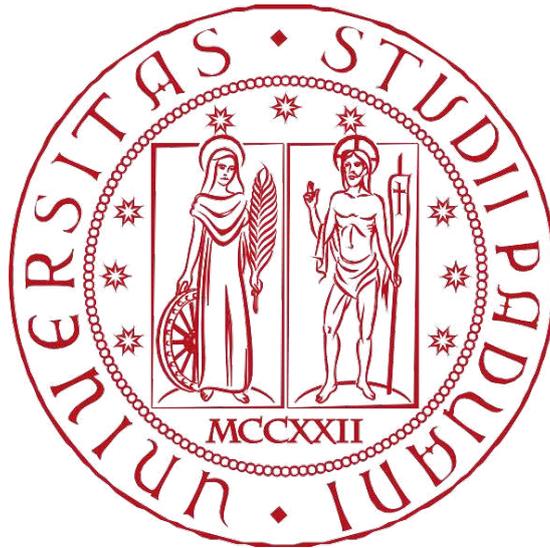


University of Padua

School of Medicine and Surgery

Master's Degree Programme in Medical Biotechnologies

Department of Molecular Medicine



Application of different methodologies for the detection of genetic variants in patients with neurodevelopmental disorders negative for single nucleotide variants

Supervisor: Prof. Antonella Russo

Department of Molecular Medicine, University of Padua

Co-Supervisor: Dr. Giorgia Beffagna

R&I Genetics, Padua

Student: Anna Teatini

Academic year: 2024/2025

INDEX

ABSTRACT	3
INTRODUCTION	4
1. NEURODEVELOPMENTAL DISORDERS	4
1.1 DEFINITION AND CLASSIFICATION OF NEURODEVELOPMENTAL DISORDERS	4
1.2 PATHOPHYSIOLOGY AND GENETIC CONTRIBUTION	5
1.3 EPILEPSY AND NEURODEVELOPMENTAL DISORDERS COMORBIDITY	6
2. EPILEPSY	7
2.1 DEFINITION OF EPILEPSY AND SEIZURES	7
2.2 CLASSIFICATION AND EPIDEMIOLOGY OF SEIZURES AND EPILEPSY	7
2.3 PATHOPHYSIOLOGY AND GENETIC CONTRIBUTION	9
3. SINGLE NUCLEOTIDE VARIANTS (SNVs)	11
3.1 DEFINITION AND MECHANISM OF FORMATION	11
3.2 PATHOLOGICAL ASPECTS OF SNVs	12
3.3 <i>RNU4-2</i> SMALL NUCLEAR RNA.....	12
4. COPY NUMBER VARIATION (CNV)	14
4.1 DEFINITION AND MECHANISMS OF FORMATION	14
4.2 PATHOLOGICAL ASPECTS OF CNVs.....	15
5. COMBINED-METHODOLOGIES APPROACH	17
AIM OF THE STUDY	18
MATERIALS AND METHODS	19
1. COHORT OF PATIENTS	19
2. DNA EXTRACTION	20
3. BIOINFORMATIC ANALYSIS	22
4. IN-DEPTH RESEARCH IN DATABASES AND LITERATURE	26
5. LABORATORY VALIDATION METHODS	27
5.1 REAL TIME PCR.....	27
5.2 ARRAY COMPARATIVE GENOMIC HYBRIDIZATION.....	30
5.3 SANGER SEQUENCING.....	36
RESULTS AND DISCUSSION	43
CONCLUSIONS	55
BIBLIOGRAPHY & SITOGRAPHY	56
APPENDIX	64

ABSTRACT

Neurodevelopmental disorders (NDDs) are a class of early-onset conditions that affect brain development and function, resulting in deficits in different areas such as cognition, communication and motor skills. These disorders affect approximately 15% of children and adolescents worldwide.

A notable feature of NDDs is their frequent comorbidity with epilepsy, a neurological disorder characterized by recurrent seizures. It is estimated that 15-26% of people diagnosed with NDDs also present epilepsy. Both NDDs and epilepsy exhibit significant genetic and clinical heterogeneity, which poses a challenge in achieving accurate molecular diagnosis.

Recent years have seen significant advances in the field of molecular diagnostics, particularly in the analysis of single nucleotide variations (SNVs) and copy number variations (CNVs). These genetic variations have been identified as key targets for improving diagnostic rates in NDDs and epilepsy, conditions that still have some of the highest rates of undiagnosed cases.

The aim of this thesis work was to improve diagnostic outcomes by assessing the impact of genetic variations in NDDs and epilepsy disorders through a combined methodological approach. A group of 100 epilepsy patients, who tested negative for SNVs in ~ 1000 known causative genes, was selected for this study. Available whole-exome sequencing (WES) data was analysed using bioinformatic tools to identify CNVs. Five CNVs were detected and validated through real-time PCR and comparative genomic hybridization (array-CGH). The findings confirmed the presence of CNVs in four patients, suggesting their potential genetic involvement.

A parallel analysis was conducted on another cohort of 100 NDDs patients negative for SNVs and CNVs. Sanger sequencing was performed to identify the presence of variants in *RNU4-2* small nuclear RNA, a key component of the major spliceosome, because recent studies have associated rare mutations in this snRNA with onset of NDDs. One patient was found to be positive for a recurrent *de novo* insertion, n.64-65insT, within an 18-base pair critical region of the *RNU4-2* small nuclear RNA, confirming that *RNU4-2* may be a causative gene for NDDs.

In conclusion, by combining WES with Real Time PCR, array-CGH and Sanger sequencing, I was able to identify rare genetic variations that were previously undetected. These findings highlight the effectiveness of a combined methodological approach in improving the molecular diagnosis for NDDs and epilepsy.

INTRODUCTION

1. NEURODEVELOPMENTAL DISORDERS

1.1 DEFINITION AND CLASSIFICATION OF NEURODEVELOPMENTAL DISORDERS

Neurodevelopmental disorders (NDDs) arise from improper brain development and function [1]. Approximately 15% of children and adolescents worldwide are affected by NDDs, which lead to impaired cognition, communication, adaptive behaviour and psychomotor skills [2]. These conditions present a highly variable phenotype, characterized by common features such as delayed cognitive (from moderate to severe), emotional and motor development [1].

DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition) and ICD-11 (International Classification of Diseases, Eleventh Edition) are two international classification systems used to diagnose and classify mental and behavioural disorders [3,4]. In the latest version of ICD-11, NDDs are included in the chapter 'Mental, behavioural or neurodevelopmental disorders' [3]. Whereas the DSM-5 defines NDDs as a group of early-onset conditions characterized by dysfunction that results in impairment of specific neurological functions [4]. Although ICD-11 and DSM-5 differ in some respects (such as structure and use), they are complementary and are often used together to ensure a correct diagnosis.

NDDs include autism spectrum disorder (ASD), intellectual disability (ID), attention deficit hyperactivity disorder (ADHD), neurodevelopmental motor disorders, such as tic disorders, and specific learning disabilities [2,4]. These disorders are usually diagnosed in the first few months or years of life [2]. While some, such as ASD and ADHD, often persist into adolescence and adulthood, others may regress with age, especially if effective therapeutic treatments are available [2]. The symptoms of the main NDDs are briefly described below:

- ASD is a neurodevelopmental disorder characterized by difficulties in communication and social interaction, repetitive behaviours and challenges in performing daily activities [5]. The term spectrum reflects the wide variability in the manifestation of the phenotype and its severity [5]. ASD is classified as a neurodevelopmental disorder because symptoms typically appear around the age of two [5].
- ID is a neurological disorder characterized by significant cognitive deficits that moderately to severely impair an individual's functional abilities [6]. The DSM-5 defines ID as a neurodevelopmental disorder of childhood onset, characterized by an individual's inability to perform basic activities of daily living [6].
- ADHD is a psychiatric disorder that causes poor concentration, disorganization, lack of attention and memory loss [7]. ADHD has a childhood-onset, with children showing developmentally inappropriate hyperactivity or impulsivity [7]. The DSM-5 classifies this disorder into three categories according to the predominant phenotype observed: predominantly inattentive, predominantly hyperactive, or combined [7].

The comorbidity of different disorders observed in patients with NDDs supports their inclusion in the single category of NDDs [4]. The co-occurrence of symptoms of different conditions has been documented in several cases [4]. For example, many studies have shown that between 22% - 83% of children with ASD, also have symptoms of ADHD, and vice versa [4,8]. In addition, NDDs can be classified as syndromic when they occur alongside additional features. Recent study has investigated the phenotypic spectrum of NDDs, highlighting their frequent association with epilepsy, motor delays, facial dysmorphisms, skeletal abnormalities, and, in few cases, congenital anomalies such as kidney and heart defects [9]. The broad phenotypic variability observed in affected individuals reflects the genetic heterogeneity that underlies NDDs.

1.2 PATHOPHYSIOLOGY AND GENETIC CONTRIBUTION

As mentioned above, comorbidity is a common feature of NDDs. This phenomenon may be explained by common pathogenic mechanisms underlying different NDDs [8]. Some of these pathways are well-characterized, while others remain unidentified. The identified pathways mainly involve synaptic processes [9] that perform crucial functions in the central nervous system (CNS). These include the regulation of presynaptic and ion channels, the control of the membrane potential in the presynaptic part and the management of the localization of neurotransmitter receptors in specialized postsynaptic membranes [10]. The identification of unknown pathways could improve the diagnostic accuracy of NDDs [1] however, their extensive genetic heterogeneity adds significant complexity to the diagnostic process.

To date, approximately 6,000 candidate risk genes have been suggested to be involved in NDDs, of which 1,500 show a significant genotype-phenotype association [11]. Recent studies have explored the wide range of genes associated with NDDs and the results shown that most are expressed either during embryonic development or after-birth [1]. In addition to the large number of genes associated to NDDs, these conditions show remarkable heterogeneity in the genetic variations observed in affected individuals. Chromosomal rearrangements, copy number variations, small insertions and deletions, and point mutations have all been identified in the NDDs genes [1]. These genetic variants mostly arise *de novo*, but sometimes can be inherited, with most alleles following a recessive inheritance pattern [1,11]. This complexity presents challenges in establishing clear associations between specific phenotypes and their corresponding genotypes [9]. Remarkably, more than half of affected individuals remain without a molecular diagnosis [1].

The main factor influencing the genotype-phenotype association in NDDs is whether they are of polygenic or monogenic origin. Polygenic NDDs are characterized by more than one genetic variation [1]. These variations, together with environmental and/or epigenetic factors, influence the phenotypic expression of the patient, increasing the diagnostic complexity [10]. In contrast, monogenic NDDs are caused by a single genetic mutation [10]. Although these disorders are more likely to be accurately diagnosed, they still pose certain challenges. Notably, clinical heterogeneity is also evident in NDDs of monogenic origin.

Currently, next-generation sequencing techniques, such as whole-exome sequencing (WES), allow the identification of all genetic variants in an individual [12]. Despite this progress, establishing a clear correlation between these variants and the phenotypic manifestations in patients with polygenic NDDs remains a challenge. In contrast, genotype-phenotype correlations are more straightforward in monogenic disorders, where a single mutation directly determines the observed phenotypic traits.

As mentioned above, the introduction of techniques such as WES has significantly improved the detection rate of genetic variants in NDDs [12]. Indeed, the yield of WES has been shown to be around 36%, much higher than the yield of chromosomal microarray (CMA), which is around 15-20% [12]. However, a significant proportion of patients remain undiagnosed, highlighting the need for new approaches and studies in the genetics of NDDs.

1.3 EPILEPSY AND NEURODEVELOPMENTAL DISORDERS COMORBIDITY

The DSM-5 defines NDDs as early-onset conditions characterized by developmental deficits. Although epilepsy and epileptic encephalopathy are not classified as NDDs, genetic epilepsy is commonly observed in individuals with these disorders [13]. In particular, about 15-26% of people with NDDs also have epilepsy, compared with a prevalence of about 0.8% in the general population [13,14]. Recent studies have investigated individuals affected by ASD, ADHD and ID, who also have comorbid epilepsy [14,15]. The link between these conditions is thought to be due to the mechanisms underlying epilepsy [14]. Notably, people with epilepsy have an increased risk of developing NDDs compared with people without epilepsy [16].

The hypothesis that epilepsy and NDDs may be influenced by the same or a related causal mechanism is supported by epidemiological studies [17]. Epilepsy results from a disruption in the balance between neuronal excitation and inhibition, which interferes with the proper transmission of signals in the brain [14]. Several epilepsy-related genes have been implicated in the development of NDDs, although there is considerable phenotypic heterogeneity in these disorders [14].

The relationship between these two disorders needs to be carefully assessed [17,18]. Considering the impact of epilepsy in patients with NDDs could improve diagnostic accuracy and lead to more personalized treatment approaches [17,18].

2. EPILEPSY

2.1 DEFINITION OF EPILEPSY AND SEIZURES

Epilepsy is a chronic neurological disorder that affects over 70 million people worldwide [19]. It is characterized by recurrent seizures which are the direct consequence of changes in the brain's electrical activity [20]. Behaviour, movement, sensation and even the level of consciousness can all be affected and changed by epilepsy. It is estimated that up to 10% of the global population experience one or more seizures during their lifetime. However, this does not necessarily mean they have epilepsy [21], which is defined by the presence of one of the following conditions: *(i)* at least two unprovoked seizures that occur more than 24 hours apart; *(ii)* a single unprovoked seizure with high probability of further seizures over the next 10 years ($\geq 60\%$); or *(iii)* features of an epilepsy syndrome [19].

Epileptic seizures occur when a group of cortical neurons, situated within the grey matter of the cerebral cortex, generates an abnormal, intense, and highly synchronized burst of electrical activity [22]. Cortical neurons are of great importance in the human brain, they are highly interconnected and communicate through electrical signals generating a synchronised network of cells [23]. When these neurons become hyper-excited, favouring membrane depolarization, can occur an imbalance between excitatory and inhibitory states, which may result in an epileptic seizure [22]. Most seizure episodes are unpredictable [24] and patients experiencing them are typically unable to control seizures onset or cessation due to their involuntary nature.

The malfunction of one of the mechanisms involved in either the excitation phase, initiating the seizure, or the relaxation phase responsible for its termination, leads to a condition known as status epilepticus (SE) [25]. SE is characterized by the persistence of a seizure over a period of 20 to 30 minutes, or the occurrence of frequent epileptic seizures without the possibility of recovery between episodes [20].

Epilepsy diagnosis is very complicated due to the absence of common signatures, consequences of the high phenotypic heterogeneity of this disease [19]. To date, two techniques are widely used to allow distinction between different types of epilepsy and to check the diagnosis to exclude or confirm it [26]: *(i)* magnetic resonance imaging (MRI) is the standard neuroimaging tool for the detection of potential brain lesions or abnormalities caused by epilepsy, allowing precise diagnosis [19,26]; *(ii)* electroencephalography (EEG) instead, is a test that measures brain's electrical activity by placing small metal discs, called electrodes, on the patient scalp [27].

2.2 CLASSIFICATION AND EPIDEMIOLOGY OF SEIZURES AND EPILEPSY

In 2017, the International League Against Epilepsy (ILAE) published an updated classification of epilepsy and seizures, which is done at three levels: seizure types, epilepsy types and epilepsy syndromes [20]. Because of their extreme importance, causes and comorbidities should be identified for the diagnosis and for the selection of an appropriate medical therapy [19]. There are six main causes that have been shown to play a role in seizures and epilepsy: genetic, structural, metabolic, infectious, immune and unknown (**Figure 1**) [19].

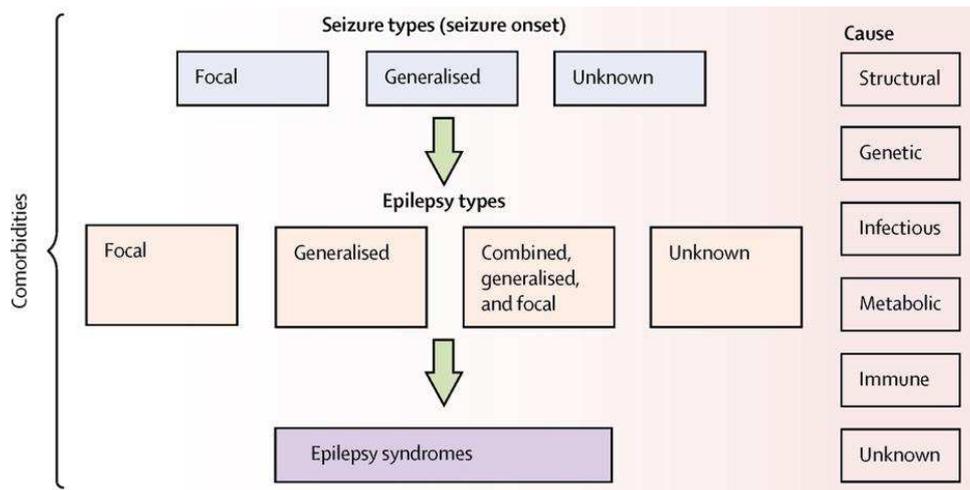


Figure 1. Epilepsy classification and causes (from Ingrid E Scheffer et al., 2017).

Seizures are classified as focal, generalized or unknown depending on the region of onset [19]. They are defined as focal if they start in a limited area of the brain [28] from which they can spread to other regions. Normally the patients remain conscious during focal seizures [28]. If both sides of the brain are affected [29] from the beginning, the seizure is classified as generalized. Because they affect the entire cerebral cortex, they cause the patient to lose consciousness [28]. If the onset of the seizure is not known, for example when it occurs at night, the seizure is defined as unknown [28]. With more information from the patient, a seizure of unknown onset can be classified as focal or generalized [28].

Generalized seizures are characterized by (i) motor symptoms, which may include persistent jerky movements (clonic), brief muscle spasms (myoclonic) or tightening (tonic) of muscles, and (ii) non-motor symptoms, usually known as absence seizures (short and sudden loss of consciousness) [29]. In focal onset seizures, (i) motor symptoms resemble those of generalized seizures, while (ii) non-motor symptoms are characterized by seesawing emotions, alterations in involuntary functions and behaviour arrest (temporary cessation of all body movements) [28]. For unknown onset seizures (i) motor symptoms are primarily tonic-clonic movements and (ii) non-motor symptoms involve behaviour arrest (**Figure 2**) [30,31].

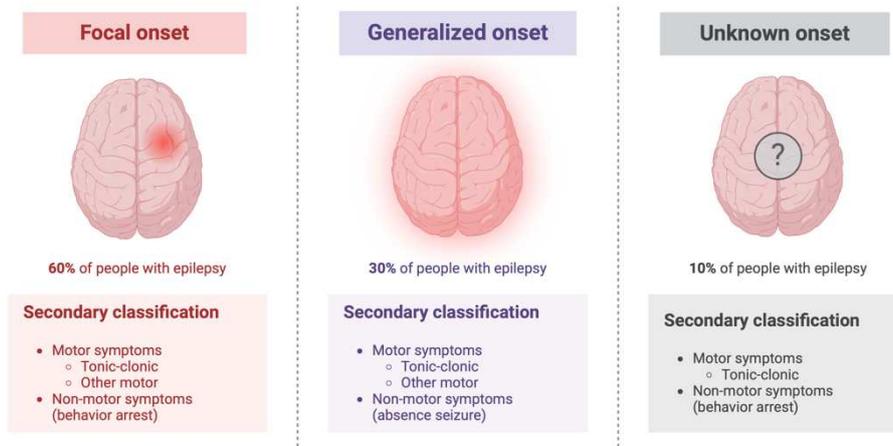


Figure 2. Seizures onset classification. Focal, generalized and unknown onset seizures with incidence rates and main associated symptoms (modified template from Nashed M, 2023; BioRender.com).

Epilepsy predominantly affects infants and children, with an incidence of 0.5-1% worldwide, making it one of the most common chronic neurological disorders in this age group [20]. Although mortality is not a major consequence, there are nearly 125,000 deaths worldwide each year [32].

Epilepsy-associated risk factors are strongly related with age [19]. Malformations that occur during brain development, leading to several neurological defects, are typically more prevalent in childhood [19]. The onset of epilepsy can be the result of pre-existing conditions or chronic conditions as trauma, infections and tumours in the patient [19].

2.3 PATHOPHYSIOLOGY AND GENETIC CONTRIBUTION

The ILAE classification divides epilepsies into three etiological categories: idiopathic, symptomatic, and cryptogenic [24]. Idiopathic epilepsy is thought to result from a genetic defect, although no specific cause has been identified [24]. Structural disorders such as malformation, tumour or trauma are responsible for symptomatic epilepsies [24]. If the cause is unknown, epilepsy is classified as cryptogenic [24].

Epileptogenesis is defined as the dynamic process by which epilepsy develops [33], resulting from an imbalance between excitatory and inhibitory states of neurons [34]. It is associated with disruption of ion homeostasis, alterations in energy metabolism and receptor functionality, resulting in inappropriate neurotransmitter uptake and excessive neurotransmitter release [35,36]. Specifically, in the Central Nervous System (CNS), nerve cells interact and communicate through axons, by sending out electrical impulses which are generated by the release of chemicals, known as neurotransmitters [23]. Homeostasis in the CNS is maintained by neurons and glial cells, including astrocytes [36,37]. These are responsible for maintaining the correct ion concentration through ion pumps and sequestering neurotransmitters from the extracellular space to restore a balanced brain environment [38]. Even in the presence of these recovery mechanisms, an irreversible consequence of seizures is the neuronal death [39]. Usually, a single, isolated epileptic seizure, does not cause serious damage; however, frequent seizures (e.g. SE) can have serious consequences for the brain, including nerve cell death [39].

Although epilepsy can be caused by multiple factors, genetic is the most significant in terms of diagnosis [41]. It is estimated that 40–60% of all epilepsies have a genetic origin [40]. Epilepsy is considered to be of genetic origin if there is a known or suspected specific disease-causing variant [41].

As discussed above, the term "epilepsy" refers to a group of disorders with different features such as seizure types, age of onset and potential comorbidities with other diseases [42,43]. This wide clinical and phenotypic diversity reflects the genetic heterogeneity underlying these disorders. To date, 1,506 genes have been linked to epilepsy [43]. Of these, 168 are associated with epilepsy as an isolated condition, 364 with comorbid epilepsy and NDDs, and 974 with epilepsy associated with severe physical and/or systemic abnormalities [43].

Genetic epilepsies can be classified as monogenic, caused by a single gene mutation (either inherited or *de novo*), or polygenic, where the phenotype results from mutations in multiple genes [41]. Approximately 40% of severe epilepsies are monogenic [44]. While single-gene mutations can simplify genotype-phenotype correlations, epilepsy remains clinically highly heterogeneous [43]. This heterogeneity means that a single gene can produce several phenotypes with different degrees of penetrance in the population. The genetic spectrum of epilepsy-associated mutations is broad: in some individuals, epilepsy is caused directly by a specific genetic mutation, whereas in others it results from an underlying genetic predisposition

(e.g. a mutation in an epilepsy-associated gene) that is later triggered by second hits such as environmental factors or additional genetic variations [19].

The primary genetic mutations identified in epilepsy range from single nucleotide variants (SNVs) to large copy number variations (CNVs) [41]. Most variations are *de novo*, but some are inherited and often lead to familial epilepsies [45]. These are a group of disorders defined by a shared genetic cause within a family. Familial epilepsies can be either monogenic or polygenic [45].

Recently, the advent of next-generation sequencing (NGS) technologies has enabled the discovery of new genes involved in epilepsy [45]. Of these technologies, WES has proved particularly valuable, as it allows the analysis of all coding regions of the genome [46]. By combining WES data with epilepsy-specific virtual genetic panels, it is possible to filter identified variants and achieve faster molecular diagnosis [46]. However, establishing clear genotype-phenotype correlations remains a major challenge. In this context, monogenic epilepsy offers a more straightforward route to accurate correlation and molecular diagnosis.

3. SINGLE NUCLEOTIDE VARIANTS (SNVs)

3.1 DEFINITION AND MECHANISM OF FORMATION

One of the most common types of genetic mutation is the single-nucleotide variant (SNV). It is a variation of a single nucleotide in the DNA sequence, occurring in a defined position (**Figure 3**) [47]. SNVs are a major contributor to the genetic diversity among individuals of the same species [48]. On average, a human individual has around 5,000,000 SNVs when compared to a reference genome [49]. Variants with a population frequency greater than 1%, are defined as single nucleotide polymorphisms (SNPs) [48].

