ANNALS OF LABORATORY MEDICINE

Rapid Targeted Sequencing Using Dried Blood Spot Samples for Patients With Suspected Actionable Genetic Diseases

Man Jin Kim, M.D.[1,2,*], Soo Yeon Kim, M.D.[1,2,*], Jin Sook Lee, M.D.[3], Sanggoo Kang, M.B.A.[4], Lae-Jeong Park, Ph.D.[5], Wooyong Choi, B.S.[4], Ju Yeol Jung, M.S.[4], Taehyung Kim, M.S.[4], Sung Sup Park, M.D.[6], Jung Min Ko, M.D.[2,3], Moon-Woo Seong, M.D., Ph.D.[6], and Jong Hee Chae, M.D., Ph.D.[1,2,7]

[1]Department of Genomic Medicine, Seoul National University Hospital, Seoul, Korea; [2]Rare Disease Center, Seoul National University Hospital, Seoul, Korea; [3]Department of Pediatrics, Department of Genome Medicine and Science, Gil Medical Center, Gachon University College of Medicine, Incheon, Korea; [4]MedySapiens, Inc., Seoul, Korea; [5]Department of Electrical Engineering, Gangneung-Wonju National University, Gangneung, Korea; [6]Department of Laboratory Medicine, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Korea; [7]Department of Pediatrics, Seoul National University Children's Hospital, Seoul National University College of Medicine, Seoul, Korea

Background: New genome sequencing technologies with enhanced diagnostic efficiency have emerged. Rapid and timely diagnosis of treatable rare genetic diseases can alter their medical management and clinical course. However, multiple factors, including ethical issues, must be considered. We designed a targeted sequencing platform to avoid ethical issues and reduce the turnaround time.

Methods: We designed an automated sequencing platform using dried blood spot samples and a NEOseq_ACTION panel comprising 254 genes associated with Mendelian diseases having curable or manageable treatment options. Retrospective validation was performed using data from 24 genetically and biochemically confirmed patients. Prospective validation was performed using data from 111 patients with suspected actionable genetic diseases.

Results: In prospective clinical validation, 13.5% patients presented with medically actionable diseases, including short- or medium-chain acyl-CoA dehydrogenase deficiencies (N=6), hyperphenylalaninemia (N=2), mucopolysaccharidosis type IVA (N=1), alpha thalassemia (N=1), 3-methylcrotonyl-CoA carboxylase 2 deficiency (N=1), propionic acidemia (N=1), glycogen storage disease, type IX(a) (N=1), congenital myasthenic syndrome (N=1), and citrullinemia, type II (N=1). Using the automated analytic pipeline, the turnaround time from blood collection to result reporting was <4 days.

Conclusions: This pilot study evaluated the possibility of rapid and timely diagnosis of treatable rare genetic diseases using a panel designed by a multidisciplinary team. The automated analytic pipeline maximized the clinical utility of rapid targeted sequencing for medically actionable genes, providing a strategy for appropriate and timely treatment of rare genetic diseases.

Key Words: Neonatal screening, High-throughput nucleotide sequencing, Metabolism, Inborn errors, Dried blood spot

Kim MJ, et al.
Rapid targeted sequencing for genetic diseases

ANNALS OF
LABORATORY
MEDICINE

INTRODUCTION

Next-generation sequencing (NGS)-based approaches for clinical evaluation of infants with suspected genetic diseases have expanded substantially in recent years [1-3]. In particular, whole-genome sequencing (WGS) and whole-exome sequencing (WES) have improved the diagnosis of rare genetic diseases in critically ill neonates and infants. Compared with conventional genetic testing, rapid genomic testing (RGT) permits accurate genetic diagnosis of many patients [4-6]. Recent studies have suggested that RGT is faster and more reliable than standard WES and WGS [7-9]. NGS-based RGT enables rapid and timely diagnosis of rare genetic diseases.

Early diagnosis of incurable genetic diseases and profound disability in newborns raises several ethical issues, such as parents making premature decisions as the bonding with the baby is not fully established [10]. Moreover, RGT poses a substantial clinical and laboratory burden [8]. A successful diagnosis implementation requires changing the existing workflow by a multidisciplinary team, including intensive-care therapists, clinical geneticists, and clinical pathologists. Furthermore, the cost of rapid WGS and WES is two to four times that of standard WGS and WES [10]. The cost of analysis of the large amount of data generated by WGS or WES is inevitably higher than that of conventional gene sequencing [11]. Many related findings and variants of uncertain significance (VUSs) increase the risk of inaccurate diagnosis and the turnaround time (TAT) of result interpretation. Therefore, WGS or WES in infants and children with genetic diseases is not a viable option, especially in developing countries.

If actionable genes are targeted, ethical approval can be waived, which reduces the TAT. Applying variant interpretation and pathogenicity prediction models can further reduce the TAT and the risk of inaccurate diagnosis. Therefore, we designed an RGT platform to specifically sequence 254 medically actionable genes associated with preventable or curable Mendelian diseases using dried blood spot (DBS) samples of patients.

MATERIALS AND METHODS

Patient enrollment
Patients were prospectively enrolled based on the following criteria: (1) abnormalities in routine neonatal screening testing (NST); (2) unexplained neonatal hypotonia or neonate-onset seizures without exponential causes such as prematurity, malformations, and perinatal insults; (3) unexplained and abnormal laboratory findings, such as liver enzymes, metabolic acidosis, and hyperammonemia, suggesting hidden metabolic diseases; (4) skeletal dysplasia or joint problems suggestive of storage diseases; and (5) neurodevelopmental delay with abnormal results in metabolic screening testing. This study was approved by the Institutional Review Board of Seoul National University Hospital, Seoul, Korea (No. H-2002-069-1101). Informed consent for genetic testing was obtained from all patients and their guardians in the case of minor patients.

