Technische Universität München
TUM School of Medicine and Health

Identification of novel genetic variation causing epilepsy using linkage analysis and next generation sequencing techniques

Monika Valeria Schlögl

Vollständiger Abdruck der von der TUM School of Medicine and Health der Technischen Universität München zur Erlangung einer Doktorin der Medizin genehmigten Dissertation.

Vorsitz: apl. Prof. Dr. Bernhard Haslinger

Prüfer der Dissertation:

1. Prof. Dr. Thomas Meitinger
2. Prof. Dr. Marcus Deschauer

Die Dissertation wurde am 20.03.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Medicine and Health am 06.02.2024 angenommen.

Contents

I. List of Figures

II. List of Tables

III. List of Abbreviation

ACMG	American College of Medical Genetics and Genomics
AD	Alzheimer's dementia
BP	Base pair
CI	Confidence interval
cM	Centimorgan
cMRI	Cerebral magnetic resonance imaging
DNA	Deoxyribonucleic acid
dNTP	Desoxynukleosidtriphosphate
ECM	Extracellular matrix
EEG	Electroencephalogram
GABA	Gamma-aminobutyric acid
GASH/Sal	Genetic audiogenic seizure hamster from Salamanca
GEE	Generalized epileptiform activity
gnomAD	Genome Aggregation Database
GTEx	Genotype-tissue expression
GWAS	Genome-wide association study
HGMD	Human Gene Mutation Database
ID	Identifier
ILAE	International League Against Epilepsy
Indel	Insertion or deletion
LOD	Logarithm of the odds
LOEUF	Loss-of-function observed/expected upper bound fraction
OD	Optical density
MAF	Minor Allele Frequency
Mb	Megabases
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
mtDNA	Mitochondrial DNA
mTOR	Mechanistic target of rapamycin
NCBI	National Center for Biotechnology Information
NGS	Next-generation sequencing
OMIM	Online Mendelian Inheritance in Man

OR	Odds Ratio
PCR	Polymerase chain reaction
pLI	Probability for loss-of-function intolerance
RNA	Ribonucleic acid
SE	Status epilepticus
SNP	Single nucleotide polymorphism
TBI	Traumatic brain injury
TLE	Temporal lobe epilepsy
TSC	Tuberous sclerosis
TUM	Technische Universität München
UTR	Untranslated regions
WES	Whole exome sequencing

1 Introduction

1.1 Definition

Epilepsy is a generic term for neurological disorders characterized by hyperexcitability of neurons of the cerebral cortex or by a lack of inhibition of excitation. (Stafstrom & Carmant, 2015) It describes a brain condition characterized by a persistent predisposition to epileptic seizures. (Stafstrom & Carmant, 2015) This is manifested by seizure-like synchronized neuronal discharges. Depending on whether the potential discharges affect only parts of one hemisphere or both hemispheres of the brain, a distinction is made between focal and generalized seizures. (Fisher et al., 2014)

According to the definition of the ILAE (International League Against Epilepsy) 2014 (Fisher et al., 2014), epilepsy is present, if one of the following three conditions applies: First, at least two unprovoked seizures must occur more than 24 hours apart. Second, epilepsy can be diagnosed when an unprovoked or reflex seizure occurs and there is an increased likelihood of future seizures within the next ten years, corresponding to the general risk of recurrence of at least 60%. Third, epilepsy can be diagnosed when an epilepsy syndrome is present. (Fisher et al., 2014)

The terms “seizure” and “epilepsy” must be distinguished. An epileptic seizure is an isolated clinical event that is a transient occurrence of symptoms and/or signs as a result of abnormal hyperactivity or synchronization of neuronal activity in the cerebrum. (Fisher et al., 2014) Epilepsy, on the other hand, is a disease characterized by a persistent tendency to have epileptic seizures and by the cognitive, psychological, social, and neurobiological consequences of the disease. The diagnosis of epilepsy requires the occurrence of at least one epileptic seizure. (Fisher et al., 2014)

1.2 Epidemiology

About 2 to 4% of all people experience a single isolated epileptic seizure during their lifetime (World Health Organization et al., 2005) and about 1% of the population develops epilepsy. Two age peaks can be distinguished in early childhood and later life. (Hauser et al., 1993)

Adults are more likely to suffer from focal epilepsies, mostly due to cerebrovascular events and neoplasms. (Lemke, 2019) In children, non-genetic syndromes such as absence epilepsies, Rolando epilepsy, juvenile myoclonic epilepsy (JME) predominate on the one

hand, and on the other hand, congenital genetic causes are responsible for the development of generalized epilepsies in addition to (multi)focal epilepsies. (Lemke, 2019) Genetic epilepsies are often accompanied by other psychiatric and/or neurological problems. Because of the great clinical and genetic heterogeneity of epilepsies, etiologic-genetic workup of these patients has long been very difficult. Antiepileptic therapy was and still is mainly based on empirical, partly individual experience with the respective phenotype. (Lemke, 2019)

1.3 Pathophysiology

Basically, two main types of seizures and epilepsies are distinguished: epilepsies of focal origin and primary (idiopathic) generalized epilepsies. (Scheffer et al., 2017) In epilepsies of focal origin, the seizure occurs in an epileptogenic pacemaker zone, usually associated with an epileptogenic lesion. From there, the seizure spreads through synaptically connected cortical structures. In contrast, in generalized seizures and epilepsies, epileptogenic potentials develop simultaneously in both hemispheres. (Scheffer et al., 2017) Cortico-thalamic interactions seem to be the main cause. (van Luijtelaar & Sitnikova, 2006)

The complex mechanisms of the onset and development of epilepsy are still not fully understood. The hyperexcitability of neurons and simultaneous abnormal electrical activity of larger neuronal networks leads to the occurrence of epileptic seizures. It is assumed that the imbalance between excitation (increased) and inhibition (decreased) in these neuronal networks leads to epileptic seizures. (Barker-Haliski & White, 2015; Engelborghs et al., 2000; Filakovszky et al., 1999; Vezzani, 2014)

Possible causes include altered membrane properties of neurons, e.g., due to defective ion channels, or abnormalities in excitation transmission between neurons due to an imbalance or incorrect composition of neurotransmitters. Biochemically, epilepsy is probably caused by a disequilibrium between increased excitatory glutamate or aspartate action and decreased inhibitory GABA (gamma-aminobutyric acid) effect on neurons. (Barker-Haliski & White, 2015; Engelborghs et al., 2000; Filakovszky et al., 1999; Vezzani, 2014)

Certain types of seizures, such as absences, may also be triggered by excessive inhibition. (Hosford & Wang, 1997) Therefore, substances that increase GABA concentration in the synaptic cleft (e.g., vigabatrin) may be effective against seizures of focal origin, but may also lead to worsening of absences. (Hosford & Wang, 1997)

Smaller neuronal cell assemblies that are not involved in physiological functions may also act as initial pacemakers of epileptic seizures, such as in cortical dysplasia, and thus may be part

of the pathophysiology. These may be partly responsible for drug resistance. (Barkovich et al., 2015; Blümcke et al., 2011) Pacemaker cells with very high discharge rates (over 1,000 Hz) could play an important role here. (Usui et al., 2010)

The medio-temporal structures of the hippocampus represent a special feature in the pathogenesis of epilepsy. (Routbort et al., 1999) Early trauma or epileptic seizure activity leads to secondary changes with neuronal cell death in this area. At the same time, regenerative mechanisms such as moss fiber sprouting and neuronal neogenesis from progenitor cells begin to build and restore the physiological network. This can contribute to hyperexcitability. (Routbort et al., 1999) Thus, after a seizure, there is increased formation of neurons in the hippocampus. An acute seizure is usually associated with increased neurogenesis and migration of the newly generated neurons to ectopic regions, which in turn may contribute to the pathogenesis of epilepsy. (Jessberger et al., 2005; Kuruba et al., 2009)

1.4 Classification of epilepsy

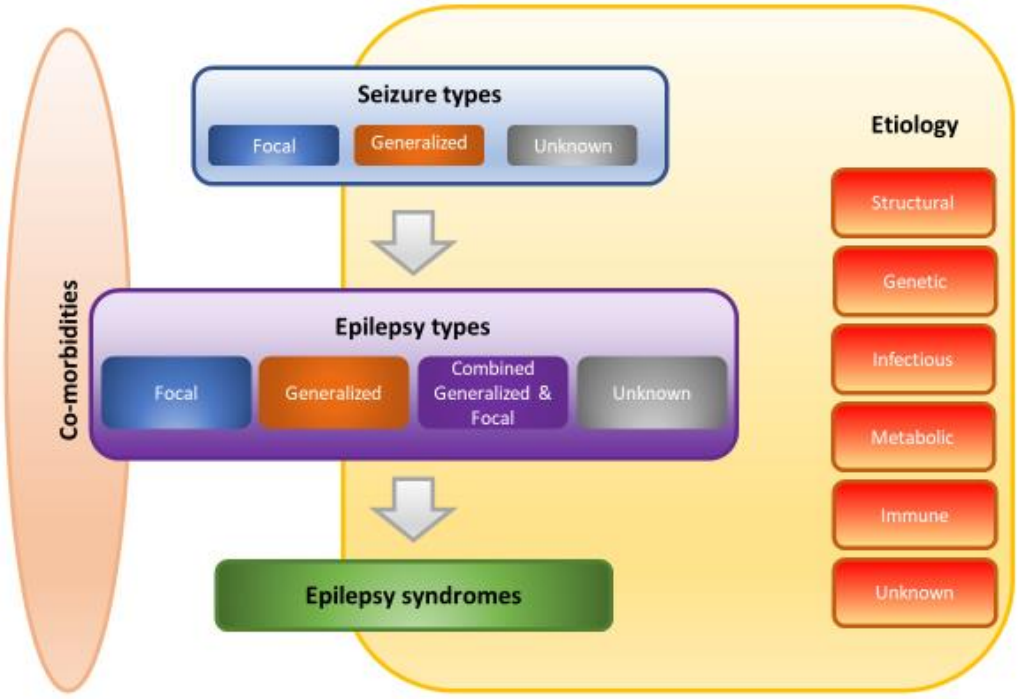

Figure 1.1.: Framework for classification of epilepsy (Modified from: ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology, Ingrid E. Scheffer, 2017 (Scheffer et al., 2017))

According to the ILAE classification of epilepsies of 2017, epilepsies are classified on three levels (Seizure types, Epilepsy types and Epilepsy syndromes, see Figure 1.1). (Scheffer et

al., 2017) This multi-level approach offers the clinician various diagnostic options for a specific case. Optimally, classification is performed at all three levels. (Scheffer et al., 2017)

1.4.1 Seizure classification

Epileptic seizures, as mentioned earlier, are divided into focal seizures, in which overexcitation is confined to a single area, generalized seizures, in which seizure signs extend across both hemispheres, and seizures of unknown spread (see Figure 1.2). (Scheffer et al., 2017)

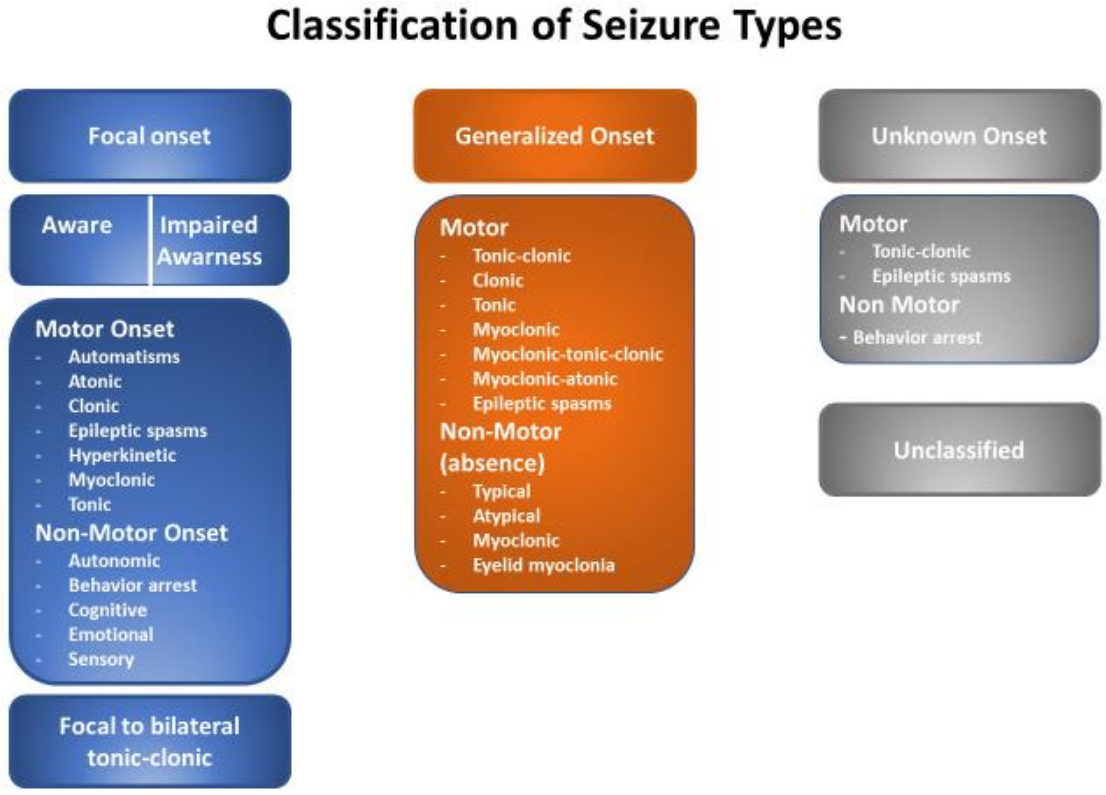

Figure 1.2.: Classification of Seizure Types Basic Version ILAE 2017 (Modified from: Instruction manual for the ILAE 2017 operational classification of seizure types, Robert S. Fisher, 2018, (Fisher et al., 2018))

Seizure classification first defines whether the first manifestation is a focal or generalized seizure. (Fisher et al., 2018) Classification to an unknown onset occurs when there is a seizure in which the time of onset has either been missed or cannot be identified. (Fisher et al., 2018) Certain symptoms and signs that occur during a focal seizure can provide information about the brain areas responsible, which may be involved in the onset and propagation of the seizure. (Fisher et al., 2018)

For focal seizures, a distinction is made between whether consciousness is maintained, i.e., whether the person is aware of himself and his surroundings during the seizure, or whether unconsciousness occurs. (Fisher et al., 2018) Seizures should be classified according to the first, dominant motor or nonmotor feature at onset, except in the case of a focal seizure, in which pausing is the prominent sign. The seizure form "focal to bilateral tonic-clonic" characterizes the pattern of the seizure and is therefore less a separate seizure form. In this case the seizure begins focally and generalizes secondarily. (Fisher et al., 2018)
A generalized seizure originates in bilaterally distributed networks and spreads rapidly. The entire cortex may be affected, alternatively only cortical and subcortical structures. (Williamson et al., 2009) Generalized seizures are divided into motor and non-motor seizures (i.e., atonic, static, myoclonic) each of which may also manifest asymmetrically. (Fisher et al., 2018) A seizure with unknown onset can be divided into a motor and non-motor seizure. The category "unclassified" includes seizures that cannot be further classified because there is insufficient information or the seizure cannot be classified into other categories. (Fisher et al., 2018)

1.4.2 Epilepsy classification

To assign a patient to the second level "epilepsy types", the patient must have been diagnosed with epilepsy in the sense of the three criteria as defined in 2017 (see 1.1). A focal epilepsy includes unifocal, multifocal (starting in different networks) and hemispheric seizures. (Scheffer et al., 2017) These can be divided into the same groups as listed in section 1.4.1. Patients with focal epilepsy have focal seizure types and can therefore have localized seizures as well as seizures across both hemispheres. Thus, unifocal epilepsy may be associated with generalized seizures. Patients with generalized epilepsy typically show generalized spike-wave complexes on the EEG (electroencephalogram) and may have different seizure types, such as absences, myoclonic, tonic, atonic and tonic-clonic seizures. (Scheffer et al., 2017) The clinical diagnosis of generalized epilepsy is supported by typical epileptogenic potentials in the EEG. The group of "combined generalized and focal" epilepsies includes those patients who suffer from both generalized and focal seizures. This diagnosis is also clinically established and supported by an EEG. (Scheffer et al., 2017) The term "unclassified" is used for those cases in which clinically an epilepsy is present, but due to missing information (e.g.: EEG) it cannot be determined more precisely whether it is a focal or generalized epilepsy. (Scheffer et al., 2017)

1.4.3 Epilepsy syndromes

The term syndrome is used to describe a disease with a typical combination of symptoms and features (especially seizure forms). (Camfield & Camfield, 2002) When differentiating epileptic syndromes, the cause of the disease plays less of a role than common characteristics. These include, in addition to the type and course of seizures, many other factors, such as age at onset or typical EEG abnormalities and / or imaging findings. (Camfield & Camfield, 2002) Diagnosis is often related to age at onset and remission of epilepsy. Prognosis, temporal variations, and seizure precipitating factors also play important roles in epilepsy syndromes. (Dreifuss et al., 1985; Epilepsy, 1989; Scheffer et al., 2017)

Classification of syndromes is useful to identify underlying causes and appropriate antiepileptic therapy. (Camfield & Camfield, 2002) Epilepsy syndromes are more frequently diagnosed in children and infants. (Neligan et al., 2012)

These syndromes cover a very broad spectrum, ranging from barely affected (Rolandic epilepsy or childhood absence epilepsy) to severely affected patients (Lennox-Gastaut syndrome, West syndrome and Dravet syndrome). Some important and more common epilepsy syndromes are explained in more detail in the following sections.

Rolandic epilepsy is usually a self-limiting epilepsy that reaches its peak age at seven to eight years of age. (Danielsson & Petermann, 2009) The seizures, which usually occur at night, are brief and hemifacial and may progress from focal to bilateral tonic-clonic seizures. Neurology, imaging, and cognition are usually unremarkable. EEG shows centrotemporal sharp waves with high amplitude activated by sleep and drowsiness. The epilepsy usually resolves by late adolescence without permanent residual effects. (Gündüz et al., 1999; Lindgren et al., 2004)

Childhood absence epilepsy is a genetic or idiopathic generalized epilepsy that occurs in otherwise healthy children, primarily by the age of six to seven years. (Tenney & Glauser, 2013) It is characterized by brief periods of absence seizures that can occur very frequently per day and are often triggered by hyperventilation. Characteristic features include 3-Hz spike-and-wave discharges on the EEG in addition to unremarkable development, neurology, and cognition. (Galli et al., 2018; Sadleir et al., 2006)

The syndromes listed below belong to the group of epileptic encephalopathies. In the past, it was thought that epileptic activity itself may lead to cognitive and neurological disorders. Nowadays, it is rather assumed that both occur independently of each other. So the encephalopathy exists in addition to epilepsy and is not necessarily a consequence of it.

(Scheffer & Liao, 2020) The predominant symptoms of epileptic encephalopathies are diffuse brain dysfunction, severe cognitive impairment and frequent seizures resistant to therapy. (Jain et al., 2013)

One of these epileptic encephalopathies is West syndrome, which usually occurs around the first year of life. Thereby flash-like myoclonias or tonic flexion spasms occur (Dulac, 2001) and the EEG often shows hypsarrhythmia (high delta waves with single irregular spikes and sharp waves). (Hrachovy & Frost Jr, 2003) Prognosis is usually poor, as global developmental impairment (with or without regression) is observed. (Appleton, 2001) Common causes include prenatal developmental disorders, hypoxic-ischemic encephalopathies, tuberous sclerosis (TSC), or metabolic diseases. Associated genes include *ARX*, *CDKL5*, *SPTAN1*, and *STXBP1*. (Dulac, 2001)

Dravet syndrome also belongs to the group of epileptic encephalopathies and typically begins in the first year of life, with the first seizure being febrile in most cases. (Dravet, 2011) Variants in the sodium channel gene *SCN1A* are the cause in 70-80% of all patients. Children experience focal (usually hemiclonic) and generalized tonic-clonic seizures, among others, which are usually refractory to therapy. As the disease progresses, affected individuals develop cognitive and behavioral disturbances, and ataxias and pyramidal signs have also been described in many patients. (Dravet, 2011)

The last syndrome mentioned here that belongs to this group is Lennox-Gastaut syndrome. It occurs in early childhood and is characterized by a coexistence of different seizure patterns (tonic, astatic and myoclonic seizures or atypical absences). (Epilepsy, 2020) Developmental delays and cognitive and behavioral abnormalities are common, as is status epilepticus (SE). Mainly structural brain abnormalities are causally responsible, but genetic predisposition is also suspected. (Epilepsy, 2020)

In summary, epilepsy syndromes are very broad and can range from mild self-limiting seizures to severe epileptic encephalopathies with early death.

1.5 Etiology

1.5.1 Non-genetic etiology

There are numerous reasons for epilepsy or seizures besides altered genes. When the first epileptic seizure occurs, the cause should be determined directly. A magnetic resonance imaging (MRI) scan of the brain (cMRI) is often performed to identify or rule out a possible

structural lesion as a causative factor. (Scheffer et al., 2017) In the case of Rolando or absence epilepsies, for example, there is no need for MRI, because the seizure semiology and the EEG are clear. If no abnormality is found, genetic, infectious, metabolic, or immune-mediated changes may also be the origin of the epilepsy (see Figure 1.1). (Scheffer et al., 2017) In many cases, no clear etiology can be found. The epilepsy is then termed idiopathic, and this is most common in epilepsy syndromes. About 75% of all epilepsies are of non-genetic origin (Syvertsen et al., 2015) and therefore some non-genetic causes will be discussed in more detail below. According to Syvertsen et. al. (2015), the etiology could not be determined in another 5% due to lack of information.

Structural abnormalities in the brain significantly increase the risk of epilepsy. (Berg et al., 2010) Neuronal changes, scars, and malformations of brain structures prevent electrical potentials from propagating physiologically, which increases seizure susceptibility. (Herman, 2002) A further explanation is that brain damage also affects GLT-1 transporters (astrocyte glutamate transporters) resulting in glutamate accumulation in the nervous system, which also leads to increased susceptibility because glutamate is an excitatory transmitter. (Tanaka et al., 1997) Imaging techniques are then used to detect these anomalies and, in conjunction with the clinical findings, identify them as the likely cause of the patient's seizures. The abnormality may occur, for example, after a stroke, bleeding, trauma, infection, but may also be genetic. (Scheffer et al., 2017) In the latter case, seizures are often the result of a disorder of cortical development. Although such malformations have genetic causes, epilepsy is caused by the structural correlate. In summary, there are multiple correlations between epilepsy and structural lesions. (Scheffer et al., 2017)

Excluding febrile seizures, infections are the most common cause of non-genetic epilepsies worldwide. (Vezzani et al., 2016) This means that epilepsy is a direct consequence of a cerebral infection. Infections of the central nervous system can lead to acute symptomatic seizures that are related to the initial infection and can also cause epilepsies. (Scheffer et al., 2017) Sometimes the infections show a structural correlation, but the epilepsy is mainly caused by the infection process. Examples include tuberculosis, HIV (Human Immunodeficiency Virus), or cerebral malaria. (Scheffer et al., 2017)

An immune-mediated etiology can be assumed if the disease is a direct consequence of immune system dysregulation. (Scheffer et al., 2017) Autoimmune encephalopathies with seizures should be distinguished from those without seizures (for example, multiple sclerosis). Epileptic seizures can be triggered in association with autoimmune-mediated inflammation and, consequently, disruption of the blood-brain barrier (Specchio et al., 2010) of the nervous

system. (Scheffer et al., 2017) An example of autoimmune-mediated epilepsy caused by antibodies is the limbic encephalitis. (Malter et al., 2010) In this case, antibodies target proteins of voltage-gated potassium channels, forming an antibody complex. This causes limbic encephalitis, which can lead to a loss of short-term memory accompanied by focal sensory and motor seizures and psychiatric symptoms, etc. (Bien & Elger, 2007; Malter et al., 2010) West syndrome has also been associated with an autoimmune-mediated etiology, as steroid therapy in many cases leads to significant improvement of epilepsy in these patients, even if no inflammatory changes could be detected so far (no cells in the cerebrospinal fluid, no specific antibodies, no protein elevation, etc.). (Geva-Dayan et al., 2012; Palace & Lang, 2000)

Metabolic disorders are mostly of genetic origin. These include congenital metabolic disorders. The inborn errors of metabolism are a heterogeneous group that can be inherited or result from spontaneous mutations. These disorders involve disturbances in metabolic pathways such as Gaucher disease or phenylketonuria. (Jeanmonod et al., 2021; Scheffer et al., 2017) Metabolic derailment sometimes causes seizures and sometimes brain damage occurs primarily. Non-genetic metabolic disorders, on the other hand, include seizures in hypoglycemia. (Jeanmonod et al., 2021; Scheffer et al., 2017)

1.5.2 Genetic etiology

Most forms of epilepsy have a strong genetic etiology, as indicated by the frequent positive family history. (Zimprich et al., 2015) It is estimated that approximately 20% (Syvertsen et al., 2015) of all epilepsies in adults and children are caused by monogenic (about 1-2% of all epilepsy cases) or polygenic (about 18-19% of all epilepsy cases) genetic alterations. (Hildebrand et al., 2013; Myers & Mefford, 2015) Heritability estimates the extent of variation in a phenotypic trait in a population that is attributable to genetic variation among individuals (Visscher et al., 2008) and the value is highly dependent on the experimental setting. Heritability has been estimated to be approximately 80% in twin studies by comparing concordance rates in monozygotic and dizygotic twins, whose ages ranged from twelve to 41 years. (Kjeldsen et al., 2001) They suffered from epilepsy, febrile seizures, other seizure types or staring spells. (Kjeldsen et al., 2001) Family studies show that the relative risk for epilepsy increases three- to sixfold when one parent is affected compared with the general population. (Zimprich, 2009) The relative risk with an affected sibling is about 2.5 to 3.5 times higher. (Zimprich, 2009)

Monogenic (rare variants with strong effects in single genes) and polygenic genetic changes (frequent genetic variants with weak effect in multiple genes) need to be distinguished. The concept behind monogenic or Mendelian epilepsies is that the seizures are the direct consequence of a genetic alteration. (Zimprich et al., 2015) Monogenic etiology is most clearly exemplified in rare monogenic epilepsies (see 1.5.2.1) with a positive family history. However, a positive family history may be missing in case of *de novo* mutations or may remain hidden under certain circumstances (low penetrance, recessive inheritance). (Hildebrand et al., 2013; Poduri & Lowenstein, 2011; Zimprich et al., 2015)

In the majority of all patients with epilepsy, the underlying genetics are much more complex. In this context, the terms "common", “complex”, or “polygenic” epilepsy are often used. (Berg et al., 2010; Sisodiya & Mefford, 2011; Zimprich et al., 2015) The least genetic burden is seen in epilepsies caused by structural changes, such as posttraumatic or tumor associated. However, in these cases, a partial genetic predisposition is also discussed, but has not yet been explored. (Peljto et al., 2014)

Figure 1.3 gives an overview of the genetic variants that contribute to human epilepsy: common variants are variants that are frequently found in the population but confer a low risk for a disease. Rare variants are associated with a strong increased risk of epilepsy. So common genetic variants are usually associated with a weak effect and rare variants with a strong effect. (Helbig et al., 2016) Mendel's rules describe the inheritance process for traits whose expression is determined by only one gene at a time. Such variants are rare but have a high risk of disease whereas common variants with a high risk of pathogenicity are very rare. Rare variants with a low risk are difficult to detect. (Helbig et al., 2016)

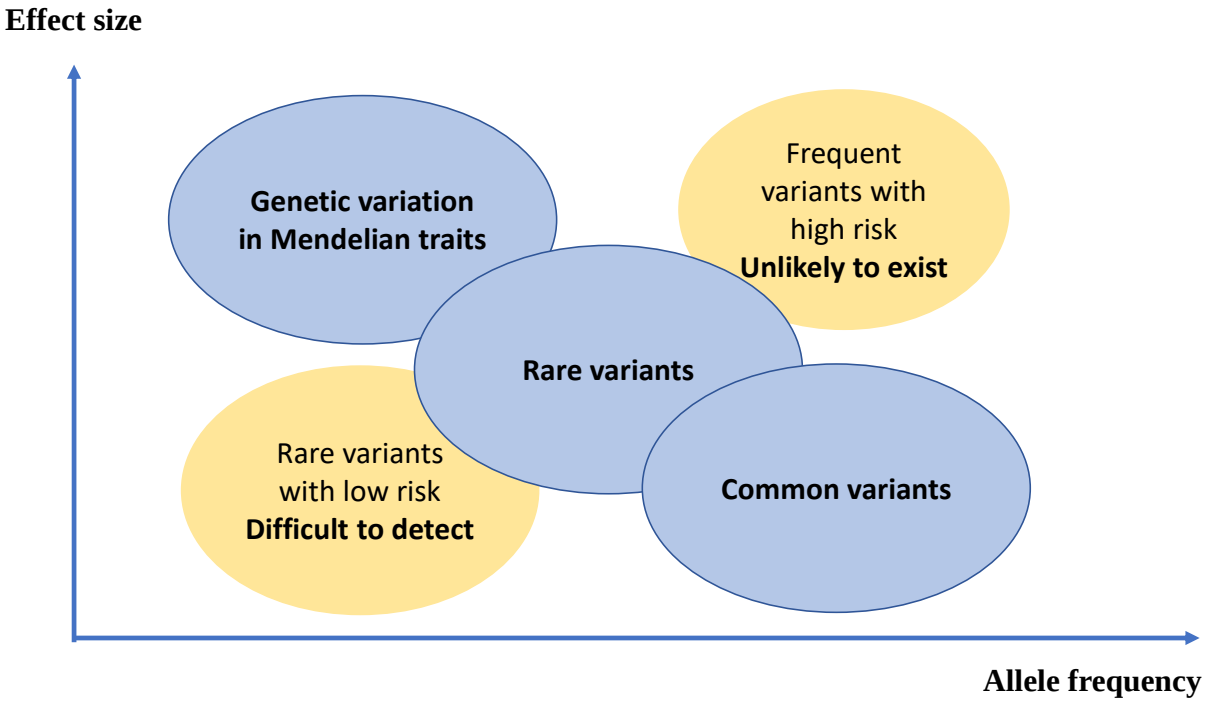

Figure 1.3.: Allele Frequency of different variants and their effect size. Modified from: Primer Part 1 – The building blocks of epilepsy genetics. (Helbig et al., 2016)

More than three million SNPs (single nucleotide polymorphisms) are known in the approximately 3.3 billion base pairs (bp) of the human genome. SNPs are variations at a single position in the genome that occur worldwide with a frequency of 1% or more. (Brookes, 1999) Most of these variants have no known effect on the phenotype or a trait. Traits are a specific characteristic of a human determined by genes, environmental factors, or a combination of both. (Evans et al., 2002) Traits can be quantitative (e.g., blood pressure or pulse) or qualitative (e.g., hair color). A particular trait is part of an individual's entire phenotype. (Serpico, 2020) The variants that do affect our phenotype are part of the normal variability and are responsible for individuality. (Strachan & Read, 2010d)

1.5.2.1 Monogenic epilepsy forms

The next sections explain the generalities and prevalence of monogenic epilepsies and what the possible causes and examples are. Monogenic epilepsies represent a clinically and genetically heterogeneous group of disorders. (Krenn et al., 2021) They are rare and represent only about 1-2% of all epilepsies. They result from mutations in single genes, which mostly code for voltage- or ligand-gated ion channels (potassium, sodium, chloride channels) and thus regulate neuronal excitability. (Weber & Lerche, 2008) This leads to dysfunction or even complete loss of function of the affected gene, resulting in the development of epilepsy.

Monogenic epilepsies follow the Mendelian inheritance rules (autosomal dominant, autosomal recessive, or X-linked recessive/dominant). (Weber & Lerche, 2008) A monogenic cause should be considered in the case of syndromal epilepsy, in seizures that are refractory to treatment, or if other causes (for example non-genetic) have already been excluded. In most cases of monogenic epilepsy, penetrance is not 100 and carriers may remain asymptomatic. (Gourfinkel-An et al., 2004) With decreasing penetrance, the overlap to polygenic epilepsies becomes therefore fluid. The following section explains when a monogenic cause of epilepsy can be assumed and what the most common monogenic genetic changes are.

Three main pathways leading to the development of epilepsy are distinguished. These are divided into ion channel mutations, mutations in enzyme/enzyme modulator genes and finally neurodevelopment-associated epilepsy genes which may lead to changes in receptor binding, cell adhesion, signal transcription, membrane trafficking, cytoskeleton and nucleic acid binding. (Wang et al., 2017)

Ion channel mutations represent the most frequent group among monogenic forms of epilepsy. (Wang et al., 2017) An example at this point are *SCN1A* variants, which are associated with a broad phenotypic spectrum, ranging from GEFS+ (generalized epilepsy with febrile seizure plus) to Dravet syndrome, in which cognitive development is also impaired. (see 1.4.3). (Wang et al., 2017) So far, more than 400 genes coding for ion channels are known. Ion channels are important for establishing a delicate balance between inhibitory and excitatory effects. Epilepsies caused by ion channel disorders often occur episodically, so that the symptoms in patients are temporary. (Steinlein, 2010) Some well-known ion channel genes are for example those coding voltage-dependent channels for potassium (KCNQ2), for sodium (SCN1A), and receptors for acetylcholine (CHRNA2) or GABA (GABRA1, GABRA6, GABRB2). (Conti et al., 2015; Guazzi & Striano, 2019; Lee et al., 2019) The rare autosomal dominant inherited benign familial neonatal convulsions are caused by variants in two subunits of voltage-dependent potassium channels (*KCNQ2, KCNQ3*). (Biervert et al., 1998) Still another example is ADNFLE (autosomal dominant nocturnal frontal lobe epilepsy), which is characterized by alterations in a subunit of the neuronal nicotinic acetylcholine receptor (*CHRNA2*, *CHRNA4*, *CHRNB2*). Patients with this familial epilepsy have short nocturnal focal seizures of the frontal lobe, usually beginning in childhood. (Combi et al., 2004)

The enzyme/enzyme modulator genes mainly cause severe forms of epilepsy, as they are important in the development of neurons and play a significant role in epileptogenesis. They represent the second most common group among monogenic forms of epilepsy. (Wang et al.,

2017) One example of this group is certain mutations in the *ALDH7A1* gene that cause, among others, pyridoxine-dependent epilepsy. (Mills et al., 2010) ALDH7A1 plays a role in lysine catabolism as well as in the detoxification of aldehydes produced in alcohol metabolism and lipid oxidation. (Brocker et al., 2011) Variants located further away from the active center have a greater impact on the structure and function of the protein. They affect catalytic function and thus prevent formation of the active tetramer. (Korasick & Tanner, 2021)

To be distinguished from the first two groups are the neurodevelopmentally associated epilepsy genes. These are genes that are associated with both epilepsy and brain maldevelopment. Variants in these genes result in gross neurodevelopmental malformations (without the intermediate stage of altered enzyme activity). (Jansen et al., 2006; Nellist et al., 2009) One example is tuberous sclerosis, which is caused by changes in genes such as *TSC1* and *TCS2* and is associated with severe epilepsy. (Jansen et al., 2006; Nellist et al., 2009) The gene products of these genes mainly have a central function within the mTOR (mechanistic target of rapamycin) signal transduction pathway such that pathogenic *TSC1* or *TSC2* variants lead to overactivation, resulting in increased proliferation and cellular growth in characteristic TSC lesions (e.g., periventricular calcification). (Wong, 2010) mTOR is a protein kinase that serves as a central regulator of various processes and regulates cell proliferation and differentiation. (Sarbassov et al., 2005) Another example is DEE2 (developmental and epileptic encephalopathy 2) with mutations in the cyclin-dependent kinase-like 5 gene (*CDKL5*). Variants in this gene lead to altered activity in the CDKL5 protein, which functions as a kinase. (Kilstrup-Nielsen et al., 2012) Pathogenic variants have been found in young patients with early infantile epileptic encephalopathy, among others. In addition to recurrent seizures, developmental delays or loss of speech occur. (Kilstrup-Nielsen et al., 2012; Weaving et al., 2004) The altered enzyme activity leads to neuronal changes and is thus responsible for epilepsy. Other examples include phenotypes caused by mutations in *DEPDC5* (DEP domain containing protein 5), *NPRL2* and *NPRL3* (nitrogen permease regulator-like 2 and 3). These genes encode the GATOR1 complex (GAP activity towards rags complex 1) - a negative regulator of the mTORC1 pathway (mechanistic target of rapamycin complex 1). (Baldassari et al., 2019) In turn, removal of the inhibitory influence leads to increased activation of the mTOR pathway. Patients with variants in these genes mainly suffer from focal seizures. (Baldassari et al., 2019)

There are various approaches to identify such monogenic causes. The individual pathogenic alterations in the genome of affected family members can be localized by linkage analysis and

subsequent fine mapping of the candidate regions and their sequencing. Further approaches to identify monogenic causes include genome-wide sequencing methods. These are divided into exome and genome sequencing. In exome sequencing, all protein-coding regions are sequenced (exons), in genome sequencing, additionally all non-coding regions (introns and intergenic regions) are sequenced.

1.5.2.2 Oligogenic epilepsy forms

Between monogenic and polygenic forms of epilepsy, there are also oligogenic forms. Traits that are influenced by a moderate number of genes are called oligogenic. (Rieger et al., 1976) Support for a genetic basis for epilepsies is also provided by studies of concordance rates for epilepsy in twin pairs. (Bourgeois et al., 2007) The concordance rate for idiopathic generalized epilepsy in monozygotic twins is up to 95%. (Samuel F. Berkovic et al., 1998) The risk of occurrence of idiopathic generalized epilepsy in first-degree relatives of affected individuals is approximately 8-12% and decreases to 1-2% in second-degree relatives. These risk scores strongly suggest an oligogenic inheritance pattern. (Steinlein, 2004)
A next example is GEFS+ (Generalized Epilepsy with Febrile Seizures Plus). In this epilepsy type, the mode of inheritance remains controversial. (Bourgeois et al., 2007) A dominant pattern is observed in some and an oligogenic effect in others, explaining a wide variation in clinical phenotype. (Bourgeois et al., 2007) Experimental evidence for oligogenic epilepsy has hardly been documented so far.

1.5.2.3 Polygenic epilepsy forms

The vast majority of idiopathic epilepsies are not caused by a single hereditary defect, but are due to an increased susceptibility to epilepsy caused by variants in many different genes. (Neubauer & Hahn, 2011) (S. F. Berkovic et al., 1998; Helbig et al., 2008) In polygenic epilepsy, according to current concepts, variation in hundreds of genes together with environmental factors contribute to epilepsy. (Speed et al., 2014) Each individual has a more or less strong tendency to develop seizures due to his or her genetic background. (Zimprich, 2009)
The basic model behind a polygenic inheritance is, that the combination of several variants in susceptibility genes increases neuronal hyperexcitability above a certain seizure threshold in an individual in his or her environment. (Dreifuss et al., 1985) As previously mentioned, the

respective variants have a comparatively high allele frequency (see Figure 1.3) and are associated with a comparably lower functional impairment not leading to a complete inactivation of the gene or protein function. Most frequently, these variants modify protein expression. (Epilepsies, 2014; Zimprich et al., 2015)

The most important approach to identify and characterize the molecular mechanisms that lead to the expression of a phenotype in a polygenic context are genome-wide association studies (GWAS). The aim is to identify genetic polymorphisms associated with a specific disease or trait thereby linking a particular haplotype (or allele) to a particular phenotype. Two groups are required to perform GWAS. (Begum et al., 2012; Manolio, 2010) The first group carries the trait or disease of interest, and the second group represents the control group (the healthy reference population) and needs to be the same age, sex, and region. The DNA (deoxyribonucleic acid) of all probands is analyzed for numerous SNPs. A significant enrichment of a certain marker in the case group represents an association. Most loci of the marker SNPs used are not located in a protein-coding region, but either in non-coding regions (e.g. regulatory regions) or in introns. (Begum et al., 2012; Manolio, 2010)

A SNP usually has two alleles: a more frequent one, called the major allele, and a less common one, called the minor allele. (Bush & Moore, 2012; Germer et al., 2000) With respect to every individual SNP, the population is divided into three groups: Depending on diploidy, a person may be a homozygous carrier of the more common or less common allele for a SNP, or a heterozygous carrier of both alleles. The term minor allele frequency (MAF) indicates the frequency with which the second most common allele occurs in a population. (Bush & Moore, 2012; Germer et al., 2000)

Association studies are used to identify genetic risk factors for polygenic epilepsies. In 2014, the ILAE conducted a meta-analysis examining over 8,500 epilepsy cases and over 26,000 controls from twelve cohorts. (Epilepsies, 2014) A total of three loci with genome-wide significance were identified (2q24.3, 4p15.1 and 2p16.1). Meta-analysis of the entire epilepsy cohort identified loci at 2q24.3, a region including *SCN1A* (a voltage-gated sodium channel, which is associated with some monogenic epilepsy forms, see 1.5.2.1) with an OR (Odds Ratio) of 0.89 (95%-CI (confidence interval): 0.86-0.93) and at 4p15.1, involving *PCDH7* (not previously associated with epilepsy; encodes a protocadherin molecule) with an OR of 0.88 (95%-CI: 0.82-0.93). (Epilepsies, 2014) A significant signal at 2p16.1 where the *VRK2* (vaccinia-related kinsae 2) and *FANCL* (Fanconi anemia complementation group L, coding a RING type E3 ubiquitin ligase of the Fanconi anemia pathway) genes are located, was

identified in the cohort of genetic generalized epilepsy with an OR of 1.23 (95%-CI: 1.16-1.31). (Epilepsies, 2014; Myers & Mefford, 2015)

In 2018 ILAE again performed a GWAS encompassing over 15,000 individuals with epilepsy and almost 30,000 controls. This revealed 16 significant loci across the genome, eleven of which were novel. (Abou-Khalil et al., 2018) Analysis of all epilepsy cases strengthened two previous associations at 2p16.1 (*FANCL, BCL11A*) with a p-value of $8.1x10^{-9}$ and 2q24.3 *(SCN3A, SCN2A, TTC21B, SCN1A*) with a p-value of $1.7x10^{-13.}$ (Abou-Khalil et al., 2018; Epilepsies, 2014) In addition, GWAS revealed a new genome-wide significant locus at 16q12.1 (*HEATR3, BRD7*) with a p-value of $4x10^{-8}$. Affected patients suffered from focal, genetically generalized, and unclassified epilepsy forms. That the group under study is highly clinically heterogeneous is a major weakness of the study, because different loci presumably play a role in each subgroup. The analyses suggest that the associated variants are involved in the regulation of gene expression in the brain and increase the overall risk for the most common epilepsies and certain epilepsy syndromes. (Abou-Khalil et al., 2018)

Allelic series refers to a set of different alleles in a gene or pathway, each causing varying degrees of impact on a trait or phenotype. In the context of genetic epilepsy, understanding the allelic series can provide insights into the spectrum of genetic variations influencing the manifestation and severity of epileptic disorders. (McCaw et al., 2023)

The importance of GWAS has increased significantly in recent years due to low-cost genotyping methods. (Annas & Elias, 2014) Through GWAS, numerous associations between diseases and the genetic variation of an organism's genome have been identified in recent years. The most important limitation of GWAS results is the low percentage of variability explained. (Tam et al., 2019) A causal correlation between allele and phenotype can only be investigated by molecular biological and biochemical methods. (Begum et al., 2012) Molecular biological effects are not so easy to measure and often no conclusions can be drawn for individual patients.

