Traditional and next-generation genetic methods in the molecular diagnosis of epilepsy

PhD thesis

ÁGNES TILL MD

Doctoral school leader:

Ferenc Gallyas MD PhD DSc

Supervisors:

Kinga Hadzsiev MD PhD **Béla Melegh MD PhD DSc**

University of Pécs

Medical School

Interdisciplinary Medical Sciences Doctoral School

Clinical Center, Department of Medical Genetics

2020, Pécs

1. INTRODUCTION

1.1. The history of epilepsy genetics and the new classification of epilepsies

The advent of the next-generation genetic methods has transformed the field of clinical practice of epilepsy genetics. Just two decades ago only 12 to 20 genes have been known to play a role in causing epilepsy. Today it has become a challenge for the specialists to keep pace with the rapid rate of ongoing gene discovery. The genetic basis of epilepsy has long been recognized as a major underlying cause for those with the condition. The recurrence risk of generalized epilepsy for first-degree relatives are increased 5-to 10-fold compared to the average population. Estimates of heritability based on twin studies further support a contribution of genetic variation. The increased familial risk has traditionally been thought to be the consequence of multifactorial inheritance. However, there are families segregating epilepsy as an autosomal dominant disease with high penetrance and a within-family concordance for seizure type. These extended, multigenerational families, led to the initial cloning of epilepsy genes by positional methods such as *SCN1A, CHRNA4, KCNQ2, KCNQ3* and *SCN2A.* Genome wide linkage analysis and Sanger sequencing were highly successful strategies in the context of a milder, dominant-acting epilepsy, but linkage studies can be of little use for the neonate-, lacking a positive family history, and with intractable seizures and poor long-term outcome.

Historically, up to 70% of epilepsy etiology was „idiopathic", but with the advent of widely available genetic testing, an etiology in subsets of epilepsy patients may be obtainable in more than 30%. The major scientific advances that have taken place in the last years-, have contributed to the updated classification of the epilepsies (International League Against Epilepsy, ILAE 2017).Updated ILAE classification emphasizes genetic classification while discouraging the use of the designation „idiopathic". The new classification incorporates etiology along each stage of diagnosis, emphasizing the need to consider etiology and comorbidities at each step of diagnosis, as it often carries significant treatment implications. Etiolgy is divided into six partly overlapping subgroups: structural, genetic, infectious, metabolic, immune and unknown. New terminologies are introduced such as developmental and epileptic encephalopathy.

1.2. The developmental and epileptic encephalopathies

The greatest success in epilepsy gene discovery has come from the study of the epileptic encephalopathies (EE). The term EE means that the epileptic activity itself contributes to severe cognitive and behavioral impairments above and beyond what might be expected from the

underlying pathology alone (e.g.,cortical malformation). Global or selective impairments can worsen over time. The concept of the EE may be applicable to epilepsies at all ages. Many epilepsy syndromes associated with encephalopathy have a genetic etiology, but aquired causes may also be in the background, such as hypoxic-ischaemic encephalopathy. It is suggested to use the term „developmental and epileptic encephalopathy" in those severe genetic disorders, which also have developmental consequences arising directly from the effect of the genetic mutation, in addition to the effect of the frequent epileptic activity on development. Often it may not be possible to clarify whether the epileptic or developmental component is more important in contributing to a patient's presentation. To keep it simple, I use the term EE in my dissertation. The EEs include many age-related electroclinical syndromes with specific seizure types and EEG features. Comorbidities are common, including autism spectrum disorder and behavioral and movement disorders. The outcome is often poor. With the molecular revolution, the number of known monogenic determinants underlying the EEs has grown rapidly. *De novo* dominant mutations are frequently identified; somatic mosaicism and recessive disorders are also seen. Several genes can cause one electroclinical syndrome, and conversely, one gene might be associated with phenotypic pleiotropy. Diverse genetic causes and molecular pathways have been implicated, involving ion channels, proteins needed for synaptic, regulatory and developmental functions. A genetic cause for an EE was first recognized in 2001, with the finding that all seven children in a study of Dravet syndrome (DS) had a *de novo SCN1A* mutation (type 1 alpha subunit of voltage-gated sodium channel).

Copy number variations are important molecular causes of EE, with up to 8-23,5 % of cases showing a causative or potentially contributing copy number variant (CNV). Chromosomal microarray studies searching for pathogenic CNVs, such as microdeletions and microduplications, are now a standard early investigation for all patients with EE. Chromosomal microarrays are also important for novel gene discovery. Increased efficiency and reduced cost of next-generation sequencing (NGS) technology have enabled different experimental designs to discover new EE genes. The whole exome sequencing (WES) of the proband and both parents (trio-WES) has proven to be helpful in the diagnosis of EE. In the majority of patients with EE the genetic background nowadays can be identified, we know that *de novo* mutations account for many sporadic cases. Less commonly, one or two affected relatives in the wider family suggest complex inheritance with a combination of risk variants across multiple genes, like in DS.