NEOseq_ACTION panel design
The NEOseq_ACTION panel was designed to diagnose genetic diseases in neonates or during early infancy that require a rapid and precise diagnosis to ensure prompt and efficient treatment. The NEOseq_ACTION is a comprehensive NGS panel that includes 254 genes that were selected based on consultations with clinical and metabolic geneticists. The selected genes include potentially actionable genes associated with diseases exposed by neonatal metabolic screening performed in Korea and diseases listed on the US Recommended Uniform Screening Panel Core and Secondary Conditions of Advisory Committee on Heritable Disease in Newborns and Children (https://www.hrsa.gov/advisory-committees/heritable-disorders/rusp/index.html).

The diseases targeted in this panel can be classified as: glycogen storage diseases; adrenal insufficiency; galactosemia; inborn errors of amino acid, organic acid, and fatty acid metabolism; treatable or curable lysosomal storage diseases; immunodeficiency; hemoglobinopathy; genetic hyperbilirubinemia; congenital and genetic sensorineural hearing loss; and other genetic diseases, such as cystic fibrosis, mineral metabolism disorders, and congenital myasthenic syndromes (curable diseases: Table 1; full list: Supplemental Data Table S1).

Retrospective validation of the NEOseq_ACTION panel
The validation set included data from 24 patients with known pathogenic single nucleotide variants (SNVs) and small insertion or deletions (indels) across 22 genes (Supplemental Data Table S2). All validation samples were confirmed through biochemical testing. In total, 99.8% and 98.96% of bases were covered by at least 1- and 20-fold coverage, respectively. The median percentage of on-target reads across samples was 78.2%. The mean depth of coverage was 142×. All 40 variants in the 24 cases were detected using the NEOseq_ACTION panel.

Sample processing, target enrichment, and sequencing
Genomic DNA was extracted from two out of three DBSs (Fig. 1)

ANNALS OF
LABORATORY
MEDICINE

Kim MJ, et al.
Rapid targeted sequencing for genetic diseases

Table 1. Overview of the genes included in the NEOseq_ACTION panel, with associated diseases and curative medical actions

| Category	Disease	Gene symbol	Action
Inborn errors of metabolism of amino acids, organic acids, and fatty acids	Biotinidase deficiency	*BTD*	Biotin supplementation
	Carnitine deficiency	*SLC22A5*	L-carnitine supplementation
	Holocarboxylase synthetase deficiency	*HLCS*	Biotin supplementation
	Homocystinuria	*MTHFR*	Vitamin B6, folic acid
	Hyperphenylalaninemia	*GCH1, PTS, QDPR*	Tetrahydrobiopterin, L-DOPA, 5-hydroxytryptophan supplementation
	Isolated methylmalonic acidemia	*MMAA, MMAB, MMADHC*	Hydroxycobalamin supplementation
	Methylmalonic acidemia with homocystinuria	*MMACHC, MMADHC*	Hydroxycobalamin supplementation
	N-acetylglutamate synthetase deficiency	*NAGS*	Carglumic acid (Carbaglu) treatment
Lysosomal storage diseases	Cystinosis	*CTNS*	Cysteamine treatment
	α-Mannosidosis	*MAN2B1*	Stem cell transplantation
	Aspartylglucosaminuria	*AGA*	Stem cell transplantation
	Other lysosomal storage diseases	*ARSB, GAA, GALNS, GBA, GLA, IDS, IDUA, LIPA, TPP1*	Enzyme replacement therapy
Immunodeficiency	Immunodeficiency	*ADA, AK2, CD3D, CD3E, DCLRE1C, FOXN1, IL2RA, IL2RG, IL7R, JAK3, LCK, LIG4, MTHFD1, NFKBIA, NHEJ1, PEPD, PNP, PTPRC, RAG1, RAG2, STK4, TCN2, ZAP70*	Stem cell transplantation
Other diseases	Acrodermatitis enteropathica	*SLC39A4*	Zinc supplementation
	Hypophosphatasia	*ALPL*	Enzyme replacement therapy
	Pyridoxine-dependent epilepsy	*ALDH7A1*	Pyridoxine supplementation
	Thiamine metabolism dysfunction syndrome 2	*SLC19A3*	Biotin and thiamine supplementation
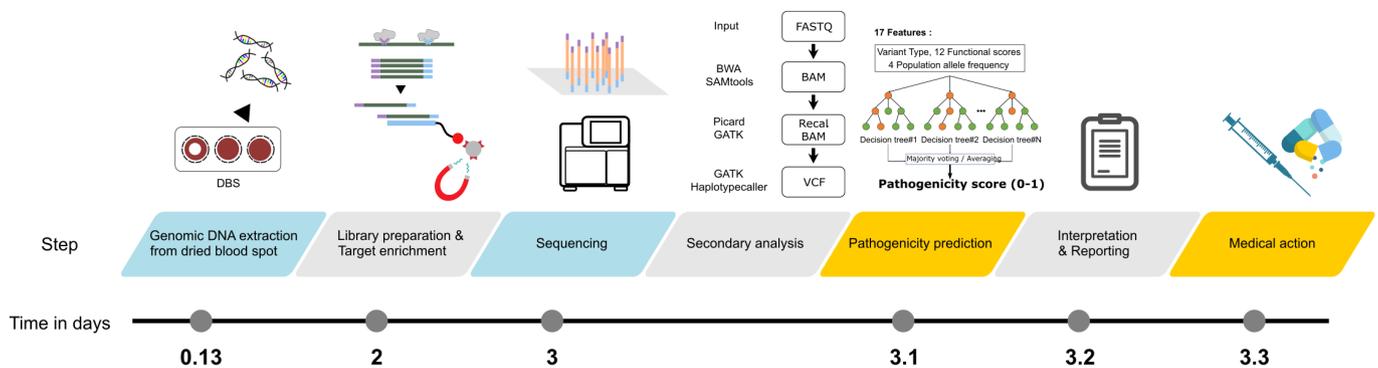
Fig. 1. Workflow for NEOseq_ACTION testing. Genomic DNA is isolated from DBSs, which are commonly used for neonatal screening. Libraries are prepared using the Nextera kit (Illumina, San Diego, CA, USA). The target DNA is enriched using the Twist Bioscience capture kit (Twist Bioscience, South San Francisco, CA, USA) and sequenced on the MiSeq platform (Illumina). Secondary analysis and variant interpretation are carried out using the MedyCVi (MedySapiens, Seoul, Korea) bioinformatics pipeline. Variant pathogenicity is predicted using the MedyCVi module, MedyPatho (MedySapiens). The time from DNA extraction to medical intervention was 3.4 days.
Abbreviations: DBS, dried blood spot; BWA, Burrows–Wheeler aligner; GATK, genome analysis tool kit; BAM, binary alignment map; VCF, variant call format.

