



## A New Integrated Newborn Screening Workflow Can Provide a Shortcut to Differential Diagnosis and Confirmation of Inherited Metabolic Diseases

Jung Min Ko<sup>1</sup>, Kyung Sun Park<sup>2</sup>, Yeeok Kang<sup>2</sup>, Seong-Hyeuk Nam<sup>2</sup>, Yoonjung Kim<sup>3</sup>, Inho Park<sup>2</sup>, Hyun Wook Chae<sup>4</sup>, Soon Min Lee<sup>4</sup>, Kyung-A Lee<sup>3</sup>, and Jong-Won Kim<sup>5</sup>

<sup>1</sup>Department of Pediatrics, Seoul National University Children's Hospital, Seoul National University College of Medicine, Seoul;

<sup>2</sup>SD Genomics Co., Ltd., Seoul;

Departments of <sup>3</sup>Laboratory Medicine and <sup>4</sup>Pediatrics, Yonsei University College of Medicine, Seoul;

<sup>5</sup>Department of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.

**Purpose:** We developed a new workflow design which included results from both biochemical and targeted gene sequencing analysis interpreted comprehensively. We then conducted a pilot study to evaluate the benefit of this new approach in newborn screening (NBS) and demonstrated the efficiency of this workflow in detecting causative genetic variants.

**Materials and Methods:** Ten patients in Group 1 were diagnosed clinically using biochemical assays only, and 10 newborns in Group 2 were diagnosed with suspected inherited metabolic disease (IMD) in NBS. We applied NewbornDiscovery (SD Genomics), an integrated workflow design that encompasses analyte-phenotype-gene, single nucleotide variant/small insertion and deletion/copy number variation analyses along with clinical interpretation of genetic variants related to each participant's condition.

**Results:** A molecular genetic diagnosis was established in 95% (19/20) of individuals. In Group 1, 13 and 7 of 20 alleles were classified as pathogenic and likely pathogenic, respectively. In Group 2, 11 and 6 of 17 alleles with identified causative variants were pathogenic and likely pathogenic, respectively. There were no variants of uncertain significance. For each individual, the NewbornDiscovery and biochemical analysis results reached 100% concordance, since the single newborn testing negative for causative genetic variant in Group 2 showed a benign clinical course.

**Conclusion:** This integrated diagnostic workflow resulted in a high yield. This approach not only enabled early confirmation of specific IMD, but also detected conditions not included in the current NBS.

**Key Words:** Newborn screening, inherited metabolic disease, dried blood spot, targeted gene panel sequencing, next-generation sequencing

### INTRODUCTION

Newborn screening (NBS) for inherited metabolic diseases (IMDs) is a representative public healthcare project conducted worldwide. In Korea, about 80% of newborn babies under-

go NBS for approximately 40 to 50 IMDs. However, only 6 IMDs (congenital hypothyroidism, congenital adrenal hyperplasia, phenylketonuria, maple syrup urine disease, galactosemia, and homocystinuria) have been performed free of charge by the government and other IMDs are not financially

**Received:** January 12, 2018 **Revised:** April 16, 2018 **Accepted:** April 19, 2018

**Co-corresponding authors:** Kyung-A Lee, MD, PhD, Department of Laboratory Medicine, Yonsei University College of Medicine, 211 Eonjuro, Gangnam-gu, Seoul 06273, Korea.

Tel: 82-2-2019-3531, Fax: 82-2-2019-4822, E-mail: KAL1119@yuhs.ac and

Jong-Won Kim, MD, PhD, Department of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Korea.

Tel: 82-2-3410-2705, Fax: 82-2-3410-2719, E-mail: kimjw@skku.edu

•The authors have no financial conflicts of interest.

© Copyright: Yonsei University College of Medicine 2018

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

covered by the government.<sup>1</sup> Traditionally, abnormal NBS result warrants differential diagnosis of the suspected disease, based on an analysis of specific enzymatic or biochemical features in the blood or urine specimens (Fig. 1A). The American College of Medical Genetics and Genomics (ACMG) has developed and established action sheets and algorithms that provide next steps for differentiation of conditions associated with the same or similar metabolic profiles (<https://www.ncbi.nlm.nih.gov/books/NBK55827/>). However, it is not easy to differentiate IMDs based on biochemical analysis of metabolites alone, which is occasionally impracticable in a clinical setting.<sup>2</sup> Furthermore, even different diseases that affect the same metabolic component have distinct and specific treatments and prognoses. In this scenario, a molecular genetic analysis and identification of causative genetic variants are often required to confirm specific disease. Sanger sequencing may require several weeks to distinguish multiple diseases or genes. Therefore, empirical treatment is usually provided while awaiting the final results.<sup>3</sup>

Recently, several multi-gene panels based on next generation sequencing (NGS) have been developed and validated to facilitate simultaneous and cost-effective analysis of multiple genes related to IMDs or severe primary immunodeficiency. In particular, satisfactory NGS results were obtained with only