Traditional base by base sequencing (Sanger sequencing) is time consuming and targets one gene at a time but is highly accurate and generally less expensive per test ordered than new alternatives. NGS, or massively parallel sequencing allows for rapid sequencing of large numbers of DNA segments that are broken into small pieces, sequenced, and than realigned and analyzed computationally. NGS has made large gene panels, whole exome sequencing (WES), and even whole genome sequencing (WGS) possible. Gene panels sequence a list of genes known to be associated with a specific phenotype. As research increases our understanding, new genes are often added to the list. WES offers a broad evaluation for genetic variation by sequencing most of the protein-encoding exons and splice junctions in a patient's genome, as it is not limited to a list of genes known to be associated with a phenotype. The data obtained from NGS often includes many variants of unclear significance (VUS) that require interpretation but may not aid in the diagnosis. Parental samples are often required to further classify a VUS. WGS evaluates most of the DNA content of the entire genome but it is not clinically available at this time.

Despite our expanding knowledge, the steps from patient evaluation to genetic diagnosis remain complicated. Recent ILAE recommendations indicate that genetic evaluation should be undertaken in case of drug resistant epilepsy in patients of any age, including EEs also.

1.3. The genetic generalized epilepsies and their genetic background

For those individuals that present with generalized epilepsy, there are a number of well-recognized, non-syndromic, genetic generalized epilepsy (GGE) syndromes. These include childhood absence epilepsy, juvenile myoclonic epilepsy, juvenile absence epilepsy and epilepsy with tonic-clonic seizure alone. Patients may initially present with one GGE syndrome and later evolve into a second one. Individuals with GGE syndromes typically have normal intelligence or only mild impairment. The GGE-related genes may contribute to pathogenesis of GGE by an oligogenic or susceptibility inheritance mechanism. This is often the case with the more common, less severe group of epilepsies with an older age of onset. They often result from the multifactorial or complex inheritance of many liability genes with low-penetrance effects. Mutations in genes encoding T-type calcium channel subunits, such as *CACNA1H*, EF-hand domain-containing protein 1 (*EFHC1*), gamma aminobutyric acid (GABA) receptor subunits (*GABRG2, GABRA1*), chloride channel (*CLCN2*) and sodium channel genes (*SCN1A, SCN1B*) have also been shown to be responsible for GGE.

2. AIMS OF THE STUDY

The topic of the dissertation is the molecular diagnosis of epileptic patients using traditional and new-generation genetic methods. Our institute has a long tradition reaching back to decades ago examining and diagnosing children with developmental delay and intellectual disabilities. Given that epilepsy is a common comorbidity in this group of patients, I choose it as the topic of my research. Patients from all over the country visit our genetic outpatient clinic because nationally unique methods are available in our laboratory. Thus, we also see patients with rare diseases relatively often, so we have the opportunity to explore genotype-phenotype relationships more accurately. With the advent of the new-generation genetic methods our knowledge about the genetic background of epilepsy has expanded significantly, but the traditional cytogenetic and molecular genetic methods also have their significance in the diagnostic algorithm. Knowing of the advantages and limitations of the various methods, we can instigate an appropriate diagnostic work-up, which is cost-effective and does not lead to unnecessary testing. Exploring more accurate genotype-phenotype relationships, in addition to research aims, may ultimately provide patients and their families with more detailed knowledge about prognosis, the expected comorbidities and also possibly lead to therapeutic consequences.

We determined the following goals in the three groups of epileptic patients examined in our study:

1. Identification of breakpoints and exact genomic content of supernumerary marker chromosomes in patients with isodicentric (15) syndrome (idic (15) syndrome) diagnosed by conventional cytogenetic methods (karyotyping, fluorescence in situ hybridization) using array comparative genomic hybridization (array CGH) method. We searched for phenotype-genotype relationships, primarily regarding to the development of epilepsy. Analysis of the function of the affected genes and genomic regions using data from the scientific literature.

2. Investigation of the mutation spectrum of the *SCN1A* gene in the Hungarian population in case of DS and GEFS+ syndrome by Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA). We set out to explore genotype-phenotype relationships. Our aim was to investigate the segregation of the *SCN1A* mutations in the affected families.

3. We aimed to demonstrate the importance of WES testing in the GGE disease group. Interpretation of a variant classified as VUS.