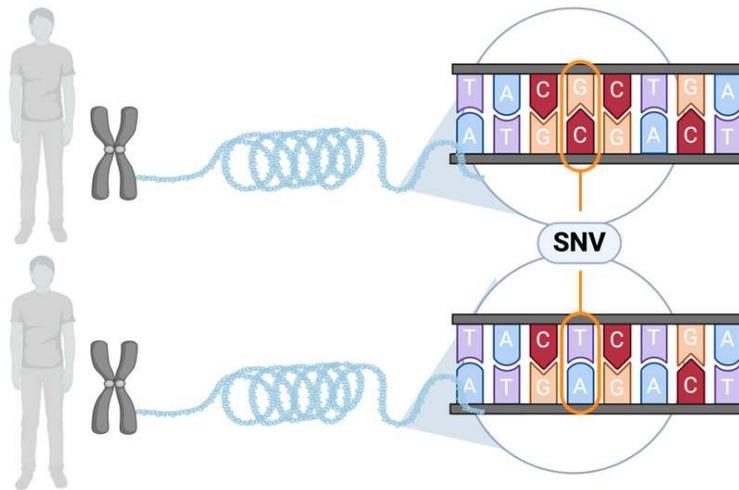


Figure 3. Single nucleotide variant (SNV). Each SNV is a point mutation involving the substitution of a single nitrogenous base in the DNA. The image shows an example where one base differs between two individuals, highlighting the genetic variation. Created with [BioRender.com](https://www.biorender.com).

SNVs can either be inherited, if they are present as germline mutations in one or both parents, or they can occur *de novo*, as germline or somatic mutations [50]. SNVs can be found in both non-coding and coding regions of the genome [51]. The latter can be classified into two categories:

- Synonymous SNVs: when the substituted base does not involve the coding of a different amino acid in the protein [52].
- Non-synonymous SNVs: when the substituted base causes the triplet to code for a different amino acid (missense substitution) or a stop codon (nonsense substitution) [52].

SNVs can be the result of the action of both biological and environmental factors, or they can occur by chance. DNA replication is one of the mechanisms by which SNVs can be introduced. Indeed, the DNA polymerase can occasionally insert the wrong nucleotide causing a mismatch that, if left unresolved, can lead to a permanent mutation in the genome [53]. Mutations can also arise from environmental factors such as radiation and chemicals [54,55] These factors affect DNA by modifying the chemical groups on nitrogenous bases, potentially leading to SNVs if not properly repaired [55].

Although most SNVs are classified as polymorphisms, some have been associated with several different diseases [54], highlighting their potential role in genetic diagnostics.

3.2 PATHOLOGICAL ASPECTS OF SNVs

In recent decades, improvements in genetic analysis techniques have led to the association of SNVs with human diseases [56]. Indeed, SNVs represent the most common form of human genetic variation and are often the primary cause of disease [57]. Clinically, these variants are important because they alter the sequence and consequentially the activity of the affected gene products [58], often leading to the functional impairments responsible for the clinical manifestations observed in NDDs, including epilepsy [59,60].

Remarkably, approximately 98.5% of the genome consists of non-coding regions, encouraging research into their potential association with NDDs [61]. One interesting finding concerns small non-coding RNAs (sncRNAs). They account for 37.4% of RNA sequences [61] and, although many of their functions are still unclear, they are known to play a central role in regulating chromatin expression, modulating gene transcription and controlling splicing [62]. A subcategory of small non-coding RNA (sncRNAs) is represented by small nuclear RNAs (snRNAs). They are involved in the formation of the spliceosome complex, which plays an essential role removing introns for the maturation of the messenger RNA (mRNA) [63].

3.3 *RNU4-2* SMALL NUCLEAR RNA

Recent studies have identified rare variants in *RNU4-2* in NDDs patients, suggesting its potential role as a causative gene for NDDs [64,65]. *RNU4-2* is located on chromosome 12 and it encodes U4, a small nuclear RNA (snRNA) which is a key component of the major spliceosome [61]. The U4 mature snRNA binds to the U6 snRNA through complementary base pairing and, together with the U5 snRNA, they form the U4/U6.U5 tri-snRNP major spliceosome [61]. U1 and U2 snRNAs bind to the 5' splice site (5'SS) pre-mRNA [64] and recruit the U4/U6.U5 tri-snRNP, forming the precatalytic complex [66]. While bound to U4 snRNA, U6 snRNA is maintained into inactive conformation, which makes the spliceosomal complex inactive as well [65]. Activation is mediated by a helicase that dissociates the U1 snRNA from the 5'SS and transfers the 5'SS to the ACAGAGA box of the U6 snRNA, forming the B complex [65]. Subsequently, the BRR2 helicase unwinds the U4/U6 snRNA duplex, causing the dissociation of U4 snRNA [66]. This allows U6 snRNA to fold and bind to U2 snRNA [65], triggering the rearrangement that leads to the formation of the active catalytic spliceosome [67].

Genetic studies comparing individuals with and without NDDs, have identified single nucleotide insertions, deletions and variations in a specific 18 bp region in the middle of *RNU4-2* in some patients [61]. This region is located between stem I, where U4 and U6 snRNAs pair, and the 3' stem-loop structures (**Figure 4**) [61]. Variants in this specific 18bp region disrupt the proper positioning of the U6 ACAGAGA box, which is critical for 5'SS binding, and destabilize the interaction between the BRR2 helicase and U4 [61,64].

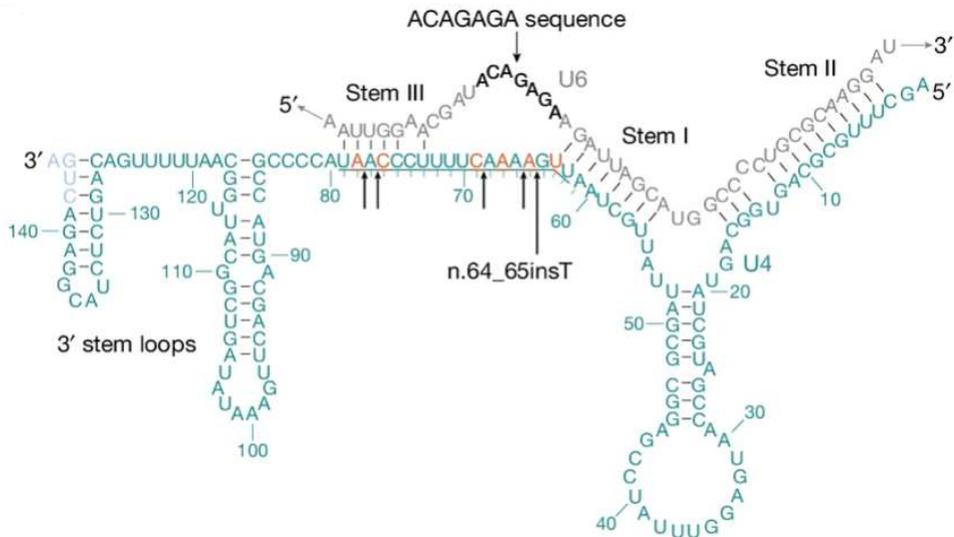


Figure 4. Duplex U4/U6 snRNAs. Schematic representation of the secondary structure of U4 (blue) and U6 (grey) RNAs highlighting their complementary pairing regions and stem structures. The 18 bp region, spanning from nucleotide n.62 to n.79, is underlined. Arrows mark the locations of base insertions, while bases involved in SNVs are highlighted in orange (Adapted from Chen Y. et al., 2024)

As previously mentioned, these studies identified an 18 bp region within the *RNU4-2* gene (GRCh38: chr. 12: 120291825-120291842) where the most significant variants are located [61]. It is estimated that defects in this region could account for approximately 0.4% of NDDs [61]. The insertion of a thymine between nucleotide n.64 and n.65 was found to be the most recurrent variant among patients (GRCh38:chr:12:120291839:T:TA; n. 64_65insT) [61,64,65]. Other pathogenic mutations that were found in NDDs affected individuals are the following: n.77_78insT, n.64_65insG, 76C>T, 76C>G, 67A>G, 65A>G [61,64].

The correlation between NDDs and *RNU4-2* variants is supported by the high expression of this gene in the brain [61]. While U4 snRNA is expressed by several genes, only two, *RNU4-2* and *RNU4-1*, are active in the brain [61]. Despite their 97.2% sequence similarity, no variants associated with *RNU4-1* have been linked to NDDs to date [61]. In addition, *RNU4-2* shows higher expression in the brain than *RNU4-1*, suggesting that the U4 snRNA encoded by *RNU4-2* is the predominant transcript in this tissue [61].

Characterization of the phenotypic spectrum of individuals with variants in the critical 18 bp region of *RNU4-2* revealed several common features [68]. All individuals have moderate to severe intellectual disability, often associated with developmental delay [61,68]. Almost all cases have epileptic seizures, with focal epilepsy being the most observed seizure type [61]. Microcephaly is a shared feature in almost all patients, together with characteristic dysmorphic features. [68]. Other identified phenotypes include growth delay, hypotonia and structural brain abnormalities detected by imaging studies [68].

4. COPY NUMBER VARIATION (CNV)

4.1 DEFINITION AND MECHANISMS OF FORMATION

Copy number variation (CNV) defines a molecular phenomenon characterized by differential dosage of specific genomic segments [69]. The number of repeats in these segments can vary among individuals within the same species [70]. CNVs usually manifest as duplications or deletions, leading to gains or losses of genomic material [69]. Unlike small variations in the genome (indels), CNVs involve larger insertions or deletions, ranging in size from 50 base pairs (bp) to several megabases (Mb) [71]. Compared to single nucleotide variants, CNVs affect a larger fraction of the genome, approximately 4.8-9.5% [70].

To date, several different molecular mechanisms have been involved in the development of CNVs, including processes related to DNA recombination, replication, and repair [70]. In addition, environmental factors are thought to contribute to the formation of CNVs; chemical mutagens (e.g., hydroxyurea) act on specific regions of the genome, whereas physical factors (e.g., ionizing radiation) have more random targets [70]. Different outcomes are observed depending on the genomic region affected [71]. Indeed, small variations, either intergenic or in genes that can tolerate them, account for benign CNVs [71]. In contrast, CNVs that potentially cause disease or phenotypic changes, may involve one or more genes, regulatory regions or elements [71].

CNVs are classified as recurrent or non-recurrent (**Figure 5**) [70], based on the mechanisms that cause them and their frequency in the population. Recurrent CNVs are genomic alterations having the same size and clustering of breakpoints, among unrelated individuals [70]. Breakpoints occur in proximity of specific regions known as low copy repeats (LCRs), also termed “segmental duplications”, which are blocks of DNA between 10 and 400 kb with more than 95% sequence identity [72].

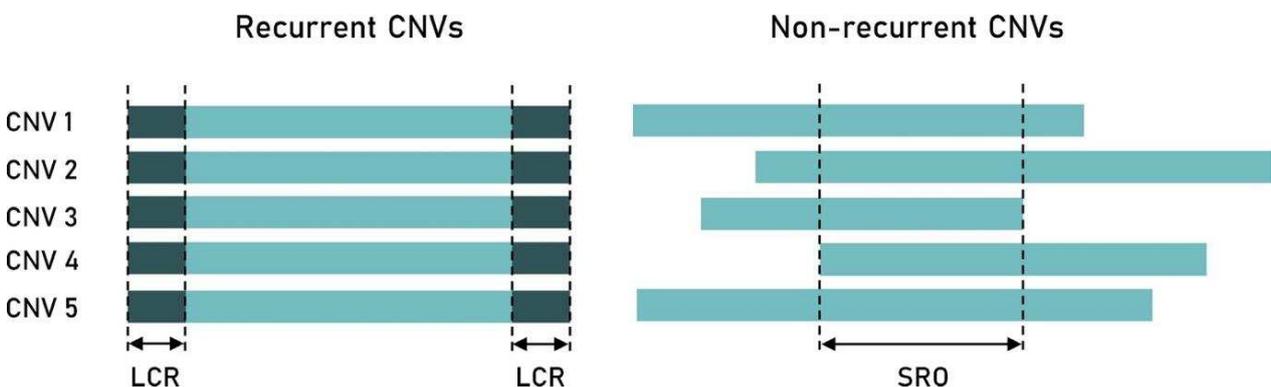


Figure 5. Recurrent and non-recurrent CNVs; the dashed lines highlight low copy repeats (LCRSs) in recurrent CNVs, and the smallest region of overlap (SRO) in non-recurrent CNVs [70].

The formation of recurrent CNVs is driven by several mechanisms, of which non-allelic homologous recombination (NAHR) is particularly relevant [73]. NAHR is a general recombination mechanism between homologous DNA sequences situated on different alleles [74]. Because of their high degree of similarity, non-allelic LCRs may align during mitosis or meiosis triggering NAHR, and the resulting crossing-over leads to genomic rearrangements in the daughter cells [74].

Recurrent rearrangements mediated by NAHR, can introduce simple (such as a deletion, duplication or inversion) or complex CNVs [75]. These events result from a combination of different types of rearrangements occurring within the same cluster (e.g., an inversion combined with a duplication) (**Figure 6; d**) [75]. Simple and complex CNVs also result from non-recurrent recombination events [76].

The type of CNV, resulting from the rearrangement, is determined by the location and orientation of the LCRs [77]. Deletions and duplications result from rearrangements between LCRs with direct orientation on different chromatids of different chromosomes (interchromosomal) (**Figure 6; a**) or on chromatids of the same chromosome (intrachromosomal) (**Figure 6; b**); whereas misalignment of LCRs on the same chromatid (intrachromatid) (**Figure 6; c**) results in deletion [77]. Inversion events occur when LCRs have opposite orientations (**Figure 6; e**) [77,78]. NAHR between misaligned interchromosomal/intrachromosomal LCRs can induce translocations. [78].

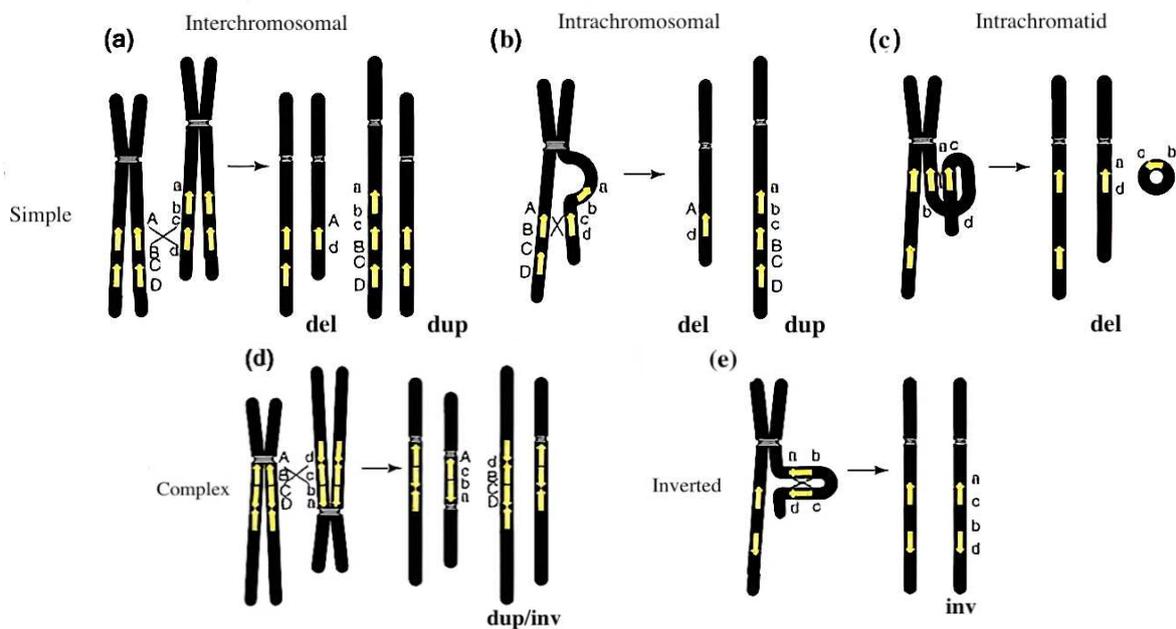


Figure 6. CNVs chromosomal rearrangements. The upper- and lower-case letters flanking the yellow arrows refer to the unique sequences included in the rearrangement. Abbreviations: duplications (dup), deletions (del), inversions (inv); (Adapted from Stankiewicz & Lupski, 2002).

Non-recurrent CNVs are unique or rare genomic alterations that occur at different loci in different individuals [69]. They have variable sizes due to interspersed breakpoints, but they may share a common region known as smallest region of overlap (SRO) [69]. It is rare to find the same non-recurrent CNV in more than one patient, but since SRO is shared between similar CNVs, it may be of interest to link the affected gene(s) to a possible phenotype(s) [70].

4.2 PATHOLOGICAL ASPECTS OF CNVs

Over the past two decades, CNVs have been acknowledged for their important role in disease development [79]. This occurs when deletions, duplications or other genomic rearrangements disrupt gene dosage,

resulting in loss or gain of genetic material [69]. In addition to dosage imbalance, CNVs can affect coding sequences or regulatory regions of genes, leading to altered gene expression or function [70]. The structural properties of CNVs, including the type of mutation, the size of the event, and the function of the involved gene(s), are key factors in the development of various diseases [71,80].

In recent years, an increasing number of unbalanced chromosomal rearrangements (such as deletions and duplications) have been associated with neurocognitive disorders, including epilepsy, strengthening the role of these mutations in the genetic aetiology of this group of conditions [80,81]. Indeed, studies suggest that 1-4% of individuals with epilepsy have identifiable diagnostic CNVs [82]. This means that in a small but significant percentage of epilepsy cases, the underlying cause can be traced back to these structural genomic changes [82]. Therefore, analysing epilepsy patients for the presence of CNVs may play a crucial role in achieving a possible molecular diagnosis [81,83].

5. COMBINED-METHODOLOGIES APPROACH

Identifying genetic variations in NDDs, including epilepsy, requires the use of different tools. In recent years, the increasing use of databases has greatly improved our understanding of these conditions. These systematically organised data collections contain genetic, phenotypic and clinical information, making them a valuable tool for researchers [83]. By allowing in-depth study of NDDs associated genetic variants, databases help to define genotype-phenotype relationships [83]. A large proportion of this data has been generated by the extensive use of next-generation sequencing (NGS) in patient analysis [1]. Among NGS techniques, whole-exome sequencing (WES) is currently the most commonly used approach for genetic diagnosis of NDDs, including epilepsy [84]. WES enables the identification of SNVs and small indels in genes associated with Mendelian disorders [84].

Recent studies have demonstrated that integrating multiple diagnostic methodologies can improve the molecular diagnosis of NDDs, including epilepsy [84]. For instance, a study from Tsuchida et al. (2018), investigated a group of epilepsy patients, who had previously tested negative for SNVs, focusing on the detection of CNVs. Samples were analysed with WES and identified CNVs were confirmed with quantitative PCR. This combined approach provided a more complete understanding of the genetic factors contributing to the disease.

Another study by Goh J., et al. (2025) employed a combined diagnostic approach using WES and Sanger sequencing in SNV-negative patients with NDDs. This strategy was employed to detect single nucleotide insertions in the *RNU4-2* snRNA, which has recently been proposed as a candidate gene for NDDs [61,64].

In conclusion, the implementation of a combined diagnostic approach is essential to improve the molecular diagnosis of NDDs, including epilepsy.

AIM OF THE STUDY

Establishing a precise genotype-phenotype correlation is currently one of the major challenges in the study of NDDs, including epilepsy. In recent years, the advent of next-generation sequencing (NGS) techniques has made it possible to analyse the entire genome and expand genetic databases. However, the analysis of SNVs is often insufficient, leaving many patients without a molecular diagnosis. Recent studies have highlighted the crucial role of CNVs in NDDs, recommended their inclusion in diagnostic analysis.

During my thesis work at the R&I Genetics Laboratory in Padua, I am evaluating the impact of CNVs in a cohort of 100 epilepsy patients who were negative in previous SNVs screenings. Using bioinformatics tools, I search for CNVs and prioritise the identified variants by consulting the scientific literature and major genetic databases. Then I validate the presence of CNVs by real-time PCR or array-CGH, depending on the size of the event, and confirm the actual presence of the genetic variation. I also analyse samples from available family members to determine the inheritance pattern or *de novo* nature of the identified CNVs.

At the same time, I am conducting a study on a second group of 100 patients with NDDs, independent from the first cohort, who are negative for SNVs and CNVs, to identify possible mutations involving the *RNU4-2* snRNA, a key component of the major spliceosome. Recent research suggests a potential role for *RNU4-2* in the pathogenesis of NDDs, making it a compelling candidate for further investigation. As this region of interest is non-coding and undetectable by WES, I am analysing the samples using Sanger sequencing technique.

By using these complementary approaches, this study aims to increase the diagnostic yield and improve our understanding of the genetic basis of epilepsy and NDDs by providing insight into the contribution of CNVs and non-coding mutations in these conditions.

MATERIALS AND METHODS

1. COHORT OF PATIENTS

This thesis was conducted on two independent patient cohorts.

The first cohort consisted of 100 patients analysed between 2018 and 2024. All individuals were diagnosed with epilepsy and had previously undergone WES for SNVs screening, with yielded negative results, leaving them without a molecular diagnosis. This group of 100 patients included 64 females and 36 males, ranging in age¹ from 0 to 51 years, with the majority clustered around an average age of 15 years. All individuals were unrelated, and in addition, samples from one or both parents were available for 40 patients. Samples from this cohort of patients were analysed using VarSeq software to detect CNVs. The most relevant events were genotype-phenotype correlated through extensive database and literature searches. Finally, CNVs with a confirmed genotype-phenotype correlation were validated using real-time PCR or array-CGH.

A parallel analysis was conducted on a second cohort of 100 patients with NDDs. These individuals had previously undergone WES, followed by VarSeq bioinformatic analysis, for both SNVs and CNVs in NDDs known genes. However, no significant variants were identified, resulting in a lack of molecular diagnosis. The cohort included 53 females and 47 males, ranging in age¹ from 0 to 35 years, with the majority clustered around an average age of 10 years. For 67 probands there was the availability of one or both parents. Samples from this cohort of patients were analysed by Sanger sequencing to detect variants in *RNU4-2* snRNA, recently identified as a potential causative gene in NDDs [64,65].

¹ Patient's age at time of sample collection.

2. DNA EXTRACTION

DNA extraction from blood samples was performed using the QIASymphony® SP instrument exclusively on the parental samples of the four epilepsy patients with CNVs of interest, as the proband samples had already been extracted for a previous analysis. It enables the purification of nucleic acids, providing high-quality products, essential for downstream applications, such as next-generation sequencing (NGS) [85]. The QIASymphony® SP can process up to 96 samples [85], significantly reducing the processing time. In addition, by eliminating manual pipetting steps, that are usually prone to human error, the system ensures reliability and reproducibility [85]. In this work, the QIASymphony® DNA Midi Kit was used in combination with Blood 400 DSP protocol. This protocol requires a starting volume of 400 μ L of blood sample and produces a 200 μ L elution volume containing the purified DNA. The principal components of the kit are reported in the **Table 1** [86].

Table 1. QIASymphony® DNA Midi Kit active ingredients of the principal components.

Reagent	Component	Concentration (w/w) %
Reagent Cartridge (RC)	Maleic acid	≥ 0.1 to <1
	Guanidine hydrochloride	≥ 30 to <50
	Nonionic detergent	≥ 1 to <25
	Ethanol	≥ 10 to <90
	Isopropanol	≥ 30 to <50
	Lithium chloride	≥ 1 to <10
	Guanidinium thiocyanate	≥ 20 to <30
Enzyme Rack (ER)	Proteinase K	≥ 1 to <10

The other components of the kit are the following:

- Buffer ATE (ATE): is a low EDTA elution buffer containing 10 mM Tris-Cl pH 8.3, 0.1 mM EDTA and 0.04% NaN₃ (sodium azide);
- Piercing Lid (PL);
- Reuse Seal Set (RSS): contains 8 Reuse Seal Strips.

