의학박사 학위논문

엑솜 시퀀싱 분석을 통한 리 증후군에서의 원인 유전자의 발견

Whole exome sequencing-based discovery of causative genes in Leigh syndrome

2016년 8월

서울대학교 대학원

의학과 뇌신경과학 전공

이 진 숙

엑솜 시퀀싱 분석을 통한
리 증후군에서의 원인 유전자의 발견

지도교수 김 기 중

이 논문을 의학박사 학위논문으로 제출함

2016 년 **4** 월

서울대학교 대학원

의학과 뇌신경과학전공

이 진 숙

이진숙의 박사학위논문을 인준함

2016 년 **6** 월

위 원 장 _____ **(인)**

부 위 원 장 _____ **(인)**

위 원 _____ **(인)**

위 원 _____ **(인)**

위 원 _____ **(인)**

Abstract

Leigh syndrome (LS) is one of the most common mitochondrial disorders, which is characterized by various neurodegenerative features and bilateral symmetric central nervous system lesions. About 30% of LS have mitochondrial DNA mutations, while the remainder is thought to have nuclear DNA mutations. To date, mutations in more than 75 genes have been identified in both mitochondrial and nuclear genome. The aim of this study was to discover the causative nuclear genes in LS by using whole exome sequencing. We included 35 patients (34 families) who were diagnosed with LS at the Seoul National University Children's Hospital from 2001 to 2015. Diagnosis was based on characteristic clinical presentation and neuroradiological findings, which was supported by biochemical features suggestive of mitochondrial dysfunction. They had no mitochondrial DNA mutations identified. Pathogenic variants in genes previously shown to cause LS were identified in 14 patients (14/34=41.2%). The results showed the genetic complexity: *NDUFS1* (1), *NDUFV1* (1), *NDUFAF6* (2), *SURF1* (2), *SLC19A3* (2), *ECHS1* (2), *PNPT1* (1), *IARS2* (2), and *NARS2* (1). All the pathogenic variants identified in this study, except three, have never been reported. Of note, mutations in genes associated with mitochondrial translation were frequently observed (4/14=28.6%). Additionally there were two novel genes identified in 3 patients (3/34=8.8%), of which the functions have not been discovered in human: *APOA1BP* (1) and *VPS13D* (2). There was poor genotype-phenotype correlation in view of clinical, neuroradiological, and biochemical features, even after reviewing the reported cases

with the same causative genes. Our findings suggest that LS has a genetic heterogeneity, showing mutations in mitochondrial translation and valine metabolism as well as oxidative phosphorylation system subunits or assembly factors. Whole exome sequencing broadened the genetic spectrum of LS, showing the clinical utility to identify the causative genes in patients with LS. Discovery of molecular defects in Korean LS might prompt us to develop the gene panel for Korean LS.

Keywords: Leigh syndrome, mitochondrial disorder, mitochondrial DNA, nuclear DNA, whole exome sequencing

Student Number: 2007-30568

Contents

Abstract ---------- i

Contents ---------- iii

List of tables ---------- iv

List of figures ---------- v

List of abbreviations ---------- vi

Introduction ---------- 1

Materials and Methods ---------- 3

Results ---------- 17

Discussion ---------- 29

Conclusion ---------- 35

References ---------- 36

국문초록 ---------- 40

List of tables

Table 1. Exome sequencing metrics for all patients ---------- 10

Table 2. PCR primers for the causative variants of LS ---------- 12

Table 3. List of assay reagents used for mitochondrial respiratory chain enzyme complex analysis ---------- 15

Table 4. The clinical and molecular findings of all patients in this study ---------- 22

Table 5. Pathogenicity of causative variants in known genes of LS ---------- 25

Supplementary Table 1. Variants list of all patients identified with analysis pipeline based on the recessive inheritance ---------- 26

Supplementary Table 2. Only one heterozygous variant list of LS candidate gene-28

List of figures

Figure 1. Whole-exome sequencing and analysis ---------- 8

Figure 2. Variants-filtering analysis pipeline ---------- 9

List of abbreviations

LS – Leigh syndrome

NGS – next generation sequencing

OXPHOS – oxidative phosphorylation

MRC – mitochondria respiratory chain

MRI – magnetic resonance imaging

MRS – magnetic resonance spectroscopy

mtDNA – mitochondrial DNA

Introduction

Leigh syndrome (LS) is one of the most common mitochondrial disorders, which was first described in 1951 as subacute necrotizing encephalomyelopathy by Dr. Denis Leigh[1]. LS typically presents with various neurodegenerative symptoms before 2 years of age and rapidly progresses to die before 5 years of age[2-4]. It is characterized by bilateral symmetric central nervous system lesions neuroradiologically. However, more patients with atypical presentations expanded the clinical spectrum since the first report of LS. The genetic diversity in LS has been also revealed with the recent advances in next generation sequencing (NGS). Genetic heterogeneity, combined with clinical heterogeneity and overlapping features, complicate the molecular diagnosis of LS[5,6].

Mitochondrial oxidative phosphorylation (OXPHOS) system is controlled by dual genome, mitochondrial and nuclear genome[6]. About 30% of LS have mitochondrial DNA mutations, while the remainder is thought to have nuclear DNA mutations[6,7]. To date, mutations in more than 75 genes have been identified in both mitochondrial and nuclear genome[5]. Previous genetic testing for nuclear mutations was performed gene by gene, based on the clinical findings and mitochondria respiratory chain (MRC) enzyme assays, which made it difficult to identify molecular etiology. Currently, high-throughput NGS approaches can provide a cost-effective analysis for mitochondrial disorders[8,9]. Whole exome sequencing can identify not just known nuclear genes of LS, but also novel genes which have not been reported to be causative of LS or newly discovered genes in human. Although common causes of

LS include genes encoding OXPHOS subunits and assembly factors, there have been cases with defects in genes for mitochondrial translation reported recently[5,6,10]. The aim of this study was to discover the causative nuclear genes in LS by using whole exome sequencing and determine a correlation between genotype and phenotype. The results may broaden/expand the genetic spectrum of LS, help understanding the pathophysiology, and in future be helpful for development of gene panel of LS.

Materials and Methods

Patients

We recruited 35 patients (34 families) with LS who did not have a molecular diagnosis, including two affected siblings from one family. They were diagnosed with LS at the Seoul National University Children's Hospital from 2001 to 2015. The diagnosis was established based on the following characteristic clinical features and brain magnetic resonance imaging (MRI) findings: (1) neurodegenerative symptoms, including encephalopathy, psychomotor retardation, hypotonia or spasticity, dystonia, dyskinesia, seizures, ataxia, and brainstem dysfunctions such as dysphagia and ptosis; and (2) bilateral symmetric lesions in one or more areas of the central nervous system, including the basal ganglia, brainstem, thalamus, and cerebellum. The diagnosis was further supported by biochemical features, including MRC defects and increased lactate levels in serum or cerebrospinal fluid. Analyses of arterial blood gas, serum amino acids, urine organic acids, carnitine profile, tandem mass screening, serum copper and ceruloplasmin, and general laboratory tests, including serum ammonia, were performed to exclude other causes of metabolic disorders. None of the patients had any history of perinatal asphyxia or kernicterus. Additionally we performed a thorough medical evaluation to exclude Wernicke encephalopathy, acute necrotizing encephalopathy, Reye syndrome, acute disseminated encephalomyelitis, other types of encephalitis, vasculitis, cerebral infarction, and pantothenate kinase-associated neurodegeneration.

Of the patients, patients with mitochondrial DNA mutations identified were excluded from this study. Mitochondrial genetic analysis was performed with targeted and/or whole mitochondrial genome sequencing in our previous study[11].