3. PATIENTS

3.1. Patients with Idic (15) syndrome

In the first part of our research we performed accurate phenotypic analysis and additional genetic testing of five patients with idic (15) syndrome who were previously diagnosed by conventional genetic methods in our institute. Idic (15) syndrome is the most common genetic disease caused by a supernumerary chromosome with abnormal structure.

About half of the supernumerary marker chromosomes (SMCs) in humans are of 15 origins due to instability of 15q11.2-q13 genomic region. In the background of frequent rearrangements, there are five clusters of low copy repeats which are the basis of recurrent breakpoints known as BP1-BP5 being detected in the derivative chromosomes arising through different recombination events (non-allelic homologous recombination). U-type exchange is one of the crossovers which can result in a supernumerary isodicentric chromosome 15 (idic(15)) showing remarkable structural heterogeneity. The most frequently described breakpoints involved in large idic(15) are BP4 and BP5 with two extra copies of genomic regions between BP2-BP3 (partial tetrasomy). To determine the exact genomic content and breakpoints of marker chromosomes of 15 origin, we performed an array CGH assay on the samples of all five patients. The differences in genotypes were compared with the differences in phenotypes. From the review of the literature we attempted to explain the differences of the epilepsy syndromes of the patients, especially the presence or absence of their epilepsies.

3.2. Patients with Dravet syndrome and Generalized epilepsy febrile seizure plus syndrome

In the second part of our work we examined patients with fever-provoked epileptic seizures over a five year period (January 2012 and December 2017). A total of 183 Hungarian individuals were referred for genetic examination by neurologists and child neurologists. A total of 63 patients met the clinical and EEG diagnostic criteria of DS or GEFS+ syndrome. DNA samples were examined for *SCN1A* gene mutations using Sanger sequencing analysis and MLPA method. In case an *SCN1A* mutation was found, segregation analysis was performed to determine the mutation's *de novo* or inherited origin. The aim of this study was to investigate the mutational spectrum of the *SCN1A* gene in Hungarian patients with DS and GEFS+ syndrome phenotype.

3.3. Patient with genetic generalized epilepsy

In the third part of the study, we examined a paediatric male patient with normal intellect and an extended area of skin hypopigmentation suffering from generalized epilepsy displaying a switch in epilepsy syndrome during the course of the disease. Targeted genetic examination of the patient towards a neurocutaneous syndrome was unsuccessful. In the framework of a foreign collaboration, the sample of the patient was tested by WES. The segregation analysis of the variant interpreted by the laboratory as VUS and the interpretation of the results were performed in our institute.

4. METHODS

4.1. Taking and storage of samples

In the genetic counselling of the Department of Medical Genetics, Clinical Centre, University of Pécs, a detailed phenotype analysis of the patients participating in the study was performed; we provided genetic counseling to our patients/legal representatives of patients in accordance with the Human Genetic Law (XXI/2008). The patient/patient's legal representative has given written informed consent to perform genetic tests. After information and consent to the investigation, 2 ml of blood with Na-heparinate and/or 5-7 ml of blood with ethylene-diamine-tetraacetic acid (EDTA) anticoagulant was taken from each selected individuals for laboratory investigations. Samples were stored in a biobank after processing. The collection and usage of DNA samples and management of data during publication followed the Helsinki Declaration of 1975 and was in accordance with the Hungarian law (XXI/2008) for genetic examination, research and biobanking. The sudy degign was approved by the HRB National Ethics Committee.

4.2. Isolation of DNA

Genomic DNA was extracted from peripheral blood cells of the patients using the E.Z.N.A. Blood DNA Maxi Kit (Omega Bio-tek, USA) according to the protocol of the manufacturer.

4.3. GTG banding

Karyotyping was carried out by Giemsa–Trypsin (GTG) banding from peripheral blood lymphocytes using standard procedures. In order to exclude mosaicism, 100 cells were analyzed in each case.

4.4. FISH

The *UBE3A* locus (Prader-Willi/Angelman Critical Region) specific probe was applied for FISH examination containing controls in 15p11.2 and 15q22 regions, respectively (Vysis, Abbott Laboratories, Abbott Park, Illinois, U.S.A.). The protocol used was in accordance with the manufacturer's instructions.

4.5. Uniparental disomy

Uniparental disomy of chromosomes 15 was investigated in Patient 2. and 3. using polymorphic STR markers: D15S11, D15S122, D15S128, D15S210, D15S97, D15S113, GABRB3, D15S165, and D15S659, respectively. After PCR amplification of the markers, the

resulting products were separated on a polyacrylamide gel and detected by silver staining. The parents of the other three patients did not consent to the study.