1.6 Aim of the study

The aim of the study was to find a new gene for a previously undescribed form of monogenic epilepsy. This included a linkage analysis in a large pedigree with childhood onset epilepsy and a more detailed analysis of the candidate gene *FERMT2*.

Prior to the project of the doctoral thesis, exome sequencing was performed in a diagnostic context at the Institute of Human Genetics in Munich in a patient with epilepsy from a family

originating in Poland (family #1), which had identified a variant in the *FERMT2* gene as a possible cause of the index patient's epilepsy (02/2019). Subsequently, the IHG in-house database (status 02/2019) with 25,000 exomes was screened for matching cases without success. A search via the matching portal GeneMatcher provided three additional patients with epilepsy to have a *FERMT2* variant. One family was from Portugal (family #2) with 41 family members, of whom 23 DNA samples were available (including ten affected and 13 unaffected) and the other family from Iceland (family #3) with one index patient. Sanger sequencing identified seven additional individuals (affected and unaffected) in family #2 with this variant. Because this variant did not segregate completely in the pedigree of the Portuguese family and the family was large enough, linkage analysis was performed (05/2019) in individuals of family #2 with available DNA to identify possible other disease loci.

Early childhood epilepsies can lead to cognitive, psychological, and motor developmental delays and thus represent a great burden for the patient and his family. Even the daily life of an epilepsy patient (e.g., driving a car) is impaired. (Noeker et al., 2005; Sillanpää & Cross, 2009) Epilepsies also cause high costs for the public health system or health insurance companies. (Begley & Beghi, 2002) Therefore, it is an important task to provide support to patients and to identify a possible genetic cause of the disease. For example, searching for mutations in epilepsy could help to further understand the underlying pathomechanisms, which are still not fully understood. Knowing that a particular variant leads to a particular phenotype can provide information about cell and tissue physiology. Based on the underlying mechanism of the epilepsy, further diagnostic algorithms can be developed through research and a treatment could eventually be found that targets the causative mutation in each case.

2 Patients, Materials and Methods

2.1 Patients and controls

Three different families were examined in the present work. In family #1 (see Figure 3.1), exome sequencing was performed at the Institute of Human Genetics in Munich (02/2019) and then reanalyzed in the course of this thesis to determine the genetic background. In family #2 (see Figure 3.2), genome-wide linkage analysis and haplotyping were carried out in 05/2019 after re-analysis of exome sequencing data from Portugal, which was performed in two individuals by the Behavioral and Molecular Lab, Universidade de Minjo, Braga, Portugal. Therefore, the main part of the methods refers to family #2. In family #3, the identified variant in the *FERMT2* gene via GeneMatcher (explanation see 3.1.1) was visualized in a 3D model, as in the other two families.

In family #1, blood samples were available from a total of three individuals. For family #2, there is cooperation with Portugal. The patients were phenotyped there and the DNA samples from 23 family members for genetic testing were also obtained from this institution. All patients were clinically examined and diagnosed by a neurologist in Portugal. The patients' DNA was sent to the "Helmholtz Zentrum" in Munich and processed there with the patients' consent form. The work was approved by the Ethics Committee of the Technical University of Munich (TUM).

First, the DNA isolation followed by molecular genetic methods used for exome sequencing are explained in more detail. Then, the bioinformatic methods are described. This is followed by the molecular genetics and bioinformatic methods for the SNP array according to the same scheme.

2.2 DNA-Isolation

2.2.1 Automated DNA-Isolation

DNA isolation from whole blood was performed using the Chemagic 360 instrument from PerkinElmer (Waltham, Massachusetts, USA) according to the manufacturer's recommendations (see https://chemagen.com/products/ivd-products/chemagic-360-ivd/). Magnetic separation was achieved with metal rods immersed in a process solution. First, the cells were lysed and then beads, small magnetic balls that efficiently bind DNA, were added

to the sample. A metal rod covered with plastic was magnetized by an electromagnet and immersed in the sample. The beads and the DNA bound to them were attracted to the needle. The advantage of magnetic beads over liquids is that there is no cross-contamination of samples during the isolation process. The DNA was then purified with special buffers, leaving the washed DNA fixed to the needle. The DNA was subsequently dissolved in the elution buffer, the empty beads were removed with the magnetic needle, and the dissolved DNA was finally collected in the collection vessel. This resuspension technique allows the isolation and purification of long DNA fragments.

2.2.2 DNA quality control

Next, after DNA isolation, the DNA was examined for quality and quantity purity using the *NanoDrop 1000* spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The aromatic rings of the bases absorb light, the absorbance is determined at a wavelength of 260 nm (= OD260 (optical density)) and 280 nm (= OD280). These values are used to calculate the extinction coefficient (OD260/OD280), which should be between 1.8 and 2.0. If it is above this range, contamination with RNA (Ribonucleic acid) is likely, if it is below it, proteins are present in the sample.
The DNA concentration of a sample can be calculated by rule of three. A measured OD260 value of 1 corresponds to a DNA concentration of about 50µg/ml. It follows that:

DNA [µg/ml] = measured OD260 [µg/ml] x 50 x dilution factor

In addition, a measurement using Quant-iT technology (Thermo Fisher Scientific, Waltham, Massachusetts, USA) was followed (see https://www.thermofisher.com/de/de/home.html). This uses specific fluorophores that fluoresce when bound to DNA, RNA, or proteins. The larger the amount of the target molecule in the sample, the higher the fluorescence intensity. This results in a very high signal-to-background ratio and from this the amount of DNA can be determined. This method is more precise than UV absorption because only the molecule of interest is labeled and there is no interference from other free nucleotides or contaminants.
All DNA samples used for exome sequencing and SNP array met the quality criteria and had an extinction coefficient ratio between 1.8 and 2.0 and had measurable amounts of DNA and were therefore sufficient for further analysis.

2.3 Exome sequencing

Since the didesoxy method developed by Sanger in 1977, enormous technological progress has been made in the field of sequence analysis. Today, it is possible to sequence the entire human genome or large subunits of it. These sequencing methods are also referred to as "next-generation sequencing" (NGS). (Kuß, 2014) Whole-exome sequencing (WES) represents one form of NGS. In this process, the entire exome, i.e., as far as possible all protein-coding regions of the DNA, is enriched and sequenced. In general, the sequencing process can be divided into library preparation, sequencing and imaging, and data analysis.

2.3.1 Wet lab

The Agilent SureSelect Human All Exon 60 Mb (Megabases) V6 Kit (enrichment of protein-coding regions, Santa Clara, California, USA) was utilized for capturing and library preparation and the Novaseq 6000 system from Illumina (San Diego, California, USA) was used for exome sequencing. With this high-throughput sequencing device, a DNA fragment can be sequenced from both ends, thus achieving higher coverage. The following steps were performed according to the manufacturer's recommendations.

2.3.1.1 Capturing and library preparation

Common to NGS methods is the need to produce a DNA fragment "library" prior to sequencing, consisting of enriched DNA fragments with platform-specific adapters. Two methods exist for the enrichment of exons. One is solution-based capture, and the other is array-based capture. In array-based capture, single-stranded oligonucleotides are bound to the surface of a chip. (Harakalova et al., 2011; Wagner, 2018) The Agilent SureSelect Human All Exon 60 Mb V6 Kit (Santa Clara, California, USA) was used in this work for exome enrichment which is in-solution capture based. Then, a pool of oligonucleotides bound to beads (small magnetic balls that efficiently bind DNA) are added to the sample and they can hybridize with the target region. (Bashiardes et al., 2005) 200 ng of a DNA sample was used as input. First, the isolated DNA was prepared for sequencing. For this purpose, it was mechanically (via ultrasound) split into fragments of about 300 bp with the LE220-plus Focused-ultrasonicator (Covaris, Massachusetts, USA). This step is also called fragmentation. Adapters (oligonucleotides) were then attached to both ends of the DNA fragments. These

include sequencing binding sites and an additional complementary sequence which could then hybridize to magnetic beads (see Figure 2.1) The sequences were complementary to the primers used later and the oligonucleotides were then ferromagnetically attracted and washed. The next steps were polymerase chain reaction amplification and sequencing. (Bashiardes et al., 2005) Figure 2.1 gives an overview about this process.

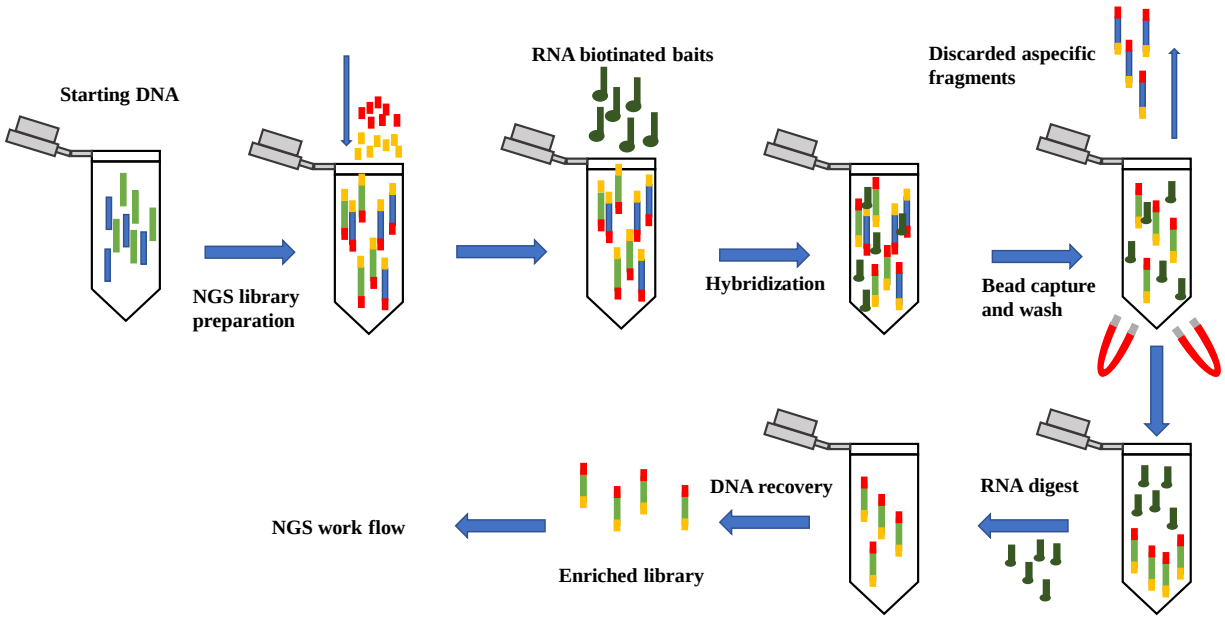

Figure 2.1.: Agilent SureSelect sequence capture workflow. In solid-sequence, capture is performed by hybridization between NGS library sequences, and the capture probes immobilized on an array surface. Modified from: https://www.researchgate.net/figure/Agilent-SureSelect-sequence-capture-workflow-In-solid-sequence-capture-is-performed-by_fig5_225373000

2.3.1.2 Cluster Generation

The Novaseq 6000 system from Illumina (San Diego, USA) was used for cluster generation and exome sequencing. The following processes run automatically: The first step on the flow cell is the DNA amplification by PCR. A flow cell is a thick glass slide with channels or lanes. (Piro, 2020) These nanowells are coated with a lawn of oligonucleotides to which the complementary adapters attach (see Figure 2.2, part 1.). (Piro, 2020) The complementary sequence to the single-stranded DNA fragment is then synthesized and this process is done simultaneously for all DNA fragments on the entire flow cell. Then the DNA is denatured and the single strand which is not attached to the oligonucleotides of the flow cell gets washed away. This is followed by the so-called bridge PCR (see Figure 2.2, part 2.). (Albert et al., 2007; Mertes et al., 2011; Schober, 2002; Turner et al., 2009)

There are two different types of oligonucleotides, and the second type will now also hybridize to one of the oligonucleotides on the plate. This results in a "bridging" of the DNA, which is

why this procedure is called bridge PCR. As in normal PCR, a DNA polymerase synthesizes a complementary strand starting from the primer (see Figure 2.2, part 3. and 4.). This process is done simultaneously for all DNA fragments on the whole flow cell. This strand is then split off again by denaturation, resulting in two single strands, but now both fragments stick to the plate. They are called forward strand and reverse strand (see Figure 2.2, part 5.). The whole procedure will be repeated, and the DNA fragments build a bridge again and again (bridge amplification). By repeating this process, so-called clusters are formed on the flow cell, which consist of DNA strands with identical sequences (see Figure 2.2, part 6.). (Albert et al., 2007; Mertes et al., 2011; Schober, 2002; Turner et al., 2009).

2.3.1.3 Sequencing

Illumina sequencing works on the principle of sequencing-by-synthesis. To perform sequencing which works automatically, primers, DNA polymerase and four dNTPs (deoxynucleoside triphosphate) are added, to which different fluorophores that emit at different wavelengths, and to which a blocking group are attached. In each cycle, the strand is extended by one base by the polymerase and DNA synthesis stops by the blocking group. A laser detects the fluorescent labels. The image is stored by a camera and the clusters are registered by a program. The fluorescent label is then enzymatically cleaved from the nucleotide so that the next dNTP can be inserted. This process is repeated until the DNA fragments are fully synthesized. Through PCR, there are multiple copies, so many light signals are obtained simultaneously. The final step is the data analysis, where the reads are overlaid and compared with a reference genome and through this overlay the whole genome sequence can be generated. (Albert et al., 2007; Mertes et al., 2011; Schober, 2002; Turner et al., 2009)

2.3.2 Bioinformatical workflow

FASTQ file formats are obtained, which combine both the sequence and an associated quality score per base. They were established for the exchange of sequencing data. (Cock et al., 2010) Now, each read sequence must be mapped to a reference genome. This process is called alignment. There are a number of software tools available for this purpose, which differ, however, e.g. with regard to the accuracy of the alignment or with regard to the runtime when processing large genomes. (Musich et al., 2021) Several requirements must be met for

merging data sets and for imputing variant data. For example, to impute genotypes using reference panels, all alleles must be characterized by the forward or plus strand allele of the same reference genome. The alignment of the reads was done to the Human Genome Assembly GRCh37 (hg19) and analyzed using programs that work with different algorithms. This was done by the Burrows-Wheeler Aligner (BWA) 0.7.5.a using the Mem algorithm (Wagner et al., 2019) on a custom reference based on the above mentioned reference sequence.

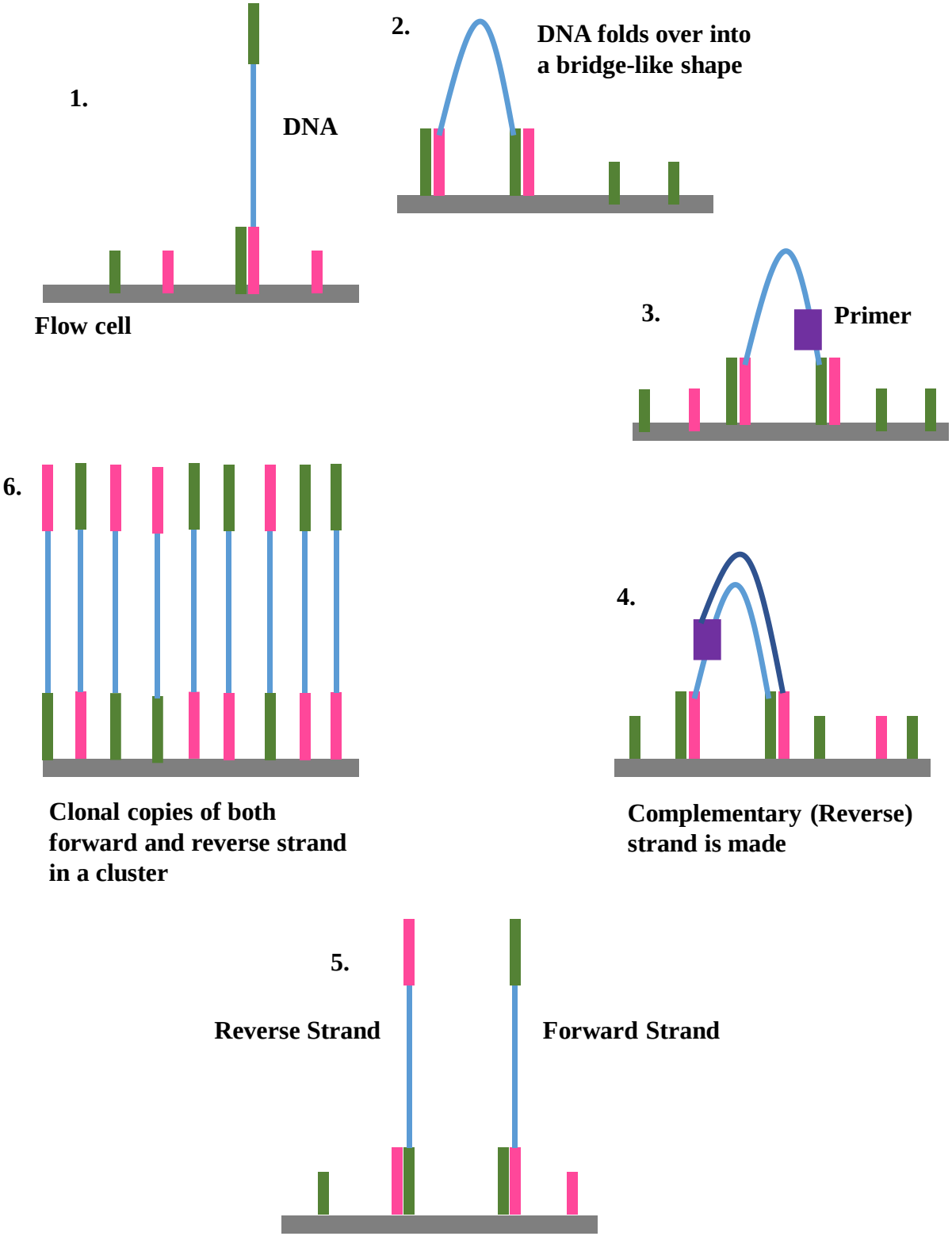

Figure 2.2.: Illumina Sequencing: The grey bar represents the flow cell. The purple rectangle shows the briar. The pink and green ends are adapters. The blue lines represent individual DNA fragments. The DNA attaches itself to the flow cell via complementary sequences (1.). By bending the DNA strand and attaching it to a second oligo, bridging occurs (2.). Next, a polymerase synthesizes the reverse strand (3. and 4.). Now the two strands break up and straighten up (5.). Each strand then forms a new bridge (bridge amplification). A cluster of clones of DNA forward and reverse strands is formed (6.). Modified from: https://en.wikipedia.org/wiki/Illumina_dye_sequencing

SAMtools v. 0.1.7 was used to create a sorted BAM file. To remove duplicate reads and thus reduce the number of false-positive heterozygous calls, Picard was used. (Wagner, 2018) It is a program that includes Java-based command-line utilities for processing SAM files. SAMtools was also used to create a BAI file, and SNVs were called by SAMtools, while indels (insertion or deletion) were called by DeepVariant. (Wagner, 2018) CNVs were called using ExomeDepth and mtDNA (mitochondrial DNA) variants were called using GAT. (Wagner et al., 2019) Calls were filtered for variants that were "on-target" (bed-files from the Agilent V6 kit +/- 500 bp), occurred on both DNA strands, and had a minimum coverage of 5. SNVs found in <80% of the reads were declared heterozygous, when found in >80%, they were considered homozygous. So, for the analysis of the sequence variants SAMtools v. 0.1.7 was used and lastly, Annovar was implemented for gene-based annotation of the changes. (Wagner, 2018) The analysis is based on autosomal recessive and autosomal dominant inheritance. Rare relevant variants were compared using the in-house database (> 25,000 exomes) and databases available on the internet such as dbSNP identifiers (https://www.ncbi.nlm.nih.gov/projects/SNP/), Genome Aggregation Database (exome/genome of over 135,000 individuals, gnomAD, http://gnomad.broadinstitute.org) and evaluated using prediction programs such as SIFT scores (Li & Homer, 2010), CADD (http://cadd.gs.washington.edu), PolyPhen-2 scores (Adzhubei et al., 2013), LRT scores (Chun & Fay, 2009), PhyloP conservation scores (Cooper et al., 2005), GERP++ conservation scores and other information.

2.3.3 Data analysis

To reduce this large number of variants to some candidate variants, several filtering steps were used. The following filter systems were used to detect potentially pathogenic variants:

1. HGMD (Human Gene Mutation Database)/gnomAD search: This search lists variants with a MAF <1%, which were found in one of the two databases (HGMD/gnomAD). In this way, already published variants could also be found.
2. Autosomal recessive/X-linked search: This search filters variants that are either homozygous, hemizygous (X-linked) or predictably heterozygous.
3. Autosomal dominant search (based on OMIM (Online Mendelian Inheritance in Man) phenotype): A full-text OMIM search with specific phenotype-based terms (e.g., myopathy) is performed to create candidate gene lists. Variants of genes in these lists with a MAF <1%

were listed and evaluated for plausibility. An autosomal dominant search is also performed without the use of OMIM, with the aim of presenting new candidate genes.

4. Search for variants with copy number variation (CNV): using ExomeDepth and Pindel, deletions and duplications can be identified.

5. Search for mtDNA.: this search enables the identification of disease-causing variants in the mtDNA.

6. Specific search options: It is also possible to search for variants in specific genes (gene search). In addition, filtering for regions with LOH (loss of heterozygosity) that may indicate consanguinity (large homozygous regions) is also available. In the case of trio-exome sequencing, a specific *de novo* search is used to consider variants that are missing in both parents.

Variants are then classified using the American College Medical Genetics and Genomics (ACMG) criteria. (Abou Tayoun et al., 2018; Ellard et al., 2019; Richards et al., 2015) The ACMG approach applies to the interpretation of variants in patients with presumed hereditary (primarily Mendelian/monogenic) disease. It is important to distinguish whether a variant is pathogenic, i.e., disease-causing, or merely disruptive/damaging to the encoded protein. Pathogenicity should be determined based on the totality of evidence, including all cases studied. If a variant meets neither the criteria for pathogenicity nor benignity or results in conflicting assessments, it is termed a variant of unclear significance. (Richards et al., 2015)

2.4 3D modelling of a protein

The website AlphaFold (https://alphafold.ebi.ac.uk/) can visualize a protein in a 3D model. It predicts its structure from its amino acid sequence. The protein name is entered in the input window, the organism Homo sapiens is selected and then the 3D presentation of the protein and the positions of the individual amino acids as well as the respective chemical interactions are presented. UniProt (https://www.uniprot.org/) can also be used to search for the protein sequence and its functional information in the Homo sapiens organism. It helps to identify, in which domain the searched variant is located, or whether it is in the region of a beta leaflet or an alpha helix for example.

2.5 SNP-Array

SNP-Arrays (containing 14668 SNPs) of 23 unaffected and affected individuals from the Portuguese family were performed by Helmholtz Zentrum München (Munich, Germany). DNA microarrays are used for analysis of gene expression changes as well as for the typing of genetic material. The microarray technology is based on the hybridization of nucleic acids. (Vora et al., 2004) DNA obtained from whole blood is usually not sufficient for a microarray and, moreover, is not labeled. Therefore, amplification and labeling are required for the sample to be processed on the microarray.

2.5.1 Wet lab

The following processes were done according to the recommendations of the manufacturer Illumina (San Diego, USA). The samples were amplified several thousand times and then enzymatically fragmented on the next day. Then the DNA was precipitated with alcohol, resuspended, and finally the samples were hybridized to the Infinium Global Screening Array-24 v3.0 BeadChip (Illumina, San Diego, USA). Each bead was covered with hundreds of thousands of copies of a specific oligonucleotide, which acted as capture sequences in a specific Illumina assay. Complementary oligonucleotides present in the sample bound to the beads by hybridization. Non-hybridised DNA samples were washed away. On the third day, the samples underwent enzymatic base extension and fluorescent staining. Once the microarrays dried by being centrifugated, they were stored in the dark until scanned. Then computer image analysis was performed to process the data before bioinformatics analysis of the target genes was performed. For sample analysis, the BeadChip was then imaged with the iScan system (Illumina, San Diego, USA), which uses a laser to excite the fluorophore of the single base extension product on the beads. This was done by inserting the chip into the bead array reader and scanning the barcode of the BeadChip. The fluorophores emitted light, and the scanner recorded high-resolution images of them.

2.5.2 Data analysis

Finally, Illumina software automatically performed analysis and genotype calling. The data sets were merged using the PLINK software, which is a toolset for whole-genome association analysis. It covers a variety of analysis steps required for GWAS (e.g., data management,

population stratification, summary statistics, association analysis, and estimation of identity by ancestry). (Chang et al., 2015; Purcell et al., 2007)

The reduction of the large number of SNPs (to a random value of 14668 SNPs) is performed in several steps by a quality control with PLINK. To reduce the large SNP array dataset, LD pruning (pruning for linkage disequilibrium = removal of loci based on high pairwise LD to restrict the number to informative SNPs) was used. Samples with a call rate <90% were removed from the analysis and unmapped SNPs were discarded. Markers with a MAF >1% were also deleted.

2.5.3 Linkage analysis

2.5.3.1 The Concept – Linkage analysis

In genetics, gene linkage refers to the fact that genes that are physically close to each other are inherited together. Due to gene linkage, certain traits are always inherited in combination with others by the next generation. (Teare & Barrett, 2005)

Linkage analysis is a bioinformatic statistical method that attempts to link phenotypic traits, such as diseases, to genes or genetic markers and thus localize them chromosomally. (Arnemann, 2019) Linkage analysis tests whether certain alleles of a polymorphism within a pedigree with a phenotypic trait or a disease are cosegregated, i.e., are inherited together or linked, or whether they are inherited independently. (Strachan & Read, 2010c) The segregating allele does not necessarily cause the disease. The disease locus (in the region of a neighboring allele) may also be located on the chromosome in relative proximity and thus the segregating allele may also act as a marker. (Strachan & Read, 2010c) The linkage between the two chromosomal loci is then not separated by recombination events in meiosis, even over several generations. The further apart two loci on a chromosome are, the more likely it is that they will be separated in a crossover during meiosis. (Strachan & Read, 2010c) This probability is called the recombination frequency and is a measure of the distance between the two loci. Two loci with a recombination rate of 0.01 and thus 1% recombination between them are defined on a genetic map as 1 centimorgan (cM) apart. The unit cM is therefore a mathematical quantity and corresponds on average to a physical distance of 10^6 bp (depending on the region, this physical distance varies slightly). (Strachan & Read, 2010c) A single recombination results in two recombinant and two non-recombinant chromatids. If the loci on the chromosome are further apart, there may be more than one overlap between them.

Double overlaps may involve two, three or all four chromatids, but the total effect averaged over all double overlaps never yields more than 50% recombinants. Thus, the recombination fractions never exceed 0.5, no matter how far apart the loci are. (Strachan & Read, 2010c) The recombination frequency θ is therefore the quotient of the amount of meiosis with observed recombination and the total number of all meiosis examined. The closer the gene loci are located on the same chromosome (the more the recombination frequency θ=0 approaches), the greater the degree of linkage between the gene loci. (Strachan & Read, 2010c)

The detection of linkage in affected families therefore allows a statement about the chromosomal location of a disease gene. In the statistical analysis of linkage data, a fundamental distinction is made between parametric (model-dependent) and non-parametric (model-free) methods. (Strachan & Read, 2010e) In parametric linkage analysis, the trait or disease under consideration is assumed to follow a specific mode of inheritance. It requires a precise genetic model with specific parameters: Information on the penetrance of each genotype, mode of inheritance, and gene frequencies. (Strachan & Read, 2010e)

The OR is calculated from the probability that the observed data are compatible with linkage and the probability that they are not linked. Odds is a chance or a ratio of chances. It indicates the strength of a link between two characteristics and is therefore a measure of association in which two chances are compared. (Bland & Altman, 2000) The logarithm of this ratio, the LOD score (logarithm of the odds), is used as a measure of the probability of linkage. (Strachan & Read, 2010c) The LOD score compares the probability of obtaining the test data when the two loci are linked to the probability of obtaining the same data by chance only. (Morton, 1955) For monogenic characteristics or diseases, a LOD score of 3.0 is considered a strong indication of linkage (i.e., the probability that linkage exists would be 1,000 times greater than the probability that it does not exist), while a LOD score of -2 is considered to exclude linkage. (Strachan & Read, 2010c) However, the fact of multiple testing must be considered. For genetically complex diseases, the evaluation of linkage results is more difficult, since here the disease does not follow a clear pattern of inheritance. (Strachan & Read, 2010c)

2.5.3.2 Methods – Linkage analysis

Because both males and females in family #2 had epilepsy over several generations, an autosomal dominant inheritance was assumed for this family (see Figure 3.2). In individuum

II. 3, II.5 and IV.8 (see Figure 3.2) no clear epilepsy could be diagnosed. Due to the symptoms (e.g.: epileptic activity in the EEG, febrile convulsions) they were nevertheless considered "affected" in the further course of the work. The program Merlin was used to perform linkage analysis.

In Merlin, four different files were required to perform the analyses: a data file (*parametric.dat*, see Table 13), a map file (*parametric.map*, see Table 14), a pedigree file (*parametric.ped*, see Table 15) and a model file (*parametric.model*, see Table 16). The pedigree file contains information on all patients about their affection status, parents, sex and generation. A model file containing data on estimated allele frequency, prevalence, and expected penetrance for heterozygotes and homozygotes for epilepsy is used to calculate the LOD score for different inheritance patterns. The data file contains all SNPs (approximately five markers per cM, a total of nearly 14,700 markers) on which the linkage analysis is based. These markers are linked to the positions on the individual chromosomes in cM in the map file. For an example of the used files see 8.1.

In the used model file, a prevalence of 1% of epilepsy was ultimately calculated if no allelic copy of the disease allele was present in affected individuals, 90% if one copy was present, and 100% if two copies were present. The values 0% - 100% - 100% for a dominant inheritance pattern were not assumed in the model file because, on the one hand, phenocopies (the phenotype of another allele is imitated by the influence of non-hereditary, exogenous factors (Graw, 2020)) may also be present and because, among other things, febrile convulsions are basically common and reduced penetrance may also be present.

First, for each analysis the input files had to be validated. Pedstats is a program, which performs a basic quality control of the input files. Basic statistics, such as heterozygosity and completeness of genotyping can be provided. Pedstats also can generate graphical summaries of allele and genotype frequencies. (Wigginton & Abecasis, 2005)

The following command in DOS (Disk operating system) is used by Pedstats:

```
prompt> pedstats -d parametric.dat -p parametric.ped
```

In the next step, Merlin was used to perform a parametric linkage analysis by entering the following command:

```
prompt> merlin -d parametric.dat -p parametric.ped -m parametric.map
--model parametric.model --pdf
```

By adding --pdf at the end of the command, a PDF file is created. All input files are created individually for each chromosome.

2.5.4 Haplotyping

2.5.4.1 The Concept – Haplotyping

A haplotype is a variant of a nucleotide sequence on one and the same chromosome in the genome of a living being. A particular haplotype can be specific to an individual, a population or even a species. If a diploid organism possesses the genotype AaBb regarding two alleles A and B, it may be based on the haplotypes AB|ab or Ab|aB. (Gibbs et al., 2003)
Two cases can be distinguished (in the following, the term "allele" refers to the different nucleotides A, C, G and T, but the number of repetitions of a particular microsatellite, for example, can also define an allele). (Vieira et al., 2016)
In humans, haploid refers to the set of chromosomes in the nucleus of an egg or sperm cell. At fertilization, the two sets of haploid chromosomes in these germ cells combine to give the double set of chromosomes in the zygote, a diploid cell. (Strachan & Read, 2010a)
If the degree of ploidy of the species under consideration is at least 2, the genome is composed of two or more homologous sets of chromosomes, one half usually originating from the maternal and the other from the paternal parent. (Strachan & Read, 2010a) The alleles compared may be individual combinations of SNPs, as in the International HapMap (Haplotype Map) Project, which can be used as genetic markers. (Gibbs et al., 2003) Different types of SNPs are distinguished:
When a maternal and a paternal homologous chromosome set in an individual differ in the nucleotide positions of the DNA, these SNPs become visible when the corresponding chromosomes of the individual are sequenced. Such a SNP is called heterozygous SNP in the corresponding individual. (Lynce & Marques-Silva, 2006; Strachan & Read, 2010b)
If in an individual a maternal and a paternal homologous set of chromosomes is identical in a considered gene locus, no SNPs will be visible when sequencing the DNA of the individual. Only if a different allele is found at the same locus in at least one second individual can one speak of a SNP at the corresponding nucleotide position. Such a SNP is called a homozygous SNP in the first individual but may represent a heterozygous SNP in another individual. (Lynce & Marques-Silva, 2006; Neigenfind et al., 2008; Strachan & Read, 2010b)

2.5.4.2 Methods – Haplotyping

To further narrow down the identified regions of the linkage analysis, haplotyping was performed using the Merlin program.

The following command was run to estimate the haplotypes:

```
prompt> merlin -d haplo.dat -p haplo.ped -m haplo.map –best -- horizontal
```

For the .dat, .ped, and .map files, the same files were used as for the linkage analysis. The haplotypes were created for all individuals with existing DNA. The option --best provides haplotypes that correspond to the most likely pattern of gene flow. The --horizontal command produces a horizontal output format with a single haplotype per line. Estimated haplotypes are shown in the output file merlin.chr (see an excerpt of the chr file in 8.2, Table 18). A merlin.flow file is also created (see an excerpt of the flow file in 8.2, Table 17, for an example), which summarizes the descent of estimated haplotypes through the pedigree.

The flow file uses a unique designation for each founder haplotype (e.g.: A, B, C, etc.) and helps to identify the descent of founder alleles through the pedigree to further narrow down specific loci. The .Chr file lists the two haplotypes for each individual, always listing the maternal haplotype for non-founders first, followed by the paternal haplotype. (Genetics & Michigan, 2020)

If affected individuals have the same haplotype at a particular locus but healthy individuals do not, the disease locus can be further narrowed down in this way. This is done by comparing the existing data of all individuals in the .chr file at the corresponding positions. This was done for the individual three candidate regions on chromosome 2 and 18 as well as for the *FERMT2* region on chromosome 14.

3 Results

3.1 Clinical findings and pedigrees

3.1.1 Family #1

The family members of family #1 (see Figure 3.1) are from Poland. They were examined in Germany in Vogtareuth in the Department of Neuropediatrics and Neurological Rehabilitation of the Schön-Klinik. According to the available clinical data, the affected girl (III.1) has been suffering from severe, therapy-refractory focal right occipital epilepsy since she was eight years old. Seizure frequency increased from an interval of several times daily to every few seconds. EEG showed occipital deceleration but without the typical potentials of epilepsy. The cMRI showed subtle signal enhancements on both sides of the occipital lobe (right > left), which are not considered pathological. Ictal SPECT (single photon emission computed tomography) performed in 2015 showed decreased enhancement in the right occipital lobe. The sister (III.2) is reported to be healthy. Both parents (II.1 and II.2) have myopia, and the maternal grandmother (I.1) was diagnosed with migraine. No other neurological disorders - especially febrile seizures and other seizure types or developmental disorders - are known in the family.

Trio-exome sequencing was performed in this family at the Institute of Human Genetics in Munich in 02/2019 and a potential pathogenic de novo variant had been discovered. The exome data were subsequently reanalyzed as part of this doctoral thesis to look for pathogenic variants in known epilepsy disease genes, but no additional pathogenic variants could be identified.

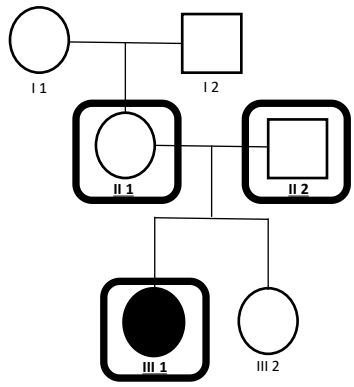

Figure 3.1.: Pedigree Family #1. Squares symbolize male family members, circles female family members. Individuals with diagnosed epilepsy are highlighted in black. White symbols indicate healthy individuals. Roman numerals (I-III) indicate the generations, Arabic numerals (1-2) the individual persons in the generations. The three individuals included in trio exome sequencing are framed in black.

To clarify a potential causal association between the variant identified in the index patient (see 3.2.1) and her epilepsy, additional variants in this gene were searched for in the in-house database. For this purpose, the in-house exome database was queried for variants with a MAF < 0.1% according to the in-house database (Klinikum Rechts der Isar, Munich) in 06/2022 with >25,000 exome datasets. But no matching cases were identified. However, two other patients with different *FERMT2* variants in a family from Portugal (family #2, see 3.1.2) and one patient in a family from Iceland (family #3, see 3.1.3) were identified via GeneMatcher. Sanger sequencing identified seven additional individuals in family #2 with a potentially pathogenic *FERMT2* variant. GeneMatcher (see https://genematcher.org/) is a website designed to make connections between families, patients, clinicians, and researchers worldwide. The site's primary purpose is to help to solve previously "unsolved" exomes. As of Sep. 28, 2022, there are already 12,964 submitters from 102 different countries with 65,314 gene submissions. Among them, there have already been 10,068 matches detected between genes from different patients (https://genematcher.org/statistics/).

3.1.2 Family #2

The family members of family #2 (see Figure 3.2) included in this study are all from Portugal. The family covers four generations, with the first generation being already deceased. DNA samples were available from 23 of 41 individuals, who were neurologically examined and diagnosed by a neurologist in Portugal. In the family, there are a total of ten affected individuals with available DNA. Individuals II.3, II.5 and IV.8 (see Table 1) could not be diagnosed with epilepsy although they have been thoroughly examined and questioned. Due to the symptomatology (e.g.: epileptic activity in the EEG, febrile seizures) they were nevertheless classified as "affected" in the further course of the work. To determine the genetic background, exome sequencing data from patient III.14 and IV.1 were reanalyzed in the beginning of this work. To determine, whether additional individuals in family #2 carry the *FERMT2* variant, Sanger sequencing was performed in Portugal. Next, SNP-arrays of 23 unaffected and affected individuals were performed at the Helmholtz Zentrum München (Munich, Germany). Genome-wide parametric linkage analysis was then conducted in the further course of this work.

Individuals II.3 and II.5 with the subsequent generations were not included in the linkage analysis because the subsequent generations of these individuals are unaffected by epilepsy and thus less informative than the left part of the pedigree. The DNA was provided by Prof.

Dr. Patrícia Espinheira de Sá Maciel from Behavioral and Molecular Lab, Universidade de Minjo, Braga, Portugal. In the further course, there was a regular exchange via emails and Zoom meetings with Portugal about new results and further progress.

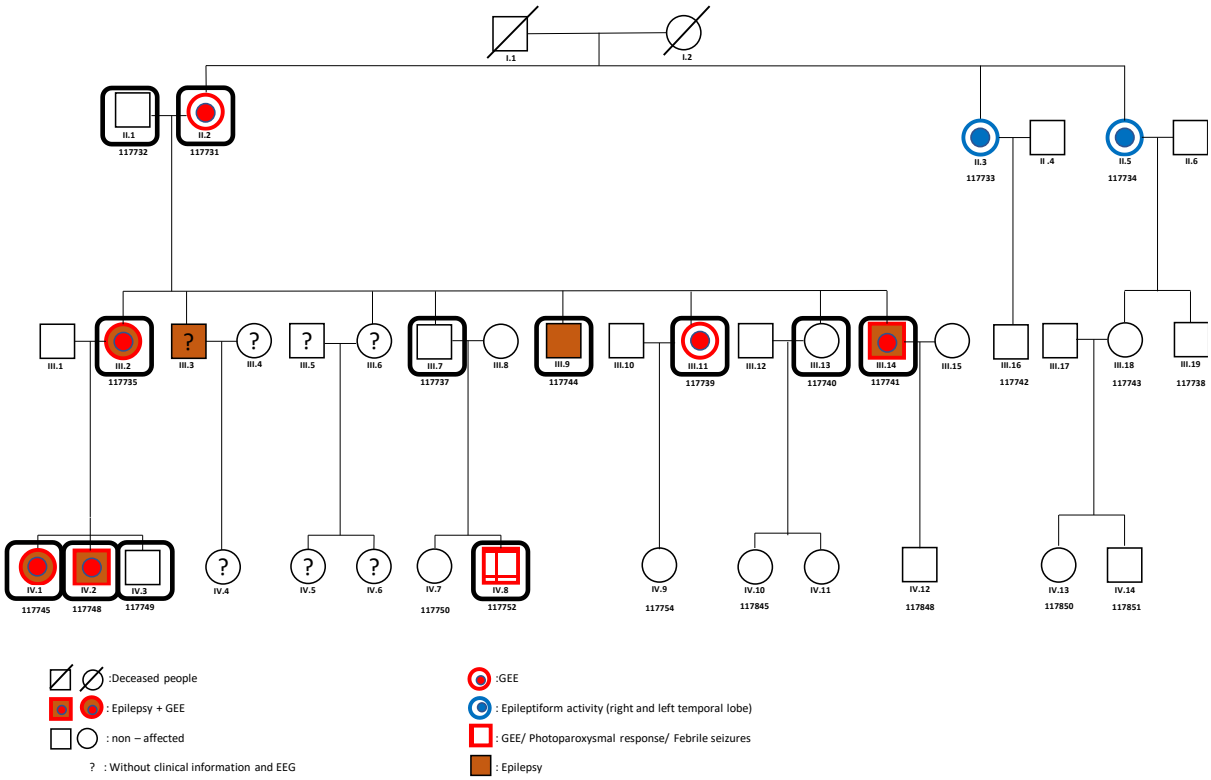

Figure 3.2.: Pedigree Family #2. Squares symbolize male family members, circles female family members. Deceased individuals are marked with a crossed-out symbol. Individuals with diagnosed epilepsy are highlighted in brown, those with GEE are outlined in red. Individuals with epileptic activity (without diagnosed epilepsy and without generalized epileptiform activity) in the EEG are marked blue. White symbols indicate healthy individuals. Roman numerals (I-IV) indicate the generations, Arabic numerals (1-19) the individual persons in the generations. The numbers below the Roman numerals represent the DNA codes of individuals with available DNA. The twelve individuals included in the genome-wide linkage analysis are framed in black. Abbreviations: GEE = generalized epileptiform activity; EEG = Electroencephalogram.

