Integrating Genotype-Phenotype Correlations into Precision Medicine

Towards Individualized Patient Management, Preventive Strategies, and Advanced Diagnostics in Mitochondrial Diseases

Ph.D. Thesis

Dr. Idris János Jimoh

University of Szeged

Doctoral School of Experimental and Preventive Medicine

Medical Genetics and Genomics program

Program Director. Prof. Márta Széll

Ph.D. supervisor: Prof. Dr. Mária Judit Molnár

1. Introduction

Despite the advances of modern medicine, both life expectancy and, in particular, the “health-span” have increased only modestly. This highlights the limitations of the traditional “one-size-fits-all” model of care. Conventional healthcare frequently overlooks individual variability, which can lead to suboptimal therapeutic responses, prolonged waiting times, and rising healthcare expenditures.

In contrast, precision and personalized medicine offer treatments and clinical management tailored to the unique biological characteristics of each individual. Genetics and genomics play a pivotal role in elucidating disease pathomechanisms, predicting drug responses, and supporting risk assessment. The clinical integration of genetic information improves the diagnostic yield for rare diseases and can also enhance risk stratification for multifactorial conditions or disorders associated with multimorbidity.

Across Europe, significant diagnostic delays persist in rare and monogenic diseases. The European Reference Networks (ERNs) aim to enhance diagnostic capacity and harmonize standards of care.

In Hungary, the Institute of Genomic Medicine and Rare Disorders at Semmelweis University was among the first national members of the ERNs, contributing to the EURO-NMD and ERN-RND networks.

Mapping patient pathways and assessing the healthcare impact of genetic diseases at the national level can support more effective resource planning. Structured data collection through patient registries and the integration of clinical and genomic datasets—providing real-world evidence—improve patient stratification, care management, and therapeutic development efforts. This approach is particularly relevant in mitochondrial disorders, where extreme genetic and phenotypic heterogeneity, along with biological complexity, make these diseases an exemplary model for demonstrating the potential and benefits of precision and personalized medicine.

2. Objectives

Mitochondrial disorders are associated with considerable clinical and genetic heterogeneity, with significant variability in both prevalence and phenotypic expression. The primary objective of our research was to investigate the prevalence and clinical variability of selected

mitochondrial disorders. To create patient cohorts, two model diseases were selected based on their relative frequency among mitochondrial disorders and the anticipated availability to identify a sufficient number of patients in Hungary to support meaningful analysis, with a solid foundation of well-characterized patient registries.

1. Evaluate the genetic epidemiology of SPG7-associated disorders in Hungary
2. Observe the phenotypic spectra of SPG7 and mt-TL1-associated conditions
3. Establish genotype-phenotype correlations (including zygosity and heteroplasmy levels) affecting penetrance, age of onset, and phenotypic variability
4. Assessing the therapeutic effectiveness of cariprazine based on its molecular effects in a pilot patient with MELAS syndrome
5. Identify underrepresented patient groups to improve diagnostics and patient pathways
6. Building an electronic health record (EHR)-based framework for creating trial-ready patient cohorts, registries and biobanks to support genotype–

phenotype correlation studies, as well as the discovery of clinical, digital, and molecular biomarkers

3. Methods

Ethical Statement: ethical approval required for the research was issued by TUKEB (44599-2/2013/EKU 535/2013) and SE RKEB (241/2024). All participants provided written informed consent.

Collection of Patient Data and Development of the EHR-based Data Structure: A retrospective and/or prospective analysis was conducted. Patients were pseudonymized and then registered in the ERN and GENOMIT registry.

Patient Inclusion Criteria: (A) SPG7 cohort: The baseline cohort included four major patient subgroups: classical HSP; cerebellar ataxia; motor neuron lesion; and a group presenting with a mitochondrial phenotype. (B) mt-TL1 cohort: The second patient group consisted of individuals diagnosed with the m.3243A>G variant at the Institute of Genomic Medicine and Rare Diseases. In one patient, a case report involving cariprazine therapy was published.

Clinical Assessments: Patients underwent detailed phenotypic evaluations, along with electrophysiological

tests (ENG/EMG), imaging studies (MRI and DAT-Scan), performed in selected cases.

Genetic Diagnostic Methods: The following methods were used in the genetic analyses: DNA isolation, long-range PCR, RFLP analysis, Sanger sequencing, targeted NGS panel, and whole-exome sequencing (WES). Identified variants were classified according to ACMG guidelines.

Myopathological Examinations: NADH, modified SDH, cytochrome-c-oxidase, and hematoxylin-eosin staining.