using a Chemagic 360 instrument (PerkinElmer, Baesweiler, Germany) and fragmented using a Nextera DNA library preparation kit (Illumina, San Diego, CA, USA). Libraries were prepared from 100 ng of total genomic DNA using the Nextera DNA library preparation kit, according to the manufacturer's instructions. Paired-end 150-bp sequencing was performed on the MiSeqDx plat-

form (Illumina). Raw sequencing data were obtained in the FASTQ format.

Pipeline for variant calling
The sequencing reads were aligned to the human genome reference sequence (GRCh37) using the Burrow–Wheeler Aligner

Kim MJ, et al.
Rapid targeted sequencing for genetic diseases

ANNALS OF
LABORATORY
MEDICINE

(BWA-MEM, version 0.7.5) [12]. The aligned sequence reads (SAM file) were sorted, indexed, and converted to BAM files using SAMtools (version 1.9) [13]. Duplicate sequences in the BAM files were removed using Picard tools (version 2.18.2) (https://broadinstitute.github.io/picard/). Next, local realignment and base recalibration of the BAM files were performed using the Base Recalibrator in GATK (version 4.0.3.0) [14] based on the dbSNP (version 138), Mills, and 1000G gold standard indel databases. SNVs and indels were identified using the GATK Haplotype Caller tool. Called variants were annotated using SnpEff [15].

Variant interpretation and pathogenicity prediction model

Called variants were interpreted using the MedyCVi software (MedySapiens, Seoul, Korea) based on InterVar [16]—a bioinformatics tool for the interpretation of genetic variants according to the American College of Medical Genetics and Genomics and Association for Molecular Pathology guidelines [17] based on databases of disease-associated genetic variants such as ClinVar [18] and Human Genome Mutation Database (HGMD) [19]. A machine learning-based prediction model was developed to support clinical decisions in variant interpretation based on estimation of the pathogenicity of each VUS.

The MedyPatho (MedySapiens) prediction model predicts the pathogenicity of non-synonymous SNVs based on nine functional scores (PolyPhen [20], MutPred [21], Sorting Intolerant From Tolerant [22], Protein Variant Effect Analyzer [23], Variant Effect Scoring Tool [24], Likelihood Ratio Test [25], MutationAssessor [26], MutationTaster [27], Fathmm [28]), two conservation scores (GERP [29] and PhastCons [30]), and the population allele frequency from four databases (Genome Aggregation Database [31], Exome Aggregation Consortium [32], Korean Reference Genome [33], and Korean Variant Archive [34].

The prediction model was trained and evaluated using a dataset compiled from the variant databases of ClinVar and HGMD. Specifically, only missense variants were selected, and for the ClinVar database, variants with a review status labeled as "criteria provided, multiple submitters, no conflicts; criteria provided, single submitter; or reviewed by expert panel" were considered. In addition, pathogenic variants marked as disease-causing mutation or disease-associated polymorphism in the HGMD were included. The final dataset comprised 143,555 pathogenic and 86,326 benign variants.

The variants added to ClinVar since 2020 were considered as test data. Variants for which protein sequences were similar to those in the training set were excluded using the CD-HIT pro-
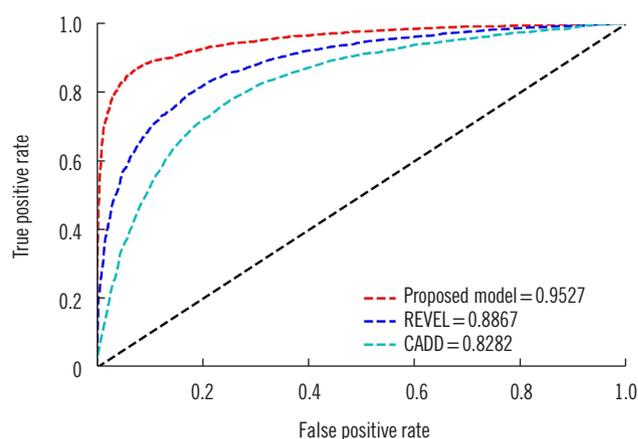

Fig. 2. Performance comparison of pathogenicity prediction tools using test data from ClinVar release Nov 29, 2020.

gram [35] based on a sequence identity threshold of 80%. The prediction model was trained with 43,590 pathogenic and 26,499 benign variants and was subsequently evaluated with 8,295 pathogenic and 8,630 and benign variants.