3.1.3 Family #3

In Reykjavik, Iceland, another potentially pathogenic *FERMT2* de novo variant was identified via GeneMatcher (06/2022) in a patient at deCODE genetics (correspondence with Ms. Telma Sulem) in family #3. The index is a male patient, born in 2015 with autism, moderate intellectual disability, macrocephaly, and speech and social developmental delay. No further clinical information on the parents or other affected individuals in the relatives was available.

3.2 Exome sequencing

3.2.1 Exome sequencing results of family #1

The trio exome analysis (genetic examination of the DNA of the index (III.1) together with the mother (II.1) and the father (II.2)) had identified a heterozygous *de novo* missense variant NM_006832.3:c.392A>G, p.(Asn131Ser) at the genomic position chr14:g.53360145T>C in the gene *FERMT2* (Fermitin Member 2) during routine molecular diagnostics in the index prior to this work. In a reanalysis of this family, this variant in *FERMT2* could be confirmed. In addition, sequence variation was searched for in a list of 977 known epilepsy genes (Wang et al., 2017). This focused search did not reveal any known or potentially disease causing variants.

FERMT2 (OMIM: #607746) encodes a scaffold protein and enhances integrin activation mediated by *TLN1* (Talin1) and/or *TLN2* (Talin2). (Bouvard et al., 2013) Integrins have an important function in cell signal transduction. The gene enhances integrin-mediated cell adhesion to the extracellular matrix (ECM) and thus signal propagation. (Keck et al., 2018) The ECM consists of particles formed and secreted by glial cells and neurons, which in turn form stable aggregates in the extracellular space. (Berezin et al., 2014) ECM signalling also performs regulatory functions in neuronal processes such as proliferation, cell motility, differentiation, and axonal growth. (Wu & Reddy, 2012) Figure 3.3 gives an overview of the tertiary structure of the *FERMT2* protein in three different levels.

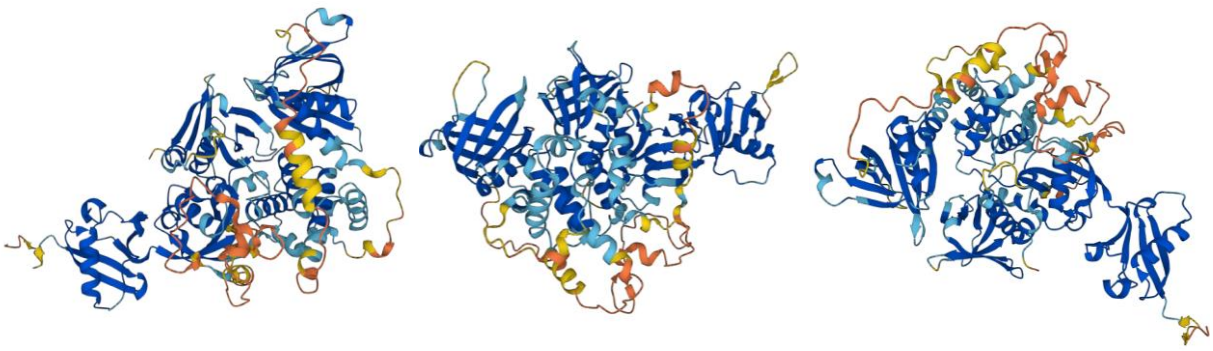

Figure 3.3.: Tertiary structure of the *FERMT2* protein. Copied from: https://alphafold.ebi.ac.uk/entry/Q96AC1

To date, an association between *FERMT2* and Alzheimer's dementia (AD) has been demonstrated. (Chapuis et al., 2017) In this context, variants in *FERMT2* were significantly associated with alteration in Aβ-peptide levels in the cerebrospinal fluid in patients with AD.

It has been shown that under-expression of *FERMT2* increases Aβ-peptide production, and that *FERMT2* thus modulates AD risk by regulating APP (amyloid-β precursor protein) metabolism and Aβ-peptide production. (Chapuis et al., 2017) The gene has not been previously associated with monogenic diseases.

On the other hand, the expression of *Fermt2* was greatly increased after SE (status epilepticus) in a rat model, especially in the hippocampus, suggesting a role of the protein in the regression of the ECM structure and cell adhesion after an epileptic seizure. (Keck et al., 2018) The protein Fermt2 can bind intracellularly to part of the ß-integrin subunit, activating integrin receptors and thus playing a crucial role in the exchange between intracellular and extracellular components and signaling events. (Keck et al., 2018) *Fermt2* may consequently have an important contribution in the development of epilepsy and haploinsufficiency may therefore increase the risk for epilepsy. (Keck et al., 2018)

The gnomAD database records a Z-score of 1.96 for the gene *FERMT2*, which means that this gene statistically has a moderate intolerance to missense variants. In addition, this database gives a pLI (probability for loss-of-function intolerance) of 1.00 and a LOEUF value of 0.13, which means that *FERMT2* statistically has a high intolerance to loss-of-function. A more detailed explanation of pLI and LOEUF follows under section 3.5. The position of the variant was covered 33 times. The variant was found in 7/33 reads, which is an indication, but no proof for a mosaic. Based on the variant c.392A>G, the replacement of a highly conserved amino acid and at the same time the loss of the splice acceptor site of exon 4 is predicted.

Figure 3.4 shows the position of the *FERMT2* variant and the chemical connections or interactions to neighboring structures in the protein. The variant c.392A>G, p.(Asn131Ser) is not located in any specific domain. Starting from the Asn131-residue, there are hydrogen bonds to two alpha helices and one beta leaflet, so changes within this region could lead to alterations in protein structure and function. Replacing the large amino acid asparagine by the small amino acid serine can lead to a collapse of the protein structure and thus be locally disruptive.

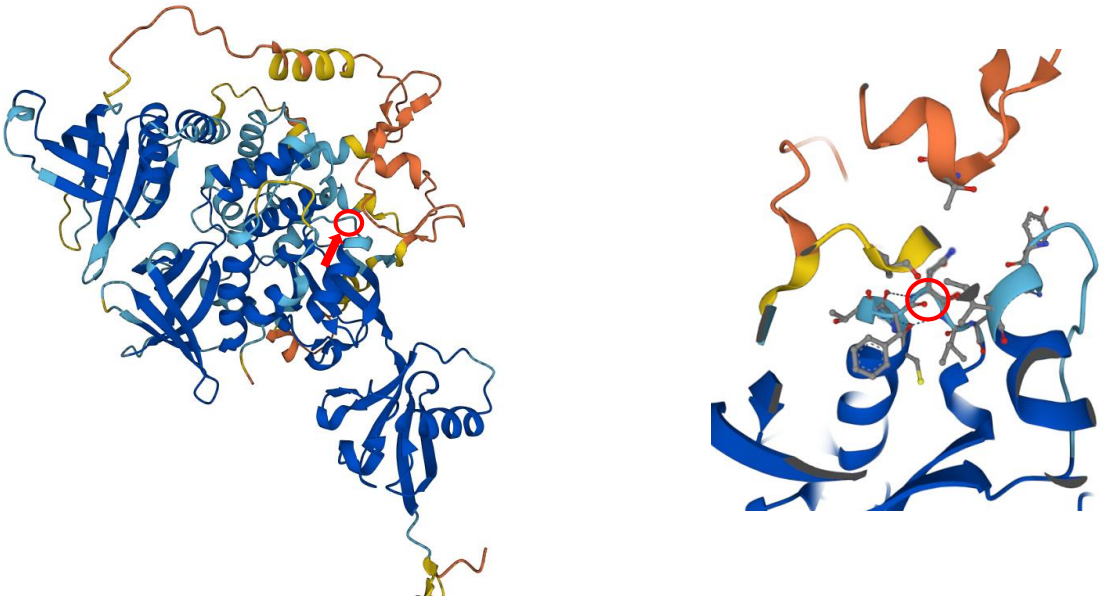

Figure 3.4.: Tertiary structure of the *FERMT2* protein with the variant of family #1. In the left image, the red arrow and the red circle show the *FERMT2* variant c.392A>G, p.(Asn131Ser). The right image shows the chemical connections at this position to other regions in the protein. Copied from: https://alphafold.ebi.ac.uk/entry/Q96AC1

In gnomAD and in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) this variant was not listed. To obtain more information about these and the other variants, the search engine VarSome (https://varsome.com/) was used. It is a tool used to share worldwide expertise on human variants and to analyze the impact of human genetic variations.

According to the ACMG classification, this variant was classified as "benign". The REVEL (rare exome variant ensemble learner) score for the variant c.392A>G is about 0.1749 and is classified “benign”. The REVEL program provides an ensemble method which predicts the pathogenicity of missense variants based on a combination of results from 13 individual tools. (N. M. Ioannidis et al., 2016) These include MutPred, PolyPhen-2, SIFT, MutationTaster, LRT, GERP++, phyloP, among others. (N. M. Ioannidis et al., 2016) The score for a single missense variant can range from 0 to 1, with higher scores reflecting the greater likelihood that the variant is disease-causing. (Nilah M Ioannidis et al., 2016) The phyloP score measures evolutionary conservation at individual alignment sites and is 5.953 for this variant. The higher the score, the more conserved the site. (Pollard et al., 2010) In summary, the family #1 variant is situated in a very highly conserved location, with the most comprehensive pathogenicity program predicting a benign effect. However, replacing the large amino acid asparagine with the small amino acid serine may have a disruptive effect.

3.2.2 Exome sequencing results of family #2

Via GeneMatcher a candidate variant in *FERMT2* in a family from Portugal was found. This is a family (#2) with 41 family members, of which there are ten affected individuals with available DNA samples. In two affected individuals (III.14 and IV.1), exome sequencing was performed prior to this work by the Behavioral and Molecular Lab, Universidade de Minjo, Braga, Portugal, which detected the variant. In both individuals, an identical variant was found in the *FERMT2* gene: NM_006832.3:c.1538C>T p.(Thr513Met) at genomic position chr14:g.53331183G>A. Thereafter, Sanger sequencing was performed in all individuals for whom DNA was available. This variant was also present in an additional six affected individuals (II.2, II.5, III.2, III.9, IV.2, IV.8) and one non-affected (III.7), see Figure 3.5. Family members with the variant c.1538C>T p.(Thr513Met) have been marked with a blue vertical arrow. This variant is present in 689 alleles on gnomAD (12/2022) for a total allele number of 282814 (= approx. 0.24% of all alleles). This corresponds to an allele frequency of 0.24% and is thus rather frequent and therefore predicted to be non-pathogenic. In ClinVar this variant is not listed. According to the ACMG classification, this variant has to be classified as "benign". It is located in a moderate conserved site (Phylop = 3.791) and mutation prediction tools are predicting again a benign effect (REVEL = 0.071).

No other pathogenic variants were found in this family in genes associated with epilepsy. Figure 3.6 shows the position of this *FERMT2* variant in the tertiary structure on the left and the chemical links or interactions to neighboring structures in the protein on the right. There are hydrogen bonds from this site to an alpha helix and the variant could therefore again lead to alterations in protein function. The variant c.1538C>T p.(Thr513Met) is located in the FERM domain (F = 4.1 protein, E = ezrin, R = radix, and M=moesin). (Chishti et al., 1998) This is a protein module involved in the localization of cytoplasmic protein to the plasma membrane. (Chishti et al., 1998; Pearson et al., 2000) These domains are found primarily in proteins associated with the cytoskeleton, as is the case of *FERMT2*. (Chishti et al., 1998; Pearson et al., 2000) The replacement of the polar and neutral amino acid threonine with the nonpolar and hydrophobic methionine could lead to altered chemical interactions and thus to modified protein function. Since the variant is located in the surface region and thus on the external side, the structure of the protein could be changed in such a way that the protein function is compromised.

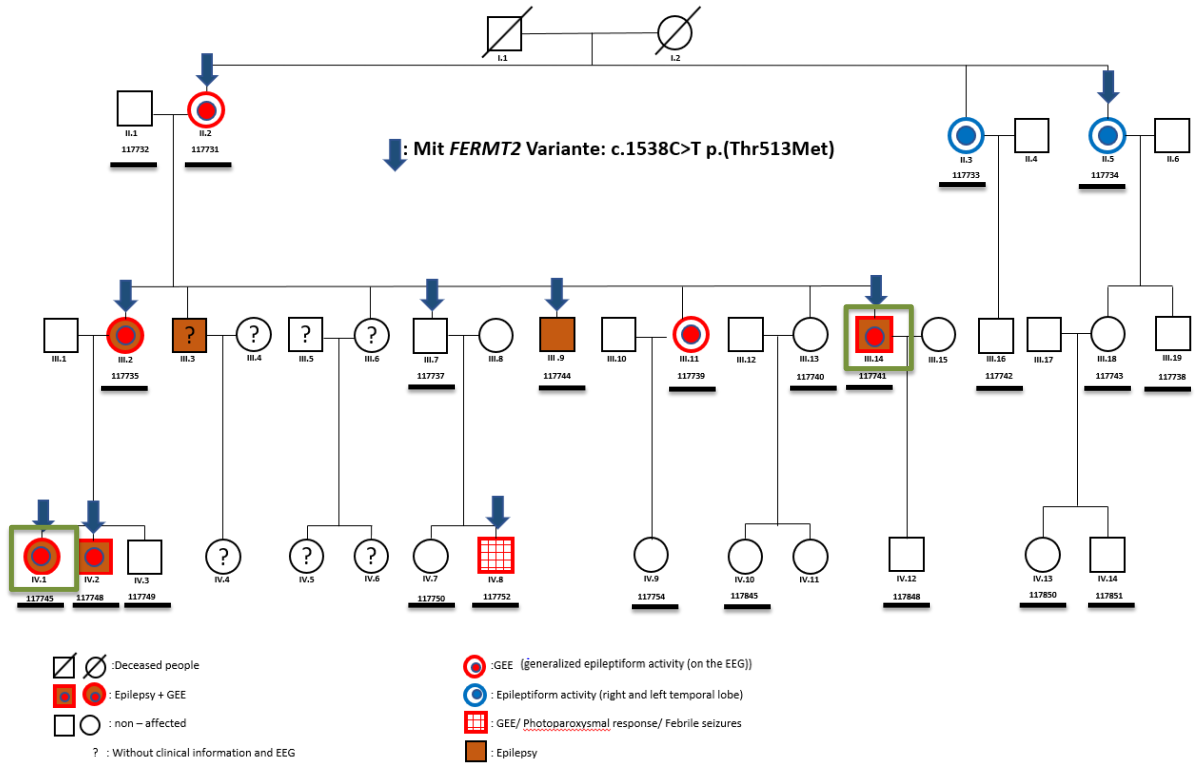

Figure 3.5.: Family members with *FERMT2* variant of family #2. Squares symbolize male family members, circles female family members. Deceased individuals are marked with a crossed-out symbol. Individuals with diagnosed epilepsy are highlighted in brown, those with GEE are outlined in red. Individuals with epileptic activity (without diagnosed epilepsy and without generalized epileptiform activity) in the EEG are marked blue. White symbols indicate healthy individuals, roman numerals (I-IV) the generations and Arabic numerals (1-19) indicate the individual persons in the generations. The numbers below the roman numerals represent the DNA codes of individuals with available DNA. The blue arrow marks patients with the *FERMT2* variant c.1538C>T p(Thr513Met). The green square shows the two patients in whom exome sequencing was performed in Portugal. The black underlines mark all individuals tested in Sanger sequencing. Abbreviations: GEE = generalized epileptiform activity; EEG = Electroencephalogram.

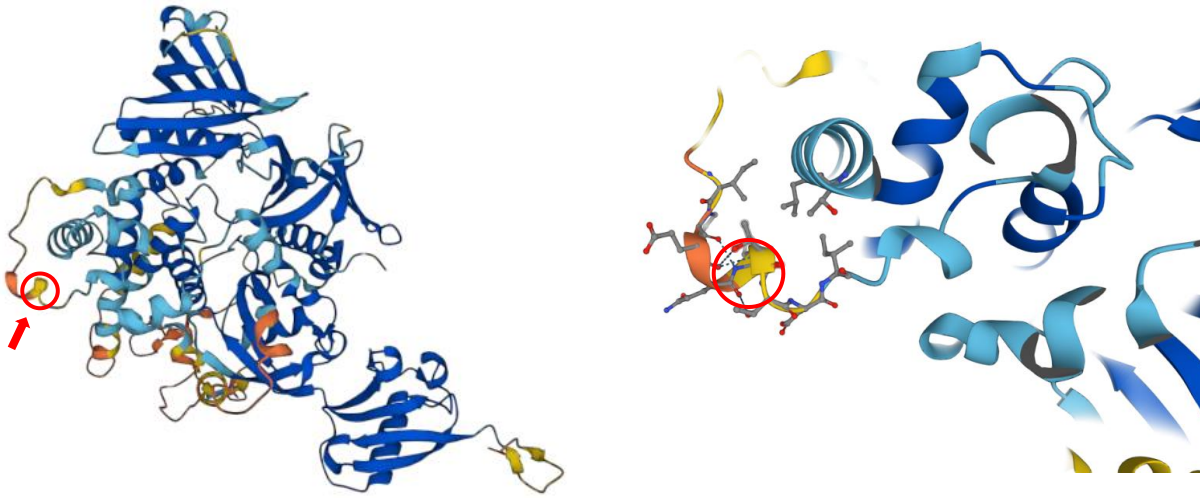

Figure 3.6.: Tertiary structure of the *FERMT2* protein with the variant of family #2. In the left image, the red arrow and the red circle show the *FERMT2* variant c.1538C>T p.(Thr513Met). The right image shows the chemical connections at this position to other regions in the protein. Copied from: https://alphafold.ebi.ac.uk/entry/Q96AC1

Table 1 shows all patients diagnosed with epilepsy carrying the *FERMT2* variant. Those with generalized epileptiform activity on EEG (GEE), epileptiform activity, febrile seizures, or photoparoxysmal response carry the variant in only three out of five cases. Patients II.3 and III.11 would also have been expected to have this variant, as these individuals have epileptiform potentials in the EEG. Patient III.7 carries the variant without being affected. The reason for this could be an incomplete/reduced penetrance since his affected son (IV.8) is also a carrier of the variant.

Based on the present results, linkage of *FERMT2* is not certain because the variant does not segregate. The question arises whether there is another locus that was missed by exome sequencing. For segregation, the variant would be expected to be present in all affected patients and absent in all unaffected patients. The allele frequency is also rather high and thus likely non-pathogenic. Therefore, linkage analysis was subsequently performed on family #2 (see 3.3).

Variant detected	Epilepsy	GEE/ epileptiform activity/febrile seizures/ photoparoxysmal response	Non-affected
FERMT2:	III.2	II.2	II.1
c.1538C>T	III.9	II.3	III.7
p.(Thr513Met)	III.14	II.5	III.13
	IV.1	III.11	III.16
	IV.2	IV.8	III.18
			III.19
			IV.3
			IV.7
			IV.9
			IV.10
			IV.12
			IV.13
			IV.14

Table 1.: Segregation analyses of array comparative genomic hybridization (aCGH) findings and Sanger sequencing. Individuals carrying the *FERMT2* variant are highlighted in red.

3.2.3 Exome sequencing results of family #3

The *de novo* missense variant NM_006832.3:c.1976G>A, p.(Arg659His) at the genomic position chr14:g.53325162C>T, which is not listed in gnomAD or in ClinVar, was identified as part of a diagnostic process by deCODE genetics in Reykjavik, Iceland, and matched via GeneMatcher in 06/2022. Communication with Reykjavik was by email.

The variant is located in a different domain than in the described index patient from family #1, but also in the FERM domain as described in family #2 (see 3.1.2 and 3.2.2). Figure 3.7 shows the variant of the index patient in the tertiary structure of the *FERMT2* protein. There are chemical bonds to two alpha helices and one beta leaflet that can lead to modifications in protein function. The location of the variant in the FERM domain, as well as the exchange of the basic larger amino acid arginine for the smaller basic amino acid histidine could lead to a change in protein function due to altered chemical bonds. The variant is located quite centrally in the protein and could therefore also act locally disruptive.

The phenotypic spectrum in the index patient of family #3 is clearly different from the other patients with detected *FERMT2* variants. Nevertheless, an association to the *FERMT2* gene cannot be excluded in this case either.

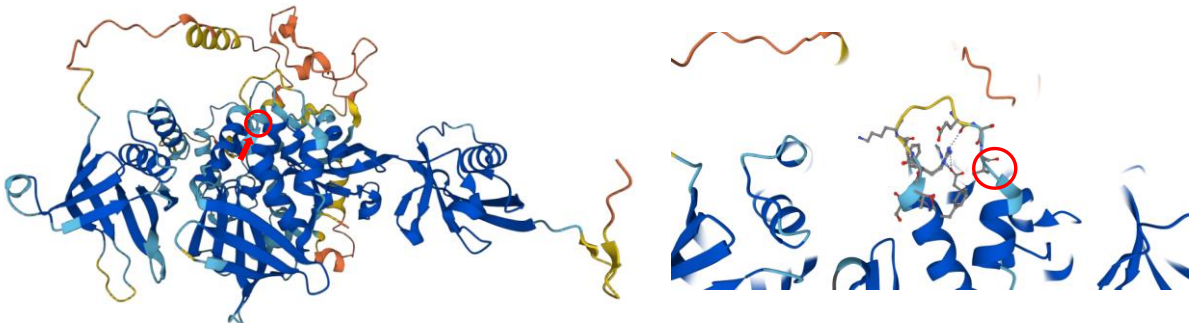

Figure 3.7.: Tertiary structure of the *FERMT2* protein with the variant of family #3. In the left image, the red arrow and the red circle show the *FERMT2* variant c.1976G>A, p.(Arg659His). The right image shows the chemical connections at this position to other regions in the protein. Copied from: https://alphafold.ebi.ac.uk/entry/Q96AC1

According to the ACMG classification, this variant has to be classified as " variant of uncertain significance".

The following argues for a pathogenic effect of this variant: The PhyloP score is 7.799 and the REVEL is 0.5619 with an uncertain impact on the pathogenicity of this missense variant. So, this variant is located at a very highly conserved site. An uncertain or even pathogenic effect is likely due to the location of the variant in the FERM domain, as well as the replacement of the larger amino acid with a smaller one.

Against a pathogenic effect speaks the fact, that mutation prediction tools predominantly predict an uncertain effect.

3.3 Linkage analysis

What LOD score could theoretically be achieved with family #2 for an autosomal dominant inheritance mechanism was calculated using the following approximation:

$$\text{Z (LOD score)} = \log_{10}\left(\frac{1}{0{,}5^{n}}\right)$$

where n corresponds to the number of informative meioses from which it can be deduced which allele was inherited by the offspring when one parent is heterozygous for both the marker and the disease-causing allele. Individuals II.3 and II.5 from family #2 were not included in the linkage analysis, because none of the offspring were affected and the clinical data were inconclusive. Only twelve of the total 23 individuals with available DNA were included because the computational power was limited. Therefore, a reduced dataset was used. This resulted in ten informative meiosis (seven affected and three unaffected) and using this setting, the estimation of the theoretically achievable LOD score results in a value of 3.01 which is considered significant (Lathrop et al., 1984; Wiltshire et al., 2001):

$$\text{Z (LOD score)} = \log_{10}\left(\frac{1}{0{,}5^{10}}\right) \approx 3.01$$

Including only the affected individuals of family #2 results in a total of seven informative meioses and thus a LOD score of 2.11:

$$\text{Z (LOD score)} = \log_{10}\left(\frac{1}{0{,}5^{7}}\right) \approx 2.11$$

In the first case, the LOD score is significant (>3), while in the last case no significant value can theoretically be reached.

Based on the SNP genotyping data, as described in chapter 2.5, a parametric linkage analysis was performed. The Merlin program was used to perform the analysis as described in 2.5.3.2

(see 8.3, supplemental Table 19 gives an overview about the family structure and the affection status of the family members, which was also used as an input file).

Individual family members were added or removed in the analyses to see how this affects the LOD score. The individuals framed in black in Figure 3.2 were used in the further analyses (12 in total). Table 2 lists the chromosomes examined, their length in cM and the number of markers used.

Chrom.	Chrom. Length (cM)	Number of used markers	Chrom.	Chrom. Length (cM)	Number of used markers
1	274,76	1160	12	171,21	724
2	266,80	1126	13	129,36	563
3	221,41	862	14	125,40	496
4	205,41	899	15	133,64	502
5	206,29	909	16	130,74	534
6	189,60	848	17	137,39	562
7	192,18	757	18	121,51	520
8	165,82	698	19	110,80	445
9	159,79	676	20	114,59	464
10	178,63	680	21	76,79	266
11	153,59	693	22	75,67	284

Table 2.: Information about the chromosomes examined. The individual chromosomes are listed in the left column, the length of each chromosome is shown as cM in the middle column, and the number of markers used per chromosome is presented on the right.

No region exceeded the LOD score of 3.0, (LOD scores > 3 are considered significant). (Lathrop et al., 1984; Wiltshire et al., 2001) However, LOD scores above two were achieved in two regions.

Figure 3.8 and Figure 3.9 show the most significant results of the linkage analysis. A total of three regions on two chromosomes could be identified which indicate linkage. Region 1 (on chromosome 2 at around 124.5-127.0 cM) has a LOD score of about 2.7 and region 2 (on chromosome 2 at around 238.7-253.5 cM) a LOD score of about 1.7. The third identified region (on chromosome 18 at around 81.4-91.5 cM) has a LOD score of more than 2.5. As expected, the *FERMT2* region on chromosome 14 (at around 49.2-51.2 cM) showed no significant results (LOD score 0.6, see Figure 3.10).

Chromosome 2:

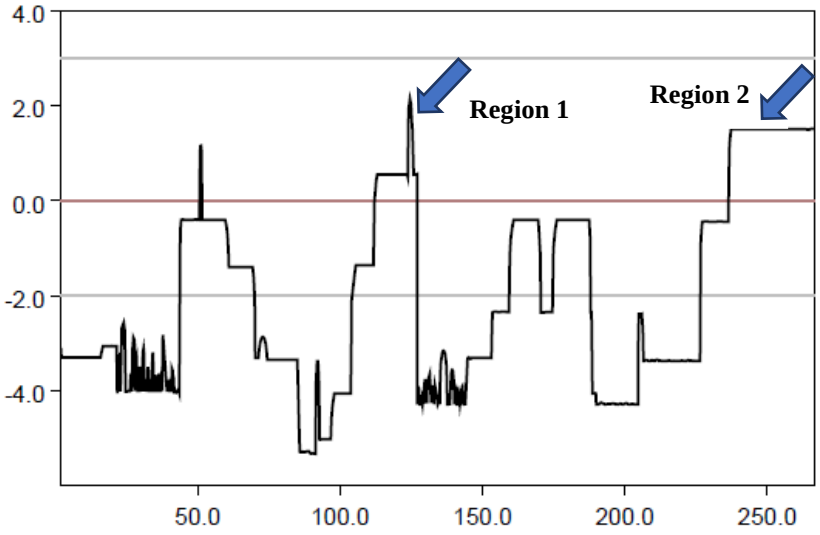

Figure 3.8.: Parametric linkage analysis of family #2 for the dominant model of chromosome 2. The position on the chromosome is given in cM on the x axis and the corresponding LOD score on the y axis. The lower grey line marks a LOD score of -2, the middle purple bar a LOD score of 0 and the top grey line a LOD score of 3. The blue arrows point to the regions with the highest LOD scores of all chromosomes examined.

Chromosome 18:

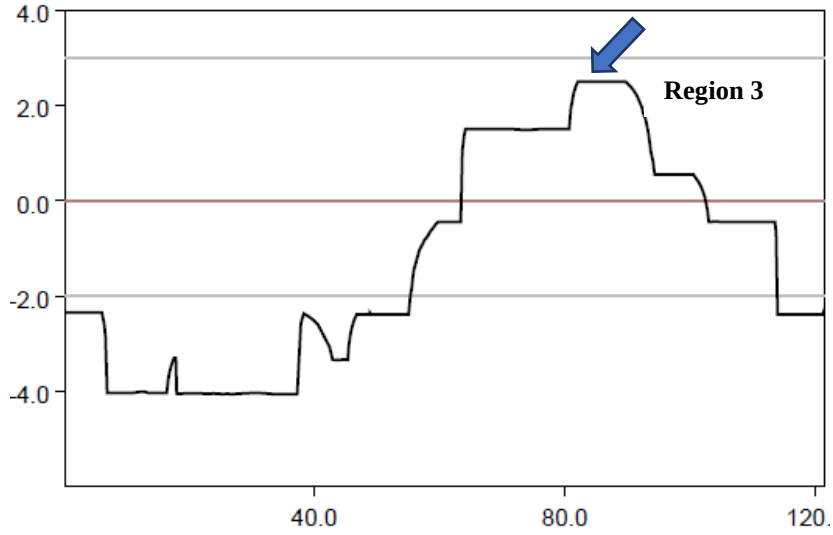

Figure 3.9.: Parametric linkage analysis of family #2 for the dominant model of chromosome 18. The position on the chromosome is given in cM on the x axis and the corresponding LOD score on the y axis. The lower grey line marks a LOD score of -2, the middle purple bar a LOD score of 0 and the top grey line a LOD score of 3. The blue arrow points to the regions with the highest LOD scores of all chromosomes examined.

Chromosome 14:

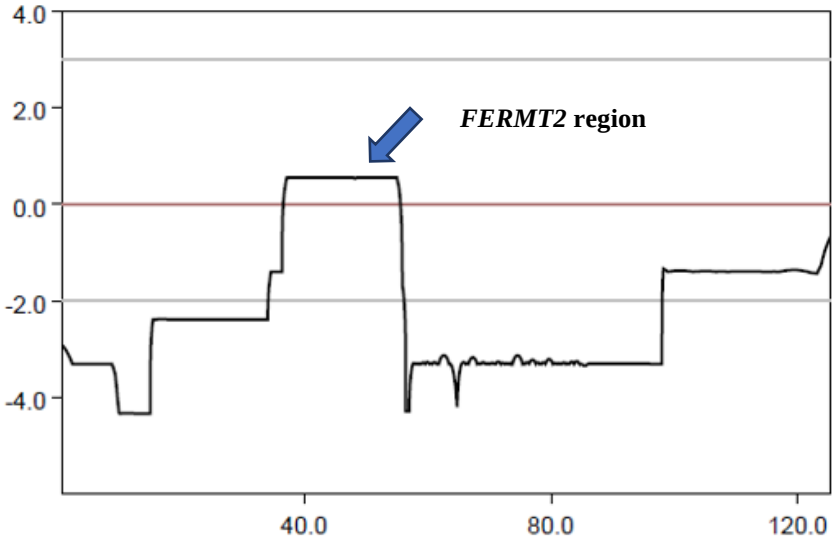

Figure 3.10.: Parametric linkage analysis of family #2 for the dominant model of chromosome 14. The position on the chromosome is given in cM on the x axis and the corresponding LOD score on the y axis. The lower grey line marks a LOD score of -2, the middle purple bar a LOD score of 0 and the top grey line a LOD score of 3. The blue arrow indicates the *FERMT2* region.

3.4 Haplotyping

To verify and narrow down the possible linkage to the above-mentioned regions, a haplotyping of the three areas and the *FERMT2* region was performed following the genome-wide linkage analysis. In addition, genotyping data from eleven further family members (which could not be included in the linkage analysis before because of the limited computing power) was integrated in the haplotyping to adequately trace the segregation of the alleles in the pedigree.

The results of haplotyping were presented for the segregation analyses (see 2.5.4.2). The alleles were sorted according to their parental origin, resulting in specific haplotypes. In this way, the haplotypes of affected and healthy family members in each analyzed region could be compared. The reference genome refers to hg19 in the following sections. Figure 3.11 shows the haplotypes of all family members in the *FERMT2* region. All individuals with the *FERMT2* variant identified in the Sanger and exome sequencing also carry the same haplotype at this position. Affected individuals II.2, II.5, III.2, III.9, III.14, IV.1, IV.2, and IV.8 have the same haplotype in the region of the *FERMT2* gene. II.3 and III.11 do not have this haplotype, although they are affected and the healthy person III.7 has inherited this haplotype to the diseased son (IV.8).

In the following, the results of the haplotyping of the three regions on chromosomes 2 and 18 are described in more detail. Given the high linkage on chromosome 2 and 18 for family #2 showing a dominant inheritance pattern, the haplotypes of the family members were reconstructed to further map the disease locus. No haplotype could be identified for any region, which occurs in all affected patients but not in healthy individuals. However, incomplete clinical information for individuals complicates the tracing of haplotypes. Figure 3.12, Figure 3.13 and Figure 3.14 show the results of these regions in the pedigree. Haplotypes that could be detected in over more than one generation or in at least two individuals are highlighted in colour to follow segregation across generations.

In region 1 (see Figure 3.12) on chromosome 2, a common haplotype (rs68012394 – GSA-rs1835328, see Table 24) composed of 12 SNP markers was detected in most affected individuals (in II.2, II.3, III.2, III.9, III.11, III.14, IV.1 and IV.2) and was marked in yellow. Other affected individuals (II.5, IV.8) do not carry this haplotype. III.7 and IV.9 carry the haplotype and were not affected at the time of neurologic examination. This region (containing twelve markers) could not be further delineated and corresponds to approximately

2.5 Mb in length. The haplotype does not segregate at this point. Discrepancies exist in several individuals: In II.5, the haplotype would also be expected. III.7 carries the same haplotype but does not pass it on to the affected patient IV.8, so there are discrepancies here as well. IV.9 does not match the presumed inheritance since she is healthy.

In region 2 (see Figure 3.13) on chromosome 2, a common haplotype (rs28688824 – rs6716187) composed of 65 SNP markers was detected in most of the affected individuals (II.2, II.5, III.2, III.9, III.11, III.14, IV.1, IV.2 and IV.8) and was again marked in yellow, whereas the affected individual II.3 did not have this haplotype. But this part of the pedigree is not very informative because, on the one hand, the clinical data are incomplete and, on the other hand, none of the subsequent generations has seizures or epilepsy. Individuals without epilepsy, seizures, or epileptiform activity (III.7, IV.3, IV.12) also carry this haplotype completely. However, III.7 passes it on to the affected child implying reduced penetrance. IV.3 and IV.12 may have been too young for clinical manifestation at the time of examination. Again, clinical re-evaluation would be useful. III.18, IV.7, IV.9, and IV.14 are also reported healthy and carry the haplotype incompletely. III.18 is only consistent for the first four markers of the haplotype (rs28688824 – rs2344309), IV.7 has all but the last five markers equal (rs28688824 – rs61544914). In IV.9 only the last five are the same (rs12479385 – rs6716187) and in IV.14 the first three (rs28688824 – rs2853386). Thus, recombination of alleles occurred here. Region 2 (see Table 25) could be narrowed down from 113 markers to 65 markers (approximately 14.8 Mb) by analysis of the FLOW file. In summary, there are discrepancies regarding haplotype segregation in this region as well. Nevertheless, due to the distribution of haplotypes, region 2 is probably the most interesting and best for the identification of possible candidate genes.

On chromosome 18, region 3 (see Figure 3.14), a common haplotype (II.2, III.2, III.9, III.11, III.14, IV.1 and IV.2) (rs11874251 – rs7229020) with 45 SNP-markers could be detected in most affected individuals and was again marked in yellow. Affected individuals II.5 and IV.8 carry this haplotype incompletely. II.5 has only the first and last five markers identical (rs11874251 – rs7244658 and GSA-rs605818 – rs7229020), IV.8 has all but the last five alleles equal (rs11874251 – rs2156175). Patient II.3 with epileptiform activity does not carry this haplotype (a phenocopy could be an explanation) and unaffected individuals carry it either completely (III.7, IV.7: rs11874251 – rs7229020) or partially (III.13: GSA-rs605818 – rs7229020), (III.19: rs11874251 – rs7244658 and GSA-rs605818 – rs7229020), (IV.3:

rs11874251 – rs2156175) and (IV.9: rs11874251 – rs7244658)). Again, recombination of alleles occurred. Finally, region 3 (see Table 26) could also be narrowed down from 218 markers to 45 (approximately 10 Mb) by analyzing all haplotypes, but again, the haplotype does not segregate.

3.4.1 *FERMT2* region, hg19, (approx. 49,2-51,2 cM)

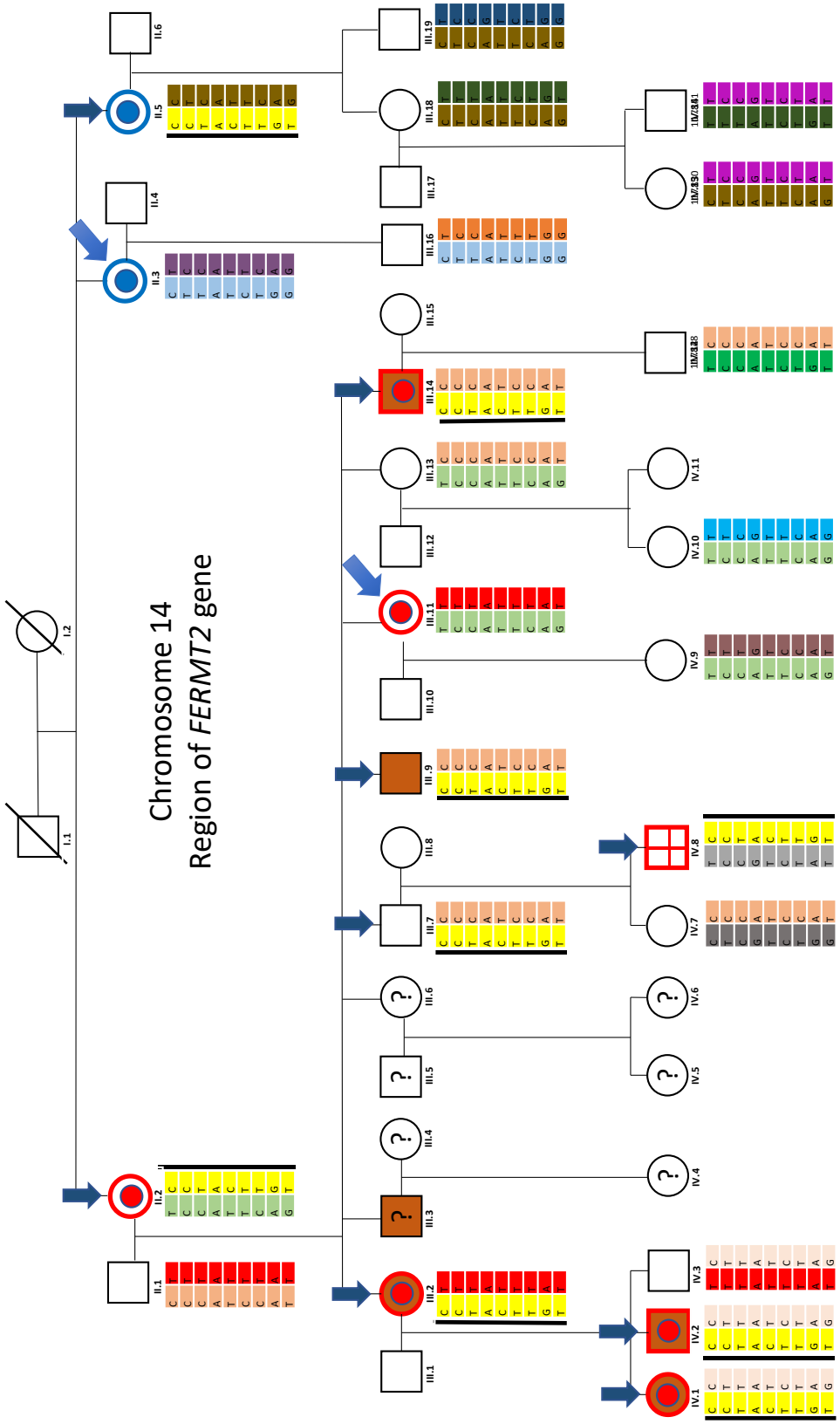

Figure 3.11.: Results of haplotyping of the *FERMT2* gene region (chromosome 14). Maternal alleles are noted on the left, paternal alleles are noted on the right. The haplotype found in all individuals with *FERMT2* variant are highlighted in yellow and marked with a black line on the side. Haplotypes that could only be found in one individual are highlighted in light grey. Individuals not matching the suspected autosomal dominant inheritance pattern are marked with an oblique light blue arrow ⬇: With *FERMT2* variant

3.4.2 Region 1 on chromosome 2, hg19, (124,52-126,98 cM)

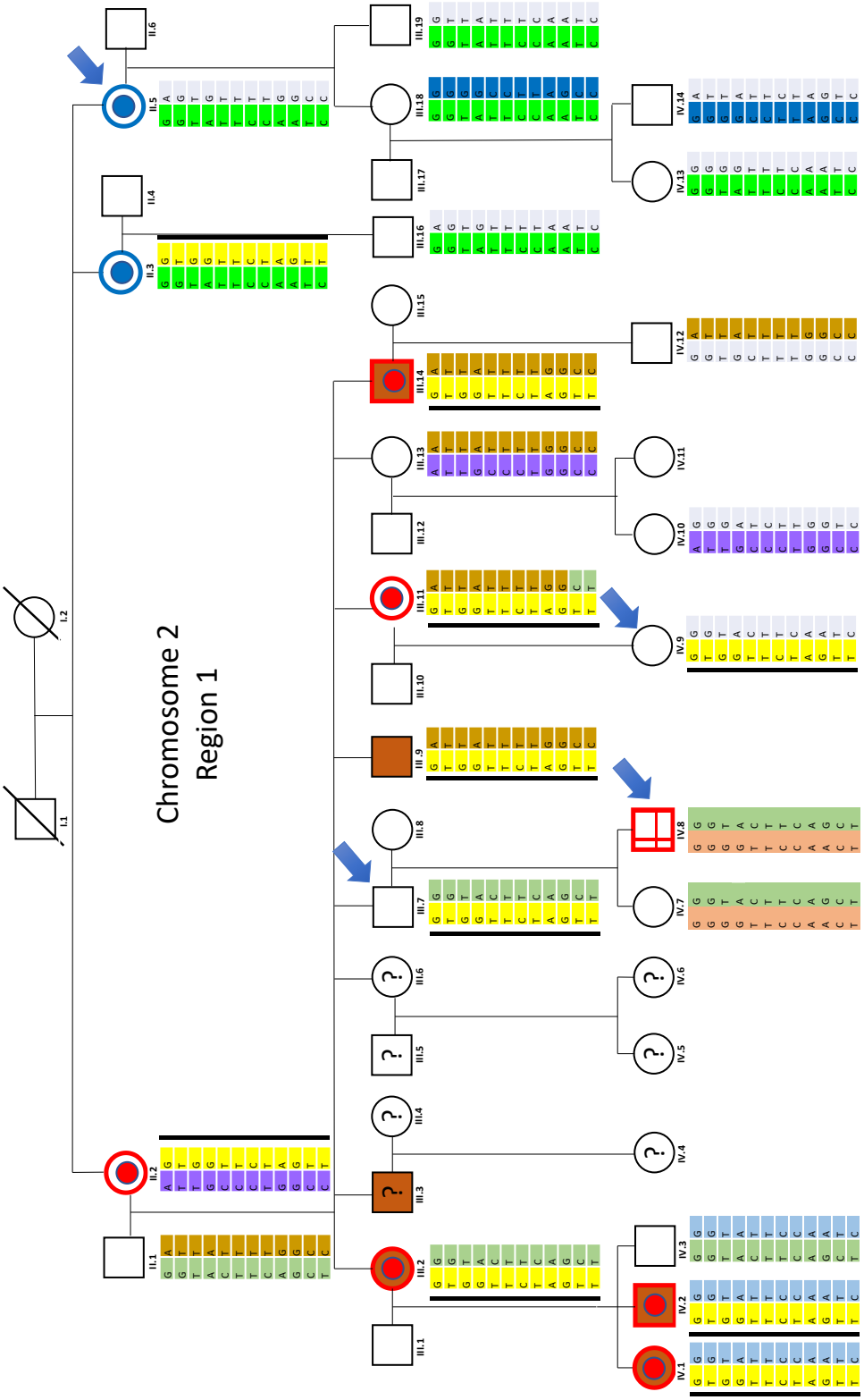

Figure 3.12.: Results of haplotyping of region 1 (chromosome 2). Maternal alleles are noted on the left, paternal alleles are noted on the right. The haplotype found in all affected persons (except for individual IV.8 and II.5) is highlighted in yellow and marked with a black line on the side. Haplotypes that could only be found in one individual are highlighted in light grey. Individuals not matching the suspected autosomal dominant inheritance pattern are marked with a blue arrow.