4.6. Array CGH

Array CGH was performed using Agilent Human Genome Unrestricted G3 ISCA v2 Sureprint 8x60K oligo-array (Amadid 021924; it is a microarray with high resolution containing 18,851 60-mer oligo probes in ISCA regions (International Standards for Cytogenomic Arrays Consortium) and 40,208 backbone probes with an average 60 KB overall median probe spacing in coding and non-coding genomic regions, respectively) (Agilent, Santa Clara, CA) [13]. DNA was purified from peripheral blood using the NucleoSpin®Dx Blood DNA Purification Kit (Thermo-Fisher Scientific, Waltham, MA) according to the protocol of the manufacturer. NanoDrop spectrophotometer was applied for calculation of the concentration and purity of the isolated DNA. Labeling and hybridization of the samples was prepared according to the Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis – Enzymatic Labeling Protocol. The patient's DNA and a sex-matched reference DNA (1 μg from each) were digested with *Alu*I and *Rsa*I enzymes for 2.5 hrs at 37 °C. The digested DNA was labeled via random priming (Agilent Genomic DNA Labeling Kit; Agilent, Santa Clara, CA) using Cy5-dUTP for patient samples and Cy3-dUTP for control DNA, respectively. Purification after labeling was performed by Amicon Ultra AU-30 filters. The patient and reference samples with 50 μg Human Cot-1 DNA together were cohybribized at 65 °C for 24 hrs. Washing was performed following the instructions of Agilent Protocol v7.2. Array image was obtained by Agilent dual laser scanner G2565CA and analyzed with Agilent Feature Extraction software (v10.10.1.1.). Agilent Cytogenomics software (v2.5.8.11) was used for visualization of the results. DNA sequence information refers to the public UCSC (University of California, Santa Cruz, Genome Browser) database. The copy number variations detected were compared to known aberrations available in public databases like DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources), the Database of Genomic Variants, Clingen Dosage Sensitivity Map, ClinVar, and Ensembl (Ensembl GRCh37 Release 97 (July 2019)).

4.7. Automated Sanger sequencing

Amplification of exons in *SCN1A*, *TSC1*, *TSC2* and *SLC12A5* genes was performed by polymerase chain reaction (PCR) reaction, using exon-specific primer pairs designed in our department. The analysis of the PCR products was made by bidirectional Sanger sequencing on ABI 3500 Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA) automatic sequencer, using BigDye Terminator reagent. The resulting sequences were evaluated by

comparison with the corresponding reference sequence with application of the Winstar program (DNASTAR Inc., Madison, WI, USA). In the case of novel *SCN1A* variants, interpretation of the results was performed with the help of Mutation Taster, PolyPhen-2 and Mutation Assessor prediction softwares. The identified mutations were validated on a second sample obtained from the patients. All identified variants were classified by the standards and guidelines set by the American College of Medical Genetics and Genomics (ACMG) standards and guidelines.

4.8. Multiplex ligation-dependent probe amplification

The major rearrangements of the *SCN1A*, *TSC1* and *TSC2* genes were tested with MLPA method. MLPA was performed with the SALSA MLPA Kit P-137 Probemix (MRC-Holland, Netherlands), with the MLPA kit P124-C3 Probemix, (MRC-Holland) and with the MLPA kit P046-D1 Probemix (MRC-Holland), in accordance with the manufacturer's instructions. For each reaction, 50-250 ng of DNA sample was used in a reaction mixture of 5 µl. DNA was hybridized to specific probes for 16 hours, then, after ligation, PCR amplification was used in the presence of fluorescently labeled primers. The produced DNA fragments were separated by capillary electrophoresis. Coffalyser software (MRC Holland, Amsterdam, The Netherlands) was applied for evaluation of the results.

4.9. Next generation sequencing, whole exomsequencing

Whole exomeseqencing of the sample of the patient with GGE was performed at a genetic laboratory abroad (Centogene AG, Germany) as part of a research collaboration. Following fragmentation of the genomic DNA, the regions to be analyzed were enriched using Agilent SureSelect Human All Exon Kit V6. After library preparation, sequencing was performed on Illumina platform (coverage 100x). Bioinformatic processing of the raw data was made using CentoGene's in house developed programs (and algorithms), the reported data refer to GRCh37/hg19.

5. RESULTS

5.1. The phenotype and the structural differences in patients with idic (15) syndrome

5.1.1. Phenotype of the patients

The clinical features of our five patients are characteristic for idic(15) syndrome: central hypotonia in early age, developmental delay (DD), moderate to severe intellectual disability (ID), seizures, autistic behaviour, and absent or very poor speech, like in the vast majority of patients. However, the severity of the symptoms varies among the patients. Uncharacteristic facial dysmorphy and the absence of major malformations are also a common feature of our patients.