Before starting the run, samples (barcoded tubes) were kept at room temperature (15–25°C). The samples were manually loaded into the loading lane, that was gently inserted into the instrument to allow the scan of the barcodes and to ensure proper identification and tracking. DNA purification was performed in 4 steps: lysis, binding, washing and elution. For each sample, the instrument collected 400 microliters of blood using specially designed tips that employ conductivity to determine the optimal collection level [86]. This system allows the tips to be lowered just enough to take the desired volume while simultaneously limiting the risk of contamination. The volume taken from each sample was transferred to a dedicated well for the lysis step that was achieved by adding the solution containing proteinase K.

QIAGEN magnetic particles were used for the isolation of the DNA from the blood (**Figure 7**). These particles have a silica-coated surface [86], which allows nucleic acids to bind to the beads. The magnetic particles were added to each lysed sample and the solution was transferred to sample preparation cartridges. Next, a magnetic rod protected by a lid, enters the well containing the sample and attracts the magnetic particles that bind to its surface [86]. The magnetic bar is then placed over another well containing a washing buffer and the magnetic particles are released. The binding and washing steps are repeated several times. For elution of the purified DNA, the magnetic bar with the particles is placed over a well containing the elution buffer ATE, which allows separation of the DNA from the magnetic particles, that instead remain attached to the magnetic bar [86].

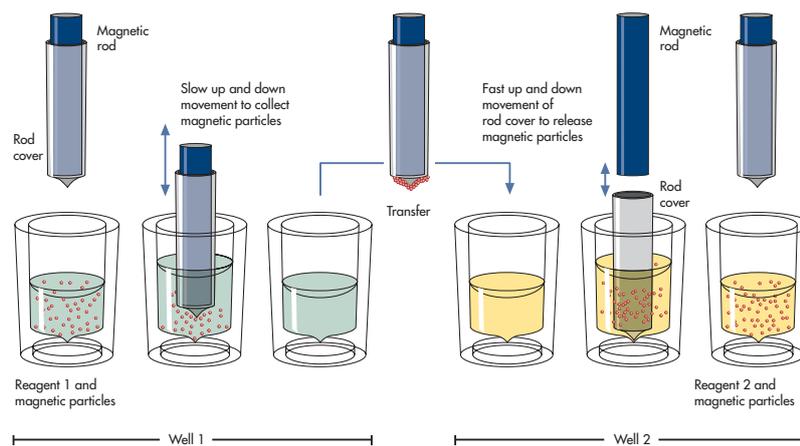


Figure 7. QIASymphony Instrument DNA Purification Workflow. Reagent 1 contains the lysate solution and magnetic particles. The magnetic particles bind to the DNA and then attach to the magnetic bar. The magnetic bar releases the particles into Reagent 2 (Wash Buffer). Steps 1 and 2 are repeated several times during the sample purification process [86].

The purified DNA obtained from each sample was then quantified using Thermo Scientific™ NanoDrop™ Lite Plus spectrophotometer. The quantification was performed as follows:

- 1) Selection of Nucleic Acid Type: dsDNA was selected as the nucleic acid type on the NanoDrop™ Lite Plus Home screen [87];
- 2) Blank Reading: 1 μL of water was added to the pedestal for the blank reading by selecting “Blank” [87];
- 3) Sample Measurement: after cleaning the pedestal, 1 μL of sample was loaded and the concentration was measured by selecting 'Measure' [87];
- 4) Repeat for Each Sample: Step 3 was repeated for each sample and the measured concentrations were noted on the respective sample tubes.

3. BIOINFORMATIC ANALYSIS

Whole-exome sequencing (WES) of the DNA samples was performed previously for the analysis of SNVs. The sequencing reads obtained from these patients were then aligned to a reference genome (GRCh37/hg19) to identify any variants present in the samples. The data that had been obtained from the previous alignment were collected in BAM (Binary Alignment Map) and VCF (Variant Call Format) files, which were then imported into the VarSeq™ 2.6.2 software (Golden Helix, Inc.) VarSeq identifies CNVs from NGS coverage data by comparing test samples with reference samples. Coverage represents the average number of reads in a specific exonic region. A minimum of 10 and a maximum of 50 reference samples were used for normalization and calling of the events to ensure a balance between quality and dataset size. Reference samples with percentage differences greater than 20% compared to the analysed samples, were excluded to improve accuracy. In addition, only same-sex controls were used for sex chromosome normalization. The coverage of each region within a sample were compared with that of other exonic regions to normalize any variation.

The CNVs identified by the software were filtered using two gene panels: one containing 1,059 primary genes linked to epilepsy, and another containing 655 candidate genes whose potential association with epilepsy requires further validation. The events were then filtered according to the following criteria (**Figure 8**):

- *CNV type*: the event categories considered were duplications, heterozygous deletions and deletions;
- *Flags*: Samples were selected only if no quality control warnings (flags) were associated with them. The categories of flags are detailed in **Table 2**;
- *#Samples*: number of samples in which the same event is found, without distinction between deletion and duplication, but only focusing on the region involved. In the present thesis work, the number of samples was ≤ 3 , accounting for the possibility of the same event being shared within a trio;
- *#Match type*: summary of the number of CNV events from a pool of 393,355 CNVs across 492 samples that overlap the query CNV by at least 20% (similarity coefficient $\geq 20\%$), counting only those events that share the same CNV type. This can result in different overlap types (reported in **Table 3**).

Samples with a Sample % 100x (the percentage of bases with 100x coverage or higher over the entire region) greater than 70 were considered reliable for CNV analysis. In addition, the average percentage difference between the sample and paired controls for both the autosomal and allosomal regions of the tested samples, was selected with a maximum value of 10.

For each sample, the VarSeq software allows the CNV event and the exonic region involved to be visualized graphically. For each CNV, the following information was reported:

- *Region*: genomic coordinates (start and end) of the event;
- *Span*: size of the event (bp);
- *#Targets*: number of exon(s) involved in the event;
- *Gene Names*: gene(s) involved in the event;
- *Overlapping Exons (Clinically Relevant)*: refers to the exons involved in the event, the term 'all' is used when the CNV overlaps to the whole gene (s) (RefSeq Genes 105.20220307, NCBI);

- *Transcript Name* (Clinically Relevant): refers to the specific transcript of a gene (RefSeq Genes 105.20220307, NCBI);
- *Disorders*: Description of the associated disorders to the genes involved in the event, followed by the OMIM ID, if available (OMIM Genes 2024-09-01);
- *Disorders Inheritance*: inheritance information of the disorders (autosomal recessive, autosomal dominant, X-linked recessive and X-linked dominant). “Missing” is displayed if the inheritance information is not available (OMIM Genes 2024-09-01);
- *CNV state*: type of the CNV (Duplication, Deletion or Het Heletion) (RefSeq Genes 105.20220307, NCBI);
- *Avg Z Score*: average Z-score of the targets in the event. The Z score is the number of standard deviations a value is from the reference samples. It was calculated as (normalized target depth – references average depth)/standard deviation. Ideal Z score values are as follows:
 - Z score=0 if the region is wild type;
 - Z score ≥ 3 for duplication;
 - Z score ≤ -3 for deletion.
- *Avg Ratio*: average ratio of the targets in the event: This value indicates the ratio between the coverage of the sample under analysis and the average coverage of the reference samples. It is obtained by dividing the normalized target depth by the mean control depth. Ideal Ratio values for autosomic regions are as follows:
 - Ratio=1 for wild type;
 - Ratio=1.5 for heterozygous duplication;
 - Ratio=0.5 for heterozygous deletion;
 - Ratio=0 for homozygous deletion;
 - Ratio=2 for the presence of 4 copies.

The Ideal Ratio values for allosomal regions will differ due to the X chromosome dosage difference between males and females, which is 1:2. Therefore, the expected ratios are as follows:

- In a female subject:
 - Ratio = 1 for wild type;
 - Ratio = 1.5 for heterozygous duplication;
 - Ratio = 0.5 for heterozygous deletion;
 - Ratio = 0 for homozygous deletion;
- In a male subject:
 - Ratio = 1 for a single-copy duplication;
 - Ratio = 1.5 for a two-copy duplication;
 - Ratio = 0.5 for wild type;
 - Ratio = 0 for deletion.
- *p value*: p-value: probability that Z-scores, as extreme as those of the event, occurs by chance in a diploid region. If the Z-score is very high or very low (i.e. very far from the mean), the p-value will

be very small, indicating that it is very unlikely to occur by chance in a normal distribution. Thus, values tending towards zero indicate good call reliability.

The values given for the Z score and the Ratio are ideal values; for a complete and correct interpretation, reference has been made to the values given in **Table 2**. Following the identification of the events with the most reliable values, genotype-phenotype correlation was performed to highlight CNVs that might be involved in the specific pathology of the patient being analysed.

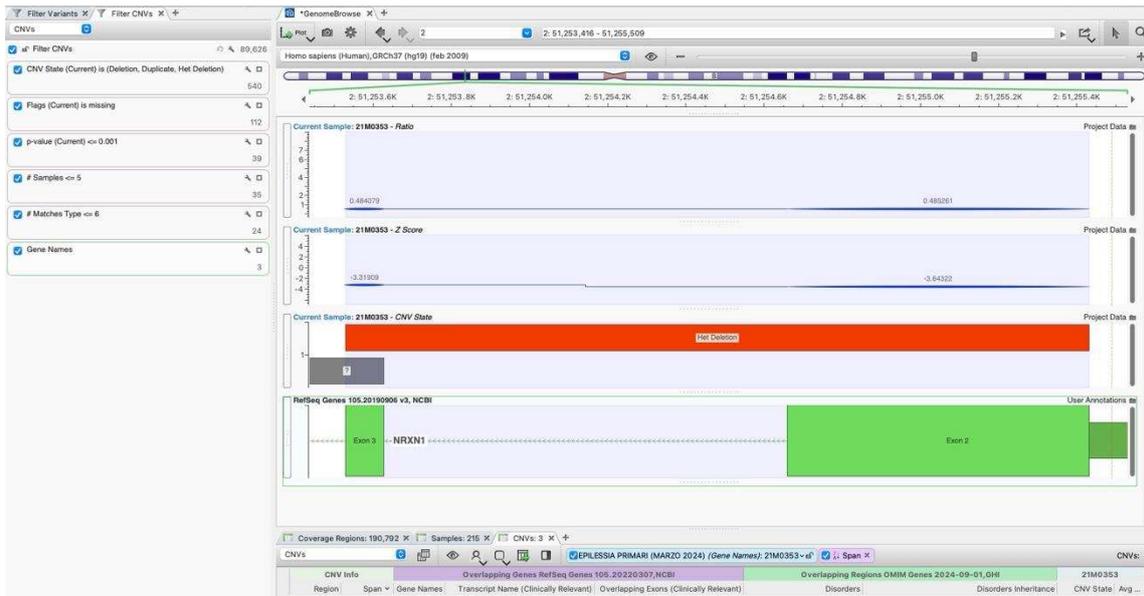


Figure 8. VarSeq home screen. On the left are the applied filters, while on the right is a graphical representation of the affected region (in this case, a deletion of *NRXN1* exons 2 -3 in patient 1), along with the ratio and z-score.

Table 2. Flags for Low Z Score, Insufficient Ratio, Deletion Contains Heterozygous Variants and Extreme GC Content.

Extreme GC Content	GC Content is below 0.30 or above 0.70
Deletion Contains Heterozygous Variants	Every exon of the deletion contains multiple heterozygous variants.
Low Z Score	Event has a low average z-score. Flag Thresholds: <ul style="list-style-type: none"> ● Hom². Deletion: Target Count ≤ 7 and z-score > -0.65 ● Het³. Deletion: <ul style="list-style-type: none"> - Target Count ⁴= 1 and Z-score > -2.5 - Target Count = 2 and Z-score > -1.4 - Target Count < 7 and Z-score > -1.2

² Homozygous

³ Heterozygous

⁴ Number of exons involved in the event

	<ul style="list-style-type: none"> ● Duplication: <ul style="list-style-type: none"> - Target Count = 1 and Z-score < 2.5 - Target Count = 2 and Z-score < 2.0 - Target Count < 7 and Z-score < 1.25
Insufficient Ratio	<p>Event has an average ratio that is inconsistent with the CNV state.</p> <p>Flag Thresholds:</p> <ul style="list-style-type: none"> ● Het. Deletion and Hom. Deletion: <ul style="list-style-type: none"> - Target Count \leq 2 and Ratio > 0.6 - Target Count < 8 and Ratio > 0.65 ● Duplication: <ul style="list-style-type: none"> - Target Count = 1 and Ratio < 1.42 - Target Count \leq 3 and Ratio < 1.35 - Target Count < 8 and Ratio < 1.28

Table 3. "Overlapping types" classification.

Contain region	The CNV fully contains the region
Within Region	The CNV is fully contained by the region
Identical	The CNV is identical to the region
Partial Overlap	The CNV partially overlaps the region

4. IN-DEPTH RESEARCH IN DATABASES AND LITERATURE

Genotype-phenotype correlation was performed to identify significant CNVs. Databases and scientific literature were consulted to analyse the genes involved and the types of CNVs they presented. The databases consulted were the following:

- OMIM (Online Mendelian Inheritance in Man) (<https://www.omim.org>);
- NSYDD (Neurodevelopmental Syndromes with Dysmorphic Features Database) (<https://sysnidd.dbmr.unibe.ch>);
- DECIPHER (<https://www.deciphergenomics.org>);
- Orphanet (<https://www.orpha.net/it>);
- GeneCards (<https://www.genecards.org>);
- HPO (Human Phenotype Ontology) (<https://hpo.jax.org>);
- HGMD (Human Gene Mutation Database) (<https://www.hgmd.cf.ac.uk/ac/index.php>).

Together with the literature search, it was possible to identify events associated with the most significant phenotypes linked to epilepsy. Particular attention was given to associations that included not only the gene, but also the specific exon(s) affected by the duplication or deletion.

The patients' medical records were reviewed to verify and potentially confirm the relationship between the observed clinical phenotype and the genetic variant identified by the analysis. Again, the use of databases and literature was essential to verify the presence of an effective genotype-phenotype correlation.

As they are based on probabilistic data, VarSeq calls with phenotypic matches in patients were confirmed by real-time PCR or array-CGH. These techniques made it possible to verify and confirm the presence of the CNV (duplication or deletion) in each sample and to delineate the ends of the event(s) with specificity.

5. LABORATORY VALIDATION METHODS

5.1 REAL TIME PCR

Real-time polymerase chain reaction (real-time PCR), also known as quantitative PCR, is a molecular biology technique in which target amplification is measured in real time as the reaction progresses [88]. The main advantage of real-time PCR is that it allows very accurate quantification of the initial DNA copy number, from the amount of amplicon generated after a series of PCR cycles [88]. Therefore, both quantitative and qualitative assays can be performed by using this technique [88]. The chosen protocol used in this thesis work employs SYBR Green, a dye that binds to double-stranded DNA (dsDNA) without being sequence-specific during the PCR process. To assess the specificity of the fluorescent signal obtained, an additional step was performed at the end of the real-time PCR: the melting curve analysis.

The copy number of a specific genomic region in a sample of interest, was determined using a relative quantification method which consists of comparing the test sample with a reference sample validated for the absence of CNVs in the region of interest. The results were expressed as the fold change (increase or decrease) in expression of the test sample compared to the reference one. For the actual quantification, the $\Delta\Delta C_T$ method was used.

In this work, real-time PCR was used to validate the presence of small-sized CNVs previously detected through VarSeq software. The genes involved in the CNV events were visualized on Integrative Genomics Viewer (IGV, <https://igv.org>), which allows to obtain the chromosomal coordinates of the start and the end of the exon(s) of interest. The coordinates were then copied on UCSC Genome Browser on Human using the genome version GRCh37/hg19 (<https://genome.ucsc.edu/index.html>) from which it was possible to obtain the nucleotide sequence. The sequence was then uploaded on Primer3Plus (<https://www.primer3plus.com/index.html>). The primer pairs were designed according to the following criteria:

- Localization: region of the sequence where primers pair;
- Primer minimum and optimal size of 20 nucleotides;
- Max Poly X: The maximum allowable length of a mononucleotide repeat is 3;
- Ideal Melting temperature (T_m) is 60°C. It is tolerated a maximum difference of 1°C between Forward and Reverse primers;
- GC content: maximum of 50-60%. Only a slight difference between Forward and Reverse primers is tolerated;
- Verify the presence of secondary structures (such as hairpins, self-dimers or cross-dimers). The T_m of the secondary structures must be below the annealing temperature, which is 60°C, to avoid amplification interference;
- The T_m of the primers must be below 95°C that is DNA denaturation temperature;
- The amplicon length can range between 60-100 bp or 140-180 bp. Housekeeping will be chosen between 80 bp or 160 bp respectively, depending on the length of the product.

Primers were designed to target the exon(s) involved in the CNV events. If the CNVs were confirmed, the primers were designed also for the flanking exon(s). This was done to identify the extremes of the event. The primers were designed with Primer3Plus (<https://www.primer3plus.com/index.html>), that associates each pair of primers with a penalty that varies according to their pairing efficiency. The chosen primers had a penalty below 1, whenever possible. The reverse and forward primers were then copied on the “In-silico PCR” section of the UCSC Genome Browser on Human (<https://genome.ucsc.edu/index.html>), showing

the pairing region for both primers. This additional step was used as confirmation of the correct primer pairing. The primers designed are reported in **Table 4**.

Table 4. Primers designed for real-time PCR.

Gene	Exon	Primer pair	Product size
<i>NRXN1</i>	1	F: AAGCATGCATCGGTCAAAGC R: TTGCCGGTGCTGTTACATGA	90 bp
	2	F: TGAACACCGTCATGTCCCTG R: CTTTCATCGACCAGGTGGAGG	68 bp
	3	F: TTCTGTTAGAGGCTTTGCTGT R: TTGCCGGTGCTGTTACATGA	72 bp
	4	F: AGAAAGCACCCACCTTCCAC R: AAGGTCCAGTCTGCATGCTG	100bp
<i>TMTC3</i>	4	F: ACAGGAGTTGTTGGAAGAGCA R: ACTTATGGAATTGTCTGGTCCTT	99 bp
	5	F: TGGACTCCAATTGCCTTGACA R: CACACAGCAAATTCCTACAAC	90bp
	7	F: AAGCCCAACTCCTACAAGGC R: GGTCCAATCACAGCAGAGCT	97bp
	10	F: TGTGGGTCATGCTCTGGAAA R: GCTGAACATGGGTAGCCTGT	77 bp
	11	F: TGGTGCCCATATGAATGTAGGA R: TGAGGCATCAGTGATTTAGCCA	96 bp
<i>SCN1A</i>	14	F: TGAAGGTCCACCAACCAAGG R: GTAGGTCATCCCGGATGCTG	98 bp
	15	F: ATTGCTCGTTGCCTTTGGGA R: AGGGAACAACCACTGAAACT	100 bp
	19	F: TGTGCATCCTATCCACAGCA R: AACCTTGCAGCCACTGATGA	67 bp
	20	F GCGTAAATTTGTTTGCTGGC R TCGATGTCAAACCTGTCACCA	71 bp
<i>FBNI</i> (HK)	25	F: ATTCCTCAGTACCCAGGCT R: GGAGTGCACCCTGCCTATTG	80 bp

Prior to real-time PCR, primers were resuspended in deionized (DI) water to obtain a final concentration of 50 ng/ μ L, followed by working aliquots of 5 ng/ μ L. The real-time PCR was performed as follows: DNA samples of the probands and the four references (two male and two female), were diluted in DI water with a final concentration of 1 ng/ μ L. The master mix was then prepared for each pair of primers, including those for the housekeeping, coding for *FBNI* gene, according to the following single-well volumes:

- 0,2 μ L Primer Forward (5 ng/ μ L);
- 0,2 μ L Primer Reverse (5 ng/ μ L);
- 3,6 μ L DI water;
- 5 μ L 2X AceQ SYBR® qPCR Master Mix.

The total volume of the master mix for each well was 9 μL , to which 1 μL of DNA sample (from the proband or from the references) was then added. Both proband and references were loaded in triplicate. For the blanks, 10 μL of the master mix of each target, was added to the plate. The amplification reaction has been performed with the Applied Biosystems QuantStudio™ 5 Dx Real-Time PCR System, using QuantStudio™ 5 Dx Software v1.0.2 for the analysis. The plate was placed in the thermal cycler and run with the program shown in **Table 5**.

Table 5. Real-time PCR thermal program.

Step	Processing Phase	Repetitions	Temperature	Time
Stage 1	Pre-denaturation ⁵	Reps: 1	95°C	5 min
	Denaturation		95°C	10 sec
Stage 2	Annealing + Extension	Reps: 40	60°C	30 sec
			95°C	15 sec
Stage 3	Melting curve	Reps:1	60°C	60 sec
			95°C	15 sec
Stage 4	Holding	Reps: 1	4°C	∞

At the end of the reaction, the amplification curves of the blanks (no amplification expected), proband and controls, were analysed. The melting curves of the targets and the housekeeping were then evaluated to assess the specificity of the amplification reaction. The C_T values were exported and copied into an excel sheet, and $\Delta\Delta C_T$ was calculated. To verify the actual presence of CNV, the values obtained for the samples were compared with the values of the reference given below:

- Wild type: 1;
- Duplication: 1,5;
- Heterozygous deletion: 0,5.

However, for genes located on chromosome X, it is important to consider the sex of the proband. For instance, a male proband affected by a duplication on the X chromosome will have a value of 1, because he has only one X chromosome, unlike a female who has two. Conversely, a deletion on the X chromosome in a male will result in a value of zero.

⁵ The AceQ SYBR® qPCR Master Mix contains a chemically modified hot-start DNA polymerase, that requires pre-denaturation at 95°C for at least 5 min

5.2 ARRAY COMPARATIVE GENOMIC HYBRIDIZATION

Array comparative genomic hybridization (array-CGH) is a high-resolution technique which allows genome-wide detection of unbalanced chromosomal abnormalities, such as gains (duplications) and losses (deletions) of genetic material leading to CNVs [89,90]. To achieve this, the patient (Proband) genomic DNA is labelled with a red fluorophore, mixed 1:1 with a green labelled control DNA (Reference) and the two co-hybridize to a microarray, a solid support spotted with thousands oligo probes (single-stranded DNA segments) representative of the human genome [90]. Proband and Reference samples compete for binding to microarray probes and the ratio between the red/green fluorescence intensities enables the detection of duplications and/or deletions.

In this thesis work, array-CGH analysis was conducted to confirm the presence of large CNVs previously detected using VarSeq software. Array-CGH experiments were performed using Agilent GenetiSure Cyto CGH 4×180K format, consisting of one slide which contains four arrays, each with 180.000 probes. The single slide allowed the simultaneous analysis of four patients samples, each coupled with a sex-matched reference gDNA. The array-CGH protocol is articulated in several steps and **Table 6** provides the materials and reagents utilized.

Table 6. Array-CGH materials and reagents kit [91].

Processing Step	Reagents Kit	Contents
Quantitative analysis of genomic DNA	Qubit 1X dsDNA Broad-Range Assay Kit	Working solution, DNA standards, assay tubes
Qualitative analysis of genomic DNA	Genomic DNA Reagents and Genomic DNA ScreenTape	Sample buffer, ladder, ScreenTapes
Sample Labelling	SureTag Complete DNA Labelling Kit	Labelling reagents and dyes, purification columns, male and female Reference DNA
Purification of labelled samples	TE Buffer 1X (pH 8.0)	TE Buffer 1X (pH 8.0)
Sample Hybridization	Oligo aCGH/ChIP-on-chip Hybridization Kit Human Cot-I DNA	Hybridization Buffer and Blocking Agent Human nonspecific hybridization blocking reagent

5.2.1 QUANTITATIVE AND QUALITATIVE ANALYSIS OF GENOMIC DNA (gDNA)

Quantitative analysis was performed using the Invitrogen™ Qubit™ 4 fluorometer with the dsDNA Broad-Range Assay Kit.

- Standards were prepared by mixing 190 µL of 1X dsDNA Working Solution with 10 µL of each of the two pre-warmed standard solutions with known concentrations of 0 ng/µL (blank) and 100 ng/µL [92], respectively.
- Samples were prepared by adding 1 µL of patients DNA to 199 µL of 1X dsDNA Working Solution.

Both standards and samples were then vortexed and incubated in the dark for 2-3 minutes, allowing the fluorescent dye in the Working Solution to bind to the DNA. The fluorescence emitted by the two standards was used to generate the calibration curve. The fluorescence intensities measured from the patient gDNA samples were then plotted against this curve to determine their respective concentrations.