3.4.3 Region 2 on chromosome 2, hg19, (238,71-253,50 cM)

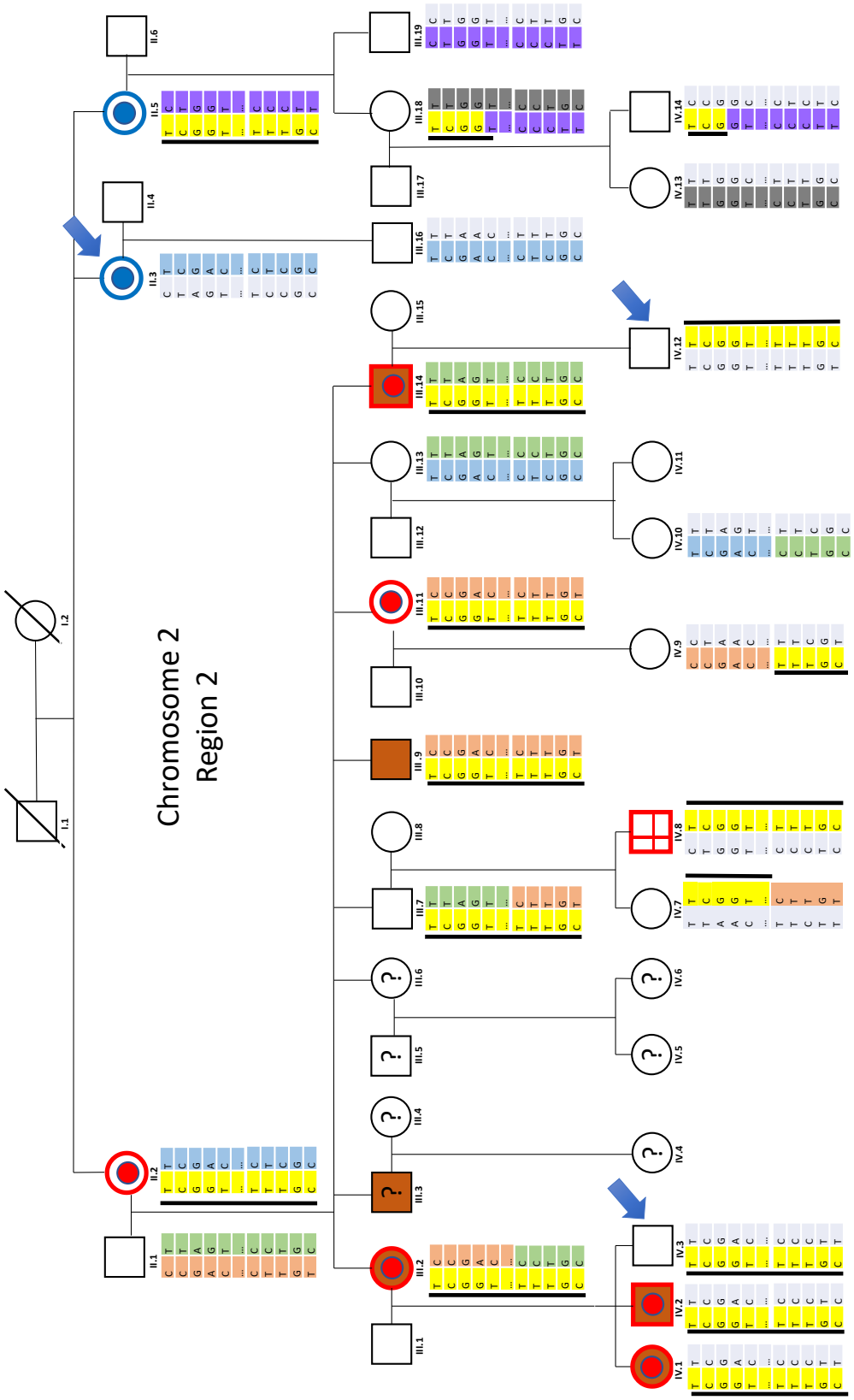

Figure 3.13.: Results of haplotyping of region 2 (chromosome 2). Maternal alleles are noted on the left, paternal alleles are noted on the right. The haplotype found in all affected persons (except for individual II.3) is highlighted in yellow and marked with a black line on the side. Haplotypes that could only be found in one individual are highlighted in light grey. Individuals not matching the suspected autosomal dominant inheritance pattern are marked with a blue arrow.

3.4.4 Region 3 on chromosome 18, hg 19, (81,42-91,51 cM)

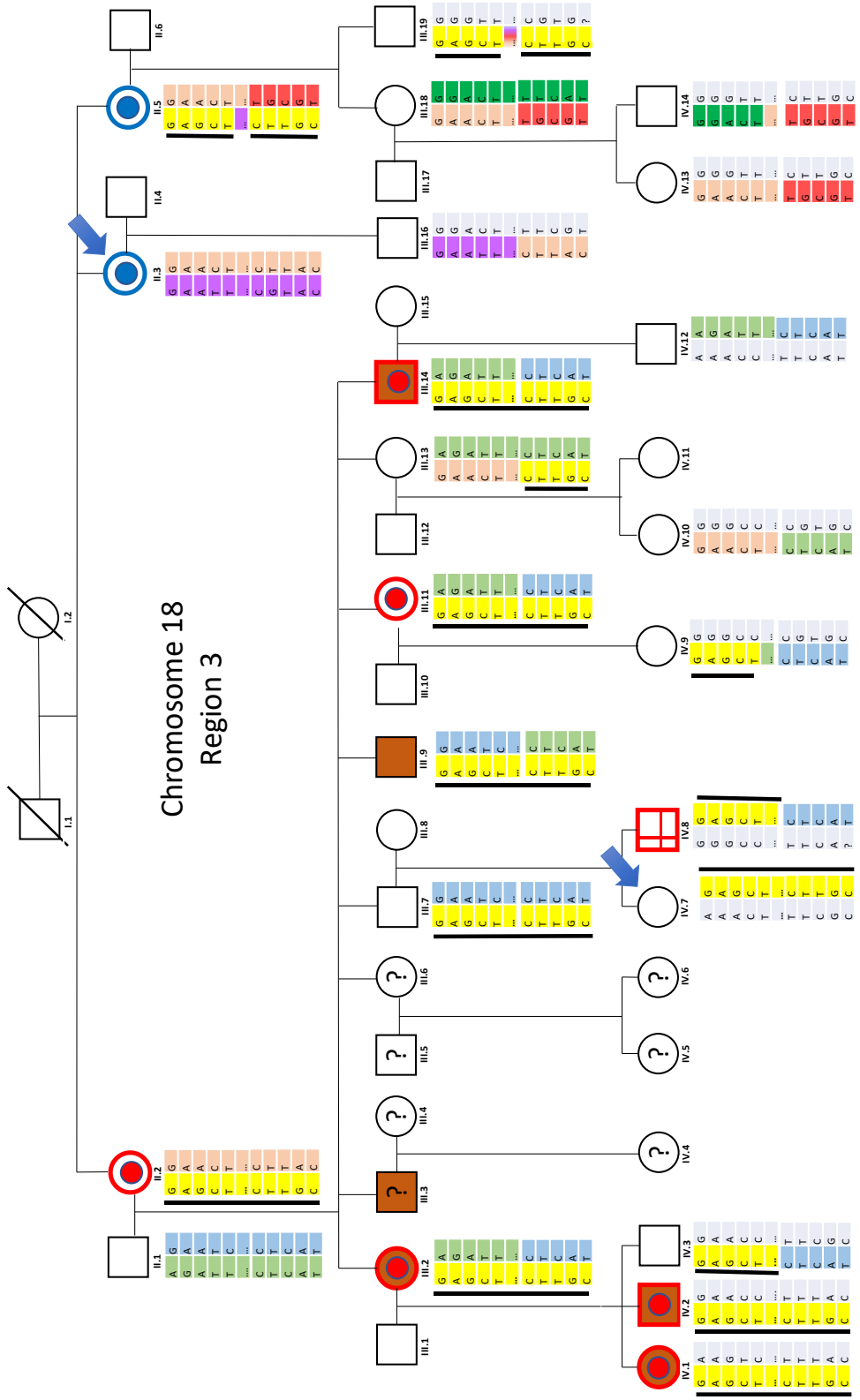

Figure 3.14.: Results of haplotyping of region 3 (chromosome 18). Maternal alleles are noted on the left, paternal alleles are noted on the right. The haplotype found in all affected persons (except for individual II.3) is highlighted in yellow and marked with a black line on the side. Haplotypes that could only be found in one individual are highlighted in light grey. Individuals not matching the suspected autosomal dominant inheritance pattern are marked with a blue arrow.

3.5 Search for candidate genes

To elicit possible candidate genes, DECIPHER (https://www.deciphergenomics.org/) was used to identify all genes located in each region that were narrowed down by haplotyping. Genes that were listed in OMIM or had a LOEUF (Loss-of-function observed/expected upper bound fraction) score ≤ 0.35 were examined in more detail to determine whether an association with epilepsy might exist. The LOEUF score is an estimate of the observed/expected ratio. (Karczewski et al., 2020) It allows each gene to be ranked within a continuous range of tolerance to inactivation. Low LOEUF scores indicate low tolerance to inactivation, whereas high LOEUF scores indicate higher tolerance. The advantage over the pLI is that the LOEUF can be used as a continuous measure and not as a dichotomous one (e.g., pLI > 0.9). (Karczewski et al., 2020; Karczewski et al., 2019) Protein expression and localization information was described according to either the Human Protein Atlas (HPA) (http://www.proteinatlas.org/) or the Genotype-Tissue Expression (GTEx) database (https://gtexportal.org/home/). A Z-score provides information about the tolerance to missense variants in the genes of interest. (Samocha et al., 2017) High Z-scores indicate low tolerance to missense variants, whereas low scores suggest high tolerance. (Samocha et al., 2017)

In the next sections, all genes are listed in tables for which there might be an association of epilepsy or seizures in the families described. The appendix lists all genes (including those not previously associated with seizures or epilepsy) located in each region that could be narrowed down by haplotyping (see 8.4).

3.5.1 Candidate genes of the locus on chromosome 2 – Region 1

A total of six genes were detected in region 1 that could be related to the seizures and epilepsy in family #2: *ACTR3*, *CKAP2L*, *DPP10*, *IL1A*, *IL1B* and *IL1RN*. For each gene, information on LOEUF score and Z-score is included in a table, whether the genes are listed in OMIM and associated with diseases, information on expression, protein function, studies associated with the gene, and gene-gene interactions of interest.

Gene	***ACTR3* (Actin Related Protein 3)**
LOEUF score	0.12
Z-score	3.61
OMIM listed (1) and phenotypes (2)	(1): yes (2): /
Expression	- in almost every Organ (heart, gastrointestinal tract, brain,...); located mainly in the cytosol, nucleus and cytoskeleton
Protein function	- important component of the ARP2/3 complex located at cell surfaces > essential for cell shape and motility through lamellipodial actin formation and protrusion (Welch et al., 1997)
Associated studies	- Examination of the proteome of brain samples from epilepsy and control cases in humans > *ACTR3* involved in epithelial junction remodeling during an epileptic seizure > role in the reorganization of cortical microarchitecture; implications for remodeling of neuronal circuits or brain vasculature (Pires et al., 2021)
Relevant gene interactions	*STXBP1* > associated with seizures and epilepsies; involved in fusion of synaptic vesicles; associated with "Developmental and epileptic encephalopathy 4" (OMIM #612164) > early onset of tonic seizures in childhood, severely impaired psychomotor development, quadriplegia and poor or absent speech,… (Deprez et al., 2010)

Table 3.: Gene information about *ACTR3*.

Gene	***CKAP2L* (cytoskeleton associated protein 2 like)**
LOEUF score	0.63
Z-score	-0.09
OMIM listed (1) and phenotypes (2)	(1): yes (2): "Filippi syndrome" (OMIM #272440): seizures, microcephaly, facial dysmorphia or growth retardation (Hussain et al., 2014)
Expression	- very low levels in the brain
Protein function	- essential for division of neural stem and progenitor cells (T. Yumoto et al., 2013) - component of the human centrosome; located in the spindle, midbody, and spindle pole (Jakobsen et al., 2011)
Associated studies	- knockdown mouse models > Overexpression of *Ckap2l* > increased mitosis in ventricular/subventricular zones; excess of *Ckap2l* reduced proliferation rate of neural progenitor cells and increased their cell cycle exit > apoptosis/abnormal differentiation of neural progenitor cells could lead to reduction in the size of embryonic ventricular/subventricular zones - Deregulated levels could affect neurons formed during prenatal brain development > phenotype with reduced brain size or impaired neuronal function (Takahito Yumoto et al., 2013)
Relevant gene interactions	/

Table 4.: Gene information about *CKAP2L*.

Gene	***DPP10* (Dipeptidyl peptidase)**
LOEUF score	0.15
Z-score	1.22
OMIM listed (1) and phenotypes (2)	(1): yes (2): /
Expression	- In almost every organ (heart, gastrointestinal tract, brain,...); located mainly in the cytosol, nucleus, and cytoskeleton
Protein function	- component of the neuronal Kv4 channel - enhances cell surface expression of potassium channel KCND2 and modulates its activity and gating properties (Bezerra et al., 2015; Jerng et al., 2004)
Associated studies	- subthreshold transient K+ current affects neuronal excitability; Reduction or loss of this current may contribute to higher excitability and susceptibility to the development of epileptiform activity (Bezerra et al., 2015) - *DPP10* as a potential therapeutic target for certain forms of epilepsy, such as temporal lobe epilepsy (TLE) (Bezerra et al., 2015; Singh et al., 2006)
Relevant gene interactions	(1): *KCND3* > encodes voltage-gated potassium channels (Smets et al., 2015) (2): *KCND2* > involved in potassium channel formation and detected in twins associated with infantile-onset severe refractory epilepsy (Lee et al., 2014) (3): *KCNQ3* > expressed in the brain; encodes M channels; is associated with autosomal dominant inherited "Seizures, benign neonatal, 2" (OMIM #121201) (Fister et al., 2013; Rogawski, 2000); the M channel affects potassium conductance and thus the excitability of neurons (Soh et al., 2014)

Table 5.: Gene information about *DPP10*.

Gene	***IL1A* (Interleukin 1 Alpha) (*1)** ***IL1B* (Interleukin 1 Beta) (*2)** ***IL1RN* (Interleukin 1 Receptor Antagonist) (*3)**
LOEUF score	(*1): 1.31, (*2): 0.65, (*3): 0.85
Z-score	(*1): 0.54, (*2): 1.1.5, (*3): 0.01
OMIM listed (1) and phenotypes (2)	(1): yes (2): - (*1): / - (*2): "Gastric cancer risk after H. pylori infection" (OMIM #613659) - (*3): "Interleukin 1 receptor antagonist deficiency" (OMIM #612852) - (*3): "Gastric cancer risk after H. pylori infection" (OMIM #613659) - (*3): "Microvascular complications of diabetes 4" (OMIM #612628)
Expression	- especially in the esophagus and lungs
Protein function	- (*1): involved in various immune reactions and inflammatory processes (Cohen et al., 2015) - (*2): involved in inflammatory reactions and cellular processes (cell proliferation, differentiation, and apoptosis) (Van Damme et al., 1985) - (*3): inhibits activities of *IL1A* and *IL1B*; modulates a variety of interleukin-1-related immune and inflammatory reactions, particularly in the acute phase of infection and inflammation (Pinheiro et al., 2021)
Associated studies	(*1/*2/*3): - mouse model: Investigation of the role of interleukin-1 signalling after post-traumatic epilepsy for pediatric brain injury (Semple et al., 2017) - Mice: sham surgery or controlled cortical beat at postnatal day 21; then treated with either an IL-

	1 receptor antagonist (IL-1Ra) or vehicle: > convulsant pentylenetetrazol > mice with traumatic brain injury (TBI) > increased seizure response compared to sham mice at two weeks and three months after injury; associated with abnormal sprouting of mossy fibers in the hippocampus; Seizure susceptibility was reduced two weeks after TBI compared with vehicle by treatment with IL-1Ra susceptibility - treatment resulted in a reduction of astrogliosis in the hippocampus - IL-1Ra-TBI mice had fewer evoked seizures six months later compared with vehicle controls > IL-1 signalling as a mediator of posttraumatic astrogliosis and seizure susceptibility (Semple et al., 2017)
Relevant gene interactions	/

Table 6.: Gene information about *IL1A, IL1B, IL1RN*.

3.5.2 Candidate genes of the locus on chromosome 2 – Region 2

A total of four candidate genes could be identified for region 2: *ATG16L1*, *COL6A3*, *DGKD* and *KCNJ13*.

Gene	***ATG16L1* (Autophagy Related 16 Like 1)**
LOEUF score	0.32
Z-score	2.63
OMIM listed (1) and phenotypes (2)	(1): yes (2): "inflammatory bowel disease" (OMIM #611081)
Expression	- predominantly in the brain - particularly in the cerebellum; mainly located in lysosomes, endosomes in the cytosol, cytoskeleton, and nucleus
Protein function	- part of a large protein complex necessary for autophagy (Boada-Romero et al., 2013) - Related pathway: mTOR-signalling (Wang et al., 2018)
Associated studies	Investigation of the expression of *Atg16l1* in the hippocampus of rats with SE (Gan et al., 2017) - miR-96 was associated with significantly altered expression (overexpression) in SE - *Atg16l1* is a direct target of miR-96 > miR-96 expression is inverse to *Atg16l1* expression in the hippocampus of rats with SE-> miR-96 inhibits *Atg16l1* expression in the hippocampus of rats with SE -> inhibiting autophagosome formation - Intracerebral injection of miR-96 mimetics showed that overexpression of miR-96 and consequent decrease in *Atg16l1* attenuated SE-induced brain injury > MiR-96 has a protective role during the development of SE (by inhibiting *atg16l1* among others) and thus may be a novel potential therapy in SE (Gan et al., 2017)
Relevant gene interactions	(1): *WIPI2*: "Intellectual developmental disorder with short stature and variable skeletal anomalies" (OMIM #618453): seizures in some case reports (Maroofian et al., 2021) (2): *ACSF3*: "Combined malonic and methylmalonic aciduria" (OMIM #614265) > seizures in affected patients (Sloan et al., 2011) (3): *CLTC*: "Intellectual developmental disorder 56" (OMIM # 617854) cases with epilepsies and seizures (Thomas et al., 2022)

Table 7.: Gene information about *ATG16L1*.

Gene	***COL6A3* (Collagen Type VI Alpha 3 Chain)**
LOEUF score	0.56
Z-score	-0.61
OMIM listed (1) and phenotypes (2)	(1): yes (2): - "Bethlem myopathy 1" (OMIM #158810) - "Dystonia 27" (OMIM #616411) - "Ullrich congenital muscular dystrophy 1" (OMIM #254090)
Expression	- mainly expressed extracellularly, but only scarcely in the brain
Protein function	- encodes one of the three alpha chains of type VI collagen (in most connective tissues) - an important component for the organization of the ECM (Pan et al., 2013)
Associated studies	- experimentally induced SE of rats > ECM signalling pathway was upregulated compared to sham rats; *Col6a3* was increased expressed in hippocampal tissues > ECM signalling remained significantly activated for at least ten days after the onset of SE > COL6A3 appears to play a role in the ECM pathway in neuronal damage in the hippocampus during epileptogenesis (Han et al., 2019)
Relevant gene interactions	/

Table 8.: Gene information about *COL6A3*.

Gene	***DGKD* (Diacylglycerol-Kinase-Delta)**
LOEUF score	0.37
Z-score	2.96
OMIM listed (1) and phenotypes (2)	(1): yes (2): /
Expression	- in the brain
Protein function	- encodes a cytoplasmic enzyme > production of two lipids acting as messengers in signaling cascades (Sakane et al., 2002; Sato et al., 2013) > important component of cellular signal transduction (Sakane et al., 2002)
Associated studies	- Expression studies in mice and Drosophila > *dgkd* is involved in embryonic development of the midbrain and forebrain > important role in neuronal development and brain pathophysiology (Leach et al., 2007)
Relevant gene interactions	*EPT1:* "Spastic paraplegia 81" (OMIM#618768) > seizures in some patients, facial dysmorphia, delayed motor development, and language impairment (Ahmed et al., 2017; Horibata et al., 2018)

Table 9.: Gene information about *DGKD*.

Gene	***KCNJ13* (PotassiumyInwardly Rectifying Channel Subfamily J Member 13)**
LOEUF score	0.86
Z-score	1.93
OMIM listed (1) and phenotypes (2)	(1): yes (2): - "Snowflake vitreoretinal degeneration" (OMIM #193230) - "Leber congenital amaurosis 16" (OMIM #614186)
Expression	- neuronally expressed at low levels
Protein function	- encodes an inward potassium channel (Kir 7.1) through which potassium ions enter a cell (Sergouniotis et al., 2011)
Associated studies	(1): - Study GASH/Sal (genetic audiogenic seizure hamster from Salamanca) animals (Díaz-Rodríguez et al., 2020; Muñoz et al., 2017) = experimental model of reflex epilepsy of audiogenic origin (resulting from an autosomal recessive disorder) - GASH/Sal - animals exhibit tonic-clonic seizures after acoustic stimulation (short latency period) (Muñoz et al., 2017) - acoustic stimulation > Overexpression of *Kcnj13* in the transcriptome of GASH/Sal animals compared with the respective control group > overexpression of this gene after an epileptic seizure as a compensatory mechanism for Kir channel deregulation associated with epileptogenicity (Díaz-Rodríguez et al., 2020) (2): - *KCNJ13* is overregulated in chronic epilepsy (Winden et al., 2011)
Relevant gene interactions	/

Table 10.: Gene information about *KCNJ13*.

3.5.3 Candidate genes of the locus on chromosome 18 – Region 3

In the final candidate region, two additional genes were detected as possible candidate genes. Those are *BCL2* and *MC4R*.

Gene	***BCL2* (BCL2 -apoptosis regulator)**
LOEUF score	0.71
Z-score	1.4
OMIM listed (1) and phenotypes (2)	(1): yes (2): /
Expression	- expressed in the brain, but predominantly in the thyroid gland or female reproductive organs
Protein function	- inhibits apoptotic death of certain cells, such as lymphocytes (Schimmer et al., 2003)
Associated studies	(1): - apoptosis study of protein levels in neocortex samples from patients with TLE of BCL-2 family (Henshall et al., 2000) > Compared with autopsy controls, these patients had significantly higher levels of anti-apoptotic BCL-2 (Engel & Henshall, 2009) (2): - Study of elevated serum BCL-2 levels in children with TLE > higher in patients with TLE than in controls; level also correlated with seizure frequency, disease severity, and disease duration > upregulation of this gene may be an adaptive response to seizure stress that should contribute to limiting neuronal cell death (Kilany et al., 2012)
Relevant gene	*PPP3CA:* involved in synaptic vesicle recycling and associated with Developmental and epileptic

interactions	encephalopathy 91 (OMIM #617711) (Mizuguchi et al., 2018) > hypertelorism, muscular hypotonia, epileptic encephalopathy, delayed psychomotor development, poor or absent speech, etc. (Rydzanicz et al., 2019)

Table 11.: Gene information about *BCL2*.

Gene	***MC4R* (melanocortin-4 receptor)**
LOEUF score	1.60
Z-score	-1.01
OMIM listed (1) and phenotypes (2)	(1): yes (2): "Obesity" (OMIM #155541)
Expression	- majority is expressed in the brain - particularly in the hypothalamus and basal ganglia
Protein function	- encoded protein is mediated by G proteins and interacts with adrenocorticotropic and MSH (Melanocyte-stimulating hormone) hormones (Vaisse et al., 2000) - *MC4R* is involved in regulation of energy balance and body weight; influences energy metabolism and activity of endocrine and autonomic nervous systems (Adamska-Patruno et al., 2019) - *MC4R* is the receptor for melatonin (Paprocka et al., 2016)
Associated studies	(1): - characteristic of patients with epilepsy are increased phase shifts of melatonin release - melatonin can improve sleep disturbances in children with epilepsy (decrease in wakefulness and sleep latency after falling asleep) (Jain et al., 2015; Paprocka et al., 2016) (2): - case study: index patient with a *MC4R* variant > melatonin rhythm was significantly disturbed (low levels at night that increased again during the day) > sleep disturbances at night, several naps during the day, early awakenings - he had therapy refractory focal epilepsy and sleep disturbances also persisted (Ge et al., 2021) - there is evidence of an association of refractory epilepsy and abnormal *MC4R* gene expression (X. Liu et al., 2016) - therapeutic trial with melatonin in animal models: anticonvulsant effect (Banach et al., 2011) > *MC4R* variants could lead to sleep disorders, obesity, and epilepsy - in the index patient, melatonin treatment in addition to anticonvulsant therapy improved sleep quality and epilepsy (earlier falling asleep, later waking, less daytime sleep, absence of epileptic discharges in the EEG) (Ge et al., 2021)
Relevant gene interactions	(1): *GNB5* (2): *GNB2* (3): *GNB1* > all associated with neurological or intellectual developmental delay, leading to seizures and epilepsy in affected individuals (Endo et al., 2020; Fukuda et al., 2020; Poke et al., 2019)

Table 12.: Gene information about *MC4R*.

4 Discussion

4.1 Epilepsy as a genetically heterogeneous disease

To determine the genetic cause(s) of epilepsy in the studied families, exome sequencing (in all families) and linkage analysis followed by recombinant haplotype reconstruction (in family #2) was performed to identify possible causative variants. In family #1, a variant in a candidate gene of interest (*FERMT2*) had been identified via exome sequencing in the index patient prior to this work and additional variants in this gene were detected via GeneMatcher in two additional individuals in family #2 and in one individual in family #3. By Sanger sequencing, the same variant from the first two individuals of family #2 was detected in seven additional persons in this pedigree. Thus, a total of eleven individuals with a *FERMT2* variant could be analysed. Because the variant did not segregate completely, and *FERMT2* could not be proven to be a causal factor for their epilepsy, linkage analysis was performed in family #2, revealing three candidate regions of interest. These were subsequently further narrowed down by haplotyping. Finally, candidate genes within the regions of interest were analysed (twelve in total).

In the present work, no clear or causal monogenic alteration could be identified that could have adequately explained the epilepsy of the affected individuals. Most epilepsies have non-genetic causes (about 75%, see 1.5.1). Among genetic epilepsies (about 20% see 1.5.2), the majority are polygenic (see 1.5.2.3), however. As of January 2023, the OMIM online database lists more than 200 monogenic diseases whose genetic causes have been understood and in which epileptic seizures are part of the clinical picture. Since all these conditions are relatively rare, they account for only about 1-2% of all epilepsy disorders. An autosomal dominant mode of inheritance is most common. (Habermann-Horstmeier, 2008)

In the linkage analysis performed in family #2, no significantly linked loci (LOD score > 3.0) were found (see 3.3). The LOD score for candidate region 1 (chromosome 2) is around 2.7, for candidate region 2 (chromosome 2) about 1.7 and for candidate region 3 on chromosome 18 this score is 2.5. The LOD score in the *FERMT2* region does not exceed 0.5. In the subsequent haplotyping of these regions, there were no common haplotypes present in all affected individuals that were absent in all healthy individuals. Accordingly, no segregation occurred. In the following, the possible genetic explanations are discussed, which in turn are likely, and which could be the most probable cause. *FERMT2* as a possible genetic cause is discussed in detail later in the thesis (see 4.2.4). In summary, the *FERMT2* analyses suggest

that, first, the causal variants are not located in this gene. And secondly, that based on the linkage results, in three candidate regions (loci on chromosomes 2 and 18) a more detailed investigation should be carried out.

4.2 Possible genetic causes for epilepsy in the families

The most likely explanation for a role of genetics in epilepsy are variants in multiple genes which act together with other environmental factors. This means a multifactorial inheritance mechanism is behind it. This scenario cannot be subject to this work because it would require a large number of samples due to the small effects to be expected in each individual case. Two other scenarios are being further discussed here. The first one are genes in the candidate regions provided by linkage analysis in the extended family from Portugal (family #2). The second one is about *FERMT2* as the prime candidate gene in families #1 and #3.

4.2.1 Candidate genes in the locus of region 1 (chromosome 2)

Linkage analysis identified three possible loci with LOD scores of 2.7 (region 1), 1.7 (region 2), and 2.5 (region 3). Chromosome 2 contains two regions. This first region ranges from rs68012394 to GSA-rs1835328 (twelve markers), includes 61 genes in total (27 of these are listed in OMIM and eight in Morbid), encompassing 2,5 Mb (hg19 chromosome 2.: 112,424,375-115,298,930), and contains six potential candidate genes (see 3.5.1): *ACTR3*, *CKAP2L*, *DPP10*, *IL1A*, *IL1B*, *IL1RN*.

ACTR3 has a very low LOEUF score of 0.12, indicating that this gene has a low tolerance to inactivation. The Z-score is also elevated at 3.61, indicating an increased likelihood of intolerance to missense variants. One study showed that *ACTR3*, a cytoskeleton-associated protein is involved in epithelial junction remodeling during an epileptic seizure. (Pires et al., 2021) This suggests that it has a role in the reorganization of cortical microarchitecture and has implications for remodeling of neuronal circuits or brain vasculature. (Pires et al., 2021) The low LOEUF score, the role of the hippocampus in the pathogenesis of epilepsy in general, and the neuronal synaptic dysfunction in the hippocampus caused by variants in *ACTR3* suggest that this gene may be associated with epilepsy in those individuals carrying the same haplotype in this region (see Figure 3.12). Nevertheless, to date, there are no association studies proving a significant link between a disease and this gene.

CKAP2L has a higher LOEUF of 0.63 and thus this gene is much more tolerant to loss-of-function mutations. Variants in this gene in in vitro and in vivo knockdown resulted in reduced brain size and impaired neuronal function in mouse models. (Hussain et al., 2014; Takahito Yumoto et al., 2013) Although this gene has previously been described in the context with epilepsy (autosomal recessive Filippi syndrome), it is not certain to play a role in the family #2 studied. Besides seizures, phenotypic features such as microcephaly, facial dysmorphia or growth retardation dominate in Filippi syndromes. (Hussain et al., 2014) These have not been described in the patient population at the time of clinical examination and history taking. In addition, *CKAP2L* is significantly more tolerant to missense variants (Z-score = -0.09) and is expressed at very low levels in the brain. However, it cannot be excluded as it could be an autosomal dominantly inherited allelic disease.

Of particular interest is *DPP10* as a possible candidate gene. With a LOEUF of 0.15, it exhibits a low tolerance to loss-of-function variants. On the one hand, it is involved in the assembly of the potassium channel KCND2 with its coded protein (Bezerra et al., 2015) and ion channels have been described in many cases as an important cause in the pathogenesis of epilepsy (see 1.3 and 1.5.2.1). Variants in this gene can lead to changes in potassium current and thus reduce the membrane potential, resulting in increased cell excitability (Bezerra et al., 2015) and this in turn contributes to the development of epileptiform activity. Second, *DPP10* interacts directly or indirectly via other genes with genes related to epilepsies or seizures. In individuals from family #2 who have epilepsy and carry the same haplotype in this region, *DPP10* may be a significant contributor among all possible candidate genes for epileptogenesis due to its low LOEUF score, its expression, and its numerous gene interactions.

The *IL1A*, *IL1B*, and *IL1RN* genes have higher LOEUF scores, decreasing the likelihood that variants in these genes are disease-causing. In addition, these variants have previously been described only in the context of epilepsy after childhood TBI. In the present patient population, there is no evidence of a traumatic genesis in the medical history. Therefore, these genes do not seem to play an important role in the patients of the present work and will not be discussed further.

4.2.2 Candidate genes in the locus of region 2 (chromosome 2)

The same approach as in region 1 was taken for region 2 to find further genes that might be associated with epilepsy in affected individuals. The region detected by linkage analysis could

be narrowed down by haplotyping to the range rs28688824 to rs6716187 with 65 markers. This region includes 103 genes (55 of these are listed in OMIM and 14 in Morbid) and measuring about 15 Mb (hg19 chromosome 2.: 232,779,515-239,503,407). Four candidate genes are located in this region: *ATG16L1*, *COL6A3*, *DGKD* and *KCNJ13*. More detailed information on corresponding studies and further gene information are listed under 3.5.2.

The first gene that needs to be discussed is *ATG16L1*. It encodes a protein that is involved in the process of autophagy. It is suppressed by miR-96, thereby reducing brain damage following an epileptic seizure. (Gan et al., 2017) *ATG16L1* has no high tolerance to loss-of-function mutations (LOEUF = 0.32). In summary, *ATG16L1* would be a good candidate gene based on its function, expression, and low LOEUF score. A counter argument is that diseases that are caused by variation in interacting genes of *ATG16L1* do not fit the phenotypic spectrum of the studied individuals in this work.

The second gene in region 2 is *COL6A3*, which appears to have a protective effect on the hippocampus during epileptic seizures via the ECM pathway. (Han et al., 2019) This gene appears to be relatively tolerant to missense and loss-of-function variants (LOEUF = 0.56, Z-score = -0.61). It interacts with genes that are associated with diseases where affected individuals may experience some seizures, but these are not the main symptom. (Birman et al., 2012) In summary, *COL6A3* is rather an unlikely candidate gene.

DGKD is a possible candidate gene, first, because of its low LOEUF score (0.37), its expression in the brain, its role as a component of cellular signal transduction (Howles et al., 2019) and the studies performed on this gene (see 3.4.3). This gene could also contribute in a small degree to a polygenic or complex inheritance mechanism of the affected individuals from family #2.

The last possible candidate gene in region 2 is *KCNJ13*. It encodes an inward potassium channel through which potassium ions can enter a cell. (Sergouniotis et al., 2011) *KCNJ13* has been shown in experimental studies to be overregulated in epileptic seizures and appears to be a compensatory mechanism for Kir channel deregulation associated with epileptogenicity. (Leach et al., 2007) However, this gene is relatively tolerant to missense and loss-of-function mutations (LOEUF = 0.86 and Z-score = 1.93), and it is neuronally expressed at low levels. Interaction with genes already associated with seizures or epilepsy has not yet been discovered. Therefore, *KCNJ13* is probably not a suitable candidate gene for family #2.

Based on the segregation of haplotypes (see 3.4.3) and some interesting candidate genes, region 2 seems to be the most attractive as a possible genetic cause for epilepsy in family #2.

4.2.3 Candidate genes in the locus of region 3 (chromosome 18)

The list of genes within the conceivable disease region for chromosome 18 were mapped by genotype analysis based on SNP genotyping array and linkage analysis data followed by recombinant haplotype analysis to a region of about 10 Mb (see Figure 3.9 and Figure 3.14), limited by the SNP-markers rs11874251 and rs7229020 (45 markers, hg19 chromosome 18: 57,124,276-63,275,241). None of the 50 genes listed in Table 22 (24 are listed in OMIM and eight in Morbid) is a conspicuous candidate based on gene function, yet two genes were detected that could contribute at least one risk factor for the development of the epilepsies. These genes have already been described under 3.5.3 in detail but will now be analyzed again regarding their significance as a possible cause or risk factor of the epilepsies of the family members from the Portuguese family.

The integral protein of the outer mitochondrial membrane encoded by *BCL2* inhibits apoptotic death of certain cells, such as lymphocytes. (Schimmer et al., 2003) *BCL2* has been shown in studies to be upregulated after the onset of an epileptic seizure and may contribute to limiting neuronal cell death through its anti-apoptotic partial function. (Engel & Henshall, 2009; Kilany et al., 2012) With a LOEUF of 0.71 and a Z-score of 1.4, this gene is most likely tolerant to loss-of-function and missense variants. It interacts with a gene associated with "Developmental and epileptic encephalopathy 91" (OMIM #617711). (Mizuguchi et al., 2018) The phenotype of the Portuguese family does not correspond to the described symptoms.

Another possible candidate gene is *MC4R*. The protein encoded by MC4R is a receptor for melatonin and melatonin treatment has been shown to improve sleep and sleep quality in studies. (Ge et al., 2021; Jain et al., 2015; Paprocka et al., 2016) Although *MC4R* variants have no clear pathogenic effect, this gene may also be involved in epilepsies because of its expression in the brain and some interactions with epilepsy-associated genes. Many affected individuals from family #2 have the same haplotype at this position on chromosome 18 and thus *MC4R* might play a role.

4.2.4 Is *FERMT2* responsible for epilepsy in the families studied?

Since variants were identified in the *FERMT2* gene in eleven individuals (one via exome sequencing in family #1: c.392A>G, p.(Asn131Ser), nine via exome and Sanger sequencing in family #2: c.1538C>T p.(Thr513Met) and one via GeneMatcher in family #3: c.1976G>A, p.(Arg659His), the haplotype on chromosome 14 of the individuals from family #2 was also

reconstructed in this region including the 2 Mb area surrounding the gene: rs8020209 to rs7151053 (chromosome 14.: 52,073,301-54,172,855). Because the *FERMT2* variant cannot be proven causal but could be co-segregating with a causal variant in a neighboring gene, the environment of the *FERMT2* gene was analyzed in more detail. There are 25 protein coding genes in this region (see Table 23), among which 13 are listed in OMIM or Morbid. Some of these genes, including *PSMC6* and *PTGDR*, have already been analyzed in more detail in some case reports or studies and have been associated with epilepsies and seizures. (Dimassi et al., 2014) For example, a variant in *PSMC6* was discovered in a patient with Rolandic epilepsy. (Dimassi et al., 2014) Nevertheless, when alterations in these genes occurred, other genetic syndromes and symptoms such as microcephaly or mental retardation were prominent or had a high LOEUF, so none of these genes, except *FERMT2*, will be discussed in more detail here.

In a rat model, as mentioned above, it was shown that the orthologous gene appears to play an important role in the regression of ECM structure and cell adhesion after an epileptic seizure. (Keck et al., 2018) It is then upregulated and thus increasingly expressed. It enhances integrin activation and consequently cell signal transduction. Disruption in this pathomechanism could therefore make a significant contribution to the development of epilepsy. (Keck et al., 2018) With its LOEUF score of 0.08 it is likely intolerant to loss-of-function. Figure 4.1 shows the three variants of family #1 ((c.392A>G, p.(Asn131Ser)), family #2 (c.1538C>T p.(Thr513Met)), and family #3 (c.1976G>A, p.(Arg659His)). The three respective protein residues have chemical bonds to different alpha chains and ß-sheets. These interactions or the individual localizations of the variants in the protein could lead to alterations in the protein function or its stability for example. This could result in decreased regression of the ECM and cell adhesion after SE. However, there are no more experimental studies on these families or on other affected individuals worldwide with these or other variants so far, so this is only a hypothesis (06/2022). None of these variants have been described in PubMed (https://www.ncbi.nlm.nih.gov/pmc/) or in ClinVar in the context of epilepsy yet (06/2022). The variant of family #2 (c.1538C>T p.(Thr513Met)) has a relatively high allele frequency of about 0.24% and this rather speaks against the variant being pathogenic in monogenic forms of epilepsy. Unfortunately, based on the limited data and experiments available to date on variants in *FERMT2* and the gene in general, no definitive conclusions can be drawn.