This study was approved by Seoul National University Hospital institutional review board (IRB No.1406-081-588), and blood and muscle samples were obtained from enrolled patients whose parents provided informed consent.

Genetic Analysis

We performed whole exome sequencing of the proband in 33 patients and trio (proband and parents) in 2 patients. Genetic variants were validated using Sanger sequencing. Segregation studies were done with parental DNA samples whenever possible.

Whole exome sequencing and Analysis

After preparation of genomic DNA, exome capture was performed using SureSelect^{XT2} Human All Exon v4 + UTRs kit (Agilent Technologies, Santa Clara, CA, USA) with capture size 70 Mb. Mitochondrial genome was not included in the capture array. Sequencing was performed by Illumina HiSeq 2500 with 100 paired-end reads (Illumina, San Diego, CA, USA). Exome run quality (metrics) for all samples of the patients was summarized in Table 1. Then, the reads were aligned to GRCh37 by BWA (Burrows-Wheeler Alignment) v.2.2.0 and PCR duplicates were

removed by SAMtools v.1.9.0. Variants were called by SAMtools v.1.9.0. Quality-based filtering was done (QS>80).

The identified variants were annotated based on novelty, impact on the encoded protein, evolutionary conservation, and expression using the in-house automated pipeline. PhyloP score and amino acid conservation were extracted from the UCSC genome annotation database and clustalW. Among the called variants, the common variants that are listed in public databases (dbSNP build 137; 1000 Genomes Project release 10.31.2012; NHLBI Exome Sequencing Project; ExAC) were excluded (minor allele frequency > 0.2%), and only rare variants were considered as potential causative variants. *In silico* mutation prediction programs such as SIFT, Polyphen-2, and MutationTaster as well as conservation among different species were used to access the pathogenicity of novel variants (Figure 1).

Autosomal recessive model analysis was first done to search for homozygous or compound heterozygous variants based on the expected inheritance pattern in LS (Figure 2). Then, if no homozygous or compound heterozygous variants were identified, we searched for one heterozygous variant on autosomes or hemizygous variants on the X chromosome, which are known to cause LS or other mitochondrial disorders. Then we tried to discover the second allelic variants for the cases with only one variant identified, by performing Sanger sequencing for the poorly or uncovered regions in whole-exome sequencing (WES). Additionally the analysis for loss of heterozygosity or copy number variation was done, with an effort to discover any homozygous deletions.

Sanger sequencing

Direct Sanger sequencing of compound heterozygous mutations of each gene was performed by standard methods following polymerase chain reaction (PCR) amplifications using specific primers (Table 2). The PCR condition included one cycle of pre-denaturation at 95°C for 2 min, 35 cycles of denaturation at 95°C for 20 s, annealing 59°C for 40 s, elongation at 72°C for 45 s, and one cycle of post-elongation at 72°C for 6 min. Sanger sequencing reactions were run on an ABI-3730XL DNA analyzer (Applied Biosystems). Sequence variations were analyzed via comparison with the wild-type sequence. The mutation nomenclature followed the recommendations of the Human Genome Variation Society. Segregation analysis of pathogenic variants was performed by sequencing parental DNA samples when possible.

Orthologs

Full-length orthologous protein sequences in vertebrate were identified by a BLAST search of the LS candidate genes in human and extracted from GenBank. Orthologs were confirmed based on BLAST searches of the protein sequence against the human protein sequence, with the requirement that the LS candidate genes in human be the top hit and protein sequences were aligned using the ClustalW algorithm.

Interpretation of sequence variants

According to the ACMG guidelines in 2015 (rules for combing criteria to classify sequence variants), we classified the variants as pathogenic, benign, or uncertain significance[12].

Medical record review

Medical records were reviewed with a focus on the age of onset, age of death, initial manifesting symptom or sign, family history, and clinical course of the disease. Initial and serial follow-up MRI findings were reviewed. Biochemical features were also reviewed. A lactate peak in the brain magnetic resonance spectroscopy (MRS) was assessed. Lactate and pyruvate levels were measured in the serum and/or cerebrospinal fluid. As described in the previous study, the activities of mitochondrial respiratory chain complexes in the supernatants from muscle homogenates were measured using a spectrophotometer (DU-730, Beckman Coulter, Fullerton, CA, USA) (Table 3)[11]. The activity was expressed relative to that of citrate synthase, which is a marker of mitochondrial mass. A reduction in enzyme activity below 15% was considered to be a significant decrease in activity.

- **1st step:**

sample preparation, hybridization, sequencing

- **2nd step:**

mapping, variant calling, annotation

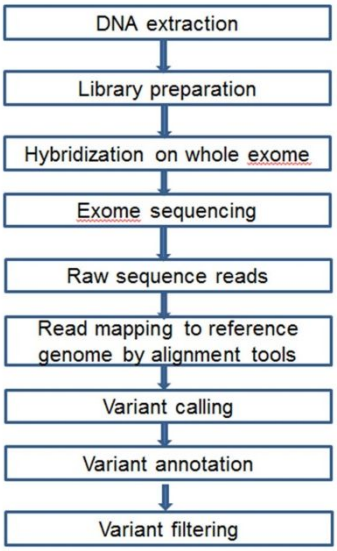

Figure 1. Whole-exome sequencing and analysis. After high-throughput sequencing of DNA coding regions, WES data analysis comprises aligning raw sequencing reads to reference genome sequence, calling, and filtering of variants, using the analysis algorithm.

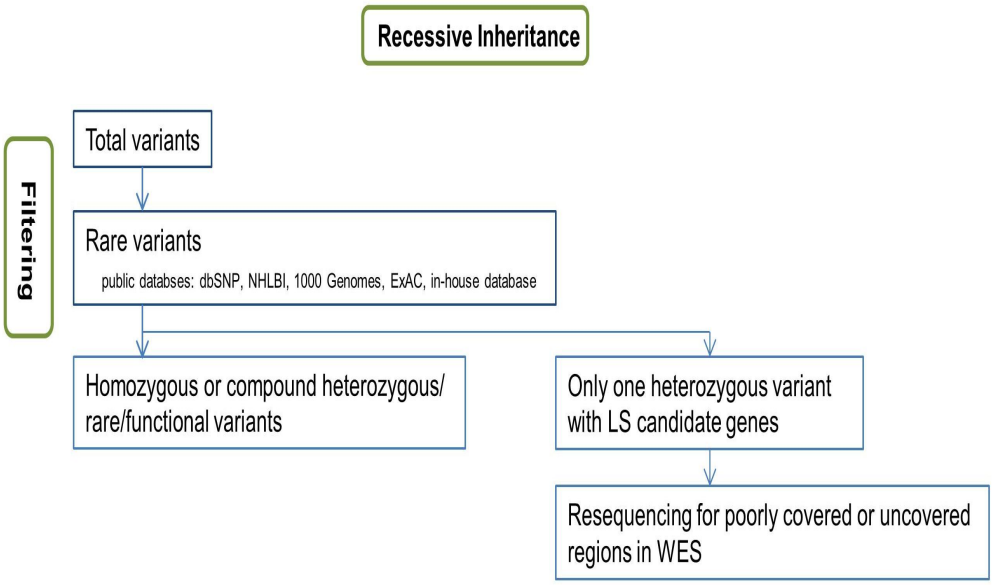

Figure 2. Variants-filtering analysis pipeline. Autosomal recessive model analysis was first done to search for homozygous or compound heterozygous variants. In cases with one heterozygous variant identified, resequencing for poorly covered or uncovered regions in WES was performed.