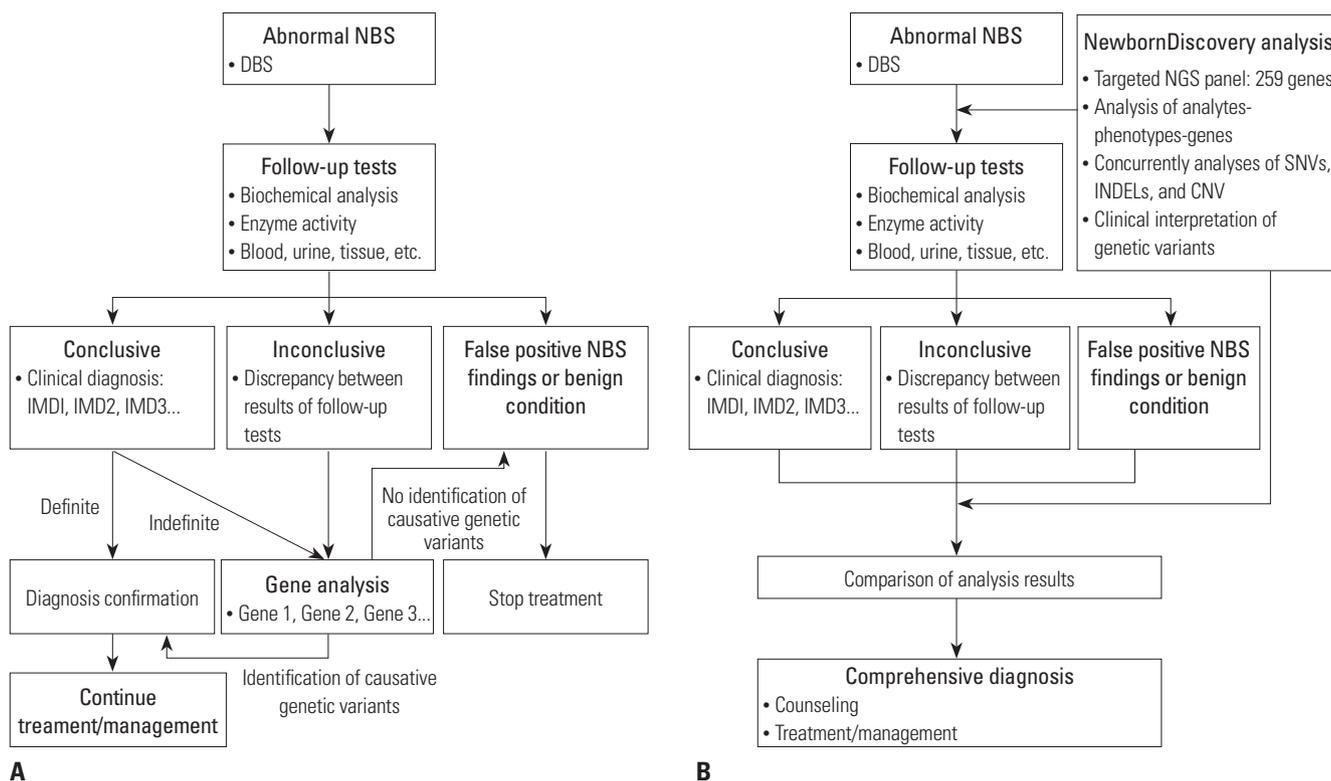
a small amount of DNA, utilizing minimally invasive procedures with dried blood spots (DBS) or saliva within 1–3 weeks.<sup>4-6</sup>

Here, we introduce a new analytical workflow which includes results from both NBS/biochemical assays and targeted gene panel analysis interpreted simultaneously. We demonstrated the efficiency of this workflow for the detection of causative genetic variants in clinically diagnosed patients.

## MATERIALS AND METHODS

### Study participants and study design

The Institutional Review Boards of Seoul National University Hospital approved and monitored this study (H-1601-079-734). All study participants (parent, legal guardian, and/or child) provided written informed consent to participate in the study. Twenty participants were enrolled into 2 groups (Tables 1 and 2) with each group comprising 10 subjects. No participant had a family history of consanguineous marriage. Group 1 included patients who had been clinically diagnosed with specific IMD based on clinical and biochemical criteria, and were managed prior to enrollment in this study (Table 1). Six of the 10 patients had a type of lysosomal storage disease (LSD) that was not screened by the current NBS in Korea. We analyzed the detec-



**Fig. 1.** Algorithms for diagnosis of IMDs in newborns. Conventional (A) and newly integrated (B) algorithms are presented for the diagnosis of IMDs in newborns suspected with a specific disease on NBS. In contrast to the conventional approach, the newly integrated approach is based on standard biochemical and targeted gene panel analyses concurrently, and data from each analysis are compared and reaffirmed to enable rapid disease diagnosis. IMD, inherited metabolic disease; NBS, newborn screening; DBS, dried blood spot; NGS, next generation sequencing; SNV, single nucleotide variant; INDEL, small insertions and deletions; CNV, copy number variations.

Table 1. Samples from Clinically Diagnosed Patients in Group 1 Validated by NewbornDiscovery

Patient ID	Previous diagnostic study	Subjective symptom or sign	Clinical diagnosis	NewbornDiscovery			Final diagnosis
				Gene	Refseq	Genetic alteration [zygosity]	
CD_1	AC (C3↑), urine OA (MMA↑)	No	Methylmalonic acidemia	<i>MUT</i>	NM_000255.3	c.322C>T, p.(Arg108Cys) [Het] c.1481T>A, p.(Leu494*) [Het]	Methylmalonic acidemia P
CD_2	Plasma AA (Met↑), normal HCY	No	Methionine adenosyltransferase I/III deficiency	<i>MAT1A</i>	NM_000429.2	c.746G>A, p.(Arg249Gln) [Het] c.895C>T, p.(Arg299Cys) [Het]	Methionine adenosyltransferase I/III deficiency (AR type) LP
CD_3	Gal6sulfatase activity↓, urine GAG↑	Yes	Morquio A	<i>GALNS</i>	NM_000512.4	c.281G>A, p.(Arg94His) [Het] c.707A>C, p.(His236Pro) [Het]	Morquio A LP
CD_4	Gal6sulfatase activity↓, urine GAG↑	Yes	Morquio A	<i>GALNS</i>	NM_000512.4	c.725C>G, p.(Ser242Cys) [Het] c.1019G>A, p.(Gly340Asp) [Het]	Morquio A P
CD_5	AC (C5↑), urine OA (IVG↑)	No	Isovaleric acidemia	<i>IVD</i>	NM_002225.3	c.466-3_466-2delinsGG, skipping of exon 5 [Hom]	Isovaleric acidemia P
CD_6	Acid-a-glucosidase activity↓	Yes	Pompe disease	<i>GAA</i>	NM_000152.4	c.1225dup, p.(Asp409Glyfs*97) [Het] c.1316T>A, p.(Met439Lys) [Het]	Pompe disease LP
CD_7	Acid-a-glucosidase activity↓	Yes	Pompe disease	<i>GAA</i>	NM_000152.4	c.118C>T, p.(Arg40*) [Het] c.2015G>A, p.(Arg672Gln) [Het]	Pompe disease P
CD_8	Acid-a-glucosidase activity↓	Yes	Pompe disease	<i>GAA</i>	NM_000152.4	c.1579_1580del, p.(Arg527Glyfs*3) [Het] c.1857C>G p.(Ser619Arg) [Het]	Pompe disease P
CD_9	Beta-glucocerebrosidase activity↓	Yes	Gaucher disease	<i>GBA</i>	NM_001005741.3	c.254G>A, p.(Gly85Glu) [Het] c.887G>A, p.(Arg296Gln) [Het]	Gaucher disease P
CD_10	AC (C4↑), urine OA (EMA↑, MSA↑)	No	Short-chain acyl-CoA dehydrogenase deficiency	<i>ACADS</i>	NM_000017.3	c.164C>T, p.(Pro55Leu) [Hom]	Short-chain acyl-CoA dehydrogenase deficiency LP

AC, acylcarnitine; OA, organic acid; MMA, methylmalonic acid; AA, amino acid; Met, methionine; HCY, homocysteine; GAG, glycosaminoglycan; IVG, isovalerylglycine; EMA, ethylmalonic acid; MSA, methylsuccinic acid; Het, heterozygous; Hom, homozygous; LP, likely pathogenic; P, pathogenic; AR, autosomal recessive; ACMG, American College of Medical Genetics and Genomics.