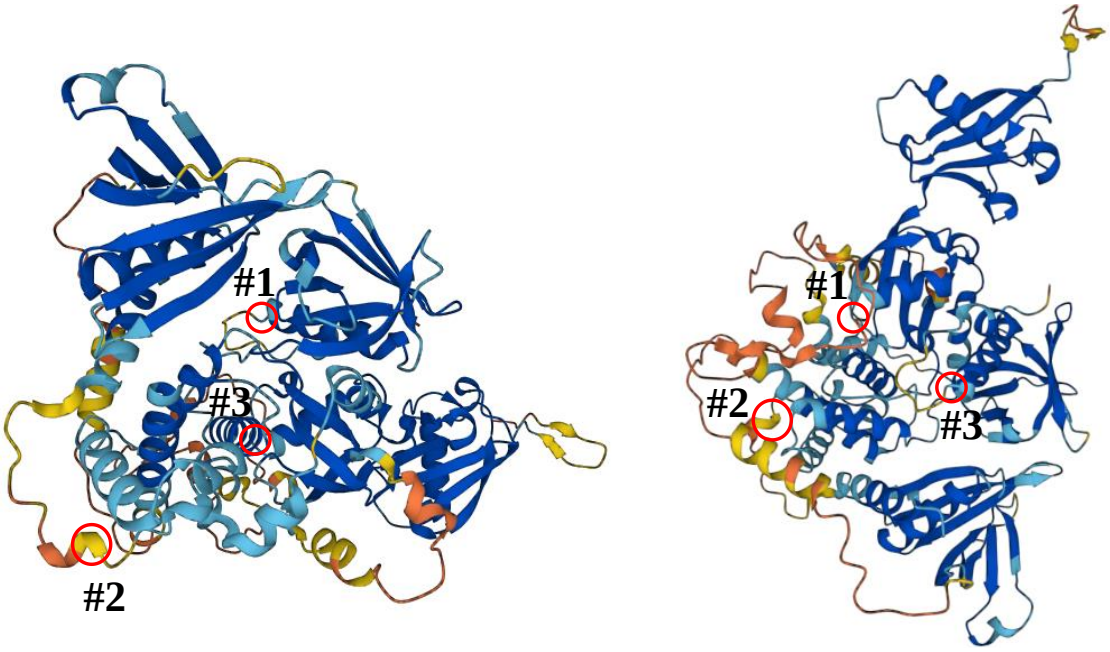

Figure 4.1.: Overview of the detected variants in the *FERMT2* protein. View of the tertiary structure in two perspectives. The positions of each variant have been circled in red and the adjacent digit (#1, #2, and #3) indicates the associated family.

Table 1 (see 3.2.2) shows that all patients from family #2 who clearly have epilepsy carry the same *FERMT2* variant. Among those who have seizures or epilepsy-typical EEG abnormalities, three out of five carry the variant. In II.3 or III.11 (not carrying the variant), other variants, structural lesions, or even phenocopies could explain the phenotype. It is not unlikely that phenocopies are present in such a common phenotype such as epilepsy. The phenocopy rate has been estimated to be 5% in studies on febrile convulsions and generalized epilepsy, because the frequency of febrile convulsions in the general population is about 3%-5%. (Baulac et al., 1999) Considering the left part of the pedigree of family #2 with eight affected and genotyped individuals, a rate of 5% would result in 0.4 phenocopies. In another study, the phenocopy rate was estimated to be 3% (Wallace et al., 1998) and that would be 0.24 patients out of eight in this case. In summary, it is not unlikely that there is at least one phenocopy in this family. Reevaluation of the patient history of III.11 and II.3 would have been very helpful, but follow-up attempt had failed.

The unaffected individual III.7, carries the *FERMT2* variant but is unaffected. This could be due to reduced penetrance, meaning that the trait (here epilepsy) is not expressed in every case. Reasons for this may include compensating genes or environmental influences. In summary, *FERMT2* may play a certain role in epileptogenesis in the patients studied in family #1 and #3. In family #2, *FERMT2* may contribute to a multifactorial event. Because of the

high allele frequency of this variant, *FERMT2* is very unlikely to be a monogenic cause of the epilepsies of the individuals in the Portuguese family.
However, more attention should be paid to this gene in the future even in cases of unclear epilepsy. Further exome sequencing of many individuals and controls may provide more information about the importance of this gene in this context in the future. Also, further animal models or other experimental approaches could contribute to further understanding. In retrospect, further unclear cases with epilepsy could also be rechecked for variants in *FERMT2* and revised.

4.2.5 Summary of variant assessment in candidate genes

Following the detailed discussion of the possible candidate genes, a summary of the possible causes in the affected families now follows. Assuming genetic causes, there may be a monogenic or polygenic inheritance mechanism behind it. For a polygenic inheritance would speak that "common is common". Also, in epilepsies many different genes with environmental factors are predominantly causative. Moreover, no haplotype could be identified for any region, which occurs in all affected patients but not in healthy individuals. This indicates that it is rather not a monogenic disease in the family studied. However, it cannot be ruled out, because incomplete penetrance or other genetic variants may play an important role. In addition, there could be a mistake in the linkage analysis because one or two or three individuals have phenocopies. As described in 4.2.4, it is not at all unlikely (in 0.4 out of 8 patients) that at least one phenocopy is present. For a monogenic disease speaks (based on family #2) that both males and females are affected over four generations and that suggests an autosomal dominant inheritance. Another argument for a monogenic inheritance is that in all three families a gene has been identified (*FERMT2*) via exome sequencing or Sanger sequencing that has already been linked to epilepsy in an animal model. In addition, three loci of interest were found via linkage analysis. Therefore, *FERMT2* in all three families or an unidentified variant in a linked region in family #2 could be disease-causing.

Linkage analysis provides a strong tool for monogenic disease gene identification and provides at least preliminary evidence for a candidate gene in three chromosomal regions, two on chromosome 2 and one on chromosome 18. It can be concluded that there are genes that are more likely and those that are less likely to lead to the epilepsy of the affected individuals in the present study. *DPP10* in region 1 on chromosome 2 seems to be the most interesting candidate gene for several reasons. First, it is itself involved in the formation of a potassium

channel and potassium channels play an important role in the pathophysiology of epilepsy. Second, it in turn interacts with other genes that are also involved in the formation of other potassium channels. With a LOEUF of 0.15, *DPP10* has a low tolerance to loss-of-function mutations. Also, of interest is *ACTR3* in region 1, which functions as an important component of a complex essential for cell shape and cell motility and thus is also an important component of electrophysiological cell-to-cell transmission of epileptogenic potentials. Variants in the genes can lead to dysfunction in the remodeling of epithelial junction sites during an epileptic seizure. Last, the low LOEUF score may be another indication that variants in *ACTR3* are associated with epilepsy in family #2. *DGKD* in region 2 would be the third gene in a priority list. It has a role in signal transduction and variants have been associated with epilepsy cases. In a mouse model, *DGKD* has been shown to play a role in forebrain and midbrain development. Moreover, this gene has a low LOEUF score (0.37). Furthermore, in region 2 the haplotypes "fit best" in family #2, meaning that most of those who have epilepsy carry the same haplotype and a large proportion of healthy individuals do not carry it.

FERMT2 on chromosome 14 remains a very interesting candidate gene in all three families. In most individuals, the genotype matches the phenotype. Individual III.7, who does not have epilepsy, may have reduced penetrance, III.14, who is diseased but does not have the *FERMT2* variant, possibly has a phenocopy, and II.3 who also does not have the genotype did not have a definite epilepsy diagnosis. In the rest of family #2, the affected individuals all carry the *FERMT2* variant (c.1538C>T p.(Thr513Met)) and the index patients in family #1 (c.392A>G, p.(Asn131Ser)) and in family #3 (c.1976G>A, p.(Arg659His)) also have a *FERMT2* variant. The important role in regression of the ECM structure after an epileptic seizure also suggests that pathogenic variants could lead to increased seizure frequency. The low tolerance to loss-of-function (LOEUF score = 0.08) also argues for this gene as the cause for the epilepsies in the studied patients. Last, all variants could also lead to instability in the protein and thus to loss-of-function due to their alterations. But, as described above, the high allele frequency of the variant in family #2 argues against a causal link and against the cause of epilepsy in the affected Portuguese family members.

In summary, based on haplotyping, region 2 is the most likely cause of epilepsy in family #2. *FERMT2* remains a very interesting candidate gene in family #1 and #3 but seems very unlikely as a cause for the epilepsy of family #2 because of the high allele frequency. However, a polygenic inheritance mechanism is probably just as likely in all three families. In patients with unclear epilepsy, re-evaluation of exome data from our own databases regarding

variants in these genes and more experimental studies would certainly be very interesting, informative, and necessary.

4.3 How disease-causing variants can be missed by WES

4.3.1 Repeat expansions

Because exome sequencing of all members of family #1 and of two members of family #2 and of the index patient of family #3 did not reveal causative candidate genes or a clear genetic cause, it must be discussed which genetic variants could be causing the epilepsy in the studied families but have been overseen by this technology. One possible cause would be repeat expansions in either coding or noncoding regions. Repeat expansions cannot be detected by the methods used in the present work (Bahlo et al., 2018) (Illumina HiSeq 6000) because the additional bases inserted cause too many mismatches compared to the reference genome. (Wagner, 2018)

One form of epilepsy in which repeat expansions have already been confirmed as a cause is for example the progressive myoclonus epilepsy (EPM1). (Lalioti et al., 1997) This is a rare disorder characterized by the occurrence of moyclonia, seizures, and progressive neurologic deficits. (Zupanc & Legros, 2004) In this epilepsy, dodecamer repeat expansions occur in the cystatin B gene. This gene encodes a cysteine proteinase inhibitor and its protein is thought to play a role in protecting against proteases emerging from lysosomes. (Alakurtti et al., 2005) The dodecamer repeats (12-mers) are located approximately 70 nucleotides upstream from the transcription start site closest to the 5′ end of the *CSTB* gene. (Lalioti et al., 1997) Normal alleles contain two or three copies of this repeat, whereas mutant alleles contain more than 60 such repeats. The altered spacing of promoter elements due to expansion of the dodecamer repeats contributes to the decreased expression of the cystatin B gene in EPM1. The levels of messenger RNA (mRNA) encoded by this gene were found to be decreased in cells from affected individuals. The results indicate that seizure activity leads to a rapid and widespread increase in *CSTB* synthesis in forebrain neurons at normal allele number. Accordingly, upregulation of *CSTB* after seizures may counteract apoptosis by binding cysteine proteases. With increased allele number, there is no upregulation of the gene and thus no decreased apoptosis. (Lalioti et al., 1997)

4.3.2 Non-coding variants

Another conceivable mechanism in which genetic defects can cause epilepsy are variants that are non-coding. Noncoding DNA sequences do not code for protein sequences via transcription into mRNA. They are difficult to identify because they are not enriched in WES. Part of this DNA is transcribed into functional non-coding RNA (e.g. ribisomal/ regulatory or transfer RNA). (Pagni et al., 2022) The resulting RNAs are not used for translation. In addition, the 5'- and 3'-untranslated regions (UTRs =margin region of the mRNA that does not code for the actual protein) located at both ends of the mRNA also belong to the non-coding DNAs. Non-coding variants are often found in the region of the promoter, in a splice site, in the RNA polymerase binding site, or in regulatory elements. (Li et al., 2016; Ling et al., 2015) In one study, noncoding promoter and intronic regions were sequenced in some patients who did not have loss-of-function variants in *SLC2A1,* which encodes glucose transporter-1 (GLUT-1). (Y.-C. Liu et al., 2016) Mutations in *SLC2A1* lead to glucose transport disorders across the blood-brain barrier, resulting in various forms of epilepsy. (Y.-C. Liu et al., 2016) One proband had a de novo splice site mutation five bp away from the intron-exon boundary. Three patients out of 55 studied had deep intronic *SLC2A1* variants. Among them, two patients had a recurrent variant that produced less *SLC2A1* mRNA transcript. (Y.-C. Liu et al., 2016) The results suggest that low cerebrospinal fluid glucose levels may be associated with pathogenic *SLC2A1* mutations, including deep intronic *SLC2A1* variants. An outlook of the study is that by extending genetic screening to noncoding regions, it will be possible to diagnose more patients with GLUT1 deficiency and improve epilepsy by introducing the ketogenic diet if necessary. (Y.-C. Liu et al., 2016)

Long noncoding RNAs have also been shown to have regulatory functions involved in a variety of biological processes. Some studies report that these long noncoding RNAs are involved in the regulation of pathological processes of epilepsy and are dysregulated during epileptogenesis. (Villa et al., 2019)

In the *SCN1A* gene (see 1.5.2.1 and 1.5.2.3), a microdeletion in the 5'-promoter region was found in patients with Dravet syndrome (Nakayama et al., 2010), and another heterozygous variant in the promoter region was identified in a patient with febrile seizures and focal epilepsy (Gao et al., 2017). The heterozygous variant in the promoter region resulted in reduced *SCN1A* promotor activity by approximately 40% compared with the wild-type variant. (Gao et al., 2017) This may explain the relatively mild phenotypic impairment caused

by this noncoding variant compared with the phenotype of *SCN1A*-coding variants, which can lead to non-expression. (van Loo & Becker, 2019)

There are also mechanisms that lead to altered transcriptional regulation in non-coding sequences. This is the case when a genetic variant is located within the binding site of an activating or inhibiting transcription factor. This mechanism has been detected, for example, in a group of patients suffering from childhood absence epilepsy. (Urak et al., 2006) These have a variant in the *GABRB3* promoter. The *GABRB3* gene encodes the β3 subunit of the GABA-A receptor, which mediates phasic (synaptic) and tonic (perisynaptic) inhibition. (Farrant & Nusser, 2005; Hirose, 2014) Functional analysis of this variant revealed reduced transcriptional activity of the *GABRB3* promoter in childhood absence epilepsy patients compared to the *GABRB3* promoter in controls. (Urak et al., 2006) This resulted in reduced expression of the *GABRB3* gene (Urak et al., 2006) and consequently reduced ß3 levels. This may lead to a loss of the inhibitory properties of the receptor, which finally causes the seizure activity. (van Loo & Becker, 2019)

These genetic mechanisms should also be considered as possible causes of epilepsy.

4.3.3 Locus heterogeneity

Locus heterogeneity occurs when causal variants in two or more genes underlie the etiology of a Mendelian trait within a pedigree. (Atteeq U Rehman et al., 2015) In one study ten consanguineous Pakistani families with autosomal recessive hearing loss were presented with mutations in two or more genes. (Atteeq U. Rehman et al., 2015; Wagner, 2018) The results showed that in the cohort studied, familial locus heterogeneity occurs in 15% of the families. In these cases, linkage analysis would result in lower LOD scores similarly to the presence of phenocopies. (Atteeq U. Rehman et al., 2015)

In summary, detection of causal genes for diseases affected by locus heterogeneity is difficult using genetic analysis methods such as linkage analysis and genome sequencing. These methods rely on comparing affected family members. However, if different family members have different disease-causing genes, such genes may not be accurately identified. This could consequently affect linkage and segregation analysis.

4.3.4 Technical issues

Exome sequencing has not detected any convincing candidate variants other than in *FERMT2* in the present study. Therefore, the question arises how disease-causing variants can be overlooked with this technology. On average, modern kits currently enrich 98 % of exomes (Cronn et al., 2012), as it remains difficult to develop hybridization probes for specific coding regions, especially GC-rich sequences, and repetitive elements. Additionally, coverage of targeted coding exons is not 100% precise, as regions exist that are poorly or not covered at all. (Li et al., 2015; Parla et al., 2011; Rehm et al., 2013) The main reason for insufficient coverage, as mentioned earlier, is GC-rich regions that pose a difficulty for certain exome enrichment kits. This usually includes the first exons of protein-coding genes. Because PCR is required for these technologies, AT- and GC-rich target sequences may be underrepresented in genome alignments and assemblies. This may result in low coverage. (Dohm et al., 2008; Harismendy et al., 2009; Metzker, 2010; Wagner, 2018)
In summary, these reasons may explain why no sufficiently explanatory genetic cause for epilepsy was found in this work. Linkage analysis and haplotyping were able to compensate to some extent for these limitations in Portuguese family #2.

4.3.5 Bioinformatic difficulties

A further reason why WES methods fail to identify the causative genetic variants or muatations may be that the inherited disease is not detected due to difficulties in bioinformatic analysis of exome sequencing data. The choice of variant caller and annotation tools may lead to various discrepancies. (McLaren et al., 2016; Pabinger et al., 2014) For example, one study has shown that the combination of different software tools with read aligners for variant detection has variable power. (Hwang et al., 2015) In another study, the agreement for annotation of suspected loss-of-function variants was only 44% when RefSeq transcript sets were used as the basis for annotation compared to Ensembl transcript sets. (McCarthy et al., 2014) This study also showed that 87% of all exonic variants and only 65% of loss-of-function variants were found to match when the results of the annotation software tools annovar and Variant Effect Predictor (McLaren et al., 2010) were compared. Consequently, this means that there are a lot of false negative variants.

In summary, it is still not possible to perform an optimal bioinformatic analysis. The combination of software tools always has strengths and weaknesses and is not always able to detect all existing variants.

4.4 Possibilities and limitations of linkage analysis

4.4.1 Comparison of parametric and non-parametric linkage analysis

Linkage analysis has been an established method for many years to identify disease-causing genes when the mode of inheritance is monogenic. The mode of inheritance of the forms of epilepsy may be monogenic to a small extent, i.e. that a single gene mutation leads to the disease. More often, a more complex di- or polygenic pattern of inheritance is suspected, in which several genes together with environmental factors are causative.
Parametric linkage analysis is well suited for the analysis of a genome if the genetic model is known. It must be based on a monogenic dominant or monogenic recessive inheritance. Parametric linkage analysis is opposed to non-parametric linkage analysis, which is a method that works model-free, so that no specific mode of inheritance is assumed for the disease under investigation. (Kruglyak et al., 1996) The disadvantage of non-parametric analysis is that only genetic information of affected individuals is processed. The parametric methods therefore have the greater test strength compared to the non-parametric methods - provided that the correct genetic model is used. (Abreu et al., 1999) The advantage of the fundamentally greater statistical power of parametric linkage analyses is offset by the disadvantage of having to create a speculative genetic model. However, to keep the alpha error as low as possible, the combination of both methods is recommended. (Abreu et al., 1999) In the context of the present work, parametric methods were predominantly applied.

4.4.2 Possibilities

Linkage analyses have the advantage over association studies that significant results can be obtained even with a relatively small collective if the families have a suitable structure. Linkage analyses have been successfully used to identify genetic variants that contribute to rare diseases such as Huntington disease (Myers et al., 1989) and to monogenic forms a common disorders. (Erdmann et al., 2013) In summary, linkage analyses of single appropriate families can further help to determine disease-defining loci.

4.4.3 Limitations

The parametric linkage analysis requires the assumption of precise genetic models, including penetrance, disease gene frequency, and the clear classification of individuals as affected or unaffected. Therefore, LOD scores change significantly when penetrance is reduced. Similarly, misdiagnosis can greatly affect the LOD score at the calculated site of the disease gene. The limited or partly inaccurate clinical information in family #2 of the individuals (when exactly the patients were examined or was there a re-evaluation of the clinic and were there subsequent changes of the affection status from unaffected to affected for example) are limitations.

Therefore, linkage analyses rely on pedigrees of entire families that are as complete as possible.

Linkage analyses benefit from highly polymorphic genetic markers, as this decreases the likelihood that both parents carry the same allele. (Kruglyak et al., 1996; Read & Strachan, 1996)

The information content of segregation analysis increases with the number of markers used. Basically, a disadvantage of linkage analysis is that one can only indirectly reach conclusions about the disease-causing genes via the genetic markers used. Sequencing of candidate genes obtained by linkage analysis is subsequently necessary to detect disease-causing variants. As mentioned earlier, phenocopies such as in a phenotype as common as epilepsy are not unlikely and, consequently, are also a limiting factor for linkage analyses.

Association studies provide the advantage, in contrast to linkage analyses, of not having to rely on pedigrees of entire families that are as complete as possible. Association studies also do not require family studies, but only subjects with a well-defined phenotype. (Cordell & Clayton, 2005; Cui et al., 2010; Manolio, 2010)

Most genetic diseases are based on complex inheritance mechanisms, which are therefore difficult or impossible to detect by linkage analysis.

4.4.4 Development of methods for the identification of genetic variants

Nowadays, and also in the future, alternative methods are increasingly simplifying the identification of the causative genetic factors of complex hereditary diseases. Current high-throughput sequencing techniques, in contrast to conventional Sanger sequencing, allow parallel processing of up to millions of DNA sequences. The cost of sequencing a person's

entire genome has dropped strongly in recent years, and whole genome sequencing is already commercially available. (Goodwin et al., 2016) The ability to assess non-coding sections of the genome outside of genes and their regulatory sequences is considered to be an advantage. (Goodwin et al., 2016) A disadvantage is the relatively high price of producing, processing and storing whole genome sequences. More affordable is whole exome sequencing (WES). In this case, only the protein-coding regions of genes, i.e., exons, are sequenced. The resulting high data volumes of high-throughput sequencing place correspondingly high demands on data processing and statistical analysis, but also create a better and cause-related overview of the sequence changes in the genes of affected individuals. (Goodwin et al., 2016) It is possible to examine the entire genome of single persons or entire families for detection of disease-causing variants. One approach would be WES of these people within a collective of families that are homogeneous with respect to the disease to identify epilepsy causing variants and then test for concordance and segregation of these variants within families. Other approaches include regular reanalyses of exomes at intervals of about two years and further searches via GeneMatcher. Otherwise, there are options such as long-read genome sequencing at Oxford Nanopore Technologies (ONT), a third-generation sequencer that can sequence DNA at remote locations and generate ultra-long reads or PacBio (Pacific Biosciences of California, https://www.pacb.com/), which develops and creates systems for gene sequencing and some novel real-time biological observations as single-molecule real-time sequencing (SMRT). (Rhoads & Au, 2015)

Another approach to identifying further causes is to study the total amount of proteins (=proteomics) produced in an organism (=proteome). (Aslam et al., 2017) The proteome changes throughout life and varies from cell to cell. Proteomics can be used, for example, to study when and where proteins are expressed, how they are modified by posttranslational modifications, or which proteins are involved in which metabolic pathways. The interactions of the individual proteins with each other can also be investigated. (Cho, 2007; Hanash, 2003)

RNA sequencing can also continue to help uncover unsolved cases. This is a sequencing technique that uses NGS to detect the presence of RNA and the total amount of RNA in a biological sample on the one hand, and to analyse the constantly changing cellular transcriptome on the other. (Chu & Corey, 2012; Wang et al., 2009)

A last-mentioned approach at this point to reveal more cases is "multi-OMICs". An "OMIC" is a meta-genome and/or a meta-transcriptome, depending on how it is sequenced. (Bock et al., 2016) In this method, datasets from multiple "OMICs" (e.g.: transcriptome, genome, proteome, or microbiome, …) are combined and analysed to find coherent matching geno-

pheno-envirotype relationships or associations between biological entities, among others. (Bersanelli et al., 2016; Bock et al., 2016; Tarazona et al., 2018)

5 Summary

Epilepsy is a predominantly non-hereditary neurological disorder, which nevertheless has genetic - especially polygenic - causes in many cases. There are also many cases in which the epilepsy is monogenic. The latter was originally suspected in this work in the family members studied and affected. The phenotypic spectrum ranges from focal seizures to multiple daily generalized seizures accompanied by myoclonias, cognitive deficits, and neurodevelopmental disorders. This clinical picture has numerous causes - including genetic ones. Despite intensive efforts, many genetic causes of epilepsy have not yet been identified.

The aim of this work was to identify variants that cause epilepsy in the affected patients in the three families. Exome sequencing was performed at the Institute of Human Genetics, TUM, in the patient of family #1, which identified a variant in the *FERMT2* gene on chromosome 14 as a possible cause for the epilepsy. In a rat model, this gene showed to play an important role in the regression of ECM structure and cell adhesion after an epileptic seizure. Subsequently, the in-house database of 25 000 patients was searched for further matches, however unsuccessfully. Then, two different *FERMT2* variants were found via GeneMatcher in a total of three additional patients with epilepsy in two other families from Portugal and Iceland (family #2 and #3) and Sanger sequencing identified seven additional family members of the Portuguese family with the *FERMT2* variant. Thus, a total of eleven individuals in the three families had a variant in this gene.

Although *FERMT2* variants could not be excluded to be disease causing, there is nevertheless rather limited evidence for them to be pathogenic. In family #1 the most comprehensive pathogenicity program predicts a benign effect for the variant, in family #2 the variant has a relatively high allele frequency in the general population. The linkage analysis in family #2, although inconclusive, suggests alternative candidate regions. In family #3 the variant has occurred *de novo*, but mutation prediction tools predict predominantly an uncertain effect.

Using SNP-array technology, linkage analysis followed by haplotyping was performed in family #2, which revealed high LOD scores for markers on chromosome 2 and chromosome 18 for dominant inheritance. This method was performed because the *FERMT2* variant did not segregate completely in family #2 and provided a means to discover additional possible regions missed by exome sequencing. The candidate gene search of the three regions of the linkage analysis revealed several genes that could play a role in the pathogenesis of epilepsy. *DPP10* and *ACTR3* in region 1 on chromosome 2, *ATG16L1* and *DGKD* in region 2 on chromosome 2 and finally *MC4R* in region 3 on chromosome 18 stood out as particular

candidate genes. This is because they have already been described in clinical studies and case studies in connection with epileptic seizures and/or epilepsy and appear to play a role in epileptogenesis in experimental studies, among other things. Region 2 is a 15 Mb region and remains of particular interest due to the distribution of haplotypes and because the result of the linkage analysis suggests a candidate locus there.

Finally, the possible explanations why the underlying genetic defect of epilepsy could not be identified were reviewed, despite linkage analysis was performed in one and exome sequencing in three families. These include locus heterogeneity, repeat expansions and non-coding variants. The limitations of linkage analysis and exome-sequencing must also be considered. All these points could be the reason why the genetic cause of epilepsy in these family members has not been identified with certainty, and follow-up studies using more advanced methods such as transcriptome and genome sequencing are needed to understand the complex genetic mechanisms responsible.

6 Zusammenfassung

Epilepsie ist eine überwiegend nicht erblich bedingte neurologische Erkrankung, die jedoch in zahlreichen Fällen auch genetische - insbesondere polygene - Ursachen haben kann. Dennoch gibt es auch viele Fälle, in denen die Epilepsie monogen bedingt ist. Dies wurde in dieser Arbeit ursprünglich bei den untersuchten und betroffenen Familienmitgliedern vermutet. Das phänotypische Spektrum reicht von fokalen Anfällen über mehrmals täglich generalisierte Anfälle begleitet von Myoklonien, kognitiven Defiziten und neurologischen Entwicklungsstörungen. Dieses Krankheitsbild hat zahlreiche Ursachen - unter anderem genetische. Trotz intensiven Bemühungen konnten viele genetische Ursachen der Epilepsie noch nicht identifiziert werden.

Ziel dieser Arbeit war es, Varianten zu identifizieren, die Epilepsie bei den betroffenen Patienten in den drei Familien verursachen. Am Institut für Humangenetik der TUM wurde beim Index der Familie #1 eine Exom-Sequenzierung durchgeführt, die eine Variante im *FERMT2*-Gen auf Chromosom 14 als mögliche Ursache für die Epilepsie identifizierte. In einem Rattenmodell hatte sich gezeigt, dass dieses Gen eine wichtige Rolle bei der Rückbildung der extrazellulären Matrixstruktur und der Zelladhäsion nach einem epileptischen Anfall spielt. Anschließend wurde die hauseigene Datenbank mit 25 000 Patienten nach weiteren Übereinstimmungen ohne Erfolg durchsucht. Dann wurden über GeneMatcher zwei verschiedene *FERMT2*-Varianten bei insgesamt drei weiteren Patienten mit Epilepsie in zwei anderen Familien aus Portugal und Island (Familie #2 und #3) gefunden. Die anschließende Sanger-Sequenzierung identifizierte in sieben weiteren Familienmitgliedern der portugiesischen Familie die *FERMT2*-Variante. Somit hatten insgesamt elf Personen in den drei Familien eine Variante in diesem Gen.

Obwohl nicht ausgeschlossen werden konnte, dass *FERMT2*-Varianten krankheitsverursachend sind, gibt es dennoch nur begrenzte Hinweise darauf, dass sie pathogen sind. In Familie #1 sagen Pathogenitätsprogramme für die Variante einen benignen Effekt voraus und in Familie #2 hat die Variante eine relativ hohe Allelfrequenz in der allgemeinen Bevölkerung. Die Kopplungsanalyse in Familie #2 ist zwar nicht schlüssig, legt aber alternative Kandidatenregionen nahe. In Familie #3 ist die Variante *de novo* aufgetreten, aber Pathogenitätsprogramme sagen überwiegend einen unsicheren Effekt dieser Variante voraus.

Mittels SNP-Array wurde in Familie #2 eine Kopplungsanalyse mit anschließender Haplotypisierung durchgeführt, die erhöhte LOD scores für Marker auf Chromosom 2 und 18

für den dominanten Erbgang ergab. Diese Methoden wurden durchgeführt, da die *FERMT2*-Variante in Familie #2 nicht vollständig segregierte und eine Möglichkeit bot, zusätzliche mögliche Regionen zu entdecken, die durch die Exomsequenzierung übersehen wurden. Die Kandidatengensuche der drei Regionen der Kopplungsanalyse ergab einige Gene, die eine Rolle bei der Pathogenese der Epilepsie spielen könnten. *DPP10* umd *ACTR3* in Region 1 auf Chromosom 2, *ATG16L1* und *DGKD* in Region 2 auf Chromosom 2 und zuletzt *MC4R* in Region 3 auf Chromosom 18 stachen dabei als besondere Kandidatengene hervor. Denn sie wurden unter anderem bereits in klinischen Studien und Fallstudien in Zusammenhang mit epileptischen Anfällen und/oder Epilepsien beschrieben und scheinen in experimentellen Untersuchungen eine Rolle im Rahmen der Epileptogenese zu spielen. Region 2 ist eine 15 Mb große Region und bleibt aufgrund der Verteilung der Haplotypen und weil das Ergebnis der Kopplunsganalyse dort einen Kandidatenlocus vermuten lässt, insbesondere interessant.

Schließlich wurden die möglichen Erklärungen geprüft, warum der der Epilepsie zugrunde liegende Gendefekt nicht identifiziert werden konnte, obwohl in einer Familie eine Kopplungsanalyse und in drei Familien eine Exom-Sequenzierung durchgeführt wurde. Dazu gehören Locus-Heterogenität, Repeat-Expansionen und nichtkodierende Varianten, die die Krankheit verursachen könnten. Auch die Grenzen der Kopplungsanalyse und der Exom-Sequenzierung müssen berücksichtigt werden. Zusammenfassend gibt es einige mögliche Ursachen, weshalb die genetische Ursache der Epilepsie bei diesen Familienmitgliedern nicht mit Sicherheit identifiziert werden konnte. Folgestudien mit weiteren Untersuchungstechniken wie Transkriptom- und Genomsequenzierung werden erforderlich, um die ursächlichen genetischen Veränderungen in den betroffenen Individuen zu identifizieren.

7 Bibliography

Abou-Khalil, B., Auce, P., Avbersek, A., Bahlo, M., Balding, D. J., Bast, T., Baum, L., Becker, A. J., Becker, F., Berghuis, B., Berkovic, S. F., Boysen, K. E., Bradfield, J. P., Brody, L. C., Buono, R. J., Campbell, E., Cascino, G. D., Catarino, C. B., Cavalleri, G. L., Cherny, S. S., Chinthapalli, K., Coffey, A. J., Compston, A., Coppola, A., Cossette, P., Craig, J. J., de Haan, G.-J., De Jonghe, P., de Kovel, C. G. F., Delanty, N., Depondt, C., Devinsky, O., Dlugos, D. J., Doherty, C. P., Elger, C. E., Eriksson, J. G., Ferraro, T. N., Feucht, M., Francis, B., Franke, A., French, J. A., Freytag, S., Gaus, V., Geller, E. B., Gieger, C., Glauser, T., Glynn, S., Goldstein, D. B., Gui, H., Guo, Y., Haas, K. F., Hakonarson, H., Hallmann, K., Haut, S., Heinzen, E. L., Helbig, I., Hengsbach, C., Hjalgrim, H., Iacomino, M., Ingason, A., Jamnadas-Khoda, J., Johnson, M. R., Kälviäinen, R., Kantanen, A.-M., Kasperavičiūte, D., Kasteleijn-Nolst Trenite, D., Kirsch, H. E., Knowlton, R. C., Koeleman, B. P. C., Krause, R., Krenn, M., Kunz, W. S., Kuzniecky, R., Kwan, P., Lal, D., Lau, Y.-L., Lehesjoki, A.-E., Lerche, H., Leu, C., Lieb, W., Lindhout, D., Lo, W. D., Lopes-Cendes, I., Lowenstein, D. H., Malovini, A., Marson, A. G., Mayer, T., McCormack, M., Mills, J. L., Mirza, N., Moerzinger, M., Møller, R. S., Molloy, A. M., Muhle, H., Newton, M., Ng, P.-W., Nöthen, M. M., Nürnberg, P., O'Brien, T. J., Oliver, K. L., Palotie, A., Pangilinan, F., Peter, S., Petrovski, S., Poduri, A., Privitera, M., Radtke, R., Rau, S., Reif, P. S., Reinthaler, E. M., Rosenow, F., Sander, J. W., Sander, T., Scattergood, T., Schachter, S. C., Schankin, C. J., Scheffer, I. E., Schmitz, B., Schoch, S., Sham, P. C., Shih, J. J., Sills, G. J., Sisodiya, S. M., Slattery, L., Smith, A., Smith, D. F., Smith, M. C., Smith, P. E., Sonsma, A. C. M., Speed, D., Sperling, M. R., Steinhoff, B. J., Stephani, U., Stevelink, R., Strauch, K., Striano, P., Stroink, H., Surges, R., Tan, K. M., Thio, L. L., Thomas, G. N., Todaro, M., Tozzi, R., Vari, M. S., Vining, E. P. G., Visscher, F., von Spiczak, S., Walley, N. M., Weber, Y. G., Wei, Z., Weisenberg, J., Whelan, C. D., Widdess-Walsh, P., Wolff, M., Wolking, S., Yang, W., Zara, F., Zimprich, F., & The International League Against Epilepsy Consortium on Complex, E. (2018). Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nature Communications*, *9*(1), 5269. https://doi.org/10.1038/s41467-018-07524-z

Abou Tayoun, A. N., Pesaran, T., DiStefano, M. T., Oza, A., Rehm, H. L., Biesecker, L. G., Harrison, S. M., & Group, C. S. V. I. W. (2018). Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Human mutation*, *39*(11), 1517-1524.

Abreu, P. C., Greenberg, D. A., & Hodge, S. E. (1999). Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. *The American Journal of Human Genetics*, *65*(3), 847-857.

Adamska-Patruno, E., Goscik, J., Czajkowski, P., Maliszewska, K., Ciborowski, M., Golonko, A., Wawrusiewicz-Kurylonek, N., Citko, A., Waszczeniuk, M., Kretowski, A., & Gorska, M. (2019). The MC4R genetic variants are associated with lower visceral fat accumulation and higher postprandial relative increase in carbohydrate utilization in humans. *Eur J Nutr*, *58*(7), 2929-2941. https://doi.org/10.1007/s00394-019-01955-0

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, *76*(1), 7.20. 21-27.20. 41.

Ahmed, M. Y., Al-Khayat, A., Al-Murshedi, F., Al-Futaisi, A., Chioza, B. A., Pedro Fernandez-Murray, J., Self, J. E., Salter, C. G., Harlalka, G. V., Rawlins, L. E., Al-Zuhaibi, S., Al-Azri, F., Al-Rashdi, F., Cazenave-Gassiot, A., Wenk, M. R., Al-Salmi, F., Patton, M. A., Silver, D. L., Baple, E. L., McMaster, C. R., & Crosby, A. H. (2017). A mutation of EPT1 (SELENOI) underlies a new disorder of Kennedy pathway phospholipid biosynthesis. *Brain*, *140*(3), 547-554. https://doi.org/10.1093/brain/aww318

Alakurtti, K., Weber, E., Rinne, R., Theil, G., Haan, G.-J. d., Lindhout, D., Salmikangas, P., Saukko, P., Lahtinen, U., & Lehesjoki, A.-E. (2005). Loss of lysosomal association of cystatin B proteins representing progressive myoclonus epilepsy, EPM1, mutations. *European Journal of Human Genetics*, *13*(2), 208-215.

Albert, T. J., Molla, M. N., Muzny, D. M., Nazareth, L., Wheeler, D., Song, X., Richmond, T. A., Middle, C. M., Rodesch, M. J., Packard, C. J., Weinstock, G. M., & Gibbs, R. A. (2007). Direct selection of human genomic loci by microarray hybridization. *Nature Methods*, *4*(11), 903-905. https://doi.org/10.1038/nmeth1111

Annas, G. J., & Elias, S. (2014). 23andMe and the FDA. *New England Journal of Medicine*, *370*(11), 985-988.

Appleton, R. E. (2001). West syndrome: long-term prognosis and social aspects. *Brain and Development*, *23*(7), 688-691. https://doi.org/https://doi.org/10.1016/S0387-7604(01)00264-9

Arnemann, J. (2019). Kopplungsanalyse. In A. M. Gressner & T. Arndt (Eds.), *Lexikon der Medizinischen Laboratoriumsdiagnostik* (pp. 1386-1386). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-48986-4_3516

Aslam, B., Basit, M., Nisar, M. A., Khurshid, M., & Rasool, M. H. (2017). Proteomics: Technologies and Their Applications. *Journal of Chromatographic Science*, *55*(2), 182-196. https://doi.org/10.1093/chromsci/bmw167

Bahlo, M., Bennett, M. F., Degorski, P., Tankard, R. M., Delatycki, M. B., & Lockhart, P. J. (2018). Recent advances in the detection of repeat expansions with short-read next-generation sequencing. *F1000Research*, *7*(736), 736.

Baldassari, S., Picard, F., Verbeek, N. E., van Kempen, M., Brilstra, E. H., Lesca, G., Conti, V., Guerrini, R., Bisulli, F., Licchetta, L., Pippucci, T., Tinuper, P., Hirsch, E., de Saint Martin, A., Chelly, J., Rudolf, G., Chipaux, M., Ferrand-Sorbets, S., Dorfmüller, G., Sisodiya, S., Balestrini, S., Schoeler, N., Hernandez-Hernandez, L., Krithika, S., Oegema, R., Hagebeuk, E., Gunning, B., Deckers, C., Berghuis, B., Wegner, I., Niks, E., Jansen, F. E., Braun, K., de Jong, D., Rubboli, G., Talvik, I., Sander, V., Uldall, P., Jacquemont, M.-L., Nava, C., Leguern, E., Julia, S., Gambardella, A., d'Orsi, G., Crichiutti, G., Faivre, L., Darmency, V., Benova, B., Krsek, P., Biraben, A., Lebre, A.-S., Jennesson, M., Sattar, S., Marchal, C., Nordli, D. R., Lindstrom, K., Striano, P., Lomax, L. B., Kiss, C., Bartolomei, F., Lepine, A. F., Schoonjans, A.-S., Stouffs, K., Jansen, A., Panagiotakaki, E., Ricard-Mousnier, B., Thevenon, J., de Bellescize, J.,

Catenoix, H., Dorn, T., Zenker, M., Müller-Schlüter, K., Brandt, C., Krey, I., Polster, T., Wolff, M., Balci, M., Rostasy, K., Achaz, G., Zacher, P., Becher, T., Cloppenborg, T., Yuskaitis, C. J., Weckhuysen, S., Poduri, A., Lemke, J. R., Møller, R. S., & Baulac, S. (2019). The landscape of epilepsy-related GATOR1 variants. *Genetics in Medicine*, *21*(2), 398-408. https://doi.org/10.1038/s41436-018-0060-2

Banach, M., Gurdziel, E., Jędrych, M., & Borowicz, K. K. (2011). Melatonin in experimental seizures and epilepsy. *Pharmacol Rep*, *63*(1), 1-11. https://doi.org/10.1016/s1734-1140(11)70393-0

Barker-Haliski, M., & White, H. S. (2015). Glutamatergic Mechanisms Associated with Seizures and Epilepsy. *Cold Spring Harbor perspectives in medicine*, *5*(8), a022863-a022863. https://doi.org/10.1101/cshperspect.a022863

Barkovich, A. J., Dobyns, W. B., & Guerrini, R. (2015). Malformations of cortical development and epilepsy. *Cold Spring Harbor perspectives in medicine*, *5*(5), a022392-a022392. https://doi.org/10.1101/cshperspect.a022392

Bashiardes, S., Veile, R., Helms, C., Mardis, E. R., Bowcock, A. M., & Lovett, M. (2005). Direct genomic selection. *Nature Methods*, *2*(1), 63-69.

Baulac, S., Gourfinkel-An, I., Picard, F., Rosenberg-Bourgin, M., Prud'homme, J.-F., Baulac, M., Brice, A., & LeGuern, E. (1999). A Second Locus for Familial Generalized Epilepsy with Febrile Seizures Plus Maps to Chromosome 2q21-q33. *The American Journal of Human Genetics*, *65*(4), 1078-1085. https://doi.org/https://doi.org/10.1086/302593

Begley, C. E., & Beghi, E. (2002). The economic cost of epilepsy: a review of the literature. *Epilepsia*, *43*, 3-9.