To assess the purity and integrity of the gDNA, a qualitative analysis was performed to detect potential contaminants (proteins, carbohydrates and traces of organic solvents) and to determine the degree of gDNA fragmentation. DNA purity was evaluated by measuring the A260/A280 and A260/A230 absorbance ratios using the Thermo Scientific™ NanoDrop™ Lite Plus spectrophotometer during initial quantification of gDNA extracts. DNA integrity was instead assessed by capillary electrophoresis on the Agilent 4150 TapeStation System using Genomic DNA Reagents and Genomic DNA ScreenTape. The following protocol was used:

- According to the number of tubes needed (equal to the number of samples to be analysed plus one reserved for the Genomic DNA Ladder), a whole optical 8-tube strip or part of it was used;
- 10 µL of Genomic DNA Sample Buffer were dispensed into all the strip tubes. After that, 1 µL of Ladder was added in the first tube, while the remaining tubes were filled with 1 µL of DNA from each sample;
- The strip was closed with the cap, centrifuged at 2000 rpm for one minute and then loaded into the instrument. A pre-warmed ScreenTape (a device composed of 16 channels where the electrophoretic separation of DNA fragments takes place) was inserted in the appropriate nest;
- The TapeStation Controller software was started to run the analysis. After 10-15 minutes, the TapeStation Analysis software displayed the results in the form of electropherograms with peaks corresponding to different sized groups of fragments. In addition, to each sample was assigned a DNA Integrity Number (DIN) on a scale from 1 to 10, indicating the level of DNA integrity. The higher the DIN, the lower the degree of DNA fragmentation. DINs of at least 6 were considered acceptable.

Once the quality of the gDNA samples had been assessed, 500 ng from each of them (both Proband and References) were taken and diluted as follows:

$$\frac{500 \text{ ng}}{C_s \left(\frac{\text{ng}}{\mu\text{L}}\right)} + V_{H_2O} (\mu\text{L}) = 26 \mu\text{L}$$

where C_s is the sample concentration previously calculated with Qubit and V_{H_2O} is the volume of nuclease-free water to be added to the sample to reach a final volume of 26 µL.

5.2.2 gDNA HEAT FRAGMENTATION AND DENATURATION

gDNA heat fragmentation (98°C, fragment size: 1-3 kb) and denaturation were achieved by adding 5 µL of Random Primers to each sample, for a total volume of 31 µL. The tubes were then loaded into the Applied Biosystems™ MiniAmp™ Thermal cycler and the program shown in **Table 7** was run.

Table 7. Array-CGH fragmentation and denaturation thermal program.

Step	Processing Phase	Repetitions	Temperature	Time
Stage 1	Incubation	1	98°C	10 min
Stage 2	Incubation	1	4°C	3 min
Stage 3	Holding	1	4°C	∞

5.2.3 FLUORESCENT LABELLING WITH CYANINE

Denatured gDNA fragments are used as templates for primer extension reaction. Exo Klenow (a fragment of DNA Polymerase I which exhibits 5'→3' polymerase activity but lacks exonuclease activity) incorporates cyanine-labelled dUTPs, giving a fluorescent labelled DNA product. Specifically, two cyanine were used to achieve differential labelling:

- Cy5, used for Proband labelling. It emits red light in the excited state (emission peak of 670 nm);
- Cy3, used for Reference labelling. it emits green light in the excited state (emission peak of 570 nm).

The labelling mixtures were prepared as reported in **Table 8** (values refer to four experiments on a single array-CGH slide). Next, 19 µL of Cy5 mix was added to Proband samples and 19 µL of Cy3 mix to Reference samples, bringing the final volume of each tube to 50 µL. Finally, the tubes were incubated in a thermal cycler with the profile shown in **Table 9**.

Table 8. Labelling mixtures with cyanine 5 and 3.

Reagents	Volume
5X gDNA Buffer	42 µL
10X dNTPs	21 µL
dUTP with cyanine (5 or 3) 1 mM in TE Buffer (pH 7.5)	12,6 µL
Exo Klenow (5 U/µL)	4,2 µL

Table 9. Thermal program for the labelling of DNA fragments with cyanine.

Step	Processing Phase	Repetitions	Temperature	Time
Stage 1	Incubation	1	37°C	120 min
Stage 2	Incubation	1	65°C	10 min
Stage 3	Holding	1	4°C	∞

5.2.4 LABELLED gDNA PURIFICATION

For the purification of cyanine-labelled gDNA, a clean-up column was placed into a 2 mL collection tube for each sample; then, the full volume of the samples was transferred into the respective columns. The procedure consisted of two washing steps:

- An initial wash was performed by pipetting 430 μL of 1X TE buffer (pH 8.0) to each column. The columns were centrifuged at 14000 rpm for 10 minutes and the flow-through was discarded;
- For the second wash, 480 μL of 1X TE buffer (pH 8.0) were added and the columns were centrifuged again (14000 rpm for 10 minutes).

Once finished, the columns were inverted into a new set of tubes and centrifuged at 1000 rpm for one minute. The liquid collected at the bottom of the tubes contains the purified labelled gDNA.

To achieve a uniform volume in all samples, the tubes were placed in an Eppendorf 5301 Concentrator and centrifuged for 23 minutes at 250 rpm and 45°C to allow the liquid part to evaporate. Each pellet was then resuspended in 21 μL of 1X TE buffer (pH 8.0). The concentration of the purified samples and the molarity of the cyanine (in terms of pmol of cyanine per μL of sample) were determined with the Nanodrop. These values were necessary for the calculation of:

- The sample yield (μg):

$$\frac{C_s \left(\frac{\text{ng}}{\mu\text{L}} \right) \cdot V_s (\mu\text{L})}{100}$$

where C_s is the sample concentration and V_s is the sample volume. Optimal results should range 8-13 μg ;

- The specific activity (pmol/ μg):

$$\frac{M_c \left(\frac{\text{pmol}}{\mu\text{L}} \right)}{C_s \left(\frac{\text{ng}}{\mu\text{L}} \right)} \cdot 1000$$

where M_c is the molarity of the cyanine and C_s is the sample concentration. Optimal specific activity should range 20-60 pmol/ μg .

The Probands and References were then sorted according to decreasing values of yield and specific activity. Each Proband was then paired with the Reference (of the same sex) occupying the same position in the previously determined order, giving priority to specific activity in case of discordance between the two parameters. Finally, each Proband-Reference pair was mixed in a single tube.

5.2.5 HYBRIDIZATION

Proband-Reference mixtures were subjected to a preliminary treatment to increase fluorescent signal intensity, reduce at minimum background noise and optimize the kinetics of the subsequent hybridization

on the microarray slide. The Hybridization Master Mix was prepared as follows (volumes refer to a single array-CGH slide):

- 21 μL of Cot-I DNA (1 mg/ μL), useful in preventing hybridization of centromeric and telomeric repeat sequences to microarray probes;
- 46,2 μL of 10X Blocking Agent, which reduces non-specific hybridization;
- 231 μL of 2X HI-RPM Hybridization Buffer. With its content of lithium salts and ethers, it further increases the stringency of hybridization by preventing non-specific binding.

A volume of 71 μL of Hybridization Master Mix was added to each Proband-Reference mixture. The tubes were then loaded into the thermocycler and run with the program shown in **Table 10**. At the end of the incubation, the hybridization chamber was assembled in the following steps:

- The backing slide was placed on the hybridization chamber base. The slide surface displays four rectangular wells (as many as the array-CGH experiments) delimited by silicon gaskets;
- 100 μL of each Proband-Reference mixture were dispensed onto the relevant gasket slide wells;
- The microarray slide was then placed on top in a face-down position, ensuring the probes were in contact with the mixtures;
- The chamber cover was put on the slide sandwich and the clamp was firmly tightened to close the assembly.

The assembled hybridization chamber was finally loaded into the oven and incubated with rotation at 67°C for 24 hours. Additionally, a slide-staining dish and two flasks containing Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 2 were placed in the incubator at 37°C for 24 hours before use (see Section 5.2.6, “Microarray Wash”).

Table 10. Array-CGH hybridization thermal program.

Step	Processing Phase	Repetitions	Temperature	Time
Step 1	Incubation	1	98°C	3 min
Step 2	Incubation	1	37°C	30 min
Step 3	Holding	1	37°C	∞

5.2.6 MICROARRAY WASH

Short before the end of the hybridization, two slide-staining dishes were prepared as follows:

- Dish #1 was filled up to about three-quarters of its volume with Agilent Oligo aCGH/ChIP-on-Chip Wash Buffer 1 at room temperature;
- Dish #2 was filled with the same buffer and a magnetic stir bar was added. This dish was then placed on a magnetic stir plate.

Once the 24-hour hybridization was over, the hybridization chamber was removed from the oven and disassembled on a clean surface; the slide sandwich was submerged into Dish #1, separating the microarray slide from the gasket slide. The slide containing the probes then underwent two washes:

- First wash: the microarray slide was placed in a designated rack and transferred to Dish #2. Here it was incubated for 5 minutes at 350 rpm stirring speed; in the last minute of the first wash, the dish and the two Buffer 2 flasks previously incubated at 37°C (see Section 5.2.5, “Hybridization”) were removed from the incubator. The dish was placed on a magnetic stir plate heated at 37°C and filled with Buffer 2. A magnetic stir bar was finally added;
- Second wash: the rack was then transferred to the Buffer 2-filled dish and incubated for 1 minute with magnetic stirring.

At the end of the second wash, the slide was removed from the rack and the excess liquid was wiped off the edges. Finally, the microarray slide was placed in the slide holder for scanning with the Agilent SureScan Dx Microarray Scanner System.

5.2.7 MICROARRAY SCANNING AND ANALYSIS

The first step was to start the scanner (SureScan Dx) and the image acquisition software (Agilent Microarray Scan Control) on the PC. The assembled slide holder was then put into the scanner cassette, starting thereafter the scanning process of the four arrays.

The acquired images, with their characteristic arrangement of green/yellow/red-coloured spots, were examined using the Agilent Feature Extraction software. Here we could check for major issues, such as unbalanced labelling, insufficient volume of Sample/Reference mixtures, inappropriate washing and physical damages to the slide surface.

Finally, the Agilent CytoGenomics software was started to convert the data extracted from images into CNV calls. For each CNV in the list are shown its:

- Type (loss or gain);
- Start and stop coordinates;
- Size;
- Number of probes included in the call;
- Associated \log_2 of Proband/Reference fluorescent signals ratio (in short, \log_2 ratio).

The reference values for the interpretation of the results are given in **Table 11**.

Table 11. Reference values for the analysis of the array-CGH results.

Description	Average Proband CN ⁶	Reference CN	Ratio (P/R) ⁷	Ideal \log_2 ratio	Actual \log_2 ratio
Homozygous deletion	0	2	0	$-\infty$	- 4
Heterozygous deletion	1	2	0,5	-1	-0,9
Diploid	2	2	1	0	0
Duplication (1 copy gain)	3	2	1,5	+0,58	+0,53
Amplification (2 copies gain)	4	2	2	+1	+0,87

⁶ Copy Number

⁷ Proband/Reference

5.3 SANGER SEQUENCING

Sanger sequencing is a method of DNA sequencing based on chain termination [93]. It allows the nucleotide bases of a sequence to be determined with 99.99% accuracy [94]. Sanger is defined as a "chain termination" method because it uses four dideoxynucleoside triphosphates (ddNTPs: ddATP, ddGTP, ddCTP and ddTTP) in addition to the four deoxynucleoside triphosphates (dNTPs: A, G, C and T) for normal sequence extension [94]. During the sequencing reactions, ddNTPs compete with dNTPs to insert a chain-terminating dideoxynucleotide. ddNTPs, according to the type of base, are also linked to different fluorophores which, once synthesis is blocked, allows the determination of the base at the end of the nucleotide chain [94]. At the end of the sequencing step, the dsDNA fragments, which differ incrementally by one nucleotide from their common 5' end, are separated by size by capillary electrophoresis [94]. As the DNA fragments migrate, they pass a laser that excites the fluorophores causing them to emit fluorescence at a specific wavelength. The fluorescence signals are recorded, and an intensity profile is given for the differently coloured fluorophores.

5.3.1 POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION

Polymerase chain reaction (PCR) was performed prior to the sequencing phase, to obtain a higher number of DNA copies, which is essential to increase the efficiency of the Sanger protocol [95].

The primers were designed to target *RNU4-2*, a small nuclear RNA (**Table 12**).

Table 12. *RNU4-2* primer design.

Gene	Primer pair	Product size
<i>RNU4-2</i>	F: ACCCAGTTCCCAACATAGTGT R: AGCCGTGTTCTCCATTGAGA	770bp

For each sample in the cohort, an aliquot was prepared with the following components:

- Sample DNA;
- 0,5 µL Primer Forward (50 ng/µL);
- 0,5 µL of Primer Reverse (50 ng/µL).

The volume of DNA to be added was calculated according to the concentration of each sample. The analysed samples had concentrations ranging from 45 to 65 ng/µL; therefore, the volumes collected were as follows:

- 2.2 µL for samples with concentration of 45 ng/µL;
- 2 µL for samples with concentration of 50 ng/µL;
- 1.81 µL for samples with concentration of 55 ng/µL;
- 1.6 µL for samples with concentration of 60 ng/µL;
- 1.54 µL for samples with concentration of 65 ng/µL.

The master mix was then prepared according to the quantities given in **Table 13** and aliquoted into each tube. Prior to the PCR run, each tube was spinned and vortexed then loaded into the Applied Biosystems™ MiniAmp™ Thermal Cycler and run with the shown in **Table 14**.

Table 13. Master mix components and respective volumes for PCR amplification.

Reagents	Volume (x 1 tube)
Buffer 10x w/o Mg	2,5 μ L
MgCl ₂ 25 mM	2 μ L
dNTPs mix 20 mM	0.75 μ L
HOT FIREPol® (5 U/ μ L)	0.3 μ L
Deionised H ₂ O	15.2 μ L
DMSO+form (1:1)	1,25 μ L
Total	22 μL

Table 14. PCR Thermal Cycler program.

Step	Processing Phase	Repetitions	Temperature	Time
Stage 1	Pre-denaturation ⁸	Reps: 1	95°C	15 min
	Denaturation		95°C	30 sec
Stage 2	Annealing	Reps: 35	62°C	30 sec
	Elongation		72°C	30 sec
Stage 3	Final elongation	Reps: 1	72°C	7 min
Stage 4	Holding	Reps: 1	4°C	∞

5.3.2 AGAROSE GEL ELECTROPHORESIS

Agarose gel electrophoresis is a technique that allows the separation of DNA fragments ranging in size from 100 bp to 25 kb [96]. The gel acts as a filter, consisting of a network of pores that allow molecules to be separated according to their size. The rate of migration of the DNA fragments is inversely proportional to their size and it is influenced by the type and concentration of agarose gel, DNA conformation and the applied voltage [96].

In this work, gel electrophoresis was performed to assess the effectiveness of the PCR amplification of the samples. The first step consisted of preparing the buffer for the electrophoresis run. The PanReacAppliChem © ITW Reagents Tris-Borate-EDTA (TBE) Buffer 10X was diluted to give a final concentration of 1X.

The gel was prepared as follows:

⁸ The HOT FIREPol® (5 U/ μ L) is a hot-start DNA polymerase which requires incubation step at 95°C for 12-15 minutes at the beginning of the PCR cycle.

- 1) 1 g of agarose was weighed into an Erlenmeyer flask and 1X TBE buffer was added to obtain a final volume of 100 mL;
- 2) The agarose/buffer mixture was melted in the microwave. When the mixture boiled, the flask was removed and swirled to mix the contents. The process was repeated every 20-30 seconds until the mixture was transparent;
- 3) The flask was then cooled to a temperature of approximately 55°C;
- 4) 10 µL of BioSigma™ Midori green Advance DNA Stain (dye that binds to DNA with a high degree of sensitivity and emits fluorescence when excited by UV transilluminators) was added to the agarose/buffer mixture and the flask was swirled to obtain a homogeneous solution;
- 5) Two combs were placed into the gel mold to create the wells. The melted agarose was placed in the gel mold and allowed to set at room temperature;
- 6) Once the gel had set, the combs were removed, and the gel was transferred to the electrophoresis running bath (the buffer hat to cover the surface of the gel).

Sample loading was performed as follows: for each sample, a mix was made by combining 3 µL of PCR amplification product with 1.9 µL of loading buffer 6X (it indicates the migration of the DNA and allows the sample to sink into the gel wells) and transferred into the gel wells. The loading buffer was produced in the laboratory as follows:

- 1) 25 ml of glycerol (50% v/v in H₂O) was added into a flask;
- 2) 25 ml of H₂O was added to the glycerol and the solution was manually mixed to equalize;
- 3) 250 mg of bromophenol blue (0.25%) was then added to the flask. The final solution was mixed by magnetic stirring using IKA® Plate (RCT digital);
- 4) The obtained buffer was aliquoted and then stored at -20°C.

During the loading phase, an empty well was left between each proband, to which 2.5 µL of Solis BioDyne™ 100 bp DNA Ladder Ready to Load was added.

The gel box was closed with the lid, on which the electrodes for the generation of the current were plugged: the anode (black lead) and the cathode (red lead). The myVolt™ Touch power supply was set to 120V. The electrophoresis run was stopped when the samples had migrated a sufficient distance. The DNA bands on the gel were examined by UV transillumination.

5.3.3 PCR CLEAN-UP

The PCR clean-up was achieved by enzymatic purification. The aim was to remove any excess reagents, such as dNTPs and primers. The enzyme used was Applied Biosystems™ ExoSAP-IT™ PCR Product Clean-up Reagent. This enzyme hydrolyses excess primers and dNTPs, thus cleaning the PCR product.

For each sample, a test tube containing a total reaction volume of 7 µL was prepared as follows:

- 5 µL of PCR reaction product;
- 2 µL of Applied Biosystems™ ExoSAP-IT™ PCR Product Clean-up Reagent.

The tubes were spinned and then loaded into the Applied Biosystems™ MiniAmp™ thermal cycler and run with the program reported in **Table 15**.

Table 15. Thermal cycler PCR clean-up program.

Step	Processing Phase	Repetitions	Temperature	Time
Stage 1	Purification	Reps: 1	37°C	15 min
Stage 2	Enzyme inactivation	Reps: 1	80°C	15 min
Stage 3	Holding	Reps: 1	4°C	∞

5.3.4 CYCLE SEQUENCING

Once purified, the target DNA sequences are amplified by chain termination PCR. In this thesis work, the Sanger sequencing automatic method, in which the four ddNTPs are added in the same reaction mixture, was employed.

A 96-well plate was used for cycle sequencing. The master mix for chain termination PCR was prepared as shown in **Table 16** and then aliquoted into the 96-well plate. The components of the final mixture for each well are detailed in **Table 17**.

Table 16. Chain termination PCR master mix components and respective volumes.

Reagents	Volume (x 1 tube)
Deionised H ₂ O	5 µL
BrilliantDye™ Terminator v3.1 premix	1 µL
BrilliantDye™ Terminator 5X Sequencing Buffer	2 µL
Total	8 µL

Table 17. Chain termination PCR components and relative volumes.

Reagents	Volume (x 1 tube)
Master mix	8 µL
Post-PCR clean-up product	1 µL
Primer ⁹ (5 ng/µL)	1 µL
Total	10 µL

The wells were sealed with caps and any bubbles that formed during the loading step were manually removed. The plate was then loaded into the Applied Biosystems™ MiniAmp™ thermal cycler and run with the following program (**Table 18**):

⁹ For each well only one between forward and reverse primer was added.

Table 18. Thermal cycler chain termination PCR program.

Step	Processing Phase	Repetitions	Temperature	Time
Stage 1	Pre-denaturation	Reps: 1	96°C	3 min
	Denaturation		96°C	10 sec
Stage 2	Annealing	Reps: 25	55°C	5 sec
	Elongation		62°C	4 min
Stage 3	Holding	Reps: 1	4°C	∞

5.3.5 SEQUENCING CLEAN-UP

The clean-up of the sequencing reaction was achieved by enzymatic purification. This step was performed to remove unincorporated ddNTPs, dNTPs, primers, salt ions and enzymes. The sequencing clean-up allows to remove any contaminants and increase the quality of the sequencing reading.

The BigDye® XTerminator™ Purification Kit is composed by:

- 1) XTerminator Solution – removes unincorporated ddNTPs terminators and salts from the post cycle-sequencing reaction [97];
- 2) SAM™ Solution – Improves XTerminator solution performance and stabilizes the sample after purification [97].

The master mix containing BigDye® XTerminator™ Purification Solutions was prepared using the following volumes per single well:

- 45 µL of XTerminator Solution;
- 10 µL of SAM™ Solution.

The tube containing the BigDye® XTerminator™ Purification Master Mix was vortexed and 55 µL was added to each well of a 96-well plate, then 10 µL of chain-termination PCR reaction products were transferred to the new wells containing the master mix using a multichannel pipette. Empty wells were filled with 40 µL of Applied Biosystems™ Hi-Di™ Formamide. The wells were closed with caps and the plate was placed on an IKA® MS 3 digital shaker at 1800 rpm for 30 minutes. Vortexing of BigDye XTerminator reagents allows immobilization of unincorporated and excess components [97]. The plate was then centrifuged at 1000 rpm for 1 minute to displace the immobilized components at the bottom of the well. The dye-labelled extension products were instead transferred to the supernatant during centrifugation [97].

5.3.6 CAPILLARY ELECTROPHORESIS

After centrifugation, the cap strips were removed from the wells and the plate was prepared for the loading into the Applied Biosystems™ SeqStudio™ 8 Flex Genetic Analyzer. The plate assembly required three steps:

- A septum membrane was placed over the wells;
- The plate was then transferred onto a suitable base plate holder;
- The SeqStudio™ Flex series plate retainer (lid) was placed on the top of the plate and secured to the bottom base.

The assembled plate was loaded into the instrument. The plate file model was created using the SeqStudio™ Plate Manager 2.0 software and subsequently imported into the instrument. This software, supplied by Thermo Fisher, allowed the selection of injection groups for which the application mode (sequencing), dye set and run module was then set. Each well of the plate was then given a name to facilitate the subsequent reading of the results. Once the plate had been loaded and the appropriate paddle file had been selected, the capillary electrophoresis (CE) was carried out by SeqStudio™ 8 Flex Genetic Analyzer as follows:

- 1) The DNA sample is injected into fine capillary, filled with polymer matrix for the electrophoretic run, at low voltage [98];
- 2) The voltage increased to a constant level causing the DNA fragments to migrate from the negative to the positive pole [98]. As in agarose gel electrophoresis, smaller fragments migrate faster than larger ones, reaching the detector first [98];
- 3) In the detection cell, the dyes attached to the DNA fragments are excited by a narrow beam of laser light. A small amount of laser light is absorbed by the dyes and each of the four dyes emitted a different wavelength light that determined the 4 bases [98];
- 4) The emitted light captured by a CCD camera and the registered fluorescence is translated into multidye-data [98], that in this work represent the sequence information.

5.3.7 DATA ANALYSIS

Applied Biosystems™ Sequencing Analysis Software v7 was used to analyse the obtained data. This software generates an electropherogram where each peak corresponds to a base (A, T, C or G) which is identified by its associated fluorescence (**Figure 9**). By applying specific parameters, the sequence can be edited to remove peaks at the beginning or at the end of the sequence that may be imprecise or difficult to evaluate.

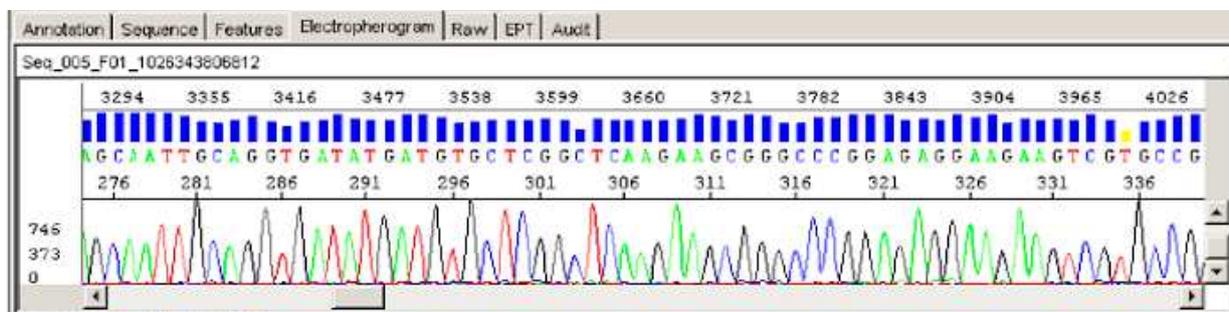


Figure 9. Sanger sequencing electropherogram. The sequence at the top shows the nucleotides corresponding to each fluorescence peak. The yellow square highlights a low-quality signal, while the blue square indicates a high-quality signal. From Applied Biosystems Sequencing Analysis Software Quick Reference Card (<https://assets.thermofisher.com>).