The prediction model was built using LightGBM [36]—a tree-based gradient boosting model. Performance was evaluated in comparison with two *in silico* prediction tools (Rare Exome Variant Ensemble Learner [REVEL] [37] and Combined Annotation-Dependent Depletion [CADD] [38]). Fig. 2 shows the performance comparison of the pathogenicity prediction tools based on test data from ClinVar release Nov 29, 2020.

All mutations filtered by MedyPatho (MedySapiens) were manually curated by genetic experts.

RESULTS

Clinical patient characteristics

To evaluate the usefulness of the NEOseq_ACTION panel for patients, prospective panel testing was conducted in 111 patients. The clinical characteristics of the patients are shown in Table 2. The mean age at symptom or sign recognition was 10.4 months (range, 1–188 months). Around three quarters of the patients presented with symptoms or showed laboratory abnormalities within a month after birth.

Seventy patients were clinically asymptomatic, and 57 patients were referred to the outpatient clinic because of initial NST abnormality. Patients with NST abnormality were directly enrolled without further investigation, although most of them had simultaneous metabolic screening or NST repetition. Thirteen patients visited the clinic for incidentally detected laboratory abnormalities, such as elevated liver enzymes, creatinine phosphokinase

ANNALS OF
LABORATORY
MEDICINE

Kim MJ, et al.
Rapid targeted sequencing for genetic diseases

Table 2. Demographics of the prospectively enrolled patients

Variable	Total patients, N (%) (N = 111)
Male : Female	58 : 53
Age at symptom/sign recognition	
≤ 1 month	83 (74.8)
1–12 months	10 (9.0)
> 12 months	18 (16.2)
Age at testing	
≤ 1 month	55 (49.5)
1–12 months	23 (20.7)
> 12 months	33 (29.7)
Major symptoms	
Asymptomatic	
NST abnormality	57 (51.4)
LFT abnormality	5 (4.5)
CPK elevation	3 (2.7)
Acidosis	2 (1.8)
Other laboratory abnormalities	3 (2.7)
Symptomatic	
Neonatal hypotonia/respiratory failure	14 (12.6)
Early-onset seizures	3 (2.7)
Developmental delay and additional findings	10 (9.0)
Joint pain/skeletal abnormality	5 (4.5)
Others	9 (8.1)

Abbreviations: NST, neonatal screening testing; LFT, liver function testing; CPK, creatine phosphokinase.

(CPK), or unexplained acidosis. The patients had no perinatal history and showed normal development.

Forty-one patients had one or more clinical symptoms. Neonatal hypotonia and unexplained respiratory failures were the most frequent presentations. However, NST could not provide reliable data because the patients often required mechanical ventilation and total parenteral nutrition because of frequent infections. Among 10 patients with developmental delay, four showed abnormal metabolic screening results, and another four had additional features, including coarse face, skeletal abnormalities, or extensive Mongolian spots suggestive of storage diseases.

Evaluation and validation of the NEOseq_ACTION panel

The average time from patient visit to result reporting was 8.3 days, and the average time from sample preparation to result reporting was 3.4 days. As the MiSeq Reagent Kit v2 Nano was used for less than three samples, sequencing was performed

even if only one sample was submitted. We identified significant variants in 10 genes in 15 patients; detailed information is provided in Table 3. Among them, 11 patients were asymptomatic (10 with NST abnormality, and one with abnormal liver function). Four symptomatic patients (neonatal hypotonia, early-onset seizures, skeletal abnormalities, or abdominal distension) were genetically confirmed. Suspected compound heterozygous or homozygous variants were detected in 13 patients (genes related to autosomal recessive diseases), and two hemizygous variants were detected in two patients. Among the 15 patients with suspected recessive diseases, both alleles were initially classified as pathogenic or likely pathogenic in nine patients (P2, P5, P14, P41, P45, P46, P72, P74, and P83). The other patients carried one pathogenic or likely pathogenic allele and one VUS (P6, P10, P25, P36, P40, and P111). Parental testing was performed in four families, and further biochemical testing was performed for all except two patients. Therefore, the final classification was changed for four patients (P6, P25, P36, and P40). The second variant in P10 and P111 remained of unknown significance after retesting. However, the patients started treatment as soon as the biochemical testing results provided strong evidence for the final diagnosis. As P10 and P111 showed a mildly elevated serum phenylalanine level when compared with actual patients with phenylketonuria on repeated testing, their second variants were considered not significant.

Representative cases

All positive results in 15 cases affected the treatment decisions (Table 4). Two patients received a final diagnosis, although they had received appropriate treatment based on clinical suspicion. Eleven patients started medication or changed their daily diet. All families received counseling based on the patients' results. The average time from testing to medical decision making was 20.8 days (2–68 days). Below, we describe some representative cases.