5.1.2. GTG banding and FISH

GTG-banding of the five patients (P.1. - P.5.) was performed at a 550-band level resolution. The karyotype of each patient contained a supernumerary acrocentric marker chromosome of G-group size. Based on the evaluation of 100 cells in each case, mosaicism could be excluded. Due to the clinical symptoms and the presence of the SMC, metaphase FISH analysis of the UBE3A locus (Prader-Willi/Angelman Critical Region) was performed in each case which showed the presence of both, the D15Z1 and UBE3A regions on the SMCs (in addition to the normal chromosomes 15), thereby confirming that the extra chromosome is of 15 origin.

5.1.3. Uniparental disomy

Uniparental disomy could be excluded only in Patient 2. and 3. The parents of Patients 1., 4. and 5. did not consent to the study.

5.1.4. Array CGH

The copy numbers and breakpoints of the different genomic regions in the SMCs 15 are presented in Table 1. (according to ISCN 2016). In addition, common benign variants were detected. The base pair positions of the genomic imbalances were designated according to the February 2009 Assembly (GRCh37/hg19).

Table 1.: **Molecular subtypes and the array CGH results of the idic(15) chromosomes**

Patient	Molecular subtype	Array CGH result
1.	A	arr [GRCh37] 15q11.1q13.2(20102541_30322138)x4
2.	C	arr [GRCh37] 15q11.2q13.2(22765628_31183907)x4, arr [GRCh37] 15q13.3(31261835_32861626)x3
3.	C	arr [GRCh37] 15q11.2q13.3(22765628_30178222)x4, arr [GRCh37] 15q13.1q13.3(30226187_32445252)x3
4.	D	arr [GRCh37] 15q11.1q13.3(20102541_30078386)x4, arr [GRCh37] 15q13.1q13.3(30251859_32510863)x3
5.	B	arr [GRCh37] 15q11.1q13.2(20102541_31077833)x4, arr [GRCh37] 15q13.2q13.3(31123186_33009483)x3

A. Large idic(15) – symmetrical breakpoints BP1-BP4:BP4-BP1.

B. Large idic(15) – asymmetrical breakpoints BP1-BP5:BP4-BP1.

C. Large idic(15) – asymmetrical breakpoints BP2-BP5:BP4-BP2.

D. Large idic(15) – asymmetrical breakpoints BP1-BP5:BP3-BP1

Based on the results of the array CGH we classified our patients into three molecular subtypes. The breakpoints of the Patient 2 and Patient 3 are identical. The breakpoints of Patient 1 are symmetrical, those of the others are asymmetrical.

5.2. Mutation spectrum of the *SCN1A* gene in the Hungarian patients with epilepsy

A total of 12 previously described *SCN1A* alterations (in 15 patients and three relatives) and 15 previously unknown pathogenic mutations (in 15 patients and two relatives) were identified by Sanger sequencing analysis. MLPA testing detected gross deletions of the *SCN1A* gene in three additional patients. Altogether, different types of *SCN1A* mutations were identified in 33 patients from our cohort. The inheritance patterns could not be determined in all cases as parental samples were not available in some families. The mutations proved to be inherited in six cases. The parents either have developed GEFS+ syndrome or suffered from febrile seizures in childhood. A couple of parents have remained unaffected.

Among the novel, previously undescribed *SCN1A* mutations 12 missense variants, two frameshift causing and one in-frame deletions were identified. Using the ACMG guidelines for the interpretation of sequence variants, 2 of 15 novel variants were classified as “pathogenic”, 12 were classified as “likely pathogenic” and one remained a variant of uncertain significance.

The diagnosis of DS was confirmed with MLPA method in three additional patients. Phenotypes of two of them were not significantly different from those with point mutations. On the other hand, the third patient with a large heterozygous deletion of exon 1-17 had an unexpectedly mild DS phenotype.

5.3. A rare form of ion channel gene mutation identified as underlying cause of generalized epilepsy

Examination of the DNA sample of a paediatric male patient with an extended area of skin hypopigmentation showing myoclonic-atonic seizures in infancy and switching to absence seizures in childhood proved to be negative when the *TSC1* and *TSC2* genes were analysed by Sanger sequencing and MLPA method in our laboratory. In the framework of a collaboration the patient’s DNA sample was sent to a foreign laboratory for WES testing. The WES identified a heterozygous c.1417G>A missense variant in the *SLC12A5* gene, resulting in a valine-isoleucine exchange at amino acid position 473 of the protein (p.Val473Ile). The detected variant is unknown in the literature, it is classified as VUS in the report of the laboratory because of the incongruent predictions attained from the different prediction softwares. The identified heterozygous missense mutation in a potassium chloride cotransporter gene together with the phenotype underscores the diagnosis of an epilepsy syndrome known in the literature as idiopathic generalized epilepsy type 14. Parental examination was performed in our laboratory. The mother proved to be of normal genotype for the mutation, while the asymptomatic father also carries the heterozygous c.1417G>A mutation in the *SLC12A5* gene.