SNVs in heterozygosis are marked by a red square at their respective positions. The system detects SNVs by identifying two overlapping peaks at the same specific position during analysis. However, because the system does not rely on a reference sequence, but it analyses the signals independently at each position, homozygous variations are not detected.

In contrast, insertions are indicated by a shift in the electropherogram. Following the insertion, the subsequent peaks fail to align correctly with the expected reference sequence.

Each individual sequence was analysed. Using the Integrative Genomics Viewer (IGV, <https://igv.org>) with the GRCh38/hg38 genome reference, the sequences for the *RNU4-2* were displayed. This approach facilitated a comparison between the reference sequence and the Sanger-obtained sequence to identify any homozygous variants and visualize those reported by the software.

With the use of *IGV* we determine the precise genomic position of the variation or insertion. This approach allowed the identified variant to be compared with those in genomic databases, which classify mutations into the following categories:

- Pathogenic;
- Likely pathogenic;
- Conflicting pathogenicity classifications;
- Probably benign;
- Benign;
- Uncertain significance.

The databases consulted included: gnomAD browser (<https://gnomad.broadinstitute.org>), ClinVar - Clinical Genome Resource (<https://www.clinicalgenome.org>), Varsome (<https://varsome.com>), HGMD (Human Gene Mutation Database) (<https://www.hgmd.cf.ac.uk/ac/index.php>) and dbSNP (Single Nucleotide Polymorphism Database) (<https://www.ncbi.nlm.nih.gov>).

RESULTS AND DISCUSSION

NDDs are a group of early-onset conditions that affect brain development and function, leading to deficits in areas such as cognition, communication and motor skills [1,2]. These disorders affect approximately 15% of children and adolescents worldwide [13,14]. Notably, 15-26% of individuals with NDDs also have epilepsy, compared to a 0.8% prevalence in the general population, highlighting epilepsy as a common feature of these conditions [1]. Epilepsy is a chronic neurological disorder characterized by recurrent seizures, that affects over 70 million people worldwide [19,20].

A defining characteristic of both NDDs and epilepsy, is their high genetic heterogeneity [11,43]. To date, around 6,000 candidate risk genes have been proposed to play a role in NDDs, of which 364 are associated with epilepsy [11,43], representing a significant challenge in establishing an accurate molecular diagnosis for affected patients. As a result, focusing only on SNVs analysis is often insufficient when investigating the molecular basis of NDDs, including epilepsy.

The current thesis work, carried out at the R&I Genetics laboratory in Padua, focused on the genetic analysis of NDDs patients, who had undergone prior investigations but had not yet received a definitive molecular diagnosis. Specifically, DNA samples had already been sequenced through WES and analysed for the presence of SNVs in NDDs known genes, though no positive results were obtained. In this context, a secondary analysis was undertaken to explore the possible presence of CNVs, that in recent years have emerged as major contributors to neurocognitive disorders, further highlighting their importance in the genetic aetiology of this group of conditions. [80,81]. In addition, given the recent discovery of rare variants in *RNU4-2* in patients with NDDs, a secondary analysis was performed to investigate the presence of variants in this snRNA. *RNU4-2* belongs to the class of small sncRNAs, that play key roles in chromatin regulation, modulation of gene transcription and control of splicing [62]. In particular, *RNU4-2* is an essential component of the major spliceosome [61], which is responsible for the splicing of approximately 99% of introns, and disruptions in this pathway have been implicated in several human diseases [99]. Notably, variants in this snRNA have been shown to cause specific abnormalities in the 5' splice site (5'SS), a critical region for the first phase of splicing [100].

The chosen approach for sample analysis combined multiple methods, including WES, real-time PCR, array-CGH, and Sanger sequencing, with the goal of enhancing the molecular diagnosis of NDDs, including epilepsy.

The analysis focused on investigating the presence of CNVs in 100 exomes from epilepsy patients (see section 1, "Cohort of Patients", Materials and Methods). These patients were chosen because their samples had already been sequenced and analysed for the presence of SNVs. However, this analysis did not yield any positive results, resulting in a lack of a molecular diagnosis.

The exomes of the 100 patients were analysed using VarSeq bioinformatic software, which detects CNVs through a probabilistic approach based on NGS coverage data, comparing test samples with reference samples. CNVs were selected based on the metrics provided by VarSeq, according to the reference values shown in **Table 2**. Specifically, deletions with a ratio of 0.5 (for heterozygous deletions) or a ratio of 0 (for homozygous deletions) and Z-score values ≤ -3 were selected. While duplications with a ratio of 1.5 (for heterozygous duplications) or a ratio of 2 (for four copies) and Z-score values ≥ 3 were selected. From this selection, thirteen candidate patients were identified, for whom a genotype-phenotype correlation was conducted. Specifically, the clinical phenotype reported by physicians was compared with the variant-associated phenotype documented in the scientific literature. Based on this analysis, nine patients were excluded due to lack of agreement between their clinical presentation and the variant-associated phenotype.

This process ultimately led to the identification four patients with significant CNVs. The clinical phenotypes reported by physicians for the patients carrying the CNVs are shown in **Table 19**.

Table 19. Age, Sex and clinical phenotypes of patients with significant CNVs.

Patient ¹⁰	Age ¹¹	Sex	Clinical phenotype(s)
1	19	F	Generalized epilepsy, drug resistant seizures.
2	20	F	Epileptic encephalopathy, dysmorphism, intellectual disability and severe pshyco-motor delay.
3	18	F	Normal psycho-motor development. Myoclonic epileptic seizures, sometimes with short 'absences'. Recently shows an increase in nocturnal seizures.
4	1	F	Motor and speech delay with febrile seizures poorly responsive to pharmacological treatment.

The five CNVs identified in four patients through bioinformatics analysis (main features shown in **Table 20** and **Table 21**), were subsequently validated in the laboratory using real-time PCR or array-CGH methodologies.

Table 20. Main features of the small CNVs identified by VarSeq.

Patient	Gene	Exon(s)	Span (bp)	CNV state	Avg Z score	Avg Ratio	p-value
1	<i>NRXN1</i>	2-3	1903	Het. ¹² deletion	-3,48	0,48	9,07E-06
3	<i>TMTC3</i>	5-10	16206	Het. deletion	-3,00	0,50	3,33E-11
4	<i>SCN1A</i>	15-19	6377	Duplication	3,23	1,29	1,97E-13

Table 21. Main features of the large CNVs identified by VarSeq associated to each patient

Patient	Genomic segment	Span	CNV state	Avg Z score	Avg Ratio	p-value
1	16p13.11 (chr16 15,489,755-16,481,355del)	991.6 kb	Het deletion	-3,69	0,54	1,66E-112
2	10p15.3 (chr10:1,087,127-1,175,284dup)	8.9 kb	Duplication	5,04	1,80	2,23E-09

¹⁰ For simplification, the patients were numbered

¹¹ Refers to the patient's actual age at the time of blood collection

¹² Heterozygous

Real-time PCR confirmed smaller CNVs (results reported in **Table 22**), while larger events, involving more than one gene, were validated by array-CGH (results presented in **Table 23**). Notably, three CNVs were found to be larger than initially reported by VarSeq, likely due to the probabilistic approach of the software used for variant detection. Consequently, to precisely define the boundaries of the CNVs confirmed through real-time PCR, additional primers were designed for the exons adjacent to the event highlighted by the software. The ratio obtained for each analysed exon are provided in the Appendix in **Table I**, **Table II** and **Table III**. Among the four patients with positive findings, three had heterozygous deletions and two had duplications.

In addition, the availability of parental samples enabled the analysis of variant segregation.

Table 22. Real-time PCR results.

Patient	Gene	Exon	Obtained Ratio	Interpretation	Segregation
1	<i>NRXN1</i>	1	0,47	Heterozygous Deletion	Maternal inheritance
		2	0,43	Heterozygous Deletion	
		3	0,45	Heterozygous Deletion	
		4	1,03	Diploid	
3	<i>TMTC3</i>	4	1,00	Diploid	Parents not available
		5	0,38	Heterozygous Deletion	
		7	0,51	Heterozygous Deletion	
		10	0,44	Heterozygous Deletion	
		11	1,06	Diploid	
4	<i>SCN1A</i>	14	1,07	Diploid	<i>De novo</i>
		15	1,40	Heterozygous duplication	
		19	1,51	Heterozygous duplication	
		20	1,03	Diploid	

Table 23. Array-CGH results.

Patient	Genomic segment	Size	<i>MeanLogRatio</i> ¹³	Segregation
1	2p16.3 (51,225,530-51,343,363)	117.83 kb	-0,802	Maternal inheritance
	16p13.11 (15,053,284-16,289,746)	1.4 Mb	-0,928	Non-informative ¹⁴
2	10p15.3 (1,088,683-1,199,411)	110.73 kb	0,833	Paternal inheritance

The four positive patients will now be individually analysed and discussed in detail.

¹³ Reference values on Table 11

¹⁴ In the absence of paternal sample, variant segregation cannot be established with certainty

Patient 1

Patient 1 is a 19-year-old woman diagnosed with drug-resistant generalized epilepsy, a form of epilepsy characterized by seizures affecting both sides of the brain. According to the ILAE classification of epilepsies, generalized epilepsy of genetic origin is classified as idiopathic generalized epilepsy (IGE) [99]. Patients with IGE experience normal cognitive development, and brain magnetic resonance imaging (MRI) typically appears normal [100]. Recent studies have shown that approximately 10-15% of patients with IGE experience drug-resistant seizures, although the mechanisms responsible for these seizures remain undefined [101].

Bioinformatic analysis of patient 1 using VarSeq software identified two CNVs. The first is a heterozygous deletion of 1903 bp affecting exons 2 and 3 of the *NRXN1* gene. The second is a larger heterozygous deletion of 991.6 kb affecting the 16p13.11 genomic region (chr16:15,489,755-16,481,355del). The CNV involving *NRXN1* was validated by real-time PCR, which revealed that the deletion also included exon 1. In VarSeq, the first exon is often not detected or, conversely, many false positives are reported. This happens because the sequences at the beginning of a gene can be more challenging to analyse accurately, especially if there are low quality sequences (in this case, the DNA samples were extracted several years ago), repetitive regions or poorly defined areas in the reference genome.

The larger deletion was confirmed by array-CGH, showing a 1.4 Mb deletion encompassing the 16p13.11 cytoband and involving the following genes: *PDXDC1* (OMIM: 614244), *NTANI* (OMIM: 615367), *MPV17L* (OMIM: 618100), *MARF1* (OMIM: 614593), *FOPNL* (OMIM: 617149), *ABCC1* (OMIM: 158343), *ABCC6* (OMIM: 603234), *NDE1* (OMIM: 609449), and *MYH11* (OMIM: 160745) (**Figure 10, A**). The reason why both events were not initially validated by array-CGH is that the deletion involving *NRXN1*, identified by VarSeq, affected only two exons, making it too small to be detected by the probes. However, when array-CGH was performed to confirm the larger deletion, it also detected the smaller CNV. Specifically, the deletion at cytoband 2p16.3 affecting *NRXN1* was found to be 117.83 kb (**Figure 10, B**), larger than the size predicted by VarSeq. Indeed, the small deletion encompassed the first three exons of *NRXN1*, and part of a non-coding region associated with the *LOC730100* gene, which encodes a lncRNA with an unclear function. This region had not been previously reported by VarSeq, as the software analyses data from WES and therefore detects only variants in exonic regions. The inclusion of *LOC730100* increased the size of the small CNV, allowing the CGH array to detect it.

Additionally, the proband's mother sample was analysed using real-time PCR and array-CGH (for the reason explained above), revealing the presence of the same heterozygous deletion of exons 1, 2, and 3 of the *NRXN1* gene and confirming the maternal inheritance of this mutation (**Figure 11**). The large CNV involving the 16p13.11 locus was not detected in the mother, but the absence of the paternal sample led to the variant being classified as non-informative regarding its segregation.

The *NRXN1* gene encodes neurexin-1, a member of the family of transmembrane proteins known as neurexins [102,103]. These proteins are involved in the correct development and function of neuronal synapses by acting as presynaptic cell adhesion molecules [102,103]. Neurexin-1 produces two distinct mRNAs from two promoters: the longer mRNA encodes alpha-neurexins, which are involved in synapse formation and neurotransmitters regulation, while the shorter mRNA encodes beta-neurexins, which regulate the specificity of synaptic connections [104,105].

Heterozygous exonic deletions of *NRXN1* have been linked to a variety of NDDs, including epilepsy [102]. However, several studies have also detected deletions in healthy controls and apparently healthy parents of affected individuals [106,107]. This suggests that additional factors—whether genetic, epigenetic, or environmental—may influence the penetrance and expressivity of *NRXN1*, thereby explaining the broad phenotypic spectrum observed in patients [102,106]. Further research into the role of *NRXN1* in epilepsy has revealed that heterozygous microdeletions affecting the promoter region and the first three exons, which disrupt the production of the alpha-isoform, are variants associated with an elevated risk of developing IGE

[108]. In the study by Møller et al. (2013), heterozygous *NRXNI* microdeletions were found in 5 of 1,569 individuals with IGE (patients with severe psychiatric disorders or moderate to severe intellectual disability were excluded) and in 2 of 6,201 healthy controls, giving frequencies of 0.3% and 0.03%, respectively [108]. Family history analysis of the five affected individuals indicated that two had inherited the mutation from a healthy parent [108]. This, along with the presence of the microdeletion in two healthy controls, suggests that *NRXNI* microdeletions may function as susceptibility variants for IGE [108].

The variant identified in patient 1 involving *NRXNI* is a loss-of-function deletion that spans the promoter and the first three exons, resulting in the disruption of the alpha-isoform. The variant is documented in the literature and has been observed in IGE patients. However, it has also been found in a small percentage of healthy individuals. In accordance with the ACMG guidelines (Appendix **Table IV**), this suggests that the variant should be classified as potentially pathogenic.

Møller et al. (2013) study also identified a large CNV (>500 kb) in three of the five patients with *NRXNI* deletions, which was absent in both the healthy controls and the apparently healthy carrier parents [108]. This large CNV was identified in patient 1, whereas it was absent in the healthy mother. This finding, consistent with reports in the literature, helps to explain the normal phenotype of the mother versus the pathological outcome observed in the daughter.

The presence of both the heterozygous deletions of *NRXNI* and the large CNV, supports the two-hit hypothesis, which suggests that a single mutation is often insufficient to cause disease and that a second mutation is required for the condition to manifest [109]. In the context of the IGE study, this model proposes that the *NRXNI* microdeletion in heterozygosity acts as the initial predisposing event for IGE development, while a second genetic variant (>500 kb) affecting genes involved in synaptic transmission, ultimately triggers the onset of the disease [108,109].

In patient 1, a heterozygous deletion involving *NRXNI*, inherited from the healthy mother, was identified along with a 1.4 Mb deletion at the 16p13.11 locus. Since this larger deletion was absent in the mother, it may provide a possible explanation for the pathological phenotype observed in the proband. Recent studies have shown a clear association between deletions involving 16p13.11 and IGE [110,111]. Specifically, De Kovel et al. (2010) identified the 16p13.11 deletion in 6 out of 1234 patients with IGE (0.5%), compared to only 2 out of 3022 healthy controls (0.07%). The significant statistics reported in this study confirm the association between this deletion and IGE. Additionally, some of the 16p13.11 deletions observed in IGE patients were inherited, while others occurred *de novo* [111]. Interestingly, in some instances, parents carrying the deletion did not show clinical signs of epilepsy, suggesting incomplete penetrance and variable clinical expression [110,111]. Despite this variability, the results support the 16p13.11 deletion as a risk factor for IGE [110,111]. Among the genes affected by this deletion, *NDEI* (Nuclear Distribution Element 1), which was also found to be deleted in patient 1, stands out as a potential candidate for association with epilepsy, given its critical role in human brain development, particularly in neuronal migration and cortical formation [110,111].

Since the 16p13.11 deletion was found at a higher frequency in IGE patients compared to controls, and the deletion of *NDEI* (which is also deleted in the patient 1) has been associated with a loss of function [112], it is proposed to classify the variant as potentially pathogenic in accordance with the ACMG guidelines (Appendix **Table IV**) and the scientific literature.

In conclusion, the co-presence of the heterozygous microdeletion involving the *NRXNI* gene and the heterozygous deletion of 16p13.11 in patient 1, both recognized risk factors associated with IGE, is consistent with the two-hit hypothesis. According to this model, the combined effect of two potentially pathogenic variants, is sufficient to result in the manifestation of the clinical phenotype. This genetic interaction supports the role of these variants in developing IGE.

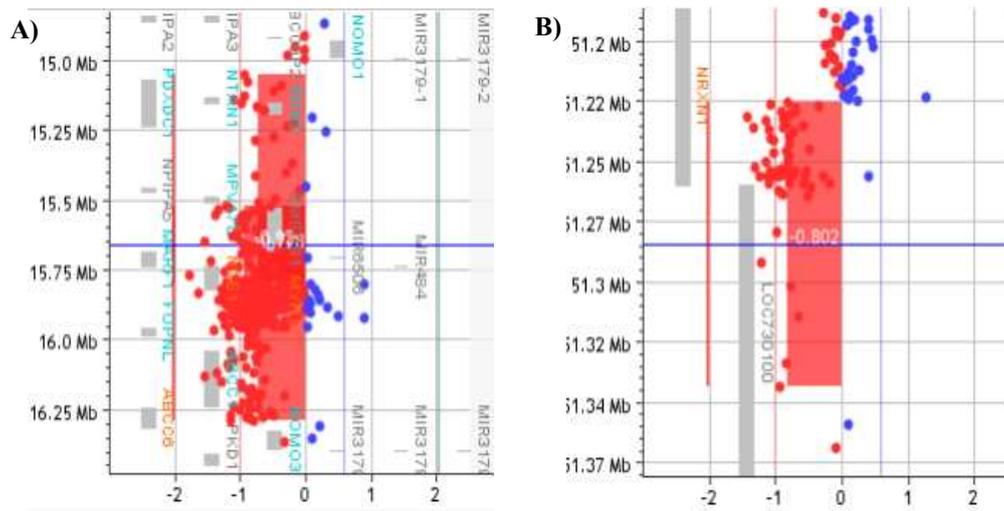


Figure 10. Array-CGH analysis of patient 1. The X-axis shows the change in signal intensity ratio between reference and proband samples, while the Y-axis shows the genomic position. The red dots represent the proband probes, while the blue dots represent the reference probes. The region highlighted in red marks the area where the genomic variation is observed. **(A)** 16p13.11 deletion with MeanLogRatio of 0.928. **(B)** 2p16.3 deletion with MeanLogRatio of 0.802.

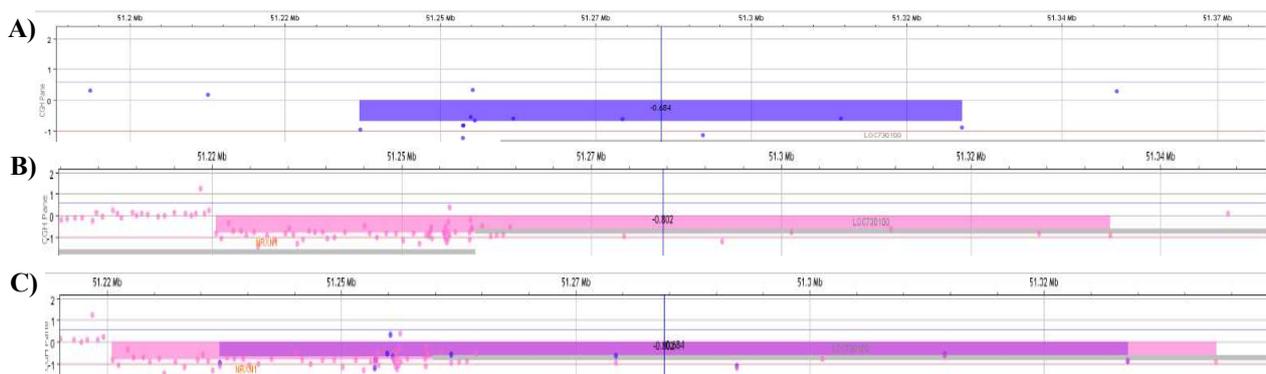


Figure 11. Comparison of proband and maternal samples in array-CGH analysis. **(A)** Upper panel: 2p16.3 deletion detected in maternal sample (purple). **(B)** Middle panel: 2p16.3 deletion identified in the proband sample (pink). **(C)** Bottom panel: Overlapping region showing the common deletion between proband and maternal sample. The alignment of the signals in both samples support the hypothesis of maternal transmission of this CNV.

Patient 2

Patient 2 is 20-year-old woman presenting epileptic encephalopathy, dysmorphic features, intellectual disability and severe psycho-motor delay. Epileptic encephalopathies are a group of disorders characterized by the deterioration in cognitive, sensory and/or motor function due to abnormal neuronal activity [113]. The presence of both epileptic encephalopathy and intellectual disability defines a condition known as

developmental epileptic encephalopathy (DEE) [114]. DEE affected individuals often experience frequent and variable type of seizures, intellectual disability and significant developmental delay [115].

Bioinformatic analysis of patient 2 using VarSeq identified an 8,9 kb duplication encompassing the entire *WDR37* gene (OMIM: 618586) on chromosome 10. This duplication was subsequently confirmed by array-CGH, which revealed an amplification of two-gain copies, giving a total of four copies. The CNV also included the genes *IDI2-AS1* (OMIM: 615391) and *IDII* (OMIM: 604055), spanning 110.73 kb. Analysis of the parental samples showed the presence of the same variant in the father, supporting paternal inheritance (**Figure 12**).

IDII gene encodes for a protein involved in cholesterol biosynthesis and other metabolic pathways, although it has not yet been linked to any specific human disease [116]. Similarly, *IDI2-AS1* is an antisense RNA of unclear function and does not appear to be involved in disease. *WDR37* is a ubiquitously expressed WD40 repeat (WDR) protein [117]. WDR proteins are involved in a variety of cellular processes including gene expression, signal transduction, protein metabolism, cell cycle regulation and immune responses [118]. Although the exact role of *WDR37* remains to be defined, the study of Sorokina et al. (2021) have shown that it interacts with the proteins PCAS1 and PACS2. These proteins are involved in intracellular protein trafficking, regulation of cell survival and calcium homeostasis [119,120].

WDR37 mutations are associated with Neuro-oculo-cardio-genito-urinary syndrome (OMIM: 618652), also known as *WDR37* syndrome, which is a multisystem disorder characterized by poor growth and anomalies of the ocular, craniofacial, neurologic, cardiovascular, skeletal, and gastrointestinal systems [117]. Affected individuals show a highly variable phenotype mainly characterized by intellectual disability, moderate to severe psycho-motor delay, facial dysmorphisms, epilepsy and abnormal cerebellum morphology [117].

A recent study by Agarwala, P., et al. (2023), showed that individuals with *WDR37* mutations have a clinical phenotype similar to that seen in DEEs, suggesting that *WDR37* alterations may play a role in the pathogenesis of these disorders. This association may be due to impaired interaction of *WDR37* with its major partners, the PACS1 and PACS2 proteins [118]. Furthermore, a significant phenotypic overlap has been observed between individuals with mutations in *PACSI*, *PACS2* and those with *WDR37* mutations [118]. Notably, alterations in *PACSI* and *PACS2* have already been associated with the development of DEEs [118].

WDR37 variants associated with the disease are primarily heterozygous missense mutations or deletions that lead to loss of protein function [117,118]. However, a few duplications of varying sizes affecting more than one gene, including *WDR37*, have been reported in databases [121]. In particular, a duplication involving six genes, including *IDII* and *WDR37* (also duplicated in patient 2), has been detected in an individual with intellectual disability and short stature [121].