P2 was a 3.5-year-old girl with no perinatal history. At 12 months, her growth was retarded. At 2 years, she complained of bilateral knee pain, which was followed by gross deformation of several joints. She underwent some tests and received regular rehabilitation, but without clinical improvement. She was referred to our clinic 2 years after the onset of the symptoms and underwent simultaneous NEOseq_ACTION panel sequencing and enzyme testing on suspicion of mucopolysaccharidosis (MPS) type IVA. Compound heterozygous variants in *GALNS* were identified in 5 days, and glucose 6-sulfatase activity was 0.37 nmol/17 hr/mg protein (reference range [ref], 18.6–61.8). The patient was con-

Table 3. Significant variants with pathogenicity classification detected using the NEOseq_ACTION panel

Case No.	Gene symbol	Transcript	Nucleotide change	Amino acid change	Zygosity	Variant classification by MedyPatho	Segregation	Clinical validation	Final classification	ACMG evidences*
P2	GALNS	NM_000512.4	c.319G>A	p.Ala107Thr	Het	LP	Trans	Serum galactose-6-sulfatase 0.37 nmol/17 hr/mg protein (ref, 18.6–61.8)	LP	PM2, PM3, PP3, PP5
	GALNS	NM_000512.4	c.451C>A	p.Pro151Thr	Het	LP			LP	PM1, PM2, PM5, PP3, PP5
P5	ACADS	NM_000017.2	c.1031A>G	p.Glu344Gly	Het	LP	Not done	Serum C4 1.995 µmol/L (cut off, 0.41)	LP	PM1, PM2, PP3, PP5
	ACADS	NM_000017.2	c.164C>T	p.Pro55Leu	Het	LP			LP	PS3, PM2, PP5
P6	PCCA	NM_000282.3	c.1846-2A>G	p.?	Het	LP	Trans	Urine propionylglycine 80.0 (ref, not detected)	LP†	PVS1, PM2
	PCCA	NM_000282.3	c.938G>C	p.Arg313Pro	Het	VUS†			LP†	PM2, PM3, PM5, PP3
P10	PAH	NM_000277.1	c.975C>G	p.Tyr325*	Het	P	Not done	Serum phenylalanine 163 µmol/L (ref, 13–91)	P	PVS1, PM2, PP5
	PAH	NM_000277.1	c.158G>A	p.Arg53His	Het	VUS			VUS	BS2, PM5
P14	SLC25A13	NM_014252.3	c.852_855delTATG	p.Met285Profs*2	Het	P	Not done	Serum citrulline 378 µmol/L (ref, 8–36)	P	PVS1, PM2, PP5
	SLC25A13	NM_014252.3	c.1177+1G>A	p.?	Het	P			P	PVS1, PM2, PP5
P25	ACADM	NM_000016.4	c.1189T>A	p.Try397Asn	Het	LP	Trans	Serum C6 0.145 µmol/L (cut off <0.1), C8 0.26 µmol/L (cut off <0.18), C10:1 0.16 µmol/L (cut off < 0.13)	LP	PM2, PM5, PP2, PP3, PP4, PP5
	ACADM	NM_000016.4	c.1231G>T	p.Val411Leu	Het	VUS†			LP†	PM2, PM3, PP2, PP4
P36	ACADM	NM_000016.4	c.1189T>A	p.Tyr397Asn	Het	LP	Not done	Serum C6 0.64 µmol/L (cut off <0.1), C8 1.59 µmol/L (cut off <0.12), C10:1 0.45 µmol/L (cut off <0.08)	LP	PM2, PM5, PP2, PP3, PP4, PP5
	ACADM	NM_000016.4	c.91C>T	p.Arg31Cys	Het	VUS†			LP†	PM2, PM5, PP2, PP4, PP5
P40	MCCC2	NM_022132.4	c.838G>T	p.Asp280Tyr	Het	LP	Not done	Urine 3-hydroxyisovaleric acid 1,731.1 mmol/mol Cr (ref, <18)	LP	PM2, PP2, PP3, PP4, PP5
	MCCC2	NM_022132.4	c.1342G>A	p.Gly448Arg	Het	VUS†			LP†	PM2, PP2, PP3, PP4, PP5
P41	PREPL	NM_006036.4	c.1979_1980del	p.Leu660Glnfs*9	Het	LP†	Trans	Repetitive nerve stimulation test, insignificant	P†	PVS1, PM2, PM3
	PREPL	NM_006036.4	c.1747-9_1747del	p.?	Het	LP†			P†	PVS1, PM2, PM3
P45‡	HBA2	NM_000517.4	c.427T>C	p.*143Qext*31	Hem	LP	Not done	Not done	LP	PM2, PM3, PM4
P46	PHKA2	NM_000292.2	c.1246-2A>G	p.?	Hem	LP	Not done	GOT/GPT 59/49 IU/L (ref, 1–40/1–40)	LP	PVS1, PM2
P72	ACADS	NM_000017.2	c.164C>T	c.Pro55Leu	Hom	LP	Not done	Serum C4 0.805 µmol/L (ref, <0.41)	LP	PS3, PM2, PP5
P74	ACADS	NM_000017.2	c.1031A>G	p.Glu344Gly	Het	LP	Not done	Serum C4 0.905 µmol/L (ref, <0.41)	LP	PM1, PM2, PP3, PP5
	ACADS	NM_000017.2	c.1130C>T	p.Pro377Leu	Het	LP			LP	PM1, PM2, PP3, PP5
P83	ACADS	NM_000017.2	c.1031A>G	p.Glu344Gly	Hom	LP	Not done	Serum C4 1.965 µmol/L (ref, <0.41)	LP	PM1, PM2, PP3, PP5
P111	PAH	NM_000277.1	c.1068C>G	p.Tyr356*	Het	P	Not done	Serum phenylalanine 173 µmol/L (ref, 25–74)	P	PVS1, PM2, PP5
	PAH	NM_000277.1	c.158G>A	p.Arg53His	Het	VUS			VUS	BS2, PM5