Begum, F., Ghosh, D., Tseng, G. C., & Feingold, E. (2012). Comprehensive literature review and statistical considerations for GWAS meta-analysis. *Nucleic Acids Research*, *40*(9), 3777-3784. https://doi.org/10.1093/nar/gkr1255

Berezin, V., Walmod, P. S., Filippov, M., & Dityatev, A. (2014). Targeting of ECM molecules and their metabolizing enzymes and receptors for the treatment of CNS diseases. *Progress in brain research*, *214*, 353-388.

Berg, A. T., Berkovic, S. F., Brodie, M. J., Buchhalter, J., Cross, J. H., Van Emde Boas, W., Engel, J., French, J., Glauser, T. A., Mathern, G. W., Moshé, S. L., Nordli, D., Plouin, P., & Scheffer, I. E. (2010). Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005–2009. *Epilepsia*, *51*(4), 676-685. https://doi.org/10.1111/j.1528-1167.2010.02522.x

Berkovic, S. F., Howell, R. A., Hay, D. A., & Hopper, J. L. (1998). Epilepsies in twins: genetics of the major epilepsy syndromes. *Ann Neurol*, *43*(4), 435-445. https://doi.org/10.1002/ana.410430405

Berkovic, S. F., Howell, R. A., Hay, D. A., & Hopper, J. L. (1998). Epilepsies in twins: Genetics of the major epilepsy syndromes. *Annals of Neurology*, *43*(4), 435-445. https://doi.org/https://doi.org/10.1002/ana.410430405

Bersanelli, M., Mosca, E., Remondini, D., Giampieri, E., Sala, C., Castellani, G., & Milanesi, L. (2016). Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinformatics*, *17 Suppl 2*(Suppl 2), 15. https://doi.org/10.1186/s12859-015-0857-9

Bezerra, G. A., Dobrovetsky, E., Seitova, A., Fedosyuk, S., Dhe-Paganon, S., & Gruber, K. (2015). Structure of human dipeptidyl peptidase 10 (DPPY): a modulator of neuronal Kv4 channels. *Scientific Reports*, *5*(1), 8769. https://doi.org/10.1038/srep08769

Bien, C. G., & Elger, C. E. (2007). Limbic encephalitis: A cause of temporal lobe epilepsy with onset in adult life. *Epilepsy & Behavior*, *10*(4), 529-538. https://doi.org/https://doi.org/10.1016/j.yebeh.2007.03.011

Biervert, C., Schroeder, B. C., Kubisch, C., Berkovic, S. F., Propping, P., Jentsch, T. J., & Steinlein, O. K. (1998). A potassium channel mutation in neonatal human epilepsy. *Science*, *279*(5349), 403-406. https://doi.org/10.1126/science.279.5349.403

Birman, H., Gülay, Ü., Dar, K. A., Kapucu, A., & Samet, A. (2012). Effects of luteolin on liver, kidney and brain in pentylentetrazol-induced seizures: involvement of metalloproteinases and NOS activities. *Balkan medical journal*, *2012*(2), 188-196.

Bland, J. M., & Altman, D. G. (2000). The odds ratio. *Bmj*, *320*(7247), 1468.

Blümcke, I., Thom, M., Aronica, E., Armstrong, D. D., Vinters, H. V., Palmini, A., Jacques, T. S., Avanzini, G., Barkovich, A. J., Battaglia, G., Becker, A., Cepeda, C., Cendes, F., Colombo, N., Crino, P., Cross, J. H., Delalande, O., Dubeau, F., Duncan, J., Guerrini, R., Kahane, P., Mathern, G., Najm, I., Ozkara, C., Raybaud, C., Represa, A., Roper, S. N., Salamon, N., Schulze-Bonhage, A., Tassi, L., Vezzani, A., & Spreafico, R. (2011). The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission. *Epilepsia*, *52*(1), 158-174. https://doi.org/10.1111/j.1528-1167.2010.02777.x

Boada-Romero, E., Letek, M., Fleischer, A., Pallauf, K., Ramón-Barros, C., & Pimentel-Muiños, F. X. (2013). TMEM59 defines a novel ATG16L1-binding motif that promotes local activation of LC3. *Embo j*, *32*(4), 566-582. https://doi.org/10.1038/emboj.2013.8

Bock, C., Farlik, M., & Sheffield, N. C. (2016). Multi-Omics of Single Cells: Strategies and Applications. *Trends Biotechnol*, *34*(8), 605-608. https://doi.org/10.1016/j.tibtech.2016.04.004

Bourgeois, B. F., Dodson, E., Nordli Jr, D. R., Pellock, J. M., & Sankar, R. (2007). *Pediatric epilepsy: diagnosis and therapy*. Demos Medical Publishing.

Bouvard, D., Pouwels, J., De Franceschi, N., & Ivaska, J. (2013). Integrin inactivators: balancing cellular functions in vitro and in vivo. *Nature reviews Molecular cell biology*, *14*(7), 430-442.

Brocker, C., Cantore, M., Failli, P., & Vasiliou, V. (2011). Aldehyde dehydrogenase 7A1 (ALDH7A1) attenuates reactive aldehyde and oxidative stress induced cytotoxicity. *Chemico-biological interactions*, *191*(1-3), 269-277.

Brookes, A. J. (1999). The essence of SNPs. *Gene*, *234*(2), 177-186. https://doi.org/https://doi.org/10.1016/S0378-1119(99)00219-X

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLOS Computational Biology*, *8*(12), e1002822. https://doi.org/10.1371/journal.pcbi.1002822

Camfield, P., & Camfield, C. (2002). Epileptic Syndromes in Childhood: Clinical Features, Outcomes, and Treatment. *Epilepsia*, *43*(s3), 27-32. https://doi.org/https://doi.org/10.1046/j.1528-1157.43.s.3.3.x

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1). https://doi.org/10.1186/s13742-015-0047-8

Chapuis, J., Flaig, A., Grenier-Boley, B., Eysert, F., Pottiez, V., Deloison, G., Vandeputte, A., Ayral, A.-M., Mendes, T., Desai, S., Goate, A. M., Kauwe, J. S. K., Leroux, F., Herledan, A., Demiautte, F., Bauer, C., Checler, F., Petersen, R. C., Blennow, K., Zetterberg, H., Minthon, L., Van Deerlin, V. M., Lee, V. M.-Y., Shaw, L. M., Trojanowski, J. Q., Albert, M., Moghekar, A., O'Brien, R., Peskind, E. R., Malmanche, N., Schellenberg, G. D., Dourlen, P., Song, O.-R., Cruchaga, C., Amouyel, P., Deprez, B., Brodin, P., Lambert, J.-C., & Adgc, A. s. D. N. I. (2017). Genome-wide, high-content siRNA screening identifies the Alzheimer's genetic risk factor FERMT2 as a major modulator of APP metabolism. *Acta Neuropathologica*, *133*(6), 955-966. https://doi.org/10.1007/s00401-016-1652-z

Chishti, A. H., Kim, A. C., Marfatia, S. M., Lutchman, M., Hanspal, M., Jindal, H., Liu, S. C., Low, P. S., Rouleau, G. A., Mohandas, N., Chasis, J. A., Conboy, J. G., Gascard, P., Takakuwa, Y., Huang, S. C., Benz, E. J., Jr., Bretscher, A., Fehon, R. G., Gusella, J. F., Ramesh, V., Solomon, F., Marchesi, V. T., Tsukita, S., Tsukita, S., Hoover, K. B., & et al. (1998). The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane. *Trends Biochem Sci*, *23*(8), 281-282. https://doi.org/10.1016/s0968-0004(98)01237-7

Cho, W. C. (2007). Proteomics technologies and challenges. *Genomics, proteomics & bioinformatics*, *5*(2), 77-85.

Chu, Y., & Corey, D. R. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic Acid Ther*, *22*(4), 271-274. https://doi.org/10.1089/nat.2012.0367

Chun, S., & Fay, J. C. (2009). Identification of deleterious mutations within three human genomes. *Genome research*, *19*(9), 1553-1561.

Cock, P. J., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, *38*(6), 1767-1771.

Cohen, I., Rider, P., Vornov, E., Tomas, M., Tudor, C., Wegner, M., Brondani, L., Freudenberg, M., Mittler, G., Ferrando-May, E., Dinarello, C. A., Apte, R. N., & Schneider, R. (2015). IL-1α is a DNA damage sensor linking genotoxic stress

signaling to sterile inflammation and innate immunity. *Sci Rep, 5*, 14756. https://doi.org/10.1038/srep14756

Combi, R., Dalprà, L., Tenchini, M. L., & Ferini-Strambi, L. (2004). Autosomal dominant nocturnalfrontal lobe epilepsy. *Journal of Neurology, 251*(8), 923-934. https://doi.org/10.1007/s00415-004-0541-x

Conti, V., Aracri, P., Chiti, L., Brusco, S., Mari, F., Marini, C., Albanese, M., Marchi, A., Liguori, C., Placidi, F., Romigi, A., Becchetti, A., & Guerrini, R. (2015). Nocturnal frontal lobe epilepsy with paroxysmal arousals due to CHRNA2 loss of function. *Neurology, 84*(15), 1520-1528. https://doi.org/10.1212/WNL.0000000000001471

Cooper, G. M., Stone, E. A., Asimenos, G., Green, E. D., Batzoglou, S., & Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome research, 15*(7), 901-913.

Cordell, H. J., & Clayton, D. G. (2005). Genetic association studies. *The Lancet, 366*(9491), 1121-1131.

Cronn, R., Knaus, B. J., Liston, A., Maughan, P. J., Parks, M., Syring, J. V., & Udall, J. (2012). Targeted enrichment strategies for next-generation plant biology. *American journal of botany, 99*(2), 291-311.

Cui, Y., Li, G., Li, S., & Wu, R. (2010). Designs for linkage analysis and association studies of complex diseases. *Statistical Methods in Molecular Biology*, 219-242.

Danielsson, J., & Petermann, F. (2009). Cognitive deficits in children with benign rolandic epilepsy of childhood or rolandic discharges: A study of children between 4 and 7 years of age with and without seizures compared with healthy controls. *Epilepsy & Behavior, 16*(4), 646-651. https://doi.org/https://doi.org/10.1016/j.yebeh.2009.08.012

Deprez, L., Weckhuysen, S., Holmgren, P., Suls, A., Van Dyck, T., Goossens, D., Del-Favero, J., Jansen, A., Verhaert, K., & Lagae, L. (2010). Clinical spectrum of early-onset epileptic encephalopathies associated with STXBP1 mutations. *Neurology, 75*(13), 1159-1165.

Díaz-Rodríguez, S. M., López-López, D., Herrero-Turrión, M. J., Gómez-Nieto, R., Canal-Alonso, A., & Lopéz, D. E. (2020). Inferior colliculus transcriptome after status epilepticus in the genetically audiogenic seizure-prone hamster GASH/Sal. *Frontiers in Neuroscience, 14*, 508.

Dimassi, S., Labalme, A., Lesca, G., Rudolf, G., Bruneau, N., Hirsch, E., Arzimanoglou, A., Motte, J., de Saint Martin, A., Boutry-Kryza, N., Cloarec, R., Benitto, A., Ameil, A., Edery, P., Ryvlin, P., De Bellescize, J., Szepetowski, P., & Sanlaville, D. (2014). A subset of genomic alterations detected in rolandic epilepsies contains candidate or known epilepsy genes including GRIN2A and PRRT2. *Epilepsia, 55*(2), 370-378. https://doi.org/https://doi.org/10.1111/epi.12502

Dohm, J. C., Lottaz, C., Borodina, T., & Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research, 36*(16), e105.

Dravet, C. (2011). The core Dravet syndrome phenotype. *Epilepsia*, *52*, 3-9.

Dreifuss, F., Martinez-Lage, M., & Johns, R. A. (1985). Proposal for classification of epilepsies and epileptic syndromes. *Epilepsia*, *26*(3), 268-278.

Dulac, O. (2001). What is West syndrome? *Brain and Development*, *23*(7), 447-452. https://doi.org/https://doi.org/10.1016/S0387-7604(01)00268-6

Ellard, S., Baple, E. L., Owens, M., Eccles, D. M., Abbs, S., Deans, Z. C., & McMullan, D. (2019). ACGS best practice guidelines for variant classification 2019. *ACGS Guidelines*.

Endo, W., Ikemoto, S., Togashi, N., Miyabayashi, T., Nakajima, E., Hamano, S.-i., Shibuya, M., Sato, R., Takezawa, Y., & Okubo, Y. (2020). Phenotype–genotype correlations in patients with GNB1 gene variants, including the first three reported Japanese patients to exhibit spastic diplegia, dyskinetic quadriplegia, and infantile spasms. *Brain and Development*, *42*(2), 199-204.

Engel, T., & Henshall, D. C. (2009). Apoptosis, Bcl-2 family proteins and caspases: the ABCs of seizure-damage and epileptogenesis? *International journal of physiology, pathophysiology and pharmacology*, *1*(2), 97-115. https://pubmed.ncbi.nlm.nih.gov/21383882
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3047241/

Engelborghs, S., D'hooge, R., & De Deyn, P. (2000). Pathophysiology of epilepsy. *Acta neurologica belgica*, *100*(4), 201-213.

Epilepsies, T. I. L. A. E. C. o. C. (2014). Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol*, *13*(9), 893-903. https://doi.org/10.1016/s1474-4422(14)70171-1

Epilepsy, C. o. C. a. T. o. t. I. L. A. (1989). Proposal for Revised Classification of Epilepsies and Epileptic Syndromes. *Epilepsia*, *30*(4), 389-399. https://doi.org/10.1111/j.1528-1157.1989.tb05316.x

Epilepsy, I. L. A. (2020). *LENNOX GASTAUT SYNDROM*. Retrieved 09.04.2020 from epilepsydiagnosis.org/syndrome/lgs-overview.html

Erdmann, J., Stark, K., Esslinger, U. B., Rumpf, P. M., Koesling, D., De Wit, C., Kaiser, F. J., Braunholz, D., Medack, A., & Fischer, M. (2013). Dysfunctional nitric oxide signalling increases risk of myocardial infarction. *Nature*, *504*(7480), 432-436.

Evans, D. M., Gillespie, N., & Martin, N. (2002). Biometrical genetics. *Biological psychology*, *61*(1-2), 33-51.

Farrant, M., & Nusser, Z. (2005). Variations on an inhibitory theme: phasic and tonic activation of GABAA receptors. *Nature Reviews Neuroscience*, *6*(3), 215-229.

Filakovszky, J., Gerber, K., & Bagdy, G. (1999). A serotonin-1A receptor agonist and an N-methyl-D-aspartate receptor antagonist oppose each others effects in a genetic rat epilepsy model. *Neurosci Lett*, *261*(1-2), 89-92. https://doi.org/10.1016/s0304-3940(99)00015-4

Fisher, R. S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J. H., Elger, C. E., Engel, J., Jr., Forsgren, L., French, J. A., Glynn, M., Hesdorffer, D. C., Lee, B. I., Mathern, G. W., Moshe, S. L., Perucca, E., Scheffer, I. E., Tomson, T., Watanabe, M., & Wiebe, S. (2014). ILAE official report: a practical clinical definition of epilepsy. *Epilepsia, 55*(4), 475-482. https://doi.org/10.1111/epi.12550

Fisher, R. S., Cross, J. H., French, J. A., Higurashi, N., Hirsch, E., Jansen, F. E., Lagae, L., Moshé, S. L., Peltola, J., Roulet Perez, E., Scheffer, I. E., & Zuberi, S. M. (2018). Operationale Klassifikation der Anfallsformen durch die Internationale Liga gegen Epilepsie: Positionspapier der ILAE-Klassifikations- und Terminologiekommission. *Zeitschrift für Epileptologie, 31*(4), 272-281. https://doi.org/10.1007/s10309-018-0216-8

Fister, P., Soltirovska-Salamon, A., Debeljak, M., & Paro-Panjan, D. (2013). Benign familial neonatal convulsions caused by mutation in KCNQ3, exon 6: a European case. *Eur J Paediatr Neurol, 17*(3), 308-310. https://doi.org/10.1016/j.ejpn.2012.10.007

Fukuda, T., Hiraide, T., Yamoto, K., Nakashima, M., Kawai, T., Yanagi, K., Ogata, T., & Saitsu, H. (2020). Exome reports a de novo GNB2 variant associated with global developmental delay, intellectual disability, and dysmorphic features. *European Journal of Medical Genetics, 63*(4), 103804.

Galli, J., Micheletti, S., Malerba, L., Fazzi, E., & Giordano, L. (2018). Childhood absence epilepsy evolving to eyelid myoclonia with absence epilepsy. *Seizure, 61*, 1-3.

Gan, J., Cai, Q., Qu, Y., Zhao, F., Wan, C., Luo, R., & Mu, D. (2017). miR-96 attenuates status epilepticus-induced brain injury by directly targeting Atg7 and Atg16L1. *Scientific Reports, 7*(1), 1-13.

Gao, Q.-W., Hua, L.-D., Wang, J., Fan, C.-X., Deng, W.-Y., Li, B., Bian, W.-J., Shao, C.-X., He, N., & Zhou, P. (2017). A point mutation in SCN1A 5′ genomic region decreases the promoter activity and is associated with mild epilepsy and seizure aggravation induced by antiepileptic drug. *Molecular neurobiology, 54*(4), 2428-2434.

Ge, W.-R., Wan, L., & Yang, G. (2021). Melatonin for an obese child with MC4R gene variant showing epilepsy and disordered sleep: A case report. *World journal of clinical cases, 9*(11), 2688-2695. https://doi.org/10.12998/wjcc.v9.i11.2688

Genetics, C. f. s., & Michigan, U. o. (2020). *MERLIN Tutorial -- Haplotyping*. Retrieved 17.04.2020 from csg.sph.umich.edu/abecasis/merlin/tour/haplotyping.html

Germer, S., Holland, M. J., & Higuchi, R. (2000). High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. *Genome research, 10*(2), 258-266.

Geva-Dayan, K., Shorer, Z., Menascu, S., Linder, I., Goldberg-Stern, H., Heyman, E., Lerman-Sagie, T., Ben Zeev, B., & Kramer, U. (2012). Immunoglobulin Treatment for Severe Childhood Epilepsy. *Pediatric Neurology, 46*(6), 375-381. https://doi.org/https://doi.org/10.1016/j.pediatrneurol.2012.03.015

Gibbs, R. A., Belmont, J. W., Hardenbol, P., Willis, T. D., Yu, F., Yang, H., Ch'ang, L.-Y., Huang, W., Liu, B., Shen, Y., Tam, P. K.-H., Tsui, L.-C., Waye, M. M. Y., Wong, J.

T.-F., Zeng, C., Zhang, Q., Chee, M. S., Galver, L. M., Kruglyak, S., Murray, S. S., Oliphant, A. R., Montpetit, A., Hudson, T. J., Chagnon, F., Ferretti, V., Leboeuf, M., Phillips, M. S., Verner, A., Kwok, P.-Y., Duan, S., Lind, D. L., Miller, R. D., Rice, J. P., Saccone, N. L., Taillon-Miller, P., Xiao, M., Nakamura, Y., Sekine, A., Sorimachi, K., Tanaka, T., Tanaka, Y., Tsunoda, T., Yoshino, E., Bentley, D. R., Deloukas, P., Hunt, S., Powell, D., Altshuler, D., Gabriel, S. B., Zhang, H., Zeng, C., Matsuda, I., Fukushima, Y., Macer, D. R., Suda, E., Rotimi, C. N., Adebamowo, C. A., Aniagwu, T., Marshall, P. A., Matthew, O., Nkwodimmah, C., Royal, C. D. M., Leppert, M. F., Dixon, M., Stein, L. D., Cunningham, F., Kanani, A., Thorisson, G. A., Chakravarti, A., Chen, P. E., Cutler, D. J., Kashuk, C. S., Donnelly, P., Marchini, J., McVean, G. A. T., Myers, S. R., Cardon, L. R., Abecasis, G. R., Morris, A., Weir, B. S., Mullikin, J. C., Sherry, S. T., Feolo, M., Altshuler, D., Daly, M. J., Schaffner, S. F., Qiu, R., Kent, A., Dunston, G. M., Kato, K., Niikawa, N., Knoppers, B. M., Foster, M. W., Clayton, E. W., Wang, V. O., Watkin, J., Gibbs, R. A., Belmont, J. W., Sodergren, E., Weinstock, G. M., Wilson, R. K., Fulton, L. L., Rogers, J., Birren, B. W., Han, H., Wang, H., Godbout, M., Wallenburg, J. C., L'Archevêque, P., Bellemare, G., Todani, K., Fujita, T., Tanaka, S., Holden, A. L., Lai, E. H., Collins, F. S., Brooks, L. D., McEwen, J. E., Guyer, M. S., Jordan, E., Peterson, J. L., Spiegel, J., Sung, L. M., Zacharia, L. F., Kennedy, K., Dunn, M. G., Seabrook, R., Shillito, M., Skene, B., Stewart, J. G., Valle, D. L., Clayton, E. W., Jorde, L. B., Belmont, J. W., Chakravarti, A., Cho, M. K., Duster, T., Foster, M. W., Jasperse, M., Knoppers, B. M., Kwok, P.-Y., Licinio, J., Long, J. C., Marshall, P. A., Ossorio, P. N., Wang, V. O., Rotimi, C. N., Royal, C. D. M., Spallone, P., Terry, S. F., Lander, E. S., Lai, E. H., Nickerson, D. A., Abecasis, G. R., Altshuler, D., Bentley, D. R., Boehnke, M., Cardon, L. R., Daly, M. J., Deloukas, P., Douglas, J. A., Gabriel, S. B., Hudson, R. R., Hudson, T. J., Kruglyak, L., Kwok, P.-Y., Nakamura, Y., Nussbaum, R. L., Royal, C. D. M., Schaffner, S. F., Sherry, S. T., Stein, L. D., Tanaka, T., †The International HapMap, C., Genotyping centres: Baylor College of, M., ParAllele, B., Chinese HapMap, C., Illumina, McGill, U., Génome Québec Innovation, C., University of California at San, F., Washington, U., University of, T., Riken, Wellcome Trust Sanger, I., Whitehead Institute, M. I. T. C. f. G. R., Community engagement/public, c., sample-collection groups: Beijing Normal, U., Beijing Genomics, I., Health Sciences University of Hokkaido, E. E. I., Shinshu, U., Howard, U., University of, I., University of, U., Analysis Groups: Cold Spring Harbor, L., Johns Hopkins University School of, M., University of, O., University of Oxford, W. T. C. f. H. G., Health, U. S. N. I. o., Ethical, L., Social Issues: Chinese Academy of Social, S., Genetic Interest, G., Howard, U., Kyoto, U., Nagasaki, U., University of, M., University of, O., Vanderbilt, U., Wellcome, T., Medicine, S. N. P. D. B. C. o., Washington, U., Scientific Management: Chinese Academy of, S., Chinese Ministry of, S., Technology, Genome, C., Génome, Q., Japanese Ministry of Education, C. S. S., Technology, The, S. N. P. C., Initial Planning Groups, P., Ethical, L., Social Issues, G., & Methods, G. (2003). The International HapMap Project. *Nature*, *426*(6968), 789-796. https://doi.org/10.1038/nature02168

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333-351.

Gourfinkel-An, I., Baulac, S., Nabbout, R., Ruberg, M., Baulac, M., Brice, A., & LeGuern, E. (2004). Monogenic idiopathic epilepsies. *The Lancet Neurology*, *3*(4), 209-218.

Graw, J. (2020). Was ist Genetik? In *Genetik* (pp. 1-24). Springer.

Guazzi, M., & Striano, P. (2019). GABA strikes down again in epilepsy. *Annals of translational medicine, 7*(3), 57-57. https://doi.org/10.21037/atm.2018.12.55

Gündüz, E., Demirbilek, V., & Korkmaz, B. (1999). Benign rolandic epilepsy: neuropsychological findings. *Seizure, 8*(4), 246-249. https://doi.org/https://doi.org/10.1053/seiz.1999.0293

Habermann-Horstmeier, L. (2008). Epilepsie und Vererbung. *Dt. Gesellschaft für Epileptologie e.V.*

.

http://www.dgfe.org/home/showdoc,id,433,aid,219.html#:~:text=Bis%20heute%20sind%20etwa%20160,%25%20aller%20Epilepsie%2DErkrankungen%20aus.

Han, C.-L., Zhao, X.-M., Liu, Y.-P., Wang, K.-L., Chen, N., Hu, W., Zhang, J.-G., Ge, M., & Meng, F.-G. (2019). Gene expression profiling of two epilepsy models reveals the ECM/integrin signaling pathway is involved in epiletogenesis. *Neuroscience, 396*, 187-199.

Hanash, S. (2003). Disease proteomics. *Nature, 422*(6928), 226-232.

Harakalova, M., Mokry, M., Hrdlickova, B., Renkens, I., Duran, K., Van Roekel, H., Lansu, N., Van Roosmalen, M., De Bruijn, E., & Nijman, I. J. (2011). Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective targeted next-generation sequencing. *Nature protocols, 6*(12), 1870-1886.

Harismendy, O., Ng, P. C., Strausberg, R. L., Wang, X., Stockwell, T. B., Beeson, K. Y., Schork, N. J., Murray, S. S., Topol, E. J., & Levy, S. (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome biology, 10*(3), 1-13.

Hauser, W. A., Annegers, J. F., & Kurland, L. T. (1993). Incidence of epilepsy and unprovoked seizures in Rochester, Minnesota: 1935-1984. *Epilepsia, 34*(3), 453-468. https://doi.org/10.1111/j.1528-1157.1993.tb02586.x

Helbig, I., Heinzen, E. L., Mefford, H. C., & Commission, t. I. G. (2016). Primer Part 1—The building blocks of epilepsy genetics. *Epilepsia, 57*(6), 861-868. https://doi.org/10.1111/epi.13381

Helbig, I., Scheffer, I. E., Mulley, J. C., & Berkovic, S. F. (2008). Navigating the channels and beyond: unravelling the genetics of the epilepsies. *Lancet Neurol, 7*(3), 231-245. https://doi.org/10.1016/s1474-4422(08)70039-5

Henshall, D. C., Clark, R. S., Adelson, P. D., Chen, M., Watkins, S. C., & Simon, R. P. (2000). Alterations in bcl-2 and caspase gene family protein expression in human temporal lobe epilepsy. *Neurology, 55*(2), 250-257. https://doi.org/10.1212/wnl.55.2.250

Herman, S. T. (2002). Epilepsy after brain insult. *Targeting epileptogenesis, 59*(9 suppl 5), S21-S26. https://doi.org/10.1212/WNL.59.9_suppl_5.S21

Hildebrand, M. S., Dahl, H. H., Damiano, J. A., Smith, R. J., Scheffer, I. E., & Berkovic, S. F. (2013). Recent advances in the molecular genetics of epilepsy. *J Med Genet*, *50*(5), 271-279. https://doi.org/10.1136/jmedgenet-2012-101448

Hirose, S. (2014). Mutant GABAA receptor subunits in genetic (idiopathic) epilepsy. *Progress in brain research*, *213*, 55-85.

Horibata, Y., Elpeleg, O., Eran, A., Hirabayashi, Y., Savitzki, D., Tal, G., Mandel, H., & Sugimoto, H. (2018). EPT1 (selenoprotein I) is critical for the neural development and maintenance of plasmalogen in humans. *J Lipid Res*, *59*(6), 1015-1026. https://doi.org/10.1194/jlr.P081620

Hosford, D. A., & Wang, Y. (1997). Utility of the Lethargic (lh/lh) Mouse Model of Absence Seizures in Predicting the Effects of Lamotrigine, Vigabatrin, Tiagabine, Gabapentin, and Topiramate Against Human Absence Seizures. *Epilepsia*, *38*(4), 408-414. https://doi.org/10.1111/j.1528-1157.1997.tb01729.x

Howles, S. A., Wiberg, A., Goldsworthy, M., Bayliss, A. L., Gluck, A. K., Ng, M., Grout, E., Tanikawa, C., Kamatani, Y., & Terao, C. (2019). Genetic variants of calcium and vitamin D metabolism in kidney stone disease. *Nature Communications*, *10*(1), 1-10.

Hrachovy, R. A., & Frost Jr, J. D. (2003). Infantile epileptic encephalopathy with hypsarrhythmia (infantile spasms/West syndrome). *Journal of clinical neurophysiology*, *20*(6), 408-425.

Hussain, M. S., Battaglia, A., Szczepanski, S., Kaygusuz, E., Toliat, M. R., Sakakibara, S.-i., Altmüller, J., Thiele, H., Nürnberg, G., & Moosa, S. (2014). Mutations in CKAP2L, the human homolog of the mouse Radmis gene, cause Filippi syndrome. *The American Journal of Human Genetics*, *95*(5), 622-632.

Hwang, S., Kim, E., Lee, I., & Marcotte, E. M. (2015). Systematic comparison of variant calling pipelines using gold standard personal exome variants. *Scientific Reports*, *5*(1), 1-8.

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., & Karyadi, D. (2016). REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *The American Journal of Human Genetics*, *99*(4), 877-885.

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., Powell, I. J., Cussenot, O., Cancel-Tassin, G., Giles, G. G., MacInnis, R. J., Maier, C., Hsieh, C. L., Wiklund, F., Catalona, W. J., Foulkes, W. D., Mandal, D., Eeles, R. A., Kote-Jarai, Z., Bustamante, C. D., Schaid, D. J., Hastie, T., Ostrander, E. A., Bailey-Wilson, J. E., Radivojac, P., Thibodeau, S. N., Whittemore, A. S., & Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet*, *99*(4), 877-885. https://doi.org/10.1016/j.ajhg.2016.08.016

Jain, P., Sharma, S., & Tripathi, M. (2013). Diagnosis and Management of Epileptic Encephalopathies in Children. *Epilepsy Research and Treatment*, *2013*, 501981. https://doi.org/10.1155/2013/501981

Jain, S. V., Horn, P. S., Simakajornboon, N., Beebe, D. W., Holland, K., Byars, A. W., & Glauser, T. A. (2015). Melatonin improves sleep in children with epilepsy: a randomized, double-blind, crossover study. *Sleep Med*, *16*(5), 637-644. https://doi.org/10.1016/j.sleep.2015.01.005

Jakobsen, L., Vanselow, K., Skogs, M., Toyoda, Y., Lundberg, E., Poser, I., Falkenby, L. G., Bennetzen, M., Westendorf, J., Nigg, E. A., Uhlen, M., Hyman, A. A., & Andersen, J. S. (2011). Novel asymmetrically localizing components of human centrosomes identified by complementary proteomics methods. *Embo j*, *30*(8), 1520-1535. https://doi.org/10.1038/emboj.2011.63

Jansen, A. C., Sancak, O., D'Agostino, M. D., Badhwar, A., Roberts, P., Gobbi, G., Wilkinson, R., Melanson, D., Tampieri, D., Koenekoop, R., Gans, M., Maat-Kievit, A., Goedbloed, M., van den Ouweland, A. M., Nellist, M., Pandolfo, M., McQueen, M., Sims, K., Thiele, E. A., Dubeau, F., Andermann, F., Kwiatkowski, D. J., Halley, D. J., & Andermann, E. (2006). Unusually mild tuberous sclerosis phenotype is associated with TSC2 R905Q mutation. *Ann Neurol*, *60*(5), 528-539. https://doi.org/10.1002/ana.21037

Jeanmonod, R., Asuka, E., & Jeanmonod, D. (2021). Inborn errors of metabolism. In *StatPearls [Internet]*. StatPearls Publishing.

Jerng, H. H., Qian, Y., & Pfaffinger, P. J. (2004). Modulation of Kv4.2 channel expression and gating by dipeptidyl peptidase 10 (DPP10). *Biophys J*, *87*(4), 2380-2396. https://doi.org/10.1529/biophysj.104.042358

Jessberger, S., Römer, B., Babu, H., & Kempermann, G. (2005). Seizures induce proliferation and dispersion of doublecortin-positive hippocampal progenitor cells. *Experimental Neurology*, *196*(2), 342-351. https://doi.org/https://doi.org/10.1016/j.expneurol.2005.08.010

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., Walters, R. K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J. X., Samocha, K. E., Pierce-Hoffman, E., Zappala, Z., O'Donnell-Luria, A. H., Minikel, E. V., Weisburd, B., Lek, M., Ware, J. S., Vittal, C., Armean, I. M., Bergelson, L., Cibulskis, K., Connolly, K. M., Covarrubias, M., Donnelly, S., Ferriera, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M. E., Aguilar Salinas, C. A., Ahmad, T., Albert, C. M., Ardissino, D., Atzmon, G., Barnard, J., Beaugerie, L., Benjamin, E. J., Boehnke, M., Bonnycastle, L. L., Bottinger, E. P., Bowden, D. W., Bown, M. J., Chambers, J. C., Chan, J. C., Chasman, D., Cho, J., Chung, M. K., Cohen, B., Correa, A., Dabelea, D., Daly, M. J., Darbar, D., Duggirala, R., Dupuis, J., Ellinor, P. T., Elosua, R., Erdmann, J., Esko, T., Färkkilä, M., Florez, J., Franke, A., Getz, G., Glaser, B., Glatt, S. J., Goldstein, D.,

Gonzalez, C., Groop, L., Haiman, C., Hanis, C., Harms, M., Hiltunen, M., Holi, M. M., Hultman, C. M., Kallela, M., Kaprio, J., Kathiresan, S., Kim, B.-J., Kim, Y. J., Kirov, G., Kooner, J., Koskinen, S., Krumholz, H. M., Kugathasan, S., Kwak, S. H., Laakso, M., Lehtimäki, T., Loos, R. J. F., Lubitz, S. A., Ma, R. C. W., MacArthur, D. G., Marrugat, J., Mattila, K. M., McCarroll, S., McCarthy, M. I., McGovern, D., McPherson, R., Meigs, J. B., Melander, O., Metspalu, A., Neale, B. M., Nilsson, P. M., O'Donovan, M. C., Ongur, D., Orozco, L., Owen, M. J., Palmer, C. N. A., Palotie, A., Park, K. S., Pato, C., Pulver, A. E., Rahman, N., Remes, A. M., Rioux, J. D., Ripatti, S., Roden, D. M., Saleheen, D., Salomaa, V., Samani, N. J., Scharf, J., Schunkert, H., Shoemaker, M. B., Sklar, P., Soininen, H., Sokol, H., Spector, T., Sullivan, P. F., Suvisaari, J., Tai, E. S., Teo, Y. Y., Tiinamaija, T., Tsuang, M., Turner, D., Tusie-Luna, T., Vartiainen, E., Vawter, M. P., Ware, J. S., Watkins, H., Weersma, R. K., Wessman, M., Wilson, J. G., Xavier, R. J., Neale, B. M., Daly, M. J., MacArthur, D. G., & Genome Aggregation Database, C. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature, 581*(7809), 434-443. https://doi.org/10.1038/s41586-020-2308-7

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., Walters, R. K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J. X., Samocha, K. E., Pierce-Hoffman, E., Zappala, Z., O'Donnell-Luria, A. H., Minikel, E. V., Weisburd, B., Lek, M., Ware, J. S., Vittal, C., Armean, I. M., Bergelson, L., Cibulskis, K., Connolly, K. M., Covarrubias, M., Donnelly, S., Ferriera, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M. E., Neale, B. M., Daly, M. J., & MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*, 531210. https://doi.org/10.1101/531210

Keck, M., van Dijk, R. M., Deeg, C. A., Kistler, K., Walker, A., von Rüden, E.-L., Russmann, V., Hauck, S. M., & Potschka, H. (2018). Proteomic profiling of epileptogenesis in a rat model: Focus on cell stress, extracellular matrix and angiogenesis. *Neurobiology of Disease, 112*, 119-135. https://doi.org/https://doi.org/10.1016/j.nbd.2018.01.013

Kilany, A., Raouf, E. R. A., Gaber, A. A., Aloush, T. K., Aref, H. A., Anwar, M., Henshall, D. C., & Abdulghani, M. O. (2012). Elevated serum Bcl-2 in children with temporal lobe epilepsy. *Seizure, 21*(4), 250-253.

Kilstrup-Nielsen, C., Rusconi, L., La Montanara, P., Ciceri, D., Bergo, A., Bedogni, F., & Landsberger, N. (2012). What we know and would like to know about CDKL5 and its involvement in epileptic encephalopathy. *Neural plasticity, 2012*.

Kjeldsen, M. J., Kyvik, K. O., Christensen, K., & Friis, M. L. (2001). Genetic and environmental factors in epilepsy: a population-based study of 11 900 Danish twin pairs. *Epilepsy research, 44*(2-3), 167-178.

Korasick, D. A., & Tanner, J. J. (2021). Impact of missense mutations in the ALDH7A1 gene on enzyme structure and catalytic function. *Biochimie*, *183*, 49-54. https://doi.org/https://doi.org/10.1016/j.biochi.2020.09.016

Krenn, M., Schloegl, M., Pataraia, E., Gelpi, E., Schröder, S., Rauscher, C., Mayr, J. A., Kotzot, D., Zimprich, F., & Meitinger, T. (2021). Delineation of epileptic and neurodevelopmental phenotypes associated with variants in STX1B. *Seizure*, *87*, 25-29.

Kruglyak, L., Daly, M. J., Reeve-Daly, M. P., & Lander, E. S. (1996). Parametric and nonparametric linkage analysis: a unified multipoint approach. *American journal of human genetics*, *58*(6), 1347.

Kuruba, R., Hattiangady, B., & Shetty, A. K. (2009). Hippocampal neurogenesis and neural stem cells in temporal lobe epilepsy. *Epilepsy & Behavior*, *14*(1, Supplement 1), 65-73. https://doi.org/https://doi.org/10.1016/j.yebeh.2008.08.020

Kuß, A. (2014). Next Generation Sequencing in der Humangenetik.

Lalioti, M. D., Scott, H. S., Buresi, C., Rossier, C., Bottani, A., Morris, M. A., Malafosse, A., & Antonarakis, S. E. (1997). Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. *Nature*, *386*(6627), 847-851. https://doi.org/10.1038/386847a0

Lathrop, G. M., Lalouel, J. M., Julier, C., & Ott, J. (1984). Strategies for multilocus linkage analysis in humans. *Proceedings of the National Academy of Sciences*, *81*(11), 3443-3446. https://doi.org/doi:10.1073/pnas.81.11.3443

Leach, N. T., Sun, Y., Michaud, S., Zheng, Y., Ligon, K. L., Ligon, A. H., Sander, T., Korf, B. R., Lu, W., & Harris, D. J. (2007). Disruption of diacylglycerol kinase delta (DGKD) associated with seizures in humans and mice. *The American Journal of Human Genetics*, *80*(4), 792-799.

Lee, H., Lin, M. C., Kornblum, H. I., Papazian, D. M., & Nelson, S. F. (2014). Exome sequencing identifies de novo gain of function missense mutation in KCND2 in identical twins with autism and seizures that slows potassium channel inactivation. *Hum Mol Genet*, *23*(13), 3481-3489. https://doi.org/10.1093/hmg/ddu056

Lee, I.-C., Chang, T.-M., Liang, J.-S., & Li, S.-Y. (2019). KCNQ2 mutations in childhood nonlesional epilepsy: Variable phenotypes and a novel mutation in a case series. *Molecular genetics & genomic medicine*, *7*(7), e00816-e00816. https://doi.org/10.1002/mgg3.816

Lemke, J. R. (2019). Diagnostik genetisch bedingter Epilepsien. *medizinische genetik*, *31*(3), 303-312. https://doi.org/10.1007/s11825-019-00255-6

Li, H., & Homer, N. (2010). A survey of sequence alignment algorithms for next-generation sequencing. *Briefings in bioinformatics*, *11*(5), 473-483.

Li, M. H., Abrudan, J. L., Dulik, M. C., Sasson, A., Brunton, J., Jayaraman, V., Dugan, N., Haley, D., Rajagopalan, R., Biswas, S., Sarmady, M., DeChene, E. T., Deardorff, M. A., Wilkens, A., Noon, S. E., Scarano, M. I., Santani, A. B., White, P. S., Pennington,

J., Conlin, L. K., Spinner, N. B., Krantz, I. D., & Vetter, V. L. (2015). Utility and limitations of exome sequencing as a genetic diagnostic tool for conditions associated with pediatric sudden cardiac arrest/sudden cardiac death. *Human Genomics*, *9*(1), 15. https://doi.org/10.1186/s40246-015-0038-y

Li, W., Notani, D., & Rosenfeld, M. G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nature Reviews Genetics*, *17*(4), 207-223.

Lindgren, Å., Kihlgren, M., Melin, L., Croona, C., Lundberg, S., & Eeg-Olofsson, O. (2004). Development of cognitive functions in children with rolandic epilepsy. *Epilepsy & Behavior*, *5*(6), 903-910. https://doi.org/https://doi.org/10.1016/j.yebeh.2004.08.004

Ling, H., Vincent, K., Pichler, M., Fodde, R., Berindan-Neagoe, I., Slack, F. J., & Calin, G. A. (2015). Junk DNA and the long non-coding RNA twist in cancer genetics. *Oncogene*, *34*(39), 5003-5011.

Liu, X., Ou, S., Xu, T., Liu, S., Yuan, J., Huang, H., Qin, L., Yang, H., Chen, L., Tan, X., & Chen, Y. (2016). New differentially expressed genes and differential DNA methylation underlying refractory epilepsy. *Oncotarget*, *7*(52), 87402-87416. https://doi.org/10.18632/oncotarget.13642

Liu, Y.-C., Lee, J. W. A., Bellows, S. T., Damiano, J. A., Mullen, S. A., Berkovic, S. F., Bahlo, M., Scheffer, I. E., Hildebrand, M. S., & Group, C. (2016). Evaluation of non-coding variation in GLUT1 deficiency. *Developmental Medicine & Child Neurology*, *58*(12), 1295-1302. https://doi.org/https://doi.org/10.1111/dmcn.13163

Lynce, I., & Marques-Silva, J. (2006). *Efficient haplotype inference with boolean satisfiability* Proceedings of the 21st national conference on Artificial intelligence - Volume 1, Boston, Massachusetts.

Malter, M. P., Helmstaedter, C., Urbach, H., Vincent, A., & Bien, C. G. (2010). Antibodies to glutamic acid decarboxylase define a form of limbic encephalitis. *Annals of Neurology*, *67*(4), 470-478. https://doi.org/https://doi.org/10.1002/ana.21917

Manolio, T. A. (2010). Genomewide association studies and assessment of the risk of disease. *New England journal of medicine*, *363*(2), 166-176.