Table 1. Exome sequencing metrics for all patients

sample ID	read number	average read length(bp)	coverage depth	% of reads on target	% of bases covered at least 8x	mean error rate
Case 1	111,263,512	90	77.77x	64.11%	97.27%	0.26%
Case 2	150,506,268	90	98.58x	60.84%	98.19%	0.26%
Case 3	126,325,602	90	87.66x	65.83%	98.05%	0.26%
Case 4	103,139,932	90	72.23x	67.47%	96.21%	0.22%
Case 5	101,716,872	90	71.35x	66.01%	97.43%	0.26%
Case 6	115,055,922	90	80.56x	65.11%	97.49%	0.23%
Case 7	102,157,304	90	72.6x	66.94%	97.03%	0.25%
Case 8	148,187,864	90	97.51x	61.36%	98.22%	0.26%
Case 9	132,442,072	90	90.29x	63.70%	97.99%	0.26%
Case 10	126,037,812	90	85.4x	62.62%	97.85%	0.26%
Case 11	95,740,966	90	68.17x	66.49%	96.92%	0.25%
Case 12	119,727,630	90	84.42x	66.11%	97.82%	0.23%
Case 13	158,748,400	90	109.17x	66.33%	98.78%	0.25%
Case 14	115,451,066	90	75.59x	61.67%	97.52%	0.26%
Case 15	143,343,344	90	93.62x	62.96%	98.18%	0.24%
Case 16	77,147,276	90	54.88x	66.59%	95.59%	0.25%
Case 17	99,108,126	90	72.83x	69.48%	96.96%	0.25%
Case 18	120,514,074	90	83.11x	67.06%	97.50%	0.22%
Case 19	119,895,572	90	83.32x	66.81%	97.46%	0.23%
Case 20	103,295,068	90	76.17x	70.65%	97.54%	0.25%

Case 21	78,369,032	90	56.72x	67.43%	95.77%	0.25%
Case 22	97,392,792	90	65.93x	64.45%	95.48%	0.24%
Case 23	94,254,746	90	66.56x	65.41%	96.58%	0.25%
Case 24	87,226,368	90	63.27x	67.04%	95.87%	0.24%
Case 25	87,226,368	90	74.8x	65.27%	97.40%	0.25%
Case 26	124,005,264	90	77.89x	62.84%	97.64%	0.24%
Case 27	106,874,590	90	75.5x	67.74%	96.91%	0.22%
Case 28	143,149,726	90	94.23x	61.62%	98.14%	0.26%
Case 29	142,113,175	90	123.80x	67.49%	98.78%	0.29%
Case 30	130,894,534	90	91.49x	65.01%	97.86%	0.23%
Case 31	111,785,458	90	78.46x	66.50%	96.61%	0.23%
Case 32	119,909,868	90	82.07x	66.39%	97.42%	0.23%
Case 33	119,909,868	90	64.42x	60.70%	96.80%	0.26%
Case 34	69,635,830	100	80.6x	78.23%	98.89%	0.16%
Case 35	112,845,706	101	103.76x	69.81%	98.75%	0.22%

Table 2. PCR primers for the causative variants of LS

Gene		Nucleotide Sequences Forward	Reverse	Product size (bp)
NDUFS1				
	207012294	GGGAGGTGAATGTGATCTGC	aacagcatccctcttctccc	368
	207009734	gcaccatagtaaacaccgtatg	gcaactcagattccagtagtct	294
PNPT1				
	55894157	aggttgagaacttcctgggag	caccatatgacaaaacctacagc	292
	55906850	agatggggttgaagccatga	ggctaccgacaaaacatacaca	299
NDUFAF6				
	96044290, 96044247	tgtgcagcttggaaagtgaa	agcagggaaaagaaatcactagc	290
	96064405	ttccccttaccacagagctg	caatctggctgcaaagtccg	571
	96070035	agaaaaccatctcagggaaact	TGTGAGCTACTTGACATCCCA	383
APOA1BP				
	156563742	cttcctgaacaccaccctct	AGAATACCCACCCACCTTCC	250
	156562233	gagagacagctgggccaata	ttcaaccccagatcccagag	246
SLC19A3				
228563833, 228563982, 228564166		tgcttcgtggtcatgcaatt	AAGCCCACCATAGAGACCAG	743

SURF1				
	136218824	cttcctccccttcagcctag	CCAGTAGCACATGATCCAGC	203
	136220750	gctcggccactgttcttatc	gctcccacatgtcctactca	242
VPS13D				
	12382648	ggatgtaaaaccaatgcagct	agcaagactgagtgggtttt	302
	12428569	tggatggcatcttgtctcct	accctcccacacatgaaaca	232
IARS2				
	220273991	GCAGATTTTGAAATACATGGC	TTCTGACTATTTCAAGGACCTGAG	293
	220310171	cagtgagccgacatcatgc	gaccacaatcccagcctct	500
	220279361	acaagttaaaggggaagctgg	acacaaacctgagcatttcaca	444
	220311261	agcccgcatttcattacagt	tggccaacacaggagattct	422
ECHS1				
	135179506, 135179512	tcactcagcactcaggaagg	tcatttggagccagaaagcc	380
	135183498	ataggatgggtgggctctct	tccttcaaccacgagaggag	231
	135180470	aagccccttggaatttgctg	cacgagggagacaactgg	379
NDUFV1				
	67379663	tttaacctcctccccaccac	CGAAAGTGGCGGATCAGAC	371

	67378498	cctagcagccaccagttct	ttagtcagcccctagaccca	364
NARS2				
	78204200	ggacaagacagaggcacaca	aacacataaacgtccgaggt	330
	78147799	ggtctatcacagtgaggtaaaca	GGGGGTGCTTTTCCTTAACC	299
	78279749	agggcaattgtacaacctctt	cacacaccacagagatgcaa	600
	78180291	aggagtacgctttttctggtt	acccttttagtcactgctattcca	396

Table 3. List of assay reagents used for mitochondrial respiratory chain enzyme complex analysis

Complex	Assay reagents
I	1 ml of 250-mM potassium phosphate, 50-mM MgCl2, pH 7.2 1 ml of 25-mg/ml BSA 1 ml of 13-mM NADH 20 ml of 1-M KCN 20 ml of 1-mg/ml antimycin A 10 ml of 65-mM ubiquinone1
II	1 ml of 250-mM potassium phosphate, 50-mM MgCl2, pH 7.2 200 ml of 1-M sodium succinate 100 ml of 5-mM DCPIP 20 ml of 1-M KCN 20 ml of 1-mg/ml antimycin A 20 ml of 1-mg/ml rotenone
I+III	25-mM potassium phosphate buVer, 5-mM MgCl2, pH 7.2 2.5-mg/ml BSA 2-mM KCN 2-mg/ml rotenone 0.6-mM n-dodecyl-b-D-maltoside 15-mM cytochrome c (oxidized) 35-mM ubiquinol2
II+III	1 ml of 250-mM potassium phosphate, 50-mM MgCl2, pH 7.2 200 ml of 1-M sodium succinate 20 ml of 1-M KCN

	10 ml of 1-mg/ml rotenone
IV	20-mM potassium phosphate, pH 7.0 0.45-mM n-dodecyl-b-D-maltoside 15-mM cytochrome c (reduced)

Results

Totally 56 patients from 54 families (including two families with each two siblings) were diagnosed with LS during the study period. Mitochondrial genome sequencing was performed in all patients, which revealed mitochondrial DNA (mtDNA) mutations in 17 patients of 16 families, including two siblings in one family (16/54=29.6%). They included *MT-ND3* (5): T10158C (1), T10191C (1), G10197A (2), G10197C (1); *MT-ND5* (2): G13513A (2); *MT-ND6* (4): G14459A (2), T14487C (2); *MT-ATP6* (5): T8993G (2), T8993C (2), T9176C (1). Additionally we performed the *PDHA1* sequencing in the patients who had the typical clinical features and laboratory findings with lactate/pyruvate ratio less than 20. Two patients with *PDHA1* mutations were identified, who were all males (2/54=3.7%).