Statistical Analysis: Analyses were performed using Prism GraphPad V7.0b and SigmaPlot (2015).

4. Results

4.1 Results from the SPG7 screening cohort

SPG7 RDV was detected in 13.5% (50/437) of the cohort, excluding probands' relatives. RDV prevalence varied across four partly overlapping cohorts, with the highest diagnostic yield seen in the ataxia group[Figure-1]. Among patients carrying biallelic RDVs, 48% initially presented with paraparesis, 48% with ataxia, and 12% with visual impairment. For those with monoallelic RDVs, presenting symptoms included spasticity (36%), paraparesis (21.4%), ataxia (35.7%), visual complaints

(3.6%), and fasciculations or mitochondrial-related symptoms such as myalgia, fatigue, exercise intolerance, or PEO (14.2%).

4.1.1. Identified disease-associated SPG7 variants

In the SPG7 gene, 16 rare damaging variants (RDVs) were identified among 58 symptomatic individuals in mono- or biallelic forms. Based on ACMG criteria, 10 variants were classified as pathogenic or likely pathogenic. The most common RDVs were p.Leu78Ter (n=23) and p.Ala510Val (n=21)[Table1]. None of our patients harbored secondary AFG3L2 variants associated with digenic etiology.

Variant	Exon	Variant Type	CADD	GERP	ACMG	GnomAD 2.1 all AF	All affected cases / Families	Biallelic / Monoallelic cases	Bi / monoallelic cases with spasticity	Bi / monoallelic cases with ataxia	Bi / monoallelic cases with LMNL	Bi / monoallelic cases with mitochondrial dysfunction	Asymptomatic carriers
p.Leu78Ter	2	Nonsense	33	5.32	PATH	0.00039	23 / 18	14 / 9	11 / 7	14 / 9	1 / 0	2 / 2	1
p.Lys340Glu	8	Missense	25.9	5.85	VUS_LP	0	1 / 1	0 / 1	0 / 0	0 / 1	0 / 0	0 / 1	1
p.Gly344Asp	8	Missense	25.7	5.85	LP	0.00001	2 / 1	1 / 1	1 / 1	1 / 1	0 / 0	0 / 0	0
p.Gly352AlafsTer87	8	Frameshift	NA	5.85	LP	0	1 / 1	1 / 0	1 / 0	1 / 0	0 / 0	0 / 0	0
p.Gly352Ser	8	Missense	29.4	5.85	LP	0.00001	1 / 1	1 / 0	0 / 0	1 / 0	0 / 0	0 / 0	0
p.Val379Met	8	Missense	27.9	5.85	VUS_LP	0.00003	1 / 1	0 / 1	0 / 1	0 / 1	0 / 0	0 / 0	0
p.Arg398Ter	9	Nonsense	41	5.63	PATH	0.00002	2 / 2	2 / 0	1 / 0	2 / 0	1 / 0	2 / 0	0
p.Tyr406Cys	9	Missense	29.5	5.63	LP	0.00001	1 / 1	0 / 1	0 / 1	0 / 1	0 / 0	0 / 0	0
p.Arg486Gln	11	Missense	22.6	5.42	RISK FACTOR	0.004722	5 / 5	0 / 5	0 / 2	0 / 4	0 / 1	0 / 0	0
p.Gln507Ter	11	Nonsense	49	5.42	PATH	0	1 / 1	1 / 0	0 / 0	1 / 0	0 / 0	0 / 0	0
p.Ala510Val	11	Missense	27.8	5.42	PATH	0.00289	21 / 19	12 / 9	9 / 5	12 / 7	1 / 2	2 / 2	3
c.1552+1G>T	11	Splice	34	5.42	PATH	0.00002	4 / 4	2 / 2	1 / 1	2 / 1	1 / 1	0 / 0	0
p.Val540Met	12	Missense	21.6	5.47	VUS	0.00002016	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0
p.Ser645Thr	14	Missense	14.82	5.93	VUS	0.00076	2 / 2	0 / 2	0 / 1	0 / 2	0 / 0	0 / 1	0
p.Asn739fs	17	Frameshift	NA	5.59	PATH	0.00001	2 / 2	2 / 0	1 / 0	2 / 0	1 / 0	2 / 0	0
p.Tyr740Cys	17	Missense	25.7	5.59	VUS	0.00004599	1 / 1	0 / 1	0 / 0	0 / 1	0 / 0	0 / 0	0