Maroofian, R., Gubas, A., Kaiyrzhanov, R., Scala, M., Hundallah, K., Severino, M., Abdel-Hamid, M. S., Rosenfeld, J. A., Ebrahimi-Fakhari, D., Ali, Z., Rahim, F., Houlden, H., Tooze, S. A., Alsaleh, N. S., & Zaki, M. S. (2021). Homozygous missense WIPI2 variants cause a congenital disorder of autophagy with neurodevelopmental impairments of variable clinical severity and disease course. *Brain Commun*, *3*(3), fcab183. https://doi.org/10.1093/braincomms/fcab183

McCarthy, D. J., Humburg, P., Kanapin, A., Rivas, M. A., Gaulton, K., Cazier, J.-B., & Donnelly, P. (2014). Choice of transcripts and software has a large effect on variant annotation. *Genome medicine*, *6*(3), 1-16.

McCaw, Z. R., O'Dushlaine, C., Somineni, H., Bereket, M., Klein, C., Karaletsos, T., Casale, F. P., Koller, D., & Soare, T. W. (2023). An allelic-series rare-variant association test

for candidate-gene discovery. *Am J Hum Genet*, *110*(8), 1330-1342. https://doi.org/10.1016/j.ajhg.2023.07.001

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., & Cunningham, F. (2016). The ensembl variant effect predictor. *Genome biology*, *17*(1), 1-14.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., & Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, *26*(16), 2069-2070.

Mertes, F., ElSharawy, A., Sauer, S., van Helvoort, J. M. L. M., van der Zaag, P. J., Franke, A., Nilsson, M., Lehrach, H., & Brookes, A. J. (2011). Targeted enrichment of genomic DNA regions for next-generation sequencing. *Briefings in Functional Genomics*, *10*(6), 374-386. https://doi.org/10.1093/bfgp/elr033

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature Reviews Genetics*, *11*(1), 31-46.

Mills, P. B., Footitt, E. J., Mills, K. A., Tuschl, K., Aylett, S., Varadkar, S., Hemingway, C., Marlow, N., Rennie, J., & Baxter, P. (2010). Genotypic and phenotypic spectrum of pyridoxine-dependent epilepsy (ALDH7A1 deficiency). *Brain*, *133*(7), 2148-2159.

Mizuguchi, T., Nakashima, M., Kato, M., Okamoto, N., Kurahashi, H., Ekhilevitch, N., Shiina, M., Nishimura, G., Shibata, T., & Matsuo, M. (2018). Loss-of-function and gain-of-function mutations in PPP3CA cause two distinct disorders. *Human molecular genetics*, *27*(8), 1421-1433.

Morton, N. E. (1955). Sequential tests for the detection of linkage. *American journal of human genetics*, *7*(3), 277.

Muñoz, L. J., Carballosa-Gautam, M. M., Yanowsky, K., García-Atarés, N., & López, D. E. (2017). The genetic audiogenic seizure hamster from Salamanca: The GASH:Sal. *Epilepsy & Behavior*, *71*, 181-192. https://doi.org/https://doi.org/10.1016/j.yebeh.2016.03.002

Musich, R., Cadle-Davidson, L., & Osier, M. V. (2021). Comparison of Short-Read Sequence Aligners Indicates Strengths and Weaknesses for Biologists to Consider. *Front Plant Sci*, *12*, 657240. https://doi.org/10.3389/fpls.2021.657240

Myers, C. T., & Mefford, H. C. (2015). Advancing epilepsy genetics in the genomic era. *Genome Med*, *7*, 91. https://doi.org/10.1186/s13073-015-0214-7

Myers, R., Leavitt, J., Farrer, L., Jagadeesh, J., McFarlane, H., Mastromauro, C., Mark, R., & Gusella, J. (1989). Homozygote for Huntington disease. *American journal of human genetics*, *45*(4), 615.

Nakayama, T., Ogiwara, I., Ito, K., Kaneda, M., Mazaki, E., Osaka, H., Ohtani, H., Inoue, Y., Fujiwara, T., & Uematsu, M. (2010). Deletions of SCN1A 5′ genomic region with promoter activity in Dravet syndrome. *Human mutation*, *31*(7), 820-829.

Neigenfind, J., Gyetvai, G., Basekow, R., Diehl, S., Achenbach, U., Gebhardt, C., Selbig, J., & Kersten, B. (2008). Haplotype inference from unphased SNP data in heterozygous polyploids based on SAT. *BMC Genomics*, *9*(1), 356. https://doi.org/10.1186/1471-2164-9-356

Neligan, A., Hauser, W. A., & Sander, J. W. (2012). Chapter 6 - The epidemiology of the epilepsies. In H. Stefan & W. H. Theodore (Eds.), *Handbook of Clinical Neurology* (Vol. 107, pp. 113-133). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-52898-8.00006-9

Nellist, M., van den Heuvel, D., Schluep, D., Exalto, C., Goedbloed, M., Maat-Kievit, A., van Essen, T., van Spaendonck-Zwarts, K., Jansen, F., Helderman, P., Bartalini, G., Vierimaa, O., Penttinen, M., van den Ende, J., van den Ouweland, A., & Halley, D. (2009). Missense mutations to the TSC1 gene cause tuberous sclerosis complex. *Eur J Hum Genet*, *17*(3), 319-328. https://doi.org/10.1038/ejhg.2008.170

Neubauer, B. A., & Hahn, A. (2011). Genetik der Epilepsien. *Monatsschrift Kinderheilkunde*, *159*(8), 714-720. https://doi.org/10.1007/s00112-011-2395-5

Noeker, M., Haverkamp-Krois, A., & Haverkamp, F. (2005). Development of mental health dysfunction in childhood epilepsy. *Brain and Development*, *27*(1), 5-16.

Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., Krabichler, B., Speicher, M. R., Zschocke, J., & Trajanoski, Z. (2014). A survey of tools for variant analysis of next-generation genome sequencing data. *Briefings in bioinformatics*, *15*(2), 256-278.

Pagni, S., Mills, J. D., Frankish, A., Mudge, J. M., & Sisodiya, S. M. (2022). Non-coding regulatory elements: Potential roles in disease and the case of epilepsy. *Neuropathology and Applied Neurobiology*, *48*(3), e12775.

Palace, J., & Lang, B. (2000). Epilepsy: an autoimmune disease? In (Vol. 69, pp. 711-714): BMJ Publishing Group Ltd.

Pan, T.-C., Zhang, R.-Z., Markova, D., Arita, M., Zhang, Y., Bogdanovich, S., Khurana, T. S., Bönnemann, C. G., Birk, D. E., & Chu, M.-L. (2013). COL6A3 protein deficiency in mice leads to muscle and tendon defects similar to human collagen VI congenital muscular dystrophy. *Journal of Biological Chemistry*, *288*(20), 14320-14331.

Paprocka, J., Kijonka, M., Pęcka, M., & Sokół, M. (2016). Melatonin in Epilepsy: A New Mathematical Model of Diurnal Secretion. *Int J Endocrinol*, *2016*, 3861461. https://doi.org/10.1155/2016/3861461

Parla, J. S., Iossifov, I., Grabill, I., Spector, M. S., Kramer, M., & McCombie, W. R. (2011). A comparative analysis of exome capture. *Genome Biol*, *12*(9), R97. https://doi.org/10.1186/gb-2011-12-9-r97

Pearson, M. A., Reczek, D., Bretscher, A., & Karplus, P. A. (2000). Structure of the ERM Protein Moesin Reveals the FERM Domain Fold Masked by an Extended Actin Binding Tail Domain. *Cell*, *101*, 259-270.

Peljto, A. L., Barker-Cummings, C., Vasoli, V. M., Leibson, C. L., Hauser, W. A., Buchhalter, J. R., & Ottman, R. (2014). Familial risk of epilepsy: a population-based study. *Brain*, *137*(Pt 3), 795-805. https://doi.org/10.1093/brain/awt368

Pinheiro, L. H. M., Guimarães, L. S., Antunes, L. S., Küchler, E. C., Kirschneck, C., & Antunes, L. A. A. (2021). Genetic variation involved in the risk to external apical root resorption in orthodontic patients: A systematic review. *Clinical Oral Investigations*, *25*(10), 5613-5627.

Pires, G., Leitner, D., Drummond, E., Kanshin, E., Nayak, S., Askenazi, M., Faustin, A., Friedman, D., Debure, L., Ueberheide, B., Wisniewski, T., & Devinsky, O. (2021). Proteomic differences in the hippocampus and cortex of epilepsy brain tissue. *Brain Communications*, *3*(2). https://doi.org/10.1093/braincomms/fcab021

Piro, R. M. (2020). Chapter 7 - Sequencing technologies for epigenetics: From basics to applications. In D. Kabelitz & J. Bhat (Eds.), *Epigenetics of the Immune System* (Vol. 16, pp. 161-183). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-817964-2.00007-1

Poduri, A., & Lowenstein, D. (2011). Epilepsy genetics--past, present, and future. *Current opinion in genetics & development*, *21*(3), 325-332. https://doi.org/10.1016/j.gde.2011.01.005

Poke, G., King, C., Muir, A., de Valles-Ibáñez, G., Germano, M., Moura de Souza, C. F., Fung, J., Chung, B., Fung, C. W., & Mignot, C. (2019). The epileptology of GNB5 encephalopathy. *Epilepsia*, *60*(11), e121-e127.

Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R., & Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome research*, *20*(1), 110-121.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, *81*(3), 559-575. https://doi.org/10.1086/519795

Read, A. P., & Strachan, T. (1996). *Molekulare Humangenetik*. Spektrum Akademischer Verlag.

Rehm, H. L., Bale, S. J., Bayrak-Toydemir, P., Berg, J. S., Brown, K. K., Deignan, J. L., Friez, M. J., Funke, B. H., Hegde, M. R., & Lyon, E. (2013). ACMG clinical laboratory standards for next-generation sequencing. *Genetics in Medicine*, *15*(9), 733-747.

Rehman, A. U., Santos-Cortez, R. L. P., Drummond, M. C., Shahzad, M., Lee, K., Morell, R. J., Ansar, M., Jan, A., Wang, X., & Aziz, A. (2015). Challenges and solutions for gene identification in the presence of familial locus heterogeneity. *European Journal of Human Genetics*, *23*(9), 1207-1215.

Rehman, A. U., Santos-Cortez, R. L. P., Drummond, M. C., Shahzad, M., Lee, K., Morell, R. J., Ansar, M., Jan, A., Wang, X., Aziz, A., Riazuddin, S., Smith, J. D., Wang, G. T., Ahmed, Z. M., Gul, K., Shearer, A. E., Smith, R. J. H., Shendure, J., Bamshad, M. J.,

Nickerson, D. A., Hinnant, J., Khan, S. N., Fisher, R. A., Ahmad, W., Friderici, K. H., Riazuddin, S., Friedman, T. B., Wilch, E. S., Leal, S. M., & University of Washington Center for Mendelian, G. (2015). Challenges and solutions for gene identification in the presence of familial locus heterogeneity. *European Journal of Human Genetics*, *23*(9), 1207-1215. https://doi.org/10.1038/ejhg.2014.266

Rhoads, A., & Au, K. F. (2015). PacBio sequencing and its applications. *Genomics, proteomics & bioinformatics*, *13*(5), 278-289.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., & Spector, E. (2015). 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. *Genetics in Medicine*, *17*(5), 405-423.

Rieger, R., Michaelis, A., & Green, M. M. (1976). *Glossary of genetics and cytogenetics: classical and molecular*. Springer.

Rogawski, M. A. (2000). KCNQ2/KCNQ3 K+ channels and the molecular pathogenesis of epilepsy: implications for therapy. *Trends in neurosciences*, *23*(9), 393-398.

Routbort, M. J., Bausch, S. B., & McNamara, J. O. (1999). Seizures, cell death, and mossy fiber sprouting in kainic acid-treated organotypic hippocampal cultures*M. J. Routbort and S. B. Bausch contributed equally to the work described in this paper.*. *Neuroscience*, *94*(3), 755-765. https://doi.org/https://doi.org/10.1016/S0306-4522(99)00358-9

Rydzanicz, M., Wachowska, M., Cook, E. C., Lisowski, P., Kuźniewska, B., Szymańska, K., Diecke, S., Prigione, A., Szczałuba, K., & Szybińska, A. (2019). Novel calcineurin A (PPP3CA) variant associated with epilepsy, constitutive enzyme activation and downregulation of protein expression. *European Journal of Human Genetics*, *27*(1), 61-69.

Sadleir, L., Farrell, K., Smith, S., Connolly, M., & Scheffer, I. (2006). Electroclinical features of absence seizures in childhood absence epilepsy. *Neurology*, *67*(3), 413-418.

Sakane, F., Imai, S., Yamada, K., Murakami, T., Tsushima, S., & Kanoh, H. (2002). Alternative splicing of the human diacylglycerol kinase delta gene generates two isoforms differing in their expression patterns and in regulatory functions. *J Biol Chem*, *277*(45), 43519-43526. https://doi.org/10.1074/jbc.M206895200

Samocha, K. E., Kosmicki, J. A., Karczewski, K. J., O'Donnell-Luria, A. H., Pierce-Hoffman, E., MacArthur, D. G., Neale, B. M., & Daly, M. J. (2017). Regional missense constraint improves variant deleteriousness prediction. *bioRxiv*, 148353.

Sarbassov, D. D., Ali, S. M., & Sabatini, D. M. (2005). Growing roles for the mTOR pathway. *Current opinion in cell biology*, *17*(6), 596-603.

Sato, M., Liu, K., Sasaki, S., Kunii, N., Sakai, H., Mizuno, H., Saga, H., & Sakane, F. (2013). Evaluations of the selectivities of the diacylglycerol kinase inhibitors R59022 and R59949 among diacylglycerol kinase isozymes using a new non-radioactive assay method. *Pharmacology*, *92*(1-2), 99-107. https://doi.org/10.1159/000351849

Scheffer, I. E., Berkovic, S., Capovilla, G., Connolly, M. B., French, J., Guilhoto, L., Hirsch, E., Jain, S., Mathern, G. W., Moshé, S. L., Nordli, D. R., Perucca, E., Tomson, T., Wiebe, S., Zhang, Y.-H., & Zuberi, S. M. (2017). ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*, *58*(4), 512-521. https://doi.org/10.1111/epi.13709

Scheffer, I. E., & Liao, J. (2020). Deciphering the concepts behind "Epileptic encephalopathy" and "Developmental and epileptic encephalopathy". *European Journal of Paediatric Neurology*, *24*, 11-14. https://doi.org/https://doi.org/10.1016/j.ejpn.2019.12.023

Schimmer, A. D., Munk-Pedersen, I., Minden, M. D., & Reed, J. C. (2003). Bcl-2 and apoptosis in chronic lymphocytic leukemia. *Current treatment options in oncology*, *4*(3), 211-218.

Schober, D. (2002). Microarrays, Genexpressionsanalyse und Bioinformatik. *BIOspektrum*, *8*(3), 307-310.

Semple, B. D., O'Brien, T. J., Gimlin, K., Wright, D. K., Kim, S. E., Casillas-Espinosa, P. M., Webster, K. M., Petrou, S., & Noble-Haeusslein, L. J. (2017). Interleukin-1 Receptor in Seizure Susceptibility after Traumatic Injury to the Pediatric Brain. *J Neurosci*, *37*(33), 7864-7877. https://doi.org/10.1523/jneurosci.0982-17.2017

Sergouniotis, P. I., Davidson, A. E., Mackay, D. S., Li, Z., Yang, X., Plagnol, V., Moore, A. T., & Webster, A. R. (2011). Recessive mutations in KCNJ13, encoding an inwardly rectifying potassium channel subunit, cause leber congenital amaurosis. *The American Journal of Human Genetics*, *89*(1), 183-190.

Serpico, D. (2020). Beyond quantitative and qualitative traits: three telling cases in the life sciences. *Biology & Philosophy*, *35*(3), 1-26.

Sillanpää, M., & Cross, J. H. (2009). The psychosocial impact of epilepsy in childhood. *Epilepsy & Behavior*, *15*(2), S5-S10.

Singh, B., Ogiwara, I., Kaneda, M., Tokonami, N., Mazaki, E., Baba, K., Matsuda, K., Inoue, Y., & Yamakawa, K. (2006). A Kv4. 2 truncation mutation in a patient with temporal lobe epilepsy. *Neurobiology of Disease*, *24*(2), 245-253.

Sisodiya, S. M., & Mefford, H. C. (2011). Genetic contribution to common epilepsies. *Curr Opin Neurol*, *24*(2), 140-145. https://doi.org/10.1097/WCO.0b013e328344062f

Sloan, J. L., Johnston, J. J., Manoli, I., Chandler, R. J., Krause, C., Carrillo-Carrasco, N., Chandrasekaran, S. D., Sysol, J. R., O'Brien, K., & Hauser, N. S. (2011). Exome sequencing identifies ACSF3 as a cause of combined malonic and methylmalonic aciduria. *Nature Genetics*, *43*(9), 883-886.

Smets, K., Duarri, A., Deconinck, T., Ceulemans, B., van de Warrenburg, B. P., Züchner, S., Gonzalez, M. A., Schüle, R., Synofzik, M., & Van der Aa, N. (2015). First de novo KCND3 mutation causes severe Kv4. 3 channel dysfunction leading to early onset cerebellar ataxia, intellectual disability, oral apraxia and epilepsy. *BMC medical genetics*, *16*(1), 1-7.

Soh, H., Pant, R., LoTurco, J. J., & Tzingounis, A. V. (2014). Conditional deletions of epilepsy-associated KCNQ2 and KCNQ3 channels from cerebral cortex cause differential effects on neuronal excitability. *Journal of Neuroscience, 34*(15), 5311-5321.

Specchio, N., Fusco, L., Claps, D., & Vigevano, F. (2010). Epileptic encephalopathy in children possibly related to immune-mediated pathogenesis. *Brain and Development, 32*(1), 51-56. https://doi.org/https://doi.org/10.1016/j.braindev.2009.09.017

Speed, D., O'Brien, T. J., Palotie, A., Shkura, K., Marson, A. G., Balding, D. J., & Johnson, M. R. (2014). Describing the genetic architecture of epilepsy through heritability analysis. *Brain, 137*(Pt 10), 2680-2689. https://doi.org/10.1093/brain/awu206

Stafstrom, C. E., & Carmant, L. (2015). Seizures and epilepsy: an overview for neuroscientists. *Cold Spring Harbor perspectives in medicine, 5*(6), a022426. https://doi.org/10.1101/cshperspect.a022426

Steinlein, O. K. (2004). Genes and mutations in human idiopathic epilepsy. *Brain and Development, 26*(4), 213-218. https://doi.org/https://doi.org/10.1016/S0387-7604(03)00149-9

Steinlein, O. K. (2010). Genetik der Epilepsien. *BIOspektrum, 16*, 16-18.

Strachan, T., & Read, A. (2010a). *Chromosome Structure and Function*

Strachan, T., & Read, A. (2010b). *Genes in Pedigrees and Populations*.

Strachan, T., & Read, A. (2010c). *Genetic Mapping of Mendelian Characters*.

Strachan, T., & Read, A. (2010d). *Human Genetic Variability and Its Consequences*

Strachan, T., & Read, A. (2010e). *Mapping genes conferring susceptibility of complex diseases*.

Syvertsen, M., Nakken, K. O., Edland, A., Hansen, G., Hellum, M. K., & Koht, J. (2015). Prevalence and etiology of epilepsy in a Norwegian county—A population based study. *Epilepsia, 56*(5), 699-706. https://doi.org/https://doi.org/10.1111/epi.12972

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics, 20*(8), 467-484. https://doi.org/10.1038/s41576-019-0127-1

Tanaka, K., Watase, K., Manabe, T., Yamada, K., Watanabe, M., Takahashi, K., Iwama, H., Nishikawa, T., Ichihara, N., Kikuchi, T., Okuyama, S., Kawashima, N., Hori, S., Takimoto, M., & Wada, K. (1997). Epilepsy and Exacerbation of Brain Injury in Mice Lacking the Glutamate Transporter GLT-1. *Science, 276*(5319), 1699-1702. https://doi.org/10.1126/science.276.5319.1699

Tarazona, S., Balzano-Nogueira, L., & Conesa, A. (2018). Chapter Eighteen - Multiomics Data Integration in Time Series Experiments. In J. Jaumot, C. Bedia, & R. Tauler (Eds.), *Comprehensive Analytical Chemistry* (Vol. 82, pp. 505-532). Elsevier. https://doi.org/https://doi.org/10.1016/bs.coac.2018.06.005

Teare, M. D., & Barrett, J. H. (2005). Genetic linkage studies. *The Lancet*, *366*(9490), 1036-1044.

Tenney, J. R., & Glauser, T. A. (2013). The current state of absence epilepsy: can we have your attention? the current state of absence epilepsy. *Epilepsy currents*, *13*(3), 135-140.

Thomas, Q., Motta, M., Gautier, T., Zaki, M. S., Ciolfi, A., Paccaud, J., Girodon, F., Boespflug-Tanguy, O., Besnard, T., & Kerkhof, J. (2022). Bi-allelic loss-of-function variants in TMEM147 cause moderate to profound intellectual disability with facial dysmorphism and pseudo-Pelger-Huët anomaly. *The American Journal of Human Genetics*, *109*(10), 1909-1922.

Turner, E. H., Ng, S. B., Nickerson, D. A., & Shendure, J. (2009). Methods for Genomic Partitioning. *Annual Review of Genomics and Human Genetics*, *10*(1), 263-284. https://doi.org/10.1146/annurev-genom-082908-150112

Urak, L., Feucht, M., Fathi, N., Hornik, K., & Fuchs, K. (2006). A GABRB3 promoter haplotype associated with childhood absence epilepsy impairs transcriptional activity. *Human molecular genetics*, *15*(16), 2533-2541.

Usui, N., Terada, K., Baba, K., Matsuda, K., Nakamura, F., Usui, K., Tottori, T., Umeoka, S., Fujitani, S., Mihara, T., & Inoue, Y. (2010). Very high frequency oscillations (over 1000Hz) in human epilepsy. *Clinical Neurophysiology*, *121*(11), 1825-1831. https://doi.org/https://doi.org/10.1016/j.clinph.2010.04.018

Vaisse, C., Clement, K., Durand, E., Hercberg, S., Guy-Grand, B., & Froguel, P. (2000). Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *The Journal of clinical investigation*, *106*(2), 253-262.

Van Damme, J., De Ley, M., Opdenakker, G., Billiau, A., De Somer, P., & Van Beeumen, J. (1985). Homogeneous interferon-inducing 22K factor is related to endogenous pyrogen and interleukin-1. *Nature*, *314*(6008), 266-268. https://doi.org/10.1038/314266a0

van Loo, K. M. J., & Becker, A. J. (2019). Transcriptional Regulation of Channelopathies in Genetic and Acquired Epilepsies. *Front Cell Neurosci*, *13*, 587. https://doi.org/10.3389/fncel.2019.00587

van Luijtelaar, G., & Sitnikova, E. (2006). Global and focal aspects of absence epilepsy: the contribution of genetic models. *Neuroscience & Biobehavioral Reviews*, *30*(7), 983-1003.

Vezzani, A. (2014). Epilepsy and Inflammation in the Brain: Overview and Pathophysiology: Epilepsy and Inflammation in the Brain. *Epilepsy currents*, *14*(2_suppl), 3-7.

Vezzani, A., Fujinami, R. S., White, H. S., Preux, P.-M., Blümcke, I., Sander, J. W., & Löscher, W. (2016). Infections, inflammation and epilepsy. *Acta Neuropathologica*, *131*(2), 211-234. https://doi.org/10.1007/s00401-015-1481-5

Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. d. F. (2016). Microsatellite markers: what they mean and why they are so useful. *Genetics and molecular biology*, *39*(3), 312-328. https://doi.org/10.1590/1678-4685-GMB-2016-0027

Villa, C., Lavitrano, M., & Combi, R. (2019). Long Non-Coding RNAs and Related Molecular Pathways in the Pathogenesis of Epilepsy. *International Journal of Molecular Sciences*, *20*(19), 4898. https://www.mdpi.com/1422-0067/20/19/4898

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era — concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255-266. https://doi.org/10.1038/nrg2322

Vora, G. J., Meador, C. E., Stenger, D. A., & Andreadis, J. D. (2004). Nucleic acid amplification strategies for DNA microarray-based pathogen detection. *Applied and Environmental Microbiology*, *70*(5), 3047-3054.

Wagner, M. (2018). *Mapping the chromosomal locus of oculopharyngodistal myopathy with microsatellite markers and next generation sequencing* lmu].

Wagner, M., Berutti, R., Lorenz-Depiereux, B., Graf, E., Eckstein, G., Mayr, J. A., Meitinger, T., Ahting, U., Prokisch, H., Strom, T. M., & Wortmann, S. B. (2019). Mitochondrial DNA mutation analysis from exome sequencing—A more holistic approach in diagnostics of suspected mitochondrial disease. *Journal of Inherited Metabolic Disease*, *42*(5), 909-917. https://doi.org/https://doi.org/10.1002/jimd.12109

Wallace, R. H., Wang, D. W., Singh, R., Scheffer, I. E., George, A. L., Phillips, H. A., Saar, K., Reis, A., Johnson, E. W., Sutherland, G. R., Berkovic, S. F., & Mulley, J. C. (1998). Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel ß1 subunit gene SCN1B. *Nature Genetics*, *19*(4), 366-370. https://doi.org/10.1038/1252

Wang, J., Lin, Z.-J., Liu, L., Xu, H.-Q., Shi, Y.-W., Yi, Y.-H., He, N., & Liao, W.-P. (2017). Epilepsy-associated genes. *Seizure - European Journal of Epilepsy*, *44*, 11-20. https://doi.org/10.1016/j.seizure.2016.11.030

Wang, S., Huang, Y., Zhou, C., Wu, H., Zhao, J., Wu, L., Zhao, M., Zhang, F., & Liu, H. (2018). The role of autophagy and related microRNAs in inflammatory bowel disease. *Gastroenterology Research and Practice*, *2018*.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, *10*(1), 57-63. https://doi.org/10.1038/nrg2484

Weaving, L. S., Christodoulou, J., Williamson, S. L., Friend, K. L., McKenzie, O. L., Archer, H., Evans, J., Clarke, A., Pelka, G. J., & Tam, P. P. (2004). Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *The American Journal of Human Genetics*, *75*(6), 1079-1093.

Weber, Y., & Lerche, H. (2008). *Genetic mechanisms in idiopathic epilepsies.* .

Welch, M. D., Iwamatsu, A., & Mitchison, T. J. (1997). Actin polymerization is induced by Arp2/3 protein complex at the surface of Listeria monocytogenes. *Nature*, *385*(6613), 265-269. https://doi.org/10.1038/385265a0

Wigginton, J. E., & Abecasis, G. R. (2005). PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics*, *21*(16), 3445-3447. https://doi.org/10.1093/bioinformatics/bti529

Williamson, R., Hanif, S., Mathews, G. C., Lagrange, A. H., & Abou-Khalil, B. (2009). Generalized-onset seizures with secondary focal evolution. *Epilepsia*, *50*(7), 1827-1832.

Wiltshire, S., Hattersley, A. T., Hitman, G. A., Walker, M., Levy, J. C., Sampson, M., O'Rahilly, S., Frayling, T. M., Bell, J. I., & Lathrop, G. M. (2001). A genomewide scan for loci predisposing to type 2 diabetes in a UK population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *The American Journal of Human Genetics*, *69*(3), 553-569.

Winden, K. D., Karsten, S. L., Bragin, A., Kudo, L. C., Gehman, L., Ruidera, J., Geschwind, D. H., & Engel Jr, J. (2011). A systems level, functional genomics analysis of chronic epilepsy. *PLoS One*, *6*(6), e20763.

Wong, M. (2010). Mammalian target of rapamycin (mTOR) inhibition as a potential antiepileptogenic therapy: From tuberous sclerosis to common acquired epilepsies. *Epilepsia*, *51*(1), 27-36. https://doi.org/10.1111/j.1528-1167.2009.02341.x

World Health Organization, Global Campaign against Epilepsy, Programme for Neurological Diseases, Neuroscience, International Bureau for Epilepsy, World Health Organization. Department of Mental Health, Substance Abuse, International Bureau of Epilepsy, & International League against Epilepsy. (2005). *Atlas: epilepsy care in the world*. World Health Organization.

Wu, X., & Reddy, D. S. (2012). Integrins as receptor targets for neurological disorders. *Pharmacology & therapeutics*, *134*(1), 68-81.

Yumoto, T., Nakadate, K., Nakamura, Y., Sugitani, Y., Sugitani-Yoshida, R., Ueda, S., & Sakakibara, S.-i. (2013). Radmis, a novel mitotic spindle protein that functions in cell division of neural progenitors. *PLoS One*, *8*(11), e79895.

Yumoto, T., Nakadate, K., Nakamura, Y., Sugitani, Y., Sugitani-Yoshida, R., Ueda, S., & Sakakibara, S. (2013). Radmis, a novel mitotic spindle protein that functions in cell division of neural progenitors. *PLoS One*, *8*(11), e79895. https://doi.org/10.1371/journal.pone.0079895

Zimprich, F. (2009). Grundlagen und Zahlen zur genetischen Beratung bei komplexen Epilepsien. *Journal für Neurologie, Neurochirurgie und Psychiatrie*, *10*.

Zimprich, F., Baumgartner, C., Dobesberger, J., von Oertzen, J., Feichtinger, M., Rauscher, C., Pataraia, E., Luef, G., Unterberger, I., & Schlachter, K. (2015). Österreichische Gesellschaft für Epileptologie. *Mitteilungen der Österreichischen Gesellschaft für Epileptologie*, *2*(1), 2-7.

Zupanc, M. L., & Legros, B. (2004). Progressive myoclonic epilepsy. *The Cerebellum*, *3*(3), 156. https://doi.org/10.1080/14734220410035356

8 Anhang

8.1 Merlin input files

M	rs811047
M	rs4788477
M	rs9937452
M	rs4788693
M	rs11075971
M	rs1423732
M	rs12149821
M	rs8051981
M	rs705894
M	rs12920607
M	rs737672
M	rs2663152
M	rs2526050
M	rs7187327
M	rs11640870
M	rs4888327
M	rs36097424
M	rs100912
M	rs7202210
M	rs7185108
M	rs8058474
M	rs4485401
M	rs28546976
M	rs7500451
M	rs3843959
M	rs2454935
M	GSA-rs7193915
M	GSA-rs10163466
M	rs4887918
M	GSA-rs7189359
M	rs10781989
M	rs1947771
M	rs8058686
M	rs6564499
M	rs2287972
M	rs3764293
M	rs12448571
M	rs2548861
M	rs2656635
M	rs10492910
M	rs12923991
M	rs1110891
M	rs1875940
M	rs17643319
M	rs11643445
M	rs17648647
M	rs4888925
M	rs11150140
M	rs270434
M	rs4581714
M	rs4888967
M	rs2720445

Table 13.: Exemplary extract from the DAT-File. This is an exemplary extract from the .dat Merlin input file used. The "M" in the left column indicates that the respective SNPs are markers that should be included into the program. In the right column, the SNPs, that have been genotyped, are listed.

18	rs1445088	74,09923
18	rs1445093	74,25858
18	rs8093706	74,54618
18	rs1144049	74,70353
18	rs9950125	74,80829
18	rs12457160	75,09109
18	rs9954439	75,29861
18	rs1553108	75,45439
18	rs34163044	75,68093
18	rs28377497	75,8719
18	rs949031	76,09814
18	rs2872041	76,32831
18	exm2268153	76,38
18	rs1506632	76,71249
18	rs2588482	76,91538
18	rs4801018	77,12669
18	GSA-rs10513889	77,22047
18	GSA-rs12454365	77,50038
18	rs4801076	77,6064
18	rs12150819	77,88623
18	rs8093842	78,51064
18	rs319439	78,71609
18	rs11877445	78,83174
18	rs4941159	79,01905
18	rs4528665	79,28938
18	rs1893480	79,4786
18	rs12051934	79,71942
18	rs4940694	79,87178
18	rs9947399	80,11844
18	exm2268154	80,25471
18	rs4940757	80,40197
18	rs7231686	80,682
18	GSA-rs561894	80,9306
18	rs11874251	81,42175
18	rs1789475	81,47918
18	rs2564489	81,71146
18	rs2195534	82,02397
18	rs7244658	82,03665
18	rs1791330	82,2942
18	rs66952881	82,38962
18	rs11152195	82,6989
18	rs28534746	82,71854
18	rs2331841	83,18513
18	rs9946888	83,44933
18	rs1943241	83,90685
18	rs530676	84,15061
18	GSA-rs1478526	84,48793
18	rs11152269	84,98
18	rs242622	85,0724
18	rs12953685	85,58305
18	rs757620	85,65647
18	rs2535357	85,87887

Table 14.: Exemplary extract from the MAP-File. This is an exemplary extract from the .map input file used. “18” in the left column illustrates that this is chromosome 18. The markers from the .dat file (middle row) are assigned to a specific position on the chromosome in cM (right column).

A	B	C	D	E	F	G	H
1	1	0	0	1	0	0 0	0 0
1	2	0	0	2	0	0 0	0 0
1	5	0	0	1	1	0 0	0 0
1	10	0	0	2	1	0 0	0 0
1	11	0	0	1	1	0 0	0 0
1	12	0	0	1	1	0 0	0 0
1	13	0	0	2	1	0 0	0 0
1	15	6	7	2	0	0 0	0 0
1	117731	1	2	2	2	C A	G T
1	117732	0	0	1	1	A A	G T
1	117735	117732	117731	2	2	A A	T T
1	117737	117732	117731	1	1	A A	T T
1	117739	117732	117731	2	2	A A	G T
1	117740	117732	117731	2	1	A A	G T
1	117741	117732	117731	1	2	C A	G G
1	117744	117732	117731	1	2	A A	T T
1	117745	5	117735	2	2	C A	T T
1	117748	5	117735	1	2	A A	G T
1	117749	5	117735	1	1	A A	G T
1	117752	117737	10	1	2	A A	G T

Table 15.: Exemplary extract from the PED-File. This is an exemplary sample of the .ped file. Column "A" represents the whole family #2. Since only one family has been considered, the number 1 appears here throughout. "B" is the single individual. Patients without DNA ID (identifier) were numbered starting with 1 (e.g., 1-15). The other patients received their DNA IDs as a designation (e.g. 117732, 117744,…). In the third column the fathers were noted, in the fourth column the mothers of each person. If one parent was unknown, this family member was given the number 0 in “C”/”D”. Column "E" indicates the sex (1: male; 2: female). Column "F" indicates the affection status (0: unknown, 1: not affected; 2: affected). In column "G" and “H” the SNP data are added after everyone according to the order of the SNPs in the .map file. If a member of the pedigree has not been genotyped, 0 0 is used.

Epilepsy	0.0001	0.01,0.9,1.0	Dominant_model

Table 16.: MODEL-File. Exemplary .model file used for the parametric linkage analysis in Merlin. This file consists of four columns in total: First the affection status label (matching the .dat file). Second a disease allele frequency. Third the probability of being affected for individuals with 0,1 and 2 copies of the disease allele (penetrances), and finally a label for the analysis model.

8.2 Merlin output files

117741 (MATERNAL)	117741 (PATERNAL)	117744 (MATERNAL)	117744 (PATERNAL)	117737 (MATERNAL)	117737 (PATERNAL)
A	Q	A	Q	C	Q
A	Q	A	Q	C	Q
A	Q	A	Q	C	Q
A	Q	A	Q	C	Q
A	Q	A	Q	C	Q
A	Q	A	Q	C	Q
A	Q	A	Q	C	Q
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
A	Q	A	Q	C	R
C	Q	A	Q	C	R
C	Q	A	Q	C	R
C	Q	A	Q	C	R

Table 17.: Exemplary extract from the FLOW Merlin output file. In the first row is the DNA-ID of each individual. Maternal haplotypes are listed on the left, paternal haplotypes on the right.

117741 (MATERNAL)	117741 (PATERNAL)	117744 (MATERNAL)	117744 (PATERNAL)	117737 (MATERNAL)	117737 (PATERNAL)
A	A	A	A	G	A
A	A	A	A	G	A
A	G	A	G	G	G
G	T	G	T	G	T
T	T	T	T	C	T
C	C	C	C	C	C
A	A	A	A	A	A
T	T	T	T	T	T
G	G	G	G	G	A
C	C	C	C	T	C
T	T	T	T	T	T
T	C	T	C	T	T
G	G	G	G	A	A
T	C	T	C	C	T
C	C	C	C	A	A
G	G	G	G	A	A
T	T	T	T	G	T
A	A	A	A	G	A
A	A	A	A	C	A
G	G	G	G	T	G
C	C	C	C	C	C
C	C	C	C	T	T
G	A	G	A	G	G
G	G	G	G	A	G
A	A	A	A	C	C
A	A	A	A	A	G
A	G	A	G	A	G
A	G	A	G	A	A
C	C	C	C	T	T
A	C	A	C	A	C
T	C	T	C	T	T
G	A	G	A	G	G
A	A	G	A	A	A
G	A	A	A	G	G
A	A	A	A	A	G

Table 18.: Exemplary extract from the CHR Merlin output file. In the first row is the DNA-ID of everyone. Maternal alleles are listed on the left, paternal alleles on the right.

8.3 Family structure and affection status of the family members of family #2

ID Family member	ID father	ID mother	Sex (1: male; 2: female)	Affection status (0: unknown; 1: unaffected; 2: affected)
I.1	0	0	1	0
I.2	0	0	2	0
117732	0	0	1	1
117731	I.1	I.2	2	2
117733	I.1	I.2	2	2
II.4	0	0	1	1
117734	I.1	I.2	2	2
II.6	0	0	1	1
III.1	0	0	1	1
117735	117732	117731	2	2
III.3	117732	117731	1	2
III.4	0	0	2	0
III.5	0	0	1	0
III.6	117732	117731	2	0
117737	117732	117731	1	1
III.8	0	0	2	1
117744	117732	117731	1	2
III.10	0	0	1	1
117739	117732	117731	2	2
III.12	0	0	1	1
117740	117732	117731	2	1
117741	117732	117731	1	2
III.15	0	0	2	1
117742	II.4	117733	1	1
III.17	0	0	1	1
117743	II.6	117734	2	1
117738	II.6	117734	1	1
117745	III.1	117735	2	2
117748	III.1	117735	1	2
117749	III.1	117735	1	1
IV.4	III.3	III.4	2	0
IV.5	III.5	III.6	2	0
IV.6	III.5	III.6	2	0
117750	117737	III.8	2	1
117752	117737	III.8	1	2
117754	III.10	117739	2	1
117845	III.12	117740	2	1
IV.11	III.12	117740	2	1
117848	117741	III.15	1	1
117850	III.17	117743	2	1
117851	III.17	117743	1	1

Table 19.: Family structure and affection status of the family members of family #2.
Column “ID family member”: Roman numbers stand for the generation; Arabic numbers for the arrangement in the pedigree (see Figure 3.2) The six-digit numbers represent DNA codes for individuals with available DNA. Column “ID father”/ “ID mother”: 0 (presumed parents whose DNA was not available and therefore not

included in the investigation). Affected persons are highlighted in red. Individuals who were included in the linkage analysis are highlighted in grey.