After excluding two patients whose DNAs were not available, we performed whole exome sequencing in 35 patients from 34 families (including two siblings in one family). In two cases, we performed trio exome sequencing in the probands and their parents. The mean depth for exome sequencing in all patients was 80.7X. Mean read numbers 114.03 Mb, mean read length 90.6 bp, mean % of reads on target 65.7%, mean % of bases covered at least 8x 97.38%, mean error rate 0.24%. Detailed metrics of exome sequencing were summarized in Table 1.

As shown in Table 4, pathogenic variants in genes known to cause LS were identified in 14 patients. They included *NDUFS1* (1), *NDUFV1* (1), *NDUFAF6* (2), *SURF1* (2), *SLC19A3* (2), *ECHS1* (2), *PNPT1* (1), *IARS2* (2), and *NARS2* (1) (See Supplementary Table 1. for all the variants identified in 35 patients based on the autosomal recessive

model analysis). Among 14 patients with the variants, parental tests were performed in 7 available families. According to the ACMG guidelines for the interpretation of sequence variants, we classified the variants as pathogenic or likely pathogenic. Genetic variants identified in twelve patients were pathogenic, while the variants of two patients were likely pathogenic (Table 5). These causative genes identified in the present study have been reported to cause LS or be associated with LS, while all variants except three (*ECHS1*: c.713C>T, p.A238V; *SURF1*: c.845delCT, p.S282Cfs*9; *SURF1*: c.54+1G>T) were novel (Table 5). Of note, mutations in genes associated with mitochondrial translation were frequently observed in the present study. They all were the reported genes identified by whole exome sequencing in LS or other mitochondrial disorders recently. Additionally, novel genes were identified in 3 patients (3/34=8.8%): *APOA1BP* (1) and *VPS13D* (2). They were meaningful because they might be possibly pathogenic variants through in vitro studies in future. *APOA1BP*, the gene associated with mitochondrial function, encodes apolipoprotein A-I-binding protein, which is included in the MitoCarta[13]. *VPS13D* variants were identified in two unrelated patients with LS, although the function of the gene encoding vacuolar protein sorting 13, yeast, homolog D, is unknown in human. It has scores of 4.7513 in MitoCarta and false discovery rate of 6.4%, which means that the gene can be associated with mitochondrial disorders. With the ACMG guidelines for the pathogenicity of the sequence variants, *VPS13D* variants were variants of unknown significance, while *APOA1BP* variants were likely pathogenic. Other variants included *DUOX2*, *QRFPR*, *TAF1L*, and *UNC13C* (Supplementary Table 1), although the possibility of functions associated with mitochondria is very low.

UNC13C variants were identified in two siblings in one family. However, the functions of the gene remained unknown in LS.

We searched for the cases with one heterozygous variant which are known to cause LS or other mitochondrial disorders and predicted to be damaging. Six cases were found to carry one heterozygous variant in *ECHS1*, *NDUFAF6*, *NDUFS7*, or *SURF1* (Supplementary Table 2). After we identified the coverage depth of the candidate gene in each patient, we performed Sanger sequencing for the exonic regions which were poorly covered in whole exome sequencing. Through this approach, we identified other *SURF1* variant in another allele in one patient (Case 32), which had only one read in exome sequencing.

The clinical and molecular findings of the 35 patients are summarized in the Table 4. They included male 15 and female 20. The mean/median age of onset was two/twelve months (range, one month to 21 years). Family history of LS or other mitochondrial disease was present in four patients (4/34=11.8%). Fifteen patients were lost to follow-up and two patient died during follow-up at 17 and 20 months, respectively. The mean/median age in the surviving eighteen patients was 10.3 and 9.5 years of age (range, 14 months to 24 years). There was no correlation between age of onset and the presence of pathogenic variants. The most common presenting symptom was developmental delay (14, 14/35=40%). Other initial manifestations included developmental regression, ataxia, irritability, altered consciousness, and seizures. Clinical features also included non-neurological symptoms such as cardiomyopathy, ophthalmoplegia and nephropathy.

Initial brain MRI findings were available for all patients. Most of them revealed the

typical findings of LS, involvement of the basal ganglia (globus pallidus, putamen, caudate nucleus), thalamus, white matter, midbrain, brainstem, and cerebellum. MRS was performed for 23 patients; a lactate peak was present in all cases except one. One case (Case 27) with LS/MELAS overlap showed multiple cerebral cortical infarction. There were two cases which had bilateral symmetric lesions in basal ganglia and other areas initially and extended to progressive leukodystrophy (Case 11 with *SURF1* variants and Case 25 with *NDUFV1* variants). There was another case (Case 6) with unilateral basal ganglia infarct initially to bilateral symmetric basal ganglia lesions during follow-up. Additionally there was one case in which nonspecific brain atrophy initially developed when the presenting symptoms such as developmental delay with poor eye contact and eyeball roving movements developed at 3 months of age, and persisted until 3 years of age, thereafter basal ganglia lesions developed after 6 years of age (Case 2 with *PNPT1* variants). Of note, there were two patients with compound heterozygous *SLC19A3* variants. They showed difference in clinical severity. Case 10 presented with infantile spasms at the age of 4 months, who also had developmental delay and involuntary movements. At his current age of 29 months, he had severe psychomotor retardation and medically intractable seizures, showing significant brain atrophy with lesions in caudate nucleus, basal ganglia, thalamus, brainstem, and subcortical regions. On the other hand, Case 15 presented with irritability and developmental delay at the age of 5 months. During follow-up, she showed ataxic gait and dystonia with seizures. Her neuroimagings at the age of 20 years showed lesions in caudate nucleus, basal ganglia, and cortical regions, including sulcus prominence and gyral atrophy in cerebral

hemisphere. At her current age of 24 years, she could walk without assistance and her seizures were well controlled with one antiepileptic medication. Neuroradiological findings of two cases with *SLC19A3* variants were not so different from those of other patients with LS.

Biochemical findings with evidence of mitochondrial dysfunction were found in all patients, which were supportive of diagnosis of LS. They included increase in serum or cerebrospinal fluid lactate and/or lactate peaks in MRS. Although the mitochondrial respiratory chain enzyme activities were measured with spectrophotometer in 26 patients, the results were not always compatible with the genetic defects.