*All variants were assessed based on the ACMG or Association for Molecular Pathology guidelines [17]; †Initial classification using MedyPatho was changed at final classification; ‡Given the high homology between *HBA1* and *HBA2* and that common deletions account for 85% of pathogenic variants in alpha thalassemia, Sanger sequencing and multiplex ligation probe amplification were performed to confirm the variant. This patient had a heterozygous common—SEA deletion; thus, the zygosity of the c.427T>C variant was determined to be hemizygous.

Abbreviations: ACMG, American College of Medical Genetics and Genomics; Het, heterozygous; Hom, homozygous; Hem, hemizygous; LP, likely pathogenic; P, pathogenic; VUS, variant of unknown significance; ref, reference range.

ANNALS OF
LABORATORY
MEDICINE

Kim MJ, et al.
Rapid targeted sequencing for genetic diseases

Table 4. Final diagnosis and medical decision for positive cases

Case No.	Age at symptom onset (months)	Major complaint	Final diagnosis	Time to result reporting (days)	Time to medical decision (days)	Treatment plan
P2	40	Joint pain	Mucopolysaccharidosis IVA	5	12	Enzyme replacement therapy
P5	0	NST abnormality	SCAD deficiency	3	18	Avoid fasting, frequent meals
P6	8	Seizure	Propionic acidemia	3	NA[†]	L-carnitine, low-isoleucine/valine diet
P10	0	NST abnormality	Hyperphenylalaninemia	3	11	Avoid unnecessary treatment, further testing
P14	0	NST abnormality	Citrullinemia, type II	2	NA[†]	Low-carbohydrate and high-protein diet
P25	149	LFT abnormality	MCAD deficiency	2	68	Low-fat diet, avoid fasting
P36	0	NST abnormality	MCAD deficiency	5	19	Low-fat diet, avoid fasting
P40	0	NST abnormality	3-Methylcrotonyl-CoA carboxylase 2 deficiency	2	NA[†]	L-carnitine, avoid fasting
P41	1	Neonatal hypotonia	Congenital myasthenic syndrome	2	2	Acetylcholine esterase inhibitor, salbutamol
P45	0	NST abnormality (incidental finding*)	Alpha thalassemia	10	22	Waiting for stem cell transplantation
P46	21	Abdominal distension	Glycogen storage disease IXa1	10	22	Corn starch, low-fat diet, avoid fasting
P72	0	NST abnormality	SCAD deficiency	4	28	Avoid fasting, frequent meals
P74	0	NST abnormality	SCAD deficiency	9	17	Avoid fasting, frequent meals
P83	0	NST abnormality	SCAD deficiency	4	24	Avoid fasting, frequent meals
P111	0	NST abnormality	Hyperphenylalaninemia	1	8	Avoid unnecessary treatment

*The patient was enrolled because of NST abnormality (elevated C18, C20–C24) and was diagnosed as having alpha thalassemia, unrelated to the initial NST findings. We concluded that the initial NST finding and diagnosis were incidental; [†]Medical action was started before the genetic diagnosis in P6, P14, and P40 based on their symptoms or NST findings.

Abbreviations: NST, neonatal screening testing; LFT, liver function testing; SCAD, short-chain acyl-CoA dehydrogenase; MCAD, medium-chain acyl-CoA dehydrogenase; NA, not applicable.

firmed to have MPS IVA and started enzyme replacement therapy 3 weeks after the first visit.

P14 was a 3-week-old boy whose NST results indicated low arginine (92.6 µmol/L; ref, 120–600) and high citrulline (172.6 µmol/L; ref, 8–47) levels. The patient was admitted to the neonatal intensive care unit and started a diet with special milk and carglumic acid for 2 weeks. The patient was referred for prompt diagnosis, and panel sequencing demonstrated pathogenic compound heterozygous variants in *SLC25A13*. He was diagnosed as having citrullinemia type II and started oral arginine therapy. He is currently 13 months old and shows normal development.

P41 was a 4-month-old boy who presented poor suckling and respiratory difficulty after birth. He was lethargic and depended on the nasogastric tube. He could make eye contact with his family, but not control his head. Panel testing indicated pathogenic compound heterozygous variants in *PREPL* associated with congenital myasthenic syndrome. We attempted repetitive nerve stimulation tests; however, the results were unreliable because of extremely low muscle mass and immature myelination.

He started on an acetylcholine esterase inhibitor and salbutamol and showed remarkable clinical improvement. A month after starting the treatment, he could roll and control his head. He started to crawl at 6 months and could stand without support at 13 months. He has speech and eats ordinary meals orally.

P111 was a 1-month-old boy who showed NST abnormality indicating an elevated phenylalanine level (240 µmol/L; ref, 38–137). He underwent NST thrice because of abnormal results at different hospitals. The panel test indicated a pathogenic variant and a VUS in *PAH*. As detailed metabolic testing indicated slightly increased phenylalanine (100 µmol/L; ref, 25–74) and normal tyrosine (89 µmol/L; ref, 13–91) levels, we concluded that the patient had hyperphenylalaninemia, rather than phenylketonuria. The patient continued on breast milk and showed normal development until the last follow-up.