6. DISCUSSION

6.1. The role of array CGH in the diagnosis of patients with idic (15) chromosome

A high prevalence of idic (15) chromosome, out of the SMCs, are detected by fluorescence in situ hybridization (FISH) that targets the Prader-Willi/Angelman region of idic (15) as the next diagnostic procedure. After karyotyping array CGH is a powerful method for detailed characterization of the genomic content involved in idic (15), possibly providing additional data for genotype-phenotype correlation, especially in cases with atypical clinical signs. There is a wide range of severity in the clinical signs experienced by individuals with idic (15) syndrome, our patients show this heterogeneity. Various genetic mechanisms have been hypothesized to explain this clinical heterogeneity. Only one of the five examined patients (Patient 1) has a SMC which shows a characteristic symmetrical structure with a single breakpoint. This mechanism of formation resulted in the presence of four identical copies within the segmental aneuploidy. Unlike the others, this patient showed no symptoms of epilepsy or autism up to 24 months of age. However, it must be pointed out that there is a possibility he may develop epilepsy and/or autism at a later age. Among our patients there are two subjects with identical, asymmetrical breakpoints (Patient 2. and Patient 3.) whose epilepsy started in infancy with epileptic spasms leading to the electro-clinical diagnosis of WS. The severity of their DD/ID is much more pronounced than in the other three cases with different molecular mechanisms.

Several theories have been published analysing the relationship between epilepsy observed in idic (15) patients and the genes involved in the supernumerary chromosome. Some of them highlighted the extra copy numbers of the affected genes, others the various mechanisms leading to altered gene expression, as well as imprinting as possible causes of seizures. The electro-clinical phenotype, course, and response to therapy of epilepsy are very heterogeneous in idic (15), despite the presence of the extra marker chromosome as common causative factor.

The possible role of the *CHRNA7* gene in autism and epilepsy has repeatedly emerged in the literature; however, its clinical significance is still debated. Analysis of cases with different sizes of microduplications involving *CHRNA7* - which encodes the alpha 7 subunit of nicotinic acetylcholine receptors - rather suggests that it might represent a risk factor for neurobehavioral disorders. With regard to our patients, *CHRNA7* was present in the idic (15) chromosome of all patients with the exception of Patient 1., the only one who has neither seizures nor autistic features. This patient has an idic (15) in which *CHRNA7* is not involved.

It appears, that the differences in the breakpoints and affected genomic regions do not provide sufficient explanation for the observed clinical heterogeneity. Earlier studies of similar patients raised the possibility that extra copies of the 15q11.2-13 region may cause an imbalance of homologous pairing of the alleles involved in segmental aneuploidy. This theory may lead to the identification of another mechanism in this complex region containing several imprinted and biallelically expressed genes leading to the formation of an abnormal phenotype. The cluster of three $GABA_A$ receptor subunit genes (*GABRB3*, *GABRA5*, and *GABRG3* which encode the receptor subunits β3, α5, and γ3, respectively) is of particular importance for neurodevelopmental disorders with epilepsy and autism because of GABA (Gamma amino-butiric acid) being the main inhibitory neurotransmitter in the brain. It is possible, that *trans* interactions between 15q11-13 paternal and maternal homologues may have significance in optimal biallelic expression of the genes within the $GABA_A$ receptor domain. It is conceivable, that the phenotype of the four patients reported here with both epilepsy and autistic features may be related to the asymmetric marker structure. Among the possible pathomechanisms leading to the symptoms, one should also consider that the lower copy number of a section of the marker interferes with the *trans* interaction of homologues originating from the two parents.

6.2. The phenotypic and genotypic heterogeneity of DS and GEFS+ syndrome

Most of the identified mutations proved to be missense mutations that probably alter but do not abolish the ion channel's function. Among the discovered mutations there were only two previously known, recurrent mutations that were identified in more than one patient in our cohort (p.Thr1174Ser and p.Arg1245*). In accordance with our data, previous observations also show that out of more than 1,200 reported *SCN1A* mutations, only 18% are recurrent. Apart from the 30 patients with *SCN1A* point mutations, MLPA method revealed three cases of *SCN1A* gene deletion. The frequency of MLPA-detected anomalies were 9,09 % in our cohort which is similar to that published in the literature. Based on literature data, the average frequency of MLPA-detected deletions and duplications is approximately 10-12% among *SCN1A*-mutation negative patients; therefore, we recommend this method as a second-tier of screening.