Table 4. The clinical and molecular findings of all patients in this study

Patient No	Sex	Age, S/D/L	Age of onset	Initial symptoms	FHx	Other Sx	Gene	Variants	Inheritance
1	M	S, 18Y	11Y	Ataxia/dysarthria	(+)		*NDUFS1*	c.512 C >A: p.T171N c.754G>T: p.D252Y	Mat Pat
2	F	S, 11Y	2mo	DD		Epilepsy	*PNPT1*	c.1145G>A: p.G382E c.646T>G: p.L216V	Mat Pat
3	M	S, 9Y	15mo	DD					
4	M	L at 4Y	15mo	DD	(+)				
5	F	S, 16Y	5Y	Gait disturbance (ataxia)			*NDUFAF6*	c.265G>A: p.A89T c.223_231 del9: p.C76Sfs*12	Mat Pat
6	F	S, 12Y	4Y	Hemiplegia					
7	F	S, 12Y	7mo	DD, hypotonia		Epilepsy	*NDUFAF6*	c.820A>G: p.R274G c.874-2A>C	Pat Mat
8	F	S, 18Y	30mo	Gait disturbance		r/o FSGS	*APOA1BP*	c.733A>C: p.K245Q c.368A>T: p.D123V	Pat Mat
9	M	S, 6Y	15mo	DD					
10	M	S, 29mo	4mo	Seizure/DD		Dyskinesia	*SLC19A3*	c.449C>T: p.A150V c.265A>C: p.S89R	Mat Pat
11	M	S, 4Y	3Y	Ataxia		Involuntary movements	*SURF1*	c.845_846delCT: p.S282Cfs*9 c.367delA: p.R123Gfs*12	Mat Pat
12	F	S, 5Y	10mo	DD					
13	M	S, 10Y	15mo	DD					
14	M	S, 9Y	5mo	Rigidity					
15	F	S, 24Y	5mo	Irritability/DD		Ataxia/ Dystonia/	*SLC19A3*	c.449C>T: p.A150V c.597_598insT:	Mat

						Epilepsy		p.H200Sfs*25
16	F	S, 17Y	5mo	Dystonia		Ataxia/ Involuntary m./Epilepsy	*VPS13D*	c.7760T>C: p.L2587P c.10420C>T: p.R3474*
17	F	S, 9Y	18mo	Ataxia/weakness		Hypotonia/ Dystonia	*IARS2*	c.550G>A: p.A184T c.1967T>C: p.F656S
18	M	L at 31Y	21Y	DM/dysarthria		Weakness		
19	F	L at 7Y	8mo	Status epilepticus		DD/regress/ Ataxia/ Nystagmus	*ECHS1*	c.713C>T: p.A238V c.324C>G: p.F108L
20	F	L at 3Y	6mo	Dev regression				
21	F	L at 2Y	6mo	Irritability				
22	F	L at 3Y	12mo	DD	(+)			
23	F	L at 4Y	6mo	DD		Dystonia/ Dev regress	*ECHS1*	c.707T>C: p.V236A c.542G>A: p.R181H
24	F	L at 1Y	1mo	Dystonia				
25	F	L at 2Y	9mo	Dev regression/ Ptosis/ encephalopathy			*NDUFV1*	c.1235T>C: p.L412P c.733G>A: p.V245M
26	M	L at 7Y	12mo	DD				
27	M	L at 17Y	8Y	Seizure				
28	M	L, < 1Y	1mo	Seizure				
29	M	L at 4Y	6mo	DD				
30	F	L at 12Y	12mo	DD		Ataxia/ Epilepsy/ Involuntary m.	*VPS13D*	Hom c.7760T>C: p.L2587P
31	F	D at 17mo	11mo	DD	(+)	CMP/ Hypotonia/	*IARS2*	c.1195A>G: p.M399V

						Encephalopathy		c.2052delT: p.Q685Kfs*15
32	F	L at 6Y	4Y	DD	(+)	Rigidity/ Invol. m./ Dystonia/ Dysarthria	*SURF1*	c.54+1G>T c.367delA: p.R123Gfs*12
33	M	D at 20mo	15mo	Ataxia				
34	M	S, 30mo	12mo	DD/ encephalopathy				
35	F	S, 13mo	3mo	Seizure		Involuntary m.	*NARS2*	c.731C>G: p.A244G c.1351C>T: p.R451C

Table 5. Pathogenicity of causative variants in known genes of LS

Gene	Variants	Segregation	Reported	Pathogenicity
NDUFS1	c.512C>A: p.T171N	Maternal	Novel	Pathogenic
	c.754G>T: p.D252Y	Paternal	Novel	
NDUFV1	c.1235T>C: p.L412P		Novel	Pathogenic
	c.733G>A: p.V245M		Novel	
NDUFAF6	c.265G>A: p.A89T	Maternal	Novel	Pathogenic
	c.223_231del9: p.C76Sfs*12	Paternal	Novel	
NDUFAF6	c.874-2A>C	Maternal	Novel	Pathogenic
	c.820A>G: p.R274G	Paternal	Novel	
SURF1	c.845_846delCT:	Maternal	Reported	Pathogenic
	p.S282Cfs*9	Paternal	Novel	
	c.367delA: p.R123Gfs*12			
SURF1	c.54+1G>T		Reported	Pathogenic
	c.367delA: p.R123Gfs*12		Novel	
SLC19A3	c.449C>T: p.A150V	Maternal	Novel	Pathogenic
	c.265A>C: p.S89R	Paternal	Novel	
SLC19A3	c.449C>T: p.A150V	Maternal	Novel	Pathogenic
	c.597_598insT:		Novel	
	p.H200Sfs*25			
ECHS1	c.713C>T: p.A238V		Reported	Pathogenic
	c.324C>G: p.F108L		Novel	
ECHS1	c.707T>C: p.V236A		Novel	Pathogenic
	c.542G>A: p.R181H		Novel	
PNPT1	c.1145G>A: p.G382E	Maternal	Novel	Likely pathogenic
	c.646T>G: p.L216V	Paternal	Novel	
IARS2	c.1195A>G: p.M399V		Novel	Pathogenic
	c.2052delT: p.Q685Kfs*15		Novel	
IARS2	c.550G>A: p.A184T		Novel	Likely pathogenic
	c.1967T>C: p.F656S		Novel	
NARS2	c.731C>G: p.A244G	Maternal	Novel	Pathogenic
	c.1351C>T: p.R451C	Paternal	Novel	

Supplementary Table 1. Variants list of all patients identified with analysis pipeline based on the recessive inheritance

Patient No.	Type	SNV/indel	Gene	Amino acid change
1	comHet	SNV/SNV	**NDUFS1**	D252Y/T171N
2	comHet	SNV/SNV	**PNPT1**	G382E/ L216V
3	Hom	SNV	EGFL6	A534T
4	comHet	SNV/SNV	UNC13C	K345E/I1096M
5	comHet	SNV/Indel	**NDUFAF6**	A89T/L74
	comHet	SNV/SNV	TTI1	P796R/M722T
6	comHet	SNV/SNV	TRPM6	E1644A/T1370N
7	comHet	SNV/SNV	**NDUFAF6**	R274G/2bp upstream of exon 9
	comHet	SNV/SNV	ALMS1	T1374A/S2048T/Q2940R
	comHet	SNV/SNV	NAT8	K155E/I44M
	comHet	SNV/SNV	SHANK2	R1293K/R758H
8	comHet	SNV/SNV	**APOA1BP**	K245Q/D123V
9	None			
10	comHet	SNV/SNV	**SLC19A3**	A150V/S89R
	comHet	SNV/SNV	ZNF469	P2665S/G3416A
	comHet	SNV/SNV	FAM149A	R268W/T450I
	Hom	SNV	CNKSR2	I161V
	Hom	SNV	GDPD2	R105Q
	Hom	SNV	MAP7D2	G353V
11	comHet	Indel/Indel	**SURF1**	A174/R123
12	comHet	SNV/SNV	TATDN2	E222G/R472C
13	Hom	SNV	CXorf58	T315M
	Hom	SNV	EGFL6	D174G
	Hom	SNV	GPR112	Y695H
	Hom	SNV	MAGEE1	R14G
	Hom	SNV	RNF128	I46T
	Hom	SNV	SUPT20HL2	N359K
14	comHet	SNV/SNV	IGF2R	I284V/D472E
	Hom	SNV	ARMCX6	Q116R
	Hom	SNV	DNASE1L1	R194C
	Hom	SNV	FAM155B	R167W
15	comHet	SNV/Indel	**SLC19A3**	A150V/H200
	comHet	SNV/SNV	AP1G2	R491Q/E24V
16	comHet	SNV/SNV	**VPS13D**	L2587P/R3474X
17	comHet	SNV/Indel	**IARS2**	M399V/D684
18	comHet	SNV/SNV	EP300	G322A/V456I
	Hom	SNV	COL4A5	P1314S
	Hom	SNV	COL4A6	P1221S
	Hom	SNV	FRMPD4	V1256M
19	comHet	SNV/SNV	**ECHS1**	A238V/F108L
	comHet	SNV/SNV	ZNF804B	G733D/Q1051X
20	comHet	SNV/Indel	EFR3A	H490Q/P230
	comHet	SNV/SNV	KIAA1731	D1088N/P562A
	comHet	SNV/SNV	KIAA1671	K439R/V793I