DISCUSSION

In this proof-of-concept study, we demonstrated the performance

of rapid targeted sequencing using a NEOseq_ACTION panel. Actionable genes were identified in 15 patients (13.5%). For actionable genetic diseases, early therapeutic intervention can allow neonates and infants to have a normal life or can minimize the consequences of the disease. We used the NEOseq_ACTION panel to confirm clinically or biochemically suspected diseases. The panel allowed for the diagnosis of various curable or treatable diseases from a few DBSs obtained during routine NST or for critically ill neonates in intensive care unit settings.

Obstacles to the application of rapid WES and WGS in neonates and young infants, including insufficient sequencing depth, limited bioinformatics, and lack of data interpretation experts, have been reported [8]. Multiple factors have to be considered, such as genetic and clinical heterogeneity, including the appearance of formes frustes or incomplete presentation of classical phenotypes in neonates and young infants, and comorbidity because of the fragility of infants. Therefore, a targeted approach covering 100–300 disease-associated genes is cost-effective for rapid detection of medically actionable diseases. A short TAT is essential for the clinical effectiveness of rapid genome sequencing, particularly in neonates.

VUS interpretation can be a bottleneck in reducing the TAT [39]. In this study, rare VUSs with a population frequency of ≤0.6% appeared 25 times per patient (data not shown). As most inborn errors of metabolism are autosomal recessive genetic diseases, clinical geneticists struggle with many VUSs. Therefore, VUSs are the major cause of a prolonged TAT. The initial classification of four variants in four families (P6, P25, P36, and P40) by MedyPatho was changed in the final classification; two VUSs in P6 and P25 were reclassified as likely pathogenic variants based on the results of segregation testing, and two VUSs in P36 and P40 were reclassified as likely pathogenic variants according to biochemical testing results. MedyPatho has shown clinically acceptable accuracy in the classification of clinically relevant variants. Variant classifiers will become more sophisticated as they include accumulating data from specialized panels, such as NEOseq_ACTION, thereby reducing the burden of variant interpretation and the time required for manual curation. Therefore, automatic variant classifiers, such as MedyPatho, are essential to reduce the TAT.

We conducted a prospective study in patients with suspected actionable diseases. More than half of the patients were asymptomatic neonates with NST abnormalities. Although the standard option was to repeat the testing, the patients may have had several unnecessary tests, diet modification, and extra medical expenses. Moreover, we could take immediate, appropriate medi-

cal action for confirmed cases. Rapid screening and not missing any treatable diseases are critical to achieve qualitative practice, and the panel provided valuable information to clinicians for patients in the intensive care unit.

This study demonstrated that simultaneous biochemical or functional validation reduced the uncertainty of VUSs. Diseases recently covered in NST panels, such as genetic hearing loss and critical arrhythmias, may require regular check-ups for confirmative diagnosis if patients harbor a VUS. Although numerous ethical and social issues are to be considered, these diseases can be diagnosed and appropriately treated before symptom onset. Accumulating data on VUSs and long-term phenotypes may provide valuable clues on pathogenicity and can be utilized for the advancement of automated analysis programs. For example, P10, referred for NST abnormality, carried missense variants c.975C>G and c.158G>A in *PAH*. The c.975C>G variant was classified as pathogenic and c.158G>A as a VUS. P10 was diagnosed as having hyperphenylalaninemia with one pathogenic missense variant because serial biochemical testing consistently indicated a mild elevation of the serum phenylalanine level. The c.158G>A variant had no clinical significance. P40 had two variants in *MCCC2*: c.838G>T was classified as likely pathogenic and c.1342G>A as a VUS. His urine 3-hydroxyisovaleric acid level was 173.1.1 mmol/mol Cr, which is 10 times higher than the normal level, and thus, he was confirmed to have 3-methyl-crotonyl-CoA carboxylase 2 deficiency with reclassification of the VUS.

This study demonstrated the importance of collaboration among field experts during panel design. Targeted RGT does not simply require a collection of genes of interest. A multidisciplinary team including clinical experts in pediatrics, clinical geneticists, clinical pathologists, and bioinformaticians should participate in gene selection for confirmative diagnosis. In particular, the details of medical actions should be clearly defined. Such a combined effort can maximize the clinical utility of targeted RGT of medically actionable genes. Moreover, predefining clear medical actions avoids ethical issues.

One of the limitations of this study was the relatively low detection rate. However, considering the benefits of detecting and managing actionable diseases at the presymptomatic stage, our diagnostic yield was acceptable.

NEOseq_ACTION panel sequencing requires only a small amount of DNA, e.g., from DBSs, and the TAT from blood collection to result reporting using the automated analytic pipeline was less than 4 days. The cost of NEOseq_ACTION sequencing is substantially lower than that of rapid WGS or WES, and be-

cause automated variant classification systems are embedded in bioinformatics pipelines, fewer experts are required. NEOseq_ACTION can be particularly advantageous in regions where resources and clinical geneticists are scarce, but rare genetic diseases are prevalent.

In conclusion, this prospective pilot study evaluated the clinical feasibility of rapid and timely diagnosis of treatable rare genetic diseases using a NEOseq_ACTION panel. Targeted RGT using the NEOseq_ACTION panel in acutely unwell neonates with suspected genetic diseases represents an example of the clinical application of genomic medicine. Considering the fragility of neonates in medical and ethical terms, the NEOseq_ACTION panel may provide a valuable opportunity for appropriate treatment of presymptomatic or early-stage diseases, even for NST in asymptomatic neonates.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Chae JH contributed to study conception and design; Chae JH, Kim SY, Lee JS, and Ko JM were involved in clinical evaluation; Kim MJ interpreted the results; Kim MJ and Kim SY drafted the manuscript; and Chae JH and Seong MW supervised the study. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

None declared.