Table 1 - Identified rare damaging variants in the cohorts

4.1.2 Genotype–phenotype correlations

The average age at onset in affected individuals was as follows: 41.12 (±9.97) years for patients with biallelic variants (39 ± 9.73 years for homozygous, 44.3 ± 9.95 years for compound heterozygous), 45.86 (±13.37) years for monoallelic heterozygous patients, and 51.6 (±19.65) years for patients with the risk allele. Asymptomatic relatives with monoallelic RDVs had a relatively young age (mean age 30.75 ± 2.9 years). The type and location of the SPG7 variant did not significantly affect the age of onset or the symptoms. However, homozygous and compound heterozygous patients appeared to differ in terms of clinical phenotype, particularly in the severity of spasticity, muscle strength, and SARA scores. Monoallelic patients exhibited similar symptoms to those with biallelic variants, but they tended to have a milder phenotype [Figure-2].

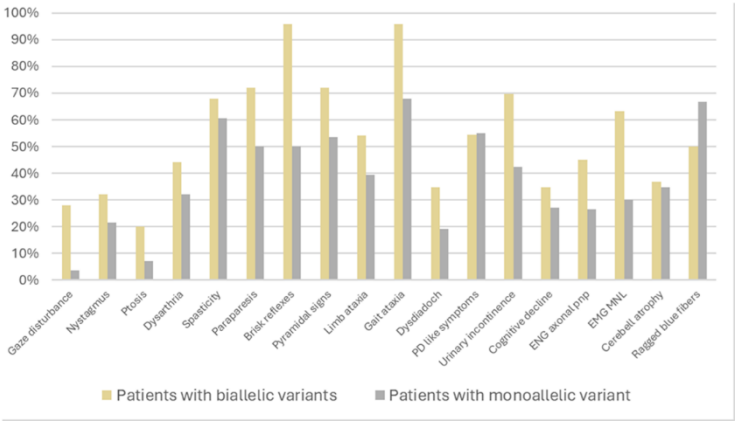

Figure -2 - Genotype–phenotype observations

The figure compares the frequency of SPG7-related symptoms between patients with biallelic and monoallelic non-risk factor RDVs.

Although multiple investigations have been conducted, we did not find a significant correlation between the mutation type, age of onset, or symptom severity. mtDNA deletion screening was performed on 52 blood and 6 muscle samples, revealing multiple deletions in 15.38% of blood samples (8/52) and 66% of muscle samples (4/6). Integrating histology and mtDNA data, mitochondrial dysfunction was identified in 33.3% of heterozygous and 16.0% of biallelic patients.

4.2 Results from the MT-TL1 m.3243A>G cohort

The mean age of the cohort at the last follow up was 41 ± 13.15 years. In addition to detailed phenotyping and clinical assessment, longitudinal follow-up data from 18 patients were analyzed over a mean period of 5.99 ± 4.41 years, including retrospective data. From this cohort, a single pilot patient was selected to assess the potential beneficial effect of Cariprazine in mitochondrial disease.

During the initial and follow-up clinical evaluations, all patient presented MELAS-associated symptoms. In 5 cases, childhood onset was observed. Based on our observations, HP showed significant effect on the onset. Spearman correlation: r -0,6126; 95% CI: -0.8123 to -0.2843; two-tailed P: <.001

4.2.1. Clinical phenotype of MT-TL1 m.3243A>G patients

Given the multisystemic nature of MT-TL1 m.3243A>G-related MELAS, individuals in our cohort exhibited a broad and heterogeneous spectrum of clinical manifestations, reflecting the systemic impact of mitochondrial dysfunction[Figure-3]. Sensorineural hearing loss (57.69%) and diabetes mellitus (42.31%) were the most common features. Stroke-like episodes occurred in 30.77%, and migraine in 15.38%. Cardiac

involvement affected 38.46% of patients, including hypertrophic (19.23%) and dilated cardiomyopathy (7.69%), as well as arrhythmias (7.69%). Myopathic symptoms were present in 30.77%. CIPO was diagnosed in 19.23%. Brain MRI frequently showed nonspecific T2 hyperintensities (80%). Neuropsychiatric assessment revealed high rates of anxiety (40%), cognitive complaints (35.71%), depression (33.33%), and less commonly neurodevelopmental disorders (13.33%) and schizophrenia spectrum disorders (13.33%).