Phenotype exhibited large variability in our patient cohort, and we could not detect any strong correlation between genotype and phenotype. The phenotype might not necessarily be determined only by the SCN1A protein itself, but by a number of auxiliary proteins.

The early clinical diagnosis of DS may be difficult because the typical clustering of symptoms become apparent only during follow-up. An infant with prolonged febrile seizures and a confirmed *SCN1A* mutation has an *SCN1A* gene-related disorder. Due to the lack of consistent genotype-phenotype correlations, it is unpredictable whether the disorder may lead to the evolution toward GEFS+ syndrome or DS.

6.3. The role of WES in the genetic diagnosis of generalized epilepsy

The *SLC12A5* gene encodes the human neuronal KCC2 channel that is a major extruder of intracellular chloride in mature neurons. It is a member of the cation-chloride cotransporter gene family. Normal functioning of KCC2 results in low intracellular Cl^- concentrations in mature neurons which is essential for the development of adequate synaptic inhibition. It has long been hypothesized that defects in the *SLC12A5* gene and consequently changes in KCC2 protein expression and function are responsible for the inadequacy of the otherwise precisely regulated rapid postsynaptic GABAergic inhibition and reduced hyperpolarization associated with various neurological and psychiatric disorders (e.g., schizophrenia). However, to date only a few cases of monogenic disease due to mutations of the *SLC12A5* gene have been reported in the literature. In recent years the variants identified in the *SLC12A5* gene have been associated with two distinctive epilepsy syndromes: mutations of both alleles resulted in a severe infantile-onset pharmacoresistant epilepsy syndrome, *epilepsy of infancy with migrating focal seizures* (EIFMS), while mutation of one allele is expected to result in a much milder epilepsy syndrome, a GGE type. In vitro functional studies have shown a decrease in chloride efflux capacity of the mutated cells. The fact, that the mutation that can also be detected in the asymptomatic father of our patient can be explained by the incomplete penetrance and by the complex inheritance of the GGE syndromes.

7. CONCLUSIONS

The next-generation molecular genetic methods (NGS, array CGH) have revolutionized the genetics of epilepsy, they do not, however, replace traditional cytogenetic (karyotyping, FISH) and molecular genetic (Sanger sequencing, MLPA) methods but complement them. Most of the limitations of traditional and new methods can be eliminated by their parallel use. In addition to detecting new variants, the next-generation methods are suitable for confirming and refining the results of traditional technologies due to their high resolution. Conversely, variants that are identified by NGS are validated by conventional bidirectional (Sanger) sequencing in all cases. We performed our studies in two groups of diseases with severe EEs, in patients with idic (15) syndrome who presented with WS-LGS in almost all cases and in patients with GEFS+ - DS spectrum disease. In the third part of our research we turned out attention to the study of the background of GGEs, a group of disease that unlike the previous ones, is not a rare disease and affects a significant proportion of patients with epilepsy. Here I summarize our results topic by topic.

1. Idic (15) should be investigated in patients presenting with an epilepsy difficult to treat such as WS and LGS without a structural brain lesion, and in patients with early central hypotonia even in the absence of dysmorphic features. Karyotyping with FISH examination has diagnostic value in this disease. However, array CGH is a powerful method for detailed characterization of the genomic content involved in idic (15), even if the chromosomal aberrations were previously detected. Since this technique provides valuable information on the copy number of each section involved in idic (15), its use can contribute to the better understanding of the genotype-phenotype relationship. Thereby we can learn more about the molecular mechanisms of EEs. The exact knowledge of the marker's composition is important for the clinical management of the patient as well.

2. Our work is the first genetic study that investigates the mutations of the *SCN1A* gene in patients with GEFS+ or DS phenotype in the Hungarian population by Sanger sequencing and MLPA method. Targeted sequencing of the *SCN1A* gene revealed 15 novel point mutations in addition to the 12 previously described point mutations. Although close genotype-phenotype relationship could not be revealed, some important aspects of the disease spectrum were highlighted by our results. Some cases proved to be hereditary, and in these cases with familial occurence the phenotypic heterogeneity of the family members may be related to the complex inheritance and the disease-modifying effects of other genes. In three patients with DS a large

heterozygous gene deletion was confirmed by the MLPA method, which underscores the need to perform the study in cases where sequencing yields a negative result. Like others, we have found that the GEFS+ phenotype can occur as a *de novo* mutation and than passed on to the offspring to develop a clinical picture of DS. So we think, it is essential to perform a genetic test in patients and families with both phenotypes. Based on other authors and our own experience, we believe that the separation of GEFS+ and DS phenotypes is not always completely possible, especially at the onset of the disease, so it is more appropriate to use the term *SCN1A*-related epilepsy. Our future goal is to follow the patients with confirmed *SCN1A* mutation to clarify their phenotype and to get to know the natural course of the disease as thoroughly as possible. We would like to perform segregation studies in as many families as possible.