RESEARCH FUNDING

ORCID

Man Jin Kim   https://orcid.org/0000-0002-9345-6976
Soo Yeon Kim   https://orcid.org/0000-0003-2240-3647
Jin Sook Lee   https://orcid.org/0000-0002-3652-3570
Sanggoo Kang   https://orcid.org/0000-0001-8777-1550
Lae-Jeong Park   https://orcid.org/0000-0002-2672-7270
Wooyong Choi   https://orcid.org/0000-0002-6981-899X
Ju Yeol Jung   https://orcid.org/0000-0003-3642-5345
Taehyung Kim   https://orcid.org/0000-0002-8063-8148
Sung Sup Park   https://orcid.org/0000-0003-3754-4848
Jung Min Ko   https://orcid.org/0000-0002-0407-7828
Moon-Woo Seong   https://orcid.org/0000-0003-2954-3677
Jong Hee Chae   https://orcid.org/0000-0002-9162-0138

REFERENCES

1. Meng L, Pammi M, Saronwala A, Magoulas P, Ghazi AR, Vetrini F, et al. Use of exome sequencing for infants in intensive care units: ascertainment of severe single-gene disorders and effect on medical management. JAMA Pediatr 2017;171:e173438.
2. Petrikin JE, Willig LK, Smith LD, Kingsmore SF. Rapid whole genome sequencing and precision neonatology. Semin Perinatol 2015;39:623-31.
3. Willig LK, Petrikin JE, Smith LD, Saunders CJ, Thiffault I, Miller NA, et al. Whole-genome sequencing for identification of Mendelian disorders in critically ill infants: a retrospective analysis of diagnostic and clinical findings. Lancet Respir Med 2015;3:377-87.
4. Anazi S, Maddirevula S, Faqeih E, Alsedairy H, Alzahrani F, Shamseldin HE, et al. Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. Mol Psychiatry 2017;22:615-24.
5. Stark Z, Tan TY, Chong B, Brett GR, Yap P, Walsh M, et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. Genet Med 2016;18:1090-6.
6. Vissers LELM, van Nimwegen KJM, Schieving JH, Kamsteeg EJ, Kleefstra T, Yntema HG, et al. A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. Genet Med 2017;19:1055-63.
7. Farnaes L, Hildreth A, Sweeney NM, Clark MM, Chowdhury S, Nahas S, et al. Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization. NPJ Genom Med 2018;3:10.
8. Stark Z, Lunke S, Brett GR, Tan NB, Stapleton R, Kumble S, et al. Meeting the challenges of implementing rapid genomic testing in acute pediatric care. Genet Med 2018;20:1554-63.
9. van Diemen CC, Kerstjens-Frederikse WS, Bergman KA, de Koning TJ, Sikkema-Raddatz B, van der Velde JK, et al. Rapid targeted genomics in critically ill newborns. Pediatrics 2017;140:e20162854.
10. Gyngell C, Newson AJ, Wilkinson D, Stark Z, Savulescu J. Rapid challenges: ethics and genomic neonatal intensive care. Pediatrics 2019;143:S14-21.
11. Mardis ER. The $1,000 genome, the $100,000 analysis? Genome Med 2010;2:84.
12. Li H and Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25:1754-60.
13. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.
14. 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.
15. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide

polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80-92.

16. Li Q and Wang K. InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. Am J Hum Genet 2017;100:267-80.

17. 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.

18. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;44:D862-8.

19. Stenson PD, Mort M, Ball EV, Chapman M, Evans K, Azevedo L, et al. The Human Gene Mutation Database (HGMD®): optimizing its use in a clinical diagnostic or research setting. Hum Genet 2020;139:1197-207.

20. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9.

21. Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics 2009;25:2744-50.

22. 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.

23. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One 2012;7:e46688.

24. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics 2013;14(S3):S3.

25. Chun S and Fay JC. Identification of deleterious mutations within three human genomes. Genome Res 2009;19:1553-61.

26. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res 2011;39: e118.

27. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 2010;7:575-6.

28. 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.

29. Cooper GM, Stone EA, Asimenos G, NISC Comparative Sequencing Program, Green ED, Batzoglou S, et al. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res 2005;15:901-13.

30. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 2005;15:1034-50.

31. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434-43.

32. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res 2017;45:D840-5.

33. Jung KS, Hong KW, Jo HY, Choi J, Ban HJ, Cho SB, et al. KRGDB: the large-scale variant database of 1722 Koreans based on whole genome sequencing. Database (Oxford) 2020;2020:baz146.

34. Lee S, Seo J, Park J, Nam JY, Choi A, Ignatius JS, et al. Korean Variant Archive (KOVA): a reference database of genetic variations in the Korean population. Sci Rep 2017;7:4287.

35. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 2012;28:3150-2.

36. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst 2017;30:3146-54.

37. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet 2016;99:877-85.

38. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res 2019;47:D886-94.

39. Navarrete R, Leal F, Vega AI, Morais-López A, Garcia-Silva MT, Martín-Hernández E, et al. Value of genetic analysis for confirming inborn errors of metabolism detected through the Spanish neonatal screening program. Eur J Hum Genet 2019;27:556-62.