**Table 2.** NBS-Positive Samples in Group 2 Confirmed by Combined Diagnostic Approach Using NewbornDiscovery

Patient ID	Biochemical tests			Subjective symptoms or sign	Initial diagnosis	Gene	RefSeq	NewbornDiscovery		
	Abnormal metabolites on NBS	Plasma	Urine					Genetic alteration [zygosity]	ACMG category	Final diagnosis
UD_1	C5DC1, C6↑, C8↑, C10↑, C12↑, C14↑, C14:1↑, C16↑, C18↑	C5DC1, C6↑, C8↑, GA	C10↑, C12↑, C14↑, C14:1↑, C16↑, C18↑	No	Multiple acyl CoA dehydrogenase deficiency	<i>ETFDH</i>	NM_004453.3	Exon 1-7 deletion [Het] c.831+3A>C [Het]	P LP	Multiple acyl CoA dehydrogenase deficiency
UD_2	C4↑	EMA, MSA		No	Short-chain acyl-CoA dehydrogenase deficiency	<i>ACADS</i>	NM_000017.3	c.1031A>G, p.(Glu344Gly) [Hom]	LP	Short-chain acyl-CoA dehydrogenase deficiency
UD_3	C5↑	C4-OH↑, C5↑	IVG, 3-OH-IVA	Yes	Isovaleric acidemia	<i>IVD</i>	NM_002225.3	c.466-3_466-2delinsGG, skipping of exon 5 [Hom]	P	Isovaleric acidemia
UD_4	C4↑, C5↑, C4-OH↑, C5-DC↑, C12↑, C14↑, C14:1↑, C16↑, C18↑	C4↑, C5↑, C4-OH↑, GA, EMA, IVG, 2-OH-GA	C5-DC↑, C12↑, C14↑, C14:1↑, C16↑, C18↑	Yes	Multiple acyl CoA dehydrogenase deficiency	<i>ETFDH</i>	NM_004453.3	Exon 7-8 deletion [Het] c.1354A>G, p.(Arg452Gly) [Het]	P LP	Multiple acyl CoA dehydrogenase deficiency
UD_5	C5↑	C5↑, C3-DC↑, C4-OH↑	IVG	No	Isovaleric acidemia	<i>IVD</i>	NM_002225.3	c.466-3_466-2delinsGG, skipping of exon 5 [Hom]	P	Isovaleric acidemia
UD_6	Met↑	Met↑		No	Methionine adenosyltransferase I/III deficiency	<i>MATTA</i>	NM_000429.2	c.746G>A, p.(Arg249Gln) [Het]	LP	Methionine adenosyltransferase I/III deficiency (AD type)
UD_7	Cit↑	Cit↑		No	Argininosuccinic aciduria	<i>ASL</i>	NM_000048.3	c.467C>T, p.(Pro156Leu) [Het] Exon 15 deletion [Het]	LP P	Argininosuccinic aciduria
UD_8	C3↑	C3↑	PG, 3-OH-PA, MCA	Yes	Propionic acidemia	<i>PCCB</i>	NM_000532.4	c.331C>T, p.(Arg111*) [Hom]	P	Propionic acidemia
UD_9	C0↓, AC↓	AC↓, normal CO, CO/C16+, C18↑		No	Systemic carnitine deficiency, Carnitine palmitoyltransferase 1A deficiency	<i>SLC22A5</i>	NM_003060.3	c.289C>G, p.(Leu97Val) [Het] c.396G>A, p.(Trp132*) [Het]	LP P	Systemic carnitine deficiency
UD_10	Pro↑	Pro↑		No	Hyperprolinemia				ND ND	ND ND

NBS, newborn screening; ACMG, American College of Medical Genetics and Genomics; GA, glutaric acid; EMA, ethylmalonic acid; MSA, methylsuccinic acid; IVG, isovalerylglycine; IVA, isovaleric acid; PG, propionylglycine; PA, propionic acid; MCA, methylcitric acid; Met, methionine; Cit, citrulline; Pro, proline; AC, acylcarnitine; Het, heterozygote; Hom, homozygote; LP, likely pathogenic; P, pathogenic; AD, autosomal dominant; ND, not determined.

tion rate of causative genetic variants and the concordance rate between clinical and molecular diagnosis in this group. Group 2 included newborns referred with abnormal NBS tests in an actual clinical setting (Table 2). We performed a pilot prospective clinical study in which this new analytical workflow (NewbornDiscovery, SD Genomics, Seoul, Korea) was applied simultaneously during follow-up standard biochemical tests of plasma and urine metabolites (Fig. 1B).