Patient 2 carries an amplification of *WDR37*, *IDI2-AS1* and *IDII* with a gain of two copies, giving a total of four copies. To date, no duplications - and therefore no direct evidence of *WDR37* overexpression contributing to the pathogenesis of these disorders - have been described in the literature. This lack of data may reflect the rarity of this disorder and the limited number of documented cases in both scientific publications and databases. In addition, patient 2 inherited the variant from his unaffected father, suggesting incomplete penetrance, a finding also reported in patients with *WDR37* missense mutations or deletions [118].

In conclusion, the phenotype of patient 2, characterized by intellectual disability, severe psychomotor delay, dysmorphic features and epileptic encephalopathy, is consistent with that observed in patients with *WDR37* loss-of-function mutations [2–4]. However, the lack of information on the potential pathogenic role of *WDR37* gain-of-function mutations must be taken into account. Therefore, according to the ACMG guidelines (Appendix **Table IV**), this variant is classified as a VUS.

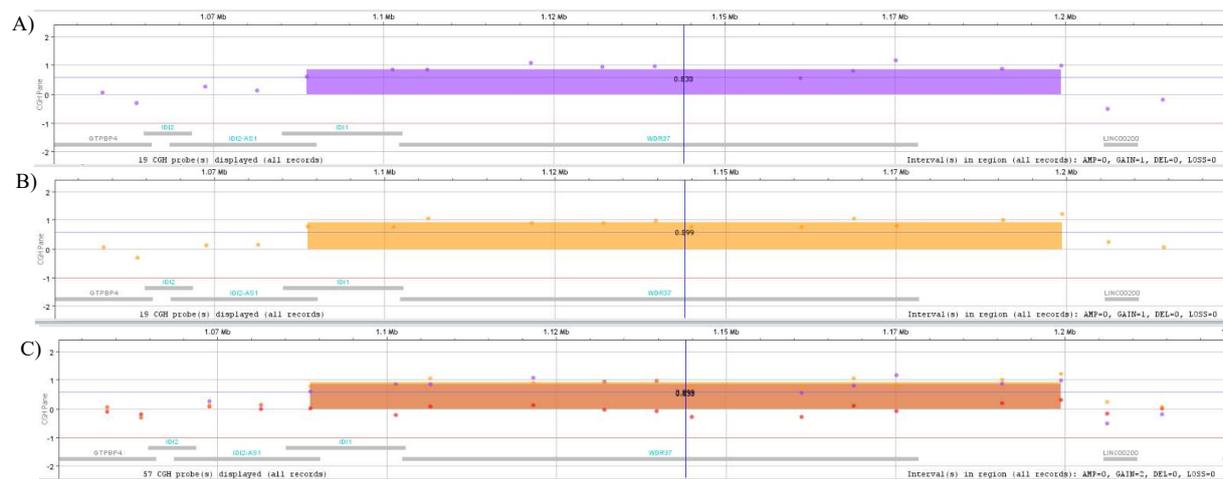


Figure 12. Comparison of proband and paternal samples in array-CGH analysis. (A) Upper panel: 10p15.3 duplication detected in the proband (purple). **(B)** Middle panel: 10p56.3 duplication identified in paternal sample (yellow). **(C)** Bottom panel: Overlapping region showing the common duplication between proband and paternal sample supporting the hypothesis of paternal inheritance of the event.

Patient 3

Patient 3 is an 18-year-old woman affected by myoclonic epileptic seizures, occasionally accompanied by brief absence episodes. Recently, there has been a noticeable increase in nocturnal seizures. Her psychomotor development remains normal. Myoclonic seizures are a type of seizure characterized by sudden, involuntary muscle jerks [124]. Individuals with myoclonic seizures often experience other types of seizures, including absence seizures [124]. Nocturnal seizures are defined by their occurrence during sleep and their potential to disrupt normal resting patterns [125,126].

Bioinformatic analysis of patient 3 using VarSeq software identified a heterozygous deletion of 16206 kb encompassing exons 5 to 10 of the *TMTC3* gene (OMIM: 617218), on chromosome 12. This variant was validated and confirmed by real-time PCR. Analysis of the exons adjacent to the event (exon 4 and exon 11) yielded wild-type results, supporting the CNV size predicted by the VarSeq software. The lack of parental samples from patient 3 prevented the evaluation of variant segregation.

TMTC3 encodes an O-mannosyl-transferase protein that belongs to a family of four transmembrane O-mannosyl-transferases found in the endoplasmic reticulum. The *TMTC3* gene contains two primary domains: the transmembrane domain (TMD), which plays a role in protein integration and positioning within membranes, and the tetratricopeptide repeat (TPR) domain, which is primarily involved in protein interactions and cellular regulatory processes [127]. Bi-allelic mutations in *TMTC3* have been associated with the occurrence of Lissencephaly-8 (LIS8) (OMIM: 617255) [126]. This is an autosomal recessive neurological disorder characterised by delayed psycho-motor development, intellectual disability with poor or absent speech, early-onset seizures and hypotonia [127]. The link between mutations in *TMTC3* and the development of neurocognitive disorders and epilepsy was clarified by showing that *TMTC3* localises in synaptic regions where it plays a role in regulating GABAergic inhibitory synapses, highlighting its importance in maintaining proper neuronal function [128].

Individuals with homozygous or compound heterozygous mutations in the *TMTC3* gene show variable expressivity in their clinical phenotype [127,129]. However, common features include moderate to severe

intellectual disability and seizures, typically myoclonic and nocturnal [127,129]. Patient 3 presents with nocturnal myoclonus seizures, consistent with the phenotype seen in individuals with bi-allelic mutations, but demonstrates normal cognitive and motor development, in contrast to other patients who all exhibit intellectual disability. To date, only missense point mutations, frameshift mutations and the introduction of a premature stop codon, within the TMD and TPR regions, have been reported [127–129]. Currently, *TMTC3* mutations have been identified and reported in the literature in 15 affected individuals from eight families, seven of which are consanguineous [127-129].

Patient 3 has a heterozygous deletion spanning from exon 5 to exon 10, which includes both the TMD and TRD domains. Although this is the only mutation identified in patient 3, there may be unidentified variants in non-coding regions of *TMTC3*. To date, mutations in the intronic regions or the 5'UTR of *TMTC3* are classified as VUS [130], yet they provide a valuable starting point for further research.

In conclusion, the lack of evidence in databases and literature regarding the effects of *TMTC3* deletions may be explained by the small number of cases observed, due to the rarity of the disorder. However, the loss of function mechanism is known to be associated with disease, suggesting that variant should be classified as a VUS, according to ACMG guidelines (Appendix **Table IV**).

Patient 4

Patient 4 is a 17-month-old girl suffering from febrile seizures and delayed motor and speech development, along with poor response to drugs. Febrile seizures are the most common type of seizure in young children between 12 and 18 months of age, affecting 2-5% of children in Western countries [131,132]. Children with febrile seizures often have psycho-motor delays, mainly due to specific mutations that can contribute to brain dysfunction and disrupt normal neurodevelopment [133].

Bioinformatic analysis of patient 4 using VarSeq software identified a 6377 bp duplication spanning exons 15 to 19 of the *SCN1A* gene (OMIM: 182389), located on chromosome 2. The presence of the variant was confirmed by real-time PCR. In addition, wild-type results from analysis of the exons adjacent to the event (exon 14 and exon 20) supported the CNV size predicted by the VarSeq software. Analysis of parental samples allowed the variant to be classified as *de novo*.

The *SCN1A* gene encodes the alpha subunit of voltage-gated sodium channels, which are essential for regulating sodium exchange between the intracellular and extracellular spaces [134,135]. These channels play a critical role in the generation and propagation of action potentials in both neurons and muscle cells [134]. Indeed, *SCN1A* is essential for maintaining proper neuronal excitability and signalling [135]. Mutations in *SCN1A* gene are associated with developmental and epileptic encephalopathy 6B, non-Dravet (OMIM: 619317), Dravet syndrome (OMIM: 607208), Febrile seizures, familial, 3A (OMIM: 604403), Generalized epilepsy with febrile seizures plus, type 2 (OMIM: 604403) and Migraine, familial hemiplegic, 3 (OMIM: 609634).

Based on the number of patients in whom mutations have been identified, *SCN1A* is predominantly associated with Dravet syndrome (DS) [136], accounting for approximately 80% affected individuals [137]. DS is a severe neurological disorder that typically begins in the first year of life and it is characterized by the onset of seizures that are often triggered by fever and are frequently resistant to treatment [137,138]. By the second year of life, signs of psychomotor delay often become evident, followed by the onset of behavioural challenges and learning disabilities [138,139].

Although *SCN1A* mutations can be inherited, most occur *de novo* [140]. The majority of these mutations are either deletions or point mutations [136,141,142]. However, Marini et al. (2009) study investigated chromosomal rearrangements involving *SCN1A* in patients with DS and identified duplications in four patients in a cohort of 269 individuals [136]. This cohort included 126 patients with DS, 97 with generalised

epilepsy with febrile seizures plus and 66 with other severe encephalopathic epilepsies [136]. Among the four patients identified, one presented with early-onset epileptic encephalopathies and carried a duplication of exon 26, while the remaining three were diagnosed with DS — one with a duplication of exon 26 and two with duplications spanning exons 8 to 16. The clinical manifestations in these patients included the onset of febrile seizures that later evolved into other seizure types, along with motor and/or learning delays and intellectual disability.

Patient 4 presented with treatment-resistant febrile seizures and delays in speech and movement. Her phenotype aligns with that typically seen in DS patients. Specifically, at 17 months old, she is within the age range of those reported in the literature, where febrile seizures in DS patients typically begin in infancy and evolve into different types over time. In addition, patient 4 shows delays in speech and movement, symptoms commonly observed in DS [133,137].

The analysis of patient 4 revealed a duplication of exons 15 to 19 in the *SCN1A* gene. Notably, exons 15 and 16 were also found to be duplicated in two patients with DS [136].

In conclusion, the *de novo* occurrence of the duplication, and the presence of the exon 15-16 duplication in affected individuals reported in the literature suggest that the variant should be classified as potentially pathogenic in accordance with the ACMG guidelines (Appendix **Table IV**) and the literature.

In conclusion, analysis of CNVs in the cohort of 100 patients, who had previously tested negative for SNVs screening, enabled the identification of four patients with five CNVs of interest. The following variants were analysed in detail and classified, according to the ACMG guidelines (Appendix **Table IV**), as follows: three classified as potentially pathogenic and two as VUS. This result points to a possible genetic involvement of the CNVs that were detected in the four patients. This finding is consistent with data reported in the literature, which indicate that diagnostic CNVs can be identified in 1-4% of individuals with epilepsy and in more than 10% of individuals with seizures and NDDs [82]. Thus, this thesis work further supports the crucial role of CNVs as genetic causes of NDDs, including epilepsy.

However, the fact that 96% of patients remain undiagnosed is also a significant finding that warrants careful consideration. During bioinformatic analysis with VarSeq software, the prioritisation process, with the use of specific filters, may have led to the exclusion of some variants. In addition, only calls with high quality metrics were considered; however, the DNA samples were extracted a few years ago. This may have affected the sequencing quality and therefore the accuracy of the CNV calls made by VarSeq as they were compared to more recently sequenced DNA samples. It is also possible that some patients carry variants in non-coding regions of the genome or in genes that have not yet been linked to epilepsy. Although this analysis focused on the monogenic causes of epilepsy, it is well known that the condition can also have a polygenic origin. Therefore, variants in other genomic regions, or non-genetic factors, may also contribute to the pathological phenotype observed in these patients.

As already mentioned, bioinformatic analysis led to the identification of thirteen patients with CNVs showing good-quality metrics. Subsequent genotype-phenotype correlation, based on clinical information provided by physicians, led to additional prioritization of four variants. Notably, nine patients displayed phenotypes that were highly inconsistent with those reported in the literature for the identified mutations and were therefore excluded from further consideration (details on these patients, including metrics and genotype-phenotype associations, are provided in Appendix **Table V**).

A parallel analysis was conducted on a second cohort of 100 NDDs patients (see section 1, "Cohort of Patients", Materials and Methods). This group of NDDs patients was obtained from previously sequenced samples that were analysed using bioinformatic methods to identify SNVs and CNVs. No significant variants were detected, leaving the patients without a confirmed molecular diagnosis. This outcome is consistent with findings reported in the literature, where 40-60% of patients with genetic NDDs remain without a definitive molecular diagnosis [142]. The recent discovery of *RNU4-2* snRNA variants in NDDs suggested its potential role as a causative gene in these disorders [61,64]. Building on this, a secondary analysis focusing on *RNU4-2* was performed in the cohort of 100 NDDs patients without a molecular diagnosis.

The 100 samples were analysed with Sanger sequencing which led to the identification of a variant in the *RNU4-2* gene in one patient. Specifically, the patient has a thymine insertion between nucleotides 64 and 65 (GRCh38: chr.12) (**Figure 13**), located in a critical 18 bp region of the *RNU4-2* gene (GRCh38: chr.12: 120291825-120291842) [61]. This region lies between stem I, where the U4 and U6 snRNAs pair, and the 3' stem-loop structures [61]. Variants within this critical sequence disrupt the proper positioning of the U6 ACAGAGA box, which is essential for binding to the 5' splice site (5'SS) and weaken the interaction between the BRR2 helicase and U4 [61,66].

The patient with the identified mutation is a 12-year-old male presenting with microcephaly, severe intellectual disability, reduced bone mineralization, and febrile convulsions. EEG analysis also revealed abnormal findings. The phenotypic characterization of patients with the 64_65insT *RNU4-2* variant revealed significant variability in clinical manifestations [61,68]. Some features were consistently observed across all patients, including moderate to severe intellectual disability, often accompanied by developmental delay, and epilepsy, including febrile seizures [61,68]. EEG analysis showed abnormalities in most patients [68]. Microcephaly, characterized by a reduced head circumference, was present in almost all patients, often accompanied by distinct facial dysmorphisms [68]. Other frequently observed phenotypes involved multiple organ systems, including skeletal abnormalities like osteopenia, recurrent fractures, and spinal deformities. [61]. The clinical phenotype of the patient with the *RNU4-2* variant matches the manifestations described in the literature.

To verify the segregation of the variant in the positive patient, parental samples were also analysed using Sanger sequencing. The results revealed the absence of the insertion in both parents, confirming the *de novo* origin of the variant (**Figure 14**), consistent with findings reported in the scientific literature [61,64,65]. Chen, Y., et al. (2024) identified this variant in 46 individuals out of 8,841 probands with undiagnosed NDDs (0.52%), all with *de novo* occurrence. Notably, the variant was absent in 3,408 probands with diagnosed NDDs, 21,817 probands with non-NDDs-related phenotypes, and 33,122 healthy individuals. The insertion identified in the patient is the most common *de novo* variant identified in *RNU4-2* snRNA, accounting for 0.52% of individuals with NDD [61].

Indeed, the insertion of a T at position 64 displaces the G at position 65. The G at position 64 of the U4 snRNA is thought to be essential for stabilising the ACAGAGA loop of the U6 snRNA, which binds to the 5' splice sites and triggers splicing after U4-U6 unwinding [64]. The displacement of the G from position 64, caused by the T insertion, is likely to disrupt the proper functioning of the major spliceosome, leading to the onset of disease in individuals harbouring this variant [61].

According to ACMG guidelines (Appendix **Table IV**), the 64_65insT variant can be classified as pathogenic due to its documented presence in the literature, its *de novo* occurrence in patient 1, and its frequency in individuals with NDDs.

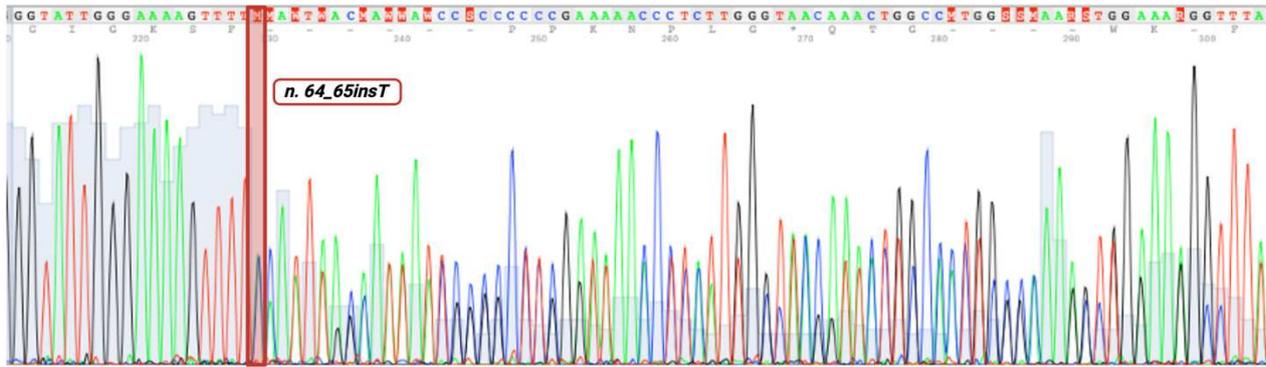


Figure 13. Electropherogram of the *RNU4-2* variant in the Proband. The red rectangle highlights the insertion of an A (corresponding to a T in the complementary genome strand) which induces a sequence shift.

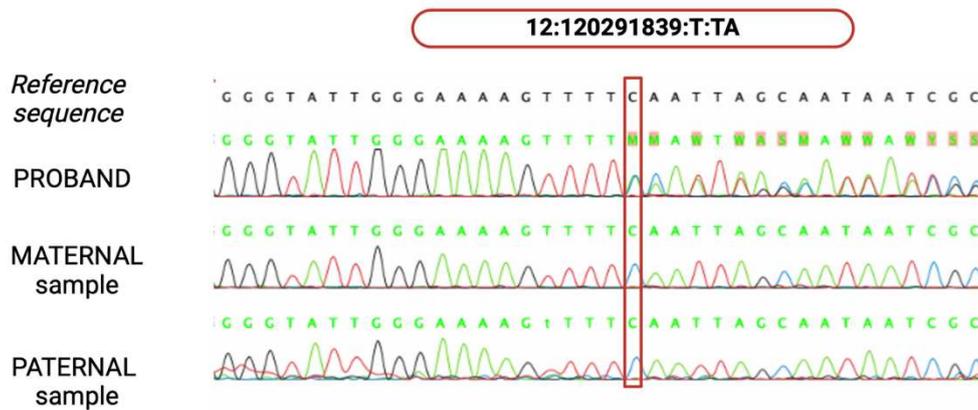


Figure 14. Comparison of proband and parents electropherograms. Sanger sequencing electropherograms comparing the proband, maternal and paternal samples against a reference sequence. The variant 12:120291839:T:TA is highlighted, indicating an insertion of an adenine (A) at this position, that is not observed in the parents.

In the study by Chen, Y., et al. (2024), the 64_65insT *RNU4-2* variant was detected in 46 individuals out of 8,841 probands with undiagnosed NDD, corresponding to an occurrence rate of 0.52%.

In this thesis work, Sanger sequencing analysis of the *RNU4-2* gene in a cohort of 100 NDD patients, all negative for SNVs and CNVs, led to the identification of 64_65insT *RNU4-2* variant in one patient, suggesting a 1% increase in the detection rate. The higher rate observed in this study compared to that reported by Chen, Y., et al. (2024) may be a result of the smaller sample size in this study (100 individuals) compared to theirs (8,841). In addition, all 100 patients had previously been tested for SNVs and CNVs in genes known to be associated with NDDs, with negative results.

In conclusion, these results are consistent with previously reported data and provide further evidence for the role of *RNU4-2* in NDDs, highlighting its potential for future use in genetic screening for these disorders.

CONCLUSIONS

In conclusion, the thesis work carried out at R&I Genetics laboratory in Padua supports that the analysis of different types of genetic variants can significantly enhance the rate of molecular diagnosis in patients with NDDs, including epilepsy. The routine analysis of SNVs, which is already well established in genetic testing, was combined with the investigation of CNVs, a more recently recognized contributor to these disorders. This integrated approach offers a greater chance of identifying a potential genetic cause, particularly in patients who have not undergone more recent, comprehensive analyses.

Furthermore, SNVs and CNVs analysis can be complemented by examining variants in non-coding regions, which may also play a crucial role in achieving a molecular diagnosis. For instance, the involvement of *RNU4-2* in the genetic determination of NDDs, that have been recently discovered, highlights the importance of widening the focus of genetic investigations.

Future prospects lie in expanding the analysis of different types of genetic variants, enabling a more comprehensive diagnostic approach. Indeed, WES allows the investigation of coding regions, which can then be confirmed through various techniques, such as real-time PCR and array-CGH. However, WES does not allow the detection of variants in intronic or other non-coding regions, which is why, in this thesis work, it has been complemented by Sanger sequencing.

To overcome these limitations and enable the detection of a broader range of genetic variants, an important next step would be the introduction of whole-genome sequencing (WGS), which would allow for the simultaneous analysis of both coding and non-coding regions. Unlike other techniques that require enrichment steps, WGS provides more homogeneous whole-genome reads, enabling more accurate detection of CNVs. In addition, WGS can identify exonic and intronic variants, triplet expansions and chromosomal rearrangements with a single assay. Nevertheless, variant confirmation still requires the use of techniques tailored to the specific type of variant identified. This approach could greatly improve genetic diagnostics, particularly for patients with only one identified variant in a recessive gene and where limitations of current techniques prevent the detection of a possible second variant.

BIBLIOGRAPHY & SITOGRAPHY

- [1] Parenti I, Rabaneda LG, Schoen H, Novarino G. Neurodevelopmental Disorders: From Genetics to Functional Pathways. *Trends Neurosci* 2020;43:608–21. <https://doi.org/10.1016/j.tins.2020.05.004>.
- [2] Gidziela A, Ahmadzadeh YI, Michelini G, Allegrini AG, Agnew-Blais J, Lau LY, et al. A meta-analysis of genetic effects associated with neurodevelopmental disorders and co-occurring conditions. *Nat Hum Behav* 2023;7:642–56. <https://doi.org/10.1038/s41562-023-01530-y>.
- [3] Zigic N, Pajevic I, Hasanovic M, Avdibegovic E, Aljukic N, Hodzic V. EPV0196 Neurodevelopmental disorders in ICD-11 classification n.d. <https://doi.org/10.1192/j.eurpsy.2023.1547>.
- [4] Morris-Rosendahl DJ, Crocq MA. Neurodevelopmental disorders-the history and future of a diagnostic concept. *Dialogues Clin Neurosci* 2020;22:65–72. <https://doi.org/10.31887/DCNS.2020.22.1/macrocq>.
- [5] Autism Spectrum Disorder - National Institute of Mental Health (NIMH) n.d. <https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd> (accessed February 21, 2025).
- [6] Boat TF, Wu JT. Clinical Characteristics of Intellectual Disabilities - Mental Disorders and Disabilities Among Low-Income Children. National Academies Press (US) 2015:169–79.
- [7] Magnus W, Anilkumar AC, Shaban K. Attention Deficit Hyperactivity Disorder. *StatPearls* 2023.
- [8] Bonti E, Zerva IK, Koundourou C, Sofologi M. The High Rates of Comorbidity among Neurodevelopmental Disorders: Reconsidering the Clinical Utility of Distinct Diagnostic Categories. *Journal of Personalized Medicine* 2024, Vol 14, Page 300 2024;14:300. <https://doi.org/10.3390/JPM14030300>.
- [9] Dingemans AJM, Jansen S, van Reeuwijk J, de Leeuw N, Pfundt R, Schuurs-Hoeijmakers J, et al. Prevalence of comorbidities in individuals with neurodevelopmental disorders from the aggregated phenomics data of 51,227 pediatric individuals. *Nat Med* 2024;30:1994–2003. <https://doi.org/10.1038/S41591-024-03005-7>.
- [10] Nagata KI. Pathophysiological Mechanism of Neurodevelopmental Disorders—Overview. *Cells* 2022, Vol 11, Page 4082 2022;11:4082. <https://doi.org/10.3390/CELLS11244082>.
- [11] Leblond CS, Le TL, Malesys S, Cliquet F, Tabet AC, Delorme R, et al. Operative list of genes associated with autism and neurodevelopmental disorders based on database review. *Molecular and Cellular Neuroscience* 2021;113:103623. <https://doi.org/10.1016/J.MCN.2021.103623>.
- [12] Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet Med* 2019;21:2413–21. <https://doi.org/10.1038/S41436-019-0554-6>.
- [13] Shankar R, Perera B, Thomas RH. Epilepsy, an orphan disorder within the neurodevelopmental family. *J Neurol Neurosurg Psychiatry* 2020;91:1245–7. <https://doi.org/10.1136/JNNP-2020-324660>.
- [14] Chakraborty S, Parayil R, Mishra S, Nongthomba U, Clement JP. Epilepsy Characteristics in Neurodevelopmental Disorders: Research from Patient Cohorts and Animal Models Focusing on Autism Spectrum Disorder. *Int J Mol Sci* 2022;23:10807. <https://doi.org/10.3390/IJMS231810807>.
- [15] Heyne HO, Singh T, Stamberger H, Abou Jamra R, Caglayan H, Craiu D, et al. De novo variants in neurodevelopmental disorders with epilepsy. *Nat Genet* 2018;50:1048–53. <https://doi.org/10.1038/S41588-018-0143-7>.
- [16] Chow J, Jensen M, Amini H, Hormozdiari F, Penn O, Shifman S, et al. Dissecting the genetic basis of comorbid epilepsy phenotypes in neurodevelopmental disorders. *Genome Med* 2019;11:1–14. <https://doi.org/10.1186/S13073-019-0678-Y/FIGURES/5>.
- [17] Watkins L V., Linehan C, Brandt C, Snoeijen-Schouwenaars F, McGowan P, Shankar R. Epilepsy in adults with neurodevelopmental disability - what every neurologist should know. *Epileptic Disord* 2022;24:9–25. <https://doi.org/10.1684/EPD.2021.1366>.
- [18] Shimizu H, Morimoto Y, Yamamoto N, Tayama T, Ozawa H, Imamura A. Overlap Between Epilepsy and Neurodevelopmental Disorders: Insights from Clinical and Genetic Studies. *Epilepsy* 2022;41–54. <https://doi.org/10.36255/EXON-PUBLICATIONS-EPILEPSY-NEURODEVELOPMENTAL-DISORDERS>.
- [19] Thijs RD, Surges R, O'Brien TJ, Sander JW. Epilepsy in adults. *Lancet* 2019;393:689–701. [https://doi.org/10.1016/S0140-6736\(18\)32596-0](https://doi.org/10.1016/S0140-6736(18)32596-0).