4.2.2. The effect of HP on comorbidities

Patients with low heteroplasmy (10–25%) showed mainly adult-onset hearing loss and diabetes, without pediatric cases or severe multisystem disease. The intermediate group (25–30%) displayed broader involvement, including myopathy, cardiac abnormalities, and occasional stroke-like episodes, with earlier onset than the low-HP group. Moderate-to-high levels (30–35%) were associated with markedly more severe disease, including epilepsy, earlier cardiac symptoms, and childhood or adolescent onset. In the high heteroplasmy group (≥35%),

patients exhibited the full MELAS spectrum with frequent childhood onset and multisystem involvement. Overall, higher blood heteroplasmy correlated with earlier onset and increasing clinical severity.

4.2.3. Effect of clinical symptoms on diagnostic time

The mean diagnostic delay was 6.96 ± 6.81 years. Epilepsy led to the fastest diagnosis (3.0 ± 4.0 years; $p = 0.04$), while hypothyreosis (12.33 ± 10.02 years), cardiomyopathy (8.56 ± 8.38 years), and CIPO (7.8 ± 8.2 years) showed the longest, though not statistically significant, delays. Diabetes, myopathy, hypacusis, stroke-like episodes, and migraine were associated with intermediate intervals. Overall, the data highlight substantial variability across specialties and emphasize the need for greater mitochondrial disease awareness in endocrinology, cardiology, and gastroenterology.

4.3. Experience with Cariprazine in a pilot m.3243A>G patient

At age 27, the patient presented with symptoms and MRI findings consistent with Wernicke–Korsakoff syndrome, along with a family history later associated with the

diagnosis of MELAS. By age 32, the clinical picture was dominated by a severe psychotic syndrome with prominent negative symptoms and cognitive decline. Before treatment, the patient exhibited disorganized thinking, auditory hallucinations, blunted affect, apathy, psychomotor slowing, and poor self-care. The PANSS score was 164, and the ACE score was 81/100 (MMSE 29/30). Aripiprazole (7.5 mg twice daily) was initiated, but after one month without therapeutic improvement, treatment was switched to cariprazine (3 mg/day), chosen for its suitability in mitochondrial disease. This resulted in substantial clinical improvement without adverse effects. After three months, the patient returned to work, resumed social functioning, and showed cognitive gains. At six months, the ACE score improved to 86/100 and the PANSS score decreased to 59. No extrapyramidal signs, elevations in resting lactate, or other significant side effects were observed.

5. Conclusions and novelty of the thesis

The thesis establishes national epidemiological benchmarks for mitochondrial disorders in Hungary, including one of the world's largest SPG7 cohorts and a major regional MELAS cohort, highlighting substantial underdiagnosis.

It created previously nonexistent patient registries and demonstrates that combining structured registry data with unstructured data-lake sources yields robust real-world evidence. This framework contributed to ERN and GENOMIT databases, supporting Europe-wide diagnostic harmonization.

The thesis identifies clinically actionable genotype-phenotype correlations that refine genetic screening strategies, including SPG7 testing in motor neuron disease–like presentations without amyotrophy and m.3243A>G screening in isolated hearing loss, cardiomyopathy, or diabetes mellitus. It also expands evidence for dominant SPG7 variants and improves diagnostic algorithms.

Finally, it highlights that integrating more accessible diagnostic methods - such as myopathology and mtDNA deletion analyses - into the diagnostic algorithm significantly improves the effectiveness of patient evaluation, and that deep phenotyping may reshape

current inheritance models and pathogenic thresholds in SPG7 and m.3243A>G-related diseases.

6. List of publications

First-author articles:

Expanding the Phenotypic Spectrum of SPG7 Rare Damaging Variants: Insights From a Hungarian Cohort. Jimoh, I. J., et al. (2025). *Clinical Genetics*.

Wernicke–Korsakoff syndrome associated with mtDNA disease. Jimoh, IJ, et al. Therapeutic advances in neurological disorders 13 (2020): 1756286420938972.

A mitochondrialis betegségek diagnosztikai és kezelési lehetőségei 2021-ben Jimoh, IJ ; Molnár, MJ (2021) Orvostovábbképző szemle

Other relevant co-author articles:

Cornealis polymegathismus és retinalis pigmenthám-eltérések MELAS-szindrómában. István, L., Benyó, F., Csorba, A., Jimoh, IJ., Gál, A., Molnár, M. J., ... & Szabó, V. (2022). Szemészet, 159(2)

Broadening the phenotype of the TWNK gene associated Perrault syndrome. Fekete, B., Pentelényi, K., Rudas, G., Gál, A., Grosz, Z., Illés, A., Jimoh IJ., ... & Molnar, M. J. (2019). *BMC Medical Genetics*, *20*, 1-8.