21	comHet	SNV/SNV	DUOX2	P1391A/A649E
	comHet	SNV/SNV	FUK	T478M/A902P
22	comHet	SNV/SNV	UNC13C	K345E/I1096M
23	comHet	SNV/SNV	**ECHS1**	V236A/R181H
24	None			
25	comHet	SNV/SNV	**NDUFV1**	L412P/V245M
	comHet	SNV/SNV	USP43	S961R/P1017A
26	comHet	SNV/SNV	NFRKB	T1039S/P778L
	Hom	SNV	CDR1	V97M
	Hom	SNV	QRFPR	A271V
27	Hom	SNV	KIF4A	R985Q
28	comHet	SNV/SNV	KIAA0947	Q984E/K152R
	comHet	SNV/Indel	MYH4	M826V/E1493
	Hom	Indel	ICAM3	A352
	Hom	SNV	ARMCX1	K379E
	Hom	SNV	FAM47A	A787V
	Hom	SNV	GPR112	Y695H
29	None			
30	Hom	SNV	**VPS13D**	L2587P
	comHet	SNV/SNV	ATP1A4	V306M/I595K
31	comHet	SNV/SNV	**IARS2**	A184T/A184T
32	None			
33	comHet	SNV/SNV	TAF1L	F1224S/T81M
	comHet	SNV/SNV	TAOK2	H352Y/A1065V
	Hom	SNV	CDX4	P78L
	Hom	SNV	TMEM187	S256A
	Hom	SNV	VWA5B1	T147P
34	comHet	SNV/SNV	ANO7	Q714R/E757Q
	comHet	SNV/SNV	DNTT	D83E/W87S
	comHet	SNV/SNV	RYR3	M1625L/V4369M
35	comHet	SNV/SNV	**NARS2**	R224C/A17G
	comHet	SNV/SNV	ARHGAP10	R345W/S668R
	comHet	SNV/SNV	MUC19	G5238S/G5278S/S5296G/C5297W
	comHet	SNV/SNV	ZNF91	A338T/A310T

comHet, compound hetereozygous; Hom, homozygous; SNV, single nucleotide variant; Indel, insertion or deletion
LS4 and LS22 are siblings
These variants are compound heterozygous or homozygous variants before filtering with rarity or functions of variants. Bold forms indicate the pathogenic variants in genes known to cause LS or variants in novel genes with the possibility of pathogenicity although being not known to be associated with LS or mitochondrial diseases.

Supplementary Table 2. Only one heterozygous variant list of LS candidate genes

Patient No.	Type	SNV/indel	Gene	Amino acid change
3	Het	SNV	**ECHS1**	F287L
9	Het	SNV	**NDUFAF6**	R274G
24	Het	SNV	**ECHS1**	G175S
32	Het	Indel	**SURF1**	R123
21	Het	SNV	**ECHS1**	L8P
33	Het	SNV	**NDUFS7**	R59W

Het, hetereozygous

Discussion

Mitochondrial disorder is caused by defects in dual genome: mitochondrial and nuclear genome. Mitochondrial DNA composes of 13 genes for OXPHOS subunits and 22 tRNA genes / 2 rRNA genes for mitochondrial translation[14]. In view of nuclear DNA, there are more than 1,000 nuclear genes associated with mitochondrial metabolism and maintenance as well as OXPHOS subunits/assembly factors[6,9,13]. To date, pathogenic mutations have been identified in more than 75 genes in dual genome as an increasing number of nuclear DNA mutations are being reported[5]. In children with mitochondrial disorders, nuclear DNA mutations are the predominant causes compared to mitochondrial DNA mutations. Recently many studies in mitochondrial disorders have been done to search for nuclear genetic defects, with the help of NGS[8,9,15-18].

LS is the most common mitochondrial disorder in childhood with a prevalence of 1:40,000 live births, which has several characteristic clinical and neuroradiological findings[19]. It is well known that LS is genetically heterogeneous[5]. However, there have been just a few case reports of which the causative mutations in LS were identified through high-throughput sequencing[20-22]. This might be partly due to difficulty of collecting a large enough set of patients with LS. Herein we performed WES in 36 patients with LS and identified nuclear causative genes of LS, in order to understand molecular defects and furthermore discover novel genes in LS.

The present study identified the genetic etiology in 14 of 34 families (41.2%). They included definite (8, 23.5%) and likely pathogenic causes (6, 17.7%). This yield rate was relatively high compared to previous studies[15,23]. Although many WES studies in other

mitochondrial disorders have been reported to date, there are none in WES study for LS alone, except some case reports. Other various diseases associated with secondary mitochondrial dysfunctions could be included in the previous studies, thereby resulting in relatively lower yield rate. The present study with WES in clinically homogeneous group, LS, showed the higher yield rate and helped us to understand the genetic basis and pathophysiology of LS and novel causative genes of LS. Targeted NGS has been widely used for identification of causes in mitochondrial disease in that it is less costly, lower false positive rate, shorter turnaround time, and easier to interpret, compared to WES[15,18]. However, WES can give us additional information of novel causative genes in LS despite these limitations, as there have been many nuclear genes which are unknown but associated with mitochondrial functions.

Of note, there is another reason for high yield rate in this study. By understanding the possibility that 3-5% of exomes or more can be missed due to very high/low GC regions or repetitive sequences during exome capture procedure in WES, we performed resequencing for the poorly covered regions. In some cases with one heterozygous known variant causing LS, we made an effort to search for other allele variant. We found one case with compound heterozygous variants in *SURF1*. Even with good quality of the exome sequencing in the present study, there are still limitations such as no detection of mutations in the non-coding regions or uneven and incomplete sequencing coverage of some exonic regions.

This study shows the genetic heterogeneity of LS in both mitochondrial and nuclear genome. In 2008, small-sized study with 16 patients was performed in Korean LS and revealed mitochondrial DNA mutations in 31.3%[11]. Additionally relatively large-sized

cohort study with 39 patients was performed, but mitochondrial DNA mutations alone were identified in 28%, without nuclear DNA sequencing performed (unpublished). As reported previously in other studies, Korean studies with LS identified mitochondrial DNA mutations in about 30%[7,11]. The present study is the largest-sized genetic study with Korean LS patients and furthermore revealed both mitochondrial and nuclear genetic defects. Additionally, all the pathogenic variants identified in this study, except one, have never been reported. These findings expanded the genetic spectrum of LS. Discovery of molecular defects in Korean LS might prompt us to develop the gene panel for Korean LS and treatment for each molecular defect in LS in future. This would lead to cost-effective diagnostic and therapeutic approach for patients with LS. Clinical trial with drugs such as rapamycin can be done in future, although there are none with *NDUFS4* mutations in this study[24].