- [20] Bastos F, Cross JH. *Epilepsy*. *Handb Clin Neurol* 2020;174:137–58. <https://doi.org/10.1016/B978-0-444-64148-9.00011-9>.
- [21] *Epilepsy* n.d. <https://www.who.int/news-room/fact-sheets/detail/epilepsy> (accessed February 21, 2025).
- [22] Penderis J. Pathophysiology of epileptic seizures. *In Pract* 2014;36:3–9. <https://doi.org/10.1136/INP.G5098>.
- [23] Lovinger DM. Communication Networks in the Brain: Neurons, Receptors, Neurotransmitters, and Alcohol. *Alcohol Research & Health* 2008;31:196.
- [24] Shneker BF, Fountain NB, Orlowski JM. *Epilepsy*. *Disease-a-Month* 2003;49:426–78. [https://doi.org/10.1016/s0011-5029\(03\)00065-8](https://doi.org/10.1016/s0011-5029(03)00065-8).
- [25] Trinka E, Cock H, Hesdorffer D, Rossetti AO, Scheffer IE, Shinnar S, et al. A definition and classification of status epilepticus--Report of the ILAE Task Force on Classification of Status Epilepticus. *Epilepsia* 2015;56:1515–23. <https://doi.org/10.1111/EPI.13121>.
- [26] Elmi A, Ibrahim A, Hassan M, Osman F, Çelik C, Dirie A, et al. Magnetic Resonance Imaging Findings and Their Association with Electroencephalographic Data in Children with Epilepsy at Tertiary Care Hospital in Mogadishu Somalia. *Int J Gen Med* 2024;17:253–61. <https://doi.org/10.2147/IJGM.S448291>.
- [27] Drenthen GS, Jansen JFA, Gommer E, Gupta L, Hofman PAM, van Kranen-Mastenbroek VH, et al. Predictive value of functional MRI and EEG in epilepsy diagnosis after a first seizure. *Epilepsy Behav* 2021;115. <https://doi.org/10.1016/J.YEBEH.2020.107651>.
- [28] Types of Seizures | Epilepsy Foundation n.d. <https://www.epilepsy.com/what-is-epilepsy/seizure-types> (accessed February 21, 2025).
- [29] Absence Seizures | Johns Hopkins Medicine n.d. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/epilepsy/absence-seizures> (accessed February 21, 2025).
- [30] Fisher RS, Cross JH, French JA, Higurashi N, Hirsch E, Jansen FE, et al. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 2017;58:522–30. <https://doi.org/10.1111/EPI.13670>.
- [31] Paige AL, Cavanna AE. Generalized Tonic-Clonic Seizure. *Neuroimaging of Consciousness* 2023:81–97. https://doi.org/10.1007/978-3-642-37580-4_6.
- [32] Chen Z, Brodie MJ, Ding D, Kwan P. Editorial: Epidemiology of epilepsy and seizures. *Frontiers in Epidemiology* 2023;3. <https://doi.org/10.3389/FEPID.2023.1273163>.
- [33] Pitkänen A, Lukasiuk K. Mechanisms of epileptogenesis and potential treatment targets. *Lancet Neurol* 2011;10:173–86. [https://doi.org/10.1016/S1474-4422\(10\)70310-0](https://doi.org/10.1016/S1474-4422(10)70310-0).
- [34] Bromfield EB, Cavazos JE, Sirven JI. *An Introduction to Epilepsy*. *An Introduction to Epilepsy* 2006:4.
- [35] Abood W, Bandyopadhyay S. Postictal Seizure State. *StatPearls* 2023.
- [36] Deyn D, Paul P. *Pathophysiology of epilepsy*. vol. 100. 2000.
- [37] Ahtainen A, Genocchi B, Tanskanen JMA, Barros MT, Hyttinen JAK, Lenk K. Astrocytes Exhibit a Protective Role in Neuronal Firing Patterns under Chemically Induced Seizures in Neuron-Astrocyte Co-Cultures. *International Journal of Molecular Sciences Article* 2021. <https://doi.org/10.3390/ijms222312770>.
- [38] Vargas-Sánchez K, Mogilevskaya M, Rodríguez-Pérez J, Rubiano MG, Javela JJ, González-Reyes RE. Astroglial role in the pathophysiology of status epilepticus: an overview. *Oncotarget* 2018;9:26954. <https://doi.org/10.18632/ONCOTARGET.25485>.
- [39] Dingledine R, Varvel NH, Dudek FE. When and how do seizures kill neurons, and is cell death relevant to epileptogenesis? *Adv Exp Med Biol* 2014;813:109–22. https://doi.org/10.1007/978-94-017-8914-1_9.
- [40] Albuz B, Ozdemir O, Silan F. The high frequency of chromosomal copy number variations and candidate genes in epilepsy patients. *Clin Neurol Neurosurg* 2021;202. <https://doi.org/10.1016/J.CLINEURO.2021.106487>.
- [41] Rastin C, Schenkel LC, Sadikovic B. Complexity in Genetic Epilepsies: A Comprehensive Review. *Int J Mol Sci* 2023;24. <https://doi.org/10.3390/IJMS241914606>.
- [42] Myers CT, Mefford HC. Advancing epilepsy genetics in the genomic era. *Genome Med* 2015;7:1–11. <https://doi.org/10.1186/S13073-015-0214-7/TABLES/1>.
- [43] Zhang MW, Liang XY, Wang J, Gao L Di, Liao HJ, He YH, et al. Epilepsy-associated genes: an update. *Seizure: European Journal of Epilepsy* 2024;116:4–13. <https://doi.org/10.1016/J.SEIZURE.2023.09.021>.
- [44] Guerrini R, Conti V, Mantegazza M, Balestrini S, Galanopoulou AS, Benfenati F. Developmental and epileptic encephalopathies: from genetic heterogeneity to phenotypic continuum. *Physiol Rev* 2022;103:433. <https://doi.org/10.1152/PHYSREV.00063.2021>.

- [45] Helbig I, Lowenstein DH. Genetics of the epilepsies: where are we and where are we going? *Curr Opin Neurol* 2013;26:179–85. <https://doi.org/10.1097/WCO.0B013E32835EE6FF>.
- [46] De Wachter M, Schoonjans AS, Weckhuysen S, Van Schil K, Löfgren A, Meuwissen M, et al. From diagnosis to treatment in genetic epilepsies: Implementation of precision medicine in real-world clinical practice. *Eur J Paediatr Neurol* 2024;48:46–60. <https://doi.org/10.1016/J.EJPN.2023.11.003>.
- [47] Zou H, Wu LX, Tan L, Shang FF, Zhou HH. Significance of Single-Nucleotide Variants in Long Intergenic Non-protein Coding RNAs. *Front Cell Dev Biol* 2020;8. <https://doi.org/10.3389/FCELL.2020.00347>.
- [48] Zhu T, Salmeron J. High-definition genome profiling for genetic marker discovery. *Trends Plant Sci* 2007;12:196–202. <https://doi.org/10.1016/J.TPLANTS.2007.03.013>.
- [49] Human Genomic Variation n.d. <https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genomic-variation> (accessed February 21, 2025).
- [50] Gómez-Romero L, Palacios-Flores K, Reyes J, García D, Boege M, Dávila G, et al. Precise detection of de novo single nucleotide variants in human genomes. *Proc Natl Acad Sci U S A* 2018;115:5516–21. https://doi.org/10.1073/PNAS.1802244115/SUPPL_FILE/PNAS.1802244115.SAPP.PDF.
- [51] Marwal A, Gaur RK. Molecular markers: tool for genetic analysis. *Animal Biotechnology: Models in Discovery and Translation* 2020:353–72. <https://doi.org/10.1016/B978-0-12-811710-1.00016-1>.
- [52] Chu D, Wei L. Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. *BMC Cancer* 2019;19:1–12. <https://doi.org/10.1186/S12885-019-5572-X/FIGURES/5>.
- [53] Strachan T, Lucassen A. Genetics and Genomics in Medicine. *Genetics and Genomics in Medicine* 2022. <https://doi.org/10.1201/9781003044406>.
- [54] Krokkan HE, Bjørås M. Base excision repair. *Cold Spring Harb Perspect Biol* 2013;5:1–22. <https://doi.org/10.1101/CSHPERSPECT.A012583>.
- [55] DNA Damage – Biotech Khan n.d. <https://biotechkhan.wordpress.com/2014/10/14/dna-damage/> (accessed February 21, 2025).
- [56] Zou H, Wu LX, Tan L, Shang FF, Zhou HH. Significance of Single-Nucleotide Variants in Long Intergenic Non-protein Coding RNAs. *Front Cell Dev Biol* 2020;8:536059. <https://doi.org/10.3389/FCELL.2020.00347/PDF>.
- [57] Capriotti E, Altman RB, Bromberg Y. Collective judgment predicts disease-associated single nucleotide variants. *BMC Genomics* 2013;14 Suppl 3:1–9. <https://doi.org/10.1186/1471-2164-14-S3-S2/FIGURES/5>.
- [58] Tamargo-Gómez I, Fernández ÁF, Mariño G. Pathogenic Single Nucleotide Polymorphisms on Autophagy-Related Genes. *Int J Mol Sci* 2020;21:1–42. <https://doi.org/10.3390/IJMS21218196>.
- [59] Koh HY, Smith L, Wiltrout KN, Podury A, Chourasia N, D’Gama AM, et al. Utility of Exome Sequencing for Diagnosis in Unexplained Pediatric-Onset Epilepsy. *JAMA Netw Open* 2023;6:e2324380–e2324380. <https://doi.org/10.1001/JAMANETWORKOPEN.2023.24380>.
- [60] Hamanaka K, Miyake N, Mizuguchi T, Miyatake S, Uchiyama Y, Tsuchida N, et al. Large-scale discovery of novel neurodevelopmental disorder-related genes through a unified analysis of single-nucleotide and copy number variants. *Genome Med* 2022;14. <https://doi.org/10.1186/S13073-022-01042-W>.
- [61] Chen Y, Dawes R, Kim HC, Ljungdahl A, Stenton SL, Walker S, et al. De novo variants in the RNU4-2 snRNA cause a frequent neurodevelopmental syndrome. *Nature* 2024;632:832. <https://doi.org/10.1038/S41586-024-07773-7>.
- [62] Mattick JS, Makunin I V. Non-coding RNA. *Hum Mol Genet* 2006;15:R17–29. <https://doi.org/10.1093/HMG/DDL046>.
- [63] Sperling R. Small non-coding RNA within the endogenous spliceosome and alternative splicing regulation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 2019;1862:194406. <https://doi.org/10.1016/J.BBAGRM.2019.07.007>.
- [64] Greene D, Thys C, Berry IR, Jarvis J, Ortibus E, Mumford AD, et al. Mutations in the U4 snRNA gene RNU4-2 cause one of the most prevalent monogenic neurodevelopmental disorders. *Nature Medicine* 2024 30:8 2024;30:2165–9. <https://doi.org/10.1038/s41591-024-03085-5>.
- [65] Schot R, Ferraro F, Geeven G, Diderich KEM, Barakat TS. Re-analysis of whole genome sequencing ends a diagnostic odyssey: Case report of an RNU4-2 related neurodevelopmental disorder. *Clin Genet* 2024;106:512–7. <https://doi.org/10.1111/CGE.14574>.
- [66] Charenton C, Wilkinson ME, Nagai K. Mechanism of 5’ splice site transfer for human spliceosome activation. *Science* 2019;364:362–7. <https://doi.org/10.1126/SCIENCE.AAX3289>.

- [67] Nguyen THD, Galej WP, Bai XC, Savva CG, Newman AJ, Scheres SHW, et al. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* 2015 523:7558 2015;523:47–52. <https://doi.org/10.1038/nature14548>.
- [68] Barbour K, Bainbridge MN, Wigby K, Besterman AD, Chuang NA, Tobin LE, et al. The Face and Features of RNU4-2: A New, Common, Recognizable, Yet Hidden Neurodevelopmental Disorder. *Pediatr Neurol* 2024;161:188–93. <https://doi.org/10.1016/J.PEDIATRNEUROL.2024.09.015>.
- [69] Shaikh TH. Copy Number Variation Disorders. *Curr Genet Med Rep* 2017;5:183–90. <https://doi.org/10.1007/S40142-017-0129-2>.
- [70] Pös O, Radvanszky J, Buglyó G, Pös Z, Rusnakova D, Nagy B, et al. DNA copy number variation: Main characteristics, evolutionary significance, and pathological aspects. *Biomed J* 2021;44:548–59. <https://doi.org/10.1016/J.BJ.2021.02.003>.
- [71] Hovhannisyan G, Harutyunyan T, Aroutiounian R, Liehr T. DNA Copy Number Variations as Markers of Mutagenic Impact. *Int J Mol Sci* 2019;20:4723. <https://doi.org/10.3390/IJMS20194723>.
- [72] Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, et al. Segmental Duplications and Copy-Number Variation in the Human Genome. *The American Journal of Human Genetics* 2005;77:78–88. <https://doi.org/10.1086/431652>.
- [73] Sankaranarayanan K, Taleei R, Rahmanian S, Nikjoo H. Ionizing radiation and genetic risks. XVII. Formation mechanisms underlying naturally occurring DNA deletions in the human genome and their potential relevance for bridging the gap between induced DNA double-strand breaks and deletions in irradiated germ cells. *Mutat Res* 2013;753:114–30. <https://doi.org/10.1016/J.MRREV.2013.07.003>.
- [74] Gu W, Zhang F, Lupski JR. Mechanisms for human genomic rearrangements. *Pathogenetics* 2008;1. <https://doi.org/10.1186/1755-8417-1-4>.
- [75] Stankiewicz P, Lupski JR. Genome architecture, rearrangements and genomic disorders. *Trends Genet* 2002;18:74–82. [https://doi.org/10.1016/S0168-9525\(02\)02592-1](https://doi.org/10.1016/S0168-9525(02)02592-1).
- [76] Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res* 2013;2:130–43. <https://doi.org/10.3978/J.ISSN.2218-676X.2013.04.02>.
- [77] Bursed B, Zamariolli M, Bellucco FT, Melaragno MI. Mechanisms of structural chromosomal rearrangement formation. *Mol Cytogenet* 2022;15. <https://doi.org/10.1186/S13039-022-00600-6>.
- [78] Kim S, Peterson SE, Jasin M, Keeney S. Mechanisms of germ line genome instability. *Semin Cell Dev Biol* 2016;54:177–87. <https://doi.org/10.1016/J.SEMCDB.2016.02.019>.
- [79] Liu P, Carvalho CMB, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* 2012;22:211. <https://doi.org/10.1016/J.GDE.2012.02.012>.
- [80] Mefford HC. CNVs in Epilepsy. *Curr Genet Med Rep* 2014;2:162. <https://doi.org/10.1007/S40142-014-0046-6>.
- [81] Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human genome. *Nat Rev Genet* 2015;16:172–83. <https://doi.org/10.1038/NRG3871>.
- [82] Montanucci L, Lewis-Smith D, Collins RL, Niestroj LM, Parthasarathy S, Xian J, et al. Genome-wide identification and phenotypic characterization of seizure-associated copy number variations in 741,075 individuals. *Nat Commun* 2023;14. <https://doi.org/10.1038/S41467-023-39539-6>.
- [83] Mirzaa GM, Millen KJ, Barkovich AJ, Dobyns WB, Paciorkowski AR. The Developmental Brain Disorders Database (DBDB): A Curated Neurogenetics Knowledge Base With Clinical and Research Applications. *Am J Med Genet A* 2014;0:1503. <https://doi.org/10.1002/AJMG.A.36517>.
- [84] Zhai Y, Zhang Z, Shi P, Martin DM, Kong X. Incorporation of exome-based CNV analysis makes trio-WES a more powerful tool for clinical diagnosis in neurodevelopmental disorders: A retrospective study. *Hum Mutat* 2021;42:990–1004. <https://doi.org/10.1002/HUMU.24222>.
- [85] QIAasymphony SP/AS instruments n.d. <https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/instruments-equipment/qiasymphony-spas-instruments> (accessed February 25, 2025).
- [86] QIAasymphony® DNA Handbook. n.d.
- [87] NanoDrop Micro-UV Spectrophotometer NanoDrop Lite Plus User Guide S120 NanoDrop Lite Plus UG Revision Date 2022.
- [88] Artika IM, Dewi YP, Nainggolan IM, Siregar JE, Antonjaya U. Real-Time Polymerase Chain Reaction: Current Techniques, Applications, and Role in COVID-19 Diagnosis. *Genes (Basel)* 2022;13. <https://doi.org/10.3390/GENES13122387>.

- [89] Weiss MM, Hermsen MAJA, Meijer GA, Van Grieken NCT, Baak JPA, Kuipers EJ, et al. Comparative genomic hybridisation. *Molecular Pathology* 1999;52:243. <https://doi.org/10.1136/MP.52.5.243>.
- [90] Ahmad E, Ali A, Nimisha, Kumar Sharma A, Ahmed F, Mehdi Dar G, et al. Molecular approaches in cancer. *Clinica Chimica Acta* 2022;537:60–73. <https://doi.org/10.1016/J.CCA.2022.09.027>.
- [91] For Research Use Only. Not for Use in Diagnostic Procedures n.d.
- [92] Incubation time n.d.
- [93] Sanger Sequencing: Principle, Steps, Applications, Diagram n.d. <https://microbenotes.com/sanger-sequencing/> (accessed February 25, 2025).
- [94] Sanger Sequencing: Introduction, Principle, and Protocol | CD Genomics Blog n.d. <https://www.cd-genomics.com/blog/sanger-sequencing-introduction-principle-and-protocol/> (accessed February 25, 2025).
- [95] Sequenziamento con metodo Sanger, procedura e passaggi n.d. https://www.sigmaaldrich.com/IT/it/technical-documents/protocol/genomics/sequencing/sanger-sequencing?srsltid=AfmBOor9d_GRc3TcugVUdrAnMeiwjwFIT5DJI_lwi23F7GzCwXKMawBi (accessed February 25, 2025).
- [96] Lee PY, Costumbrado J, Hsu CY, Kim YH. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp* 2012. <https://doi.org/10.3791/3923>.
- [97] Applied Biosystems BigDye® X Terminator™ Purification Kit Protocol 2007.
- [98] Fisher Scientific T. When used for purposes other than Human Identification the instruments cited are for Research Use Only. Not for use in diagnostic procedures. SeqStudio™ Flex Series Genetic Analyzer with Instrument Software v1.1.1. n.d.
- [99] Perez-Velasquez JL, Bertram EH, Carter Snead O. Idiopathic (Genetic) Generalized Epilepsy. *Neurobiol Dis* 2024;289–96. <https://doi.org/10.1016/B978-012088592-3/50030-X>.
- [100] Devinsky O, Elder C, Sivathamboo S, Scheffer IE, Koepp MJ. Idiopathic Generalized Epilepsy: Misunderstandings, Challenges, and Opportunities. *Neurology* 2024;102. <https://doi.org/10.1212/WNL.0000000000208076>.
- [101] Gesche J, Beier CP. Drug resistance in idiopathic generalized epilepsies: Evidence and concepts. *Epilepsia* 2022;63:3007–19. <https://doi.org/10.1111/EPI.17410>.
- [102] Al Shehhi M, Forman EB, Fitzgerald JE, McInerney V, Krawczyk J, Shen S, et al. NRXN1 deletion syndrome; phenotypic and penetrance data from 34 families. *Eur J Med Genet* 2019;62:204–9. <https://doi.org/10.1016/J.EJMG.2018.07.015>.
- [103] Montalbano S, Krebs MD, Rosengren A, Vaez M, Hellberg KLG, Mortensen PB, et al. Analysis of exonic deletions in a large population study provides novel insights into NRXN1 pathology. *Npj Genomic Medicine* 2024 9:1 2024;9:1–10. <https://doi.org/10.1038/s41525-024-00450-8>.
- [104] Südhof TC. Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. *Cell* 2017;171:745–69. <https://doi.org/10.1016/J.CELL.2017.10.024>.
- [105] Hu Z, Xiao X, Zhang Z, Li M. Genetic insights and neurobiological implications from NRXN1 in neuropsychiatric disorders. *Mol Psychiatry* 2019;24:1400–14. <https://doi.org/10.1038/S41380-019-0438-9>.
- [106] Ching MSL, Shen Y, Tan WH, Jeste SS, Morrow EM, Chen X, et al. Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. *Am J Med Genet B Neuropsychiatr Genet* 2010;153B:937–47. <https://doi.org/10.1002/AJMG.B.31063>.
- [107] Dabell MP, Rosenfeld JA, Bader P, Escobar LF, El-Khechen D, Vallee SE, et al. Investigation of NRXN1 deletions: clinical and molecular characterization. *Am J Med Genet A* 2013;161A:717–31. <https://doi.org/10.1002/AJMG.A.35780>.
- [108] Møller RS, Weber YG, Klitten LL, Trucks H, Muhle H, Kunz WS, et al. Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. *Epilepsia* 2013;54:256–64. <https://doi.org/10.1111/EPI.12078>.
- [109] Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, et al. A recurrent 16p12.1 microdeletion suggests a two-hit model for severe developmental delay. *Nat Genet* 2010;42:203. <https://doi.org/10.1038/NG.534>.
- [110] Heinzen EL, Radtke RA, Urban TJ, Cavalleri GL, Depondt C, Need AC, et al. Rare Deletions at 16p13.11 Predispose to a Diverse Spectrum of Sporadic Epilepsy Syndromes. *Am J Hum Genet* 2010;86:707. <https://doi.org/10.1016/J.AJHG.2010.03.018>.
- [111] De Kovel CGF, Trucks H, Helbig I, Mefford HC, Baker C, Leu C, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 2010;133:23–32. <https://doi.org/10.1093/BRAIN/AWP262>.