8.4 Genes listed in region 1,2 and 3

8.4.1 Genes in region 1, Chr. 2 grch37:2:112,424,375-115,298,930

Name / Description	Location	LOEUF	OMIM
ACRP1 / ACR pseudogene 1	2:113,667,478-113,668,621	-	-
ACTR3 / actin related protein 3	2:113,890,063-113,962,596	0.12	OMIM
ANAPC1 / anaphase promoting complex subunit 1	2:111,611,639-111,884,690	0.30	OMIM Morbid
CBWD2 / COBW domain containing 2	2:113,437,691-113,496,204	0.82	OMIM
CDK8P2 / cyclin dependent kinase 8 pseudogene 2	2:112,932,625-112,933,171	-	-
CENPNP2 / CENPN pseudogene 2	2:111,742,586-111,744,662	-	-
CHCHD5 / coiled-coil-helix-coiled-coil-helix domain containing 5	2:112,584,240-112,589,275	1.08	OMIM
CKAP2L / cytoskeleton associated protein 2 like	2:112,736,349-112,764,664	0.63	OMIM Morbid
DDX11L2 / DEAD/H-box helicase 11 like 2 (pseudogene)	2:113,599,036-113,601,261	-	-
DPP10 / dipeptidyl peptidase like 10	2:114,442,299-115,845,780	0.15	OMIM
EEF1E1P1 / eukaryotic translation elongation factor 1 epsilon 1 pseudogene 1	2:111,887,914-111,888,741	-	-
FAM138B / family with sequence similarity 138 member B	2:113,577,382-113,578,852	-	-
FBLN7 / fibulin 7	2:112,138,385-112,188,218	0.96	OMIM
FOXD4L1 / forkhead box D4 like 1	2:113,498,665-113,501,155	1.34	OMIM
HMGN2P23 / high mobility group nucleosomal binding domain 2 pseudogene 23	2:112,981,166-112,981,434	-	-
IGKV1OR2-108 / immunoglobulin kappa variable 1/OR2-108 (non-functional)	2:113,406,396-113,406,872	-	OMIM
IL1A / interleukin 1 alpha	2:112,773,925-112,784,493	1.31	OMIM
IL1B / interleukin 1 beta	2:112,829,751-112,836,816	0.65	OMIM Morbid
IL1F10 / interleukin 1 family member 10	2:113,067,970-113,075,843	1.85	OMIM
IL1RN / interleukin 1 receptor antagonist	2:113,107,214-113,134,016	0.85	OMIM Morbid (3)
IL36A / interleukin 36 alpha	2:113,005,459-113,008,044	1.30	OMIM
IL36B / interleukin 36 beta	2:113,022,089-113,052,867	1.20	OMIM
IL36G / interleukin 36 gamma	2:112,973,203-112,985,658	1.74	OMIM
IL36RN / interleukin 36 receptor antagonist	2:113,058,638-113,065,382	1.90	OMIM Morbid
IL37 / interleukin 37	2:112,912,971-112,918,882	1.73	OMIM
LINC01191 / long intergenic non-protein coding RNA 1191	2:113,970,719-114,007,304	-	-

LINC01961 / long intergenic non-protein coding RNA 1961	2:113,512,222-113,515,488	-	-
MERTK / MER proto-oncogene, tyrosine kinase	2:111,898,607-112,029,561	**0.75**	OMIM Morbid
MIR1302-3 / microRNA 1302-3	2:113,582,959-113,583,096	-	-
MIR4771-2 / microRNA 4771-2	2:111,771,061-111,771,134	-	-
MIR4782 / microRNA 4782	2:113,721,290-113,721,368	-	-
NDUFB4P6 / NADH:ubiquinone oxidoreductase subunit B4 pseudogene 6	2:112,286,057-112,286,422	-	-
NT5DC4 / 5'-nucleotidase domain containing 4	2:112,721,486-112,742,879	**0.93**	-
PAX8 / paired box 8	2:113,215,997-113,278,921	**0.20**	OMIM Morbid
PAX8-AS1 / PAX8 antisense RNA 1	2:113,211,421-113,276,581	-	-
PGM5P4 / phosphoglucomutase 5 pseudogene 4	2:113,541,937-113,563,298	-	-
PGM5P4-AS1 / PGM5P4 antisense RNA 1	2:113,526,836-113,542,694	-	-
POLR1B / RNA polymerase I subunit B	2:112,541,915-112,579,818	**0.25**	OMIM Morbid
POLR2DP1 / POLR2D pseudogene 1	2:112,995,517-112,995,937	-	-
PSD4 / pleckstrin and Sec7 domain containing 4	2:113,157,325-113,209,396	**0.37**	OMIM
RABL2A / RAB, member of RAS oncogene family like 2A	2:113,627,229-113,643,396	**1.56**	OMIM
RGPD8 / RANBP2 like and GRIP domain containing 8	2:112,368,369-112,434,488	**1.07**	OMIM
RN7SL297P / RNA, 7SL, cytoplasmic 297, pseudogene	2:111,930,175-111,930,467	-	-
RNU2-41P / RNA, U2 small nuclear 41, pseudogene	2:114,420,125-114,420,302	-	-
RNU6-1180P / RNA, U6 small nuclear 1180, pseudogene	2:113,086,805-113,086,911	-	-
RNU6-744P / RNA, U6 small nuclear 744, pseudogene	2:113,681,111-113,681,217	-	-
RPL23AP7 / ribosomal protein L23a pseudogene 7	2:113,610,502-113,627,090	-	-
RPL34P8 / ribosomal protein L34 pseudogene 8	2:111,675,026-111,675,343	-	-
RTRAFP1 / RTRAF pseudogene 1	2:111,940,302-111,941,036	-	-
SEPHS1P7 / selenophosphate synthetase 1 pseudogene 7	2:114,260,050-114,261,226	-	-
SLC20A1 / solute carrier family 20 member 1	2:112,645,939-112,663,825	**0.48**	OMIM
SLC30A6P1 / SLC30A6 pseudogene 1	2:111,952,144-111,952,671	-	-
SLC35F5 / solute carrier family 35 member F5	2:113,705,011-113,756,693	**0.83**	-
SNRPA1P1 / SNRPA1 pseudogene 1	2:113,657,908-113,658,675	-	-
TMEM87B / transmembrane protein 87B	2:112,055,269-112,119,318	**1.03**	OMIM
TTL / tubulin tyrosine ligase	2:112,482,156-112,541,739	**0.52**	OMIM
VINAC1P / vinculin/alpha-catenin family member 1, pseudogene	2:112,439,312-112,469,687	-	-
WASH2P / WASP family homolog 2, pseudogene	2:113,588,550-113,599,043	-	-
XIAPP3 / X-linked inhibitor of apoptosis pseudogene 3	2:112,853,328-112,854,122	-	-
ZC3H6 / zinc finger CCCH-type containing 6	2:112,275,597-112,340,063	**0.41**	-
ZC3H8 / zinc finger CCCH-type containing 8	2:112,211,529-112,255,136	**0.56**	-

Table 20.: Genes in region 1. In the left column, the individual genes are listed. In the adjacent column is their exact position on the respective chromosome. The third column shows the LOEUF score of the genes. The fourth column shows whether the gene is already listed in OMIM or not. The thickly highlighted genes have already been associated with seizures or epilepsies and have been described in more detail in 3.5.1.

8.4.2 Genes in region 2, Chr. 2 grch37:2:232,779,515-239,503,407

Name / Description	Location	LOEUF	OMIM
ACKR3 / atypical chemokine receptor 3	2:236,567,787-236,582,354	0.68	OMIM
AGAP1 / ArfGAP with GTPase domain, ankyrin repeat and PH domain 1	2:235,494,043-236,131,800	0.23	OMIM
AGAP1-IT1 / AGAP1 intronic transcript 1	2:235,505,751-235,507,566	-	-
ALPG / alkaline phosphatase, germ cell	2:232,406,844-232,410,714	1.31	OMIM
ALPI / alkaline phosphatase, intestinal	2:232,456,125-232,460,753	1.49	OMIM
ALPP / alkaline phosphatase, placental	2:232,378,724-232,382,889	1.29	OMIM
ARL4C / ADP ribosylation factor like GTPase 4C	2:234,493,041-234,497,081	0.83	OMIM
ASB1 / ankyrin repeat and SOCS box containing 1	2:238,426,742-238,452,250	0.94	OMIM
ASB18 / ankyrin repeat and SOCS box containing 18	2:236,194,872-236,264,409	1.64	-
ATG16L1 / autophagy related 16 like 1	2:233,210,051-233,295,674	0.32	OMIM Morbid
CEP19P1 / CEP19 pseudogene 1	2:235,147,016-235,147,500	-	-
CHRND / cholinergic receptor nicotinic delta subunit	2:232,525,993-232,536,667	0.98	OMIM Morbid (4)
CHRNG / cholinergic receptor nicotinic gamma subunit	2:232,539,692-232,548,115	1.31	OMIM Morbid (2)
COL6A3 / collagen type VI alpha 3 chain	2:237,324,003-237,414,207	0.56	OMIM Morbid (3)
COPS8 / COP9 signalosome subunit 8	2:237,085,882-237,100,474	0.49	OMIM
DGKD / diacylglycerol kinase delta	2:233,354,494-233,472,104	0.37	OMIM
DIS3L2 / DIS3 like 3'-5' exoribonuclease 2	2:231,961,245-232,344,350	0.33	OMIM Morbid
DIS3L2P1 / DIS3 like 3'-5' exoribonuclease 2 pseudogene 1	2:232,442,034-232,447,418	-	-
DNAJB3 / DnaJ heat shock protein family (Hsp40) member B3	2:233,742,750-233,744,015	-	-
ECEL1 / endothelin converting enzyme like 1	2:232,479,827-232,487,834	0.83	OMIM Morbid
ECEL1P1 / endothelin converting enzyme like 1 pseudogene 1	2:232,415,818-232,419,816	-	-
ECEL1P2 / endothelin converting enzyme like 1 pseudogene 2	2:232,385,750-232,387,457	-	-
ECEL1P3 / endothelin converting enzyme like 1 pseudogene 3	2:232,349,754-232,351,309	-	-
EEF1B2P7 / eukaryotic translation elongation factor 1 beta 2 pseudogene 7	2:232,729,523-232,730,198	-	-
EFHD1 / EF-hand domain family member D1	2:232,606,057-232,682,780	1.49	OMIM
EIF4E2 / eukaryotic translation initiation factor 4E family member 2	2:232,550,674-232,583,644	0.49	OMIM
ERFE / erythroferrone	2:238,158,970-238,168,900	1.28	OMIM
ESPNL / espin like	2:238,100,340-238,133,287	0.94	-
GBX2 / gastrulation brain homeobox 2	2:236,165,236-236,168,386	0.56	OMIM
GIGYF2 / GRB10 interacting GYF protein 2	2:232,697,299-232,860,605	0.06	OMIM Morbid
HES6 / hes family bHLH transcription factor 6	2:238,238,267-238,240,662	1.38	OMIM
HJURP / Holliday junction recognition protein	2:233,833,416-233,854,566	0.79	OMIM

HSPE1P9 / heat shock protein family E (Hsp10) member 1 pseudogene 9	2:234,455,157-234,455,783	-	-
ILKAP / ILK associated serine/threonine phosphatase	2:238,170,402-238,203,708	0.30	OMIM
INPP5D / inositol polyphosphate-5-phosphatase D	2:233,059,967-233,207,903	0.23	OMIM
IQCA1 / IQ motif containing with AAA domain 1	2:236,324,147-236,507,535	0.69	-
IQCA1-AS1vIQCA1 antisense RNA 1	2:236,391,074-236,392,388	-	-
KCNJ13 / potassium inwardly rectifying channel subfamily J member 13	2:232,765,802-232,776,565	0.86	OMIM Morbid (2)
KLHL30 / kelch like family member 30	2:238,138,668-238,152,947	1.22	-
KLHL30-AS1 / KLHL30 antisense RNA 1	2:238,152,889-238,155,994	-	-
LINC01107 / long intergenic non-protein coding RNA 1107	2:238,510,690-238,555,054	-	-
LINC01173 / long intergenic non-protein coding RNA 1173	2:234,682,668-234,717,764	-	-
LINC01891 / long intergenic non-protein coding RNA 1891	2:234,444,590-234,454,595	-	-
LINC01937 / long intergenic non-protein coding RNA 1937	2:238,554,541-238,740,849	-	-
LINC02610 / long intergenic non-protein coding RNA 2610	2:238,224,552-238,231,699	1.95	-
LRRFIP1 / LRR binding FLII interacting protein 1	2:237,627,587-237,813,682	0.56	OMIM
MIR5001 / microRNA 5001	2:232,550,474-232,550,573	-	-
MIR562 / microRNA 562	2:232,172,653-232,172,747	-	-
MIR6811 / microRNA 6811	2:237,510,931-237,510,988	-	-
MLPH / melanophilin	2:237,485,428-237,555,322	0.90	OMIM Morbid
MROH2A / maestro heat like repeat family member 2A	2:233,775,679-233,833,423	0.75	-
MSL3P1 / MSL complex subunit 3 pseudogene 1	2:233,865,437-233,868,444	-	-
NEU2 / neuraminidase 2	2:233,032,672-233,035,057	1.93	OMIM
NGEF / neuronal guanine nucleotide exchange factor	2:232,878,701-233,013,256	0.39	OMIM
NPPC / natriuretic peptide C	2:231,921,809-231,926,396	1.57	OMIM
NRBF2P6 / NRBF2 pseudogene 6	2:232,343,116-232,343,903	-	-
PER2 / period circadian regulator 2	2:238,244,044-238,290,102	0.35	OMIM Morbid
PPFIA1P1 / PPFIA1 pseudogene 1	2:233,545,482-233,547,032	-	-
PRLH / prolactin releasing hormone	2:237,566,574-237,567,175	1.86	OMIM
PRSS56 / serine protease 56	2:232,520,388-232,525,716	-	OMIM Morbid
RAB17 / RAB17, member RAS oncogene family	2:237,574,322-237,601,614	0.95	OMIM
RAMP1 / receptor activity modifying protein 1	2:237,858,893-237,912,106	1.12	OMIM
RBM44 / RNA binding motif protein 44	2:237,798,389-237,842,808	0.51	-
RN7SL204P / RNA, 7SL, cytoplasmic 204, pseudogene	2:236,072,295-236,072,556	-	-
RN7SL32P / RNA, 7SL, cytoplasmic 32, pseudogene	2:233,205,199-233,205,479	-	-
RN7SL359P / RNA, 7SL, cytoplasmic 359, pseudogene	2:232,649,849-232,650,136	-	-
RNU1-31P / RNA, U1 small nuclear 31, pseudogene	2:236,210,373-236,210,548	-	-
RNU6-1051P / RNA, U6 small nuclear 1051, pseudogene	2:237,015,204-237,015,305	-	-
RNU6-107P / RNA, U6 small nuclear 107, pseudogene	2:232,782,665-232,782,766	-	-
RNU6-1140P / RNA, U6 small nuclear 1140, pseudogene	2:237,563,165-237,563,270	-	-

RNU6-1333P / RNA, U6 small nuclear 1333, pseudogene	2:238,028,971-238,029,063	-	-
RNU6-234P / RNA, U6 small nuclear 234, pseudogene	2:238,412,642-238,412,745	-	-
RNU7-127P / RNA, U7 small nuclear 127 pseudogene	2:236,037,250-236,037,311	-	-
RPL17P11 / ribosomal protein L17 pseudogene 11	2:233,721,522-233,722,065	-	-
RPL3P5 / ribosomal protein L3 pseudogene 5	2:236,407,846-236,409,077	-	-
RPS20P12 / ribosomal protein S20 pseudogene 12	2:234,329,483-234,329,838	-	-
SAG / S-antigen visual arrestin	2:233,307,816-233,347,055	**1.44**	OMIM Morbid (2)
SCARNA5 / small Cajal body-specific RNA 5	2:233,275,727-233,276,002	-	OMIM
SCARNA6 / small Cajal body-specific RNA 6	2:233,288,676-233,288,940	-	OMIM
SCLY / selenocysteine lyase	2:238,060,924-238,099,413	**1.15**	OMIM
SH3BP4 / SH3 domain binding protein 4	2:2349,52,017-235,055,714	**0.57**	OMIM
SNORC / secondary ossification center associated regulator of chondrocyte maturation	2:232,857,270-232,878,708	**1.02**	-
SPP2 / secreted phosphoprotein 2	2:234,050,679-234,077,134	**1.89**	OMIM
TIGD1 / tigger transposable element derived 1	2:232,543,883-232,550,557	**1.16**	OMIM
TMSB10P1 / thymosin beta 10 pseudogene 1	2:235,882,362-235,882,493	-	-
TRAF3IP1 / TRAF3 interacting protein 1	2:238,320,441-238,400,897	**0.74**	OMIM Morbid
TRPM8 / transient receptor potential cation channel subfamily M member 8	2:233,917,373-234,019,522	**1.04**	OMIM
UBE2F / ubiquitin conjugating enzyme E2 F (putative)	2:237,966,827-238,042,782	**0.38**	OMIM
UBE2F-SCLY / UBE2F-SCLY readthrough (NMD candidate)	2:237,967,014-238,099,412	-	-
UGT1A1 / UDP glucuronosyltransferase family 1 member A1	2:233,760,270-233,773,300	**1.08**	OMIM Morbid (5)
UGT1A10 / UDP glucuronosyltransferase family 1 member A10	2:233,636,454-233,773,305	**1.74**	OMIM
UGT1A11P / UDP glucuronosyltransferase family 1 member A11, pseudogene	2:233,603,553-233,604,401	-	-
UGT1A12P / UDP glucuronosyltransferase family 1 member A12, pseudogene	2:233,585,439-233,586,291	-	-
UGT1A13P / UDP glucuronosyltransferase family 1 member A13, pseudogene	2:233,647,926-233,649,026	-	-
UGT1A2P / UDP glucuronosyltransferase family 1 member A2, pseudogene	2:233,747,214-233,748,079	-	-
UGT1A3 / UDP glucuronosyltransferase family 1 member A3	2:233,729,108-233,773,299	**1.33**	OMIM
UGT1A4 / UDP glucuronosyltransferase family 1 member A4	2:233,718,736-233,773,300	**1.13**	OMIM
UGT1A5 / UDP glucuronosyltransferase family 1 member A5	2:233,712,992-233,773,299	**1.29**	OMIM
UGT1A6 / UDP glucuronosyltransferase family 1 member A6	2:233,691,607-233,773,300	**1.06**	OMIM
UGT1A7 / UDP glucuronosyltransferase family 1 member A7	2:233,681,938-233,773,299	**1.46**	OMIM
UGT1A8 / UDP glucuronosyltransferase family 1 member A8	2:233,617,645-233,773,310	**1.62**	OMIM
UGT1A9vUDP glucuronosyltransferase family 1 member A9	2:233,671,898-233,773,300	**1.11**	OMIM
USP40 / ubiquitin specific peptidase 40	2:233,475,520-233,566,782	**0.91**	OMIM

Table 21.: Genes in region 2. In the left column, the individual genes are listed. In the adjacent column is their exact position on the respective chromosome. The third column shows the LOEUF score of the genes. The fourth column shows whether the gene is already listed in OMIM or not. The thickly highlighted genes have already been associated with seizures or epilepsies and have been described in more detail in 0.

8.4.3 Genes in region 3, Chr. 18 grch37:18:57,124,276-63,275,241

Name / Description	Location	LOEUF	OMIM
ACTBP9 / ACTB pseudogene 9	18:62,442,028-62,443,126	-	-
ATP5MC1P6 / ATP synthase membrane subunit c locus 1 pseudogene 6	18:63,496,989-63,497,400	-	-
BCL2 / BCL2 apoptosis regulator	18:63,123,346-63,320,128	**0.71**	OMIM Morbid
CCBE1 / collagen and calcium binding EGF domains 1	18:59,430,939-59,697,662	**1.01**	OMIM Morbid
CDH20 / cadherin 20	18:61,333,430-61,555,779	**0.41**	OMIM
CTBP2P3 / CTBP2 pseudogene 3	18:60,663,115-60,664,369	-	-
ENTR1P1 / ENTR1 pseudogene 1	18:60,009,994-60,011,143	-	-
GLUD1P4 / glutamate dehydrogenase 1 pseudogene 4	18:59,778,572-59,778,989	-	-
HMGN1P31 / high mobility group nucleosome binding domain 1 pseudogene 31	18:60,794,613-60,794,916	-	-
HMSD / histocompatibility minor serpin domain containing	18:63,949,301-63,981,774	**1.67**	OMIM
KDSR / 3-ketodihydrosphingosine reductase	18:63,327,726-63,367,228	**0.73**	OMIM Morbid
LINC00305 / long intergenic non-protein coding RNA 305	18:64,079,989-64,149,088	-	OMIM
LINC01538 / long intergenic non-protein coding RNA 1538	18:64,181,314-64,260,055	-	-
LINC01544 / long intergenic non-protein coding RNA 1544	18:61,748,082-61,758,464	-	-
LINC01916 / long intergenic non-protein coding RNA 1916	18:65,423,958-65,448,167	-	-
LINC01924 / long intergenic non-protein coding RNA 1924	18:64,041,555-64,423,601	-	-
MC4R / melanocortin 4 receptor	18:60,371,062-60,372,775	**1.60**	OMIM Morbid
MRPS5P4 / mitochondrial ribosomal protein S5 pseudogene 4	18:60,519,841-60,520,165	-	-
NFE2L3P1 / nuclear factor, erythroid 2 like 3 pseudogene 1	18:59,969,438-59,971,396	-	-
PHLPP1 / PH domain and leucine rich repeat protein phosphatase 1	18:62,715,541-62,980,433	**0.26**	OMIM
PIGN / phosphatidylinositol glycan anchor biosynthesis class N	18:61,905,255-62,187,118	**1.02**	OMIM Morbid
PMAIP1 / phorbol-12-myristate-13-acetate-induced protein 1	18:59,899,996-59,904,305	**1.85**	OMIM
RELCH / RAB11 binding and LisH domain, coiled-coil and HEAT repeat containing	18:62,187,255-62,310,249	**0.27**	OMIM
RN7SL342P / RNA, 7SL, cytoplasmic 342, pseudogene	18:59,972,914-59,973,207	-	-
RN7SL705P / RNA, 7SL, cytoplasmic 705, pseudogene	18:62,650,308-62,650,601	-	-
RNF152 / ring finger protein 152	18:61,808,067-61,894,247	**1.13**	OMIM
RNU4-17P / RNA, U4 small nuclear 17, pseudogene	18:60,163,567-60,163,707	-	-
RNU6-116P / RNA, U6 small nuclear 116, pseudogene	18:61,391,595-61,391,703	-	-
RNU6-142P / RNA, U6 small nuclear 142, pseudogene	18:62,731,226-62,731,332	-	-
RNU6-567P / RNA, U6 small nuclear 567, pseudogene	18:60,018,625-60,018,731	-	-
RPIAP1 / ribose 5-phosphate isomerase A pseudogene 1	18:61,957,273-61,958,014	-	-
RPL12P39 / ribosomal protein L12 pseudogene 39	18:64,072,158-64,072,653	-	-

RPL17P44 / ribosomal protein L17 pseudogene 44	18:62,415,678-62,416,227	-	-
RPL30P14 / ribosomal protein L30 pseudogene 14	18:61,404,858-61,405,205	-	-
RPS26P54 / ribosomal protein S26 pseudogene 54	18:59,761,558-59,761,905	-	-
RPS3AP49 / RPS3A pseudogene 49	18:60,149,576-60,150,365	-	-
SERPINB10 / serpin family B member 10	18:63,897,174-63,936,111	**1.47**	OMIM
SERPINB11 / serpin family B member 11	18:63,647,579-63,726,432	**1.16**	OMIM
SERPINB12 / serpin family B member 12	18:63,556,160-63,567,011	**1.95**	OMIM
SERPINB13 / serpin family B member 13	18:63,586,989-63,604,639	**1.15**	OMIM
SERPINB2 / serpin family B member 2	18:63,871,692-63,903,888	**1.53**	OMIM
SERPINB3 / serpin family B member 3	18:63,655,197-63,661,893	**1.86**	OMIM
SERPINB4 / serpin family B member 4	18:63,637,259-63,644,256	**1.88**	OMIM
SERPINB5 / serpin family B member 5	18:63,476,958-63,505,085	**0.90**	OMIM
SERPINB7 / serpin family B member 7	18:63,752,935-63,805,376	**1.31**	OMIM Morbid
SERPINB8 / serpin family B member 8	18:63,970,029-64,019,779	**1.41**	OMIM Morbid
SINHCAFP2 / SINHCAF pseudogene 2	18:60,016,778-60,017,441		-
TNFRSF11A / TNF receptor superfamily member 11a	18:62,325,287-62,391,288	**0.60**	Morbid (3)
VPS4B / vacuolar protein sorting 4 homolog B	18:63,389,190-63,422,483	**0.48**	OMIM
ZCCHC2 / zinc finger CCHC-type containing 2	18:62,523,025-62,587,709	**0.28**	-

Table 22.: Genes in region 3. In the left column, the individual genes are listed. In the adjacent column is their exact position on the respective chromosome. The third column shows the LOEUF score of the genes. The fourth column shows whether the gene is already listed in OMIM or not. The thickly highlighted genes have already been associated with seizures or epilepsies and have been described in more detail in 3.5.3.

8.4.4 Genes in the *FERMT2* region, Chr. 14 grch37:14:52073301-54172855

Name / Description	Location	LOEUF	OMIM
COX5AP2 / cytochrome c oxidase subunit 5A pseudogene 2	14:52,117,602-52,117,995	-	-
DDHD1 / DDHD domain containing 1	14:53,036,745-53,153,323	0.34	OMIM
DDHD1-DT / DDHD1 divergent transcript	14:53,153,354-53,157,528	-	Morbid -
ERO1A / endoplasmic reticulum oxidoreductase 1 alpha	14:52,639,915-52,695,900	0.37	OMIM
FERMT2 / FERM domain containing kindlin 2	14:52,857,268-52,952,435	0.08	OMIM
FRMD6 / FERM domain containing 6	14:51,489,100-51,730,727	0.49	OMIM
FRMD6-AS1 / FRMD6 antisense RNA 1	14:51,649,516-51,651,744	-	-
GNG2 / G protein subunit gamma 2	14:51,826,195-51979,342	1.35	OMIM
GNPNAT1 / glucosamine-phosphate N-acetyltransferase 1	14:52,775,193-52,791,668	1.17	OMIM
GPR137C / G protein-coupled receptor 137C	14:52,552,836-52,637,713	0.50	Morbid -
LINC02319 / long intergenic non-protein coding RNA 2319	14:52,111,122-52,129,625	-	-
LINC02331 / long intergenic non-protein coding RNA 2331	14:53,685,139-53,850,882	-	-
NID2 / nidogen 2	14:52,004,809-52,069,059	0.56	OMIM
OR7E105P	14:51,756,996-51,757,988	-	-
O / lfactory receptor family 7 subfamily E member 105 pseudogene			
OR7E106P / olfactory receptor family 7 subfamily E member 106 pseudogene	14:51,763,056-51,764,076	-	-
OR7E159P / olfactory receptor family 7 subfamily E member 159 pseudogene	14:51,771,789-51,772,086	-	-
PSMC6 / proteasome 26S subunit, ATPase 6	14:52,707,178-52,728,590	0.25	OMIM
PTGDR / prostaglandin D2 receptor	14:52,267,698-52,276,724	1.50	OMIM
PTGER2 / prostaglandin E receptor 2	14:52,314,305-52,328,598	0.98	Morbid OMIM
RNA5SP385 / RNA, 5S ribosomal pseudogene 385	14:51,692,752-51,692,861	-	Morbid -
RNU6-301P / RNA, U6 small nuclear 301, pseudogene	14:51,724,103-51,724,209	-	-
RPS3AP46 / RPS3A pseudogene 46	14:53,612,724-53,613,405	-	-
RTRAF / RNA transcription, translation and transport factor	14:51,989,514-52,010,694	0.59	OMIM
STYX / serine/threonine/tyrosine interacting protein	14:52,730,166-52,774,989	0.51	OMIM
TXNDC16 / thioredoxin domain containing 16	14:52,430,596-52,552,522	0.63	OMIM

Table 23.: Genes in the *FERMT2* region. In the left column, the individual genes are listed. In the adjacent column is their exact position on the respective chromosome. The third column shows the LOEUF score of the genes. The fourth column shows whether the gene is already listed in OMIM or not. The thickly highlighted genes have already been associated with seizures or epilepsies.

8.5 Candidate regions defined by narrowed down haplotypes

2	**rs68012394**	124.5194
2	**rs10178648**	124.5441
2	**GSA-rs1400323**	124.8221
2	**rs6755850**	124.9221
2	**rs1545133**	125.2063
2	**GSA-rs1018317**	125.3382
2	**GSA-rs11123180**	125.4178
2	**rs80314553**	125.8193
2	**GSA-rs2900745**	125.9723
2	**rs13016677**	126.1775
2	**rs10208402**	126.8527
2	**GSA-rs1835328**	126.9769

Table 24.: Region 1, MAP-File. In bold are the regions to which the region of interest could be narrowed down by haplotyping. “2” in the left column illustrates that this is chromosome 2. The markers from the .dat file (middle row) are assigned to a specific position on the chromosome in cM (right column).

2	rs4355094	237.25
2	rs2289233	237.2747
2	rs1060780	237.5052
2	GSA-rs7589292	237.5809
2	rs7570052	238.1896
2	rs4973413	238.3177
2	rs4973473	238.5054
2	**rs28688824**	238.714
2	**rs2853362**	239.0265
2	**rs2853386**	239.1801
2	**rs2344309**	239.3472
2	**rs10193825**	239.4983
2	**rs6704768**	239.7191
2	**GSA-rs10211596**	240.0074
2	**rs6705372**	240.1261
2	**rs7425956**	240.2454
2	**rs36180605**	240.546
2	**rs6760912**	240.7048
2	**GSA-rs706848**	241.0088
2	**GSA-rs882428**	241.1137
2	**rs6759892**	241.6414
2	**rs4663969**	241.775
2	**rs1500481**	241.8923
2	**rs968402**	242.0772
2	**rs7605850**	242.3339
2	**rs1395835**	242.4808
2	**rs250949**	242.6223
2	**rs10929015**	242.9582
2	**rs4663369**	243.0852
2	**rs12478222**	243.5009
2	**rs10176266**	243.5954
2	**rs11685523**	243.788
2	**rs4663150**	243.9266
2	**rs4663483**	244.1541
2	**rs13008181**	244.3266
2	**GSA-rs1946950**	244.5396
2	**rs73126170**	245.0399
2	**GSA-rs10190780**	245.1392
2	**GSA-rs6741246**	245.2314
2	**rs6719254**	245.4687
2	**rs2116399**	246.0019
2	**rs4381737**	246.102
2	**rs11676121**	246.2568
2	**rs1356334**	246.646
2	**rs7559293**	246.8748
2	**rs11676226**	247.2674
2	**rs12991774**	247.4316
2	**rs1806688**	247.5263
2	**rs10181993**	247.7105
2	**rs7599602**	247.9559
2	**rs3768921**	248.0879

2	**rs6735079**	248.3627
2	**rs2046281**	248.4622
2	**GSA-rs2720100**	248.8088
2	**rs11686416**	248.9355
2	**rs1435847**	249.1461
2	**rs4641891**	249.8923
2	**rs4663708**	249.9288
2	**rs4663726**	250.1257
2	**rs6431548**	250.8137
2	**rs2029771**	250.9889
2	**rs10460281**	251.1254
2	**rs1198819**	251.5235
2	**GSA-rs12995100**	251.6665
2	**rs6431564**	251.7164
2	**rs6713087**	251.8251
2	**rs61544914**	252.2248
2	**rs12479385**	252.2976
2	**rs6740485**	252.5591
2	**rs56294817**	252.5882
2	**rs9287631**	253.3438
2	**rs6716187**	253.4986
2	GSA-rs4663946	253.7732
2	rs9711321	253.9971
2	rs36178247	254.1736
2	rs13011030	254.3439
2	rs13010905	254.518
2	rs3791370	254.7777
2	rs3791442	254.8855
2	rs519165	255.1375
2	rs1467296	255.3237
2	rs11124194	255.5205
2	rs13392399	255.7036
2	rs62180956	255.9458
2	GSA-rs4328655	256.0311
2	rs11124220	256.3621
2	rs11904223	256.5425
2	rs2412024	256.7798
2	rs138289134	256.8813
2	rs12695027	257.0567
2	rs10933566	257.5375
2	GSA-rs2368502	257.6715
2	rs13017098	257.9607
2	rs1109411	258.1348
2	rs4676406	258.4326
2	rs2060188	258.5249
2	rs4676381	258.83
2	rs9288742	259.9404
2	rs7559834	260.5004
2	rs4270338	260.5041
2	rs10200024	260.6105
2	rs10183172	261.0196

2	rs6746124	261.1212
2	rs13019391	261.5028
2	rs11683447	261.9456
2	rs4234101	262.5834
2	rs7570235	264.1071
2	rs7598295	264.2169
2	rs9288728	264.2173
2	rs6437270	264.6727
2	rs34767042	265.3321
2	rs10755057	265.4712
2	rs56095230	266.797

Table 25.: Region 2, MAP-File. In bold are the regions to which the region of interest could be narrowed down by haplotyping. “2” in the left column illustrates that this is chromosome 2. The markers from the .dat file (middle row) are assigned to a specific position on the chromosome in cM (right column).

18	rs1450804	63.9141
18	rs8095304	64.12957
18	rs1421201	64.29998
18	rs2543025	64.56029
18	rs12455675	64.69201
18	GSA-rs9959453	64.90079
18	rs3744998	65.16325
18	rs590694	65.2749
18	rs9807399	65.41121
18	rs1562000	65.74
18	GSA-rs1452277	65.954
18	rs1462981	66.05379
18	rs3893060	66.28659
18	GSA-rs4121690	66.33689
18	GSA-rs2048523	66.98895
18	GSA-rs1560900	67.07252
18	rs1553804	67.32924
18	rs10502885	67.43015
18	rs2852945	67.698
18	rs7504740	68.07092
18	rs4300413	68.17641
18	rs11082646	68.28183
18	rs9961721	68.51854
18	rs4940355	68.81077
18	GSA-rs7239114	68.91924
18	rs9960934	69.39113
18	rs9947954	69.48047
18	rs2046241	69.64619
18	rs4939827	69.92409
18	rs895667	70.04939
18	rs9967417	70.57242
18	rs12963212	70.66685
18	rs11875988	70.8875
18	rs630665	71.14144
18	rs4398183	71.30917
18	rs74839251	71.5406
18	rs8091509	71.67328
18	GSA-rs11662142	71.91
18	rs9949687	72.06473
18	rs1838191	72.29314
18	rs2255610	72.47333
18	rs12455399	72.81071
18	rs5001894	72.89069
18	GSA-rs321858	73.17381
18	rs1368151	73.35667
18	rs748317	73.52309
18	rs2953258	73.79208
18	GSA-rs4312399	73.93714
18	rs1445088	74.09923
18	rs1445093	74.25858
18	rs8093706	74.54618

18	rs1144049	74.70353
18	rs9950125	74.80829
18	rs12457160	75.09109
18	rs9954439	75.29861
18	rs1553108	75.45439
18	rs34163044	75.68093
18	rs28377497	75.8719
18	rs949031	76.09814
18	rs2872041	76.32831
18	exm2268153	76.38
18	rs1506632	76.71249
18	rs2588482	76.91538
18	rs4801018	77.12669
18	GSA-rs10513889	77.22047
18	GSA-rs12454365	77.50038
18	rs4801076	77.6064
18	rs12150819	77.88623
18	rs8093842	78.51064
18	rs319439	78.71609
18	rs11877445	78.83174
18	rs4941159	79.01905
18	rs4528665	79.28938
18	rs1893480	79.4786
18	rs12051934	79.71942
18	rs4940694	79.87178
18	rs9947399	80.11844
18	exm2268154	80.25471
18	rs4940757	80.40197
18	rs7231686	80.682
18	GSA-rs561894	80.9306
18	**rs11874251**	81.42175
18	**rs1789475**	81.47918
18	**rs2564489**	81.71146
18	**rs2195534**	82.02397
18	**rs7244658**	82.03665
18	**rs1791330**	82.2942
18	**rs66952881**	82.38962
18	**rs11152195**	82.6989
18	**rs28534746**	82.71854
18	**rs2331841**	83.18513
18	**rs9946888**	83.44933
18	**rs1943241**	83.90685
18	**rs530676**	84.15061
18	**GSA-rs1478526**	84.48793
18	**rs11152269**	84.98
18	**rs242622**	85.0724
18	**rs12953685**	85.58305
18	**rs757620**	85.65647
18	**rs2535357**	85.87887
18	**rs9945515**	86.1014
18	**rs4941052**	86.28721

18	**rs1942990**	86.47124
18	**rs930171**	86.5826
18	**rs4369774**	86.91499
18	**rs694419**	87.03765
18	**rs595286**	87.20095
18	**rs547690**	87.5726
18	**GSA-rs9945969**	87.70436
18	**GSA-rs12457893**	87.88612
18	**rs1355196**	88.09502
18	**rs6567383**	88.42972
18	**rs12959692**	88.53383
18	**rs1720853**	88.56751
18	**rs6567401**	88.84978
18	**rs4940617**	89.46069
18	**rs1155993**	89.48473
18	**rs7239193**	89.68194
18	**rs1395170**	89.75704
18	**rs9956331**	90.47855
18	**rs2156175**	90.54607
18	**GSA-rs605818**	90.67138
18	**GSA-rs1587962**	90.83627
18	**GSA-rs607293**	91.09609
18	**rs681307**	91.14826
18	**rs7229020**	91.50667
18	GSA-rs2193298	92.30563
18	rs8091835	92.60558
18	rs17728286	92.72561
18	rs11665452	92.88915
18	rs1351407	93.17113
18	rs894909	93.36361
18	rs4306630	93.70705
18	rs11661918	94.04276
18	rs1850558	94.30309
18	rs6566351	94.57392
18	rs562065	94.68719
18	rs9789218	94.91586
18	rs574539	95.06039
18	rs1876978	95.17209
18	rs595533	95.4918
18	GSA-rs4290560	95.60725
18	GSA-rs1354967	96.12653
18	GSA-rs12967871	96.32055
18	rs10084094	96.56815
18	GSA-rs8097743	96.62204
18	GSA-rs10871650	97.19918
18	GSA-rs9963608	97.23955
18	rs2575177	97.72993
18	rs1007822	97.796
18	rs7230661	98.09509
18	rs12969215	98.31754
18	rs1528881	98.60765

18	GSA-rs12954585	98.72969
18	GSA-rs996919	98.86375
18	rs1378146	99.08882
18	rs11660894	99.28865
18	rs948665	99.35351
18	rs10514031	99.86567
18	rs4268850	100.189
18	GSA-rs8088832	100.2315
18	rs17352438	100.4764

Table 26.: Region 3, MAP-File. In bold are the regions to which the region of interest could be narrowed down by haplotyping. “18” in the left column illustrates that this is chromosome 18. The markers from the .dat file (middle row) are assigned to a specific position on the chromosome in cM (right column).

9 Votum der Ethikkommission

Technische Universität München

Technische Universität München · Fakultät für Medizin · Ethikkommission
Ismaninger Str. 22 · 81675 München · Germany

Herrn
PD Dr. Tim M. Strom
Institut für Humanmedizin
Ismaninger Str. 22

81675 München

Fakultät für Medizin
Ethikkommission

Prof. Dr.
Albert Schömig
Vorsitzender

Prof. Dr.
Franz B. Hofmann
Stellvertretender Vorsitzender

Prof. Dr.
Georg Schmidt
Geschäftsführendes Mitglied

Ismaninger Str. 22
81675 München
Germany

Tel +49.89.4140.4371
Fax +49.89.4140.4199

info@ek.med.tum.de
www.ek.med.tum.de

06.02.2013
Projektnummer: 5360/13 (bitte bei jedem Schriftwechsel angeben)

Genomweite Sequenzierung bei genetischen Syndromen

Sehr geehrter Herr Kollege Strom,

in der Sitzung der Ethikkommission vom Dienstag, dem 05.02.2013 wurde das o.g. noch einmal Projekt besprochen.

Zur Begutachtung lagen Prüfplan und Elterninformation/-einverständniserklärung vor.

Die Ethikkommission, die sich aus den Mitgliedern
Herrn Prof. Dr. G. Schmidt, Herrn Dr. R. Haubenthaler,
Herrn Prof. Dr. B. Hemmer, Herrn Prof. Dr. Chr. Peschel,
Herrn Prof. Dr. G. H. Schlund und Frau Prof. Dr.Dr. R. Senekowitsch-Schmidtke zusammensetzte, kam zu folgendem Votum:

Die Ethikkommission erhebt keine Einwände gegen das geplante Forschungsprojekt.

In der Informationsschrift für die Eltern ist deutlich zu machen, daß genetische Zufallsbefunde, die auf ein bisher unbekanntes Krankheitsrisiko beim Studienteilnehmern oder seinen Eltern hinweisen, nur dann mitgeteilt werden, wenn sie

1) *bei der initialen Genomsequenzierung auffallen und wenn*
2) *es sich um eine behandelbare Erkrankung handelt.*

Der diesbezügliche Hinweis auf die Beratung durch die Ethikkommission sollte gestrichen werden.

In der Patienteninformation sollte klargestellt werden, wer das Projekt finanziert. Medizinischer Jargon ist durch laienverständliches Deutsch zu ersetzen. Unvermeidliche Fremdwörter sind gegebenenfalls in Klammern zu erläutern.

In der Information/Einverständniserklärung sind beide Eltern anzusprechen; sie ist von beiden zu unterschreiben. Dies gilt nicht für alleinerziehende Eltern. Im Abschnitt über den Datenschutz ist auch der Schutz elterlicher Daten zu erwähnen.

Mit freundlichem Gruß

[redacted]

Prof. Dr. G. Schmidt
Vorsitzender

Die Ethikkommission der Fakultät für Medizin der Technischen Universität München arbeitet gemäß den nationalen gesetzlichen Bestimmungen und den ICH-GCP-Richtlinien.

Mitteilungen über schwerwiegende oder unerwartete unerwünschte Ereignisse sind mit einer Stellungnahme des Prüfarztes zum Nutzen/Risiko-Verhältnis des Vorhabens einzureichen (§ 40, Abs. 1, Satz 4 AMG)

Bei Vorlage von Amendments sind Änderungen oder Ergänzungen deutlich zu kennzeichnen. Der Prüfarzt sollte die Protokolländerungen (aufgeteilt nach „wesentlichen" und „nicht wesentlichen" Änderungen) einzeln auflisten und mitteilen, ob die Änderungen nach seiner Ansicht ethisch relevant sind. Falls erforderlich, ist eine revidierte Patienteninformation/ Einverständniserklärung einzureichen.

Nach Publikation der Studie bittet die Ethikkommission um Zusendung eines Sonderdruckes.

Technische Universität München

Technische Universität München · Fakultät für Medizin · Ethikkommission
Ismaninger Str. 22 · 81675 München · Germany

Herrn
PD Dr. Tim M. Strom
Institut für Humanmedizin
Ismaninger Str. 22

81675 München

Fakultät für Medizin
Ethikkommission

Prof. Dr.
Georg Schmidt
Vorsitzender

Prof. Dr.
Franz B. Hofmann
Stellvertretender Vorsitzender

Geschäftsstelle
Ismaninger Str. 22
81675 München
Germany

Tel +49.89.4140.4371
Tel +49.89.4140.7737
Fax +49.89.4140.4199

info@ek.med.tum.de
www.ek.med.tum.de

21.02.2013
Projektnummer: 5360/13 (bitte bei jedem Schriftwechsel angeben)

Genomweite Sequenzierung bei genetischen Syndromen

Sehr geehrter Herr Dr. Strom,

hiermit bestätigen wir den Eingang der geänderten Elterninformation/-einverständniserklärung zu o.g. Studie.

Die Ethikkommission erteilt nun eine zustimmende Bewertung zu o.g. Forschungsprojekt.

Mit freundlichem Gruß

[redacted]

Prof. Dr. G. Schmidt
Vorsitzender

10 Acknowledgments

It is now time to thank the people who accompanied me on the way to my dissertation. Without them, this work would not have been possible.

First and foremost, I would like to thank my doctoral advisor Prof. Dr. Thomas Meitinger for accepting me into his research group and for his scientific and methodological expertise.

I would like to especially thank my mentor Dr. Matias Wagner for his constant support, time and advice and for allowing me to discuss all my questions with him at any time. He helped me to grow with new challenges and to successfully overcome obstacles. I would also like to thank him for all the encouraging words during the last years.

A special thanks also goes to Patrícia Espinheira de Sá Maciel and her team for the neurological examination and providing the DNA samples of the patients.

Furthermore, I would like to thank Professor M. Staudt, Dr. K. Weber and J. Anderl for the examination and DNA provision of individual patients.

I would also like to thank the entire team of the Human Genetics Department of the Klinikum Rechts der Isar for their ideas and advice.

My very special thanks, however, go to my family, who have made my path in life possible so far and to whom I dedicate this work.