Among the cases with molecular defects identified (59.3%), mtDNA and nDNA mutations were 16 and 16, although the remainder unidentified causes (40%) might be nuclear genetic defects. In total, complex I subunits and assembly factors were 13 and 0, complex IV subunits and assembly factors 0 and 4, complex V subunits 5, pyruvate dehydrogenase complex subunits and cofactors 2 and 2, mtRNA translation 4, and valine metabolism defects 2. Complex I (NADH:ubiquinone oxidoreductase) is the largest multimeric enzyme complex of OXPHOS and its dysfunction is the most common enzyme defect in mitochondrial disorders[5,14]. Considering each genetic mutations, *ATP6* and *ND3* variants were identified most frequently in this study.

Of note, there are many genes associated with mitochondrial translation identified: *PNPT1* (OMIM*610316, encoding a polyribonucleotide nucleotidyltransferase), *IARS2*

(OMIM*612801, encoding a mitochondrial isoleucyl-tRNA synthetase), and *NARS2* (OMIM*612803, encoding a mitochondrial asparaginyl-tRNA synthetase). Recently increasing number of cases with nuclear defects associated with mitochondrial translation has been published[5,6,10,25]. Additionally there are two cases with *SLC19A3* variants, one of the preventable causes in LS. Their clinical and neuroradiological findings of cases with *SLC19A3* variants did not help us to differentiate them from other LS-like disorders[26]. *SLC19A3* (OMIM*606152) encodes a thiamine transporter 2 (hTHTR2), the defects of which result in biotin-responsive basal ganglia disease (BBGD). Supplementation of biotin and thiamine (both 10-15 mg/kg/day) can prevent disease progression, otherwise fatal. Our cases were not suspected of BBGD, but were partially treated with multivitamin, including thiamine. One of them was surviving at 24 years of age, although she did not show developmental progression. In cases with LS, empirical treatment with high dose biotin and thiamine is recommended together with targeted sequencing. In Korean population with LS, we should consider the possibility of *SLC19A3* mutations, because *SLC19A3*-associated LS is a treatable or preventable neurometabolic disease. The present study also identified defects in valine metabolism as a cause of LS. *ECHS1* (OMIM*602292) encodes short-chain enoyl-CoA hydratase, the enzyme upstream of 3-hydroxyisobutyryl-CoA hydrolase. *ECHS1* mutations have been reported to be associated with LS recently, although the exact mechanism remains uncertain. No abnormal organic acid profiles in urine were observed in our two cases with *ECHS1* defects. Further metabolite studies in the cases can be needed if possible. As the previous studies for MRC deficiency with WES identified causative variants associated with not only intramitochondrial protein synthesis but also mitochondrial

replication and expression and also searched for other neurometabolic disorders or the treatable/preventable causes mimicking MRCD, we identified the molecular defects in valine metabolism or thiamine transporter as well as the genetic variants in OXPHOS subunits or assembly factors[5,6,27].

In cases with pathogenic variants identified, there were no characteristic neuroradiological findings observed differently from typical LS. As reported in the previous studies rarely, cases with variants in *SURF1* and *NDUFV1* showed progressive leukodystrophy during follow-up[28,29]. Although LS is defined by bilateral symmetric lesions in basal ganglia or other CNS areas, meaning systemic disease, rarely it occurs with unilateral lesions initially and extends to bilateral symmetric lesions.

In the present study, we attempted to determine a correlation between genotype and phenotype by reviewing the previous reports together, because the numbers of patients with the same genotype were very low in this study. Even with effort for comparing with the cases with reported genetic mutations, there was poor genotype-phenotype correlation. No significant differences in clinical severity were found between patients with mutations in assembly factor genes and patients with mutations in nuclear-encoded subunit genes[30]. In view of most of mitochondrial DNA mutations except T8993G/C, clinical expression of a mtDNA mutation is influenced not only by the pathogenicity of the mutation itself but also by the heteroplasmic mutant load, the variation in mutant load among different tissues, and the energy requirements of brain and other tissues, which may vary with age[31]. For nuclear DNA mutations, it remains uncertain why it is difficult to determine genotype-phenotype correlation, although there might be some individual modifying factors[3]. Together with genetic diversity, broad clinical spectrum, including

atypical presentation or overlapping clinical features, makes the targeted genetic testing for phenotype difficult.

We discovered candidate novel genes of LS through WES, although the functional studies are needed. There were two genes, *APOA1BP* and *VPS13D*, with unknown significance. The variants might have pathogenicity of LS in that *APOA1BP* has Mitocarta score of 16 and *VPS13D* is the gene identified in two unrelated patients with LS. There is one example, *PET100*, which was originally not localized in mitochondrial membrane and included in MitoCarta[13,32]. However, after analyzing mouse heart mitochondria, it was shown to be localized in mitochondrial membrane and discovered to be a founder mutation in Lebanese patients with LS. For the two genes, *APOA1BP* and *VPS13D*, we are planning to perform mitochondrial rescue experiment in vitro.

Conclusion

Leigh syndrome (LS) is one of the common mitochondrial disorders which have genetic and clinical complexity. In the present study, whole exome sequencing truly showed genetic heterogeneity in LS, which enhanced our understanding of genetic basis of LS. This study allowed us to investigate the mutational spectrum and distribution of causative genes in Korean patients with LS, and analyze genotype-phenotype correlations in LS. It is known that mitochondrial DNA mutations account for about 30% of LS, while others might be caused by nuclear DNA defects. About half of the cases without mtDNA mutations were found to carry nuclear genome defects: mutations in mitochondrial translation and valine metabolism as well as oxidative phosphorylation system subunits or assembly factors. Taken together in both genome, complex I defects were the most common cause of LS in Korean populations. Nuclear genes associated with mitochondrial RNA translation accounted for a substantial proportion of genetic causes of LS. Treatable or preventable causes in LS like *SLC19A3* mutations or *ECHS1* mutations with valine metabolism defects should be always considered. It was difficult to determine genotype-phenotype correlations in this study. Together with genetic diversity, broad clinical spectrum, made the targeted genetic testing difficult. Whole exome sequencing showed the clinical utility to identify the causative genes in patients with LS, although some cases with LS still remain unsolved etiologically and therefore whole genome sequencing is needed in future.

References

1. Leigh D. Subacute necrotizing encephalomyelopathy in an infant. Journal of neurology, neurosurgery, and psychiatry 1951;14:216-21.

2. Finsterer J. Leigh and Leigh-like syndrome in children and adults. Pediatric neurology 2008;39:223-35.

3. Baertling F, Rodenburg RJ, Schaper J, et al. A guide to diagnosis and treatment of Leigh syndrome. Journal of neurology, neurosurgery, and psychiatry 2014;85:257-65.

4. Koenig MK. Presentation and diagnosis of mitochondrial disorders in children. Pediatric neurology 2008;38:305-13.

5. Lake NJ, Compton AG, Rahman S, Thorburn DR. Leigh syndrome: One disorder, more than 75 monogenic causes. Annals of neurology 2016;79:190-203.

6. DiMauro S, Schon EA, Carelli V, Hirano M. The clinical maze of mitochondrial neurology. Nature reviews Neurology 2013;9:429-44.

7. Naess K, Freyer C, Bruhn H, et al. MtDNA mutations are a common cause of severe disease phenotypes in children with Leigh syndrome. Biochimica et biophysica acta 2009;1787:484-90.

8. Dinwiddie DL, Smith LD, Miller NA, et al. Diagnosis of mitochondrial disorders by concomitant next-generation sequencing of the exome and mitochondrial genome. Genomics 2013;102:148-56.