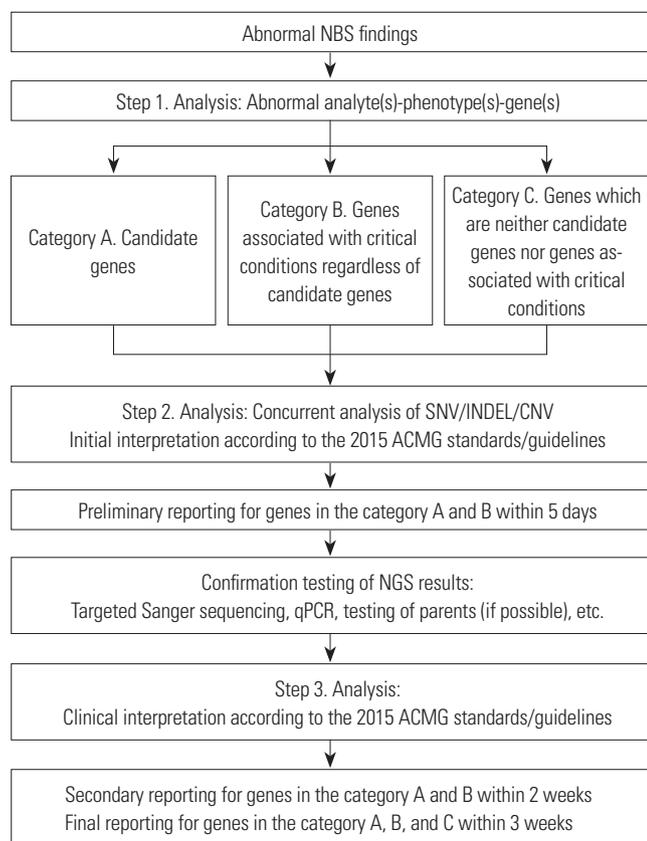
### New analytical workflow for NBS

NewbornDiscovery is an integrated workflow design that encompasses concurrent analyte-phenotype-genes (step 1) with single nucleotide variant (SNV)/small insertion and deletion (INDEL)/copy number variation (CNV) analyses (step 2), as well as the clinical interpretation of genetic variants according to the 2015 ACMG standards/guidelines (step 3) (Fig. 2A).<sup>7</sup> An NGS panel for NewbornDiscovery comprised 259 genes related to 196 IMDs (including endocrinopathies), 9 hemoglobinopathies, 4 congenital non-hemolytic hyperbilirubinemias, 23 immunodeficiencies, and 17 hearing losses (Supplementary Table 1, only online). After the step 1 (abnormal analyte-phenotype-gene analysis), based on the associated phenotypes, clinical urgency, and incidental findings, NewbornDiscovery categorized the genes as A (candidate genes directly associated with phenotype), B [genes associated with critical conditions according to the 2014 Society for Inherited Metabolic Disorders position statement (<http://www.simd.org/Issues/>) regardless of candidate genes], or C (neither candidate genes nor genes associated with critical conditions) (Fig. 2A). A classification of identified and reported genetic variants was used differentially according to the gene category (Fig. 2B). In the present study, a preliminary report of genes in categories A and B was generated within 5 days. After additional testing to confirm the NGS results, a secondary report for genes in categories A and B was generated within 2 weeks, and final reporting was completed within 3 weeks.

notype-gene analysis), based on the associated phenotypes, clinical urgency, and incidental findings, NewbornDiscovery categorized the genes as A (candidate genes directly associated with phenotype), B [genes associated with critical conditions according to the 2014 Society for Inherited Metabolic Disorders position statement (<http://www.simd.org/Issues/>) regardless of candidate genes], or C (neither candidate genes nor genes associated with critical conditions) (Fig. 2A). A classification of identified and reported genetic variants was used differentially according to the gene category (Fig. 2B). In the present study, a preliminary report of genes in categories A and B was generated within 5 days. After additional testing to confirm the NGS results, a secondary report for genes in categories A and B was generated within 2 weeks, and final reporting was completed within 3 weeks.

### Analysis and interpretation of variants, and validation of NGS data

A customized panel using a hybridization capture approach was designed to enrich DNA fragments spanning the genes of interest. This panel covered whole coding exons (CDS) and



**A**

**Fig. 2.** An integrated workflow design for NBS. The NewbornDiscovery workflow features a 3-step analysis (A) for identification of the reported range of genetic variants (B) in the present study. The reported time and range are determined according to the gene category after step 1 analysis. Category B includes genes associated with critical conditions, according to the 2014 Society for Inherited Metabolic Disorders position statement (<http://www.simd.org/Issues/>), regardless of candidate genes. NBS, newborn screening; SNV, single nucleotide variant; INDEL, small insertions and deletions; CNV, copy number variations; ACMG, American College of Medical Genetics and Genomics; NGS, next generation sequencing; qPCR, quantitative PCR.

Category A

Inheritance patterns	No. of genetic variants identified according to the classification		
	Pathogenic or likely pathogenic	Uncertain significance	Likely benign or benign
Autosomal dominant	≥1	≥1	Not reported
Autosomal recessive	≥1	≥1	Not reported
X-linked dominant	≥1	≥1	Not reported
X-linked recessive	≥1	≥1	Not reported

Category B and Category C

Inheritance patterns	No. of genetic variants identified according to the classification (zygosity)		
	Pathogenic or likely pathogenic	Uncertain significance	Likely benign or benign
Autosomal dominant	≥1	Not reported	Not reported
Autosomal recessive	≥1 (homozygote) ≥2 (heterozygote)	Not reported	Not reported
X-linked dominant	≥1	Not reported	Not reported
X-linked recessive	≥1 (male) ≥1 (female: homozygote) ≥2 (female: heterozygote)	Not reported	Not reported

**B**

known pathogenic variants in UTR, introns, or promoters in corresponding genes. In order to adequately cover the CDS regions and detect CNV with high accuracy, the probes of all target regions were tiled at a density higher than 3× on average. Genomic DNA was extracted from DBS (total 2 circles measuring 1 cm in diameter; total 4 paper punches with 3.0–3.5 mm in diameter in detail). DNA was isolated on a Chemagic 360-D instrument (PerkinElmer, Baesweiler, Germany), based on magnetic bead technology, using a Chemagic STAR DNA Blood200 kit (PerkinElmer, Waltham, MA, USA) according to the manufacturer's instructions. DNA concentrations were determined using Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). The total genomic DNA yield of each sample was approximately 100–200 ng.

Isolated genomic DNA was fragmented, bar-coded with library adapters, and incubated with oligonucleotide probes. Sequencing was performed with 2×150 bp NextSeq runs (Illumina, San Diego, CA, USA). The NGS quality control matrix was at least 200× mean reads on target with a minimum of 20× coverage on >99.9% target bases. We used an integrated clinical interpretation platform, based on the NewbornDISCOVERY workflow for variant analysis, according to the 2015 ACMG standards/guidelines.<sup>7</sup>

The sequence reads were aligned to the human reference genome (GRCh37/hg19) with decoy sequences (hs37d5) using BWA-MEM (version 0.7.12) with the default option.<sup>8</sup> The SNVs and INDELS were called using the Genome Analysis Toolkit HaplotypeCaller (version 3.4.46).<sup>9</sup> Variant calling was performed with a minimum of 10 reads. Minimum mapping and base quality scores of 20 and 10, respectively, were required to call a variant. The variants were annotated using ANNOVAR.<sup>10</sup>

CNVs were called using DeviCNV (version 1.0, SD Genomics, manuscript in preparation) developed for exon level CNV detection from targeted gene panel sequencing data, based on read depth analysis. DeviCNV provides a score summarizing quality (confidence intervals of copy number ratios, number of probes supporting the candidates, regression qualities, etc.) for each CNV candidate, and plots for target genes containing copy number ratio with confidence interval of each target probe to help users select clinically relevant CNV candidates for further validation.