- [112] Liu JYW, Kasperavičiute D, Martinian L, Thom M, Sisodiya SM. Neuropathology of 16p13.11 Deletion in Epilepsy. *PLoS One* 2012;7:e34813. <https://doi.org/10.1371/JOURNAL.PONE.0034813>.
- [113] Dulac O. Epileptic Encephalopathy. *Epilepsia* 2001;42:23–6. <https://doi.org/10.1046/J.1528-1157.2001.042SUPPL.3023.X>.
- [114] Poke G, Stanley J, Scheffer IE, Sadleir LG. Epidemiology of Developmental and Epileptic Encephalopathy and of Intellectual Disability and Epilepsy in Children. *Neurology* 2023;100:E1363–75. <https://doi.org/10.1212/WNL.0000000000206758>.
- [115] Developmental Epileptic Encephalopathy | Epilepsy Foundation n.d. <https://epilepsyfoundation.org.au/understanding-epilepsy/rare-and-genetic-epilepsies/developmental-and-epileptic-encephalopathy/> (accessed March 11, 2025).
- [116] Clizbe DB, Owens ML, Masuda KR, Shackelford JE, Krisans SK. IDI2, a second isopentenyl diphosphate isomerase in mammals. *Journal of Biological Chemistry* 2007;282:6668–76. <https://doi.org/10.1074/jbc.M610922200>.
- [117] Kanca O, Andrews JC, Lee PT, Patel C, Braddock SR, Slavotinek AM, et al. De Novo Variants in WDR37 Are Associated with Epilepsy, Colobomas, Dysmorphism, Developmental Delay, Intellectual Disability, and Cerebellar Hypoplasia. *Am J Hum Genet* 2019;105:413–24. <https://doi.org/10.1016/J.AJHG.2019.06.014>.
- [118] Sorokina EA, Reis LM, Thompson S, Agre K, Babovic-Vuksanovic D, Ellingson MS, et al. WDR37 syndrome: identification of a distinct new cluster of disease-associated variants and functional analyses of mutant proteins. *Hum Genet* 2021;140:1775–89. <https://doi.org/10.1007/S00439-021-02384-Y>.
- [119] Schuurs-Hoeijmakers JHM, Oh EC, Vissers LELM, Swinkels MEM, Gilissen C, Willemsen MA, et al. Recurrent de novo mutations in PACS1 cause defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome. *Am J Hum Genet* 2012;91:1122–7. <https://doi.org/10.1016/J.AJHG.2012.10.013>.
- [120] Olson HE, Jean-Marçais N, Yang E, Heron D, Tatton-Brown K, van der Zwaag PA, et al. A Recurrent De Novo PACS2 Heterozygous Missense Variant Causes Neonatal-Onset Developmental Epileptic Encephalopathy, Facial Dysmorphism, and Cerebellar Dysgenesis. *Am J Hum Genet* 2018;102:995–1007. <https://doi.org/10.1016/J.AJHG.2018.03.005>.
- [121] “wdr37”[GENE] - ClinVar - NCBI n.d. <https://www.ncbi.nlm.nih.gov/clinvar> (accessed March 19, 2025).
- [122] Clinical Synopsis - #618652 - NEUROOCULOCARDIOGENITOURINARY SYNDROME; NOCGUS - OMIM n.d. <https://omim.org/clinicalSynopsis/618652?highlight=wdr37> (accessed March 14, 2025).
- [123] Hay E, Henderson RH, Mansour S, Deshpande C, Jones R, Nutan S, et al. Expanding the phenotypic spectrum consequent upon de novo WDR37 missense variants. *Clin Genet* 2020;98:191–7. <https://doi.org/10.1111/CGE.13795>.
- [124] Myoclonic Seizure: What It Is, Symptoms & Treatment n.d. <https://my.clevelandclinic.org/health/diseases/23172-myoclonic-seizure> (accessed March 11, 2025).
- [125] Absence Seizures | Johns Hopkins Medicine n.d. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/epilepsy/absence-seizures> (accessed March 11, 2025).
- [126] Nocturnal Seizures | Cedars-Sinai n.d. <https://www.cedars-sinai.org/health-library/diseases-and-conditions/n/nocturnal-seizures.html> (accessed March 11, 2025).
- [127] Jerber J, Zaki MS, Al-Aama JY, Rosti RO, Ben-Omran T, Dikoglu E, et al. Biallelic Mutations in TMTC3, Encoding a Transmembrane and TPR-Containing Protein, Lead to Cobblestone Lissencephaly. *Am J Hum Genet* 2016;99:1181–9. <https://doi.org/10.1016/J.AJHG.2016.09.007>.
- [128] Farhan SMK, Nixon KCJ, Everest M, Edwards TN, Long S, Segal D, et al. Identification of a novel synaptic protein, TMTC3, involved in periventricular nodular heterotopia with intellectual disability and epilepsy. *Hum Mol Genet* 2017;26:4278–89. <https://doi.org/10.1093/HMG/DDX316>.
- [129] Hana S, Karthik D, Shan J, Hayek S El, Chouchane L, Megarbane A. A Report on a Family with TMTC3-Related Syndrome and Review. *Case Rep Med* 2020;2020:7163038. <https://doi.org/10.1155/2020/7163038>.
- [130] “TMTC3”[GENE] - ClinVar - NCBI n.d. <https://www.ncbi.nlm.nih.gov/clinvar/?term=%22TMTC3%22%5BGENE%5D&redir=gene> (accessed March 12, 2025).
- [131] Febrile Seizures | National Institute of Neurological Disorders and Stroke n.d. <https://www.ninds.nih.gov/health-information/disorders/febrile-seizures> (accessed March 11, 2025).

- [132] Corsello A, Marangoni MB, Macchi M, Cozzi L, Agostoni C, Milani GP, et al. Febrile Seizures: A Systematic Review of Different Guidelines. *Pediatr Neurol* 2024;155:141–8. <https://doi.org/10.1016/J.PEDIATRNEUROL.2024.03.024>.
- [133] SCN1A Seizure Disorders - GeneReviews® - NCBI Bookshelf n.d. <https://www.ncbi.nlm.nih.gov/books/NBK1318/> (accessed March 11, 2025).
- [134] SCN1A sodium voltage-gated channel alpha subunit 1 [Homo sapiens (human)] - Gene - NCBI n.d. <https://www.ncbi.nlm.nih.gov/gene/6323> (accessed March 11, 2025).
- [135] Goldin AL, Snutch T, Lubbert H, Dowsett A, Marshall J, Auld V, et al. Messenger RNA coding for only the alpha subunit of the rat brain Na channel is sufficient for expression of functional channels in *Xenopus* oocytes. *Proc Natl Acad Sci U S A* 1986;83:7503–7. <https://doi.org/10.1073/PNAS.83.19.7503>.
- [136] Marini C, Scheffer IE, Nabbout R, Mei D, Cox K, Dibbens LM, et al. SCN1A duplications and deletions detected in Dravet syndrome: implications for molecular diagnosis. *Epilepsia* 2009;50:1670–8. <https://doi.org/10.1111/J.1528-1167.2009.02013.X>.
- [137] Ding J, Li X, Tian H, Wang L, Guo B, Wang Y, et al. SCN1A Mutation—Beyond Dravet Syndrome: A Systematic Review and Narrative Synthesis. *Front Neurol* 2021;12:743726. <https://doi.org/10.3389/FNEUR.2021.743726/FULL>.
- [138] Sugawara T, Mazaki-Miyazaki E, Fukushima K, Shimomura J, Fujiwara T, Hamano S, et al. Frequent mutations of SCN1A in severe myoclonic epilepsy in infancy. *Neurology* 2002;58:1122–4. <https://doi.org/10.1212/WNL.58.7.1122>.
- [139] Shbarou R, Mikati MA. The Expanding Clinical Spectrum of Genetic Pediatric Epileptic Encephalopathies. *Semin Pediatr Neurol* 2016;23:134–42. <https://doi.org/10.1016/J.SPEN.2016.06.002>.
- [140] SCN1A Seizure Disorders - GeneReviews® - NCBI Bookshelf n.d. <https://www.ncbi.nlm.nih.gov/books/NBK1318/> (accessed March 11, 2025).
- [141] Shi X, Wang J, Kurahashi H, Ishii A, Higurashi N, Kaneko S, et al. On the likelihood of SCN1A microdeletions or duplications in Dravet syndrome with missense mutation. *Brain Dev* 2012;34:617–9. <https://doi.org/10.1016/J.BRAINDEV.2011.11.005>.
- [142] Sun H, Zhang Y, Liu X, Ma X, Yang Z, Qin J, et al. Analysis of SCN1A mutation and parental origin in patients with Dravet syndrome. *Journal of Human Genetics* 2010 55:7 2010;55:421–7. <https://doi.org/10.1038/jhg.2010.39>.
- [143] Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* 2009;10:741–54. <https://doi.org/10.1038/NRM2777>.
- [144] Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 2013;15:733–47. <https://doi.org/10.1038/GIM.2013.92>.
- [145] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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;17:405–24. <https://doi.org/10.1038/GIM.2015.30>.
- [146] Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 2013;15:565–74. <https://doi.org/10.1038/GIM.2013.73>.
- [147] Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, et al. Sherlock: a comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med* 2017;19:1105–17. <https://doi.org/10.1038/GIM.2017.37>.
- [148] Glass IA, Swindlehurst CA, Aitken DA, McCrea W, Boyd E. Interstitial deletion of the long arm of chromosome 2 with normal levels of isocitrate dehydrogenase. *J Med Genet* 1989;26:127–30. <https://doi.org/10.1136/JMG.26.2.127>.
- [149] Lewis H, Samanta D, Örsell JL, Bosanko KA, Rowell A, Jones M, et al. Epilepsy and Electroencephalographic Abnormalities in SATB2-Associated Syndrome. *Pediatr Neurol* 2020;112:94. <https://doi.org/10.1016/J.PEDIATRNEUROL.2020.04.006>.
- [150] Larti F, Kahrizi K, Musante L, Hu H, Papari E, Fattahi Z, et al. A defect in the CLIP1 gene (CLIP-170) can cause autosomal recessive intellectual disability. *European Journal of Human Genetics* 2015 23:3 2014;23:331–6. <https://doi.org/10.1038/ejhg.2014.13>.

- [151] Calame DG, Guo T, Wang C, Garrett L, Jolly A, Dawood M, et al. Monoallelic variation in DDX9, the gene encoding the DExH-box helicase DDX9, underlies neurodevelopment disorders and Charcot-Marie-Tooth disease. *Am J Hum Genet* 2023;110:1394–413. <https://doi.org/10.1016/j.ajhg.2023.06.013>.
- [152] Van Reeuwijk J, Janssen M, Van Den Elzen C, Beltran-Valero De Bernabé D, Sabatelli P, Merlini L, et al. POMT2 mutations cause alpha-dystroglycan hypoglycosylation and Walker-Warburg syndrome. *J Med Genet* 2005;42:907–12. <https://doi.org/10.1136/JMG.2005.031963>.
- [153] Yanagisawa A, Bouchet C, Van Den Bergh PYK, Cuisset JM, Viollet L, Leturcq F, et al. New POMT2 mutations causing congenital muscular dystrophy: Identification of a founder mutation. *Neurology* 2007;69:1254–60. <https://doi.org/10.1212/01.WNL.0000268489.60809.C4>.
- [154] Conti V, Carabalona A, Pallesi-Pocachard E, Parrini E, Leventer RJ, Buhler E, et al. Periventricular heterotopia in 6q terminal deletion syndrome: role of the C6orf70 gene. *Brain* 2013;136:3378–94. <https://doi.org/10.1093/BRAIN/AWT249>.
- [155] Stiles AR, Simon MT, Stover A, Eftekharian S, Khanlou N, Wang HL, et al. Mutations in TFAM, encoding mitochondrial transcription factor A, cause neonatal liver failure associated with mtDNA depletion. *Mol Genet Metab* 2016;119:91–9. <https://doi.org/10.1016/j.ymgme.2016.07.001>.
- [156] Kang I, Chu CT, Kaufman BA. The mitochondrial transcription factor TFAM in neurodegeneration: Emerging evidence and mechanisms. *FEBS Lett* 2018;592:793. <https://doi.org/10.1002/1873-3468.12989>.
- [157] Oda H, Nakagawa K, Abe J, Awaya T, Funabiki M, Hijikata A, et al. Aicardi-Goutières Syndrome Is Caused by IFIH1 Mutations. *Am J Hum Genet* 2014;95:121. <https://doi.org/10.1016/j.ajhg.2014.06.007>.
- [158] Cananzi M, Wohler E, Marzollo A, Colavito D, You J, Jing H, et al. IFIH1 loss-of-function variants contribute to very early-onset inflammatory bowel disease. *Hum Genet* 2021;140:1299–312. <https://doi.org/10.1007/S00439-021-02300-4>.
- [159] Rutsch F, Macdougall M, Lu C, Buers I, Mamaeva O, Nitschke Y, et al. A specific IFIH1 gain-of-function mutation causes singleton-merten syndrome. *Am J Hum Genet* 2015;96:275–82. <https://doi.org/10.1016/j.ajhg.2014.12.014>.

APPENDIX

Table I. Patient 1 real-time PCR results.

Target	Sample	Target C _T	HK C _T	Ratio ¹⁵
NRXN1 Exon 1	<i>Patient 1</i>	22,66940117	21,96571287	0,47157188
	CTRL1	22,14614105	22,45970154	0,954487617
	CTRL3	22,54939938	22,93016148	
	CTRL4	22,23366547	22,85890388	1,028918105
NRXN1 Exon 2	<i>Patient 1</i>	22,81145922	21,96571287	0,432554635
	CTRL1	22,15573502	22,45970154	0,959707819
	CTRL3	22,5882384	22,93016148	0,985292314
	CTRL4	22,49560452	22,85890388	
NRXN1 Exon 3	<i>Patient 1</i>	23,54020882	21,96571287	0,450956516
	CTRL1	22,79715252	22,45970154	1,062973184
	CTRL2	23,10631657	22,68076038	
	CTRL3	23,2683382	22,93016148	1,062438591
	CTRL 4	23,27815914	22,85890388	1,004377021
NRXN1 Exon 4	<i>Patient 1</i>	20,80791855	21,5175066	1,03981912
	CTRL1	21,43440946	22,21858406	0,99849115
	CTRL2	21,44180298	22,08670171	0,9942243
	CTRL4	21,35779095	22,01104641	

Table II. Patient 3 real-time PCR results.

Target	Sample	Target C _T	HK C _T	Ratio ¹⁵
TMTC3 Exon 4	<i>Patient 3</i>	21,72789001	22,27228874	1,005456256
	CTRL1	23,18676567	23,72333926	1,000017452
	CTRL3	23,849624	24,38617241	
	CTRL4	24,47357273	24,99249554	0,987857188
TMTC3 Exon 5	<i>Patient 3</i>	24,10319805	22,74240685	0,382731872
	CTRL1	22,65528552	22,68242607	1,001621691
	CTRL3	22,00226307	22,21333445	1,137816998
	CTRL4	22,53320599	22,55800883	
TMTC3 Exon 7	<i>Patient 3</i>	23,48530324	23,14660263	0,513630501
	CTRL1	22,0613718	22,68273799	0,999216758
	CTRL2	21,66541386	22,28791046	
	CTRL4	22,23635292	22,84281826	0,988949484
TMTC3 Exon 10	<i>Patient 3</i>	23,79228274	22,74240685	0,436034368
	CTRL2	22,06981277	22,17995644	0,974364391
	CTRL3	21,97803752	22,24333445	1,084993643
	CTRL4	22,41039848	22,55800883	
TMTC3 Exon 11	<i>Patient 3</i>	22,1848774	22,58292395	1,05980683
	CTRL1	23,64996529	23,95333926	0,992492943
	CTRL3	23,84478283	24,16707481	1,005593173
	CTRL4	24,67825031	24,99249554	

¹⁵ In the empty box, Ratio=1

Table III. Patient 4 real-time PCR results.

Target	Sample	Target C_T	HK C_T	Ratio¹⁵
SCN1A Exon 14	<i>Patient 4</i>	22,2440815	22,56068207	1,06800198
	<i>Patient 4 M¹⁶</i>	21,97550011	22,28438187	1,062303138
	<i>Patient 4 P¹⁷</i>	22,16216373	22,37142754	0,991426381
	CTRL1	22,30264378	22,71417522	1,066630505
	CTRL 2	21,9685936	22,19027985	0,935114735
	CTRL 3	21,48439217	21,80286312	
SCN1A Exon 15	<i>Patient 4</i>	21,44263363	22,62647438	1,40167429
	<i>Patient 4 M</i>	21,85544872	22,54017448	0,991741585
	<i>Patient 4 P</i>	21,91610648	22,62330334	1,007309655
	CTRL1	21,17118168	21,86787128	
	CTRL 2	21,78219891	22,52053134	1,029285227
	CTRL 3	21,5005312	22,21728069	1,014001573
SCN1A Exon 19	<i>Patient 4</i>	22,53180504	22,62647438	1,507918259
	<i>Patient 4 M</i>	22,71901894	22,30566692	1,060347297
	<i>Patient 4 P</i>	22,82464155	22,2983873	0,980530652
	CTRL1	22,50434494	21,9956398	0,992530777
	CTRL 2	23,01842022	22,52053134	
	CTRL 3	22,4751683	22,02215385	1,031593389
SCN1A Exon 20	<i>Patient 4</i>	22,08481598	22,655	1,026849613
	<i>Patient 4 M</i>	21,90147877	22,41478729	0,987110288
	<i>Patient 4 P</i>	21,79963684	22,33040962	0,999132171
	CTRL1	21,80132548	22,43955421	1,076391868
	CTRL 2	21,4603672	22,0453064	1,037357994
	CTRL 4	21,6807003	22,21272564	

¹⁶ Maternal sample

¹⁷ Paternal sample

Table IV. Pathogenicity criteria adopted for the interpretation of genetic variants [144–147].

Classification	Pathogenicity criteria according to ACMG			Literature analysis	
Pathogenic	A	+ at least one between B C D	-	Described in the literature	Not described in literature = potentially pathogenic
	A	+ at least two between F G H I	-		
	A	+ one between F G H I	+ one between J K L		
	A	+ at least two between J K L	-		
	At least two between B C D E	-	-		
	One between B C D E	+ at least three between F G H I	-		
	One between B C D E	+ at least two between F G H I	+ at least two between J K		
	One between B C D E	+ one between F G H I	+ all J K L		
Potentially pathogenic	A	+ one between F G H I	-		
	One between B C D E	+ at least one between F G H I	-		
	One between B C D E	+ at least two between J K L	-		
	At least three between F G H I	-	-		
	Two between F G H I	+ at least two between J K L	-		
	One between F G H I	+ all J K L	-		

Pathogenicity criteria:

- A: Loss-of-function variants (LOF);
- B: Variant located in the same codon and resulting in the same substitution amino acid substitution known to be pathogenic;
- C: De novo variant confirmed on the proband's parents;
- D: Variant for which functional in vitro studies have been performed demonstrating harmful effect on the protein;
- E: Variants for which there is a known prevalence, in terms of frequency, in individuals affected;
- F: Variant located in a hot spot or in a functional protein zone biochemically characterized
- G: Variant absent from the population database or present with a frequency compatible or lower than the incidence of the associated phenotype;
- H: For recessive phenotypes, variant located in trans to a second variant in the same gene;
- I: Variant located in the same codon in which a pathogenic variant is described;
- J: Variant that co-segregates with the pathological phenotype;
- K: Missense variant in a highly conserved gene;
- L: Variants for which in silico computational analysis indicates with high probability a damaging effect on protein structure or activity.

Benign and potentially benign variant

Variant for which one of the following conditions is fulfilled:

- It has an allele frequency higher than the disease frequency in the reference population;
- It is reported in the literature as not being associated with the pathological phenotype;
- It is found in cis with a pathogenic variant (in the case of recessive phenotypes);
- It is found in an unaffected relative (in cases of dominant phenotypes in the absence of incomplete penetrance).

Variant of uncertain significance

Variant for which one or more of the following conditions are fulfilled:

- Not classifiable as pathogenic, potentially pathogenic, potentially benign or benign;
- It is a variant reported in a single affected individual in the absence of segregation and satisfactory clinical information;
- It is a variant for which conflicting evidence is reported;
- It is a variant located at a position adjacent to the canonical splicing site, when similar variants are described in dedicated databases.

Table V. Affected genes, variants with associated VarSeq metrics, and genotype-phenotype associations of excluded patients.

Patient	Gene Exon	Span bp	CNV state	Avg Z Score	Avg Ratio	p-value	Gene associated disorder	Patient synopsis
1	<i>SATB2</i> Exon 6	124	Heterozygous Deletion	-3,30	0,58	6,91E-04	Glass syndrome (OMIM: 612313): autosomal dominant, all patients show intellectual disability with absent or limited speech development, behavioural problems, and dysmorphic facial features [148,149]. Approximately 17% of patients have also seizures [149]. Deletion of single exon 6 have never been reported.	Suspected cortical dysplasia, with normal pshyco-motor development.
2	<i>CLIP1</i> Exon 20	53	Duplication	3,46	1,50	3,81E-04	To date, there are only a few individuals with recessive nonsense variants [150]. The clinical phenotype is characterized by developmental delay and dysmorphism [150]. Some individuals also have seizures, but the pathogenic mechanism is still unclear.	Headaches since the age of 3, with critical episodes occurring at the age of 14. Normal brain MRI. Suspected hereditary epilepsy (father had similar symptoms). VarSeq bioinformatic analysis of the father's sample

								did not reveal any CNV shared with the son, including the exon 20 duplication of the <i>CLPI</i> gene.
3	<i>DHX9</i> Exon 19	74	Duplication	3,48	1,47	6,04E-04	Intellectual developmental disorder (OMIM: 620988): Autosomal dominant, loss of function mutations. Characterized by developmental delay, mild to severely impaired intellectual development, accompanied by neuropsychiatric disorders [151]. Seizures observed in few individuals which also present ID and microcephaly [151].	Febrile convulsions with focal seizure clusters. Normal psychomotor and neurological development. Normal brain MRI.
4	<i>POMT2</i> Exon. 13	152	Heterozygous deletion	-4,51	0,57	1,11E+09	Muscular dystrophy-dystroglycanopathy (congenital with brain and eye anomalies), type A, 2 (OMIM: 613150), Type B, 2 (OMIM: 613156), Muscular dystrophy-dystroglycanopathy (limb-girdle), type C, 2 (OMIM: 613158). Recessive missense and nonsense mutations can lead to three forms of muscular dystrophy-dystroglycanopathy (MDDG), autosomal recessive diseases [152,153].	Suspected cortical dysplasia, with normal psychomotor development.
6	<i>ERMARD</i> Exon 7	137	Duplication	3,76	1,43	3,36E-04	Periventricular nodular heterotopia type 6 (PVNH6) (OMIM: 615544), a disorder resulting from a defect in the pattern of neuronal migration [154]. PVNH6 results in	Suspected focal epilepsy.

							delayed psychomotor development, delayed speech, strabismus, and onset of seizures with hypsarrhythmia in early infancy [154].	
8	<i>TFAM</i> Exon 2	119	Duplication	4,37	1,45	2,41E-04	Mitochondrial DNA depletion syndrome 15 (hepatocerebral type) (MTDPS) (OMIM: 617156). MTDPS is a disease caused by defects in mitochondrial DNA maintenance and leads to liver failure and neurological complications during infancy [155]. <i>TFAM</i> have been also associated to neurodegenerative disorders such as Alzheimer, Huntington, Parkinson and others [156].	Generalized idiopathic epilepsy. Normal psychomotor development, no dysmorphic features. EEG shows epileptiform abnormalities in bilateral posterior regions.
.9	<i>IFIH1</i> Exon 13	152	Duplication	3,63	1,46	2,92E-04	Aicardi-Goutieres syndrome 7 (AGS7) (OMIM: 615846), Immunodeficiency 95 (IMD95) (OMIM: 619773) and Singleton-Merten syndrome 1 (SGMRT) (OMIM: 182250). AGS7 is an autosomal dominant inflammatory disorder characterized by severe neurologic impairment [157]. IMD95 is an autosomal recessive disorder characterized predominantly by the onset of recurrent and severe viral respiratory infections in infancy or early childhood [158]. SGMRT is an uncommon	Epileptic encephalopathy. From 4 months of age, cluster epileptic seizures. Normal development. MRI: no abnormalities in cortical development.

							autosomal dominant disorder characterize by abnormalities of blood vessels, teeth, and bones [159].	
--	--	--	--	--	--	--	--	--