9. Carroll CJ, Brilhante V, Suomalainen A. Next-generation sequencing for mitochondrial disorders. British journal of pharmacology 2014;171:1837-53.

10. Boczonadi V, Horvath R. Mitochondria: impaired mitochondrial translation in

human disease. The international journal of biochemistry & cell biology 2014;48:77-84.

11. Chae JH, Lee JS, Kim KJ, Hwang YS, Hirano M. Biochemical and genetic analysis of Leigh syndrome patients in Korea. Brain & development 2008;30:387-90.

12. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine 2015;17:405-24.

13. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic acids research 2016;44:D1251-7.

14. DiMauro S, Schon EA. Mitochondrial respiratory-chain diseases. The New England journal of medicine 2003;348:2656-68.

15. Lieber DS, Calvo SE, Shanahan K, et al. Targeted exome sequencing of suspected mitochondrial disorders. Neurology 2013;80:1762-70.

16. Vasta V, Merritt JL, 2nd, Saneto RP, Hahn SH. Next-generation sequencing for mitochondrial diseases: a wide diagnostic spectrum. Pediatrics international : official journal of the Japan Pediatric Society 2012;54:585-601.

17. Wong LJ. Next generation molecular diagnosis of mitochondrial disorders. Mitochondrion 2013;13:379-87.

18. Calvo SE, Compton AG, Hershman SG, et al. Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. Science translational medicine 2012;4:118ra10.

19. Lake NJ, Bird MJ, Isohanni P, Paetau A. Leigh syndrome: neuropathology and pathogenesis. Journal of neuropathology and experimental neurology 2015;74:482-92.

20. Simon M, Richard EM, Wang X, et al. Mutations of human NARS2, encoding the mitochondrial asparaginyl-tRNA synthetase, cause nonsyndromic deafness and Leigh syndrome. PLoS genetics 2015;11:e1005097.

21. Schwartzentruber J, Buhas D, Majewski J, et al. Mutation in the nuclear-encoded mitochondrial isoleucyl-tRNA synthetase IARS2 in patients with cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, and peripheral neuropathy or with Leigh syndrome. Human mutation 2014;35:1285-9.

22. Tetreault M, Fahiminiya S, Antonicka H, et al. Whole-exome sequencing identifies novel ECHS1 mutations in Leigh syndrome. Human genetics 2015;134:981-91.

23. Ohtake A, Murayama K, Mori M, et al. Diagnosis and molecular basis of mitochondrial respiratory chain disorders: exome sequencing for disease gene identification. Biochimica et biophysica acta 2014;1840:1355-9.

24. Johnson SC, Yanos ME, Kayser EB, et al. mTOR inhibition alleviates mitochondrial disease in a mouse model of Leigh syndrome. Science (New York, NY) 2013;342:1524-8.

25. Vedrenne V, Gowher A, De Lonlay P, et al. Mutation in PNPT1, which encodes a polyribonucleotide nucleotidyltransferase, impairs RNA import into mitochondria and causes respiratory-chain deficiency. American journal of human genetics 2012;91:912-8.

26. Alfadhel M, Almuntashri M, Jadah RH, et al. Biotin-responsive basal ganglia disease should be renamed biotin-thiamine-responsive basal ganglia disease: a retrospective review of the clinical, radiological and molecular findings of 18 new cases. Orphanet journal of rare diseases 2013;8:83.

27. Rahman S. Emerging aspects of treatment in mitochondrial disorders. Journal of

inherited metabolic disease 2015;38:641-53.

28. Rahman S, Brown RM, Chong WK, Wilson CJ, Brown GK. A SURF1 gene mutation presenting as isolated leukodystrophy. Annals of neurology 2001;49:797-800.

29. Ortega-Recalde O, Fonseca DJ, Patino LC, et al. A novel familial case of diffuse leukodystrophy related to NDUFV1 compound heterozygous mutations. Mitochondrion 2013;13:749-54.

30. Koene S, Rodenburg RJ, van der Knaap MS, et al. Natural disease course and genotype-phenotype correlations in Complex I deficiency caused by nuclear gene defects: what we learned from 130 cases. Journal of inherited metabolic disease 2012;35:737-47.

31. Thorburn DR, Rahman S. Mitochondrial DNA-Associated Leigh Syndrome and NARP. In: Pagon RA, Adam MP, Ardinger HH, et al., eds. GeneReviews(R). Seattle (WA): University of Washington, Seattle

University of Washington, Seattle. All rights reserved.; 1993.

32. Lim SC, Smith KR, Stroud DA, et al. A founder mutation in PET100 causes isolated complex IV deficiency in Lebanese individuals with Leigh syndrome. American journal of human genetics 2014;94:209-22.

국문초록

리 증후군은 가장 흔한 미토콘드리아 질환 중 하나로, 다양한 신경퇴행 증상 및 양측 대칭성의 중추신경계 병변을 특징으로 하는 질환이다. 약 30%에서 미토콘드리아 유전자 돌연변이가 관찰되지만, 나머지 70%는 핵 유전자 돌연변이에 의해 발병하는 것으로 생각된다. 현재까지 미토콘드리아 및 핵 게놈에서 75개 이상의 유전자 돌연변이가 발견되었다. 본 연구에서는 엑솜 시퀀싱을 이용하여 리 증후군의 원인 핵 유전자 돌연변이를 발견하고자 한다. 서울대학교 어린이 병원에서 리 증후군으로 진단받은 환자들 중에서 미토콘드리아 유전자 돌연변이가 발견되지 않은 35명 (34 가족)의 환자들에 대해 연구를 진행하였다. 리 증후군은 특징적 임상 양상과 뇌신경 영상 소견에 근거하고, 미토콘드리아 기능 이상을 시사하는 생화학적 이상 소견 등을 고려하여 진단하였다. 리 증후군을 일으키는 것으로 알려져 있는 유전자들의 변이는 총 14명에서 발견되었다 (총 14가족, 41.2%). 발견된 원인 유전자는 *NDUFS1* (1명), *NDUFV1* (1명), *NDUFAF6* (2명), *SURF1* (2명), *SLC19A3* (2명), *ECHS1* (2명), *PNPT1* (1명), *IARS2* (2명), and *NARS2* (1명) 이었다. 흥미롭게도 미토콘드리아 번역에

관련된 핵 유전자 돌연변이가 28.6%를 차지하였다. 한편, 3명의 환자에서는 사람에서의 기능이 아직 확인이 되지 않은 새로운 유전자 2개가 발견되었다 (*APOA1BP*, *VPS13D*). 원인 유전자 돌연변이가 발견된 증례들의 임상 양상, 뇌 신경 영상 소견, 생화학적 이상 소견 등을 이전에 보고된 증례들과 비교해 보았으나 유전형-표현형 사이의 특정 상관관계는 없었다. 본 연구를 통해 리 증후군의 유전적 복합성을 보여주었고, 그 원인으로는 잘 알려진 산화적 인산화 시스템의 서브유닛이나 어셈블리 요인뿐 아니라 미토콘드리아 번역에 관련된 유전자 돌연변이와 발린 대사 과정에 관련된 유전자 돌연변이까지 다양하였다. 엑솜 시퀀싱은 리 증후군의 원인 유전자 스펙트럼을 넓혀주었고, 유전적 원인을 찾는 데 있어서 임상적 유용성을 보여주었다. 뿐만 아니라, 한국인 리 증후군 환자들의 원인 유전자에 대한 유전자 패널을 개발하여 빠르고 정확한 진단이 가능할 것이다.

주요어: 리 증후군, 미토콘드리아 질환, 미토콘드리아 유전자, 핵 유전자, 엑솜 시퀀싱

학 번: 2007-30568