3. Our presented case is a good example of the fact that the GGE disease group is a complex inherited disease. Although the laboratory evaluated the variant in our patient as VUS, and it was also detected in the asymptomatic father, this does not exclude the role of the variant as a pathogen, as the disease-modifying effect of other genes must be taken into account. The „gold standard" to confirme that a variant is pathogen is to prove the changes in protein function (in this case a decrease in chloride efflux capacity), but we were not able to investigate this. Based on data from the literature (the role of *SLC12A5*-gene mutation in EIMFS and IGE confirmed by functional studies) and the phenotypic similarity, we believe that the variant, even if does not by itself causes, but contributes to the patient's GGE disease. Genetic counseling is very important in our case, given that homozygous and compound heterozygous mutations in the *SLC12A5*- gene are associated with a very severe early-onset EE. NGS techniques, including WES testing have diagnostic value in the diagnostic algorithm of epileptic patients.

8. LIST OF PUBLICATIONS

Publications related to the dissertation

Czakó M, **Till Á**, Szabó A, Ripszám R, Melegh B, Hadzsiev K. Possible Phenotypic Consequences of structural differences in idic (15) in a small cohort of patients. Int J Mol Sci 2019; 20: 4935.

IF:4.183 (2018)

Till Á, Zima J, Fekete A, Bene J, Czakó M, Szabó A, Melegh B, Hadzsiev K. Mutation spectrum of the *SCN1A* gene in a Hungarian population with epilepsy. Seizure 2020; 74:8-13.

IF:2.765 (2018)

Till Á, Szalai R, Hegyi M, Kövesdi E, Büki G, Hadzsiev K, Melegh B. Generalizált epilepszia hátterében azonosított ioncsatorna-génmutáció ritka formája. Orv Hetil 2019; 160 (21): 835-838.

IF:0.564 (2018)

Other publications in this topic

Düh A, **Till Á**, Bánfai Zs, Hegyi M, Melegh B, Hadzsiev K. Súlyos epilepsziás encephalopathia hátterében azonosított *MECP2*-gén-mutáció fiúbetegben. Orv Hetil 2019; 160 (51): 2036-2039.

IF:0.564 (2018)

Other publications

Zima J, Eaton A, Pál E, **Till Á**, Ito YA, Warman-Chardon J, Hartley T, Cagnone G, Melegh BI, Care4Rare Canada, Boycott KM, Melegh B, Hadzsiev K. Intrafamilial variability of limb-girdle muscular dystrophy, LGMD1D type (online megjelent, nyomtatás alatt). Eur J Med Genet 2019: 103655.

IF:2.022 (2018)

Hadzsiev K, Gyorsok Z, **Till Á**, Czakó M, Bartsch O. Rubinstein-Taybi syndrome 2 with cerebellar abnormality and neural tube defect. Clin Dysmorphol 2019; 28 (3): 137-141.

IF:0.760 (2018)

Fekete A, Hadzsiev K, Bene J, Nászai A, Mátyás P, **Till Á**, Melegh B. Két generációban megfigyelhető mitokondriális DNS A8344G mutáció. Orv Hetil 2017; 158 (12): 468-471.

IF:0.322

Till Á, Hadzsiev K, Lőcsei-Fekete A, Czakó M, Duga B, Melegh B. A 22-es csapdája? A 22q11 deletiós szindróma változatos klinikai megjelenése két eset kapcsán. Orv Hetil 2015; 156 (45): 1834-1838.
IF:0.291

Till Á, Rózsai B, Kovács N, Storcz J, Somogyvári F, Deák J, Ertl T, Decsi T. Újszülöttkori Cytomegalovírus-fertőzés kezelése ganciclovirrel. Orv Hetil 2006; 147 (13): 609-612.

Quotable abstracts

Till Á, Hadzsiev K, Szabó A, Bánfai Z, Melegh B, Berente Bene J, Czakó M. Co-occurrence of two recurrent copy number alterations in a child with complex phenotype. Eur J Hum Genet 2019; 27: 1-688.
IF: 3,650 (2018)

Zima J, **Till Á**, Ripszám R, Szabó A, Bánfai Z, Laufer Z, Hadzsiev K. Bannayan-Riley-Ruvalcaba syndrome due to a new *PTEN* gene mutation. Eur J Hum Genet 2019; 27: 1-688.
IF: 3,650 (2018)

Cumulative impact factor: 11,471

Cumulative impact factor with quotable abstracts: 18,771