To obtain the frequencies of variants in large populations, we used five population databases: the Exome Aggregation Consortium (ExAC; <http://exac.broadinstitute.org/>), the Exome Variant Server (EVS; <http://evs.gs.washington.edu/EVS/>), the 1000-Genome Project (1000 GP; <http://www.internationalgenome.org/>), the Korean Reference Genome Database (KRGDB; <http://152.99.75.168/KRGDB/menuPages/intro.jsp>), and the dbSNP (<https://www.ncbi.nlm.nih.gov/SNP/>). Locus-specific databases, such as the ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), the Human Genome Mutation Database (HGMD; <http://www.hgmd.org/>), and the Leiden Open Variation Database (LOVD; <http://www.lovd.nl/3.0/home>), were

used to assess the supporting evidence of benign impact or pathogenicity (BP6 or PP5 according to the 2015 ACMG standards/guidelines). To determine possible structural or functional damage due to a missense mutation, we used multiple lines of *in silico* software such as SIFT,<sup>11</sup> PolyPhen-2,<sup>12</sup> MutationTaster,<sup>13</sup> and FATHMM.<sup>14</sup> We also used dbSNV to predict whether a variant affects alternative splicing.<sup>15</sup>

We reviewed previously published literature for functional analysis of variants (BS3, PM1, or PS3), population frequencies (BS1 or PS4), or segregation (BS4 or PP1), using an integrated interpretation platform based on expert manual reviews. Finally, we classified variants into pathogenic, likely pathogenic, of uncertain significance, likely benign, or benign categories.

All suspicious SNVs or INDELS associated with patients' disorders were validated with Sanger sequencing. In the case of abrupt decrease in the sequencing depth of specific exons using DeviCNV, we performed exon-based quantitative PCR to detect and validate large deletions containing one or several exons. Quantitative real-time PCR (qPCR) was performed using a QuantStudio 3 real-time PCR system (Applied Biosystems, Foster City, CA, USA). Specific primers were designed for targeted exons and *GAPDH* (as a control housekeeping gene). All qPCR reactions were performed in a 20 µL final volume with 500 nM each of PCR primers, 2 µL template DNA, and POWER UP SYBR master mix (Applied Biosystems). The qPCR was carried out with an initial denaturation step at 95°C for 10 min, followed by 40 cycles of denaturation at 95°C for 15 s and annealing/extension at 60°C for 1 min. The segments of the targeted exons and *GAPDH* were amplified in triplicate for each DNA sample derived from both patient and control subjects. The relative quantity of targeted exons was determined using the 2<sup>-ΔΔCt</sup> method,<sup>16</sup> and the results were normalized against *GAPDH* as an endogenous control.

## RESULTS

### NGS analysis

The average reads on target of NewbornDISCOVERY gene panel were 414×. In addition, the NewbornDISCOVERY gene panel yielded 99.9% coverage at 20× coverage and 95.4% coverage at 100× coverage. NGS analysis results are shown in Supplementary Table 2 (only online).

### Clinical and biochemical characteristics of Group 1: the clinical diagnosis group

Ten patients in Group 1 were previously diagnosed via biochemical analysis (Table 1). Six of them presented with typical symptoms and signs in accordance with underlying LSDs, three (CD<sub>6–8</sub>) with Pompe's disease (OMIM#232300), two (CD<sub>3, 4</sub>) with Morquio A disease (OMIM#253000), and one (CD<sub>9</sub>) with Gaucher's disease (OMIM#230800). They were clinically diagnosed, based on an analysis of specific enzyme

activity and/or pathologic findings of affected tissues at the mean age of 52±49 months (range 8 months–11 years). All patients were undergoing treatment with enzyme replacement therapy at the time of enrollment. The other four patients showed abnormal NBS results without manifesting symptoms or signs, and were clinically diagnosed, based on biochemical studies of plasma and urine metabolites at the mean age of 29±11 days (range 20–42 days). Two patients diagnosed with organic acidemia, CD\_1 with methylmalonic acidemia (OMIM# 251000) and CD\_5 with isovaleric acidemia (IVA, OMIM# 243500), were managed with a low protein diet with special milk. CD\_2 with methionine adenosyltransferase I/III deficiency (MATD, OMIM#250850) was followed without methionine restriction. CD\_10 with short chain acyl-CoA dehydrogenase deficiency (SCADD, OMIM#201470) was managed with riboflavin supplementation together without prolonged fasting. All the 10 patients were followed for 58±28 months and showed normal neuropsychological development.

### Molecular genetic analysis of group 1 patients based on NewbornDiscovery

We identified causative genetic variants in all 20 alleles derived from the Group 1 patients (20/20), and the detection rate of NewbornDiscovery in this group was 100% (Table 1). The clinical significance of all identified causative genetic variants was classified as pathogenic (13/20) or likely pathogenic (7/20) according to the 2015 ACMG standards/guidelines.<sup>7</sup> No allele in this group was classified as a variant of uncertain significance in the genes under category A.

Eight patients were compound heterozygotes, and all the identified variants were confirmed *trans* by parental testing. The other two patients were homozygous for either c.466-3\_466-2delinsGG on *IVD* (CD\_5) or p.(Pro55Leu) on *ACADS* (CD\_10).

### Clinical and biochemical characteristics of group 2 undiagnosed in the pilot study group

All 10 subjects enrolled in this group were newborns with abnormal NBS results (Table 2). The mean age at enrollment was 10.4±5.5 days (range 3–21 days), the mean gestational age was 38.9±1.3 weeks, and no premature infants were enrolled. Their mean birth weight was 3.2±0.4 kg, and all newborns except one (UD\_8 with a birth weight of 2.52 kg) showed an appropriate weight for their gestational age. We assessed these patients biochemically using conventional clinical and molecular diagnostic approaches simultaneously according to the protocol indicated in Fig. 1B.

