57**: 994–1003.

258 Lal D, Reinthaler EM, Schubert J *et al.* DEPDC5 mutations in genetic focal epilepsies of childhood. *Ann Neurol* 2014; **75**: 788–792.

259 Ferreira P, Luco SM, Sawyer SL, Davila J, Boycott KM, Dyment DA. Late diagnosis of cerebral folate deficiency: Fewer seizures with folinic acid in adult siblings. *Neurol Genet* 2016; **2**: e38.

260 Ramaekers V, Sequeira JM, Quadros E V. Clinical recognition and aspects of the cerebral folate deficiency syndromes. *Clin Chem Lab Med* 2013; **51**: 497–511.

261 Seltzer LE, Ma M, Ahmed S *et al.* Epilepsy and outcome in FOXG1-related disorders. *Epilepsia* 2014; **55**: 1292–1300.

262 Brunetti-Pierri N, Paciorkowski AR, Ciccone R *et al.* Duplications of FOXG1 in 14q12 are associated with developmental epilepsy, mental retardation, and severe speech impairment. *Eur J Hum Genet* 2011; **19**: 102–107.

263 Pontrelli G, Cappelletti S, Claps D *et al.* Epilepsy in patients with duplications of chromosome 14 harboring FOXG1. *Pediatr Neurol* 2014; **50**: 530–535.

264 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* 2002; **70**: 530–536.

265 Pierson TM, Yuan H, Marsh ED *et al.* GRIN2A mutation and early-onset epileptic encephalopathy: personalized therapy with memantine. *Ann Clin Transl Neurol* 2014; **1**: 190–198.

266 Lemke JR, Lal D, Reinthaler EM *et al.* Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet* 2013; **45**: 1067–1072.

267 Smigiel R, Kostrzewa G, Kosinska J *et al.* Further evidence for GRIN2B mutation as the cause of severe epileptic encephalopathy. *Am J Med Genet A* 2016; **170**: 3265–3270.

268 Li D, Yuan H, Ortiz-Gonzalez XR *et al.* GRIN2D Recurrent De Novo Dominant Mutation Causes a Severe Epileptic Encephalopathy Treatable with NMDA Receptor Channel Blockers. *Am J Hum Genet* 2016; **99**: 802–816.

269 Weckhuysen S, Ivanovic V, Hendrickx R *et al.* Extending the KCNQ2 encephalopathy spectrum: clinical and neuroimaging findings in 17 patients. *Neurology* 2013; **81**: 1697–1703.

270 Bearden D, Strong A, Ehnot J, DiGiovine M, Dlugos D, Goldberg EM. Targeted treatment of migrating partial seizures of infancy with quinidine. *Ann Neurol* 2014; **76**: 457–461.

271 Mikati MA, Jiang YH, Carboni M *et al.* Quinidine in the treatment of KCNT1-positive epilepsies. *Ann Neurol* 2015; **78**: 995–999.

272 Le Meur N, Holder-Espinasse M, Jaillard S *et al.* MEF2C haploinsufficiency caused by either microdeletion of the 5q14.3 region or mutation is responsible for severe mental retardation with stereotypic movements, epilepsy and/or cerebral malformations. *J Med Genet* 2010; **47**: 22–29.

273 Bienvenu T, Diebold B, Chelly J, Isidor B. Refining the phenotype associated with MEF2C point mutations. *Neurogenetics* 2013; **14**: 71–75.

274 French JA, Lawson JA, Yapici Z *et al.* Adjunctive everolimus therapy for treatment-resistant focal-onset seizures associated with tuberous sclerosis (EXIST-3): a phase 3, randomised, double-blind, placebo-controlled study. *Lancet (London, England)* 2016; **388**: 2153–2163.

275 Mirzaa GM, Campbell CD, Solovieff N *et al.* Association of MTOR Mutations With Developmental Brain Disorders, Including Megalencephaly, Focal Cortical Dysplasia, and Pigmentary Mosaicism. *JAMA Neurol* 2016; **73**: 836–845.

276 Lotte J, Bast T, Borusiak P *et al.* Effectiveness of antiepileptic therapy in patients with PCDH19 mutations. *Seizure* 2016; **35**: 106–110.

277 Hoffmann GF, Schmitt B, Windfuhr M *et al.* Pyridoxal 5'-phosphate may be curative in early-onset epileptic encephalopathy. *J Inherit Metab Dis* 2007; **30**: 96–99.

278 Mills PB, Camuzeaux SSM, Footitt EJ *et al.* Epilepsy due to PNPO mutations: genotype, environment and treatment affect presentation and outcome. *Brain* 2014; **137**: 1350–1360.

279 Guerrini R. Dravet syndrome: the main issues. *Eur J Paediatr Neurol* 2012; **16 Suppl 1**: S1-4.

280 Wallace A, Wirrell E, Kenney-Jung DL. Pharmacotherapy for Dravet Syndrome. *Paediatr Drugs* 2016; **18**: 197–208.

281 Ceulemans B, Schoonjans A-S, Marchau F, Paelinck BP, Lagae L. Five-year extended follow-up status of 10 patients with Dravet syndrome treated with fenfluramine. *Epilepsia* 2016; **57**: e129-34.

282 Chiron C. Current therapeutic procedures in Dravet syndrome. *Dev Med Child Neurol* 2011; **53 Suppl 2**: 16–18.

283 Brunklaus A, Ellis R, Reavey E, Forbes GH, Zuberi SM. Prognostic, clinical and demographic features in SCN1A mutation-positive Dravet syndrome. *Brain* 2012; **135**: 2329–2336.

284 Howell KB, McMahon JM, Carvill GL *et al.* SCN2A encephalopathy: A major cause of epilepsy of infancy with migrating focal seizures. *Neurology* 2015; **85**: 958–966.

285 Larsen J, Carvill GL, Gardella E *et al.* The phenotypic spectrum of SCN8A encephalopathy. *Neurology* 2015; **84**: 480–489.

286 Boerma RS, Braun KP, van den Broek MPH *et al.* Erratum to: Remarkable Phenytoin Sensitivity in 4 Children with SCN8A-related Epilepsy: A Molecular Neuropharmacological Approach. Neurotherapeutics. 2016; **13**: 238.

287 Barker BS, Ottolini M, Wagnon JL, Hollander RM, Meisler MH, Patel MK. The SCN8A encephalopathy mutation p.Ile1327Val displays elevated sensitivity to the anticonvulsant phenytoin. *Epilepsia* 2016; **57**: 1458–1466.

288 Klepper J. GLUT1 deficiency syndrome in clinical practice. *Epilepsy Res* 2012; **100**: 272–277.

289 Pasca L, De Giorgis V, Macasaet JA, Trentani C, Tagliabue A, Veggiotti P. The changing face of dietary therapy for epilepsy. *Eur J Pediatr* 2016; **175**: 1267–1276.

290 Gumus H, Bayram AK, Kardas F *et al.* The Effects of Ketogenic Diet on Seizures, Cognitive Functions, and Other Neurological Disorders in Classical Phenotype of Glucose Transporter 1 Deficiency Syndrome. *Neuropediatrics* 2015; **46**: 313–320.

291 Hussain S. Developing a PPI inhibitor-based therapy for STXBP1 haploinsufficiency-associated epileptic disorders. *Front Mol Neurosci* 2014; **7**: 6.