Based on the results of biochemical assays, five patients were diagnosed with possible organic acidemias: UD\_8 with propionic academia (OMIM#606054), UD\_3 and 5 with IVA, and UD\_1 and 4 with multiple acyl-CoA dehydrogenase deficiency (MADD, OMIM#231680). Three patients tested positive for a diagnosis of aminoacidopathies: UD\_7 with argininosuccinic aciduria (ASLD, OMIM#207900), UD\_6 with MATD, and

UD\_10 with hyperprolinemia (OMIM#239500, #239510). Two patients were suspected to have fatty acid oxidation disorders: UD\_2 with SCADD and UD\_9 with systemic carnitine deficiency (SCD, OMIM#212140) or carnitine palmitoyltransferase 1A deficiency (CPT1D, OMIM#255120). In case of UD\_9, the findings in the results obtained from initial NBS varied from the follow-up quantitative plasma carnitine profiles. Although this female patient was referred with an abnormally low free carnitine level on NBS, her quantitative plasma carnitine profiles showed a low acylcarnitine level, but normal total and free carnitine levels. These results were inconsistent with the diagnosis of SCD, but suggested the possibility of CPT1D.

Although 7 of 10 newborns showed no subjective symptoms or signs related to specific IMD, 3 newborns presented with clinical symptoms and signs of metabolic crisis even before NBS results were obtained (60 h after birth for UD\_3, 18 h for UD\_4, and 58 h for UD\_8). These 3 patients required neonatal intensive care for 5–17 days, and UD\_4 with MADD died at 20 days due to unresolved metabolic acidosis, hyperammonemia, respiratory difficulty, and loss of normal reflexes. We initiated and continued with general and disease-specific management according to presumptive IMDs, based on biochemical findings, in order to prevent the onset or worsening of complications, until confirmation or exclusion of specific IMDs. Except for UD\_4 who died in the neonatal period, the other newborns manifested normal developmental milestones without any acute metabolic crisis, although the average duration of follow-up was only 9 to 12 months in this group.

### Molecular genetic analysis by NewbornDiscovery of group 2 patients

We found causative genetic variants in 17 of 20 alleles in 9 of 10 Group 2 patients, and the disease confirmation rate of NewbornDiscovery in this group was 90% (Table 2). UD\_6 was diagnosed with an autosomal dominant form of MATD carrying a single heterozygous variant. The other eight subjects (UD\_1–5, UD\_7–9) were diagnosed with autosomal recessive inherited disorders with two independently segregating mutant alleles (1 from each parent). Parental testing revealed 4 variants involving UD\_1 (c.831+3A>C in *ETFDH*), UD\_4 [p.(Arg452Gly) in *ETFDH*], UD\_7 [p.(Pro156Leu) in *ASL*], and UD\_9 [p.(Leu97Val) in *SLC22A5*], which were confirmed *trans* with a pathogenic variant on the other allele. Accordingly, the ACMG categories were adjusted for uncertain significance to likely pathogenic status. Therefore, the clinical significance of all identified causative genetic variants was classified as pathogenic (11/17) or likely pathogenic (6/17).<sup>7</sup> Three patients (UD\_1, UD\_4, and UD\_7) showed a deletion allele across one to seven exons of the causing gene, which was confirmed by quantitative PCR. The clinical significance of these CNVs was classified as pathogenic.<sup>7</sup>

Two IVA patients (UD\_3 and 5) in group 2 carried the same homozygous pathogenic variant in *IVD*, c.466-3\_466-2delin-

sGG, as CD\_5 from group 1. Furthermore, the UD\_6 harbored a variant, p.(Arg249Gln), which was also identified in CD\_2.

There was only one patient (UD\_10) who was not molecularly diagnosed via NewbornDiscovery analysis in the present study. Hyperprolinemia type 1 or 2 was suspected based on elevated proline levels (2.1–2.4 fold higher than the cut-off value) in repeated NBSs. However, no meaningful variant in *PRODH* or *ALDH4A1* was detected using NewbornDiscovery, and no suspicious evidence of large deletions affecting both alleles was found from NewbornDiscovery data. Interestingly, the plasma proline levels measured in our hospital showed a tendency to decrease without any abnormality in other metabolic and chemistry studies. The high proline levels on NBS might have been transient and not pathologically correlated.

## DISCUSSION

Traditionally, the medical screening protocol for newborns with abnormal NBS is shown in Fig. 1A. Recurrent abnormal results detected on NBS warrant follow-up biochemical tests for diagnosis and management. Follow-up biochemical tests usually include plasma amino acid analysis, urine organic acid analysis, specific enzyme assays, and measurements of toxic metabolites such as ammonia, ketones, and lactic acid. However, discrepancies between the results of biochemical tests have often been observed, and biochemical enzyme assays are labor-intensive and tedious, and yield semiquantitative results.<sup>2,17</sup> In such cases, Sanger sequencing is the next step employed to confirm and differentiate the diagnosis. However, Sanger sequencing is also expensive and is hampered by the genetic heterogeneity of IMDs, which may result in delayed confirmation of diagnosis. In addition, although tandem mass spectrometry (TMS)-based NBS facilitates simultaneous detection of approximately 40–50 metabolites, new therapeutic agents have been developed for the treatment of various IMDs, many of which are not currently included in the NBS lists.<sup>18</sup>

Multi-gene panels based on NGS techniques represent a potential alternative and complement the current NBS. They are associated with minimal ethical implications related to whole exome/genome sequencing by avoiding secondary findings or findings that cannot be interpreted. NBS should be targeted to diseases that are clinically well defined and medically actionable. Due to possible secondary findings associated with genome-based tests that are independent of the original purpose of NBS, a targeted gene analysis involving multi-gene panel is recommended to integrate NGS-based tests with the NBS.<sup>19</sup> The NewbornDISCOVERY analysis is focused on both diagnostic and proactive applications, and provides uncertain to positive results of suspected disorders (in the gene category A) for the diagnostic use. Incidental findings, which are variants in the genes under categories B and C, are reported with

clear results for proactive use: 1 or more pathogenic or likely pathogenic variants for autosomal dominant disorders, 1 homozygous pathogenic/likely pathogenic variant or 2 heterozygous pathogenic/likely pathogenic variants for autosomal recessive disorders.

The application of NGS techniques to NBS is technically feasible. Recent studies show that DNA can be successfully extracted from DBS for conventional NBS,<sup>4,6</sup> and one of the studies reported an analytical sensitivity of 99.8%.<sup>4</sup> However, the NGS analysis and interpretation are still insufficient and difficult. Above all, the difficulty in interpreting variants of uncertain significance (VUS) should be considered, because the database of genotype-phenotype correlation is insufficient at present. Furthermore, these VUS are detected frequently (Supplementary Table 2, only online). In addition, it is well known that clinical phenotype varies even among family members affected with the same genotype.

The general price of clinical NGS testing (e.g., approximately 500–3000 USD for clinical NGS panel testing) depends on the number of genes analyzed, the range of interpretation or consultation, or inclusion of secondary validation testing. It exceeds 2000 USD for clinical whole exome sequencing and is higher than NBS (approximately 100–200 USD), biochemical assays (approximately 100–500 USD), or targeted Sanger sequencing (approximately 500–1500 USD for whole coding exons) alone. However, if a patient is finally diagnosed with a specific IMD, the new integrated approach (NBS combined with NGS testing) reduces the total diagnostic cost compared with the conventional approach. It is possible to confirm the specific disease earlier, if NGS is implemented at an appropriate stage of the diagnostic process. In addition, it covers most IMDs at low cost, compensates for false-positive results of current NBS, and leads to the identification of new causative genetic variants.<sup>4,5,20</sup>

In this study, we used a newly introduced analysis workflow, NewbornDiscovery, which comprises 259 genes including those associated with newly treatable or actionable diseases (Supplementary Table 1, only online). In sequence data obtained through multi-gene panel sequencing, many heterozygous variants that are not associated with the patient's actual disease are concurrently identified. Biochemical data are indispensable for the interpretation of the clinical significance of these identified variants and also for the exclusion of or focus on specific diseases. By executing these two different elements of the workflow in parallel and comparing the results of analysis with each other, we not only achieved a higher rate of disease confirmation in a shorter time frame, but also initiated management of a specific IMD earlier and provided accurate genetic counseling to each patient.

In the present study, both patients diagnosed with MADD (UD\_1 and 4) harbored different novel sequence variants in an *ETFDH* allele. The c.831+3A>C from UD\_1 was predicted to affect RNA splicing by *in silico* analyses,<sup>15</sup> although analysis

of cDNA constructed from the patient's RNA was necessary for confirmation. The p.(Arg452Gly) from UD\_4 is a missense variant. In a previous report, a MADD type C patient harbored a variant at the same amino acid, p.(Arg452Lys). However, the clinical significance of this variant has yet to be identified.<sup>21</sup> Therefore, based on sequence data alone, these two variants were initially interpreted as VUS.<sup>7</sup> However, abnormal biochemical findings to suggest MADD may be considered concurrently using the NewbornDiscovery approach. The NGS data were meticulously investigated focusing on *ETFDH*, and a large deletion (exons 1–7 for UD\_1 and exons 7–8 for UD\_4) was quickly found on the other allele using DeviCNV. Finally, the variants c.831+3A>C and p.(Arg452Gly) were re-classified as likely pathogenic after confirmation of their location in *trans* to the deletion allele by parental testing.<sup>7</sup> The p.(Pro156Leu) on *ASL* from UD\_7 was a novel variant, and on another variant in the same amino acid, p.(Pro156Arg) has been reported, in an ASLD patient.<sup>22</sup> By identifying a deletion in the other allele using the NewbornDiscovery, we established the molecular diagnosis of ASLD. In total, 3/37 mutant alleles in this study (8.1%) were identified as deletion CNVs.

In the present study, we found additional benefits of NewbornDiscovery, shown in the UD\_9 case. The results of the patient's TMS and plasma carnitine profiles were not consistent, and the two suspected conditions, SCD and CPT1D, required different medical management. L-carnitine supplementation is the mainstay of management for SCD,<sup>23</sup> whereas CPT1D requires a low-fat and high-carbohydrate diet instead of L-carnitine supplementation.<sup>24</sup> Based on NewbornDiscovery, we found two causative genetic variations in *SLC22A5* and confirmed the diagnosis of SCD within 2 weeks (and a preliminary report within 5 days), thus facilitating prompt disease-specific treatment for UD\_9. The likely pathogenic variant, p.(Leu97Val) located in the first extracellular loop, existed in *trans* to the pathogenic variant, p.(Trp132\*). The Asp91 site near Leu97 is a protein glycosylation site, and normal glycosylation plays an important role in carnitine recognition. Furthermore, a known pathogenic variant near Asp91, p.(Arg83Leu), impairs glycosylation and transfer of the mature protein to the plasma membrane.<sup>25</sup>

Persistent hypermethioninemia is a unique IMD, characterized by both autosomal dominant and autosomal recessive inheritance depending on the type of causative variants. In the autosomal dominant type, p.(Arg264His) is the most frequently identified variant worldwide,<sup>25</sup> whereas p.(Arg249Gln) is commonly found in Korean patients.<sup>26</sup>

A majority of patients, especially those carrying the autosomal dominant type, show no clinical abnormalities. Mean plasma methionine concentrations show a positive correlation with brain abnormalities on imaging and regular monitoring of plasma methionine levels is recommended.<sup>25</sup> In the present study, CD\_2 and UD\_6 were diagnosed with autosomal recessive and autosomal dominant forms of persistent hyperme-

thioninemia, respectively. The p.(Arg249Gln) was identified in both patients, however, CD\_2 carried another recessively inherited variant, p.(Arg299Cys),<sup>27</sup> and showed higher methionine levels than those of UD\_6. An understanding of the genetic subtypes via NewbornDiscovery facilitates appropriate individual genetic counseling, and prognostic evaluation of the disease.

In the present study, we identified recurrent IMD-related causative variants in Koreans. The c.466-3\_466-2delinsGG is the only individual *IVD* variant found in this study, and all three IVA patients were homozygous for this variant. The c.466-3\_466-2delinsGG on *IVD* was first detected in Korean IVA patients and has been shown to be the most common *IVD* causative variant identified in Korea.<sup>28</sup> Considering that each patient is independent of the others and born out of non-consanguineous marriage, it may represent a founder effect. Two variants associated with *ACADS* [p.(Pro55Leu) and p.(Glu344Gly)] and detected as homozygotes in CD\_10 and UD\_2, have also been recurrently identified in Korean and Japanese SCADD patients.<sup>29-31</sup>

In conclusion, the newly developed multi-gene panel, NewbornDiscovery, and the new diagnostic approach combined with biochemical data yielded satisfactory results in both disease confirmation and exclusion. A small amount of blood obtained from a minimally invasive procedure is sufficient to complete the analyses, even in newborns. This approach not only compensates for the disadvantages of current NBS, but also reduces medical costs and socioeconomic losses associated with delayed diagnosis of IMDs. Considering the developmental direction of NBS expansion, we expect that simultaneous testing of TMS and multi-gene panel will become the first-tier NBS in the future.

## ACKNOWLEDGEMENTS

The study was supported by grant A120030 from the Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea.

We express our gratitude to the patients and their parents for participation in this study.

## ORCID

Jung Min Ko <https://orcid.org/0000-0002-0407-7828>  
Kyung-A Lee <https://orcid.org/0000-0001-5320-6705>  
Jong-Won Kim <https://orcid.org/0000-0002-0708-9242>

## REFERENCES

1. Park KJ, Kim JW. Perspectives on next-generation newborn screening. *Lab Med Online* 2015;5:169-75.
2. Mak CM, Lee HC, Chan AY, Lam CW. Inborn errors of metabolism and expanded newborn screening: review and update. *Crit Rev Clin Lab Sci* 2013;50:142-62.
3. Schulze A, Lindner M, Kohlmüller D, Olgemöller K, Mayatepek E,

- Hoffmann GF. Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics* 2003;111(6 Pt 1):1399-406.
4. Bhattacharjee A, Sokolsky T, Wyman SK, Reese MG, Puffenberger E, Strauss K, et al. Development of DNA confirmatory and high-risk diagnostic testing for newborns using targeted next-generation DNA sequencing. *Genet Med* 2015;17:337-47.
  5. Bodian DL, Klein E, Iyer RK, Wong WS, Kothiyal P, Stauffer D, et al. Utility of whole-genome sequencing for detection of newborn screening disorders in a population cohort of 1,696 neonates. *Genet Med* 2016;18:221-30.
  6. Hollegaard MV, Grauholm J, Nielsen R, Grove J, Mandrup S, Hougaard DM. Archived neonatal dried blood spot samples can be used for accurate whole genome and exome-targeted next-generation sequencing. *Mol Genet Metab* 2013;110:65-72.
  7. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405-24.
  8. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589-95.
  9. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297-303.
  10. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
  11. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073-81.
  12. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-9.
  13. Schwarz JM, Cooper DN, Schuelke M, Seelow D. Mutation-Taster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014;11:361-2.
  14. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat* 2013;34:57-65.
  15. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res* 2014;42:13534-44.
  16. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>( $\Delta\Delta C_T$ ) Method. *Methods* 2001;25:402-8.
  17. Kim YM, Kim JH, Choi JH, Kim GH, Kim JM, Kang M, et al. Determination of autosomal dominant or recessive methionine adenosyltransferase I/III deficiencies based on clinical and molecular studies. *Mol Med* 2016;22:147-55.
  18. Chien YH, Abdenur JE, Baronio F, Bannick AA, Corrales F, Couce M, et al. Mudd's disease (MAT I/III deficiency): a survey of data for MAT1A homozygotes and compound heterozygotes. *Orphanet J Rare Dis* 2015;10:99.
  19. Fernández-Irigoyen J, Santamaría E, Chien YH, Hwu WL, Korman SH, Faghfoury H, et al. Enzymatic activity of methionine adenosyltransferase variants identified in patients with persistent hypermethioninemia. *Mol Genet Metab* 2010;101:172-7.
  20. Lee YW, Lee DH, Vockley J, Kim ND, Lee YK, Ki CS. Different spectrum of mutations of isovaleryl-CoA dehydrogenase (IVD) gene in Korean patients with isovaleric acidemia. *Mol Genet Metab* 2007;92:71-7.
  21. Kim YM, Cheon CK, Park KH, Park S, Kim GH, Yoo HW, et al. Novel and recurrent ACADS mutations and clinical manifestations observed in Korean patients with short-chain acyl-coenzyme A dehydrogenase deficiency. *Ann Clin Lab Sci* 2016;46:360-6.
  22. Shirao K, Okada S, Tajima G, Tsumura M, Hara K, Yasunaga S, et al. Molecular pathogenesis of a novel mutation, G108D, in short-chain acyl-CoA dehydrogenase identified in subjects with short-chain acyl-CoA dehydrogenase deficiency. *Hum Genet* 2010;127:619-28.
  23. Magoulas PL, El-Hattab AW. Systemic primary carnitine deficiency: an overview of clinical manifestations, diagnosis, and management. *Orphanet J Rare Dis* 2012;7:68.
  24. Bennett MJ, Santani AB. Carnitine palmitoyltransferase 1A deficiency. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, et al., editors. *GeneReviews*(R). Seattle (WA): University of Washington; 1993.
  25. Goodman SI, Binard RJ, Woontner MR, Frerman FE. Glutaric acidemia type II: gene structure and mutations of the electron transfer flavoprotein:ubiquinone oxidoreductase (ETF:QO) gene. *Mol Genet Metab* 2002;77:86-90.
  26. Filippo CA, Ardon O, Longo N. Glycosylation of the OCTN2 carnitine transporter: study of natural mutations identified in patients with primary carnitine deficiency. *Biochim Biophys Acta* 2011;1812:312-20.
  27. Wiltink RC, Kruijshaar ME, van Minkelen R, Onkenhout W, Verheijen FW, Kemper EA, et al. Neonatal screening for profound biotinidase deficiency in the Netherlands: consequences and considerations. *Eur J Hum Genet* 2016;24:1424-9.
  28. Therrell BL, Padilla CD, Loeber JG, Kneisser I, Saadallah A, Borrajo GJ, et al. Current status of newborn screening worldwide: 2015. *Semin Perinatol* 2015;39:171-87.
  29. Howard HC, Knoppers BM, Cornel MC, Wright Clayton E, Sénécal K, Borry P, et al. Whole-genome sequencing in newborn screening? A statement on the continued importance of targeted approaches in newborn screening programmes. *Eur J Hum Genet* 2015;23:1593-600.
  30. Berg JS, Powell CM. Potential uses and inherent challenges of using genome-scale sequencing to augment current newborn screening. *Cold Spring Harb Perspect Med* 2015;5:a023150.
  31. Kim SH, Park HD, Sohn YB, Park SW, Cho SY, Ji S, et al. Mutations of ACADS gene associated with short-chain acyl-coenzyme A dehydrogenase deficiency. *Ann Clin Lab Sci* 2011;41:84-8.